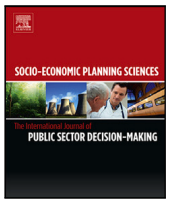
Socio-Economic Planning Sciences 105 (2026) 102469

Contents lists available at ScienceDirect

Socio-Economic Planning Sciences

journal homepage: www.elsevier.com/locate/seps



Exploring the sustainable fixed income market: A clustering-based approach with evidence from Italy and France☆

Giovanna Ferraro [a], Alessandro Ramponi [b,*], Saverio Storani [a]

[a] Department of Enterprise Engineering, University of Rome Tor Vergata, Italy
[b] Department of Economics and Finance, University of Rome Tor Vergata, Italy

ARTICLE INFO

Keywords:
GSS+ bonds
SDGs
Gower distance
K-medoid
Regional analysis

ABSTRACT

In the context of the global climate change and sustainable development, Green, Social, Sustainability, Sustainability-Linked, and Transition bonds (GSS+) have emerged as crucial instruments for financing projects with environmental and social objectives. This paper proposes a methodological framework to analyze GSS+ bonds by considering their financial and non-financial characteristics and their alignment with the United Nations Sustainable Development Goals (SDGs). It introduces a reproducible and scalable clustering approach to study this segment of the market. We perform an initial international exploratory analysis using a global dataset of more than 50,000 bond issuances and then a detailed case study focused on Italy and France. The Partitioning Around Medoids (PAM) algorithm is applied for clustering, and bond dissimilarities are calculated using the Gower distance, which is appropriate for mixed data types. The Rand index is used to assess the robustness of the clustering structure. The findings show different national characteristics: Italy's market is smaller and more concentrated, with a corporate focus on infrastructure and energy, while France presents a wider issuance landscape with a more interconnected use of proceeds and SDG structure.

1. Introduction

The urgency of raising funds for sustainable development has been highlighted in recent years by the acceleration of climate change and the rise of global social issues. The United Nations (UN) adopted the 2030 Agenda for Sustainable Development in 2015, which outlined 17 Sustainable Development Goals (SDGs) as part of a comprehensive plan to address issues of institutional fragility, poverty, inequality, and environmental degradation. Significant financial resources are needed to achieve these goals, and financial markets are increasingly asked to facilitate the transition to a sustainable economy [1–4]. This transition is further facilitated by technological innovations, including digital platforms and AI-driven solutions that enhance circular economy opportunities and resource allocation efficiency [5]. Within this context, sustainable fixed-income instruments, namely Green, Social, Sustainability, Sustainability-Linked, and Transition bonds, or GSS+ bonds[1] have emerged as key components of the sustainable finance architecture [6,7]. These instruments enable issuers to finance projects with measurable environmental and social impacts [8–10], while allowing investors to align financial returns with sustainability objectives. Since the first green bond was issued in 2008 by the World Bank and the Nordic Investment Bank, the GSS+ market has grown rapidly in both volume and diversity [11].

The GSS+ bonds include five main categories, each designed to serve a specific sustainability purpose. (i) Green Bonds fund projects aimed at achieving positive environmental impacts, such as renewable energy development, energy efficiency improvements, pollution prevention, and climate change adaptation. (ii) Social Bonds finance projects that address social challenges, particularly those benefiting vulnerable populations. (iii) Sustainability Bonds combine elements of both green and social bonds, financing projects that deliver environmental and social benefits. (iv) Sustainability-Linked Bonds incentivize improvements in the issuer's overall sustainability performance, rather than earmarking proceeds. Their financial characteristics, such as the coupon rate, adjust depending on whether sustainability targets are met. (v) Transition Bonds finance emission-reduction efforts in hard-to-abate sectors, such as heavy industry or aviation, enabling the shift towards more sustainable industrial processes.

☆ This article is part of a Special issue entitled: 'Statistical Measurement of Economic Analysis' published in Socio-Economic Planning Sciences.

* Corresponding author.
E-mail addresses: giovanna.ferraro@uniroma2.it (G. Ferraro), alessandro.ramponi@uniroma2.it (A. Ramponi), saverio.storani@uniroma2.it (S. Storani).

[1] The acronym GSS+ was introduced in World Bank Group report from October 2022, *Sovereign Green, Social and Sustainability Bonds: Unlocking the Potential for Emerging Markets and Developing Economies*, https://thedocs.worldbank.org/en/doc/4de3839b85c57eb958dd207fad132f8e-0340012022/original/WB-GSS-Bonds-Survey-Report.pdf.

https://doi.org/10.1016/j.seps.2026.102469
Received 25 July 2025; Received in revised form 10 January 2026; Accepted 6 March 2026
Available online 7 March 2026

These instruments are issued by various entities, including financial institutions, corporations, governments, and supranational organizations, reflecting the broad institutional involvement in sustainable finance [12,13].

Despite the expanding literature on sustainable finance, existing studies largely focus on aggregate issuance volumes, individual bond categories, or specific financial dimensions. As a result, limited attention has been paid to how financial characteristics, use of proceeds (UoP), issuer profiles, and SDG alignment jointly shape the internal structure of the GSS+ bond market. Consequently, the latent heterogeneity of GSS+ instruments and systematic differences across national contexts remain insufficiently explored.

The literature on sustainable finance and GSS+ bonds can be broadly divided into two strands: a conceptual stream that establishes the institutional and theoretical foundations of sustainable fixed-income markets [14], and an empirical stream that analyzes bond characteristics and market behavior [15].

While recent contributions have begun to address multidimensional aspects of GSS+ bonds [16], significant gaps persist, particularly with respect to national specificities and methodological approaches capable of capturing the joint role of financial and non-financial attributes. To address these gaps, this paper adopts a clustering-based approach tailored to mixed-type data and focuses on both a global exploratory analysis and a comparative case study of Italy and France. Against this background, the study addresses the following research questions:

RQ1: What latent structures characterize the GSS+ bond market when financial and non-financial attributes are jointly considered, and to what extent do these structures reflect a thematic segmentation aligned with the SDGs?

RQ2: How do differences in institutional settings and market maturity influence the internal composition and thematic integration of GSS+ bond clusters in the Italian and French markets?

By addressing these questions, the paper aims to contribute to the sustainable finance literature by providing a more granular understanding of the internal architecture of the GSS+ market and its policy-relevant implications.

In particular, this paper adds to the expanding literature on sustainable finance by presenting a methodological framework based on clustering that is specifically designed to handle the multidimensional and heterogeneous nature of GSS+ bonds. Given that no ex-ante classification fully captures the jointly variation of financial and non-financial attributes, an unsupervised learning approach provides a natural framework to uncover latent structures in the market.

Through the implementation of the Partitioning Around Medoids (PAM) algorithm (see [17]), and the Gower distance [18], a well-established metric appropriate for mixed-type datasets, we are able to detect more homogeneous groups of GSS+ bonds and reveal latent patterns across various market segments. This approach allows us to subsequently observe SDG alignment, revealing latent patterns and synergies across various market segments. In addition, the Rand index [19] is used to evaluate the robustness of the clustering solution, enabling us to confirm that the results remain stable under different weighting schemes.

The empirical analysis is carried out in two stages. First, we used data from Environmental Finance Data, a leading supplier of sustainable debt tracking, to perform a global exploratory overview of over 50,000 GSS+ bonds. In this preliminary analysis, key trends in issuance, thematic concentrations, and regional differences are highlighted. Second, we focus on a comparative case study of Italy and France, two major European economies that share common sustainability commitments but differ in institutional settings and market maturity, using a clustering-based methodology (see, e.g., [20]). This dual perspective enables us to combine a global overview with a detailed national-level investigation of market structures.

Our results show significant variations between the two countries, starting with the scale of the market (272 issuances in Italy and 1475 in France). In Italy, GSS+ bonds are predominantly issued by corporate entities, with a higher average issuance size and shorter maturities. These bonds are primarily focused on energy and infrastructure, resulting in a more concentrated pattern of both UoP and SDG targeting. In contrast, the French market appears more mature and systemic, characterized by a larger role for financial institutions and a more diversified issuer base. The resulting clusters effectively highlight these structural differences. Overall, the French market shows more intricate bond structures and higher thematic integration, indicating a more advanced stage of development in sustainable finance compared to Italy.

The paper is organized as follows. Section 2 introduces the data and the variables. Section 3 presents the methodological framework, including the dissimilarity metric, the clustering algorithm, and the robustness checks. Section 4 reports the results of the global exploratory analysis and the comparative case study of Italy and France. Section 5 concludes with a discussion of the implications, limitations, and directions for future research.

2. Data source and collection

2.1. Data source

The dataset related to GSS+ used in this research project is provided by Environmental Finance Data (EFD),[2] a company specialized in reporting on sustainable investment and green finance. The dataset covers more than 95% of the sustainable bond market and includes various types of bonds, such as corporate bonds, green asset-backed and mortgage-backed securities, green project bonds, sovereign and supranational bonds, state and municipal bonds, and green guarantees. The database incorporates a wider range of sustainability-focused financial instruments and contains more than 50,000 issuances, covering the period from 2007 to the present.

2.2. Data collection

For the purpose of this study, we selected a specific subset of features considered the most relevant to our analytical goals. The extracted information includes key structural, financial, and geographical characteristics. Each observation in the dataset corresponds to a single *bond issuance*, and multiple issuances may originate from the same issuer.

A relevant selected variable is the *issuance label*, which classifies each bond according to its environmental or social objectives. Five main categories are identified: Green bonds, Social bonds, Sustainability bonds, Sustainability-Linked bonds, and Transition bonds. This variable is instrumental in distinguishing the issuance orientation. Then, among the financial characteristics selected, we consider the *dollar value*, converted to US dollars for comparability, the *settlement date*, and the *maturity date* used to determine the bond tenor. We also consider the *coupon rate*, which is originally expressed as a percentage, which has been grouped into the floating and fixed rates (from now on define as *coupon type*). In addition, geographical and institutional variables play a central role in our analysis. These include the issuer's *country code* and *region*, categorized into ten macro-areas: Europe, Asia, North America, South America, Central America, Middle East, Africa, Oceania, Caribbean, and Supranational.

[2] The data are available on the website https://efdata.org, which is a subscription-based service that ensures high-quality and reliable information.

Table 1
Issuer sectors categories.

Issuer sector (Column 1)	Issuer sector (Column 2)
Government — municipal/local	Housing finance institution
Financial	Public transportation
Utilities	Real Estate — development and management
Water	Energy — renewable: other
Energy — renewable: solar	Education
Energy — renewable: wind	Energy — non-renewable
Real Estate — construction and construction materials	Government — national
Other	Waste/Recycling
Real Estate — REIT	Housing association
Healthcare	Automotive
Manufacturing — other	Technology
Food/beverage	Telecommunications
Industrial machinery and engineering	Chemicals, plastics, and rubber
Paper and packaging	Logistics — shipping and maritime
Mining/metals	Logistics — warehousing
Oil and Gas	Agriculture/Fisheries
Aviation	Financial — insurance
Retail	Textiles and apparel
Tourism and hospitality	Pharmaceuticals

Table 2
Bond issuance Use of Proceeds categories.

Use of proceeds (Column 1)	Use of proceeds (Column 2)
Affordable housing	Green Buildings
Access to essential services	Sustainable water management
Socioeconomic advancement and empowerment	Energy efficiency
Renewable energy	Clean transportation
Affordable basic infrastructure	Pollution prevention and control
Climate change adaptation	Sustainable management of living natural resources
Employment generation including through the potential effect of SME financing and microfinance	Terrestrial and aquatic biodiversity conservation
Eco-efficient products	Production technologies and processes
Food security	General Corporate Purposes
Covid-19 response	Other eligible green categories
GB5 — Clean energy	Other eligible social categories
GB1 — Energy saving	GB3 — Resource Conservation and Recycling
GB4 — Clean transportation	GB6 — Ecological protection and climate change adaptation
GB2 — Pollution Prevention and Control	

We also distinguish between different *issuer types*, representing a highly heterogeneous set of entities: Municipal, Agency, Corporate, Financial Institution, Supranational,[3] and Sovereign. Finally, three additional variables enrich the dataset from sectoral, target, and sustainability perspectives. First, the *issuer sector* captures the economic activity of the bond issuer, ranging from energy and finance to telecommunications and agriculture (see Table 1 for a complete overview). Second, the *use of proceeds* (UoP) describes the intended purpose of the funds raised, including categories such as renewable energy, green buildings, clean transportation, and adaptation to climate change, among others (see Table 2). Lastly, the *SDGs* variable indicates the specific United Nations Sustainable Development Goals (SDGs) targeted by each bond, based on the 17 global goals established in the 2030 Agenda for Sustainable Development.

This set of variables allows us to explore the GSS+ bond market from different perspectives and serves as the foundation for the descriptive analysis in the next section.

3. Research design and methodology

This section outlines the methodological framework used to identify common components between GSS+ bonds. In our context, clustering techniques offer a valuable approach to discover hidden patterns within complex and high-dimensional datasets, such as those found in fixed income markets. The use of clustering to analyze heterogeneous populations has been successfully applied in other contexts, such as health insurance policyholders [21]. In the case of GSS+ bonds, issuers and instruments show relevant heterogeneity across various dimensions, such as sector, geography, UoP, and financial characteristics, among others. By creating representative groupings or *centroids*, clustering can support unsupervised classification, enable meaningful comparisons, inform benchmarking practices, and contribute to the empirical design of portfolio strategies (see [22]).

In order to apply the clustering procedure, we first introduce a weighted dissimilarity measure to capture the differences between the issues according to their characteristics. A range of methods exists for computing dissimilarities, including random forest proximity [23,24], distance metric learning [25,26], PCA mixed data [27], and generalized dissimilarity modeling [28]. However, many of these approaches either require supervised settings, impose strong parametric assumptions, or lack interpretability. In contrast, we adopt the Gower distance [18], a well-established metric designed to handle mixed-type data by computing normalized pairwise dissimilarities. Gower's metric has been extensively used in clustering applications and remains a standard approach in unsupervised settings due to its flexibility, scalability, and interpretability [17,18,29]. This enables a coherent integration of heterogeneous features in the construction of the dissimilarity matrix. Ensuring methodological rigor is crucial when analyzing such complex sustainability-related datasets, as emphasized in recent studies [30].

Next, we apply the numerical algorithm to cluster the data set issues based on these dissimilarities. The chosen clustering method is PAM (Partitioning Around Medoids). Although there are alternatives such as

[3] A supranational organization, as a bond issuer, refers to an international institution formed by multiple countries that issues debt (bonds) to finance projects with cross-border or global objectives. Examples: World Bank, European Investment Bank, Nordic Investment Bank.

DBSCAN or hierarchical agglomerative clustering, [31,32], in this paper we adopt the PAM approach as it results more robust than k-means when dealing with non-Euclidean distance measures and less sensitive to outliers (see, [17,33–35]). In fact, unlike k-means, which requires purely numerical data and assumes spherical cluster shapes due to its reliance on Euclidean distance and centroids [36], and fuzzy c-means, which similarly assumes numerical input and produces soft partitions based on centroids [37], PAM operates directly on any dissimilarity matrix and selects actual data points (medoids) as cluster representatives. Then, we determine the optimal number of clusters by evaluating different configurations and selecting the one that maximizes cluster quality. Finally, we compare cluster partitions under different choices for the weights of the dissimilarity measure exploiting the Rand Index. These steps are detailed in the following sections.

3.1. Clustering process

To cluster GSS+ issues based on their characteristics, we consider a dissimilarity measure that accounts for both numerical and categorical variables. Given the heterogeneity of the dataset, comprising numerical and categorical features, see Section 2.2, we employ a distance metric suitable for mixed-type data, in particular the Gower distance [18].

In light of this, we define X our dataset with n observations (issues), each described by p variables, where the first h variables are numerical, and the remaining $p-h$ variables are categorical. Let $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ip})$ and $\mathbf{x}_j = (x_{j1}, x_{j2}, \dots, x_{jp})$ denote two different issues. To account for the potentially different relevance of each variable, we introduce a set of non-negative weights w_k, for $k = 1, \dots, p$, where p is the total number of variables. These weights determine how much each variable contributes to the calculation of the total distance between two issuers. The weighted Gower dissimilarity between two issuers i and j is then defined as follows:

$$d(\mathbf{x}_i, \mathbf{x}_j) = \frac{\sum_{k=1}^{p} w_k \cdot \delta_{i,j,k} \cdot d_k(x_{ik}, x_{jk})}{\sum_{k=1}^{p} w_k \cdot \delta_{i,j,k}}. \tag{1}$$

The coefficients $\delta_{i,j,k}$ represent an extension of Gower's dissimilarity introduced by [17], which properly handles missing values in the distance computation. They are defined as follows:

$$\delta_{i,j,k} = \begin{cases} 1, & \text{if the variable } k \text{ is observed for both } i \text{ and } j \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

The partial dissimilarity $d_k(x_{ik}, x_{jk})$ is computed depending on the nature of variable k:

$$d_k(x_{ik}, x_{jk}) = \begin{cases} \frac{|x_{ik} - x_{jk}|}{R_k}, & \text{if variable } k \text{ is numerical} \\ 1, & \text{if variable } k \text{ is categorical and } x_{ik} \neq x_{jk} \\ 0, & \text{if variable } k \text{ is categorical and } x_{ik} = x_{jk}. \end{cases} \tag{3}$$

Here, $R_k = \max_i(x_{ik}) - \min_i(x_{ik})$ represents the range of observed values for the numerical variable k. This normalization ensures that the differences between the values of variable k are scaled to the interval $[0, 1]$, allowing all numerical variables to contribute comparably to the overall dissimilarity, regardless of their original scales.

Once the dissimilarity matrix is calculated, we apply a clustering algorithm to detect clusters of issues with similar characteristics. For this purpose, we use the PAM algorithm. The goal is to partition the data into an apriori fixed number $C \in \mathbb{N}$ of groups in such a way that intra-cluster dissimilarity is minimized, and inter-cluster dissimilarity is maximized. The PAM algorithm minimizes the following objective function:

$$\min_{\mathbf{U},\mathbf{M}} \sum_{l=1}^{C} \sum_{q=1}^{N} u_{ql} d(q, \mu_l), \tag{4}$$

where $\mathbf{U} = (u_{ql} : q = 1, \dots, N; l = 1, \dots, C)$ is the membership matrix, indicating the assignment of issue q to cluster l (with $u_{ql} = 1$ if issue q belongs to cluster l, and $u_{ql} = 0$ otherwise). The matrix $\mathbf{M} = (\mu_l : l = 1, \dots, C)$ represents the cluster prototypes (called medoids), where μ_l is the medoid of the cluster l. A medoid is a data point within a cluster that is most centrally located based on the chosen dissimilarity metric. The term $d(q, \mu_l)$ denotes the dissimilarity between cluster member q and the lth cluster medoid.

Having set a number C of clusters, the algorithm alternates between assigning issues to the closest medoid and recalculating the medoids based on the current assignments, iterating until convergence [38]. This process results in the optimal partition of issues into C communities for the dataset at hand.

3.1.1. Cluster optimization and partitioning

Additionally, rather than predefining the number C of clusters, we adopt an iterative approach to determine their optimal number. For each possible number of clusters $C = 2, \dots, \bar{C}$, we compute the clustering solution and evaluate its quality using the Average Silhouette Width (ASW) criterion [39,40]. The ASW provides a measure of how well-separated the clusters are, with values close to +1 indicating well-defined clusters. The optimal number corresponds to the configuration that maximizes the ASW, ensuring that the clustering solution is coherent and meaningful. In practice, for each possible number of clusters, we calculate the silhouette for each issue and then compute the average silhouette width. The configuration with the highest ASW value is selected as the final clustering solution. This method helps identify the best partition of issues based on the balance between cohesion within clusters and separation between clusters.

Finally, we want to evaluate the stability between two partitions of the same set of elements when the weight w_k changes. We use the Rand Index [19], a commonly adopted measure in community detection and clustering analysis [41].

Let C_1 and C_2 be two partitions of the same set X of n elements. The Rand Index is defined as

$$\mathrm{RI}(C_1, C_2) = \frac{a + b}{\binom{n}{2}}, \tag{5}$$

where a is the number of pairs of elements assigned to the same community in both partitions, b is the number of pairs assigned to different communities in both partitions, and $\binom{n}{2} = \frac{n(n-1)}{2}$ is the total number of possible pairs. The Rand Index ranges from 0 to 1, where 1 indicates identical partitions (perfect agreement), 0 indicates completely discordant partitions, and intermediate values represent varying degrees of similarity between the two partitions. Values above 0.5 can be interpreted as moderate to high similarity, while lower values indicate less agreement.

4. Empirical results: From global overview to case studies

This section presents the empirical findings of the study, structured in two steps. Firstly, we provide a global exploratory overview of the GSS+ bond market to identify broad thematic and regional trends. Afterward, we conduct a granular comparative case study of France and Italy using a clustering-based methodological framework described in Section 3.

4.1. Exploratory analysis

In this subsection, we provide an overview of the GSS+ bond market, focusing on a selected set of key variables. This step offers insights into the overall composition, size, and structure of the market.

Fig. 1 provides a comprehensive overview of the geographical distribution of bond issuances. In Panel (a), the circular bar plot highlights the dominance of a few countries, with the United States standing out from the others. Additionally, European countries are significant players in terms of the number of issuances, reflecting an increasing attention to the GSS+ market. Furthermore, in Asia, we can observe countries like China, Japan, and South Korea, which have a high number of issuances, denoting a GSS+ market that is increasingly capable

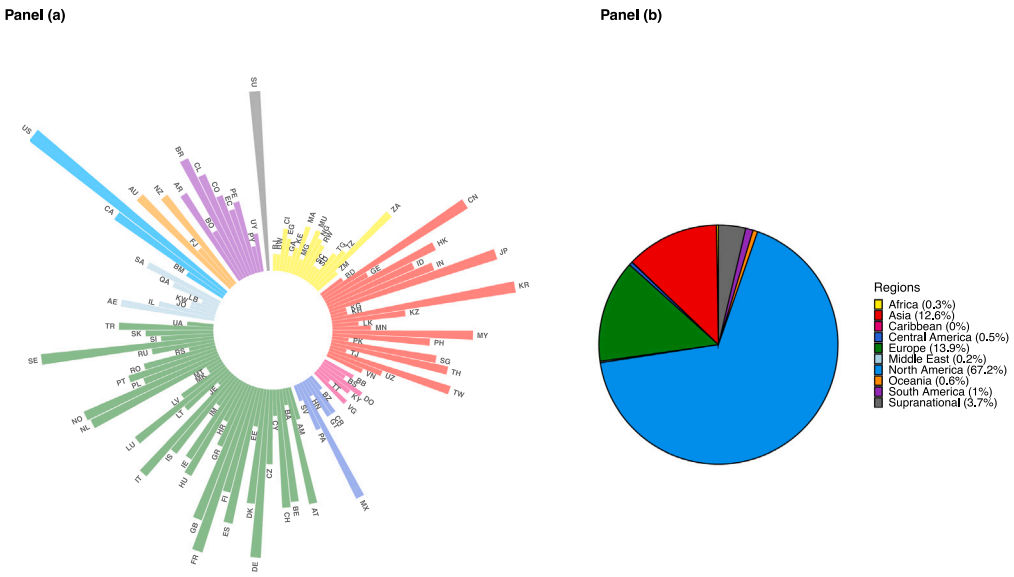


Fig. 1. Panel (a) displays a circular bar plot representing all issuer countries, where the height of each bar corresponds to the logarithm of the number of bonds issued by country. Panel (b) shows a pie chart illustrating the global market configuration, where each slice represents the percentage share of total GSS+ bond issuance by regions. In both panels, colors indicate the division of countries into world regions.

of competing with European players in terms of issuance volume and market activity. Panel (b) complements the country-level analysis by summarizing bond issuances at the regional level through a pie chart, revealing a highly uneven geographical distribution. North America dominates the market, accounting for 34,857 issuances (67.2%), followed by Europe with 7183 issuances (13.9%) and Asia with 6524 issuances (12.6%). Supranational entities represent a smaller yet notable share of the market, with 1903 issuances (3.7%). In contrast, all other regions contribute marginally to the total volume: South America counts 519 issuances (1.0%), Oceania follows with 335 (0.6%), Central America with 240 (0.5%), Africa with 137 (0.3%), and the Middle East with 102 (0.2%). The Caribbean region, with just 22 issuances, accounts for less than 0.1% of global activity.

Fig. 2 shows in Panel (a) a Venn diagram illustrating the intersections among the five issuance labels described in Section 2.2. The diagram shows that most bonds are associated with a single label, suggesting that issuers tend to focus on a specific category. However, some overlaps are observed, particularly between Green and Social Bonds. This visual representation highlights the degree of co-occurrence among labels within the dataset, providing an initial overview of how frequently bonds are classified under multiple labels. In the same figure, Panel (b) shows the distribution of GSS+ issuances with bars representing their log-transformed dollar value. The logarithmic scale accommodates the wide range of issuance sizes, enabling meaningful comparisons between dominant categories like Green and Social Bonds and smaller ones such as Sustainability-Linked Bonds. Europe, Asia, and North America lead in both volume and value, reflecting mature markets with robust financial infrastructures. In contrast, Africa and South America exhibit lower issuance values, highlighting emerging markets with significant growth potential.

The boxplot in Fig. 3 reveals the differences in bond maturity profiles among different types of issuers. Sovereign and Municipal issuers tend to have longer average maturities, which may reflect their public sector nature and the need to finance long-term infrastructure or public projects with extended debt horizons. In contrast, Financial Institutions generally issue bonds with shorter maturities, possibly due to their focus on more flexible funding needs and regulatory capital management. Agencies, Corporate, and Supranational issuers display intermediate tenor profiles with similar average durations, suggesting a balance between long-term investment goals and market flexibility.

Table 3 presents the distribution of bond issuances across different economic sectors, distinguishing between issuers classified within multiple sectors (“combined”) and those uniquely associated with a single sector (“alone”). Notably, sectors such as Government — municipal/local and Housing finance institution exhibit the highest total occurrences, reflecting their significant presence in the market. Renewable energy sectors, including solar and wind, also show substantial activity, highlighting the growing emphasis on sustainable finance. In contrast, sectors like Pharmaceuticals and Tourism and hospitality have relatively low counts, suggesting either emerging involvement or disinterest about GSS+ market.

Table 4 reports the 17 SDGs defined by the UN 2030 Agenda, while Fig. 4 presents the frequency distribution of the 17 SDGs across the dataset. The histogram, which can serve as an indicator of thematic focus within the sustainable debt market, reveals an uneven distribution of the SDGs. Goal 11 (Sustainable Cities and Communities) is the most frequently referenced, suggesting a strong focus on urban development, infrastructure, and housing within sustainable bond frameworks; Goal 1 (No Poverty) and Goal 7 (Affordable and Clean Energy) also show high representation, indicating that social inclusion and renewable energy effectively represent major priorities; Goals 10 (Reduced Inequalities) and 8 (Decent Work and Economic Growth) follow closely, reflecting attention to social justice and employment. In contrast, Goals 16 (Peace, Justice and Strong Institutions), 17 (Partnerships for the Goals), 5 (Gender Equality), and 7 (No hunger) appear rarely, suggesting less emphasis of such goals in sustainable finance instruments.

Finally, Fig. 5 concludes the exploratory analysis depicting a heat map for the UoP variable. The heatmap illustrates the co-occurrence frequencies among UoP categories. Darker red tiles represent stronger associations between categories that are frequently pursued together within the same bond issuance. The most prominent co-occurrences involve social-focused categories such as Socioeconomic advancement and empowerment, Affordable housing, and Access to essential services, highlighting a strong interlinkage among social development goals. Environmental categories like Green Buildings, Energy efficiency, and Renewable energy also show high levels of overlap, suggesting their frequent integration within climate-related financing frameworks. This pattern points to common thematic pairings in GSS+ structuring.

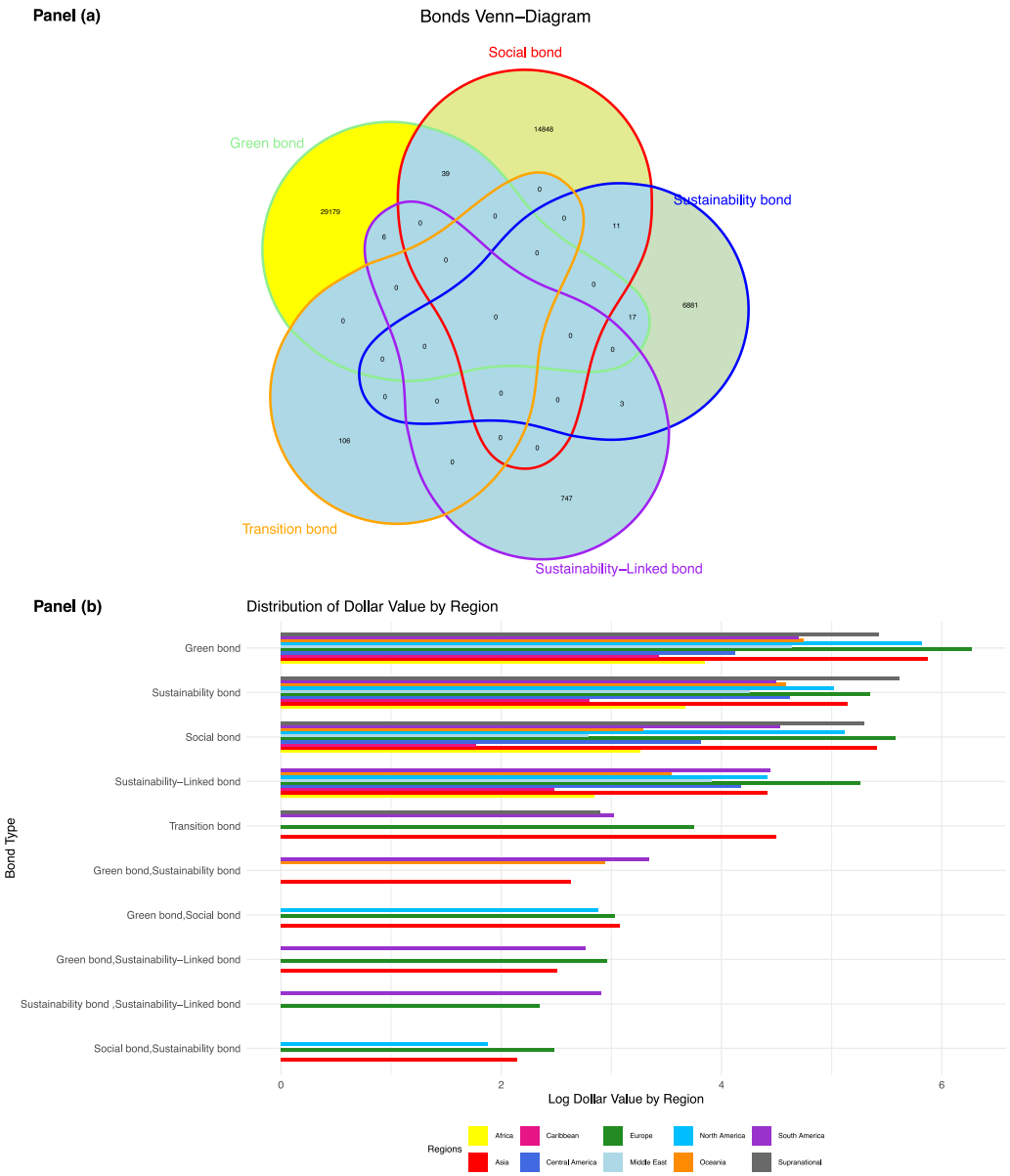


Fig. 2. Panel (a) is a Venn diagram showing the overlap among five bond labels. Each area indicates the count of bonds shared between sets. Color intensity (from light blue to yellow) reflects the number of overlapping items. Panel (b) contain the distribution of labeled GSS+ bond issuances by type and geographic region. The bars represent the log-transformed dollar value of bond issuances across regions, grouped by bond labels. Log scale is used to account for large disparities in issuance volumes between categories.

Towards a case study analysis. The global exploratory analysis highlights a fragmented GSS+ bond landscape, characterized by significant regional disparities and considerable heterogeneity across both issuance types and issuer profiles. These findings suggest that while broad global trends exist, the structural drivers of sustainable finance are deeply rooted in specific regional and national contexts. To better disentangle these complexities, we narrow our focus to the European market — a global leader in sustainability standards — by conducting a comparative analysis of France and Italy. These two countries represent a particularly interesting case study: as founding EU members and major economic players, they share a common commitment to the 2030 Agenda, yet they exhibit distinct issuance patterns and varying degrees of market maturity. By comparing these two pivotal countries, we aim to reveal how different institutional landscapes and market advancements influence the internal composition of the GSS+ universe, providing a granular test for our clustering methodology.

4.2. Case study

In this section, to deepen the global analysis, a focused geographical study was conducted on Italy and France using the clusterization methodology. This choice was motivated by the need to understand how the Italian financial and economic sectors are responding to sustainability objectives, particularly through the issuance of GSS+ bonds. In contrast, France stands out as one of the most active European countries in this domain (the number of issues in our dataset for Italy and France is 272 and 1475, respectively), with a well-established regulatory framework and a robust pipeline of green and sustainable issuances. Comparing these two countries could offer valuable insights into different national approaches to sustainable finance within the broader European context.

An initial insight emerges from the distribution of issuer types presented in Fig. 6, which highlights a clear structural divergence between the two countries. In Italy, corporate issuers dominate the market,

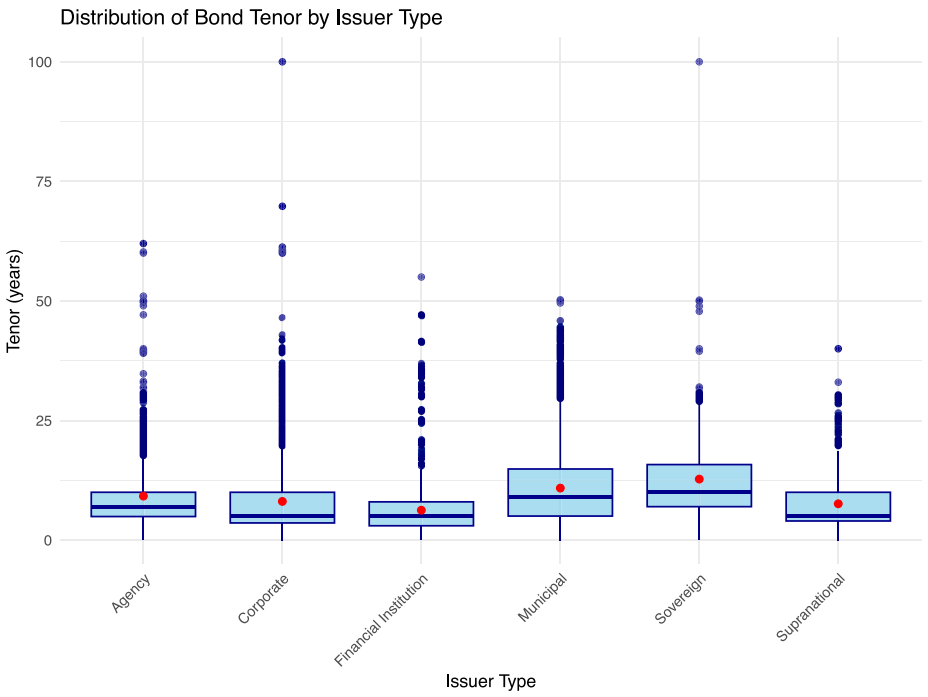


Fig. 3. Boxplot illustrating the distribution of bond tenors (in years) across different issuer types. The red dots represent the mean maturity for each issuer type. This visualization highlights differences in maturity profiles among issuer types, providing insights into their typical debt horizon.

Table 3
Number of occurrences by issuer sector, categorized into "combined", "alone", and "total". The table shows the distribution of occurrences across various economic sectors, with "combined" referring to issuers associated with multiple sectors, "alone" referring to issuers associated with a single sector, and "total" representing the overall occurrences for each sector.

	Issuer sector	Combined	Alone	Total
1	Agriculture/Fisheries	52	40	92
2	Automotive	46	194	240
3	Aviation	5	73	78
4	Chemicals, plastics, and rubber	34	104	138
5	Education	789	152	941
6	Energy — non-renewable	605	93	698
7	Energy — renewable: other	839	258	1097
8	Energy — renewable: solar	667	363	1030
9	Energy — renewable: wind	747	73	820
10	Financial	814	8348	9162
11	Financial — insurance	5	69	74
12	Food/beverage	65	132	197
13	Government — municipal/local	8284	15448	23732
14	Government — national	2	554	556
15	Healthcare	224	35	259
16	Housing association	102	201	303
17	Housing finance institution	5144	10738	15882
18	Industrial machinery and engineering	65	74	139
19	Logistics — warehousing	82	19	101
20	Logistics — shipping and maritime	66	42	108
21	Manufacturing — other	87	146	233
22	Mining/metals	19	83	102
23	Oil and Gas	26	73	99
24	Other	352	164	516
25	Paper and packaging	40	81	121
26	Pharmaceuticals	0	28	28
27	Public transportation	965	776	1741
28	Real Estate — REIT	99	376	475
29	Real Estate — development and management	551	922	1473
30	Real Estate — construction and construction materials	473	172	645
31	Retail	19	18	37
32	Technology	111	106	217
33	Telecommunications	54	133	187
34	Textiles and apparel	18	18	36
35	Tourism and hospitality	12	17	29
36	Utilities	1064	496	1560
37	Waste/Recycling	452	45	497
38	Water	1243	187	1430

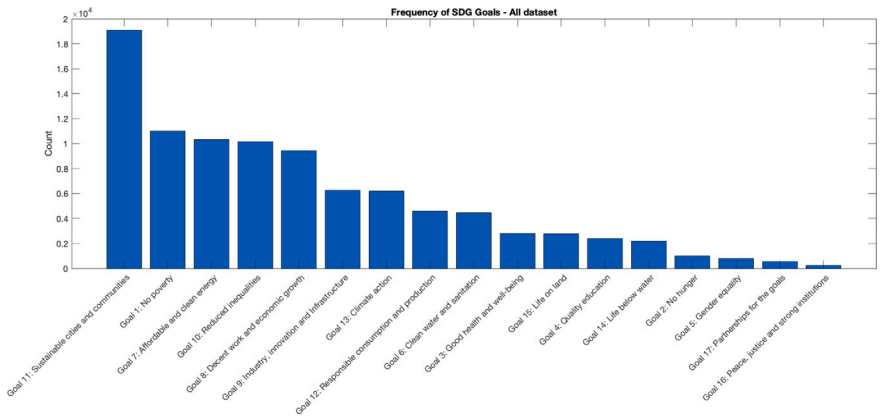


Fig. 4. Frequency distribution of the SDGs across the GSS+ dataset.

Table 4
The 17 Sustainable Development Goals (SDGs) of the UN 2030 Agenda with brief descriptions.

SDG	Goals	Description
1	No Poverty	Eradicate extreme poverty in all its forms everywhere, ensuring social protection and economic inclusion for all.
2	Zero Hunger	Achieve food security, improved nutrition, and promote sustainable agricultural practices globally.
3	Good Health & Well-Being	Ensure healthy lives and promote well-being for all at all ages, strengthening healthcare systems.
4	Quality Education	Ensure inclusive, equitable, and quality education, and promote lifelong learning opportunities for all.
5	Gender Equality	Achieve gender equality and empower all women and girls, eliminating discrimination and barriers.
6	Clean Water & Sanitation	Ensure availability and sustainable management of water and sanitation for everyone.
7	Affordable & Clean Energy	Ensure access to affordable, reliable, sustainable, and modern energy while improving energy efficiency.
8	Decent Work & Economic Growth	Promote sustained, inclusive, and sustainable economic growth, full productive employment, and decent work for all.
9	Industry, Innovation & Infrastructure	Build resilient infrastructure, promote inclusive and sustainable industrialization, and foster innovation.
10	Reduced Inequalities	Reduce inequality within and among countries, ensuring social, economic, and political inclusion for all.
11	Sustainable Cities & Communities	Make cities and human settlements inclusive, safe, resilient, and sustainable, improving urban living conditions.
12	Responsible Consumption & Production	Ensure sustainable consumption and production patterns, improving resource efficiency and reducing waste.
13	Climate Action	Take urgent action to combat climate change and its impacts through mitigation, adaptation, and resilience building.
14	Life Below Water	Conserve and sustainably use the oceans, seas, and marine resources for sustainable development.
15	Life on Land	Protect, restore, and promote sustainable use of terrestrial ecosystems, manage forests, and halt biodiversity loss.
16	Peace, Justice & Strong Institutions	Promote peaceful and inclusive societies, provide access to justice for all, and build accountable institutions.
17	Partnerships for the Goals	Strengthen the means of implementation and revitalize the global partnership for sustainable development worldwide.

accounting for more than half of total issues, whereas in France, the leading role is held by financial institutions.

On the other hand, analysis of SDG frequency distributions in both the Italian and French datasets reveals broadly similar patterns, marked by a clear prioritization of environmental and infrastructure-related goals, see Fig. 7. In both countries, Goals 7 (Affordable and Clean Energy), 13 (Climate Action), and 11 (Sustainable Cities and Communities) emerge as the most frequently addressed, underscoring the strong alignment of green bond issuances with climate mitigation, energy transition, and sustainable urban development objectives. Goal 9 (Industry, Innovation and Infrastructure) also features prominently in the Italian dataset, further reflecting the emphasis on technological and industrial modernization. In contrast, social and governance-related goals, such as Goal 5 (Gender Equality), Goal 16 (Peace, Justice and Strong Institutions), and Goal 17 (Partnerships for the Goals), are consistently underrepresented in both contexts. This imbalance suggests that, while green finance instruments are effectively mobilized towards environmental objectives, there remains a significant gap in the integration of social equity and institutional resilience dimensions. The remaining goals, appearing with intermediate frequency (typically

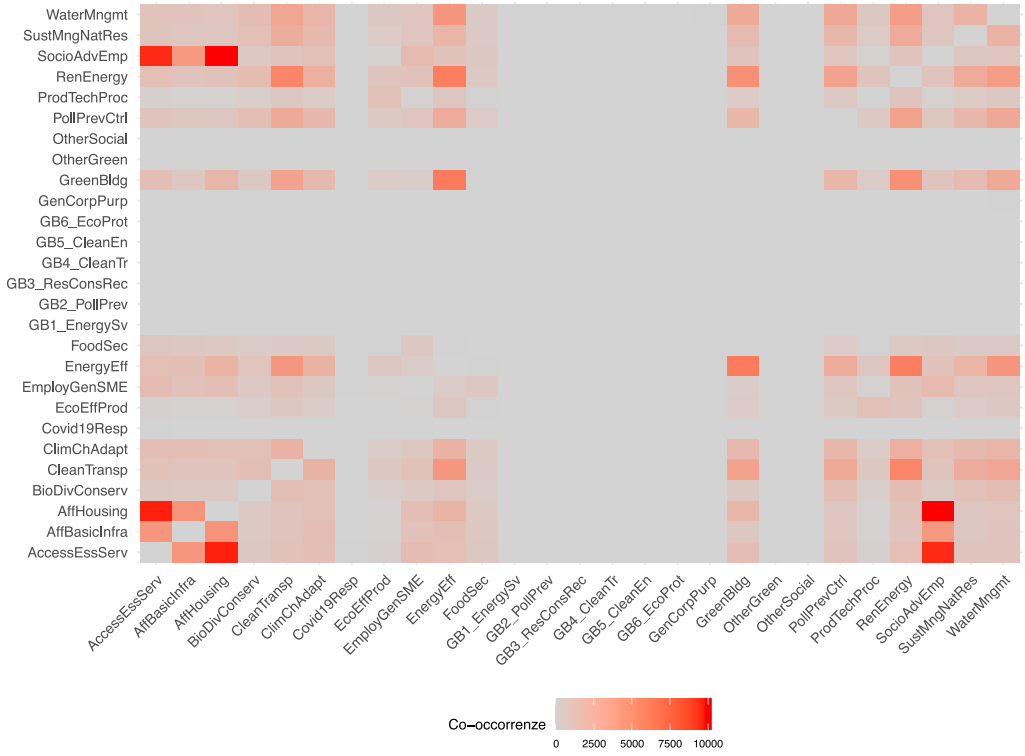
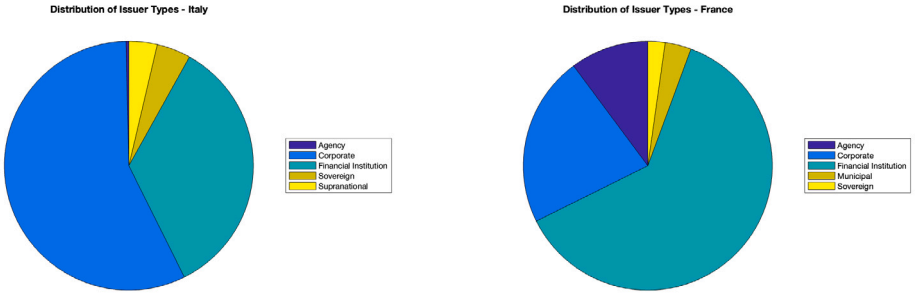


Fig. 5. Heatmap of co-occurrences between different use of proceeds (UoP) categories. The color scale ranges from light gray (low co-occurrence frequency) to red (high frequency), indicating how often pairs of categories appear together in the analyzed issuances. Axes represent abbreviated UoP categories for better readability. The legend is positioned at the bottom to aid interpretation of co-occurrence density.



(a) Italy (b) France

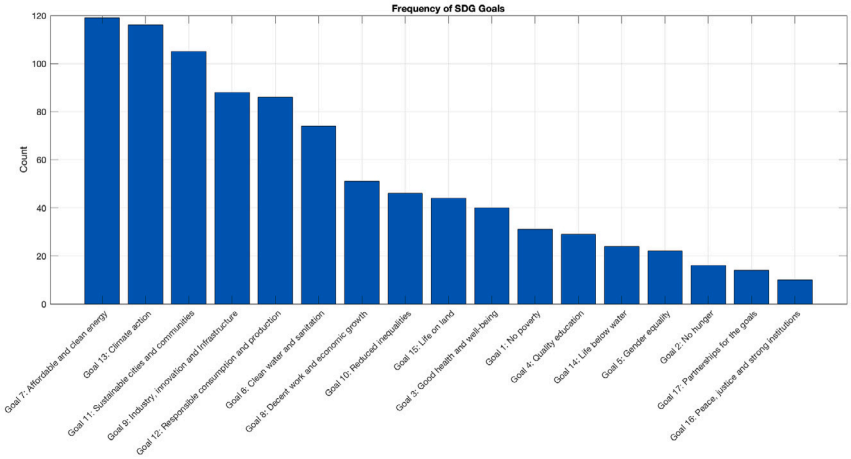
Fig. 6. Distribution of issuer types in the Italian (a) and French (b) datasets. This comparison highlights differences in the institutional composition of the GSS+ markets in the two countries.

between 100 and 200 occurrences), indicate areas of emerging but still less consolidated attention within the broader sustainable development agenda. It is relevant to note that this distribution is different from that of the entire dataset (see Fig. 4), where on the contrary Goal 11 is the one most present followed, in addition to Goal 7, by Goals 1, 10 and 8.

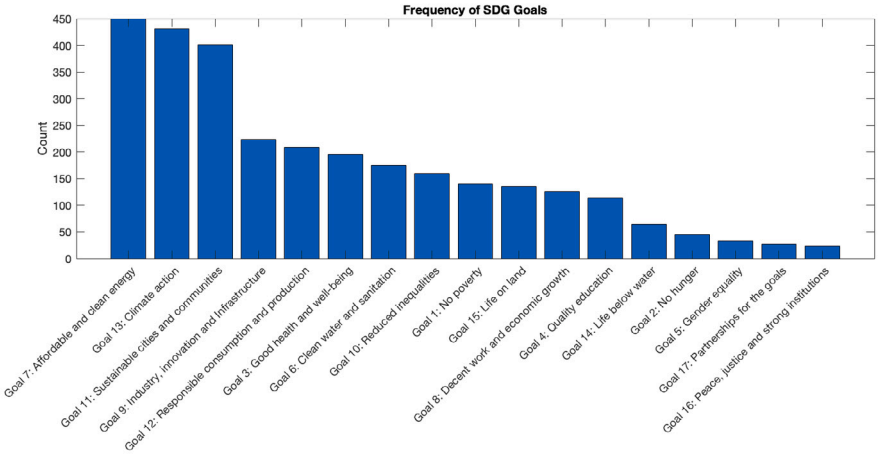
A deeper analysis of the SDGs co-occurrence network (Fig. 8) reveals, for the Italian case, a densely interconnected subset of goals, particularly Goals 6 (Clean Water and Sanitation), 7 (Affordable and Clean Energy), 11 (Sustainable Cities and Communities), 12 (Responsible Consumption and Production), 13 (Climate Action), and 15 (Life on Land). These SDGs frequently co-occur, suggesting that GSS+ bond frameworks often support integrated environmental transitions, including sustainable infrastructure, resource efficiency, and urban resilience. The observed network structure underscores the tendency of GSS+ bonds to address multi-goal synergies, particularly those related to environmental protection and infrastructure, while highlighting the relative under-representation of governance and institutional targets.

Differently, the SDGs co-occurrence network derived from the French subset reveals a highly interconnected network of goals, indicating an integrated approach in which GSS+ bond allocations address a broad spectrum of sustainability targets. Notably, Goals 6 (Water and Sanitation), 11 (Sustainable Cities), 13 (Climate Action), 3 (Health and Well-being), 8 (Decent Work), and 10 (Reduced Inequalities) emerge as central hubs, suggesting that projects financed by GSS+ bonds often integrate environmental sustainability with socio-economic inclusion. Unlike previous networks, no SDG is structurally isolated and even governance-related goals such as Goal 16 and Goal 17 exhibit non-negligible co-occurrence patterns.

It is important to note that bonds issued by financial institutions are not typically earmarked for a single project, but rather used to finance

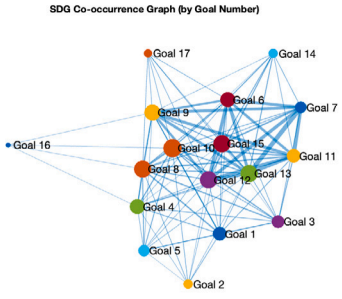


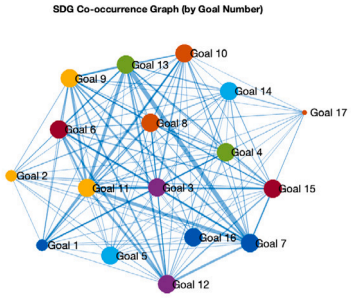
(a) Italy



(b) France

Fig. 7. Absolute frequency of Sustainable Development Goals (SDGs) in the Italian (a) and French (b) datasets. The plots illustrate the relative prominence of specific SDGs among the labeled bond issuances in the two countries.





(a) Italy (b) France

Fig. 8. Co-occurrence networks of SDGs for the Italian (a) and French (b) datasets. The co-occurrence matrix was filtered to discard frequencies lower than 10. Edge thickness is scaled proportionally to the joint frequency of SDG pairs.

Table 5
Use of proceeds categories with corresponding identifiers. This table reports the set of UoP categories identified across Italian and France bond emissions, each associated with a unique identifier. These identifiers are used throughout the analysis to standardize and visualize the categories within the co-occurrence networks of the two countries, as illustrated in Figs. 12 and 13.

ID	Use of proceeds	Italy	France
P01	Renewable energy	123 (14.49%)	410 (16.41%)
P02	Energy efficiency	103 (12.13%)	320 (12.81%)
P03	Clean transportation	86 (10.13%)	331 (13.25%)
P04	Sustainable water management	76 (8.95%)	141 (5.64%)
P05	Green Buildings	67 (7.89%)	285 (11.40%)
P06	Pollution prevention and control	64 (7.54%)	107 (4.28%)
P07	Sustainable management of living natural resources	50 (5.89%)	95 (3.80%)
P08	Socioeconomic advancement and empowerment	35 (4.12%)	108 (4.32%)
P09	Access to essential services	34 (4.00%)	161 (6.44%)
P10	Employment generation including through the potential effect of SME financing and microfinance	34 (4.00%)	63 (2.52%)
P11	Eco-efficient products	32 (3.77%)	84 (3.36%)
P12	Production technologies and processes	32 (3.77%)	84 (3.36%)
P13	Affordable housing	24 (2.83%)	93 (3.72%)
P14	Terrestrial and aquatic biodiversity conservation	24 (2.83%)	66 (2.64%)
P15	Affordable basic infrastructure	22 (2.59%)	44 (1.76%)
P16	Climate change adaptation	21 (2.47%)	64 (2.56%)
P17	Food security	10 (1.18%)	20 (0.80%)
P18	Covid-19 response	5 (0.59%)	10 (0.40%)
P19	GB5 — Clean energy	4 (0.47%)	0 (0%)
P20	GB3 — Resource Conservation and Recycling	1 (0.12%)	0 (0%)
P21	Other eligible green categories	1 (0.12%)	6 (0.24%)
P22	Other eligible social categories	1 (0.12%)	7 (0.28%)

a portfolio of initiatives involving multiple companies. As such, they can target a wider range of SDGs simultaneously. This characteristic helps explain the high degree of co-occurrence of SDGs observed in the French dataset, where financial institutions represent a significant share of issuers.

Finally, the distribution of UoP frequencies, when expressed as percentages, exhibits a broadly similar pattern across the compared groups, showing a clear prevalence of energy and transportation themes (see Table 5). On the contrary, Fig. 9 shows the co-occurrence networks (see, e.g., [42]) related to UoP characteristic for Italy and France. In these networks, each node corresponds to a specific UoP, and links indicate that two categories were included together in at least one bond issuance. The thickness and color intensity of the links reflect how frequently these co-occurrences occur: the more frequent, the thicker and darker the link. Node positions are determined by the Fruchterman–Reingold algorithm (see [43]), which arranges nodes based on connection strength—bringing more strongly linked categories closer together. The full list of UoP ID is provided in Table 5.

The co-occurrence network for Italy (Panel a) appears relatively sparse and modular, with the strongest connections observed between P01 (Renewable Energy), P02 (Energy Efficiency), P03 (Clean Transportation), P04 (Sustainable Water Management), and P05 (Green Buildings). Secondary linkages emerge with P11 (Eco-efficient Products) and P06 (Pollution Prevention and Control), suggesting a focused environmental agenda with some extensions towards circular economy themes. The overall structure reflects a tendency among Italian issuers to concentrate on a core set of environmental priorities with limited thematic diversification per bond. In contrast, the French network (Panel b) shows a more integrated topology, nevertheless having an evident co-occurrence of similar themes, with a clearly identifiable clique formed by P01, P02, P03 and P05 and a strong connection with P04. The presence of additional peripheral linkages indicates a broader thematic reach and a higher incidence of multi-objective bond structures, likely influenced by issuer diversity and evolving market standards. These differences underscore some peculiarities between patterns in the GSS+ bond design: while Italian issuers appear to adopt a more targeted and sector-specific approach, French issuers tend to structure green bonds around interconnected environmental themes, reflecting a more comprehensive sustainability strategy.

Table 6
Distribution of entities in each cluster for France and Italy.

Country	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
France	283	193	58	42	143
Italy	75	47	33	26	8

4.3. Cluster analysis and discussions

To implement the clustering procedure outlined in Section 3, we selected the following set of features: *labels*, *dollar value*, *use of proceeds*, *issuer sector*, *tenor*, and *coupon type*. These variables, which comprise numerical and categorical data, encapsulate the key attributes of each bond issuance.

In particular, labels, use of proceeds, and issuer sector describe the sustainability focus and institutional profile of the instrument, which are central to understanding how GSS+ bonds operationalize environmental and social objectives and align with the SDGs. Dollar value, tenor, and coupon type represent the key financial characteristics of the bond, reflecting issue size, investment horizon, and risk-sharing mechanisms that are standard in fixed-income analysis. Other variables available in the dataset were not included in the clustering step to avoid redundancy, excessive sparsity, or dominance of highly correlated features. Geographic information, in particular, was intentionally used ex post in the comparative analysis of Italy and France rather than as a clustering driver.

The application of the PAM algorithm, coupled with the optimal selection of the number of clusters based on the Average Silhouette Width (ASW), results in a five-cluster solution for both the Italian and French datasets. The size of each cluster is reported in Table 6.

The cluster analysis of the Italian subsample reveals the following key characteristics, as illustrated in Figs. 10 and 12 along with Tables 7 and 8.

Cluster 1 represents the largest group within the Italian sample, comprising 75 bond issuances. The predominant sectors include utilities, the energy sector (both renewable and non-renewable), and public transportation, outlining a typical profile of actors engaged in infrastructural and systemic projects. All bonds in this cluster are fixed-rate coupons, and the average tenor is approximately 8 years, reflecting an intermediate financial structure aligned with the time-windows of

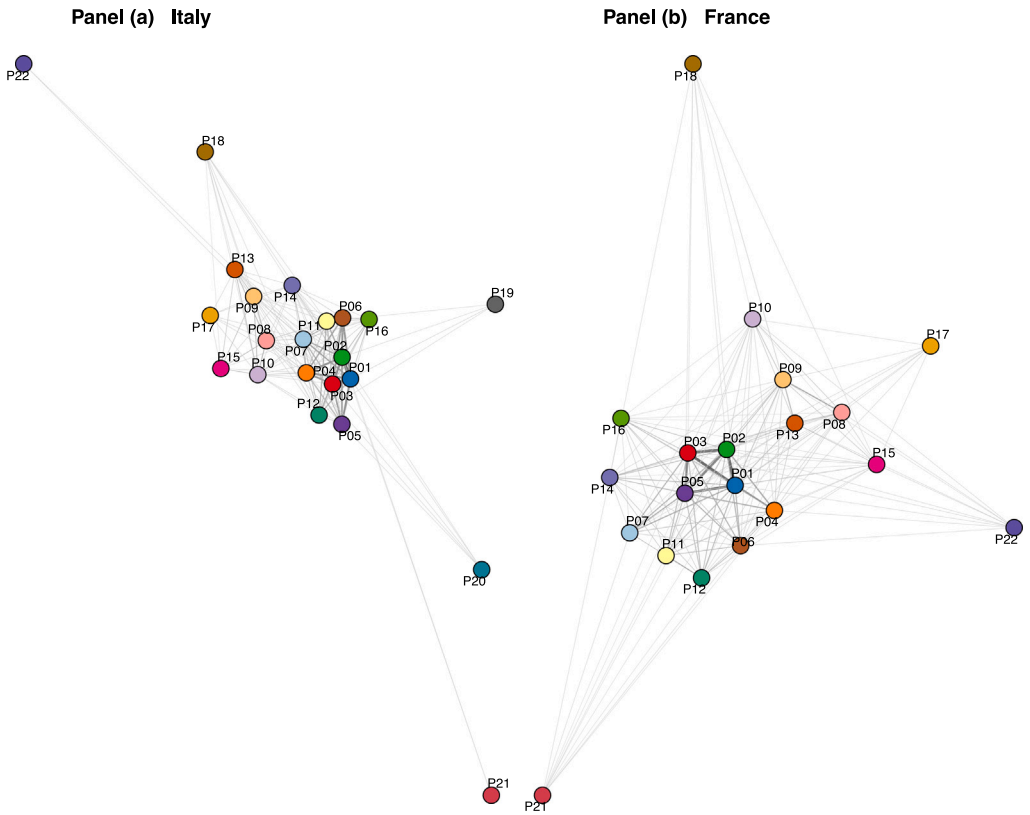


Fig. 9. Co-occurrence networks of use of proceeds categories for Italy and France. Panel (a) displays the network for Italian issuers, while Panel (b) refers to French issuers. Node labels correspond to use of proceeds categories as listed in Table 5. Links between nodes indicate that two specific categories appeared together in at least one bond issuance. Both the color and thickness of the edges reflect the strength of co-occurrence: darker and thicker edges denote a higher frequency of joint appearances, thus revealing stronger thematic connections between categories.

many sustainable infrastructure investments. The co-occurrence analysis of UoP, as illustrated by the network linking ID-codes P01, P02, P03, P04 and P06, reveals a systematic association between environmental and energy objectives. This proximity between nodes in the co-occurrence network indicates an integrated approach to the allocation of raised capital, where environmental initiatives are often implemented in synergy with energy-related projects.

Cluster 2 comprises 47 bond issues, all from the financial sector. In contrast to Cluster 1, these bonds are characterized by a higher average value. The average tenor is about 6 years, while the coupon type remains exclusively fixed-rate. The network of UoP co-occurrences reveals a closeness between IDs P01 to P07. This pattern reflects the high absolute frequencies of these categories in the Italian sample (see Table 5) and suggests a broad but thematically consistent focus on sustainability among financial issuers.

Cluster 3 consists exclusively of issuers from the financial sector (33 overall), like Cluster 2. However, it differs considerably in terms of label, structure and objectives. The medoid bond is a social bond, marking a shift from the environmental focus of the previous clusters towards more socially oriented objectives. The average duration is the shortest of all clusters, about five years, while the issue size is intermediate. Furthermore, this cluster contains the highest proportion of variable-rate bonds, indicating a greater inclination towards flexible or adaptive financial structures. The UoP co-occurrence network shows strong associations between P08, P09 and P10. Although the issuers belong to the financial sector, the selected UoP categories reveal a strategic allocation of capital towards social development objectives.

Cluster 4 comprises 26 issues: while the median value is in line with Cluster 1, the cluster's medoid is about three times larger, indicating a small number of high-value bonds. The average tenor is comparable to Cluster 1, suggesting a similar investment horizon. The label of this cluster is Sustainability Bond, which distinguishes it from the green and social-labeled bonds observed in the other clusters. The coupon structure is predominantly fixed, but variable-rate bonds are also present, reflecting a moderate degree of structural heterogeneity. The issuing sectors include agriculture and fisheries, utilities, renewable and non-renewable energy, as well as isolated cases of finance, government and telecommunications. This composition suggests a hybrid public–private profile, often associated with projects that intersect environmental and social dimensions. Moreover, the network of UoP co-occurrences shows strong associations between P04, P07–P10, P13–P15 and P17. This broad thematic spread is consistent with the integrated nature of sustainability bonds, which are designed to finance projects with both environmental and social impacts.

Finally, Cluster 5 comprises the smallest number of issues in the Italian sample (8 issues), but represents a distinct segment of the GSS+ market. The cluster is mainly composed of national government issuers and, to a lesser extent, public transport entities, highlighting the sovereign nature of these instruments. All bonds in this cluster are labeled as green, with long maturities and large issue sizes, in line with the characteristics of strategic environmental infrastructure projects. The coupon structure is predominantly fixed-rate, although a small presence of variable-rate bonds indicates some variability in the design of the instruments. The UoP co-occurrence network reveals a strong thematic focus on P01, P02, P03, P06 and P14. This set of objectives reflects a coherent environmental policy agenda, characteristic of sovereign issuers that aim to address multiple sustainability priorities within a unified framework.

The cluster analysis of the France subsample shows the following key characteristics, as illustrated in Figs. 11 and 13 along with Tables 7 and 9. Cluster 1 is the largest in the French sample and comprises 283 bond issues, with median tenor of 8 years. All bonds in this cluster

Table 7
Selection of central cluster issues. Each row shows the reference medoids of the 5 clusters. Italian and French cases.

Labels	Dollar Value	Use of Proceeds	Issuer Sector	Tenor	Coupon Type
Italian Case					
Green bond	359.87	Clean transportation, Energy efficiency, Green Buildings, Pollution prevention and control, Renewable energy	Aviation	8.17	fix
Green bond	821.19	Clean transportation, Eco-efficient products, Energy efficiency, Green Buildings, Pollution prevention and control, Renewable energy, Sustainable natural resources, Water management	Financial	6.00	fix
Social bond	590.57	Covid-19 response, Employment generation including SME financing and microfinance	Financial	5.00	fix
Sustainability bond	1221.51	Access to essential services, Covid-19 response, Eco-efficient products, Energy efficiency, Renewable energy	Telecommunications	8.00	fix
Green bond	9705.24	Clean transportation, Energy efficiency, Pollution prevention and control, Renewable energy, Sustainable natural resources, Biodiversity conservation	Government — national	13.44	fix
French Case					
Green bond	593.200	Clean transportation, Eco-efficient products, Energy efficiency, Green Buildings, Renewable energy, Terrestrial and aquatic biodiversity conservation	Aviation; Real Estate — construction and construction materials	8.02	fix
Green bond	20.000	Clean transportation, Eco-efficient products, Green Buildings, Pollution prevention and control, Sustainable water management	Financial	7.00	fix
Sustainability bond	586.400	Access to essential services, Energy efficiency, Green Buildings	Healthcare; Other	7.00	fix
Green bond	16.074	Clean transportation, Eco-efficient products, Energy efficiency, Pollution prevention and control, Renewable energy, Sustainable management of living natural resources, Sustainable water management, Terrestrial and aquatic biodiversity conservation	Financial	4.04	float
Social bond	817.425	Access to essential services, Other eligible social categories, Socioeconomic advancement and empowerment	Financial	6.50	fix

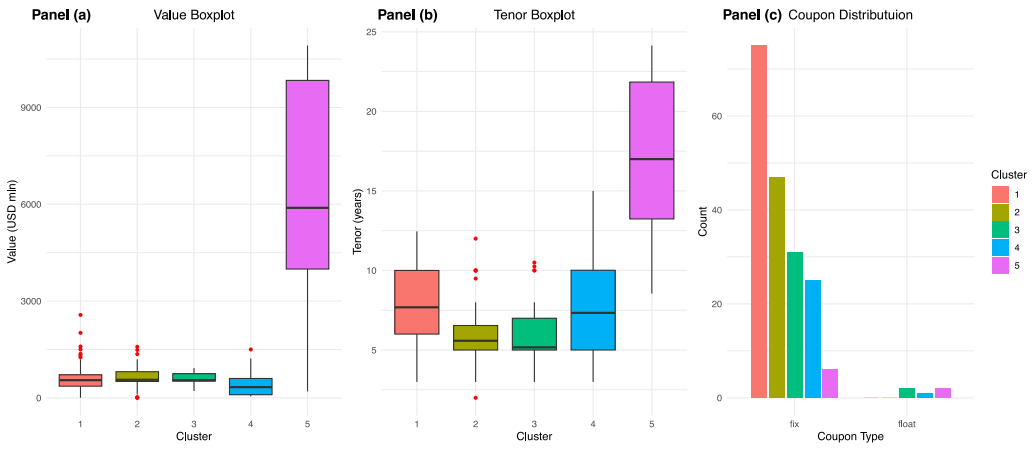


Fig. 10. Italian cluster characteristics. Panel (a) and Panel (b) show boxplots illustrating the distribution of numeric variables across clusters. Specifically, Panel (a) displays the distribution of dollar values (expressed in dollars), while Panel (b) presents the distribution of bond tenors (expressed in years). Panel (c) shows the distribution of coupon type (fixed and floating rate) across clusters. Each color corresponds to a different cluster.

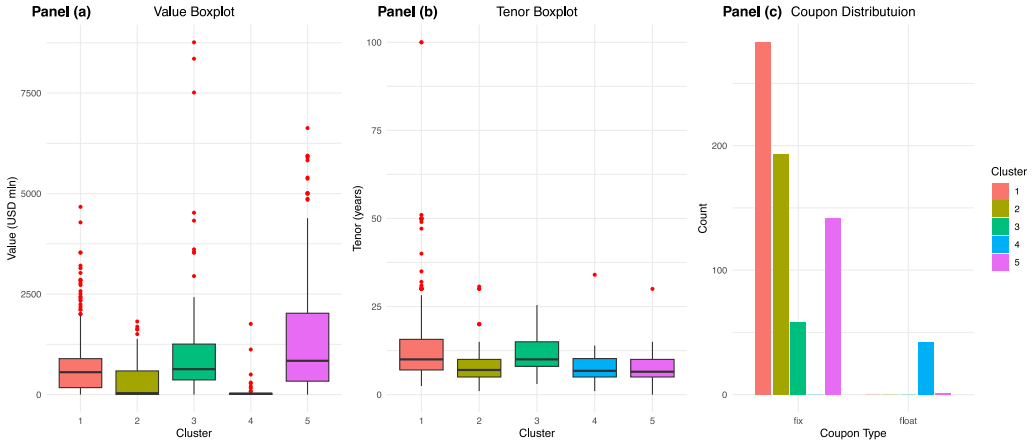


Fig. 11. France cluster characteristics. Panel (a) and Panel (b) show boxplots illustrating the distribution of numeric variables across clusters. Panel (c) shows the distribution of coupon type across clusters. For details see Fig. 10.

have fixed-rate coupons, indicating a standardized, risk-averse financial structure. The sectors of the issuers are very diverse: public transport, energy (both renewable and non-renewable), various real estate segments, government agencies (national and municipal), utilities and manufacturing. This heterogeneity indicates a wide adoption of GSS+ bonds in both the public and private sectors. The network of UoP co-occurrences reveals strong associations among these IDs: P01–P06, P11–P12, P14 and P16. This scheme reflects a comprehensive and multidimensional environmental agenda, in line with the complexity of the projects funded in this cluster.

Cluster 2 includes 193 bond issues, making it the second largest cluster in the French sample. Similar to Cluster 1, the medoid bond is marked with the green label, confirming the prominence of environmental issues in this clusters. This cluster is characterized by a significantly lower average issue size, centered around 20 mln, and a moderate tenor of around 7 years. All bonds are characterized by fixed-rate coupons, indicating a standard financial structure despite their small size. Cluster 2 is composed exclusively of financial sector issuers, suggesting a specialized green finance segment driven by banks and insurance companies. The network of UoP co-occurrences shows a close alignment around the main environmental categories, P01–P05. This pattern indicates a targeted application of green finance instruments towards standardized and well-recognized sustainability goals.

Cluster 3 comprises 58 bond issues. Unlike the previous clusters, where green labels are predominant, the cluster 3 medoid is a sustainability bond, indicating a more balanced focus on environmental and social objectives. All bonds are characterized by fixed-rate coupons and the average duration is broadly in line with Clusters 1 and 2. Cluster 3 shows greater sectoral diversity, including municipal governments, housing associations, healthcare, catering, telecommunications. This reflects a market segment where public or semi-public actors mobilize capital for integrated sustainability programs. The network of UoPs shows a broad thematic spread, with strong associations involving both environmental objectives, P01, P02, P03, P05, P14, P15, P16 and social objectives such as P09 and P10.

Cluster 4 is the smallest of the five clusters, with 42 issues. It is characterized by a smaller average issue size than the other clusters and a high prevalence of variable-rate bonds. Moreover, this cluster has the shortest tenor, which is another distinguishing feature. From a sectoral perspective, the cluster is strongly dominated by the financial sector, which accounts for 38 of the 42 issues. Despite its small size and the prevalence of low-value, and floating-rate bonds, Cluster 4 UoP co-occurrence network is divided into two separated components: one mainly social and the other more green/environmental. The first component comprises emissions related to social objectives such as P08, P09, P10, P13, P15. The second component groups emissions with UoP that are more aligned with green or environmental sustainability objectives, consistent with sectors such as renewable and non-renewable energy, utilities, and waste management.

Finally, Cluster 5 has a medium size, with 143 issuances, and its medoid is classified as a social bond. It has the highest average issue

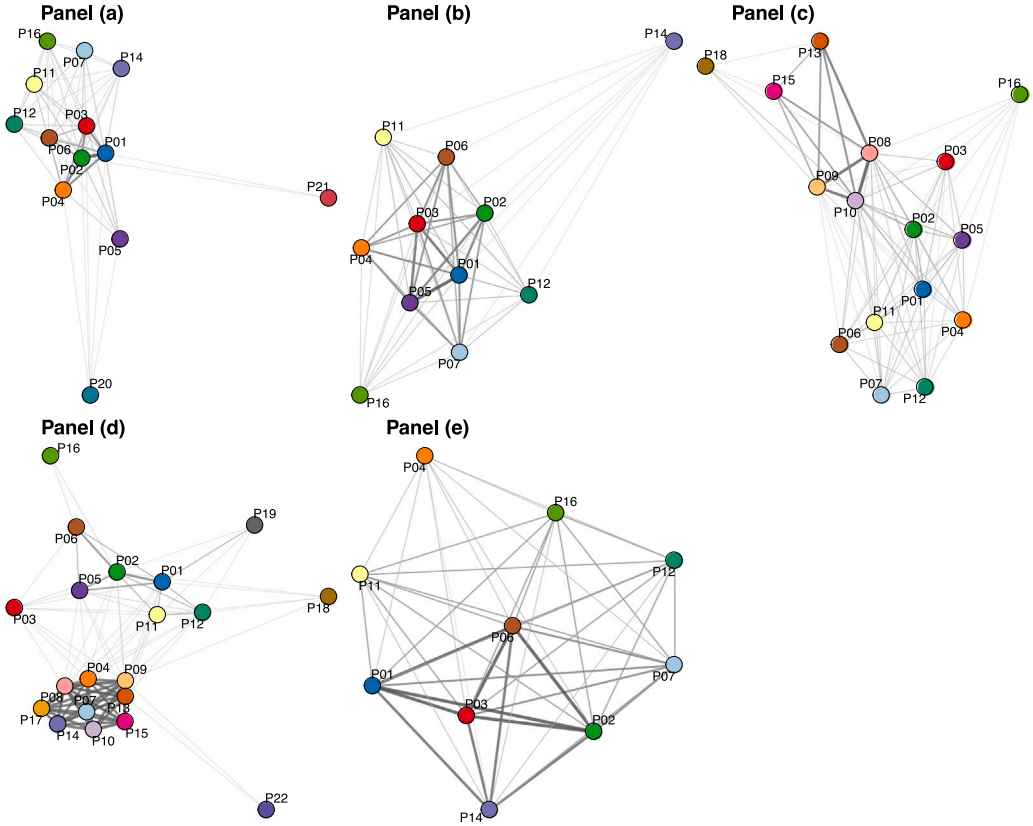


Fig. 12. Co-occurrence networks of use of proceeds among Italian issuers by cluster. Panel (a) displays the co-occurrence network for issuers in Cluster 1, while Panels (b)–(e) correspond respectively to Clusters 2–5. Node labels refer to use of proceeds categories as defined in Table 5. Links between nodes represent instances in which two specific use of proceeds categories appear together in the same bond issuance. Both the color and thickness of the links indicate the strength of the co-occurrence: darker and thicker edges correspond to higher co-occurrence frequency, thus highlighting stronger associations between categories.

size among the clusters, while its tenor distribution closely recalls that of Clusters 2 and 4. Only a small fraction of the issues are variable-rate bonds. The issue sector is exclusively financial, and the UoP co-occurrence network is highly connected, with the exception of UoP categories P14 and P18.

SDGs alignment. To enhance the policy relevance of the clustering results and ensure conceptual consistency with the UN 2030 Agenda, we align the characteristics of the cluster medoids reported in Table 7 with the SDGs. This mapping is based on the official SDG definitions and targets and relies on the dominant UoP categories associated with each medoid.

In the Italian sample, the SDG alignment reveals a relatively specialized and goal-focused structure, with each cluster primarily contributing to a limited and well-defined set of SDGs. Cluster 1 is aligned with SDGs 7, 11, and 13, reflecting the infrastructural and decarbonization-oriented nature of the financed projects, with secondary spillovers to SDG 9 through investments in sustainable and resilient infrastructure. Similarly, Cluster 2 is primarily associated with SDGs 7, 11, and 13, while its focus on eco-efficient products and water management also supports SDGs 6 and 12. Cluster 3 is dominated by social objectives, including Covid-19 response measures, employment generation, and SME financing. These activities align most directly with SDGs 3 and 8, in line with UN targets on health resilience, employment, and economic inclusion, and also contribute to SDGs 1 and 10 through their emphasis on vulnerable groups and access to financial resources. Cluster 4 combines social infrastructure with clean energy deployment. Access to essential services and pandemic-related interventions primarily support health and education outcomes (SDGs 3 and 4), while investments in renewable energy and energy efficiency advance SDG 7. Equity and sustainability dimensions further link this cluster to SDGs 10, 12, and 13. Finally, Cluster 5 focuses on large-scale environmental infrastructure, biodiversity conservation, and clean transportation. These activities are most closely associated with SDGs 13, 14, and 15, with complementary contributions to SDG 9 through infrastructure development.

By contrast, the French sample exhibits a more integrated and multi-goal SDG alignment, consistent with a more mature and institutionally diversified sustainable finance ecosystem. Cluster 1 emphasizes clean transportation, energy efficiency, green buildings, and biodiversity conservation, aligning primarily with SDGs 9 and 11, while also supporting SDGs 12, 14, and 15. Cluster 2, which focuses on green buildings, pollution prevention, and water management, contributes mainly to SDGs 6, 11, and 12, with additional support for SDG 9. Cluster 3 combines access to essential services with energy efficiency, thereby linking SDG 3 through strengthened health and social services with SDG 7 via energy-related improvements; its urban and infrastructural dimension also connects it to SDG 11. Cluster 4 displays the highest thematic diversification, with medoids covering a wide range of environmental UoPs associated with circular economy practices. It aligns most strongly with SDG 12, which explicitly promotes resource efficiency and waste reduction, while also contributing to SDGs 13, 14, and 15 through climate- and ecosystem-related investments. Finally, Cluster 5 is dominated by social objectives, including socioeconomic empowerment and access to essential services, corresponding primarily to SDGs 1, 8, and 10, with additional contributions to SDGs 3 and 4.

Overall, the French GSS+ bond universe exhibits a higher degree of thematic integration, with individual instruments often designed to address multiple, overlapping SDGs. This pattern reflects a more advanced and systemic sustainable finance architecture compared to the more specialized cluster structure observed in the Italian market.

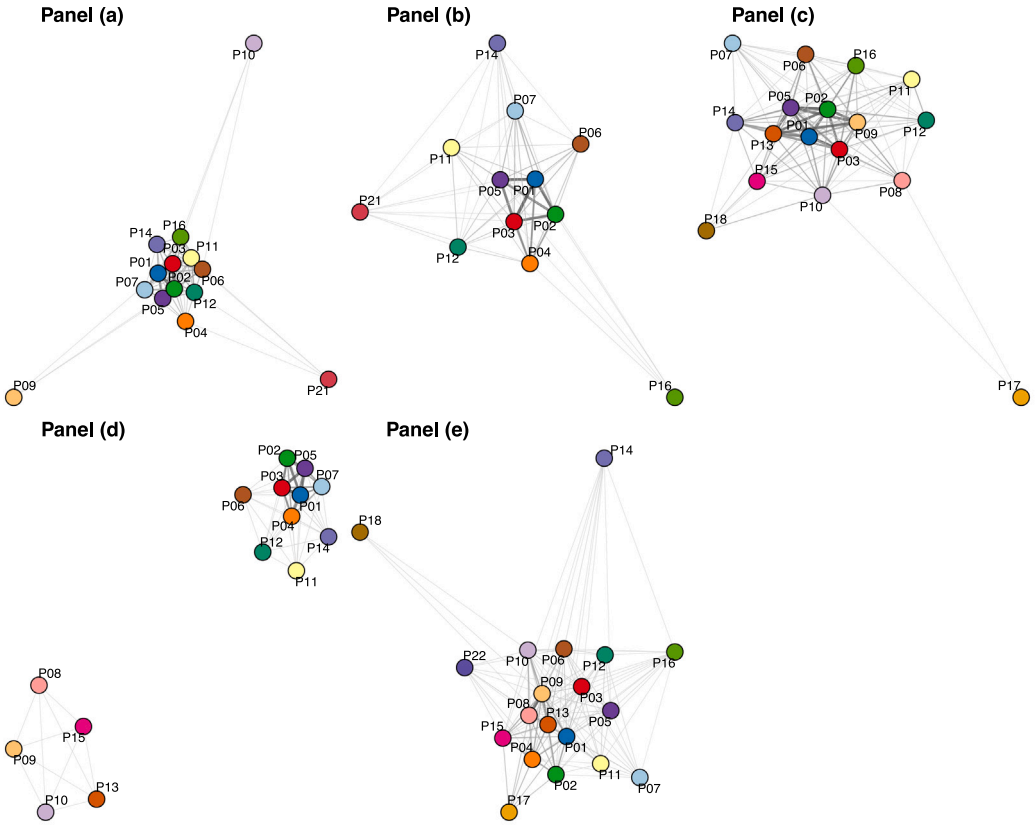


Fig. 13. Co-occurrence networks of use of proceeds among French issuers by cluster. Panels (a)–(e) represent Clusters 1–5, respectively. The structure and visual encoding of the networks follow the same criteria described in Fig. 12.

Table 8
Issuer sectors by cluster (Italy). This table reports the distribution of issuer sectors across clusters for the Italian sample. Each row corresponds to a specific sector (or combination of sectors) and indicates the number of issuers *n* within the corresponding cluster.

Cluster	Issuer sector	n
1	Utilities	15
1	Energy — non-renewable; Energy — renewable: other; Utilities; Water	9
1	Financial — insurance	8
1	Energy — renewable: other; Energy — renewable: solar; Energy — renewable: wind; Utilities	7
1	Government — national	6
1	Public transportation	6
1	Energy — non-renewable; Energy — renewable:other; Energy — renewable: solar; Energy — renewable: wind	5
1	Energy — renewable: solar	4
1	Telecommunications; Utilities; Water	4
1	Energy — non-renewable; Energy — renewable: other; Energy — renewable: solar; Energy — renewable: wind; Utilities	3
1	Energy — renewable: other; Utilities	3
1	Automotive	2
1	Aviation	1
1	Energy — renewable: other; Energy — renewable: solar; Energy — renewable: wind; Industrial machinery and engineering	1
1	Energy — renewable: wind	1
2	Financial	47
3	Financial	33
4	Agriculture/Fisheries	10
4	Utilities	8
4	Energy — non-renewable; Energy — renewable: other; Energy — renewable: solar; Energy — renewable: wind; Utilities	4
4	Financial	1
4	Financial — insurance	1
4	Government — national	1
4	Telecommunications	1
5	Government — national	6
5	Public transportation	2

Table 9
Issuer sectors by cluster (France). See Table 8 for further details.

Cluster	Issuer sector	n
1	Public transportation	82
1	Energy — non-renewable; Energy — renewable: other; Energy — renewable: solar; Energy — renewable: wind; Utilities	49
1	Real Estate — REIT	36
1	Government — national	30
1	Government — municipal/local	22
1	Real Estate — development and management	17
1	Energy — renewable: other; Energy — renewable: solar; Energy — renewable: wind	9
1	Utilities; Waste/Recycling; Water	7
1	Waste/Recycling	5
1	Other	4
1	Automotive	3
1	Industrial machinery and engineering; Manufacturing — other	3
1	Utilities	3
1	Energy — renewable: solar; Energy — renewable: wind	2
1	Financial — insurance	2
1	Real Estate —construction and construction materials	2
1	Aviation; Real Estate — construction and construction materials	1
1	Chemicals, plastics, and rubber	1
1	Real Estate — development and management; Real Estate — REIT	1
1	Real Estate — development and management; Retail	1
1	Technology	1
1	Telecommunications	1
1	Water	1
2	Financial	193
3	Government — municipal/local	17
3	Other	15
3	Housing association; Other	8
3	Food/beverage	4
3	Government — national	3
3	Healthcare	3
3	Energy — renewable: solar	2
3	Financial — insurance	2
3	Telecommunications	2
3	Automotive	1
3	Healthcare; Other	1
4	Financial	38
4	Government — municipal/local	2
4	Energy — non-renewable; Energy — renewable: other; Energy — renewable: solar; Energy — renewable: wind; Utilities	1
4	Waste/Recycling	1
5	Financial	143

4.4. Clustering robustness analysis

To assess the robustness of the clustering results, we conducted a sensitivity analysis using the Rand Index described in Section 3.1.1. We recognize that the Rand Index may be susceptible to sample imbalance or variations in cluster size, as larger clusters could disproportionately influence the stability metric. However, in our study, although cluster sizes vary reflecting the inherent heterogeneity of the Italian and French markets, the consistently high values of the Rand Index suggest that even smaller, more specific clusters maintain their structural integrity under perturbation. In fact, the Gower metric permits one to assign different weights to each variable included in the clustering process. This allows us to test whether variations in the importance attributed to individual features significantly affect the resulting cluster partitions. We compare a baseline scenario where all variables are equally weighted with alternative configurations in which each variable, in turn, receives double weight relative to the others. This systematic doubling of weights serves as a stress test to ensure that no single feature among the six considered variables dominates the clustering structure, providing a flexible yet rigorous sensitivity analysis for mixed data types.

This procedure is applied to both the Italian and French datasets. As illustrated in Fig. 14, the Rand Index values for all alternative configurations remain consistently high, generally at or above 0.7 for both countries. This indicates a strong alignment between the baseline and perturbed cluster structures, suggesting that the clustering outcomes are overall stable with respect to changes in variable weights. However, we notice that the largest deviations occur when double weight is assigned to the *label* feature, implying that this attribute has the greatest influence on the partitioning structure. In contrast, features such as *issue amount* and *coupon type* appear to have relatively limited impact on the clustering results. These findings confirm the robustness of the methodological approach and provide additional insight into the relative importance of the variables considered.

5. Conclusions

This study contributes to the expanding field of sustainable finance by offering a thorough analysis of the global GSS+ bond market. We presented a methodological framework that combines financial and non-financial bond characteristics using the Gower distance and PAM clustering algorithm, utilizing a dataset provided by Environmental Finance Data.

In particular, we investigated the internal structure of the GSS+ bond market and its cross-country heterogeneity by addressing two main research questions. With reference to **RQ1**, which asked whether latent structures emerge when financial and non-financial bond characteristics are jointly considered, our results provide clear evidence that the GSS+ universe is far from homogeneous. The clustering analysis reveals well-defined market segments characterized by distinct

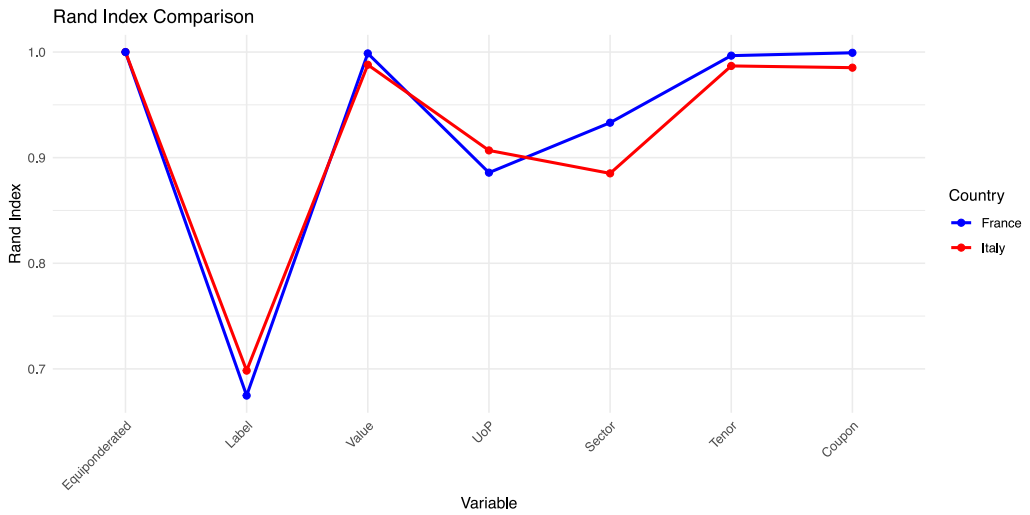


Fig. 14. Rand Index comparison for Italy and France across different Gower distance weighting scenario. "Equiponderated" refers to equal weights across all variables, while each of the other labels indicates that the corresponding variable was given double weight relative to the others. The Rand Index measures the similarity between each configuration and the equiponderated baseline.

combinations of use of proceeds, issuer profiles, financial features, and sustainability orientations. Importantly, these latent structures are meaningfully aligned with specific SDGs, confirming that the interaction between financial design and sustainability intent plays a crucial role in shaping the internal architecture of the sustainable fixed-income market.

Regarding **RQ2**, which focused on the role of institutional context and market maturity, the comparative analysis of Italy and France reveals substantial differences in the composition and integration of GSS+ clusters. The Italian market appears more specialized, with clusters that tend to concentrate on a narrower set of SDGs through clearly delineated instruments. On the other hand, the French market exhibits a higher degree of thematic integration, with clusters characterized by broader use-of-proceeds profiles and bonds designed to simultaneously address multiple sustainability objectives. These differences suggest that institutional frameworks, policy environments, and market development stages significantly influence how sustainable finance instruments are structured and deployed.

Theoretical implications. This study adds to the body of knowledge on sustainable finance in a number of ways. In particular, with regard to analysis of the market structure, it provides a multifaceted segmentation of the GSS+ bonds by combining quantitative and qualitative characteristics, going beyond conventional volume-based or issuer type classifications. We also show how unsupervised learning techniques allow for a soft assignment of observations, which can guide the interpretation of group structures within the data and reveal latent structures by using a Gower–PAM clustering approach. Furthermore, our study advances knowledge of how the SDGs are operationalized in various national contexts through sustainable finance instruments. Beyond individual bond characteristics, the co-occurrence analysis of SDGs reveals thematic synergies and underlying strategic priorities.

Finally, this study contributes methodologically by introducing a reproducible and scalable clustering framework tailored to mixed financial and non-financial data. Unlike standard descriptive or single-label approaches, the proposed methodology enables latent market structures to emerge endogenously, offering a flexible analytical tool that can be applied across countries and over time to support evidence-based policy design and market monitoring.

Practical implications. From a practical perspective, the findings offer relevant implications for policymakers, issuers, and investors involved in the GSS+ bond market. The observed differences between Italy and France highlight how institutional settings and market maturity shape the internal structure of sustainable finance. In particular, the more integrated and diversified clustering structure observed in France suggests that well-established regulatory frameworks and the prominent role of financial institutions can foster multi-objective bond designs and stronger thematic integration. Conversely, the more concentrated patterns identified in Italy point to the need for policy measures aimed at broadening issuer participation and encouraging greater diversification in the allocation of use of proceeds and SDG targeting. For issuers and investors, the clustering results provide a data-driven tool to benchmark bond characteristics against homogeneous market segments. Issuers can use these insights to design instruments that align more effectively with prevailing sustainability practices or to strategically differentiate new issuances, while investors can exploit cluster information to improve portfolio diversification, risk assessment, and impact-oriented allocation strategies.

Limitations and future research. This study has some limitations. The pre-selected variables that form the basis of the clustering method may not encompass all pertinent aspects of sustainability performance. Integrating additional datasets, such as issuer-level Environmental, Social, and Governance (ESG) scores, could provide a more comprehensive understanding of GSS+ bonds by linking instrument characteristics to the sustainability profiles of their issuers [44].

Furthermore, only two countries are included in the comparative case study. Extending the analysis to other European or non-European contexts would improve the findings' generalizability.

Future studies could evaluate the impact performance of various types of bonds, investigate temporal dynamics in clustering behavior, or incorporate textual data from bond prospectuses to enhance the feature set. In addition, linking clustering results with market returns and investor behavior would benefit scholars and practitioners alike.

CRediT authorship contribution statement

Giovanna Ferraro: Writing – original draft, Visualization, Validation, Supervision, Software, Methodology, Data curation, Conceptualization. **Alessandro Ramponi:** Writing – original draft, Visualization, Validation, Supervision, Software, Methodology, Data curation, Conceptualization. **Saverio Storani:** Writing – original draft, Visualization, Validation, Supervision, Software, Methodology, Data curation, Conceptualization.

Acknowledgments

The author A. Ramponi acknowledges financial support from the Italian Ministry of University and Research (MUR) through the PRIN 2022 PNRR Project "The Resilience of Sustainable Finance" (Grant No. P2022ENNYP – CUP: B53D23026490001), co-funded by the European Union–Next GenerationEU. All authors reviewed and approved the final version.

Data availability

The authors do not have permission to share data.

References

[1] Soergel B, Kriegler E, Weindl I, Rauner S, Dirnaichner A, Ruhe C, et al. A sustainable development pathway for climate action within the UN 2030 Agenda. Nat Clim Chang 2021;11(8):656–64.
[2] Huynh TLD, Hille E, Nasir MA. Diversification in the age of the 4th industrial revolution: the role of artificial intelligence, green bonds and cryptocurrencies. Technol Forecast Soc Chang 2020;159:120188.
[3] Carnevali E, Deleidi M, Pariboni R, Passarella MV. Cross-border financial flows and global warming in a two-area ecological SFC model. Socio-Economic Plan Sci 2021;75:100819.
[4] Yuan X, Zhang J, Shi J, Wang J. What can green finance do for high-quality agricultural development? Fresh insights from China. Socio-Economic Plan Sci 2024;94:101920.
[5] Dutta PK, El-kenawy SM, Abotaleb M, Eid MM. AI-driven marketplaces and price prediction tools for rag pickers: Enhancing economic opportunities in Africa's circular economy. Babylon J Artif Intell 2023;2023:33–42.
[6] Maltais A, Nykvist B. Understanding the role of green bonds in advancing sustainability. J Sustain Financ Invest 2020;1–20.
[7] Flammer C. Corporate green bonds. J Financ Econ 2021;142:499–516.
[8] Yumei T, Xihui HC, Provash KS, Sarra B. Asymmetric effects of geopolitical risks and uncertainties on green bond markets. Technol Forecast Soc Chang 2023;189:122348.
[9] Changsong L, Xiaojing C, Zeyu W, Jiali Z, L. Huan. The impact of green bond issuance on corporate green innovation: A signaling perspective. Int Rev Financ Anal 2025;102:104113.
[10] Shah SS, Nakouwo SN, Sobirjonovna GM, Khan A. Exploring the sustainability impact of green bonds on ecological and resource capacities. Renew Energy 2025;243:122590.
[11] Rongnan L, Qi Z, Weijun X. Do green bonds have an advantage in information disclosure? Int Rev Financ Anal 2025;101:104008.
[12] Agliardi E, Agliardi R. Financing environmentally-sustainable projects with green bonds. Environ Dev Econ 2019;24(6):608–23.
[13] Rasoulinezhad E, Taghizadeh-Hesary F. Role of green finance in improving energy efficiency and renewable energy development. Energy Effic 2022;15(2):14.
[14] Bartolini N, Romagnoli S, Santini A. Understanding climate risk in Europe: Are transition and physical risk priced in equity and fixed-income markets? J Empir Financ 2025;84:101672.
[15] Jiang C, Li W, Wang Y, Ding R. A complex network analysis of risk spillover and linkage effects between green bonds and commodities under the characteristics of the Chinese market. Int Rev Econ Financ 2026;10491.
[16] Bhutta US, Tariq A, Farrukh M, Raza A, Iqbal MK. Green bonds for sustainable development: Review of literature on development and impact of green bonds. Technol Forecast Soc Chang 2022;175:121378.
[17] Kaufman L, Rousseeuw PJ. Finding groups in data: An introduction to cluster analysis. John Wiley & Sons; 2009.
[18] Gower JC. A general coefficient of similarity and some of its properties. Biometrics 1971;857–71.
[19] Rand WM. Objective criteria for the evaluation of clustering methods. J Amer Statist Assoc 1971;66(336):846–50.
[20] Cerqueti R, Deffains-Crapsky C, Storani S. Green finance instruments: Exploring minibonds issuance in Italy. Corp Soc Responsib Environ Manag 2023;30(4):1965–86.
[21] Khairi FA, Farhan L, Adelaja OA. Data mining driven segmentation of health insurance policyholders using k-means clustering. Mesopotamian J Artif Intell Heal 2025;2025:187–96.
[22] Jeyapaulraj J, Desai D, Mehta D, Chu P, Pasquali S, Sommer P. Supervised similarity learning for corporate bonds using random forest proximities. In: Proceedings of the third ACM international conference on AI in finance. 2019.
[23] Breiman L. Random forests. Mach Learn 2001;45(1):5–32.
[24] Englund C, Verikas A. A novel approach to estimate proximity in a random forest: An exploratory study. Expert Syst Appl 2012;39(17):13046–50.
[25] Ying Y, Li P. Distance metric learning with eigenvalue optimization. J Mach Learn Res 2012;13(1):1–26.
[26] Desai D, Mehta D. On robustness of mutual funds categorization and distance metric learning. J Financ Data Sci 2021;3(4):130–50.
[27] Chavent M, Kuentz V, Labenne A, Saracco J. Multivariate analysis of mixed data. the R package PCAmixdata. Electron J Appl Stat Anal 2022;15(3):606–45.
[28] Ferrier S, Manion G, Elith J, Richardson K. Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. Divers Distrib 2007;13(3):252–64.
[29] Huang Z. Extensions to the k-means algorithm for clustering large data sets with categorical values. Data Min Knowl Discov 1998;2(3):283–304.
[30] Abdalkareem ZA, Amir A. Mapping the intersection of artificial intelligence and the sustainable development goals: A bibliometric and scientometric analysis (2003–2025). ESTIDAMAA 2025;2025:1–12.
[31] Ester M, Kriegel HP, Sander J, Xu X. A density-based algorithm for discovering clusters in large spatial databases with noise. In: Proceedings of the second international conference on knowledge discovery and data mining. 1996.
[32] Ward Jr JH. Hierarchical grouping to optimize an objective function. J Amer Statist Assoc 1963;58(301):236–44.
[33] Ranalli M, Rocci R. A comparison between methods to cluster mixed-type data: Gaussian mixtures versus gower distance. In: Scientific meeting of the classification and data analysis group of the Italian statistical society. 2019.
[34] Foss AH, Markatou M, Ray B. Distance metrics and clustering methods for mixed-type data. Int Stat Rev 2019;87(1):80–109.
[35] Garcia-Escudero LA, Gordaliza A. Robustness properties of k means and trimmed k means. J Amer Statist Assoc 1999;94(447):956–69.
[36] MacQueen J. Some methods for classification and analysis of multivariate observations. In: Proceedings of the fifth Berkeley symposium on mathematical statistics and probability. 1967.
[37] Bezdek JC, Ehrlich R, Full W. FCM: The fuzzy c-means clustering algorithm. Comput Geosci 1984;10(2–3):191–203.
[38] Rdusseeun LKPJ, Kaufman P. Clustering by means of medoids. In: Proceedings of the statistical data analysis based on the l1 norm conference. 1987.
[39] Arbelaitz O, Gurrutxaga I, Muguerza J, Pérez JM, Perona I. An extensive comparative study of cluster validity indices. Pattern Recognit 2013;46(1):243–56.
[40] Batool F, Hennig C. Clustering with the average silhouette width. Comput Statist Data Anal 2021;158:107190.
[41] Cerqueti R, Ferraro G, Mattera R, Storani S. Mapping socio-environmental policy integration in the European union: A multilayer network approach. J Clean Prod 2025;144792.
[42] Storani S, Falkenberg M, Quattrociocchi W, Cinelli M. Relative engagement with sources of climate misinformation is growing across social media platforms. Sci Rep 2025;15:18629.
[43] Fruchterman TM, Reingold EM. Graph drawing by force-directed placement. Softw: Pr Exp 1991;21(11):1129–64.
[44] Landi GC, Turriziani L, Muto V, Ricciardi I. Exploring the nexus between ESG risk variations and investment preferences: Insights from sustainable ETFs during the COVID-19 era. Socio-Economic Plan Sci 2024;95:102039.

Giovanna Ferraro, Ph.D. is Assistant Professor in Management Engineering at the Department of Enterprise Engineering, University of Rome Tor Vergata. Her research interests are related to complex networks, Social Network Analysis, technology transfer, innovation management and eco-innovation. She is involved in several European projects regarding technology transfer and business innovation management. She is the author of research papers published in international journals, conference proceedings as well as chapters of books.

Alessandro Ramponi, Ph.D. is Associate Professor at the Department of Economics and Finance, University of Rome Tor Vergata. At this institution, he is actively involved in both the Ph.D. and Master of Science programs. He regularly participates in international conferences as a speaker and serves as a referee for a variety of journals. He is member of AMASES (Association for Mathematics Applied to Social and Economic Sciences) and INDAM-GNAMPA. His research interests are in the area of computational finance, stochastic models, risk assessment and management, Monte Carlo methods, neural networks and machine learning. He has published papers in international journals and conference proceedings.

Saverio Storani, Ph.D. is a postdoctoral researcher at the Department of Enterprise Engineering, University of Rome Tor Vergata. His research focuses on complex networks, econometrics, financial forecasting, and sustainability issues related to SMEs. He has published in internationally recognized journals. Currently, he is involved in a study concerning credit risk in SMEs with a focus on ESG factors.