Ecological Indicators 157 (2023) 111151

Contents lists available at ScienceDirect

Ecological Indicators

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

A novel approach to rating SMEs' environmental performance: Bridging the ESG gap

Seben Ozkan [a], Silvia Romagnoli [b,*], Pietro Rossi [a,b]

[a] Prometeia, Piazza Trento e Trieste 3, 40137 Bologna, Italy
[b] University of Bologna, Department of Statistics, Via Belle Arti 41, 40126 Bologna, Italy

ARTICLE INFO

Keywords:
Satellite observations
ESG rating
SME
Neural Network

ABSTRACT

Given the increasing significance of sustainability in investment decisions and regulatory frameworks, Environmental Social and Governance (ESG) ratings for companies are becoming increasingly relevant in the decision-making processes of stakeholders. While large listed companies are mandated to disclose ESG information, the same cannot be said for Small and Medium Enterprises (SMEs). SMEs are not obligated to provide either sustainability information or their own ESG ratings, leaving them susceptible to potential disadvantages in securing capital and attracting investments. Moreover, ESG rating agencies source all the necessary data from the very companies they are meant to assess, leading to an evident conflict of interest.

In this paper, we propose a comprehensive solution to urgently address this gap in ESG disclosure. Leveraging the unique capabilities of Neural Networks (NN) to comprehend and replicate intricate patterns, we train a NN using available environmental and rating data from large companies. The NN learns how to replicate ratings based on the available information. Once the network is adequately trained, we employ it to generate ratings for SMEs that would otherwise lack any form of rating. Another point of innovation is represented by the type of data used, i.e. we utilize data acquired through satellite observations within the European Union (EU) Copernicus Program, ensuring an impartial means of gathering information on environmental activities. Our NN is fed with satellite observations, with the target being the ratings recognized by supervisory agencies. Once the network has been satisfactorily trained and can accurately reproduce the target set of ratings, it is directly applied to the same dataset for a group of SME companies. In doing so, we establish a methodology for consistently rating SMEs' environmental performance in alignment with the methodology used for larger companies.

1. Introduction

Since the adoption of the 2015 Paris Agreement, the Sustainable Development Goals (SDGs) outlined in the UN's 2030 agenda, and the European Green Deal in 2019, the concept of sustainable finance has garnered increasing attention and importance. Sustainable finance pertains to any financial activity that takes into consideration the environmental (E), social (S), and governance (G) aspects or impacts of that activity (EC, 2021). Within this framework, the "E" encompasses all activities related to environmental concerns such as greenhouse gas (GHG) emissions, water usage, waste and pollution management, and biodiversity preservation. "S" focuses on working conditions both within and outside the company, employee rights, and treatment in the workplace, while "G" encompasses management issues like board composition and executive decisions. In essence, the goal of sustainable finance is to promote investment decisions and company evaluations that align with targets for a low-carbon, green, and sustainable economy.

Within the realm of sustainable finance and its objectives, the concept of ESG (Environmental, Social, and Governance) rating has emerged to define and assess companies' positions and stances regarding sustainability goals. These goals are of significance to investors, the companies themselves, and society at large. As the demand for identifying and evaluating companies' sustainability, particularly their ESG performance, continues to grow, driven by both investment opportunities and, more importantly, compliance with European Commission (EC) regulations, numerous rating agencies have emerged.

The common approach adopted by these agencies involves considering the three overarching pillars: E (Environmental), S (Social), and G (Governance). Subsequently, they incorporate relevant positive or negative activities or initiatives undertaken by companies within each of

* Corresponding author.
E-mail addresses: seben.ozkan@prometeia.com (S. Ozkan), silvia.romagnoli@unibo.it (S. Romagnoli), pietro.rossi3@unibo.it (P. Rossi).

https://doi.org/10.1016/j.ecolind.2023.111151
Received 28 July 2023; Received in revised form 17 October 2023; Accepted 22 October 2023

these pillars (Escrig-Olmedo et al., 2019). Nevertheless, as pointed out in Lopez et al. (2007), there is no singular definition of sustainability, nor is there a universally standardized method for measuring it.

Efforts are underway to coordinate and align the understanding of sustainability. For instance, the European Union (EU) is striving to identify and categorize business activities that align with its objectives to achieve the targets outlined in the 2030 agenda and the European Green Deal. To achieve this, the EU has established the EU Taxonomy, which serves as a classification system defining and listing activities that can be deemed sustainable economic activities (EC, 2021). This enables stakeholders to understand the criteria that activities must meet to be in line with the taxonomy. Furthermore, stakeholders can assess how their activities contribute to six EU environmental objectives, with climate change mitigation and adaptation taking precedence among these objectives. Lastly, stakeholders can evaluate whether their activities adhere to the "do no harm" principle, meaning that while they align with certain objectives, they should not harm the remaining ones simultaneously.

While establishing such a framework aids in defining standardized and objective criteria and references for investment decisions and a sustainable economy, there are still areas of incompleteness and ongoing debates surrounding the taxonomy. First and foremost, the list of sustainable activities is limited to specific activities rather than entire sectors. Consequently, there is the possibility that a company's activities may not align with those listed in the taxonomy, as not all activities within every sector are currently covered. Despite the assertion that the taxonomy should not be the sole reference for financing decisions, it remains incomplete.

Secondly, the evolution and development of the Taxonomy is an ongoing process, subject to change. For instance, there has been debate about which activities should be considered taxonomy-aligned. One recent dispute revolved around whether to include activities related to nuclear energy and natural gas production or usage in the list (SCHEER Scientific Committee on Health, 2021). In 2022, it was approved that under specific conditions, certain activities in these two sectors could be considered transitional activities related to the climate change mitigation objective of the Taxonomy (Spinaci, 2022).

Amidst the ambiguity surrounding the notion of sustainability, the International Organization of Securities Commissions (IOSCO) has revealed that rating agencies and data providers evaluate and assess companies' ESG performances using different indicators, frameworks, and methodologies. They subsequently provide stakeholders with ESG information that corresponds to the companies accordingly (EC, 2021). Billio et al. (2020) have demonstrated that this heterogeneity in rating criteria leads to different rating agencies having opposing views on the ESG performances of the same companies. Furthermore, Berg et al. (2022) have discovered that the divergence in ESG ratings for the same companies can be attributed to 56% from differences in measurement methods employed by agencies, 38% from variations in scope, and 6% from the weights assigned by agencies in their methodologies.

For instance, MSCI (Morgan Stanley Capital International) rates companies based on their exposure to ESG risks and opportunities, as well as their management of these factors. In contrast, Refinitiv incorporates the level of disclosure and controversial activities conducted by companies in addition to their ESG performances during the ESG rating process. Meanwhile, the Financial Times Stock Exchange (FTSE) Russell considers factors such as companies' exposure to ESG risks, the relevance of these risks to their operations and sector, and the quality of their management skills in dealing with these risks when determining the weights for various categories. On the other hand, Sustainalytics takes into account the level of ESG risk exposure compared to the industry average for the same risks. It also distinguishes between manageable risks and idiosyncratic risks that are unexpected or irrelevant within the company's sector when assigning weights to categories. In summary, the categories established under each ESG pillar and the methodologies used to calculate individual pillar scores and overall ESG scores vary from one rating agency to another due to these diverse calculation methods in use.

IOSCO also emphasized that rating agencies use different data products, and most of these data sources are not shared by the rating agencies or by the companies themselves. As an illustration of this, a study found that in the 2021 Carbon Disclosure Project (CDP) Questionnaire, only 49% of private companies reported their Scope 1 and 2 emissions (Lino et al., 2022). Furthermore, the data used as input for the ratings is sourced from various places, including company surveys, annual reports, and reports from non-governmental organizations. Consequently, rating agencies heavily rely on disclosures made by the companies themselves to gather the relevant ESG information. For instance, MSCI indicates that 45% of the ESG data it utilizes for ratings comes from alternative sources, while the majority of data sources consist of voluntary and mandatory company disclosures (Eric Moen, 2020).

In addition to the challenges associated with different methodologies and data sources, there are also issues related to the disclosure of sustainability data. Although disclosure is a crucial aspect of ESG ratings, a study revealed that many companies improved their ratings by increasing their disclosure without necessarily improving their ESG performance (Rogge and Ohnesorge, 2022). Furthermore, the disclosure rates for ESG data and performance are even lower among U.S. operating companies. A recent 2022 report found that only 57% of the 1000 largest companies in the Russell 1000 index disclosed their Scope 1 and 2 emissions. In contrast, due to strengthened directives and regulations in the European Union (EU), 100% of the companies covered by these directives reported their greenhouse gas emissions (Thornton et al., 2022).

In addition to the aforementioned issues in the current rating process, as suggested by IOSCO, the current method of obtaining input data may lead to conflicts of interest and contribute to non-objective ESG rating outcomes. Moreover, collecting the necessary ESG data from companies often involves waiting for sustainability or annual reports to be disclosed. This waiting and data collection period may not align with the more immediate analysis needs of investors. Consequently, the Commission considers timeliness, accuracy, and reliability of the ratings as significant concerns regarding the current approach to ESG ratings (Delaney and Stewart, 2020).

As a result of these discrepancies between the taxonomy, data sources, and methodologies, ESG ratings become less precise and comparability between the ESG performances of different companies becomes challenging. This, in turn, diminishes the reliability and accuracy of the rating results.

To promote sustainable finance and enhance ESG rating processes, further steps and strategies have been taken to align with both European and international policy commitments, primarily for the transition to a low-carbon and greener economy. One of the recent measures to stimulate financing for sustainable activities is the introduction of the new Corporate Sustainability Reporting Directive (CSRD). This directive builds upon what was initially proposed by the Non-Financial Reporting Directive (NFRD) and aims to expand sustainability reporting requirements to a broader range of companies, regardless of their size (EC, 2021). With the new CSRD, efforts are directed at establishing a new set of rules and obligations to ensure that companies provide reliable, transparent, and comparable information essential for gathering and processing during the ESG rating process, thus enhancing the quality of ESG ratings. Furthermore, while the NFRD's scope was limited to large listed companies with over 500 employees, the CSRD extends its coverage to all large companies, whether listed or not, irrespective of employee numbers. This expansion means that sustainability reporting and disclosure of sustainability data now encompass nearly 50,000 companies, up from 11,000 under the previous directive (EC, 2021).

While the new Corporate Sustainability Reporting Directive (CSRD) expands the scope of covered companies, there remains a lack of mandatory sustainability reporting requirements, standardization, and proportionality, especially for unlisted small and medium-sized enterprises (SMEs) and micro-enterprises.

It's worth noting that 64% of negative environmental impacts, such as energy use, GHG emissions, and waste disposal, are attributed to SMEs, a majority of which are unlisted. Therefore, the exclusion of unlisted SMEs and micro-enterprises remains a concern even with the new CSRD (Constantinos, 2010). Additionally, many disclosure requirements, including the disclosure of companies' taxonomy-aligned activities or those not meeting the taxonomy criteria, remain voluntary for SMEs. Considering that SMEs account for 99% and micro-enterprises for 92.7% of all businesses in the EU (European Commission Press Corner, 2021), and are crucial contributors to Europe's economy, it becomes apparent that SMEs should be included in mandatory sustainability reporting and, consequently, in the ESG rating process. This inclusion is essential for SMEs to keep pace with developments in sustainable finance, secure capital support for transitioning to a sustainable economy, and align their business strategies accordingly.

Furthermore, banks and larger companies that invest in or collaborate with SMEs require access to relevant sustainability information from these SMEs to understand associated risks and opportunities. This underscores the need to make ESG rating for SMEs mandatory rather than voluntary. Additionally, mandating ESG performance disclosure compels SMEs to allocate resources and time for data collection, analysis, and reporting. This can help to reduce or eliminate the company size bias, a term indicating that larger companies tend to receive more favorable ESG ratings compared to smaller ones due to their data availability, disclosure practices, and adoption of sustainable management frameworks in their operations (Drempetic et al., 2020).

Despite the well-documented drawbacks of current ESG rating approaches, such as issues related to materiality, reliability, accuracy, comparability, and timeliness, neither the literature nor practical applications have presented direct solutions or guidance. Several initiatives, such as the introduction of a taxonomy, establishing a single database for both financial and sustainability-related information like the European Single Access Point (ESAP), proposing sustainability reporting frameworks such as the Global Reporting Initiative (GRI), mandating sustainability reporting, expanding the scope of reporting through the CSRD, and establishing regulatory and supervisory bodies for rating agencies, have been discussed in the literature (Kotsantonis and Serafeim, 2019; Capizzi et al., 2021; Sipiczki, 2022; Rogge and Ohnesorge, 2022). However, no direct solutions or recommendations have been provided to address the challenge of collecting input data for ESG rating methodologies.

In this paper, a novel approach is introduced to address two key limitations of the current ESG rating system, with a focus on the Environmental (E) pillar. First, a new method for directly collecting environmental data from Copernicus satellite observations for the assessment of the E pillar in ESG ratings for large companies is proposed. This approach is expected to resolve issues of inaccuracy, incomparability, and conflicts of interest that arise when accessing sustainability input data. The new data collection method relies solely on observable raw data from a single satellite database source.

It's important to note that this paper specifically addresses the E pillar and not the Social (S) and Governance (G) pillars. This is because satellite observations provide access to environmental data without requiring companies to provide data, which is the case for the S and G pillars. Currently, satellite observations are used in various applications, including earth observatory programs and in situ observations to detect environmental problems, monitor natural disasters' effects, forecast climate variables for the agriculture and insurance sectors, and track commodities. For example, Elkind et al. (2020) discuss the use of satellite observations to detect methane leakage, especially in the oil and gas sector. Yang and Broby (2020) suggest the use of satellite imagery for monitoring air pollutant emissions, water pollution, waste management, and natural resource management in alignment with GRI indicators. Patterson et al. (2016) argue for the use of satellites to observe mining operations in Brazil. However, there is currently no study or practice demonstrating the integration of satellite data into the ESG rating processes, particularly for evaluating the E pillar, as proposed and implemented in this paper.

Secondly, in the current landscape where investors are actively seeking raw sustainability data for SMEs to develop their own assessment methodologies, even accessing and making this raw data available to stakeholders remain problematic (Fernandez et al., 2021). Despite efforts to involve SMEs in sustainability reporting by providing simplified and tailored frameworks and encouraging them to adopt more sustainable business models (Barbagila et al., 2021), to our knowledge, there has been no study proposing potential environmental sustainability data sources and methods that can be employed to assess SME activities. Consequently, the novel approach introduced in this paper provides a new resource and method for both SMEs and rating agencies to effortlessly access the necessary environmental data for SMEs, making the data more accessible to both parties. This approach reduces the burden on SMEs to gather and report data, resulting in a more equitable ESG rating process, particularly for the Environmental (E) rating of large companies. This method of reaching and assessing SMEs effectively averts the issue of shadow rating for SMEs.

The general framework of the methodology presented in this paper begins by selecting large companies whose ESG, specifically E, ratings are publicly disclosed. Subsequently, categories to be included under the E pillar are defined. Relevant data for these categories, observed in the locations of the selected companies, is collected from the satellite databases of the Copernicus program, which is detailed in the following sections. Both the environmental data and the disclosed ESG rating data of the companies are fed as input into a Neural Network (NN) to establish the relationship between the data and the final ESG scores. With this approach, supported by a more objective and reliable dataset, it becomes possible to measure and evaluate the environmental performance of SMEs.

The paper's structure is as follows: Section 2 explains the rationale behind data selection, describes the data used in the paper, and elaborates on the methodology employed. Section 3 covers the explanation of results obtained from the NN, the application of the NN model to SMEs, and the corresponding results for SMEs. It concludes with a discussion on sustainability performance differences among branches of the same company as well as suggestions for the improvement and further research. The final section, Section 4, provides a conclusion.

2. Data and method

2.1. Copernicus

Copernicus is a program coordinated and managed by the European Commission (EC), with the aim of providing both satellite and in situ observations, analyses, and forecasts to end users for various purposes, including sustainable development, agricultural planning, city infrastructure management, and traffic management (Copernicus, 2021). Copernicus encompasses six services, each delivering databases for different fields and objectives. These services offer historical observation data as well as forecasting models.

One of these services is the Copernicus Atmosphere Monitoring Service (CAMS), which focuses on atmospheric monitoring and provides information primarily in five key areas: air quality and atmospheric composition, ozone layer and ultraviolet radiation, emissions and surface fluxes, solar radiation, and climate forcing (Copernicus. 2021). CAMS offers access to databases containing information such as greenhouse gas emissions, other toxic gas emissions, and radiative forcing. Radiative forcing refers to the energy changes in the atmosphere caused by climate change, observed in various atmospheric layers.

The Copernicus Climate Change Service (C3S) is responsible for providing information related to climate change and its potential impacts. This service offers data on variables like sea surface temperature, precipitation, snowfall amounts, and wind patterns.

Another service within the Copernicus program is the Copernicus Land Monitoring Service (CLMS), which monitors the current status and changes in bio-geophysical structures using both quantitative and qualitative methods. It provides information about vegetation types and changes in land structures, such as urbanization or deforestation in specific areas.

The Copernicus Marine Environment Monitoring Service (CMEMS) offers data related to wind patterns, sea ice conditions, and ocean currents, supporting the shipping sector and offshore activities. Additionally, it provides information about the physical, chemical, and biological characteristics of water systems, aiding in the monitoring of climate change impacts on marine ecosystems.

The Copernicus Emergency Management Service (CEMS) is designed for monitoring humanitarian crises and issuing early warnings for potential disasters such as floods, forest fires, and droughts.

The final service provided by the Copernicus Program is the Copernicus Security Service (CSS), which oversees borders to detect possible illegal immigration and collaborates with CMEMS in monitoring marine systems.

Since the scope and suitability of the other databases did not align with the requirements of this paper, only the databases from CAMS, C3S and CLMS of the Copernicus Program are used[1]. As mentioned earlier, this approach aims to provide an alternative method to solely relying on companies' self-reported data for the environmental (E) component of ESG ratings.

2.2. Data from the services of copernicus program

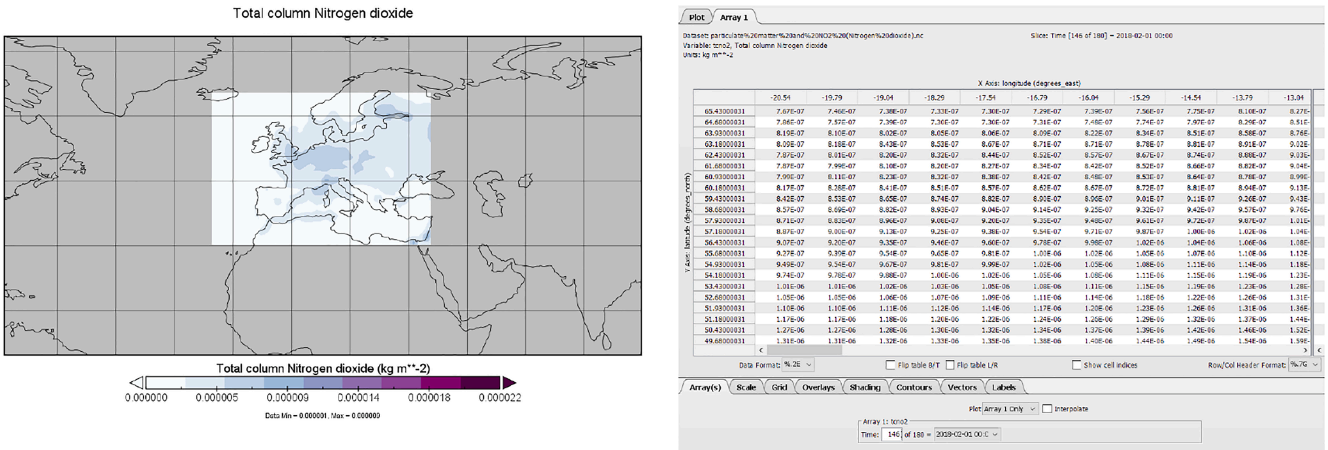
As explained in the introduction, rating agencies predominantly depend on companies' disclosures to obtain input data for their evaluations. This reliance not only reduces the reliability of the gathered data but also hinders standardization in rating large companies. Additionally, since SMEs are not currently obligated to report their ESG data, assessing the sustainability performance of SMEs for investment purposes or inclusion in supply chains becomes challenging.

In this paper, a novel methodology is introduced which relies on satellite data to assess companies falling under the environmental (E) pillar. The decision to focus exclusively on the E pillar is twofold. First, sustainability issues and the associated data under this pillar are key parameters in climate change research and have been observable and analyzable through satellite remote sensing tools since the first space observation conducted by the Vanguard-2 satellite in 1959 (Yang et al., 2013). Second, the parameters and data to be evaluated for the E pillar are part of the concept known as Essential Climate Variables (ECVs), developed by the Global Climate Observing System (GCOS). ECVs define a standardized set of variables that significantly contribute to describing the Earth's climate and are intended to be observed and measured to yield climate-related results (Petiteville et al., 2015). Moreover, as outlined in the COP21 handbook, ECVs such as atmospheric composition (carbon dioxide, methane, ozone, and aerosols), vegetation types and evolution, and soil moisture are variables whose corresponding information is primarily collected through satellite observations and is included in the evaluation criteria of the E pillar scoring methodology used by rating agencies.

The initial step involves the selection of companies whose data will be observed. This selection is closely linked to the sectors targeted in the paper. Identifying the sectors is crucial for determining which data will be observed via satellite and used for ESG rating. The choice of sectors is made by considering those whose E scores are predominantly defined by observable and measurable data collected from satellites. This approach ensures a more accurate establishment of the relationship between satellite data and E scores, as there are no significant additional factors or issues influencing the final E score. Therefore, the primary factor to consider when selecting sectors is the relevance of sustainability issues and data to the companies' respective industries and the environment. Many rating agencies incorporate material sustainability issues into their ESG rating methodology. This means that sustainability issues are identified, assessed, and prioritized based on the companies' operating industries or sectors. This approach aids in identifying which factors or criteria are the most significant ones in assessing the impact of the companies' sectors on the environment or society, resulting in more meaningful assessments that eliminate irrelevant or unimportant topics.

Bloomberg has introduced its Environmental and Social (ES) Scores, which provide performance measurement criteria and related quantitative data related to companies' operations. The process begins by categorizing companies based on their operating industries using the Bloomberg Industry Classification System (BICS). Following this categorization, Bloomberg defines sustainability issues and corresponding sub-issues that it will consider when assessing companies' sustainability performance. For example, sub-issues related to GHG emissions management, such as GHG emissions themselves and the presence of GHG emissions policies and regulations, are essential when evaluating the environmental (E) performance of companies in the Metals & Mining and Steel sector. However, these issues are not even part of the ESG assessment criteria for companies in the Household Products sector. Similarly, hazardous waste management and sub-issues like the percentage of hazardous waste generated and recycled are crucial environmental issues for the Chemicals industry but are not included as assessment criteria for companies in the Packaged Food and Beverages sector. Following the definition of sustainability issues and sub-issues, Bloomberg conducts sector-specific issue prioritization. This prioritization takes into account three components: the probability, magnitude, and timing of the sustainability issue's impact, both for the company and the broader system. For instance, the issue of sustainable products and its related sub-issues may receive a priority score of 7 for the Chemicals sector and 1 for the Household Products sector. This indicates that this issue will carry more weight in the ESG rating process for companies in the Chemicals sector compared to those in the Household Products sector. Following this prioritization, sustainability issues and sub-issues undergo ESG rating scoring based on their weights and the extent of disclosure by the companies when sharing the required data. Bloomberg's approach for making sector-specific categorizations for sustainability issues in its ESG scoring methodology serves as the basis for data collection in this paper. It allows for the selection of industries to be covered, focusing on those with significant and relevant impacts to be considered for the environmental (E) pillar. Taking Bloomberg's ESG scoring methodology as a reference, the sectors considered in this paper are Chemicals, Containers & Packaging, Construction Materials, Metals & Mining, and Steel. The choice of these sectors, as previously mentioned, is based on the fact that when evaluating companies in these

[1] The proposed methodology is adaptable to any type of jurisdiction, provided the correct explanatory variables and a consistent dataset are selected. For example the proposed methodology is aligned with several United Nations SDGs including SDG 6 (Clean Water and Sanitation), SDG 7 (Affordable and Clean Energy), SDG 12 (Responsible Consumption and Production), SDG 13 (Climate Action), SDG 14 (Life Below Water) and SDG 15 (Life on Land). Concerning satellite data resources, NASA's Earth Science Division and NASA's Earth Observation program are the key counterparts to the EU's Copernicus program in the US, providing Earth observation data and research for environmental monitoring, natural disaster and resources management.



(a) The map view in Panoply showing NO_2 emissions observed in February, 2018 for Europe

(b) The array holding the NO_2 emissions data observed in February, 2018 for Europe

Fig. 1. Example showing map views and array structure in Panoply for NO_2 emissions observed in February, 2018 in Europe.

sectors for ESG rating, the priority and weight assigned to sustainability issues observable from satellites are high.[2].

After selecting the sectors to cover, the next step is to determine which companies within these sectors will be considered for the analysis of their ESG scores. While the names of the companies are kept confidential, the selection criteria are as follows. Initially, the focus is on the largest companies in terms of market capitalization that have their headquarters and the majority of their operational sites in Europe. Subsequently, it is verified whether these companies have associated E scores available in the Bloomberg Terminal, starting from 2015 up to 2020 (excluding 2020). The selection of 2015 as the starting year is due to Bloomberg's provision of E scores for companies from that year, and 2019 is chosen because most of the companies' most recent E scores are evaluated for the year 2019 by Bloomberg. Companies meeting both of these criteria and also present in the Bloomberg Industry Classification System (BICS) are then chosen. As a result, 15 companies are selected for the assessment of their sustainability performance and used as proxies for scoring SMEs.

As mentioned in the introduction, the data from satellite observations are accessed through the databases of the three services of the Copernicus Program. While these services offer forecasting capabilities, this paper utilizes databases containing only historical observations. The data obtained from these databases are stored in the form of NetCDF (Network Common Data Form), which is a widely used file format in fields such as climate studies, meteorology, bioscience, and chemistry. NetCDF allows for the storage of multidimensional scientific data in arrays and includes explanations of the data. Fig. 1.

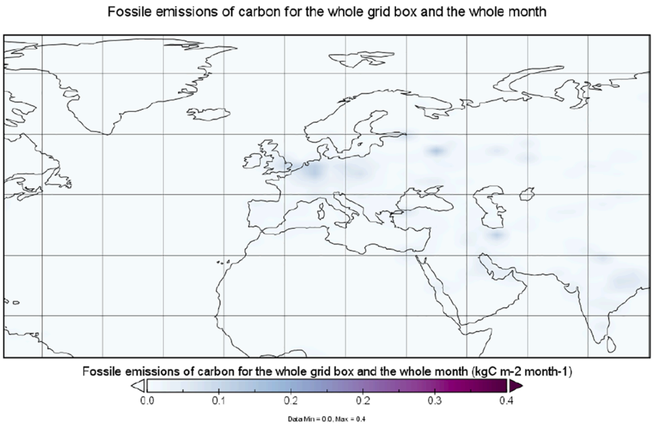
To visualize and work with the NetCDF data in an Excel-like format, a software program called Panoply is employed. Panoply categorizes the data based on coordinates and corresponding data, enabling users to view the data on a geographical map. It also provides arrays of time, allowing users to set and modify their desired time intervals and observe data changes over time on the same map. Furthermore, the scale bar in Panoply facilitates the observation of data magnitude and changes across different locations and timeframes. It also enables the conversion of map data into an array format, which includes coordinates and related data as per the user's choice. SubFigs. 1a and 1b illustrate the map view and array format of the same data, NO_2 emissions observed in February 2018, within a limited area of Europe.

Within the first service, CAMS, several data sets are utilized, including "CAMS Global Inversion-Optimized Greenhouse Gas Fluxes and Concentrations", "CAMS Global Emission Inventories", "CAMS Global Reanalysis Monthly Averaged Fields," and "CAMS Global Greenhouse Gas Reanalysis Monthly Averaged Fields." The first data set is employed to collect observations related to the emissions of carbon dioxide (CO_2) and nitrous oxide (N_2O). These two gases are of paramount importance for assessing companies' environmental impacts since they are responsible for trapping heat in the atmosphere and contributing to climate change. Moreover, in the Bloomberg ESG rating methodology, the emissions of these greenhouse gases, both Scope 1 (direct emissions from operations) and Scope 2 (indirect emissions from sources like electricity usage, heating, or cooling), are prioritized as the primary sustainability issue.

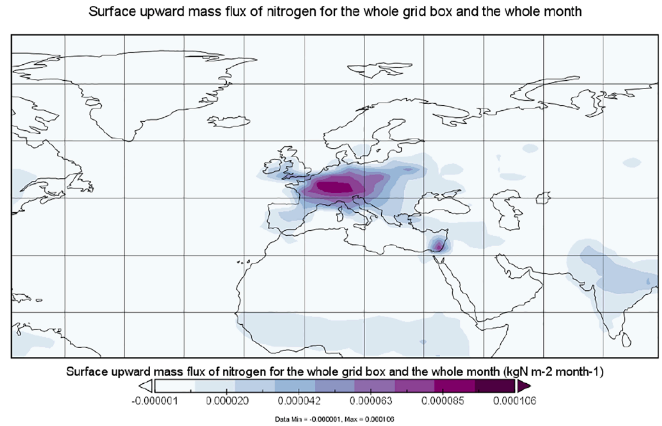
After collecting raw data, CAMS employs the atmospheric inversion technique to convert CO_2 and N_2O concentration data into corresponding net flux values from the surface to the atmosphere. This process involves modeling atmospheric factors such as winds, vertical diffusion of heat and humidity, and convection. Consequently, CAMS determines the surface upward net flux in units of $kgCm^{-2}month^{-1}$ (where C stands for carbon) and $kgNm^{-2}month^{-1}$ (where N stands for nitrogen) for these two gases[3]. Additionally, CAMS distinguishes

[2] We propose a methodology focused on the E pillar which is based on objective and open satellite data while we did not include any S and G contribution due to the unavailability of not-self-reported information. As a matter of fact among the data needed to appreciate the role of S and G pillars we can mention the employee turnover and their satisfaction with various aspects of their job, work environment and company culture, the workplace safety records, diversity metrics including gender, age, ethnicity and other relevant demographics, the community engagement metrics, the Ownership structure and the Board composition, the risk metrics as well as the transparency in financial reporting, the compliance with regulations including ethics and code of conduct, the stakeholder engagement records and the internal controls. These specific data and metrics may vary depending on the industry, location and nature of the company's operations and are mainly based on qualitative insights captured by conducting surveys, interviews and site visits. Therefore to validate our proposal we selected the sectors having the highest E weight and the lowest S and G impact and compared our ESG proxy to the Bloomberg ESG rating. Hence finally a sector-specific criticality is certainly represented by the weight of the E pillar in the ESG rating because it would make the validation by comparison null and void.

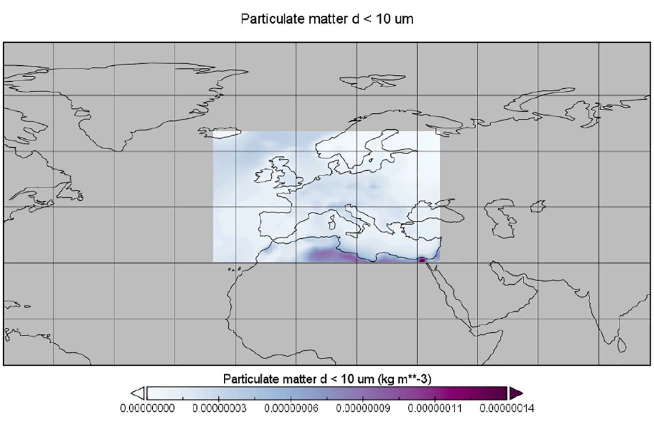
[3] $kgCm^{-2}month^{-1}$ and $kgNm^{-2}month^{-1}$ can be read as "kilograms of carbon/nitrogen per square meter per month." It's a common unit for expressing the rate of carbon/nitrogen flow or emissions over a specific area and time period, often used in environmental and climate science to quantify carbon/nitrogen-related processes.



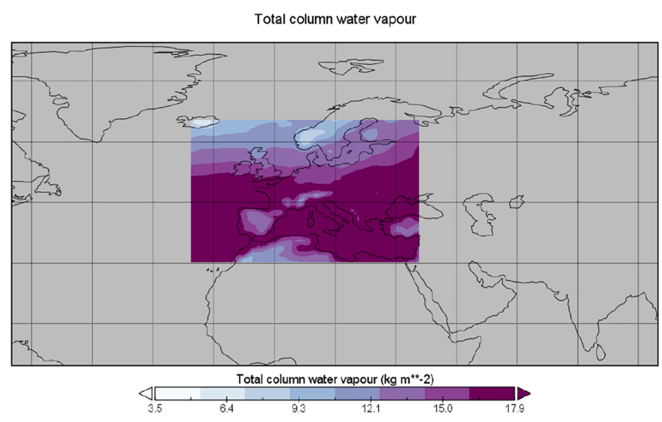
(a) The map showing CO_2 emissions observed in November, 2018



(b) The map showing N_2O emissions observed in April, 2017



(c) The map showing PM10 emissions observed in January, 2017



(d) The map showing water vapour amounts observed in May, 2015

Fig. 2. Example map views in Panoply for the data accessed from CAMS.

between sources of emissions for these gases.

For the purposes of this paper, data related to emissions resulting from fossil fuel combustion are selected. Visualization of CO_2 data from November 2018 and N_2O data from April 2017 over the map of Europe can be observed in Fig. 2 with subFigs. 2a and 2b.

From the second data set, various data points are collected, including emissions of another greenhouse gas, methane (CH_4), as well as emissions of other toxic gases such as carbon monoxide (CO), sulphur dioxide (SO_2), formaldehyde (CH_2O) and volatile organic compounds (VOCs), all expressed in units of $kgm^{-2}s^{-1}$, referring to the flux of these substances. This data set provides information about atmospheric pollutants and air quality, serving as indicators of hazardous gas emissions. The selected gases are common industrial emissions and play a crucial role in defining companies' efforts in hazardous waste management, which is a prioritized sustainability issue under the E pillar in Bloomberg's ESG scoring methodology. In this data set, the emissions for these gases are specified as anthropogenic emissions, meaning that they originate from activities such as power generation, industrial operations, solid waste, wastewater, solvents, and fugitive emissions. This specification enhances the consistency and reliability of inferring data related to the companies under the paper's scope, allowing for meaningful analysis of the relevance of emissions data to companies' operations while minimizing the influence of other factors, such as natural sources of these gases.

The third data set, similar to the others, provides information about atmospheric composition, with a primary focus on air quality. From this data set, information related to the quantity of particulate matter with a diameter smaller than 10μm (PM10), expressed in units of kgm^{-3}, is collected. PM10 is considered one of the most critical elements in air pollution (Cholakian et al., 2019). Additionally, data related to the quantity of nitrogen dioxide (NO_2), with units in kgm^{-2}, is gathered. NO_2 plays a significant role in absorbing solar radiation and has the potential to react with other gases to form ozone (O_3) (Solomon et al., 1999). An example map of Europe displaying the PM10 concentrations observed in January 2017 is presented in Fig. 2c.

Lastly, from the final data set within CAMS, data pertaining to the quantity of water vapor, measured in units of kgm^{-2}, is collected. Water vapor is one of the major greenhouse gases. A visual representation of the distribution of water vapor over Europe in May 2015 is displayed in subFig. 2d..

The second service of the Copernicus Program utilized in this paper is the CCS. From this service, two data sets are retrieved: "Ozone monthly gridded data from 1970 to present derived from satellite observations" and "ERA5 monthly averaged data on single levels from 1979 to present" [4]. From the first data set, information regarding ozone (O_3) levels in the troposphere is collected. Ozone is naturally generated in the

[4] ERA5 is the fifth generation reanalysis of climate and weather run by ECMWF (European Center for Medium-Range Weather Forecasts).

atmosphere through the interaction of oxygen and solar radiation and serves as a protective shield against harmful UV radiation from the Sun. However, excessive ozone accumulation in the troposphere is considered a greenhouse gas and can be hazardous to both human health and ecosystem productivity and biodiversity (Archibald et al., 2020). Therefore, data regarding ozone levels, measured in terms of mole content ($molm^{-2}$), specifically in the troposphere, is gathered. An example observation from July 2019 is illustrated in subFig. 3a. The second data set is employed to obtain information about groundwater usage, particularly represented by volumetric soil water content, which is an indicator for one of the Essential Climate Variables (ECVs), soil moisture. The data is collected from layer 4, expressed in $m^{-3}m^{-3}$, which corresponds to the water content found 100–298 cm below the surface[5]. This specific layer is chosen because it corresponds to the storage of groundwater reservoirs (Prakash Khedun et al., 2014). This data, as shown in Fig. 3b, is crucial to gather as one of the prioritized sustainability issues for the selected sectors pertains to water management, including freshwater withdrawals and groundwater usage.

The last service accessed is the CLMS, which is used for collecting data concerning the monitoring of the land evolution over time. From the data set "Leaf Area Index" data about another ECV that is the leaf area index (LAI) is collected. This variable is an indicator for the thickness of the vegetation cover expressed in the unit m^2m^{-2} and it is an alternative way to express vegetation cover in an area. As seen from the sub Figs. 4a and 4b, it gives an indication of the plant covers, that is, the colours in the maps are higher in summer time when compared to winter which may be explained by the shedding of the leaves of the trees in winter time. Additionally, LAI is gathered only for the vegetation that falls under the "high vegetation" that includes various leaves typed trees, forests and woodlands and excludes "low vegetation" typed structures such as crops and grasses. In this way, it is believed that a more accurate conclusion about the presence of a possible disruption of the biodiversity and ecosystem can be reached. Furthermore, it can be used as an indication for the sustainable sourcing especially for the companies working in Paper Containers & Packaging.

After collecting all the data from the Copernicus services as monthly data, the yearly averages of these data are taken since the ESG rating of the companies are given yearly according to the yearly sustainability performances of the companies. The sustainability data can be averaged or collected in a different way according to the raters' ESG scoring frequency or updates of the ESG scores. Table 1 reports the environmental variables collected from the corresponding Copernicus datasets.

2.3. Method

After selecting the companies to be covered in this paper, further research is conducted to identify their branches in Europe. During this process, the focus is on identifying manufacturing sites or production plants rather than city center offices. This emphasis stems from the belief that sustainability data associated with the companies is more closely linked to the operations occurring at production facilities rather than structures like administrative buildings.

To obtain the addresses of these company branches, official company websites are consulted to identify the facilities owned and operated by the companies. Once the addresses are collected, the next step is to determine the precise coordinates that will be used in the subsequent stages of data retrieval. These coordinates are acquired using Google Maps and are presented in latitude and longitude format. Although they may not cover the entire facility area, they serve as point representatives of these locations. Collecting this coordinate data is crucial for ensuring the most accurate input for the sustainability data obtained from the Copernicus Program's services and transferred from Panoply to Excel.

2.3.1. 2D interpolation

With the data collected from the service datasets and the coordinates of the branches, the next step involves retrieving the corresponding emissions and other sustainability data observed at the branch coordinates. To facilitate this, it's important to standardize the spatial resolution of the satellites used for these services. Spatial resolution measures the smallest object or area on the ground that can be resolved by the sensor or represented by each pixel (Liang and Wang, 2020). A satellite sensor with higher spatial resolution can capture more pixels, thus providing more detailed observations in an image. The datasets utilized in this paper are derived from different satellite observations, resulting in varying spatial resolutions ranging from 2000 m^2 to 10^5 km^2. This means that the smallest areas monitored by the satellites providing sustainability data can range from 2000 m^2 to 10^5 km^2. To ensure that the data can be compared in terms of the level of detail captured, 2D interpolation (or bilinear interpolation) is applied to the coordinates of the branches, as described in Press et al. (1992). 2D interpolation is commonly used in applications like image processing when a smoother image is desired rather than pixelated images. In this paper, with the exception of two data sets, "volumetric soil water" and "high vegetation LAI," 2D interpolation is employed for all the other data sets. This interpolation method is used to gather sustainability information from the nearest four coordinates offered by the data sets and then make inferences about the data observed over the exact desired coordinates, which correspond to the branches (Schowengerdt, 2007). In the case of the two mentioned data sets, "volumetric soil water" and "high vegetation LAI," the spatial resolution of the satellites monitoring these data is higher than the satellites used for the other data, with a magnitude on the order of approximately $10^3\,m^2$, therefore there is no need to apply the interpolation. Additionally, the nature of these two data types allows for observations in the surrounding area, making it more realistic to measure data for the branches. For instance, measuring the volumetric soil water usage in the surrounding four pixels and taking an average of these four data points, which is then attributed to the branch, makes more sense because it's likely that the branch utilizes groundwater from the surrounding area, not just the exact coordinate point. This reasoning also applies to high vegetation LAI, as monitoring deforestation in the surrounding area is more practical compared to focusing solely on a single point corresponding to the branch's coordinate. This is why, for these two data sets, instead of the application of 2D interpolation the average of the data from the four pixels surrounding the desired point is taken and assigned to the branch.

The concept of 2D interpolation can be better understood through the illustration provided in Fig. 5.

At first, a linear interpolation is done in the x-axis, i.e..

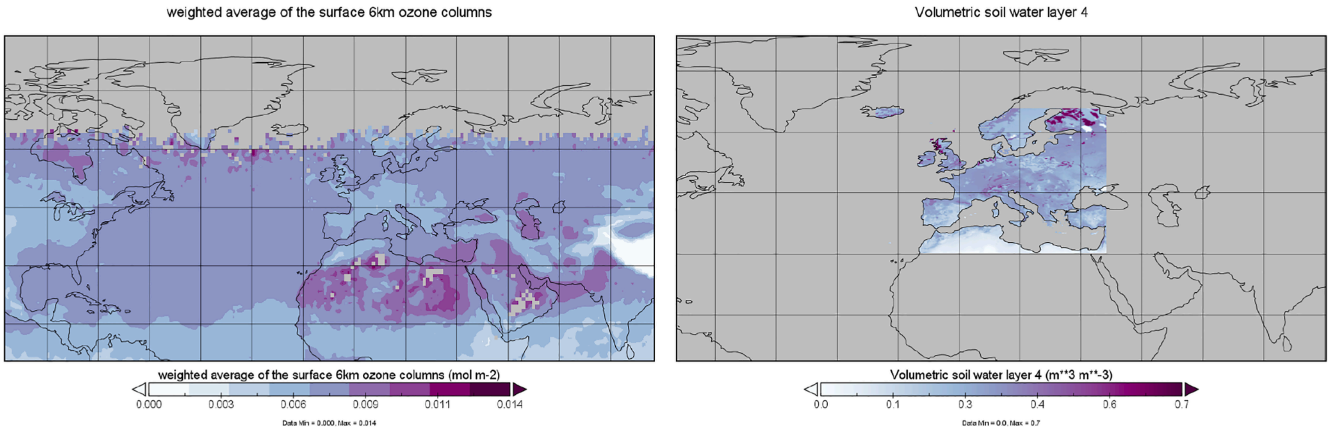
$$\begin{aligned} f(x,y_1) &= \frac{x_2-x}{x_2-x_1}f(Q_{11}) + \frac{x-x_1}{x_2-x_1}f(Q_{21}) \\ f(x,y_2) &= \frac{x_2-x}{x_2-x_1}f(Q_{12}) + \frac{x-x_1}{x_2-x_1}f(Q_{22}). \end{aligned} \quad (1)$$

It corresponds to assign f values linearly changing along x-axis between 2 points $f(Q_{11})$, $f(Q_{21})$ and $f(Q_{12})$, $f(Q_{22})$. One can see that values of function f in (1) correspond to $f(R_1)$ and $f(R_2)$ on Fig. 5. Then, they are substituted in the following equation:

$$f(x,y) = \frac{y_2-y}{y_2-y_1}f(x,y_1) + \frac{y-y_1}{y_2-y_1}f(x,y_2), \quad (2)$$

which corresponds to assigning f values linearly changing along y-axis between 2 points $f(R_1)$ and $f(R_2)$. After the substitutions and simplifications, the desired unknown function value becomes:

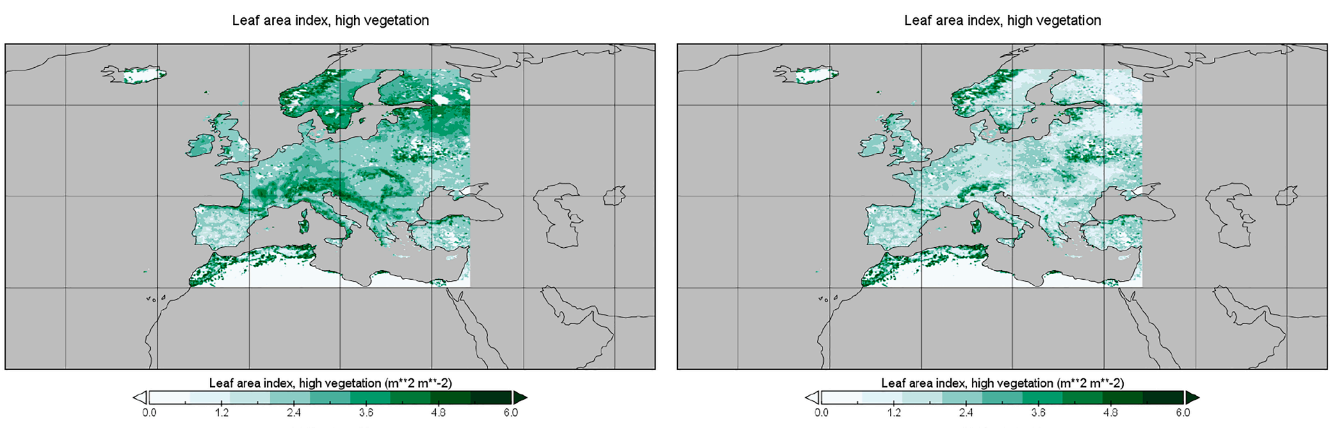
[5] $m^{-3}m^{-3}$ expresses certain density or concentration values, where the numerator (m^{-3}) represents the total volume, and the denominator (m^{-3}) represents a specific subset or concentration of that volume. It's commonly used in various scientific and engineering contexts to describe characteristics such as the density of a substance in a given volume or the concentration of particles or entities within a specific volume.



(a) The map showing ozone amount observed in troposphere level in July, 2019

(b) The map showing soil moisture monitored in September, 2016

Fig. 3. Example map views in Panoply for the data accessed from CCS.



(a) The map showing high vegetation LAI observed in June, 2016

(b) The map showing high vegetation LAI observed in December, 2016

Fig. 4. Example map views in Panoply for the data accessed from CLMS.

$$f\left(x,y\right)=\frac{1}{(x_2-x_1)(y_2-y_1)}\left[x_2-x \quad x-x_1\right]\begin{bmatrix} f(Q_{11}) & f(Q_{12}) \\ f(Q_{21}) & f(Q_{22}) \end{bmatrix}\begin{bmatrix} y_2-y \\ y-y_1 \end{bmatrix} \tag{3}$$

As an example, 2015 VOC emission data of a company branch having the coordinates (x, y) = (50.8262, 7.0292) is found by first finding the nearest four coordinates which hold the emission data. These coordinates are: Q_{11} = (x_1, y_1) = (50.7500, 6.9500), Q_{21} = (x_2, y_1) = (50.7500, 7.0500), Q_{12} = (x_1, y_2) = (50.8500, 6.9500) and Q_{22} = (x_2, y_2) = (50.8500, 7.0500). The known VOC emission values corresponding to these coordinates are: f(Q_{11}) = 2.83×10^{-8}, f(Q_{21}) = 6.19×10^{-8}, f(Q_{12}) = 1.21×10^{-9}, f(Q_{22}) = 2.77×10^{-8}. After substituting these values in Eq. (3), VOC emission at the desired point (x, y) = (50.8262, 7.0292) is found to be f(x,y) = 4.71×10^{-8}.

For the "high vegetation LAI" data set, the difference between the LAI values of two consecutive years is calculated to determine whether deforestation, reforestation, or no significant change has occurred.

After obtaining the coordinates of the company branches and the corresponding sustainability data, the next step involves preparing the data for the training phase of the neural network (NN). Since the raw sustainability data consists of different units and a wide range of scales, it is essential to normalize the original data. Data normalization is a critical step before feeding data into the NN. It serves multiple purposes, such as preventing extremely large outputs resulting from large ordered input values, ensuring balanced weighting for each input to minimize bias, and expediting the training phase (Sola and Sevilla, 1997; Nayak et al., 2012; Jayalakshmi and Santhakumaran, 2011). Among various normalization techniques, Min–Max Normalization is chosen for this purpose. This technique aims to scale the data within the range of $[0,1]$. Therefore, all the raw sustainability data (x) is treated as follows:

$$x_{normal}=\frac{x-x_{min}}{x_{max}-x_{min}}, \tag{4}$$

where x_{min} and x_{max} are the minimum and maximum values of the data found in the data set for the same attribute (same sustainability data), respectively.

After preparing the data for the training phase, it's essential to split the data into training, validation, and test sets. To ensure a more objective evaluation of the model's performance, a specific company and its corresponding branches are excluded from both the training and validation datasets. The remaining data is shuffled, so each row in the dataset now contains information about a branch, its emission profile for a given year, and its E score for that year. The first 75% of these rows, equivalent to 925 data points, are allocated to the training phase, while the remaining 25%, which amounts to 309 data points, are set aside for the validation phase. Furthermore, the excluded company, along with its

Table 1
The table showing the variables' names used in the analysis and their corresponding datasets from the Copernicus Program.

Variable Name	Copernicus Dataset	Copernicus Service
carbon dioxide (CO2), nitrous oxide (N2O)	CAMS Global Inversion-Optimised Greenhouse Gas Fluxes and Concentrations	Copernicus Atmosphere Monitoring Service (CAMS)
methane (CH4), carbon monoxide (CO), sulphur dioxide (SO2), formaldehyde (CH2O), volatile organic compounds (VOCs)	CAMS Global Emission Inventories	Copernicus Atmosphere Monitoring Service (CAMS)
PM10 (Particulate Matter diameter smaller than 10 micrometers) nitrogen dioxide (NO2)	CAMS Global Re-analysis Monthly Averaged Fields	Copernicus Atmosphere Monitoring Service (CAMS)
water vapour	CAMS Global Greenhouse Gas Reanalysis Monthly Averaged Fields	Copernicus Atmosphere Monitoring Service (CAMS)
ozone (O3)	Ozone monthly gridded data from 1970 to present derived from satellite observations	Copernicus Climate Change Service (C3S)
volumetric soil water	RA5 monthly averaged data on single levels from 1979 to present	Copernicus Climate Change Service (C3S)
Leaf Area Index (LAI)	Leaf Area Index	Copernicus Land Monitoring Service (CLMS)

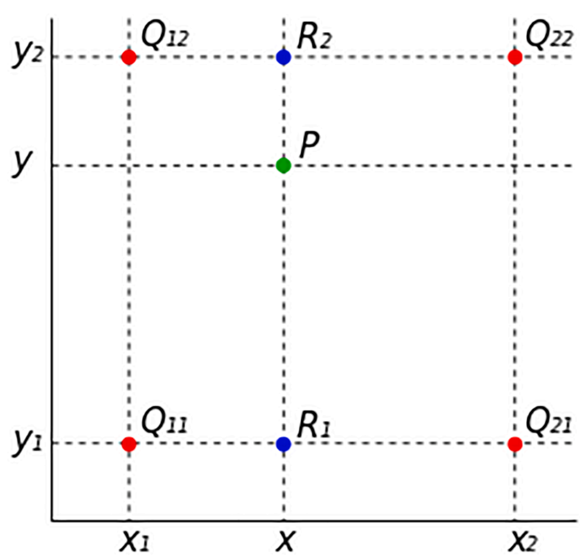


Fig. 5. Configuration of the points whose values are known (f(Q_{11}), f(Q_{12}), f(Q_{21}), f(Q_{22})) and unknown (f(x,y)).

10 branches and their sustainability profiles, comprising 50 data points, is utilized in the test phase. The model's predictive performance is evaluated based on this company's data for the years 2015 to 2019. Since the model hasn't been exposed to this particular company's data previously, it is assumed that any error observed in this prediction phase accurately reflects the model's true performance.

2.3.2. Neural network application

After completing the 2D interpolation, we proceed with the implementation of the neural network (NN). The rationale behind using a NN lies in the fact that many rating agencies follow a general approach to determine the final ESG rating score, which operates in a manner similar to that of a neural network. While the methodologies for ESG rating may differ, there is a common approach shared among agencies, which is succinctly described in their methodology disclosure reports. This approach typically starts with defining various categories under each pillar, as previously explained in this paper. For instance, under pillar E, categories such as carbon and toxic emissions, land and water usage, biodiversity impact, and waste management are considered. Once these categories are established, specific weights are assigned to each of them, which are then multiplied by the raw data scores corresponding to the categories. Ultimately, the weighted sum of the categories under each pillar is used to compute individual pillar scores, as well as the overall ESG score. The application of a neural network aligns with this practical methodology. By establishing a connection between the raw data and the final ESG scores through the NN, we can apply the trained NN results to the environmental data of selected SMEs, which have been collected from the Copernicus program. For the NN implementation, we employ the PyTorch library in Python. PyTorch offers several advantages, including being a free and open-source library, as well as enabling the saving of optimized models and learned parameters (Pytorch, n.d).

Training of the NN

Starting with the training phase, we begin by defining and experimenting with various neural network (NN) architectures. We build NN models with differing numbers of hidden layers, the quantity of neurons in these layers, learning rates, and epochs. For example, we test networks with a single hidden layer consisting of 500 neurons, two hidden layers with 200 neurons each, two hidden layers with 300 neurons each, four hidden layers with 100, 70, 70, and 70 neurons, and networks with six and eight hidden layers, each having 50 neurons. We vary the epoch numbers within a range of 200 to 5000 with increments of 200 and use learning rates of $10^{-4.2}$, 10^{-4}, $10^{-3.8}$, $10^{-3.5}$, 10^{-3}, and $10^{-2.5}$. The chosen activation function is the sigmoid function. In these NN models, the first layer is composed of the sustainability data obtained from the satellites. Each neuron in this layer corresponds to different sustainability attributes (such as CO_2 emissions, volumetric soil water, VOC emissions, etc.) for a given branch and year. The initial weights for each neuron are assigned automatically by PyTorch using randomization, following the normalized Xavier initialization method. This method is recommended for initializing weights when employing the sigmoid activation function. After each epoch, we calculate the error by comparing the NN's outputs with the expected results, which are the E scores collected from Bloomberg. We utilize the quadratic loss function (mean squared error) to measure the error, and weights are updated during the backpropagation phase to minimize this error. By the end of the training part, we obtain optimized weights. However, other hyperparameters and the number of hidden layers remain to be validated in the subsequent steps.

Validation of the NN

In the validation phase, we utilize the remaining 25% of the data set, which consists of 309 data points, to evaluate the candidate models developed during the training phase. Subsequently, we calculate the error by computing the mean squared error between the actual E values provided by Bloomberg and the model's predictions. This phase is carried out for models featuring various hyperparameters, such as epoch number and learning rate. The validation phase is crucial for two key reasons. First, although it is common to observe a decrease in error as the epoch number increases during the training phase, it is possible that, in the validation step, the error may begin to rise after surpassing a certain threshold of epochs. This is primarily because the model begins to memorize the training data, losing its ability to accurately assign E scores to new data points beyond the training set. Second, the learning rate has a significant impact on the size of steps taken to find the minimum error point on the error surface. Large learning rate values may lead to overshooting the minimum points, while small learning rate values can result in slow optimization due to excessively small step sizes. Consequently, it is imperative to experiment with different learning rate values to identify an optimal setting.

Testing of the NN

The final phase of the neural network (NN) analysis is the test phase. In this phase, the model that was chosen based on its performance in the validation part is assessed in terms of its error magnitude. A new and separate data set, consisting of the branches belonging to the excluded company for the years 2015 to 2019, is employed. During this phase, the model, selected as the best performer in the validation phase, predicts the E scores for these unseen branches. Subsequently, the mean squared error is computed between these predictions and the actual E values, and this error serves as an indicator of the model's accuracy.

3. Results and discussion

3.1. NN model results

The model chosen based on the minimum error in the validation phase comprises six hidden layers, each with 50 neurons, a learning rate of $10^{-3.2}$, and a total of 5000 epochs. As seen in Fig. 6, predictions made with this model are both precise and accurate. The utilization of such a large epoch number is attributed to the limited scope of sustainability data collected in this paper, which doesn't encompass all the parameters considered by ESG rating agencies. The aim is to compensate for the predictions made with this limited information by increasing the number of epochs in the neural network (NN). However, it's important to note that this can lead to a "memorized" model with reduced generalization capability on other datasets. Nonetheless, as demonstrated in Fig. 6, this selected model manages to capture the general trend in the ESG (or E) ratings in the independent validation dataset. When this model is tested in the final stage, the outcomes are illustrated in Fig. 8, with an associated error of 0.0217. It's important to acknowledge that while some predictions may appear slightly divergent from the actual values, it is not expected for the prediction points to align precisely with the actual values. This discrepancy arises from the fact that the actual values represent the entire company's E score, whereas our model's predictions pertain to individual operating sites. In real-life scenarios, the two sustainability performances are challenging to directly compare, constituting the primary source of inherent error in this model.

A crucial point of consideration relates to the availability and quality of data. It's essential to acknowledge that data gaps or inaccuracies may introduce biases into the representation of ESG performance, as the neural network relies heavily on the data it is trained on. Therefore, if certain ESG aspects are underrepresented or entirely missing, the NN might not offer a comprehensive overview, potentially leading to incorrect conclusions and resulting in inaccurate ESG ratings that misrepresent a company's sustainability endeavors. Furthermore, the quality of data used for training the NN is of paramount importance. Inaccurate or low-quality data can yield erroneous predictions and ratings. The NN's algorithms may struggle to differentiate between reliable and unreliable information, which can have significant implications. Another point that warrants attention is the evolution of ESG metrics over time. When dealing with older datasets that do not align with current ESG reporting standards, data gaps and inaccuracies might be more pronounced. To address these issues, it is crucial to maintain a continuous process of updating and improving the dataset used to train the NN. Regular updates, potentially on a weekly basis, can help ensure the dataset remains relevant and aligned with current standards. In cases of data gaps, imputation or data enhancement techniques can be employed to mitigate potential distortions. Moreover, it is advisable to routinely validate the NN's outputs against ground truth data or human assessments to enhance the credibility and reliability of the model. Lastly, maintaining transparency in the NN's methodology and providing clear explanations for the ESG ratings it generates can foster trust among users and stakeholders. Fig. 7.

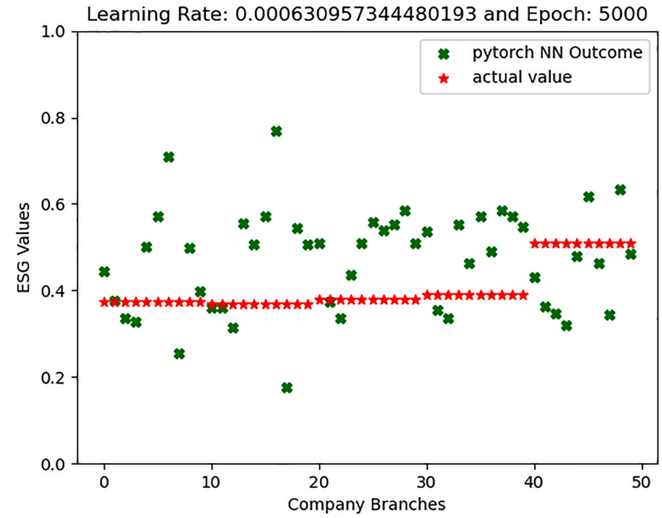


Fig. 8. The figure made by predictions on the test dataset of the selected model in the validation phase. The total error is 0.0217.

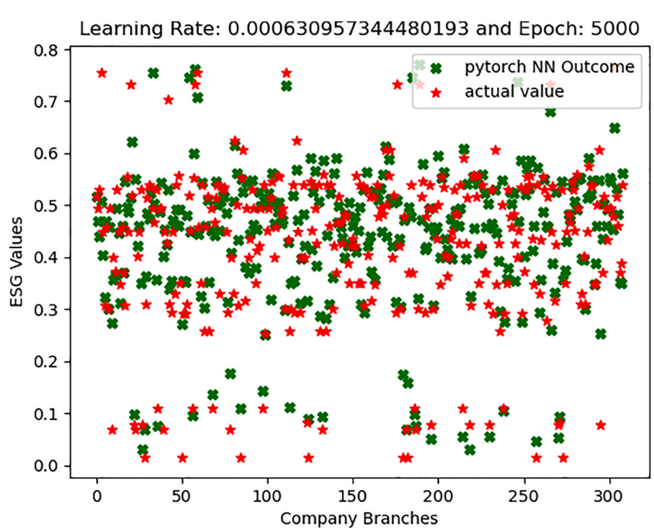


Fig. 6. The predictions of the NN model on the validation data set which belongs to the 25% of the data set.

3.2. Application on SMEs

Following the application of the NN model to the E scores of large companies retrieved from Bloomberg, the model is employed to assess its performance in predicting E scores for SMEs. To achieve this, 10 SME companies that meet the European Commission's criteria for SME classification are carefully chosen. Furthermore, these companies are exclusively drawn from the sectors outlined in the Data section, ensuring alignment between the model developed for large companies and its suitability for SMEs. Specifically, four companies originate from the Packaging sector, four from the Chemicals sector, and three from Steel Manufacturing. Once the SMEs are selected, the same methodology is applied. Initially, the coordinates of these companies are determined, and the corresponding sustainability data is sourced from the Copernicus Program, mirroring the process used for the large companies. The subsequent steps encompass interpolation and normalization. These prepared data for the SMEs are then input into the NN model, resulting in E scores for these SMEs spanning the years 2015, 2016, 2017, 2018, and 2019.

The E score outcomes that the NN model results for the SMEs are displayed in Fig. 9.

A last comment concerning the scalability of the algorithm is in

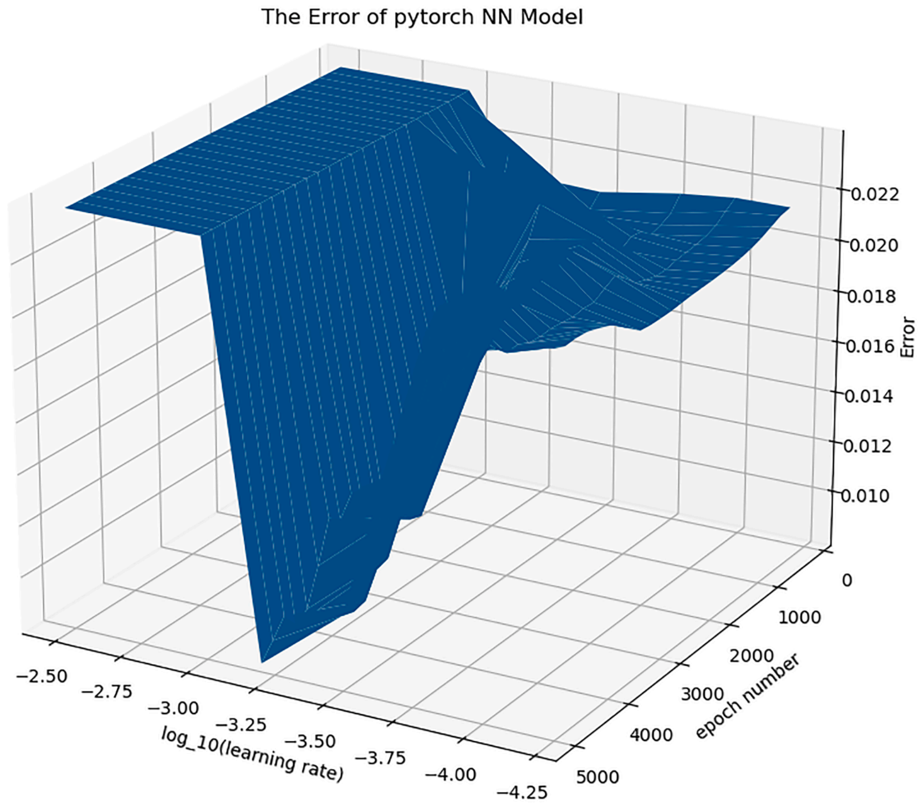


Fig. 7. 2D Error Plot. Notice 5000 epoch and $10^{-3.2}$ learning rate corresponds to the minimum of error.

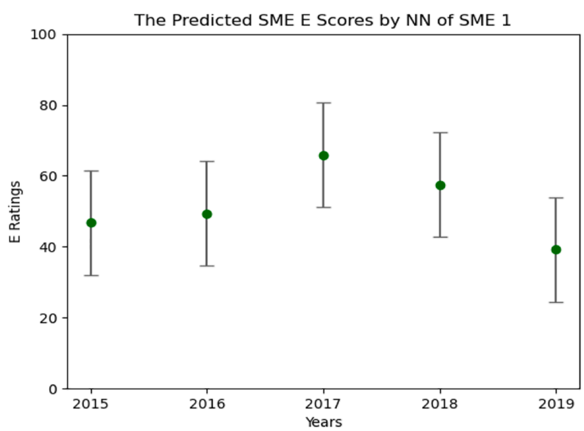
order. When deploying a trained NN to assess a larger cohort of SMEs, several prospective challenges related to scalability may arise. These challenges can impact the efficiency and effectiveness of the assessment process. Among them the increased data volume can strain computational resources and lead to longer processing times for assessments. Scalability may require more powerful processing demands, i.e. high-performance computing clusters or cloud computing resources, adequate data storage solutions to ensure data availability/accessibility and efficient preprocessing pipelines. Additionally, ongoing monitoring and quality control of data, model performance, and interpretability are key components of successfully deploying a trained NN for assessing a larger cohort of SMEs.

3.3. Heterogeneity among company branches

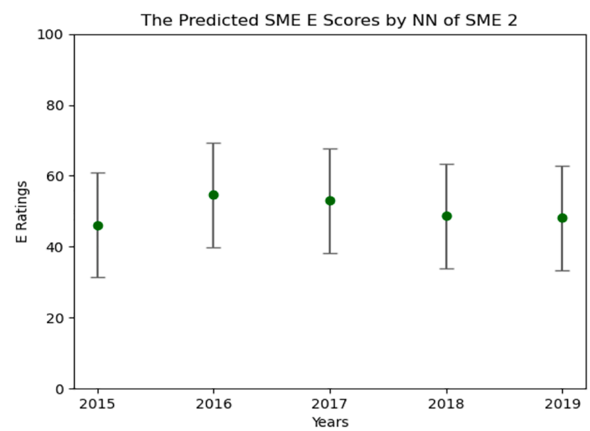
As shown in Fig. 8, it's evident that the NN model assigns different E scores to branches of the same company in a given year, even if the actual E scores, marked by the red stars, are the same. This discrepancy arises because ESG rating agencies typically evaluate companies based on the disclosure information provided by the company as a whole. For instance, they consider the total amount of CO_2 emissions data collected from all branches and aggregate this data, without distinguishing individual branches' emissions data. It is, therefore, reasonable to assume that the weighted averages of the NN model's E score predictions for branches would yield the actual E scores. This tendency is partially visible in Fig. 8. In other words, it is observed that the averages of the NN model's predictions, which represent the weighted average E scores of the branches, are quite close to the actual E scores. This practice results in an unfair and inconsistent scoring among branches, as both sustainably and unsustainably performing branches receive the same final ESG (or E) scores. In contrast, the NN model takes into account the localized coordinates of the branches and their corresponding sustainability data accessed from the Copernicus Program. This localized perspective in the rating process is more realistic as it assigns different ratings to the branches based on their individual performances, rather than evaluating the entire company's performance as a whole. If this localized rating process were adopted, stakeholders and the company itself would gain insight into the variations in the sustainability performances of different branches. They could make investment decisions and take improvement actions accordingly, resulting in more tailored and informed money management decisions. For example, an investor might be interested in investing in a company with an environmentally friendly branch in their hometown. In this case, considering only the branches in their hometown and assigning an E score based on these branches would be more suitable for the investor, compared to considering the overall ESG rating of the entire company, which might include branches of less relevance to the investor's preferences. Additionally, by taking into account the sizes and operation capacities of the branches, weighted averages of local performances could be calculated to arrive at a final score. Currently, there is no rating agency that provides localized ESG scores for branches, so quantitative comparisons with the proposed NN model assessments cannot be made. However, as argued earlier, the NN model's results do indicate distinctions in environmental profiles among branches.

3.4. Further remarks

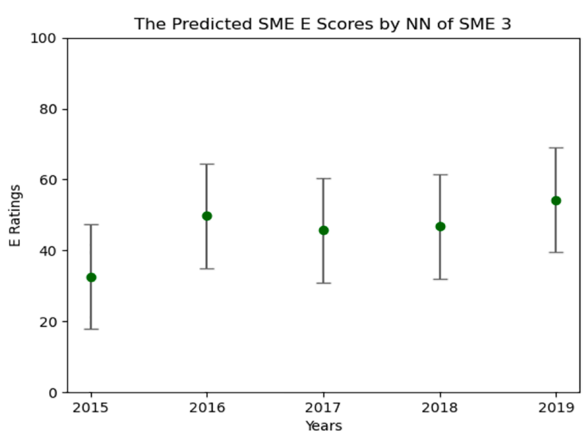
As a final point of reflection, we wish to compare the proposed method with traditional ESG rating to underscore the desirability of integrating the two within the current regulatory framework, encouraging more robust ESG practices and fostering a culture of transparency and accountability. Traditional ESG rating is currently based on a dual information channel that we could classify as external and internal data, where objectivity is the differentiating factor. If external data is non-influenceable (as is the case with information regarding carbon emissions, water usage, and biodiversity and natural resources), internal data comes from measurements related to the efficiency of the production process in term of efforts to reduce energy consumption and transition to renewable energy sources, to minimize waste and promote recycling and circular economy practices, to prevent and mitigate pollution and environmental harm, to be compliant with environmental regulations and permits, to develop sustainable eco-friendly products, and to ensure sustainable sourcing and procurement practices. Therefore it is closely tied to the transparency and completeness of a firm's disclosures. The proposed NN-approach has high potential for integrating with the traditional rating in determining an objective metric that would enable an assessment independent of the choice of a specific proprietary model for environmental impact of companies, regardless of size, with rating convergence guaranteed by the use of a single, open-source dataset untainted by disclosure mechanisms. This objective metric should be seen as a first step of evaluation, to be followed by a second step where internal data (self-disclosed) are taken into account, acting for a refinement of the objective rating; so the main driver of the rating is



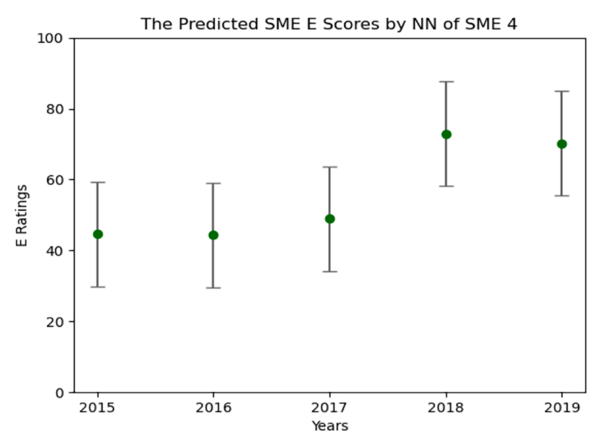
(a) E scores outcomes for the first SME



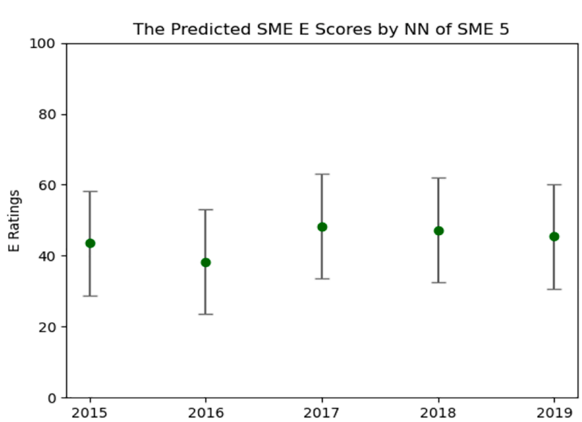
(b) E scores outcomes for the second SME



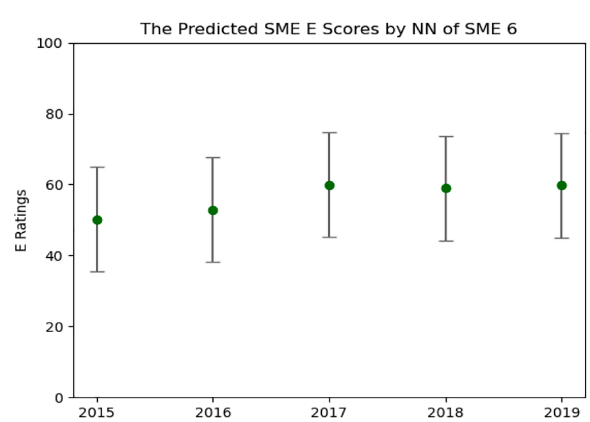
(c) E scores outcomes for the third SME



(d) E scores outcomes for the fourth SME

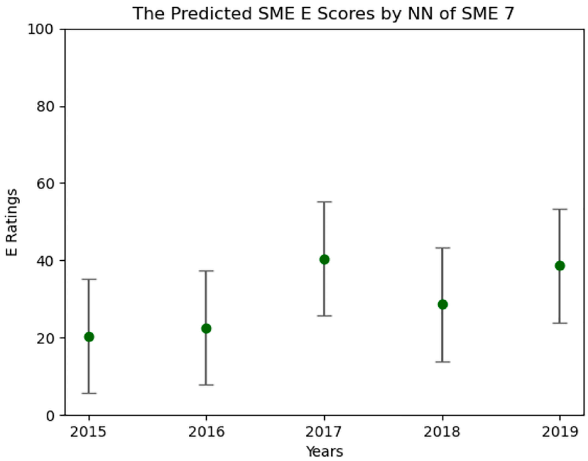


(e) E scores outcomes for the fifth SME

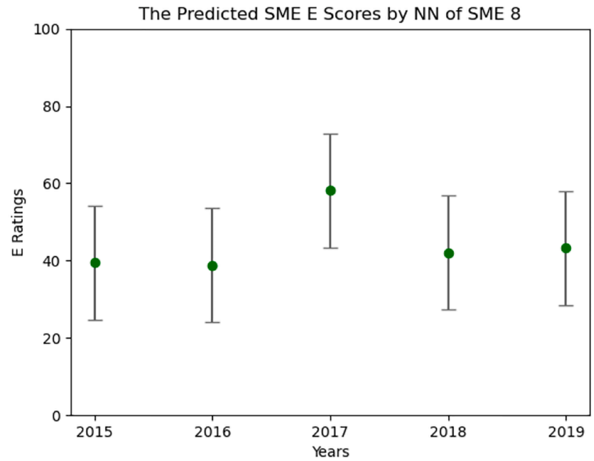


(f) E scores outcomes for the sixth SME

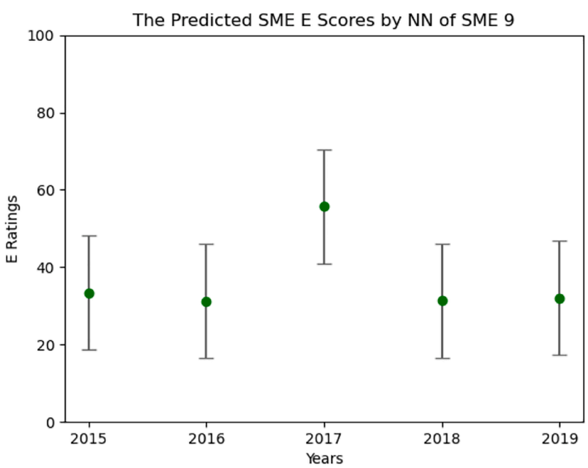
Fig. 9. The graphs showing the E scores resulted from the NN model for the SME companies from year 2015 to 2019. The error bars represent the error calculated in the test phase.



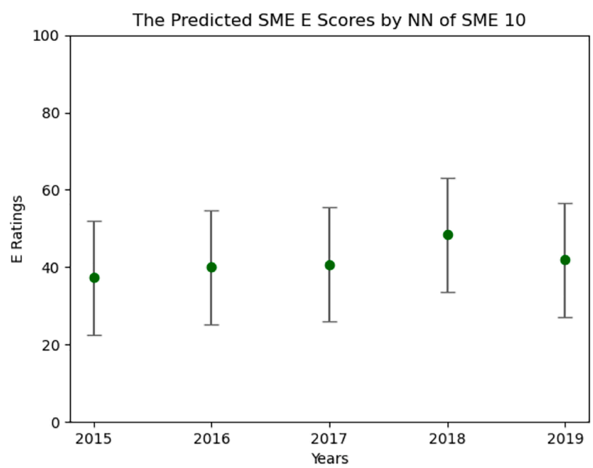
(g) E scores outcomes for the seventh SME



(h) E scores outcomes for the eighth SME



(i) E scores outcomes for the ninth SME



(j) E scores outcomes for the tenth SME

Fig. 9. (*continued*).

ensured to be objective, at least until a level of reliability in the quality of internal reporting is reached, allowing us to equate external data with internal ones. The proposed NN-based approach is entirely innovative and leverages the capabilities of NNs to infer SMEs' ratings from the complex network representing publicly traded companies, thus stimulating a positive ripple effect in the SME sector. In addition to the warnings to mitigate the limitations posed by data gaps, sector-specific challenges, regional disparities, computational demands, scalability, and evolving ESG factors discussed in the preceding sections, it is worth mentioning some recommendations for future research and improvements. First of all, the number of years or time points taken as input data can be increased. Since there are limited number of years in which the companies are assigned an ESG score (or E score), the input data set used in NN is limited to years 2015–2019. Similarly, the number of companies and their branches can be increased to broaden the sample size and gather a more substantial insight about the companies' sustainability performances. It is already stated that the overall E score of a company can be represented as the weighted average of each of its individual branches. Therefore, including more branches of a company and averaging their E score predictions would yield a more accurate E score for the whole company. Secondly, although it is tried to cover the majority and the most significant environmental parameters contributing to E score, there are still several parameters used by the rating agencies which cannot be reached via satellite observations. So, if the extent of the satellite observations can be expanded, more accurate results for future E scores can be achieved. Third, although it is seeked to work with satellites having higher spatial resolutions, the branch locations are considered as pixels representing the average of the surrounding area. If the satellites' spatial resolutions are enhanced further the operation sites can be observed more accurately and closely. For example, for the GHG emissions data, the observation areas can be focused on the chimneys of the manufacturing facilities. Finally, as also stated before, if the ESG scoring methodology is changed so that the individual scores are admitted to the branches, the model constructed in this paper can be used directly to assess the branches separately.

4. Conclusion

Considering the growing importance of sustainability and the European Commission's efforts to promote sustainability in the business world, this paper aims to make a valuable contribution to the existing literature on ESG (Environmental, Social, and Governance) ratings for companies. Specifically, this paper addresses the challenge of objectively assessing companies and the lack of regulations for ESG evaluations concerning Small and Medium-sized Enterprises (SMEs). The proposed methodology introduces an innovative solution based on a Neural Network (NN). This NN is calibrated to calculate E scores for branches of large companies based on their environmental profiles. The model is then leveraged to assign E scores to SMEs in an objective manner. The focus on the E score is driven by the observation that the overall ESG scores of the selected companies heavily rely on the Environmental (E) pillar. This observation aligns with the SASB (Sustainability Accounting Standards Board) standards and the materiality map, as well as data obtained from the Bloomberg terminal. To access the necessary input data for the NN, satellite observations from the Copernicus Program are utilized. These observations encompass 14 different categories, including greenhouse gas emissions, devegetation, groundwater usage, and air pollutants. The data is extracted from satellite databases corresponding to the locations of selected company branches, spanning from the years 2015 to 2019. This novel approach to sourcing data from satellites mitigates the risk of incomplete or inaccurate sustainability data disclosures by companies. It also enables timely assessments of companies without having to rely on company reports or other external sources, thus facilitating the ESG rating of SMEs. After applying the NN model to large companies, including the validation and test phases, the model is then extended to assess 10 selected SMEs operating in the same sectors. The results demonstrate that the NN model effectively aligns with the changes in the sustainability data of the SMEs, accurately capturing year-by-year shifts in emissions, pollution, and other sustainability metrics, and assigning E scores accordingly. A noteworthy observation is that the NN model assigns varying E scores to different branches of the same company, in contrast to the uniform ESG scores typically assigned by rating agencies. This approach of assigning diversified ESG scores to branches is seen as more transparent and informative. It allows for the evaluation of a company's performance at different branches and locations in terms of sustainability. This innovative methodological approach is expected to make a substantial contribution to the field of sustainability, particularly concerning the SME rating process. Furthermore, it provides a solid foundation for future studies and regulatory developments within the scope of EU resolutions.

CRediT authorship contribution statement

Seben Ozkan: Conceptualization, Methodology, Software, Data curation, Writing - original draft. **Silvia Romagnoli:** Supervision, Writing - review & editing. **Pietro Rossi:** Software, Validation, Supervision.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Silvia Romagnoli reports financial support, administrative support, and statistical analysis were provided by University of Bologna. Silvia Romagnoli reports a relationship with University of Bologna that includes: employment.

Data availability

Data will be made available on request.

References

Archibald, A.T., Neu, J.L., Elshorbany, Y.F., Cooper, O.R., Young, P.J., Akiyoshi, H., Zeng, G., 2020. Tropospheric Ozone Assessment Report. Elementa: Science of the Anthropocene, 8(1). https://doi.org/10.1525/elementa.2020.034.

Barbagila, M., Butticè, V., Giudici, G., Mendy, J., Sarker, T., Sharma, G.D., Thomas, A., Department, A.Z., 2021. Supporting Smes in Sustainable Strategy Development Post-Covid-19 : Challenges and Policy Agenda for the G20. Italy.

Berg, F., Kolbel, J.F., Rigobon, R., 2022. Aggregate confusion: The divergence of ESG Ratings. Rev. Finance. https://doi.org/10.1093/rof/rfac033.

Billio, M., Costola, M., Hristova, I., Latino, C., Pelizzon, L., 2020. Inside the ESG ratings: (dis)agreement and performance. SSRN Electron. J. https://doi.org/10.2139/ssrn.3659271.

Capizzi, V., Gioia, E., Giudici, G., Tenca, F., 2021. The Divergence of ESG Ratings: An Analysis of Italian Listed Companies. J . Finan. Manage., Markets Instit. 9 (2), 1–21. https://doi.org/10.1142/S2282717X21500067.

Cholakian, A., Colette, A., Coll, I., Ciarelli, G., Beekmann, M., 2019. Future climatic drivers and their effect on PM10 components in Europe and the Mediterranean Sea. Atmos. Chem. Phys. 19 (7), 4459–4484. https://doi.org/10.5194/acp-19-4459-2019.

Constantinos, C., 2010. SMEs and the environment in the European Union. Retrieved 2022, from https://op.europa.eu/en/publication-detail/-/publication/aa507ab8-1a2a-4bf1-86de-5a60d14a3977.

Copernicus in detail. 2021, from https://www.copernicus.eu/en/about-copernicus/copernicus-detail.

Delaney, D., Stewart, P., 2020. Report on Sustainability-Related Ratings, Data and Research (Publication). Retrieved 2022, from ERM website:https://op.europa.eu/en/publication-detail/-/publication/d7d85036-509c-11eb-b59f-01aa75ed71a1/language-en/format-PDF/source-183474104.

Drempetic, S., Klein, C., Zwergel, B., 2020. The influence of firm size on the ESG score: Corporate Sustainability Ratings under review. J. Bus. Ethics 167 (2), 333–360. https://doi.org/10.1007/s10551-019-04164-1.

EC, 2021. FAQ: What is the EU Taxonomy and how will it work in practice. Retrieved 2022, from https://ec.europa.eu/info/sites/default/files/business_economy_euro/banking_and_finance/documents/sustainable-finance-taxonomy-faq_en.pdf.

Eric Moen., 2020. MSCI ESG Ratings [Brochure]. Author. Retrieved 2022, fromhttps://www.msci.com/documents/1296102/21901542/MSCI+ESG+Ratings+Brochure-cbr-en.pdf.

Elkind, J., Blanton, E., Denier Van Der Gon, H., Kleinberg, R.L., Leemhuis, A., 2020. Nowhere to Hide: the implications of satellite-based methane detection for policy, industry and finance. Center on Global Energy Policy, Columbia.

Escrig-Olmedo, E., Fernandez-Izquierdo, M., Ferrero-Ferrero, I., Rivera-Lirio, J., Munoz-Torres, M., 2019. Rating the raters: Evaluating how ESG Rating Agencies Integrate Sustainability principles. Sustainability 11 (3), 915. https://doi.org/10.3390/su11030915.

European Commission Press Corner, 2021. Retrieved 2021, fromhttps://ec.europa.eu/commission/presscorner/detail/en/ip_21_1804.

Fernandez, E., Friedrich, H.-J., O'Gorman, O., Huemer, G., Mazanec, L., Nieto-Márquez, J.G., Vismara, A., Plejić, M., Plattner, L., Österberg, L., 2021. Empowering EU Capital Markets - Making listing cool again Final report of the Technical Expert Stakeholder Group (TESG) on SMEs. European Commission https://finance.ec.europa.eu/system/files/2021-05/210525-report-tesg-cmu-smes_en.pdf.

Jayalakshmi, T., Santhakumaran, A., 2011. Statistical normalization and back propogation for classification. Int. J. Comput. Theory Eng. 89–93 https://doi.org/10.7763/ijcte.2011.v3.288.

Kotsantonis, S., Serafeim, G., 2019. Four things no one will tell you about ESG data. J. Appl. Corporate Finan. 31 (2), 50–58.

Liang, S., Wang, J., 2020. Spatial resolution. In: Advanced remote sensing: Terrestrial information extraction and applications. Academic Press, London, United Kingdom. https://doi.org/10.1016/C2017-0-03489-4.

Lino, M., Doolan, P., Divgi, P., Mehrotra, R., 2022. Closing the Public-Private Environmental Transparency Gap (Rep.). Retrieved 2022, from https://www.bain.com/insights/closing-the-public-private-environmental-transparency-gap/.

Lopez, M.V., Garcia, A., Rodriguez, L., 2007. Sustainable Development and corporate performance: A study based on the dow jones sustainability index. J. Bus. Ethics 75 (3), 285–300. https://doi.org/10.1007/s10551-006-9253-8.

Nayak, S.C., Misra, B.B., Behera, H.S., 2012. Evaluation of normalization methods on neuro-genetic models for stock index forecasting. In: 2012 World Congress on Information and Communication Technologies. https://doi.org/10.1109/wict.2012.6409147.

Petiteville, I., Lecomte, P., Ward, S., Dyke, G., Steventon, M., Harry, J. (Eds.)., 2015. Satellite Earth Observations in Support of Climate Information Challenges (Rep.). Retrieved 2021, from http://eohandbook.com/cop21/files/CEOS_EOHB_2015_COP21.pdf.

Patterson, S.R., Kozan, E., Hyland, P., 2016. An Integrated Model of an Open-Pit Coal Mine: Improving Energy Efficiency Decisions. Int. J. Prod. Res. 54, 4213–4227.

Prakash Khedun, C., Sanchez Flores, R., Rughoonundun, H., Kaiser, R., 2014. World Water Supply and use: Challenges for the future. Encycl. Agricul. Food Syst. 450–465 https://doi.org/10.1016/b978-0-444-52512-3.00083-8.

Rogge, E., Ohnesorge, L., 2022. The Role of ESG Rating Agencies and Market Efficiency in Europe's Climate Policy. Hastings Environ. Law J. 28, 139. Retrieved 2022, from https://repository.uchastings.edu/cgi/viewcontent.cgi?article=1620&;context=hastings_environmental_law_journal.

Schowengerdt, R.A., 2007. 7.2.3 Resampling. In: Remote Sensing: Models and methods for Image Processing. Elsevier, New Delhi. https://doi.org/10.1016/B978-012369407-2/50010-3.

Sipiczki, Agnes, 2022. A critical look at the ESG market. In: CEPS Papers, 36111. Centre for European Policy Studies.

Sola, J., Sevilla, J., 1997. Importance of input data normalization for the application of neural networks to complex industrial problems. IEEE Trans. Nucl. Sci. 44 (3), 1464–1468. https://doi.org/10.1109/23.589532.

Solomon, S., Portmann, R.W., Sanders, R.W., Daniel, J.S., Madsen, W., Bartram, B., Dutton, E.G., 1999. On the role of nitrogen dioxide in the absorption of solar radiation. J. Geophys. Res.: Atmos. 104 (D10), 12047–12058. https://doi.org/10.1029/1999jd900035.

SCHEER (Scientific Committee on Health, Environmental and Emerging Risks) (2021, June 29). SCHEER review of the JRC report on Technical assessment of nuclear energy with respect to the do no significant harm criteria of Regulation (EU) 2020/852 (Taxonomy Regulation) (Rep.). Retrieved 2022, from https://ec.europa.eu/i

nfo/sites/default/files/business_economy_euro/banking_and_finance/documents/210629-nuclear-energy-jrc-review-scheer-report_en.pdf.
Spinaci, S. (2022), 'EU Taxonomy: Delegated Acts on Climate, and Nuclear and Gas', European Parliament Think Tank, https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI.
Thornton, L., Khan, S., Cabral, S., Stutzman, M., 2022. The Current State of Environment Disclosure in Corporate America: Assessing What Data Russell 1000 Companies Publicly Share (Rep.). Retrieved 2022, from https://justcapital.com/wp-content/uploads/2022/04/JUST-Capital_Environment-State-of-Disclosure-Report_2022.pdf.
Yang, J., Gong, P., Fu, R., Zhang, M., Chen, J., Liang, S., Dickinson, R., 2013. The role of Satellite Remote Sensing in climate change studies. Nature Climate Change 3 (10), 875–883. https://doi.org/10.1038/nclimate1908.
Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 1992. Numerical recipes in C: the art of scientific computing, (2nd ed.). Cambridge University Press, New York, NY, USA, pp. 123–128.
Pytorch. PyTorch. (n.d.).https://pytorch.org/features/.
Yang, Z., Broby, D., 2020. Sustainable Finance: AI Applications in Satellite Imagery and Data. Discussion paper, University of Strathclyde. [Online]. Retrieved: https://strathprints.strath.ac.uk/id/eprint/73828.