UNIVERSITY OF TWENTE.

Optimising Sustainable Development in Europe: DEA-Based Efficiency Assessment with SME Integration

by

Ward de Lange

A thesis submitted to the
Faculty of Electrical Engineering, Mathematics and Computer Science (EEMCS)
in partial fulfilment of the requirements for the degree of

MSc in Business Information Technology

Faculty of Electrical Engineering, Mathematics and Computer Science (EEMCS)

University of Twente

Enschede, Overijssel, The Netherlands

July 2025

Abstract

Sustainable Development Goal (SDG) are central to the European Union (EU)'s policy framework, yet assessing progress remains complex due to the multidimensional nature of SDG data and national disparities. This thesis investigates the efficiency of EU member states in achieving SDGs 8 (Decent Work and Economic Growth), 9 (Industry, Innovation, and Infrastructure), and 12 (Responsible Consumption and Production) by applying a Slacks-Based Measure (SBM) Data Evelopment Analysis (DEA) model under Variable Return to Scale (VRS) asssumption. The study evaluates how efficiently countries convert inputs such as resource consumption and emissions into desirable outputs like waste recycling and renewable energy capacity. A key contribution is the integration of Small and Medium-sized Enterprises (SME) data into the analysis, exploring their role in shaping SDG performance. The research also applies Findable, Accessible, Interoperable, and Reusable (FAIR) data principles to assess the quality and usability of the United Nations (UN) SDG database. Spearman rank analysis is employed to assess the relationship between SMEs and SDG indicators. Findings reveal significant disparities in efficiency across EU countries, highlight specific areas for improvement through input-output slack analysis, and show the relevance of SMEs in achieving SDG targets. This study contributes a novel approach to SDG assessment by combining DEA with SME integration, offering actionable insights for policy-makers to enhance SDG performance across the EU.

AUTHOR'S DECLARATION

I hereby declare that this thesis consists of original work of which I have authored. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I authorize the University of Twente to lend this thesis to other institutions or individuals for the purpose of scholarly research. I further authorize University of Twente to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research. I understand that my thesis will be made electronically available to the public.

Ward de Lange

ACKNOWLEDGEMENTS

I would like to sincerely thank my supervisors, Dr. Marcos Machado and Dr. Patrica Rogetzer, for their invaluable guidance, support, and feedback throughout the development of this thesis.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

Abbreviations

BCC Banker,Charnes and Cooper.

CCR Charnes, Cooper and Rhodes.

CO_2 Carbon Dioxide.

CRISP-DM Cross-Industry Standard Process For Data Mining.

CRS Constant Return to Scale.

DEA Data Evelopment Analysis.

DMU Decision-Making Units.

EU European Union.

FAIR Findable, Accessible, Interoperable, and Reusable.

GDP Gross Domestic Product.

R&D Research and Development.

SBM Slacks-Based Measure.

SDG Sustainable Development Goal.

SLR Structured Litertature Review.

SME Small and Medium-sized Enterprises.

UN United Nations.

VRS Variable Return to Scale.

1

INTRODUCTION

Sustainable development has become a global priority as nations strive to balance economic growth, environmental protection, and social well-being. The UN SDGs offer a structured framework to address these challenges by 2030 [4]. In the EU, the pursuit of these goals is integrated into the broader policy landscape, including the European Green Deal [5] and national sustainability strategies. However, assessing the progress towards the SDGs remains challenging due to the complexity and variability of data across regions. Accurate assessments require analytical methods capable of managing diverse indicators across differently-sized countries.

Among the methodologies available for performance assessment, DEA [6, 7] stands out as a non-parametric technique which is widely used in various fields including sustainability and policy evaluation Liu et al. [8]. DEA enables the comparison of multiple Decision-Making Units (DMU)s, such as countries, by examining how efficiently they convert inputs like emission and resource use into outputs such as innovation and economic growth. DEAs applicability to sustainability assessments has led to a surge in its use across different studies focused on energy use [9], environmental impact [10], economic productivity [11], and well-being Ehrenstein et al. [12]. Many studies have utilised DEA to compare countries [13], regions [14] and organisations [15] based on their resource efficiency and progress towards SDG targets. While some research focuses on individual SDGs [16, 17], others attempt to construct overarching indices integrating multiple sustainability dimensions [18, 19]. Yet there are considerable variations in the DEA models used, the choice of performance indicators, and the scope of SDGs considered.

A particularly underexplored aspect of the SDG efficiency in the EU is the role of SMEs. While SMEs constitute about 99% of all EU businesses and account for a substantial share of employment and innovation [20], their impact on sustainability is largely overlooked in current research.

This thesis aims to address this gap by applying DEA to assess the efficiency of EU countries in achieving SDGs 8 (Decent Work and Economic Growth), 9 (Industry, Innovation, and Infrastructure), and 12 (Responsible Production and Consumption), while integrating SME into the analysis. Moreover, the study evaluates the FAIRness of the primary dataset to ensure data quality and usability. The study is guided by the following main research question:

How can DEA-based efficiency analysis of SDG implementation in Europe be enhanced by considering the role of SMEs in shaping key sustainability indicators?

Supporting sub-research questions include:

1. How can DEA be applied to quantify SDG efficiency for EU countries?
2. What are the key inefficiencies for EU countries' progress towards SDGs 8, 9 and 12?
3. What is the impact of SMEs on the SDG indicators that drive DEA efficiency outcomes in EU countries?

This research makes a scientific contribution by combining DEA modelling with SME integration and FAIR data principles in the context of SDG assessment. By focusing on SDGs 8, 9, and 12, the study targets goals that are directly influenced by enterprise activity and innovation. Additionally, the study assesses the quality and accessibility of the UN's SDG data through a FAIR data lens and employs Spearman rank correlation to investigate the statistical association between SME numbers and SDG indicators.

From a practical standpoint, the research delivers actionable findings for policymakers and sustainability professionals. The DEA model reveals relative efficiency scores across EU member states and identifies specific input-output slacks, providing guidance where specific improvements can be made. By quantifying the correlation of SMEs on SDG indicators, the study offers an additional perspective for sustainability strategies. Furthermore, the FAIR assessment identifies data gaps and opportunities for improvement in dataset usability.

The structure of this thesis is as follows: Chapter 2 presents the SLR, detailing existing work on DEA in the context of SDG assessments and identifying research gaps. Chapter 3 outlines the research design and analytical methods used, including the Cross-Industry Standard Process For Data Mining (CRISP-DM) framework, DEA modelling strategy, and preparation techniques. Chapter 4 describes the experimental set-up, including data collection, preparation, and modelling steps. Chapter 5 presents the results of the FAIR data assessment, DEA efficiency analysis, and SME correlations studies, followed by the discussion of findings, limitations, and recommendations for future research. Chapter 6 concludes the thesis with key insights and answers to the research questions.

2

Literature Review

2.1. SLR Methodology

The methodology for this SLR is based on the framework proposed by Varsha P S et al. [1], which outlines eight essential steps for conducting a structured and reproducible SLR. Figure 2.1 illustrates these steps, which provide a robust foundation for addressing the following knowledge questions related to SRQ 1:

- What are key performance indicators relevant to assessing SDG efficiency?
- Which DEA models are best suited for evaluating SDG implementation efficiency?

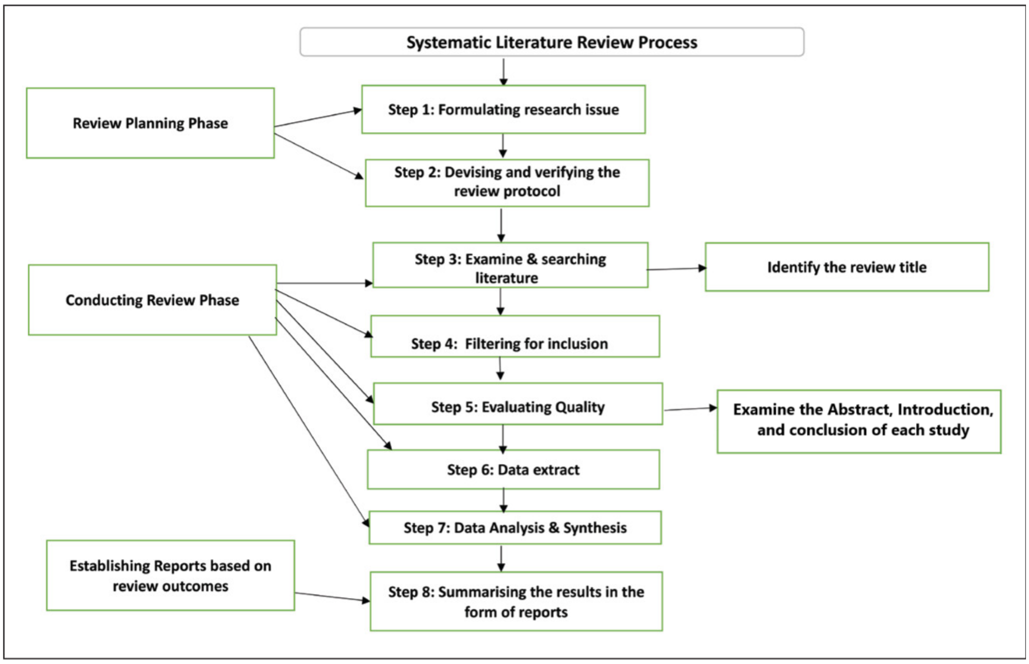


Figure 2.1: The SLR process stages adapted from [1]

The first step involves formulating the research issue and corresponding questions. These questions, specified in Chapter 1, guided the subsequent development of a systematic methodology for identifying, screening, and analysing relevant literature. The second step concentrates on the creation and validation of the research methodology that will be used for the search and analysis of relevant data/research papers, which is done in this section.

Step three involves examining and searching for relevant literature, for this study that was done using the Scopus literature database [1], selected for its extensive coverage of multidisciplinary research and its advanced search functionalities. The search was conducted using search queries designed to capture variations in terminology related to the research questions, targeting words found in the title, abstract, or keywords:

- (”Data Envelopment Analysis” OR ”DEA ” OR ”efficiency metric” OR ”efficiency indicator” OR ”efficacy metric” OR ”efficacy indicator”) AND (”environmental metrics” OR ”social development indicators” OR ”economic growth measures” OR ”environmental indicator” OR ”social development metrics” OR ”social development indicator”)
- (”Data Envelopment Analysis” OR ”DEA” OR ”efficiency analysis” OR ”efficacy analysis”) AND (”Sustainable Development Goals” OR ”SDG” OR ”global goals”)

These queries yielded a total of 407 results. Following this broad initial selection, steps 4 and 5 apply a filtering process to refine the literature set. Figure 2.2 shows the number of papers filtered out by each phase as well as the number of papers left.

Step four involves filtering these results based on inclusion/exclusion criteria. This process begins by assessing relevance, selecting only papers published in journals and conference proceedings after 2020. The rest of the criteria used for inclusion and exclusion are summarised in Table 2.1, ensuring only relevant and accessible studies are analysed.

Table 2.1: article inclusion/exclusion criteria.

Criteria	**Decision**
Inclusion of pre-defined keywords in title, abstract, or keyword list	Inclusion
Article published in a scientific journal or conference	Inclusion
Article written in English	Inclusion
Article published before 2020	Exclusion
Duplicates of an original article	Exclusion
Irrelevant subject area	Exclusion
Unavailability of the article online for free	Exclusion

[1] https://www.scopus.com/home.uri

The fifth step focuses on evaluating the quality and relevance of selected articles, which essentially takes place in the last phase of the paper selection. This phase involves a systematic manual screening of the articles by reading the title, abstract, introduction, and conclusion of each paper and checking if they focussed on at least two of the following main subjects: any form of DEA, at least one SDG, and Europe. The full selection process is summarized in Figure 2.2

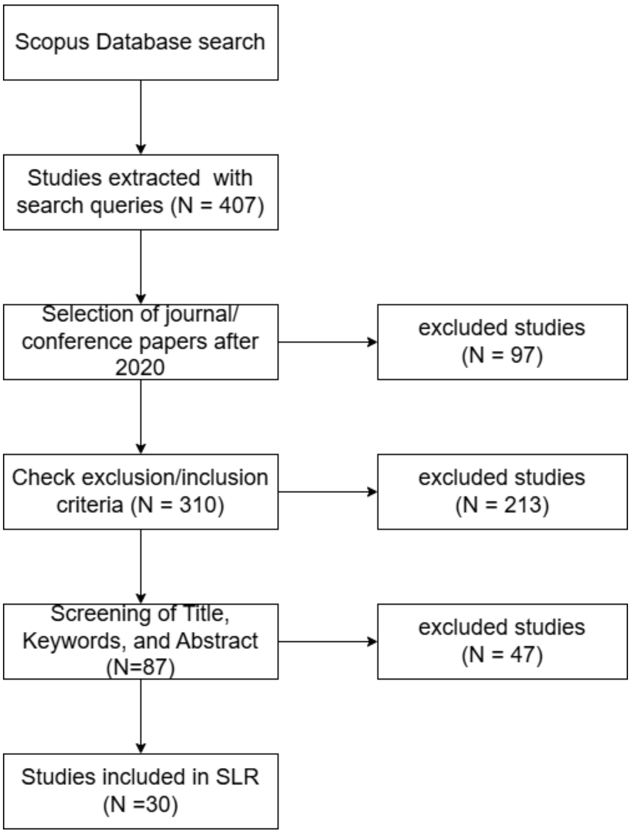


Figure 2.2: Illustration of the article selection process

After the relevant papers are selected and screened, we start step 6, which refers to methodically extracting the relevant data from each selected paper. Table A.1 in the appendix shows the extracted data: main focus, SDG focus, analysis method used, dataset, input variables, and output variables used for the DEA analysis for each paper. The seventh step, data analysis and synthesis, allows for integration of findings and identify new insights. The last step is to condense it all into a comprehensive report that shows the findings of this study in a concise and clear manner.

These last steps serve as a basis for the summarization and discussion of the papers presented in Sections 2.2, 2.3, and 2.4 which discuss a quantitative analysis of the literature, the themes in the findings of the literature, and the conclusions of this study.

2.2. DECONSTRUCTION OF LITERATURE LANDSCAPE

This section provides an overview of the collected research papers by analysing quantitative data. The goal is to identify publication trends, shifts in academic focus, and growth in this research field. The analysis is divided into three key subsections:

- **Temporal distribution** - Illustrates how contributions in this research field have evolved over time. Temporal analysis reveals publication trends, shifts in research priorities, and historical context for the development of this research domain.
- **Journal distribution** - This subsection examines the distribution of journals where relevant literature has been published, which highlights the most influential and significant contributors to the topic, as well as the multidisciplinary interest in the field.
- **Subject area distribution** - The last subsection analyses the subject areas of the selected studies, identifying which disciplines are most engaged with this research topic. This may provide insight into which SDG receives the most attention and which areas remain underexplored

For transparency and clarity, all articles included in this study are indexed in Table A.1. This table provides a comprehensive summary of key attributes such as the primary focus of the study, the SDG it addresses, the type of DEA model used, the dataset applied, and the input/output features examined.

2.2.1. TEMPORAL DISTRIBUTION

Figure 2.3 illustrates the number of publications per year, highlighting how interest in this research domain has evolved over the last five years.

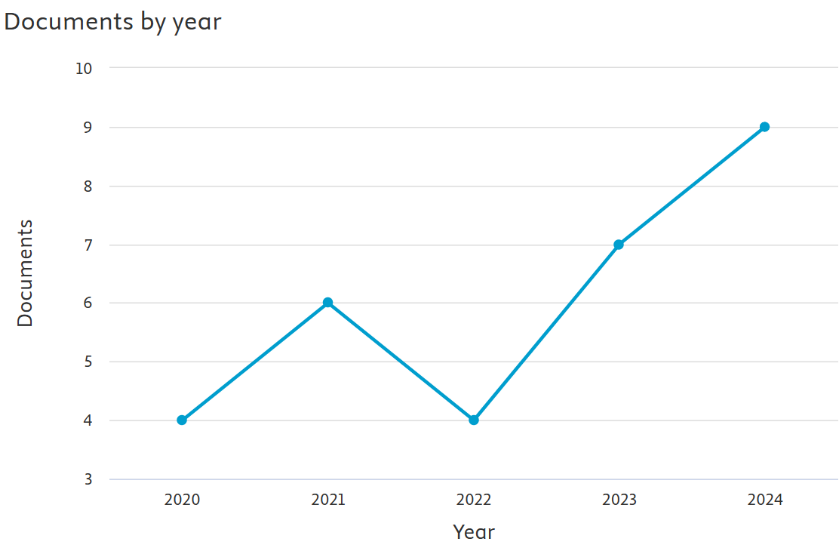


Figure 2.3: The research landscape over the years

The data reveals a steady increase in research activity, with a notable rise in publications in recent years. However, there is a slight decline in 2022, which may be attributed to shifting research priorities during that period. No relevant studies have been published in 2025 so far, which can be attributed to the search being conducted in January 2025. Overall, this historical trend highlights the growing interest in SDG progress analysis using DEA methods. The upward trajectory in publications suggests a strong potential for future research in this domain, reinforcing the importance of ongoing academic contributions and highlighting the need for an SLR on the topic.

2.2.2. JOURNAL DISTRIBUTION

Figure 2.4 presents the distribution of journals publishing research on this topic. The analysis reveals that 27% of the selected papers were published in Sustainability Switzerland[2], emphasising its prominence as a leading outlet for this research. Other significant contributors include Sustainable Development[3] and the International Journal of Energy Economics and Policy[4], which collectively account for another 17% of publications. Beyond these key journals, research on this topic is dispersed across a variety of other publications, indicating a broad interdisciplinary interest. This distribution suggests that while sustainability remains a central focus, the field also attracts contributions from diverse academic disciplines, reflecting its multidisciplinary nature and widespread relevance.

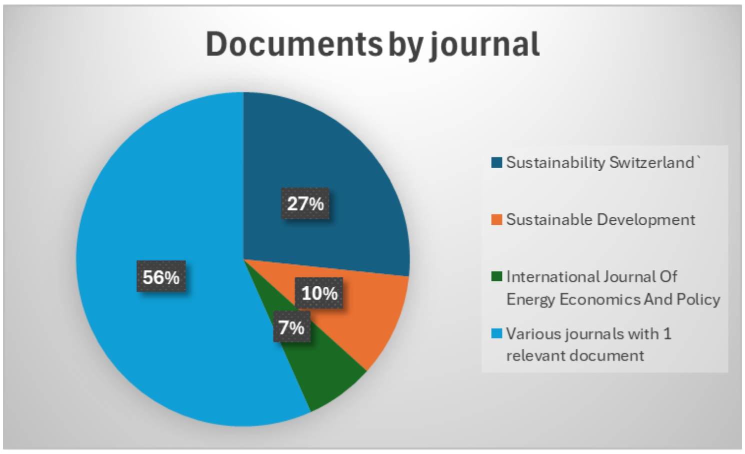


Figure 2.4: Document publication distribution

2.2.3. SUBJECT AREA DISTRIBUTION

Figure 2.5 provides an overview of the primary subject areas of the selected studies. It shows that majority of research falls within energy, social sciences, and environmental sciences, highlighting a strong alignment with sustainability-related topics. This distribution suggests that research in this domain primarily focuses on specific SDGs, mainly:

- **SDG 7 (Affordable and Clean Energy)** – due to the high number of studies focused on the

[2] https://www.mdpi.com/
[3] https://onlinelibrary.wiley.com/journal/10991719
[4] https://econjournals.com/index.php/ijeep/index

energy domain Maya et al. [9], Alam [21], Rusydiana et al. [22].

- **SDG 1 (No Poverty)** – given the emphasis on social studies and financial indicators Roy et al. [23], Wang et al. [24].
- **SDG 13 (Climate Action)** – reflecting the research focus on environmental sciences Ehrenstein et al. [12], Al-Ayouty and Hassaballa [25].

The dominance of these subject areas suggests that while there is considerable research in energy and environmental sustainability, other SDGs may be comparatively underrepresented. More research may be needed to explore how DEA methods can be applied to analyse other SDGs.

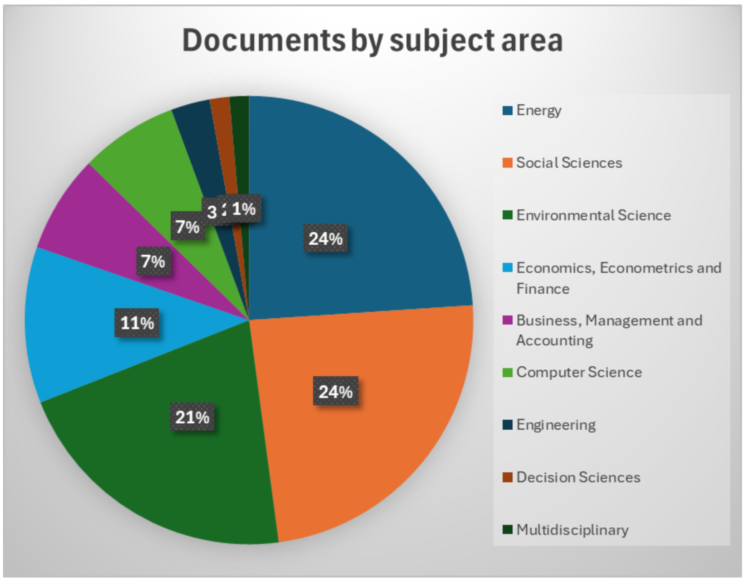


Figure 2.5: Distribution of papers by subject area

By analysing the temporal, journal, and subject area distributions, this study provides a comprehensive overview of the research landscape. The findings illustrate a growing academic interest in using DEA methods to analyse SDG progress, with sustainability and energy-related topics emerging as dominant themes. This analysis also highlights the potential for expanding research into additional SDGs, ensuring a more holistic understanding of global sustainability challenges.

2.3. PREDOMINANT THEMES IN THE LITERATURE

The reviewed literature provides insights into the application of DEA for assessing SDG progress. This section presents qualitative findings from these studies regarding the research questions stated in Section 1. To maintain clarity, this section is divided into three subsections. The first examines the performance indicators used to assess the efficiency of various SDGs, the second focuses on different DEA methods applied for efficiency assessments of SDG progress, and the third explores the relation of SME-related data with DEA-SDG assessments.

2.3.1. PERFORMANCE INDICATORS USED FOR SDG PROGRESS ASSESSMENT

Performance indicators are quantifiable measures used to evaluate progress toward achieving specific goals[5]. In the context of the implementation of the SDGs, these indicators serve as benchmarks to evaluate the national and regional progress toward each goal. Effective performance indicators should be relevant, measurable, and aligned with the UN SDG targets [6]. Table A.2 shows a comprehensive list of all the SDGs and a description of their specific goals. To be able to effectively compare efficiency within a group, whether it be cities, provinces, regions, or entire countries, common indicators need to be found that are available for all the members of the group.

The UN maintains an official list[7] of indicators for each goal, updated annually. The current list comprises 231 distinct indicators, some applicable for multiple goals, such as installed renewable energy-generating capacity (Watts per capita), which is used for SDG 7 and SDG 12. However, not all countries have data for every indicator, complicating cross-group comparisons. To address this, researchers typically select a subset of indicators relevant to their target group, sometimes supplementing them with additional indicators.
Despite variations in study focus, SDG performance indicators generally fall into three broad categories Hu and Cheng [19], Roy et al. [26]:

- **Environmental Indicators** - Measures that mostly reflect overall environmental protection and sustainability progress, such as CO_2 emissions per capita (metric tons), renewable energy share in total energy consumption (%), and biodiversity loss %. These indicators primarily assess the progress of SDG 7 (affordable clean energy) and SDG 13 (Climate action).
- **Social Indicators** - Measures related to education, healthcare, poverty, and equality, including literacy rates (% of population over 15 years old), life expectancy (years), and gender equality indices. These are commonly linked to SDG 1 (no poverty), SDG 3 (Good health and well-being), and SDG 4 (quality education).
- **Economic Indicators** - Metrics assessing economic performance such as GDP (USD) , unemployment rates (%), labour productivity (GDP per hour worked) and innovation indices. These indicators are essential to monitoring SDG 8 (decent work and economic growth) and SDG 9 (industry innovation and infrastructure).

The reviewed literature highlights significant variability in performance indicator selection based on the availability of common indicators for the group of countries or regions being investigated and the specific focus of the research [9, 25]. Researchers often use a combination of UN and custom indicators [11, 27] to better fit their study objective. For example, D'Adamo et al. [27] use the recycling rate (UN indicator) and circular material use (custom indicator) to assess

[5] kpi.org/
[6] sdgresources.relx.com
[7] unstats.un.org

the efficiency of recycling in Europe in relation to SDG 12. The following breakdown shows common indicators grouped by SDG:

- **SDG 1 (No Poverty)** - Studies by Roy et al. [23], Wang et al. [24] evaluate poverty but use completely different indicators. Roy et al. [23] includes education (mean years of schooling for adult population), gender equality (female labour force % of total), and food supply (kcal per person per day) due to it's socioeconomic development focus, while Wang et al. [24] incorporates agricultural GDP (RMP) and crop-sown area (hectares) as well as infrastructure indicators such as internet penetration (%) and road area per capita (metre) to emphasize rural development.

- **SDG 3 (Good Health and Well-being)** - Chien et al. [16] employs social protection benefits (million €), healthcare expenditure (million €), and population, which, while not listed by the UN, provide insights into health-related SDG efficiency.

- **SDG 4 (Quality Education)** - Studies by Matulová [28], Perović and Kosor [29] focus on university rankings to compare SDG 4 progress between counties by using different metrics. Matulová [28] applies 39 sustainably-related metrics, while Perović and Kosor [29] uses tertiary expenditure (% GDP) and amount of tertiary teachers per capita.

- **SDG 7 (Affordable and Clean Energy)** - Maya et al. [9], Huang et al. [18], Alam [21], Rusydiana et al. [22] all assess energy efficiency in different regions using GDP (USD), emissions (metric tons), labour force (number of people), and total energy consumption (TWh or coe).

- **SDG 8 (Decent Work and Economic Growth)** - Ibujés-Villacís and Franco-Crespo [11], Blanco et al. [30], Singpai and Wu [31] use varying approaches, Blanco et al. [30] focuses solely on graduate employability (# graduates employed), Ibujés-Villacís and Franco-Crespo [11] examines salaries (% of spending) and profit, and Singpai and Wu [31] adheres strictly to official UN indicators.

- **SDG 9 (Industry, Innovation, and Infrastructure)** - Studies by Mura et al. [15], Wang et al. [24], Al-Ayouty and Hassaballa [25], Pishdar et al. [32] primarily use economic related indicators for SDG assessment. Mura et al. [15] made use of total assets (USD) and operating costs (USD) as indicators to evaluate the SDG commitments of companies. While Pishdar et al. [32] employed revenue (USD per year) and capital for their benchmarking tool. Furthermore, Wang et al. [24] used income per rural capita (RMP) and government financial support for agriculture (RMP) to assess sustainable rural development in China. Lastly, Al-Ayouty and Hassaballa [25] implemented employees per year and value of energy output as indicators of the environmental efficiency of Egyptian industries.

- **SDG 11 (Sustainable Cities and Communities)** - Maricut et al. [17], Li et al. [33] both focus on urban development with slightly different approaches. Maricut et al. [17] includes urban area per capita (m^2), employment rate (%), share of green areas in urban centres (%). While Li et al. [33] emphasises urban innovations through amount of R&D

investment, personnel (# persons), and patents granted (# patents).

- **SDG 12 (Responsible Consumption and Production)** - D'Adamo et al. [27] focuses on circular economy efficiency, using indicators related to investment in Circular Economy (CE), employment rates in CE sectors (% of total employment), recycle rate (% of total waste), and circular material use (% total material use).

- **Overall SDG Assessments** - Some studies create overarching indices to measure multiple or all the SDGs collectively. Such studies by Blancas and Contreras [34], Issever Grochová and Litzman [35], Cristóbal et al. [36] strictly use UN indicators that their entire research group has data for, while Li et al. [14], Hu and Cheng [19] employ custom environmental, social, and economic measures to develop their indices.

The literature suggests that most studies that focus on a single or a few related SDGs select a mix of UN and custom indicators based on data availability and research scope [21, 37]. Studies assessing multiple SDGs tend to rely on a subgroup of available UN indicators for their target groups [34, 35]. A recurring challenge is the lack of consistent data across different countries, leading researchers to tailor their indicator selection based on accessibility and relevance.

2.3.2. USE OF DEA IN SDG EFFICIENCY ASSESSMENTS

Data Evelopment Analysis (DEA) is a non-parametric linear programming technique used to evaluate the efficiency of a DMU, such as a country or region, in converting multiple inputs into desirable outputs [38]. DEA is particularly useful in contexts where multiple input and output variables exist.

TYPES OF DEA

DEA models can be categorised on their orientation and return to scale assumptions. An input-oriented model seeks to minimise inputs while maintaining a constant level of outputs. Alternatively, output-oriented DEA models aim to optimise outputs while keeping the input variables stable [27]. Two of the most commonly used DEA models are the Constant Return to Scale (CRS) and VRS models.

The first is the CRS model introduced by Charnes et al. [6] (Charnes, Cooper and Rhodes (CCR) model) which assumes any proportional increase in the inputs leads to a proportional increase or decrease in the outputs. Mathematically the efficiency of a DMU under CRS is expresses as:

$$max\ h_0 = \frac{\sum_{r=1}^{s} u_r y_{r0}}{\sum_{i=1}^{m} v_i x_{i0}} \tag{2.1}$$

subject to

$$\frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \leq 1\ ;\ j = 1, ..., n \tag{2.2}$$

$$u_r, v_i \geq 0\ ;\ r = 1, ..., s\ ; i = 1, ..., m. \tag{2.3}$$

where y_{i0}, x_{ij} are the inputs and outputs of the jth DMU, u_r, v_i are the weights of the output r

and input i, n is the number of DMU, s is the number of outputs and m is the number of inputs. h_0 represents the efficiency score, which falls between 0 and 1, where 1 indicates most efficient

The VRS model, introduced by Banker [7] (Banker,Charnes and Cooper (BCC) model), extends the CRS model by accounting for economies of scale. This means that a DMU should not be directly compared to a much bigger or smaller DMU. The BCC model modifies CCR equation 2 by adding a convexity constraint:

$$\sum_{j=1}^{n} \lambda_j = 1 \tag{2.4}$$

Where λ_j represents the weight assigned to compare each peer DMU in the set, this constraint ensures that a DMU is only compared to other DMUs of similar size, preventing efficiency scores from being influenced by significant differences in scale. As a result, the BCC model is more suitable for assessing efficiency among DMUs of different sizes, such as region or countries with vastly different economic scales.

Beyond the CCR and BCC model, the SBM model, introduced by Tone [39], is frequently used in DEA efficiency comparisons. Unlike the previous models, SBM accounts for slack in inputs and outputs, removing the assumption that inputs or outputs can always be adjusted proportionally.
Lastly, some studies develop customized DEA models to better match research needs by incorporating novel variations to the standard models. These include network DEA models, Three-stage DEA, or MILP DEA

DEA APPLICATIONS

The literature suggests no single DEA model is universally optimal for SDG efficiency assessments [12, 13]. However, different models tend to be preferred based on the type of SDG being analysed, scale of the DMUs, and chosen performance indicators. Table A.1 provides a detailed breakdown of all the literature reviewed in this SLR, specifying the DEA model used in each study, the SDGs addressed, the datasets utilised, and the input/output indicators applied in the DEA analysis. The following section explores emerging patterns in the literature.

The CCR model is most used in studies where input-output relations are more linear. For example, Alam [21] applied CCR-DEA to SDG 7, using energy consumption, labour, and population as inputs and CO_2 emission and GDP as outputs. Similarly, studies such as [11, 15, 25] employed this model to assess industrial and economic efficiency, where capital and costs serve as inputs and the output is added value. In the case of SDG 12, D'Adamo et al. [27] uses CCR-DEA to analyse circular economy efficiency by incorporating investments and employment rates as inputs and recycling rates and circular material use as outputs.

In contrast, the BCC model is more frequently applied to SDGs where efficiency is influenced by resource availability. For instance, Roy et al. [23] used BCC-DEA to assess SDG 1 by analysing globalisation efficiency through the availability of energy, food, and water sanitation. Blanco et al. [30] applied this model to SDG 8, evaluating graduate employability based on student enrolment and teaching staff. Additionally, Maricut et al. [17] used BCC-DEA for SDG 11 to assess sustainable urban development by considering the proportion of green spaces in urban areas.

The SBM model is proven particularly effective in multidimensional SDG performance assessments. Studies such as [13, 19, 36] employed SBM-DEA for composite SDG indices incorporating economic, social, and environmental dimensions. This model is useful when efficiency assessments require measuring slack variables to provide deeper insights. Additionally, SBM-DEA has been applied to individual SDGs, such as SDG 9, where Pishdar et al. [32] uses a network SBM model to compare DMU, based on their circular economy and sustainability performance, facilitating rankings beyond just simple efficiency scores.

Various custom models have been proposed for more complex SDG analysis, for example, Chien et al. [16] developed a three-stage network DEA model to be able to assess DEA advancements over time, Blancas and Contreras [34] used a customised DEA model that aggregates indicators into composite rankings, making comparisons between DMUs more straightforward. Ehrenstein et al. [12] introduces a Mixed Integer Linear Programming (MILP) DEA model to identify which DMU is closest to a specific target level.

The literature suggests that different models are suited to specific types of SDG analyses. CCR-DEA is most effective for energy, industry, and production-related SDGs, where input-output relations are largely proportional [15, 21, 27]. BCC-DEA is particularly useful for social and economic SDGs, where efficiency depends on resource availability and differences in DMU scale [17, 23, 30]. SBM-DEA is well-suited for multidimensional assessments requiring flexible slack adjustments [19, 32, 36]. Finally, customised DEA models provide enhanced insights for complex assessments involving interconnected indicators [12, 16, 34]. Ultimately, selecting the appropriate DEA model depends on the specific SDG being analysed, the nature of the input/output variables, and the research objectives. The literature highlights that while DEA is a powerful tool for evaluating SDG efficiency, careful model selection is essential to ensure accurate and meaningful assessments.

2.3.3. CONNECTING DEA-SDG ASSESSMENT WITH SMES

A SME is a business that falls below a defined threshold for employees, revenues, and assets, which varies by country and industry [40]. The EU classifies SMEs as firms with fewer than 250 employees and either a turnover below €50 million turnover or a balance sheet total below €43 million [41]. SMEs are the backbone of most economies, accounting for more than 90% of businesses worldwide and generating over 50% of global employment [42]. Improving their ability to operate efficiently and sustainably is crucial to achieving SDG targets.

As SDG implementation requires efficient resource utilisation across economic sectors, assessing the contribution of SMEs to sustainability efforts is essential. SME contribute to multiple SDG by fostering employment (SDG 8), driving innovation (SDG 9), and promoting responsible production (SDG 12) [40]. While this study shows that DEA is widely used to evaluate national and regional SDG efficiency, few studies have explored how SMEs contribute to these efficiencies. Given their role in the economy, integrating SME-specific data into DEA-SDG assessments could provide a new framework for evaluating their sustainability impact. This approach could help SMEs identify best practices, optimise resource allocation, and improve their contribution to the SDGs.

One relevant study from the SLR, Ibujés-Villacís and Franco-Crespo [11], applies DEA to measure the efficiency of the manufacturing industry in Equador concerning SDG 8, 9, and 12, using SME-related data as part of the inputs. The study evaluates how factors such as profits, wages, and assets influence sustainable industrial growth and the differences between SMEs and large enterprises. Their findings suggest that SMEs are more efficient in production, but that the influence on SDGs grows with the size of the company.

To further investigate the intersection of DEA-SDG assessments and SME efficiency, another literature search was conducted across Scopus[8], IEEE Xplore [9], and Web of Science[10] using the following query:

- (”Data Envelopment Analysis” OR ”DEA” OR ”efficiency analysis” OR ”efficacy analysis”) AND (”Sustainable Development Goals” OR ”SDG” OR ”global goals”) AND (”Small and Medium Enterprise” OR ”SME” OR ”SMEs”)

Despite searching across multiple databases, no additional studies were found that link SME performance data to DEA-based SDG assessments. This finding highlights a critical research gap. While DEA is widely used in national and regional SDG efficiency analysis, the addition of SME-related data remains largely unexplored. Addressing this gap could provide insight into how SMEs contribute to SDG implementation and how they can improve their efficiency in achieving sustainability goals.

2.3.4. FAIRIFICATION OF DATA IN EU-BASED DATASETS

The integration of FAIR principles into EU-based datasets has gained growing attention in recent years, largely driven by the European Commission’s agenda for open science and data-driven policymaking [43]. A key focus of initiatives like the European Open Science Cloud (EOSC)[11] is the FAIRification of research and statistical datasets. This process reflects the growing need for FAIR data infrastructures for responsible data reuse and long-term reusability of research outputs.

[8] https://www.scopus.com
[9] https://ieeexplore.ieee.org
[10] https://www.webofscience.com
[11] https://research-and-innovation.ec.europa.eu/strategy/strategy-research-and-innovation/our-digital-future/open-science/european-open-science-cloud-eosc_en

Despite substantial EU investment in FAIR-related tools, such as FAIRsFAIR[12] and FAIR-IMPACT[13], evaluations of the FAIRness of core EU datasets remain scarce. Although multiple FAIR assessment tools are available, there is little evidence of systematic studies focused on FAIRness of the databases managed by Eurostat and national statistical offices. This lack of analysis presents a clear research gap. As the EU continues to advocate for open and interoperable research data, future studies should evaluate FAIRness of existing datasets and identify steps to bridge any gaps.

2.4. FINAL CONSIDERATION OF THE LITERATURE REVIEW

This study aimed to perform a SLR to explore the application of Data Evelopment Analysis (DEA) in evaluating sustainable development metrics related to Sustainable Development Goal (SDG) progress. The methodology followed a structured framework to identify relevant research, extract key insights, and analyse DEA models employed in assessing SDG efficiency. Through a comprehensive review of existing literature, this study identified relevant performance indicators and assessed their applicability across various DEA models to evaluate sustainability efforts effectively. The structured approach provides the ability to address the research questions asked at the start of this review:

- **Key Sustainable Development Goal (SDG) performance indicators relevant to assessing sustainability efficiency**
  The research identified that SDG assessments rely on a combination of UN-prescribed indicators and custom metrics, largely dictated by the availability of shared indicators among the target research group and specific research focus. Studies focusing on a single or small set of SDGs tend to use custom indicators supplemented with some UN indicators, while broader SDG index studies primarily select a subgroup of UN indicators that are available for their entire target group.

- **Best suited DEA models for evaluating SDG progress efficiency**
  The analysis suggests that DEA model selection depends on the specific SDG being analysed and the aim of the individual study. The CCR model is most suited for SDGs with linear input-output relationships, such as energy and industrial efficiency. The BCC model is better suited to social and economic SDGs, where variations in resource availability and regional differences affect efficiency outcomes. The SBM model is particularly effective for multidimensional SDG assessments, as it accounts for slack in various indicators. Custom DEA models incorporating network analysis and ranking methodologies are emerging to improve over-time assessments and comparative evaluations.

This literature review contributes by systematically reviewing DEA applications in SDG assessments, identifying the most relevant performance indicators, and categorising DEA methodologies based on their effectiveness for different SDGs. By providing a structured framework,

[12] https://www.fairsfair.eu/

[13] https://fair-impact.eu/

this research offers practical guidance for policymakers, researchers, and industry professionals seeking to optimise sustainability assessments. Additionally, this study highlights the growing role of network-based DEA and ranking approaches in improving SDG efficiency evaluations, offering insights into emerging analytical approaches.

2.4.1. LIMITATIONS AND FUTURE RESEARCH FOR THE LITERATURE REVIEW

While this SLR provides a comprehensive overview of DEA applications in assessing SDG efficiency, several limitations should be acknowledged. First, the review was limited to English studies indexed in Scopus, which may have excluded relevant research published in other languages or databases. In addition, accessibility constraints and predefined search terms could have inadvertently omitted studies using alternative methodologies or domain-specific terminology.

A further limitation lies in the uneven distribution of focus across SDGs. The majority of reviewed studies concentrate on energy, climate, and economic dimensions, with relatively limited attention given to goals involving social development or institutional quality. This reflects a broader trend in the literature and presents opportunities for expanding the scope of future assessments.

Notably, this review confirms that the role of SMEs in SDG efficiency remains significantly underexplored. Despite the economic importance of SMEs in the EU, there is a lack of empirical integration of SME-specific data into DEA-based sustainability evaluations. Only a limited number of studies have attempted to assess their contribution in relation to SDGs 8, 9, and 12, suggesting a clear research gap that warrants further investigation.

In summary, this literature review highlights not only the current methodological landscape but also several areas for further exploration. Future research would benefit from broader data inclusion, deeper analysis of underrepresented SDGs, and the integration of SMEs into DEA frameworks. These directions can support the development of more comprehensive and policy-relevant sustainability assessments.

3

METHODOLOGY

To assess the efficiency performance of the different EU nations in achieving SDG 8, 9, and 12 and the impact of SMEs, this study adopts a combination of data-driven methodologies and guiding frameworks. This chapter outlines the frameworks applied to ensure transparency and reproducibility throughout the research process.

The chapter is divided into three sections: the first section focuses on the CRISP-DM framework that provides an overarching structure for the analytical workflow, providing a systematic approach to data analysis tasks. The second section focuses on the technical applications of the study, outlining the data-driven techniques and models used to assess the efficiency of EU nations and the impact of SMEs on that efficiency. The final section dives into the qualitative methods that can be used to verify the results of the models. By dividing this chapter into two sections, the study aims to provide a complete overview of the research framework.

3.1. CROSS-INDUSTRY STANDARD PROCESS FOR DATA MINING

CRISP-DM is a well-established methodology for structured data analysis and data mining. Though it was originally developed for data mining in 1998 [44], CRISP-DM has since been used in all kinds of machine learning and DEA tasks [45]. The framework was chosen for its structure and compatibility with various data analysis tasks, making it suitable for guiding this multi-step SDG-SME analysis.

Figure 3.1 shows the structured approach of CRISP-DM, which consists of 6 phases that somewhat follow each other, but some back-and-forth iterations between phases are sometimes necessary. The six phases of CRISP-DM [2] are:

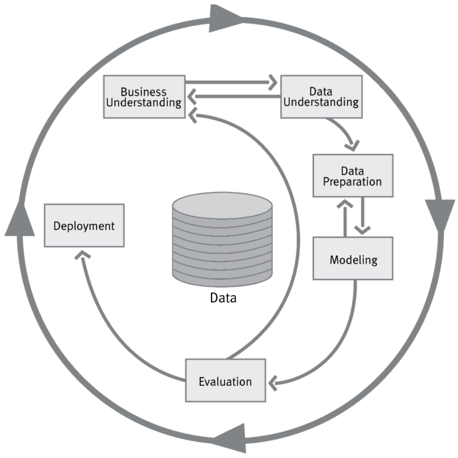


Figure 3.1: CRISP-DM Process Model (Source: [2])

1. **Business understanding**: Determine the project objectives and requirements of the analysis. Establish the project scope, overall planning, and the end goal.

2. **Data understanding**: Identify data sources, check for data completeness and quality issues. Insights from this phase can then be used in subsequent preprocessing steps.

3. **Data preparation**: Transformed the gathered data into a suitable format for the intended analysis model. This phase usually involves normalisation, missing value handling, restructuring tables, and selecting relevant features.

4. **Modelling**: Apply the selected analytical models to the final dataset obtained from the data preparation. This could be paired with careful calibration and changes to the model.

5. **Evaluation**: Assess the data analysis models for technical accuracy and validity. This stage may include comparative diagnostics and external validation to determine if the model achieved valid and useful results.

6. **Deployment**: Share the results in the preferred format agreed upon in the first phase, depending on the purpose of the project. This could include dashboards, tables, or a monitoring application.

Thus, CRISP-DM ensures a structured approach for any data-driven project that guarantees the appropriate steps are taken to ensure the proper development of an analysis.

3.2. ANALYTICAL METHODS

3.2.1. SBM-DEA

DEA is a non-parametric analysis method used to evaluate relative efficiency between multiple comparable entities, known as DMU, based on their ability to transform multiple inputs into outputs. The SBM model developed by Tone [39] is an advanced variation on the classic DEA models [6, 7] that incorporates slacks that represent inefficiencies in input use or output shortfalls. Some more key features of SBM-DEA and DEA in general are:

1. **Orientation**: The option in any DEA model is the orientation of the model. Which is either input-oriented, which focuses on minimising the inputs to reach the desired level of outputs or output-oriented, which tries to maximise the desired outputs with the given inputs. However, SBM adds the option of non-oriented analysis, where both maximising outputs and minimising inputs are possible.

2. **Return to scale**: The second parameter to take into account for DEA analysis is the Return to scale. Originally, the CCR model [6] assumed a CRS, which assumes any proportional increase in the inputs leads to a proportional increase or decrease in the outputs. Alternatively, a model with VRS as first introduced with the BCC model [7] takes economies of scale into account, where an increase in inputs might not mean an equal increase in outputs for all DMU's as not all entities operate at optimal size and capacity.

3. **Non-radial modelling**: Where the traditional CCR and BCC models always assume that all the inputs and outputs can be proportionally increased or decreased, SBM evaluates slacks, such as excesses in inputs or shortfalls in outputs, directly [46].

SBM-DEA is particularly suited for SDG indicators, where both overuse of resources and underperformance of outcomes are relevant inefficiencies.

3.2.2. FAIR

The FAIR data principles guide responsible and sustainable data management. Originally developed for the scientific community [47], these principles are increasingly applied across disciplines to ensure data integrity and transparency, while also increasing ease of usability. The core principles are as follows:

1. **Findable**: All data and metadata should be easily findable and indexed in searchable repositories. Proper naming conventions and metadata are, therefore, essential to increase discoverability.

2. **Accessible**: Data should be retrievable using open and standardised protocols. Any conditions for access, such as licenses or other limitations, should be indicated as well.

3. **Interoperable**: Data should be in a standardised format so that it can be integrated with other datasets.

4. **Reusable**: Data has to be be richly described with metadata so it can be reused by other researchers.

To accurately assess the FAIRness of a data set, specific evaluation criteria were adapted from the GO FAIR initiative's[1] interpretation of the FAIR principles. Focusing on metadata quality, use of persistent identifiers, open protocols, and standardised vocabularies.

[1] https://www.go-fair.org/fair-principles/

Table 3.1: FAIR Principles and Assessment Criteria [3]

FAIR Principle	GO FAIR Criteria
Findable	(F1) Data are assigned persistent identifiers (e.g., DOIs); (F2) data are distributed with rich metadata; (F3) metadata includes data identifiers; (F4) (Meta)data indexed in searchable resource
Accessible	(A1.1) Data retrievable via standardised open protocols (e.g., HTTPS, API);(A1.2) The protocol allows for authentication where necessary; (A2) metadata accessible even if the data are not.
Interoperable	(I1) Data use formal, accessible, and shared formats and vocabularies; (Meta)data use vocabularies that follow FAIR principles;(I3) data includes semantic references to other data.
Reusable	(R1.1)(Meta)data include detailed licensing information; (R1.2) (Meta)data associated with detail provenance;(R1.3) (Meta)data aligned with domain standards.

3.2.3. SPEARMAN RANK CORRELATION ANALYSIS

The last analysis method that is relevant for this study is the Spearman correlation coefficient [48]. It is a non-parametric measure of the strength and direction of a monotonic relationship between two variables. Spearman Correlation works by:

1. Ranking the data from each variable individually.
2. Computing the difference in ranks between the variables of each observation.
3. applying equation 3.1 to determine the correlation (ρ) between the chosen variables.

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)} \tag{3.1}$$

Where d_i is the difference in rank between the variables of each observation and n is the number of observations. The resulting coefficient ρ ranges from -1 to 1, indicating perfect negative to perfect positive association. A ρ of 0 means that the variables have no effect on each other. // This method is more appropriate than Pearson or Kendall for this study because it does not assume linearity or normality in the data distribution, both of which are often violated in environmental datasets.

3.3. QUALITATIVE METHODS

To ensure the validity of research findings, qualitative methods can often be employed to ensure the results are not just technically sound but also contextually meaningful and relevant. Common qualitative validation techniques include expert interviews [49], focus groups [50], and expert surveys [51]. These approaches enable the collection of expert feedback on model results, particularly useful in multidimensional fields such as sustainability.

Among these various options for validation, this thesis applies a structured expert survey, which is especially useful when input is needed from multiple domain experts in a standard format. They typically include both closed-ended and open-ended questions to capture overall and specific feedback [52]. Compared to interviews and focus groups, surveys offer greater scalability and anonymity, making them suitable for validating results in a time-efficient way. The feedback obtained this way can then be used to validate the plausibility and relevance of the research results.

4

EXPERIMENTAL SET-UP

This chapter details the process used to assess the efficiency of EU countries in achieving SDGs 8,9, and 12, and investigates the influence of SMEs on this efficiency. The methodology follows the CRISP-DM framework (introduced in Section 3.1), ensuring a systematic and replicable approach. Figure 4.1 illustrates the process flow from data collection to deployment.

4.1. BUSINESS UNDERSTANDING

The main objective of this research is to evaluate the relative efficiency of EU member states in progressing toward selected SDG targets using DEA. A secondary aim is to explore how SMEs influence that efficiency. The scope covers all 27 EU countries and their data related to SDG 8, 9, and 12. Ultimately, this study seeks to identify the most efficient countries, quantify the impact of SMEs on efficiency, and generate policy-relevant insights by analysing slack values, thereby identifying specific input excesses and output shortfalls that can be addressed to enhance SDG performance.

4.2. DATA UNDERSTANDING

The primary dataset was retrieved from the UN SDG Global Database[1], selected for its thorough and comprehensive records compared to Eurostat's SDG monitoring datasets. SME data,defined by the number of SMEs per country, was obtained from Eurostat's business demography statistics[2].

The raw dataset included a wide array of indicators across SDG sub-goals and breakdowns (age, gender, disability, etc.). Initial inspection, by hand and visualisations, revealed challenges such as missing values, varied measurement units, and binary indicators representing policy adop-

[1] `https://unstats.un.org/sdgs/dataportal`

[2] `https://ec.europa.eu/eurostat/databrowser/view/SBS_SC_SCA_R2__custom_16600735/default`

tion. a rigorous preprocessing phase was required to ensure consistency and comparability. The SME data needed minimal formatting. After the initial inspection, the primary database was evaluated for compliance with FAIR data practices.

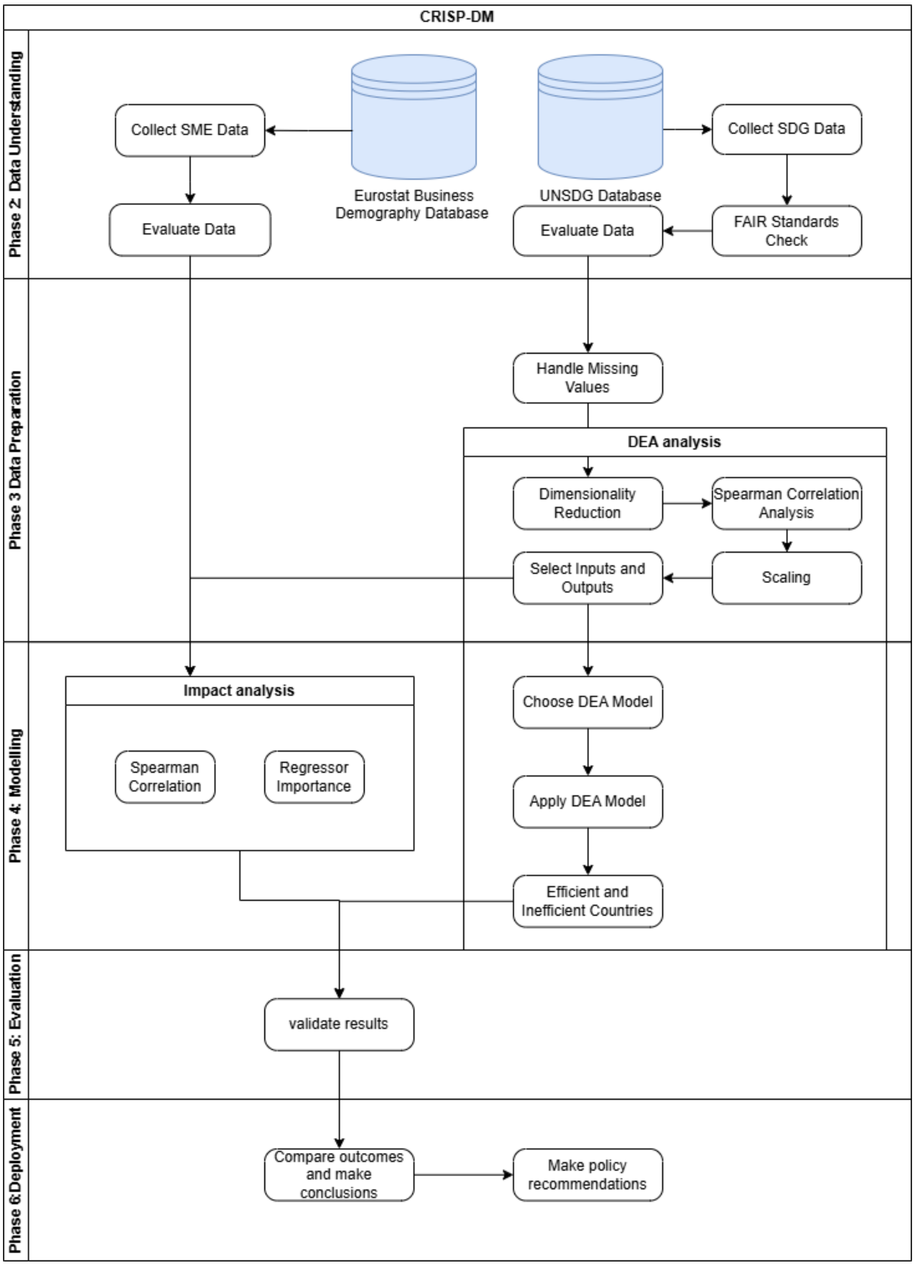


Figure 4.1: Process overview of experimental set-up

4.3. DATA PREPARATION

To ensure the dataset is suitable for DEA modelling, several preprocessing steps are applied to address missing data, reduce dimensionality, and prepare the final input-output structure.

4.3.1. HANDLING MISSING VALUES

Binary indicators for policy adoption were treated such that all positive values (1) prior to the measurement year were retained. Countries with no positive values were assigned "0". Non-binary indicators with missing values were removed to ensure consistency across countries. Figure 4.2 illustrates the amount missing data per country, quantifying the gaps in data and highlighting where data loss is most prevalent.

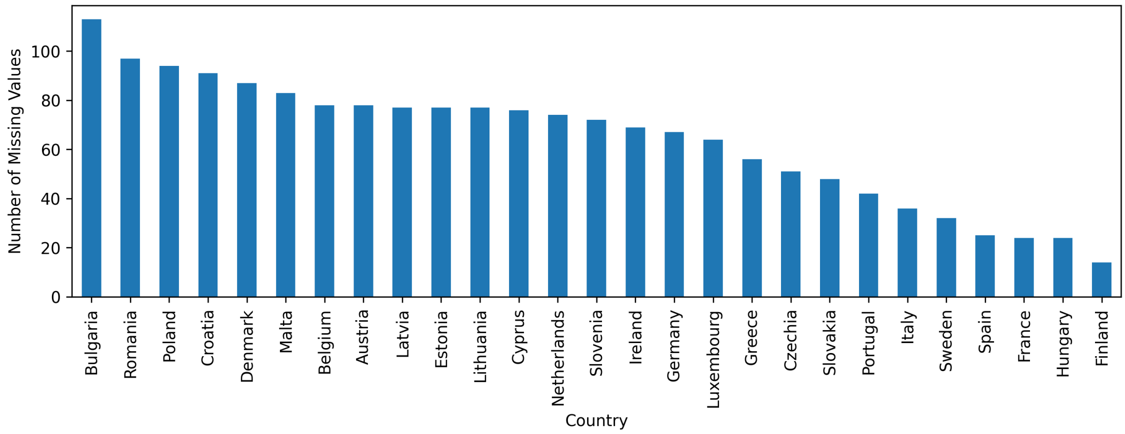


Figure 4.2: Missing values per country

4.3.2. DIMENSIONALITY REDUCTION

According to Cooper et al. [53] (equation 4.1), the number of DMU should be at least three times the number of indicators. Given 27 DMUs, approximately nine indicators per goal should be retained.

$$\text{Number of DMU} \geq Max\{m * t, 3(m+t)\} \tag{4.1}$$

Following missing value handling, 170 total indicators remain, meaning dimensionality reduction is required. Dimensionality was reduced by removing high-granularity breakdowns with higher granularity. Furthermore, some sub-indicators contained multiple measures such as total, per capita or per unit of GDP. In those cases, per capita measures were selected over total values to ensure comparability. Figure 4.3 provides an overview of this reduction process, showing the step-by-step narrowing of the initial 347 indicators to the final 22 used for modelling (see Table 4.1).

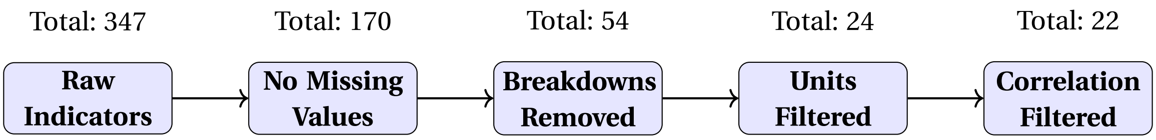
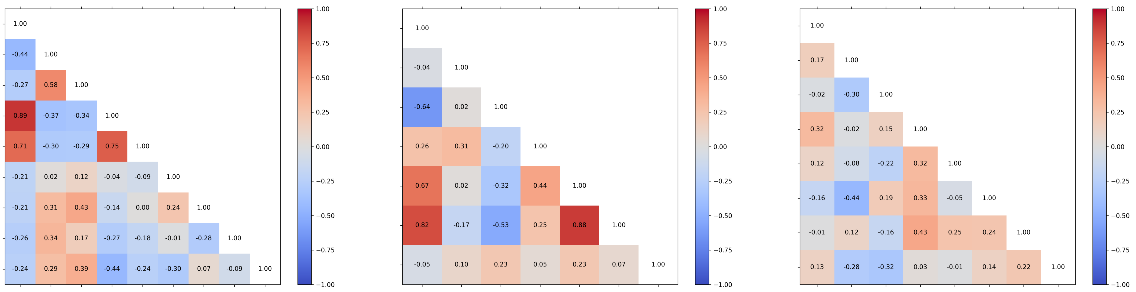


Figure 4.3: Dimension Reduction Pipeline for SDG Indicators

4.3.3. CORRELATION ANALYSIS

A Spearman correlation analysis identified redundant indicators ($\rho > 0.85$). Spearman's rank correlation is selected over Pearson's due to its suitability for non-linear and non-normally distributed data, which aligns with the characteristics of the dataset used in this study (See Figure 4.5). Figure 4.4 shows the heatmaps of the indicators for each goal, revealing that both *Gross Domestic Product (GDP) growth per worker* and *GDP growth per capita* were highly correlated, meaning that *GDP growth per worker* was removed. Similarly, *researchers per million* was excluded due to its strong correlation with *Research and Development (R&D) expenditure*.



(a) SDG 8 (b) SDG 9 (c) SDG 12

Figure 4.4: Spearman correlation heatmaps for the different goals (detailed version in Figure B.1)

4.3.4. FEATURE SCALING

According to Sarkis [54] both negative values and high differences in data magnitudes is detrimental to the performance of DEA models. Figure 4.5 shows the current difference in range between the indicators, where some range from 0 to 5, like *Labour Rights*, while others range from 1000-30000, such as *Manufacturing Value Added*. Therefore, some scaling is needed to make sure the DEA model functions best. Initially, standard scaling was implemented, but then negative values were still present. So, Min-Max Scaling was used for the final dataset.

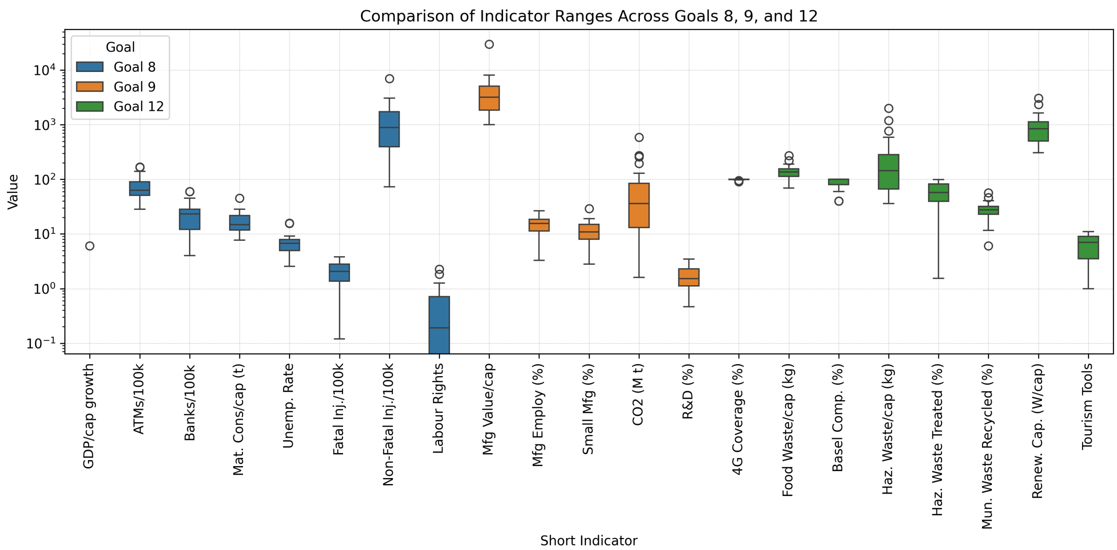


Figure 4.5: Indicator ranges (log scale)

4.3.5. SELECTION OF INPUTS AND OUTPUTS

Following the final selection of indicators, each indicator was classified as either an input (to be minimised) or an output (to be maximised) for DEA. Table 4.1 provides a comprehensive overview of categorised indicators across the three SDGs. Notably, *Material Consumption per Capita* appears under both SDG 8 and 12, reflecting its relevance to both economic and environmental dimensions. Additionally, SDG 9 is characterised by a single input and five outputs, illustrating an output-intensive structure in its efficiency assessment.

Table 4.1: DEA Inputs and Outputs Grouped by SDG Goals 8, 9, and 12

SDG	**Inputs**	**Outputs**
SDG 8	• Material Consumption per capita (Tonnes) • Unemployment Rate (%) • Fatal Injuries (per 100k) • Non-Fatal Injuries (per 100k)	• GDP Growth per capita (%) • ATMs per 100k population • Bank Branches per 100k • Labour Rights Compliance (%)
SDG 9	• Carbon Dioxide (CO_2) Emissions (Million tonnes)	• Manufacturing Employment (%) • Manufacturing Value Added per Capita • Small-Scale Manufacturing (%) • R&D Employment (%) • 4G Network Coverage (%)
SDG 12	• Material Consumption per capita (Tonnes) • Food Waste per Capita (KG) • Hazardous Waste per Capita (KG)	• Basel Compliance (%) • Hazardous Waste Treated (%) • Municipal Waste Recycled (%) • Renewable Energy Capacity (Watts per Capita) • Tourism Accounting Tools (Number of Tables)

4.4. MODELLING

The finalised dataset consisted of 27 DMUs and the set of carefully selected scaled indicators that meet the dimensionality constraint. Inputs and outputs were exported to use for the DEA model using the SBM approach under the assumption of VRS, configured with an output orientation to reflect the research objective of maximising sustainability outcomes given the current use of resources. Modelling was performed via the "deaR Shiny"[3] web application developed by Benítez et al. [55], which supports SBM-DEA models, as well as a plethora of others such as the BCC and CCR models.

Each country was treated as a DMU, and the model generated efficiency scores, slacks values, and target benchmarks for each. Efficient countries formed the production frontier, while inefficient ones were projected onto this frontier to identify performance gaps.

The SBM model was selected for its ability to incorporate slacks, offering insights into the sources of inefficiency. The VRS assumption was chosen to accommodate the difference in scale, resources, and infrastructure across EU countries.

To investigate the secondary research objective, investigating SME impact on SDG efficiency, a Spearman correlation matrix was computed between SME counts and DEA indicators. Given the non-normal distribution of most variables, Spearman's rank correlation coefficient, Random forest feature importance, and linear regression coefficients were applied using sklearn[4] and scipy[5] libraries.

4.5. EVALUATION

Validation is carried out through both direct and indirect methods to ensure model robustness. Indirect validation involves experimenting with different model configurations, such as varying the DEA parameters and testing alternative returns-to-scale assumptions. Direct validation includes the use of different preprocessing techniques, particularly comparing the effects of Min-Max scaling versus standard mean scaling on the DEA outcomes.

Additionally, external validation is conducted through a structured expert survey with domain specialists in sustainability and researcher experts on the use of DEA. Experts are asked about the plausibility of the DEA results on SDGs 8, 9, and 12, as well as the meaningfulness of the correlation between SME amounts and CO_2 emissions. The survey also includes optional questions to rank countries for each SDG, although the results of these questions were only used for qualitative cross-checking.

Their responses are analysed both descriptively (via medians and modes) and statistically using a Friedman test to determine the consistency and significance of expert agreement. These combined efforts help assess alignment of the DEA results with expert judgement and practical expectations, reinforcing the model's validity.

[3] https://rbensua.shinyapps.io/deaR/#
[4] https://scikit-learn.org
[5] https://scipy.org/

4.6. DEPLOYMENT

The final step in the CRISP-DM cycle is to present the results in accessible formats for policymakers and stakeholders. The outputs of the models included the following data:

1. Efficiency scores and rankings for each DMU.
2. Slack values indicating input excesses and outputs shortfalls.
3. Target values indicating where improvements can be made for inefficient DMUs.
4. Reference sets of peer countries used for benchmarking.
5. Spearman correlation coefficients between the DEA indicators and number of SMEs per country.

Visualisations, summary tables, and policy recommendations are presented in Chapter 5.

5

RESULTS AND DISCUSSION

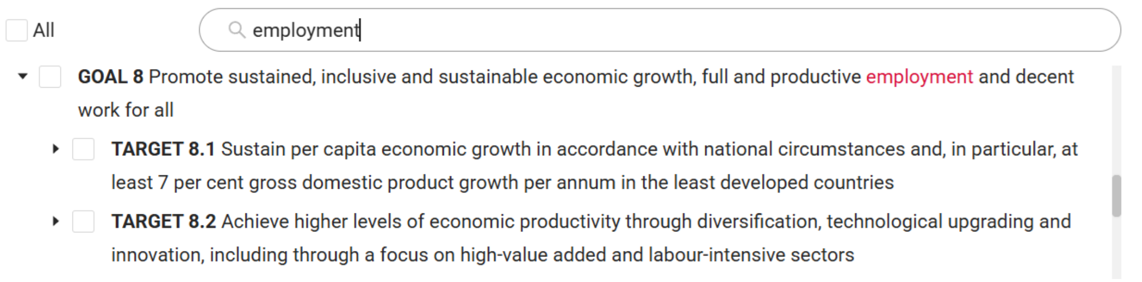
This chapter presents the main analytical results of the study. Section 5.1 begins with an evaluation of the FAIRness of the UN SDG database. Section 5.2 follows with the DEA efficiency scores and slack analysis for SDGs 8, 9, and 12. Section 5.1 explores the correlation between SME presence and the indicators used in the DEA model. The chapter concludes with the limitations of this study and any suggestions for future research in Section 5.4.

5.1. FAIR DATA ASSESSMENT

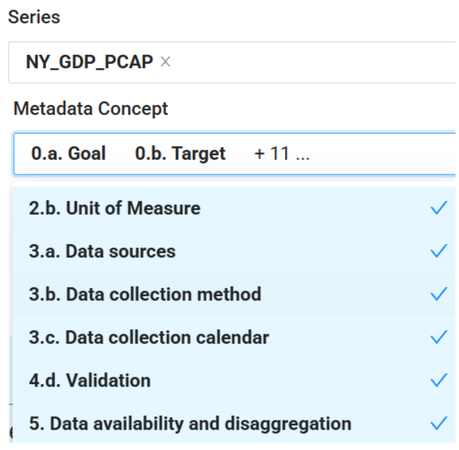
To evaluate how well the UN SDG database complies with the FAIR principles [47], this assessment applies criteria based on the GO FAIR initiative. Each principle is scored on a scale from 0 to 4, where 0 indicates non-compliance (poor performance) and 4 indicates full compliance (excellent performance).

5.1.1. FINDABILITY

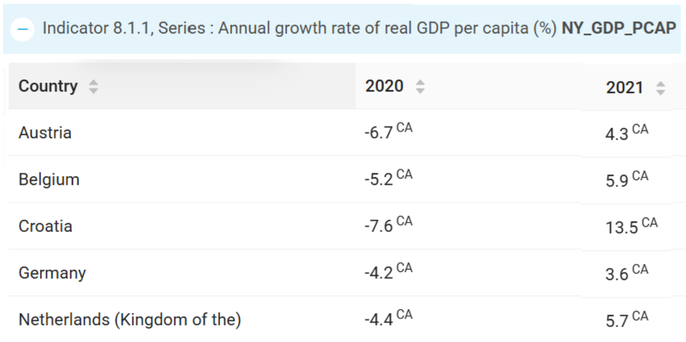
The database is highly discoverable via the UN's data portal (see Figure 5.1a), which supports keyword-, SDG-, and indicator-based searches (F4). Figure 5.1b shows that rich metadata is provided for each indicator (F3), including definitions, data sources, and methodology (F2). However, no globally unique and persistent identifiers are assigned to individual data points; instead, a combination of indicator, year and country is used (F1). As a result, while the data is well-indexed and described, individual data points are not traceable or findable. As the dataset succeeds in 3 of the 4 requirements for Findability, it gets a score of 3 for this dimension.



(a) Discoverability through UNs data portal's search (F4)



(b) Rich metadata (F2, F3)



(c) Data points without unique identifiers (F1)

Figure 5.1: Screenshots from the UN SDG database

5.1.2. Accessibility

Data from the UN portal is retrievable in several standardised open protocols, including HTTPS and an API, where at least HTTPS is a free and universally implementable protocol (A1.1). Datasets are downloadable in multiple machine-readable formats such as CSV and JSON. Access is unrestricted, and no registration is required for data retrieval (A1.2). Lastly, the metadata [1] is stored separately from the data [2] it references, meaning it should remain accessible even if the datasets are deprecated or deleted (A2). Overall, the database is fully compliant with the accessibility criteria, giving it a full score of 4.

5.1.3. Interoperability

Although the database supports several data formats such as JSON and SDMX, which are widely used in the official statistics domain (I1), it lacks semantic web and metadata formats such as RDF or OWL (I2). This means there is no easy way to make connections between indicators or external elements, limiting the potential for automatic integration with other datasets (I3). As it uses some standard formats but fails to use formats that allow linking and integration with other datasets, the interoperability of the dataset gets a score of 2.

[1] https://unstats.un.org/sdgs/dataportal/SDMXMetadataPage
[2] https://unstats.un.org/sdgs/dataportal/database

5.1.4. REUSABILITY

The database includes detailed metadata on data sources, collection methods, and update times (R1.2). These features support reproducibility and readability for informed reuse. However, the absence of a data usage license (R1.1) could be a limitation for reuse, though it could be considered public domain coming from the UN. Lastly, as it uses SDMX formats, it does adhere to domain-specific data standards (R1.3). Therefore, while the metadata and methodological rigour are high, the lack of clear licensing information reduces the reusability score to 3.

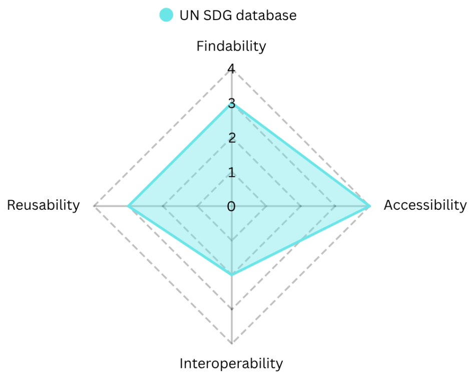


Figure 5.2: FAIR assessment of UN SDG Global database

The FAIRness profile of the UN SDG database reflects a strong commitment to accessibility and metadata quality, but reveals opportunities for improvement in interoperability and clarity in licensing. Figure 5.2 provides a visual overview of the results.

5.2. DEA RESULTS

The DEA analysis is split into three parts, one for each SDG goal. Results include efficiency scores for each DMU, slacks that indicate the shortfalls of inefficient DMUs in each indicator, and the reference DMU that each inefficient DMU was compared to.

5.2.1. SDG 8: DECENT WORK AND ECONOMIC GROWTH

The DEA efficiency scores for SDG 8 reveal some disparity in performance among EU member states. While 15 countries achieved a perfect efficiency score of 1.000, indicating optimal use of indicators to generate the desired outputs, 12 countries were identified as inefficient to varying degrees. The most inefficient performers include Sweden (0.0022), Latvia (0.0028), Slovakia (0.0029), Italy (0.0053), and Spain (0.0117), highlighting challenges in converting the input resources into sustainable economic outcomes.

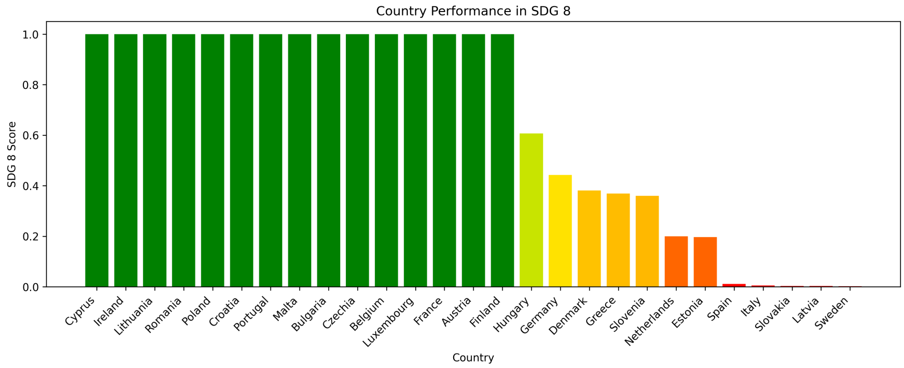


Figure 5.3: Bar chart of SDG 8 efficiency scores

Figure 5.3 presents these efficiency scores, and Figure 5.4 illustrates the slack value of each indicator for each inefficient country. These values highlight the distance between current performance and the target set by the model to reach the efficient frontier. These slack values provide insight into each country's specific shortfalls. While higher slack values indicate greater inefficiency and thus carry more weight in interpreting performance gaps, any non-zero slack value still reflects a measurable deviation from optimal outcomes. Even when slack is relatively low, it represents an area where improvements are still needed and can accumulate into broader inefficiencies when occurring across multiple indicators or countries.

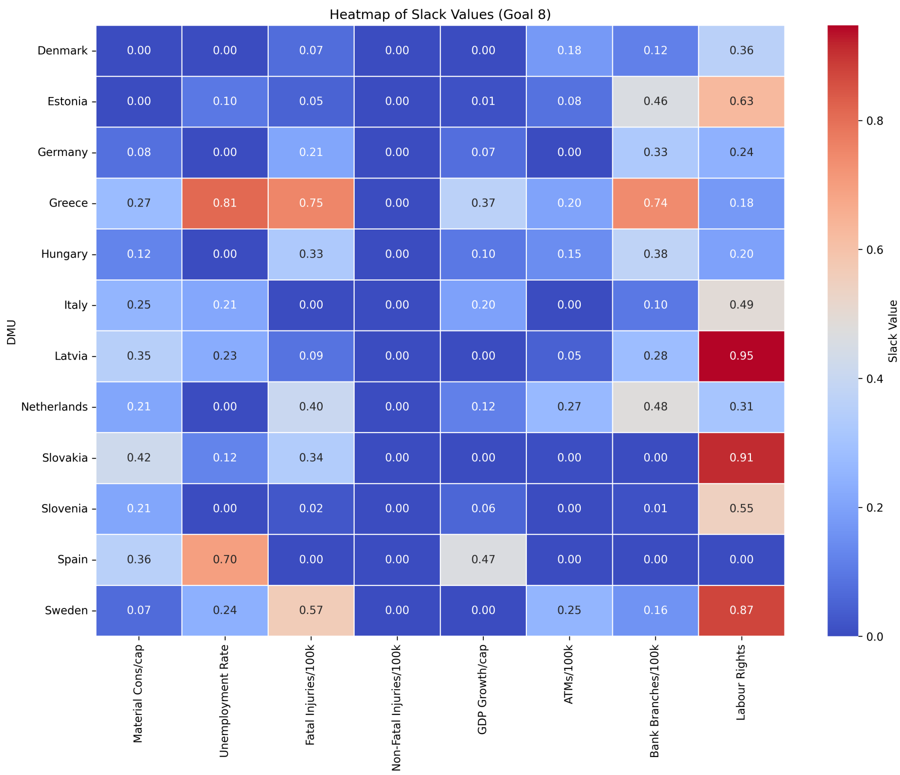


Figure 5.4: Heatmap of SDG 8 input and output slacks for each inefficient country

A closer examination of the slacks reveals that several indicators consistently contribute to inefficiencies among almost all DMUs. The most significant ones are *Compliance with Labour Rights*, *Material Consumption*, and *Fatal Occupational Injuries*. For example, Latvia, Sweden, and Slovakia exhibit significant slack values in *Labour Right Compliance* with a smaller but still notable shortfall in *Fatal Occupational Injuries*. Together, these patterns suggest broader systemic issues in workplace regulation and safety.

Greece, though ranking fifth among the inefficient countries, shows some slack across almost all indicators. This suggests that its inefficiency is not due to extreme underperformance in any single area, but rather a moderate shortfall across the board. In contrast, the *Non-fatal Occupational Injuries* indicator shows no slack for any DMU, indicating that no country had room to improve in this domain according to the DEA model. This signifies that attempting to enhance performance for this indicator would likely lead to deterioration in other areas.

Meanwhile, countries such as Hungary, Germany, and Denmark that are closer to the efficiency frontier need only minor improvements. Particularly in *Compliance with Labour Rights* and access to financial infrastructure as measured by *Number of Bank Branches*. Addressing these targeted inefficiencies could help them achieve full efficiency.

These findings show the importance of examining not just the efficiency score but also the accompanying slack variables. While an efficiency score provides a useful metric, slack analysis offers the granular insights need to make specific policy recommendations. This dual outcome enables countries to pinpoint specific weaknesses in policy and implement focused interventions that drive meaningful progress towards SDG 8.

5.2.2. SDG 9: INDUSTRY, INNOVATION, AND INFRASTRUCTURE

The DEA of SDG 9 reveals a similarly uneven landscape of efficiency across EU states as observed for SDG 8. Of the 27 countries analysed, 15 achieved perfect efficiency, while the remaining 12 demonstrated varying degrees of inefficiency. Among the most inefficient performers were Bulgaria (0.0005), Romania (0.0037), and Luxembourg (0.0163), indicating significant shortfalls related to industrial development, innovation, and infrastructure capabilities.

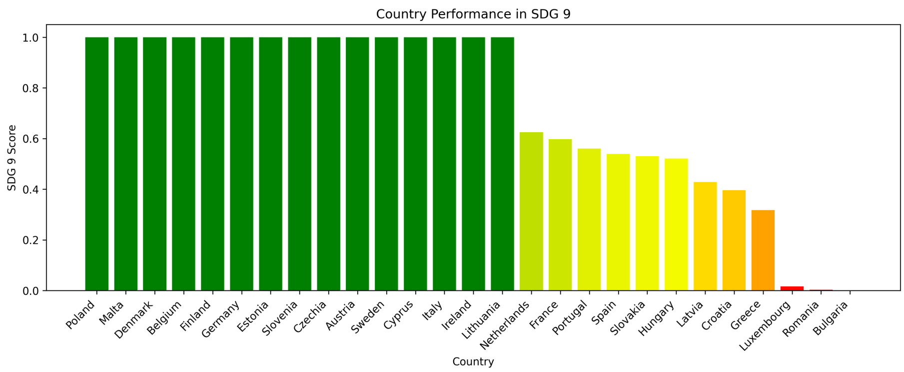
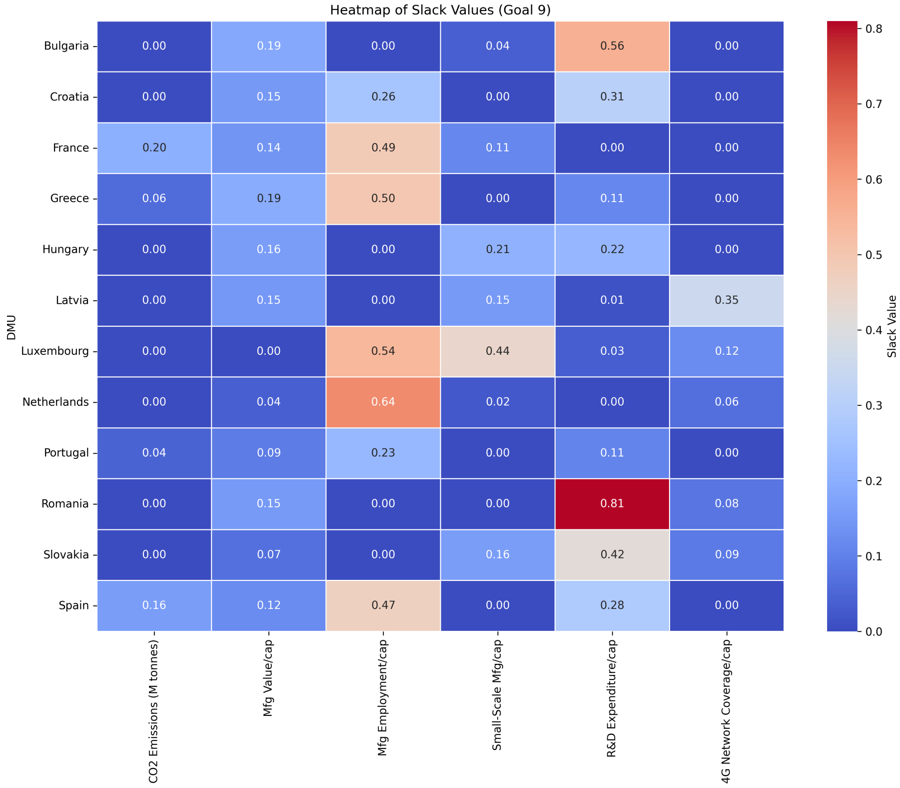


Figure 5.5: Bar chart of SDG 9 efficiency scores

Figure 5.5 presents the efficiency scores, while Figure 5.6 illustrates the slack values for each inefficient country. These slack values help identify where each nation is falling short and where improvements can be made to enhance SDG 9 performance.



DMU	CO2 Emissions (M tonnes)	Mfg Value/cap	Mfg Employment/cap	Small-Scale Mfg/cap	R&D Expenditure/cap	4G Network Coverage/cap
Bulgaria	0.00	0.19	0.00	0.04	0.56	0.00
Croatia	0.00	0.15	0.26	0.00	0.31	0.00
France	0.20	0.14	0.49	0.11	0.00	0.00
Greece	0.06	0.19	0.50	0.00	0.11	0.00
Hungary	0.00	0.16	0.00	0.21	0.22	0.00
Latvia	0.00	0.15	0.00	0.15	0.01	0.35
Luxembourg	0.00	0.00	0.54	0.44	0.03	0.12
Netherlands	0.00	0.04	0.64	0.02	0.00	0.06
Portugal	0.04	0.09	0.23	0.00	0.11	0.00
Romania	0.00	0.15	0.00	0.00	0.81	0.08
Slovakia	0.00	0.07	0.00	0.16	0.42	0.09
Spain	0.16	0.12	0.47	0.00	0.28	0.00

Figure 5.6: Heatmap of SDG 9 input and output slacks for each inefficient country

A detailed examination of the slack data reveals several common areas of inefficiency across multiple countries. Bulgaria, Romania, and Slovakia, for example, struggle particularly with *R&D expenditure*, which is a driver of innovation and technological advancement.

More broadly, *Manufacturing Employment per Capita* appears to have a strong influence on multiple inefficient DMUs, while *Manufacturing Value Added per Capita* has a more modest effect, affecting almost all DMUs to a small extent. These areas are essential for building resilient infrastructure and promoting domestic industrialisation. Their underperformance suggests that several EU countries may not adequately support their domestic manufacturing sectors or investing in job creation within these industries.

CO_2 Emissions seems to be less of a concern overall, except in larger countries such as France and Spain. These countries might face challenges in decoupling economic activity from carbon emissions, underscoring the need for cleaner industrial practices. In contrast, slack in *4G Network Coverage* was primarily observed in smaller countries like Latvia and Luxembourg, though its overall impact also appears limited. This suggests that gaps in digital infrastructure are more likely to affect smaller EU member states.

These findings highlight the importance of not only promoting general industrial growth but also targeting specific innovation and infrastructure components that can drive meaningful efficiency improvements.

5.2.3. SDG 12: Responsible Consumption and Production

The efficiency results for SDG 12 underscore pronounced variability in sustainable consumption and production performance among EU member states. Only 13 countries achieved a perfect efficiency score, while the remaining 14 exhibited a range of inefficiencies. Among the most inefficient performers were Belgium (0.0064), Romania (0,0069), Malta (0.0083), and Hungary (0.0135), all of which demonstrated extensive slack across several indicators.

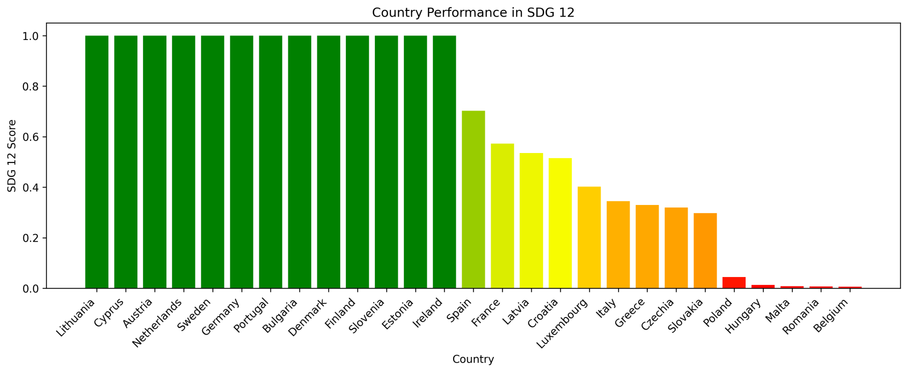


Figure 5.7: Bar chart of SDG 12 efficiency scores

Figure 5.7 displays the efficiency scores, and Figure 5.8 visualises the slack values, highlighting areas of underperformance. A close analysis of the slack data reveals that inefficiencies for SDG 12 generally involve a combination of multiple moderate to high slacks, rather than a single dominant factor. Most inefficient countries show overlapping shortfalls across various indicators, suggesting broader structural or policy-related shortcomings in sustainable governance.

Belgium, Greece, Italy, and Malta show substantial slack in the implementation of *Tourism Accounting Tools*, suggesting underutilization of mechanisms that track and promote sustainable tourism development. These tools are used for monitoring the environmental and economic impacts of tourism and guiding corrective policies.

Spain, Luxembourg, and Romania all exhibit slack in *Basel Convention Compliance*, and each also shows a distinct inefficiency in a separate waste-related indicator. Specifically, Spain demonstrates slack in *Food Waste per Capita*. Romania underperforms in *Municipal Waste Recycling*, while Luxembourg displays slack in *Hazardous Waste Treatment*. These findings emphasise that while all three countries struggle with Basel compliance, their additional inefficiencies are unique, reflecting diverse national challenges in waste management.

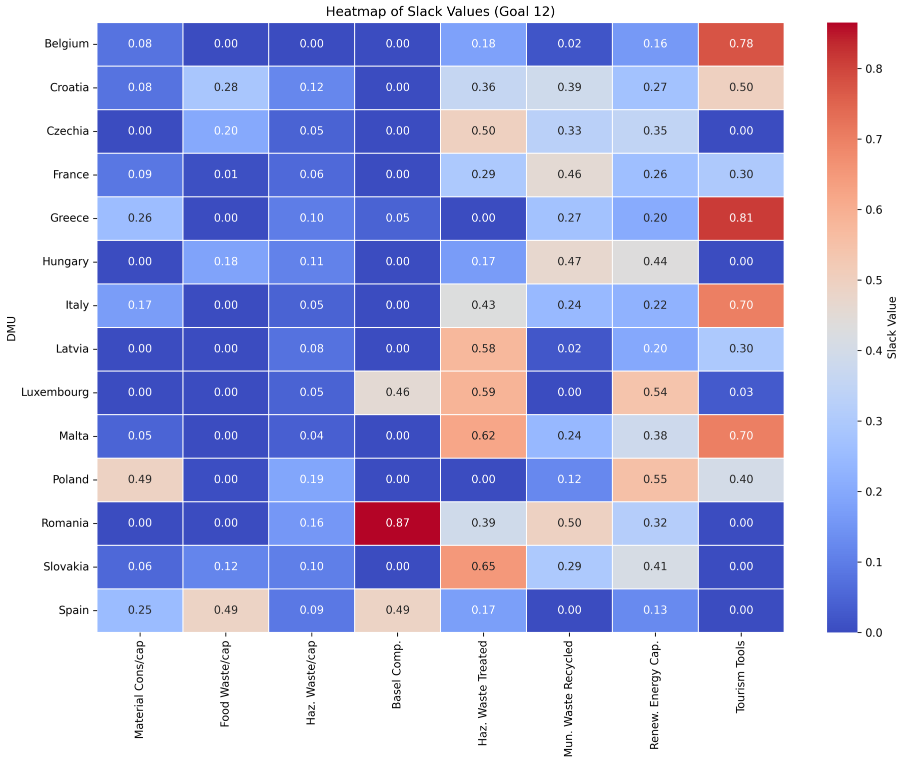


Figure 5.8: Heatmap of SDG 12 input and output slacks for each inefficient country

Across the broader spectrum, *Hazardous Waste Treatment*, *Municipal Waste Recycling*, and *Renewable Energy Capacity* are common areas of slacks among almost all countries. These persistent inefficiencies highlight systematic issues in the adoption and deployment of clean technologies and the circular economy practices throughout most of the EU.

The results emphasise that consumption and production is a multidimensional challenge requiring systematic interventions. As with SDGs 8 and 9, analysing both efficiency scores and indicator level slack provides critical insights for policy makers aiming to enhance performance. Tailored strategies, focusing on specific weak points such as waste management or renewable energy capacity, can significantly reduce inefficiency and accelerate progress towards SDG 12 goals.

5.2.4. Comparative analysis across SDGs

The integration of results from SDGs 8, 9 and 12 reveals both convergence and divergence in efficiency patterns across the EU. Figure 5.9 presents an overview of efficiency scores across all three SDGs for each country. Austria, Cyprus, Ireland, Finland, and Lithuania consistently performed well across all three goals, representing an efficiency frontier across economic inclusivity, innovation capacity, and sustainable consumption practices.

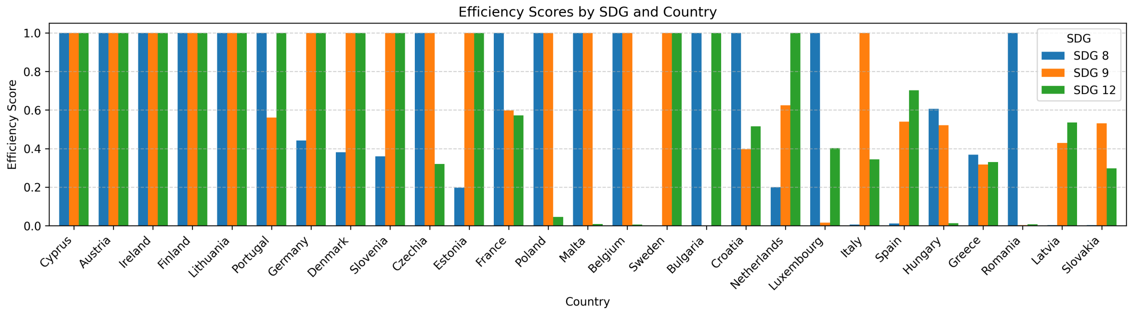


Figure 5.9: Combined bar chart of all SDG efficiency scores per country

At the other end of the spectrum, several countries consistently underperformed. Slovakia, Latvia, Greece, Hungary, and Spain stood out for their inefficiency across all SDGs. Slovakia, Latvia, and Spain showed some of the weakest performances in SDG 8 while also lagging in innovation and production efficiency (SDGs 9 and 12). Greece showed middling performance on all fronts, whereas Hungary ranked among the lowest in innovation and also underperformed in SDGs 8 and 9.

A number of countries exhibited notable asymmetries in their SDG efficiency profiles. For example, Germany, Denmark, Slovenia, and Estonia demonstrated top-tier performance in SDG 9 and 12, but scored poorly in SDG 8, pointing to strength in innovation and sustainable consumption, but weakness in employment standards and economic growth. Conversely, Poland, Malta, and Belgium achieved full efficiency in SDGs 8 and 9, yet were among the lowest performers in SDG 12, suggesting that gains in economic and technical dimensions have not yet translated into sustainable production and consumption. Bulgaria presented another pattern, attaining perfect efficiency scores in SDGs 8 and 12, but ranking lowest in SDG 9 due to gaps in R&D and industrial development.

Overall, these results demonstrate that high efficiency in one SDG does not necessarily guarantee strong performance in others. While a handful of countries have achieved balanced and integrated progress across all three dimensions, many others display sectoral imbalances. These disparities emphasise the importance of tailored, multidimensional policy responses to ensure that progress in one area does not come at the expense of another.

5.3. CORRELATION ANALYSIS

The results of the correlation analysis examine the relationship between the number of SMEs and the SDG indicators used in the DEA model. The analysis focused on identifying patterns in the influence of SME presence in each country on the socioeconomic and environmental indicators.

For SDG 8, the correlations were generally weak. The strongest association was *Material Consumption* (ρ = -0.46), which negatively correlates with SME amount, indicating a potential contribution by SMEs to more efficient resource usage. Other indicators, such as *Fatal* and *Non-*

fatal Occupational Injuries, showed little correlation, indicating limited direct linkage between SME numbers and labour protection influence.

For SDG 9, the analysis found a strong positive correlation with *CO_2 Emissions* (Spearman ρ = 0.95), suggesting that countries with more SMEs tend to have higher CO_2 output. This indicates a substantial contribution of SMEs to national emission levels and raises concerns about their environmental footprint. Fei et al. [56] supports this finding, emphasising that SMEs often contribute significantly to CO_2 emissions, partly due to their frequent exception from environmental regulation.

R&D Expenditure also positively correlates with SME presence and ranked highly in the regression analysis, highlighting the link between SMEs and innovations. Other indicators, such as *Manufacturing Employment* and *Manufacturing Added Value*, showed moderate positive correlations. In contrast, *Small-Scale Manufacturing* exhibited a weak and slightly negative Spearman correlation, although it has notable importance in the Random Forest model.

Indicators for SDG 12, such as *Renewable Energy Capacity* and *Hazardous Waste Treated*, were slightly correlated with SME amount, which may imply that countries with more SMEs achieve slightly more sustainable energy and waste practices, though the relationships are generally weak and varied across methods.

Overall, the correlation results reveal that SME presence is most closely linked to indicators associated with industry and innovation, rather than environmental performance or social impact. These patterns suggest that policies related to SMEs will most likely have more impact on innovation and emissions, but will not necessarily contribute meaningfully to SDGs 8 and 12.

5.4. LIMITATIONS AND FUTURE RESEARCH

This section outlines several limitations of the study and proposes directions for future research to address these constraints and build upon the findings presented.

First, several indicators from the original dataset had missing values for at least one of the relevant countries, requiring omission of those indicators, which may influence the results of the analysis. Additionally, the majority of the data was from the year 2020, because it had the fewest missing values across indicators and countries, making it the most suitable for comprehensive analysis. Although this ensures consistency, more recent data might better reflect current trends and policy.

Secondly, the indicators selected may not fully capture the multidimensional nature of the SDGs, because of the omission of some of the UN indicators due to missing data, some of the social and environmental dimensions of SDGs 8 and 12 may be underrepresented, leading to an incomplete picture of country performance.

Third, while the VRS output-oriented SBM-DEA model was suitable for this study's goals, alternative DEA specifications or frontier models could yield different efficiency rankings and provide further insights.

Fourth, the correlation analysis between the number of SMEs and SDG indicators identifies statistical associations, but it can't determine whether changes in SME numbers *cause* changes in SDG indicators. Random Forest and regression might improve the robustness of the assessment, but remain susceptible to variable bias.

Future research could build on the current research in several ways. Incorporating time series data would allow researchers to examine changes over time in SDG performance and SME development. Expanding the range of indicators could better capture the whole scope of the SDGs. Evaluating the effects of specific SME-focused policies with case studies could provide deeper insights on their impact and could find some causation instead of correlations. Additionally, using alternative DEA models, such as MCDA or customised DEA, would test the robustness of the results. Finally, further research should consider regional or subnational analysis to uncover variations within countries that national data may not fully catch.

5.5. VALIDATION OF RESULTS

To assess the plausibility and contextual relevance of the quantitative results, a structured expert survey was conducted with seven domain specialists in sustainability. The survey aims to validate the DEA efficiency results and the relevance of the identified correlation with SMEs. The survey contains four Likert-scale questions addressing the results of the study, and four optional questions about how the respondent would rank them.

5.5.1. DESCRIPTIVE FINDINGS

Figure 5.10 presents the frequency of expert ratings for each survey item, including median and mode statistics.

Regarding SDG 8, most experts evaluated the efficiency outcomes as either somewhat plausible or very plausible. Although three indicated that it was to some extent implausible, the median and mode of 3 suggest a moderate level of agreement about the plausibility of these results.

For SDG 9, the feedback also indicated a median of 3, though this time with just one somewhat implausible response. This reflects that the experts generally view the findings as plausible, although not with overwhelming conviction.

The assessment of SDG 12 yielded higher confidence among respondents, with both the median and mode reaching 4. Which suggests a stronger agreement with the model's efficiency ranking for responsible consumption and production.

Finally, in relation to the correlation between SMEs and CO_2 emissions, all experts agree that the relationship is either meaningful or very meaningful, with the responses again landing at a median and mode of 3, reinforcing the relevance of this finding.

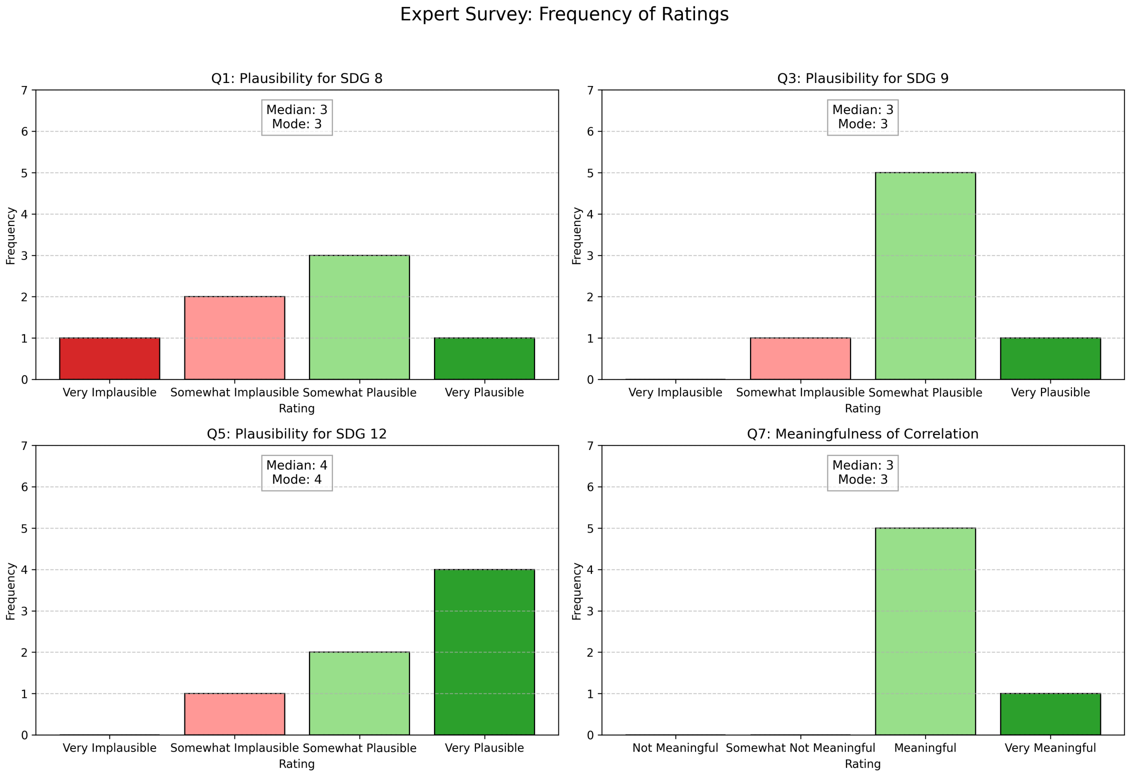


Figure 5.10: Expert Survey Results: Frequency of ratings for each key result

5.5.2. STATISTICAL TESTING OF SURVEY RESULTS

To support the qualitative findings with inferential evidence, a statistical test is conducted on the four Likert scale responses. The chosen test is the Friedman test for k samples, with Bonferroni correction [57], selected for its appropriateness in handling ordinal data and small data sizes. The test helps to check if the differences in judgement are significantly different (alternative hypothesis) or if we should assume the null hypothesis (the samples come from the same population). The resulting p-value is 0.297, which is above the designated alpha, which means we should assume these four samples are not significantly different. Given that the observations in the four samples are not significantly different, it adds evidence that experts are consistent in their positive opinion about the four indicators.

6

CONCLUSION

This thesis set out to explore how DEA-based methods can be employed to evaluate the efficiency of EU countries in achieving selected SDGs, with a particular focus on SDGs 8 (Decent Work and Economic Growth), 9 (Industry, Innovation, and Infrastructure), and 12 (Responsible Consumption and Production). A key innovation in this study is the integration of SME-related data into the DEA assessment, which is often overlooked in existing SDG assessments. Additionally, the FAIRness of the primary data source is evaluated to ensure the quality and accessibility of the data employed.

6.1. ANSWERS TO THE RESEARCH QUESTIONS

This section will answer each of the sub-research questions before addressing the main research question.

DEA was effectively used to measure how efficiently EU countries transform sustainability-related inputs into desirable outputs across the three selected SDGs. Based on the insights from the SLR, the SBM model with VRS and an output-oriented approach was chosen, which enabled a nuanced assessment of relative efficiency among the 27 EU countries. Each country was treated as a DMU, and indicators were selected based on data quality, availability, and the absence of significant correlation between them. The inclusion of slacks made it possible to identify which countries underperformed, but also in which specific indicators (both inputs and outputs) those inefficiencies occurred. This approach allowed for the quantification of how far each inefficient country deviates from the efficiency frontier, offering clear targets for improvement.

The analysis revealed distinct inefficiencies across the three SDGs. For SDG 8, several countries demonstrated inefficiencies primarily due to high unemployment rates, insufficient labour rights compliance, and a high occurrence of lethal workplace-related injuries. These indicators reflect fundamental shortcomings in economic inclusion and worker protection, which are central to the goal of promoting decent work and sustained economic growth. Some countries showed limited access to financial services as indicated by a lower number of bank branches per capita, which further hinders inclusive financial participation and thus economic growth.

In the case of SDG 9, inefficiencies stemmed from low levels of R&D investments, weak performances in domestic manufacturing employment, and to a lesser extent, CO_2 emissions and digital infrastructure, such as limited 4G coverage. These weaknesses suggest a lag in technological advancement and industrial innovation capacity in certain EU member states, which undermines the pursuit of resilient infrastructure and sustainable industrialisation.

SDG 12 inefficiencies were largely driven by high levels of food and hazardous waste generation and poor performance in recycling and circular economy practices. Some countries lacked comprehensive systems for monitoring and treating hazardous waste or tourism accounting tools. Together, these findings highlight significant areas where each specific EU nation needs to reduce environmental burden and adopt more sustainable consumption patterns.

The results of the correlation analysis revealed a complex relationship between the number of SMEs and the indicators used to measure SDG performance in EU countries. For SDG 8 the correlation between SME presence and performance indicators was generally weak. The only notable relationship was a negative correlation with material consumption (ρ = -0.46), suggesting that countries with more SMEs may achieve slightly more efficient resource use. However, no significant relationships were found between SMEs and labour indicators, indicating limited influence of SMEs on employment safety.

For SDG 9, SME amounts showed a much stronger association. A particularly strong positive relation was observed between the number of SMEs and CO_2 emissions, indicating that countries with more SMEs tend to have a higher level of emissions. This finding points to a potentially concerning environmental footprint of SMEs. At the same time, a positive correlation with R&D expenditure highlights the importance of SMEs in innovation ecosystems. Moderate positive associations were observed with indicators related to manufacturing such as employment and added value of the manufacturing industry, although the correlation with small-scale manufacturing was unexpectedly weak and slightly negative. This suggests that while SMEs are drivers of innovation and industrial activity, not all industrial outputs scale directly with SME numbers.

Regarding SDG 12, the correlations were generally weak across the board, just like with SDG 8, the most notable was the negative correlation with material consumption. Beyond that, slight positive relationships exist with indicators like renewable energy capacity and hazardous waste treatment, hinting at a minor impact of SMEs in sustainable production practices.

Overall, the findings suggest that the influence of SMEs on SDG efficiency in the EU is most prominent in areas related to Industry and Innovation (SDG 9), with a dual impact of both increased innovation and increased emissions. Their influence on SDGs 8 and 12 appears less direct and more limited. This nuanced impact underscores the importance of designing SME policies that not only foster innovation but also mitigate environmental harm and contribute to broader sustainability objectives.

Integrating SME-related data into DEA-based SDG assessments offers an important contextual layer that enhances the interpretive richness of the analysis. This thesis has shown that SME numbers can be linked to several key SDG indicators, most notably within SDG, where SMEs are sources of increased CO_2 emissions and contributors to R&D expenditure. This dual role underscores the value of incorporating SME metrics when in SDG-DEA analyses, as it shows an additional layer of influence that may exist beneath the initial data.

While SMEs did not show strong influence on many of the direct performance indicators under SDGs 8 and 12, their clear association within SDG9 reveals them as critical, if complex actors in the sustainability landscape. Their inclusion in the DEA analysis may offer policymakers a more thorough understanding of which economic structures contribute to or hinder SDG progress, allowing for better-targeted interventions.

6.2. ACADEMIC AND PRACTICAL CONTRIBUTIONS

From an academic perspective, this thesis contributes to the existing body of literature by integrating DEA methodology with SME and FAIR data perspectives. This approach has been underexplored in SDG analysis. It confirms that model configurations (SBM-VRS, output-oriented) can be tailored to reflect SDG contexts more accurately and provides a methodological blueprint for similar studies.

Practically, the findings support more informed policymaking at the EU and national levels. By identifying country-specific inefficiencies and linking them to the SME landscape, the study offers actionable insights for improving SDG performance. It also highlights the necessity of maintaining and enhancing the FAIRness of statistical data on UN and EU level, especially in terms of licensing, metadata, and interoperability.

References

[1] Varsha P S, A. Chakraborty, A. K. Kar. How to Undertake an Impactful Literature Review: Understanding Review Approaches and Guidelines for High-impact Systematic Literature Reviews. South Asian Journal of Business and Management Cases 13 (2024) 18–35. doi:10.1177/22779779241227654.

[2] L. Wuttke, CRISP-DM: das Standard-Vorgehensmodell für Data Mining, 2023. URL: https://datasolut.com/crisp-dm-standard/.

[3] GO FAIR, FAIR Principles - GO FAIR, 2019. URL: https://www.go-fair.org/fair-principles/.

[4] United Nations, Transforming our world: the 2030 Agenda for Sustainable Development | Department of Economic and Social Affairs, 2014. URL: https://sdgs.un.org/2030agenda.

[5] C. Fetting. The European Green Deal. ESDN report (2020). URL: https://www.esdn.eu/fileadmin/ESDN_Reports/ESDN_Report_2_2020.pdf.

[6] A. Charnes, W. W. Cooper, E. Rhodes. Measuring the efficiency of decision making units. European Journal of Operational Research 2 (1978) 429–444. doi:10.1016/0377-2217(78)90138-8.

[7] R. D. Banker. Estimating most productive scale size using data envelopment analysis. European Journal of Operational Research 17 (1984) 35–44. doi:10.1016/0377-2217(84)90006-7.

[8] X. Liu, J. Miao, H. Xia, A. Qiu, J. Chen. Trends and Policy Implications of Data Envelopment Analysis Method in the Process of Environment Sustainable Development. International Business Research 13 (2020) p25. doi:10.5539/IBR.V13N11P25.

[9] N. X. Maya, M. M. Prieto, L. Megido. Energy efficiency in middle-income countries with DEA: An approach for Latin America. Sustainable Futures 8 (2024). doi:10.1016/j.sftr.2024.100345.

[10] Y. Zhu, F. Yang, F. Wei, D. Wang. Measuring environmental efficiency of the EU based on a DEA approach with fixed cost allocation under different decision goals. Expert Systems with Applications 208 (2022) 118183. doi:10.1016/J.ESWA.2022.118183.

[11] J. Ibujés-Villacís, A. Franco-Crespo. Relationship between Productivity and Efficiency with Sustainable Development Goals: The Case of the Manufacturing Industry in Pichincha, Ecuador; [Relación entre productividad y eficiencia con los Objetivos de Desarrollo Sostenible: el caso de la industria manufacturera en Pichincha, Ecuador]. Revista de Metodos Cuantitativos para la Economia y la Empresa 35 (2023) 34 – 56. doi:10.46661/revmetodoscuanteconempresa.5475.

[12] M. Ehrenstein, R. Calvo-Serrano, Galán-Martín, C. Pozo, P. Zurano-Cervelló, G. Guillén-Gosálbez. Operating within Planetary Boundaries without compromising well-being? A Data Envelopment Analysis approach. Journal of Cleaner Production 270 (2020). doi:10.1016/j.jclepro.2020.121833.

[13] I. Łącka, Brzezicki. JOINT ANALYSIS OF NATIONAL ECO-EFFICIENCY, ECO-INNOVATION AND SDGS IN EUROPE: DEA APPROACH. Technological and Economic Development of Economy 28 (2022) 1739 – 1767. doi:10.3846/tede.2022.17702.

[14] W. Li, J. Zhang, H. Zhu, T. Hao, L. Mei, Y. Su. Green Marketing and the Path to Realizing Local Sustainable Development—Joint Dynamic Analysis of Data Envelopment Analysis (DEA) and Fuzzy Set Qualitative Comparative Analysis (fsQCA) Based on China's Provincial Panel Data. Sustainability (Switzerland) 16 (2024). doi:10.3390/su16114644.

[15] M. Mura, M. Longo, F. Boccali, F. Visani, S. Zanni. From outcomes to practices: Measuring the commitment to sustainability of organisations. Environmental Science and Policy 160 (2024). doi:10.1016/j.envsci.2024.103868.

[16] H.-F. Chien, L.-C. Lu, S.-Y. Chiu, T.-Y. Lin, Y.-H. Chiu. Assessing Social Sustainability Efficiency in European Countries: Focusing on Health and Well-Being. Sustainable Development (2024). doi:10.1002/sd.3295.

[17] A. Maricut, G.-I. Gradinaru, E. Marin, V. Vasile. Efficiency of urban development: gaps in continental profile. Management and Marketing 18 (2023) 398 – 410. doi:10.2478/mmcks-2023-0021.

[18] J. Huang, X. Wang, H. Liu, S. Iqbal. Financial Consideration of Energy and Environmental Nexus with Energy Poverty: Promoting Financial Development in G7 Economies. Frontiers in Energy Research 9 (2021). doi:10.3389/fenrg.2021.777796.

[19] J.-L. Hu, Y.-C. Cheng. ASEAN's efficiency scores in achieving the sustainable development goals. Sustainable Development 32 (2024) 5499 – 5512. doi:10.1002/sd.2981.

[20] Eurostat, Micro & small businesses make up 99% of enterprises in the EU - News articles - Eurostat, 2024. URL: https://ec.europa.eu/eurostat/web/products-eurostat-news/w/ddn-20241025-1.

[21] T. Alam. Assessing the Energy Efficiency of Saudi Arabia's Relatively Middle Eastern Countries in the Context of Sustainable Development Goal Seven (SDG7). International Journal of Energy Economics and Policy 14 (2024) 690 – 696. doi:10.32479/ijeep.17347.

[22] A. S. Rusydiana, N. Laila, N. Tubastuvi, M. A. Ibrahim, L. Marlina. Energy efficiency in oic countries: Sdg 7 output. International Journal of Energy Economics and Policy 11 (2021) 74 – 81. doi:10.32479/ijeep.10499.

[23] A. Roy, A. Basu, X. Dong. Achieving socioeconomic development fuelled by globalization: An analysis of 146 countries. Sustainability (Switzerland) 13 (2021). doi:10.3390/su13094913.

[24] J. Wang, Z. You, P. Song, Z. Fang. Rainfall's impact on agricultural production and government poverty reduction efficiency in China. Scientific Reports 14 (2024). doi:10.1038/s41598-024-59282-2.

[25] I. Al-Ayouty, H. Hassaballa. Towards sustainable development: Measuring environmental total factor productivity in Egypt. European Journal of Sustainable Development 9 (2020) 55 – 68. doi:10.14207/ejsd.2020.v9n2p55.

[26] A. Roy, N. Garai, J. K. Biswas. Exploration of urban sustainability in India through the lens of sustainable development goals. Discover Sustainability 4 (2023). doi:10.1007/s43621-023-00158-2.

[27] I. D'Adamo, C. Daraio, S. Di Leo, M. Gastaldi, E. N. Rossi. Driving EU sustainability: Promoting the circular economy through municipal waste efficiency. Sustainable Production and Consumption 50 (2024) 462 – 474. doi:10.1016/j.spc.2024.08.022.

[28] M. Matulová. RANKING OF EUROPEAN UNIVERSITIES BY DEA-BASED SUSTAINABILITY INDICATOR. Journal on Efficiency and Responsibility in Education and Science 16 (2023) 287 – 298. doi:10.7160/eriesj.2023.160403.

[29] L. M. Perović, M. M. Kosor. The efficiency of universities in achieving sustainable development goals. Amfiteatru Economic 22 (2020) 516 – 532. doi:10.24818/EA/2020/54/516.

[30] M. Blanco, L. Bares, M. Ferasso. Efficiency Analysis of Graduate Alumni Insertion into the Labor Market as a Sustainable Development Goal. Sustainability (Switzerland) 14 (2022). doi:10.3390/su14020842.

[31] B. Singpai, D. Wu. Using a DEA–AutoML approach to track SDG achievements. Sustainability (Switzerland) 12 (2020) 1 – 26. doi:10.3390/su122310124.

[32] M. Pishdar, M. Danesh Shakib, J. Antucheviciene, A. Vilkonis. Interval type-2 fuzzy super SBM network DEA for assessing sustainability performance of third-party logistics service providers considering circular economy strategies in the era of Industry 4.0. Sustainability (Switzerland) 13 (2021). doi:10.3390/su13116497.

[33] F. Li, H. Zhang, D. Zhang, H. Yan. Structural Diffusion Model and Urban Green Innovation Efficiency—A Hybrid Study Based on DEA-SBM, NCA, and fsQCA. Sustainability (Switzerland) 15 (2023). doi:10.3390/su151712705.

[34] F. J. Blancas, I. Contreras. Global SDG composite indicator: A new methodological proposal that combines compensatory and non-compensatory aggregations. Sustainable Development (2024). doi:10.1002/sd.3109.

[35] L. Issever Grochová, M. Litzman. The efficiency in meeting measurable sustainable development goals. International Journal of Sustainable Development and World Ecology 28 (2021) 709 – 719. doi:10.1080/13504509.2021.1882606.

[36] J. Cristóbal, M. Ehrenstein, A. Domínguez-Ramos, Galán-Martín, C. Pozo, M. Margallo, R. Aldaco, L. Jiménez, Irabien, G. Guillén-Gosálbez. Unraveling the links between public spending and Sustainable Development Goals: Insights from data envelopment analysis. Science of the Total Environment 786 (2021). doi:10.1016/j.scitotenv.2021.147459.

[37] V. Skvarciany, I. Lapinskaitė, V. Stasytytė. Efficiency of Digital Economy in the Context of Sustainable Development: DEA–Tobit Approach. Prague Economic Papers 32 (2023) 129 – 158. doi:10.18267/j.pep.824.

[38] S. Thore, R. Tarverdyan. Beyond Gross Domestic Product. Measuring Sustainable Development Goals Performance (2022) 101–114. doi:10.1016/B978-0-323-90268-7.00001-3.

[39] K. Tone. A slacks-based measure of efficiency in data envelopment analysis. European Journal of Operational Research 130 (2001) 498–509. doi:10.1016/S0377-2217(99)00407-5.

[40] D. Morina, P. Gashi. The Role of SMEs on the Economic Development: Kosova's Case. SSRN Electronic Journal (2016). doi:10.2139/SSRN.2820980.

[41] European Commision, SME definition - European Commission, 2003. URL: https://single-market-economy.ec.europa.eu/smes/sme-fundamentals/sme-definition_en.

[42] World Bank, World Bank SME Finance, 2019. URL: https://www.worldbank.org/en/topic/smefinance.

[43] European Commission for Research and Innovation, Turning FAIR into reality – Final report and action plan from the European Commission expert group on FAIR data, Publications Office, 2018. doi:doi/10.2777/1524.

[44] C. Shaerer. The CRISP-DM model: the new blueprint for data mining. Journal of data warehousing 5 (2000) 13–22. URL: https://www.bibsonomy.org/bibtex/5d61c7884c690a61453e519ef8f5de38.

[45] Y. Chen, Y. Chen, A. Oztekin. A hybrid data envelopment analysis approach to analyse college graduation rate at higher education institutions. INFOR: Information Systems and Operational Research 55 (2017) 188–210. doi:10.1080/03155986.2016.1262584.

[46] K. Tone. On the Consistency of Slacks-based Measure-Max Model and Super-Slacks-based Measure Model. Universal Journal of Management 5 (2016) 160 – 165. doi:10.13189/ujm.2017.050307.

[47] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J. W. Boiten, L. B. da Silva Santos, P. E. Bourne, J. Bouwman, A. J. Brookes, T. Clark, M. Crosas, I. Dillo, O. Dumon, S. Edmunds, C. T. Evelo, R. Finkers, A. Gonzalez-Beltran, A. J. Gray, P. Groth, C. Goble, J. S. Grethe, J. Heringa, P. A. t Hoen, R. Hooft, T. Kuhn, R. Kok, J. Kok, S. J. Lusher, M. E. Martone, A. Mons, A. L. Packer, B. Persson, P. Rocca-Serra, M. Roos, R. van Schaik, S. A. Sansone, E. Schultes, T. Sengstag, T. Slater, G. Strawn, M. A. Swertz, M. Thompson, J. Van Der Lei, E. Van Mulligen, J. Velterop, A. Waagmeester, P. Wittenburg, K. Wolstencroft, J. Zhao, B. Mons. Comment: The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data 3 (2016) 1–9. doi:10.1038/SDATA.2016.18;SUBJMETA=479,648,697,702,706;KWRD=PUBLICATION+CHARACTERISTICS,RESEARCH+DATA.

[48] J. H. Zar. Spearman Rank Correlation. Encyclopedia of Biostatistics (2005). doi:10.1002/0470011815.B2A15150.

[49] C. Von Soest. Why Do We Speak to Experts? Reviving the Strength of the Expert Interview Method Why Do We use Expert Interviews? (2001). doi:10.1017/S1537592722001116.

[50] B. Subaeki, A. A. Rahman, K. Manaf, R. H. Laluma, A. Wahana, N. Lukman. Assessing Tax Online System Success: A Validation of Success Model with Focus Group Study. 2022 IEEE 8th International Conference on Computing, Engineering and Design, ICCED 2022 (2022). doi:10.1109/ICCED56140.2022.10010597.

[51] S. Cohen, N. Mayer, J. Rovny, R. Lindstädt. How reliable and how valid are expert surveys? 5 (2022). URL: https://sciencespo.hal.science/hal-04724274v1.

[52] C. Maestas. Expert Surveys as a Measurement Tool: Challenges and New Frontiers. The Oxford Handbook of Polling and Polling Methods (2015) 583–608. doi:10.1093/OXFORDHB/9780190213299.013.13.

[53] W. W. Cooper, L. M. Seiford, K. Tone, Data envelopment analysis: A comprehensive text with models, applications, references and DEA-solver software: Second edition, Springer US, 2007. doi:10.1007/978-0-387-45283-8/COVER.

[54] J. Sarkis. Preparing Your Data for DEA. Modeling Data Irregularities and Structural Complexities in Data Envelopment Analysis (2007) 305–320. doi:10.1007/978-0-387-71607-7{_}17.

[55] R. Benítez, V. Coll-Serrano, V. J. Bolós. Dear-shiny: An interactive web app for data envelopment analysis. Sustainability (Switzerland) 13 (2021) 6774. doi:10.3390/SU13126774/S1.

[56] H. Fei, M. H. Hanif, I. Hanif. How flare-up of small and medium enterprises intensifies carbon emissions in Asian and European regions: a panel analysis. Environmental Science and Pollution Research 30 (2023) 104742–104752. doi:10.1007/S11356-023-29755-4/TABLES/9.

[57] J. D. Gibbons, Nonparametric Statistics: An Introduction , volume 07-090, Sage Publications, 1993.

[58] A. Łozowicka, B. Lach. CI-DEA: A Way to Improve the Discriminatory Power of DEA—Using the Example of the Efficiency Assessment of the Digitalization in the Life of the Generation 50+. Sustainability (Switzerland) 14 (2022). doi:10.3390/su14063610.

[59] M. C. López-Penabad, A. Iglesias-Casal, J. F. S. Neto, J. M. Maside-Sanfiz. Does corporate social performance improve bank efficiency? Evidence from European banks. Review of Managerial Science 17 (2023) 1399 – 1437. doi:10.1007/s11846-022-00579-9.

[60] M. A. Pereira, R. C. Marques. Technical and scale efficiency of the brazilian municipalities' water and sanitation services: A two-stage data envelopment analysis. Sustainability (Switzerland) 14 (2022). doi:10.3390/su14010199.

A

Appendix A: Systematic Literature Review

Table A.1: Table reporting all the articles that have been examined to conduct the research and highlighting their main features

Paper	Main focus	SDG	Analysis technique	Dataset	Input variables	Output variables
[23]	Efficiency of socioeconomic development and globalization on national and global scale	1	BCC-DEA, Malmquist productivity index, and Spearman correlation	-	KOF Globalization index (43 economic, social and political indicators) and the global connectedness index	Education, Employment, Energy, Food, Gender Equality, Health, Networks, Justice, Political voice, and Water & sanitation
[19]	Identify efficient policies from ASEAN and closely related countries to best tackle the SDGs before 2030	All	SBM-DEA	Data on the SDG performance of 21 countries (ASEAN, peers, India, and US)	Economic, Environmental, and Social Variables	Efficiency index for each SDG

Continued on next page

Paper	Main focus	SDG	Analysis technique	Dataset	Input variables	Output variables
[16]	This study investigates the challenges of achieving social sustainability efficiency in European countries, especially in health and well-being	3	An innovative modified circular dynamic three-stage systematic network DEA model	Eurostat database.	Labour Productivity, Number Of Employees, Social Protection Expenditures, Hospital Room Beds, Medical Technology, Healthcare Expenditure, Household Expenditure	Social Protection Benefits, Treatable And Preventable Mortality Rates, And Population
[21]	This study assesses and compares the energy efficiency of Saudi Arabia with that of other Middle Eastern countries in the context of SDG 7	7	CCR-DEA model	World Bank Indicators, 2024.	Energy Consumption, Total Labour Force, Population Total	CO_2 Emissions and GDP
[58]	Article proposes a new DEA model that uses composite indicators for correlated variables: CI-DEA	-	DEA, CI-DEA, PCA-DEA models	-	-	-
[59]	Analyse the impact of corporate social performance on bank efficiency in Europe	-	BCC-DEA	Thomson Reuters Eikon Asset 4 ESG database	Personnel Expenses, Deposits, Fixed Assets, And Average Cost Of Labour	Loans, Earning Assets, And Non-Interest Income
[27]	Paper presents an SLR and DEA of waste management efficiency and recycling in Europe	12	CCR-DEA	Eurostat(2017-2021)	MSW Generation, Bottom of the Waste Disposal, Private Investment related to CE, Persons Employed in CE, and GHD emissions	Recycle Rate and Circular Material Use
[30]	Comparative analysis of the efficiency of labour insertion rate of alumni from Latin American universities	8	BCC-DEA	-	Undergraduate Students, Graduate Students, And Teaching Staff	Qs Graduate Score
[37]	This paper aims to measure the efficiency of the digital economies of EU countries.	1, 2, 7, 13	CCR-DEA	SDG index	Human Capital, Connectivity, Integration of Digital Technology, and Digital Public Services	SDG Index
[17]	The paper aims to identify the degree of efficiency of economic development in the context of sustainable development in urban settings	11	BCC-DEA	OECD data	Urban Area Per Capita, Employment Rate, Labour Productivity, Share of Green Areas in Urban Centres, And Working-Age Population	GDP per Capita and Population Exposure to Air Pollution
[9]	This study aims to assess energy efficiency of Latin American countries	7	BCC-DEA	World bank	Economically Active Population, Gross Fixed Capital Formation, And Total Energy Consumption	GDP
[22]	This study aims to measure the effectiveness of energy access in 50 countries that are members of the OIC	7	BCC-DEA	SESRIC	Labour and Capital	Universal Access to Modern Energy Service, Share of Renewable Energy, and Energy Efficiency
[26]	This paper analyses the efficiency of SDG performance of 56 Indian cities to make policy suggestions	All except 14,15, and 17	BCC-DEA	NITI Aayog Indian city SDG dataset	Environmental SDG's	Economical and Societal SDGs
[18]	Provides analysis of G7 nations by creating energy poverty indexes	7	MCDA and DEA-like model	-	Energy Self-Sufficiency, Energy Dependency, Diversification of Energy Dependency, Energy Consumption, Carbon Emissions Index/Intensity, Renewable Energy, Gdp per Capita, Human Development Index, Forest Area	Composite Index Score

Continued on next page

Paper	Main focus	SDG	Analysis technique	Dataset	Input variables	Output variables
[15]	Making a DEA-based index to assess a company's environmental and social sustainability commitment	9	CCR-DEA	AIDA database	Total Assets, Operating Costs, Environmental Certification, Social Certification	CSR, Production, Water, and Energy
[34]	Propose a new SDG composite indicator that combines compensatory and non-compensatory aggregation rules	All	Customized non-linear DEA	UN SDG indicators, FAO, ILO, OECD, UNICEF, WHO, World Bank	95 SDG Indicators from the UN	SSG Composite Indicator
[14]	Proposes the use of green marketing components to contribute to achieving SDGs on a local level aided by a local SDG index	All	SBM-DEA and Malmquist exponential model	China Statistical Yearbook and China Energy Statistical Yearbook	Full-time R&D Personnel, R&D Funding, and Electricity Consumption	Number of Patent Applications, Green Product Sales Revenue, and Regional Green Land Coverage
[32]	Develop a benchmarking tool for third-party logistic service providers to improve on circular economy and sustainability	9	SBM-NDEA	-	Management Commitment and Capital Employed	Revenue per Year and Trust in Brand
[13]	Perform a combined analysis of eco-efficiency, eco-innovation, and SDGs of 27 EU countries	6 and 7	SBM-DEA and Dynamic Divisional Malmquist index	Eurostat, OECD, Sustainable Development Report, Eco-Innovation Scoreboard	Labour Force, Energy Consumption, GDP, GHG Emissions, and Resource Efficiency	Eco-efficiency, Eco-innovation, and SDG Indices
[12]	Assess of nations in converting their environmental impact into a happy populace with the planetary boundary framework	3 and 13	Custom MILP DEA	World Happiness Report and planetary boundaries report	CO_2 Emissions, Nitrogen and Phosphorous Flows, Blue Water Use, eHANPP, Ecological Footprint, and Material Footprint	Happiness Level
[24]	Evaluates sustainable rural development in China and climate change impacts on this industry.	1 and 9	Custom network DEA	China Statistical Yearbook, the China Agricultural Yearbook, and the China Rural Statistics Yearbook	Sown Area of Crops, Legal Entities in Agriculture, and Financial Support	Agricultural GDP, Deposable Income per Rural Capita, Sustainable Agricultural Development Infrastructure, and Amount of People Covered by Minimum Subsistence Guarantee
[28]	Develop a novel approach to university ranking based on their contribution to sustainable development in European countries	4	Multiple custom MCDM-DEA	UI GreenMetric World University Ranking (2022)	-	6 UI-GMR Sub-indicators
[11]	Evaluate the behaviour of productivity and efficiency of the manufacturing industry of Ecuador in relation to the achievement of some of de SDGs	8, 9, 12	CCR-DEA	Large and SME company balance sheets	Current Assets, Non-current Assets, Cost of Sales, Expenses in Salaries, Operating Expenses, and Non-operational Expenses	Yearly Income and Profit
[33]	Evaluate the impact of various factors on the efficiency of urban green innovation in 101 Chinese cities	11	SBM-DEA, NCA, and fsQCA	Surveys	R&D Investment, R&D Personnel, Electricity Consumption by Cities, and Internet Access	Patents Granted, Sewage Discharge, and Sulphur Dioxide Emissions
[60]	Make an efficiency analysis of water and sanitation services in Brazilian municipalities	6	BCC-DEA	SNIS water service database	OPEX per Year	Water Connections, Sewage Connections, Consumed Water, Collected Wastewater, Treated Wastewater, Length of Water Supply Network, Length of Wastewater Network

Continued on next page

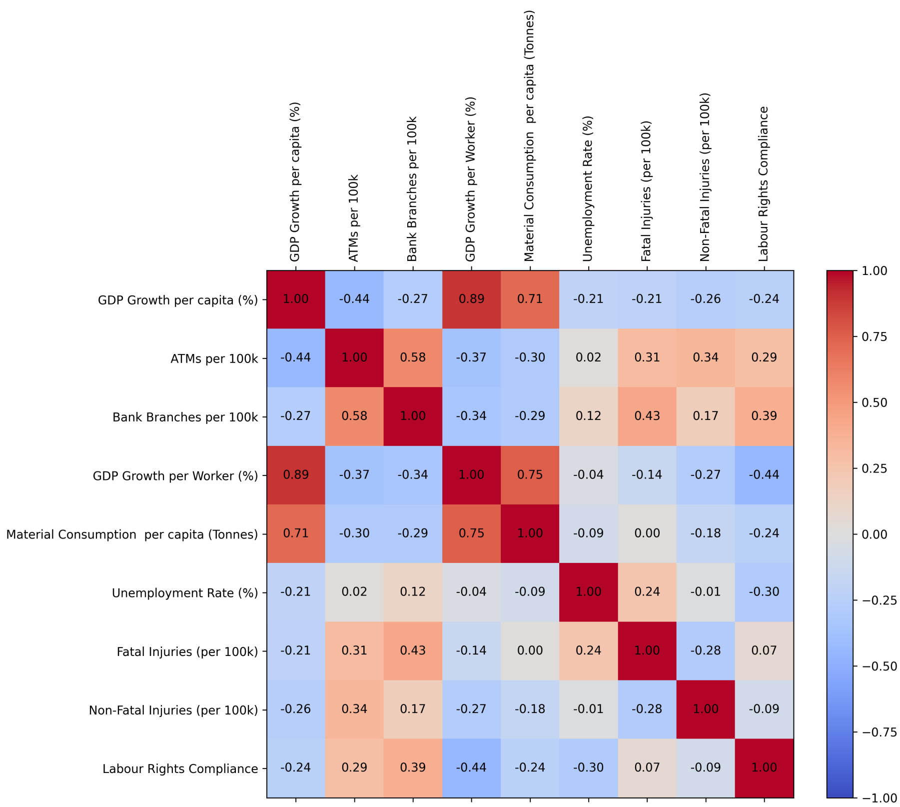
Paper	Main focus	SDG	Analysis technique	Dataset	Input variables	Output variables
[35]	Assess the level of SDG achievement with DEA of each country and the chances that they will achieve their SDG goals	All	BCC-DEA	UN dataset	SDG Indicators that should be minimized (mainly SDG 3)	SDG Indicators that should be maximized (mainly SDG 15)
[29]	Investigate the efficiency of public spending on tertiary education on country level and the efficiency of resource use in universities	4 and 17	CCR-DEA	Times Higher Education University Impact Rankings	Tertiary Education Expenditure Percentage of GDP and Teachers in Tertiary Education per Capita	University Impact Ranking Score and SDG 17 Score
[25]	Analyses the efficiency in environmental total factor productivity of 10 energy-intensive industries in Egypt	9	CCR-DEA	CAPMAS, World Bank	Employees by year, Value of fixed assets, and Expenditure on energy	Value of energy output and CO_2 Emissions
[36]	Propose a mapping to link public spending with achieving the SDGs and assess public spending efficiency	All	SBM-DEA	2019 SDGID	UN SDG Indicators and Government Expenditure	SDG Performance
[31]	Integrating DEA with AutoML to assess and predict SDG performance	7, 8, 9	BCC-DEA	The Belt and Road Initiative	Energy Consumption, Population, Labour Force, Adult Population, Total Employment, and Gross Value Added	Access to Clean Fuel, Access to Electricity, Renewable Energy Consumption, GDP, Domestic Material Consumption, Unemployment, Air Freight, Air Passengers, Manufacturing Unemployment, Manufacturing Value Added, CO_2 Emissions, Internet Users, and Mobile Phone S ubscriptions

Table A.2: Goal of each SDG, source: https://sdgs.un.org/goals

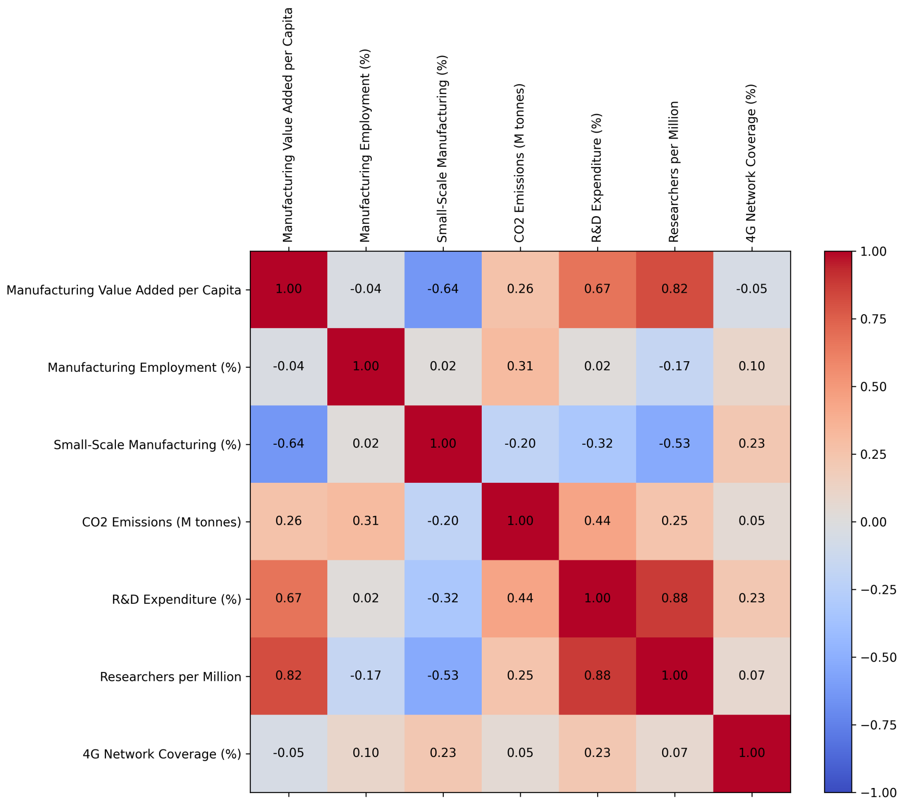
SDG	Name	Goal
1	No poverty	End poverty in all its forms everywhere.
2	Zero hunger	End hunger, achieve food security and improved nutrition and promote sustainable agriculture.
3	Good health and well being	Ensure healthy lives and promote well-being for all at all ages.
4	Quality education	Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all.
5	Gender equality	Achieve gender equality and empower all women and girls.
6	Clean water and sanitation	Ensure availability and sustainable management of water and sanitation for all.
7	Affordable and clean energy	Ensure access to affordable, reliable, sustainable and modern energy for all.
8	Decent work and economic growth	Promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all.
9	Industry, innovation, and infrastructure	Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation.
10	Reduced inequalities	Reduce inequality within and among countries.
11	Sustainable cities and communities	Make cities and human settlements inclusive, safe, resilient and sustainable.
12	Responsible consumption and production	Ensure sustainable consumption and production patterns.
13	Climate action	Take urgent action to combat climate change and its impacts.
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, sustainably manage forests, combat desertification, and halt and reverse land degradation and halt biodiversity loss.
16	Peace, justice, and strong institutions	Promote peaceful and inclusive societies for sustainable development, provide access to justice for all and build effective, accountable and inclusive institutions at all levels.
17	Partnership for goals	Strengthen the means of implementation and revitalize the Global Partnership for Sustainable Development.

B

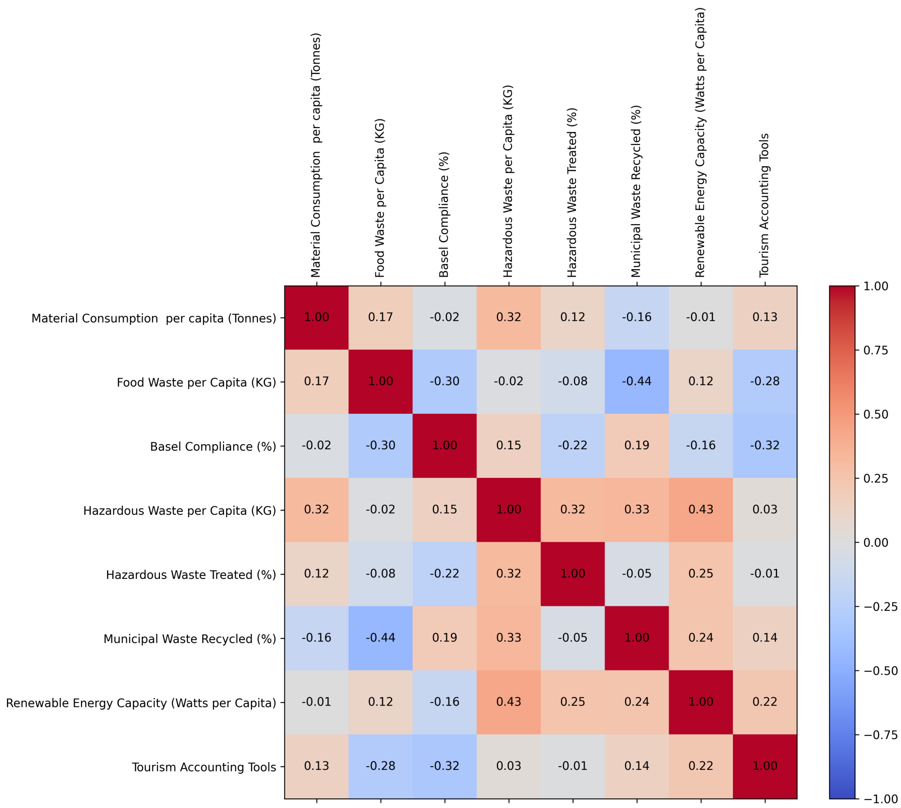
Appendix B: Experimental Set-up



(a) SDG 8



(b) SDG 9



(c) SDG 12

Figure B.1: Detailed Spearman heatmaps for SDGs 8 (a), 9 (b), and 12 (c)