SCUOLA DI DOTTORATO

UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA

Department of Economics, Management and Statistics

Ph.D. program in Economics, Statistics and Data Science, XXXVII cycle

Curriculum of Economics

Sustainability by Proximity: How Geographic Closeness Drives Corporate and Regional Carbon Emissions

Caterina Morelli

Registration number: 885798

Supervisor: Prof. Paolo Maranzano

Co-supervisor: Prof. Philipp Otto

Tutor: Prof. Lucia Dalla Pellegrina

Coordinator: Prof. Matteo Manera

Academic Year 2024/2025

To Angela.

Abstract

This thesis brings together four studies that follow a common thread: measuring and modeling environmental performance in an interconnected world, and building tools that turn rich spatiotemporal data into actionable insight. The work progresses from empirical economics to methodological statistics, showing how sustainability is economically material, spatially structured, and dynamically heterogeneous. The first study examines global equity reactions to the Russia–Ukraine war and finds that, while markets fell broadly, firms with stronger, industry-adjusted environmental performance suffered milder losses. Sustainability operates as a practical risk buffer under geopolitical stress. The second study maps the environmental pillar of corporate assessments for EU-based firms using multidimensional spatiotemporal clustering. It reveals coherent groups with clear geographic structure—transnational pockets of low performers in emissions-intensive activities and clusters of stronger performers in services—providing a readable geography to guide investment screening and policy targeting. The third study shifts to the emissions themselves, combining decades of regional greenhouse-gas data with spatial proximity to uncover persistent clusters of decarbonization and stagnation. A diagnostic measure, Joint Inertia, quantifies the added value of geography in explaining emission trajectories, highlighting where place-based policy and cross-regional coordination are most effective. The final study develops a cluster-wise spatial dynamic panel model that unifies interdependence and heterogeneity, allowing spillovers and responses to vary across latent groups. An estimation strategy and simulations demonstrate feasibility and accuracy, and the framework generalizes to settings where local interactions meet global shocks. Collectively, the studies offer evidence, diagnostics, and models that turn granular data into decisions—supporting resilient portfolios, credible corporate transition paths, and targeted, high-impact public policy for a low-carbon future.

Contents

Introduction

Environmental sustainability has emerged over the past decades as a defining challenge of our time. The accelerating concentration of greenhouse gases in the atmosphere—driven by production, energy use, transport, and intensive agriculture—is altering the Earth's climate system, stressing ecosystems, and amplifying the frequency and intensity of extreme weather events. In parallel, air pollution imposes severe health and economic costs worldwide. Because many of the drivers of local air pollution are also the sources of climate-warming emissions, policies that curb pollution can deliver immediate health benefits while advancing long-term decarbonization.

Firms play a central role in this picture. Across major European economies, companies account for the bulk of territorial CO_2 emissions, far outweighing the contribution of households. While aggregate emissions have declined in some countries in recent decades, corporate activities still dominate the emissions mix, underscoring the importance of production technologies, energy systems, supply chains, and corporate governance for both climate and air-quality outcomes. In response, sustainability has moved from the periphery to the core of how firms operate and how markets evaluate them. Investors, customers, workers, and regulators increasingly rely on standardized assessments of environmental, social, and governance performance to inform decisions, rewarding transparency and credible progress while pressuring laggards.

At the same time, we are experiencing a data revolution. Satellites, remote sensors, GPS-enabled devices, administrative registers, and corporate disclosures generate unprecedented volumes of spatiotemporal information. These data reveal how environmental pressures and responses evolve across space and time, how shocks propagate through networks of regions and firms, and how policies and technologies diffuse. Harnessing this abundance is both an opportunity and a challenge: it demands models that are sufficiently rich to capture complex dynamics and interconnections, yet flexible and interpretable enough to guide action.

Two features of real-world systems are especially important. First, interdependence: outcomes in one place are influenced by what happens in connected places through trade, supply chains, mobility, and shared infrastructure. Second, heterogeneity: regions and firms differ in their structures, capabilities, and constraints, so they do not respond identically to the same shock or policy. Ignoring either feature risks misleading inference and blunt policy. This motivates approaches that combine spatial and temporal information, that detect groups with similar behaviors while respecting geography, and that measure the contribution of spatial linkages beyond what is explained by observed characteristics.

This thesis grows out of that motivation. Over the last four years, my research has focused on environmental sustainability in an interconnected world, leveraging spatiotemporal methodologies—and, where needed, extending them—to meet the scale and nuance of modern data. I study emissions alongside concentrations to place attention on the sources of environmental pressure; I use clustering and dynamic modeling to uncover patterns that are invisible to average effects; and I develop tools to quantify how much spatial information actually adds to our understanding of complex phenomena. Taken together, the work aims to illuminate where and why environmental performance diverges across places and sectors, how shocks and policies ripple through economic and environmental networks, and how data-driven methods can support targeted, effective action.

This thesis brings together four studies that follow a common thread: measuring and modeling environmental performance in an interconnected world, and building tools that turn rich spatiotemporal data into actionable insight.

The first chapter, "ESG Performance and Stock Market Responses to Geopolitical Turmoil: Evidence from the Russia–Ukraine War," coauthored with Simone Boccaletti, Elisa Ossola, and Paolo Maranzano and forthcoming in Economic Modelling, asks whether sustainability helps firms withstand major shocks. Examining global equity reactions to the February 2022 invasion, it shows that markets fell broadly but not uniformly: firms with stronger ESG—especially environmental performance—faced milder losses. Sustainability emerges as a practical buffer against geopolitical risk rather than a mere reputational label.

The second chapter, "Multidimensional Spatiotemporal Clustering—An Application to Environmental Sustainability Scores in Europe," coauthored with Simone Boccaletti, Paolo Maranzano,

and Philipp Otto and published in Environmetrics, shifts from market reactions to measurement and mapping. Focusing on the environmental pillar of corporate assessments for EU-based firms, it identifies coherent groups of companies that share similar environmental profiles and reveals clear geographic patterns of leadership and lagging. The result is a readable map of corporate environmental performance that can inform investment decisions and policy priorities.

The third chapter, "Spatiotemporal Clustering of GHGs Emissions in Europe: Exploring the Role of the Spatial Component," with Paolo Maranzano and Philipp Otto and currently under review, turns to regional greenhouse-gas emissions themselves. Using long-run European data, it uncovers persistent clusters of decarbonization and stagnation and introduces a simple diagnostic to gauge how much geography shapes these patterns. The findings point to where place-based policies and cross-border coordination can be most effective.

The final chapter, "Cluster-Wise Spatiotemporal Panel Model," a work in progress, develops a modeling framework that captures both interdependence among regions and differences in how they respond to common shocks. It allows spillovers and responses to vary across latent groups, providing an interpretable tool for systems—such as environmental emissions—where local interactions and global forces intertwine.

Taken together, the thesis offers a concise arc from economics to statistics. The first study is primarily economic in focus but grounded in rigorous quantitative analysis. The second blends economic insight with statistical clustering to reveal the spatial organization of corporate sustainability. The third again combines the two, but with a stronger methodological emphasis anchored in evidence about the role of space. The fourth is largely methodological, proposing a new econometric approach motivated by real environmental data needs. Across chapters, the common aim is to turn rich spatiotemporal information into clear, actionable understanding of how sustainability unfolds in markets and across regions.

Contribution in Economics

The thesis advances economic understanding of sustainability in capital markets and the real economy along three dimensions. First, it demonstrates that firms' environmental strength can act as a shock absorber when geopolitical risk spikes. By documenting heterogeneous stock-market reactions to the Russia–Ukraine invasion and showing that higher ESG—especially environmental performance—attenuates downside risk, the first chapter reframes sustainability as a component of risk management and financial resilience, not only as a disclosure or reputational attribute. This has direct implications for asset pricing under stress, corporate risk governance, and the design of regulatory guidance that encourages credible environmental practices.

Second, it brings geographic structure into the analysis of corporate sustainability. The second chapter shows that environmental performance is not randomly distributed: it clusters across space and industry in ways that align with local regulation, resource endowments, and sectoral composition. This yields an economic map of "green" leadership and lagging that is useful for investors' screening and stewardship, for benchmarking across jurisdictions, and for policymakers seeking to calibrate incentives where they can be most effective.

Third, it connects firm-level behavior to territorial emission outcomes. The third chapter uncovers how regional greenhouse-gas trajectories coevolve across sectors and gases, revealing persistent pockets of decarbonization and stagnation. By quantifying the incremental role of geography in shaping these trajectories, it clarifies where place-based policies and interregional coordination are likely to yield the highest returns. Together, these chapters move from market responses to corporate practices and finally to regional emissions, offering a coherent economic narrative about how sustainability diffuses, how shocks propagate through networks, and where targeted intervention can accelerate transition.

Contribution in Statistics

Methodologically, the thesis contributes new tools and evidence for modeling high-dimensional spatiotemporal data under interdependence and heterogeneity. The second chapter adapts hierarchical clustering to fuse feature similarity with geographic proximity and extends it to a multidimensional, time-aware setting. This provides an interpretable way to detect groups that are coherent both in their environmental trajectories and in space—an approach well suited to modern, multivariate sustainability data.

The third chapter deepens the statistical perspective by introducing a diagnostic—Joint Inertia—that measures the marginal explanatory contribution of spatial distances beyond non-spatial dissimilarities in the clustering process. This moves spatial analysis from assumption to measurement: rather than imposing that space matters, the method quantifies how much it matters in a given application. A simulation experiment and an extensive European case study illustrate how this diagnostic guides model choice and cluster interpretation.

The fourth chapter advances the econometrics of interdependent panels by proposing a cluster-wise spatial dynamic panel model that allows both marginal effects and spillover intensities to vary within and across latent groups. It provides a quasi–maximum likelihood estimator for heterogeneous spatial and temporal dependence and a two-step procedure that jointly recovers parameters and endogenous cluster membership, freeing the analyst from predefined partitions. Simulations assess parameter recovery and the identification of direct and indirect effects under varying heterogeneity, demonstrating robustness and practical feasibility. Although motivated by environmental applications, the framework is general and can be deployed wherever local interactions and global shocks coexist.

Across these statistical contributions, the common thread is flexibility with interpretability: methods that are rich enough to capture complex spatial and temporal linkages, yet structured enough to yield transparent diagnostics, policy-relevant clusters, and estimable dynamic effects. In this way, the thesis not only analyzes sustainability with today's data, but also equips researchers and practitioners with methodological building blocks for tomorrow's questions.

Chapter 1

ESG Performance and Geopolitical Turmoil

from

ESG Performance and Stock Market Responses to Geopolitical Turmoil: evidence from the Russia-Ukraine War

Coauthored with: Simone Boccaletti, Elisa Ossola, Paolo Maranzano

Published in: *Economic Modelling*

1.1 Introduction

In recent years, Environmental, Social, and Governance (ESG) considerations have emerged as a pivotal and guiding force in redirecting financial investment towards sustainable activities. Although the focus on sustainable investing dates back to the early 2000s, it was not until the 2015 Paris Agreement that investors started to place significant weight on sustainability practices when assessing companies. This has encouraged many companies to reconsider their approach to issues of ethics, the environment and social responsibility. This movement aims to go beyond the minimum financial returns required, and to make a positive contribution to the planet and society as a whole, and it has permeated financial markets under many different perspectives (Rau & Yu, 2024). The Russian invasion of Ukraine has prompted a wide-ranging debate on the relationship between political events and responsible investment. The two main sectors affected by this event are the energy and commodity markets, given the global changes in the demand and supply market structure for electricity and raw materials. Policymakers are primarily attempting to address whether and how investments in the energy and ecological transition can mitigate the risks associated with such events. From this perspective, the conflict has highlighted the connection between political events, social stability, and the ethical implications embedded in investment decisions. In the immediate aftermath of the Russian invasion, stock markets around the world fell sharply and suddenly. This was exacerbated by the fact that the global economy was still recovering from the effects of the SARS-CoV-2 pandemic. Furthermore, Western countries promptly imposed sanctions against the Russian economy, with varying degrees of severity and differing in nature and scope. For instance, on February 27th, the EU and the US took measures against Russian airlines (e.g., banning aircraft from their airspace) and banks (e.g., blocking access to the SWIFT payment system). Another significant sanction was the ban on fuel exports and the imposition of import restrictions aimed at hampering economic activity within Russia's borders. Additionally, most large, listed companies initiated divestment and suspended their activities within the Russian market. With regard to the Organisation for Economic Co-operation and Development's (OECD) Responsible Business Conduct (RBC) principles, "Russia's war against Ukraine [...] has also revealed areas where greater clarity might be needed in how RBC principles and standards can be applied consistently in decision making by both businesses and policymakers, especially in contexts in which conflict or human

rights abuses are widespread" (OECD, 2023). The impact of the Russian invasion on sectors at a global level was evident in many ways. For example, the energy sector experienced two significant facts. First, the conflict led to a surge in investment in renewable energy sources, particularly in Europe, as a means of mitigating the rising cost of fossil fuels. Second, some countries chose to reopen coal mines and power plants, a decision that triggered major environmental debates. The study examines the impact of the event on a number of stock markets on a global scale. To gain a deeper understanding of this phenomenon, the research will also consider companies' sustainability performance as proxied by ESG ratings. The study aims to answer the following research questions: first, whether the Russian invasion generated a significant impact on worldwide companies' stock returns; second, to assess which factors were the most relevant in determining the significant influence on the stock market response at company level. To answer the first research question, we employ an event study (ES) methodology on a sample of listed companies from around the world. Specifically, we use a set of cross-sectionally adjusted and unadjusted test statistics, employing either parametric or non-parametric specifications to determine whether the event had a significant impact on stock markets. Using multiple adjusted tests is consistent with the analysis of Pelagatti and Maranzano (2021), who point out that cross-correlation adjustment is necessary when considering highly correlated time series to ensure consistent test size and power. Moreover, as many test statistics have comparable statistical properties, selecting a single test statistic is challenging. For this reason, as a by-product of the analysis, we propose a rule-of-thumb based on a majority voting approach (see Hastie et al., 2009, Section 10.1) to decide whether the event produced a significant effect or not. By applying a majority rule, we are able to identify which scenarios (i.e., country, sector, and sustainability levels and their pair combinations) are most affected by the event. The ES approach is indeed applied at the country, sector, and sustainability dimensions. Furthermore, we consider the pair combinations of country, sector, and sustainability level to account for the different degrees of specificity. To address the second research question, we examine the factors influencing the Cumulative Abnormal Returns (CARs) in the period around the event date. This investigation aims to determine whether there is a correlation between the CARs and a number of financial information of the companies, and whether the market reactions were influenced by the characteristics of the companies in question, including their sustainability performance. To the best of our knowledge, the only paper that considers ESG performance and the Russian invasion of Ukraine in an ES setting is the one by Tsang et al. (2024). The authors find evidence suggesting

that companies with higher ESG scores exhibit superior stock market performance. However, while they focus exclusively on the top 100 companies in the S&P 500 index, our study has a broader scope. Indeed, we adopt a global perspective, conduct a comprehensive country-sector analysis, and examine the factors influencing the CARs at the company level. Furthermore, in addition to standard t-statistics, we employ a range of statistical tests, accounting for the various concerns pertaining to assumptions about stock returns. Our findings can be summarized as follows. First, despite the pervasive nature of the negative consequences of the offensive, our findings suggest that the impacts were limited and significant, with notable heterogeneity observed across countries and sectors. In particular, applying ES methodology across multiple frameworks defined at country, sector, and sustainability levels enables the most relevant level of impact of the event studied to be identified. Secondly, we show that larger companies and those with lower debt ratios were able to resist the negative impact of Russian invasion on the stock markets. Finally, we find that companies with superior sustainability ratings, particularly for environmental performance, were better able to mitigate potentially negative effects than those with minimal sustainability. Our paper makes several contributions to the academic literature on the impact of the invasion on financial markets. Firstly, the large size of the sample and the level of granularity of the analysis (i.e., at sector and country levels) enable us to identify the characteristics and mechanisms through which the event had its greatest impact. These are more exhaustive and general than the impact occurred on a restricted but economically relevant sample of companies (see, for example, Tsang et al., 2024) or on global market indicators (Granát et al., 2023). Secondly, we detect the main determinants of stock market reactions, focusing on ESG performance. Our findings show that the potential negative effects of the invasion at company levels is reduced by firms' environmental performance. Eventually, we also consider the fact that event study statistics may exhibit different statistical properties. Indeed, in the common framework of ES, researchers only rely on the well-established statistics (e.g, Patell and BMP) ignoring their potential drawback as well as the existence of more robust statistics (see, for example, Obi et al., 2023 Granát et al., 2023) Conversely, this abundance of available statistics raises an issue concerning the selection of the most suitable specifications for the specific case under study. In light of the issue, we propose a rule-of-thumb approach –based on consensus of the test statistics (i.e., the majority rule) – to identify the significance of the event in the stock market. In particular, we evaluate how many statistics reject the null hypothesis of no impact of the considered event, allowing to characterize and to compare markets according to

the degree of consensus of statistics across countries, industries, and ESG performance. The rest of the paper is organized as follows. Section 1.2 revises the academic literature relevant to the paper; Section 1.3 describes the sample and outlines the empirical strategy. Section 1.4 gathers the empirical results from the event study. Section 1.5 presents the analysis of the relevant factors at company levels. Section 1.6 proposes a discussion and implications of the empirical results. Finally, Section 1.7 concludes.

1.2 Background and related literature

The impact of armed conflicts on the economies and financial markets has long been a topic of interest to academics. As Schneider and Troeger (2006) observe, the impact of conflicts on financial markets is manifold and depends on the severity of the conflict and the ability of economic agents to predict cooperative and conflictive events. Since 2022, a growing body of academic literature has attempted to analyse the impact of the Russia–Ukraine conflict on the stock market. By examining the returns of global stock market indices, several studies have identified the heterogeneous effects of the Russian invasion on the financial markets of different countries. These effects have been found to vary according to geographical proximity to the war zone (e.g., Boungou & Yatie, 2022; Yousaf et al., 2022), market efficiency (Granát et al., 2023), the level of globalization (e.g., Assaf et al., 2023; Boubaker et al., 2022; Kumari et al., 2023; Obi et al., 2023, and the economic relationship with, and dependence on, Russia and Ukraine (see, for example, Diaconaşu et al., 2023; Ferrández-Serrano & Angosto-Fernández, 2023). Among others, Yousaf et al., 2022 show that while on aggregate a significant and negative impact of the war can be identified European and Asian countries (e.g., France, Germany, Italy, Spain, India, Japan) were particularly and adversely affected. Moreover, Abbassi et al., 2023; Ahmed et al., 2023; Sun & Zhang, 2023 corroborate these findings through firm-level analysis, based on samples of global, European, and G7 firms, respectively. They also provide a more detailed analysis of the effects by country and by sector. Other studies focus on specific countries. For example, Pandey et al., 2023; Saini et al., 2023 show a negative impact of the war on the Indian financial market by examining Indian firms across different sectors. Keleş, 2023 provides evidence of a significant adverse reaction of the Turkish stock market, that began prior to the official announcement of war. In contrast, Kamal et al., 2023 find a negative impact on

Australian firms. Delving deep into the impact of the conflict on specific sectors, Boubaker et al., 2023 and Martins et al., 2024 analyze the market reactions of the banking sector at the global and European level, while Martins et al., 2023 consider the world's largest insurance companies. Moreover, Pandey & Kumar, 2023 and Le et al., 2023 focus on tourism companies and airline companies respectively, while Yudaruddin et al., 2023 show a significant negative impact of the offensive on the global consumer staples market. Many other papers focus on the energy sector. For instance, Diaconaşu et al., 2023 provide evidence of an increase in the oil price at the beginning of the war, Mohammed et al., 2023 show that renewable energy markets have a positive and significant effect, while Si Mohammed et al., 2023 find that natural gas and clean energy prices are less affected than traditional energy and metals markets. Aloui et al., 2023 describe the significant impact of the ongoing Russia–Ukraine conflict on the energy commodity markets. Overall, this has contributed to an increase in energy prices and volatility. Consequently, companies operating in the energy sector have experienced significant changes in financial returns Febriandika et al., 2023; Nerlinger & Utz, 2022, as well as the stock market indices related to the same sector Du et al., 2022; Saad, 2024; M. Umar et al., 2022. In such a delicate political context, many countries and companies have taken a stand adopting sanctions policies toward Russia. Some interesting studies show how these decisions have influenced investors' decisions and thus the impact on the market value of companies, finding controversial results. Indeed, according to Tee et al., 2023 and Nicolau et al., 2024, companies have benefited from these decisions, while Ayoub & Qadan, 2023 and French et al., 2023 show that firms involved in these actions have experienced significant stock price declines. Sanctions and political actions implemented against Russian institutions and companies are also closely linked to ESG issues. To the best of our knowledge, few papers have examined the market reactions in a scenario of geopolitical turmoil with a specific focus on sustainable practices.

Fiorillo et al. (2024) investigate the mediating role of the ESG factors in the effect of geopolitical risk on the risk of stock price crashes, examining a sample of listed global companies between 2010 and 2021. In particular, they observe that the negative impact is less severe for issuers with high ESG rates, especially for firms with high scores on the environmental and social dimensions. Ricci et al. (2024) find that the stock returns of European listed firms with high ESG scores are less affected than those of their peers during the escalation of the war and the announcement of gas import shocks from Russia in 2021-2022. Companies with higher environmental scores have higher returns.

Katsampoxakis et al. (2024) examine the performance of ESG stock indices during periods of crisis and suggest that stock indices of leading ESG companies in North America and Europe represent safe investments. Clancey-Shang and Fu (2024), using an event-study approach, consider a sample of US-listed firms and find that better ESG performance is associated with lower range volatility of stock returns, during the outbreak of the war. Tsang et al. (2024) analyze the impact of the conflict on different sectors using a sample consisting of the top 100 companies in the S&P 500 index. They find that more responsible firms were able to withstand the supply chain disruptions caused by the conflict. Alnafrah (2024) shows that non-ESG committed companies have been more vulnerable to geopolitical risk with respect to ESG-committed firms between 2021 and 2024. In light of the literature, it is reasonable to argue that ESG performance plays an important role in reducing the exposure of companies to both internal and external market events, such as the Russian invasion of Ukraine, preserving their market value (see Section 2.3 of Fiorillo et al., 2024). Nevertheless, a comprehensive and detailed study of this phenomenon would undoubtedly be beneficial in clarifying how ESG assessments have mitigated the impact of such a disruptive event in different countries and across economic sectors, particularly with regard to individual listed companies.

1.3 Data and Methodology

1.3.1 Data

The dataset includes the daily prices of all firms listed on the main stock exchanges around the world, as shown in Table 1.A.1 in the Appendix 1.A. These prices are collected from FACTSET. The dataset comprises daily prices for listed companies from July 30, 2021, to February 28, 2022, as well as the daily prices of the main stock exchange market index for each country. Companies with missing values during the specified period, and those with a constant price, which leading to a zero return, have been excluded from the sample. The final sample consists of 17,635 companies.[1]

[1]Since companies are listed on stock exchanges in more than one country, our time series are unavailable on non-weekend public holidays according to the national holidays celebrated in each country. To control for non-weekend public holidays, if there are no observations on a given day for all companies listed in the same country, the price from the previous day is used instead. According to this methodology, financial returns are set to zero when the stock exchange is closed.

We retrieved the ESG rating and score available on February 24, 2022 (i.e., the starting date of the Russian military invasion) from MSCI ESG ratings.[2] Sustainability ratings are a form of specialized assessment that evaluates a company's performance based on a set of indicators related to sustainable practices, which are broadly defined. MSCI ESG ratings are expressed as a letter grade and a numerical score adjusted for the sector in which the firm operates. Consequently, ESG ratings range from AAA (the highest level of performance) to CCC (the lowest level of performance), while ESG scores range from 0 (the worst performance) to 10 (the best performance). MSCI provides ESG ratings and scores for approximately 14,000 companies worldwide. As ESG ratings are not mandatory, only around 5,600 companies (i.e., 33% of those in the sample) have a rating, as shown in Table 1.A.1. On average, ESG scores are available for 32% of companies by country. The US is the largest market in the sample, representing around 26% of listed companies and around 43% of available ESG scores. Japan has the second largest representation of companies, but ESG coverage in Japan (20%) is below average. The listed companies are heterogeneous by country, including those accounting for less than 1% of observations (i.e., Austria, Ireland, the Netherlands, New Zealand, and Portugal). In terms of ESG availability, Switzerland, Ireland, and the Netherlands appear to be the countries where ESG reporting is available for more than 55% of companies listed on their respective stock markets. Conversely, Greece has the lowest availability rate of ESG data (8%).

1.3.2 Empirical Strategy

Let R_t be the observed return for company i (with $i = 1, \ldots, n$) at time t (with $t = 1, , T$). In the estimation phase, we assume independence across the companies. Thus, we model the returns of company i using the following firm-specific linear factor model:

$$R_t = a + \mathbf{b}' \mathbf{f}_t + \epsilon_t \tag{1.1}$$

where a is the firm-specific constant coefficient, f_t is a vector of K observable factors common to

[2]ESG ratings are usually issued once a year for each company, with a variable release date. Therefore, ESG ratings can be relatively older (in terms of months) from one company to another.

every company in a given stock market, b is the vector that collects the firm-specific factor loadings, and ϵ_t is the error term with classic assumptions.

To study the stock-price reactions to the Russia-Ukraine war, we compute the sequence of abnormal returns AR_t for each company i, i.e., the residual of returns regressed on the constant and K factors, and defined as $AR_t = R_t - [\hat{a} + \hat{b}f_t]$, where $\hat{a}$ and $\hat{b}$are estimated from the OLS regression on Eq. 1.1, using daily returns observed within the estimation window, that is, from August 3rd, 2021 to January, 31st, 2022. We implement the Capital Asset Pricing Model (Sharpe, 1964), for which the observable systematic factor ft corresponds to the excess return on the value-weighted market portfolio over the risk-free rate (see, e.g., Ramelli et al., 2021).

We adopt a short-run perspective about the event's impact on the considered markets; specifically, we define the event window for the ES as the the days immediately following or preceding February 24th, 2022. The main event window comprises the five trading days between February 21st, 2022, when news of the Russian invasion of the Donbass began to circulate, and February 25th, 2022. Therefore, the main event window Ω for our analysis is [-3;+1], effectively observing one week of the trading days. As we are interested in the overall effect of the event window, we compute the Cumulative Abnormal Returns (CARs), defined as $CAR_i = \sum_{(t)} AR_t$, with $i = 1, , n$. We then compute the average of the CARs at country, sector, and different ESG levels and as well as at their combinations, to obtain the Cumulative Average Abnormal Returns (CAARs). We also estimate CARs and average CARs over different event windows to provide additional insights and robustness checks. Table 1.A.2 in Appendix 1.A presents summary statistics for the CARs calculated over several event windows: (i) the [-3;+1] event window; (ii) the event windows including the three and one days before the event (i.e. [-3;0] and [-1;0], respectively); (iii) the day of the event (i.e. CAR at [0]); (iv) the event window including the day after the event (i.e. [0,+1]). On average, the CARs across the different event windows are negative. However, we observe that the negative sign changes when we consider the event window including the day after the event.

In our ES setting, we use a set of eighteen event study test statistics to infer the impact of the event on the stock market returns. These include the classical tests by Patell (Dodd & Warner, 1983; Patell, 1976), as well as the BMP (Boehmer et al., 1991; Sanders Jr. & Robins, 1991). However, it is worth noting that commonly applied ES statistics do not account for the cross-

sectional dependence in returns and are based on the assumption that data are Gaussian. Recently, Pelagatti and Maranzano (2021) showed that even weak cross-correlation can strongly bias the results from ES statistics, characterized by high rejection rates, and thus unreliable estimates. Accordingly, we consider an extended set of ES test statistics, including both unadjusted and cross-sectionally adjusted specifications, as well as contrasting parametric and non-parametric rank-based statistics to account for the potential non-Gaussian distribution of the data. Table 1.1 summarizes the eighteen test statistics employed in the empirical analysis. We kindly refer the reader to the paper by Pelagatti and Maranzano (2021) for a comprehensive discussion about definitions and statistical properties of the test statistics, as well as for an overview related to the issues of cross-correlation and non-Gaussianity in event studies.

Table 1.1: Overview of the statistical tests

		Specification	
		Parametric	**Non-parametric**
Cross-sectional	**Not-Adjusted**	t_{CS} t_{CD} t_{SKEW} Z_{PAT} Z_{BMP}	CumRank_Z Z_{grank}
	Adjusted	$Z_{PAT,adj}$ $Z_{BMP,adj}$	T_{grank} $Z_{grank,adj}$ CumRank CumRank_{mod} CumRank_T $\text{CumRank}_{Z,adj}$ P_1 P_2 $\text{CorradoTukey}_{adj}$

Note: The table compiles the test statistics discussed in Pelagatti and Maranzano (2021) to assess the statistical significance of CARs. We distinguish between cross-sectional *adjusted* and *not-adjusted* tests, and between *parametric* and *non-parametric* procedures.

Using multiple test statistics to determine whether an event has had a statistically significant impact on the stock market poses the challenge of selecting an appropriate assessment criterion. To address this challenge, we propose a heuristic rule based on a majority vote approach (see Hastie et al., 2009). Specifically, at a given level of statistical significance $\alpha = 0.05$, we count

the number of statistics indicating the rejection of the null hypothesis (i.e., the presence of a statistically significant effect of the event). We consider the event to be relevant -(i.e., the War in Ukraine generated a market reaction) if the majority of the tests reject the null hypothesis (i.e., if it is rejected by at least nine statistics). We believe that this consensus strategy is supported by the presence of heterogeneity in the statistical features embedded in the selected statistics. Notably, the combination of parametric and distribution-free estimators, as well as adjusted and unadjusted cross-sectional methods, encompasses a broad spectrum of empirical scenarios (e.g., heterogeneity in distribution shapes at the country or sector level or varying degrees of correlation between stocks) that a single statistic might not capture. In any case, some of the statistics on the list are inherently more robust (e.g., rank-based statistics), so they should be given greater attention and credibility.

To address the second research question - namely, whether stock market reactions to the Russian invasion of Ukraine vary across companies - we propose examining the role of several potential firm-specific determinants of the observed CARs. In particular, we are interested in studying the correlation between the CARs and financial information extracted from the balance sheets of the companies to understand whether market reactions were influenced by the firms' characteristics and by their sustainability performance. Let $CAR_(i, s, c)$ be the cumulated abnormal return during the [-3;1] event window for company i, belonging to sector s, and country c, we estimate OLS the following cross-sectional regression model

$$CAR_{(i,s,c)} = \alpha + \beta' X_i + \gamma_s' D_s + \gamma_c' D_c + \gamma_{(s,c)}' \left(D_s \times D_c\right) + \epsilon_{(i,s,c)}. \tag{1.2}$$

where X_i is the vector of selected company-specific variables,namely, Size, (natural logarithm of total assets) Leverage (the debt-to asset ratio), Profitability (the ROA), Tobin's Q and the sector-adjusted ESG score. These variables reflect fundamental aspects of firm performance and risk. Smaller firms often yield higher returns due to their greater growth potential and risk premiums. Leverage affects financial stability, while profitability signals operational efficiency and long-term value creation, which often attracts new investors. The Tobin's Q indicates market expectations versus asset replacement cost—higher values may suggest overvaluation or strong growth prospects. Lastly, strong environmental scores can enhance a firm's reputation, reduce its overall risk and attract ESG-focused capital. All variables are winsorized at the 1% level and standardized. Table

1.A.2 in Appendix 1.A reports descriptive statistics (prior to the standardization) of all the variables used in the analysis. In addition to the main covariates, we include the fixed-effect dummy variables according to the sector (D_s), the country (D_c), and the interaction between sector and country ($D_s D_c$) allowing us to capture the sector and country specificity across different dimensions. The regression is estimated through OLS and run on different event windows to understand the possible heterogeneous effect of the specific companies' characteristics in different periods.

1.4 Results of the event study

This section presents the empirical results of the analysis carried out at three layers: country, sector, and ESG. Additionally, we present results for combinations of two layers: country-sector, country-ESG, and sector-ESG. To assess the significance of the CAAR over the event window, we compute the eighteen statistics described in Table 1.1 for each layer and pairwise combination.[3] The event window is defined as $[-3; +1]$ around the date of the event (i.e., February 24th, 2022). The following section is based on the majority rule outlined in the previous section, specifically the number of tests out of eighteen that reject the null hypothesis that the CAAR is equal to zero at the 5% significance level. Accordingly, with a slight abuse of notation, we conclude that the effect of the event is statistically significant if more than nine out of eighteen test statistics reject the null hypothesis of no effect. Conversely, the event is deemed not to be significant if fewer than nine tests reject the null hypothesis.

1.4.1 Results at Country, Sector, and ESG levels

Hereafter, we describe the results collected by conducting analysis at country, sector, and ESG levels. For each analysis, we provide plots illustrating the statistical results of the ES statistics. In Appendix 1.A, Tables 1.A.4, 1.A.5, and 1.A.6 provide sample descriptive statistics and summary results of the tests at country, sector, and ESG levels, respectively. For completeness, in Appendix 1.B, Figures 1.B.1-1.B.6 show the effective number of statistically significant tests for each of the

[3] As described in the previous section, we have also computed the CAAR on different event windows, namely [-3;0], [-1;0], [0], and [0;+1]. Extended results are available upon request.

eighteen statistics used by country, sector, ESG, and at the combination levels, respectively. These plots clearly show that the rejection rate of the null hypothesis varies across the set of considered statistics. This evidence could support the use of majority rules, instead of single tests, in ES when choosing the proper test statistics is difficult. Focusing on the analysis at the country level, Figures 1.1 and 1.2 show the CAAR values and the number of statistically significant tests (i.e., p-value below 0.05). A predominantly negative impact is observed in all countries, except for the US, which has a positive but statistically insignificant CAAR. Looking at the negative impact, we observe heterogeneous results across countries. Notably, some European countries exhibit a strong and significant impact. In particular, the impact is highly statistically significant in a few countries: Austria, France, and Great Britain reject the null hypothesis that the CAAR is equal to zero for more than twelve test statistics. This result is confirmed when only the adjusted test statistics are considered (see Table 1.A.4). In Italy, the test statistic is rejected for at least nine test statistics. Outside Europe, the only relevant result is observed for Australian companies, for which the CAARs are negative and statistically significant values for most of the statistics. A slightly positive but statistically not significant effect in Japan and the Netherlands.

Therefore, not all countries show clear evidence of a negative impact. In fact, Greece, Japan, Norway, Sweden, and, perhaps most notably, the United States and Canada, demonstrate few, if any, strong negative effects. This pattern suggests that geographic, economic, and institutional factors play a buffering role. Countries with limited direct trade or financial exposure to Russia and Ukraine, or those with greater macroeconomic resilience, may be better positioned to withstand the secondary effects of the invasion. Earlier academic studies in the United States have pointed out that the invasion affected US-listed foreign firms and domestic firms differently, which could help to explain there has been no significant impact on this country (Clancey-Shang & Fu, 2023). Figure 1.3 and Table 1.A.5 show the aggregated results at the sector level, indicating a negative average response. In particular, the largest and most statistically significant effects are found in activities relating to sustainability. Specifically, the largest negative effect (-2.1%) is found in the "Agriculture, forestry, fishing and hunting" sector, for which we collect eight out of eighteen reject test statistics for the null hypothesis of no effect. We also observe a negative and significant impact on companies belonging to the sectors "Construction" (-1.9%), "Mining, quarrying, and oil and gas extraction" (-1.7%), and "Administrative and support service activities" (-1.6%). This could

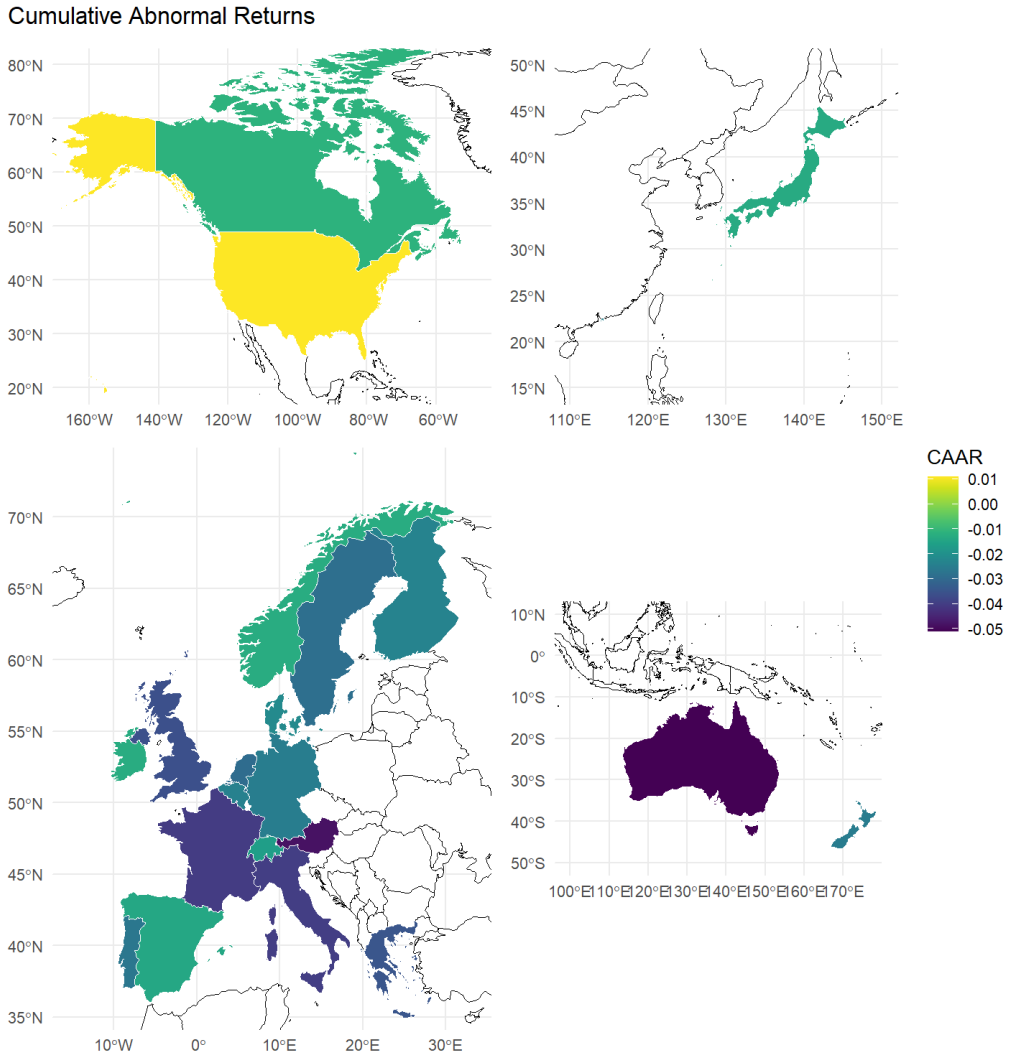


Figure 1.1: Map of CAAR by country. This map shows the intensity of stock market reactions to the Russian invasion at the country level, measured by the CAAR.

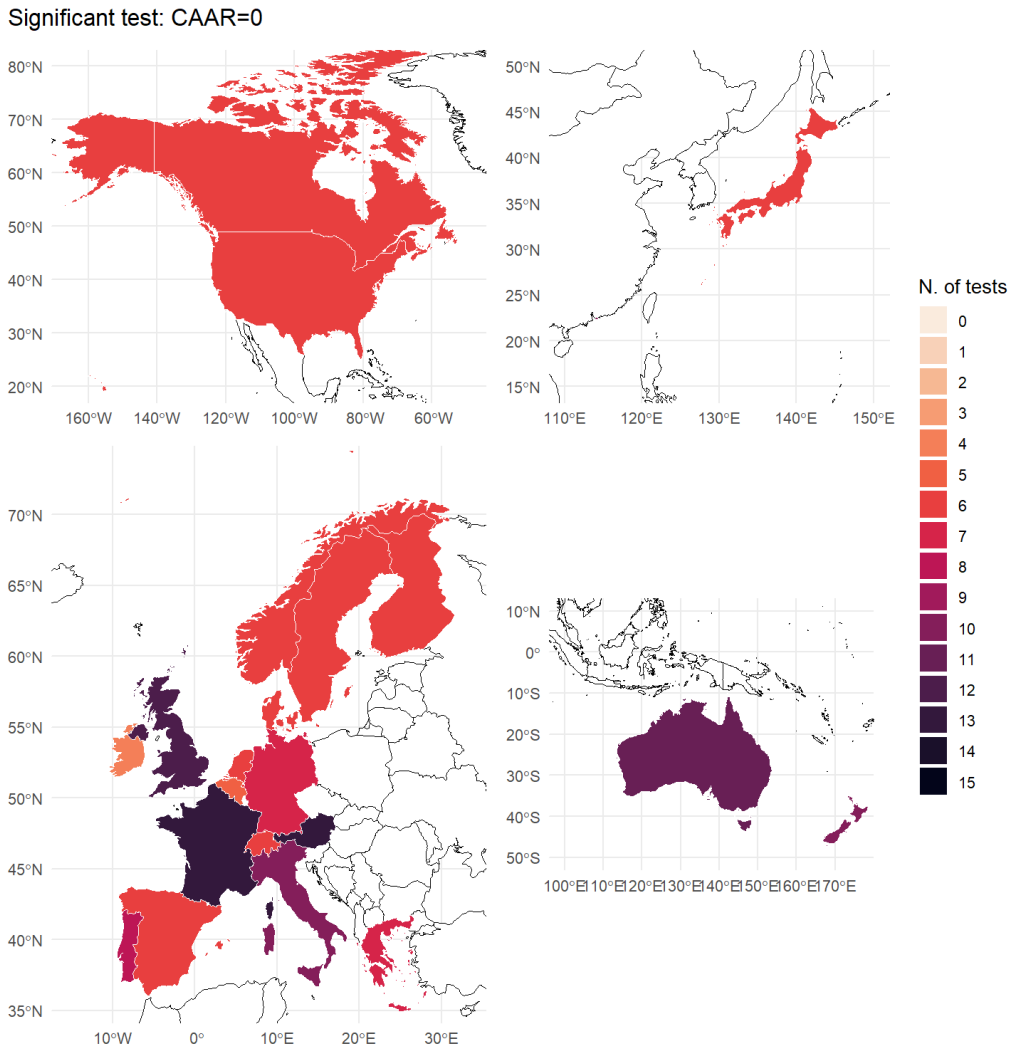


Figure 1.2: Map of the number of significant tests. This map shows, at the country level, the number of tests from Table 1.1 that reject the null hypothesis of CAAR being not statistically different from zero (i.e., the p-value is lower than 5%).

be due to the relevant role of energy and raw material markets in relation to Russia's war against Ukraine, as discussed in Aloui et al. (2023). Some effects are also observed in the sectors that indirectly received a crucial role in the energy market during this event, e.g. "Accommodation and food service", "Manufacturing" and "Retail trade", for which we get negative CAAR and a statistically significant result on six out of eighteen statistics. Focusing on the statistical tests, the adjusted test statistics do not reject zero effects in all sectors, except for 'Agriculture, forestry, fishing and hunting'. There is also a negative CAAR (-1.7%), with six tests showing a p-value lower than 0.05, in 'Financial intermediation'. This result may reflect an increase in the perceived risk of the sector, as most banks and financial institutions may have faced an increase in interest rate risk, together with an increase in the credit risk of borrowers, and thus potential losses on investments. Finally, we also observe a positive effect for companies belonging to the "Utilities" sector, but its significance is only registered for one test statistic.

Figure 1.3: CAAR and number of statistically significant test statistics by sector

The figure shows, at the sector level, the size of CAAR and the number of tests from Table 1.1 that reject the null hypothesis of CAAR being not statistically different from zero (i.e., the p-value is lower than 5%).

Considering the results in Figure 1.4 and Table 1.A.6, the analysis shows that unrated companies (no ESG rating) experience the most significant negative impact, with a number of notable results. For instance, companies with missing ESG scores due to a lack of disclosed non-financial information

appear to be associated with a loss of around 2%. Missing ESG information appears to play a crucial role in this study (see, e.g., the discussion in Sahin et al., 2022). For companies with available ESG scores, the impact is negative, greater than 1%, and not rejected for 33% of the statistics for the highest scorers (AAA-rated companies). For other ratings - i.e. from AA to CCC - the CAARs show smaller effects that are not statistically different from zero, i.e. there is no evidence of the effect of the event under review.

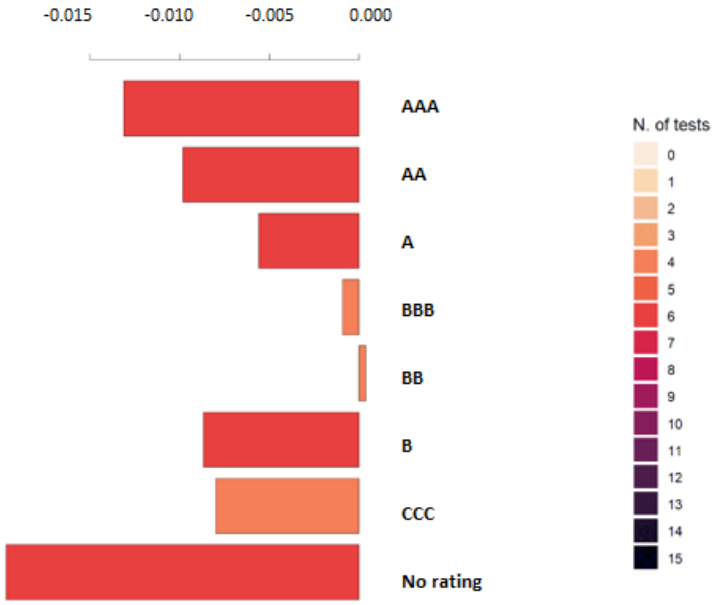


Figure 1.4: CAAR and number of statistically significant test statistics by ESG rating and availability of rating. The figure shows, at the ESG rating, the size of CAAR and the number of tests from Table 1.1 that reject the null hypothesis of CAAR being not statistically different from zero (i.e., the p-value is lower than 5%).

1.4.2 Results at country-sector level

Table 1.2 shows the CAAR for each pair of countries and sectors in our sample. This analysis reveals specific patterns not evident in the aggregate picture presented in Section 1.4.1. In particular, it allows us to identify the sectors that were the main drivers of the impact of the Russian invasion in

each country. In general, heterogeneous effects are evident; however, the vast majority of country-sector pairs show a negative response. Any positive magnitudes in the CAARs are statistically not significant. Some of the most negative responses at the country level can be traced back to specific pairs. For instance, the negative response for Austria is primarily due to the negative impact of enterprises belonging to the "Finance and Insurance" (-6.4%) and 'Manufacturing' sectors. In France, the sectors that explain its highly significant negative CAAR are mainly driven by agriculture and manufacturing. In Great Britain, however, the negative response is driven by several sectors with a negative and statistically significant impact of around -4% or more (e.g. for the enterprises in the "Health and care and social assistance"' sector). Focusing on Australia, for which the event has a large negative impact at the country level, this is mainly driven by companies belonging to the "Accommodation and food service", "Manufacturing" and "Utilities" sectors. In the case of Japan and the Netherlands, we observe a large number of positive but statistically insignificant impacts across activities, in line with the results in Figure 1.2.

Table 1.2: Results from the analysis at the country–sector level

	(1) AFS	(2) ASWMRS	(3) AFFH	(4) AER	(5) CON	(6) EDU	(7) FIN	(8) HCSA	(9) INFO
Australia	-0.104^{++}	-0.045	-0.021	-0.067^{++}	-0.034	-0.084	-0.021	-0.012	-0.057^{+}
Austria	-0.021				-0.035		-0.064^{++}		-0.031
Belgium		-0.040	0.038		-0.013		-0.004		-0.030^{+}
Canada	0.002	-0.039	-0.042	-0.028	0.013	-0.007	-0.011	-0.012	-0.028
Denmark		-0.058	0.016	0.022	-0.019		-0.036^{++}	0.011	-0.015
Finland	-0.006	-0.012			-0.021		-0.020	-0.030	-0.015
France	-0.020	-0.033	-0.087^{++}	-0.006	-0.031	-0.131	-0.048^{++}	0	-0.041^{+}
Germany	-0.044^{+}	0.005	-0.009	-0.026	0.006	0.033	-0.038^{+++}	-0.030	-0.035^{++}
Great Britain	-0.024^{+}	-0.041^{++}	-0.031	-0.046	-0.030	-0.041^{++}	-0.043^{+++}	-0.051^{++}	-0.038^{++}
Greece	0.015		-0.001	-0.068	-0.045		-0.032	-0.024	-0.061
Hong Kong	-0.010	0.008	-0.039	-0.002	-0.035^{++}	-0.071	-0.032^{+++}	-0.006	-0.030
Ireland	0.043			0.002			-0.023	-0.045^{+}	0.003
Italy	0.009	-0.062^{++}	-0.053^{+++}	-0.030	-0.069^{++}	-0.106	-0.058^{+++}	-0.037	-0.042^{++}
Japan	-0.022	-0.008	-0.010	-0.021	-0.013	-0.017	-0.031	0.006	0.001
Netherlands	-0.053	-0.053		-0.032	-0.055^{+}		-0.040	0.056	-0.055
New Zealand	-0.017	-0.126	-0.037^{++}			-0.004	-0.052^{+++}	0.011	-0.026
Norway		0.027	-0.023	-0.007	-0.005	-0.054	-0.023		-0.025
Portugal	-0.021				0.007		-0.055	-0.130	-0.053^{+}
Singapore	-0.029	-0.032	-0.005	-0.020	-0.010	0.023	-0.021^{+}	-0.043^{++}	-0.020
Spain	-0.007	-0.025	0.013		-0.008		-0.015	-0.060	-0.005
Sweden	-0.035	-0.016	-0.051	-0.014	-0.021	-0.005	-0.003	-0.028	-0.026
Switzerland	-0.027			-0.127	-0.013		-0.034^{++}	0.004	-0.063
US	-0.009	-0.001	0.028	0.004	-0.007	0.036	0.006	0.016	0.016

Note: The table reports the estimated CAAR by applying the event study methodology at the country–sector level. Each column represents a sector; each row corresponds to a country. Blank cells indicate a lack of observations for the corresponding country–sector combination. Estimates are evaluated using the majority vote approach described in Section 3.2. Statistical significance at the 5% level is denoted as follows: $^{+++}$ indicates significance in 15–18 tests from Table 1.1; $^{++}$ in 11–14 tests; and $^{+}$ in 9–10 tests. Abbreviations: AFS = Accommodation and Food Services; ASWMRS = Administrative and Support; Waste Management and Remediation Services; AFFH = Agriculture, Forestry, Fishing and Hunting; AER = Arts, Entertainment and Recreation; CON = Construction; EDU = Educational Services; FIN = Finance and Insurance; HCSA = Health Care and Social Assistance; INFO = Information.

Table 1.3: Results from the analysis at the country–sector level (continued)

	(10) MCE	(11) MAN	(12) MQOGE	(13) OSEPA	(14) PSTS	(15) RERL	(16) RET	(17) TW	(18) UTL	(19) WHT
Australia		-0.063^{+++}	-0.055	-0.040	-0.067^{+}	-0.026	-0.045	-0.017	-0.085^{++}	-0.052^{+}
Austria		-0.041^{+++}	-0.168			-0.028		-0.064	0.017	-0.127
Belgium		-0.060^{++}		-0.018	-0.062^{+++}	0.040	-0.008	0.051		0.007
Canada		-0.013	-0.009	0.001	0.011	-0.002	-0.011	-0.005	0.007	-0.027
Denmark		-0.020	-0.028		-0.014	-0.055	-0.079	-0.026	0.158	-0.015
Finland		-0.028	-0.001	-0.031	-0.017	-0.007	-0.014	-0.039	-0.158	-0.08^{++}
France		-0.049^{+++}	0.049	-0.018	-0.032	-0.028	-0.047^{+++}	-0.011	0.013	-0.044^{++}
Germany	-0.019	-0.030	0.023	0.005	-0.019	0.005	-0.021	-0.052^{+}	-0.016	-0.012
Great Britain		-0.049^{++}	-0.013	-0.025	-0.031	-0.010	-0.041^{+++}	-0.029	0.014	-0.043^{++}
Greece		-0.032	-0.036	-0.004	-0.019	-0.041	-0.042	-0.018		-0.044
Hong Kong	-0.047	-0.023^{+}	-0.007	-0.031	-0.024	-0.021^{++}	-0.023	-0.048^{+++}	-0.035	0.006
Ireland		0.008	0.169		0.038	-0.043		-0.029		-0.115
Italy		-0.040^{++}	-0.041	-0.011	-0.062^{+++}	0.002	-0.046	-0.008	0.022	-0.008
Japan	-0.005	-0.018	0.017	-0.016	0.002	-0.020	-0.012	-0.021	-0.002	-0.010
Netherlands		-0.021	-0.041	-0.001	-0.019	-0.054	-0.042	-0.076	-0.005	
New Zealand	-0.027	0.056		-0.049	-0.023	-0.005	-0.015	-0.024	-0.006	
Norway	-0.012	0.032		-0.012	-0.014	-0.023	-0.011	-0.005	-0.037	
Portugal	-0.039^{++}		0.032		-0.002		-0.029	-0.088	0.027	
Singapore	0.060	-0.007	-0.031	0.005	-0.021	-0.028	-0.019	-0.007	-0.015	-0.028
Spain	-0.008	-0.020		-0.046^{+}	-0.043^{+++}	-0.071	0.001	0.032	-0.049^{++}	
Sweden	-0.028	-0.031	0.065	-0.033	-0.041	-0.027	-0.038	-0.008	-0.065	-0.007
Switzerland	0.017	-0.008		0.017	-0.004	0.032	-0.022	0	-0.007	
US	-0.031	0.011	0.034	-0.003	0.019	0.013	0.009	0.007	0.031	-0.003

Note: See notes under Table 1.2. Abbreviations: MCE = Management of Companies and Enterprises; MAN = Manufacturing; MQOGE = Mining, Quarrying, and Oil and Gas Extraction; OSEPA = Other Services (except Public Administration); PSTS = Professional, Scientific, and Technical Services; RERL = Real Estate and Rental and Leasing; RET = Retail Trade; TW = Transportation and Warehousing; UTL = Utilities; WHT = Wholesale Trade.

The "Agriculture, forestry, fishing, and hunting" sector is confirmed as the most affected sector across countries, with statistically significant results in France, Italy, and New Zealand. Contrary to expectations, no significant results are found when also looking at the country level for the "Mining, Quarrying and Oil and Gas extraction sector" sector, which had a large negative CAAR overall (-1.7% in Table 1.A.5). Given the importance of fossil fuel markets to the conflict—especially in the context of sanctions, supply disruptions, and shifts in energy flows—this absence of significant impact is surprising. In fact, it may suggest that firms operating in this sector were already priced in geopolitical risks or possessed structural buffers. The large negative effect shown in Figure 1.3 for the "Finance and Insurance" sector is observed and statistically significant for around 50% of the countries in the sample. A similar effect is also observed in 35% of countries for enterprises in the 'Manufacturing' sector. Thus, these sectors appear to be the ones for which the event is systematically distributed across countries. The frequent statistical significance observed in the manufacturing sector likely reflects its exposure to global value chain disruptions. The war has created logistical bottlenecks, raised the cost of raw materials, and increased geopolitical uncertainty—all of which disproportionately affect manufacturing-intensive firms reliant on cross-border operations. Meanwhile, the finance sector appears to be sensitive to the reconfiguration of international capital flows, heightened risk premiums, and potential sanctions or exposure to Russian assets. The prevalence of significant results in these sectors highlights how systemic and globally interconnected areas of economic activity are particularly vulnerable to large-scale geopolitical events. Lastly, the U.S. case is intriguing: while some sectors show positive market responses, the number of significant results never exceeds eight in any one sector. This suggests a potentially fragmented or modestly beneficial effect.

1.4.3 Results at Country and ESG level

Table 1.3 shows the results at the country-ESG level. The results of Figure 1.4 are also confirmed by this analysis. As in the previous analysis, we observe a strong negative effect for the unrated companies. Looking at the EU countries, we find a strong negative significant effect for the no-rated companies in Austria, France, and Italy. This highlights that, even when the sample includes countries with the same disclosure requirements, the effects at the firm level may differ. Focusing on

non-EU countries, we observe negative effects for unrated companies in Australia, Great Britain, and Hong Kong. For the latter, the negative impact is mainly explained by the AAA-rated companies (see Figure 1.4). We do not observe any other strong impact for AAA-rated companies in the other countries. This could be due to the reduced number of observations that the analysis at this level entails. This analysis shows that ESG metrics only explain the impact of the Russian invasion in certain countries. Indeed, for 13 out of 23 countries in the sample, the results at the ESG level are not statistically significant. Finally, we do find no strong evidence that the results discriminate between more and less sustainable companies.

Table 1.4: Results from the analysis at country–ESG level

	AAA	**AA**	**A**	**BBB**	**BB**	**B**	**CCC**	**no rating**
Australia	0.021	-0.035	-0.043	-0.036	-0.020	-0.061^{+}	0.066	-0.056^{++}
Austria	-0.049	-0.103^{+++}	-0.047	-0.031			-0.021	-0.036^{++}
Belgium	-0.033	-0.020	0.033	0.004	-0.020	-0.003		-0.040^{+}
Canada	0.016	0.002	0.009	0.006	0.008	0.011	-0.010	-0.015
Denmark	0.056	-0.012	0.010	-0.005	-0.028			-0.031
Finland	-0.050	-0.014	-0.005	-0.026	-0.039			-0.023
France	-0.047	-0.020	-0.002	-0.038	-0.021	-0.024		-0.046^{+++}
Germany	-0.039^{+}	-0.028	-0.022	-0.011	0.004	0.010	-0.002	-0.027^{+}
Great Britain	-0.014	-0.020	-0.017	-0.022	-0.012	-0.147^{++}	-0.089	-0.041^{++}
Greece		-0.064		-0.024	-0.030			-0.035
Hong Kong	-0.074^{+++}	-0.053^{++}	-0.027	-0.041^{++}	-0.036^{++}	-0.036^{++}	-0.049^{++}	-0.020^{+}
Ireland	-0.039	-0.017	-0.068	-0.010	-0.068^{++}	-0.030		0.016
Italy	-0.038	-0.018	-0.029	-0.035	-0.044	-0.030		-0.043^{+++}
Japan	-0.016	-0.026	-0.023	-0.018	-0.013	-0.016	-0.023	-0.011
Netherlands	-0.019	-0.053	-0.050	-0.017	-0.044			-0.013
New Zealand	-0.036	-0.017	-0.036^{+}	-0.019	-0.038			-0.026^{+}
Norway	-0.009	-0.017	0.002	0.022	0.034	0.010		-0.016
Portugal	0.056	-0.055^{++}		-0.069^{+++}	0.015			-0.016
Singapore	-0.023	-0.048^{++}	-0.015	-0.013	-0.014	-0.030^{+}	0.011	-0.016^{+}
Spain	-0.006	-0.023	0.013	-0.004	-0.040	-0.027	0.003	-0.018
Sweden	-0.015	-0.017	-0.020	-0.029	-0.024	-0.054		-0.030
Switzerland	-0.011	-0.006	-0.008	-0.008	-0.017	-0.032		-0.026
US	0.016	0.027	0.014	0.017	0.016	0.005	0.019	0.006

Note: The table reports the estimated CAAR by applying the event study methodology at the country–ESG level. Each column represents an ESG rating class; each row corresponds to a country. Blank cells indicate a lack of observation for the corresponding country–ESG combination. Estimates are evaluated using the majority vote approach described in Section 3.2. Statistical significance at the 5% level is denoted as follows: $^{+++}$ indicates significance in 15–18 of the test statistics from Table 1.1; $^{++}$ in 11–14 tests; and $^{+}$ in 9–10 tests.

1.4.4 Results at ESG and sector level

Table 1.4 shows the results at the sector ESG level. We observe the heterogeneous response of different pairs of sustainability and sector levels to the Russian invasion of Ukraine. Both positive and negative impacts are estimated. However, we only find negative and low statistical significance (i.e., less than ten statistics are significant over the eighteen tests) for unrated companies belonging to the "Construction" and "Finance and Insurance" sectors. Thus, as in the case of the country-ESG level characteristics, no clear differences in the response of more sustainable companies compared to less sustainable companies could be identified. Nevertheless, the impact on unrated companies is still evident.

In Hong Kong, the role of ESG ratings adds another layer to the interpretation. In this context, the significant effects among firms with ESG scores are concentrated in four sectors, i.e., "Real Estate and Rental and Leasing "; "Transportation and Warehousing"; "Construction"; and "Finance and Insurance". This may reflect heightened sensitivity among ESG-rated firms to investor sentiment and geopolitical instability. Alternatively, it could suggest that ESG transparency itself acts as a conduit for market reactions, with firms that are more transparent about their governance and sustainability practices becoming focal points for portfolio reallocation in periods of uncertainty. A different dynamic is observed in France and the United Kingdom, where the results appear to be predominantly driven by firms without ESG ratings. This asymmetry suggests that the limited sustainability disclosure may correlate with higher perceived risk or lower investor confidence in times of geopolitical stress. It may also reflect the structural characteristics of these firms themselves, such as smaller size, lower international diversification, or weaker governance standards, which amplify their vulnerability in the absence of ESG signaling.

1.5 Channels at company level

Table 1.5 shows the results of estimating Eq. 1.2 for various event windows, namely, [-3;1] (the main event window), [-3;0], [-1;0], [0] (i.e., the event date only), and [0;+1]. Specifications (1)-(5) are estimated using 14,896 observations, which include the full sample but exclude the ESG score.

Table 1.5: Results from the analysis at the sector–ESG level

	AAA	**AA**	**A**	**BBB**	**BB**	**B**	**CCC**	**no rating**
Accommodation and Food Services		-0.019	-0.027	-0.015	-0.020	-0.003	-0.027	-0.020^{+}
Administrative and Support and Waste Management and Remediation Services		-0.034	-0.005	-0.023	0.005	-0.010	-0.036	-0.015
Agriculture, Forestry, Fishing and Hunting	0.048	-0.028	-0.027	0.000	-0.002			-0.024
Arts, Entertainment, and Recreation	-0.064	-0.015	-0.032	-0.008	0.076	0.008	-0.046	-0.025
Construction	-0.010	-0.022	-0.019	-0.014	-0.022	-0.010	0.039	-0.019^{+}
Educational Services	0.053	-0.035	0.070	0.009	0.083	-0.054	0.020	-0.028
Finance and Insurance	-0.041	-0.039	-0.027	-0.016	-0.006	-0.010	-0.003	-0.016^{+}
Health Care and Social Assistance		0.017	0.028	0.048	0.057	0.046	0.080	-0.021
Management of Companies and Enterprises				-0.003	-0.023			-0.020
Manufacturing	-0.010	-0.001	-0.003	0.002	-0.004	-0.022	-0.007	-0.021
Mining, Quarrying, and Oil and Gas Extraction	-0.017	-0.002	0.000	0.014	0.021	0.015	-0.069	-0.021
Information	-0.003	-0.007	0.008	0.009	0.023	0.017	0.028	-0.024
Other Services (except Public Administration)	-0.015	-0.009	0.001	0.003	0.004	-0.020		-0.019
Professional, Scientific, and Technical Services	0.000	-0.009	0.009	0.010	0.035	0.019	0.023	-0.015
Real Estate and Rental and Leasing	-0.019	0.003	-0.006	-0.003	-0.013	-0.013	-0.009	-0.011
Retail Trade	-0.032	-0.020	-0.020	-0.002	-0.005	-0.006	-0.020	-0.020
Transportation and Warehousing	-0.015	-0.023	-0.017	-0.007	-0.018	-0.015	-0.073	-0.013
Utilities	0.052	0.019	0.020	0.002	-0.016	-0.004	0.019	-0.009
Wholesale Trade	-0.027	-0.007	-0.009	-0.005	0.010	-0.020	-0.031	-0.022

Note: The table reports the estimated CAAR by applying the event study methodology at the sector–ESG level. Each column corresponds to an ESG rating; each row to a sector. Blank cells indicate a lack of observations for the corresponding sector–ESG combination. Estimates are evaluated using the majority vote approach in Section 3.2. Statistical significance at 5% is denoted as follows: $^{+++}$ indicates significance in 15–18 of the tests from Table 1.1; $^{++}$ in 11–14 tests; and $^{+}$ in 9–10 tests.

Specifications (6) to (10) and (11) to (15) introduce the ESG score and the E, S, and G pillars, respectively; therefore the analysis is performed on the subsample that includes only companies for which the sustainability information is available. All specifications include country and sector dummies, as well as the interaction term between-country and sector, including all the possible layers - studied in the previous event study analysis.

Table 1.6: Estimation results of the analysis on the CAR at the company level (models 1–15)

Specification	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**	**(13)**	**(14)**	**(15)**
Event window	$[-3;+1]$	$[-3;0]$	$[-1;0]$	$[0]$	$[0;+1]$	$[-3;+1]$	$[-3;0]$	$[-1;0]$	$[0]$	$[0;+1]$	$[-3;+1]$	$[-3;0]$	$[-1;0]$	$[0]$	$[0;+1]$
Leverage	-0.039***	-0.049***	-0.046***	-0.052***	-0.035**	-0.053***	-0.071***	-0.037**	-0.037**	-0.016	-0.054***	-0.072***	-0.038**	-0.037**	-0.017
	(0.011)	(0.011)	(0.011)	(0.010)	(0.011)	(0.019)	(0.017)	(0.017)	(0.017)	(0.018)	(0.018)	(0.017)	(0.017)	(0.017)	(0.018)
Tobin's Q	0.006	-0.043***	-0.002	-0.003	0.057***	0.069***	0.033	0.034	0.078***	0.116***	0.063***	0.028	0.029	0.076***	0.114***
	(0.012)	(0.012)	(0.012)	(0.011)	(0.011)	(0.023)	(0.023)	(0.021)	(0.021)	(0.021)	(0.023)	(0.023)	(0.021)	(0.021)	(0.021)
Profitability	0.061***	0.065***	0.055***	0.024*	0.018	-0.073*	-0.121***	-0.179***	-0.314***	-0.220***	-0.058	-0.109***	-0.167***	-0.311***	-0.214***
	(0.013)	(0.013)	(0.013)	(0.012)	(0.013)	(0.043)	(0.043)	(0.046)	(0.047)	(0.042)	(0.043)	(0.042)	(0.045)	(0.047)	(0.042)
Size	0.056***	0.046***	0.041***	0.074***	0.084***	-0.101***	-0.099***	-0.177***	-0.155***	-0.142***	-0.121***	-0.118***	-0.196***	-0.162***	-0.147***
	(0.011)	(0.011)	(0.011)	(0.010)	(0.011)	(0.024)	(0.022)	(0.022)	(0.021)	(0.023)	(0.024)	(0.023)	(0.023)	(0.021)	(0.022)
ESG Score						0.038***	0.034**	0.018	0.020	0.026*					
						(0.015)	(0.014)	(0.014)	(0.013)	(0.014)					
Environmental Score											0.066***	0.054***	0.046***	0.021	0.032**
											(0.016)	(0.015)	(0.015)	(0.014)	(0.015)
Social Score											0.023*	0.035***	0.020*	0.019*	0.006
											(0.012)	(0.012)	(0.012)	(0.011)	(0.012)
Governance Score											0.019	0.016	0.006	0.020	0.023
											(0.015)	(0.014)	(0.015)	(0.014)	(0.014)
Country dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Sector dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Country × sector dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	14,896	14,896	14,896	14,896	14,896	4,921	4,921	4,921	4,921	4,921	4,921	4,921	4,921	4,921	4,921
R-squared	0.121	0.146	0.280	0.222	0.211	0.211	0.225	0.257	0.445	0.377	0.214	0.228	0.259	0.445	0.377

Note: The table reports the estimation results of Eq. 1.2 of the cumulative abnormal returns computed on several event windows. All specifications include country and sector dummies and their interaction. Robust standard errors are shown in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Specifications (1) to (5) show that loadings for Size and Leverage are constantly significant throughout the considered event windows, whereas coefficients for Tobin's Q and Profitability are only significant in certain event windows. Specifically, the positive beta coefficient for Size indicates that larger firms are associated with "less-negative" market reactions. Larger firms are generally better able to withstand geopolitical shocks than smaller firms due to their ability to diversify their investments, to redirect capital and to access to different markets. The negative coefficient of Leverage is always statistically significant at the 1% level, indicating that firms with relatively high debt-to-asset ratio are perceived as riskier during geopolitical conflicts, resulting in lower CAR (see Tosun et al., 2024 for similar results). Profitability also seems to affect CAR positively, albeit only in the event windows [-3;+1], [-3;0] and [-1;0]. Given that geopolitical risks and events typically reduce firm profitability (see, e.g., Reyad et al., 2024), our analysis unveils that higher profitability allows firms to withstand the negative consequences of the Russian invasion of Ukraine. The introduction of the ESG score in specifications (6) to (10) is both relevant and interesting. Indeed, we observe that a higher level of sustainability has a positive effect on CARs at firm level before and after the event (i.e., specifications 6, 7, and 10). The results that include each pillar in the three metrics are of interest. The environmental pillar plays a crucial role in all event windows considered, except for the date of the event. Instead, the social pillar appears to be a positive and highly significant driver for the pre-event windows. Therefore, at the company level, better sustainability performance (specifically, better environmental performance) seems to be able to reduce the potential negative impact of Russia's invasion of Ukraine. This finding aligns with existing literature on sustainability practices and geopolitical risks (e.g., Fiorillo et al., 2024; Ricci et al., 2024). Lastly, to consider potential additional channels at the company level that might drive companies' CAR, we re-estimate model in Eq. 1.2 using the full sample (i.e., including firms without an ESG score) and by adding, one at a time, the dividend-per-share, the ratio between tangible assets and total assets and the ratio between short-term debt and long-term debt. Results are collected in Table 1.6. Dividend-per-share are typically associated with stock returns and market expectations about a company payout policy (see, e.g., Black & Scholes, 1974; Fama & French, 1988). The tangible ratio, defined as tangible assets over total assets, can typically be seen as a proxy of collateral in corporate financing. Lastly, the short-term debt-to-total debt ratio can be seen as an indicator of riskiness: the higher this ratio, the greater the company's obligation to repay or refinance short-term debt, potentially signalling liquidity concerns or heightened financial vulnerability.

Table 1.7: Estimation results of the analysis on the CAR at the company level

Specification	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
Event window:	$[-3;+1]$	$[-3;0]$	$[-1;0]$	$[0]$	$[0;+1]$
Panel A					
Dividend per share	-0.00552	0.0111	-0.0129**	-0.0228***	-0.0399***
	(0.00694)	(0.00677)	(0.00651)	(0.00666)	(0.00670)
Controls and constant term	YES	YES	YES	YES	YES
Country × sector dummies	YES	YES	YES	YES	YES
Observation	14,867	14,867	14,867	14,867	14,867
R-squared	0.121	0.147	0.147	0.281	0.223
Panel B					
Tangible assets / Total assets	0.00190	0.0240**	0.0230**	0.0211**	-0.00368
	(0.0109)	(0.0105)	(0.0104)	(0.00974)	(0.0102)
Controls and constant term	YES	YES	YES	YES	YES
Country × sector dummies	YES	YES	YES	YES	YES
Observation	13,932	13,932	13,932	13,932	13,932
R-squared	0.120	0.145	0.145	0.271	0.214
Panel C					
Short term debt / Total Debt	-0.0137	-0.0110	-0.0228**	-0.0239**	-0.0264***
	(0.0107)	(0.0106)	(0.0105)	(0.00952)	(0.00993)
Controls and constant term	YES	YES	YES	YES	YES
Country × sector dummies	YES	YES	YES	YES	YES
Observation	13,625	13,625	13,625	13,625	13,625
R-squared	0.126	0.150	0.157	0.294	0.240

Note: The table reports estimation results of Eq. 1.2 of the cumulative abnormal returns computed on several event windows. All specifications include control variables (size, leverage, profitability, and Tobin's Q) and dummy variables at the country and sector levels, and their interaction. Robust standard errors are shown in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Overall, the effect of the dividend per share on the CAR appears to be marginally significant. We found a statistically significant negative result at the 1% level only on the event date and within the event window [0,+1]. This indicates that firms with a high dividend per share experience more negative CARs, suggesting that, due to heightened geopolitical risks, these firms may fail to meet market expectations regarding their dividend policies. The tangible asset ratio appears to have a small positive effect on the CAR, indicating that more asset-heavy firms are better positioned to cope with the consequences of the Russian invasion. Lastly, the short-term debt ratio seems to have a negative effect on the CAR, but only on the day of the invasion. This aligns with expectations, as firms with higher short-term debt may face difficulties in meeting their obligations due to the increased overall risk.

1.6 Discussion and implications

Geopolitical shocks on a global scale are a major concern for society as a whole. Such events can have a significant impact on populations and human rights, as well as posing serious risks to companies and participants in financial markets. From a financial market perspective, such events increase economic uncertainty, particularly affecting international firms and those with more irreversible investments. In fact, WANG et al. (2024) document a strong negative correlation between increased geopolitical risk and corporate investment. The Russian invasion of Ukraine is one of the most significant geopolitical shocks since the Second World War (Boubaker et al., 2023), and has had various implications for the global economy from different perspective, e.g., increased in raw material prices, supply chain disruption, increased financial markets uncertainty, societal turmoil and energy supply disruptions from Russia. In addition, the relationship and interconnectedness among financial markets has changed since the war started (Z. Umar et al., 2022). Since 2022, many papers have analyzed the impact of the Russian invasion on the global economy and financial markets from different perspectives. Some studies focus on specific sectors, such as banking, energy, or food, while others take a more general approach. Our analysis aims to disentangle the various layers of specificity that a major conflict can involve, particularly with regard to country, sector, and sustainability performance. Our main finding is that the Russian invasion of Ukraine had heterogeneous and asymmetric economic effects on a global scale. Similar to Yousaf et al. (2022),

we showed that not all countries or sectors have been impacted in the same way. Specifically, the absence of significant effects in countries like Canada, Norway, and Japan implies some economies are more insulated, either due to geographic distance, energy independence, or strong institutional frameworks. This asymmetry challenges the concept of uniform global contagion suggesting that risk exposure should be modeled with greater granularity in both macroeconomic and asset-pricing frameworks. Secondly, our country-sector level results point out that sectoral interconnectedness could magnify vulnerability (see Tsang et al., 2024). The importance of the manufacturing and finance sectors in many countries underscores how global integration creates channels for the rapid transmission of shocks. Manufacturing's exposure to global supply chains and the link between finance and capital flows mean that these sectors respond quickly and visibly to geopolitical instability. In order to mitigate the impact of such instability, policymakers and firms in these sectors may need to diversify suppliers, stress-test their systems, and strengthen their scenario planning for geopolitical contingencies. Thirdly, our event study analysis indicates that, at the time of the invasion—and after controlling for companies' ESG ratings—stronger sustainability performance did not appear to mitigate the immediate market impact. Specifically, there were no significant differences in market reactions between companies with higher and lower ESG ratings. However, examining the specific transmission channels at the firm level, we observed a robust, positive, and statistically significant relationship between environmental performance and CARs. This finding, in line with Fiorillo et al. (2024), suggests that strong environmental performance can play a beneficial role in enhancing firms' resilience over time, potentially offsetting the adverse effects of heightened geopolitical risk. Lastly, we emphasize once again the importance of the consensus rule-of-thumb used to identify the statistical significance of the event of interest. ES statistics are often constructed under strong assumptions (e.g., gaussianity of the returns) or do not explicitly account for typical empirical issues of financial data (e.g., asymmetry and cross-correlation). However, for many of them, suitable statistical properties can be demonstrated (e.g., exact size and power properties for many rank-based ES statistics). In this context, it becomes difficult to reach an ideal decision and choose a single representative statistic. Therefore, we recommend leveraging the extensive list of available statistical tests as much as possible and evaluating the degree of consistency and agreement in the resulting findings. Nevertheless, it is important to note that this strategy is valid when dealing with heterogeneous statistics that are constructed taking into account the mixed characteristics of the data. In this specific case, we have used adjusted and unadjusted statistics for

cross-correlation, but also statistics based on parametric assumptions and other distribution-free statistics.

1.7 Conclusion

In this paper, we showed that the stock market reaction to the Russian invasion of Ukraine is markedly heterogeneous with respect to a variety of factors, including country and sector characteristics as well as firms' performance in sustainable practices. We focus on a short-term analysis that considers the immediate impact of the military invasion on a large set of worldwide financial and real markets. However, longer-term measurement windows (e.g., 12-month or longer) could also be adopted to capture additional insights of the war effects on stock markets (see, e.g., Kothari & Warner, 2007; Sorescu et al., 2017). Our approach grounds on the perception/assumption that the abnormal/unexpected effect of large-scale and scarcely predictable[4] events like armed conflicts between large sovereign countries are automatically and instantaneously incorporated into the prices in stock markets, which promptly react without delays. In the context, we employed a range of test statistics that directly account for the statistical properties of stock market returns. This approach allows for the potential violation of the normality assumption and the strong cross-sectional dependence to be taken into consideration. Our findings indicate that, while the event had an overall negative impact on most countries, country-sector specificities are relevant in explaining the stock market reaction to the increase in geopolitical risks. We find that ESG performance has a positive impact on cumulative abnormal returns, underscoring the positive role of sustainable investing in the current era of heightened geopolitical uncertainty. Our study also underscores the need to evaluate corporate sustainability practices from two perspectives: a broad assessment of corporate sustainability performance in an international and evolving context, and a more granular analysis that considers the differential impact of a single event (such as an escalation of geopolitical turmoil resulting from the Russia-Ukraine conflict) in different contexts (e.g., different countries or sectors). Our analysis provides some relevant insights for policymakers. By identifying the sectors in which the impact of the event was more severe, and given the mitigating effects of sustainable practices,

[4]We avoid using the word 'unpredictable' because in the case of Russia and Ukraine, as highlighted in several studies (see, e.g., Osflaten, 2025; Roberts, 2022), the preconditions for the conflict were well known and long-standing.

policymakers can use the approach presented in the paper to redirect investment in sustainable practices in specific contexts, i.e. specific countries and sectors, to withstand potential geopolitical shocks.

Abbassi, W., Kumari, V., & Pandey, D. K. (2023). What makes firms vulnerable to the Russia–Ukraine crisis? *Journal of Risk Finance*, *24*(1), 24–39. https://doi.org/10.1108/JRF-05-2022-0108

Ahmed, S., Hasan, M. M., & Kamal, M. R. (2023). Russia–Ukraine crisis: The effects on the European stock market. *European Financial Management*, *29*(4), 1078–1118. https://doi.org/10.1111/eufm.12386

Alnafrah, I. (2024). ESG practices mitigating geopolitical risks: Implications for sustainable environmental management. *Journal of Environmental Management*, *358*. https://doi.org/10.1016/j.jenvman.2024.120923

Aloui, D., Benkraiem, R., Guesmi, K., & Mzoughi, H. (2023). Managing natural resource prices in a geopolitical risk environment. *Resources Policy*, *83*. https://doi.org/10.1016/j.resourpol.2023.103628

Assaf, R., Gupta, D., & Kumar, R. (2023). The price of war: Effect of the Russia-Ukraine war on the global financial market. *Journal of Economic Asymmetries*, *28*. https://doi.org/10.1016/j.jeca.2023.e00328

Ayoub, M., & Qadan, M. (2023). Does supporting Ukraine pay well? The performance of companies that suspended their business in Russia. *Research in International Business and Finance*, *66*, 102071. https://doi.org/doi.org/10.1016/j.ribaf.2023.102071

Black, F., & Scholes, M. S. (1974). The effects of dividend yield and dividend policy on common stock prices and returns. *Journal of Financial Economics*, *1*(1), 1–22. https://doi.org/10.1016/0304-405X(74)90006-3

Boehmer, E., Masumeci, J., & Poulsen, A. B. (1991). Event-study methodology under conditions of event-induced variance. *Journal of Financial Economics*, *30*(2), 253–272. https://doi.org/10.1016/0304-405X(91)90032-F

Boubaker, S., Goodell, J. W., Pandey, D. K., & Kumari, V. (2022). Heterogeneous impacts of wars on global equity markets: Evidence from the invasion of Ukraine. *Finance Research Letters*, *48*. https://doi.org/10.1016/j.frl.2022.102934

Boubaker, S., Nguyen, N., Trinh, V. Q., & Vu, T. (2023). Market reaction to the Russian Ukrainian war: A global analysis of the banking industry. *Review of Accounting and Finance*, *22*(1), 123–153. https://doi.org/10.1108/RAF-10-2022-0294

Boungou, W., & Yatie, A. (2022). The impact of the Ukraine–Russia war on world stock market returns. *Economics Letters*, *215*. https://doi.org/10.1016/j.econlet.2022.110516

Clancey-Shang, D., & Fu, C. (2023). The Russia–Ukraine conflict and foreign stocks on the US market. *Journal of Risk Finance*, *24*(1), 6–23. https://doi.org/10.1108/JRF-07-2022-0179

Clancey-Shang, D., & Fu, C. (2024). CSR disclosure, political risk and market quality: Evidence from the Russia–Ukraine conflict. *Global Finance Journal*, *60*. https://doi.org/10.1016/j.gfj.2024.100938

Diaconaşu, D. E., Mehdian, S. M., & Stoica, O. (2023). The reaction of financial markets to Russia's invasion of Ukraine: Evidence from gold, oil, bitcoin, and major stock markets. *Applied Economics Letters*, *30*(19), 2792–2796. https://doi.org/10.1080/13504851.2022.2107608

Dodd, P., & Warner, J. B. (1983). On corporate governance. A study of proxy contests. *Journal of Financial Economics*, *11*(1-4), 401–438. https://doi.org/10.1016/0304-405X(83)90018-1

Du, J., Li, Z., & Wang, J. (2022). The reaction of energy markets to regional conflict: Evidence from event study approach. *Procedia Computer Science*, *214*(100), 935–942. https://doi.org/10.1016/j.procs.2022.11.262

Fama, E. F., & French, K. R. (1988). Dividend yields and expected stock returns. *Journal of Financial Economics*, *22*(1), 3–25. https://doi.org/10.1016/0304-405X(88)90020-7

Febriandika, N. R., Wati, R. M., & Hasanah, M. (2023). Russia's invasion of Ukraine: The reaction of Islamic stocks in the energy sector of Indonesia. *Investment Management and Financial Innovations*, *20*(1), 218–227. https://doi.org/10.21511/imfi.20(1).2023.19

Ferrández-Serrano, V., & Angosto-Fernández, P. L. (2023). A Russian carol: Impact and dependence on global equity markets during the Ukraine invasion. *Journal of Economic Studies*, *50*(7), 1399–1412. https://doi.org/10.1108/JES-06-2022-0364

Fiorillo, P., Meles, A., Pellegrino, L. R., & Verdoliva, V. (2024). Geopolitical risk and stock price crash risk: The mitigating role of ESG performance. *International Review of Financial Analysis*, *91*. https://doi.org/10.1016/j.irfa.2023.102958

French, J. J., Gurdgiev, C., & Shin, S. (2023). When doing the right thing doesn't pay: Impact of corporate decisions on Russian market participation in the wake of the Ukraine-Russia war. *Finance Research Letters*, *58*. https://doi.org/10.1016/j.frl.2023.104468

Granát, M. P., Lehmann, K., Nagy, O., & Neszveda, G. (2023). Expect the unexpected: Did the equity markets anticipate the Russo-Ukrainian war? *Finance Research Letters*, *58*. https://doi.org/10.1016/j.frl.2023.104301

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer.

Kamal, M. R., Ahmed, S., & Hasan, M. M. (2023). The impact of the Russia–Ukraine crisis on the stock market: Evidence from Australia. *Pacific Basin Finance Journal*, *79*. https://doi.org/10.1016/j.pacfin.2023.102036

Katsampoxakis, I., Xanthopoulos, S., Basdekis, C., & Christopoulos, A. G. (2024). Can ESG stocks be a safe haven during global crises? evidence from the COVID-19 pandemic and the Russia–Ukraine war with time-frequency wavelet analysis. *Economies*, *12*(4).

Keleş, E. (2023). Stock market response to the Russia–Ukraine war: Evidence from an emerging market. *Journal of East-West Business*, *29*(3), 307–322. https://doi.org/10.1080/10669868.2023.2210121

Kothari, S., & Warner, J. B. (2007). Econometrics of event studies. *Handbook of Empirical Corporate Finance SET*, *2*, 3–36. https://doi.org/10.1016/B978-0-444-53265-7.50015-9

Kumari, V., Kumar, G., & Pandey, D. K. (2023). Are the European Union stock markets vulnerable to the Russia–Ukraine war? *Journal of Behavioral and Experimental Finance*, *37*. https://doi.org/10.1016/j.jbef.2023.100793

Le, V. H., von Mettenheim, H.-J., Goutte, S., & Liu, F. (2023). News-based sentiment: Can it explain market performance before and after the Russia–Ukraine conflict? *Journal of Risk Finance*, *24*(1), 72–88. https://doi.org/10.1108/JRF-06-2022-0168

Martins, A. M., Correia, P., & Gouveia, R. (2023). Russia–Ukraine conflict: The effect on European banks' stock market returns. *Journal of Multinational Financial Management*, *67*. https://doi.org/10.1016/j.mulfin.2023.100786

Martins, A. M., Correia, P., & Gouveia, R. (2024). The impact of the Russia–Ukraine war on the world's largest listed insurance firms. *Geneva Papers on Risk and Insurance: Issues and Practice*, *49*(4), 779–803. https://doi.org/10.1057/s41288-023-00305-w

Mohammed, K. S., Usman, M., Ahmad, P., & Bulgamaa, U. (2023). Do all renewable energy stocks react to the war in Ukraine? Russo-Ukrainian conflict perspective. *Environmental Science and Pollution Research*, *30*(13), 36782–36793. https://doi.org/10.1007/s11356-022-24833-5

Nerlinger, M., & Utz, S. (2022). The impact of the Russia–Ukraine conflict on energy firms: A capital market perspective. *Finance Research Letters*, *50*. https://doi.org/10.1016/j.frl.2022.103243

Nicolau, J. L., Poretti, C., & Heo, C. Y. (2024). Solidarity actions with Ukraine and hospitality firm value. *Tourism Management*, *100*. https://doi.org/10.1016/j.tourman.2023.104813

Obi, P., Waweru, F., & Nyangu, M. (2023). An event study on the reaction of equity and commodity markets to the onset of the Russia–Ukraine conflict. *Journal of Risk and Financial Management*, *16*(5). https://doi.org/10.3390/jrfm16050256

OECD. (2023). Responsible business conduct implications of Russia's invasion of Ukraine, OECD policy responses on the impacts of the war in Ukraine. *OECD Publishing, Paris*.

Osflaten, A. (2025). Russian forecasting and pre-emption: The prelude to the invasion of Ukraine. *Scandinavian Journal of Military Studies*, *8*(1), 57–73. https://doi.org/10.31374/sjms.361

Pandey, D. K., Assaf, R., & Rai, V. K. (2023). Did the Indian stock market sail the Russia–Ukraine storm safely? *Journal of Economic Asymmetries*, *28*. https://doi.org/10.1016/j.jeca.2023.e00319

Pandey, D. K., & Kumar, R. (2023). Russia–Ukraine war and the global tourism sector: A 13-day tale. *Current Issues in Tourism*, *26*(5), 692–700. https://doi.org/10.1080/13683500.2022.2081789

Patell, J. M. (1976). Corporate forecasts of earnings per share and stock price behavior: Empirical test. *Journal of Accounting Research*, *14*(2), 246–276.

Pelagatti, M. M., & Maranzano, P. (2021). Nonparametric tests for event studies under cross-sectional dependence. *Quarterly Journal of Finance and Accounting*, *59*(1/2), 29–74.

Ramelli, S., Ossola, E., & Rancan, M. (2021). Stock price effects of climate activism: Evidence from the first global climate strike. *Journal of Corporate Finance*, *69*. https://doi.org/10.1016/j.jcorpfin.2021.102018

Rau, P., & Yu, T. (2024). A survey on ESG: Investors, institutions and firms. *China Finance Review International*, *14*, 3–33.

Reyad, H. M., Ayesha, M., Iqbal, M. M., & Zariyawati, M. A. (2024). The role of ESG in enhancing firm resilience to geopolitical risks: An eastern European perspective. *Business Strategy and Development*, *7*(4). https://doi.org/10.1002/bsd2.70027

Ricci, O., Santilli, G., Scardozzi, G., & Stentella Lopes, F. S. (2024). ESG resilience in conflictual times. *Research in International Business and Finance*, *71*. https://doi.org/10.1016/j.ribaf.2024.102411

Roberts, G. (2022). 'now or never': The immediate origins of Putin's preventative war on Ukraine. *Journal of Military and Strategic Studies*, *22*(2).

Saad, G. (2024). The impact of the Russia–Ukraine war on the United States natural gas futures prices. *Kybernetes*, *53*(10), 3430–3443. https://doi.org/10.1108/K-01-2023-0138

Saini, M., Yadav, M., Aggarwal, V., & Kumar, M. (2023). All weather friends: How did the Russia–Ukraine war impact Indian stock markets? *International Journal of Diplomacy and Economy*, *9*(1), 93–111. https://doi.org/10.1504/IJDIPE.2023.10053176

Sanders Jr., R. W., & Robins, R. P. (1991). Discriminating between wealth and information effects in event studies in accounting and finance research. *Review of Quantitative Finance and Accounting*, *1*(3), 307–329. https://doi.org/10.1007/BF02408383

Schneider, G., & Troeger, V. E. (2006). War and the world economy: Stock market reactions to international conflicts. *Journal of Conflict Resolution*, *50*(5), 623–645. https://doi.org/10.1177/0022002706290430

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, *19*(3), 425–442. https://doi.org/10.1111/j.1540-6261.1964.tb02865.x

Si Mohammed, K., Khalfaoui, R., Doğan, B., Sharma, G. D., & Mentel, U. (2023). The reaction of the metal and gold resource planning in the post-COVID-19 era and Russia-Ukrainian conflict: Role of fossil fuel markets for portfolio hedging strategies. *Resources Policy*, *83*. https://doi.org/10.1016/j.resourpol.2023.103654

Sorescu, A., Warren, N. L., & Ertekin, L. (2017). Event study methodology in the marketing literature: An overview. *Journal of the Academy of Marketing Science*, *45*(2), 186–207. https://doi.org/10.1007/s11747-017-0516-y

Sun, M., & Zhang, C. (2023). Comprehensive analysis of global stock market reactions to the Russia–Ukraine war. *Applied Economics Letters*, *30*(18), 2673–2680. https://doi.org/10.1080/13504851.2022.2103077

Tee, C.-M., Wong, W.-Y., & Hooy, C.-W. (2023). Financial sanctions and global stock market reaction: Evidence from the Russia–Ukraine conflict. *Finance Research Letters*, *58*. https://doi.org/10.1016/j.frl.2023.104398

Tosun, O. K., Eshraghi, A., & Vigne, S. A. (2024). Firms entangled in geopolitical conflicts: Evidence from the Russia–Ukraine war. *Journal of International Money and Finance*, *147*. https://doi.org/10.1016/j.jimonfin.2024.103137

Tsang, Y., Fan, Y., Feng, Z., & Li, Y. (2024). Examining supply chain vulnerability via an analysis of ESG-prioritized firms amid the Russian-Ukrainian conflict. *Journal of Cleaner Production*, *434*. https://doi.org/10.1016/j.jclepro.2023.139754

Umar, M., Riaz, Y., & Yousaf, I. (2022). Impact of Russian-Ukraine war on clean energy, conventional energy, and metal markets: Evidence from event study approach. *Resources Policy*, *79*. https://doi.org/10.1016/j.resourpol.2022.102966

Umar, Z., Polat, O., Choi, S.-Y., & Teplova, T. (2022). The impact of the Russia–Ukraine conflict on the connectedness of financial markets. *Finance Research Letters*, *48*. https://doi.org/10.1016/j.frl.2022.102976

WANG, X., WU, Y., & XU, W. (2024). Geopolitical risk and investment. *Journal of Money, Credit and Banking*, *56*(8), 2023–2059. https://doi.org/doi.org/10.1111/jmcb.13110

Yousaf, I., Patel, R., & Yarovaya, L. (2022). The reaction of G20+ stock markets to the Russia–Ukraine conflict "black-swan" event: Evidence from event study approach. *Journal of Behavioral and Experimental Finance*, *35*. https://doi.org/10.1016/j.jbef.2022.100723

Yudaruddin, R., Fitriansyah, Lesmana, D., Bintoro, R. F. A., Purnomo, A. H., Nugroho, B. A., & Santi, E. N. (2023). Does invasion russia-ukraine affect to global financial market? evidence from consumers' staples sectors. *Journal of Open Innovation: Technology, Market, and Complexity*, *9*(3), 100086. https://doi.org/https://doi.org/10.1016/j.joitmc.2023.100086

1.A Appendix: Additional Tables

This appendix provide supplementary table. Tables 1.A.1 - 1.A.3 give an overview of the data included in the analysis. Tables 1.A.4 - 1.A.6 provide details on the empirical results described in Section 1.4.

Table 1.A.1: Summary statistics

Variable	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
CAR $[-3;+1]$	14896	-0.014	0.063	-0.237	0.183
CAR $[-3;0]$	14896	-0.032	0.062	-0.252	0.145
CAR $[-1;0]$	14896	-0.014	0.048	-0.178	0.140
CAR $[0]$	14896	-0.010	0.045	-0.155	0.133
CAR $[0;+1]$	14896	0.008	0.052	-0.157	0.184
Leverage	14896	0.238	0.244	0.000	1.391
Tobin's Q	14896	1.993	3.189	0.025	21.228
ROA	14896	-7.376	38.161	-246.482	30.508
Size	14896	5.687	2.400	0.056	12.109
Dividend per Share	14867	0.384	1.057	0.000	8.033
Tangible asset / Total assets	13932	0.865	0.196	0.216	1.000
Short-term debt / Total Debt	13625	0.350	0.305	0.000	1.000
ESG Score	4921	5.310	2.135	0.000	10
Environmental Score	4921	5.129	2.227	0.100	10
Social Score	4921	4.712	1.559	0.800	8.5
Governance Score	4921	5.165	1.335	1.8	7.9

Note: The table shows the descriptive statistics of the variables involved in the analysis. Specifically, it displays the number of observations, the mean, the standard deviation, and the minimum and maximum values. All variables are winsorized at the first and the 99th percentiles.

Table 1.A.2: Distribution of the sample by countries and market indexes

Country	**Index**	**Num. of Firms**	**Share (to total sample)**	**Num. of available ESG**	**Share (to total available)**	**Avail. rate**
Australia	S&P ASX 200 Index	1,188	*6.84%*	254	*4.50%*	21%
Austria	ATX Index	56	*0.32%*	26	*0.46%*	46%
Belgium	BEL 20 Index	115	*0.66%*	45	*0.80%*	39%
Canada	S&P/TSX Index	1,501	*8.64%*	250	*4.43%*	17%
Denmark	OMX Copenhagen 20 Index	127	*0.73%*	35	*0.62%*	28%
Finland	OMX Helsinki Index	163	*0.94%*	41	*0.73%*	25%
France	CAC 40 Index	634	*3.65%*	134	*2.37%*	21%
Germany	DAX Index	643	*3.70%*	161	*2.85%*	25%
Great Britain	FTSE 100 Index	1,148	*6.61%*	344	*6.09%*	30%
Greece	Athens Stock Exch. Large Cap Index	128	*0.74%*	10	*0.18%*	8%
Hong Kong	Hang Seng Index	1,000	*5.76%*	176	*3.12%*	18%
Ireland	ISEQ All-Share	23	*0.13%*	13	*0.23%*	57%
Italy	FTSE MIB Index	352	*2.03%*	90	*1.59%*	26%
Japan	NIKKEI 225	3,729	*21.47%*	1,116	*19.77%*	30%
Netherlands	AEX Index	94	*0.54%*	54	*0.96%*	57%
New Zealand	S&P/NZX 50	96	*0.55%*	46	*0.81%*	48%
Norway	OSE all Share Index	259	*1.49%*	57	*1.01%*	22%
Portugal	PSI 20 Index	34	*0.20%*	12	*0.21%*	35%
Singapore	Straits Times Index	412	*2.37%*	68	*1.20%*	17%
Spain	Ibex 35 Index	178	*1.03%*	66	*1.17%*	37%
Sweden	OMX Stockholm 30 Index	740	*4.26%*	110	*1.95%*	15%
Switzerland	SPI SWISS Performance Index	210	*1.21%*	121	*2.14%*	58%
US	S&P 500	4,535	*26.12%*	2,417	*42.81%*	53%
Total		*17,365*	*100%*	*5,646*	*100%*	*33%*

Note: The table reports the market indexes and their corresponding country included in the analysis. The table also reports the number of observations available in our database, including information on ESG rating availability. The column "Availability rate" refers to the proportion of ESG score available at country level.

Table 1.A.3: Event study by country

Country	Num. of observation	CAAR	% of tests significant	% of signif. CS adjust test
Australia	1,188	-0.051^{++}	61.11%	36.36%
Austria	56	-0.048^{++}	72.22%	54.55%
Belgium	115	-0.024	27.78%	0.00%
Canada	1,501	-0.011	33.33%	0.00%
Denmark	127	-0.021	33.33%	0.00%
Finland	163	-0.023	33.33%	0.00%
France	634	-0.040^{++}	72.22%	54.55%
Germany	643	-0.025	38.89%	9.09%
Great Britain	1,148	-0.036^{++}	66.67%	45.45%
Greece	128	-0.035	38.89%	18.18%
Hong Kong	1,000	-0.023^{+}	55.56%	36.36%
Ireland	23	-0.013	22.22%	0.00%
Italy	352	-0.040^{+}	55.56%	27.27%
Japan	3,729	-0.013	33.33%	0.00%
Netherlands	94	-0.029	33.33%	0.00%
New Zealand	96	-0.025^{+}	55.56%	36.36%
Norway	259	-0.013	33.33%	0.00%
Portugal	34	-0.027	44.44%	27.27%
Singapore	412	-0.017	33.33%	0.00%
Spain	178	-0.014	33.33%	0.00%
Sweden	740	-0.029	33.33%	0.00%
Switzerland	210	-0.016	33.33%	0.00%
US	4,535	0.011	33.33%	0.00%

Note: The table reports the estimated CAAR by applying the event study methodology at the country level. The estimates are evaluated using the majority vote approach described in Section 3.2. Statistical significance at the 5% level is denoted as follows: $^{+++}$ indicates significance in 15–18 of the test statistics from Table 1; $^{++}$ in 11–14 test statistics; and $^{+}$ in 9–10 test statistics. The table also reports the percentage of significant results among the 18 tests, the percentage of cross-sectionally (CS) adjusted tests rejecting the null hypothesis, and the number of observations used in the analysis.

Table 1.A.4: Event study by NAICS Sector

Sector (NAICS classification)	**Num. of observation**	**CAAR**	**% of tests significant**	**% of signif. CS adjust test**
AFS	327	-0.018	33.33%	0.00%
ASWMRS	362	-0.016	27.78%	0.00%
AFFH	140	-0.021	38.89%	9.09%
AER	145	-0.018	33.33%	0.00%
CON	641	-0.019	33.33%	0.00%
EDU	75	-0.015	22.22%	0.00%
FIN	2,165	-0.017	27.78%	0.00%
HCSA	213	-0.003	0.00%	0.00%
INFO	1,727	-0.014	33.33%	0.00%
MCE	23	-0.019	0.00%	0.00%
MAN	5,889	-0.016	33.33%	0.00%
MQOGE	1,573	-0.017	33.33%	0.00%
OSEPA)	73	-0.013	22.22%	0.00%
PSTS	1,224	-0.009	33.33%	0.00%
RERL	781	-0.010	33.33%	0.00%
RET	753	-0.016	33.33%	0.00%
TW	405	-0.015	33.33%	0.00%
UTL	297	0.003	5.56%	0.00%
WHT	551	-0.018	33.33%	0.00%

Note: The table reports the estimated CAAR by applying the event study methodology at sector level. The estimates are evaluated using the majority vote approach described in Section 3.2. Statistical significance at the 5% level is denoted as follows: $^{+++}$ indicates significance in 15–18 of the test statistics from Table 1; $^{++}$ in 11–14 test statistics; and $^{+}$ in 9–10 test statistics. The table also reports the percentage of significant results among the 18 tests, the percentage of cross-sectionally (CS) adjusted tests rejecting the null hypothesis, and the number of observations used in the analysis. Abbreviations: AFS = Accommodation and Food Services; ASWMRS = Administrative and Support; Waste Management and Remediation Services; AFFH = Agriculture, Forestry, Fishing and Hunting; AER = Arts, Entertainment and Recreation; CON = Construction; EDU = Educational Services; FIN = Finance and Insurance; HCSA = Health Care and Social Assistance; INFO = Information; MCE = Management of Companies and Enterprises; MAN = Manufacturing; MQOGE = Mining, Quarrying, and Oil and Gas Extraction; OSEPA = Other Services (except Public Administration); PSTS = Professional, Scientific, and Technical Services; RERL = Real Estate and Rental and Leasing; RET = Retail Trade; TW = Transportation and Warehousing; UTL = Utilities; WHT = Wholesale Trade.

Table 1.A.5: Event study by ESG rating

Country	Num. of observation	CAAR	% of tests significant	% of signif. CS adjust test
AAA	191	-0.013	33.33%	0.00%
AA	896	-0.010	33.33%	0.00%
A	1,329	-0.006	33.33%	0.00%
BBB	1,366	-0.001	22.22%	0.00%
BB	1,079	0.000	22.22%	0.00%
B	673	-0.009	33.33%	0.00%
CCC	112	-0.008	22.22%	0.00%
no rating	11,719	-0.020	33.33%	0.00%

Note: The table reports the estimated CAAR by applying the event study methodology by ESG rating. The estimates are evaluated using the majority vote approach described in Section 3.2. Statistical significance at the 5% level is denoted as follows: $^{+++}$ indicates significance in 15–18 of the test statistics from Table 1; $^{++}$ in 11–14 test statistics; and $^{+}$ in 9–10 test statistics. The table also reports the percentage of significant results among the 18 tests, the percentage of cross-sectionally (CS) adjusted tests rejecting the null hypothesis, and the number of observations used in the analysis.

1.B Appendix: Number of significant test statistics

For each empirical exercise implemented in Section 1.4, this appendix provides plots representing the number of significant results, by the statistics listed in Table 1.1.

Figure 1.B.1: Number of significant test statistics for the analysis by country

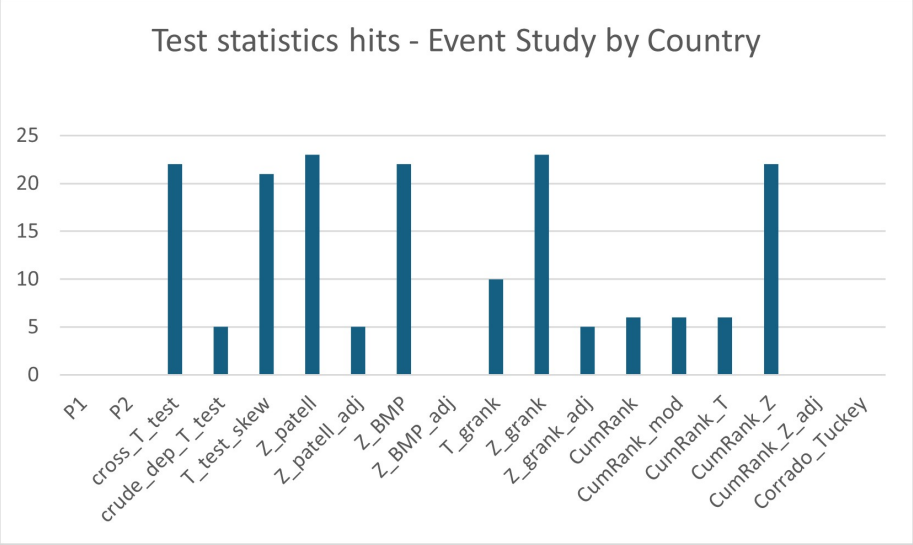


Figure 1.B.2: Number of significant test statistics for the analysis by sector

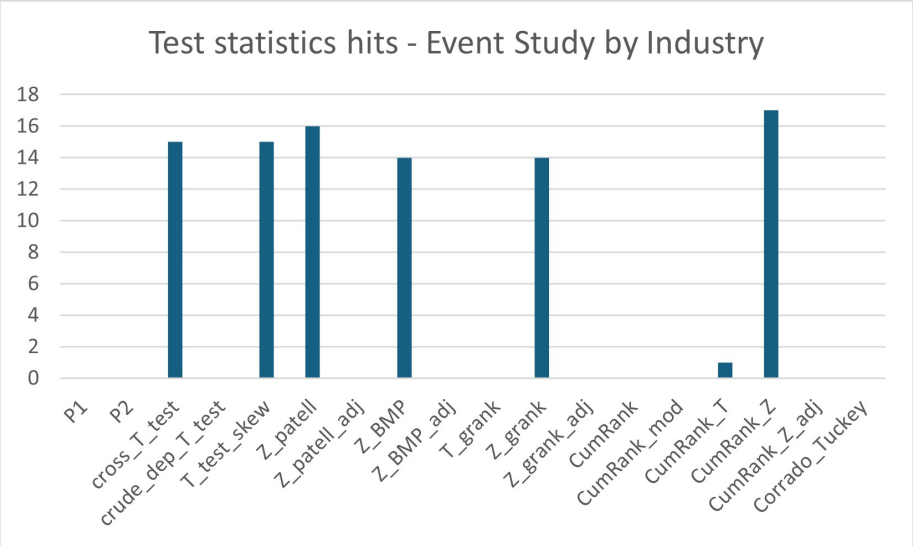


Figure 1.B.3: Number of significant test statistics for the analysis by ESG rating

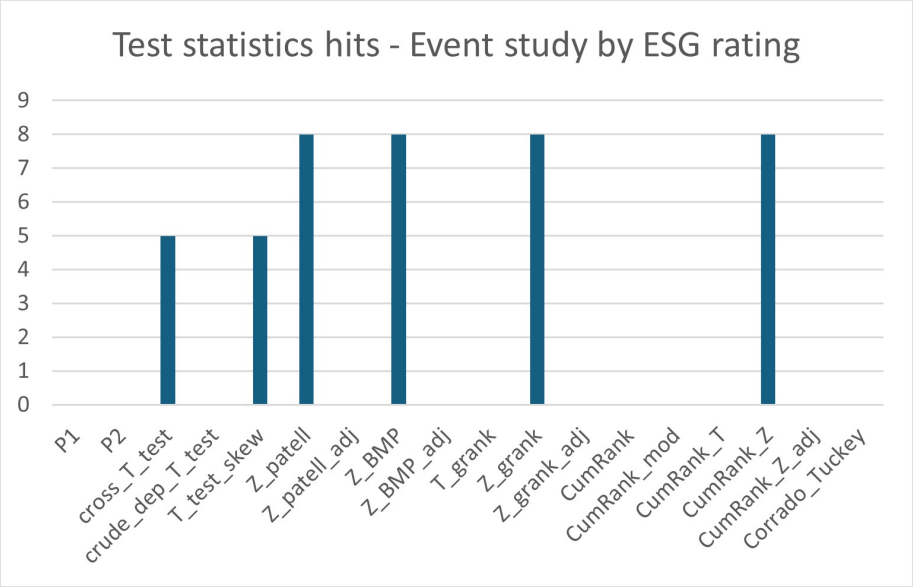


Figure 1.B.4: Number of significant test statistics for the analysis by country-sector

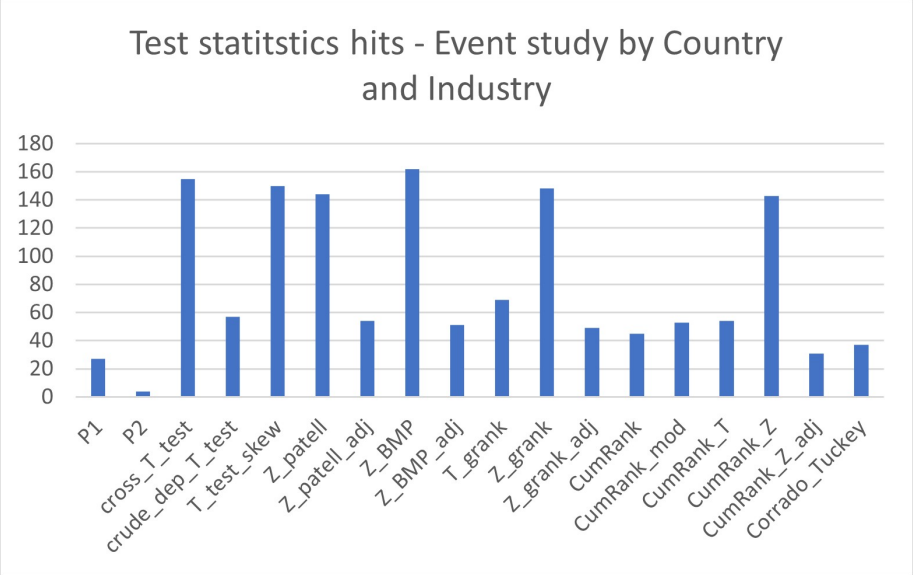


Figure 1.B.5: Number of significant test statistics for the analysis by country-ESG

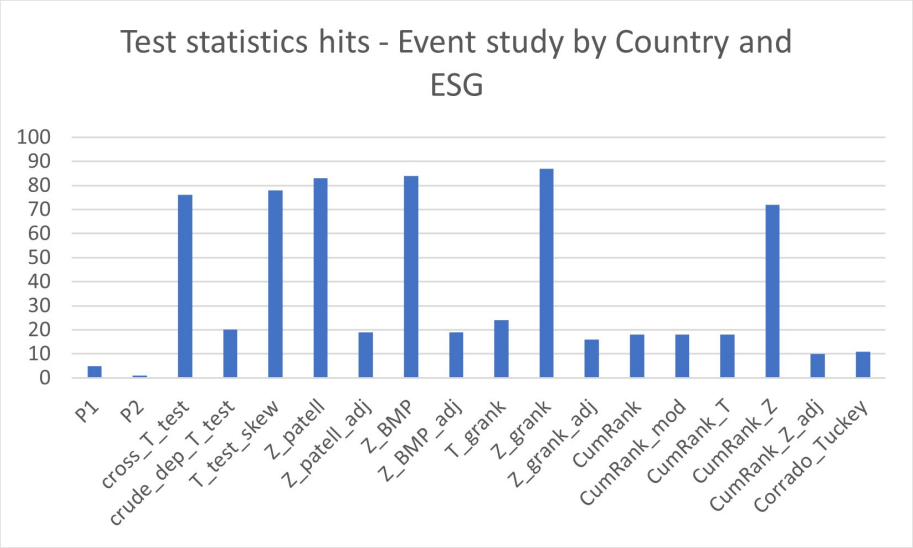
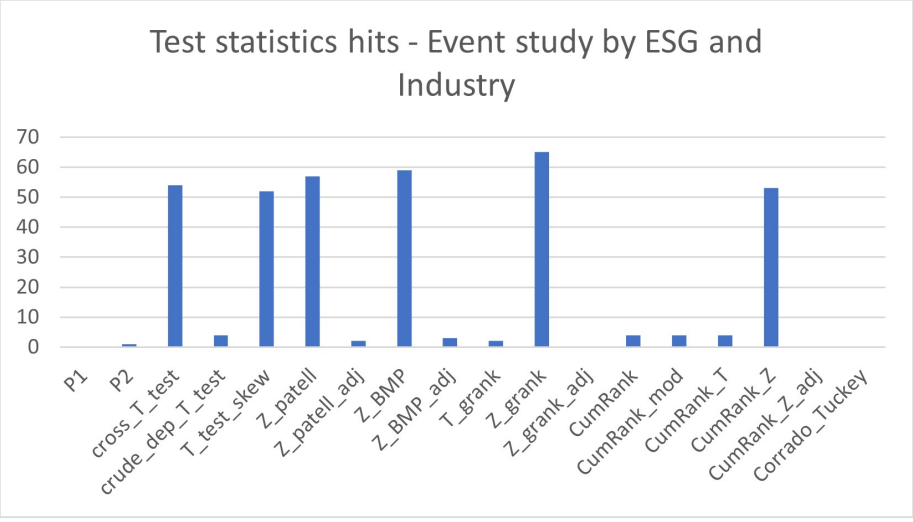


Figure 1.B.6: Number of significant test statistics for the analysis by ESG-sector



Chapter 2

Multidimensional spatiotemporal clustering of ESG scores

from

Multidimensional spatiotemporal clustering – An application to environmental sustainability scores in Europe

Coauthored with: Simone Boccaletti, Paolo Maranzano, Philipp Otto

Published in: *Environmetrics*

2.1 Introduction

As the world is facing a path toward a more sustainable, greener, and less carbon-intensive economy, Environmental, Social and Governance (ESG) practices are becoming more and more relevant from the company perspectives in mitigating sustainability-linked risks. Since the Paris Agreement of 2015, a great attention has been paid to corporate sustainability performance, especially on environmental aspects such as Greenhouse Gas (GHG) emission levels.

According to the World Health Organization (WHO), almost all of the global population breathes air that exceeds WHO guideline limits and contains high levels of pollutants. Moreover, air quality is closely linked to the earth's climate and ecosystems globally. Many of the drivers of air pollution (i.e. combustion of fossil fuels) are also sources of greenhouse gas emissions. Policies to reduce air pollution, therefore, offer a win-win strategy for both climate and health, lowering the burden of disease attributable to air pollution as well as contributing to the mitigation of climate change.

In this context, firms contribute significantly to the emission of polluting gases with respect to households. In Germany, the Environment Agency of the German Government UBA, 2024 shows that companies produced more than 85% of the CO_2 emissions in the last years. In the UK, the Department for Energy Security and Net Zero (DESNZ, 2023) affirms that in 2022, household emissions account only for 17% of the total. Moreover, the Italian National Institute of Statistics (Istat, 2022) shows that in the last 20 years, the overall volume of CO_2 has decreased by around 30%, but the emissions from companies still represent around 70% of the total.

Within this framework, Environmental, Social, and Governance issues have become crucial topics for companies' operations and stakeholders' engagement and activism. Companies have begun to put sustainability practices as a core aspect of their operations, and to disclose more and more information on their commitment to Environmental, Social, and Governance issues. Stakeholders have started to take ESG ratings and scores into consideration to make financial decisions, preferring to have relationships and interactions with companies that respect sustainability principles. ESG ratings and scores are synthetic evaluations from a specialized rater on a company's sustainability performance from many different points of view, broadly grouped under the three pillars E, S and

G. As a result, empirical evidence highlights that companies that achieve good ESG ratings are able to obtain better financial conditions, e.g., lower cost of capital and easier access to capital markets.

In this paper, we aim to investigate the spatial pattern and the time dynamics of sustainability scores assigned to firms in Western Europe. We first trace the spatial pattern of Western European firm evaluation on three different sustainability levels, i.e., the overall sustainability-ESG score, the Environmental score, and the Carbon Emission performance using a tailored version of the hierarchical spatial clustering algorithm by Chavent et al., 2018, which allows for detecting homogeneous groups of companies combining sustainability and spatial information. Furthermore, we examine the temporal dynamics of ESG-Environmental-Carbon emission evaluations in the last ten years by combining the spatial information and the similarity across the temporal series of multiple sustainability-related evaluations. The latter constitutes a multidimensional spatiotemporal extension of the spatial clustering methodology by Chavent et al., 2018.

Our findings prove that both space and time dimensions are relevant in ESG performance evaluations. The initial spatial analysis, which is carried out for ESG ratings in 2023, provided evidence of the presence of cross-national and cross-industrial groups of companies with remarkable differences in the levels of environmental performance. Specifically, clusters are differentiated according to ESG scores, and the analysis brings out a cluster of companies with very poor sustainability performance, which belong to several European countries and are mainly classified in the manufacturing and mining industries. Other clusters are less transnational, and composed mainly of companies engaged in the tertiary and service sectors. Regarding the spatiotemporal clustering, the identified groups are more prone to spatial overlapping, suggesting that the ESG scores' temporal aspect is relevant to our multidimensional approach.

The remainder of the paper is structured as follows. In Section 2.2, we briefly review the current literature on ESG evaluation and patterns in the ESG scores. Then, we describe the dataset and the data collection procedure in Section 2.3. In Section 2.4, we introduce two multidimensional hierarchical clustering algorithms, the first for the purely spatial framework and the second for the spatiotemporal case. Thereby, we focus on techniques to efficiently select the hyper-parameters (i.e., weighting parameters and the number of clusters) by proposing two tailored algorithms. In Section 2.5, we summarise the results of the cluster analyses and provide an interpretation and

discussion of the identified clusters. Eventually, in Section 2.6, we sum up the main contents of the paper and provide concluding remarks and future research perspectives.

2.2 Background

Evaluating the sustainable commitment of companies is a very complex task and the final result and its interpretation may depend on the metrics chosen and the methodology applied, as well as on the information availability. Over the years, several data providers have managed to improve methodologies, follow similar standards and increase the coverage of companies evaluated also by exploiting other sources of information. The academic literature in the economic-financial field provides increasing evidence about factors that impact firms' ESG performance and the beneficial effects that ESG commitment has on companies and stakeholders. Given its relevant role, researchers have recently begun analysing ESG patterns to provide the basis for more specific studies on this phenomenon.

2.2.1 The challenge of ESG assessment

Evaluating the sustainable commitment of companies and assigning scores regarding their impact on the different aspects that make up sustainability is a very complex task, which could be carried out following a multitude of approaches, and using different types of information systems. Considering 9 ESG rating agencies, Billio et al., 2022 offer a comprehensive description and comparison of their approaches and methodologies. Furthermore, they provide a table to offer a quick comparison of the different methodologies used by the various agencies for calculating ESG ratings. Appendix 2.B contains a summary of the table.

Following this direction, a recent branch of academic literature is in fact focusing on analysis of agreement and disagreement across different ESG rating providers. Berg et al., 2022; Christensen et al., 2022 provide a deeper understanding of the divergences the ESG rating composition, Gucciardi et al., 2024 examine the common factors in the Environmental Pillar score of different rating providers. According to their description and comparison in Billio et al., 2022, we decide to use

MSCI ESG rating in our analysis because it provides ESG rating for more than 10 years, it covers a large number of companies and the data provider has not been subject to mergers and acquisitions that may have substantially changed the methodologies underlying the data provided. Moreover, the MSCI evaluations are standardized for each industry using company-specific adjustments, according to the exposure and the challenges that each company has to face. MSCI collects available data from company disclosure, government agencies, non-governmental organizations (NGOs) and media sources monitored daily (MSCI, 2023a). As regard the carbon emission score, MSCI relies on the voluntary self-declaration provided by companies (MSCI, 2023b). While in the academic literature it is often discussed about the importance of providing corporate sustainability disclosure, especially with regard to carbon emission (Aldy et al., 2024; Bastos Neves & Semmler, 2022; Bolton et al., 2021, 2022), for the moment the institutions do not impose that it is mandatory for all enterprises, as it is necessary to define before the standardized guidelines. Although the number of companies providing such information is still very small, about a quarter according to Aldy et al., 2024, these are generally the largest in terms of size, and therefore represent the majority of economic activity in terms of total asset and sales. An accurate description of the methodology used by MSCI for the allocation of carbon emission score is given in Appendix 2.C, together with a table showing the percentage of companies that in 2022 provided the non-mandatory sustainability report and were then assigned an ESG rating.

2.2.2 The important role of ESG evaluation for firms and stakeholders

First of all, companies' commitment to sustainability principles seems to be closely linked to environmental policies. In fact Zhang, 2022 demonstrate that environmental regulation pushes companies to pay more attention to product quality and sustainability principles in production and Y. P. (Chen et al., 2022 find positive effects of environmental regulation on firm environmental investment. Also, Z. Chen et al., 2022; Qian and Yu, 2024; Xue et al., 2023 observe a positive effect of green finance policy on ESG performance and X. Wang et al., 2022 show that green finance policy encourages enterprises to develop and adopt green products and technologies. Other research focuses on the Environmental Protection Tax Law in China, finding a positive effect on the ESG performance and green technological innovation Li & Li, 2022, particularly for heavily polluting firms X. He

et al., 2023; Y. He et al., 2023. Moreover, Wu and Tham, 2023 show that executive green incentives and top management team characteristics positively impact the corporate ESG performance and Zhang et al., 2023 suggest that the disruption of environmental subsidies significantly positively affects them. Furthermore, the literature provides evidence of a positive impact of regional environmental transparency H. Chen et al., 2023, digital transformation Zhao & Cai, 2023 and digital finance Mo et al., 2023 on ESG performance.

Among the evidence on the effects of ESG performance, Fu and Li, 2023 found that it positively and significantly affects corporate financial performance, and digital transformation drives this promoting effect, Alfalih, 2023 show that social and governance dimensions of ESG influence companies' financial performance across the two measures of a firm's financial performance (ROA and Tobin's Q), while environmental dimension is significant with the Tobin's Q measure. Also, Yu and Xiao, 2022 find a significantly positive relationship between ESG composite performance and firm value and Panda and Ray, 2023 describe the positive effect of Corporate Sustainability expenditure on share prices. López-Cabarcos et al., 2023 demonstrate that the absence of CO2 equivalent emissions, the absence of incentives, and the presence of environmental investment have an impact on stock market returns. Moreover, the academic literature provides evidence of significant positive impact of ESG commitment on listed companies' stock liquidity (R. Chen et al., 2023), on productivity (Ma et al., 2022), on foreign investment flows (Chipalkatti et al., 2021), on green innovation (Lian et al., 2023; Mukhtar et al., 2023; Zheng et al., 2023) and a negative impact on company over-indebtedness (Lai & Zhang, 2022).

The above-mentioned research highlights the relevance of ESG performance and, consequently, the fundamental role that disclosure of ESG assessments plays.

2.2.3 Pattern of firms' ESG evaluation

The recent academic literature provides interesting examples of how researchers have traced and mapped different patterns of businesses of their ESG assessments throughout a cluster analysis or other methodologies for classification, sometimes considering specific aspects of sustainability or including variables related to specific features of the activities carried out by the companies them-

selves. In particular, Ronalter et al., 2023 consider a sample of firms from Europe, East Asia and North America to perform a hierarchical cluster analysis including ESG indicators, and they carry out independence tests to compare the quality management systems and environmental management systems of firms with different ESG evaluation. Gonzaga et al., 2024 employed the Kohonen Self-Organizing Map for clustering developing market companies, providing valuable evidence of the changes in ESG scores over the course of the COVID-19 pandemic. Using the same methodology, Iamandi et al., 2019 examine the sustainability profile of European companies, considering the ESG score, the scores of the pillars and the scores of the indicators composing them.

Moreover, Q. Wang, 2023 investigates the dynamics of three pillars of ESG scores among banks and observes a convergence of the evaluations in separate clusters in recent years, exploiting a specific panel data model proposed by Phillips and Sul, 2007, 2009 to represent the behaviour of economies in transition, formulated as a nonlinear time-varying factor model. Saraswati et al., 2024 use a sample of Indonesian firms and perform a K-means cluster analysis on ESG score pillars' score to clarify differences between ESG sustainability and practice and show the relationship among the distinct aspects of ESG performance. The authors identify three clusters using the Elbow method, the silhouette, and the Gap Statistical Method.

Focusing on the environmental aspect, Amores-Salvadó et al., 2023 consider a sample of public industrial firms from Europe, the United States and Canada, they classify them according to a four-position matrix based on the dichotomy environmental performance-disclosure, then they perform ANOVA tests showing differences in the groups according to nationality and sector. Ishizaka et al., 2021 propose a new hierarchical multi-criteria clustering based on PROMETHEE, they take into account uncertainty and imprecision making use of the Stochastic Multiobjective Acceptability Analysis (SMAA) and cluster ensemble methods, and they provide an interesting application on a sample US banks, considering financial variable and ESG pillar scores. Ortas et al., 2015 exploit a multidimensional HJ-Biplot technique finding evidence of how different country-specific social and institutional schemes influence ESG evaluation in a sample of firms located in Spain, France and Japan. Sariyer and Taşkın, 2022 consider a sample of companies listed in the (Borsa Istanbul) BIST sustainability index and, based on their ESG pillars scores, they perform a K-means++ algorithm which accounts for a smart centroid initialization method by assigning the first centroid randomly then selecting the rest of the centroids based on the maximum squared distance. Using

the silhouette score, the authors identify heterogeneous clusters in terms of ESG evaluation and also in terms of size and profitability.

To the best of our knowledge, only Q. Wang, 2023 explicitly account for the temporal component in the clustering algorithm. In the other research, in which data are observed in several years, the authors repeat the same analysis considering separately the different years and then interpret the evolution of clusters over time (Gonzaga et al., 2024; Ortas et al., 2015). As regards the spatial component, although some studies have identified clusters with different compositions according to the country of origin of the companies (Amores-Salvadó et al., 2023; Ortas et al., 2015; Ronalter et al., 2023), but including spatial information into the clustering algorithm is still an open question.

The primary innovation of our paper lies in the integration of spatial and temporal components to trace the patterns of ESG (Environmental, Social, and Governance) evaluations of companies. Specifically, we aim to incorporate spatial and temporal data into the clustering algorithm, ensuring that companies within the same cluster are not only geographically proximate but also exhibit similar trends over time. This approach is motivated by two main reasons. Firstly, existing literature has highlighted that sustainability commitments vary significantly based on geographical regions (e.g., see Gonzaga et al., 2024; Ortas et al., 2015; Ronalter et al., 2023) and over time (e.g., see Gonzaga et al., 2024; Q. Wang, 2023). For example, studies have demonstrated that different regions may have varying degrees of sustainability challenges and opportunities, influenced by local laws, resource availability, and cultural perspectives on sustainability. Companies situated in different areas encounter distinct environmental challenges due to these factors. By incorporating spatial and temporal data, we can better capture these regional variations and temporal trends, providing a more nuanced understanding of ESG evaluations. Secondly, our study serves as a preliminary step for future research aimed at comparing firms' environmental commitments with their actual environmental impacts. This comparison is crucial for evaluating the effectiveness of corporate sustainability initiatives. For instance, by considering air quality, soil and water pollutants, and waste production—factors often described by spatiotemporal models—we can assess whether companies' commitments to reducing CO2 emissions translate into tangible environmental benefits. This aspect of the study is particularly important, as ESG assessments typically rely on self-reported data from companies, which may not always reflect actual environmental performance. Moreover, the existing literature often focuses on the discrepancies between ratings from different ESG rating providers.

However, our research takes a step further by investigating whether companies' efforts to reduce CO2 emissions are yielding measurable results in terms of air pollution reduction. By analyzing the spatial and temporal patterns of companies' environmental impact assessments, we aim to identify specific regions, time periods, and sectors that require further scrutiny. This analysis will help pinpoint areas where companies' sustainability efforts are most effective, and where additional efforts may be needed. In summary, the inclusion of spatial and temporal components in ESG evaluation clustering not only enhances our understanding of regional and temporal variations in corporate sustainability but also lays the groundwork for future research that links corporate environmental commitments with actual environmental outcomes. Through this approach, we hope to contribute to the ongoing efforts to assess and improve the effectiveness of corporate sustainability initiatives.

2.3 Data collection and descriptive statistics

We use a unique dataset covering companies from 15 European countries. This dataset includes assessments of companies' sustainability performance and their geographical location.

2.3.1 MSCI ESG rating and Carbon Emission methodology

Companies' ESG performance is measured through ESG ratings. We collect this information on the companies' ESG rating and its components from MSCI ESG Ratings. ESG ratings are firm-level observations of their sustainability performance using different types of data, including, among others sustainability reports, media sources, and specific surveys to the clients. MSCI provides ESG ratings for more than 10k companies at the worldwide level, resulting in an overall evaluation system[1] with classes (ie. AAA, AA, A, BBB, BB, B, CCC in order from best to worst), a numerical overall score that ranges between 0 (worst) and 10 (best). The overall score is a weighted average of the scores on the three main pillars (Environmental, Social, and Governance). MSCI also provides information on key issues under the three pillars, e.g., relevant to our analysis, the carbon emission

[1]The assessments are not meant to be taken as absolute values, but need to be

scores MSCI, 2023a. The methodology that determines the aggregate pillar score from specific items, as well as the aggregation of the three pillars, E, S, and G, in the overall score and rating, are based on industry weights, reflecting the idea that valuation on key parameters is different according to the industry in which firms operate. As an example, the Environmental Pillar will weigh more on the final overall score for utilities than for firms in the Media & Entertainment industry. Lastly, ESG scores mainly cover listed firms, which are the ones that are most frequently under the attention of society, investors, and policymakers regarding sustainability issues.

Henceforth, ESG ratings are not just climate ratings. If a company's greenhouse gas emissions pose significant financial risks, its ESG rating will reflect that. For example, direct emissions pose a significant risk to power and steel companies, while emissions from their products after they have left the factory gate can pose a significant risk to automobile companies. However, for industries such as health care, the most financially relevant risks lie elsewhere, so emissions have less influence on a company's rating. In the Appendix 2.C, we report further details on the methodology used by MSCI for the computation of Carbon Emission scores and the number of companies providing the voluntary disclosure with the necessary information.

2.3.2 Sample description

Since this research focuses on the environmental aspects of sustainability scores, we include in our sample the overall ESG score, the Environmental Pillar score, and the Carbon Emission score, which are the three indicators in the set of sustainability scores that consistently draw attention from both scholars and professionals in the field. We consider observations from 2013 to 2023, and we select only companies with a weight of the Carbon Emission score greater than zero, that is, companies whose activity involves the emission of greenhouse gases. We match the ESG rating database with Orbis BvD database to link companies' ESG scores to their location. Initially, we collect the address of the Registered Office and NACE sector classification of listed companies located in Western European countries (Austria, Belgium, Denmark, France, Germany, Gibraltar, Ireland, Italy, Luxembourg, Malta, Netherlands, Portugal, Spain, Switzerland and United Kingdom). We include both active and inactive companies so that we do not exclude observations from companies that may have had an ESG rating in past years but recently have been the subject of mergers or

acquisitions or have left the market.

The focus on Western Europe is to avoid problems in the translation of companies' addresses from different alphabets since the algorithm could miss-locate firms in these cases. This henceforth excludes companies from Eastern Europe, Greece, and Scandinavian countries.

In Figure 2.3.1, we provide an overview of the sample size, in particular considering the number of firms by year and the number of firms by country in the last year. In our spatial clustering, we are using only data from 2023, thereby considering 617 companies. In the spatiotemporal cluster analysis, we use a sample of 460 companies, since we need companies with at least six observations in the time window considered, as described in the methodology section. We exclude observation from 2012 and earlier, because of the small number of available ratings.

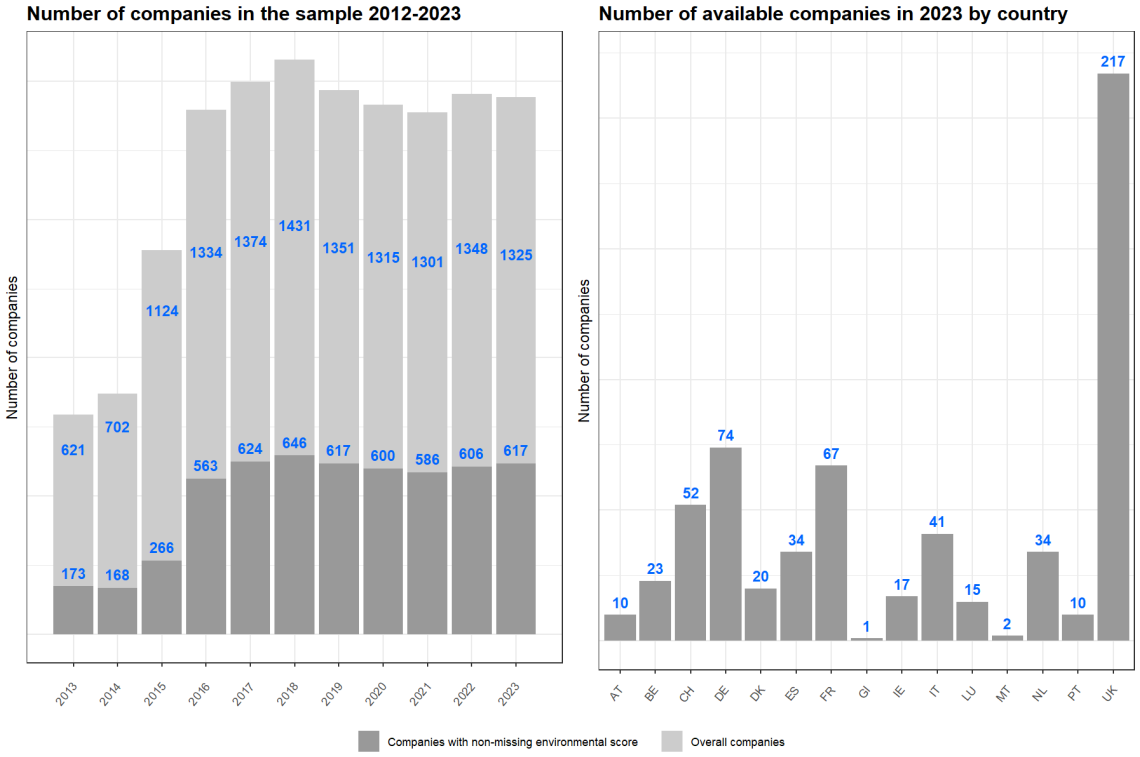


Figure 2.3.1: Left panel: Number of observations per year between 2013 and 2023. We report the total number of observations and the number of observations with a positive weight of the Carbon Emission score, which represents the observations included in our sample. Right panel: number of companies per country in 2023 with a positive Carbon Emission score weight

2.4 Methodology: hierarchical spatial and spatiotemporal clustering

The main goal of this paper is to investigate both the spatial and temporal patterns of the sustainability evaluations of Western Europe firms between 2013 and 2023. To do so, we carry out a spatial and spatiotemporal clustering analysis based on the methodology proposed by Chavent et al., 2018, which combines socio-economic features, temporal dynamics and geographical information. Specifically, we implement a modified version of their Ward-like hierarchical algorithm (Ward, 1963) that, in addition to detecting homogeneous groups under geographical constraints, selects the clustering hyperparameters such that the total proportion of explained inertia is maximized. The algorithm is then employed in a spatial framework combining cross-sectional socio-economic features and spatial information and in a spatiotemporal framework leveraging on the time dynamics of multiple sustainability distances and spatial dissimilarities.

2.4.1 Spatial hierarchical clustering

Let $D = [d_{ij}]_{i,j=1,\dots,n}$ be the dissimilarity matrix of the observations and let w_i be the weight of the i-th firm for $i = 1, \dots, n$. Without prior information, it is commonly set to $w_i = 1/n$. Alternatively, the Ward hierarchical clustering approach starts with an initial partition in n clusters of singletons, and at each step, the algorithm aggregates the two clusters such that the new partition has minimum within-cluster inertia, which measures the degree of heterogeneity within each cluster. We define $\mathcal{P}_K = (\mathcal{C}_1, \dots, \mathcal{C}_K)$ a partition of the dataset into K clusters and the pseudo-inertia of cluster $\mathcal{C}_K$ is computed as follows:

$$I(\mathcal{C}_K) = \sum_{i \in \mathcal{C}_K} \sum_{j \in \mathcal{C}_K} \frac{w_i w_j}{2 \sum_{i \in \mathcal{C}_K} w_i} {d_{ij}}^2. \tag{2.4.1}$$

The pseudo-within-cluster inertia of the partition is computed as the sum of the pseudo-inertia of each cluster. We point out that the pseudo-inertia is a generalization of the inertia when the dissimilarities can be non-Euclidean. From here on, we will always refer to pseudo-inertia, but for simplicity, we will call it inertia.

The spatial component is included by considering for the sample of n units two $n \times n$ dissimilarity matrices, namely $D_0 = [d_{0,ij}]_{i,j=1,\dots,n}$ and $D_1 = [d_{1,ij}]_{i,j=1,\dots,n}$, referring to Euclidean distances matrix of socio-economic variables under consideration and the geodetic distances matrix, respectively. Notice that, since the distances in the D_0 and D_1 matrices may belong to two very different measurement scales (e.g., socioeconomic distance in currency and physical distances in kilometres), it is necessary to scale the dissimilarity matrices with respect to their maximum values so that the distances across observations take values between 0 and 1.

In order to make the two dissimilarity matrices comparable before aggregation, we adopt a max-rescaling procedure defined as

$$\tilde{d}_{0,ij} = \frac{d_{0,ij}}{\max_{i,j} d_{0,ij}}, \qquad \tilde{d}_{1,ij} = \frac{d_{1,ij}}{\max_{i,j} d_{1,ij}}, \tag{2.4.2}$$

so that the entries of $\tilde{D}_0 = [\tilde{d}_{0,ij}]$ and $\tilde{D}_1 = [\tilde{d}_{1,ij}]$ lie in the interval $[0, 1]$. This rescaling step is crucial, since the aggregation of dissimilarities defined on different scales and/or measurement spaces may otherwise be dominated by the component with larger magnitude.

For a given mixing parameter $\alpha \in [0, 1]$, it is possible to obtain a convex combination of the of the pseudo inertia from dissimilarity matrices D_0 and D_1 and thus perform the hierarchical clustering algorithm.

Formally, the combined dissimilarity matrix is defined as

$$D_\alpha = (1 - \alpha)\tilde{D}_0 + \alpha\tilde{D}_1, \tag{2.4.3}$$

and Ward's hierarchical clustering algorithm is applied directly to the matrix D_α.

Notice that α states the importance of geographical and socio-economic information in determining the clusters. Indeed, as one set $\alpha = 0$, the geographical dissimilarities are not taken into account, while when $\alpha = 1$, the socio-economic distances are ignored and the clusters are defined according to geographical distances only.

Although the convex combination represents a simple and commonly adopted strategy for combining dissimilarity matrices, alternative approaches can also be considered, such as element-by-

element operations or the inclusion of spatial proximity penalty terms.

Given the partition $\mathcal{P}_K^\alpha = (\mathcal{C}_1^\alpha, \ldots, \mathcal{C}_K^\alpha)$, the mixed inertia for cluster $\mathcal{C}_K^\alpha$ is defined as the convex combination between the attribute inertia, and the inertia of the spatial component

$$I(\mathcal{C}_K^\alpha) = (1-\alpha) \sum_{i \in \mathcal{C}_K} \sum_{j \in \mathcal{C}_K} \frac{w_i w_j}{2 \sum_{i \in \mathcal{C}_K} w_i} \tilde{d}_{0,ij}^2 + \alpha \sum_{i \in \mathcal{C}_K} \sum_{j \in \mathcal{C}_K} \frac{w_i w_j}{2 \sum_{i \in \mathcal{C}_K} w_i} \tilde{d}_{1,ij}^2 \tag{2.4.4}$$

and the mixed within-clusters pseudo inertia is computed as the sum of the mixed pseudo inertia of its clusters, i.e.,

$$W_\alpha(\mathcal{P}_K^\alpha) = \sum_{k=1}^{K} I_\alpha(\mathcal{C}_k^\alpha). \tag{2.4.5}$$

Recall that the smaller the pseudo-inertia within the cluster, the more homogeneous the partition into K clusters. Therefore, in the spirit of Ward's criterion, at each iteration of the aggregation, the obtained cluster partition is the one that minimises $W_\alpha(\mathcal{P}_K^\alpha)$.

2.4.2 Spatial hierarchical clustering: choice of the parameters

The main issue in such a hierarchical clustering approach is the choice of the parameters α and K. Chavent et al., 2018 suggest setting a prior value for K and then providing a criterion to choose α such that it allows to explain the same proportion of the dissimilarities from both matrices, with respect to the cases in which the clusters are obtained considering only the feature matrix or the spatial matrix. They introduce the notion of the proportion of the total pseudo inertia explained by partition $\mathcal{P}_K^\alpha$ in K clusters as:

$$Q_\beta(\mathcal{P}_K^\alpha) = 1 - \frac{W_\beta(\mathcal{P}_K^\alpha)}{W_\beta(\mathcal{P}_1)} \tag{2.4.6}$$

where β can be either D_0 or D_1, depending on which dissimilarity matrix is used as a benchmark. Specifically, $Q_{D_0}(\mathcal{P}_K^\alpha)$ quantifies the proportion of socio-economic pseudo inertia (i.e., $W_{D_0}(\mathcal{P}_1)$) explained by partition $\mathcal{P}_K^\alpha$, while $Q_{D_1}(\mathcal{P}_K^\alpha)$ quantifies the amount of geographical pseudo inertia (i.e., $W_{D_1}(\mathcal{P}_1)$) explained by partition $\mathcal{P}_K^\alpha$.

To account for potential scale issues in $Q_{D_0}(\mathcal{P}_K^\alpha)$ and $Q_{D_1}(\mathcal{P}_K^\alpha)$, the $Q_\beta(\mathcal{P}_K^\alpha)$ metrics are then

normalized with respect to the baseline case of purely-geographical or purely-socio-economic clustering, that is, by computing the following ratios:

$$\tilde{Q}_{D_0}(\mathcal{P}_K^\alpha) = \frac{Q_{D_0}(\mathcal{P}_K^\alpha)}{Q_{D_0}(\mathcal{P}_K^0)} \qquad \tilde{Q}_{D_1}(\mathcal{P}_K^\alpha) = \frac{Q_{D_1}(\mathcal{P}_K^\alpha)}{Q_{D_1}(\mathcal{P}_K^1)}. \tag{2.4.7}$$

This relative formulation allows for a straightforward interpretation of the values. For instance, by considering $\tilde{Q}_{D_0}(\mathcal{P}_K^\alpha)$, for a given K and a given mixing parameter α, one is expressing the percentage improvement in the explained proportion of pseudo inertia obtained by using a mixture of geographical and socio-economic feature to generate the partition $\mathcal{P}_K^\alpha$ (i.e., $Q_{D_0}(\mathcal{P}_K^\alpha)$) with respect to the proportion of pseudo inertia it would be explained by only using socio-economic feature to generate the partition $\mathcal{P}_K^\alpha$ in K clusters (i.e., $Q_{D_0}(\mathcal{P}_K^0)$). Conversely, if one considers $\tilde{Q}_{D_1}(\mathcal{P}_K^\alpha)$, the resulting value for a specific pair of K and α expresses the improvement obtained by mixing the two dimensions instead of using a purely-geographical partitioning algorithm. Being α a measure of the trade-off between the loss of socio-economic homogeneity and the gain of geographic homogeneity, for a fixed K, increasingly values of α will correspond to higher $\tilde{Q}_{D_1}(\mathcal{P}_K^\alpha)$ and lower $\tilde{Q}_{D_0}(\mathcal{P}_K^\alpha)$. For a technical discussion about the properties of these quantities, we refer the readers to Section 3 in Chavent et al., 2018.

Chavent et al., 2018 suggest to choose α such that the normalized proportion of the explained pseudo inertia from D_0 and D_1 are as similar as possible, that is,

$$min_\alpha |\tilde{Q}_{D_0}(\mathcal{P}_K^\alpha) - \tilde{Q}_{D_1}(\mathcal{P}_K^\alpha)| \tag{2.4.8}$$

which means to identify α such that socio-economic and geographical information return as similar as possible proportion of explained pseudo inertia with respect to proportion it would be explained by only using socio-economic feature or spatial features to generate the partition. Consequently, the number of clusters can be chosen according to the dendrogram or elbow criteria. Following a similar rationale, Mattera and Franses, 2023 set an initial number of clusters K_0 considering the partition associated with D_0, then they determine α as in Chavent et al., 2018, and finally they define the optimal number of clusters based on the combined dissimilarity matrix. Notice that this selection method does not always allow to identify α such that it captures the highest possible overall dissimilarity in the data. To address such drawback, Jaya et al., 2019 start finding

α according to Mattera and Franses, 2023 while choosing a different mixing parameter in order to explain better the normalized proportion of inertia in one matrix, with a relatively small reduction of the normalized proportion of inertia from the other matrix.

Hereafter, we propose an algorithm to select the clustering hyperparameters, that is, the mixing coefficient α and the number of clusters K, that generalizes the aforementioned approaches by optimizing the *weighted average of the explained mixed pseudo inertia*, which can be expressed in several ways:

$$\begin{aligned}\bar{Q}(\mathcal{P}_K^\alpha) &= \frac{Q_{D_0}(\mathcal{P}_K^\alpha)\cdot W_{D_0}(\mathcal{P}_1) + Q_{D_1}(\mathcal{P}_K^\alpha)\cdot W_{D_1}(\mathcal{P}_1)}{W_{D_0}(\mathcal{P}_1) + W_{D_1}(\mathcal{P}_1)} \\ &= \left[1 - \frac{W_{D_0}(\mathcal{P}_K^\alpha) + W_{D_1}(\mathcal{P}_K^\alpha)}{W_{D_0}(\mathcal{P}_1) + W_{D_1}(\mathcal{P}_1)}\right].\end{aligned} \tag{2.4.9}$$

In particular, conditioning on a given K, the optimal α is given by the maximizer of $\bar{Q}(\mathcal{P}_K^\alpha)$, that is,

$$max_\alpha \bar{Q}(\mathcal{P}_K^\alpha). \tag{2.4.10}$$

Thus, while Chavent et al., 2018 defined the optimal α as the one balancing the explained inertia from socio-economic and geographical features, we are proposing to select the α, which jointly maximizes the amount of pseudo inertia explained from both the socio-economic and the geographical information (i.e., $Q_{D_0}(\mathcal{P}_K^\alpha)$ and $Q_{D_1}(\mathcal{P}_K^\alpha)$), weighted by the cumulated spatial and socio-economic pseudo inertia embedded the data (i.e., $W_{D_0}(\mathcal{P}_1) + W_{D_1}(\mathcal{P}_1)$). Further details about $\bar{Q}(\mathcal{P}_K^\alpha)$, in particular, its relationship with the other metrics presented above, are provided in Appendix 2.A.

The proposed algorithm works as follows. Having fixed a given value of K, the hierarchical clustering is performed for a sequence of α values on a regular grid from 0 to 1, with a specified constant increment $\Delta\alpha$. For any K, the best α_K^* is chosen such that the total proportion of explained inertia is maximum. The computation is iteratively repeated considering a range of K up to a defined maximum, that is, $K = 1, 2, \ldots, K_{max}$. In this way, we obtain an optimal weighting α_K^*, conditional on K, which we then optimize across a range from $K = 1, ...K_{max}$. Then, the optimal number of clusters K^* (evaluated at the optimal α_K^*) is determined according to one or more suitable criteria for hierarchical clustering (Kaufman & Rousseeuw, 1990). For instance, consider computing the increments in the weighted average proportion of explained pseudo

inertia (which can be interpreted as the increase of the explained variability) induced by a unitary increase in the number of groups or computing the values of the Silhouette index (synthesizing the average homogeneity of units within each cluster). According to the former, the optimal number of groups will match the largest value of K guaranteeing a relevant increment in the weighted average explained inertia, while according to the latter, higher Silhouette values indicate that, on average, the units are properly matched within their own cluster. The proposed algorithm is summarized in Algorithm 1.

Algorithm 1 Hierarchical Spatial Clustering: grid for choice of α and K

Define as $D_0 = [d_{0,ij}]_{i,j=1,\dots,n}$ the feature dissimilarity matrix
Define as $D_1 = [d_{1,ij}]_{i,j=1,\dots,n}$ the spatial dissimilarity matrix
Define as K_{max} the maximum number of clusters
Define as $\Delta\alpha$ the increment of α
for $K = 1, \dots, K_{max}$ **do**
  for $\alpha \in [0,1]$, by $\Delta\alpha$ **do**
    Compute the linear combination of the two dissimilarity matrices $D(\alpha) = (1-\alpha)D_0 + \alpha D_1$;
    Compute the $\mathcal{P}_K^\alpha$ = partition in K clusters according to Ward hierarchical algorithm on the combined matrix D;
    Compute the weighted average of the explained mixed pseudo inertia $\bar{Q}(\mathcal{P}_K^\alpha)$
  end for
  Select the best α for each K such that $\alpha_K^* = argmax_\alpha \bar{Q}(\mathcal{P}_K^\alpha)$
end for
Choose K^* (evaluated at the corresponding α_K^*) according to one or more hierarchical clustering criteria, such as the first difference in the weighted average proportion of explained pseudo inertia or the Silhouette index.

2.4.3 Spatiotemporal hierarchical clustering

As regards the spatio-temporal clustering, several authors proposed to adapt the methodology proposed by Chavent et al., 2018 to the case of georeferenced time series data by combining the dissimilarity matrix computed on the n time series and the spatial dissimilarity component.

To the best of our knowledge, even considering different fields of application, the literature considered time series related to only one socio-economic variable (Bucci et al., 2023; Deb & Karmakar, 2023; Mattera & Franses, 2023). We aim to extend this framework by combining together multiple dissimilarities matrices corresponding to several time series of socio-economic features in addition to the spatial distances. Specifically, we combine the four dissimilarity matrices referring to the

time series of the overall ESG score, the Environmental Pillar score and Carbon emission score, and the spatial component.

The distances among time series are computed adopting the Dynamic Time Warping (DTW) distance algorithm implemented in the function `diss.DTWARP()` from `TSclust` package in `R` Montero and Vilar, 2014. The Dynamic Programming approach using a warping function has been introduced by Sakoe and Chiba, 1971, 1978 in spoken word recognition field as a time-normalization algorithm. Berndt and Clifford, 1994 have implemented DTW distance to detect patterns in time series. DTW is a distance-minimizing temporal alignment between two time series that allows us to compute a dissimilarity measure among time series that could have different lengths and/or missing observations during the period, but having at least one overlapping time stamp. Let us consider two time series, namely $x_t(t = t_{x_1}, \ldots, T_x)$ and $y_t(t = t_{y_1}, \ldots, T_y)$, such that $t_{x_1} \lesseqgtr t_{y_1}$ and $T_x \lesseqgtr T_y$ but $T_x \geq t_{y_1}$ and $T_y \geq t_{x_1}$. Let us compute the distance between two points $d(x_s, y_r) = |x_s - y_r|$ as the DTW distance between x and y up to points s and r, which is given by the optimal alignment minimizing the following the distance:

$$\Delta(s, r) = d(x_s, y_r) + \min[\Delta(s-1, r-1), \Delta(s-1, r), \Delta(s, r-1)].$$

As shown by Berndt and Clifford, 1994, the usage of DTW distance is equivalent to minimizing the Euclidean distance between aligned time series, thus in one dimension, under all admissible temporal alignments. Therefore, time series with similar shapes will be considered similar, even if the deformation appears in different time stamps. This appears to be a proper choice in our analysis for two reasons. First of all, we know that the ESG performance of companies have varied over time, according to different events and factors, and certainly we can not assume that these follow a linear trend over time (we recall that DTW allows us to account for similar non-linear dynamics of the time series). Secondly, the time variability of ESG scores and subscores is very low; thus, we do not expect to observe time series with many fluctuations as there may be companies that have improved in a first period and then remain stable later or move toward the opposite direction; other companies could show a U shaped or U-inverted shape in their sustainability assessments. In any case, we do not impose that these potential changes in the ESG score happened for every company in the same moment, but simply that the shape of the dynamics is similar for companies belonging

to the same group.

Once the dissimilarity matrices $D_p = [d_{p,ij}]_{p=1,\dots,P;i,j=1,\dots,n}$ have been computed, where $P-1$ is the number of variables included and D_P refers to the spatial dissimilarity matrix obtained computing the geodetic distances across the observations, we can proceed in finding the parameters K and α_p for the linear combination $D(\alpha_p) = \sum_{p=1}^{P} \alpha_p D_p$. Note that $\alpha_1, \dots, \alpha_{p-1}$ are the coefficients for the temporal dissimilarity matrices, while $\alpha_P = 1 - \sum_{p=1}^{P-1} \alpha_p$ represents the weight for the spatial component. We recall that dissimilarity matrices are normalized with respect to their maximum value.

We adapt the criterion proposed in Algorithm 1 for the spatial clustering by choosing the vector α_p maximizing the weighted average of the explained mixed pseudo inertia induced by partition $\mathcal{P}_K^{\alpha_p}$.

Similarly to Algorithm 1, for a fixed number of groups K, we consider all the possible combinations for a grid of $\alpha_p (p = 1, \dots, P)$ with a constant increase of each α_p equal to $\Delta\alpha$ and such that $\sum_{p=1}^{P} \alpha_p = 1$. Then, we identify the clustering partition according to the Ward hierarchical algorithm on the combined distance matrix D. In particular, we adapt the criterion proposed in Algorithm 1 for the spatial clustering by choosing the vector α_p maximizing the weighted average of the explained mixed pseudo inertia induced by partition $\mathcal{P}_K^{\alpha_p}$, given by the following generalization of the Equation 2.4.10:

$$\bar{Q}(\mathcal{P}_K^{\alpha_p}) = 1 - \frac{\sum_{p=1}^{P} W_{D_p}(\mathcal{P}_K^{\alpha_p})}{\sum_{p=1}^{P} W_{D_p}(\mathcal{P}_1^{\alpha_p})}. \tag{2.4.11}$$

At this point, conditioning on K, we select the values of α_p for which $\bar{Q}(\mathcal{P}_K^{\alpha_p})$ is maximum. We iterate this step for a defined range of potential candidates $K = 1, 2, \dots, K_{max}$ evaluated at the corresponding optimal $\alpha^*_{K,p}$. Finally, we select the number of clusters according to the same criteria defined for the spatial clustering, that is, the increments in the weighted average proportion of explained inertia and the Silhouette index. The proposed algorithm is summarized in Algorithm 2.

Algorithm 2 Hierarchical Spatiotemporal Clustering: grid search of $\alpha_p (p = 1, \dots, P)$ and K

Define as $D_p = [d_{p,ij}]_{p=1,\dots,P;i,j=1,\dots,n}$ the feature dissimilarity matrices
Define as K_{max} the maximum number of clusters
Define as $\Delta\alpha$ the increment of α_p
for $K = 1, 2, \dots, K_{max}$ **do**
  for $\alpha_p (p = 1, \dots, P) \in [0, 1]$ by $\Delta\alpha$ such that $\sum_{p=1}^{P} \alpha_p = 1$ **do**
    Compute the linear combination of the dissimilarity matrices $D(\alpha_p) = \alpha_1 D_1 + \alpha_2 D_2 + \cdots + \alpha_P D_P$
    Compute the partition $\mathcal{P}_K^{\alpha_p}$ in K clusters according to Ward hierarchical algorithm on the combined matrix D
    Compute the weighted average of the explained mixed pseudo inertia $\bar{Q}(\mathcal{P}_K^{\alpha_p})$ for each partition
  end for
  Select the best α_{Kp} for each K such that $\alpha_{Kp}^* = argmax_{\alpha_p (p=1,\dots,P)} \bar{Q}(\mathcal{P}_K^{\alpha_p})$
end for
Choose K^* (evaluated at the corresponding α_{Kp}^*) according to one or more hierarchical clustering criteria, such as the first difference in the weighted average proportion of explained pseudo inertia or the Silhouette index.

2.5 Empirical findings

In this section, we discuss the empirical evidence offered by the spatial and spatiotemporal clustering algorithms, motivating the choice of parameters and interpreting the results.

2.5.1 Spatial Clustering

For the Spatial clustering, we considered the sample of 617 European firms with available ESG, Environmental and Carbon Emission scores for 2023. We recall that we obtain the dissimilarity matrices D_0 and D_1 from the Euclidean distances of the ESG-related variables and the geodetic distances of the coordinates of the firms, respectively.

Choice of K^* and α_K^*

We carry out the spatial cluster analysis by setting $\Delta\alpha = 0.1$ and $K_{max} = 20$. We compute the optimal hyperparameters α^* and K^* considering both the Chavent et al., 2018 procedure and the

proposed Algorithm 1. As shown in Figure 2.5.1, our algorithm leads to a higher proportion of explained pseudo inertia in the spatial component and, as a consequence, the proportion of pseudo explained inertia is lower for the socio-economic features component. Overall, the weighted average proportion is higher than the proportion reached with the Chavent methodology, meaning that we are capturing better the overall variability embedded in the data.

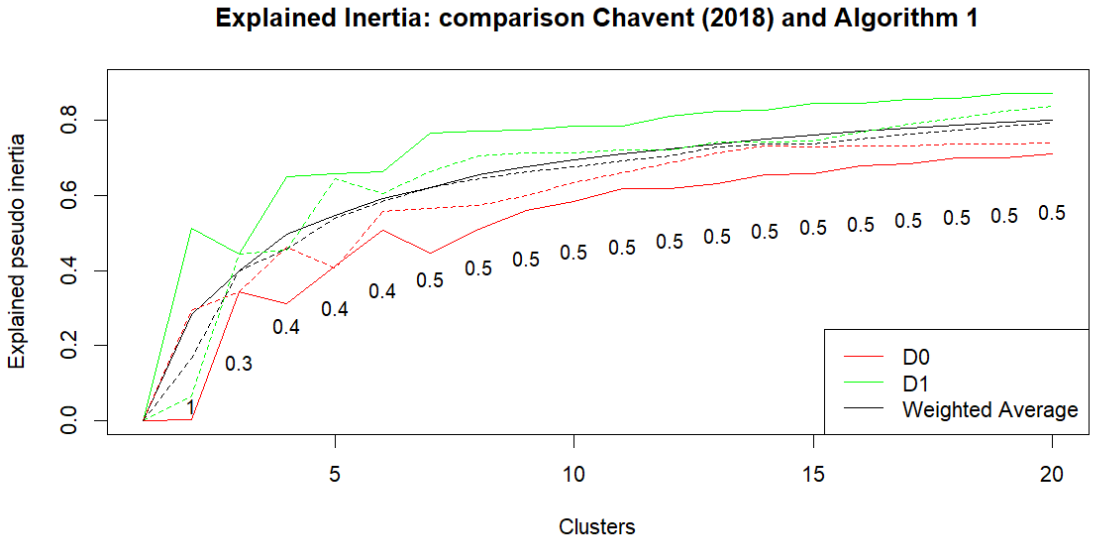
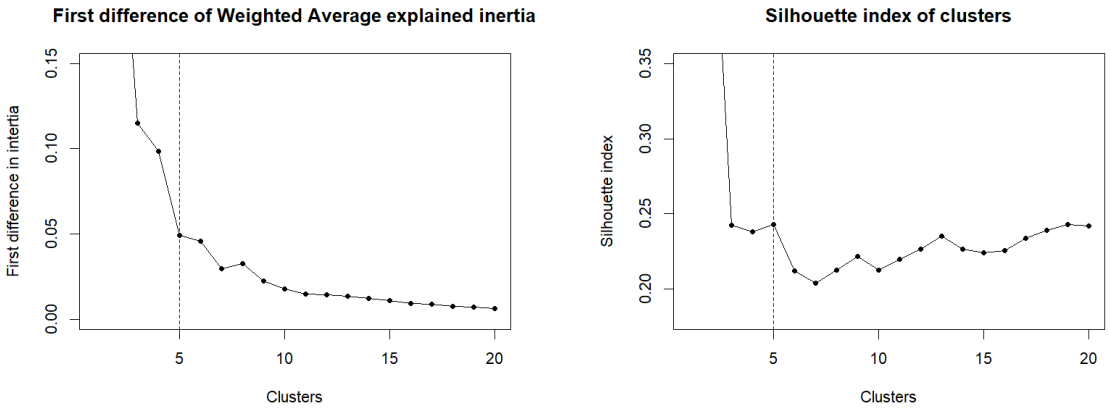


Figure 2.5.1: Comparison of proportion of explained pseudo inertia from each dissimilarity matrices D_0 and D_1 and their weighted average, using Chavent et al., 2018 method (dashed line), and Algorithm 1 (solid line) for $K = 1, \ldots, K_{max} = 20$. The values underlying the curves indicate the optimal α_K^* for Algorithm 1.

As regards the choice of the parameters, we consider the increment in the weighted average proportion of the explained inertia associated with a unitary increase in the number of clusters, and the Silhouette index, where the dissimilarity matrices are linearly combined with respect to α_K^*. The values of α_K^* are reported in Figure 2.5.1, while the aforementioned indices are described in Figure 2.5.2. We select $K^* = 5$ because it seems a good compromise, able to gain a relevant increase in the weighted average explained inertia and a reasonable value of the Silhouette index. Consequently, we have $\alpha_{K=5}^* = 0.40$.

For a more exhaustive comparison between the results obtained by the two methods, in Table



A. Gain in weighted average explained inertia

B. Silhouette index

Figure 2.5.2: Hyperparameters selection in spatial clustering. Left panel: increment in the weighted average proportion of explained inertia generated by a unitary increase in the number of clusters. Right panel: Silhouette index computed for each partition into K clusters. Recall that, for each K, we considered the best combination of the dissimilarity matrices according to the α_K^* found using the maximization inertia criterion according to Figure 2.5.1.

2.5.1 are reported the inertia, the proportion of explained inertia and the normalized proportion of the explained inertia by the two dissimilarity matrices, considering $K^* = 5$ number of clusters. With respect to Chavent et al., 2018 approach, the proportion of explained inertia increases from 0.4070 to 0.4120 in the dissimilarity matrix D_0, and the gain in D_1 is from 0.6453 to 0.6567. Consequently, the normalized proportion of explained inertia obtained with α_5^* is higher for both matrices.

Resulting spatial Clusters

Figure 2.5.3 maps the clusters produced by the spatial clustering for 2023 with parameters $K^* = 5$ and $\alpha_5^* = 0.40$. Each point represents a company location, with different colours depending on the cluster to which the company is assigned. Also, Figure 2.5.5 summarizes the distribution of firms by country (left panel) and by industry (right panel) to understand the obtained categorization better.

In terms of number of companies and country representativeness, cluster 1 is definitely the largest group, with 273 companies located in 10 out of 15 countries in the sample, while the other

		Chavent $\alpha_{K=5} = 0.50$			Algorithm 1 $\alpha^*_{K=5} = 0.40$		
	$W(\mathcal{P}_1)$	$W(\mathcal{P}_K^\alpha)$	$\mathcal{Q}(\mathcal{P}_K^\alpha)$	$\tilde{\mathcal{Q}}(\mathcal{P}_K^\alpha)$	$W(\mathcal{P}_K^\alpha)$	$\mathcal{Q}(\mathcal{P}_K^\alpha)$	$\tilde{\mathcal{Q}}(\mathcal{P}_K^\alpha)$
D_0	0.0473	0.0275	0.4070	0.5709	0.0271	0.4120	0.5780
D_1	0.0581	0.0221	0.6453	0.7802	0.0206	0.6567	0.7941

Table 2.5.1: Summary of the inertia (absolute, relative and normalized) returned by the spatial clustering. Columns 3 to 5 report results according to Chavent et al., 2018, that is, $K^* = 5$ and $\alpha^* = 0.50$. Columns 6 to 8 report results according to Algorithm 1, that is, $K^* = 5$ and $\alpha^* = 0.40$. $W(\mathcal{P}_1)$ is the total inertia from spatial dissimilarity (i.e., $W_{D_1}(\mathcal{P}_1)$) or socio-economic dissimilarity (i.e., $W_{D_0}(\mathcal{P}_1)$); $W(\mathcal{P}_K^\alpha)$ is the absolute within-cluster pseudo inertia from the mixed clustering; $\mathcal{Q}(\mathcal{P}_K^\alpha)$ is the proportion of explained pseudo-inertia; $\tilde{\mathcal{Q}}(\mathcal{P}_K^\alpha)$ is the normalized proportion of inertia.

clusters have fewer companies. Clusters 2 and 3 collect firms from 7 and 8 countries, respectively, while cluster 4 is composed only of companies from the Iberian Peninsula, and cluster 5 is mainly composed of UK and Irish companies (and one company from northern France). Moreover, cluster 3 and cluster 4 are the ones with the fewest companies, respectively 43 and 45. While cluster 4 includes all companies from Spain and Portugal (and one more from Gibraltar), the fact that cluster 3 is made up of companies from 8 countries suggests that there may be some common ESG factors that ultimately group these companies together.

The above clustering shows some degree of overlap in space, except for Cluster 4, which consists only of companies located in Spain and Portugal. In this case, the geographical distance from the Iberian peninsula to the rest of Europe seems to prevail over the distance calculated on the ESG scores, and therefore the resulting cluster seems to absorb only the spatial information, ignoring the differences between variables. In order to better examine the role of the spatial component on cluster formation, we performed several robustness checks involving the non-spatial and the spatial settings with alternative spatial constraints. In particular, for data referring to 2023 we ran the purely feature-based clustering (i.e., without spatial components) and considering a Nearest Neighbour Distance Matrix to embed the spatial information. As expected, when ignoring the spatial component, the clusters appear to be more overlapped compared to the case of spatial clustering and the Iberian companies mixed-up with other European companies. In the case of spatial clustering based on the neighbourhood matrix, although the Spanish and Portuguese companies do not form a perfectly compact cluster, it is still evident that most of them have very high ESG ratings

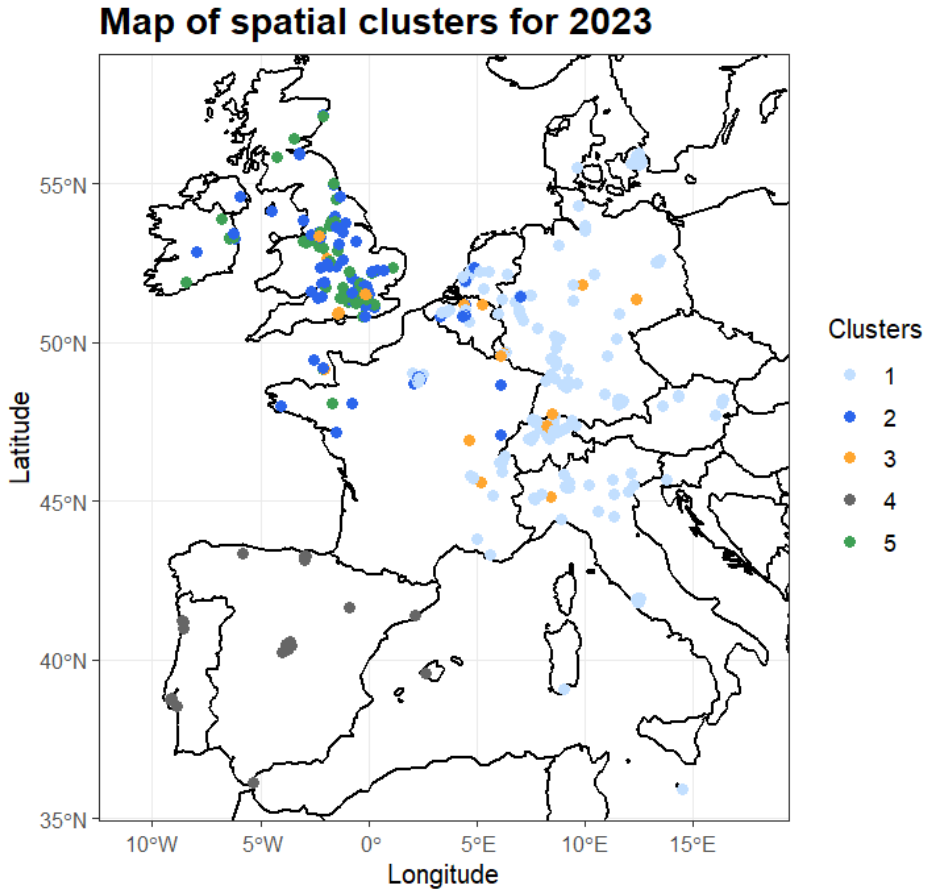


Figure 2.5.3: Map of spatial clusters for 2023 using $K^* = 5$ and $\alpha_5^* = 0.40$. Clusters are computed using two dissimilarity matrices: Geodetic spatial distance and Euclidean distance of ESG, Environmental and Carbon Emission scores.

and are homogeneous with each other. A full discussion on the two robustness checks and the corresponding results is reported in Appendix 2.D.

Concluding, in terms of geographical overlap, cluster 1 and seems to overlap with cluster 2 mainly in Belgium, the Netherlands, and France and to a minimal extent with other clusters. Instead, clusters 2, 3, and 5 show a high degree of overlapping in the UK and Ireland, while cluster 4 does not seem to overlap with other clusters.

For the environmental assessment of each cluster, Figure 2.5.4 shows the descriptive statistics

(mean, 25th and 75th percentile range) of the clusters in terms of Carbon Emission score (ce), Environmental Pillar score (env), and ESG score (esg).

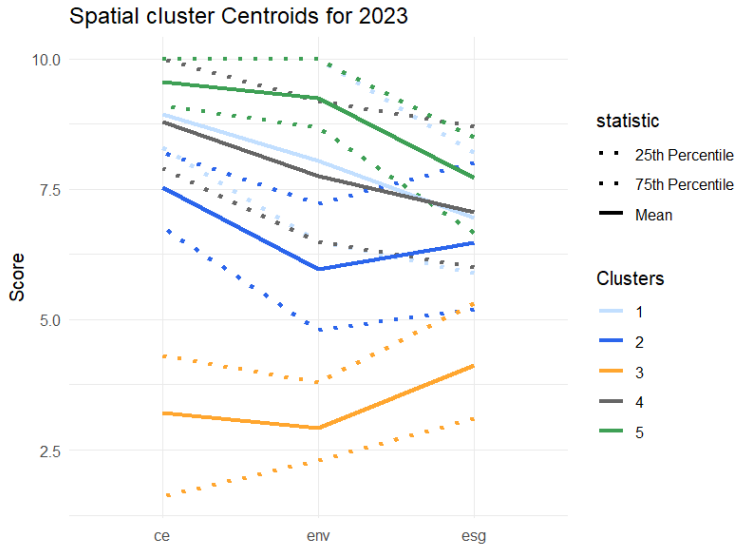


Figure 2.5.4: Cluster-specific centroids of the three economic variables used to define the spatial clusters for 2023 using $K^* = 5$ and $\alpha_5^* = 0.40$. Recall that clusters are computed using two dissimilarity matrices: Geodetic spatial distance and Euclidean distance of ESG (esg), Environmental (env) and Carbon Emission (ce) scores. The solid lines represent the means of the variables; dashed lines represent the two quartiles.

Clusters 1, 4, and 5 have similar patterns of sustainability scores; in particular, they record above-average Carbon Emission scores and Environmental scores. As their ESG score is, on average, lower than their Environmental Pillar score, the companies in this cluster have better environmental practices with respect to the Social and Governance Pillars. Cluster 5 seems to be the one with slightly larger mean scores than the other two. Cluster 2 shows a different pattern from the previous ones. Companies in this cluster seem to have higher carbon emissions than the environmental pillar scores, with their overall ESG score almost at the same level as Clusters 1 and 4. This indicates a disparity of ESG evaluation across clusters: companies in cluster 2 may lag in the overall environmental performance but exhibit relatively strong social and governance attributes, resulting in an ESG score comparable to other clusters, on average. Finally, cluster 3 seems to collect the companies with the lowest performance in terms of both carbon emissions score, environmental pillar score, and ESG score. This finding is relevant given the geographical distribution of clusters

2 and 3. Despite their considerable overlap, our algorithm is able to differentiate companies based on their environmental performance relative to the other two pillars.

Although ESG scores are industry-based, examining the distribution of sectors in the different clusters we obtain from the model is particularly relevant. Considering the distribution across industries (based on the NACE industrial classification), Figure 2.5.5 shows a remarkable cross-sectorial heterogeneity, that is, there are no industries that are fully concentrated in a few clusters. However, the concentration of industries can vary considerably between clusters. For example, while in clusters 1, 2, and 3 the manufacturing industry (NACE C) is the most relevant sector (in both absolute and relative number of firms), in cluster 4 there are more enterprises in electricity, gas, steam, and air conditioning supply (NACE D), while in cluster 5 manufacturing enterprises are outnumbered by companies in information and communication (NACE J), financial and insurance activities (NACE K) and professional, scientific and technical activities (NACE M). It is worth noting that cluster 3, which is composed of enterprises with poor environmental sustainability performance, includes enterprises from ten NACE sectors, although six of them are represented by only one enterprise.

It is well known that some sectors involve the emission of large quantities of CO_2, while other activities are less exposed to this risk. Considering the heterogeneity across cluster-sectors groups, it is interesting to observe for each cluster the Carbon Emission Exposure score, Carbon Emission Management score, which measure respectively how much each firm's activity is linked to carbon emission and the effort of each firm toward this risk, without considering any sector adjustment. Moreover, we can use consider the Environmental Pillar weight and the Carbon Emission weight to understand the importance of this component within the overall ESG assessment. In Table 2.5.2 we show the average of the ESG, Environmental Pillar, Environmental weight, Carbon Emission, Carbon Emission weight, Carbon Emission exposure, Carbon Emission management score for each spatial cluster. Additional details regarding the computation of the scores and the definition of each index are made available in Appendix 2.C.

We can note that generally, companies with worst performance in terms of carbon emissions are also those that should pay more attention to this factor, as more exposed to this risk, in particular cluster 3 has the lowest Carbon Emission score, and management score, with the highest exposure

Cluster	ESG	Env score	Env weight	CE score	CE weight	CE exp	CE manag
All	6.797	7.419	21.24	8.324	9.269	3.521	5.579
1	6.949	8.047	19.08	8.93	8.707	3.043	5.768
2	6.476	5.961	25.71	7.536	10.07	4.307	5.113
3	4.116	2.921	39.49	3.212	14.86	7.765	3.974
4	7.067	7.76	29.13	8.787	11.31	3.400	6.56
5	7.719	9.239	11.08	9.565	6.759	2.172	5.909

Table 2.5.2: Average of the ESG (*ESG*), Environmental Pillar (*Env score*), Environmental weight (*Env weight*), Carbon Emission score (*CE score*), Carbon Emission weight (*CE weight*), Carbon Emission exposure (*CE exp*), and Carbon Emission management score (*CE manag*) for each spatial cluster.

score. In line with the high exposure, the relevance of the Carbon Emission component and the Environmental Pillar are particularly high in the composition of ESG score, which is also the lowest. Similarly, cluster 2 shows an higher exposure to Carbon Emission and lower performance than the average with the Environmental and Carbon weights slightly higher than the average. Clusters 1, 4 and 5 have an average lower exposure to carbon emission risk and higher management score. So the companies that should commit themselves more in the reduction of emissions because their activity involves the emission of large quantities of pollutants, are those that are most difficult to achieve good performance.

2.5.2 Spatiotemporal Clustering

As regards spatiotemporal clustering application, by starting from the initial sample of 1045 companies with at least one rating between 2013 and 2023, we considered the subsample of 460 firms with at least six annual observations within the same window. This choice is necessary to use the DTW distance for time series, which requires, for each pair of companies, at least one overlapping time stamp. Thus, having considered a 11-years period, the minimum number of overlapping years must is set to six, that is, for every selected company we require a non-missing value for more than half of the entire period under consideration. We computed the dissimilarity matrices of the time series for ESG score, Environmental Pillar score and Carbon emission score, using the DTW algorithm described in Section 2.4.3, and the geodetic distances from the coordinates. Thus, we

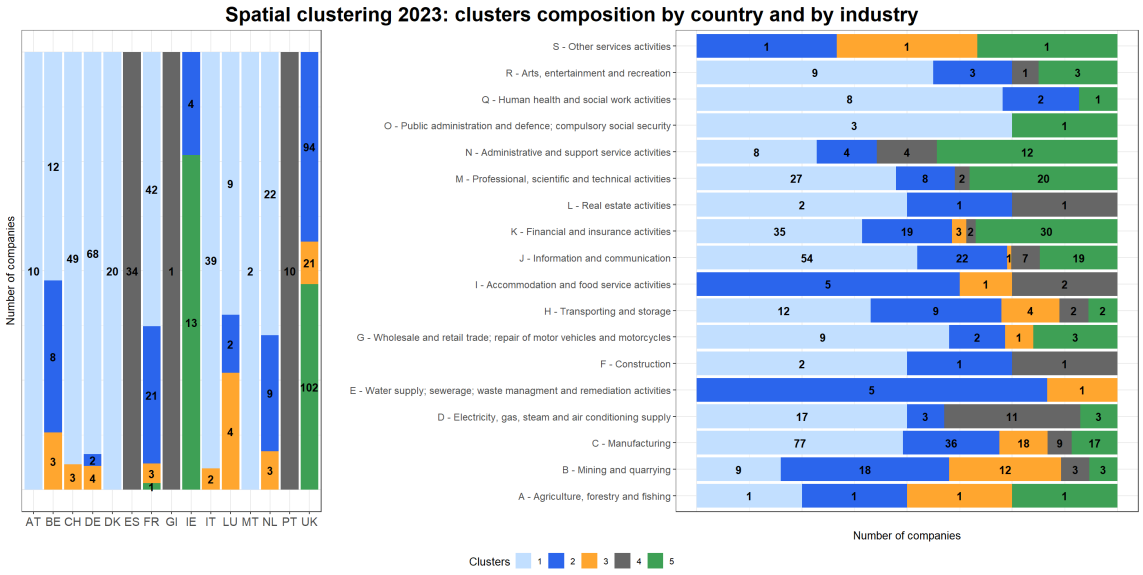


Figure 2.5.5: Spatial clusters composition by country (specified using ISO code) and industry for 2023.

obtained the dissimilarity matrices named D_{esg}, D_{env}, D_{ce} and D_{sp}.

Choice of K^* and α^*_{Kp}

To choose the optimal number of clusters K and the linear combination of weights α_p, we run the Algorithm 2 setting $K_{max} = 20$ and $\Delta\alpha = 0.05$. Having to choose a higher number of mixing parameters, we think it is more appropriate to use a smaller $\Delta\alpha$, so as to consider a greater number of combinations, including the case when the matrices all have the same weight (i.e., $\alpha_p = 0.25 \quad \forall p = 1, \ldots, P$). Notice that the latter combination would not be considered if we used a $\Delta\alpha = 0.1$ as in the purely spatial setting. In Figure 2.5.6, we show the results obtained for each number of clusters, both in terms of the proportion of explained inertia for each dissimilarity matrix and the weighted average proportion of explained inertia (top panel) and the best combination of mixing parameters.

From the plots, it is possible to notice that the spatial component is included when considering at least $K = 3$ groups. The dissimilarity matrix D_{esg} plays a less important role for a number of

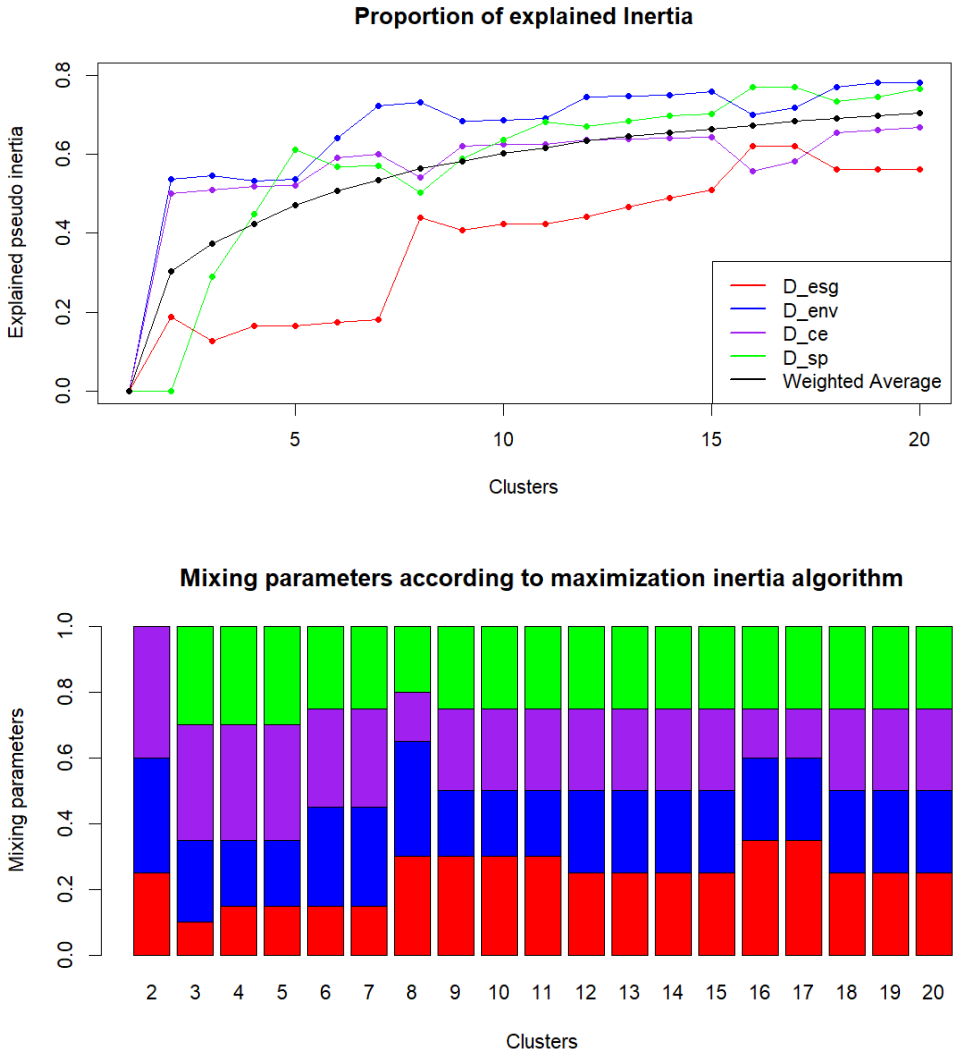
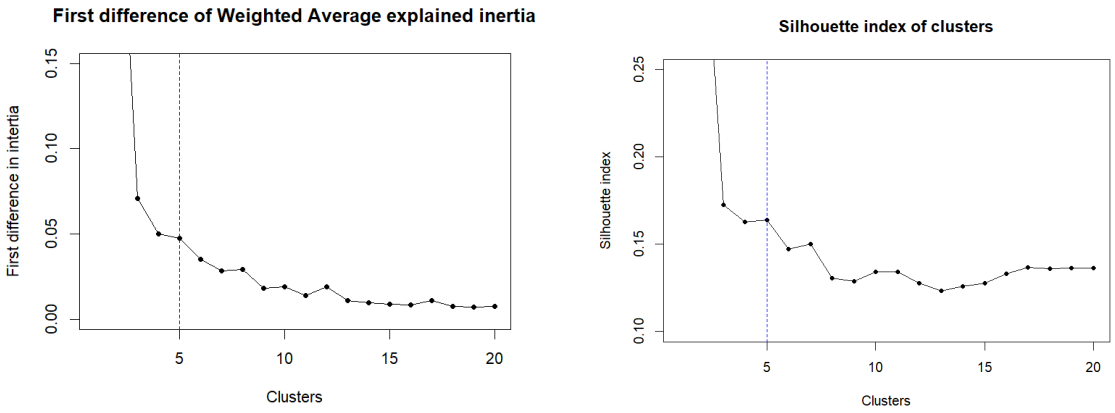


Figure 2.5.6: Top panel: proportion of explained pseudo inertia contained in each dissimilarity matrix (coloured lines) and weighted average proportion (black line). Values are computed considering for each K the optimal weighting combination $\alpha^*_{K,p}$ from Algorithm 2. Bottom panel: optimal weighting combination $\alpha^*_{K,p}$ from Algorithm 2. Colours refer to the four dissimilarity matrices used in the computation.

clusters lower than 8. For a higher number of clusters, the best combinations of the dissimilarity matrices provide similar weights. In Figure 2.5.7, we compare the gain in the weighted average of explained inertia and the silhouette index from $K = 1$ up to $K_{max} = 20$, evaluating each K at the corresponding optimal combination $\alpha^*_{K,p}$. Both plots suggest that, when choosing a number of clusters equal to $K = 5$, we manage to increase the weighted average proportion of inertia by 0.0478, and we get a silhouette index equal to 0.1640, which represents one of the highest values among those shown. Although the value of the silhouette may not seem too high, with the increasing dimensionality of the data, it becomes difficult to achieve high values because of the curse of dimensionality as the distances become more similar.



A. Gain in the weighted average of explained inertia

B. Silhouette index

Figure 2.5.7: Hyperparameters selection in spatiotemporal clustering. Left panel: increment in the weighted average proportion of explained inertia generated by a unitary increase in the number of clusters. Right panel: Silhouette index computed for each partition into K clusters. Recall that, for each K, we considered the best combination of the dissimilarity matrices according to the α^*_{Kp} found using the maximization inertia criterion according to Figure 2.5.6.

Overall, we can state that using $K^* = 5$ clusters with weights $\alpha^*_{5,esg} = 0.15$, $\alpha^*_{5,env} = 0.20$, $\alpha^*_{5,ce} = 0.35$ and $\alpha^*_{5,sp} = 0.30$ represent the best solution. In Table 2.5.3, we report the absolute explained inertia, the proportion of explained inertia and the normalized proportion of explained inertia for each dissimilarity matrix, considering our choice of hyperparameters. Overall, through cluster analysis, we are able to explain more than 50% of inertia in the dissimilarity matrices D_{env}, D_{ce}, D_{sp}, but the proportion of explained inertia in D_{esg} is lower than 25%.

	α_p	$W(\mathcal{P}_1)$	$W(\mathcal{P}_K^{\alpha_p})$	$\mathcal{Q}(\mathcal{P}_K^{\alpha_p})$	$\tilde{\mathcal{Q}}(\mathcal{P}_K^{\alpha_p})$
D_{esg}	0.15	0.0501	0.0418	0.1649	0.2195
D_{env}	0.20	0.0672	0.0311	0.5362	0.6411
D_{ce}	0.35	0.0456	0.0218	0.5206	0.6761
D_{sp}	0.30	0.0626	0.0242	0.6125	0.7453

Table 2.5.3: Summary of the inertia (absolute, relative and normalized) returned by the spatiotemporal clustering at the optimal solution $K^* = 5$. $W(\mathcal{P}_1)$ is the total inertia provided by each dissimilarity matrix; $W(\mathcal{P}_K^{\alpha_p})$ is the absolute within-cluster pseudo inertia from the mixed clustering for each matrix; $\mathcal{Q}(\mathcal{P}_K^{\alpha_p})$ is the proportion of explained pseudo-inertia; $\tilde{\mathcal{Q}}(\mathcal{P}_K^{\alpha_p})$ is the normalized proportion of inertia.

In order to achieve a substantially higher proportion of explained inertia in D_esg, it is necessary to choose a much higher number of clusters, but this would complicate the interpretation of the resulting clusters. In addition, the silhouette index assumes even lower values when the number of clusters is greater than 5, so the homogeneity of the resulting clusters seems to decrease. Thus, $K^* = 5$ is a reasonable choice in our spatiotemporal cluster analysis.

Resulting Spatiotemporal Clusters

The spatiotemporal clustering produces substantially different groups with respect to the spatial clustering results described in Section 2.5.1. Figure 2.5.8 represents the companies based on their location and on the final cluster to which they belong.

It is possible to notice the higher degree of overlap between the five clusters with respect to the spatial clustering analysis. Even in the Iberian Peninsula, we observe the presence of two clusters, namely 3 and 4, whereas in the previous analysis, only one cluster was present. This implies that the specific temporal dynamics of ESG scores could be relevant in our multidimensional approach. In other words, the inclusion of the temporal component captures new information that was missed when only the spatial component was considered. In particular, different pairs (or sometimes even triplets) of clusters could be identified in different geographical areas. Clusters 1 and 2 overlap in Italy, Denmark and Germany, and also with some enterprises from cluster 3, in Switzerland;

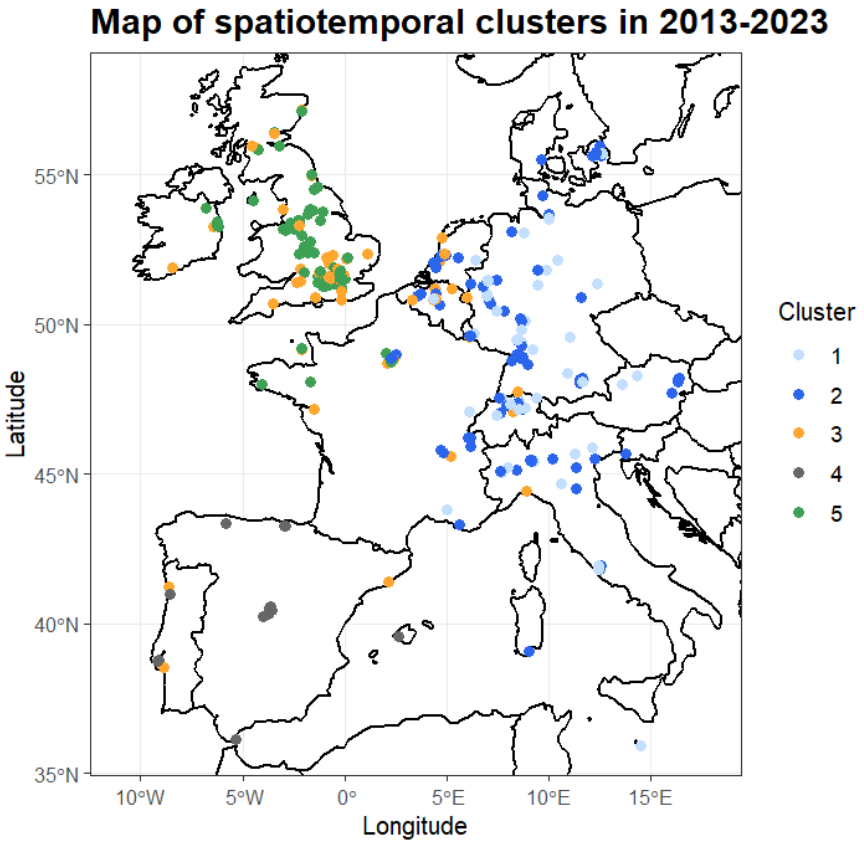


Figure 2.5.8: Map of spatiotemporal clusters between 2013 and 2023. Clusters are computed using four dissimilarity matrices: Geodetic spatial distance, DTW distance of the overall ESG scores, DTW distance of the Environmental scores and DTW distance of Carbon Emission scores.

clusters 2 and 3 overlap mainly in Belgium and the Netherlands, and they are also found together with cluster 5 in Paris; clusters 3 and 5 overlap significantly in the UK and Ireland; in Portugal there is an overlap between clusters 3 and 4, especially around the cities of Lisbon and Porto. For the environmental and sustainability assessment, we represent the mean, the 25th and the 75th percentile of the Carbon Emission Score, Environmental Pillar score and ESG score in the period observed for the different clusters.

A synthesis of the results by country and by industry is reported in Figure 2.5.9. Similarly to the previous spatial analysis, the distribution of sectors across clusters is quite heterogeneous. Two results stand out. Once again, the manufacturing sector (NACE C) is not dominant in all

clusters, but only in clusters 1, 2 and 3. In cluster 4, there are more enterprises in electricity, gas, steam and air conditioning supply (NACE D), information and communication (NACE J) and administrative and support service activities (NACE N) than in manufacturing. In cluster 5, on the other hand, manufacturing is outnumbered by information and communication (NACE J) and finance and insurance (NACE K). Finally, cluster 3, the worst-performing cluster in terms of environmental sustainability, includes a large number of manufacturing enterprises and most of the enterprises in mining and quarrying (NACE B) and in transport and storage (NACE H).

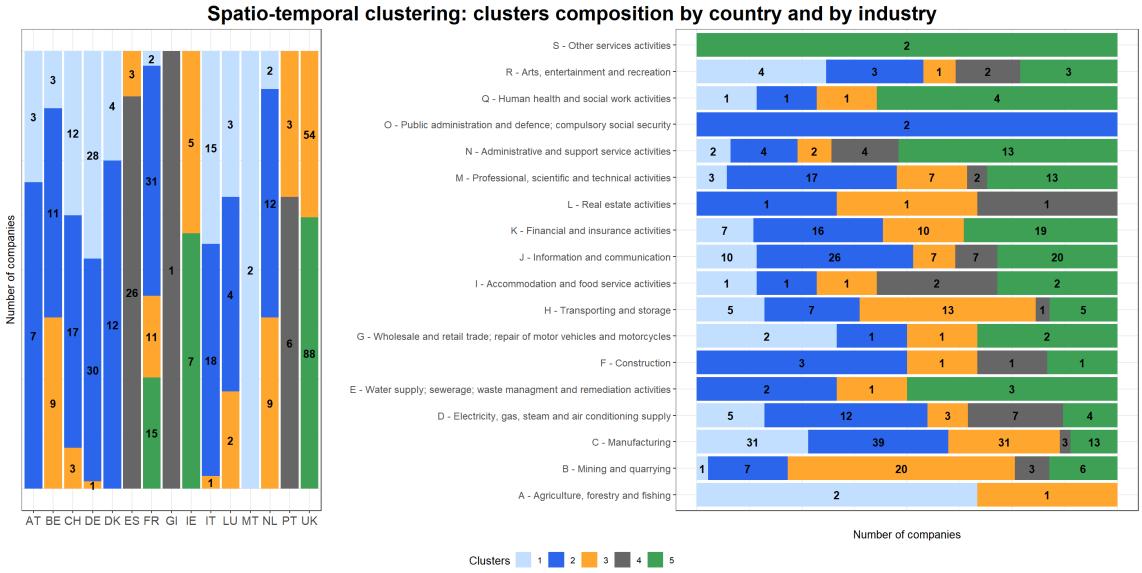


Figure 2.5.9: Spatiotemporal clusters (from 2013 to 2023) composition by country (specified using ISO code) and industry.

As in the spatial clustering analysis, some clear paths can be identified in the individual clusters, as well as some common aspects between the clusters.

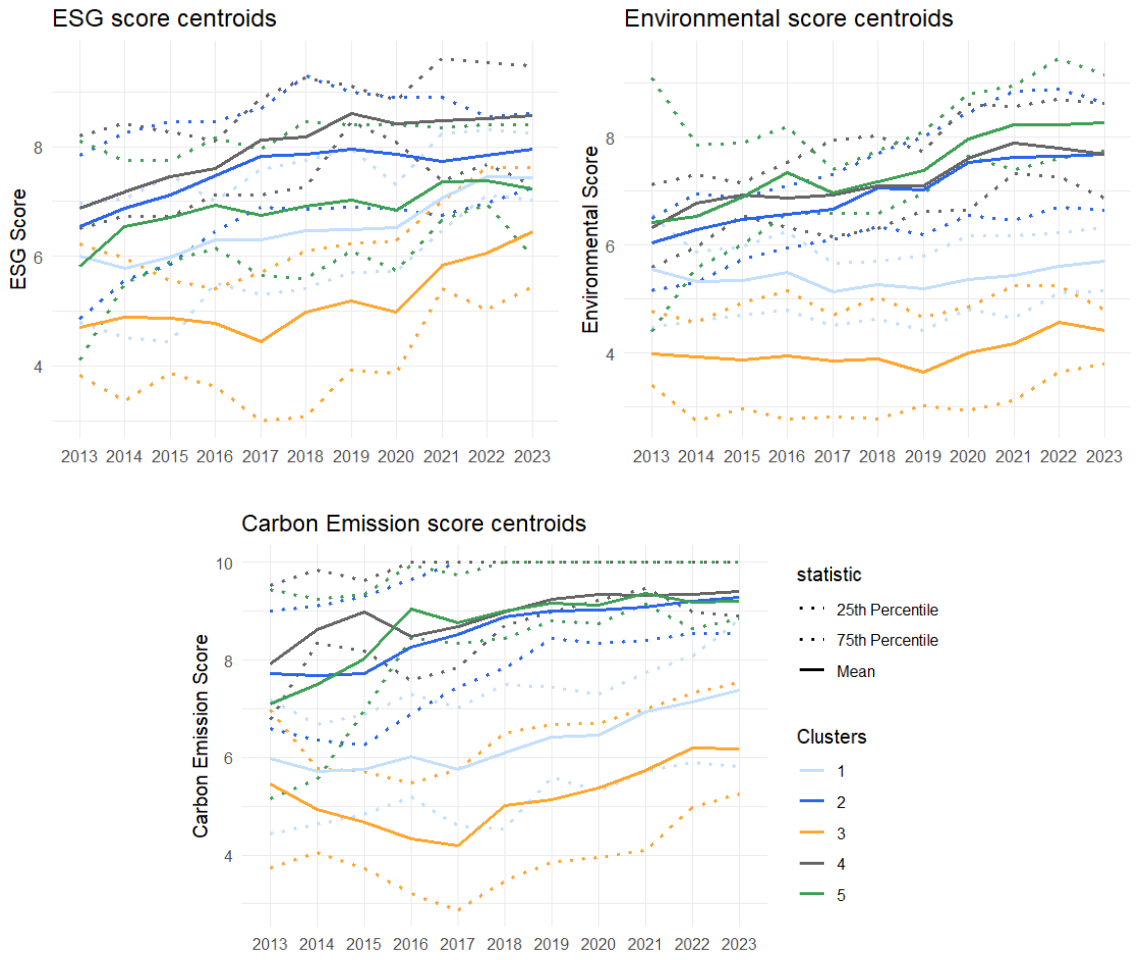


Figure 2.5.10: Time series 2013-2023 of centroids (solid lines) and quantiles (dashed lines) of the socio-economic used for the spatiotemporal clusters.

As reported in Figure 2.5.10, clusters 2, 4 and 5 seem to contain the companies with the highest performance in all three variables considered. Their scores also seem to increase over time, albeit with different magnitudes. On the contrary, with respect to the previous clusters, clusters 1 and 3 collect companies with visibly lower levels of environmental performance and carbon emission scores, with cluster 3 being the lowest-scoring cluster on average. While the pattern for the ESG score for the two clusters appears to be increasing throughout the period, as for the other three clusters, the pattern for the environmental score appears to be constant throughout the period.

Instead, the two clusters observe a very different path for the carbon emission score. For cluster 3, this variable decreases on average between 2013 and 2017, while after 2017, it seems to increase until the end of the period. For cluster 2, on the other hand, the carbon emission score seems to be constant in the period 2013-2017 and then starts to increase until 2023.

2.6 Conclusion

In this paper, we used spatial and spatiotemporal clustering methodologies to identify the main patterns that characterize the environmental performance of European companies. In particular, we examined the role of spatial location and temporal dynamics for European companies with ESG scores in the period between 2013 and 2023. For this reason, we developed a spatial and a spatiotemporal version of a hierarchical clustering technique accounting for multiple dimensions and providing alternative criteria to efficiently determine suitable values for the clustering hyperparameters, that is, the weighting combination between the considered variables and the optimal number of clusters.

Our findings suggest that both space and time matter when analyzing patterns of ESG performance. The spatial analysis for 2023 provided evidence of the presence of cross-national and cross-industrial groups of companies with remarkable differences in the levels of environmental performance. In particular, we were able to detect a subsample of companies with very poor environmental and ESG scores, which belong to several European countries and are mainly classified in the manufacturing and mining industries. Other clusters are mainly driven by companies engaged in the tertiary and service sectors, with a more regional and less transnational character (e.g., the clusters in the UK and in the Iberian peninsula), with higher levels of sustainability performance. Regarding the space-time dynamic, the identified groups are more prone to spatial overlapping (e.g., two clusters in the Iberian peninsula), suggesting that the ESG scores' temporal aspect is relevant in our multidimensional approach. As for the purely spatial analysis, also in the spatiotemporal setting, the dualism between the manufacturing sector and tertiary sectors is a key element, with the manufacturers dominating the worst environmental-performing clusters.

Finally, another outcome of interest is the gap (in some groups, considerable gaps exist) between

overall ESG score values and the scores in the environmental and carbon emissions pillars. In this sense, both approaches show a distance between environmental scores with respect to the social and governance pillars. In other words, companies with better environmental and emissions scores perform worse on average in the social and governance categories, lowering the overall ESG score. In contrast, companies with low environmental scores perform better in terms of social and governance. This difference could be due to divergence in corporate strategies, leading to specialization in internal company practices as well, and thus diverse scores.

We believe our results can help identify how local regulations, cultural factors, and economic conditions influence ESG practices and provide insights into the effectiveness of environmental policies for companies and the impact of past regulations on ESG scores. Henceforth, the results of our study provide valuable insights for companies, practitioners, investors, and stakeholders. From a managerial perspective, by the means of the cluster analysis, firms can benchmark against peers to identify best practices and areas for improvement, aiding in strategic planning and setting realistic ESG targets. Additionally, these insights enhance stakeholder communication, promoting transparency and accountability.

Moreover, the results of our study provide valuable information for professionals and investors from two perspectives. On the one hand, identifying a cluster with poor environmental performance can help assess potential risks and vulnerabilities for investments. Trough the cluster analysis, investors could assess whether firms in different clusters are exposed to different physical and transitional environmental risks. Investors can therefore use this information to reach regions or companies with better ESG and carbon performance and assess risks and vulnerabilities by making more informed decisions. On the other hand, companies that do not belong to the dominant cluster in their region are classified into other clusters based on their ESG performance. In other words, our proposed algorithm can isolate those companies for their exceptionally strong or weak sustainability performance compared to their regional counterparts. This is valuable to investors because it enables them to identify outliers in terms of ESG performance within specific regions. By isolating companies with notably high or low ESG performances, investors can make more informed decisions regarding sustainable investments in specific areas, potentially minimizing risk by avoiding companies with poor ESG records. This is also important for policy makers, especially regional authorities, who are trying to identify good and bad sustainability practices. With this

regard, our proposed algorithm could be used by policymakers to design more effective and targeted actions with the aim of reducing corporate emissions and, thus, the overall environmental impact.

Lastly, in the academic literature, this research contributes by providing new insights into the spatial pattern and temporal dynamics of Environmental performance of companies. The findings can help in developing new theoretical models and frameworks for understanding Environmental performance aiming to examine the relationship with other phenomena or changes due to significant events (e.g., policy changes, environmental incidents, economic crises), encouraging interdisciplinary research, combining environmental science, geography, economics, and data science.

In summary, the main take-home message of this paper is that the geographical location of firms is relevant for a comprehensive understanding of the time dynamics of ESG scores, particularly in explaining the ability of firms to achieve positive scores, especially in the environmental aspects of sustainability. Conversely, the transnational nature of groups, i.e., the high degree of overlap between clusters, may pose a challenge when attempting to link cluster ESG performance to different national or supranational policies that influence firms' pursuit of sustainability practices.

Aldy, J. E., Bolton, P., Halem, Z., & Kacperczyk, M. T. (2024). Behind schedule: The corporate effort to fulfill climate obligations. *Forthcoming in the Journal of Applied Corporate Finance*.

Alfalih, A. A. (2023). ESG disclosure practices and financial performance: A general and sector analysis of SP-500 non-financial companies and the moderating effect of economic conditions. *Journal of Sustainable Finance and Investment*, *13*(4), 1506–1533. https://doi.org/10.1080/20430795.2022.2150511

Amores-Salvadó, J., Martin-de Castro, G., & Albertini, E. (2023). Walking the talk, but above all, talking the walk: Looking green for market stakeholder engagement. *Corporate Social Responsibility and Environmental Management*, *30*(1), 431–442. https://doi.org/10.1002/csr.2364

Bastos Neves, J. P., & Semmler, W. (2022). A proposal for a carbon wealth tax: Modelling, empirics, and policy [Available at SSRN: https://ssrn.com/abstract=4114243].

Berg, F., Kölbel, J. F., & Rigobon, R. (2022). Aggregate confusion: The divergence of ESG ratings. *Review of Finance*, *26*(6), 1315–1344. https://doi.org/10.1093/rof/rfac033

Berndt, D. J., & Clifford, J. (1994). Using dynamic time warping to find patterns in time series. *AAAIWS'94: Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining*, 359–370.

Billio, M., Costola, M., Hristova, I., Latino, C., & L., P. (2022). Sustainable finance: A journey toward ESG and climate risk [SAFE Working Paper].

Billio, M., Costola, M., Hristova, I., Latino, C., & Pelizzon, L. (2021). Inside the ESG ratings: (dis)agreement and performance. *Corporate Social Responsibility and Environmental Management*, *28*(5), 1426–1445. https://doi.org/10.1002/csr.2177

Bolton, P., Halem, Z., & Kacperczyk, M. T. (2022). The financial cost of carbon. *Journal of Applied Corporate Finance*, *34*(2), 17–29. https://doi.org/10.1111/jacf.12502

Bolton, P., Reichelstein, S., Kacperczyk, M. T., Leuz, C., Ormazabal, G., & Schoenmaker, D. (2021). Mandatory corporate carbon disclosures and the path to net zero. *Management and Business Review*, *1*(3).

Bucci, A., Ippoliti, L., & Valentini, P. (2023). Analysing spatiotemporal patterns of Covid-19 confirmed deaths at the NUTS-2 regional level. *Regional Statistics*, *13*(2), 214–239. https://doi.org/10.15196/RS130202

Chavent, M., Kuentz-Simonet, V., Labenne, A., & Saracco, J. (2018). Clustgeo: An R package for hierarchical clustering with spatial constraints. *Computational Statistics*, *33*, 1–24. https://doi.org/ff10.1007/s00180-018-0791-1

Chen, H., Liu, S., Yang, D., & Zhang, D. (2023). Automatic air pollution monitoring and corporate environmental disclosure: A quasi-natural experiment from China. *Sustainability Accounting, Management and Policy Journal*, *14*(3), 538–564. https://doi.org/10.1108/SAMPJ-07-2022-0385

Chen, R., Liu, Y., Jiang, Y., & Liu, J. (2023). Does ESG performance promote vitality of capital market? Analysis from the perspective of stock liquidity. *Frontiers in Environmental Science*, *11*. https://doi.org/10.3389/fenvs.2023.1132845

Chen, Y. P. (, Zhuo, Z., Huang, Z., & Li, W. (2022). Environmental regulation and ESG of smes in China: Porter hypothesis re-tested. *Science of the Total Environment*, *850*. https://doi.org/10.1016/j.scitotenv.2022.157967

Chen, Z., Hu, L., He, X., Liu, Z., Chen, D., & Wang, W. (2022). Green financial reform and corporate ESG performance in China: Empirical evidence from the green financial reform and innovation pilot zone. *International Journal of Environmental Research and Public Health*, *19*(22). https://doi.org/10.3390/ijerph192214981

Chipalkatti, N., Le, Q. V., & Rishi, M. (2021). Sustainability and society: Do environmental, social, and governance factors matter for foreign direct investment? *Energies*, *14*(19). https://doi.org/10.3390/en14196039

Christensen, D. M., Serafeim, G., & Sikochi, A. (2022). Why is corporate virtue in the eye of the beholder? the case of ESG ratings. *Accounting Review*, *97*(1), 147–175. https://doi.org/10.2308/TAR-2019-0506

Deb, S., & Karmakar, S. (2023). A novel spatio-temporal clustering algorithm with applications on COVID-19 data from the United States. *Computational Statistics and Data Analysis*, *188*. https://doi.org/10.1016/j.csda.2023.107810

DESNZ. (2023). 2022 UK greenhouse gas emissions, provisional figures. *National Statistics*.

Fu, T., & Li, J. (2023). An empirical analysis of the impact of ESG on financial performance: The moderating role of digital transformation. *Frontiers in Environmental Science*, *11*. https://doi.org/10.3389/fenvs.2023.1256052

Gonzaga, B. R., Klotzle, M. C., Brugni, T. V., Rakos, I.-S., Cioca, I. C., Barbu, C.-M., & Cucerzan, T. (2024). The ESG patterns of emerging-market companies: Are there differences in their sustainable behavior after COVID-19? *Sustainability (Switzerland)*, *16*(2). https://doi.org/10.3390/su16020676

Gucciardi, G., Ossola, E., Parisio, L., & Pelagatti, M. (2024). Common factors behind companies' environmental ratings [CefES Working Paper].

He, X., Jing, Q., & Chen, H. (2023). The impact of environmental tax laws on heavy-polluting enterprise ESG performance: A stakeholder behavior perspective. *Journal of Environmental Management*, *344*. https://doi.org/10.1016/j.jenvman.2023.118578

He, Y., Zhao, X., & Zheng, H. (2023). How does the environmental protection tax law affect firm ESG? Evidence from the Chinese stock markets. *Energy Economics*, *127*. https://doi.org/10.1016/j.eneco.2023.107067

Iamandi, I.-E., Constantin, L.-G., Munteanu, S. M., & Cernat-Gruici, B. (2019). Mapping the ESG behavior of European companies. a holistic Kohonen approach. *Sustainability (Switzerland)*, *11*(12). https://doi.org/10.3390/su11123276

Ishizaka, A., Lokman, B., & Tasiou, M. (2021). A stochastic multi-criteria divisive hierarchical clustering algorithm. *Omega (United Kingdom)*, *103*. https://doi.org/10.1016/j.omega.2020.102370

Istat. (2022). I conti delle emissioni atmosferiche (aea). *Istituto Nazionale di Statistica*.

Jaya, I., Ruchjana, B., Andriyana, Y., & Agata, R. (2019). Clustering with spatial constraints: The case of diarrhea in Bandung city, Indonesia. *Journal of Physics: Conference Series*, *1397*(1). https://doi.org/10.1088/1742-6596/1397/1/012068

Kaufman, L., & Rousseeuw, P. J. (1990). *Finding groups in data: An introduction to cluster analysis*. Wiley.

Lai, X., & Zhang, F. (2022). Can ESG certification help company get out of over-indebtedness? evidence from China. *Pacific Basin Finance Journal*, *76*. https://doi.org/10.1016/j.pacfin.2022.101878

Li, J., & Li, S. (2022). Environmental protection tax, corporate ESG performance, and green technological innovation. *Frontiers in Environmental Science*, *10*. https://doi.org/10.3389/fenvs.2022.982132

Lian, Y., Li, Y., & Cao, H. (2023). How does corporate ESG performance affect sustainable development: A green innovation perspective. *Frontiers in Environmental Science*, *11*. https://doi.org/10.3389/fenvs.2023.1170582

López-Cabarcos, M. Á., Santos-Rodrigues, H., Quiñoá-Piñeiro, L., & Piñeiro-Chousa, J. (2023). How to explain stock returns of utility companies from an environmental, social and corporate governance perspective. *Corporate Social Responsibility and Environmental Management*, *30*(5), 2278–2291. https://doi.org/10.1002/csr.2483

Ma, J., Gao, D., & Sun, J. (2022). Does ESG performance promote total factor productivity? Evidence from China. *Frontiers in Ecology and Evolution*, *10*. https://doi.org/10.3389/fevo.2022.1063736

Mattera, R., & Franses, P. H. (2023). Are African business cycles synchronized? Evidence from spatio-temporal modeling. *Economic Modelling*, *128*. https://doi.org/10.1016/j.econmod.2023.106485

Mo, Y., Che, Y., & Ning, W. (2023). Digital finance promotes corporate ESG performance: Evidence from China. *Sustainability (Switzerland)*, *15*(14). https://doi.org/10.3390/su151411323

Montero, P., & Vilar, J. A. (2014). TSclust: An R package for time series clustering. *Journal of Statistical Software*, *62*(1), 1–43. https://doi.org/10.18637/jss.v062.i01

MSCI. (2023a). ESG ratings process. *MSCI ESG Research LLC*.

MSCI. (2023b). MSCI ESG ratings methodology: Carbon emissions key issue. *MSCI ESG Research LLC*.

Mukhtar, B., Shad, M. K., & Woon, L. F. (2023). Predicting the effect of environment, social and governance practices on green innovation: An artificial neural network approach. *Lecture Notes in Networks and Systems*, *550 LNNS*, 527–539. https://doi.org/10.1007/978-3-031-16865-9_42

Ortas, E., Álvarez, I., Jaussaud, J., & Garayar, A. (2015). The impact of institutional and social context on corporate environmental, social and governance performance of companies committed to voluntary corporate social responsibility initiatives. *Journal of Cleaner Production*, *108*, 673–684. https://doi.org/10.1016/j.jclepro.2015.06.089

Panda, A., & Ray, K. K. (2023). Equity market performance: The role of environmental protection and corporate social responsibility efforts. *Business Strategy and Development*. https://doi.org/10.1002/bsd2.274

Phillips, P., & Sul, D. (2007). Transition modeling and econometric convergence tests. *Econometrica*, *75*(6), 1771–1855. https://doi.org/10.1111/j.1468-0262.2007.00811.x

Phillips, P., & Sul, D. (2009). Economic transition and growth. *Journal of Applied Econometrics*, *24*(7), 1153–1185. https://doi.org/10.1002/jae.1080

Qian, S., & Yu, W. (2024). Green finance and environmental, social, and governance performance. *International Review of Economics and Finance*, *89*, 1185–1202. https://doi.org/10.1016/j.iref.2023.08.017

Ronalter, L. M., Bernardo, M., & Romaní, J. M. (2023). Quality and environmental management systems as business tools to enhance ESG performance: A cross-regional empirical study. *Environment, Development and Sustainability*, *25*(9), 9067–9109. https://doi.org/10.1007/s10668-022-02425-0

Sakoe, H., & Chiba, S. (1971). A dynamic programming approach to continuous speech recognition. https://api.semanticscholar.org/CorpusID:107516844

Sakoe, H., & Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *26*(1), 43–49. https://doi.org/10.1109/TASSP.1978.1163055

Saraswati, E., Ghofar, A., Atmini, S., & Dewi, A. A. (2024). Clustering of companies based on sustainability performance using ESG materiality approach: Evidence from Indonesia. *International Journal of Energy Economics and Policy*, *14*(2), 112–125. https://doi.org/10.32479/ijeep.15393

Sariyer, G., & Taşkın, D. (2022). Clustering of firms based on environmental, social, and governance ratings: Evidence from BIST sustainability index. *Borsa Istanbul Review*, *22*, S180–S188. https://doi.org/10.1016/j.bir.2022.10.009

UBA. (2024). Annual carbon dioxide emissions in Germany from 1990 to 2022. *the environment agency of the German government*.

Wang, Q. (2023). Herding behavior and the dynamics of ESG performance in the European banking industry. *Finance Research Letters*, *58*. https://doi.org/10.1016/j.frl.2023.104640

Wang, X., Elahi, E., & Khalid, Z. (2022). Do green finance policies foster environmental, social, and governance performance of corporate? *International Journal of Environmental Research and Public Health*, *19*(22). https://doi.org/10.3390/ijerph192214920

Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, *58*(301), 236–244. https://doi.org/10.1080/01621459.1963.10500845

Wu, Y., & Tham, J. (2023). The impact of executive green incentives and top management team characteristics on corporate value in China: The mediating role of environment, social and government performance. *Sustainability (Switzerland)*, *15*(16). https://doi.org/10.3390/su151612518

Xue, Q., Wang, H., & Bai, C. (2023). Local green finance policies and corporate ESG performance. *International Review of Finance*, *23*(4), 721–749. https://doi.org/10.1111/irfi.12417

Yu, X., & Xiao, K. (2022). Does ESG performance affect firm value? Evidence from a new ESG-scoring approach for Chinese enterprises. *Sustainability (Switzerland)*, *14*(24). https://doi.org/10.3390/su142416940

Zhang, D. (2022). Environmental regulation and firm product quality improvement: How does the greenwashing response? *International Review of Financial Analysis*, *80*. https://doi.org/10.1016/j.irfa.2022.102058

Zhang, D., Meng, L., & Zhang, J. (2023). Environmental subsidy disruption, skill premiums and ESG performance. *International Review of Financial Analysis*, *90*. https://doi.org/10.1016/j.irfa.2023.102862

Zhao, X., & Cai, L. (2023). Digital transformation and corporate ESG: Evidence from China. *Finance Research Letters*, *58*. https://doi.org/10.1016/j.frl.2023.104310

Zheng, M., Feng, G.-F., Jiang, R.-A., & Chang, C.-P. (2023). Does environmental, social, and governance performance move together with corporate green innovation in China? *Business Strategy and the Environment*, *32*(4), 1670–1679. https://doi.org/10.1002/bse.3211

2.A Appendix 2.A: Further details on the average proportion of explained mixed pseudo inertia

Let us consider a partition $\mathcal{P}_K^\alpha$ in K clusters obtained mixing the dissimilarity matrices D_0 and D_1 with the coefficient α. Also, let us denote its within-clusters mixed inertia as $W(\mathcal{P}_K^\alpha)$ and the corresponding proportion of the total pseudo inertia explained as

$$Q_0(\mathcal{P}_K^\alpha) = 1 - \frac{W_0(\mathcal{P}_K^\alpha)}{W_0(\mathcal{P}_1)} \qquad Q_1(\mathcal{P}_K^\alpha) = 1 - \frac{W_1(\mathcal{P}_K^\alpha)}{W_1(\mathcal{P}_1)},$$

where $W_0(\mathcal{P}_1)$ and $W_1(\mathcal{P}_1)$ are the total pseudo inertia under dissimilarity matrix D_0 and under dissimilarity matrix D_1, respectively.

The two previous expressions can be reformulated as follows:

$$Q_0(\mathcal{P}_K^\alpha) \cdot W_0(\mathcal{P}_1) = W_{D_0}(\mathcal{P}_1) - W_{D_0}(\mathcal{P}_K^\alpha)$$

$$Q_1(\mathcal{P}_K^\alpha) \cdot W_1(\mathcal{P}_1) = W_{D_1}(\mathcal{P}_1) - W_{D_1}(\mathcal{P}_K^\alpha)$$

Now, let us denote the *weighted average of explained mixed pseudo inertia for partition* $\mathcal{P}_K^\alpha$ as the following linear combination:

$$\bar{Q}(\mathcal{P}_K^\alpha) = \frac{Q_{D_0}(\mathcal{P}_K^\alpha) \cdot W_{D_0}(\mathcal{P}_1) + Q_{D_1}(\mathcal{P}_K^\alpha) \cdot W_{D_1}(\mathcal{P}_1)}{W_{D_0}(\mathcal{P}_1) + W_{D_1}(\mathcal{P}_1)}.$$

In practice, we linearly combine the proportion of pseudo inertia explained by partition $\mathcal{P}_K^\alpha$ under dissimilarity D_0 (i.e., $Q_{D_0}(\mathcal{P}_K^\alpha)$) and the share of pseudo inertia explained under dissimilarity D_1 (i.e., $Q_{D_1}(\mathcal{P}_K^\alpha)$) weighting the values with the total pseudo inertia provided by the geographical information (i.e., $W_{D_0}(\mathcal{P}_1)$) and the total pseudo inertia provided by the socio-economic features (i.e., $W_{D_1}(\mathcal{P}_1)$).

Alternatively, the weighted average can also be expressed as a function of the total relative gain obtained by implementing a mixed partitioning instead of a purely-spatial or purely-socio-economic

information approach, that is,:

$$\bar{Q}(\mathcal{P}_K^\alpha) = \frac{[W_{D_0}(\mathcal{P}_1) - W_{D_0}(\mathcal{P}_K^\alpha)] + [W_{D_1}(\mathcal{P}_1) - W_{D_1}(\mathcal{P}_K^\alpha)]}{W_{D_0}(\mathcal{P}_1) + W_{D_1}(\mathcal{P}_1)}$$

where $W_{D_0}(\mathcal{P}_1) - W_{D_0}(\mathcal{P}_K^\alpha)$ is the discrepancy between the overall pseudo inertia using D_0 only and the pseudo inertia induced by partition $\mathcal{P}_K^\alpha$ in K clusters (i.e., the gain in inertia obtained by using the combination of the two matrices instead of the D_0 matrix alone), while $W_{D_1}(\mathcal{P}_1) - W_{D_1}(\mathcal{P}_K^\alpha)$ is the discrepancy between the overall pseudo inertia using D_0 only and the pseudo inertia induced by partition $\mathcal{P}_K^\alpha$ in K clusters (i.e., the gain in inertia obtained by using the combination of the two matrices instead of the D_1 matrix alone). This expression allows a further interpretation of the proposed criterion for selecting α, that is, we are maximizing the relative gain in term of inertia induced by using a mixture of geographical and socio-economic information to cluster the observations instead of employing a purely spatial or purely socio-economic clustering.

Eventually, by rearranging the previous expression, one can rewrite $\bar{Q}(\mathcal{P}_K^\alpha)$ as a function of the within-cluster pseudo inertia as follows:

$$\begin{aligned}\bar{Q}(\mathcal{P}_K^\alpha) &= \frac{W_{D_0}(\mathcal{P}_1) + W_{D_1}(\mathcal{P}_1) - [W_{D_0}(\mathcal{P}_K^\alpha) + W_{D_1}(\mathcal{P}_K^\alpha)]}{W_{D_0}(\mathcal{P}_1) + W_{D_1}(\mathcal{P}_1)} \\ &= 1 - \frac{W_{D_0}(\mathcal{P}_K^\alpha) + W_{D_1}(\mathcal{P}_K^\alpha)}{W_{D_0}(\mathcal{P}_1) + W_{D_1}(\mathcal{P}_1)}.\end{aligned}$$

Notice that, the last expression allows for an easy generalization to the case of $p = 1, 2, \ldots, P$ dissimilarity matrices $D_1, \ldots, D_p, \ldots, D_P$, as in the case of spatiotemporal clustering with multiple dimensions,

$$\bar{Q}(\mathcal{P}_K^{\alpha_p}) = 1 - \frac{\sum_{p=1}^P W_{D_p}(\mathcal{P}_K^{\alpha_p})}{\sum_{p=1}^P W_{D_p}(\mathcal{P}_1^{\alpha_p})}.$$

2.B Appendix 2.B: Comparison between different ESG rating providers

	MSCI	VIGEO	REFINITIV	Suatainalitics	ISS OEKOM	RobeccoSAM	ECPI	Bloomberg
RATING SCORE	CCC to AAA	– to ++	D- to A+	0 to 100	D- to A+	0 to 100	F to EEE	0 to 100
HISTORY	1990	1983	2002	1992	1985	1995	1997	2008
SOURCES	Company disclosure, 1600+ Media sources, 100+ specialized dataset	Company disclosure, Recommendation, Conventions	Company websites, Company reports, NGO Websites, Media and news, Stock Exchange filings	Public disclosure, Media and news, NGO reports	Publicly available information, Interviews with stakeholders, information on company policies and practices, company direct contact	Survey approach	Company reports, Company screening, Media and news, Regulatory data, Bloomberg and Thomson Reuters, University networks	Company reports, Company screening, Media and news, Regulatory data, Bloomberg and Thomson Reuters, University networks
MAIN RISK FACTORS	Environmental Climate Change, Natural Resources Pollution And Waste Mgmt. Environmental Opportunities Social Product Liability Human Capital Stakeholder Needs Social Opportunities Governance Corporate Behaviour Corporate Governance	Human Resources, Human Rights Environment Business Behaviour Community Involvement Corporate Governance	Environmental Resource Use, Emission, Innovation Social Workforce, Human Rights, Community, Product Responsibility Governance Mgmt., Shareholders, Csr Strategy	Industry-Specific indicators. Factors Change According To The Industrial Group To Which A Company Belongs	Environment Climate Change Strategy, Ecoefficiency, Energy Mgmt., Env. Impact of Product Portfolio, Env. Mgmt., Water Risk And Impact Social Equal Opportunities, Freedom of Association, Health and Safety, Human Rights, Product Responsibility, Social Impact of Product Portfolio, Supply Chain Mgmt., Taxes Governance Business Ethics, Compliance, Independence of The Board, Remuneration, Shareholder Democracy, Shareholder Structure	Industry- Specific Indicators Three Main Dimensions: Economic (38/100) Environmental (27/100) Social (35/100)	Environmental Environmental Strategy Policy Environmental Mgmt. Products Production Process Social and Governance Employees and Human Capital Community Relations Markets Corporate Governance and Shareholder	Environmental Carbon Emissions, Climate Change Effect, Pollution, Waste Disposal, Renewable Energy, Resource Depletion Social Supply Chain, Political Contributions, Discrimination, Diversity, Community Relations, Human Rights, Governance Cumulative Voting, Executive Compensation, Shareholders' Rights, Takeover Defence, Staggered Boards, Independent Directors

Table 2.B.1: Key differences between ESG rating agencies as reported by Billio et al., 2021 in Table 1.

2.C Appendix 2.C: MSCI Carbon Emission Score methodology

Focusing on the Carbon Emission score, it evaluates the company's level of exposure to, and management of its risks, with regard to emissions whose sources are traceable to those indicated in scope 1 and 2 of the Greenhouse Gas Protocol. Specifically, scope 1 emissions cover direct emissions over one year from establishments that are owned or controlled by the company, while scope 2 emissions come from the generation of purchased heat, steam, and electricity consumed by the company [2]. In the approach used by MSCI, 2023b[3], the Carbon Emission score is calculated by combining information on the Carbon Emission Exposure Score, which evaluates the company's exposure to risks on this Key Issue including industry-specific adjustments, and the Carbon Emissions Management Score which evaluates the company's ability to manage its exposure to Carbon Emission risk according to company disclosures.

As mentioned above, ESG assessments are only available for those companies that voluntarily publish the necessary information. In Table 2.C.1 we provide a comparison between the number of listed companies located in Western European countries included in our sample and those with ESG rating provided by MSCI.

	N. of firms	%	Total asset	%	Sales	%
All listed firms	4.848	1	27.815	1	9.871	1
Firms with ESG rating	1266	0.26	25.915	0.93	8.714	0.88

Table 2.C.1: Comparison between all listed firms located in Western European countries and those with ESG rating provided by MSCI, considering the number of firms, the sum of the total asset in 2022 (in billion of euros), the sum of the sales revenue in 2022 (in billion of euros) and their percentages with respect to the total. Note that financial data refer to year 2022, while ESG rating have been assigned in 2023 based on the non-financial disclosure available in the previous year, thus referring to 2022. Note that the number of firms with ESG rating is slightly lower than in our sample because here we can consider only companies that are still active and still listed in July 2024, while in the sample used in the analysis we included also companies no longer listed and no longer active in order to keep as many observations as possible.

[2]See https://ghgprotocol.org/standards-guidance

[3]We recall that MSCI provides methodological details upon request through a public procedure available at the following link: https://www.msci.com/esg-and-climate-methodologies

According to the data collected from Orbis dataset, we are considering only one quarter of the companies but they represent 93% of the sum of the total asset and 88% of the sum of sales revenue. In other words, although we are considering less than half of active listed companies, these represent the majority of the invested capital and sales revenue.

2.D Appendix 2.D: The role of spatial constraint in Cluster Analysis

In Section 2.5.1, the resulting spatial clusters show some degree of overlap in space, except for Cluster 4, which consists only of companies located in Spain and Portugal. In this case, the geographical distance from the Iberian peninsula to the rest of Europe seems to prevail over the distance calculated on the ESG scores, and therefore the resulting cluster seems to absorb only the spatial information, ignoring the differences between variables. In order to better examine the role of the spatial component on cluster formation, we report in this section the results obtained from cluster analysis in cases when using only the distance matrix D_0 obtained from the Euclidean distance of the ESG score, Environmental score and Carbon Emission score in 2023 for 617 companies and when combining it with the spatial information contained in the complementary matrix of the Nearest Neighbour Distance Matrix ${D_1}^{NNDc} = 1 - {D_1}^{NND}$. The latter is computed assigning ${d_{1,ij}}^{NND} = 1$ if firm i is one of the m closest firms to j or vice versa and 0 otherwise. In other words, ${D_1}^{NNDc}$ can be interpreted as a simplified version of D_1, where m shorter distances are indicated with 0 and longer distances are equal to 1. In our case we decide to set $m = n/K$.

Examining the first case, Figure 2.D.1 and 2.D.2 illustrates the map of firms belonging to different clusters and the centroids of the variables. Clusters 1, 2 and 3 refer to firms with higher sustainability performance, and they seem to be spread evenly throughout Europe. Considering firms with worse performance, Cluster 4 exhibits the lowest values for Carbon Emission and Environmental scores, while Cluster 5 shows the lowest centroid for ESG score. Companies belonging to these clusters are mainly located in the northern, central and eastern areas of the territory concerned, while just a few of them are located in the Iberian Peninsula. We can see from this that the spatial cluster of the Iberian peninsula described in section 5 is not only due to the long distances between Spanish and Portuguese companies compared with the rest of European companies, but also reflects the fact that companies show slightly better sustainability performance, in fact Iberian companies with particularly low sustainability scores are few.

Now, we observe the spatial clusters obtained combining the Euclidean distance across the variables in 2023 with the complementary of the Nearest Neighbour distance matrix ${D_1}^{NNDc} =$

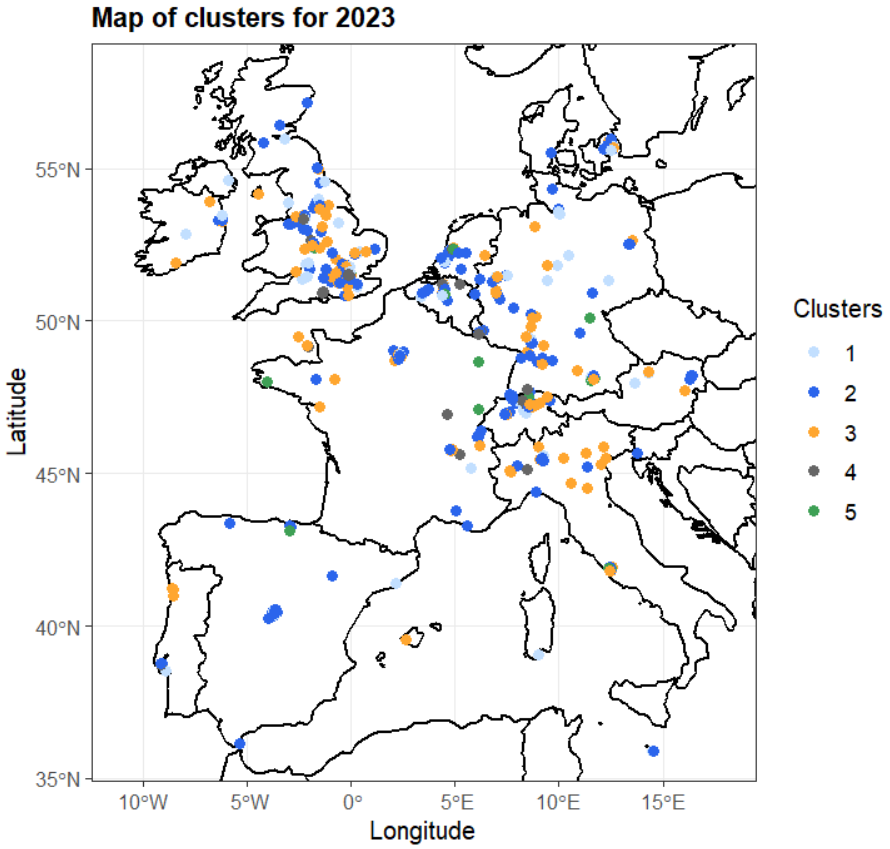


Figure 2.D.1: Map of clusters for 2023 computed using Euclidean distance of ESG, Environmental and Carbon Emission scores.

$1 - {D_1}^{NND}$, considering $m = n/K = 617/5 \approx 123$. We set $K = 5$ and $\Delta\alpha = 0.1$ and we find $\alpha^*_{K=5} = 0.1$ such that the combination of spatial and features information provide the maximum average weighted explained inertia of the dissimilarity matrices.

In Figure 2.D.3 and 2.D.4 it is possible to observe that the clusters obtained are quite similar to those shown in Section 2.5.1, where the spatial dissimilarity matrix contains geodetic distances calculated from the coordinates of firms. Also in this case, a cluster with particularly low sustainability ratings is identified and its companies are mainly located in England and central Europe. As concerns the Iberian Peninsula, although there is no geographically separate cluster from all other companies, most Spanish and Portuguese firms belong to Cluster 3, which also extends into

Figure 2.D.2: Cluster-specific centroids of the three economic variables for 2023. Clusters are computed using Euclidean distance of ESG (esg), Environmental (env) and Carbon Emission (ce) scores. The solid lines represent the means of the variables, dashed lines represent the two quartiles.

France, south of the UK and Belgium. This is a fairly compact group whose ESG, Environmental and Carbon Emission scores are higher than the other identified clusters.

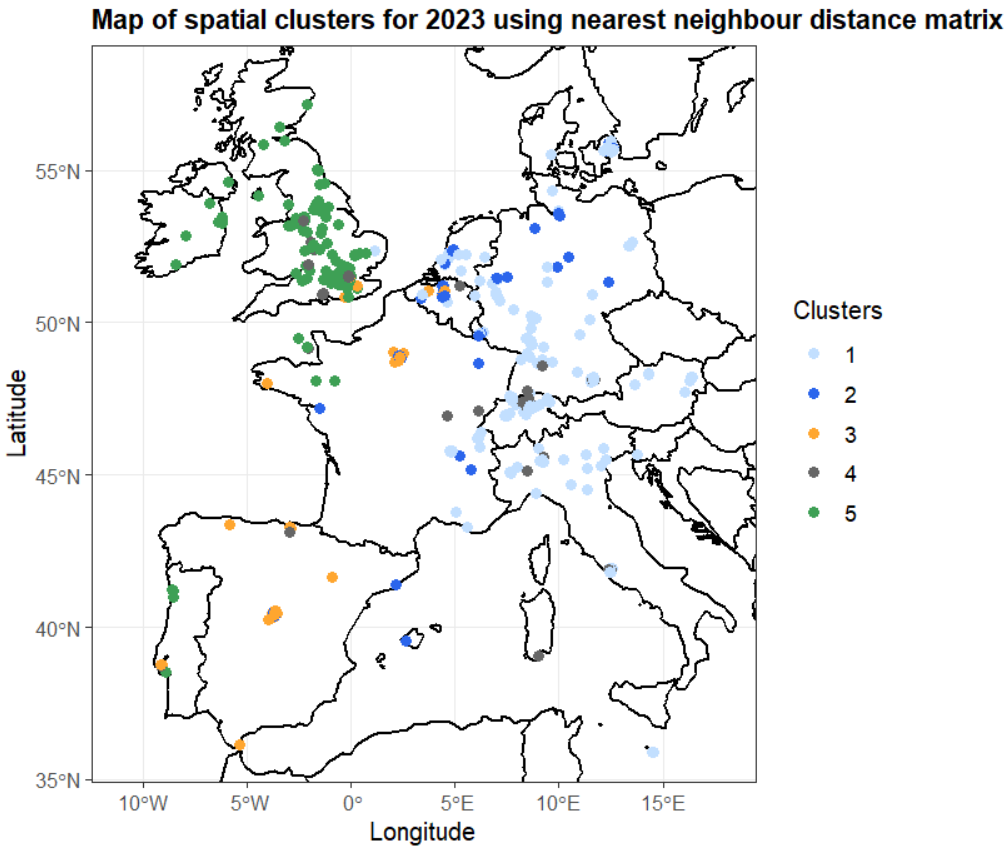


Figure 2.D.3: Map of spatial clusters for 2023 computed using Euclidean distance of ESG, Environmental and Carbon Emission scores and the complementary matrix of the Nearest Neighbour matrix of the location of firms.

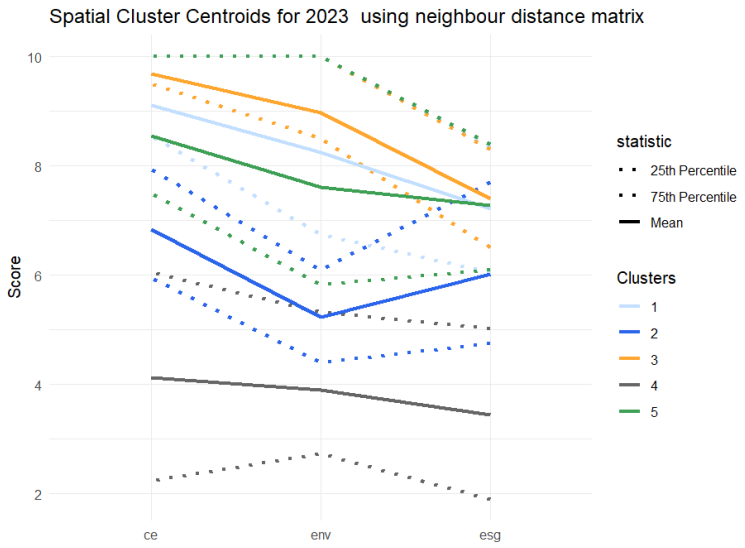


Figure 2.D.4: Cluster-specific centroids of the three economic variables for 2023. Clusters are computed using Euclidean distance of ESG (esg), Environmental (env) and Carbon Emission (ce) scores. The solid lines represent the means of the variables, dashed lines represent the two quartiles.

Chapter 3

The role of geography in clustering GHGs Emissions

from

Spatiotemporal clustering of GHGs emissions in Europe: exploring the role of spatial component

Coauthored with: Paolo Maranzano, Philipp Otto

Revised and resubmitted: *Spatial Statistics*

3.1 Introduction

The growing availability of spatiotemporal data presents both opportunities and challenges for researchers. In the environmental field, new technologies such as remote sensing, GPS devices, mobile sensors, and satellite imagery allow us to capture environmental phenomena at unprecedented spatial and temporal resolutions. These data enable detailed investigations into processes that evolve over both space and time. A large body of work has focused on developing advanced spatiotemporal models to predict pollutant concentrations and to assess their dynamics in different areas (Calculli et al., 2015; Maranzano et al., 2023; Najafabadi et al., 2020; Otto et al., 2024; Rowland, 2024; Taghavi-Shahri et al., 2020; Yu et al., 2024). Other contributions evaluate the effects of specific events (Maranzano & Fassó, 2022; Maranzano & Pelagatti, 2024) or policies (Maranzano et al., 2020; Pisoni et al., 2025; Rodeschini et al., 2024) on air quality.

Despite these advances, a critical gap remains: most studies analyze pollutant concentrations, while very few investigate pollutant emissions (Banja et al., 2025; Marongiu et al., 2024, 2025; Mutascu, 2022). Concentrations are influenced not only by emissions but also by morphological and geographical factors such as elevation, land use, and natural barriers, which affect dispersion and accumulation. Focusing exclusively on concentrations may thus obscure the fundamental role of emissions themselves—the primary drivers of pollution and climate change. Studying emissions at their source is essential to better understand the mechanisms of environmental degradation and to design effective policies.

To encourage researchers to look at air pollution from a new and challenging perspective, spatiotemporal cluster analysis offers a powerful approach to start bridging this gap, enabling the identification of spatial and temporal patterns of GHGs emissions.

In this study, spatiotemporal clustering has to be intended as the grouping of spatial units (i.e., the European regions in the empirical application) while accounting for their temporal dynamics. Each unit is described by time series of several emission variables and by its fixed geographical location. We build a single, time-invariant partition of the units by minimizing a criterion that combines (i) dissimilarities between unit-level time series (i.e., the temporal/feature information) and (ii) spatial dissimilarities between geographical coordinates. Spatial proximity increases the

chance that two regions are clustered together only insofar as it improves this combined objective; however, contiguity is not enforced.

Early works such as Oliver and Webster, 1989 and Bourgault et al., 1992 proposed the first techniques for cluster analysis of units combining features with spatial information, emphasizing the need to group units that exhibit both similar time series and geographical proximity. Several methodologies for cluster analysis of spatiotemporal data are reported in the literature review by Ansari et al., 2020, who classify them according to a taxonomy based on data types. Among the most widely used methods are density-based algorithms, such as the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) introduced by Ester et al., 1996, together with its extensions (Birant & Kut, 2007; Chen et al., 2015; Georgoulas et al., 2013; Rocha et al., 2010; Tietbohl et al., 2008). These approaches are able to detect clusters of arbitrary shape and varying sizes without requiring the number of clusters to be fixed in advance. However, they are known to be highly sensitive to input parameters such as radius, minimum number of points, or threshold values. With specific regard to time series clustering using dissimilarity measures, Piccolo, 1990 first introduced the concept by proposing a distance metric for autoregressive integrated moving-average (ARIMA) invertible models, defined as the Euclidean distance between their autoregressive expansions. Later, Pértega-Díaz and Vilar, 2010 developed several non-parametric statistics to test the equality of the log-spectra of two stochastic processes, considering both stationary and non-stationary time series, ARMA processes, and non-linear models. More recently, Montero and Vilar, 2014 described and compared a wide range of time series dissimilarity measures for clustering, including those based on raw data, extracted features, underlying parametric models, complexity levels, and forecast behaviors. They also implemented these measures in the R package `TSclust`. Other more recent methodologies explicitly incorporate spatial information. For instance, Zhou et al., 2019 extended the well known K-means algorithm to network data (like regionalized lattice data) by designing two different adjacency matrices based on the spectral decomposition of the graph information. Fouedjio, 2016 developed an agglomerative hierarchical clustering approach for spatially contiguous clusters, relying on a dissimilarity matrix built from a non-parametric kernel estimator of the multivariate spatial dependence structure. Building on this, D'Urso and Vitale, 2020 proposed a robust metric approach to neutralize the effect of possible outliers, exploiting an exponential transformation of a dissimilarity measure. More recently, Benevento and Durante, 2024, within

a hierarchical framework, suggested merging spatial and temporal information into correlation matrices that take into account the underlying geometric structure of correlation matrix space. In epidemiological applications, I. G. N. M. Jaya and Folmer, 2021 propose a two-stage spatiotemporal clustering approach to identify clusters that vary across space and time, while I. G. N. M. Jaya and Folmer, 2020 develop a Bayesian spatiotemporal random-effects model with spatiotemporal interactions and demonstrate heterogeneous temporal patterns across spatial units.

Within the same hierarchical clustering framework, which does not require assumptions on data distribution or cluster shapes, Chavent et al., 2018 proposed a spatial hierarchical clustering algorithm. This approach combines two dissimilarity matrices—one feature-based and one geographical—through a weighted scheme. Owing to its flexibility and interpretability, the method has gained popularity in subsequent applications (Bucci et al., 2023; Deb & Karmakar, 2023; I. Jaya et al., 2019; Mattera & Franses, 2023; Morelli et al., 2025; Zammarchi & Maranzano, 2024). More recently, Morelli et al., 2025 extended this approach by enabling the use of multiple dissimilarity matrices simultaneously. This extension is particularly valuable when analyzing time series of different variables, since it is possible to obtain one dissimilarity matrix for each variable (Montero & Vilar, 2014). The resulting algorithm produces a cluster partition in which the overall dissimilarity within each cluster is minimized. In other words, it allows researchers to identify groups of units that display similar temporal dynamics of the variables while also being geographically close, though not necessarily adjacent.

In this paper, we build on this framework and propose a novel measure, namely the Joint Inertia index, to quantify the contribution of spatial distances relative to non-spatial dissimilarities in the clustering process. Indeed, while the mixing weights determine how dissimilarity matrices are combined, they should not be read as the importance of space; instead, Joint Inertia measures the incremental explanatory power of spatial information beyond features alone. Our contributions are twofold: on the one side we propose a methodological improvement by introducing Joint Inertia as a tool to assess the role of space when combined with other dissimilarities and illustrating its behavior in theoretical scenarios and simulations; on the other side, we provide an empirical contribution by applying the spatio-temporal clustering and the Joint Inertia to the case of georeferenced time series of greenhouse gases (GHGs) emissions for several types of gases recorded across the European regions at the sectoral level with the primary goal to reveal geographically-relevant spatio-temporal

patterns, and evaluate how much the spatial information helps in identifying them.

The relevance of these findings extends to both policymakers and statisticians. For policymakers, the insights gained from this analysis are instrumental in formulating localized strategies to mitigate GHGs emissions, enabling more targeted and effective interventions. For statisticians in the environmental field, this study provides a baseline analysis that highlights critical factors to consider when modeling GHGs emissions data. Specifically, it underscores the importance of incorporating spatial and temporal dimensions and offers guidance on whether spatiotemporal models are necessary for accurate representation and prediction. Together, these contributions lay the groundwork for more robust and nuanced approaches to addressing pressing environmental challenges.

The rest of the paper is structured as follows: in Section 3.2, we describe the methodology proposed by Chavent et al., 2018 and Morelli et al., 2025 and we show the algorithm that summarizes the methodological framework of the paper, in Section 3.3, we provide the methodological contribution of the paper, in Section 3.4 we present the simulation experiment on the performance of the clustering algorithm and the measure to assess the role of the spatial component and in Section 3.5 we discuss the empirical findings, thus the clusters obtained from the algorithm and the contribution of the spatial distances.

3.2 Spatiotemporal hierarchical clustering

In this section, we briefly summarize the baseline methodology. First we discuss the hierarchical clustering method proposed by Chavent et al., 2018, in which a partition of clusters is obtained by considering the dissimilarities between units from two different matrices. Second, we describe the extension proposed by Morelli et al., 2025 to more than two dissimilarity matrices, which is proved particularly useful in spatiotemporal application and adds a novel algorithm for the identification of the weights for the spatial and non-spatial matrices. This builds the methodological background for our newly proposed measure, the Joint Inertia, which will be introduced in Section 3.3. Joint Inertia allow to directly assess the importance of the spatial component in the resulting clustering partition.

3.2.1 Baseline methodology: spatial hierarchical clustering

Let $D_0 = [d_{0,ij}]_{i,j=1,\dots,n}$ and $D_1 = [d_{1,ij}]_{i,j=1,\dots,n}$ refer respectively to any distance matrix of variables and the spatial distances matrix considering a sample of n units. Also, let w_i with $i = 1, \dots, n$ be the weight of the i-th cross-sectional unit to be clustered. When no prior information is available, the weight is commonly set to $w_i = 1/n$. In order to account for different measurement scales, it is necessary to scale D_0 and D_1 with respect to their maximum values so that the clustering distances across observations take values between 0 and 1.

In the spatial clustering of Chavent et al., 2018, namely the ClustGeo algorithm, feature- and space-based information are combined through a mixing parameter $\alpha \in [0, 1]$ that tunes the contribution of spatial dissimilarities: $\alpha = 0$ yields a purely feature-based solution, whereas $\alpha = 1$ yields a purely spatial one. For a partition $\mathcal{P}_K^\alpha = (\mathcal{C}_1^\alpha, \dots, \mathcal{C}_K^\alpha)$, the mixed within–cluster pseudo-inertia of $\mathcal{C}_k^\alpha$ is

$$I_\alpha(\mathcal{C}_k^\alpha) = (1-\alpha)\,\frac{1}{2\mu_k}\sum_{i\in\mathcal{C}_k^\alpha}\sum_{j\in\mathcal{C}_k^\alpha} w_i w_j\, d_{0,ij}^2 \;+\; \alpha\,\frac{1}{2\mu_k}\sum_{i\in\mathcal{C}_k^\alpha}\sum_{j\in\mathcal{C}_k^\alpha} w_i w_j\, d_{1,ij}^2, \tag{3.2.1}$$

where $\mu_k = \sum_{i\in\mathcal{C}_k^\alpha} w_i$. The total mixed within clusters pseudo-inertia of the partition is

$$W_\alpha(\mathcal{P}_K^\alpha) = \sum_{k=1}^{K} I_\alpha(\mathcal{C}_k^\alpha). \tag{3.2.2}$$

The agglomerative step merges the two clusters A and B that minimize the increase

$$\delta_\alpha(A, B) = I_\alpha(A \cup B) - I_\alpha(A) - I_\alpha(B). \tag{3.2.3}$$

Following Chavent et al., 2018, this Ward-like procedure is implemented via the Lance–Williams recursion on an aggregation matrix rather than on the raw dissimilarities. Define, for $m \in \{0, 1\}$,

$$\delta_{m,ij} = \frac{w_i w_j}{w_i + w_j}\, d_{m,ij}^2, \qquad \Delta_m = \left[\delta_{m,ij}\right]_{i,j}, \qquad \Delta_\alpha = (1-\alpha)\Delta_0 + \alpha\Delta_1, \tag{3.2.4}$$

so that for singletons $\{i\}, \{j\}$ we have $\delta_\alpha(\{i\}, \{j\}) = (1-\alpha)\delta_{0,ij} + \alpha\delta_{1,ij} = \frac{w_i w_j}{w_i + w_j}\left[(1-\alpha)d_{0,ij}^2 + \alpha d_{1,ij}^2\right]$. The Lance–Williams update applied to Δ_α yields $\delta_\alpha(A \cup B, C)$, and the hierarchy proceeds by repeatedly merging the pair with minimal $\delta_\alpha(A, B)$. This differs from running Ward directly on

$(1-\alpha)D_0 + \alpha D_1$ and ensures that the optimization targets W_α.

The partition minimizes a within–cluster criterion built from the weighted combination of inertia in D_0 and D_1. Two units tend to merge when their combined pseudo-inertia is smaller than competing pairs. As α increases, clusters become more spatially contiguous; as α decreases, feature similarity dominates. Thus D_1 modulates—rather than enforces—spatial contiguity within the objective being minimized.

The main issue in the hierarchical clustering methodology concerns the choice of the mixing parameters α and the number of clusters K. Chavent et al., 2018 suggest setting a prior value for K and then finding α such that it is the best compromises between loss of socio-economic and loss of geographical homogeneity. In order to clarify their approach, we recall the notion of the proportion of the total inertia explained by partition $\mathcal{P}_K^\alpha$ in D_0 and D_1 for K clusters, which is computed as

$$Q_{D_0}(\mathcal{P}_K^\alpha) = 1 - \frac{W_{D_0}(\mathcal{P}_K^\alpha)}{W_{D_0}(\mathcal{P}_1)} \qquad Q_{D_1}(\mathcal{P}_K^\alpha) = 1 - \frac{W_{D_1}(\mathcal{P}_K^\alpha)}{W_{D_1}(\mathcal{P}_1)} \tag{3.2.5}$$

Specifically, $Q_{D_0}(\mathcal{P}_K^\alpha)$ quantifies the proportion of inertia from features dissimilarities (i.e., $W_{D_0}(\mathcal{P}_1)$) explained by partition $\mathcal{P}_K^\alpha$, while $Q_{D_1}(\mathcal{P}_K^\alpha)$ quantifies the amount of geographical inertia (i.e., $W_{D_1}(\mathcal{P}_1)$) explained by partition $\mathcal{P}_K^\alpha$.

To account for potential scale issues in $Q_{D_0}(\mathcal{P}_K^\alpha)$ and $Q_{D_1}(\mathcal{P}_K^\alpha)$, are then normalized with respect to the baseline case of purely-geographical or purely-features clustering, that is, by computing the following ratios:

$$\tilde{Q}_{D_0}(\mathcal{P}_K^\alpha) = \frac{Q_{D_0}(\mathcal{P}_K^\alpha)}{Q_{D_0}(\mathcal{P}_K^0)} \qquad \tilde{Q}_{D_1}(\mathcal{P}_K^\alpha) = \frac{Q_{D_1}(\mathcal{P}_K^\alpha)}{Q_{D_1}(\mathcal{P}_K^1)}. \tag{3.2.6}$$

This relative formulation allows for a straightforward interpretation of the values. For a given K and a given mixing parameter α, the formulas are expressing the percentage of the explained proportion of inertia in features (or spatial) dissimilarity matrix obtained by using a mixture of features and geographical information to generate the partition $\mathcal{P}_K^\alpha$ (i.e., $Q_{D_0}(\mathcal{P}_K^\alpha)$ (or $Q_{D_1}(\mathcal{P}_K^\alpha)$)) with respect to the proportion of inertia that would be explained by only using features (or spatial) dissimilarities to generate the partition $\mathcal{P}_K^0$ (or $\mathcal{P}_K^1$) in K clusters (i.e., $Q_{D_0}(\mathcal{P}_K^0)$ (or $Q_{D_1}(\mathcal{P}_K^1)$)). For a deeper understanding and comprehensive discussion about the properties of these quantities,

we refer the readers to Section 3 in Chavent et al., 2018.

Thus, they suggest to choose α considering a trade-off between the loss of socio-economic homogeneity, thus the normalized proportion of explained inertia from the feature dissimilarities, and greater geographical cohesion measured as the normalized proportion of expalined inertia in the spatial dissimilarity.

3.2.2 Spatiotemporal hierarchical clustering

Morelli et al., 2025 introduced a generalized version of spatial hierarchical clustering that allows the use of more than two dissimilarity matrices, then allowing for a spatio-temporal formulation of the clustering task. Also, they identify the optimal combination of these matrices which maximizes the gain induced by using a mixed approach instead of a purely spatial and non-spatial approach in terms of proportion of explained inertia within the clusters without forcing the algorithm to include all the input matrices in the final combination. Such principle holds for both the spatial and non-spatial information.

The need to incorporate multiple dissimilarity matrices arose from our objective of analyzing a dataset consisting of time series for more than one variable and the geographic coordinates of individual units. Indeed, the approach proposed by Chavent et al., 2018 allows to combine the information from the spatial dissimilarity matrix and one features dissimilarity matrix containing, for example, the Euclidean distances of variables observed at a given moment or a matrix representing the distances between time series for a single variable (Bucci et al., 2023; Deb & Karmakar, 2023; Mattera & Franses, 2023). When dealing with multiple time series, it becomes necessary to compute a dissimilarity matrix for each variable observed over a given period. Therefore, the challenge lies in combining more than two matrices. Notice that the methodology is not limited to spatiotemporal data but can be applied to any scenario where high-dimensional data requires the combination of more than two dissimilarity matrices. Nevertheless, this approach has been developed specifically to address the need for identifying patterns and dynamics in phenomena where both temporal evolution and geographic location are critical.

In our setting, the temporal dimension enters the clustering exclusively through trajectory-

based dissimilarities between regions. For each emission variable, we construct a dissimilarity matrix measuring differences between full regional time series, which is combined with a spatial dissimilarity matrix within a Ward-type mixed pseudo-inertia criterion. This yields a time-invariant partition of spatial units: we cluster regions, not time points, and spatial proximity influences merges only insofar as it improves the mixed objective, without enforcing contiguity. Moreover, we adopt the ClustGeo formulation of Chavent et al., 2018, based on linear combinations of pseudo-inertias, and follow the selection strategy of Morelli et al., 2025 to determine both the weighting vector and the number of clusters.

Let $D_p = [d_{p,ij}]_{p=1,\dots,P-1;i,j=1,\dots,n}$ being the dissimilarity matrices containing the distances across time series for P variables for n units. Multiple alternative approaches can be used to compute time series dissimilarities metrics; among others, we refer the reader to Montero and Vilar, 2014 for a comprehensive review and comparison of the most recent methodologies. In our empirical application we adopt the Dynamic Time Warping (DTW) (Sakoe & Chiba, 1971, 1978), which quantifies the (dis)similarity between two time series by (non-linearly) aligning them in time and minimizing the sum of point-wise distances between matched elements subject to monotonicity and continuity constraints. Furthermore, as shown by Berndt and Clifford, 1994, the DTW distance is equivalent to the Euclidean distance computed on aligned time series.

Let us assume that the first $P-1$ matrices are associated with non-spatial features or time series and let $D_P = [d_{P,ij}]_{i,j=1,\dots,n}$ refer to the spatial (or geographical) dissimilarity matrix across the n units, which is typically computed as geodetic distances between coordinates. Moreover, we recall that dissimilarity matrices are normalized with respect to their maximum value. The main issue here is to find a suitable combination of the two clustering hyperparameters, namely the number of group K, and the weighting vector $\underline{\alpha} = [\alpha_p]_{p=1,2,\dots,P} = (\alpha_1, \alpha_2, \dots, \alpha_p, \dots, \alpha_P)$ on the $(P-1)$-simplex $\boldsymbol{\Delta}^{P-1}$, that is $\alpha_p > 0$ for all $p = 1, \dots, P$ and $\sum_{p=1}^{P} \alpha_p = 1$. In practice, this requires constructing a grid of candidate values for $\underline{\alpha}$. The grid can, in principle, be either regular or irregular, but in most applications a regular grid with constant step size $\Delta\alpha$ is preferable, since it guarantees a uniform coverage of the parameter space. Specifically, a regular grid is obtained by allowing each α_p to take values in the discrete set $\{0, \Delta\alpha, 2\Delta\alpha, \dots, 1\}$ and then retaining only the combinations that satisfy the summation constraint $\sum_{p=1}^{P} \alpha_p = 1$. In this way, the search for the optimal weighting vector is restricted to a finite number of feasible candidates, ensuring both computational tractability and

systematic exploration of the parameter space. Let us define a range of number of clusters from $K = 1$ to $K = K_{max}$ and for each $K = 1, \dots, K_{max}$ consider all the possible combinations for a grid of $\underline{\alpha}$. For each combination of K and $\underline{\alpha}$ we consider the partition $\mathcal{P}_K^{\underline{\alpha}}$ obtained applying the Ward hierarchical algorithm to the combination of pseudo-inertia from each dissimilarity matrices according to the vector of weight $\underline{\alpha}$ thus considering $I(\mathcal{C}_K^{\underline{\alpha}}) = \sum_{p=1}^{P} \alpha_p \sum_{i \in \mathcal{C}_K} \sum_{j \in \mathcal{C}_K} \frac{w_i w_j}{2 \sum_{i \in \mathcal{C}_K} w_i} d_{p,ij}{}^2$.

Note that Morelli et al., 2025 apply Ward directly to the linear combination of dissimilarity matrices, $D(\underline{\alpha}) = \sum_{p=1}^{P} \alpha_p D_p$. In contrast, we optimize a linear combination of (pseudo) within-cluster inertias in the spirit of Chavent et al., 2018; this is algebraically more involved but aligned with the mathematical foundation underlying Joint Inertia. We nevertheless adopt the hyperparameter selection strategy of Morelli et al., 2025 for $\underline{\alpha}$ and K^* to retain comparability and clarity in the selection step.

Following the criterion defined in Morelli et al., 2025, for every $K = 1, 2, \dots, K_{max}$ we choose the optimal weighting vector $\underline{\alpha} = \underline{\alpha}^*$ which maximizes the weighted average of the explained inertia[1] induced by the partition, given by

$$\bar{Q}(\mathcal{P}_K^{\underline{\alpha}}) = 1 - \frac{\sum_{p=1}^{P} W_{D_p}(\mathcal{P}_K^{\underline{\alpha}})}{\sum_{p=1}^{P} W_{D_p}(\mathcal{P}_1^{\underline{\alpha}})}. \tag{3.2.7}$$

Its value ranges between 0 and 1, reaching its maximum when partition $\mathcal{P}_K^{\underline{\alpha}}$ allows us to explain the total dissimilarity in each matrix. Notice that the maximization of Equation 3.2.7 is consistent with the underlying idea of Ward's criterion according to which the clustering partition is obtained by minimizing the total within cluster pseudo inertia (Equation 3.2.1). Indeed, Equation 3.2.7 is maximal when the numerator of the second component, which is the sum of matrix-specific total mixed pseudo-inertia, is minimum. The optimization step conditioning on K is iterated across a user-defined range of number of clusters $K = 1, 2, \dots, K_{max}$ evaluated at the corresponding optimal $\underline{\alpha}_K^*$. Finally, the optimal number of clusters $K^* | \underline{\alpha}_K^*$ is determined according to a user-defined set of clustering criteria, such as the increments in the weighted average proportion of explained

[1] We refer the reader to the detailed discussion in Morelli et al., 2025 about the analytics and the interpretation of the weighted average of explained mixed pseudo inertia, in particular to its link with the gain in using a mixed approach compared to the purely non-spatial or purely spatial cases.

inertia induced by the increase in the number of clusters and the Silhouette index (Morelli et al., 2025), the Dunn's index, the C-index, the Calinski-Harabasz's index, and the McClain-Rao's index (Zammarchi & Maranzano, 2024).

In this paper, we consider only the increments in the weighted average proportion of explained inertia as K increases, which can be interpreted as the marginal gain in explained variability attributable to adding one more cluster. This "elbow"-style criterion is commonly used in hierarchical settings and is directly aligned with the mixed-inertia objective optimized by the algorithm (Chavent et al., 2018). By contrast, homogeneity-based indices, such as the silhouette, may be less informative in our context due to the high dimensionality of the time series dissimilarities (multiple variables over a long horizon), which typically yields uniformly modest values and weak contrast across competing partitions. Nevertheless, to verify the empirical validity of the clustering results under different scenarios (i.e., spatiotemporal, temporal-only, and spatial-only), a set of internal validation indices–including the Average Silhouette, the Calinski–Harabasz index, the Within/Between ratio, and Pearson's Gamma–will be employed.

In Algorithm 3 we summarize the spatiotemporal hierarchical clustering methodology.

Algorithm 3 Hierarchical Spatiotemporal Clustering

Define as $D_p = [d_{p,ij}]_{p=1,\dots,P-1;i,j=1,\dots,n}$ the feature dissimilarity matrices
Define as $D_P = [d_{P,ij}]_{i,j=1,\dots,n}$ the spatial dissimilarity matrices
Define the weight of each unit w_i with $i = 1,\dots,n$, with no prior information $w_i = 1/n$
Set as K_{max} the maximum number of clusters
Set as $\Delta\alpha$ the increment of α_p
Compute $\Delta_p = [\delta_{p,ij}]_{p=1,\dots,P;i,j=1,\dots,n}$ as $\delta_{p,ij} = \frac{w_i w_j}{w_i + w_j} d^2_{p,ij}$
for $K = 1, 2, \dots, K_{max}$ **do**
  for $\alpha_p (p = 1,\dots,P) \in [0,1]$ by $\Delta\alpha$ such that $\sum_{p=1}^{P} \alpha_p = 1$ **do**
    Compute the aggregation matrix $\Delta(\underline{\alpha}) = \sum_{p=1}^{P} \alpha_p \Delta_p$
    Compute the partition $\mathcal{P}_K^{\alpha_p}$ in K clusters according to Ward hierarchical algorithm on the combined matrix $\Delta(\underline{\alpha})$, which correspond to apply the Ward algorithm on the combination of inertia
    Compute the weighted average explained inertia $\bar{Q}(\mathcal{P}_K^{\underline{\alpha}})$
  end for
  Select the best $\underline{\alpha}_K^*$ for each K such that $\underline{\alpha}_K^* = argmax_{\alpha_p (p=1,\dots,P)} \bar{Q}(\mathcal{P}_K^{\alpha_p})$
end for
Choose K^* (evaluated at the corresponding $\underline{\alpha}^*$) according to the first difference in the weighted average proportion of explained pseudo inertia.

3.3 The role of the spatial component in hierarchical clustering

This section presents the methodological contribution of the paper. We provide a detailed reasoning on how the spatial component influences cluster formation and propose a measure to quantify its role. The formula of the indicator is introduced and its interpretation discussed. For additional technical and mathematical details, the reader is referred to the Appendix 3.A. Previously, we mentioned that a key advantage of the multi-matrix algorithm is that the choice of the weighting vector $\underline{\alpha}$ does not necessarily require the inclusion of the spatial component but rather incorporates it only when it contributes meaningfully to explaining the overall dissimilarity in the data. However, it is important to note that the weight assigned to the spatial component cannot be interpreted as a direct measure of its relevance. While it may be intuitive to think that as the spatial weight α_P increases, the explained inertia (and thus the proportion and normalization) also increases, it remains unclear how these quantities grow relative to each other. Indeed, if the increase in inertia were proportional to the weights of the matrices, the choice of parameter α considering the case of two dissimilarity matrices, according to Chavent et al., 2018 would always result in $\alpha = 0.5$, and the optimal outcome for Morelli et al., 2025 could be any value for α.

To illustrate this unclear behavior, we provide three simulated numerical examples (or scenarios) in which the spatial component is either (a) completely irrelevant, (b) relevant but not related to the phenomenon, or (3) relevant and useful in describing the phenomenon[2]. Without losing generalities, let us consider the case with only two dissimilarity matrices, the first containing the spatial distances and the second collecting the Euclidean distances across observations for a single feature. In Figure 3.3.1, we provide a graphical representation of the three above scenarios.

In the first case, the clusters present high dissimilarities related to the features, but they perfectly overlap with respect to the spatial coordinates. As the spatial component is not playing any role

[2]Notice that the statement "the spatial component is relevant with respect to the phenomenon" refers to whether the spatial/geographical distribution of the units to be clustered is related to the phenomenon under study and thus can be useful for the purpose of adequately constructing clusters ("it is relevant for the clustering") or is unrelated to the underlying phenomenon and thus does not contribute in a useful way to the construction of clusters ("it is not relevant for the clustering").

Figure 3.3.1: Example of clusters defined under different spatial conditions. In panel (a), the spatial component is not relevant: observations from different clusters present the same spatial distribution, and clusters are perfectly overlapping in space. In panel (b), the spatial component is relevant but not related to the phenomenon: it is possible to identify groups according to spatial distances, but this is not useful in describing the phenomenon because, in each spatial cluster, we can find observations with different features. In panel (c), the spatial component is relevant and related to the phenomenon: we can identify the same clustering partition either using spatial dissimilarities or feature dissimilarities.

in describing the phenomenon and is not useful in maximizing the explained inertia, it will be not included in the final optimal combination of weights. In the second scenario, the partition obtained using only the features dissimilarity matrix is completely different from the one obtained using only the spatial dissimilarities. Here, the spatial component leads to the identification of groups that are well separated in space but are not related to the actual pattern of the phenomenon under study. Therefore, we expect the algorithm to return as optimal mixing parameter either $\alpha = 0$ or $\alpha = 1$, depending on which matrix explains a greater portion of the dissimilarity across clusters. In the third scenario, the cluster partition obtained from observing only the features is identical to the one obtained using only the spatial component. Thus, the geographical distribution is strongly related to the phenomenon of interest and will play a crucial role in building the clusters. In this latter case, the spatial component's role is clearly significant, yet any value of α would suffice to maximize the weighted average explained inertia because if the clustering algorithm on D_0 leads to the same partition obtained using only D_1, any combination of D_0 and D_1 will lead to the same results.

The above simulated example shows that the weight of the spatial component, represented by α, cannot be interpreted as a measure of the actual importance of geographical information for identifying the groups of observations that characterize the patterns of the examined variables. Indeed, it is only the solution obtained from a maximization task. On the one hand, when the spatial component is included in the combined dissimilarity matrix D, thus $1 - \alpha > 0$, or $\alpha_P > 0$ in multidimensional case, this indicates that spatial information is at least partially relevant in explaining the phenomenon. On the other hand, when the spatial information became more and more related to the dissimilarities across features, that is, D_0 and D_1 lead to similar clustering partitions, it is no longer necessary to include the spatial component because feature dissimilarities are able to capture a high proportion the inertia in the spatial dissimilarities and vice versa.

To properly understand the role of the spatial dissimilarity matrix in relation to the phenomenon under study, it is more appropriate to examine how the explained inertia in the resulting clusters changes when combining the two dissimilarity matrices with respect to the cases in which the features or spatial distances are used independently. The normalized proportion of explained inertia defined by Chavent et al., 2018 (Equation 3.2.6) is particularly useful for this purpose because it provides the percentage of explained variability in the clustering partition obtained from the combined dissimilarity matrix with respect to the case in which the partition is obtained from one

dissimilarity matrix only.

When we choose to incorporate the spatial component, we effectively renounce to a proportion of the inertia explained by the features to gain a certain amount of inertia explained within the spatial component. To clarify the mechanism, we revisit the framework of Chavent et al., 2018 describing how the proportion of explained inertia changes when the inertia from two dissimilarity matrices is combined. Let $Q_{D_0}(\mathcal{P}_K^\alpha)$ and $Q_{D_1}(\mathcal{P}_K^\alpha)$ (see 3.2.5) denote the proportion of explained inertia from D_0 and D_1, respectively, that are captured by the partition $\mathcal{P}_K^\alpha$. These quantities are then normalized relative to the proportions obtained when clustering with D_0 alone and with D_1 alone, thus we obtain the normalized proportion of explained inertia $\tilde{Q}_{D_0}(\mathcal{P}_K^\alpha)$ and $\tilde{Q}_{D_1}(\mathcal{P}_K^\alpha)$ in 3.2.6. If we consider the limit case in which $\alpha = 0$ and thus $\tilde{Q}_{D_0}(\mathcal{P}_K^\alpha) = 1$, it is intuitive that the proportion of explained inertia in D_1 will reach its minimum value with respect to all other possible values of α. When combining them using $0 < \alpha < 1$ the value of the proportion of explained inertia in D_0 will decrease (or remain constant) up to its minimum, while the proportion of explained inertia in D_1 will increase (or remain constant), up to one. Both $Q_{D_0}(\mathcal{P}_K^\alpha)$ and $Q_{D_1}(\mathcal{P}_K^\alpha)$ are non strictly monotonic with respect to α, as a consequence $\tilde{Q}_{D_0}(\mathcal{P}_K^\alpha)$ and $\tilde{Q}_{D_1}(\mathcal{P}_K^\alpha)$ are not strictly monotonic as well. However, the increase or decrease in the explained inertia with respect to α is non-constant and does not depend on α itself. Indeed including D_1 using a certain α may imply a big or small loss in the proportion of explained inertia in D_0 or it can remain unchanged and, at the same time, it may imply an high increase or zero increase in the proportion of explained inertia in D_1. Moreover, the magnitude of these variations are not related to each other; in other words a big loss (or zero loss) in $Q_{D_0}(\mathcal{P}_K^\alpha)$ is not necessary linked to big or small increase in $Q_{D_1}(\mathcal{P}_K^\alpha)$. Overall, when we take into account an additional dissimilarity matrix in a clustering partition, we are willing to lose some proportion of explained inertia from the initial matrix, hoping to gain some from the matrix we add. If the loss is smaller than the gain — that is, the decrease in one component is less than the increase in the other — then combining the two matrices is preferable.

From the above intuition, it follows that the actual contribution of the geographical information can be well approximated by the difference between the normalized proportion of inertia explained by the spatial component within the mixed cluster partition and the reduction in the normalized proportion of inertia explained by the features that results from including the spatial component. Specifically, in the case where there are only two matrices D_0 and D_1, we quantify the importance

of spatial features through a novel indicator, namely the Joint Inertia of partition $\mathcal{P}_K^{\alpha}$, which is computed as

$$JI(\mathcal{P}_K^{\alpha}) = \tilde{Q}_{D_1}(\mathcal{P}_K^{\alpha}) - (1 - \tilde{Q}_{D_0}(\mathcal{P}_K^{\alpha})) = \frac{Q_{D_1}(\mathcal{P}_K^{\alpha})}{Q_{D_1}(\mathcal{P}_K^{1})} - \left(1 - \frac{Q_{D_0}(\mathcal{P}_K^{\alpha})}{Q_{D_0}(\mathcal{P}_K^{0})}\right). \tag{3.3.1}$$

This criterion is referred to as Joint Inertia because it captures the overall gain in explained inertia that arises when the spatial and feature dissimilarity matrices are jointly incorporated in the clustering process. Rather than reflecting the contribution of each matrix in isolation, the indicator quantifies how much additional explanatory power is achieved when both sources of information are combined. In this sense, it measures the synergy between spatial and non-spatial components, emphasizing that their joint use can uncover patterns that would remain hidden if either matrix were considered separately.

The generalization to the case of multiple dissimilarity matrices, one of which relates to the spatial component, is straightforward. Recalling the notation previously introduced, let us consider the generic partition induced by a given combination of mixing parameters (i.e., $\mathcal{P}_K^{\underline{\alpha}}$) and let $D_{-P} = [d_{p,ij}]_{p=1,\dots,P-1;i,j=1,\dots,n}$ be the dissimilarity matrices containing the distances across the $P-1$ time series for n units and let $D_P = [d_{P,ij}]_{i,j=1,\dots,n}$ be the spatial dissimilarity matrix across the n units, which may be computed as geodetic distances between coordinates. Let $\underline{\alpha} = [\alpha_p]_{p=1,2,\dots,P} \in \mathbf{\Delta}^{P-1}$ be the vector of mixing weights used to compute the linear combination of the mixed within pseudo inertia from each dissimilarity matrix and let $\mathcal{P}_K^{\underline{\alpha}}$ be the resulting clustering partition. Moreover, let use define $\underline{\alpha}' = [\alpha'_p]_{p=1,2,\dots,P-1} \in \mathbf{\Delta}^{P-2}$ being the optimal vector of weights resulting from the combination of the pseudo inertia from the non-spatial dissimilarity matrices (i.e. excluding matrix P), leading to partition $\mathcal{P}_K^{\underline{\alpha}'}$. Similarly, define the partition obtained by using only the spatial information (i.e., by fixing $\alpha_P = 1$) as $\mathcal{P}_K^{\alpha_P=1}$. If we compute the optimal weighting vector obtained by the combination of the inertia from non-spatial dissimilarity matrices (i.e., $\underline{\alpha}'|\alpha_P = 0$), the resulting partition will explain as much as possible the dissimilarities in non-spatial matrices, indeed the vector of weights obtained is the one that maximizes the weighted average explained inertia thus being $\bar{Q}_{D_{-P}}(\mathcal{P}_K^{\underline{\alpha}^*|\alpha_P=0})$ greater than any other potential $\bar{Q}_{D_{-P}}(\mathcal{P}_K)$ evaluated at $\underline{\alpha}|\alpha_P = 0$.

Considering a generic combination of $\underline{\alpha}$ and K, the multi-matrix Joint Inertia for the spatial dissimilarity matrix D_P can be computed as

$$JI_P(\mathcal{P}_K^{\alpha}) = \tilde{Q}_{D_P}(\mathcal{P}_K^{\alpha}) - (1 - \tilde{Q}_{D_{-P}}(\mathcal{P}_K^{\alpha})) = \frac{Q_{D_P}(\mathcal{P}_K^{\alpha})}{Q_{D_P}(\mathcal{P}_K^{\alpha_P=1})} - \left(1 - \frac{\bar{Q}_{D_{-P}}(\mathcal{P}_K^{\alpha})}{\bar{Q}_{D_{-P}}(\mathcal{P}_K^{\alpha^*|\alpha_P=0})}\right) \tag{3.3.2}$$

where $\tilde{Q}_{D_P}(\mathcal{P}_K^{\alpha})$ is the normalized proportion of explained inertia with respect to the purely-spatial setting and $\tilde{Q}_{D_{-P}}(\mathcal{P}_K^{\alpha})$ is the normalized proportion of explained inertia with respect to the optimal non-spatial setting. Specifically it corresponds to the ratio between the weighted average proportion of inertia in the non-spatial setting, when considering the partition obtained combining the inertia from all the dissimilarity matrices $\bar{Q}_{D_{-P}}(\mathcal{P}_K^{\alpha})$ and when considering the optimal partition from the non spatial dissimilarities $\bar{Q}_{D_{-P}}(\mathcal{P}_K^{\alpha^*|\alpha_P=0})$.

As the normalized proportion of explained inertia for every dissimilarity matrix varies between 0 and 1, also the Joint Inertia $JI_P(\mathcal{P}_K^{\alpha})$ will range from 0 to 1 (i.e., $0 \leq JI_P(\mathcal{P}_K^{\alpha}) \leq 1$), being close to zero when the loss in the normalized proportion of explained inertia from the features dissimilarity due to the inclusion of the spatial component is similar to the normalized proportion of explained inertia in the spatial dissimilarity in the combined partition. Conversely, it will be close to 1 when the clustering partition obtained combining features and spatial dissimilarities allows to explain a high normalized proportion of inertia, both in spatial and non-spatial dissimilarity matrices.

We stress that the Joint Inertia can be utilized to evaluate the contribution of any dissimilarity matrix relative to all other dissimilarity matrices. In this study the Joint Inertia is specifically introduced to assess the role of the spatial component, addressing our initial research objectives. However, in other contexts, the Joint Inertia of any dissimilarity matrix can be computed to gain a comprehensive understanding of which variables exhibit shared explanatory power in terms of their dissimilarities. Further technical and mathematical details about the Join Inertia, both in the case of two matrices and when considering more than two matrices, are discussed in the Appendix 3.A.

Considering the three simulated scenarios described above, in Table 3.3.1 we provide the values of the normalized proportion of explained inertia in D_0 and D_1 and the Joint Inertia obtained for every case represented in Figure 3.3.1. When the spatial component is not related to the clustering partition obtained by features (e.g., scenario b), we obtain α being either 0 or 1. Thus, the normalized proportion of explained inertia in the features dissimilarity matrix (or in the spatial

dissimilarity matrix) will be equal to 1, while the normalized proportion of explained inertia in the spatial dissimilarity matrix (in the feature dissimilarity matrix) will be close to zero but strictly positive. The Joint Inertia index will be close to zero in cases (a) and (b), which is coherent with the fact that the spatial component plays no role in identifying the actual clustering patterns (conversely, well identified by the features space), while it is equal to 1 in case (c) because the spatial information is strongly related to the underlying phenomenon and exactly describes the groups structure[3]. When $0 < \alpha < 1$, the dissimilarity matrices are combined because their combination allows us to capture an overall higher proportion of the explained inertia, thus we expect to obtain an higher value of the presented indicator. In the following Section 3.4 we discuss the connection between the Join Inertia and the spatial information embedded in the data. In particular, we focus on how the index can be used to assess the actual relevance of the geographical dissimilarity in computing the clustering partitions.

Case (a)	**Case (b)**		**Case (c)**
Spatial component is not relevant	Spatial component is relevant but not related to the features		Spatial component is helpful
$\alpha = 0$	if $\alpha = 0$	if $\alpha = 1$	$0 < \alpha < 1$
$\tilde{Q}_{D_1}(\mathcal{P}_K^\alpha) \approx 0$	$\tilde{Q}_{D_1}(\mathcal{P}_K^\alpha) \approx 0$	$\tilde{Q}_{D_1}(\mathcal{P}_K^\alpha) = 1$	$\tilde{Q}_{D_1}(\mathcal{P}_K^\alpha) = 1$
$\tilde{Q}_{D_0}(\mathcal{P}_K^\alpha) = 1$	$\tilde{Q}_{D_0}(\mathcal{P}_K^\alpha) = 1$	$\tilde{Q}_{D_0}(\mathcal{P}_K^\alpha) \approx 0$	$\tilde{Q}_{D_0}(\mathcal{P}_K^\alpha) = 1$
$JI \approx 0$	$JI \approx 0$	$JI \approx 0$	$JI = 1$

Table 3.3.1: Choice of α, normalized proportion of explained inertia and resulting Joint Inertia in different examples of clusters under different spatial conditions. Cases are represented in Figure 3.3.1.

3.4 Simulation experiment

In this section, we present a simulation experiment that compares two rules for selecting the mixing

[3]It should be noted that the case $JI = 1$ holds when the geographical information and the feature space describe the same group structure while being strongly related to the phenomenon under study. This happens when space and features embed the same amount of information associated with the phenomenon and resulting in the same clustering partition. So, in a sense, the geography jointly determines the features and the clusters. However, geography compresses the relevant information using only two dimensions, whereas it is reasonable to expect the same degree of information to be achieved by a much larger number of features.

parameter α—the trade-off criterion of Chavent et al., 2018 and the weighted-average explained inertia maximization of Morelli et al., 2025—while using the same Ward-like hierarchical procedure throughout. We assess their performance under different scenarios for the true contribution of geographical information. In particular, Chavent et al., 2018 propose a practical (though not globally optimal) solution: for a given K, explore $\alpha \in [0, 1]$ and choose the value that best balances the loss of socio-economic and geographical homogeneity, using a plot of the proportions of explained inertia in D_0 and D_1 over the α-grid. To automate this procedure, one can select the α for which these two normalized proportions are as close as possible, i.e., the difference between them is minimized. Morelli et al., 2025 suggest to choose α such that it maximizes the overall explained inertia quantified through the weighted average explained inertia in Equation 3.2.7. We argue that the latter criterion is preferable because selecting α to maximize the total (weighted) explained inertia the procedure adapts to the actual signal present in the dataset, rather than enforcing an a priori balance between spatial and attribute contributions.

Moreover, we provide insights into the role of the spatial component by using an overlapping parameter that controls the distance across the means of the spatial distribution of the true clusters. Under the simulation design, we compare the algorithms in terms of accuracy, precision, sensitivity (Naidu et al., 2023), and the adjusted rand index (Gordon, 1999) for multi-class data.

Using a Monte Carlo simulation scheme with $N = 500$ repetitions, we simulate samples of size $n = 100$ observations from a mixture of three-dimensional Gaussian distribution in which the first dimension refers to a unique non-spatial feature and the other two dimensions refer to the spatial coordinates. We simulate data clustered into $K = 4$ different groups. We indicate with $\pi_1, .., \pi_k$ the probability to sample an observation belonging to each cluster $k = 1, \dots, 4$ and, according to the cluster, we simulate the variable Z and the spatial coordinates (X, Y). Specifically, we define the distribution of the variable as $Z_k \sim N(\mu_k, \sigma)$, where $\underline{\mu} = (\mu_1, \dots, \mu_4)$ and $\sigma = v$, and the spatial component $(X, Y)_k \sim N(\mu_{sp,k}, \Sigma_{sp})$. The means of the group-specific spatial coordinates are defined according to an overlapping parameter, namely d, which can vary between 0 and 1, and it determines the proximity across the spatial means of the clusters. In particular, $\underline{\mu}_{sp} = (\mu_{sp,1}, \mu_{sp,2}, \mu_{sp,3}, \mu_{sp,4}) = ((d, d), (-d, d), (d, -d), (-d, -d))$ while $\Sigma_{sp} = v_{sp}\mathbb{I}_2$. Notice that, given this simulation setting, as the overlapping parameter grows from 0 to 1, the centroids of the groups tend to diverge, thus generating well-distanced clusters and enhancing the relevance of the geographical information in

clustering the units.

To exemplify the contribution of the overlapping parameter d, in Figure 3.4.1, we depict some examples of the clustering partition that can be obtained for different values of the overlapping parameter and distinguishing among Chavent et al., 2018 (left column) and Morelli et al., 2025 (right column) criteria. The plots are simulated by setting $\pi_k = 1/K$, $\mu_k = (2, 4, 6, 8)$, $v = 0.4$, $v_{sp} = 0.4$, and considering the following set of values for the overlapping parameter $d \in \{0, 1/3, 2/3, 1\}$ (i.e., sorted top-down).

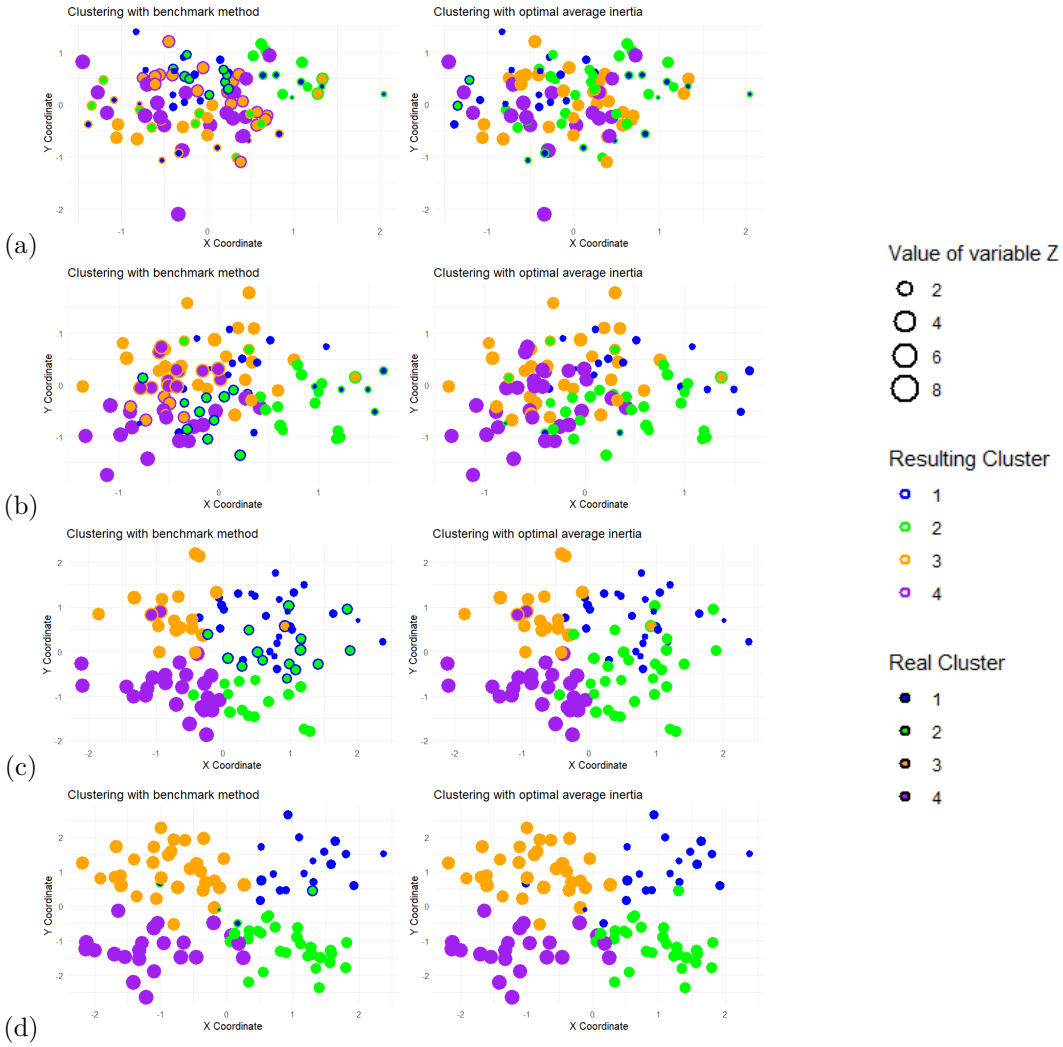


Figure 3.4.1: Clustering partition in four simulations with different overlapping parameters $d \in \{0, 1/3, 2/3, 1\}$ respectively in panel (a), (b), (c), and (d). The color of the points refers to the true cluster from which the observation is drawn, while the color of the border refers to the resulting clustering partition obtained by applying the benchmark algorithm Chavent et al., 2018 (left column) and the optimal solution for choosing α in Morelli et al., 2025 (right column).

From the figure, it is possible to observe that the maximum average criterion is able to detect a clustering partition similar to the real one, even when the spatial component is not relevant or less relevant, that is, in panel a and b. When the distance in the spatial location of the cluster is higher (e.g., in panel c and d), clustering performance improves using both methodologies. Re-

sults reported in Table 3.4.1 confirm the previous statements. Indeed, while Chavent et al., 2018 methodology allows identifying a reasonable clustering partition, the Morelli et al., 2025 methodology demonstrates improved clustering performance in terms of accuracy and adjusted rand index. This enhancement is particularly notable in scenarios where the spatial component plays a limited role, such as when clusters fully or partially overlap in space.

	Benchmark			Optimal			
d	α	Accuracy	Adj. Rand Index	α	Accuracy	Adj. Rand Index	Joint Inertia
0	0.75	0.43	0.27	0.35	0.87	0.74	0.17
1/3	0.65	0.63	0.34	0.50	0.93	0.83	0.49
2/3	0.85	0.82	0.61	0.45	0.97	0.93	0.86
1	0.75	0.96	0.91	0	1.00	1.00	0.93

Table 3.4.1: Comparison of the performance resulting from clustering partition according to the benchmark procedure from Chavent et al., 2018 and the optimal solution in Morelli et al., 2025, with Accuracy and Adjusted Rand Index and the resulting Joint Inertia to assess the role of the spatial component in the partition.

Moreover, we confirm that the parameter α can not be interpreted as a measure of assessment of the role of the spatial component as the optimized values do not reflect the degree of overlapping of the data-generating process. On the other hand, the Joint Inertia index appears effective in assessing the relevance of the spatial component in the resulting cluster distribution.

In Algorithm 4, we provide a scheme of the simulation, from the data generating process, to the choice of α according to Chavent et al., 2018 and Morelli et al., 2025, and the computation of the validity measures and the Joint Inertia.

Below, we present the results from the Monte Carlo experiment and random assignment for parameters $\pi_k \sim U(0.15, 0.35)$, $v \sim U(0.1, 0.6)$, $v_{sp} = 0.4$ and $d \sim U(0, 1)$. From Table 3.4.2, we can clearly observe that maximum average criterion by Morelli et al., 2025 is able to find clustering partitions that are very similar to the true partitions and performs remarkably better in identifying the actual mixing parameter α with respect to the Chavent et al., 2018 criterion.

In Figure 3.4.2 we provide a comparison of the resulting clustering performance with respect to the overlapping parameter d used in the simulation. It is evident that Morelli et al., 2025 algorithm performs generally better, in particular when the overlapping parameter is low and clusters

Algorithm 4 Simulation of Spatial Hierarchical Clustering

Simulate data from $K = 4$, with probabilities $\pi_1, .., \pi_4$, with variable $Z_k \sim N(\mu_k, \sigma)$ and spatial coordinates $(X, Y)_k \sim N(\mu_{sp,k}, \Sigma_{sp})$.
Set the weight of each unit $w_i = 1/n$
Compute $D_0 = [d_{0,ij}]_{i,j=1,...,n}$ the feature dissimilarity matrix as Euclidean distance across values of the variable.
Compute $D_1 = [d_{1,ij}]_{i,j=1,...,n}$ the spatial dissimilarity matrix as distance across the coordinates.
Compute $\Delta_0 = [\delta_{0,ij}]_{i,j=1,...,n}$ and $\Delta_1 = [\delta_{1,ij}]_{i,j=1,...,n}$ as $\delta_{ij} = \frac{w_i w_j}{wi+wj} d_{ij}^2$
Define as $\Delta\alpha$ the increment of α
for $\alpha \in [0, 1]$, by $\Delta\alpha$ **do**
Compute the aggregation matrix $\Delta(\alpha) = (1 - \alpha)\Delta_0 + \alpha\Delta_1$
Find the $\mathcal{P}_{K=4}^{\alpha}$ partition using Ward algorithm on the combined matrix $\Delta(\alpha)$, which correspond to apply the Ward algorithm on the combination of inertia
Compute the proportion of explained inertia $Q_{D_0,\alpha}$ and $Q_{D_1,\alpha}$
Compute the weighted average of the explained mixed pseudo inertia $\bar{Q}(\mathcal{P}_K^{\alpha})$
end for
Select $\alpha_K^c = argmin_{\alpha}(abs(Q_{D_0,\alpha} - Q_{D_1,\alpha}))$
Select $\alpha_K^* = argmax_{\alpha}\bar{Q}(\mathcal{P}_K^{\alpha})$
Compute Accuracy, Precision Sensitivity and Adjusted Rand Index, for $\mathcal{P}_{K=4}^{\alpha^c}$ and $\mathcal{P}_{K=4}^{\alpha^*}$
Compute Joint Inertia according to equation 3.3.1.

	Accuracy		**Precision**		**Sensitivity**		**Adj. Rand Index**	
	Bench	Optim	Bench	Optim	Bench	Optim	Bench	Optim
Mean	0.74	0.85	0.75	0.85	0.74	0.84	0.55	0.71
Median	0.76	0.91	0.78	0.91	0.76	0.90	0.52	0.77
SD	0.16	0.14	0.17	0.15	0.16	0.15	0.23	0.22

Table 3.4.2: Comparison of the performance resulting from clustering partition according to the benchmark procedure (columns Bench) from Chavent et al., 2018 and the optimal solution (Columns Optim) in Morelli et al., 2025, using a Monte Carlo simulation, in terms of Accuracy, Precision, Sensitivity ad Adjusted Rand Index.

are partially overlapping in space. As expected, when the overlapping parameter d is close to 1, both algorithms are able to find exactly the true partition because, as explained in the previous section, when the feature dissimilarity and the spatial dissimilarity lead to the same partition, any combination of them will lead to it.

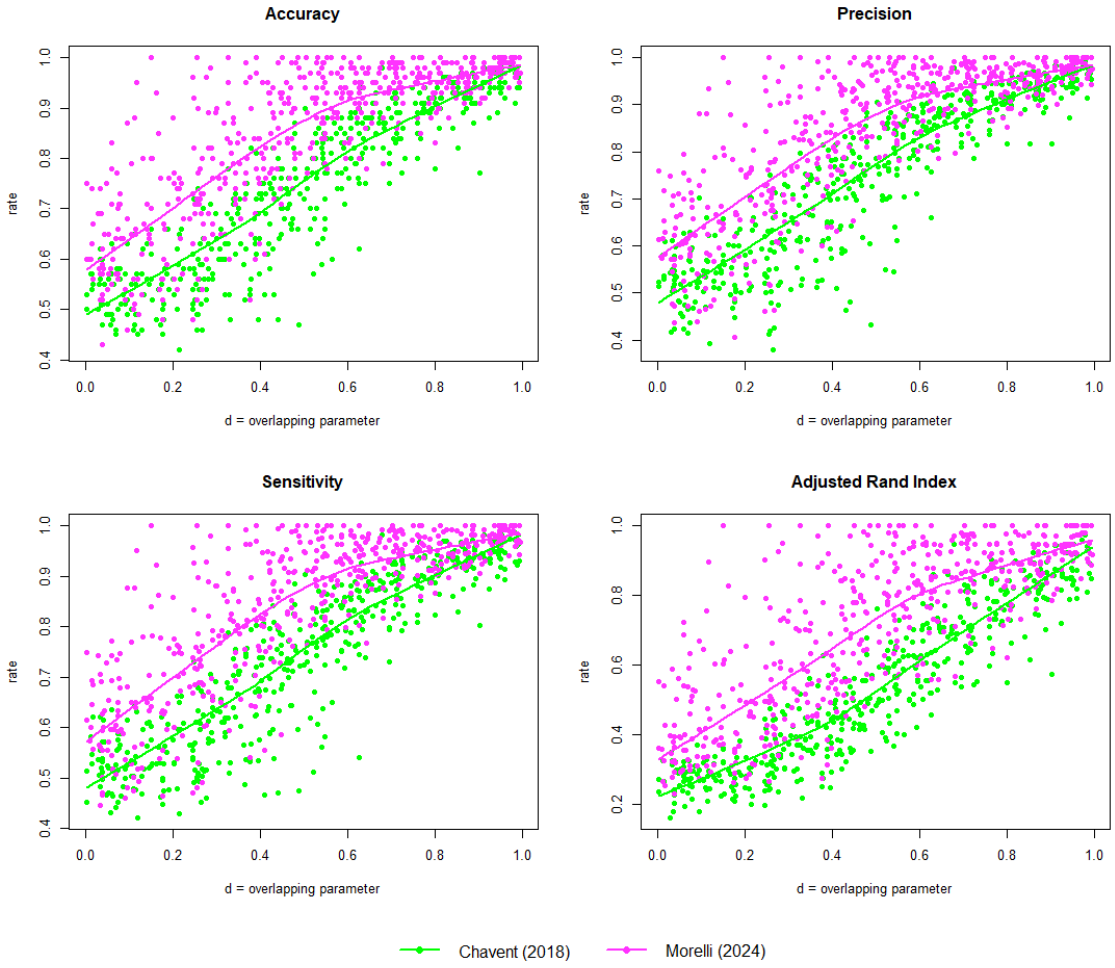


Figure 3.4.2: Clustering performance comparison between Chavent et al., 2018 and Morelli et al., 2025 in terms of Accuracy, Sensitivity, Precision, and Adjusted Rand Index, with respect to the overlapping parameter d.

In Figure 3.4.3, we provide several insights into the relationship between the spatial distance among centroids, determined by the overlapping parameter d, the mixing parameter α, and the Joint Inertia index. In the left panel, as the overlapping parameter grows (from left to right), the mixing parameter selected by both methods seems to be unaffected, confirming that the weight assigned to the spatial dissimilarity matrix does not provide information about the actual contribution of the spatial component in determining the clusters. Conversely, the right panel remarkably supports

the use of the Joint Inertia index in quantifying the importance of the spatial component. Indeed, a strong and positive correlation between the overlapping parameter d and the Joint Inertia is detected, both using $\alpha_{chavent}$ and $\alpha_{morelli}$. The positive correlation is definitely consistent with the interpretation of the overlapping parameter d given above, that is, as d grows from 0 to 1, the distance between the centroids of the groups tends to increase, increasing the informative power induced by the spatial/geographical compared to the non-spatial distance, exactly as described by the Joint Inertia.

Figure 3.4.3: In panel (a), on the left, we can observe the resulting α according to Chavent et al., 2018 and Morelli et al., 2025 when considering clusters with different overlapping degrees (i.e., when the overlapping parameter d varies between 0 and 1). In panel (b), on the right, we report the resulting Joint Inertia obtained in the simulation experiment with respect to the overlapping parameter d.

3.5 Dynamic and Regional Clustering of GHGs Emissions

In this section, we present an application of spatiotemporal clustering on environmental data concerning the annual GHGs emissions at the regional level for Europe. Our aim is to explore regional patterns in the environmental impact caused by human activities on air quality, which may also de-

pend on the geographical and morphological characteristics of the area as well as on the productive specialization of the regions.

We analyze the annual average emissions per km^2 of several GHGs, that is, methane, nitrous oxide, fluorinated gases (F-gases), and carbon dioxide produced by the main economic sectors, namely agriculture, buildings, energy, industry, transport, and waste. Data are measured in tons per km^2 of CO_2 equivalent.

Greenhouse gases in Europe originate from different activities, each of them contributing to climate change in unique ways (Crippa et al., 2024). Methane (CH_4) emissions primarily arise from agriculture, especially livestock farming and waste management, as well as fossil fuel extraction and processing. Nitrous oxide (N_2O) is largely emitted through agricultural practices, such as the use of nitrogen-based fertilizers, with smaller contributions from industrial processes and fossil fuel combustion. F-gases, synthetic in nature, are used in refrigeration, air conditioning, and industrial processes like aluminum production. Carbon dioxide (CO_2), the most abundant GHGs, is predominantly released through fossil fuel combustion for energy, transportation, and industrial activities, as well as deforestation and cement production.

We believe that the identification of patterns and dynamics in regional GHGs emissions within a spatiotemporal framework is a valuable contribution that extends the current knowledge on the status of environmental degradation in Europe, enhancing the awareness of citizens and institutions regarding the policies to be planned to contrast climate change and global warming.

All the analysis have been conducted using the `SCDA` package written in the `R` language currently available on the CRAN[4]

3.5.1 Dataset description

The data of interest relate to atmospheric emissions of sector-specific greenhouse gases measured at the regional level provided by the Emissions Database for Global Atmospheric Research (EDGAR) database (Crippa et al., 2024) within the Annual Regional Database of the European Commission

[4]https://cran.r-project.org/web/packages/SCDA/index.html (Accessed on October 9^{th}, 2025).

(ARDECO) (Bucciarelli, 2024) project.

The regional emission data are taken from EDGAR v8.0 and are modelled estimates produced with a technology-based emission-factor approach under an IPCC-consistent methodology. National and sectoral totals are downscaled to a $0.1°\times0.1°$ grid using spatial proxies (e.g. facility locations, road networks, population) and then aggregated to NUTS2 polygons (NUTS 2021). Emissions are reported in kton CO_2-eq (GWP-100, IPCC AR5). Further methodological details and discussion of assumptions and uncertainty are given in Crippa et al., 2024. In particular, we consider yearly measurements of GHGs emissions from 1990 to 2022 (i.e., T = 33 time stamps) for the N = 234 regions[5] in the European Union, in 27 countries, identified according to the Nomenclature of territorial units for statistics - Level 2 (NUTS-2) classification (European Commission, 2016). We also collect the area of NUTS-2 regions in 2022 from Eurostat [6] dataset in order to obtain the annual emissions in tons per km^2 of CO_2 equivalent, by gases and by sector.

In Figure 3.5.1 we provide descriptive statistics about our dataset, showing the long-run emission intensity with its evolution: the left panel highlights persistent high-intensity hotspots concentrated in major metropolitan/industrial corridors, whereas the right panel shows broad declines alongside localized increases from 1990 to 2022, underscoring marked spatial heterogeneity in both levels and trends.

The dataset contains annual time series of GHG emissions for four gases—CH_4, F-gases, fossil-CO_2, and N_2O—and for six emitting sectors: Agriculture, Buildings, Energy, Industry, Transport, and Waste. Averaged across all years in the sample, the global Moran's I is approximately 0.28 for CH_4, 0.13 for F-gases, 0.17 for fossil-CO_2, and 0.24 for N_2O. By sector, the mean Moran's I is highest for Agriculture and Transport which are respectively 0.48 and 0.44, followed by Industry (≈ 0.17), Waste and Buildings with the lowest annual average Moran's I, respectively 0.09 and 0.08. Complete year-by-year values and plots are reported in Appendix 3.B.

We run the spatiotemporal cluster analysis twice, considering two different datasets. In the first

[5]Notice that we excluded regions classified as extra-continental territories and those located in non-European continents, ensuring that the extremely high distances involving these regions do not render the spatial dissimilarities among the remaining regions insignificant.

[6]https://ec.europa.eu/eurostat/databrowser/view/REG_AREA3/default/table (Accessed on October 9^{th}, 2025)

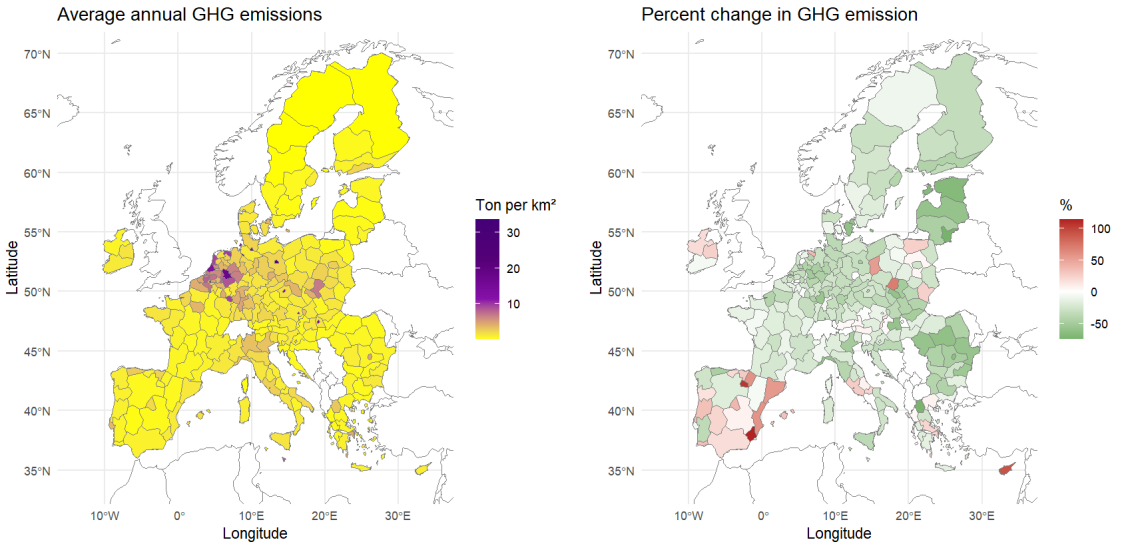


Figure 3.5.1: Descriptive statistics of GHG emission in each region at NUTS2 level: annual average tons CO_2 equivalent per km^2 between 1990-2022 (on the left) and percent change from 1990 to 2022 (on the right).

case, we implement the clustering by aggregating GHGs emissions according to the type of gas (i.e., by summing the emissions at the sectoral level for each gas). In the second case, we aggregate the emissions with respect to the sectors, thus by summing the gas-specific emissions for each sector. In both cases, we utilize the Dynamic Time Warping (DTW) strategy as the chosen distance metric. The Dynamic Programming approach using a warping function has been introduced by Sakoe and Chiba, 1971, 1978 in the spoken word recognition field as a time-normalization algorithm. Berndt and Clifford, 1994 have implemented DTW distance to detect patterns in time series. We compute the spatial dissimilarity matrix as the geodetic distance across the coordinates of the centroids of each region.

3.5.2 Spatiotemporal clusters for emissions by gas

We run the algorithm described in Morelli et al., 2025 setting $\Delta\alpha = 0.05$, for $K = 2, \ldots, 10$, combining the inertia from $P = 5$ dissimilarity matrices referring to the CH_4, F-Gases, Fossil-CO_2, N_2O emission time series and the spatial component. We select $K^* = 4$ according to the elbow rule,

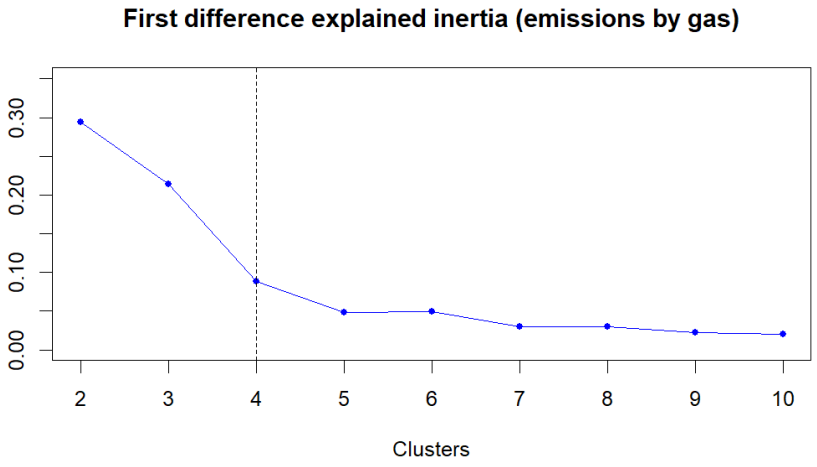


Figure 3.5.2: Increment in the weighted average proportion of explained inertia generated by a unitary increase in the number of clusters. Recall that, for each K, we considered the best combination of inertia of the dissimilarity matrices according to the α_K^* found using the maximization inertia criterion.

observing the weighted average proportion of explained inertia, which can be interpreted as the gain in explained variability due to an additional cluster. In Figure 3.5.2 we represent the first difference in the weighted average proportion of explained inertia, highlighting our choice of $K^* = 4$.

Consequently, we selected $\alpha_{4,p}^* = (0.20, 0.20, 0.15, 0.5, 0.40)$ as the optimal weighting vector for the inertia from each dissimilarity matrix. Table 3.5.1 provides a summary of the weights α_p^*, the explained inertia (both proportion and normalized), and the Joint Inertia associated with the resulting clustering partition for each dissimilarity matrix included in the algorithm.

From Table 3.5.1, it is possible to observe that the partition obtained with $K^* = 4$ clusters allows us to explain almost 40% of the DTW dissimilarities across N_2O time series and about 60% of the dissimilarities across time series of CH_4, F-gases, CO_2 and the spatial dissimilarity. The latter has a Joint Inertia of 0.64, which allows us to conclude that spatial information is crucial when determining the clusters composition, and therefore, the spatial distance is useful in the identification of patterns to explain well the data under consideration. In particular, through the resulting clustering partition $\mathcal{P}_K^{\alpha^*}$ we are able to incorporate 82% of the explained inertia in the

spatial component, relative to the case in which we consider only D_{sp} and, considering the non-spatial dissimilarities, we are still able to include 81% of the weighted average explained inertia with respect to the partition obtained excluding the spatial component. All the dissimilarity matrices have a weight greater than zero and they all seem to be relevant in the composition of the clusters, in fact they have a rather high Joint Inertia. It is possible to note that the weights assigned to the dissimilarity matrices do not always reflect the relevance of the matrix itself in relation to the clustering partition obtained, for example the matrix $D_{\text{N}_2\text{O}}$ is assigned a weight of 0.15, greater than the 0.05 assigned to matrices D_{CO_2}, yet the proportion of inertia explained in the first,which is 0.46, is less than the inertia explained in the other 0.66.

	D_{CH_4}	$D_{\text{F-gases}}$	D_{CO_2}	$D_{\text{N}_2\text{O}}$	D_{sp}
α_p^*	0.20	0.20	0.05	0.15	0.40
$\mathcal{Q}_{D_p}(\mathcal{P}_K^{\underline{\alpha}})$	0.67	0.68	0.63	0.38	0.59
$\tilde{\mathcal{Q}}_{D_p}(\mathcal{P}_K^{\underline{\alpha}})$	0.72	0.80	0.66	0.47	0.82
$\tilde{\mathcal{Q}}_{D_{-p}}(\mathcal{P}_K^{\underline{\alpha}})$	0.97	0.96	1.00	0.99	0.81
$JI_p(\mathcal{P}_K^{\underline{\alpha}^*})$	0.69	0.76	0.66	0.46	0.64

Table 3.5.1: Summary of the $\underline{\alpha}^*$ weights and the inertia (proportion and normalized) returned by the spatiotemporal clustering of emissions by gases at the optimal solution $K^* = 4$. $\mathcal{Q}_{D_p}(\mathcal{P}_K^{\underline{\alpha}})$ is the proportion of explained inertia in D_p; $\tilde{\mathcal{Q}}_{D_p}(\mathcal{P}_K^{\underline{\alpha}})$ is the normalized proportion of inertia in D_p; $\tilde{\mathcal{Q}}_{D_{-p}}(\mathcal{P}_K^{\underline{\alpha}})$ is the normalized proportion of weighted average inertia excluding D_p; and $JI_p(\mathcal{P}_K^{\underline{\alpha}^*})$ is the Joint Inertia relative to D_p.

In Figure 3.5.3 we present a map of the regions belonging to different clusters and in Figure 3.5.4 we report the average annual emission in each cluster by gas.

Cluster 1 includes regions located in East Europe, from the Scandinavian peninsula to the Balkan countries, the central and southern areas of Italy and the eastern area of Austria. It shows the lowest level of CH_4, F-gases, CO_2 and N_2O emissions compared to the other clusters and it appears to be quite stable over time, except for F-gases emissions that are slowly increasing.

Regions of the Netherlands, Malta, and several metropolitan areas in East-Central Europe belong to Cluster 2, which is characterized by emission levels that are remarkably higher—by a significant margin—than those observed in any other cluster. As for the dynamics of CH_4 emissions,

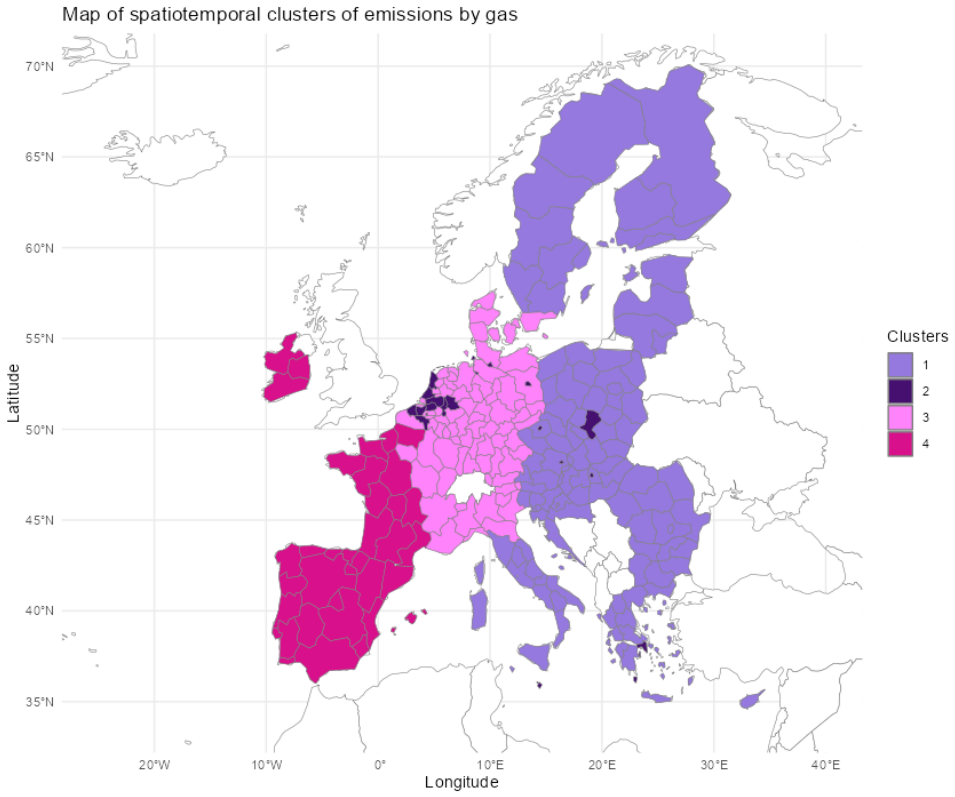


Figure 3.5.3: Map of spatiotemporal clusters obtained combining the inertia from the dissimilarity matrices of the time series of emissions of different gases between 1990 and 2022, that is D_{CH_4}, $D_{\text{F-gases}}$, D_{CO_2} and $D_{\mathrm{N}_2\mathrm{O}}$, and the geodetic distances across centroids of NUTS2 level regions D_{sp}.

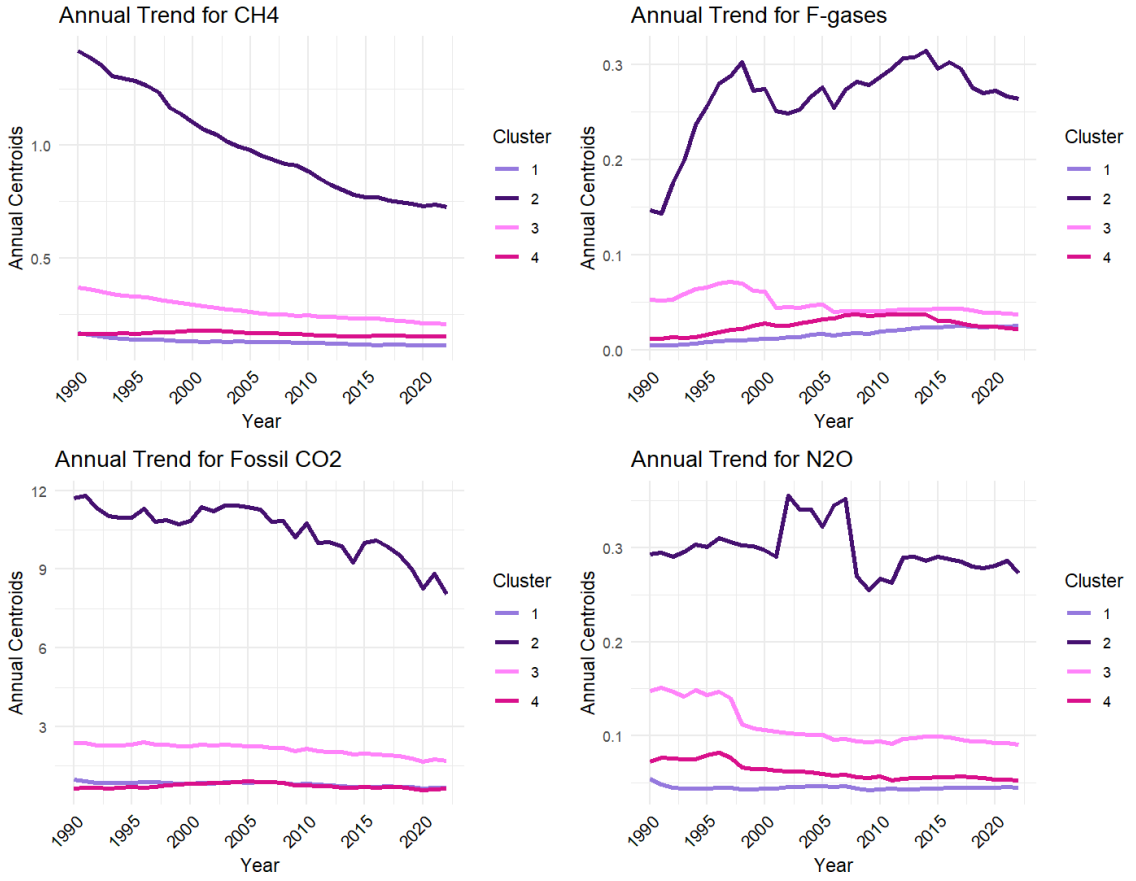


Figure 3.5.4: Annual average of emission per km^2 of gases for spatiotemporal clusters obtained combining the inertia of the dissimilarity matrices of the time series of emissions of different gases between 1990 and 2022 that is D_{CH_4}, $D_{\text{F-gases}}$, D_{CO_2} and D_{N_2O}, and the geodetic distances across centroids of NUTS2 level regions D_{sp}.

these exhibit a decreasing trend, nearly halving between 1990 and 2022. Similarly, CO_2 emissions show a decline over the years, dropping from approximately 12 tons per km^2 to about 8 tons per km^2. In contrast, emissions of F-gases experience a sharp increase during the 1990s, stabilizing thereafter at a level of approximately 0.25 tons per km^2. Meanwhile, emissions of NO_2 display fluctuations over time without any discernible trend.

Moving to Cluster 3, we can see from the map that it incorporates regions from North-central Europe, mainly Denmark, Germany, east part of France, and Northen Italy. It appears to be the second worst cluster in term of emission per km^2, showing a slowly decreasing trend in CH_4 emissions, quite stable level of CO_2 and F-gases emissions and a drop in NO_2 emissions in the last '90s.

Finally, cluster 4 consists of regions from Ireland, Spain, Portugal, and the West part of France. Regions in this cluster exhibit very low levels of emissions, comparable to those in Cluster 1. These emissions are characterized by a slightly increasing trend for F-gases and stable levels for CH_4, CO_2, and N_2O.

3.5.3 Spatiotemporal clusters for emissions by sector

The second spatiotemporal clustering considers the GHGs emissions by sector. We compute the dissimilarity matrices of the total GHGs emissions in agricultural, buildings, energy, industry, transport, and waste sectors, and we combine them with the spatial dissimilarity matrix, thus having $P = 7$. We set $\Delta\alpha = 0.1$, for $K = 2, \dots, 10$, and we select $K^* = 5$ according to the elbow rule and $\alpha^*_{5,p} = (0.2, 0, 0, 0.1, 00.3, 0.4)$ as the weight of the inertia from each dissimilarity matrix. In Figure 3.5.5, we represent the first difference in the weighted average proportion of explained inertia, highlighting our choice of $K^* = 5$.

In Table 3.5.2, we report a summary of the weights $\underline{\alpha}^*$, the explained inertia (proportion and normalized), and the Joint Inertia in the resulting clustering partition for each dissimilarity matrix included in the algorithm.

From Table 3.5.2, it is possible to observe that the partition obtained with $K^* = 5$ clusters allows

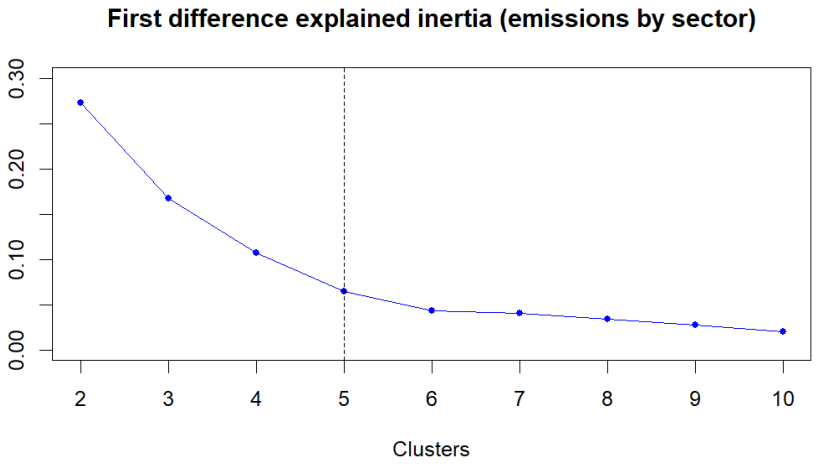


Figure 3.5.5: Increment in the weighted average proportion of explained inertia generated by a unitary increase in the number of clusters. Recall that, for each K, we considered the best combination of inertia from the dissimilarity matrices according to the α_K^* found using the maximization inertia criterion.

to explain at least 35% of the proportion of inertia in each dissimilarity matrix and the 36% of the inertia with respect to the cases in which clusters are computed only relying on one dissimilarity matrix. Each dissimilarity matrix seems to be quite helpful in obtaining the resulting partition. In particular, the spatial component presents a Joint Inertia equal to 0.62, and it incorporates 0.82 of the inertia normalized to the explained inertia in the partition obtained from the purely spatial case. On the other hand, including the spatial component, we are still able to capture 0.80 of the inertia in the non-spatial component. Although all the dissimilarity matrix D_{Tran} has weights equal to 0, thus it is not included, the resulting partition incorporates 0.61 of the inertia, which corresponds to the 67% normalized to the explained inertia from the partition obtain only considering D_{Tran}.

In Figure 3.5.6 we present a map of the regions belonging to different clusters and in Figure 3.5.7 we report the average annual emission in each cluster, for each sector.

	D_{Agri}	D_{Build}	D_{Energy}	D_{Ind}	D_{Tran}	D_{Waste}	D_{sp}
α_p^*	0.20	0.00	0.00	0.10	0.00	0.30	0.40
$\mathcal{Q}_{D_p}(\mathcal{P}_K^{\underline{\alpha}})$	0.69	0.57	0.35	0.58	0.61	0.61	0.65
$\tilde{\mathcal{Q}}_{D_p}(\mathcal{P}_K^{\underline{\alpha}})$	0.71	0.58	0.36	0.59	0.64	0.63	0.82
$\tilde{\mathcal{Q}}_{D_{-p}}(\mathcal{P}_K^{\underline{\alpha}})$	0.94	1.00	1.00	0.99	1.00	1.00	0.80
$JI_p(\mathcal{P}_K^{\underline{\alpha}^*})$	0.65	0.58	0.36	0.58	0.64	0.63	0.62

Table 3.5.2: Summary of the $\underline{\alpha}^*$ weights and the inertia (proportion and normalized) returned by the spatiotemporal clustering of emissions by gases at the optimal solution $K^* = 5$. $\mathcal{Q}_{D_p}(\mathcal{P}_K^{\underline{\alpha}})$ is the proportion of explained inertia in D_p; $\tilde{\mathcal{Q}}_{D_p}(\mathcal{P}_K^{\underline{\alpha}})$ is the normalized proportion of inertia in D_p; $\tilde{\mathcal{Q}}_{D_{-p}}(\mathcal{P}_K^{\underline{\alpha}})$ is the normalized proportion of weighted average inertia excluding D_p; and $JI_p(\mathcal{P}_K^{\underline{\alpha}^*})$ is the Joint Inertia relative to D_p.

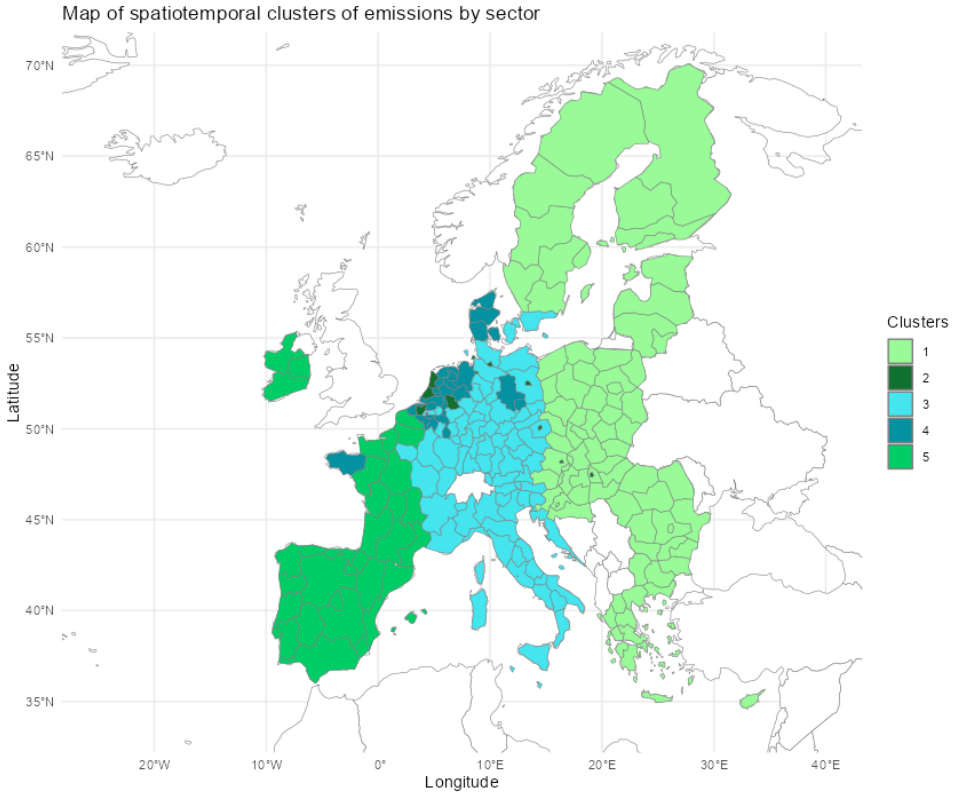


Figure 3.5.6: Map of spatiotemporal clusters obtained combining the inertia of dissimilarity matrices of the time series of emissions of different sectors between 1990 and 2022 that is D_{Agri}, D_{Build}, D_{Energy}, D_{Ind}, D_{Tran} and D_{Waste}., and the geodetic distances across centroids of nuts2 level regions D_{sp}.

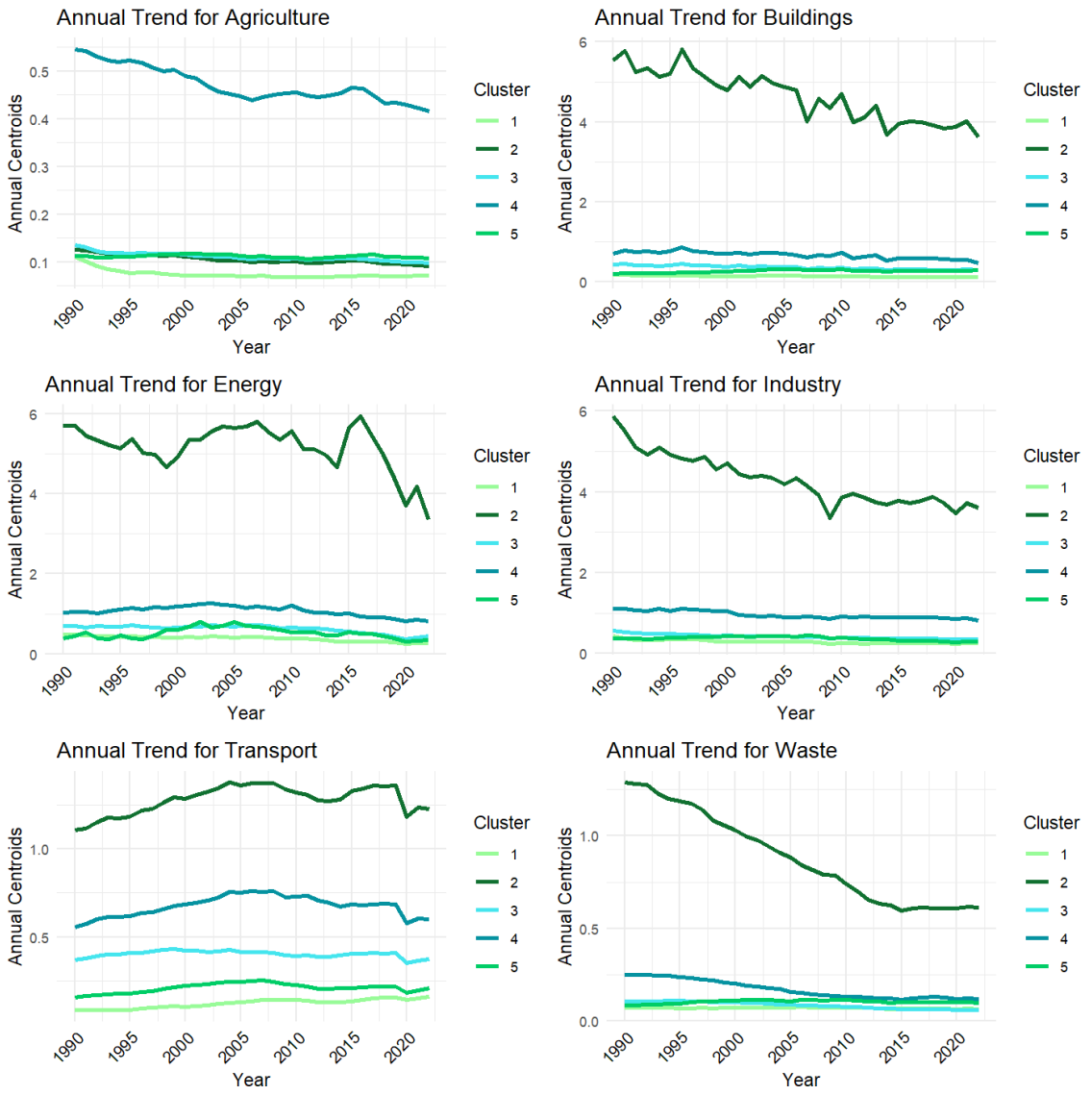


Figure 3.5.7: Annual average of emission per km^2 of sectors for spatiotemporal clusters obtained combining the inertia of the dissimilarity matrices of the time series of emissions of different sectors between 1990 and 2022 that is D_{Agri}, D_{Build}, D_{Energy}, D_{Ind}, D_{Tran} and D_{Waste}, and the geodetic distances across centroids of nuts2 level regions D_{sp}.

Cluster 1 consists primarily of regions in Eastern Europe, including Finland, Sweden, Latvia, Lithuania, Poland, Slovakia, Hungary, Romania, Bulgaria, and Greece. This cluster exhibits the

overall lowest emission levels, with stable trends in the Buildings, Energy, Industry, and Waste sectors. Emissions in the Agricultural sector show a slight decrease during the 1990s, while emissions in the Transport sector display a gradual increase over time.

Cluster 2 includes the smallest number of regions, located in Belgium and the Netherlands and some metropolitan areas in Germany and Eastern Europe. As highly industrialized and densely populated areas, it is unsurprising that this group exhibits the highest levels of emissions across all sectors except Agriculture. In particular, emissions from the Buildings and Industry sectors display a decreasing trend, declining from 6 tonnes per km^2 to less than 4 tonnes per km^2. Emissions from the Energy sector demonstrate high variability over time, with a clear reduction from 6 to 3.5 tonnes per km^2 in the last five years. Waste emissions also follow a generally decreasing trend but appear to have stabilized in recent years. Notably, only the Transport sector shows an increase in emissions over time.

Cluster 3 primarily includes regions located in Central Europe, such as Germany, the Czech Republic, Austria, Italy, Croatia, Slovenia, along with a few regions of France. This cluster exhibits steady levels of emissions, which are neither the highest nor the lowest compared to other clusters across all sectors.

Considering Cluster 4, it includes regions located in Belgium and the Netherlands, Northern France and Northern Germany and Denmark. This group is characterized by particularly high levels of emissions in the Agricultural sector, which exhibit a slowly decreasing trend. For other sectors, it ranks as the second worst cluster, following Cluster 3. In the Transport and Waste sectors, emissions follow a dynamic similar to Cluster 3, showing increasing and decreasing trends, respectively, while no clear trend is observed in the Buildings, Energy, and Industry sectors.

Finally, cluster 5 comprises regions located in Spain, Portugal, Ireland, and certain areas of France. This group is characterized by relatively low emissions across all sectors. Notably, emission levels in this group do not exhibit pronounced rising or falling trends. However, a slight upward trend is observed in the Transport sector, while the Energy sector shows a modest decline over time.

Overall, we can conclude that a relatively small number of regions, encompassing a limited por-

tion of the European territory, are responsible for exceptionally high GHG emissions. Within each breakdown, a group with particularly elevated emission levels could be distinctly identified. While it is clear that this small group contributes disproportionately to total emissions, it is equally evident that these regions also exhibit decreasing trends over time. This suggests ongoing improvements, albeit at varying rates, across individual sectors and for gases with higher emission levels.

3.5.4 Clustering Comparison and internal validation

To further investigate the role of the spatial component in cluster formation, we performed non-spatial clustering as a robustness check for emissions by gas and by sector. The resulting clusters show greater overlap compared to the spatiotemporal clustering approach, in both cases. Nevertheless, regions in Central-Northern Europe and metropolitan areas remain identifiable as those with the highest levels of emissions per km^2 across multiple sectors. A detailed discussion of this robustness check and the associated results can be found in Appendix 3.C.

In addition, Appendix 3.D reports an internal validation of the partitions. For each gas and sector we compute standard indices (Average Silhouette, Calinski–Harabasz, Within/Between ratio, Pearson's Gamma) and compare three solutions: spatiotemporal, temporal-only, and spatial-only. Results show the expected pattern—temporal-only scores higher on feature-based dissimilarities, spatial-only on the spatial one—while the spatiotemporal partition consistently lies between them, indicating a balanced trade-off. Silhouettes are generally modest (as typical with high-dimensional series), but the overall evidence supports combining temporal and spatial information.

Moreover, to assess cluster stability across environmental dimensions and to examine gas- and sector-specific patterns in greater detail, we provide in Appendix 3.E a comparative set of maps showing spatiotemporal clusters by gas (CH_4, F-gases, CO_2, N_2O) and by sector (Agriculture, Buildings, Energy, Industry, Transport, Waste). Furthermore, in order to evaluate their interaction with the spatial component, we show for each partition the resulting Joint Inertia.

3.6 Conclusion

Spatiotemporal hierarchical clustering methodology is a versatile and flexible approach to analyze patterns and dynamics of complex phenomena. In our study, we applied this methodology to greenhouse gas emissions across European regions, combining temporal emission patterns and spatial proximity into a unified analysis. Notably, we sought to assess whether the geographical information is helpful in explaining emission dynamics, providing both methodological advancements and empirical insights.

To this end, we introduced a novel measure, the Joint Inertia, to evaluate the role of the spatial component in the clustering process. This measure captures the contribution of spatial distances in generating the final clustering partition. Through a simulation experiment, we demonstrated the advantages of our methodology and highlighted how the Joint Inertia effectively quantifies the influence of spatial information on the clustering results. These findings underscore the robustness of our approach and its capacity to reveal spatial-temporal structures in the data that might otherwise remain hidden.

Our application provides compelling evidence that spatial context meaningfully enhances the analysis of emission trends. The Joint Inertia values indicate that including spatial distances enriches the clustering process, revealing region-specific patterns and dynamics that reflect Europe's inherent heterogeneity in economic, industrial, and social characteristics.

By separately analyzing emissions by gas type and sector, we uncovered distinct trends and dynamics within each cluster. This analysis highlighted how regions with high emissions are often associated with specific sectors, such as agriculture or industry, while others demonstrate relatively low emissions with stable or slowly evolving patterns. The interplay of gas-specific and sector-specific emissions varies significantly across regions, reflecting the diverse economic, industrial, and social landscapes of Europe. This detailed approach provides a more nuanced understanding of emission patterns, identifying both hotspots of high emissions and regions with notable improvements or stability. Such insights are invaluable for tailoring regional environmental policies and strategies to address sector-specific challenges effectively.

The methodological contributions and empirical findings of this study hold significant utility for both statisticians and environmental researchers. From a methodological perspective, our approach demonstrates the versatility of spatiotemporal clustering for high-dimensional data, offering a framework that can be adapted to other environmental and socioeconomic phenomena. For environmental research, this study delivers actionable insights into the spatiotemporal dynamics of greenhouse gas emissions, offering a foundation for targeted mitigation policies and resource allocation.

Future work could extend these contributions by linking cluster dynamics to socioeconomic, industrial, or policy variables, further enriching the understanding of emission drivers, exploiting proper spatiotemporal models. This would enable a more comprehensive framework for analyzing greenhouse gas emissions in a holistic spatiotemporal context, paving the way for interdisciplinary research and evidence-based policy development.

Ansari, M. Y., Ahmad, A., Khan, S. S., Bhushan, G., & Mainuddin. (2020). Spatiotemporal clustering: A review. *Artificial Intelligence Review*, *53*(4), 2381–2423. https://doi.org/10.1007/s10462-019-09736-1

Banja, M., Crippa, M., Guizzardi, D., Muntean, M., Pagani, F., & Pisoni, E. (2025). A comparative analysis of EDGAR and UNFCCC GHG emissions inventories: Insights on trends, methodology and data discrepancies. *Earth System Science Data Discussions*, *2025*, 1–39. https://doi.org/10.5194/essd-2025-385

Benevento, A., & Durante, F. (2024). Correlation-based hierarchical clustering of time series with spatial constraints. *Spatial Statistics*, *59*. https://doi.org/10.1016/j.spasta.2023.100797

Berndt, D. J., & Clifford, J. (1994). Using dynamic time warping to find patterns in time series. *AAAIWS'94: Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining*, 359–370.

Birant, D., & Kut, A. (2007). ST-DBSCAN: An algorithm for clustering spatial-temporal data. *Data and Knowledge Engineering*, *60*(1), 208–221. https://doi.org/10.1016/j.datak.2006.01.013

Bourgault, G., Marcotte, D., & Legendre, P. (1992). The multivariate (co)variogram as a spatial weighting function in classification methods. *Mathematical Geology*, *24*(5), 463–478. https://doi.org/10.1007/BF00890530

Bucci, A., Ippoliti, L., & Valentini, P. (2023). Analysing spatiotemporal patterns of COVID-19 confirmed deaths at the NUTS-2 regional level. *Regional Statistics*, *13*(2), 214–239. https://doi.org/10.15196/RS130202

Bucciarelli, G. (2024). ARDECO: Annual Regional Database of the European Commission (ARDECO) [R package version 2.1.0]. https://CRAN.R-project.org/package=ARDECO

Calculli, C., Fassò, A., Finazzi, F., Pollice, A., & Turnone, A. (2015). Maximum likelihood estimation of the multivariate hidden dynamic geostatistical model with application to air quality in Apulia, Italy. *Environmetrics*, *26*(6), 406–417. https://doi.org/https://doi.org/10.1002/env.2345

Caliński, T., & Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics*, *3*(1), 1–27. https://doi.org/10.1080/03610927408827101

Chavent, M., Kuentz-Simonet, V., Labenne, A., & Saracco, J. (2018). ClustGeo: An R package for hierarchical clustering with spatial constraints. *Computational Statistics*, *33*, 1–24. https://doi.org/ff10.1007/s00180-018-0791-1

Chen, X. C., Faghmous, J. H., Khandelwal, A., & Kumar, V. Clustering dynamic spatio-temporal patterns in the presence of noise and missing data. In: In *Ijcai international joint conference on artificial intelligence. 2015-January*. 2015, 2575–2581.

Crippa, M., Guizzardi, D., Solazzo, E., Muntean, M., Schaaf, E., Monforti-Ferrario, F., Banja, M., Olivier, J., Grassi, G., & Rossi, S. (2024). *Ghg emissions of all world countries - 2024* (Report). Publications Office of the European Union, Luxembourg. https://doi.org/10.2760/40028977,JRC138862.

Deb, S., & Karmakar, S. (2023). A novel spatio-temporal clustering algorithm with applications on COVID-19 data from the United States. *Computational Statistics and Data Analysis*, *188*. https://doi.org/10.1016/j.csda.2023.107810

D'Urso, P., & Vitale, V. (2020). A robust hierarchical clustering for georeferenced data. *Spatial Statistics*, *35*. https://doi.org/10.1016/j.spasta.2020.100407

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters a density-based algorithm for discovering clusters in large spatial databases with noise. *Proceedings - 2nd International Conference on Knowledge Discovery and Data Mining, KDD 1996*, 226–231.

European Commission, . (2016). NUTS – Nomenclature of territorial units for statistics.

Fouedjio, F. (2016). A hierarchical clustering method for multivariate geostatistical data. *Spatial Statistics*, *18*, 333–351. https://doi.org/10.1016/j.spasta.2016.07.003

Georgoulas, G., Konstantaras, A., Katsifarakis, E., Stylios, C., Maravelakis, E., & Vachtsevanos, G. (2013). "seismic-mass" density-based algorithm for spatio-temporal clustering. *Expert Systems with Applications*, *40*(10), 4183–4189. https://doi.org/10.1016/j.eswa.2013.01.028

Gordon, A. D. (1999). *Classification*. CRC Press.

Halkidi, M., Batistakis, Y., & Vazirgiannis, M. (2001). On clustering validation techniques. *Journal of Intelligent Information Systems*, *17*(2-3), 107–145. https://doi.org/10.1023/A:1012801612483

Hennig, C. (2014). How many bee species? a case study in determining the number of clusters. *Studies in Classification, Data Analysis, and Knowledge Organization*, *47*, 41–49. https://doi.org/10.1007/978-3-319-01595-8_5

Jaya, I. G. N. M., & Folmer, H. (2020). Bayesian spatiotemporal mapping of relative dengue disease risk in bandung, indonesia. *Journal of Geographical Systems*, *22*(1), 105–142. https://doi.org/10.1007/s10109-019-00311-4

Jaya, I. G. N. M., & Folmer, H. (2021). Identifying spatiotemporal clusters by means of agglomerative hierarchical clustering and bayesian regression analysis with spatiotemporally varying coefficients: Methodology and application to dengue disease in bandung, indonesia. *Geographical Analysis*, *53*(4), 767–817. https://doi.org/10.1111/gean.12264

Jaya, I., Ruchjana, B., Andriyana, Y., & Agata, R. Clustering with spatial constraints: The case of diarrhea in Bandung city, Indonesia. In: In *Journal of physics: Conference series. 1397.* (1). 2019. https://doi.org/10.1088/1742-6596/1397/1/012068

Kaufman, L., & Rousseeuw, P. J. (1990). *Finding groups in data: An introduction to cluster analysis*. Wiley.

Maranzano, P., & Fassó, A. (2022). *The impact of the lockdown restrictions on air quality during COVID-19 pandemic in Lombardy, Italy*. Springer. https://doi.org/10.1007/978-3-031-07155-3_15

Maranzano, P., Fassò, A., Pelagatti, M., & Mudelsee, M. (2020). Statistical modeling of the early-stage impact of a new traffic policy in Milan, Italy. *International Journal of Environmental Research and Public Health*, *17*(3). https://doi.org/10.3390/ijerph17031088

Maranzano, P., Otto, P., & Fassò, A. (2023). Adaptive LASSO estimation for functional hidden dynamic geostatistical models. *Stochastic Environmental Research and Risk Assessment*. https://doi.org/10.1007/s00477-023-02466-5

Maranzano, P., & Pelagatti, M. (2024). Spatiotemporal event studies for environmental data under cross-sectional dependence: An application to air quality assessment in Lombardy. *Journal of Agricultural, Biological, and Environmental Statistics*, *29*(1), 147–168. https://doi.org/10.1007/s13253-023-00564-z

Marongiu, A., Collalto, A. G., Distefano, G. G., & Angelino, E. (2024). Application of machine learning to estimate ammonia atmospheric emissions and concentrations. *Air*, *2*(1), 38–60. https://doi.org/10.3390/air2010003

Marongiu, A., Colombo, L., & Collalto, A. G. (2025). Dynamic emission inventory of ammonia in Northern Italy by machine learning. *Air Quality, Atmosphere and Health*. https://doi.org/10.1007/s11869-025-01779-4

Mattera, R., & Franses, P. H. (2023). Are African business cycles synchronized? Evidence from spatio-temporal modeling. *Economic Modelling*, *128*. https://doi.org/10.1016/j.econmod.2023.106485

Montero, P., & Vilar, J. A. (2014). TSclust: An R package for time series clustering. *Journal of Statistical Software*, *62*(1), 1–43. https://doi.org/10.18637/jss.v062.i01

Morelli, C., Boccaletti, S., Maranzano, P., & Otto, P. (2025). Multidimensional spatiotemporal clustering - an application to environmental sustainability scores in Europe. *Environmetrics*, *36*(2). https://doi.org/10.1002/env.2893

Mutascu, M. (2022). CO2 emissions in the USA: New insights based on ANN approach. *Environmental Science and Pollution Research*, *29*(45), 68332–68356. https://doi.org/10.1007/s11356-022-20615-1

Naidu, G., Zuva, T., & Sibanda, E. M. (2023). A review of evaluation metrics in machine learning algorithms. *Artificial Intelligence Application in Networks and Systems*, 15–25.

Najafabadi, A. M., Mahaki, B., & Hajizadeh, Y. (2020). Spatiotemporal modeling of airborne fine particulate matter distribution in Isfahan. *International Journal of Environmental Health Engineering (IJEHE)*, *2020*(July), 1–7. https://doi.org/10.4103/ijehe.ijehe_6_20

Oliver, M., & Webster, R. (1989). A geostatistical basis for spatial weighting in multivariate classification. *Mathematical Geology*, *21*(1), 15–35. https://doi.org/10.1007/BF00897238

Otto, P., Fusta Moro, A., Rodeschini, J., Shaboviq, Q., Ignaccolo, R., Golini, N., Cameletti, M., Maranzano, P., Finazzi, F., & Fassò, A. (2024). Spatiotemporal modelling of PM 2.5 concentrations in Lombardy (Italy): A comparative study. *Environmental and Ecological Statistics*, *31*(2), 245–272. https://doi.org/10.1007/s10651-023-00589-0

Pértega-Díaz, S., & Vilar, J. A. (2010). Comparing several parametric and nonparametric approaches to time series clustering: A simulation study. *Journal of Classification*, *27*(3), 333–362. https://doi.org/10.1007/s00357-010-9064-6

Piccolo, D. (1990). A distance measure for classifying ARIMA models. *Journal of Time Series Analysis*, *11*(2), 153–164. https://doi.org/10.1111/j.1467-9892.1990.tb00048.x

Pisoni, E., Zauli-Sajani, S., Belis, C. A., Khomenko, S., Thunis, P., Motta, C., Van Dingenen, R., Bessagnet, B., Monforti-Ferrario, F., Maes, J., & Feyen, L. (2025). High resolution assessment of air quality and health in Europe under different climate mitigation scenarios. *Nature Communications*, *16*(1). https://doi.org/10.1038/s41467-025-60449-2

Rocha, J. A. M., Times, V. C., Oliveira, G., Alvares, L. O., & Bogorny, V. DB-SMoT: A direction-based spatio-temporal clustering method. In: In *2010 ieee international conference on intelligent systems, is 2010 - proceedings*. 2010, 114–119. https://doi.org/10.1109/IS.2010.5548396

Rodeschini, J., Fassò, A., Finazzi, F., & Fusta Moro, A. (2024). Scenario analysis of livestock-related PM2.5 pollution based on a new heteroskedastic spatiotemporal model. *Socio-Economic Planning Sciences*, *96*. https://doi.org/10.1016/j.seps.2024.102053

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*(100), 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

Rowland, O. E. (2024). Comparative analysis of meteorological parameters and their relationship with NO2, PM10, PM2.5 and O3 concentrations at selected urban air quality monitoring stations in Krakow, Paris, and Milan. *Discover Environment*, *2*(1). https://doi.org/10.1007/s44274-024-00060-2

Sakoe, H., & Chiba, S. (1971). A dynamic programming approach to continuous speech recognition. *Computer Science, Linguistics*.

Sakoe, H., & Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *26*(1), 43–49. https://doi.org/10.1109/TASSP.1978.1163055

Taghavi-Shahri, S. M., Fassò, A., Mahaki, B., & Amini, H. (2020). Concurrent spatiotemporal daily land use regression modeling and missing data imputation of fine particulate matter using distributed space-time expectation maximization. *Atmospheric Environment*, *224*. https://doi.org/10.1016/j.atmosenv.2019.117202

Tietbohl, A., Bogorny, V., Kuijpers, B., & Alvares, L. O. A clustering-based approach for discovering interesting places in trajectories. In: In *Proceedings of the acm symposium on applied computing*. 2008, 863–868. https://doi.org/10.1145/1363686.1363886

Yu, C., Tan, J., Cheng, Y., & Mi, X. (2024). Data analysis and preprocessing techniques for air quality prediction: A survey. *Stochastic Environmental Research and Risk Assessment*, *38*(6), 2095–2117. https://doi.org/10.1007/s00477-024-02693-4

Zammarchi, G., & Maranzano, P. (2024). Mapping climate change awareness through spatial hierarchical clustering.

Zhou, J., Liu, T., & Zhu, J. (2019). Weighted adjacent matrix for K-means clustering. *Multimedia Tools and Applications*, *78*(23), 33415–33434. https://doi.org/10.1007/s11042-019-08009-x

3.A Appendix 3.A: Technical details of Joint Inertia

In this appendix we focus on the properties of Joint Inertia, in particular we explain why its value varies between 0 and 1 and what are the conditions that guarantee it, starting from the simplest case where we have only two matrices of dissimilarity and then extending the reasoning to the case where we want to combine more than two matrices.

Joint Inertia for the case $P = 2$ dissimilarity matrices

Let $D_0 = [d_{0,ij}]_{i,j=1,\dots,n}$ and $D_1 = [d_{1,ij}]_{i,j=1,\dots,n}$ refer respectively to any distances matrix of variables and the spatial distances matrix considering a sample of n units and let w_i with $i = 1,\dots,n$ be the weight of the i-th cross-sectional unit to be clustered. Let us consider a partition $\mathcal{P}_K^\alpha = (\mathcal{C}_1^\alpha,\dots,\mathcal{C}_K^\alpha)$ in K clusters obtained mixing the dissimilarity matrices D_0 and D_1 with the parameter α. Also, let us denote its within-clusters mixed inertia as $W(\mathcal{P}_K^\alpha)$ and the corresponding proportion of the total pseudo inertia explained as

$$Q_m(\mathcal{P}_K^\alpha) = 1 - \frac{W_m(\mathcal{P}_K^\alpha)}{W_m(\mathcal{P}_1)} = 1 - \frac{\sum_{k=1}^{K} \frac{1}{2\mu_k} \sum_{i\in\mathcal{C}_k} \sum_{j\in\mathcal{C}_k} w_i w_j d_{mij}^2}{\frac{1}{2\mu} \sum_i \sum_j w_i w_j d_{mij}^2} \qquad m = 0, 1.$$

where $\mu_k = \sum_{i\in\mathcal{C}_k} w_i$ and $W_0(\mathcal{P}_1)$ and $W_1(\mathcal{P}_1)$ are the total pseudo inertia under dissimilarity matrix D_0 and under dissimilarity matrix D_1, respectively. It is easy to observe that the possible values of this ratio vary between 0 and 1 because the sum of the square dissimilarities between the units belonging to the same clusters will not be greater than the total of the square dissimilarities within a matrix, and being quantities square, will all have positive values, for each possible value of α.

The combination in within mixed pseudo inertia will capture more the dissimilarities in D_0 (or D_1), when the value of α is closer to 0 (or close to 1), thus the resulting partition $\mathcal{P}_K^\alpha$ will be able to capture an higher proportion of explained inertia in D_0 (or D_1). Therefore, when α increases, $Q_{D_0}(\mathcal{P}_K^\alpha)$ will increase and $Q_{D_1}(\mathcal{P}_K^\alpha)$ will decrease. This is also explained in details in Section 3.2 of Chavent et al., 2018. Notice that the proportion of explained inertia is equal to 0 only when $K = 1$ and it is 1 only when $K = n$, for each value of α. Following we summarize and formalize the

concept which hold for $1 < K < n$

$$0 = Q_{D_0}(\mathcal{P}_{K=1}) < Q_{D_0}(\mathcal{P}_K^{\alpha=1}) \leq Q_{D_0}(\mathcal{P}_K^{\alpha}) \leq Q_{D_0}(\mathcal{P}_K^{\alpha=0}) < Q_{D_0}(\mathcal{P}_{K=n}) = 1$$

$$0 = Q_{D_1}(\mathcal{P}_{K=1}) < Q_{D_1}(\mathcal{P}_K^{\alpha=0}) \leq Q_{D_1}(\mathcal{P}_K^{\alpha}) \leq Q_{D_1}(\mathcal{P}_K^{\alpha=1}) < Q_{D_1}(\mathcal{P}_{K=n}) = 1.$$

Now, we consider the normalized proportion of explained inertia for dissimilarity matrix D_0 (or D_1), which correspond to the ratio between the proportion of explained inertia in the partition obtained using the combined dissimilarity matrix with parameter α, and the proportion of inertia explained using only the features (or spatial) dissimilarities, thus the maximum proportion of explained inertia.

$$\tilde{Q}_{D_0}(\mathcal{P}_K^{\alpha}) = \frac{Q_{D_0}(\mathcal{P}_K^{\alpha})}{Q_{D_0}(\mathcal{P}_K^{0})} \qquad \tilde{Q}_{D_1}(\mathcal{P}_K^{\alpha}) = \frac{Q_{D_1}(\mathcal{P}_K^{\alpha})}{Q_{D_1}(\mathcal{P}_K^{1})}.$$

These quantities are strictly positive and no greater than 1. In particular, the normalized proportion of explained inertia in D_0 (or D_1) reaches its minimum value when the feature (or spatial) dissimilarity matrix is not included in the combination, thus $\alpha = 1$ (or $\alpha = 0$), and it reaches its maximum when the clustering partition is obtained from the purely no-spatial (or spatial) case, thus in $\mathcal{P}_K^0 = (\mathcal{C}_1^0, \ldots, \mathcal{C}_K^0)$ (or $\mathcal{P}_K^1 = (\mathcal{C}_1^1, \ldots, \mathcal{C}_K^1)$).

$$0 < \tilde{Q}_{D_0}(\mathcal{P}_K^{\alpha=1}) \leq \tilde{Q}_{D_0}(\mathcal{P}_K^{\alpha}) \leq \tilde{Q}_{D_0}(\mathcal{P}_K^{\alpha=0}) = 1$$

$$0 < \tilde{Q}_{D_1}(\mathcal{P}_K^{\alpha=0}) \leq \tilde{Q}_{D_1}(\mathcal{P}_K^{\alpha}) \leq \tilde{Q}_{D_1}(\mathcal{P}_K^{\alpha=1}) = 1$$

The Joint Inertia between two dissimilarity matrices is computed using the following formula

$$JI(\mathcal{P}_K^{\alpha}) = \tilde{Q}_{D_1}(\mathcal{P}_K^{\alpha}) - (1 - \tilde{Q}_{D_0}(\mathcal{P}_K^{\alpha})),$$

It captures the net gain in explained inertia from jointly incorporating spatial and feature dissimilarities. The indicator quantifies the added explanatory power from combining both sources—i.e., the synergy between spatial and non-spatial components—revealing patterns that would be missed

if either were used alone. It can be expressed in term of dissimilarities as follows:

$$
\begin{aligned}
JI(\mathcal{P}_K^\alpha) &= \tilde{Q}_{D_1}(\mathcal{P}_K^\alpha) - (1 - \tilde{Q}_{D_0}(\mathcal{P}_K^\alpha)) \\
&= \frac{Q_{D_1}(\mathcal{P}_K^\alpha)}{Q_{D_1}(\mathcal{P}_K^1)} - \left(1 - \frac{Q_{D_0}(\mathcal{P}_K^\alpha)}{Q_{D_0}(\mathcal{P}_K^0)}\right) \\
&= \frac{1 - \frac{W_{D_1}(\mathcal{P}_K^\alpha)}{W_{D_1}(\mathcal{P}_1)}}{1 - \frac{W_{D_1}(\mathcal{P}_K^1)}{W_{D_1}(\mathcal{P}_1)}} - \left(1 - \frac{1 - \frac{W_{D_0}(\mathcal{P}_K^\alpha)}{W_{D_0}(\mathcal{P}_1)}}{1 - \frac{W_{D_0}(\mathcal{P}_K^0)}{W_{D_0}(\mathcal{P}_1)}}\right) \\
&= \frac{W_{D_1}(\mathcal{P}_1) - W_{D_1}(\mathcal{P}_K^\alpha)}{W_{D_1}(\mathcal{P}_1) - W_{D_1}(\mathcal{P}_K^1)} - \left(1 - \frac{W_{D_0}(\mathcal{P}_1) - W_{D_0}(\mathcal{P}_K^\alpha)}{W_{D_0}(\mathcal{P}_1) - W_{D_0}(\mathcal{P}_K^0)}\right) \\
&= \frac{\frac{1}{2\mu}\sum_i \sum_j w_i w_j d_{1,ij}^2 - \frac{1}{2\mu_k}\sum_{i\in\mathcal{C}_k^\alpha}\sum_{j\in\mathcal{C}_k^\alpha} w_i w_j d_{1,ij}^2}{\frac{1}{2\mu}\sum_i \sum_j w_i w_j d_{1,ij}^2 - \frac{1}{2\mu_k}\sum_{i\in\mathcal{C}_k^1}\sum_{j\in\mathcal{C}_k^1} w_i w_j d_{1,ij}^2} \\
&\quad - \left(1 - \frac{\frac{1}{2\mu}\sum_i \sum_j w_i w_j d_{1,ij}^2 - \frac{1}{2\mu_k}\sum_{i\in\mathcal{C}_k^\alpha}\sum_{j\in\mathcal{C}_k^\alpha} w_i w_j d_{0,ij}^2}{\frac{1}{2\mu}\sum_i \sum_j w_i w_j d_{0,ij}^2 - \frac{1}{2\mu_k}\sum_{i\in\mathcal{C}_k^0}\sum_{j\in\mathcal{C}_k^0} w_i w_j d_{0,ij}^2}\right).
\end{aligned}
$$

It is strictly positive because the partition obtained from the mixed dissimilarity matrix allows at least to explain as much as the partition obtained from purely non-spatial case in D_0 or the partition obtained from purely spatial case in D_1. Moreover, it is no grater than 1 because the partition $\mathcal{P}_K^0 = (\mathcal{C}_1^0, \ldots, \mathcal{C}_K^0)$ is the one that minimize the within cluster non-spatial dissimilarity with respect to other partitions, such that $\frac{1}{2\mu_k}\sum_{i\in\mathcal{C}_k^0}\sum_{j\in\mathcal{C}_k^0} w_i w_j d_{0,ij}^2 \leq \frac{1}{2\mu_k}\sum_{i\in\mathcal{C}_k}\sum_{j\in\mathcal{C}_k} w_i w_j d_{0,ij}^2$ and thus $\frac{1}{2\mu_k}\sum_{i\in\mathcal{C}_k^0}\sum_{j\in\mathcal{C}_k^0} w_i w_j d_{0,ij}^2 \leq \frac{1}{2\mu_k}\sum_{i\in\mathcal{C}_k^\alpha}\sum_{j\in\mathcal{C}_k^\alpha} w_i w_j d_{0,ij}^2$. Similarly for the spatial dissimilarities, partition $\mathcal{P}_K^1 = (\mathcal{C}_1^1, \ldots, \mathcal{C}_K^1)$ is the one that minimize the within cluster dissimilarities in the spatial component with respect to other partitions, such that $\frac{1}{2\mu_k}\sum_{i\in\mathcal{C}_k^1}\sum_{j\in\mathcal{C}_k^1} w_i w_j d_{1,ij}^2 \leq \frac{1}{2\mu_k}\sum_{i\in\mathcal{C}_k}\sum_{j\in\mathcal{C}_k} w_i w_j d_{1,ij}^2$ and thus $\frac{1}{2\mu_k}\sum_{i\in\mathcal{C}_k^1}\sum_{j\in\mathcal{C}_k^1} w_i w_j d_{1,ij}^2 \leq \frac{1}{2\mu_k}\sum_{i\in\mathcal{C}_k^\alpha}\sum_{j\in\mathcal{C}_k^\alpha} w_i w_j d_{1,ij}^2$.

Joint Inertia for the case $P > 2$ dissimilarity matrices

The generalization to the case of multiple dissimilarity matrices, one of which relates to the spatial component, is straightforward. Let us suppose that we are interested in computing the Joint Inertia for the generic variable m (with m being the generic index for the P variables of interests) compared

to the combination of all the other variables used to cluster the time series[7].

Recalling the notation previously introduced, let us consider a set of P variables observed over time at n spatial locations (or regions) and let $\underline{\alpha} = [\alpha_p]_{p=1,\dots,P} = (\alpha_1, \dots, \alpha_m, \dots, \alpha_P)$ on the $(P-1)$-simplex $\mathbf{\Delta}^{P-1}$ be the vector of mixing weights[8] used to compute the linear combination of the mixed within pseudo inertia from dissimilarity matrices $D_p = [d_{p,ij}]_{p=1,\dots,P-1;i,j=1,\dots,n}$, such that $I(\mathcal{C}_K^{\underline{\alpha}}) = \sum_{p=1}^{P} \alpha_p \sum_{i \in \mathcal{C}_K} \sum_{j \in \mathcal{C}_K} \frac{w_i w_j}{2 \sum_{i \in \mathcal{C}_K} w_i} {d_{p,ij}}^2$.

By applying the Ward-like hierarchical clustering algorithm presented in the manuscript the resulting partition is denoted as $\mathcal{P}_K^{\underline{\alpha}}$. Given the generic partition $\mathcal{P}_K^{\underline{\alpha}}$ obtained from the mixed dissimilarity matrix, the proportion of explained inertia for matrix m-th is computed as

$$Q_{D_m}(\mathcal{P}_K^{\underline{\alpha}}) = 1 - \frac{W_{D_m}(\mathcal{P}_K^{\underline{\alpha}})}{W_{D_m}(\mathcal{P}_1)}$$

and the normalized proportion of explained inertia, with respect to the case in which only variable p is used, is computed as

$$\tilde{Q}_{D_m}(\mathcal{P}_K^{\underline{\alpha}}) = \frac{Q_{D_m}(\mathcal{P}_K^{\underline{\alpha}})}{Q_{D_m}(\mathcal{P}_K^{\alpha_p=1})}$$

Recall that these quantities are both included between 0 and 1 for the same reasons and considerations as in the case of two dissimilarity matrices.

To retrieve the equation of the Joint Inertia for variable m we need to focus separately on the inertia explained in the the m-th dissimilarity matrix and the overall inertia explained in the others dissimilarity matrices different from m. Let $D_m = [d_{m,ij}]_{i,j=1,\dots,n}$ be the dissimilarity matrix for variable m across the n units and let $D_{-m} = [d_{p,ij}]_{p=1,\dots,m-1,m+1,\dots,P;i,j=1,\dots,n}$ be the dissimilarity matrices containing the distances across the $P-1$ time series different from m. Moreover, let use define $\underline{\alpha}'|\alpha_m = 0$ as the vector of weights that maximizes the weighted average explained inertia from all the dissimilarity matrices excluding m-th matrix and the resulting clustering partition as $\mathcal{P}_K^{\underline{\alpha}'|\alpha_m=0}$. Similarly, define the partition obtained by using only the information contained in the

[7] In the main manuscript, we focused on the specific case of the Joint Inertia for the spatial dissimilarity matrix, that we assumed to be the P-th variable. Hence, the results in the Appendix can be related to the results in the main manuscript by assuming that $m = P$.

[8] Notice that, although the grid can be either regular or irregular, here we consider a regular grid with constant step $\Delta\alpha$.

m-th matrix (i.e., by fixing $\alpha_m = 1$) as $\mathcal{P}_K^{\alpha_m=1}$.

If we denote the total inertia in D_{-m} as $W_{D_{-m}}(\mathcal{P}_1)$, we can compute the following quantities

$$\bar{Q}_{D_{-m}}(\mathcal{P}_K^{\underline{\alpha}}) = 1 - \frac{\sum_{p=1,p\neq m}^{P} W_{D_p}(\mathcal{P}_K^{\underline{\alpha}})}{\sum_{p=1,p\neq m}^{P} W_{D_p}(\mathcal{P}_1)} \qquad \bar{Q}_{D_{-m}}(\mathcal{P}_K^{\underline{\alpha}'|\alpha_m=0}) = 1 - \frac{\sum_{p=1,p\neq m}^{P} W_{D_p}(\mathcal{P}_K^{\underline{\alpha}'|\alpha_m=0})}{\sum_{p=1,p\neq m}^{P} W_{D_p}(\mathcal{P}_1)}$$

which represent the weighted average proportion of explained inertia in the dissimilarity matrices excluding the m-th matrix, when considering the partition obtained including the m-th matrix, thus the partition resulting from the optimal weights $\underline{\alpha}$ (left formula) and the weighted average proportion of explained inertia in the dissimilarity matrices excluding the m-th matrix, when considering the partition obtained excluding the m-th matrix, thus the partition resulting from the optimal weights $\underline{\alpha}'|\alpha_m = 0$ (right formula). [9]

The normalized proportion of explained inertia in the non-spatial combined dissimilarity matrix can be computed as

$$\tilde{Q}_{D_{-m}}(\mathcal{P}_K^{\underline{\alpha}}) = \frac{\bar{Q}_{D_{-m}}(\mathcal{P}_K^{\underline{\alpha}})}{\bar{Q}_{D_{-m}}(\mathcal{P}_K^{\underline{\alpha}^*|\alpha_m=0})}$$

By definition, all these quantities are bounded between 0 and 1. Indeed, If we employ one of the procedures in Chavent et al., 2018 or Morelli et al., 2025 to compute the optimal weighting vector for the combination of inertia from all the dissimilarity matrices different from m (i.e., $\underline{\alpha}^*|\alpha_m = 0$ subject to $\sum_{p\neq m} \alpha_p = 1$), then, the corresponding convex combination of inertia from dissimilarity matrices (i.e., $D_{-m}(\underline{\alpha}^*|\alpha_m = 0)$) will explain as much as possible the dissimilarities in the D_m matrix, thus being $Q_{D_{-m}}(\mathcal{P}_K^{\underline{\alpha}^*|\alpha_m=0})$ greater than any other potential $Q_{D_{-m}}(\mathcal{P}_K^{\underline{\alpha}|\alpha_m=0})$.

Finally, considering a generic combination of $\underline{\alpha}$ and K, the Joint Inertia for variable m when

[9] Notice that, when m is the spatial matrix, the first formula corresponds to the proportion of explained inertia in the non-spatial dissimilarity matrices when including the spatial component, while the second formula is the proportion of explained inertia induced by the non-spatial dissimilarity matrices when excluding the spatial component.

the number of dissimilarity matrices is $P > 2$ is given by

$$JI_m(\mathcal{P}_K^{\underline{\alpha}}) = \tilde{Q}_{D_m}(\mathcal{P}_K^{\underline{\alpha}}) - (1 - \tilde{Q}_{D_{-m}}(\mathcal{P}_K^{\underline{\alpha}})) = \frac{\bar{Q}_{D_m}(\mathcal{P}_K^{\underline{\alpha}})}{\bar{Q}_{D_m}(\mathcal{P}_K^{\alpha_m=1})} - \left(1 - \frac{\bar{Q}_{D_{-m}}(\mathcal{P}_K^{\underline{\alpha}})}{\bar{Q}_{D_{-m}}(\mathcal{P}_K^{\alpha^*|\alpha_m=0})}\right).$$

which can be expressed in term of dissimilarities as follows

$$\begin{aligned}
JI_m(\mathcal{P}_K^{\underline{\alpha}}) &= \tilde{Q}_{D_m}(\mathcal{P}_K^{\underline{\alpha}}) - (1 - \tilde{Q}_{D_{-m}}(\mathcal{P}_K^{\underline{\alpha}})) \\
&= \frac{\bar{Q}_{D_m}(\mathcal{P}_K^{\underline{\alpha}})}{\bar{Q}_{D_m}(\mathcal{P}_K^{\alpha_m=1})} - \left(1 - \frac{\bar{Q}_{D_{-m}}(\mathcal{P}_K^{\underline{\alpha}})}{\bar{Q}_{D_{-m}}(\mathcal{P}_K^{\alpha^*|\alpha_m=0})}\right) \\
&= \frac{1 - \frac{W_{D_m}(\mathcal{P}_K^{\alpha})}{W_{D_m}(\mathcal{P}_1)}}{1 - \frac{W_{D_m}(\mathcal{P}_K^{\alpha_m=1})}{W_{D_m}(\mathcal{P}_1)}} - \left(1 - \frac{1 - \frac{\sum_{p=1,p\neq m}^{P} W_{D_p}(\mathcal{P}_K^{\underline{\alpha}})}{\sum_{p=1,p\neq m}^{P} W_{D_p}(\mathcal{P}_1)}}{1 - \frac{\sum_{p=1,p\neq m}^{P} W_{D_p}(\mathcal{P}_K^{\underline{\alpha}'|\alpha_m=0})}{\sum_{p=1,p\neq m}^{P} W_{D_p}(\mathcal{P}_1)}}\right) \\
&= \frac{W_{D_m}(\mathcal{P}_1) - W_{D_m}(\mathcal{P}_K^{\underline{\alpha}})}{W_{D_m}(\mathcal{P}_1) - W_{D_m}(\mathcal{P}_K^{\alpha_m=1})} - \left(1 - \frac{\sum_{p=1,p\neq m}^{P} W_{D_p}(\mathcal{P}_1) - \sum_{p=1,p\neq m}^{P} W_{D_p}(\mathcal{P}_K^{\underline{\alpha}})}{\sum_{p=1,p\neq m}^{P} W_{D_p}(\mathcal{P}_1) - \sum_{p=1,p\neq m}^{P} W_{D_p}(\mathcal{P}_K^{\underline{\alpha}'|\alpha_m=0})}\right) \\
&= \frac{W_{D_m}(\mathcal{P}_1) - W_{D_m}(\mathcal{P}_K^{\underline{\alpha}})}{W_{D_m}(\mathcal{P}_1) - W_{D_m}(\mathcal{P}_K^{\alpha_m=1})} - \left(1 - \frac{\sum_{p=1,p\neq m}^{P} (W_{D_p}(\mathcal{P}_1) - W_{D_p}(\mathcal{P}_K^{\underline{\alpha}}))}{\sum_{p=1,p\neq m}^{P} (W_{D_p}(\mathcal{P}_1) - W_{D_p}(\mathcal{P}_K^{\underline{\alpha}'|\alpha_m=0}))}\right) \\
&= \frac{\frac{1}{2\mu}\sum_i \sum_j w_i w_j d_{m,ij}^2 - \frac{1}{2\mu_k}\sum_{i\in\mathcal{C}_k^{\underline{\alpha}}}\sum_{j\in\mathcal{C}_k^{\underline{\alpha}}} w_i w_j d_{m,ij}^2}{\frac{1}{2\mu}\sum_i \sum_j w_i w_j d_{m,ij}^2 - \frac{1}{2\mu_k}\sum_{i\in\mathcal{C}_k^{\alpha_m=1}}\sum_{j\in\mathcal{C}_k^{\alpha_m=1}} w_i w_j d_{m,ij}^2} \\
&\quad - \left(1 - \frac{\sum_{p=1,p\neq m}^{P}(\frac{1}{2\mu}\sum_i \sum_j w_i w_j d_{p,ij}^2 - \frac{1}{2\mu_k}\sum_{i\in\mathcal{C}_k^{\underline{\alpha}}}\sum_{j\in\mathcal{C}_k^{\underline{\alpha}}} w_i w_j d_{p,ij}^2)}{\sum_{p=1,p\neq m}^{P}(\frac{1}{2\mu}\sum_i \sum_j w_i w_j d_{p,ij}^2 - \frac{1}{2\mu_k}\sum_{i\in\mathcal{C}_k^{\underline{\alpha}'}}\sum_{j\in\mathcal{C}_k^{\underline{\alpha}'}} w_i w_j d_{p,ij}^2)}\right).
\end{aligned}$$

As in the two-matrix case, if the contribution of D_m is not useful, the resulting partition captures inertia from the other dissimilarity matrices but not from D_m (or vice versa). Consequently, the normalized proportion of explained inertia for D_m will be close to zero (though strictly positive), while the normalized proportion for the weighted average of all other matrices will be close to one. Conversely, if the partition from D_m alone and the partition from the combination that excludes D_m both coincide with the partition from the full combination, thus $\mathcal{P}_K^{\alpha_m=1} = \mathcal{P}_K^{\underline{\alpha}^*|\alpha_m=0} = \mathcal{P}_K^{\underline{\alpha}}$, then the Joint Inertia equals 1 (its maximum), indicating that D_m alone suffices to reproduce the

same sum of squared dissimilarities captured by the full set of matrices.

3.B Appendix 3.B: Spatial dependence over time

In this appendix, we provide a brief overview of spatial dependence over time via annual Moran's I; although the main paper operates on dissimilarities (not correlations), which are conceptually distinct, we consider this descriptive analysis complementary and helpful for readers to gauge the strength and evolution of spatial structure in the data. We compute the global Moran's I annually (1990–2022) for NUTS-2 emissions, separately by each gas and by each sector. We used as spatial-weights matrix $\mathbf{W}$ the k-nearest-neighbour. Significance is assessed via 999 permutations. We provide the results with $k = 6$ and, as a robustness check, we repeated the analysis with $k = 10$ and $k = 15$ obtaining very similar values over time (results are available upon request to the authors).

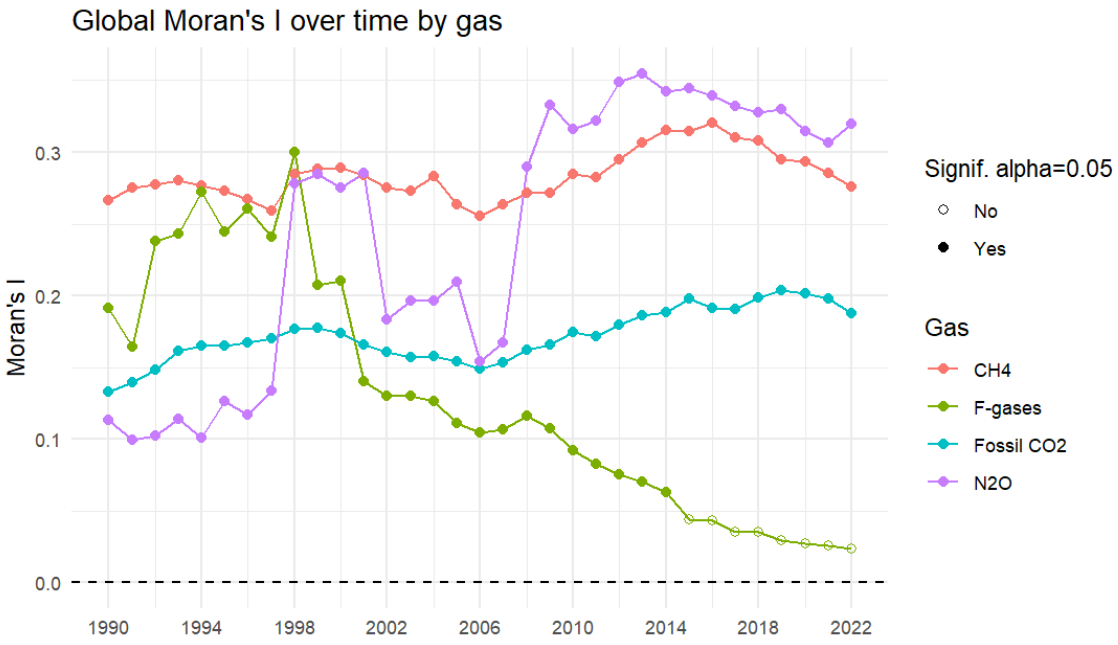


Figure 3.B.1: Global Moran's I for NUTS-2 emissions by gas (CH_4, F-gases, Fossil CO_2, N_2O), computed for each year from 1990 to 2022. Spatial weights W are k-nearest-neighbour with $k = 6$. Filled dots indicate values significantly different from zero at the 95% level (permutation test, $p < 0.05$).

In Figure 3.B.1, N_2O shows a marked increase in spatial autocorrelation in the mid-2000s (approximately 2005–2013), while F-gases decline toward near-zero spatial dependence after the early 2010s.

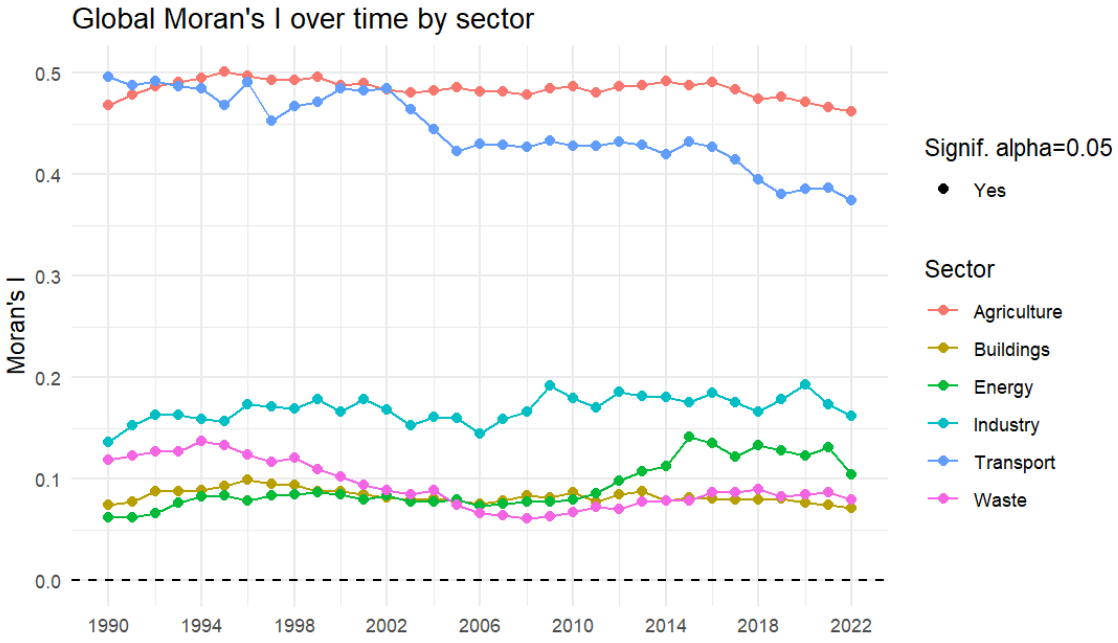


Figure 3.B.2: Global Moran's I for NUTS-2 emissions by sector (Agriculture, Buildings, Energy, Industry, Transport, Waste), computed for each year from 1990 to 2022. Spatial weights W are k-nearest-neighbour with $k = 6$. Filled dots indicate values significantly different from zero at the 95% level (permutation test, $p < 0.05$).

Figure 3.B.2 Industry and Agriculture exhibit persistent, relatively high spatial autocorrelation; Energy is moderate; Buildings, Transport, and Waste are lower, with Transport showing a mild uptick in the late 2010s.

These diagnostics document sustained—but heterogeneous—spatial dependence across regional GHG emissions. They motivate including a spatial component in our clustering framework and help interpret the partitions.

3.C Appendix 3.C: Non-spatial clustering

With the aim of better illustrating the utility of the spatiotemporal clustering methodology employed and emphasizing the role of the spatial component—central to this paper—in this appendix, we present the results obtained from performing the cluster analysis without incorporating the spatial component. This comparison underscores the added value of integrating spatial information in identifying meaningful patterns and dynamics in the data.

Temporal clustering for emissions by gas

Considering dissimilarity matrices $D_{\mathrm{CH_4}}$, $D_{\text{F-gases}}$, $D_{\mathrm{CO_2}}$ and $D_{\mathrm{N_2O}}$, as defined in Section 3.5.2, we set $\Delta\alpha = 0.05$ and $K = 4$ and we obtain ${\alpha_{4,p}}^* = (0.1, 0.35, 0.15, 0.4)$ as the optimal weighting vector for the mixed inertia from the dissimilarity matrices. In Figure 3.C.1 we show a map of the regions belonging to different clusters and in Figure 3.C.2 we provide the average annual emission in each cluster by gas.

The clustering results exhibit a high degree of spatial overlap, yet discernible spatial patterns remain evident. Notably, regions in clusters 1 and 2, which demonstrate the poorest performance, include Belgium, the Netherlands, and certain metropolitan areas in Northern Europe. These regions display a decreasing trend in emissions for the gases with the highest levels, namely $D_{\mathrm{CH_4}}$ and $D_{\mathrm{CO_2}}$. In contrast, cluster 3 consists of sparsely populated regions in Central Europe with significantly lower emissions than the previous clusters. Lastly, the cluster encompassing the largest number of regions, covering over two-thirds of the analyzed area, is characterized by remarkably low emissions and appears relatively stable over time.

Temporal clustering for emissions by sector

Moving to the emissions by sectors, we consider dissimilarity matrices D_{Agri}, D_{Build}, D_{Energy}, D_{Ind}, D_{Tran} and D_{Waste}, as defined in Section 3.5.3, we set $\Delta\alpha = 0.1$ and $K = 5$ and we obtain ${\alpha_{5,p}}^* = (0.3, 0, 0.2, 0, 0.3, 0.2)$ as the optimal weighting vector for the mixed inertia from the dissimilarity

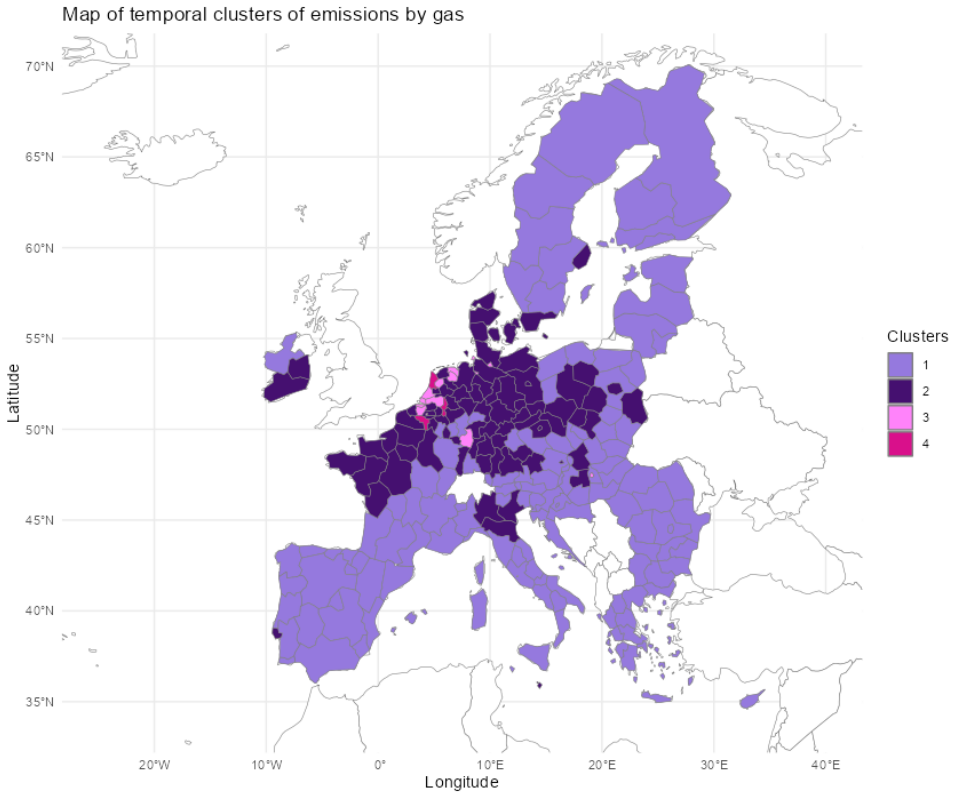


Figure 3.C.1: Map of temporal clusters obtained combining the inertia from the dissimilarity matrices of the time series of emissions of different gases between 1990 and 2022 that is D_{CH_4}, $D_{\text{F-gases}}$, D_{CO_2} and $D_{\mathrm{N}_2\mathrm{O}}$.

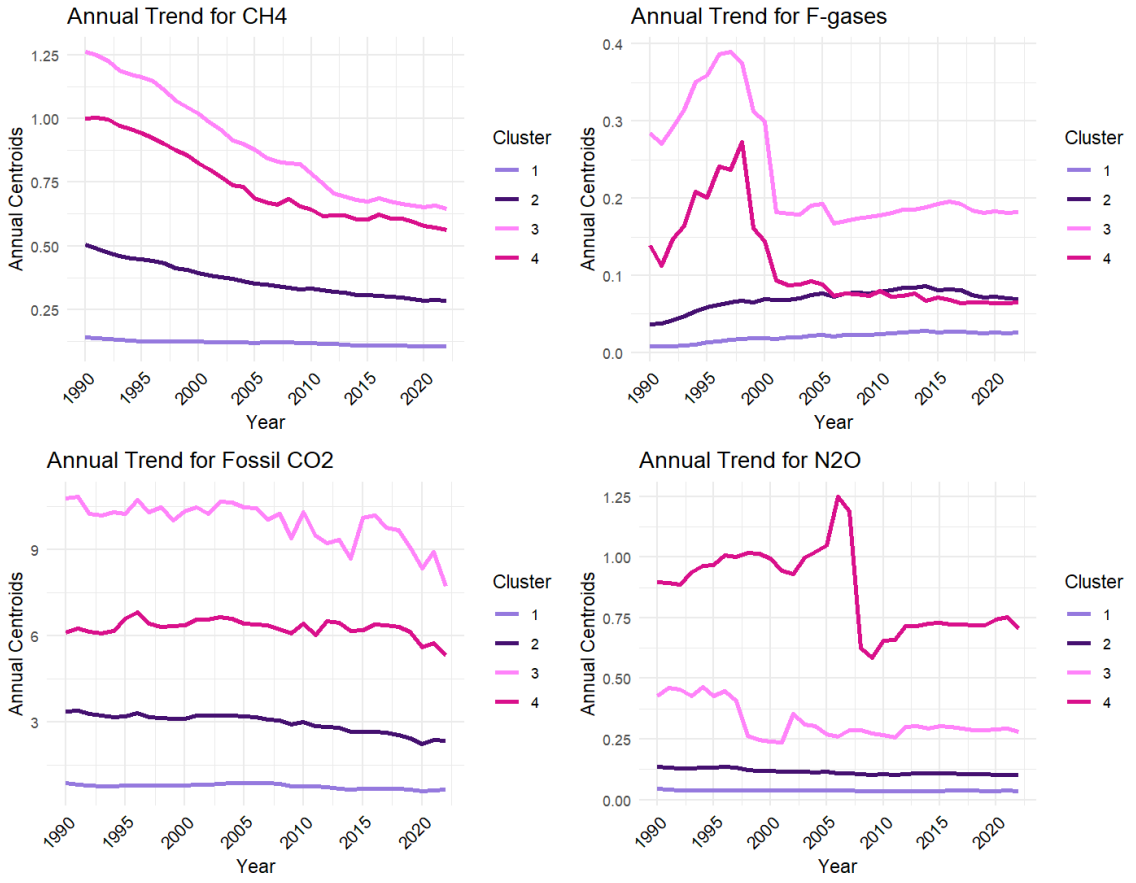


Figure 3.C.2: Annual average of emission per km^2 of sectors for temporal clusters obtained combining the inertia from the dissimilarity matrices of the time series of emissions of different gases between 1990 and 2022 that is $D_{\mathrm{CH_4}}$, $D_{\text{F-gases}}$, $D_{\mathrm{CO_2}}$ and $D_{\mathrm{N_2O}}$.

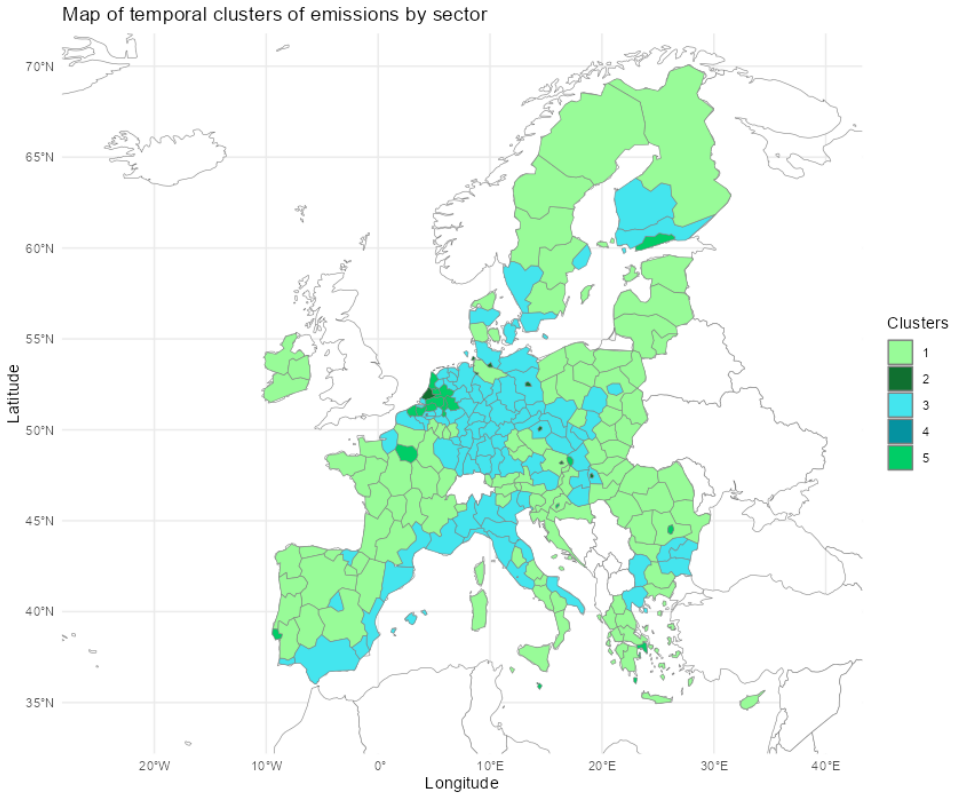


Figure 3.C.3: Map of temporal clusters obtained combining the inertia from the dissimilarity matrices of the time series of emissions of different sectors between 1990 and 2022, that is D_{Agri}, D_{Build}, D_{Energy}, D_{Ind}, D_{Tran} and D_{Waste}.

matrices. In Figure 3.C.3 we provide a map of the regions belonging to different clusters and in Figure 3.C.4 we describe the average annual emission in each cluster by gas.

From the results, we can immediately see that, even without incorporating the spatial component, there is still a degree of spatial overlap in the clustering outcomes. Notably, regions with particularly low performance are concentrated within the European Union. Specifically, Cluster 1, the largest in terms of geographic area, exhibits very low emission levels across all sectors, with no discernible increasing or decreasing trends. Cluster 2 encompasses industrial regions in Northern Europe, characterized by high emissions in the Energy and Industry sectors, which show a declining trend over time. Cluster 3 is associated with low, though not minimal, emission levels, accompanied by a slight downward trend. Cluster 4 includes several metropolitan regions and displays the

Figure 3.C.4: Annual average of emission per km^2 of sectors for temporal clusters obtained combining the inertia from the dissimilarity matrices of the time series of emissions of different sectors between 1990 and 2022, that is D_{Agri}, D_{Build}, D_{Energy}, D_{Ind}, D_{Tran} and D_{Waste}.

highest emission levels in the Buildings, Transport, and Waste sectors. Lastly, Cluster 5, primarily located in Northern Europe and Northern Italy, stands out as the largest emitter in the Agricultural sector.

Overall, our analysis confirms the significant relevance of the spatial component. Regions that are geographically close tend to exhibit similar emission patterns, both by gas type and sector. Incorporating the spatial component explicitly into the clustering algorithm produces groups that are more interpretable in terms of geographic proximity. This leads to results that are particularly valuable in environmental statistics, where spatiotemporal approaches and methodologies are essential for understanding and addressing complex phenomena.

3.D Appendix 3.D: Internal validation of clustering results

To strengthen both the methodological grounding and the practical evaluation, we provide an internal validation of the clustering partitions. For each set of dissimilarities (by gas and by sector), we compare the partition obtained with the proposed spatiotemporal approach to those produced in the purely temporal and purely spatial cases. We compute four standard indices with `fpc::cluster.stats` in R: (i) the average silhouette, which summarizes how well each unit fits its assigned cluster relative to the nearest alternative (Kaufman & Rousseeuw, 1990; Rousseeuw, 1987); (ii) the Calinski–Harabasz (CH) index (Caliński & Harabasz, 1974; Hennig, 2014), which weighs between–cluster dispersion against within–cluster dispersion; (iii) the within/between (WB) ratio of average distances, a compact measure of cluster separation versus compactness; and (iv) Pearson's gamma, a rank correlation between pairwise dissimilarities and a same/different–cluster indicator that gauges the monotonic association between the clustering and the dissimilarity structure (Halkidi et al., 2001).

The average silhouette quantifies, for each unit, how close it is to its own cluster compared with the nearest alternative; values lie in $[-1, 1]$, with larger values indicating better cohesion and separation (Kaufman & Rousseeuw, 1990; Rousseeuw, 1987). In Fig. 3.D.1a and Fig. 3.D.2a, silhouettes are generally modest—as expected for high–dimensional time–series dissimilarities—but the blue bars (spatiotemporal partition) typically fall between the orange (temporal-only) and green (spatial-only) bars. This confirms that the spatiotemporal solution strikes a compromise: it inherits part of the cohesion visible when clustering on temporal features while retaining some spatial structure, whereas the purely spatial solution yields good cohesion only with respect to the spatial dissimilarity.

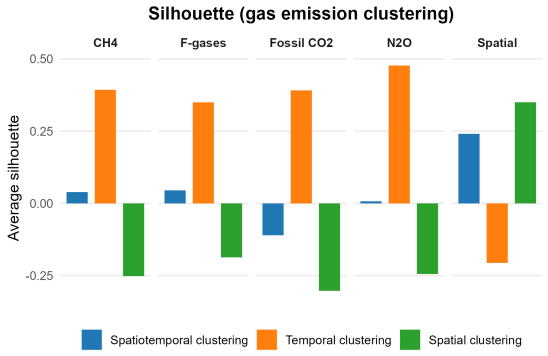
The Calinski–Harabasz (CH) index balances between–cluster dispersion against within–cluster dispersion; higher values imply tighter, well-separated clusters. Panels (b) of Figs. 3.D.1 and 3.D.2 show that temporal-only partitions tend to score higher when evaluated on feature-based dissimilarities, while the spatial-only solution dominates when the spatial dissimilarity is used. The spatiotemporal partition lies in between, reinforcing its role as a balanced allocation.

The within/between (WB) ratio summarizes compactness versus separation as the ratio of av-

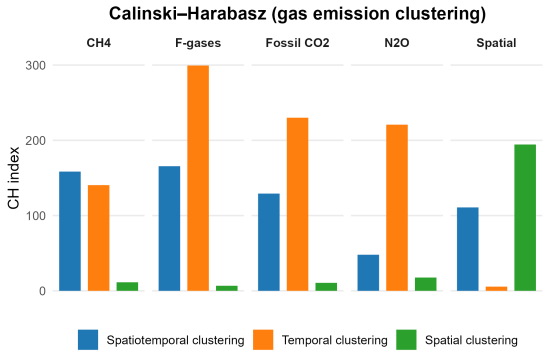
erage within-cluster to average between-cluster dissimilarities (smaller values are preferable). In Figs.3.D.1c and 3.D.2c, the blue bars again occupy an intermediate position between the temporal-only and spatial-only cases, consistent with a partition that trades off feature cohesion and geographic proximity.

Finally, Pearson's gamma measures the monotone association between pairwise dissimilarities and the same/different–cluster indicator (Halkidi et al., 2001). Larger values indicate that pairs declared "in different clusters'' are, in rank terms, more dissimilar than pairs "in the same cluster''. Panels (d) show the same qualitative pattern observed for CH and WB: the temporal-only clustering aligns best with feature-based dissimilarities, the spatial-only clustering aligns best with the spatial dissimilarity, and the spatiotemporal partition achieves a reasonable balance across both information sources.

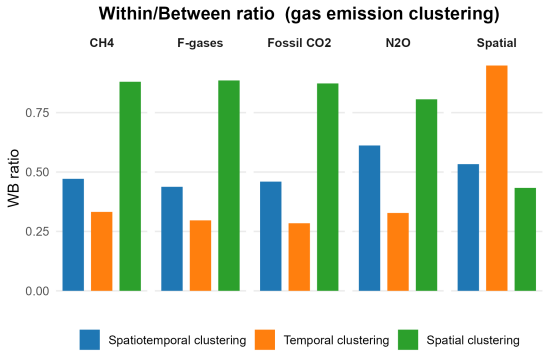
Taken together, these diagnostics (Figs. 3.D.1–3.D.2) corroborate the main message of the paper: the spatiotemporal partition provides a principled middle ground between purely temporal and purely spatial solutions, preserving meaningful geographic structure without sacrificing too much feature-based cohesion.



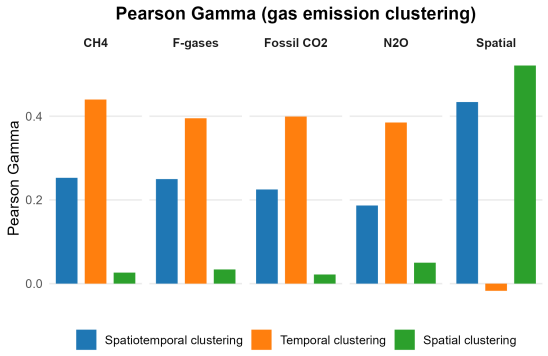
A. Average silhouette for gas-specific dissimilarities by clustering partition.



B. Calinski–Harabasz (CH) index for gas-specific dissimilarities by clustering partition.



C. Within/Between (WB) ratio for gas-specific dissimilarities by clustering partition.

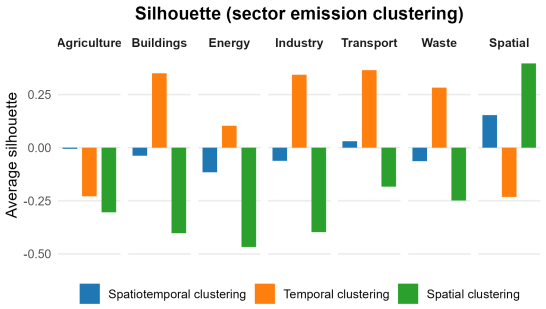


D. Pearson's gamma for gas-specific dissimilarities by clustering partition.

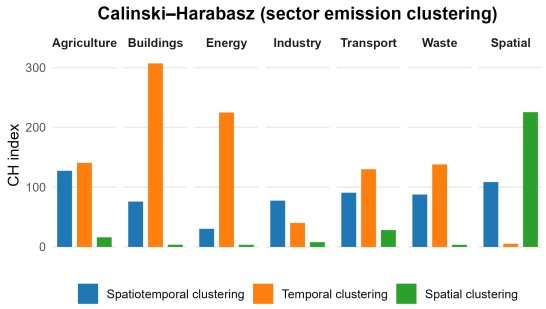
Figure 3.D.1: Internal validation indices for gas-specific dissimilarities by clustering partition (panels a–d).

3.E Appendix 3.E: Spatiotemporal clustering of single time series

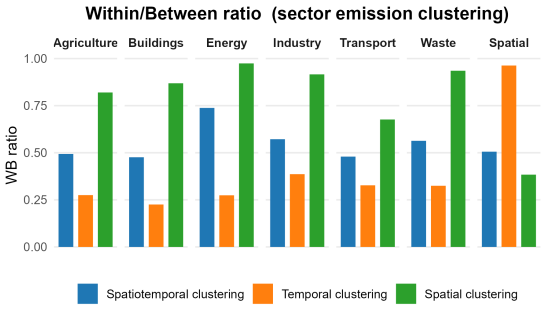
In this appendix we provide the clustering partitions obtained by combining the inertia from dissimilarity matrix of the emission time series for each gas with the spatial dissimilarity, and the dissimilarity matrix of the emission time series for each sector with the spatial dissimilarity, thus combining only two dissimilarity matrices each time. These results help assess how stable the spatial grouping is across environmental dimensions and where gas- or sector-specific patterns emerge.



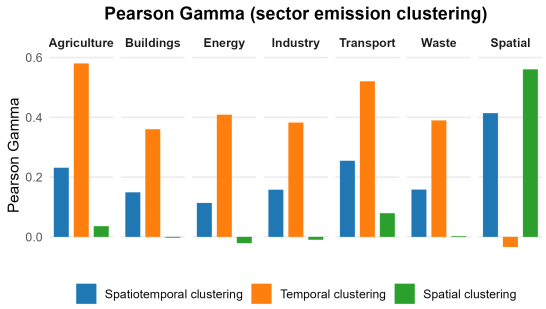
A. Average silhouette for sector-specific dissimilarities by clustering partition.



B. Calinski–Harabasz (CH) index for sector-specific dissimilarities by clustering partition.



C. Within/Between (WB) ratio for sector-specific dissimilarities by clustering partition.



D. Pearson's gamma for sector-specific dissimilarities by clustering partition.

Figure 3.D.2: Internal validation indices for sector-specific dissimilarities by clustering partition (panels a–d).

	$D_{\mathrm{CH_4}}$	$D_{\text{F-gases}}$	$D_{\mathrm{CO_2}}$	$D_{\mathrm{N_2O}}$
α	0.60	0.65	0.40	0.40
$JI(\mathcal{P}_K^{\alpha^*})$	0.58	0.70	0.60	0.25

Table 3.E.1: Summary of the α^* weights of the feature dissimilarities in the spatiotemporal clustering of the different partitions obtained combining the inertia from dissimilarity matrix of the time series of emissions by each gas between 1990 and 2022 with the inertia from spatial dissimilarity matrix, setting $K = 4$. $JI(\mathcal{P}_K^{\alpha^*})$ is the Joint Inertia resulting from each partition. Note that the joint inertia when considering only two dissimilarity matrices, remains the same whether it is calculated with respect to one matrix or the other dissimilarity matrix.

In Figure 3.E.1 CO_2, F-gases, and CH_4 display a similar backbone: three large, spatially compact clusters that roughly separate Western/South-Western, Central, and Eastern Europe, plus a fourth, partially overlapping cluster concentrated in the Northern/Central belt that gathers cross-

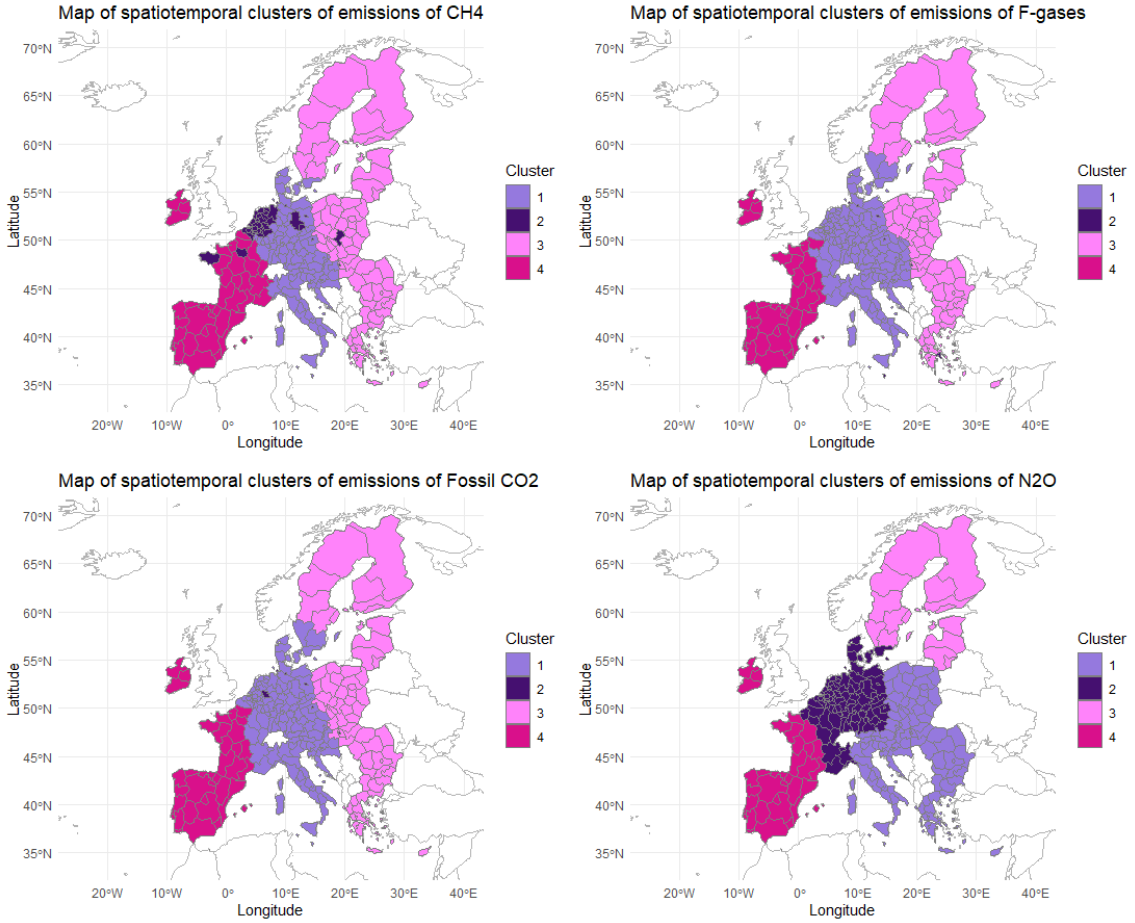


Figure 3.E.1: Map of spatiotemporal clusters obtained combining the inertia from dissimilarity matrices of the time series of emissions of each gas between 1990 and 2022 with the spatial dissimilarities. We compute the optimal α of each combination according to Morelli et al., 2025, setting $K = 4$.

country regions with comparable temporal trajectories. By contrast, N_2O departs from this pattern and yields four compact, contiguous clusters with much less overlap, producing a sharper spatial segmentation.

According to Table 3.E.1, F-gases, CH_4, and Fossil CO_2 exhibit high Joint Inertia (0.70, 0.60, and 0.60, respectively), indicating that the resulting partitions capture a large proportion of the dissimilarity from the time-series and spatial dissimilarities. By contrast, N_2O shows a low Joint Inertia (0.25), meaning the partition is not able to capture the inertia from both the spatial dissimilarity and the N_2O time series. The selected α values differ across gases, underscoring that the trade-off between temporal features and space is gas-specific.

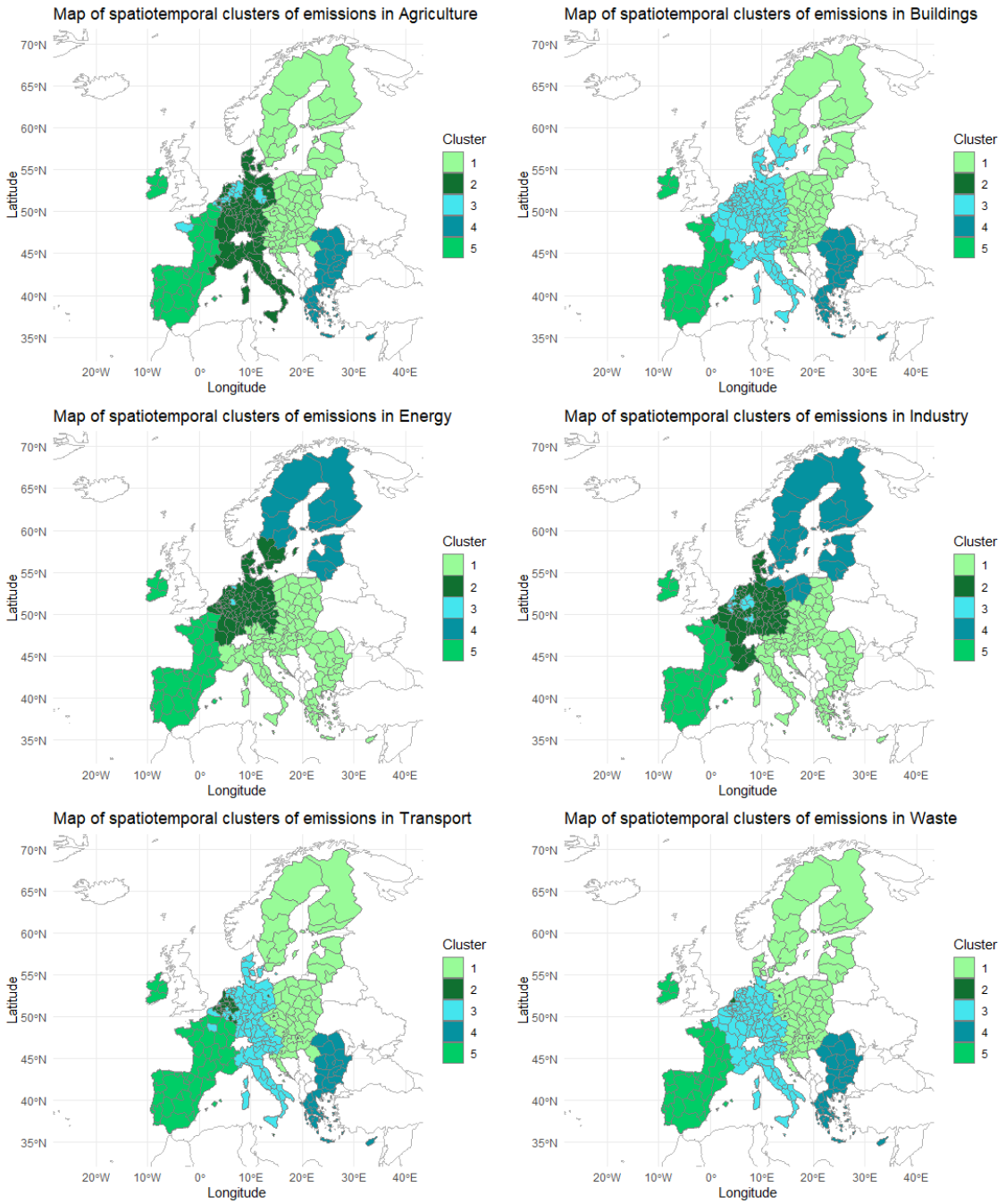


Figure 3.E.2: Map of spatiotemporal clusters obtained combining the inertia from dissimilarity matrix of the time series of emissions by each sector between 1990 and 2022 with the inertia from spatial dissimilarity matrix. We compute the optimal α of each combination according to Morelli et al., 2025, setting $K = 5$.

	D_{Agri}	D_{Build}	D_{Energy}	D_{Ind}	D_{Tran}	D_{Waste}
α	0.50	0.50	0.45	0.70	0.65	0.70
$JI(\mathcal{P}_K^{\alpha^*})$	0.65	0.63	0.54	0.55	0.73	0.60

Table 3.E.2: Summary of the α^* weights of the feature dissimilarities in the spatiotemporal clustering of the different partitions obtained combining the inertia from the dissimilarity matrix of emissions in each sector and the inertia from spatial dissimilarity matrix, setting $K = 5$. $JI(\mathcal{P}_K^{\alpha^*})$ is the Joint Inertia resulting from each partition. Note that the joint inertia when considering only two dissimilarity matrices, remains the same whether it is calculated with respect to one matrix or the other dissimilarity matrix.

In Figure 3.E.2, sectoral partitions show a common backbone: four compact groups that are largely similar across sectors, plus a fifth, overlapping cluster that typically sits in Northern/Central Europe and often captures metropolitan/high-density corridors. Despite this shared structure, boundaries and memberships differ slightly by sector—most visibly for Energy and Industry, which deviate more from the common pattern. Transport, Buildings, and Waste each display four markedly contiguous blocks plus a distinct cluster that typically includes metropolitan and other high-density regions in Central Europe. Agriculture shows a similar pattern—four compact clusters and another grouping that includes agricultural regions in Northern Europe. By contrast, Energy and Industry present a different structure, with a compact cluster spanning Germany, the Netherlands, Denmark, and eastern France, alongside a smaller cluster of regions that spatially overlaps the former.

In Table 3.E.2 Transport shows the highest Joint Inertia (0.73), pointing to a strong role of geography with respect to the dissimilarity in the time series of Transport emission, followed by Agriculture, Buildings and Waste, respectively 0.65, 0.63, 0.60. Energy and Industry display more moderate interplay between spatial information and feature information (0.54–0.55). The selected α values differ across sectors, indicating sector-specific balances between temporal features and spatial proximity.

Overall, the partitions are not identical across across gases and sectors but they share a common macro-regional backbone. Moreover, the Joint Inertia values and confirm that the relevance of spatial information is dimension-specific.

Chapter 4

Cluster-wise Spatiotemporal Panel Model

Work in progress research

I am grateful to Paolo Maranzano and Philipp Otto
for insightful comments and constructive suggestions.

4.1 Introduction

The growing interconnectedness of the global economy has made it increasingly difficult to view countries, regions, firms or individuals as isolated entities. Economic, technological, and environmental developments now spread rapidly across borders, linking markets and shaping outcomes in ways that traditional, independent-unit models struggle to capture. Such interdependence may stem from direct strategic interactions—when the decisions of one agent influence others—or from broader, shared forces, such as global technological change, shifts in regulation, or common macroeconomic shocks. At the same time, the degree and nature of these interconnections often depend on the intrinsic characteristics of the units themselves—such as their economic strength, institutional capacity, or technological readiness—which determine not only how strongly they can influence others but also how easily they can adapt to or emulate external changes.

To account for these connections, two major modeling traditions have emerged in the econometric literature. The first, the spatial approach, assumes that the behavior of one unit is directly affected by that of its neighbors, whether through geographical proximity, trade links, or economic integration. The strength of these relationships is often captured using a spatial weighting matrix that encodes how strongly each unit is connected to others. The second, the factor-based approach, models dependence as the result of unobserved common components—latent global shocks or trends—that influence all units, albeit with varying intensity. While these two frameworks were initially developed in parallel, recent advances increasingly seek to bridge them, offering unified ways to capture both local and global sources of interdependence (Chen et al., 2022; Koopman et al., 2021; Su et al., 2023). The spatial sciences, particularly geography and regional economics, provide a solid conceptual foundation for understanding spatial dependencies. A central insight in this literature is encapsulated in the well-known *First Law of Geography* (Tobler, 1970), which states that "everything is related to everything else, but near things are more related than distant things". This principle underlies the pervasive presence of spatial autocorrelation—the tendency for nearby units to exhibit similar values due to direct or indirect interactions across space. These interactions often manifest as spatial spillovers, that is, effects in one unit that influence outcomes in geographically or economically connected areas (Elhorst, 2024; Elhorst et al., 2021). Such spillovers are not limited to geographic contiguity: they may arise from structural connections, such as trade,

labor mobility, or vertical industrial integration. For instance, a shock in one region or sector may propagate to others via trade links or supply chain dependencies.

A second, equally important spatial principle is the *Second Law of Geography*, also known as spatial heterogeneity (Goodchild, 2004; Zhu & Turner, 2022; Zhu et al., 2018). This law emphasizes that the relationships between variables are not constant across space; instead, they vary with location due to differences in economic structure, social institutions, infrastructure, or policy environments. Such heterogeneity violates the assumption of identically distributed relationships across observations and renders classical regression models inadequate. In this setting, statistical models that fail to account for spatial variation may suffer from misspecification, resulting in biased or inconsistent parameter estimates (Anselin & Arribas-Bel, 2013; J. LeSage & Pace, 2009).

The importance of modeling both spatial dependence and heterogeneity has spurred the development of spatially-structured clustering techniques. These approaches, including recent extensions of clusterwise regression models (Billé et al., 2017; J. Lee et al., 2017; Sugasawa & Murakami, 2021), seek to identify groups of spatial units that not only exhibit internally similar behavior but also account for the spatial organization of the data. For example, penalized spatial clustering models integrate neighborhood structure—such as that encoded by Potts-type priors—to encourage geographically cohesive clusters, while simultaneously estimating group-specific regression coefficients. In doing so, they offer a unified framework to tackle spatial dependence and spatial heterogeneity together (Boccaletti et al., 2024; Cerqueti et al., 2025; R. et al., 2025).

Yet, many empirical applications continue to treat spatial units as homogeneous, ignoring the substantial differences in structural and behavioral characteristics that exist across regions, cities, or countries. These differences can relate to human capital, technological capacity, governance quality, and environmental exposure. For instance, the same environmental regulation may elicit different responses across regions, depending on local industrial composition or policy capacity. A technologically advanced region may quickly adopt green innovations, while a neighboring, less developed area may struggle to do so, leading to asymmetric spatial externalities.

Several recent contributions have advanced the modeling of spatial heterogeneity and cross-sectional dependence in panel data settings. Aquaro et al., 2021 develop a heterogeneous spatial autoregressive model that allows for unit-specific parameters in both the spatial lag and covariate

effects, estimated via a quasi maximum likelihood approach. Su et al., 2023 extend this framework by incorporating both spatial and temporal dynamics as well as individual and time fixed effects, offering a GMM-based estimation strategy that accommodates heteroskedasticity. In parallel, Chen et al., 2022 propose a unifying econometric framework for heterogeneous panels with spatial and factor dependence, while Koopman et al., 2021 address both global and local cross-sectional dependence in a dynamic panel context. From a different perspective, Bailey et al., 2016 propose a two-step estimation approach that controls for common factors through cross-sectional averages, allowing for heterogeneity in spatiotemporal dynamics at the unit level. Similarly, Shi and Lee, 2017 and Chen et al., 2022 explore interactive fixed effects and multifactor structures to model unobserved heterogeneity in spatial panels. Recent contributions such as Feo et al., 2024 also consider spatial dynamic panel models with heterogeneity in both spatial lags and their temporal evolution.

Despite these advances, most of the literature focuses on unit-specific heterogeneity or models with mean group estimators, without explicitly accounting for clustered heterogeneity in spatial dependence. That is, existing approaches typically do not distinguish between within-group and between-group spillovers. To the best of our knowledge, no prior study has proposed a cluster-wise framework that accommodates heterogeneous spatial interactions both within and across latent groups. Our work addresses this gap by developing a model that flexibly captures such complex interdependencies.

In this paper, we propose a novel cluster-wise spatial dynamic panel model that captures these dimensions in a unified framework. Our approach integrates (*i*) a spatial dynamic panel structure that models spillovers through spatial connectivity matrices (Elhorst, 2010; Elhorst et al., 2021); and (*ii*) a clusterwise regression design in which parameters vary across latent groups (Bonhomme & Manresa, 2015). Crucially, we allow for spatial dependence to differ both within and across clusters, thus accommodating heterogeneity in both marginal effects and spatial transmission mechanisms.

The proposed model is flexible and interpretable. It assigns each region to a cluster with group-specific dynamics, while allowing for rich spatial interconnections across all units. Spatial spillovers are weighted not only by proximity but also by the identity of the sending and receiving groups, thereby modeling both symmetric and asymmetric effects. This design is especially relevant in contexts such as greenhouse gas (GHG) emissions, where both regional disparities and spatial

externalities are critical. By embedding cluster-specific behavior into the spatial dynamics, our model offers a powerful tool to analyze spatially heterogeneous processes with interdependent units.

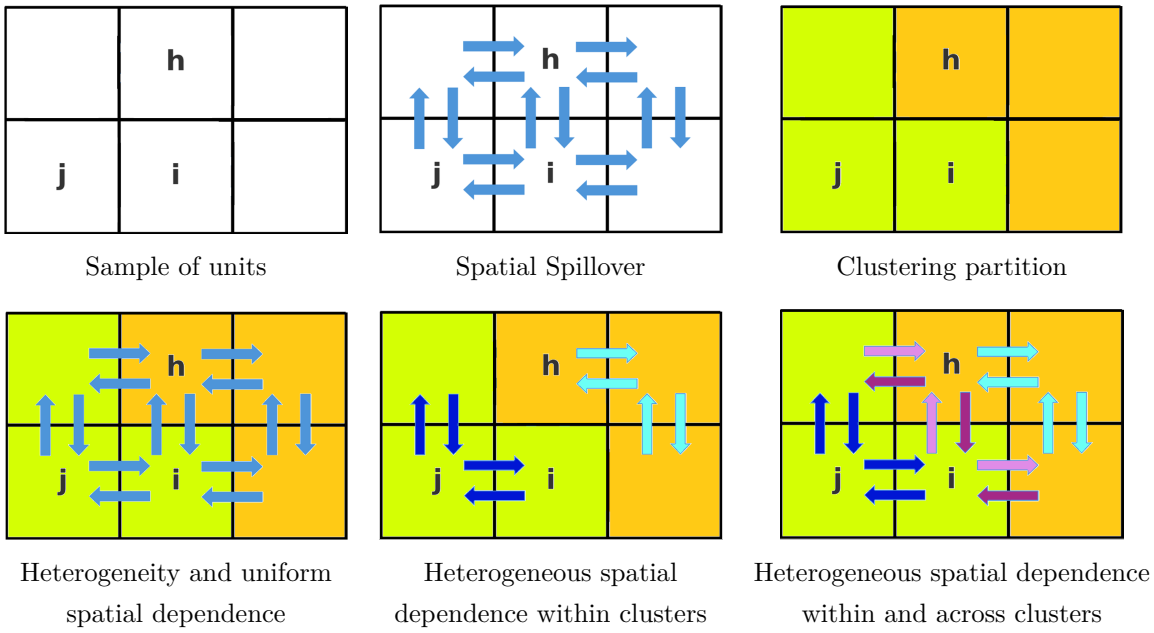


Figure 4.1.1: Conceptual representation of spatial dependence and heterogeneity by clusters.

Figure 4.1.1 provides a step-by-step conceptual illustration of the key mechanisms captured by our model. The top-left panel displays a generic spatial grid of units, labeled i, j, and h, which serves as the spatial reference structure. The top-center panel introduces the idea of spatial spillovers: here, blue arrows represent the bilateral influence that neighboring regions exert on one another, irrespective of their individual characteristics. The top-right panel adds a partitioning structure, where each unit belongs to a specific cluster (indicated by different background colors), capturing the notion of group-level heterogeneity.

The bottom-left panel combines these two aspects by illustrating uniform spatial dependence across heterogeneous clusters. Spatial spillovers (same arrows for all units) are present, but the internal structure of each group (yellow vs. green background) is heterogeneous. The bottom-center panel introduces cluster-specific spatial dependence: spillovers still occur only within each cluster, but the intensity and pattern differ across clusters, as shown by the variation in arrow color and direction. This reflects the idea that regions within the same group may interact more strongly

with each other than with regions outside their group.

Finally, the bottom-right panel captures the most general and flexible case: heterogeneous spatial dependence both within and across clusters. Arrows of different colors and directions represent varying spatial influence across unit pairs, with the structure depending not only on spatial proximity but also on the cluster identity of both the sending and receiving units. This is the most general setting captured by our proposed model, which is designed to flexibly accommodate both spatial dynamics and structured group heterogeneity.

In summary, our contribution lies in the development of a flexible spatial dynamic panel model that simultaneously captures interregional dependencies and cluster-based heterogeneity in both covariate effects and spatial spillovers. We provide a quasi–maximum likelihood (QML) estimation procedure to estimate the heterogeneous spatial and temporal dependence parameters, conditional on a given clustering partition. Furthermore, we design a comprehensive simulation experiment to assess the performance of the proposed estimation method in recovering the true parameters and in accurately identifying the direct and indirect effects implied by the model, considering different model specification and heterogeneity settings. In addition, we introduce a two-step procedure that enables the joint identification of model parameters and endogenous cluster membership, thereby eliminating the need to predefine a partition structure.

Looking ahead, we intend to apply this framework to real-world environmental data to demonstrate its practical relevance. Prior studies (Morelli et al., 2025) have revealed strong evidence of spatial clustering in GHG emission patterns and heterogeneous dynamics across regions. These findings validate our modeling approach and motivate further investigation using our new framework. By integrating spatial dynamics with group-specific heterogeneity, our work contributes to the growing literature on the spatiotemporal modeling of economic and environmental phenomena, offering valuable insights for more effective and targeted regional policy design.

The remainder of the paper is organized as follows. Section 4.2 introduces the proposed model and outlines its main features. Section 4.3 discusses the possible structures of parameter heterogeneity. In Section 4.4, we derive the expressions for the direct and indirect effects, distinguishing between cluster-specific and cross-cluster components. Section 4.5 presents the estimation procedure, while Section 4.6 describes the simulation experiment and discusses the corresponding results.

Finally, Section 4.7 concludes the paper.

4.2 Methodology

4.2.1 Cluster-wise Spatiotemporal Panel Model definition

Our model incorporates both spatial dependence and group heterogeneity in a dynamic panel data framework. Suppose we have a panel data, with N units for T time stamp, thus y_{it} is the dependent variable for unit i at time t, $\mathbf{x}_{it}$ is the vector of k explanatory variables for the same unit and time. Moreover, let $\mathcal{P} = (\mathcal{C}_1, \ldots, \mathcal{C}_G)$ being the partition in G number of clusters. Each unit in our sample belongs to one and only one cluster for the entire time period.

Also, suppose we have spatial information regarding each unit i with respect to the other units j, which can be interpreted as a measure of spatial proximity or boundaries, and is quantified and defined as w_{ij}. The spatial dependencies in the SDM are captured using a spatial weight matrix W, which is an $N \times N$ matrix where N is the number of spatial units. Each element w_{ij} in W represents the spatial proximity or influence between spatial unit i and unit j. By convention, $w_{ii} = 0$, as no unit is considered its own neighbor (Elhorst, 2010).

To ensure model stability, the spatial weight matrix W must satisfy several assumptions Elhorst, 2010; J. LeSage and Pace, 2009; J. P. LeSage, 2008. First, W should be a non-negative matrix of known constants. Additionally, W is often row-normalized, meaning the elements in each row of W sum to one, which ensures that the influence of neighboring units is proportionally scaled. For symmetric W, the eigenvalues must lie within the interval $(1/\omega_{\max}, 1)$, where $\omega_{\max}$ is the largest real eigenvalue of W. If W is not normalized, the interval instead depends on $\omega_{\min}$ and $\omega_{\max}$, the smallest and largest real eigenvalues of W, respectively. Finally, the matrices $I - \rho W$ and $I - \lambda W$ must be non-singular, where I is the identity matrix of size $N \times N$, and ρ and λ are spatial autoregressive coefficients.

Our model can be defined as follows:

$$y_{it} = \sum_{j=1}^{N}\sum_{f=1}^{G} \rho_{gf} w_{ij} y_{jt} I(j \in \mathcal{C}_f) + \tau_g y_{i,t-1} + \sum_{j=1}^{N}\sum_{f=1}^{G} \eta_{gf} w_{ij} y_{j,t-1} I(j \in \mathcal{C}_f) + \mathbf{x}'_{it}\boldsymbol{\beta}_g + \sum_{j=1}^{N}\sum_{f=1}^{G} w_{ij}\mathbf{x}'_{jt}\boldsymbol{\delta}_{gf} I(j \in \mathcal{C}_f) + \epsilon_{it}, \quad i \in \mathcal{C}_g \tag{4.2.1}$$

The parameter ρ_{gf} captures the contemporaneous spatial spillover from units in cluster f to units in cluster g, weighted by the spatial proximity defined through w_{ij}. When $g = f$, ρ_{gg} measures within-cluster endogenous spatial dependence, reflecting how the current outcomes of spatially neighboring units within the same cluster influence each other. For simplicity, in this case the notation will be simplified as $\rho_g := \rho_{gg}$. When $g \neq f$, ρ_{gf} quantifies cross-cluster spatial spillovers, that is, the extent to which units in cluster f contemporaneously affect the outcomes of neighbor units belonging to cluster g.

The parameter τ_g represents the group-specific temporal dependence. It captures the autoregressive effect of the lagged dependent variable $y_{i,t-1}$ on the current outcome y_{it} for units in cluster g. Thus, the strength of this persistence may vary by group.

The parameter η_{gf} accounts for lagged spatial dependence across clusters. Specifically, it measures the effect of past outcomes $y_{j,t-1}$ of units in cluster f on the current outcomes of units in cluster g, again weighted by the spatial structure w_{ij}. When $g = f$, this represents within-cluster dynamic spatial spillovers, similarly to the contemporaneous spatial dependence, the notation will be simplified as $\eta_g := \eta_{gg}$. When $g \neq f$, it captures delayed cross-cluster spatial interactions.

The vector $\boldsymbol{\beta}_g$ contains the group-specific coefficients associated with the set of k covariates $\mathbf{x}_{it}$. These coefficients represent the direct marginal effects of explanatory variables on the outcome for units in cluster g, allowing for heterogeneous covariate effects across groups.

The parameter vector $\boldsymbol{\delta}_{gf}$ captures the spatial spillovers of k covariates. It measures the impact of covariates $\mathbf{x}_{jt}$ from units in cluster f on the outcome y_{it} of units in cluster g, with the effect weighted by the spatial link w_{ij}. When $g = f$, this term reflects within-cluster spatial externalities of covariates. When $g \neq f$, it accounts for cross-cluster spatial spillovers of covariate effects.

Finally, the error term ϵ_{it} captures unobserved, idiosyncratic disturbances, assuming that $\epsilon_{it} \sim \mathcal{N}(0, \sigma^2)$.

Altogether, the model is designed to flexibly capture both temporal dynamics and spatial dependencies in the data while accounting for heterogeneity in behavior across different clusters. The structure permits not only endogenous and exogenous spatial effects, but also allows these effects to vary depending on whether interactions occur within the same group or across different groups.

The specification in Equation (4.2.1) represents the most general form of our proposed model, encompassing all possible sources of spatial dependence, temporal dynamics, and their cross-group interactions. However, as also emphasized in the spatial econometric literature (Anselin et al., 2008; Elhorst, 2010), such a general model may suffer from identification issues when all parameters are included simultaneously. In practice, it is common to work with restricted versions of this model, where specific parameters are set to zero based on theoretical assumptions or empirical considerations.

Therefore, in the remainder of this paper, we treat the full model (4.2.1) primarily as a theoretical benchmark, used to illustrate the conceptual framework and to describe the most general case. Nonetheless, our estimation and empirical strategy will focus on more tractable, restricted versions. A comprehensive overview of these nested specifications—including the Spatial Autoregressive Model (SAR), Spatial Durbin Model (SDM), and other dynamic extensions—is provided in Elhorst, 2001, where all potential restrictions are discussed systematically, up to the case of the standard linear regression model.

4.2.2 Restriction to Spatial Dynamic panel model

If we remove the heterogeneity in the model—namely, we assume that all units belong to a single cluster and are governed by a common set of parameters—we recover the well-known Spatial Durbin Dynamic model, as presented in its general form by Elhorst, 2001. The model is expressed as:

$$y_{it} = \rho \sum_{j=1}^{N} w_{ij} y_{jt} + \tau y_{i,t-1} + \eta \sum_{j=1}^{N} w_{ij} y_{j,t-1} + \mathbf{x}'_{it}\boldsymbol{\beta} + \sum_{j=1}^{N} w_{ij} \mathbf{x}'_{jt}\boldsymbol{\delta} + \epsilon_{it} \tag{4.2.2}$$

Equation 4.2.2 can be viewed as a special case of the more general heterogeneous model previously described, obtained by imposing the restriction that all parameters $\rho, \tau, \eta, \boldsymbol{\beta}, \boldsymbol{\delta}$ are common to all units. Several restricted versions of this model can be derived by constraining one or more of these parameters, as discussed in Elhorst, 2010; J. LeSage and Pace, 2009. In this specification, the effects of the explanatory variables and their spatial lags, as well as the spatial and temporal dependence of the dependent variable, are assumed to be homogeneous across space and time. While this allows the model to capture average spatial spillovers and dynamic effects, it does not account for potential heterogeneity in how units respond to covariates or interact with their neighbors—an important limitation when units are diverse or grouped by underlying structural differences.

4.2.3 Restriction to heterogeneous cluster model

If we restrict the model in Equation 4.2.1 by removing both spatial and temporal dependence, we obtain a cluster-wise heterogeneous model as proposed by Bonhomme and Manresa, 2015, which takes the form:

$$y_{it} = \mathbf{x}'_{it}\boldsymbol{\beta}_g + \epsilon_{it}, \quad i \in \mathcal{C}_g \tag{4.2.3}$$

In this specification, units are assumed to be independent of one another, implying that neither spatial proximity nor common economic characteristics generate interaction effects across units. Each unit evolves independently, unaffected by its neighbors. However, the model retains heterogeneity through the clustering structure: for each unit i belonging to a given cluster $\mathcal{C}_g$, a specific set of coefficients $\boldsymbol{\beta}_g$ governs the relationship between the covariates and the outcome. This allows the model to capture structural differences across groups while abstracting from spatial and dynamic

interdependence.

4.2.4 Restriction to Pooled panel model

If we further restrict the model in Equation 4.2.1 by removing spatial and temporal dependence as well as the cluster-based heterogeneity, we obtain the standard pooled regression model for panel data. This specification assumes that all units share a common set of parameters and evolve independently over time and space. The resulting model is given by:

$$y_{it} = \mathbf{x}_{it}'\boldsymbol{\beta} + \epsilon_{it}, \quad \forall i = 1, \ldots, N; \ \forall t = 1, \ldots, T \tag{4.2.4}$$

This pooled model imposes homogeneity across all cross-sectional units and time periods, thereby ignoring both interaction effects and structural heterogeneity. While computationally simple, such a specification may be too restrictive in settings where spatial spillovers, temporal dynamics, or unit-specific behaviors play a substantive role.

4.2.5 Matrix Representation of the Model

To express the model in compact matrix notation, let us define the key components across all N units and T time periods. Denote $Y_t = (y_{1t}, \ldots, y_{Nt})'$ as the $N \times 1$ vector of dependent variables at time t, and $X_t = (\mathbf{x}_{1t}', \ldots, \mathbf{x}_{Nt}')'$ as the $N \times k$ matrix of covariates. Let $W \in \mathbb{R}^{N \times N}$ be the spatial weight matrix, and $P \in \{0,1\}^{G \times N}$ the group membership matrix, where each column of P has exactly one entry equal to 1, indicating the group membership of each unit.

The group-level parameters are collected as follows:

- $\boldsymbol{\tau} \in \mathbb{R}^G$: vector of group-specific temporal autoregressive coefficients;
- $\boldsymbol{\beta} \in \mathbb{R}^{G \times k}$: matrix of group-specific covariate effects;

- $\boldsymbol{\rho}, \boldsymbol{\eta}, \boldsymbol{\delta} \in \mathbb{R}^{G \times G}$: matrices of group-to-group spatial effects for the contemporaneous dependent variable, its lag, and the covariates, respectively.

Using this notation, the full model in Equation 4.2.1 can be rewritten as:

$$\begin{aligned} Y_t = &\left(P^\top \boldsymbol{\rho} P \circ W\right) Y_t + \operatorname{diag}(P^\top \boldsymbol{\tau}) Y_{t-1} + \left(P^\top \boldsymbol{\eta} P \circ W\right) Y_{t-1} \\ &+ \operatorname{diag}_k(P^\top \boldsymbol{\beta}) X_t + \left(P^\top \boldsymbol{\delta} P \circ W\right) X_t + \epsilon_t, \end{aligned} \tag{4.2.5}$$

where:

- $\epsilon_t \sim \mathcal{N}(0, \sigma^2 I_N)$ is the vector of idiosyncratic errors;
- $\operatorname{diag}(P^\top \boldsymbol{\tau}) \in \mathbb{R}^{N \times N}$ is a diagonal matrix assigning τ_g to units in group g;
- $\operatorname{diag}_k(P^\top \boldsymbol{\beta})$ represents a block-diagonal matrix of dimension $N \times k$, where each row i takes the corresponding $\boldsymbol{\beta}_g$ according to unit i's group;
- The term $P^\top \boldsymbol{\rho} P \circ W$ is the element-wise (Hadamard) product between the matrix assigning each dyadic pair (i, j) the appropriate ρ_{gf} based on their group membership and the spatial weight matrix W. The same applies to the $\boldsymbol{\eta}$ and $\boldsymbol{\delta}$ terms.

This compact matrix formulation highlights how group-level heterogeneity enters the spatial and temporal dynamics of the model. Each unit's outcome depends on its own lag, the outcomes and covariates of its spatial neighbors, and all corresponding effects vary by the group of both the source and receiving units. This notation will be used in the following sections to describe the structure of heterogeneity and to derive the expressions for direct and indirect effects.

4.3 Types and Structure of Heterogeneity

The proposed model incorporates two distinct dimensions of heterogeneity, which can be classified according to the structure of the model parameters. Let us define the complete set of model parameters as Θ, which can be partitioned into two subsets:

- $\boldsymbol{\varphi}_g = \{\boldsymbol{\beta}_g, \tau_g\}$: parameters that depend only on the cluster membership of unit i, with a single index g, corresponding to the group $\mathcal{C}_g$ to which unit i belongs.
- $\boldsymbol{\psi}_{gf} = \{\rho_{gf}, \eta_{gf}, \boldsymbol{\delta}_{gf}\}$: parameters that capture spatial spillover effects between units in cluster g and units in cluster f, hence indexed by both origin and destination groups.

The heterogeneity in $\boldsymbol{\varphi}_g$ is standard and has been studied extensively in the literature, for example in Bonhomme and Manresa, 2015. In this case, the heterogeneity is cluster-wise, meaning that there are G distinct sets of parameters, one for each cluster. These parameters affect only the units belonging to the same group and are not influenced by spatial relationships with other units.

On the other hand, the heterogeneity in $\boldsymbol{\psi}_{gf}$ is more complex. Since these parameters depend on both the source and target cluster, there can potentially be $G \times G$ different values for each element in $\boldsymbol{\psi}_{gf}$. To better understand this structure, consider the spatial weight matrix W, which defines spatial proximity across the N units in the sample. Assuming that units are partitioned into G clusters, we can reorganize the matrix W according to the partition $\mathcal{P}^G$, such that W consists of G^2 block matrices: G square blocks on the diagonal and $G(G-1)$ rectangular off-diagonal blocks.

Each diagonal block W_{gg} captures the spatial proximity between units within the same cluster $\mathcal{C}_g$, whereas each off-diagonal block W_{gf} (with $g \neq f$) captures spatial relationships between units in different clusters. In the most general case, each submatrix we can associate a different effect, thus estimate different parameters, but specifically we can consider different cases for submatrices on the diagonal and different cases for matrices outside the diagonal, then combine the possible forms of heterogeneity.

Considering the diagonal elements W_{gg}, we can identify three main cases for within-cluster spillovers:

1. **Null case:** $\boldsymbol{\psi}_{gg} = 0$ for all g, indicating no spatial spillovers within any cluster.
2. **Homogeneous case:** $\boldsymbol{\psi}_{gg} = \psi$, constant across all clusters. This assumes that within-cluster spillovers exist but are uniform across groups.
3. **Heterogeneous case:** $\boldsymbol{\psi}_{gg} = \psi_g$, allowing each cluster to have its own level of within-cluster spatial dependence.

These cases are illustrated below, following the matrix structure shown in the diagram:

$$\psi_{gg} = \begin{bmatrix} 0 & \cdot & \cdot \\ \cdot & 0 & \cdot \\ \cdot & \cdot & 0 \end{bmatrix} \quad \text{(Null)}$$

$$\psi_{gg} = \begin{bmatrix} \psi & \cdot & \cdot \\ \cdot & \psi & \cdot \\ \cdot & \cdot & \psi \end{bmatrix} \quad \text{(Homogeneous)}$$

$$\psi_{gg} = \begin{bmatrix} \psi_1 & \cdot & \cdot \\ \cdot & \psi_2 & \cdot \\ \cdot & \cdot & \psi_3 \end{bmatrix} \quad \text{(Heterogeneous)}$$

Regarding the off-diagonal blocks W_{gf} referring to the spatial spillovers across clusters, we identify the following configurations:

1. **Null case:** $\boldsymbol{\psi}_{gf} = 0$ for all $g \neq f$, implying no cross-cluster spillovers.
2. **Homogeneous case:** $\boldsymbol{\psi}_{gf} = \bar{\psi}$, the same for all cross-cluster interactions, regardless of origin and destination group.
3. **Heterogeneous symmetric case:** $\boldsymbol{\psi}_{gf} = \boldsymbol{\psi}_{fg}$, allowing for variation in cross-cluster spillovers but assuming symmetry in the direction of influence.
4. **Heterogeneous asymmetric case:** $\boldsymbol{\psi}_{gf} \neq \boldsymbol{\psi}_{fg}$, where the strength of the spatial influence from cluster g to f differs from the reverse direction.

These configurations are illustrated below using stylized matrices, consistent with the structure provided:

$$\psi_{gf} = \begin{bmatrix} \cdot & 0 & 0 \\ 0 & \cdot & 0 \\ 0 & 0 & \cdot \end{bmatrix} \quad \text{(Null)}$$

$$\psi_{gf} = \begin{bmatrix} \cdot & \psi & \psi \\ \psi & \cdot & \psi \\ \psi & \psi & \cdot \end{bmatrix} \quad \text{(Homogeneous)}$$

$$\psi_{gf} = \begin{bmatrix} \cdot & \psi_{12} & \psi_{13} \\ \psi_{12} & \cdot & \psi_{23} \\ \psi_{13} & \psi_{23} & \cdot \end{bmatrix} \quad \text{(Heterogeneous Symmetric)}$$

$$\psi_{gf} = \begin{bmatrix} \cdot & \psi_{12} & \psi_{13} \\ \psi_{21} & \cdot & \psi_{23} \\ \psi_{31} & \psi_{32} & \cdot \end{bmatrix} \quad \text{(Heterogeneous Asymmetric)}$$

The flexibility of this framework allows for a wide range of combinations. For instance, one may consider a model with heterogeneous within-cluster spillovers (ψ_g) but null cross-cluster effects ($\psi_{gf} = 0$), which captures localized spatial dynamics without inter-cluster influence. Alternatively, one could assume homogeneous within-cluster dependence and heterogeneous asymmetric cross-cluster effects, which would model consistent internal dynamics but directional interactions across groups. At the extreme, the fully null case ($\psi_{gf} = 0$ and $\psi_{gg} = 0$) corresponds to a spatially independent model, where no spatial spillovers exist at all.

This layered approach to modeling heterogeneity enables us to explore how structural groupings (e.g., geographic regions, economic typologies) modulate the strength and direction of spatial interactions across space and time.

4.4 Direct and Indirect Effects

To quantify the impact of covariates in the presence of spatial and temporal dependence with group-level heterogeneity, we derive the matrix of partial derivatives of the dependent variable Y_t with respect to the covariates X_t. Consider the reduced form of the model:

$$\frac{\partial Y_t}{\partial X_t} = \left(I_N - \text{diag}(P^\top \boldsymbol{\tau}) - P^\top \boldsymbol{\rho} PW - P^\top \boldsymbol{\eta} PW\right)^{-1} \left[\text{diag}(P^\top \boldsymbol{\beta}) + P^\top \boldsymbol{\delta} PW\right],$$

where:

- $\boldsymbol{\tau} \in \mathbb{R}^G$: group-specific temporal autoregressive coefficients,
- $\boldsymbol{\rho}, \boldsymbol{\eta}, \boldsymbol{\delta} \in \mathbb{R}^{G \times G}$: group-to-group spatial interaction matrices for the contemporaneous dependent variable, its lag, and covariates respectively,
- $P \in \{0,1\}^{G \times N}$: group membership matrix,
- $W \in \mathbb{R}^{N \times N}$: spatial weights matrix.

The term $(I_N - \text{diag}(P^\top \boldsymbol{\tau}) - P^\top \boldsymbol{\rho} PW - P^\top \boldsymbol{\eta} PW)^{-1}$ acts as a generalized spatial-temporal multiplier, capturing the propagation of effects over both space and time. In contrast to the standard spatial Durbin model, here the feedback strength varies across units depending on their group membership and the spatial structure of inter-group dependencies.

The second term, $[\text{diag}(P^\top \boldsymbol{\beta}) + P^\top \boldsymbol{\delta} PW]$, captures the instantaneous effect of covariates and their spatial lags, adjusted according to the group of the target unit (via $\boldsymbol{\beta}$) and the group of the neighboring units (via $\boldsymbol{\delta}$).

As in J. LeSage and Pace, 2009, the matrix $\partial Y_t / \partial X_t$ can be decomposed into:

- **Direct effects**: the average of the diagonal entries, which measure how a unit's own covariate affects its own outcome, accounting for spatial and temporal feedback.

- **Indirect effects (spillovers)**: the average of the off-diagonal entries, representing the impact of a unit's covariate on the outcomes of other units through spatial interactions.

Importantly, due to the presence of the temporal component $\boldsymbol{\tau}$ and the lagged spatial component $\boldsymbol{\eta}$, the matrix of partial effects describes not only contemporaneous responses but also their propagation through time. This dynamic propagation is embedded in the structure of the multiplier matrix. The group-specific structure of the parameters allows for significant flexibility, enabling different spatial intensities and feedback mechanisms within and across groups. This leads to a rich set of patterns for both direct and indirect effects, which can be summarized by block patterns in the derivatives matrix. The diagonal blocks represent intra-group effects, while the off-diagonal blocks capture inter-group spillovers.

The full matrix of partial derivatives $\partial Y_t / \partial X_t$ has the following extended structure:

$$
\frac{\partial Y}{\partial X} = \left[\frac{\partial Y}{\partial x_1} \frac{\partial Y}{\partial x_2} \cdots \frac{\partial Y}{\partial x_N} \right] = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_N} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_2}{\partial x_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_N}{\partial x_1} & \frac{\partial y_N}{\partial x_2} & \cdots & \frac{\partial y_N}{\partial x_N} \end{bmatrix}
$$

$$
= \begin{bmatrix} I - \tau_1 I - \rho_1 W_1 - \eta_1 W_1 & -\rho_{12} W_{12} - \eta_{12} W_{12} & \cdots & -\rho_{1G} W_{1G} - \eta_{1G} W_{1G} \\ -\rho_{21} W_{21} - \eta_{21} W_{21} & I - \tau_2 I - \rho_2 W_2 - \eta_2 W_2 & \cdots & -\rho_{2G} W_{2G} - \eta_{2G} W_{2G} \\ \vdots & \vdots & \ddots & \vdots \\ -\rho_{G1} W_{G1} - \eta_{G1} W_{G1} & -\rho_{G2} W_{G2} - \eta_{G2} W_{G2} & \cdots & I - \tau_G I - \rho_G W_G - \eta_G W_G \end{bmatrix}^{-1}
$$

$$
\times \begin{bmatrix} \beta_1 I + \delta_{11} W_{11} & \delta_{12} W_{12} & \cdots & \delta_{1G} W_{1G} \\ \delta_{21} W_{21} & \beta_2 I + \delta_{22} W_{22} & \cdots & \delta_{2G} W_{2G} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{G1} W_{G1} & \delta_{G2} W_{G2} & \cdots & \beta_G I + \delta_{GG} W_{GG} \end{bmatrix} \tag{4.4.1}
$$

This matrix is block-partitioned, where each block corresponds to the effect of the covariates from one group on the outcomes of another. The diagonal blocks represent the direct effects (including spatial feedback), while the off-diagonal blocks represent the indirect (spillover) effects across groups. These spillovers are heterogeneous and direction-specific, depending on both origin and destination groups via the corresponding ρ_{gf} and δ_{gf} parameters.

As in J. LeSage and Pace, 2009 and Elhorst, 2010, we summarize the overall direct effect as the average of the diagonal elements of the full matrix $\partial Y_t / \partial X_t$, while the overall indirect effect is the average of the off-diagonal elements. However, given the group-specific nature of our model, these effects can be decomposed more finely to reflect intra- and inter-group heterogeneity:

- **Direct effects within clusters:** For each group g, the direct effect is computed as the average of the diagonal elements of the corresponding diagonal block in the derivative matrix (i.e., the g-th block on the main diagonal). This captures the impact of a region's own covariates on its own outcomes within the same cluster, including feedback through space and time.
- **Indirect effects within clusters:** Still within the diagonal blocks, the off-diagonal elements represent the intra-group spillovers—i.e., how the covariates of one region within a cluster influence the outcomes of other regions in the same cluster. Their average provides a measure of within-cluster indirect effects.
- **Indirect effects across clusters:** The off-diagonal blocks in the matrix correspond to spatial spillovers between different clusters. Averaging all elements in these off-diagonal blocks yields the cross-cluster indirect effects, capturing how a region's covariates influence the outcomes of regions in other clusters through the spatial network.

This decomposition enables a nuanced interpretation of the effects and reveals how the influence of covariates propagates not only locally but also across group boundaries, depending on both spatial proximity and structural similarity captured by the clustering.

4.5 Estimation procedure

In this section, we present the full estimation procedure for the proposed heterogeneous spatiotemporal model. In the first part, we describe a two-step procedure that jointly estimates all model parameters and identifies the group membership of each unit. This approach requires no prior assumptions about the cluster structure or knowledge of the underlying parameter values, allowing for a fully data-driven estimation of both heterogeneity and spatial dynamics.

In the second part, we focus on the estimation of the spatiotemporal dependence parameters and covariate effects, assuming the group partition is known or given. Here, we propose a quasi maximum likelihood (QML) estimation method, which extends existing approaches commonly used in spatial dynamic panel models. Unlike standard methods, which typically assume parameter

homogeneity, our procedure explicitly incorporates cluster-based heterogeneity in both spatial and temporal components.

4.5.1 Two step procedure

The parameters to be estimated in our model include both the regression coefficients and the group membership configuration. Formally, we denote the full set of regression parameters as

$$\boldsymbol{\Theta} = (\boldsymbol{\rho}, \boldsymbol{\tau}, \boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\delta}, \sigma),$$

and the group structure as $\mathcal{P} = \{\mathcal{C}_1, \ldots, \mathcal{C}_G\}$, where each $\mathcal{C}_g$ denotes a cluster and every unit belongs to one and only one cluster. Due to the complexity arising from the unknown group membership and the spatial dependence, we adopt a two-step estimation approach based on maximum likelihood, following the strategy proposed in Sugasawa and Murakami, 2021.

1. **Step 1: Estimation of model parameters given a fixed partition**

   In the first step, we assume that the group membership $\mathcal{P}$ is known and fixed. The initial partition may come from cluster analysis techniques on a set of variables or from the economic literature. Given this configuration, we estimate the structural parameters $\boldsymbol{\Theta}$ by maximizing the log-likelihood function. Specifically, the estimator is defined as:

   $$\widehat{\boldsymbol{\Theta}} = \arg\max_{\boldsymbol{\Theta}} \mathcal{L}(Y|X, W, \boldsymbol{\Theta}, \mathcal{P}^G)$$

   where $\mathcal{L}(\cdot)$ denotes the conditional density function of the sample i, conditional on covariates, spatial weights, and cluster assignment. Since the likelihood function involves spatial dependence across observations, we cannot assume independence between units, and standard MLE must be adapted. Therefore, we follow the QML framework, similar to L.-F. Lee and Yu, 2010a, 2010b; Yu et al., 2008 and as implemented by Simonovska, 2025. In the following subsection we provide an extension of their approach to the case with multiple spatial parameters.

2. **Step 2: Estimation of group membership given model parameters**

   In the second step, we consider the reverse problem: assume the model parameters $\boldsymbol{\Theta}$ are known (e.g., from Step 1), and update the group membership for each unit. The assignment of unit i to a group g_i is determined by maximizing its local log-likelihood, taking into account not only the unit's own outcome but also the outcomes of its neighbors. The estimation rule is defined as:

$$\widehat{g}_i = \arg \max_{g_i \in \{1,\dots,G\}} \left\{ \log f_i(y_i \,|\, \mathbf{x}_{i\cdot}, W, \boldsymbol{\Theta}, g_i) + \sum_{j=1}^{N} \log f_j(\cdot)\, \mathbb{I}(w_{ij} > 0) \right\}.$$

   This step reflects the spatially dependent nature of the data: units that are geographically or economically close may influence one another, and therefore their group assignments should not be estimated in isolation. The second term in the maximand explicitly incorporates the log-likelihood contributions of neighboring units (i.e., those for which $w_{ij} > 0$), thereby acknowledging the violation of conditional independence across spatial units. Intuitively, this formulation introduces local spatial smoothing in the group assignment process, reducing the likelihood of assigning a unit to a cluster that is inconsistent with the behavior of its neighbors. Compared to previous two-step methodologies in cluster-wise regression, which typically consider each unit's likelihood independently—thus neglecting cross-sectional dependence—our approach explicitly accounts for the influence of neighboring units. In particular, the likelihood contribution of unit i depends not only on its own characteristics and group assignment, but also on the likelihoods and cluster memberships of spatially connected units. Depending on the model size and the number of clusters, this step can be implemented through exhaustive enumeration (in small settings), or via more scalable methods such as a classification expectation-maximization (CEM) algorithm or a greedy assignment update, alternating with Step 1 until convergence.

This two-step procedure can be iterated until the estimated group membership and parameter estimates converge. Convergence is typically assessed by tracking changes in the log-likelihood or in the clustering assignments across iterations. This approach enables flexible modeling of group-level heterogeneity while fully accounting for spatial dependence and feedback effects in the data.

This two-step procedure can be iterated until the estimated group membership and parameter estimates converge. Convergence is typically assessed by monitoring changes in the log-likelihood or in the clustering assignments across successive iterations. This iterative approach enables flexible modeling of group-level heterogeneity while fully accounting for spatial dependence and feedback effects present in the data.

Formally, let λ denote the iteration index of the two-step algorithm. The algorithm stops when one of the following conditions is satisfied: the parameter estimates stabilize, i.e., $\boldsymbol{\Theta}^{(\lambda)} = \boldsymbol{\Theta}^{(\lambda-1)}$; or the improvement in the log-likelihood function falls below a predefined threshold, i.e., $\left|\mathcal{L}^{(\lambda)} - \mathcal{L}^{(\lambda-1)}\right| \leq \gamma$, where $\gamma > 0$; or the maximum number of iterations $\lambda_{\max}$ is reached. These stopping rules ensure convergence of the procedure while avoiding unnecessary computational burden.

4.5.2 Quasi Maximum Likelihood Estimation

To estimate the parameters of our dynamic spatial panel model with group-specific heterogeneity, we adopt a Quasi-Maximum Likelihood approach. This methodology builds on the framework proposed by L.-F. Lee and Yu, 2010a, 2010b; Yu et al., 2008 for spatial dynamic panel model, when cluster heterogeneity is not considered, and implemented by Simonovska, 2025. The model is split into two conceptual blocks: one capturing the standard (non-spatial) dependence of the dependent variable on covariates, and another focusing on the spatially lagged dependent variable and the effects of covariates on it.

In their approach, spatial dependence is treated as a correction to the primary regression, allowing for sequential estimation: first by estimating the effect of covariates without spatial feedback, and then adjusting for spatial spillovers using a residual correction term.

Our method extends this idea to the more complex setting of heterogeneous spatial dependence. Specifically, instead of a single spatial component, we consider multiple group-to-group spatial dependencies, with one spatial autoregressive coefficient for each pair of groups. We decompose the model into a non-spatial block and several spatial blocks, each corresponding to a specific group interaction. This block-wise structure allows us to first estimate the base regression, then separately estimates the effects of regressors on the spatial lag of the dependent variable for each block, and

finally recover the spatial autoregressive coefficients by maximizing a concentrated log-likelihood function that accounts for the entire spatial dependence structure.

This modular procedure not only simplifies estimation in the presence of heterogeneity but also improves computational tractability and interpretability of the estimated parameters.

Let $Y_t = (y_{1t}, \dots, y_{Nt})'$ denote the $N \times 1$ vector of dependent variables at time t, and $X_t = (\mathbf{x}'_{1t}, \dots, \mathbf{x}'_{Nt})'$ the $N \times k$ matrix of covariates. Let $W \in \mathbb{R}^{N \times N}$ be the spatial weights matrix and $P \in \{0,1\}^{G \times N}$ the group membership matrix, where each column of P contains a single entry equal to 1, indicating the group to which each unit belongs.

Regarding the spatial weights matrix W, in this paper we assume a single spatial structure that governs all spatial and spatiotemporal spillover effects. However, the framework is flexible and can accommodate multiple spatial weight matrices, such as W_x and W_y, to separately capture spatial spillovers in the independent and dependent variables, or matrices like W_0 and W_1, which distinguish between contemporaneous and temporally lagged spillovers.

We aim to estimate the set of regression parameters

$$\boldsymbol{\Theta} = (\boldsymbol{\rho}, \boldsymbol{\tau}, \boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\delta}, \sigma)$$

From the model described in Equation 4.2.5, which can be expressed as:

$$\begin{aligned}\left(I_N - P^\top \boldsymbol{\rho} P \circ W\right) Y_t = \operatorname{diag}(P^\top \boldsymbol{\tau}) Y_{t-1} + \left(P^\top \boldsymbol{\eta} P \circ W\right) Y_{t-1} \\ + \operatorname{diag}_k(P^\top \boldsymbol{\beta}) X_t + \left(P^\top \boldsymbol{\delta} P \circ W\right) X_t + \epsilon_t,\end{aligned} \tag{4.5.1}$$

where the Hadamard product "$\circ$" captures group-specific spatial interactions. We recall that we are not considering any fixed effects and that the error variance is constant over the sample. To simplify notation, we collect all regressors and lagged terms into a single matrix Z_t and rewrite the model as:

$$\left(I_N - P^\top \boldsymbol{\rho} P \circ W\right) Y_t = Z_t \mathbf{B} + \epsilon_t, \tag{4.5.2}$$

where Z_t is the $N \times q$ matrix of explanatory variables which includes the temporal lag of the dependent variable, the lagged spatial spillover of the dependent variable, the covariates and their

spillovers, all interacted by the clustering membership. Thus defined as

$$Z_t = \big[\, Y_{t-1},\ (P^\top \boldsymbol{\eta} P \circ W) Y_{t-1},\ X_t,\ (P^\top \boldsymbol{\delta} P \circ W) X_t \,\big],$$

and $\mathbf{B}$ is the stacked parameter vector, with lenght q, including all the regression coefficient of the model in Equation 4.2.5, except $\boldsymbol{\rho}$, thus defined as

$$\mathbf{B}^\top = \big[\, \boldsymbol{\tau}^\top,\ \mathrm{vec}(\boldsymbol{\eta})^\top,\ \boldsymbol{\beta}^\top,\ \mathrm{vec}(\boldsymbol{\delta})^\top \,\big].$$

Focusing on the left-hand side (LHS), we can equivalently express the spatial dependence as:

$$\left(I_N - \sum_{g=1}^{G} \sum_{f=1}^{G} \rho_{gf} W^{(gf)} \right) Y_t = Z_t \mathbf{B} + \epsilon_t, \tag{4.5.3}$$

where $W^{(gf)} \in \mathbb{R}^{N \times N}$ denotes a group-structured spatial matrix, defined elementwise as

$$w_{ij}^{(gf)} = \begin{cases} w_{ij}, & \text{if } i \in \mathcal{C}_g \text{ and } j \in \mathcal{C}_f, \\ 0, & \text{otherwise.} \end{cases}$$

Here, $\mathcal{C}_g$ denotes the set of units belonging to group g. Each matrix $W^{(gf)}$ captures spatial connections only between units in groups g and f, while maintaining the full $N \times N$ dimension.

Remark. Notice that $W^{(gf)}$ do not correspond to the submatrix W_{gf}. The latter would represent the block of W corresponding to rows in $\mathcal{C}_g$ and columns in $\mathcal{C}_f$, hence its dimensions depend on the size of clusters g and f. In contrast, $W^{(gf)}$ preserves the full $N \times N$ structure, with nonzero entries only where both group indices match, which ensures algebraic compatibility in Hadamard products and summations.

We now partition the model into its non-spatial and spatial components:

$$Y_t = Z_t \mathbf{B}_{ns} + \epsilon_{t,ns}, \tag{4.5.4}$$

$$-\rho_{gf} W^{(gf)} Y_t = Z_t \mathbf{B}_{sp}^{(gf)} + \epsilon_{t,sp}^{(gf)} \qquad \forall g, f = 1, \ldots, G \tag{4.5.5}$$

where $\mathbf{B}_{ns}$ collects the non-spatial parameters and $\mathbf{B}_{sp}^{(gf)}$ captures the contribution of the (g, f)-specific spatial effect, thus

$$\mathbf{B}_{ns}^{\top} = \left[\boldsymbol{\tau}_{ns}^{\top},\ \mathrm{vec}(\boldsymbol{\eta}_{ns})^{\top},\ \boldsymbol{\beta}_{ns}^{\top},\ \mathrm{vec}(\boldsymbol{\delta}_{ns})^{\top} \right],$$

$$\mathbf{B}_{sp}^{(gf)\top} = \left[\boldsymbol{\tau}_{sp}^{(gf)\top},\ \mathrm{vec}(\boldsymbol{\eta})_{sp}^{(gf)\top},\ \boldsymbol{\beta}_{sp}^{(gf)\top},\ \mathrm{vec}(\boldsymbol{\delta})_{sp}^{(gf)\top} \right], \quad \forall g, f = 1, \ldots, G.$$

Remark. Notice that, the sum of the RHS of Equations 4.5.4 and 4.5.5, correspond to the RHS of Equation 4.5.3, thus sum of the LHS of Equations 4.5.4 and 4.5.5, correspond to the LHS of Equation 4.5.3. By construction,

$$\mathbf{B} = \mathbf{B}_{ns} + \sum_{g=1}^{G} \sum_{f=1}^{G} \mathbf{B}_{sp}^{(gf)}, \qquad \epsilon_t = \epsilon_{t,ns} + \sum_{g=1}^{G} \sum_{f=1}^{G} \epsilon_{t,sp}^{(gf)}, \quad \forall g, f = 1, \ldots, G.$$

To facilitate estimation of the spatial blocks, we define:

$$W^{(gf)} Y_t = Z_t \mathbf{B}_{sp}^{*(gf)} + \epsilon_{t,sp}^{*(gf)}, \quad \forall g, f = 1, \ldots, G$$

such that

$$\mathbf{B}_{sp}^{(gf)} = -\rho_{gf} \mathbf{B}_{sp}^{*(gf)}, \qquad \epsilon_{t,sp}^{(gf)} = -\rho_{gf} \epsilon_{t,sp}^{*(gf)}, \quad \forall g, f = 1, \ldots, G.$$

The vectors of coefficients in the regression blocks $\mathbf{B}_{ns}$ and $\mathbf{B}_{sp}^{*(gf)}$ can then be estimated using least squares as:

$$\hat{\mathbf{B}}_{ns} = (Z'Z)^{-1} Z'Y,$$

$$\hat{\mathbf{B}}_{sp}^{*(gf)} = (Z'Z)^{-1} Z' W_T^{(gf)} Y, \quad \forall g, f = 1, \ldots, G$$

where Y is the $NT \times 1$ stacked vector of outcomes, Z is the $NT \times q$ regressor matrix, and $W_T = I_T \otimes W$, $W_T^{(gf)} = I_T \otimes W^{(gf)}$.

The corresponding residuals are:

$$\hat{\epsilon}_{ns} = Y - Z\hat{\mathbf{B}}_{ns}, \qquad \hat{\epsilon}_{sp}^{*(gf)} = W_T^{(gf)} Y - Z\hat{\mathbf{B}}_{sp}^{*(gf)}.$$

We then maximize the quasi log-likelihood function from L.-F. Lee and Yu, 2010a; Yu et al., 2008 with respect to $\boldsymbol{\rho}$, given by:

$$\mathcal{L}(Y|Z,W,\boldsymbol{\rho},\mathcal{P}^G) \propto -\frac{NT_0}{2}\log\left(\frac{1}{NT_0}\sum_{t=1}^{T_0}\epsilon_t'\epsilon_t\right) + T_0\log\left|I_N - \sum_{g=1}^{G}\sum_{f=1}^{G}\rho_{gf}W^{(gf)}\right|, \tag{4.5.6}$$

where $T_0 = T$ if the model is not dynamic, and $T_0 = T-1$ if the model is dynamic. The total residual ϵ_t can be found as the sum of the residual from the non-spatial regression block and the residuals from the spatial regression blocks, thus

$$\epsilon_t = \hat{\epsilon}_{t,ns} + \sum_{g=1}^{G}\sum_{f=1}^{G}\epsilon_{t,sp}^{(gf)} = \hat{\epsilon}_{t,ns} - \sum_{g=1}^{G}\sum_{f=1}^{G}\rho_{gf}\hat{\epsilon}_{t,sp}^{*(gf)}$$

and thus, the squared of the total residuals can be obtained as the square of the binomial as follows

$$\epsilon_t'\epsilon_t = \left(\hat{\epsilon}_{t,ns} - \sum_{g=1}^{G}\sum_{f=1}^{G}\rho_{gf}\hat{\epsilon}_{t,sp}^{*(gf)}\right)'\left(\hat{\epsilon}_{t,ns} - \sum_{g=1}^{G}\sum_{f=1}^{G}\rho_{gf}\hat{\epsilon}_{t,sp}^{*(gf)}\right)$$

$$= \hat{\epsilon}_{t,ns}'\hat{\epsilon}_{t,ns} - 2\hat{\epsilon}_{t,ns}'\sum_{g=1}^{G}\sum_{f=1}^{G}\rho_{gf}\hat{\epsilon}_{t,sp}^{(gf)} + \left(\sum_{g=1}^{G}\sum_{f=1}^{G}\rho_{gf}\hat{\epsilon}_{t,sp}^{(gf)}\right)'\left(\sum_{g=1}^{G}\sum_{f=1}^{G}\rho_{gf}\hat{\epsilon}_{t,sp}^{(gf)}\right). \tag{4.5.7}$$

Plugging the residuals in 4.5.7 in the likelihood function in 4.5.6, it is possible to estimated spatial parameters $\boldsymbol{\rho}$ as

$$\hat{\boldsymbol{\rho}} = \arg\max_{\boldsymbol{\rho}} \mathcal{L}(Y|Z,W,\boldsymbol{\rho},\mathcal{P}^G).$$

Given $\hat{\boldsymbol{\rho}}$, we update the coefficients and residuals as:

$$\hat{\mathbf{B}} = \hat{\mathbf{B}}_{ns} - \sum_{g=1}^{G}\sum_{f=1}^{G}\hat{\rho}_{gf}\,\hat{\mathbf{B}}_{sp}^{*(gf)},$$

$$\hat{\epsilon}_t = \hat{\epsilon}_{t,ns} - \sum_{g=1}^{G}\sum_{f=1}^{G}\hat{\rho}_{gf}\,\hat{\epsilon}_{t,sp}^{*(gf)}.$$

Finally, the fitted values for the entire panel are given by:

$$\widehat{Y} = \left(I_{NT} - \sum_{g=1}^{G} \sum_{f=1}^{G} \hat{\rho}_{gf} W_T^{(gf)} \right)^{-1} Z\widehat{\mathbf{B}},$$

and the error variance is computed as:

$$\widehat{\sigma}_{\varepsilon}^2 = \frac{\widehat{\epsilon}'\widehat{\epsilon}}{NT}.$$

4.6 Simulation Experiment

To evaluate the proposed estimation procedure, we design a Monte Carlo experiment based on the group-structured dynamic spatial panel data model introduced in Section 4.2. The general specification in Equation (4.2.2) includes several layers of spatial and temporal dependence. However, as discussed in Anselin et al., 2008 and Elhorst, 2010, the full model suffers from identification issues when all dependence parameters are simultaneously included. For this reason, applied research commonly relies on restricted versions that incorporate specific combinations of parameters.

Following this approach, we design a series of data-generating processes (DGPs) that consider different subsets of the full parameter space, enabling us to separately assess the performance of the estimator under various spatial and dynamic configurations. In particular, we consider five restricted models (M1–M5), each including at least one of the structural parameters of interest—spatial, temporal, or covariate effects—but never all simultaneously.

4.6.1 Spatial Structure and Heterogeneity Settings

The spatial matrix W is constructed using a k-nearest neighbors approach with $k = 10$, based on random coordinates uniformly distributed in $[0, 1]^2$. This matrix is row-standardized and has zeros on the diagonal.

We introduce heterogeneity in the spatial dependence through different configurations of the group-to-group dependence matrix $\boldsymbol{\psi} = (\boldsymbol{\rho}, \boldsymbol{\eta}, \boldsymbol{\delta})$, representing three distinct levels of cross-cluster

interaction (see Section 4.3):

- H1 - No Cross-Group Spillovers: heterogeneous within clusters but no spatial link across clusters.

$$\boldsymbol{\psi} = [\psi_{gf}] = \begin{bmatrix} \psi_1 & 0 \\ 0 & \psi_2 \end{bmatrix} \quad \text{H1}$$

- H2 - Homogeneous Cross-Group Spillovers: heterogeneous within clusters and symmetric across clusters (i.e., $\rho_{12} = \rho_{21}$).

$$\boldsymbol{\psi} = [\psi_{gf}] = \begin{bmatrix} \psi_1 & \psi_{12} \\ \psi_{12} & \psi_2 \end{bmatrix} \quad \text{H2}$$

- H3 - Asymmetric Cross-Group Spillovers: heterogeneous within clusters and asymmetric across clusters ($\rho_{12} \neq \rho_{21}$).

$$\boldsymbol{\psi} = [\psi_{gf}] = \begin{bmatrix} \psi_1 & \psi_{12} \\ \psi_{21} & \psi_2 \end{bmatrix} \quad \text{H3}$$

In all settings, within-cluster dependencies (ρ_{gg}) differ across groups. The true parameter values used in the simulations are:

$$\boldsymbol{\rho} = \begin{bmatrix} \rho_1 & \rho_{12} \\ \rho_{21} & \rho_2 \end{bmatrix} \begin{cases} \begin{bmatrix} 0.25 & 0 \\ 0 & 0.15 \end{bmatrix} \\ \begin{bmatrix} 0.25 & 0.05 \\ 0.05 & 0.15 \end{bmatrix} \\ \begin{bmatrix} 0.25 & 0.05 \\ 0.02 & 0.15 \end{bmatrix}. \end{cases}$$

Other parameters are set as follows:

$$\boldsymbol{\tau} = (\tau_1, \tau_2) = (0.3, 0.2), \quad \boldsymbol{\beta} = (\beta_1, \beta_2) = (0.4, 0.6), \quad \boldsymbol{\eta} = 0.4\boldsymbol{\rho}, \quad \boldsymbol{\delta} = 0.6\boldsymbol{\rho}.$$

4.6.2 Model Variants

We simulate data under five model configurations, each corresponding to a different restriction of the general framework:

- **M1**: Spatial lag ρ and covariate β

$$Y_t = \left(P^\top \boldsymbol{\rho} P \circ W\right) Y_t + \mathrm{diag}_k(P^\top \boldsymbol{\beta}) X_t + \epsilon_t,$$

- **M2**: Spatial lag ρ and temporal lag τ

$$Y_t = \left(P^\top \boldsymbol{\rho} P \circ W\right) Y_t + \mathrm{diag}(P^\top \boldsymbol{\tau}) Y_{t-1} + \epsilon_t,$$

- **M3**: Spatial Durbin Model with ρ, β, and spatial lag of covariates δ

$$Y_t = \left(P^\top \boldsymbol{\rho} P \circ W\right) Y_t + \mathrm{diag}_k(P^\top \boldsymbol{\beta}) X_t + \left(P^\top \boldsymbol{\delta} P \circ W\right) X_t + \epsilon_t,$$

- **M4**: Spatiotemporal autoregressive model with ρ, τ, and lagged spatial outcome η

$$Y_t = \left(P^\top \boldsymbol{\rho} P \circ W\right) Y_t + \mathrm{diag}(P^\top \boldsymbol{\tau}) Y_{t-1} + \left(P^\top \boldsymbol{\eta} P \circ W\right) Y_{t-1} + \epsilon_t,$$

- **M5**: Spatiotemporal autoregressive model with covariate, with ρ, τ, η, β

$$Y_t = \left(P^\top \boldsymbol{\rho} P \circ W\right) Y_t + \mathrm{diag}(P^\top \boldsymbol{\tau}) Y_{t-1} + \left(P^\top \boldsymbol{\eta} P \circ W\right) Y_{t-1} + \mathrm{diag}_k(P^\top \boldsymbol{\beta}) X_t + \epsilon_t,$$

Each specification is therefore a restricted but identifiable form of the general model, allowing us to evaluate how estimation accuracy changes with model complexity.

4.6.3 Monte Carlo Implementation

Each configuration is repeated over $MC = 100$ Monte Carlo iterations with $N = 200$ spatial units, $T = 50$ time periods, and $G = 2$ clusters. The relative size of the two clusters varies across simulations: for each iteration, a probability 0.5 is drawn from a uniform distribution $U(0.2, 0.8)$, and each unit i is assigned to cluster 1 with probability p_g and to cluster 2 with probability $1 - p_g$. This ensures that cluster sizes are random but balanced on average.

Simulation outputs include estimated spatial parameters (ρ_{gf}), temporal and dynamic lags (τ_g, η_{gf}), covariate effects (β_g, δ_{gf}), residual variance σ^2_ϵ. Moreover, for models that include the covariate variable X we are able to estimates direct and indirect effects over the entire sample, as well as cluster-specific direct effects and cross-clusters indirect effects.

Algorithm 5 Monte Carlo Simulation Procedure

Set $N = 100$, $T = 50$, $G = 2$, $k = 1$, $MC = 100$
Simulate a clustering partition, $\mathcal{P}^{G=2}$
Construct W via k-nearest neighbors ($k = 10$)
Set parameters $\boldsymbol{\Theta}$
Compute true values of direct and indirect effects within clusters and across clusters
for $mc = 1$ to MC **do**
    for each heterogeneity configuration (H1–H3) **do**
        Condiser the proper subset of parameters according to the heterogeneity configuration
        for each model specification (M1–M5) **do**
            Condiser the proper subset of parameters according to the model specification
            Simulate panel data (Y_t, X_t) using the DGP, such that $Y_t = f(X_t, \Theta, \mathcal{P}^2, W)$
            Estimate model parameters $\widehat{\boldsymbol{\Theta}}$ using QML procedure
            Compute estimates of direct and indirect effects within clusters and across clusters
            Store estimated coefficients, direct and indirect effects
        end for
    end for
end for

This design allows us to systematically evaluate how different sources of spatial heterogeneity and model restrictions affect estimation performance and parameter recovery.

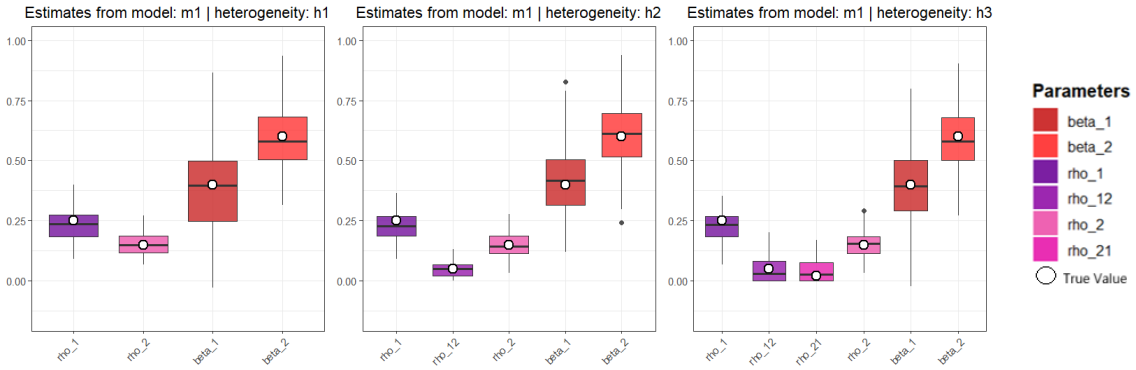
4.6.4 Discussion of Simulation Results

Overall, the simulation results confirm that the proposed estimation procedure performs well in recovering the heterogeneous spatial and dynamic parameters across different model specifications and heterogeneity settings. Figure 4.6.0 show the distribution of the estimates of spatial and temporal dependence parameters considering models M1-M5 respectively in panels a-e, and considering different heterogeneity structures of the parameters H1-H3 as described in previously. In each boxplot, the white circle indicates the true value of the parameters.

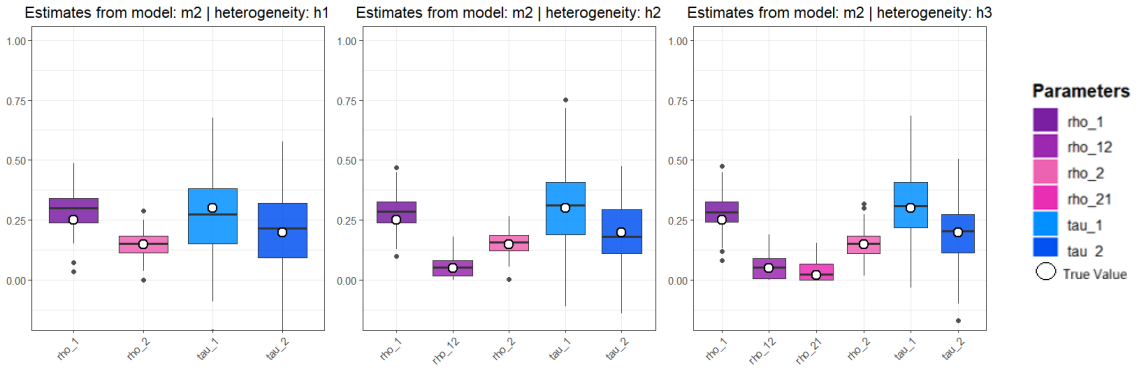
The estimates of the spatial autoregressive parameters $\boldsymbol{\rho}$ –represented in shadow of purple– are generally accurate and display low variability, closely reflecting the underlying cluster heterogeneity imposed in the data-generating process. This indicates that the estimation method is able to effectively capture group-specific spatial dependencies.

For the simpler model specifications (M1 and M2), the estimators exhibit good performance, with the majority of estimated parameters centered around their true values. Although the boxplots of $\boldsymbol{\beta}$ and $\boldsymbol{\tau}$ display higher variability (i.e., longer boxes), their medians correctly reflect the cluster-specific heterogeneity: the estimated τ_g parameters are higher for cluster 1 and lower for cluster 2, while the opposite pattern is observed for the β_g parameters, which are higher for cluster 2 and lower for cluster 1. This consistency across replications suggests that the method reliably distinguishes between cluster-specific effects even in the presence of moderate dispersion.

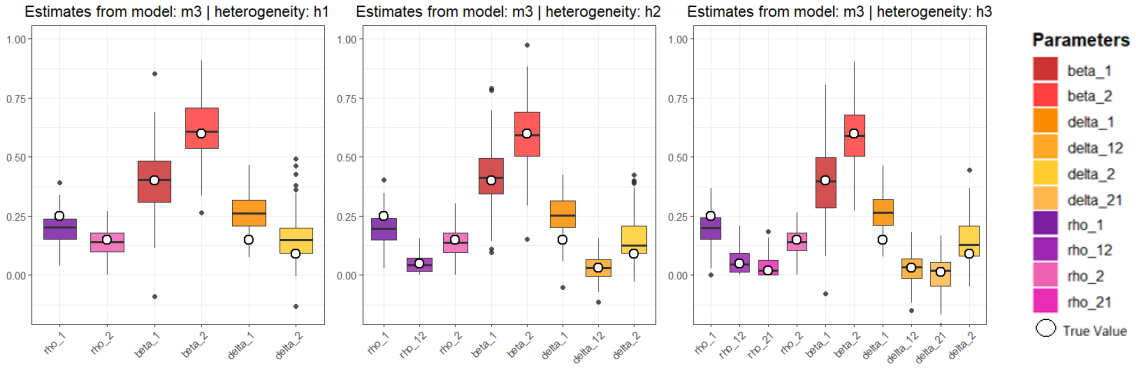
In contrast, for the more complex models (M3–M5), which include multiple spatial effects ($\boldsymbol{\rho}, \boldsymbol{\delta}$ and $\boldsymbol{\eta}$), the estimation procedure shows some limitations. In particular, the spatiotemporal coefficients δ_{gf} and the lagged spatial effects η_{gf} tend to be slightly overestimated, while the autoregressive spatial parameters ρ_{gf} are somewhat underestimated. This bias is likely due to the increased complexity of jointly identifying several interacting spatial components. Nonetheless, the model remains capable of capturing the relative heterogeneity structure across clusters: the estimated parameters consistently follow the ordering $\psi_1 > \psi_2 > \psi_{12} \geq \psi_{21}$, as in the settings, meaning that within-cluster effects in group 1 are generally stronger than those in group 2, while cross-cluster effects are smaller and slightly asymmetric, consistently with the true values of the parameters.



A. Simulation results – Model 1

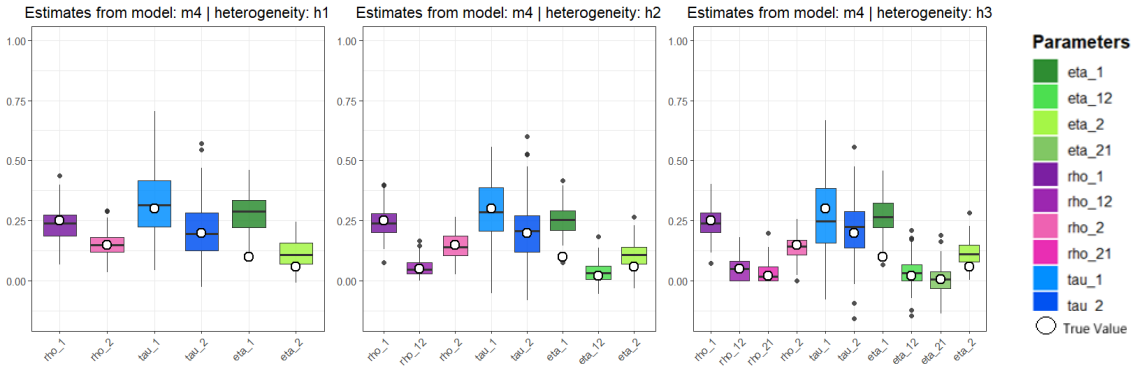


B. Simulation results – Model 2

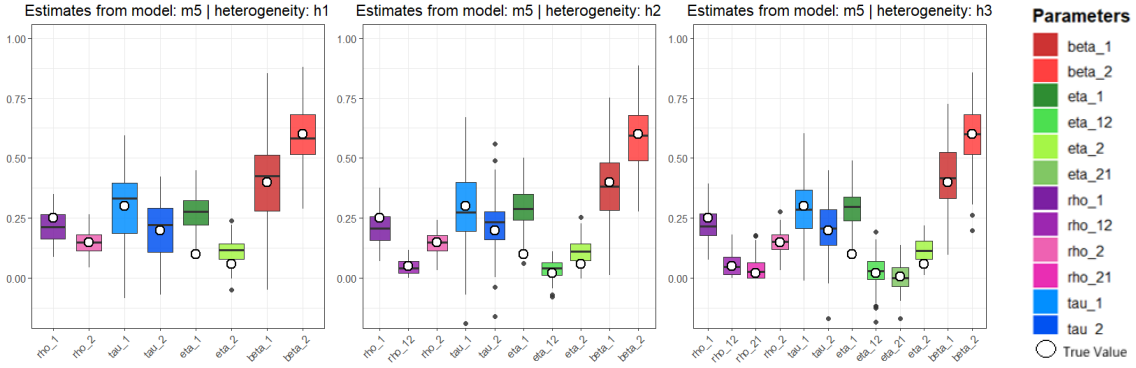


C. Simulation results – Model 3

Figure 1: Monte Carlo simulation results (a–c) for models M1–M3 under various heterogeneity settings. The boxplots show the distribution of the estimates of spatial and temporal dependence parameters, while the white circle indicates the true value of the parameters. Panels continue on next page.



D. Simulation results – Model 4



E. Simulation results – Model 5

Figure 4.6.0: (continued) Monte Carlo simulation results (d–e) for models M4–M5 under various heterogeneity settings. The boxplots show the distribution of the estimates of spatial and temporal dependence parameters, while the white circle indicates the true value of the parameters.

Now, we consider the resulting direct and indirect effects of the independent variable in models M1, M3 and M5. Direct and Indirect effects have been computed as described in Section 4.4, considering cluster-specifics effects and cross-clusters effects.

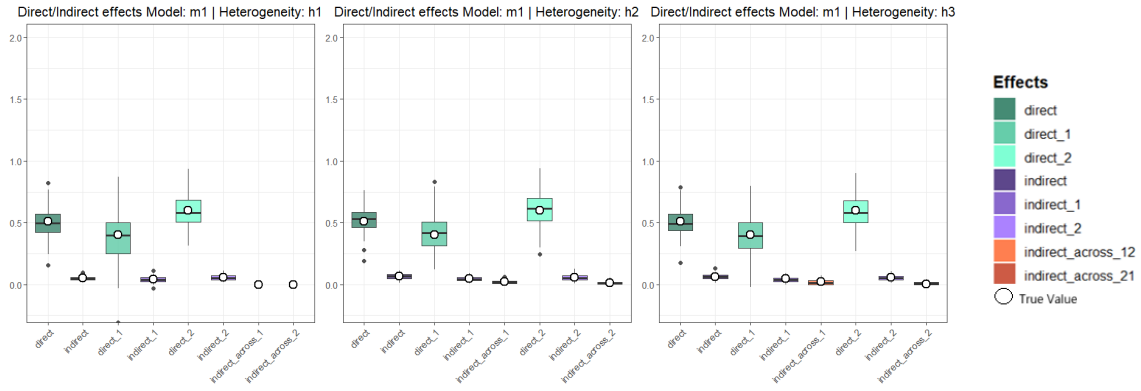
Figure 4.6.1 illustrates the distribution of the estimated effects of X on Y across Monte Carlo replications, broken down by model and heterogeneity scenario. Each boxplot corresponds to a different spatial effect component: total, cluster-specific, and cross-cluster direct and indirect effects. The horizontal axis labels denote the specific type of effect, where "direct" and "indirect" refer to the global effects, "dir_1" and "indir_1" refer to the effects in clusters 1, "ind_across_1" refers to the resulting effect on Y in units in cluster 1, from a change of X in a units in cluster 2, similarly "dir_2" and "indir_2" refer to the effects in clusters 2, "ind_across_2" refers to the resulting effect on Y in units in cluster 2, from a change of X in a units in cluster 1. The true value of each effect is overlaid as a white circle, allowing visual assessment of estimation accuracy and bias.

Overall, the estimation procedure demonstrates strong performance. In most cases, the boxplots are centered closely around the true effect values, indicating unbiasedness. For example, in Model M1, the white circles often lie near the center of the interquartile range, especially for the direct and indirect effects of clusters 1 and 2, suggesting accurate estimation and moderate variability.

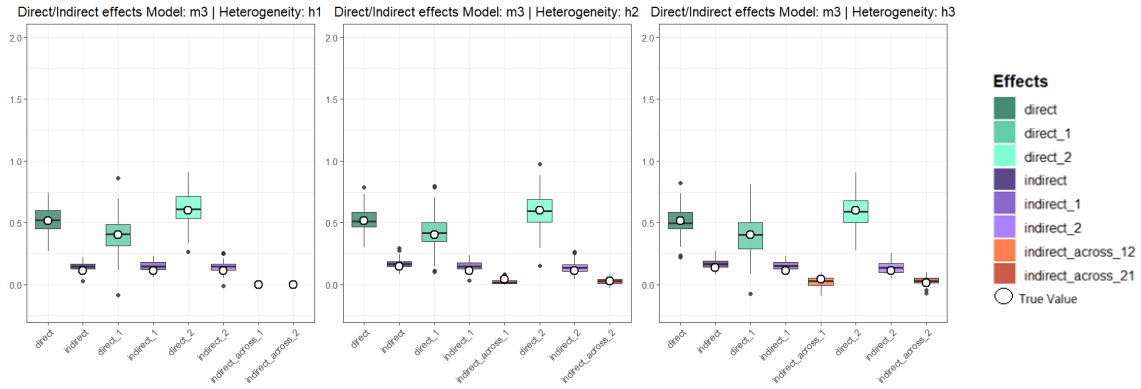
Notably, the model captures the cluster-specific heterogeneity of the effects well. Across all models and heterogeneity settings, we observe that the estimated direct and indirect effects for cluster 1 tend to be higher than those for cluster 2, consistently reflecting the true pattern. This relationship holds not only for the within-cluster effects but also for the cross-cluster spillovers, and is reflected in the estimated values.

In Models M3 and M5, the variability of some estimates increases, especially for indirect effects, as reflected in taller boxplots. Moreover, in model M5 the indirect effects in both clusters seem to be slightly overestimated. Nonetheless, even in these more complex settings, the estimates tend to preserve the correct ranking across clusters, with larger effects being associated with higher estimates. This is particularly important, as it confirms the model's ability to distinguish between different spatial dynamics across groups, even when absolute precision slightly decreases.

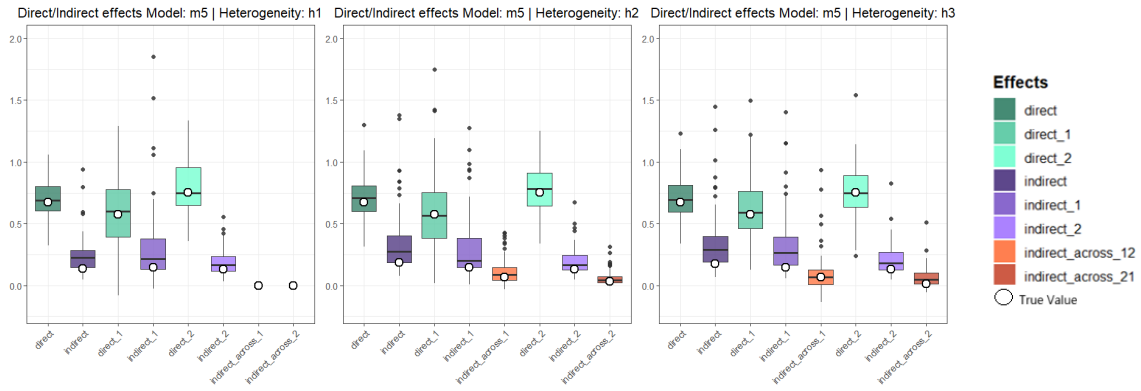
Table 4.6.1-4.6.5 reports simulation results of estimates of model M1-M5 with different hetero-



A. Estimated effects – Model 1



B. Estimated effects – Model 3



C. Estimated effects – Model 5

Figure 4.6.1: Boxplots of estimated direct and indirect effects for models M1, M3, and M5 under different heterogeneity structures. The boxplots show the distribution of the estimates of general direct and indirect effects, cluster-specific and cross-cluster direct and indirect effects. The white circle indicates the true value of the effects.

geneity structures H1-H3. For each parameter to be estimated in each model, we report its true value in the DGPs, the Mean of the estimates, the Mean Absolute Error (MAE) and the Mean Squared Error (MSE).

Simulation results for Model M1

		Heter. H1			**Heter. H2**			**Heter. H3**		
	True	**Mean**	**MAE**	**MSE**	**Mean**	**MAE**	**MSE**	**Mean**	**MAE**	**MSE**
ρ_1	0.25	0.231	0.055	0.004	0.233	0.050	0.004	0.224	0.051	0.004
ρ_{12}	0.05	–	–	–	0.048	0.027	0.001	0.048	0.041	0.002
ρ_{21}	0.02	–	–	–	–	–	–	0.041	0.036	0.002
ρ_2	0.15	0.151	0.038	0.002	0.149	0.041	0.002	0.151	0.040	0.002
β_1	0.40	0.374	0.137	0.032	0.407	0.113	0.020	0.395	0.122	0.025
β_2	0.60	0.593	0.102	0.015	0.609	0.110	0.018	0.588	0.111	0.018

Table 4.6.1: Simulation results of estimates of model M1 with different heterogeneity structures H1-H3. For each parameter to be estimated in the model, we report its true value in the DGPs, the Mean of the estimates, the Mean Absolute Error (MAE) and the Mean Squared Error (MSE).

Simulation results for Model M2

		Heter. H1			Heter. H2			Heter. H3		
	True	**Mean**	**MAE**	**MSE**	**Mean**	**MAE**	**MSE**	**Mean**	**MAE**	**MSE**
ρ_1	0.25	0.291	0.070	0.008	0.285	0.065	0.006	0.280	0.067	0.007
ρ_{12}	0.05	–	–	–	0.056	0.037	0.002	0.057	0.043	0.003
ρ_{21}	0.02	–	–	–	–	–	–	0.040	0.037	0.002
ρ_2	0.15	0.149	0.043	0.003	0.156	0.039	0.002	0.152	0.046	0.003
τ_1	0.30	0.275	0.129	0.027	0.314	0.129	0.028	0.306	0.120	0.023
τ_2	0.20	0.211	0.128	0.025	0.188	0.110	0.019	0.195	0.097	0.016

Table 4.6.2: Simulation results of estimates of model M2 under different heterogeneity structures H1–H3. For each parameter, we report its true value used in the DGP, the Mean of the estimates, the Mean Absolute Error (MAE), and the Mean Squared Error (MSE).

Simulation results for Model M3

		Heter. H1			Heter. H2			Heter. H3		
	True	**Mean**	**MAE**	**MSE**	**Mean**	**MAE**	**MSE**	**Mean**	**MAE**	**MSE**
ρ_1	0.25	0.196	0.070	0.008	0.201	0.071	0.007	0.198	0.058	0.006
ρ_{12}	0.05	–	–	–	0.047	0.030	0.001	0.058	0.043	0.002
ρ_{21}	0.02	–	–	–	–	–	–	0.039	0.037	0.002
ρ_2	0.15	0.137	0.051	0.004	0.136	0.049	0.004	0.137	0.046	0.004
β_1	0.40	0.399	0.113	0.021	0.416	0.106	0.019	0.397	0.123	0.025
β_2	0.60	0.615	0.100	0.016	0.594	0.114	0.021	0.593	0.104	0.016
δ_1	0.15	0.259	0.116	0.019	0.253	0.114	0.018	0.269	0.122	0.019
δ_{12}	0.03	–	–	–	0.030	0.041	0.003	0.028	0.051	0.004
δ_{21}	0.012	–	–	–	–	–	–	0.005	0.057	0.005
δ_2	0.09	0.154	0.084	0.014	0.148	0.083	0.013	0.144	0.078	0.011

Table 4.6.3: Simulation results of estimates of model M3 under different heterogeneity structures H1–H3. For each parameter, we report its true value used in the DGP, the Mean of the estimates, the Mean Absolute Error (MAE), and the Mean Squared Error (MSE).

Simulation results for Model M4

		Heter. H1			Heter. H2			Heter. H3		
	True	**Mean**	**MAE**	**MSE**	**Mean**	**MAE**	**MSE**	**Mean**	**MAE**	**MSE**
ρ_1	0.25	0.238	0.059	0.006	0.243	0.052	0.004	0.240	0.049	0.004
ρ_{12}	0.05	–	–	–	0.054	0.030	0.001	0.051	0.032	0.001
ρ_{21}	0.02	–	–	–	–	–	–	0.034	0.031	0.002
ρ_2	0.15	0.153	0.039	0.003	0.145	0.043	0.003	0.140	0.038	0.002
τ_1	0.30	0.324	0.113	0.020	0.297	0.108	0.018	0.268	0.131	0.024
τ_2	0.20	0.213	0.098	0.016	0.212	0.107	0.019	0.220	0.103	0.017
η_1	0.10	0.282	0.182	0.040	0.257	0.157	0.028	0.274	0.174	0.036
η_{12}	0.02	–	–	–	0.037	0.037	0.002	0.032	0.046	0.004
η_{21}	0.008	–	–	–	–	–	–	0.005	0.048	0.004
η_2	0.06	0.115	0.063	0.006	0.108	0.061	0.006	0.111	0.061	0.006

Table 4.6.4: Simulation results of estimates of model M4 under different heterogeneity structures H1–H3. For each parameter, we report its true value used in the DGP, the Mean of the estimates, the Mean Absolute Error (MAE), and the Mean Squared Error (MSE).

Simulation results for Model M5

		Heter. H1			Heter. H2			Heter. H3		
	True	**Mean**	**MAE**	**MSE**	**Mean**	**MAE**	**MSE**	**Mean**	**MAE**	**MSE**
ρ_1	0.25	0.214	0.061	0.005	0.208	0.066	0.006	0.221	0.057	0.005
ρ_{12}	0.05	–	–	–	0.046	0.027	0.001	0.059	0.040	0.002
ρ_{21}	0.02	–	–	–	–	–	–	0.042	0.037	0.002
ρ_2	0.15	0.149	0.040	0.003	0.146	0.041	0.002	0.153	0.039	0.002
τ_1	0.30	0.292	0.120	0.014	0.294	0.128	0.025	0.293	0.102	0.016
τ_2	0.20	0.205	0.102	0.014	0.221	0.084	0.012	0.206	0.099	0.015
η_1	0.10	0.274	0.175	0.035	0.291	0.192	0.041	0.293	0.193	0.043
η_{12}	0.02	–	–	–	0.037	0.036	0.002	0.027	0.050	0.004
η_{21}	0.008	–	–	–	–	–	–	0.005	0.046	0.004
η_2	0.06	0.116	0.063	0.006	0.109	0.066	0.005	0.117	0.065	0.006
β_1	0.40	0.399	0.113	0.021	0.397	0.126	0.028	0.423	0.117	0.021
β_2	0.60	0.607	0.140	0.017	0.595	0.107	0.017	0.590	0.106	0.018

Table 4.6.5: Simulation results of estimates of model M5 under different heterogeneity structures H1–H3. For each parameter, we report its true value used in the DGP, the Mean of the estimates, the Mean Absolute Error (MAE), and the Mean Squared Error (MSE).

4.7 Conclusion

This paper develops a novel spatiotemporal clustering framework for panel data that jointly accounts for dynamic dependence, spatial spillovers, and structural heterogeneity, both within and across groups of observational units. Unlike traditional models that impose homogeneity across space, our approach explicitly incorporates endogenous group formation and group-specific spatial dynamics,

providing a more realistic and flexible representation of complex interdependencies and diffusion processes over space and time. This methodological advancement is particularly important in empirical contexts where assuming uniform effects across units may lead to misleading conclusions.

Our key contribution lies in the development of a unified estimation framework that identifies both the latent cluster structure and the associated heterogeneous parameters. Specifically, we propose a two-step estimation strategy, where the first stage employs a QML approach to estimate cluster-specific spatial and temporal dependence, along with heterogeneous covariate effects and the second detects group membership. This represents a substantial generalization of existing spatial econometric techniques, allowing for richer forms of heterogeneity—including asymmetric spillovers and cross-cluster interactions—while remaining computationally tractable.

To validate the robustness and effectiveness of the proposed methodology, an extensive Monte Carlo simulation study is carried out. The simulation experiment is designed to assess the model's ability to accurately recover the true spatiotemporal parameters and to disentangle the direct and indirect effects under a variety of heterogeneity scenarios. The results confirm that our method performs well even under complex settings, including highly asymmetric and group-specific interaction structures.

Beyond its methodological contributions, the framework has important practical implications. It provides researchers and practitioners with a powerful tool for uncovering the true underlying structure of complex phenomena, particularly when interactions and dependencies are not uniform across observational units. By identifying how different variables influence one another within and across heterogeneous groups, the model can yield deeper insights into context-specific dynamics. Moreover, the ability to detect and characterize group-specific spatial behavior has significant value for policymakers, who can use such information to design more effective and targeted interventions. For example, in environmental or regional policy, identifying clusters of regions with similar emission patterns or responsiveness to environmental drivers allows for the formulation of tailored, group-specific strategies that are more likely to be effective than one-size-fits-all approaches.

We are currently applying this model to the study of GHG emissions across European regions. The empirical setting leverages a rich dataset of regional emissions, sourced from EDGAR v8.0 and harmonized at the NUTS-2 level through the European Commission's ARDECO database. The

data cover 234 regions across 27 EU countries over the period from 1990 to 2022, enabling us to capture both long-run temporal dynamics and fine-grained spatial variation in emissions behavior. Our outcome of interest is total per capita GHG emissions, aggregated over all gases and sectors, with supplementary spatial descriptors such as emissions per km^2 also available.

The spatial structure of the model is informed by a geodetic distance-based spatial weights matrix, allowing us to quantify the role of geographical proximity in shaping regional spillovers. In parallel, the inclusion of socio-economic covariates—such as income, education, sectoral composition, and innovation capacity—permits the exploration of heterogeneous drivers of emissions across different types of regions.

Our ongoing empirical analysis aims to uncover how these socio-economic variables interact with spatial linkages to explain both intra-cluster and inter-cluster spillovers in emissions, and how the marginal effects of key variables differ across structurally distinct groups of regions. Ultimately, this approach offers a more accurate and policy-relevant understanding of regional emission dynamics, which is essential for designing targeted climate policies that account for spatial externalities and structural heterogeneity across regions.

Anselin, L., & Arribas-Bel, D. (2013). Spatial fixed effects and spatial dependence in a single cross-section. *Papers in Regional Science*, *92*(1), 3–17. https://doi.org/10.1111/j.1435-5957.2012.00480.x

Anselin, L., Le Gallo, J., & Jayet, H. (2008, January). Spatial panel econometrics. https://doi.org/10.1007/978-3-540-75892-1_19

Aquaro, M., Bailey, N., & Pesaran, M. H. (2021). Estimation and inference for spatial models with heterogeneous coefficients: An application to us house prices. *Journal of Applied Econometrics*, *36*(1), 18–44. https://doi.org/10.1002/jae.2792

Bailey, N., Holly, S., & Pesaran, M. H. (2016). A two-stage approach to spatio-temporal analysis with strong and weak cross-sectional dependence. *Journal of Applied Econometrics*, *31*(1), 249–280. https://doi.org/10.1002/jae.2468

Billé, A. G., Benedetti, R., & Postiglione, P. (2017). A two-step approach to account for unobserved spatial heterogeneity. *Spatial Economic Analysis*, *12*(4), 452–471. https://doi.org/10.1080/17421772.2017.1286373

Boccaletti, S., Maranzano, P., & Viegas, M. (2024). Inequality and concentration in farmland production and size: Regional analysis for the european union from 2010 to 2020. https://arxiv.org/abs/2409.00111

Bonhomme, S., & Manresa, E. (2015). Grouped patterns of heterogeneity in panel data. *Econometrica*, *83*(3), 1147–1184. https://doi.org/10.3982/ECTA11319

Cerqueti, R., D'Urso, P., & Mattera, R. (2025). Fuzzy group fixed-effects estimation with spatial clustering. *AStA Advances in Statistical Analysis*. https://doi.org/10.1007/s10182-024-00522-1

Chen, J., Shin, Y., & Zheng, C. (2022). Estimation and inference in heterogeneous spatial panels with a multifactor error structure. *Journal of Econometrics*, *229*(1), 55–79. https://doi.org/10.1016/j.jeconom.2021.05.003

Elhorst, J. P. (2001). Dynamic models in space and time. *Geographical Analysis*, *33*(2), 119–140. https://doi.org/10.1111/j.1538-4632.2001.tb00440.x

Elhorst, J. P. (2010). Applied spatial econometrics: Raising the bar. *Spatial Economic Analysis*, *5*(1), 9–28. https://doi.org/10.1080/17421770903541772

Elhorst, J. P. (2024). Raising the bar in spatial economic analysis: Two laws of spatial economic modelling. *Spatial Economic Analysis*, *19*(2), 115–132. https://doi.org/10.1080/17421772.2024.2334845

Elhorst, J. P., Gross, M., & Tereanu, E. (2021). Cross-sectional dependence and spillovers in space and time: Where spatial econometrics and global var models meet. *Journal of Economic Surveys*, *35*(1), 192–226. https://doi.org/10.1111/joes.12391

Feo, G., Giordano, F., Milito, S., Niglio, M., & Parrella, M. L. (2024). Clustering and classification of spatio-temporal data using spatial dynamic panel data models. *Advances in Data Analysis and Classification*. https://doi.org/10.1007/s11634-024-00620-7

Goodchild, M. F. (2004). The validity and usefulness of laws in geographic information science and geography. *Annals of the Association of American Geographers*, *94*(2), 300–303. https://doi.org/10.1111/j.1467-8306.2004.09402008.x

Koopman, S. J., Schaumburg, J., & Wiersma, Q. (2021). Joint modelling and estimation of global and local cross-sectional dependence in large panels.

Lee, J., Gangnon, R. E., & Zhu, J. (2017). Cluster detection of spatial regression coefficients. *Statistics in Medicine*, *36*(7), 1118–1133. https://doi.org/10.1002/sim.7172

Lee, L.-F., & Yu, J. (2010a). Estimation of spatial autoregressive panel data models with fixed effects. *Journal of Econometrics*, *154*(2), 165–185. https://doi.org/doi.org/10.1016/j.jeconom.2009.08.001

Lee, L.-F., & Yu, J. (2010b). A spatial dynamic panel data model with both time and individual fixed effects. *Econometric Theory*, *26*(2), 564–597. https://doi.org/10.1017/S0266466609100099

LeSage, J., & Pace, R. K. (2009). *Introduction to spatial econometrics*. Taylor & Francis Inc.

LeSage, J. P. (2008). An introduction to spatial econometrics. *Revue d'Economie Industrielle*, *123*(3), 19–44. https://doi.org/10.4000/rei.3887

Morelli, C., Maranzano, P., & Otto, P. (2025). Spatiotemporal clustering of GHGs emissions in Europe: Exploring the role of spatial component [arXiv 2503.11909]. https://arxiv.org/abs/2503.11909

R., C., P., M., & R., M. (2025). Spatially-clustered spatial autoregressive models with application to agricultural market concentration in Europe. *Journal of Agricultural, Biological, and Environmental Statistics*. https://doi.org/10.1007/s13253-024-00672-4

Shi, W., & Lee, L.-f. (2017). Spatial dynamic panel data models with interactive fixed effects. *Journal of Econometrics*, *197*(2), 323–347. https://doi.org/https://doi.org/10.1016/j.jeconom.2016.12.001

Simonovska, R. (2025). Sdpdmod: An r package for spatial dynamic panel data modeling. *Computational Economics*. https://doi.org/10.1007/s10614-025-11056-2

Su, L., Wang, W., & Xu, X. (2023). Identifying latent group structures in spatial dynamic panels. *Journal of Econometrics*, *235*(2), 1955–1980. https://doi.org/10.1016/j.jeconom.2023.02.007

Sugasawa, S., & Murakami, D. (2021). Spatially clustered regression. *Spatial Statistics*, *44*. https://doi.org/10.1016/j.spasta.2021.100525

Tobler, W. R. (1970). A computer movie simulating urban growth in the detroit region. *Economic Geography*, *46*, 234–240.

Yu, J., de Jong, R., & Lee, L.-f. (2008). Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and t are large. *Journal of Econometrics*, *146*(1), 118–134. https://doi.org/doi.org/10.1016/j.jeconom.2008.08.002

Zhu, A.-X., Lu, G., Liu, J., Qin, C.-Z., & Zhou, C. (2018). Spatial prediction based on third law of geography. *Annals of GIS*, *24*(4), 225–240. https://doi.org/10.1080/19475683.2018.1534890

Zhu, A.-X., & Turner, M. (2022). How is the third law of geography different? *Annals of GIS*, *28*(1), 57–67. https://doi.org/10.1080/19475683.2022.2026467

Conclusions

This thesis set out from a simple premise: environmental sustainability is not an isolated problem to be measured once and for all, but a networked, evolving system in which markets, firms, and regions shape—and are shaped by—each other. Across the studies, three messages emerge. First, sustainability is economically material. It shows up in prices, in risk, and in resilience when shocks hit. Second, sustainability is spatial. Performance and pressure cluster across places and sectors, and geography adds explanatory power beyond what firm or sector characteristics alone can capture. Third, sustainability is dynamic and heterogeneous. Units do not respond identically to the same forces; their interactions and trajectories differ in ways that matter for both diagnosis and policy design.

Taken together, these insights argue for a pragmatic synthesis: use the richness of today's data to build models that are flexible enough to capture interdependence and heterogeneity, but disciplined and interpretable enough to inform action. The empirical results show that environmental strength can cushion firms in turbulent times, that "where" a firm is embedded helps explain "how" it performs on environmental dimensions, and that emissions trajectories coalesce into persistent, spatially structured patterns. The methodological developments then push this logic further, offering tools to detect and quantify spatial structure in complex time series and to estimate models where spillovers and responses vary by latent groups. In short, the contribution is both substantive and infrastructural: new evidence about how sustainability unfolds in markets and across regions, and new methods that make such evidence visible and usable.

The utility of this research is practical. Investors and risk managers can translate the first set of findings into portfolio construction and engagement strategies that put environmental quality at the core of resilience, not at its periphery. Corporate leaders can benchmark against spatially coherent peer groups to target improvements that are realistic given local conditions. Policymakers

can use cluster-based evidence to move beyond one-size-fits-all measures, prioritizing place-based interventions where they are likely to deliver the largest decarbonization gains and coordinating across regions where spillovers are strong. Statisticians and economists gain diagnostic tools—such as explicit measures of how much space contributes to pattern formation—and a modeling framework that reconciles local interactions with global shocks. The broader payoff is sharper targeting: resources are finite, and understanding where geography and group membership amplify or dampen change helps direct those resources wisely.

On the methodological front, several extensions invite attention. The proposed cluster-wise spatial dynamic panel model can be enriched to accommodate multi-layer connectivity—physical networks, trade ties, financial exposures—so that spillovers reflect the true channels through which shocks propagate. Endogenous network formation, where connections strengthen or weaken in response to policy or technology, is another natural step. Combining spatial dynamics with latent common factors will help parse what is global and what is local, especially in periods of synchronized shocks. Robustness to measurement error, missingness, and revisions is essential in environmental data; uncertainty quantification should be built into clustering and panel estimation so that decisions reflect not just point estimates but credible ranges. Finally, translating methods into open, well-documented software will lower adoption barriers and accelerate cumulative progress.

Policy evaluation is a particularly promising arena for future work. With spatially aware clusters and dynamic panels in hand, it becomes possible to assess whether interventions have effects that differ by group identity and to measure spillovers—intended or not—on neighboring regions. This opens the door to adaptive policy design: learning which combinations of carbon pricing, standards, subsidies, and infrastructure investments produce the steepest and most durable emissions declines in specific contexts, and iterating as evidence accumulates. The same logic applies in finance, where stewardship and disclosure requirements can be targeted at clusters most likely to deliver real-world impact rather than improvements confined to reporting.

Perhaps the most important lesson is conceptual. Sustainability cannot be managed through averages. Averages wash out the very structure—spatial, sectoral, networked—that determines who leads, who lags, and where interventions leverage the most change. By centering interdependence and heterogeneity, this thesis argues for a shift in how we diagnose problems and design solutions:

start with the map of connections and differences, then layer policies and capital where they can bend trajectories most. As data become more granular and timely, the opportunity is to maintain this balance between complexity and clarity—using sophisticated models to honor the system's richness while keeping outputs legible to the people who must act on them. If we do, we stand a better chance of aligning markets with environmental goals, accelerating decarbonization where it matters most, and building resilience in the face of shocks we cannot fully predict but can better withstand.

Acknowledgements

I would like to express my deepest gratitude to my supervisors, Prof. Paolo Maranzano and Prof. Philipp Otto, for their constant guidance, intellectual rigor, and generous support throughout this doctoral journey. Their mentorship has been fundamental in shaping both this research and my academic development. I am also sincerely grateful to my co-authors, Prof. Elisa Ossola and Prof. Simone Boccaletti, for their invaluable collaboration, insightful discussions, and scientific generosity. I would also like to thank the reviewers of this thesis, Prof. Raffaele Mattera and Prof. Roberta Pappadà, for their careful reading, constructive feedback, and valuable suggestions. My thanks extend to all colleagues of the Department of Economics at the University of Milano-Bicocca, whose stimulating environment and collegial spirit have made this path both enriching and inspiring.

I warmly thank my fellow PhD students, from my cohort and from other cycles, for the shared experiences, mutual support, and sense of community that accompanied these years.

I am profoundly grateful to my family for their unconditional support, patience, and encouragement throughout this demanding journey, and to my friends, who have always believed in me and sustained me with their constant presence.

Finally, I wish to thank Mattia, colleague, co-author, and partner, for his intellectual companionship, unwavering support, and daily encouragement, which have made this journey not only possible, but deeply meaningful.