Sustainable Substitutes for Natural and Synthetic Textiles: An Algorithmic Clustering Approach

Chukwuemeka Oguike , Sook-Hee Evans , Sejal Dua , and Luke Steele
[a]Computer Science Department, Tufts University, chukwuemeka.oguike@tufts.edu
[b]Computer Science Department, Tufts University, sook-hee.evans@tufts.edu
[c]Mechanical Engineering Department, Tufts University, sejal.dua@tufts.edu
[d]Mechanical Engineering Department, Tufts University, luke.steele@tufts.edu

Abstract

Sustainability in the textile and fashion industry has become increasingly important as apparel and footwear consumption increases on a global scale. However, there are many challenges that face an engineer or designer attempting to improve their company's carbon footprint. There is little research on different materials and their environmental impact, and all relevant work that has been done in this area is primarily consulted retroactively, to assess the degree to which a product is "eco-friendly" for marketing purposes. The goal of this project is to offer a proactive, proof-of-concept solution to this problem; to guide those in this field with suggestions on more sustainable materials and fibers to use in their projects. Thus, a tool was developed that relies on the mechanical properties and environmental impact attributes of 23 different materials. Data science techniques including K-means clustering and Principal Component Analysis were harnessed in order to recommend functionally similar, yet more sustainable alternatives to the material provided by the user. The project culminated in this tool being available on a password-protected public domain, with the hopes that some designers, engineers, and corporations can leverage insights to minimize their environmental burden.

1 Introduction

Each year, the fashion industry is responsible for approximately 4.5 billion tons of carbon dioxide pollution, 79 trillion liters of water usage, 190 thousand tonnes of microplastic pollution, and over 92 million tons of textile waste [14]. These impacts represent a substantial portion of humanity's environmental footprint, and most concerningly, they are expected to increase if significant changes are not put into place [14]. With the industry as a whole trending towards fast fashion, where high quality, long lasting clothes are being replaced with lower quality items with short lifespans, the demand for clothing has grown rapidly. Clothing brands today produce about twice as many products as they did in the year 2000, and it is projected that this demand for clothing production will increase 2 percent every year [14]. As the industry continues to grow, so does the urgency of implementing more sustainable choices.

Although individual stakeholders in the fashion industry do not have the power to end the culture of fast fashion or transform it into a circular economy, one way companies and designers can substantially reduce the impact of their products on the environment is by choosing sustainable materials for their textiles. Different materials can have vastly different impacts on the environment. For example, cotton, a natural fiber, is frequently grown with fertilizer that pollutes local waterways, a concern that is not relevant with synthetic fibers. On the other hand, polyester, a man made fiber, uses nonrenewable fossil fuels as raw materials [9]. Making substitutions in materials clearly has a substantial effect on the environmental impact of the product. However, there are currently many barriers in the clothing sector that prevent designers from effectively choosing sustainable materials. First, the supply chain of a textile is very complicated. There are many processes involved in producing a textile that are carried out by different companies before it can be used, making overall sustainability difficult to account for [1]. Second, even though there are a few tools and guidelines, such as the Higg MSI, that exist to promote clothing sustainability, there is very little research on how to use these tools to effectively make changes in a company [10]. Third, using fabrics with lower environmental impacts often require tradeoffs in the functional properties of the material, which can significantly complicate the design process. These challenges often lead to designers leaving sustainability out of their decision making processes entirely [10]. Although these barriers often affect smaller, newer companies, they can also prevent or delay larger companies from transitioning to using sustainable materials and processes. Simplifying the process of choosing materials based on their sustainability is crucial to accelerating the fashion industry's transition to being eco-friendly.

The goal of this work is to empower material designers and engineers to find more sustainable textile alternatives in an easy and intuitive way. Since there are many niche sub-fields within the larger field of materials research and development, many designers and engineers lack data and insights on sustainable options that hold the same functional properties. A simple tool or interface could potentially solve this problem. Our target population for this tool is anyone who selects materials in the creation of textiles intended for clothing and footwear, or anyone who researches types of materials to be used for different products.

2 Features

2.1 Mechanical Properties

The main purpose of including mechanical properties in the analysis was to realistically incorporate the design choices that customers and designers in the fabric industry typically make in the cluster analysis. Initially a wide range of mechanical properties were considered and subsumed under three main design choices: durability, comfort, and aesthetics. According to American Society for Testing and Materials (ASTM) International [2], tear strength, surface friction coefficient, thermal conductivity, air permeability, pilling resistance, burst strength, flex resistance, and abrasion resistance are all informative properties to consider before selecting materials for a product. However, the scarcity of data limited the mechanical properties considered to only density, tenacity, percent elongation at break, initial modulus (also known as Young's Modulus), and moisture regain. While these properties represent a smaller subset of the data hoped, they provide designers a decision space with a significant amount of functional flexibility. The density allows designers to tailor design choices to applications where material weight is important; the tenacity factors into applications where fabric strength is vital; the Young's Modulus and elongation at break measure the stretchiness of materials; and the moisture regain factors into application where liquid absorption or repulsion is important. More details about how these properties are defined and how the measurements are taken is provided below:

- **Density**: The weight per volume of a material, in g/cm^3.
- **Tenacity**: A measure of a fiber's strength to weight ratio. It is calculated by dividing the load required to break a fiber by the fiber's linear mass density. Its units are g/Denier, where 1 Denier (den) is equal to 1 gram of mass per 9,000 meters of length.
- **Elongation at Break**: A measure of how much a fiber can stretch. It is measured by stretching a fiber until it breaks and dividing the change in length by the original length. It is a unitless ratio.
- **Initial Modulus**: A measure of stiffness or resistance to stretching. It is defined as the initial slope of the stress-strain curve as a fiber is stretched. The stress is the force divided by the cross-sectional area of the fiber and the strain is the elongation of the fiber relative to its initial length [15]. The units for this quantity are GPa.
- **Moisture Regain**: A measure of how much moisture a fiber holds. It is determined by finding the mass of a fiber at a standard humidity and temperature (65% humidity, 22°C) and the mass of the fiber after all moisture is removed through an oven drying process. The moisture regain is the mass of the water content (the difference between the two masses) divided by the oven dried mass. It is a unitless ratio.

For the five properties that were included, many of the data were compiled from a range of books [8, 4, 17, 12, 13], publications [7, 11], and online databases [6, 16, 15], with some literature focusing on synthetic fibers and others focusing on natural fibers. The mechanical attributes of common fibers are frequently listed as a range given the variability in how the fibers are sourced or manufactured. These ranges were converted to singular values by taking the average of the low and high endpoints.

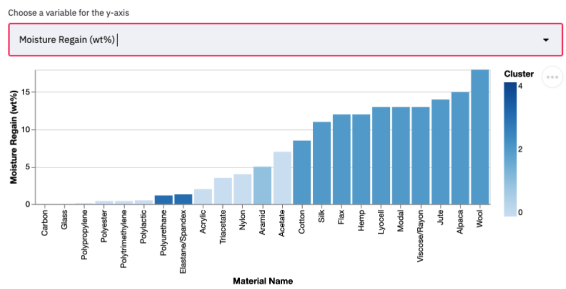

Figure 1: Example of interactive exploration of mechanical properties

An interactive tool is needed for this mechanical clustering because depending on the application, the design of products may require certain mechanical attributes to be maximized or minimized. For example, moisture regain should be minimized in apparel and sportswear products because when athletes sweat, the garment must wick moisture and dry fast. On the other hand, towels are intended to soak up moisture. The web tool allows for domain experts to interpret the results with these criteria in mind.

2.2 Environmental Properties

The Higg Material Sustainability Index (MSI) was created by the Sustainable Apparel Coalition, which is an alliance between 250 organizations including clothing brands, retailers, manufacturers, academic organizations, and nonprofits. Their goal was to develop a standardized method to measure sustainability performance and ultimately drive the industry forward [5]. In doing so, they created a product dashboard which can be utilized to assess the environmental impact of a textile material. the Higg MSI is comprised of seven different features, where lower scores indicate lower environmental impact. All of the features are then weighted and aggregated to result in a single score that represents the total overall environmental impact. These individual environmental attributes are defined as follows [5]:

- **Global Warming**: Toxicity of individual emissions released.
- **Eutrophication**: An excessive richness of nutrients in a body of water, often caused by runoff from farms. It results in dense plant growth which can harm animal life.
- **Water Scarcity**: Environmental damages of water usage.
- **Abiotic Resource Depletion (Fossil Fuels)**: Consumption of resources faster than replenishment.
- **Chemistry**: Efficiency of chemical use, toxicity and emissions.
- **Biogenic Carbon Content (kg C)**: Carbon sequestered from the atmosphere due to biomass growth.
- **Water Consumption (kg)**: The difference between water inputs and water outputs (amount of water discharged from the process)

3 Data Science Methods

Given the two disparate data sources used in this work, it was pertinent to clearly separate which data cleaning steps, data transformation steps, and data science techniques would be applied to each matrix. Ultimately, a two-fold approach was elected, wherein K-means clustering was applied to the matrix of fiber constituents and their mechanical properties to yield functionally similar materials, and each material in the resulting cluster of interest was ranked by user-weighted Higg environmental features and cost [3] to obtain suggestions. No robust data science techniques were implemented on the matrix of 195 materials and their environmental impact scores because environmental features are better suited for optimization rather than learning similarities via clustering. Additionally, this approach allows for the user's priorities to influence the suggestions they receive.

3.1 Principal Component Analysis

In order to adequately visualize the resulting material clusters, PCA was used to reduce the dimensionality of the data. Prior to performing PCA, one critical step was performed: standardization. Standardization is the process of putting different features on the same scale, with each scaled feature having a mean of 0 and a standard deviation of 1. Standardization allows for all features to be treated equally by the model. After fitting a PCA object to the standardized feature matrix, it was determined that three principal components should be used for K-means clustering because these components together account for over 80% of the variance in the data, as depicted by Figure 2.

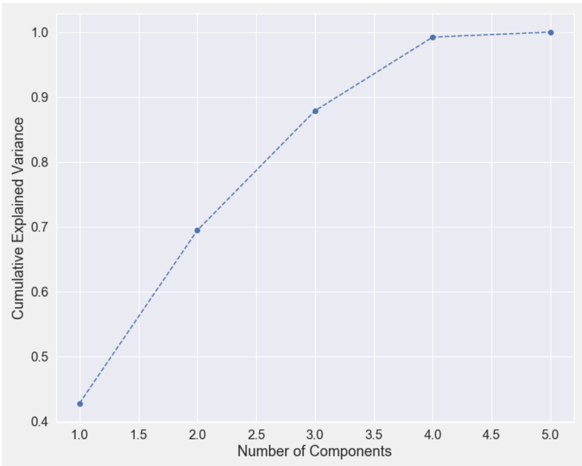

Figure 2: Explained Variance by Components

3.2 K-Means Clustering

Next, K-means clustering was used to identify similar material groups based on mechanical properties that were transformed to three principal components.

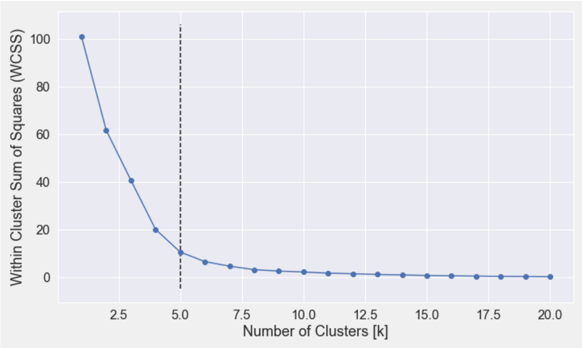

Figure 3: Optimal Number of Clusters

In order to select the number of clusters such that the bias-variance tradeoff is optimized, multiple K-means clustering algorithms are ran with varying values of K. Based on the inertia value outputted by the `sklearn` implementation of the K-means clustering algorithm, a metric called Within Cluster Sum of Squares or WCSS can be obtained for each iteration. A graph of WCSS versus K can be plotted and the elbow method can be used to identify the value of K for which the left-hand portion of the curve is steeply declining while the right-hand portion of the curve is much smoother [1]. As depicted in Figure 3, the optimal number of clusters was determined to be 5. From this point, K-means clustering was implemented on the standardized and transformed data with the specified number of clusters to obtain cluster labels for each fiber constituent in the dataset.

3.3 Filtering and Transformation

To connect the results of the clustering of mechanical properties for fiber constituents to the environmental impact scores for variants of fabrics that were obtained through the Higg portal, a naive mapping was implemented. The matrix of 195 Higg fabrics was filtered to only include fabrics which contain fiber constituents listed in the cluster of interest. If Cluster 2—which contains only natural fibers—was the identified cluster of interest, this filtering step would retain all materials containing Cotton, Flax, Hemp, Modal, Wool, etc. Interestingly, this includes blends which contain at least one fiber constituent in the cluster of interest. For example, a blend with 50% Cotton and 50% Polyester would remain in the filtered matrix. Though polyester may contribute to mechanical differences, it is up to the designer or engineer using the tool to discern from the visualization if the suggested substitute would fulfill product constraints or not.

[1]This elbow point or kink was computed programmatically using an open source Python library

Following the filtering step to map the mechanically similar fibers to a mechanically similar subspace of fabrics, the environmental impact scores were normalized. This choice was made to ensure that the weighted sum of each environmental feature and the user-selected weights or coefficients for these features did not amplify any signals caused by different features on different scales. The lowest five weighted scores were displayed to the user as top suggestions for the inputted material.

4 Web Application

This work culminated in the development of a web application which enables a user to derive their own insights. Given that the data has already been gathered and is pre-loaded into the tool, the clustering results remain consistent regardless of user input.

4.1 Clustering Results

Cluster #	Fiber Constituent(s)
0	Acetate
	Acrylic
	Nylon
	Polyester
	Polylactic Acid (PLA)
	Polypropylene
	Polytrimethylene Terephthalate (PTT)
	Triacetate
1	Aramid
2	Alpaca
	Cotton
	Flax
	Hemp
	Jute
	Lyocell
	Modal
	Silk
	Vicose/Rayon
	Wool
3	Elastane/Spandex
	Polyurethane
4	Carbon fiber
	Glass

Table 1: Similar fiber constituent(s) after performing PCA and K-means Clustering

The resulting clusters after using PCA and K-means clustering with 5 clusters are shown in Table 1. Notably, all the fibers in Cluster 0 are synthetic fibers, and all the fibers in Cluster 2 are natural fibers. Natural

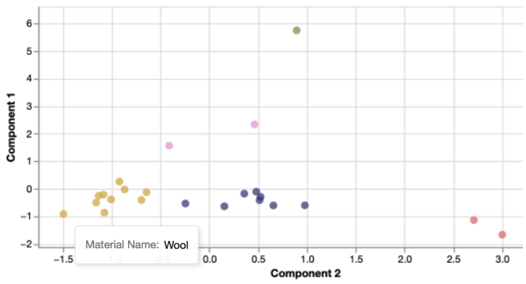

Figure 4: Two-dimensional Visualization of the fiber constituent(s) after PCA and K-Means Clustering

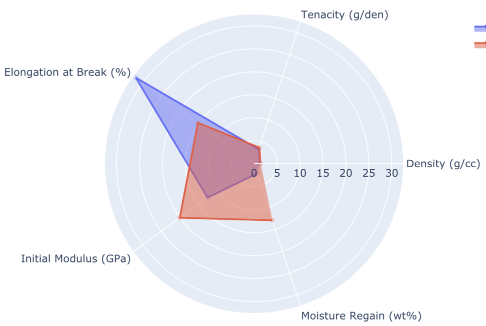

Figure 5: Visual comparison of the average mechanical properties of Cluster 1 and Cluster 2

fibers tend to be mechanically similar to other natural fibers, but not so much to synthetic fibers. Synthetic fibers tend to be mechanically similar to other synthetic fibers. These results were validating, and after further analysis by visually comparing the cluster graph shown in Figure 5, it was clear that the fibers were well-grouped together with mechanically similar fibers. The radar chart shown in Figure 4 overlays the average mechanical properties of Cluster 1 and Cluster 2, and shows how the average properties differ significantly. Despite a couple of outliers, the cluster groupings identified fiber constituents with fairly homogeneous mechanical features. The clustering approach implemented in this work yields promising results for future performance-focused innovation in the textile industry.

4.2 Use Case

An example case study was performed with Hemp fiber as the original intended material. Shown in Figure 6 are the ranked recommendations of sustainable alternatives to Hemp that are functionally similar and reside in the same cluster.

The most sustainable fiber that is mechanically similar to Hemp was Modal. Figure 7 shows the difference in environmental impacts between Hemp and Modal. Modal is shown to be a more sustainable alternative with regards to every single environmental feature. Modal is also much more budget-friendly than Hemp, as shown by the cost spire.

Sustainable Alternatives by Lowest Weighted Score

	Material Name	Fiber Constituent(s)	Weighted Score
2	Material 100	['Modal']	0.9752
168	Material BL27	['Lyocell', 'Polyester']	0.9755
104	Material 57	['Cotton']	1.0615
176	Material BL35	['Polyester', 'Viscose/Rayon']	1.0707
169	Material BL28	['Modal', 'Polyester']	1.0726

Figure 6: Top five sustainable alternatives for Hemp

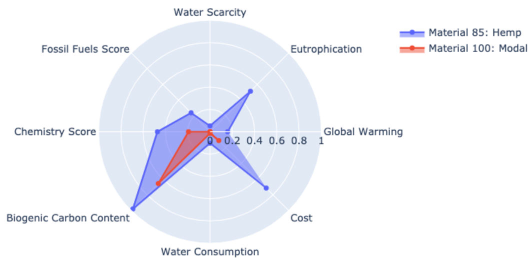

Figure 7: Environmental impact comparison of Hemp and Modal

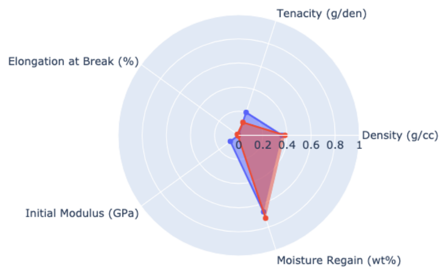

Figure 8: Mechanical comparison of Hemp and Modal

Figure 8 shows the comparison of mechanical features of Hemp and Modal. Since they are in the same cluster, it is expected that they have fairly similar functional properties. The radar chart validates that expectation, as the overlaid Hemp and Modal comparison takes almost the exact same shape. It is ultimately up to the user to determine if the mechanical properties of the top suggestion are sufficient for the desired functionality of the product.

In the fashion industry and larger textile industry, design and budget constraints vary. Sustainability is certainly emphasized in the design of many products, but certain environmental impact areas may not be of utmost concern. To account for this, an interactive slider

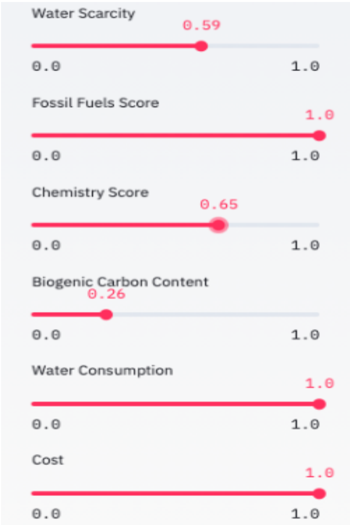

Figure 9: Interactive sliders

feature has been implemented on the left-hand side of the web interface. In Figure 9, an example is provided to demonstrate how a user might alter the priority level of the environmental features. These values ranging from 0 to 1 are then used as coefficients in the computation of the weighted environmental score which is minimized to yield the top five suggestions.

5 Discussion

5.1 Research Contribution

The tool developed is an easy to use and informative interface, designed with materials engineers and textile designers in mind. It enables users to visualize and consider the benefits and drawbacks of different options, thus raising awareness of the potentially more sustainable choices for their projects. Moreover, it allows the user to prioritize different sustainability features. For example, a company in a drought-stricken area may prefer to minimize water usage, and can maximize the weight on that, while lowering weights on other, less important features. Another variable included in the tool is the cost of staple fibers, which was provided by a domain expert [3]. Many designers and engineers will also have an interest in keeping costs low, and can prioritize a budget-friendly suggestion by simply sliding up the ‘cost’ weight to the maximum (1.0). The tool has also been launched on a password-protected public domain, which will allow any authorized users to access it, interact with it, and gain insights which help to inform their design choices.

5.2 Bias & Validity

In analyzing the efficacy of this tool, it is worth noting bias and validity limitations that were introduced in the data collection and data cleaning phases. Firstly, the different materials and their life-cycle processes were manually selected by a domain expert. This introduced some bias because there were a multitude of possible combinations for each material preparation, such as the source of the material, the knitting or weaving method, the size of yarn used, usage of certain chemicals in preparation, etc. Secondly, the only blends included were comprised of 50% one material and 50% another material, even though this perfectly even split is not necessarily common in industry. Thirdly, the ranges given for the mechanical properties were squashed into one single value which averaged the lower and upper bounds. Ideally, there would have been several different variants of fiber constituents. For example, some types of polypropylene are stretchier than others, as reflected by a higher percent elongation at break. This noise should remain in the dataset rather than introducing bias by way of simplification. Furthermore, construct validity issues were raised by implementing a naive mapping to connect mechanical fiber data to environmental scores for fabrics.

Lastly, Higg scores are also inherently biased, as the cradle-to-gate scoring methodology for the seven environmental attributes was subjectively defined by the Sustainable Apparel Coalition. The Higg MSI introduces even more bias given that the formula is subjectively pre-defined and is debated within the industry. However, this tool attempts to remove this additional bias by doing away with the MSI altogether. Instead, the interface allows the user to choose the coefficients of the seven-term equation, prioritizing which feature is most important to them. As the industry standard for environmental scoring evolves over time, this user-centric approach will help to mitigate the bias introduced by third-party organizations whose scoring methodologies may favor one particularly side of the industry.

5.3 Data Sparsity

As previously mentioned, data was gathered for 23 materials, with 5 mechanical properties for each of those. Ideally, the user would be able to choose from many more materials and blends. There were many other sources that had some mechanical properties, but full datasets that applied to other properties were not publicly available. Additionally, synthetic and natural fiber studies are different realms in the textile research space, thus it was not feasible to find a single source that reported on properties of both synthetic and natural fibers. The properties came from a handful of different sources, but the insights derived from this work would be much more reputable if the entire dataset was acquired from a single trusted source or database.

Functional properties of blends could not be inferred from the mechanical properties of each of the single constituents, as blends perform much differently than a weighted sum of the individual fibers. More data on blends would also offer useful insights for users of the tool, as blends are often used in industry. If data was available on the mechanical properties of textiles prepared in different ways, the user could select different life cycle processes and the tool could offer alternative preparation strategies, as well as alternative textiles. Preferably, the entire Higg dataset would be available for download so that every possible combination on the Higg dashboard would be included.

5.4 Future Work

Next steps for the project include gathering more comprehensive blend data and mechanical property data. This tool could potentially be used as a proof of concept demonstration for labs, start-ups, or other companies in the industry who may be able to grant access to mechanical property data, in return for a more extensive tool that could be beneficial towards their projects and employees. More fabric constituent data should also be found, so that the user can explore more varied and thorough suggestions.

References

[1] Sustainability of textiles. `ec.europa.eu/environment/industry/retail/pdf/issue_paper_textiles.pdf`. [Online; accessed 12-02-2020].

[2] Textile standards. `https://www.astm.org/Standards/textile-standards.html`. [Online; accessed 12-05-2020].

[3] John Amirtharaj. private communication, 2020. [Email correspondence; 11-24-2020].

[4] A. R. Bunsell. *Handbook of Tensile Properties of Textile and Technical Fibres*. Woodhead Publishing, 1st edition, 2009. [Online; accessed 12-05-2020].

[5] Sustainable Apparel Coalition. Higg materials sustainability index (msi) methodology. `howtohigg.org/wp-content/uploads/2020/07/Higg-MSI-Methodology-July-31-2020.pdf`. [Online; accessed 12-05-2020].

[6] CAMEO: Conservation and Art Materials Encyclopedia Online. Properties of synthetic fibers. `http://cameo.mfa.org/images/a/a2/Download_file_43.pdf`. [Online; accessed 11-27-2020].

[7] Nina Graupner, Fabrizio Sarasini, and Jörg Müssig. Ductile viscose fibres and stiff basalt fibres for composite applications – an overview and the potential of hybridisation. *Composites Part B: Engineering*, 194:108041, 2020.

[8] Sara J. Kadolph and Sara B. Marcketti. *Textiles*. Pearson, 12th edition, 2017. [Online; accessed 12-05-2020].

[9] E. Kalliala and P. Nousiainen. Life cycle assessment: Environmental profile of cotton and polyester-cotton fabrics. *Autex Research Journal*, 1, 2000.

[10] Anika Kozlowski, Michal Bardecki, and Cory Searcy. Tools for sustainable fashion design: An analysis of their fitness for purpose. *Sustainability*, 11:3581, 06 2019.

[11] Aruan Mohd Ghazali, K. Pickering, and Tan Le. A review of recent developments in natural fibre composites and their mechanical performance. *Composites Part A Applied Science and Manufacturing*, 83, 09 2015.

[12] H. V. Sreenivasa Murthy. *Introduction to Textile Fibres*. Woodhead Publishing India Pvt Limited, 2006. [Online; accessed 12-05-2020].

[13] Yasir Nawab, Munir Ashraf, Tanveer Hussain, Abher Rasheed, Khubab Shaker, Abdul Basit, Madeha Jabbar, Zulfiqar Malik, Hussain Fiaz, Khurram Shehzad, Muhammad Umair, Sheraz Ahmad, Muhammad Maqsood, Waqas Ashraf, Sajjad Baig, Talha Hamdani, and Muhammad Zohaib. *Textile Engineering. An Introduction*, pages 143–159. 07 2016. [Online; accessed 12-05-2020].

[14] Kirsi Niinimäki, Greg Peters, Helena Dahlbo, Patsy Perry, Timo Rissanen, and Alison Gwilt. The environmental price of fast fashion. *Nature Reviews Earth & Environment*, 1:189–200, 04 2020.

[15] Omnexus. Young's modulus. `https://omnexus.specialchem.com/polymer-properties/properties/young-modulus`. [Online; accessed 11-27-2020].

[16] Chemical Retrieval on the Web (CROW). Properties of synthetic fibers. `https://polymerdatabase.com/Fibers/Fibers2.html`. [Online; accessed 11-27-2020].

[17] Asim Shahzad and Sana Ullah Nasir. Mechanical properties of natural fiber/synthetic fiber reinforced polymer hybrid composites. In Mohammad Jawaid, Salit Mohd Sapuan, and Othman Y. Alothman, editors, *Green Biocomposites: Manufacturing and Properties*, pages 355–396, Cham, 2017. Springer International Publishing. [Online; accessed 12-05-2020].