AI-Driven Material Optimization: A Framework for Sustainable Fashion Product Development

Paige Leeds

23044656

A Masters Project submitted in partial fulfillment of the requirement for the award of the degree MSc Fashion Analytics and Forecasting

Fashion Business School

London College of Fashion

University of the Arts London

Total Word Count: 11,775

Submission Date: 21st November 2024

i. DECLARATION

No portion of the work referred to in this Masters Project has been submitted in support of an application for another degree or qualification of this institution or any other university or institution of learning.

In the writing of this Masters Project, I have received assistance from my supervisor, Arnab Banarjee.

I, Paige Leeds, certify that this is an original piece of work. I have acknowledged all sources and citations. No section of this MSc project has been plagiarized.

ii. ACKNOWLEDGMENTS

Firstly, I would like to thank my supervisor, Arnab Banerjee, for his guidance throughout this Masters project. He provided invaluable insights into my research, and made the entirety of the process far less daunting.

Secondly, I would like to thank my course leader, Dr. Satya Banerjee, for igniting my curiosity for AI and data analytics over the past year, creating an incredibly enriching learning experience.

Thirdly, I would like to thank Ryan Guan for teaching me Python and assisting with the code necessary to complete this study – his patience and kindness is unparalleled.

I would also like to thank my classmate-turned-friend, Alexandra Channing Kiss, for our cherished body doubling sessions via Facetime, in which we held each other accountable from over 2,500 miles away.

I would like to give an honorable mention to my lovely cats, Winifred and Cordelia, for proofreading my work and providing emotional support in exchange for treats; as well as to my wonderful brother, Ian, for channeling his distaste for run-on sentences into his endless commitment to assist in the editing process.

Finally, I would like to extend my sincerest gratitude to my grandfather and best friend, who has invested in my education time and time again without question, and never fails to express his unwavering love and support for me, without which this would not be possible.

iii. ABSTRACT

Research Context – This study explores the potential of utilizing AI-driven material optimization in product development to reduce the total environmental impact associated with garments, thereby aligning with the Sustainable Development Goals (SDGs) set forth by the United Nations. It examines how predictive algorithms and environmental key performance indicators (EKPIs) can work in tandem to create a tool for more sustainable decision-making in regard to material selection, guided by Kering's EP&L datasets.

Purpose – The aim of this research is to explore the potential of AI-driven models, particularly Random Forest, to aid in material optimization and enable sustainable fashion product development, with a focus on reducing the environmental footprint of garments.

Design/Methodology/Approach – This research adopts a mono-method quantitative approach, utilizing two Kering EP&L datasets to build and evaluate a Random Forest predictive model. The model assesses six EKPIs (GHGs, water consumption, land use, air emissions, water pollution, and waste) to predict the environmental impact of materials used throughout Kering's supply chains. Optimizing the model's performance consisted of outlier removal, logarithmic transformation, hyperparameter tuning, and cross-validation.

Findings – The Random Forest model identified GHG emissions as the most significant predictor of environmental impact, accounting for 58.07% of the model's predictive accuracy upon feature importance analysis. The model achieved an MAE of 3.29 million after logarithmic transformation and a cross-validated MAE of 0.723, indicating its generalizability.

Originality and Value – This study contributes to the limited existing body of research pertaining to the intersection between AI and sustainability in fashion, while offering a novel approach to material optimization. It builds and tests a predictive model powered by machine learning techniques and LCA principles, and provides practical recommendations to fashion brands looking to reduce their environmental footprints through material optimization.

Keywords: AI-driven material optimization, sustainable product development, environmental key performance indicators (EKPIs), sustainable fashion, circular economy, life cycle assessment (LCA), predictive algorithm, Random Forest model

TABLE OF CONTENTS

1. INTRODUCTION.. 1
1.1 Background.. 1
1.1.1 The Environmental Crisis.. 1
1.1.2 AI-Driven Material Optimization in Fashion.. 2
1.2 Rationale and Research Focus.. 4
1.3 Research Aim and Objectives.. 5
1.4 Research Design.. 6
1.5 Outline Structure.. 6
1.6 Contribution of Research.. 7
2. LITERATURE REVIEW.. 9
2.1 Introduction.. 9
2.2 Historical Context of Sustainable Fashion.. 10
2.2.1 Evolution of Sustainable Fashion.. 10
2.2.2 Key Principles of Sustainable Product Development.. 11
2.3 Theoretical Frameworks for Sustainable Fashion.. 13
2.3.1 Circular Economy Theory.. 13
2.3.2 Life Cycle Assessment (LCA).. 15
2.4 AI-Driven Material Optimization in Fashion.. 20
2.4.1 Predictive Algorithms.. 20
2.4.2 Material Optimization.. 21
2.4.3 Challenges to Adopting AI in Fashion.. 22
2.5 Conclusions and Hypotheses.. 23
3. RESEARCH DESIGN.. 25
3.1 Research Philosophy and Approach.. 25
3.1.1 Epistemological Position.. 25
3.1.2 Research Approach.. 26
3.2 Research Methods.. 26
3.2.1 Methodological Choice, Strategy, and Time Horizon.. 26
3.2.2 Data Collection Procedures.. 27
3.2.3 Benchmarks and Comparative Models.. 28
3.2.4 Novelty of the Method.. 29
3.3 Research Ethics.. 30
3.3.1 Data Rights and Permissions.. 30
3.3.2 Ethical Use of AI.. 30
3.3.3 UAL Code of Ethics.. 30
4. DATASET.. 31
4.1 Dataset Introduction.. 31
4.2 Relevance and Characteristics.. 31

4.2.1 Kering Raw Material Intensity........ 31
4.2.2 Kering EP&L Valued Result and EKPI........ 34
4.3 Variables Selected for Analysis........ 35
4.4. Limitations........ 36
5. FINDINGS & ANALYSIS........ 38
5.1 Introduction to Findings........ 38
5.2 Mean Absolute Error (MAE) and Cross-Validation Results........ 38
5.2.1 Model Selection Justification........ 40
5.3 Feature Importance Analysis........ 40
5.4 Comparison of Model Performance with Existing Literature........ 42
5.4 Comparison of Model Performance with Existing Literature........ 43
6. DISCUSSION & CONCLUSIONS........ 45
6.1 Discussion........ 45
6.1.1 Addressing Objective 1: Highlighting Gaps in AI-Driven Material Optimization........ 45
6.1.2 Addressing Objective 2: Analyzing Environmental KPIs........ 45
6.1.3 Addressing Objective 3: Evaluating AI-Driven Optimization........ 46
6.1.4 Addressing Objective 4: Practical Recommendations........ 47
6.2 Conclusions........ 48
6.2.1 Summary of Key Contributions and Fulfillment of Research Objectives........ 48
6.2.2 Research Limitations and Areas for Future Research........ 50
REFERENCES........ 52
APPENDICES........ 57
Appendix One: Individual Learning Agreement........ 57
Appendix Two: Research Ethics Form........ 59
Appendix Three: Kering Raw Material Intensity Dataset........ 61
Appendix Four: Kering Valued EP&L Result and EKPI Dataset........ 62
Appendix Five: Python Script for Random Forest Model Building........ 63
Appendix Six: Digital Consent Form........ 65

LIST OF FIGURES

Figure 1.1: Clothing and Footwear Waste per Year by Disposal Method.................................... 2
Figure 1.2: Market Value of AI in Fashion Worldwide from 2018-2027.. 3
Figure 1.3: Relative Impact on Biodiversity Along the Apparel Value Chain................................ 4
Figure 2.1: The United Nations' Sustainable Development Goals (SDGs)................................11
Figure 2.2: Circular Fashion System.. 14
Figure 2.3: Scope Covered By EP&L Measurements..16
Figure 2.4: Life Cycle Inputs & Outputs of a Product...16
Figure 2.5: Impact Pathway of Environmental Dimensions and Their Effects on Wellbeing....... 17
Figure 2.6: Breakdown of the EP&L Impact Relating to the Products' Life Cycle....................... 18
Figure 3.1: Research Onion...25
Figure 4.1: EP&L Valued Environmental Impact by Material and EKPI.....................................34

LIST OF TABLES

Table 1.1: Chapter-by-Chapter Structure...7
Table 2.1: LCA Strengths & Weaknesses.. 19
Table 2.2: Hypotheses.. 24
Table 4.1: Raw Material Intensity Variables.. 33
Table 4.2: Raw Material Intensity Descriptive Statistics..34
Table 4.3: Variables Selected for Analysis..36
Table 5.1: MAE Values at Different Stages of Model Improvement.. 39
Table 5.2: Feature Importance Analysis... 41

Chapter One
INTRODUCTION

1. INTRODUCTION

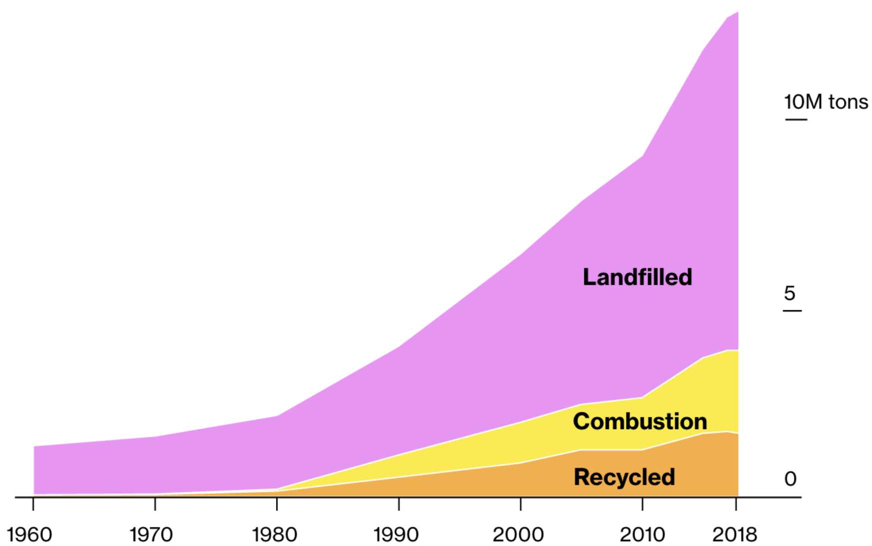
1.1 Background

1.1.1 The Environmental Crisis

The fashion industry is the second most polluting industry globally, largely due to the amount of waste it generates. The end-of-life options for garments are primarily unsustainable, with over “87% of … clothing [ultimately being] incinerated or sent to landfill” (Dottle and Gu, 2022), as illustrated by Figure 1.1. Fast fashion brands rely heavily on polyester to maintain low costs and quick production timelines. However, its frequent blending with natural fibers makes these garments nearly impossible to recycle, contributing to the fact that less than 1% of clothing is effectively recycled (Ellen MacArthur Foundation, 2021). This parallels Bloomberg’s claim that “the U.S. throws away up to 11.3 million tons of textile waste each year” (Dottle and Gu, 2022), highlighting the scale of the fashion industry’s waste problem.

Breaking into the existing cycle requires innovative solutions in numerous steps of the value chain, notably in the product development sector, during which a garment’s end-of-life can be considered prior to its construction. The material selection process is critical in determining a garment’s overall environmental impact, influencing its recyclability and, therefore, sustainability. Sustainability itself, as defined in the Brundtland Report (WCED, 1987), involves meeting the needs of the present without compromising future generations’ ability to meet their own. However, this definition has evolved to encompass ecological, economic, and social dimensions. In the context of fashion, sustainability often prioritizes ecological considerations, such as waste reduction and emissions, while economic and social factors are less emphasized. These competing priorities highlight the complexity of implementing sustainability in practice, especially within material selection processes. Frameworks such as Circular Economy Theory and Life Cycle Assessment provide valuable tools but also expose tensions between reducing environmental impacts and maintaining production efficiency. Addressing these complexities requires innovative, data-driven approaches.

In alignment with the United Nations' Sustainable Development Goal (SDG) 12 – Responsible Consumption and Production – the emphasis is on creating sustainable production processes that minimize resource use, waste, and pollution throughout the garment's lifecycle (UN, 2015). This goal underscores the need for both circular models of production and responsible disposal options to meet sustainability targets set forth by global environmental frameworks. Fortunately, the modern age encourages the use of AI-driven solutions to aid in this process. Machine learning algorithms, such as Random Forest, have the potential to transform material selection by identifying high-impact environmental variables and optimizing decision-making processes. These tools offer a scalable approach to addressing sustainability challenges, enabling brands to prioritize environmental considerations while maintaining efficiency.



Sources: Environmental Protection Agency, American Apparel & Footwear Association, Council for Textile Recycling, and International Trade Commission.

Figure 1.1: Clothing and Footwear Waste per Year by Disposal Method (Dottle and Gu, 2022)

1.1.2 AI-Driven Material Optimization in Fashion

The integration of artificial intelligence (AI) into fashion is occurring at rapid speed, and transforming current practices at every step of the value chain. According to market research firm Statista, the market value of AI in the fashion industry is forecasted to hit nearly $4.4 billion by 2027, a dramatic increase from $270 million in 2018 (The Insight Partners, 2019), as exhibited in Figure 1.2. Not only does the adoption of AI facilitate massive improvements in efficiency and personalization, but it also empowers businesses to leverage AI‘s capabilities to address sustainability concerns. This alignment with SDG 12’s objectives further positions AI as a tool that can drive efficient resource use and encourage responsible production patterns.

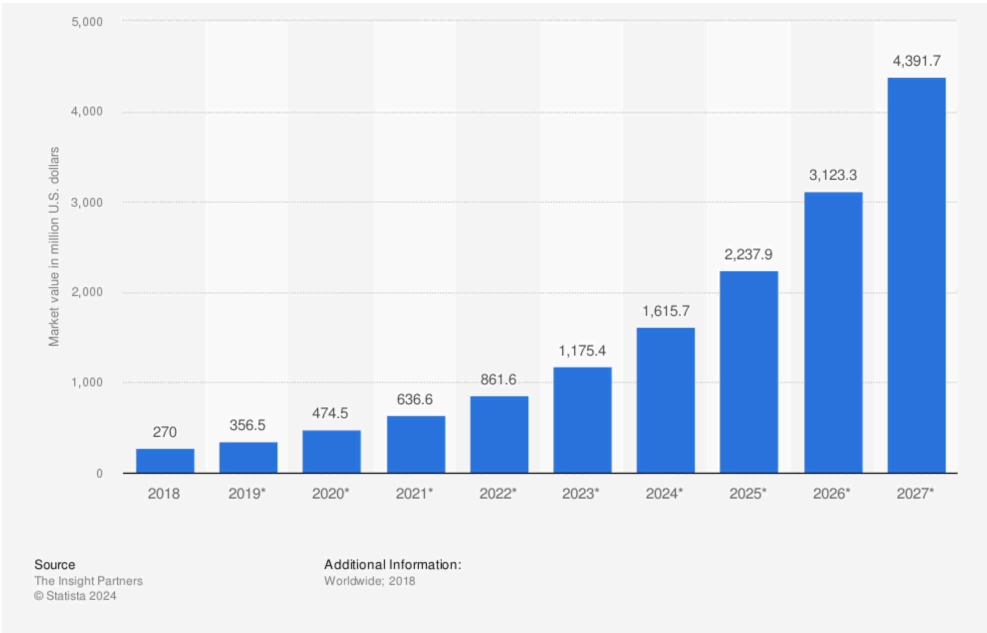


Figure 1.2: Market Value of AI in Fashion Worldwide from 2018-2027 (The Insight Partners, 2019)

The goal of material optimization is to replace unsustainable components of a garment with more sustainable ones, using as few materials as possible to aid in garments’ recyclability. Although there are tradeoffs, such as price, durability, and availability, the focus of this study is on material optimization’s influence on sustainability specifically. Each step of the value chain has environmental implications (see Figure 1.3), each of which can be significantly reduced by making sustainable material choices prior to a garment’s construction. Through considering a number of environmental key performance indicators (EKPIs), material optimization has the power to reduce the environmental impact at the product development stage, subsequently reducing the impact at the following stages. The integration of AI in this process enables the creation of predictive models, lending to its dynamic nature and allowing for real-time analysis, in order to align with the ever-evolving definition of ‘sustainability’. By using machine learning techniques, such as Random Forest regression, this research leverages AI’s predictive capabilities to assess and minimize environmental impacts, contributing to a shift towards more data-driven, sustainable material selection.

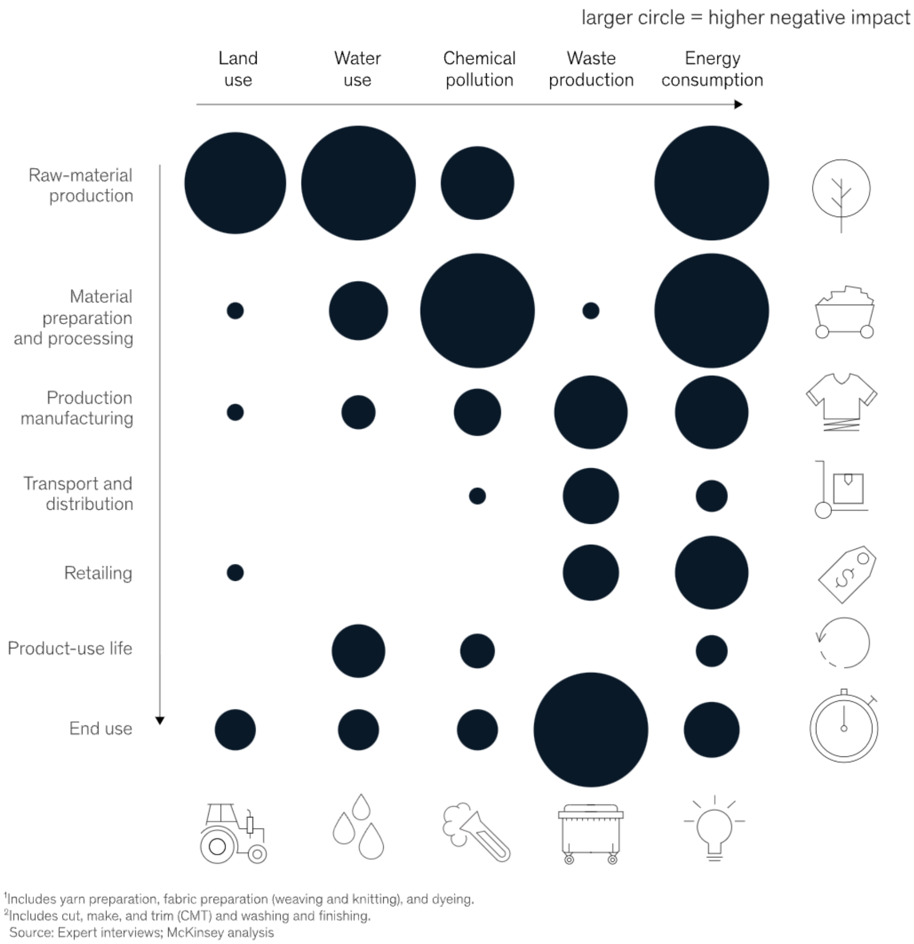


Figure 1.3: Relative Impact on Biodiversity Along the Apparel Value Chain (Granskog et al., 2019)

1.2 Rationale and Research Focus

The rationale of this study stems from the urgent need to address the environmental crisis by reducing the fashion industry’s environmental footprint. AI-driven material optimization provides a solution for fashion brands to make more informed, sustainable choices during the material selection phases of product development. By addressing sustainability concerns prior to garments’ construction – a concept informed by sustainability techniques such as lifecycle assessments – the United Nations’ Sustainable Development Goals (SDGs) could be more easily achieved by 2030. This study, specifically aligned with SDG 12, contributes to an emerging body of literature by evaluating how machine learning algorithms can aid in optimizing resource use, reducing waste, and minimizing emissions.

This research focuses on assessing the sustainability of various materials in Kering's supply chains through their Environmental Profit and Loss (EP&L) datasets in order to build a predictive model to optimize material selection. The aim of the predictive model is to utilize six environmental key performance indicators – greenhouse gasses (GHGs), water consumption, land use, air emissions, water pollution, and waste – to forecast the environmental impact of each material, providing fashion businesses with a framework to make sustainable material choices, thereby reducing the industry's environmental impact.

1.3 Research Aim and Objectives

Research Aim:

The aim of this research is to explore the potential of AI-driven material optimization in enabling sustainable fashion product development, with a focus on reducing the environmental footprint of garments.

Research Objectives:

1. To perform a comprehensive literature review to critically examine the role of AI-driven material optimization in the fashion industry, with a focus on its potential to reduce the environmental footprint of raw materials.

2. To evaluate the environmental impact of different materials used in fashion production across six environmental dimensions from Kering's EP&L report: GHGs, water consumption, land use, air emissions, water pollution, and waste.

3. To develop an AI-driven model that predicts the environmental impact of different materials, guided by insights from Kering's EP&L report, focused on uncovering patterns that optimize material selection for sustainable fashion product development.

4. To provide recommendations for fashion brands on how to integrate AI-driven material optimization into their product development strategies, with the intention of reducing their environmental footprints.

1.4 Research Design

This study takes a positivist epistemological position and deductive research approach in conducting a cross-sectional, mono-method quantitative study to evaluate the environmental impacts of various materials. It relies on secondary data from Kering, utilizing two datasets from the Kering EP&L report: ‘Raw Material Intensity 2023’ and ‘EP&L Valued Result and EKPI 2023’. These datasets provide a comprehensive breakdown of key environmental KPIs, such as GHG emissions and water pollution, across different material types. Ultimately, this data is utilized to build a predictive algorithm – using machine learning techniques such as Random Forest – in order to forecast the environmental footprints of various materials in product development. The goal is to generate data-driven insights that allow brands to make more sustainable material choices during product development stages, thereby reducing the environmental impact of garments and aiding in industry-wide sustainability efforts in compliance with the United Nations’ SDGs.

1.5 Outline Structure

The study follows a traditional chapter-by-chapter structure, with an overview of each chapter provided in Table 1.1.

Chapter	Summary
Chapter 1: Introduction	• Provides background on the environmental crisis in the fashion industry and the role of AI-driven material optimization in addressing sustainability concerns • Outlines the study's rationale and research focus, aim and objectives, research design, and intended contribution of research
Chapter 2: Literature Review	• Examines existing literature regarding the intersection of AI and sustainability in the fashion industry, focusing on the potential to reduce waste via material optimization in product development stages • Outlines the theoretical concepts and frameworks set forth by existing literature, such as Circular Economy theory and Life Cycle Assessment (LCA) principles to inform the hypotheses central to this study
Chapter 3: Research Design	• Discusses the research design and its role in addressing the research aim through the utilization of a positivist, deductive research approach • Details the machine learning techniques used to create the AI-driven model, including Random Forest, feature importance analysis, and cross-validation • Notes the ethical considerations present in the study
Chapter 4: Dataset	• Introduces the two datasets from Kering's EP&L report used for analysis, and examines their suitability and relevance for the study • Examines key features through descriptive statistics and visualizations • Notes the limitations of the dataset and how this may affect the analysis
Chapter 5: Findings & Analysis	• Presents the findings from the AI-driven model, analyzing its effectiveness for predicting the environmental footprints of each material based on their EKPIs • Notes the key features, strengths, and limitations of the predictive model
Chapter 6: Discussion & Conclusions	• Discusses the research findings, implications on the fashion business, and theoretical contributions of this study • Outlines recommendations for fashion businesses, outcomes of the hypotheses, research limitations, and potential future research avenues

Table 1.1: Chapter-by-Chapter Structure

1.6 Contribution of Research

This study aims to make several managerial and theoretical contributions, specifically in the fields of AI-driven sustainability and fashion product development. It builds on the limited existing body of research and provides an innovative, data-driven approach to material optimization. The predictive model generated in this study through the use of machine learning techniques (e.g., Random Forest), evaluates environmental KPIs to provide the environmental footprint of various materials, allowing product developers to make more informed, environmentally-conscious material choices. This novel approach to material optimization contrasts from previous techniques due to its capacity to be applied to real-time data, rather than static, historical data.

This study aims to provide a framework for fashion industry professionals to approach material selection, guided by Kering's vast amount of data, and bolstered by the advanced capabilities of the AI-driven predictive model. Ultimately, the framework has the potential to help fashion brands work toward achieving the UN's Sustainable Development Goals by 2030 by providing the data-driven insights critical to reducing their environmental impact.

Chapter Two
LITERATURE REVIEW

2. LITERATURE REVIEW

2.1 Introduction

This literature review aims to explore the intersection of sustainable product development and artificial intelligence (AI) within the fashion industry, with a focus on material optimization. Amidst growing consumer awareness and regulatory pressure to combat environmental concerns, the fashion industry is facing an urgent need to innovate and adopt sustainable practices by 2030 (UN, 2015). Sustainable product development is the primary approach to this industry-wide call to action, integrating Circular Economy Theory and Life Cycle Assessment (LCA) principles to holistically reduce waste and lower the industry's environmental impact worldwide. This review examines the evolution of the term 'sustainable development' from a historical perspective, and outlines the key principles informing current practices.

As sustainability efforts gain traction, the fashion industry's alignment with Sustainable Development Goal 12 (SDG 12) becomes critical. SDG 12 emphasizes responsible consumption and production, calling on industries to innovate and reduce their environmental impact. This context not only frames sustainable product development but also strengthens the argument for integrating environmental considerations into all stages of fashion production. The examination of SDG 12 further reinforces the necessity of sustainable practices and aligns them with global sustainability goals.

The implementation of AI technologies in sustainable product development has become increasingly relevant in recent years, appearing in a number of sectors within the fashion industry. This review assesses the key principles and frameworks underpinning sustainable product development, including the role of predictive analytics and machine learning algorithms, such as Random Forest, in optimizing material selection for reduced environmental impacts. The challenges and barriers to adopting AI within the fashion industry and its potential role in streamlining material optimization are also explored. After synthesizing the literature, two hypotheses for the study are presented.

2.2 Historical Context of Sustainable Fashion

2.2.1 Evolution of Sustainable Fashion

The term ‘sustainable development’ has appeared throughout environmental conventions over the past several decades, its definition having evolved to reflect our deepening understanding of the term ‘sustainability’ (Lozano, 2022). In 1987, the World Commission on Environment and Development (WCED) published the Brundtland Report, entitled ‘Our Common Future’, in which sustainable development is introduced and defined as “development that meets the needs of the present without compromising the ability of future generations to meet their own needs” (WCED, 1987). In 1992, the United Nations held the United Nations Conference on Environment and Development (UNCED), also known as the ‘Earth Summit’, in Rio de Janeiro, which built upon the previously established definition of sustainable development with the introduction of Agenda 21 and the Rio Declaration (UN, 1992). These documents highlighted the importance of integrating three pillars, or dimensions, of sustainability – “ecology (environment), economy, and social aspects” (Klöpffer, 2003) – into tangible principles and action plans for various industries, challenging the public to reexamine “the way [they] produce and consume, the way [they] live and work, and the way [they] make decisions” (UN, 1992).

In 2002, the United Nations held the World Summit on Sustainable Development in Johannesburg, in which the three pillars of sustainable development were reinforced with an emphasis on poverty eradication and corporate responsibility in leading this transition toward sustainable production and consumption (UN, 2002). In 2015, the United Nations announced its most recent sustainability initiative known as the 17 Sustainable Development Goals (SDGs), “a universal call to action to end poverty, protect the planet, and ensure that by 2030 all people enjoy peace and prosperity” (UN, 2015). Not only does this global framework holistically integrate the three pillars of sustainable development, but it also provides clear, measurable targets for companies to track their progress until 2030 in order to address a wide range of environmental, economical, and social issues (UN, 2015). The 17 SDGs within this ongoing initiative are illustrated in Figure 2.1.

Figure 2.1: The United Nations' Sustainable Development Goals (SDGs) (UN, 2015)

The primary focus of sustainable product development in the fashion industry in particular is addressing SDG 12, 'Responsible Production and Consumption', and SDG 13, 'Climate Action', both aided by the development of sustainable fashion products (Thakker and Sun, 2023). Based on the three pillars of sustainability, a sustainable product can be characterized as a product which takes into account the environmental footprint of its entire lifecycle from manufacture to use to disposal (Bicho *et al.*, 2023). Modifying the traditional product development process in order to embrace sustainability requires designers to consider each lifecycle stage in their material and fabrication choices – in other words, considering a product's end-of-life in the product development stages (Curwen, Park, and Sarkar, 2012).

In recent years, the global focus on sustainability has driven companies to adopt these principles more earnestly, addressing environmental concerns in response to both consumer expectations and regulatory pressures. The fashion industry, historically criticized for its environmental footprint, now faces a heightened responsibility to innovate and develop eco-friendly practices. Examining the evolution of sustainable fashion within this historical and regulatory context strengthens the theoretical foundation for sustainable product development, linking these global sustainability goals to the fashion sector.

2.2.2 Key Principles of Sustainable Product Development

There are a number of core principles that sustainable product development (SPD) relies upon to reduce the fashion industry's environmental footprint and promote sustainability, in turn serving SDGs 12 and 13. Central to SPD are green or eco-design principles, more commonly referred to as sustainable design (SD) principles in recent years (Armstrong and LeHew, 2015). Since the term ‘sustainability’ is multidimensional and allows for many interpretations, different studies elicit varying definitions for what constitutes ‘sustainable design’. Armstrong and LeHew assert that SD’s core goal is to integrate functionality with environmental friendliness, and nature should be central to the SD model, with available resources informing the design process (2015). In addition to functionality and environmental friendliness, Cai and Choi claim that the eco-design/SD procedure should also consider “ergonomics, safety, and economic elements” (2023). They cite the difference between SD and traditional design as the “use of more eco-friendly materials, patterns and techniques”, and emphasize the growing role of designing garments for easier disassembly (Cai and Choi, 2023).

Working in tandem with this design-for-disassembly (DfD) concept is the life cycle assessment (LCA), a key reporting technique that allows the entire lifecycle of a product to be quantified by “calculation parameters with real data on [the varying environmental impacts of] raw materials” (Bicho *et al*., 2023). By conducting and examining the results of an LCA, a brand can choose to remove and replace unsustainable components from their products not only to adhere with sustainability guidelines or brand goals, but also to appeal to the eco-conscious consumer of their brand (Bicho *et al*., 2023). In an Eileen Fisher case study, Curwen, Park, and Sarkar explore the role of the designer in sustainable product development and emphasize the importance of considering a product’s end-of-life prior to construction (2012). This entails choosing product components that can be more easily “disassembled and reused, as well as materials that can be processed into new products (i.e. recycled)” (Curwen, Park, and Sarkar, 2012).

There are a number of sustainable strategies that can theoretically be useful in supporting this concept, such as DfD and design for zero waste (Rathinamoorthy, 2020); however, there is a notable gap in scholarly research examining these strategies in the context of the fashion industry. A study published by Harmsen, Scheffer, and Bos examining the feasibility of different recycling methods found that even advanced sorting technologies could “only process mono-materials” and not material blends (2021). The link between mono-materials and DfD was not clearly established, and more research would be necessary in the fashion industry specifically to draw tangible conclusions (Harmsen, Scheffer, and Bos, 2021).

2.3 Theoretical Frameworks for Sustainable Fashion

2.3.1 Circular Economy Theory

Central to sustainable fashion is the concept of the circular economy, a term that has been gaining prominence to offset the magnitude of the environmental crisis. Dissanayake and Weerasinghe conducted a study involving the synthesis of 64 existing pieces of literature to develop a comprehensive definition for ‘circular fashion’ (2021): “a fashion system that moves towards a regenerative model with an improved use of sustainable and renewable resources, reduction of non-renewable inputs, pollution and waste generation, while facilitating long product life and material circulation via sustainable fashion design strategies and effective reverse logistics processes” (Dissanayake and Weerasinghe, 2021). They also assert that circular strategies can be implemented into a product’s life cycle at various stages, “from raw material selection to reuse or recycling”, as illustrated in Figure 2.2. Their proposed framework emphasizes the critical roles of four techniques in working toward fashion circularity: “resource efficiency, circular design, product life extension, and end-of-life circularity” (Dissanayake and Weerasinghe, 2021). This approach supports the creation of closed-loop systems which keep materials in circulation for as long as possible. Closed-loop systems “aim to decouple [business] growth from resource use” (Hvass, 2014) and feed materials back into the production cycle through recycling or reusing, thus reducing the need for raw material procurement and waste generation.

To further enhance circular economy principles in fashion, additional research is needed to assess long-term benefits and costs associated with implementing these systems on a larger scale. While many brands have demonstrated circular economy strategies in practice, a gap remains in fully quantifying the sustainability impact of these initiatives, both environmentally and financially. Bridging this gap could provide deeper insights into circular economy theory’s practical application in fashion and help establish a more defined framework for its long-term benefits.

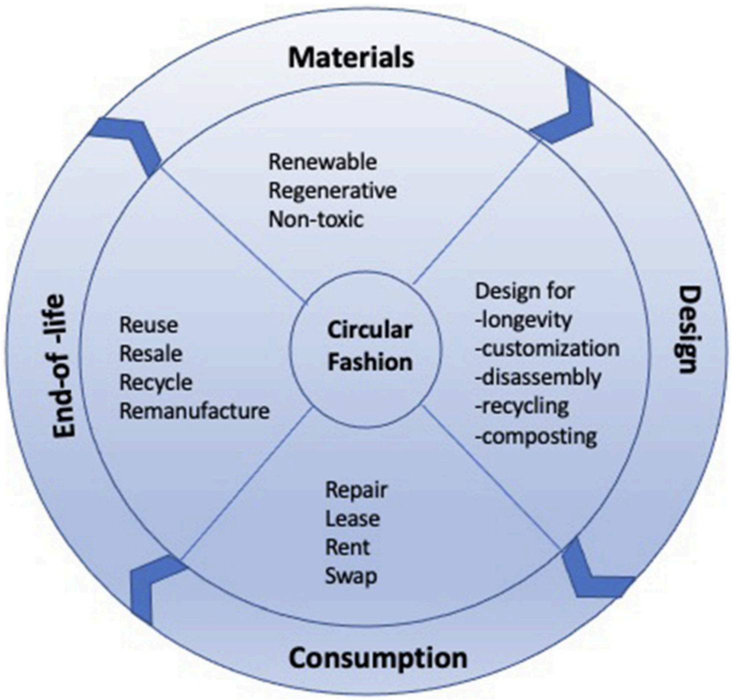


Figure 2.2: Circular Fashion System (Dissanayake and Weerasinghe, 2021)

Jacometti outlines the wasteful nature of the traditional linear economy, which follows a "take-make-dispose" or "production-use-abandonment" model and exacerbates the environmental crisis (2019). Rather than approach sustainable product development through the lens of the United Nations' SDGs, this study focuses on the "EU revised legislative framework on waste adopted within the Circular Economy Action Plan of 2015", providing a different perspective on promoting resource efficiency and waste reduction (Jacometti, 2019). The European Commission states that the goal of the circular economy is "the value of products, materials and resources [being] maintained in the economy for as long as possible, and the generation of waste [being] minimized" (2015). This aligns with the ethos of the United Nations' SDGs with an additional emphasis on waste reduction, due to the assertion that "up to 95% of textiles landfilled each year could be recycled" (Moorhouse and Moorhouse, 2017).

Numerous fashion brands have committed to actioning circularity missions by integrating closed-loop systems into their strategies. Outdoor apparel brand Patagonia has a program called Worn Wear, which incentivizes their consumers to return their used garments for repair or recycling, free of charge (Singh, Park, and Joyner Martinez, 2022). Eileen Fisher, on the other hand, approaches circularity by designing garments specifically for disassembly, ensuring that materials do not end up in landfills (Curwen, Park, and Sarkar, 2012). Even Kering's luxury brands have subscribed to circular economy principles since 2014, prior to the launch of the United Nations SDGs. Balenciaga has the "Second Life fabric initiative … to find uses for its unused fabrics", and Bottega Veneta and Gucci reuse their "waste leather scraps into shoe production or cuttings into organic fertilizer" (Moorhouse and Moorhouse, 2017).

2.3.2 Life Cycle Assessment (LCA)

A comprehensive understanding of a product's full life cycle, from raw material procurement through end-of-life, plays a critical role in sustainable product development. Life Cycle Assessments (LCA) greatly aid in this process by providing a "standardized method of tracking and reporting the environmental impacts of a product or process throughout its full life cycle" (Simonen, 2014). Outlined by the International Organization for Standardization in ISO 14044, a complete life cycle assessment consists of four phases: "goal and scope definition, inventory analysis, impact assessment, and interpretation", each of which can vary in depth and breadth depending on the desired level of detail and specific objectives (ISO, 2006). It is important to note that, although LCAs typically follow a traditional cradle-to-grave approach to align with the long-standing linear economy model, they can also be adapted to fit a cradle-to-cradle approach (in the first phase) to satisfy the objectives of a circular economy.

The stages and scope of LCA, as applied in the fashion industry, are clearly demonstrated in Figure 2.3. This figure illustrates Kering's Environmental Profit and Loss (EP&L) methodology, which evaluates sustainability from raw material procurement (Tier 4) through final assembly and retail operations (Tier 0). Tiers 0 through 4 are assessed in this study, highlighting supply chain phases where targeted interventions can significantly mitigate environmental harm. However, the subsequent use and end-of-life stages remain outside the study's scope, as they are dependent on consumer behavior. This framework emphasizes the interconnectedness of lifecycle stages and the potential for impactful sustainability measures at each step.

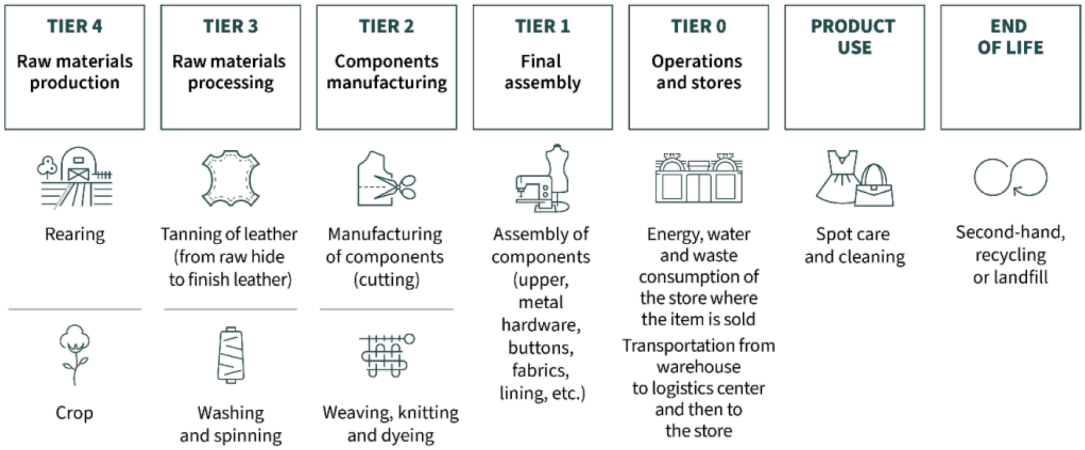


Figure 2.3: Scope Covered By EP&L Measurements (Kering, 2023c)

Regardless, the goal of an LCA is to evaluate inputs (e.g., energy, raw materials) and outputs (e.g., waste, emissions) to assign values to each stage of a product's life cycle (Simonen, 2014). Figure 2.4 complements this detailed mapping by providing a generalized view of lifecycle inputs and outputs. It highlights the broader environmental elements – such as energy use and waste generation – associated with key production phases, offering further context to the application of LCA in sustainable fashion. Together, Figures 2.3 and 2.4 provide a comprehensive perspective on lifecycle considerations. This analysis provides insights to businesses that enable them to make more informed, sustainable decisions throughout their supply chains, thereby reducing their environmental footprint (Kozlowski, Bardecki, and Searcy, 2012).

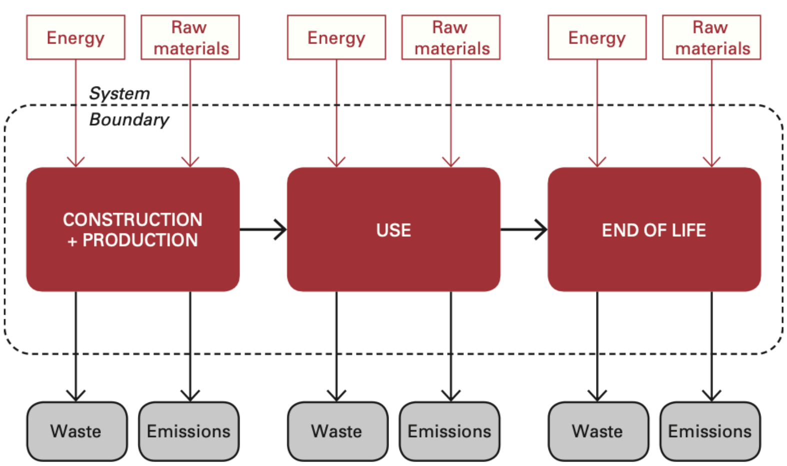


Figure 2.4: Life Cycle Inputs & Outputs of a Product (Simonen, 2014)

To further deepen this lifecycle analysis, Figure 2.5 outlines the six Environmental Key Performance Indicators (EKPIs) defined by Kering's EP&L methodology. These indicators – GHG emissions, water consumption, land use, air pollution, water pollution, and waste – translate environmental harm into measurable societal and ecological costs. For instance, GHG emissions contribute to climate change and economic losses, while water pollution degrades water quality and impacts human health (Kering, 2023c). This granular breakdown underscores how the six EKPIs align with the broader lifecycle considerations presented in Figures 2.3 and 2.4, reinforcing the need for targeted interventions at each stage.

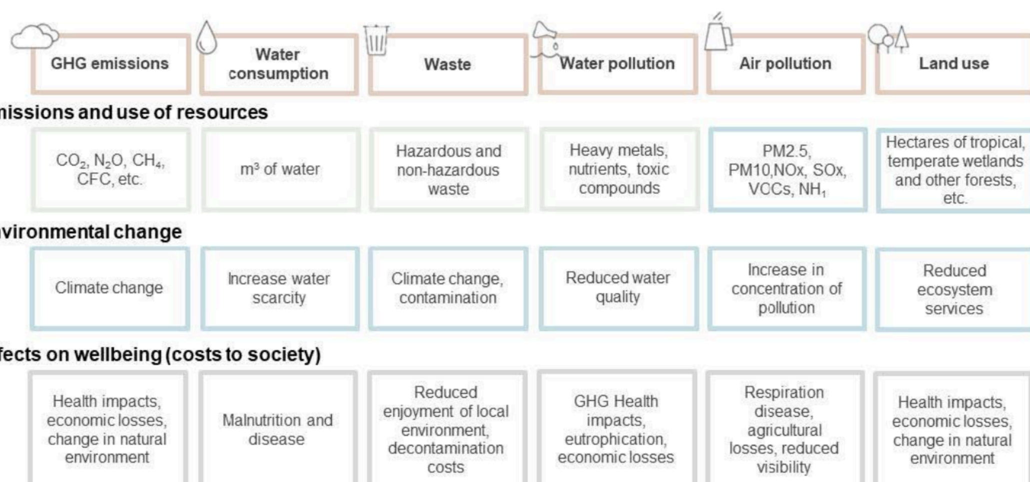


Figure 2.5: Impact Pathway of Environmental Dimensions and Their Effects on Wellbeing (Kering, 2023c)

Building on this framework, Figure 2.6 provides further quantitative insights into the environmental impact of various lifecycle stages in monetary terms. It highlights the disproportionate contribution of raw material production (65.3%) and raw material processing (8.9%) to the total environmental burden, collectively accounting for over 70% of impacts (Kering, 2023e). Such findings reinforce the importance of early-stage interventions, particularly in material sourcing and processing, to address sustainability challenges effectively. This breakdown, provided by Kering's EP&L report, serves as a critical guide for prioritizing lifecycle stages where sustainability efforts would have the most significant impact.

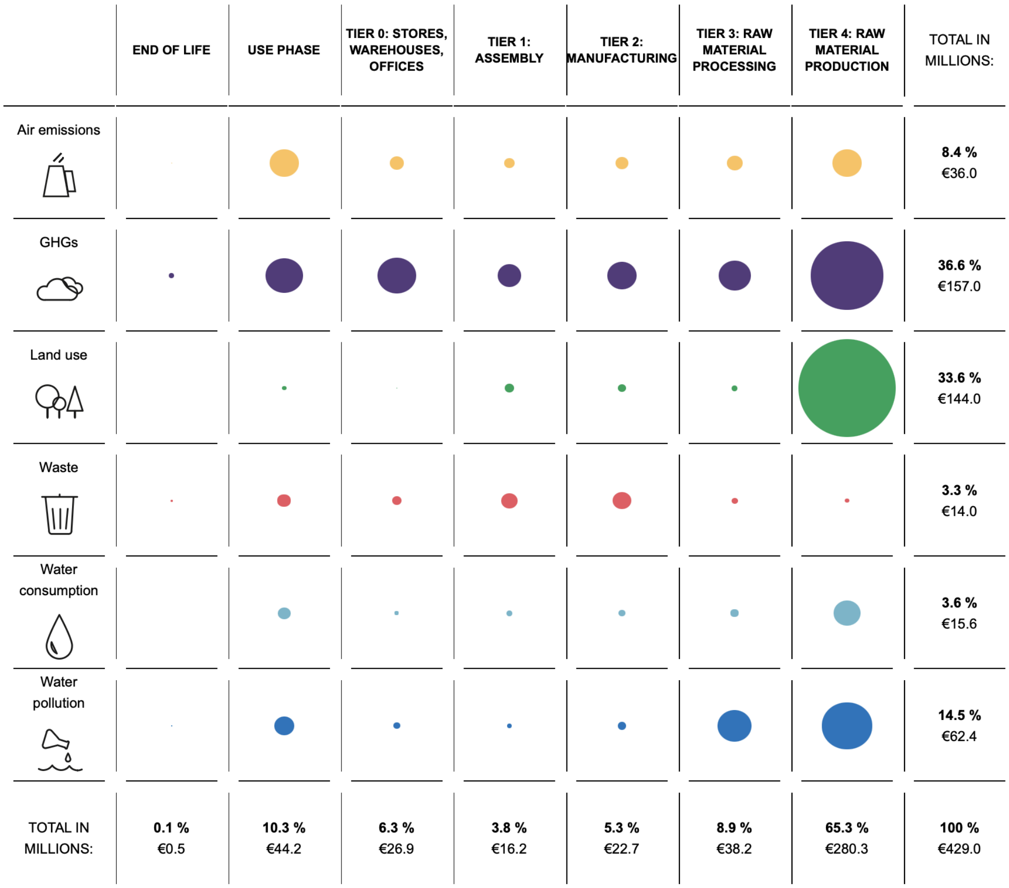


Figure 2.6: Breakdown of the EP&L Impact Relating to the Products' Life Cycle (Kering, 2023e)

Conducting an exhaustive LCA demands a thorough awareness of the environmental impacts of various materials. However, existing literature on LCA studies in the fashion industry often lacks clarity on which specific indicators need to be prioritized. Kering's methodology, guided by LCA techniques, decomposes environmental impact into six environmental dimensions, or environmental key performance indicators (EKPIs): greenhouse gasses (GHGs), water consumption, land use, air emissions, water pollution, and waste (Kering, 2023d). The inclusion of monetary metrics in Figure 2.6 adds another layer of depth, providing actionable insights into the financial implications of environmental harm at each stage. This integration underscores the value of combining qualitative and quantitative LCA approaches to inform sustainability strategies.

As such, their data is divided by material, process step, type of EKPI, and impact country, providing a comprehensive understanding of the origins of their calculations. Despite the challenges of collecting accurate data across global supply chains, LCAs have been implemented by brands like Patagonia and Levi's (Bicho *et al*., 2023) to track and reduce their environmental impacts. Not only do LCAs streamline the product development process by guiding material selection, they also enable brands to meet sustainability targets and align with broader environmental goals, such as the United Nations' SDGs. While LCAs offer numerous strengths and advantages, they also present significant challenges depending on the dataset or desired outcomes (see Table 2.1). A thorough understanding of the objectives of a particular LCA is crucial to determine whether it is the right technique for a study (Simonen, 2014). For example, LCAs require significant resources to implement effectively, making them potentially prohibitive for smaller brands with limited budgets. Addressing these challenges and integrating more accessible LCA methods could increase the adoption of LCA practices across the fashion industry and provide a more standardized framework for evaluating product sustainability.

Strengths		**Weaknesses**	
Quantifiable	Provides measurable metrics to evaluate environmental impacts systematically	Time-consuming	Requires significant resources, knowledge, and data to execute thoroughly
Comparable	Can compare different options if framed correctly, aiding decision-making	Incomplete	Certain local impacts and qualitative aspects are difficult to quantify, and some social and economic impacts aren't effectively considered
Comprehensive	Tracks impacts throughout the supply chain and product lifecycle	Requires Judgment	Judgment calls must be made on data use, inclusions, and recycling models, potentially leading to inconclusive results
Indicative	Offers insights for identifying trade-offs and opportunities for improvement	Uncertain	Assumptions and scenarios, such as changes in product lifespan, can significantly impact LCA outcomes
Motivational	Can drive manufacturers and consumers to make improvements	Faulty Precision	Industry-average data can lead to misleading precision in results
Avoids Green-Washing	Helps minimize unsubstantiated environmental claims through quantitative results	Disguises Green-Washing	Poorly documented methods can lead to misuse of LCA data to support inaccurate environmental claims

Table 2.1: LCA Strengths & Weaknesses (Adapted from Simonen, 2014)

2.4 AI-Driven Material Optimization in Fashion

Various AI technologies have demonstrated their effectiveness in product development across a range of industries. However, there is a significant gap in research regarding AI applications in fashion product development specifically. Some sources cited in this section are related to fashion design rather than product development – two distinct yet related fields – but their concepts can theoretically be applied in a similar context.

2.4.1 Predictive Algorithms

AI technologies have become increasingly prevalent in the fashion industry in the past decade, and have the potential to streamline the product development process through predictive analytics and optimized material selection. Predictive analytics, a field involving AI-driven predictive algorithms, plays a critical role in a brand's ability to forecast material demand, optimize inventory, and reduce waste throughout product development stages (Buus Lassen, la Cour, and Vatrapu, 2016). The main difference between traditional explanatory models and predictive models is that, while the former focus on "causal relationships between theoretical constructs", the latter "rely on associations between measurable variables" (Buus Lassen, la Cour, and Vatrapu, 2016).

One of the most commonly used predictive algorithms is the Random Forest model, which is characterized as a simple model, "using binary splits on predictor variables to determine outcome predictions" (Speiser *et al.*, 2019). This process involves using "randomization to create a large number of decision trees" followed by an aggregation of these trees "into a single output, using voting for classification problems or averaging for regression problems" (Rigatti, 2017). The key strength of Random Forest lies in its ability to "handle datasets with a large number of predictor variables" (Speiser *et al*., 2019), making it highly effective in uncovering meaningful interactions and non-linear effects (Rigatti, 2017), which can significantly enhance the research process.

The key strength of utilizing AI in these predictive tasks is that it has the capacity to produce more accurate, data-driven forecasts, which serve a number of functions within the supply chain: anticipating market demands, optimizing resource allocation, and enhancing sustainability by minimizing overproduction and material waste. Stitch Fix, for example, utilizes this technology to predict consumer preferences by systematically determining whether a consumer will like a particular garment based on the data stored in their profile at signup (Stitch Fix, 2020). In addition to streamlining operations, such data-driven predictions can enable brands to align their production cycles with environmental targets, supporting the United Nations' SDGs by minimizing resource consumption. This concept can and should be adapted to predict which materials or designs will best align with sustainability goals, such as the United Nations' SDGs. However, there is a gap in research regarding the intersection between predictive algorithms, sustainable product development, and the fashion industry that should be further explored.

2.4.2 Material Optimization

AI-driven algorithms are also optimizing the material selection process, a critical component of the product development stage. According to Rathore, companies address sustainability concerns in material selection by choosing to use "eco-friendly materials and methods, choosing recyclable and biodegradable resources over those which cannot break down naturally" (2019a), implying that both circular economy and LCA principles are inherent to material optimization. In a study examining the potential roles of AI in material selection, several supervised and unsupervised machine learning techniques were tested in material selection scenarios (Merayo, Rodríguez-Prieto, and Camacho, 2019). A number of insightful conclusions were made, emphasizing the role of AI in speeding up the material selection process and uncovering complex relationships between variables, favoring the results of supervised over unsupervised learning techniques (Merayo, Rodríguez-Prieto, and Camacho, 2019). However, this study fails to incorporate LCA principles into its algorithm, which is arguably a critical gap in its methodology, as the materials chosen are ultimately responsible for the environmental impacts associated with a garment's entire life cycle. Luxury conglomerate Kering, however, highlights their usage of LCA techniques in the data collection and analysis stages of their yearly EP&L studies, as noted in their open-access publication outlining their methodology (Kering, 2023c). Having a large amount of data is critical to building a model that effectively predicts material selections with accuracy (Luce, 2018); thus, Kering's extensive datasets provide a robust foundation to optimize sustainable material choices through the creation of a predictive algorithm.

As AI-driven material optimization tools become more advanced, incorporating LCA principles can greatly enhance their sustainability impact. Ensuring that algorithms consider the environmental implications of different materials at each lifecycle stage could allow brands to make more informed choices that align with global sustainability goals. Although beyond the scope of this research, it is also important to note the increasingly relevant field of generative artificial intelligence (GAI) as it pertains to material optimization. This unsupervised learning technique, although able to be implemented at nearly every stage of the product's lifecycle, is particularly useful for design and product development due to its ability to generate a wealth of creative content – such as text, images, audio, code, and other forms of media – mimicking human output (Feuerriegel *et al*., 2023). As such, it can be utilized to generate innovative, sustainable material ideas, although there is a gap in research regarding its use in the fashion industry for material design specifically. However, material scientists have been conducting studies utilizing Open AI's ChatGPT to "generate novel materials by embedding physical rules and leveraging large datasets", laying the groundwork for further exploration into the intersection between GAI and sustainable fashion product/material development (Liu *et al.*, 2023).

2.4.3 Challenges to Adopting AI in Fashion

There are a number of technical, economic, and ethical challenges to the adoption of AI, regardless of industry. Bharati Rathore asserts that the fashion industry's adoption of AI "involves overcoming technological hurdles, respecting consumer privacy, and managing the significant investment costs" (2019b). Firstly, there is a lack of sufficient, high-quality data available, making the training of AI models a challenge, especially for smaller corporations with less historical data and fewer resources. The investment of time and money required to integrate AI into a brand's operations and strategy – whether a large or small corporation – poses a significant challenge as well, since the rerouting of a brand's infrastructure and the training of its employees to accommodate AI can temporarily disrupt the flow of business (Saponaro *et al*., 2018). This barrier is further compounded by the fact that smaller businesses may lack the resources to employ robust AI capabilities, despite the fact that these systems could significantly enhance sustainability efforts.

Since product development is a creative field, it is also important to consider that integrating GAI into fashion product development can coincide with the loss of the designer's human, creative touch, inimitable by a machine (Mohiuddin Babu *et al.*, 2020); to combat this concern, it is vital that the designer oversees and intervenes throughout each step of the GAI process. As for the intersection of AI and sustainable fashion product development, it is vital to acknowledge the inherent sustainability trade-offs at play. The utilization of AI models – particularly those using large datasets and advanced machine learning techniques – require a massive amount of computational power, increasing both energy consumption and carbon emissions, thereby creating a paradox in AI's role in sustainable fashion (Mohiuddin Babu *et al*., 2020). Addressing this trade-off remains a complex challenge, as more sustainable computational practices would need to be developed to mitigate the environmental impact of high-energy AI processes. Transitioning to low-energy machine learning models may offer a partial solution to this paradox, making AI adoption more viable from a sustainability perspective.

2.5 Conclusions and Hypotheses

This literature reveals the potential of AI-driven material optimization, bolstered by the Circular Economy theory and LCA principles. The integration of AI with sustainability frameworks presents a unique opportunity to address environmental concerns in the early stages of a product's lifecycle through informed, data-driven material selection. By leveraging AI algorithms, brands can shift from reactive to proactive approaches in material selection, embedding sustainability goals directly into the design and development processes. However, there remains a notable gap in existing literature that examines and categorizes environmental impact into the six types of EKPIs and quantitatively studies the roles of AI and machine learning in this process. Addressing this gap aims to provide critical insights to facilitate the integration of sustainability into the product development processes, laying the groundwork for more sustainable decision-making.

Two hypotheses have been formulated based on the synthesis of the literature, as outlined in Table 2.1; these will be tested and analyzed in the following sections of this research.

Hypothesis 1 (H1)	AI-driven material optimization, using Random Forest models and environmental key performance indicators (EKPIs) such as GHGs, water consumption, and waste, will accurately predict the environmental impacts of different materials used in fashion production, as demonstrated through model performance metrics (e.g., mean absolute error)
Hypothesis 2 (H2)	Among the six environmental dimensions (GHGs, water consumption, land use, air emissions, water pollution, and waste), GHGs will have a significantly higher predictive importance in determining the total environmental impact of materials, as identified through feature importance analysis in the Random Forest model

Table 2.2: Hypotheses

These hypotheses aim to bridge a critical research gap by quantitatively analyzing the predictive power of environmental dimensions in sustainable fashion production. Through empirical testing, this study endeavors to validate whether AI-driven material selection can accurately reflect the environmental footprint of fashion products, thus supporting brands in achieving sustainability benchmarks.

Chapter Three
RESEARCH DESIGN

3. RESEARCH DESIGN

3.1 Research Philosophy and Approach

This study was guided by the ‘research onion’ diagram set forth by Saunders, Lewis, and Thornhill (2019), which was adapted to fit this study (see Figure 3.1). Through this lens, this chapter outlines the research design of the study, from the outermost layer to the innermost layer.

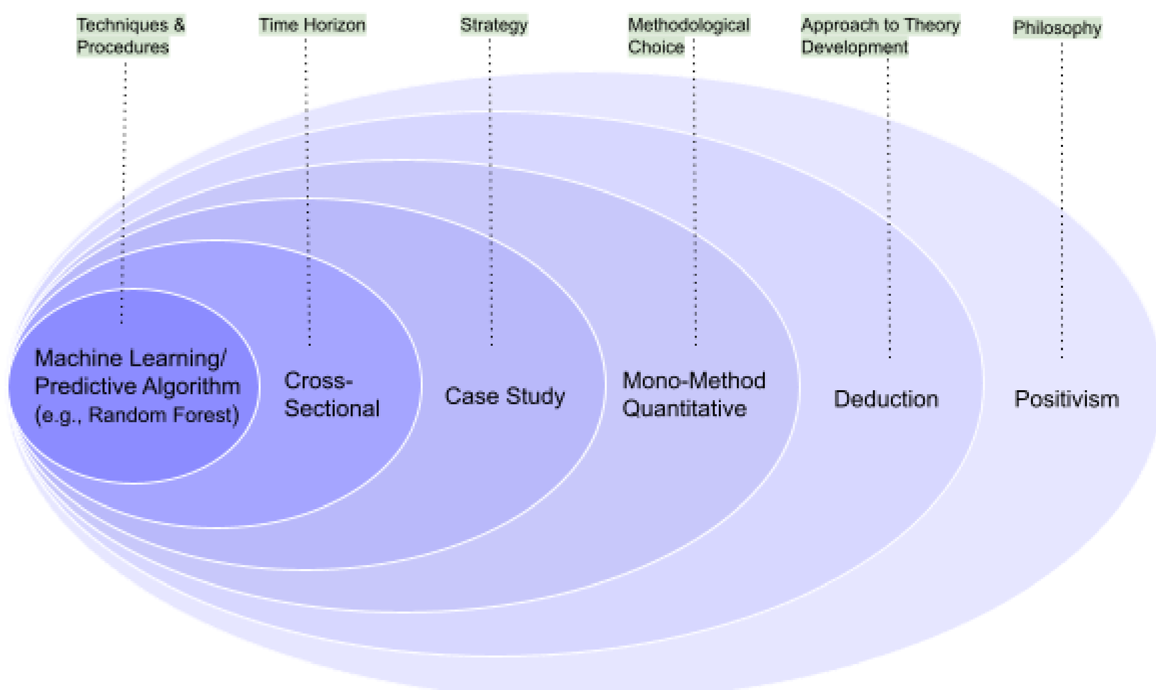


Figure 3.1: Research Onion (Adapted from Saunders, Lewis, and Thornhill, 2019)

3.1.1 Epistemological Position

This study takes a positivist epistemological stance, grounded in objectivism and based on the ontological assumption of realism (Saunders, Lewis, and Thornhill, 2019). Positivism aligns with the perspective of the natural scientist, utilizing existing theories to develop and test hypotheses without researcher biases (Saunders, Lewis, and Thornhill, 2019). This decision was made due to the study's reliance on quantitative data and statistical analyses to examine patterns in the data. By using an objective lens, the study aims to reduce subjective interpretation, thereby increasing replicability and consistency across similar studies. The analyses that emerge through a positivist lens aim to create a replicable study and form law-like generalizations for future research (Saunders, Lewis, and Thornhill, 2019). In this study, two datasets are collated for analysis with the aim of shedding light on the effectiveness of AI-driven material optimization in fashion product development.

3.1.2 Research Approach

This study utilizes a deductive research approach, as the basis of the study is rooted in existing theories or frameworks set forth by previous literature, including Circular Economy Theory and Life Cycle Assessment principles (Saunders, Lewis, and Thornhill, 2019). Deduction implies that hypotheses are formed using findings from previous studies, and these hypotheses are then rejected or accepted by the findings of the present study, a core step of the deductive research approach known as "theory falsification or verification" (Saunders, Lewis, and Thornhill, 2019). By using an objective lens, the study aims to reduce subjective interpretation, thereby increasing replicability and consistency across similar studies. Quantitative studies, in particular, rely on deduction because these have explanatory purposes rather than exploratory ones – they aim to draw connections and explain relationships between variables and are ultimately able to construct generalizations for the population (Saunders, Lewis, and Thornhill, 2019). In this study, through conducting a comprehensive literature review and collating the theories proposed by previous related studies, hypotheses relating to AI-driven material optimization were able to be created and subsequently tested via a deductive process. The application of this approach is intended to provide a structured framework that can facilitate replication by other researchers interested in sustainable product development.

3.2 Research Methods

3.2.1 Methodological Choice, Strategy, and Time Horizon

This study adopted a mono-method quantitative approach, relying on quantitative data analysis for examining the sustainability (i.e. environmental impact) of each material throughout Kering's supply chains. The Random Forest predictive model, an unsupervised machine learning technique equipped to handle large datasets and examine complex relationships between variables, was selected as the primary research method. The aim of utilizing Random Forest was predicting the environmental impact of each material based on a number of environmental KPIs (EKPIs): GHGs, water consumption, land use, air emissions, water pollution, and waste. Random Forest was chosen for its capacity to handle numerous predictor variables and interactions, effectively identifying non-linear relationships in complex environmental datasets. Additionally, its ability to minimize overfitting and provide feature importance metrics made it well-suited to the study's objectives.

Secondary data was relied upon for this study's strategy in the form of Kering's open-source datasets. Kering's ownership of a number of major luxury fashion houses allowed for extensive datasets, which would be unparalleled by any primary data within the time constraints. Their transparency and commitment to sustainability guided by the United Nations SDGs and European Commission also presented a dataset tailored specifically to the research objectives.

The time horizon of this study was cross-sectional, as the data was only studied during one particular period, rather than over an extended period of time. Due to the limited time frame of this study, a cross-sectional study was the logical approach. Although the raw data collected by Kering took place over the course of the year 2023, the analysis of these datasets for this study only took place once. Using a cross-sectional approach allowed the study to capture a snapshot of Kering's environmental performance, offering timely insights aligned with the industry's current sustainability efforts. A longitudinal study may reveal additional patterns in the data for a potential future research avenue.

3.2.2 Data Collection Procedures

Two datasets were utilized in this study, both from Kering's Environmental Profit and Loss (EP&L) report. The Raw Material Intensity dataset provided details on the environmental impact of each material, which were integrated into the analysis. The EP&L Valued Result and EKPI dataset contained environmental impact values alongside the six environmental key performance indicators (EKPIs), which were used as features in the Random Forest model to train it in assessing the total impact values based on provided EKPI values.

Data preprocessing stages were carried out in order to ensure consistency and standardization of the data prior to the creation of the predictive model. First, the two datasets were merged across the Tier, Material Group, Raw Material, Process Step, and Impact Country columns, turning the datasets into one cohesive dataframe. Ensuring consistency between datasets minimized discrepancies and facilitated cleaner data integration, which is essential for accurate model predictions. Missing or null values within the datasets were replaced with 0 in order to account for any potential outliers or skews, thus ensuring the accuracy of the predictive model. Standardization of the data was established by the usage of StandardScaler, a class within the scikit-learn library in Python, which scaled the data to each have a mean of 0 and standard deviation of 1, making the data easier to analyze.

A core step of building the model was feature selection, in which the six EKPIs were each selected and labeled as follows: eKPI_GHGs, eKPI_Water_Consumption, eKPI_Air_Emissions, eKPI_Water_Pollution, eKPI_Waste, eKPI_Land_Use. Several interaction terms were established as additional features (e.g., GHGs_Air_Emissions_Interaction, Land_Use_Water_Consumption_Interaction) to test their importance on overall environmental impact as well, due to Random Forest's affinity for examining complex relationships. Interaction terms were included to capture potential synergies or trade-offs between environmental factors, adding depth to the model's predictive capabilities. The dataset was then split into training (20%) and testing (80%) sets using the train_test_split function on Python. The training set was used to build the Random Forest model, while the testing set was used to evaluate the model's performance through its Mean Absolute Error (MAE).

3.2.3 Benchmarks and Comparative Models

Kering's EP&L report was used as a benchmark in this study, as it contained not only four extensive datasets, but also a detailed account of their methodology. They underscored the role of sustainable development principles – rooted in recommendations set forth by the United Nations and European Commission – in setting their industry-standard framework. The benchmark provides a comparative baseline, allowing for a meaningful evaluation of the Random Forest model's effectiveness relative to existing environmental impact assessments. The valued EKPI intensity results from Kering's datasets served as the static baseline to be compared with the dynamic, real-time forecasts from the predictive model, ultimately evaluating the effectiveness of the algorithm in optimizing material selection for sustainability.

Comparative models primarily consisted of historical data, which caused the results to be static in nature, as opposed to more dynamic, real-time forecasts. The methodology used to create the Kering EP&L report was rooted in a monetary approach – starting with translating each of the six environmental dimensions into a financial value based on their social and environmental implications. They then collected primary data along their operations and value chain, and used standardized LCA techniques and primary data from international databases and suppliers to create a comprehensive assessment of the corporation's entire environmental footprint. This study diverges from Kering's method by prioritizing environmental impact values directly over monetary estimations, maintaining a more granular focus on specific environmental outcomes. The main differences were their usage of monetary terms in their model, as well as the historical (i.e. static) nature of their data.

3.2.4 Novelty of the Method

The novelty of this study primarily lies in its predictive capabilities. Although it relied on insights from Kering's sustainability framework, it advanced their methodology by integrating AI, allowing for real-time forecasts. By providing material impact predictions dynamically, the model offers a proactive rather than retrospective tool, enhancing the decision-making process in sustainable product development. Reactivity in the fashion business is growing increasingly crucial; thus, predictive models can greatly aid product development teams in making more informed, timely decisions in regards to material selection. Forecasting for material selection purposes also allows for the environmental footprint of each garment to be lowered prior to the start of its lifecycle, aligning with LCA principles. Also unique to this study is its usage of interaction terms to examine each of the six environmental dimensions and uncover hidden relationships between them. The intersection of AI and sustainability, particularly in the form of AI-driven material optimization, has only recently emerged and, therefore, has not been widely studied thus far. This study serves as a pioneering effort by showcasing how machine learning can effectively support sustainable practices within the industry.

3.3 Research Ethics

3.3.1 Data Rights and Permissions

The datasets utilized for this study were open-source and publicly available, as they were sourced directly from Kering's website. Kering publishes their yearly EP&L reports with the intention of normalizing transparency within the industry, thus justifying its usage for further analysis. The datasets adhered to the General Data Protection Regulation (GDPR), as no customer data was published – only material data was present, therefore not posing a privacy issue. Additionally, there was no conflict of interest, as the researcher has no affiliation with Kering, ensuring the objectivity of this study and its outcomes.

3.3.2 Ethical Use of AI

The AI-driven models in this study were used responsibly and transparently, with the Random Forest model being used solely for academic and predictive purposes. There was no private or sensitive data being manipulated. The study considers the environmental and social implications of AI, aiming to serve as a framework for sustainable product development in accordance with the United Nations' SDGs.

3.3.3 UAL Code of Ethics

This study was conducted in accordance with the University of the Arts London's Code of Practice on Research Ethics, which is guided by the principles of "respect for persons, justice, and beneficence" (UAL, 2020). An assessment of risk was undertaken prior to research, which yielded incredibly low risk potential due to the remote nature of this study (i.e. no participants outside of the primary researcher). Academic integrity was maintained throughout the research process, and proper credits were granted to the dataset sources, collaborators, and supervisors.

Chapter Four
DATASET

4. DATASET

4.1 Dataset Introduction

The Kering Group is a French-based luxury conglomerate which owns brands such as Gucci, Saint Laurent, Bottega Veneta, Balenciaga, and Alexander McQueen, and is committed to measuring their impact to satisfy the United Nations' Sustainable Development Goals (SDGs) by 2030. With the aim of contributing to the industry-wide "transition to a responsible business model", Kering publishes their yearly Environmental Profit and Loss (EP&L) report to "measure, monetize, and monitor the environmental impact of [their] business activities" (Kering, 2023d). This report consists of four open-source datasets and interactive tools to map the corporation's progress, with the intention of establishing transparency and inspiring their "peers in the luxury and fashion industry" to adopt similar practices (Kering, 2023d). For the purposes of this study, two datasets were utilized for analysis: 'Raw Material Intensity 2023' (Kering, 2023f) and 'EP&L Valued Result and EKPI 2023' (Kering, 2023b).

4.2 Relevance and Characteristics

4.2.1 Kering Raw Material Intensity

The Raw Material Intensity dataset serves as a basis for assessing the sustainability – and thus, environmental footprint – of each material used throughout Kering's brands in each step of the supply chain, from material procurement to disposal (see Table 4.1). To note, only tiers 0 through 4 are accounted for in this dataset, as the environmental footprints through product use and end-of-life stages are dependent on the consumer's behaviors post-purchase, which is beyond the scope of this study. However, analysis of material processes in tiers 0 through 4 allow the predictive algorithm to estimate the total environmental impact of each material, which can be instrumental to material optimization in fashion product development.

This dataset contains 1,722 rows, each representing a unique material/process/country combination; and eight columns, labeled Tier, Material Group, Raw Material, Process Step, Impact Country, Environmental Impact Group, Indicator Unit, and Valued EKPI Intensity. Significant categorical variables are listed in Table 4.1, illustrating the methodology Kering uses to classify each of their materials and which countries play a role within their supply chains. The only numerical variable in this dataset is the Valued EKPI Intensity column, which is explored in more detail in the next dataset. The minimum, maximum, mean, standard deviation, and variance of Valued EKPI Intensity are displayed in Table 4.2, generated via SPSS.

Material Group	Raw Material
• Animal Fibers • Cellulose based Fibers • Leather • Metal • Paper • Plant Fibers • Plastic • Synthetic Fibers • Wood **Process Step** • Abattoir • Animal rearing • Cotton recycling • Crop farming • Dissolving pulp production • Extraction • Ginning • Processing • Production • Pulpwood harvest • Silk production • Spinning Weaving • Staple fiber production • Tanning • Tanning - Metal-free • Washing **Impact Country** • Argentina • Australia • Austria • Belgium • China • Denmark • Egypt • Finland • France • Germany • Indonesia • Italy • Mongolia • Netherlands • New Zealand • Peru • South Korea • Spain • Turkey • Ukraine • United Kingdom • United States	• Animal Fibers - cashmere - conventional • Animal Fibers - cashmere - organic • Animal Fibers - cashmere - regenerated • Animal Fibers - cashmere - restorative grazing • Animal Fibers - silk - conventional • Animal Fibers - silk - organic • Animal Fibers - silk - recycled • Animal Fibers - wool - conventional • Animal Fibers - wool - organic • Animal Fibers - wool - regenerated • Animal Fibers - wool - restorative grazing • Cellulose based Fibers - acetate • Cellulose based Fibers - viscose - certified • Cellulose based Fibers - viscose - uncertified • Leather - beef - calf - conventional • Leather - beef - calf - restorative grazing • Leather - beef - calf - salpa • Leather - goat - conventional • Leather - sheep - lamb - conventional • Leather - sheep - lamb - restorative grazing • Metal - brass - recycled • Metal - gold - 24kt - Kering responsible • Paper - paper and cardboard - recycled • Paper - paper and cardboard - uncertified • Paper - tissue • Plant Fibers - cotton - conventional • Plant Fibers - cotton - organic • Plant Fibers - cotton - recycled • Plant Fibers - hemp • Plant Fibers - linen - conventional • Plant Fibers - linen - organic • Plastic - bioplastic • Plastic - elastane - conventional • Plastic - eva • Plastic - neoprene • Plastic - nylon - eyewear - responsible • Plastic - pe • Plastic - pe - recycled • Plastic - PET - recycled • Plastic - polyurethane - conventional • Plastic - polyurethane - recycled • Plastic - pp • Synthetic Fibers - acrylic - conventional • Synthetic Fibers - elastic • Synthetic Fibers - neoprene • Synthetic Fibers - nylon - conventional • Synthetic Fibers - polyester - conventional • Synthetic Fibers - polyester - recycled • Wood - hardwood • Wood - MDF

Table 4.1: Raw Material Intensity Variables (Adapted from Kering, 2023f)

	N	Minimum	Maximum	Mean	Std. Deviation	Variance
Valued_Ekpi_intensity	1722	.000000000	1438.44595	4.44664086	60.1578261	3618.964
Valid N (listwise)	1722					

Descriptive Statistics

Table 4.2: Raw Material Intensity Descriptive Statistics (SPSS)

4.2.2 Kering EP&L Valued Result and EKPI

The EP&L Valued Result and EKPI dataset assesses the valued EKPI intensity – also present in the material intensity dataset – in greater detail (see Figure 4.1). Kering's methodology involves utilizing six environmental KPIs (key performance indicators) – "air pollution, GHG emissions, land use, waste production, water consumption and water pollution" – to calculate the total valued EKPI intensity, with higher EKPI results indicating higher environmental impact/harm (i.e. lower levels of sustainability). These six KPIs allow for the decomposition of environmental harm across the supply chain, shedding light on areas that require targeted interventions to improve sustainability.

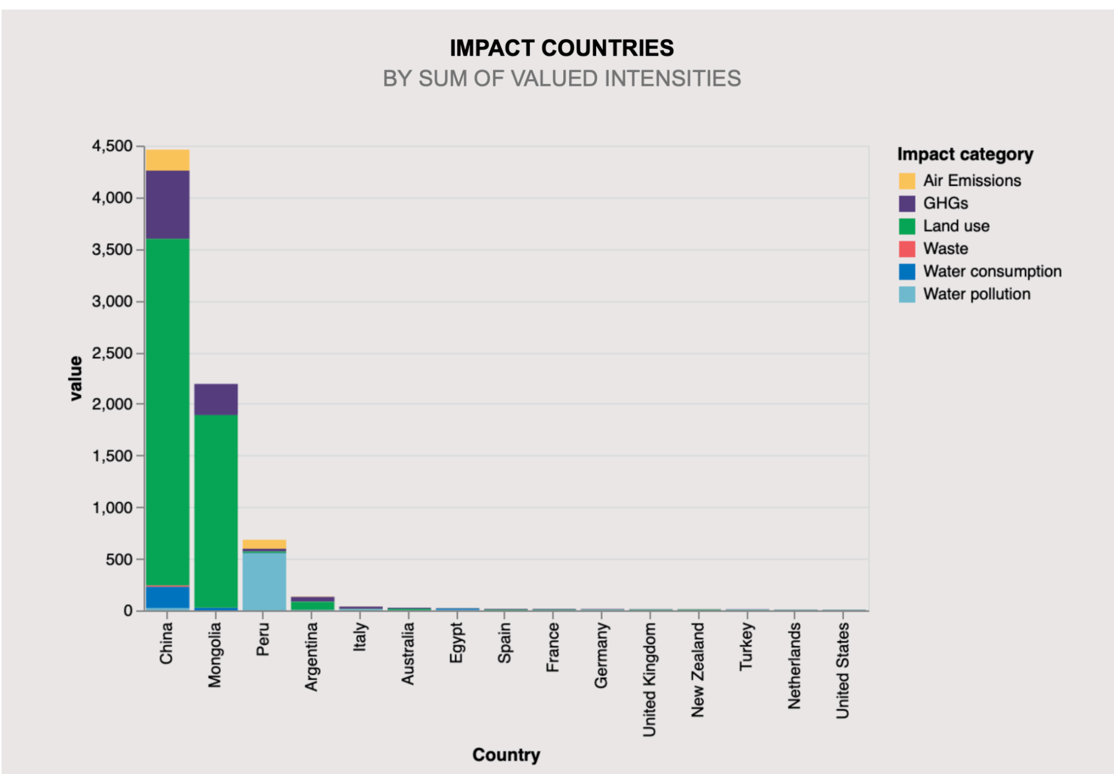


Figure 4.1: EP&L Valued Environmental Impact by Material and EKPI (Kering, 2023a)

The EP&L Valued Result and EKPI dataset contains 141,621 rows and includes 13 columns, such as Business Unit, Raw Material Group, Raw Material, Process Step, Tier, Impact Country, EKPI, EKPI Year, EKPI Result, and Valued Result. These variables provide the granularity necessary for building predictive models to optimize material selection and reduce environmental impact.

4.3 Variables Selected for Analysis

Out of the 21 columns/variables present in the datasets, 13 variables were selected for analysis, as illustrated in Table 4.3. Variables like Tier, Material Group, Raw Material, Process Step, and Impact Country were selected to enable data merging and detailed predictive modeling.

Valued EKPI Intensity (from the Raw Material Intensity dataset) and EKPI Result (from the EP&L Valued Result and EKPI dataset) serve as the primary variables for building the predictive model. The six environmental KPIs, represented as separate features (e.g., eKPI_GHGs, eKPI_Water_Consumption, eKPI_Air_Emissions, eKPI_Water_Pollution, eKPI_Waste, eKPI_Land_Use), were also included, along with interaction terms such as GHGs_Air_Emissions_Interaction and Land_Use_Water_Consumption_Interaction. These interactions allow the model to uncover complex relationships and identify which factors contribute most significantly to overall environmental harm.

Other features, such as Environmental Impact Group, Area, and Population, were excluded as they were not directly relevant to the model’s objectives. However, they may be useful for exploration in future research.

	Variable	Selected for Analysis?
Raw Material Intensity	Tier	Selected
	Material Group	Selected
	Raw Material	Selected
	Process Step	Selected
	Impact Country	Selected
	Environmental Impact Group	
	Indicator Unit	
	Valued EKPI Intensity	Selected
EP&L Valued Result and EKPI	Business Unit	
	Raw Material Group	Selected
	Raw Material	Selected
	Process Step	Selected
	Tier	Selected
	Impact Country	Selected
	EKPI	Selected
	EKPI Year	
	EKPI Result	Selected
	Year	
	Valued Result	
	Area	
	Population	

Table 4.3: Variables Selected for Analysis

4.4. Limitations

The primary limitation of these datasets is their focus on Kering's brands specifically, which are solely of luxury status. Generalizing this level of material usage and sustainability implications to fast fashion brands would need to be further studied, as fast fashion brands tend to use a higher quantity of less costly materials, such as polyester.

Additionally, although Kering emphasizes their commitment to the United Nations' SDGs and European Commission's methodology recommendations for sustainable development, it is important to note that their data is self-reported. This introduces potential biases and should be approached with a degree of skepticism.

Finally, the datasets do not account for the end-of-life phase, a critical stage in a product's lifecycle where recycling and waste management efforts could significantly impact sustainability outcomes. Incorporating these stages into future analyses could enhance the comprehensiveness of sustainability evaluations.

It is important to note that these datasets are published to shed light on overall patterns within the industry, and Kering's acclaim in the luxury fashion sector has the potential to fuel similar environmentally-conscious behaviors of other competing companies.

Chapter Five
FINDINGS & ANALYSIS

5. FINDINGS & ANALYSIS

5.1 Introduction to Findings

This chapter is chronologically organized to align with the steps of the predictive model, noting its key findings and outcomes of the analysis. Deeper implications and recommendations will be further explored in Chapter 6.

The results of the analysis indicate that the Random Forest model was moderately effective at accurately predicting the environmental impacts of the materials used in Kering's supply chains through examination of the six environmental KPIs. The model's performance was evaluated using mean absolute error (MAE), with the final model's results indicating a prediction error of approximately 3.3 million. Cross-validation was performed, further confirming the model's consistency, as evidenced by a cross-validated MAE of 0.723.

While GHG emissions emerged as the most significant predictor (58.07% feature importance), other features such as water consumption (17.62%) and air emissions (14.10%) also played critical roles in the model's predictions. Conversely, certain interaction terms and variables – such as land use and waste – contributed minimally, raising questions about their limited influence, which will be discussed in greater detail.

5.2 Mean Absolute Error (MAE) and Cross-Validation Results

The Random Forest model was applied to predict the total environmental impact based on the provided environmental key performance indicators (EKPIs). In order to evaluate the model's performance at various stages of model building, its MAE was calculated, as illustrated in Table 5.1. MAE was selected as the regression metric due to its interpretability and applicability to log-transformed data, which was a vital step of optimizing this model. The role of MAE in this analysis was to reveal how far off the model's predictions were from the true outcomes.

The table below outlines the MAE results at each stage of model building, from the initial Random Forest model to the fully optimized model. As expected, the MAE values decreased after each step of model improvement.

Step	Technique Applied	MAE Value
Initial Model	Baseline Random Forest	30,172,019.74
After Outlier Removal	Removed extreme outliers	14,950,305.07
After Logarithmic Transformation	Applied log transformation on target	3,296,096.55
Optimized Model	Hyperparameter tuning (GridSearchCV)	3,296,096.55
Cross-Validation	Random Forest with Cross-Validation	0.723 (on log-transformed data)

Table 5.1: MAE Values at Different Stages of Model Improvement (Adapted from Spyder/Python)

As shown, the baseline Random Forest model yielded a high MAE of 30,172,019.74, reflecting the challenges of working with raw, unprocessed data. Removing extreme outliers significantly improved the MAE to 14,950,305.07, suggesting that the dataset contained notable anomalies affecting prediction accuracy. Following the logarithmic transformation to the target variable (Total_Impact), the MAE was further reduced to 3,296,096.55, demonstrating the effectiveness of preprocessing techniques in improving model performance.

Following hyperparameter tuning using the GridSearchCV function in Python, the optimized model maintained the same MAE of 3,296,096.55, indicating that the logarithmic transformation had already selected the optimal parameters for the model in the preprocessing steps. This outcome highlights the importance of foundational data engineering in achieving robust predictive results.

Cross-validation confirmed the model's stability, with a low cross-validated MAE of 0.723, affirming its generalizability across subsets of the dataset. Despite these strengths, the model's limitations – such as its reliance on decision tree structures and potential overfitting to non-linear interactions – warrant further consideration, particularly for poorly performing features like land use and waste.

5.2.1 Model Selection Justification

The selection of the Random Forest algorithm was guided by its specific advantages in addressing the objectives of this study. Random Forest is a robust ensemble learning method that reduces overfitting and improves prediction accuracy by combining multiple decision trees and averaging their outputs. Its capacity to handle large datasets with high-dimensional features made it an ideal candidate for modeling the environmental KPIs in this study. Additionally, Random Forest inherently provides feature importance rankings, offering critical insights into the relative contributions of each KPI to the total environmental impact.

While alternative models like Gradient Boosting were considered, Random Forest was selected for its ability to build decision trees in parallel and average their predictions, which helps prevent overfitting and enhances generalizability. Additionally, it provides built-in feature importance metrics for straightforward interpretation, unlike Gradient Boosting, which often requires additional techniques as well as extensive fine-tuning. Random Forest proved effective with minimal hyperparameter tuning, as shown by the significant MAE reduction after logarithmic transformation. Finally, Random Forest's parallel structure ensured computational efficiency when applied to this large dataset, making it an ideal choice for predictive performance and interpretability, aligned with the study's objectives.

In conclusion, the success of Random Forest in this context highlights its potential for real-time environmental impact assessments in sustainable fashion product development. Future research could further explore alternative models, such as Gradient Boosting, to determine whether their iterative learning mechanisms yield additional benefits.

5.3 Feature Importance Analysis

Feature importance is a metric that assesses the relative significance, or power, of each feature in making predictions. As such, the Random Forest model's feature importance analysis provided critical insights into which environmental KPIs contributed the most to the predictive accuracy of the model. Table 5.2 contains the results of the feature importance analysis, including the six environmental KPIs, as well as the two interaction terms created during feature engineering.

Feature	Importance (%)
eKPI_GHGs	58.07
eKPI_Water_Consumption	17.62
eKPI_Air_Emissions	14.10
GHGs_Air_Emissions_Interaction	5.03
eKPI_Water_Pollution	4.17
eKPI_Waste	0.80
eKPI_Land_Use	0.12
Land_Use_Water_Consumption_Interaction	0.10

Table 5.2: Feature Importance Analysis (Adapted from Spyder/Python)

As illustrated in Table 5.2, greenhouse gasses (GHGs) yielded an importance of 58.07%, significantly higher than that of the other EKPIs, contributing to over half of the model's predictive accuracy. Their substantial weight in the model indicates their critical role in determining materials' environmental impacts. This finding aligns with the increasing awareness that GHG emissions play a major role in the environmental crisis; hence, reducing them is vital to working toward sustainable development goals, such as SDG 13, 'Climate Action'.

Water consumption (17.62%) and air emissions (14.10%) were the second and third most important features, respectively. The extensive use of water within the fashion industry, particularly in textile production processes, justifies its role as the second-highest predictor. Similarly, the widespread contribution of air emissions from transportation and manufacturing processes explains its significance within the model.

Interestingly, although the GHGs and air emissions interaction term only accounted for 5.03% of the model's predictive accuracy, it highlighted the importance of examining these complex relationships between variables. The compounded effect of these variables illustrates that total environmental harm is multidimensional, and can be caused by a cumulative effect of lesser environmental harms.

The remaining features – water pollution (4.17%), waste (0.80%), land use (0.12%) and the land use and water consumption interaction term (0.10%) – did not significantly contribute to the predictive accuracy of this model. While their low importance might suggest limited influence in this specific dataset, their relevance in broader sustainability contexts should not be discounted. For example, land use often has a delayed impact that may not be captured effectively within shorter production cycles. Similarly, waste primarily affects the end-of-life stage, which was excluded from the scope of this study.

These findings also highlight potential limitations of the Random Forest model in capturing non-linear or subtle relationships. Future studies might benefit from exploring alternative models such as Gradient Boosting or Neural Networks, which could better account for the interplay between underperforming features like waste and land use.

5.4 Comparison of Model Performance with Existing Literature

The results of this Random Forest model are not directly comparable to existing studies in terms of methodology, due to its novel use of predictive analytics in material optimization. Most studies, including Kering's EP&L report, focus on qualitative environmental assessments without integrating machine learning or AI-driven algorithms. However, regardless of the differences in methodology, parallels can be drawn between this research and other studies' assessments of the environmental impacts of materials within fashion supply chains.

Kering's EP&L framework laid the foundation for this study by identifying six key performance indicators, including GHG emissions and water consumption, as contributors to a material's total environmental impact throughout the product's entire lifecycle. Using Random Forest, this study's model ranked the importances of each EKPI, confirming that GHG emissions (58.07%) and water consumption (17.62%) are, in fact, the two most important predictors of a material's overall environmental impact. This finding aligns with and builds upon Kering's research by providing an added predictive capability to enable real-time, dynamic forecasts for the environmental impacts of materials.

In Bicho *et al*.'s (2023) research on sustainable fashion product development, LCA techniques were used to retrospectively assess the environmental impacts of materials and processes throughout product life cycles. Their findings suggested that end-of-life practices such as waste management and recycling heavily influenced overall environmental impact. In contrast, this study found that waste only accounted for 0.80% of the model's predictive accuracy. This discrepancy highlights the difference in focus: while Bicho *et al*. emphasized downstream impacts, this study underscores the critical need for early-stage interventions in material sourcing and production to reduce environmental footprints.

Due to the absence of comparable predictive models in existing literature, the MAE result of 3,296,096.55 lacked direct quantitative reference points. Despite this gap, the cross-validated MAE of 0.723 supports the model's generalizability and indicates its potential for broader application in sustainability-focused decision-making.

While this study introduces a novel approach powered by predictive analytics, it is important to acknowledge the limitations of the Random Forest model, particularly for features with low predictive power such as land use and waste. These variables likely require a more nuanced approach to fully capture their impacts, which could be achieved by incorporating additional datasets or alternative machine learning techniques. Addressing these gaps could provide further insight into the broader applicability of predictive modeling for sustainable product development.

5.4 Comparison of Model Performance with Existing Literature

This section will assess the validity of the two hypotheses formulated in Chapter 2 (see Table 2.2), using the quantitative findings from Chapter 5 to provide evidence for acceptance or rejection.

Hypothesis 1 (H1): AI-driven material optimization, using Random Forest models and environmental key performance indicators such as GHGs, water consumption, and waste, will accurately predict the environmental impacts of different materials used in fashion production, as demonstrated through model performance metrics (e.g., MAE).

The results of the model, including an MAE reduction from 30.17 million to 3.29 million, confirm its predictive accuracy in evaluating environmental impacts of materials. The cross-validated MAE of 0.723 further supports the generalizability of the model across datasets. These findings validate H1, demonstrating that the Random Forest model can serve as an effective tool for guiding sustainable material decisions in fashion product development.

Hypothesis 2 (H2): Among the six environmental dimensions (GHGs, water consumption, land use, air emissions, water pollution, and waste), GHGs will have a significantly higher predictive importance in determining the total environmental impact of materials, as identified through feature importance analysis in the Random Forest model.

Feature importance analysis confirmed that GHGs accounted for 58.07% of the model's predictive power, significantly outweighing other EKPIs. This validates H2 and highlights GHGs as the most critical factor in material impact assessment, aligning with Kering's EP&L methodology and broader literature emphasizing the role of GHGs in sustainability initiatives.

Chapter Six
DISCUSSION & CONCLUSIONS

6. DISCUSSION & CONCLUSIONS

6.1 Discussion

6.1.1 Addressing Objective 1: Highlighting Gaps in AI-Driven Material Optimization

The comprehensive literature review revealed a notable lack of studies exploring the intersection of AI and sustainable fashion product development. While existing research on Circular Economy (CE) theory and Life Cycle Assessment (LCA) principles provided a theoretical foundation, the application of AI-driven predictive analytics in material optimization remained underexplored.

For instance, Bicho *et al*. (2023) emphasized the importance of integrating digital technologies into CE frameworks but noted that current applications remain focused on logistics and supply chain optimization, rather than product development itself. Similarly, Dissanayake and Weerasinghe (2021) highlighted the need for more actionable, data-driven approaches in sustainable fashion but did not address the role of machine learning in material selection. This study directly contributes to addressing these gaps by leveraging Random Forest models to prioritize material choices based on environmental KPIs, bridging the divide between theory and practice. The importance of addressing these gaps extends beyond academic research, as the fashion industry increasingly faces pressure to align with Sustainable Development Goals (SDGs), particularly SDG 12 (Responsible Consumption and Production). By integrating AI-driven analytics into the material selection process, this research not only fills a critical gap in the literature but also provides a practical framework for brands seeking to meet these global sustainability targets.

6.1.2 Addressing Objective 2: Analyzing Environmental KPIs

The feature importance analysis demonstrated that greenhouse gasses (GHGs), water consumption, and air emissions collectively accounted for 89.79% of the model's predictive power. This result underscores the critical importance of prioritizing these KPIs in material selection processes to reduce the environmental footprint of garments. These findings align with Kozlowski, Bardecki, and Searcy (2012), who emphasized the significant role of GHG emissions in determining the overall sustainability performance of the fashion industry. However, while prior studies primarily highlighted the theoretical relevance of GHGs and water consumption, this research provides quantitative evidence of their predictive importance in environmental impact modeling. By applying a Random Forest model, this study bridges the gap between theoretical frameworks and actionable insights, offering a more nuanced understanding of environmental KPIs.

The prominence of GHGs as the top predictor reflects Kering's emphasis on tracking emissions across its supply chain, from raw material production to garment disposal. Similarly, water consumption and air emissions, which ranked second and third in importance, respectively, highlight critical areas for intervention in fashion product development. This finding reinforces the notion that addressing these three EKPIs can yield the most significant reductions in environmental impact, a priority echoed in Kering's EP&L methodology and broader industry sustainability frameworks. By providing these insights, the analysis emphasizes the importance of integrating data-driven decision-making into the fashion industry's sustainability strategies. The findings suggest that prioritizing materials with lower GHG emissions and water use in the design stage can significantly reduce the environmental footprint of garments, offering a pathway to achieving the United Nations' sustainability goals.

6.1.3 Addressing Objective 3: Evaluating AI-Driven Optimization

The Random Forest model demonstrated strong predictive capabilities, as evidenced by a significant reduction in mean absolute error (MAE) from 30.17 million to 3.29 million after applying logarithmic transformation. This improvement highlights the model's effectiveness in predicting the environmental impacts of materials based on their environmental KPIs. These results align with broader research on the application of machine learning in sustainability sciences, such as Dissanayake and Weerasinghe (2021), who noted the potential for AI techniques to streamline environmental impact assessments. However, while their work emphasized theoretical possibilities, this study provides practical evidence of AI's effectiveness in optimizing material choices in fashion product development. Moreover, the cross-validated MAE of 0.723 confirmed the model's generalizability, indicating its potential for application across diverse datasets without overfitting. This finding strengthens the case for adopting AI-driven models in fashion, as they offer both accuracy and adaptability, two critical components for scaling sustainable practices industry-wide.

One notable advantage of the Random Forest model is its ability to handle complex interactions between variables, such as the interaction between GHG emissions and water consumption. Prior studies, such as Kozlowski, Bardecki, and Searcy (2012), acknowledged the limitations of traditional LCA methods in capturing such complexities. This research addresses that limitation by incorporating machine learning to uncover these relationships, enabling more precise environmental impact predictions. While the model's performance is promising, it is essential to acknowledge the limitations of relying solely on historical data. For example, changes in environmental policies, material innovations, or consumer behaviors may not be fully captured in the current dataset. Future iterations of this model should consider integrating dynamic data sources to improve accuracy and relevance in real-world applications.

In conclusion, the Random Forest model represents a significant step forward in leveraging AI for sustainable fashion product development. Its ability to predict environmental impacts with high accuracy offers a practical tool for fashion brands seeking to optimize material choices, reduce environmental footprints, and align with global sustainability goals such as SDG 12 and SDG 13.

6.1.4 Addressing Objective 4: Practical Recommendations

The findings of this study provide actionable strategies for fashion brands to integrate AI-driven material optimization into their product development processes. The feature importance analysis revealed that GHGs, water consumption, and air emissions are the most influential predictors of environmental impact. This insight empowers brands to prioritize materials and production processes that minimize these impacts, enabling them to reduce their carbon footprint while enhancing resource efficiency. For example, brands could adopt low-carbon materials, explore alternative production techniques with reduced GHG emissions, and implement mono-material designs to ensure end-of-life recyclability. These strategies align with recommendations set forth by Bicho *et al.* (2023), who advocated for adopting design innovations to support Circular Economy principles. However, this research extends their work by providing data-driven guidance for material selection based on predictive analytics, offering brands a practical framework for sustainable decision-making.

To successfully integrate AI-driven material optimization, brands must invest in robust data infrastructures. This includes collecting comprehensive environmental data across their supply chains, and training predictive models tailored to their specific operations. The development of a brand-specific algorithm is a key recommendation, enabling companies to simulate the environmental impacts of their materials and processes in real-time.. Such tools could support brands in setting sustainability targets and evaluating their progress, as well as providing tangible insights into the trade-offs between different material choices. This concept builds upon Kering's EP&L methodology, which serves as a model for comprehensive data collection and analysis. As noted by Dissanayake and Weerasinghe (2021), brands often face challenges in accessing accurate data, but investments in AI and predictive analytics can mitigate these barriers. For example, AI-driven tools could integrate external factors such as evolving environmental policies or market trends to enhance their adaptability and relevance.

Despite initial investment costs, AI-driven material optimization offers significant long-term benefits. Not only can it reduce waste and production inefficiencies, but it also allows brands to target eco-conscious consumers, fostering loyalty and increasing market share. Additionally, compliance with regulatory frameworks promoting sustainability – such as SDG 12 and SDG 13 – can minimize risks and costs associated with non-compliance, further underscoring the value of adopting such innovations.

In conclusion, the integration of AI-driven material optimization represents a transformative opportunity for fashion brands. By aligning product development processes with sustainability goals through predictive analytics, brands can reduce their environmental footprints, improve operational efficiency, and position themselves as leaders in sustainable innovation.

6.2 Conclusions

6.2.1 Summary of Key Contributions and Fulfillment of Research Objectives

This study successfully achieved its aim of exploring the potential of AI-driven material optimization in enabling sustainable fashion product development. By utilizing Random Forest models and environmental KPIs, the research developed a predictive framework for assessing the environmental impacts of various materials in Kering's EP&L datasets. Each research objective was addressed as follows:

Objective 1:

This study identified a critical gap in the literature regarding the application of AI in sustainable fashion product development. Unlike previous research that largely focused on theoretical discussions, such as Bicho *et al*. (2023), this study provides practical evidence of AI's potential by applying predictive analytics to material optimization. This contribution bridges the gap between sustainability theory and real-world application.

Objective 2:

The feature importance analysis confirmed that GHGs, water consumption, and air emissions are the most significant predictors of environmental impact, accounting for 89.79% of the model's predictive power. This aligns with Kering's EP&L methodology and validates its focus on these KPIs as high-impact areas for intervention. By quantifying the predictive importance of these factors, this research offers actionable insights for prioritizing sustainability efforts.

Objective 3:

The Random Forest model demonstrated robust predictive capabilities, with an MAE reduced to 3.29 million and a cross-validated MAE of 0.723. These results underscore the feasibility of integrating AI into sustainability-focused decision-making, offering a scalable and adaptable tool for fashion brands seeking to optimize material choices.

Objective 4:

Practical recommendations were developed for integrating AI-driven material optimization into fashion brands' operations. Key suggestions included creating a brand-specific algorithm, inspired by Kering's EP&L methodology, and prioritizing low-GHG materials. These recommendations provide a pathway for brands to align product development processes with global sustainability goals, such as SDG 12 and SDG 13.

In conclusion, this research makes a significant contribution to the emerging field of AI in sustainable fashion. By combining predictive analytics with comprehensive environmental data, it offers both theoretical advancements and practical solutions, setting a precedent for future studies to further explore AI-driven sustainability.

6.2.2 Research Limitations and Areas for Future Research

Despite the significant contributions of this study, several limitations must be acknowledged to provide a balanced perspective on its findings. These limitations also present opportunities for future research to enhance the applicability and impact of AI-driven material optimization in sustainable fashion.

One key limitation was the reliance on Kering's EP&L datasets, which, while comprehensive, are specific to a luxury fashion context. As luxury brands often operate with distinct supply chain dynamics compared to fast fashion or mid-market brands, the findings may not be directly generalizable across all segments of the fashion industry. Future studies should replicate this methodology using datasets from other market segments, such as fast fashion, to identify potential differences in environmental impact predictors and material optimization strategies.

Another limitation was the novelty of the algorithm, which made it challenging to benchmark the model's MAE against similar studies. While the MAE of 3.29 million and cross-validation score of 0.723 suggest strong predictive performance, the absence of comparable models limits the ability to contextualize these results. Future research could explore alternative machine learning techniques, such as Gradient Boosting or Neural Networks, to compare performance and validate the effectiveness of Random Forest for material optimization.

The model's reliance on historical data is another limitation. As environmental policies, technological innovations, and consumer behaviors evolve, historical datasets may not fully capture the dynamic nature of sustainability challenges. Incorporating real-time or longitudinal data could enhance the model's adaptability, allowing for predictions that reflect changing conditions within the fashion industry.

The low feature importance of certain variables, such as land use (0.12%) and waste (0.80%), also warrants further investigation. While these features had minimal influence in this dataset, their broader significance in sustainability contexts, particularly in end-of-life stages or long-term environmental impacts, should not be overlooked. Future research could integrate additional data sources or frameworks, such as Circular Economy (CE) principles, to provide a more holistic view of these variables' roles.

Additionally, the study's focus on early-stage material optimization excluded downstream processes, such as garment manufacturing or end-of-life management. Expanding the scope to include these stages could offer a more comprehensive understanding of how AI-driven models can influence the entire product lifecycle, aligning with both Circular Economy goals and SDG 12 (Responsible Consumption and Production).

In summary, while this research highlights the potential of AI in sustainable fashion, its limitations emphasize the need for further exploration. By addressing these challenges through expanded datasets, advanced algorithms, and broader lifecycle analyses, future studies can build on this foundation to create even more impactful solutions for achieving sustainability in the fashion industry.

REFERENCES

Armstrong, C.M. and LeHew, M.L.A. (2011) 'Sustainable Apparel Product Development: In Search of a New Dominant Social Paradigm for the Field Using Sustainable Approaches', *Fashion Practice*, 3(1), pp. 29–62. Available at: https://doi.org/10.2752/175693811x12925927157018 (Accessed: 2 Aug 2024).

Bicho, M. *et al*. (2023) 'Sustainable Fashion Product Development: The Importance of Skills in New Materials and Processes in the Academy and Industry Context', *Advances in Fashion and Design Research II*, pp. 287-301. Available at: https://doi.org/10.1007/978-3-031-43937-7_26 (Accessed: 1 Aug 2024).

Buus Lassen, N., la Cour, L. and Vatrapu, R. (2016) 'Predictive Analytics with Social Media Data', *SAGE Publications Ltd eBooks*, pp. 328–341. Available at: https://doi.org/10.4135/9781473983847.n20 (Accessed: 20 Aug 2024).

Cai, Y.-J. and Choi, T.-M. (2020). 'A United Nations' Sustainable Development Goals Perspective for Sustainable Textile and Apparel Supply Chain Management', *Transportation Research Part E: Logistics and Transportation Review,* [online] 141, p.102010. Available at: https://doi.org/10.1016/j.tre.2020.102010 (Accessed: 2 Aug 2024).

Curwen, L.G., Park, J. and Sarkar, A.K. (2013) 'Challenges and Solutions of Sustainable Apparel Product Development', *Clothing and Textiles Research Journal*, 31(1), pp. 32-47. Available at: https://doi.org/10.1177/0887302x12472724 (Accessed: 1 Aug 2024).

Dissanayake, D.G.K. and Weerasinghe, D. (2021) 'Towards Circular Economy in Fashion: Review of Strategies, Barriers and Enablers', *Circular Economy and Sustainability*, 2(2), pp. 25–45. Available at: https://doi.org/10.1007/s43615-021-00090-5 (Accessed: 10 Aug 2024).

Dottle, R. and Gu, J. (2022). *The Global Glut of Clothing Is an Environmental Crisis, Bloomberg.com*. Available at: https://www.bloomberg.com/graphics/2022-fashion-industry-environmental-impact/?embedded-checkout=true (Accessed: 8 Aug 2024).

Ellen MacArthur Foundation (2021). *Toxin-Free, Recyclable Clothing: Napapijri Circular Series, Ellen MacArthur Foundation.* Available at: https://www.ellenmacarthurfoundation.org/circular-examples/napapijri-circular-series (Accessed 8 Aug 2024).

European Commission (2015). *Closing the Loop - An EU Action Plan for the Circular Economy*. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52015DC0614 (Accessed: 10 Aug 2024).

Feuerriegel, S. *et al.* (2023) 'Generative AI', *Business & Information Systems Engineering*, 66(1), pp. 111–126. Available at: https://doi.org/10.1007/s12599-023-00834-7 (Accessed: 22 Aug 2024).

Granskog, A. *et al.* (2020). *Biodiversity: The Next Frontier in Sustainable Fashion*. Available at: https://www.mckinsey.com/industries/retail/our-insights/biodiversity-the-next-frontier-in-sustainable-fashion (Accessed: 24 Aug 2024)

Harmsen, P., Scheffer, M. and Bos, H. (2021). 'Textiles for Circular Fashion: The Logic behind Recycling Options. Sustainability, 13(17), p.9714. Available at: https://doi.org/10.3390/su13179714 (Accessed: 2 Aug 2024).

Hvass, K.K. (2014). Exploring Business Model Innovation for Closed Loop Fashion. In: M. Charter, ed. *Proceeding of Sustainable Innovation 2013: Collaboration, Co-creation and New Business Models: Towards Sustainable Product Design*. 18th International Conference. University for the Creative Arts, pp.111-117. Available at: https://hdl.handle.net/10398/9028 (Accessed: 10 Aug 2024).

The Insight Partners (2019). *Market Value of Artificial Intelligence in Fashion Worldwide from 2018 to 2027 (in Million U.S. Dollars)*. Statista. Available at: https://www-statista-com.arts.idm.oclc.org/statistics/1070736/global-artificial-intelligence-fashion-market-size/ (Accessed: 1 Sep 2024).

International Organization for Standardization [ISO] (2006). *Environmental Management – Life Cycle Assessment – Requirements and Guidelines (ISO 14044:2006)*. Geneva: ISO. Available at: https://www.iso.org/obp/ui/en/#iso:std:iso:14044:ed-1:v1:en (Accessed: 12 Aug 2024).

Jacometti, V. (2019) 'Circular Economy and Waste in the Fashion Industry', *Laws*, 8(4), p. 27. Available at: https://www.mdpi.com/2075-471X/8/4/27 (Accessed: 10 Aug 2024).

Kering (2023a). *EP&L Index 2023: Material Intensities*. Available at: https://kering-group.opendatasoft.com/pages/material-intensities/ (Accessed: 28 Aug 2024)

Kering (2023b). *EP&L Valued Result and EKPI 2023*. Available at: https://kering-group.opendatasoft.com/explore/dataset/epl-ekpi-results/table/ (Accessed: 28 Aug 2024)

Kering (2023c). *Methodological Note Environmental Reporting*. Available at: https://www.kering.com/api/download-file/?path=2023_Kering_Methodological_note_on_environmental_reporting_d30fbbf4e4.pdf (Accessed: 28 Aug 2024).

Kering (2023d). *Methodology*. Available at: https://kering-group.opendatasoft.com/pages/methodology/ (Accessed: 28 Aug 2024).

Kering (2023e). *Our 2023 EP&L Results: Report*. Available at: https://kering-group.opendatasoft.com/pages/report/ (Accessed: 28 Aug 2024)

Kering (2023f). *Raw Material Intensity 2023*. Available at: https://kering-group.opendatasoft.com/explore/dataset/raw-material-intensity/table/ (Accessed: 28 Aug 2024)

Klöpffer, W. (2003) 'Life-Cycle Based Methods For Sustainable Product Development', *The International Journal of Life Cycle Assessment*, 8(3). Available at: https://doi.org/10.1007/bf02978462 (Accessed: 1 Aug 2024).

Kozlowski, A., Bardecki, M. and Searcy, C. (2012) 'Environmental Impacts in the Fashion Industry: A Life-Cycle and Stakeholder Framework', *The Journal of Corporate Citizenship*, 45(45), pp. 17–36. Available at: https://www.jstor.org/stable/jcorpciti.45.17 (Accessed: 12 Aug 2024).

Liu, Y. *et al*. (2023) 'Generative Artificial Intelligence and its Applications in Materials Science: Current Situation and Future Perspectives', *Journal of Materiomics*, 9(4). Available at: https://doi.org/10.1016/j.jmat.2023.05.001 (Accessed: 3 Sep 2024)

Lozano, R. (2022). *Toward Sustainable Organisations*. Springer, Cham. Available at: https://doi.org/10.1007/978-3-030-99676-5_2 (Accessed: 1 Aug 2024).

Luce, L. (2018). *Artificial Intelligence for Fashion: How AI is Revolutionizing the Fashion Industry*. [online] New York: Apress Media. Available at: https://ebookcentral.proquest.com/lib/ual/detail.action?docID=5611900 (Accessed: 3 Aug 2024).

Merayo, D., Rodríguez-Prieto, A. and Camacho, A.M. (2019) 'Comparative Analysis of Artificial Intelligence Techniques for Material Selection Applied to Manufacturing in Industry 4.0', *Procedia Manufacturing*, 41, pp. 42–49. Available at: https://doi.org/10.1016/j.promfg.2019.07.027 (Accessed: 21 Aug 2024).

Mohiuddin Babu, M. *et al*. (2022) 'The Role of Artificial Intelligence in Shaping the Future of Agile Fashion Industry', *Production Planning & Control*, pp. 1–15. Available at: https://doi.org/10.1080/09537287.2022.2060858 (Accessed: 20 Aug 2024).

Moorhouse, D. and Moorhouse, D. (2017) 'Sustainable Design: Circular Economy in Fashion and Textiles', *The Design Journal*, 20(sup1), pp. S1948–S1959. Available at: https://doi.org/10.1080/14606925.2017.1352713 (Accessed: 10 Aug 2024).

Rathinamoorthy, R. (2020) 'Sustainable Clothing Designs for Fashion: Design Strategies and Its Implementation Possibilities', *Sustainable Textiles: Production, Processing, Manufacturing & Chemistry*, pp. 127–161. Available at: https://doi.org/10.1007/978-3-030-37929-2_7 (Accessed: 22 Aug 2024).

Rathore, B. (2019a) 'Fashion Sustainability in the AI Era: Opportunities and Challenges in Marketing', *Eduzone: International Peer Reviewed/Refereed Multidisciplinary Journal*, 8(2), pp. 17–24. Available at: https://www.eduzonejournal.com/index.php/eiprmj/article/view/362 (Accessed: 23 Aug 2024).

Rathore, B. (2019b) 'From Trendy to Green: Exploring AI's Role in Sustainable Fashion Marketing', *International Journal of New Media Studies*, 06(02), pp. 12–22. Available at: https://doi.org/10.58972/eiprmj.v6i2y19.120 (Accessed: 13 Aug 2024).

Rigatti, S.J. (2017) 'Random Forest', Journal of Insurance Medicine, 47(1), pp. 31–39. Available at: https://doi.org/10.17849/insm-47-01-31-39.1 (Accessed: 21 Aug 2024).

Saponaro, M. *et al*. (2018) 'Challenges and Opportunities of Artificial Intelligence in the Fashion World', *2018 International Conference on Intelligent and Innovative Computing Applications (ICONIC)*. Available at: https://doi.org/10.1109/iconic.2018.8601258 (Accessed: 16 Aug 2024).

Saunders, M., Lewis, P. and Thornhill, A. (2019) *Research Methods for Business Students*. 8th edn. Harlow: Pearson. Available at: https://ebookcentral.proquest.com/lib/ual/detail.action?docID=5774742 (Accessed: 30 Aug 2024)

Simonen, K. (2014) *Life Cycle Assessment, Routledge eBooks*. Informa. Available at: https://doi.org/10.4324/9781315778730 (Accessed: 12 Aug 2024).

Singh, C., Park, H. and Joyner Martinez, C.M. (2022) 'Love Letters to Patagonia: Fostering Sustainable Consumption via Consumer‐Brand Relationships', *International Journal of Sustainable Fashion & Textiles*, 1(1), pp. 41–62. Available at: https://doi.org/10.1386/sft/0003_1 (Accessed: 10 Aug 2024).

Speiser, J.L. *et al*. (2019) 'A Comparison of Random Forest Variable Selection Methods for Classification Prediction Modeling', *Expert Systems with Applications*, 134(1), pp. 93–101. Available at: https://doi.org/10.1016/j.eswa.2019.05.028 (Accessed: 21 Aug 2024)

Stitch Fix (2020). *Stitch Fix Algorithms Tour, Stitchfix.com*. Available at: https://algorithms-tour.stitchfix.com/ (Accessed: 21 Aug 2024).

Thakker, A.M. and Sun, D. (2023) 'Sustainable Development Goals for Textiles and Fashion', *Environmental Science and Pollution Research*, 30(46), pp. 101989-102009. Available at: https://doi.org/10.1007/s11356-023-29453-1 (Accessed: 1 Aug 2024).

United Nations [UN] (1992). *United Nations Conference on Environment and Development, Rio de Janeiro, Brazil, 3-14 June 1992*. Available at: https://www.un.org/en/conferences/environment/rio1992 (Accessed: 1 Aug 2024).

United Nations [UN] (2002). *World Summit on Sustainable Development 26 August-4 September 2002, Johannesburg*. Available at: https://www.un.org/en/conferences/environment/johannesburg2002 (Accessed: 1 Aug 2024).

United Nations [UN] (2015). *Sustainable Development Goals*. Available at: https://www.undp.org/sustainable-development-goals (Accessed: 1 Aug 2024).

University of the Arts London [UAL] (2020). *University of the Arts London Code of Practice on Research Ethics*. Available at: https://www.arts.ac.uk/__data/assets/pdf_file/0020/43328/UAL-Code-of-Practice-on-Research-Ethics-October-2020.pdf (Accessed: 20 Aug 2024).

World Commission on Environment and Development [WCED] (1987). *Report of the World Commission on Environment and Development: Our Common Future*. Available at: https://sustainabledevelopment.un.org/content/documents/5987our-common-future.pdf (Accessed: 1 Aug 2024).

APPENDICES

Appendix One: Individual Learning Agreement

<table>
<tr><td>Student Name</td><td colspan="4">Paige Leeds (23044656)</td></tr>
<tr><td>Title of Project</td><td colspan="4">AI-Driven Material Optimization: A Framework for Sustainable Fashion Product Development</td></tr>
<tr><td>Name/s of collaborators</td><td colspan="4">N/A</td></tr>
<tr><td>Supervisor</td><td colspan="4">Arnab Banerjee</td></tr>
<tr><td>Aim of the project:</td><td colspan="4">To explore the potential of AI-driven material optimization in enabling sustainable fashion product development, with a focus on reducing the environmental footprint of garments.</td></tr>
<tr><td>Objectives (Maximum 5)</td><td colspan="4">1. To perform a comprehensive literature review to critically examine the role of AI-driven material optimization in the fashion industry, with a focus on its potential to reduce the environmental footprint of raw materials.
2. To evaluate the environmental impact of different materials used in fashion production across six environmental dimensions from Kering's EP&L report: greenhouse gas emissions (GHGs), water consumption, land use, air emissions, water pollution, and waste.
3. To develop an AI-driven model that predicts the environmental impact of different materials, guided by insights from Kering's EP&L report, focused on uncovering patterns for optimizing material selection in fashion product development.
4. To provide recommendations for fashion brands on how to integrate AI-driven material optimization into their product development strategies, with the intention of reducing their environmental footprints.</td></tr>
<tr><td>Agreed outcomes</td><td colspan="4">• A comprehensive written thesis of 10,000 words, following the requirements and structure laid out in the assessment brief
• Refinement of a predictive algorithm for AI-driven material optimization in sustainable fashion product development</td></tr>
<tr><td>Formats</td><td colspan="4">Theory-Based Dissertation (10,000 words)</td></tr>
<tr><td>Agreed contact points/ Tutorial dates</td><td>28/06/24</td><td>07/08/24</td><td>23/08/24</td><td>16/09/24</td></tr>
<tr><td>Project Timeline</td><td colspan="4">Week 34: Introduction
Week 35-36: Literature Review
Week 37-38: Coding
Week 39: Research Design
Week 40: Dataset
Week 41: Findings and Analysis
Week 42: Discussion and Conclusion
Week 43: Abstract, Figures/Tables, Formatting
Week 44-45: Edits</td></tr>
<tr><td>Ethics Form Attached (as needed)</td><td>Yes</td><td>Risk Assessment Attached</td><td colspan="2">Yes</td></tr>
<tr><td colspan="2">Learning Outcomes</td><td colspan="3">How you will evidence attainment of the outcome (max 200 words per outcome)</td></tr>
</table>

1. Execute research informed self-directed project, demonstrating an advanced level of analytical (i.e., forecasting or machine learning) and evaluative skills (Process, Enquiry, Knowledge)	This research involves the analysis of AI-driven material optimization using the Kering EP&L datasets to improve sustainability in fashion product development. It builds upon existing research in sustainability, generative AI, and data-driven material optimization in the fashion industry. The gap in research is primarily due to the reliance on historical/static data rather than on AI in forecast/predict environmental impacts of material choices during product development in fashion. The proposed methodology consists of quantitative analysis of the datasets using AI models, specifically Random Forest and feature importance analysis.
2. Apply your critical understanding of contemporary relevant theory and practice in your subject area (Enquiry, Knowledge)	Evaluation will rely on quantitative analysis of Kering's EP&L data, using machine learning models (such as Random Forest). Relevant theories to analyze the results include the Circular Economy theory, Sustainability Science, and Life Cycle Assessment (LCA) frameworks. Skills in data wrangling, predictive analytics, and understanding of environmental metrics will be required to complete the project. Collaborations will not be necessary for this study, as I am relying primarily on existing literature, secondary data, and insights from my supervisor. The research will be of Masters of Science Level by integrating academic rigor, quantitative research methodologies, and advanced machine learning applications.
3. Demonstrate effective communication of your project, context, and research design (Communication)	The research will be analyzed through a combination of statistical methods and machine learning models, most notably Random Forest. In relation to other studies, it aims to extend the body of knowledge on the intersection of AI and sustainability, particularly in regard to material optimization in product development. Theory will be related to practice by interpreting the model results through sustainability frameworks, identifying environmental impacts. The strengths and weaknesses of the preditive model will be evaluated through Mean Absolute Error (MAE) and ranking feature importance.
4. Evaluate the execution of your Masters' project against relevant academic and professional standards (Process, Realisation)	The position of this study is at the intersection of AI and sustainability, particularly regarding fashion innovation. This work is relevant to fashion companies as it aims to provide a solution for reducing their environmental impact via AI-driven material optimization. Fashion industry professionals, sustainability advocates, and AI researchers interested in innovative approaches to sustainable product development will be interested in the findings of this study.

Signed Tutor……………………………………………………… Date 13/09/24……………………………………

Signed Student……………………………………………………… Date………………………………12/09/24

Students should submit any written work for review by their supervisor **three working days** prior to a tutorial, in order to give the supervisor sufficient time to consider your work and respond to any issues. Students should **normally** expect a response to an email by your supervisor within five working days. It is your responsibility to find out when your supervisor will be on leave or out of the country.

Appendix Two: Research Ethics Form

NAME: Paige Leeds (23044656)
COLLEGE: University of the Arts London (UAL)
IF YOUR RESEARCH INVOLVES PARTICIPANTS, PLEASE COMPLETE QUESTIONS 1 TO 9. IF NOT, GO TO QUESTION 10 BELOW.
1. **Will the participants be: (please tick as appropriate)** Students at the University N/A Participants outside the University N/A
2. **How will participants be recruited and how many will be involved?** (Approximate number is OK) N/A
3. **What will the participants be asked to do?** (Explain clearly so that a non-specialist will understand) The participants will answer to a written interview that will be sent to them through email. N/A
4. **What potential risks to the interests of participants do you foresee and what steps will you take to minimise those risks?** (A participant's interests include their physical and psychological well- being, their commercial interests; and their rights of privacy and reputation). N/A
5. **What potential risks to yourself as research student do you foresee and what steps will you take to minimise those risks?** (e.g. does your research raise issues of personal safety for you or others involved in the project, especially if taking place outside working hours or off University premises) N/A
6. Please attach a copy of proposed written information /consent form to be given to participants for agreement and signing. If you are not obtaining written consent or supplying an information sheet, please explain the reasons for this. Attached: N/A
7. Does your project involve children (i.e. under 16) or vulnerable adults e.g. a person with a learning disability? NO If YES, you must refer to the Guidance Note on Informed Consent in the Code of Practice on Research Ethics at http://www.arts.ac.uk/research/researching-at-ual/researcher-support/ and obtain a Criminal Records Bureau (CRB) check. You are advised to discuss this with your supervisor. Please tick to confirm CRB check has been obtained: N/A
Please refer to the guidance note on data protection available at http://www.arts.ac.uk/research/researching-at-ual/researcher-support/ before answering the next question. Please consider the value of coding; the importance of

secure storage and disposal of personal information, particularly sensitive data (e.g. records of health, origin, criminal record etc.)

8. **Will you be obtaining personal data from any of the participants? (Y/N)**

If YES: Give full details based on the following questions:

(a) How will you store and use this information during the course of your research?
(b) What parts of this information will be confidential?
(c) Will you separate personal identifiers from other (coded) personal data, and if so, how will you safeguard the key to these data sets?
(d) Will personal data be irreversibly anonymised or, if you have separated the data, will the linking code between the two databases be destroyed?
(e) At the conclusion of your research:
    (i) Which of your data sets do you intend to retain personally for use in future research?
    (ii) Which do you intend to archive for other researchers?
    (iii) Which do you intend to destroy?
(f) Depending on your answers to (e):
    (i) If you intend to retain certain data sets for future use or to archive them:
        (i.i) How will they be stored?
        (i.ii) Will participants be informed what data will be retained, and will their consent be obtained for this?
    (ii) If you intend to destroy certain data sets at the conclusion of the research:
        (ii.i) Explain why this is appropriate
        (ii.ii) How will you ensure that the data will be disposed of in such a way that there is no risk of its confidentiality being compromised?

N/A

9. **Will payments to participants be made? (Y/N)** N/A
10. (If YES, please state amount and whether payment is for out-of-pocket expenses, or a fee. Normally no payment is made in student research projects) N/A

11. **Will any restrictions be placed on the publication of results (Y/N)** N/A
(If YES, please state the nature of the restrictions, e.g. details of any confidentiality agreement).

12. **I confirm my responsibility as a student of UAL to conduct research in accordance with the Code of Practice on Research Ethics of the University of the Arts London (the University). In signing this form I am also confirming that:**

    a) The form is accurate to the best of my knowledge and belief.
    b) There is no potential material interest that may, or may appear to, impair the independence and objectivity of researchers conducting this research.
    c) I undertake to conduct the research as set out in this form.
    d) I understand and accept that the ethical propriety of this research may be monitored by the relevant College Research body and/or the University's Research Ethics Sub-Committee.

Signature of
Student Researcher Name ____________________ Date: 12 Sep 2024

13. **I support this research and have reviewed it with the student:**
Signature of
Supervisor: ____________________ Date 13/09/2024

Appendix Three: Kering Raw Material Intensity Dataset

Tier	Material Group	Raw material	Process step	Impact country	Environmental impact group	Indicator	Valued_Ekpi_intensity
4	Animal Fibers	Animal Fibers - cashmere - conventional	Animal rearing	China	Land use	€ EP&L	1438.445946
4	Animal Fibers	Animal Fibers - cashmere - organic	Animal rearing	China	Land use	€ EP&L	1124.509578
4	Animal Fibers	Animal Fibers - cashmere - conventional	Animal rearing	Mongolia	Land use	€ EP&L	1078.345844
4	Animal Fibers	Animal Fibers - cashmere - restorative grazing	Animal rearing	China	Land use	€ EP&L	787.3824657
4	Animal Fibers	Animal Fibers - cashmere - restorative grazing	Animal rearing	Mongolia	Land use	€ EP&L	787.3824657
3	Metal	Metal - gold - 24kt - Kering responsible	Processing	Peru	Water pollution	€ EP&L	546.8952645
4	Animal Fibers	Animal Fibers - cashmere - conventional	Animal rearing	China	GHGs	€ EP&L	239.0179153
4	Animal Fibers	Animal Fibers - cashmere - organic	Animal rearing	China	GHGs	€ EP&L	239.0179153
4	Animal Fibers	Animal Fibers - cashmere - restorative grazing	Animal rearing	China	GHGs	€ EP&L	150.3026494
4	Animal Fibers	Animal Fibers - cashmere - restorative grazing	Animal rearing	Mongolia	GHGs	€ EP&L	150.3026494
4	Animal Fibers	Animal Fibers - cashmere - conventional	Animal rearing	Mongolia	GHGs	€ EP&L	150.3026494
4	Animal Fibers	Animal Fibers - cashmere - conventional	Animal rearing	China	Air Emissions	€ EP&L	97.40656833
4	Animal Fibers	Animal Fibers - cashmere - organic	Animal rearing	China	Air Emissions	€ EP&L	97.40656833
3	Metal	Metal - gold - 24kt - Kering responsible	Processing	Peru	Air Emissions	€ EP&L	87.78924623
4	Animal Fibers	Animal Fibers - cashmere - conventional	Animal rearing	China	Water consumption	€ EP&L	80.67840409
4	Animal Fibers	Animal Fibers - cashmere - organic	Animal rearing	China	Water consumption	€ EP&L	80.67840409
4	Animal Fibers	Animal Fibers - silk - conventional	Silk production	China	Water consumption	€ EP&L	27.98957265
3	Metal	Metal - gold - 24kt - Kering responsible	Processing	Peru	GHGs	€ EP&L	26.10885238
4	Animal Fibers	Animal Fibers - wool - conventional	Animal rearing	Argentina	Land use	€ EP&L	24.53568162
4	Animal Fibers	Animal Fibers - wool - organic	Animal rearing	Argentina	Land use	€ EP&L	21.70522757
4	Animal Fibers	Animal Fibers - wool - restorative grazing	Animal rearing	Argentina	Land use	€ EP&L	20.79358403
3	Metal	Metal - gold - 24kt - Kering responsible	Processing	Peru	Land use	€ EP&L	16.08912554
4	Animal Fibers	Animal Fibers - wool - regenerated	Animal rearing	Australia	Land use	€ EP&L	14.19598632
4	Animal Fibers	Animal Fibers - wool - regenerated	Animal rearing	Argentina	Land use	€ EP&L	13.4315358
4	Animal Fibers	Animal Fibers - wool - restorative grazing	Animal rearing	Argentina	GHGs	€ EP&L	13.20841853
4	Animal Fibers	Animal Fibers - cashmere - restorative grazing	Animal rearing	China	Water consumption	€ EP&L	13.0146897
4	Animal Fibers	Animal Fibers - cashmere - conventional	Animal rearing	Mongolia	Water consumption	€ EP&L	12.20375061
4	Animal Fibers	Animal Fibers - cashmere - restorative grazing	Animal rearing	Mongolia	Water consumption	€ EP&L	12.20375061
4	Animal Fibers	Animal Fibers - wool - conventional	Animal rearing	Argentina	GHGs	€ EP&L	10.91806925
4	Animal Fibers	Animal Fibers - wool - organic	Animal rearing	Argentina	GHGs	€ EP&L	10.91530385
4	Plant Fibers	Plant Fibers - cotton - conventional	Crop farming	Egypt	Water consumption	€ EP&L	9.046406255
3	Metal	Metal - gold - 24kt - Kering responsible	Processing	Peru	Water consumption	€ EP&L	6.391371736
4	Animal Fibers	Animal Fibers - cashmere - organic	Animal rearing	China	Waste	€ EP&L	6.197680564
4	Animal Fibers	Animal Fibers - cashmere - conventional	Animal rearing	China	Waste	€ EP&L	6.197680564
4	Animal Fibers	Animal Fibers - wool - regenerated	Animal rearing	Australia	GHGs	€ EP&L	5.870714442
4	Animal Fibers	Animal Fibers - wool - regenerated	Animal rearing	Argentina	GHGs	€ EP&L	5.72086639
3	Cellulose based Fibers	Cellulose based Fibers - acetate	Processing	China	Land use	€ EP&L	4.531747363
4	Animal Fibers	Animal Fibers - silk - conventional	Silk production	China	GHGs	€ EP&L	4.169227248
4	Leather	Leather - beef - calf - conventional	Animal rearing	United Kingdom	Land use	€ EP&L	3.088937121

Appendix Four: Kering Valued EP&L Result and EKPI Dataset

Business Unit	Raw material Group	Raw material	Process step	Tier	Impact country	eKPI	eKPI Year	eKPI result	Year	Valued result	Area	Population
Other	Metal	Metal - zinc - conventional	Extraction	4	Thailand	Land use	2023	0.0000229	2023	0.049068318	514000	69428524
Other	Metal	Metal - zinc - conventional	Extraction	4	Thailand	Waste	2023	0.0000657	2023	0.00000548	514000	69428524
Other	Metal	Metal - zinc - conventional	Extraction	4	Thailand	Water consumption	2023	0.133901584	2023	0.051007831	514000	69428524
Other	Metal	Metal - zinc - conventional	Processing	3	Thailand	Land use	2023	0.000015	2023	0.032171195	514000	69428524
Other	Metal	Metal - zinc - conventional	Processing	3	Thailand	Waste	2023	0.00114464	2023	0.0000953	514000	69428524
Other	Metal	Metal - zinc - conventional	Processing	3	Thailand	Water pollution	2023	0.001179047	2023	0.068518632	514000	69428524
Other	Other	Other - acetal resin	Extraction	4	Belgium	GHGs	2023	0.123191923	2023	0.010643782	30510	11422068
Other	Other	Other - acetal resin	Extraction	4	Belgium	Land use	2023	3.07E-09	2023	0.00000245	30510	11422068
Other	Other	Other - acetal resin	Extraction	4	Belgium	Waste	2023	0.001073153	2023	0.0000332	30510	11422068
Other	Other	Other - acetal resin	Extraction	4	China	Air emissions	2023	3.739993393	2023	6.206075688	9596960	1392730000
Other	Other	Other - acetal resin	Extraction	4	China	GHGs	2023	1750.06	2023	151.2049778	9596960	1392730000
Other	Other	Other - acetal resin	Extraction	4	China	Water consumption	2023	6.158531992	2023	4.699329421	9596960	1392730000
Other	Other	Other - acetal resin	Extraction	4	China	Water pollution	2023	0.014276326	2023	1.62781536	9596960	1392730000
Other	Other	Other - acetal resin	Extraction	4	France	Air emissions	2023	0.000146929	2023	0.000176577	547030	66987244
Other	Other	Other - acetal resin	Extraction	4	France	GHGs	2023	0.131713103	2023	0.011380012	547030	66987244
Other	Other	Other - acetal resin	Extraction	4	France	Waste	2023	0.001147383	2023	0.0000521	547030	66987244
Other	Other	Other - acetal resin	Extraction	4	France	Water pollution	2023	0.00000152	2023	0.000123138	547030	66987244
Other	Other	Other - acetal resin	Extraction	4	Germany	Land use	2023	6.96E-09	2023	0.00000608	357021	82927922
Other	Other	Other - acetal resin	Extraction	4	Germany	Water consumption	2023	0.000847126	2023	0.000119222	357021	82927922
Other	Other	Other - acetal resin	Extraction	4	Italy	Waste	2023	0.016555728	2023	0.000764705	301230	60431283
Other	Other	Other - acetal resin	Extraction	4	Italy	Water pollution	2023	0.000022	2023	0.001993651	301230	60431283
Other	Other	Other - acetal resin	Processing	3	Belgium	Air emissions	2023	0.00587596	2023	0.004652114	30510	11422068
Other	Other	Other - acetal resin	Processing	3	Belgium	Water pollution	2023	0.000019	2023	0.006085749	30510	11422068
Other	Other	Other - acetal resin	Processing	3	China	Air emissions	2023	74.46289296	2023	115.828461	9596960	1392730000
Other	Other	Other - acetal resin	Processing	3	China	Waste	2023	1086.19	2023	99.05830015	9596960	1392730000
Other	Other	Other - acetal resin	Processing	3	China	Water consumption	2023	84.72596862	2023	64.65099762	9596960	1392730000
Other	Other	Other - acetal resin	Processing	3	France	Land use	2023	6.12E-08	2023	0.0000477	547030	66987244
Other	Other	Other - acetal resin	Processing	3	France	Waste	2023	0.782270445	2023	0.032719486	547030	66987244
Other	Other	Other - acetal resin	Processing	3	Germany	GHGs	2023	5.880068087	2023	0.508037883	357021	82927922
Other	Other	Other - acetal resin	Processing	3	Germany	Water consumption	2023	0.019157849	2023	0.002696223	357021	82927922
Other	Other	Other - acetal resin	Processing	3	Germany	Water pollution	2023	0.0000431	2023	0.007344976	357021	82927922
Other	Other	Other - acetal resin	Processing	3	Italy	GHGs	2023	35.15861721	2023	3.037704527	301230	60431283
Other	Other	Other - ceramic	Extraction	4	China	Water pollution	2023	0	2023	0	9596960	1392730000
Other	Other	Other - ceramic	Extraction	4	France	Air emissions	2023	0.012811115	2023	0.103894036	547030	66987244
Other	Other	Other - ceramic	Extraction	4	France	Water consumption	2023	0.000505476	2023	0.0000634	547030	66987244
Other	Other	Other - ceramic	Processing	3	China	Waste	2023	0	2023	0	9596960	1392730000
Other	Other	Other - ceramic	Processing	3	France	Air emissions	2023	0.902066792	2023	9.744206046	547030	66987244
Other	Other	Other - glass	Processing	3	China	GHGs	2023	11710.06	2023	1011.75	9596960	1392730000
Other	Other	Other - glass	Processing	3	China	Land use	2023	0.000738972	2023	0.65240116	9596960	1392730000

Appendix Five: Python Script for Random Forest Model Building

```python
# Import necessary libraries
import pandas as pd
from sklearn.model_selection import train_test_split, GridSearchCV, cross_val_score
from sklearn.ensemble import RandomForestRegressor
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import mean_absolute_error
import seaborn as sns
import matplotlib.pyplot as plt
import numpy as np
import joblib

# Load the raw material intensity dataset and eKPI dataset
file_path_materials = '/Users/paigeleeds/Desktop/raw-material-intensity.xlsx'
file_path_ekpi = '/Users/paigeleeds/Desktop/epl-ekpi-results.xlsx'

# Load the Excel files into DataFrames
df_materials = pd.read_excel(file_path_materials)
df_ekpi = pd.read_excel(file_path_ekpi)

# Convert 'Tier' column to the same data type (e.g., string/object) in both datasets
df_materials['Tier'] = df_materials['Tier'].astype(str)
df_ekpi['Tier'] = df_ekpi['Tier'].astype(str)

# Merge datasets on multiple columns
merged_df = pd.merge(df_materials, df_ekpi,
                     left_on=['Raw material', 'Process step', 'Material Group', 'Tier', 'Impact country'],
                     right_on=['Raw material', 'Process step', 'Raw material Group', 'Tier', 'Impact country'],
                     how='inner')

# Pivot the data to create separate columns for each eKPI type (e.g., Land Use, Waste, GHGs)
pivot_df = merged_df.pivot_table(index=['Raw material', 'Process step', 'Material Group'],
                                 columns='eKPI',
                                 values='eKPI result',
                                 aggfunc='sum').reset_index()

# Rename the columns to match the desired format
pivot_df.columns.name = None
pivot_df.rename(columns={
    'Land use': 'eKPI_Land_Use',
    'Waste': 'eKPI_Waste',
    'GHGs': 'eKPI_GHGs',
    'Water consumption': 'eKPI_Water_Consumption',
    'Air emissions': 'eKPI_Air_Emissions',
    'Water pollution': 'eKPI_Water_Pollution'
}, inplace=True)

# Fill NaN values with 0 for missing eKPI results
pivot_df.fillna(0, inplace=True)

# Feature Engineering: create a total environmental impact score by summing all six eKPIs
pivot_df['Total_Impact'] = pivot_df[['eKPI_Land_Use', 'eKPI_Waste', 'eKPI_GHGs',
                                     'eKPI_Water_Consumption', 'eKPI_Air_Emissions',
                                     'eKPI_Water_Pollution']].sum(axis=1)

# Add interaction terms to capture relationships between different eKPIs
pivot_df['Land_Use_Water_Consumption_Interaction'] = pivot_df['eKPI_Land_Use'] * pivot_df['eKPI_Water_Consumption']
pivot_df['GHGs_Air_Emissions_Interaction'] = pivot_df['eKPI_GHGs'] * pivot_df['eKPI_Air_Emissions']
```

```
# Define the features for the model
features = ['eKPI_Land_Use', 'eKPI_Waste', 'eKPI_GHGs', 'eKPI_Water_Consumption',
            'eKPI_Air_Emissions', 'eKPI_Water_Pollution', 'Land_Use_Water_Consumption_Interaction',
            'GHGs_Air_Emissions_Interaction']

# Scale the features
scaler = StandardScaler()
pivot_df[features] = scaler.fit_transform(pivot_df[features])

# Apply log transformation to reduce the impact of large values
pivot_df['Log_Total_Impact'] = np.log1p(pivot_df['Total_Impact'])

# Proceed with the train-test split using the log-transformed target
X_train, X_test, y_train, y_test = train_test_split(pivot_df[features], pivot_df['Log_Total_Impact'], test_size=0.2, random_state=42)

# Train the Random Forest model
model = RandomForestRegressor(random_state=42)
model.fit(X_train, y_train)

# Predict and evaluate the model (reverse the log transformation for evaluation)
y_pred = np.expm1(model.predict(X_test))
mae = mean_absolute_error(np.expm1(y_test), y_pred)
print(f'Mean Absolute Error after log transformation: {mae}')

# Grid Search for hyperparameter tuning
param_grid = {
    'n_estimators': [100, 200, 300],
    'max_depth': [None, 10, 20, 30],
    'min_samples_split': [2, 5, 10],
    'min_samples_leaf': [1, 2, 4]
}

grid_search = GridSearchCV(RandomForestRegressor(random_state=42), param_grid, cv=5, scoring='neg_mean_absolute_error')
grid_search.fit(X_train, y_train)

# Best parameters from GridSearch
print(f'Best parameters: {grid_search.best_params_}')

# Use the best model to make predictions
best_model = grid_search.best_estimator_
y_pred = best_model.predict(X_test)
mae = mean_absolute_error(np.expm1(y_test), np.expm1(y_pred))
print(f'Optimized Mean Absolute Error: {mae}')

# Perform cross-validation on the optimized Random Forest model
cv_scores_rf = cross_val_score(best_model, X_train, y_train, cv=5, scoring='neg_mean_absolute_error')
print(f'Cross-validated MAE (Random Forest): {-cv_scores_rf.mean()}')

# Save the best model for future use
joblib.dump(best_model, 'final_random_forest_model.pkl')

# Feature Importance
importances = best_model.feature_importances_
feature_importance_df = pd.DataFrame({'Feature': features, 'Importance': importances}).sort_values(by='Importance', ascending=False)
print(feature_importance_df)

# Boxplot to visualize outliers in Total Impact
sns.boxplot(pivot_df['Total_Impact'])
plt.show()
```

Appendix Six: Digital Consent Form

Consent form to allow a Master's Project to be stored in the library

I hereby give my consent for my dissertation / report to be stored in the library. *

Please note:

- Master's Project will be kept on the open shelves and are accessible to all library users i.e. students, staff, alumni and visitors.
- Master's Project are clearly labelled informing other library users that they should not be photocopied, scanned or photographed.
- Master's Project given to the library in digital form will be printed; the original digital copy will not be stored by the library.
- Master's Project may be withdrawn from the library.
- If your report contains any confidential commercial information, please ensure you seek permission from the company concerned for this to be lodged in the library. Written consent from companies may be forwarded onto the library separately.

Name:	Paige Leeds
UAL Email address:	p.leeds0720231@arts.ac.uk
Non-UAL Email address:	paigeeleedss@gmail.com
Course and level:	MSc. Fashion Analytics and Forecasting; Level 7
Title of Master's Project:	AI-Driven Material Optimization: A Framework for Sustainable Fashion Product Development
Signature:	Paige Leeds
Date:	12 Sep 2024

Consent form to allow a copy of your Master's Project to be used as an exemplar

I hereby give my consent for my Master's Project to be used as an exemplar for other students.

Please note:

- Master's Project will have any reflective statement/self-evaluation removed.
- Other students will not be allowed to print, copy or distribute your work.

Name:	Paige Leeds
UAL Email address:	p.leeds0720231@arts.ac.uk
Non-UAL Email address:	paigeeleedss@gmail.com
Course and level:	MSc. Fashion Analytics and Forecasting; Level 7
Title of Master's Project:	AI-Driven Material Optimization: A Framework for Sustainable Fashion Product Development
Signature:	Paige Leeds
Date:	12 Sep 2024