Proceedings of The 2021 ASME International Design Engineering Technical Conferences & Computers and Information in Engineering Conference
IDETC/CIE 2021
August 17-20, 2021, Virtual, Online

DETC2021-70990

A PROBABILISTIC APPROACH FOR ESTIMATING THE ENVIRONMENTAL IMPACT OF NOVEL PRODUCT CONCEPTS

Vincenzo Ferrero, Chris Hoyle, Bryony DuPont
Design Engineering Laboratory
School of Mechanical, Industrial and Manufacturing Engineering
Oregon State University
Corvallis, Oregon, 97331
Email: Ferrerov@oregonstate.edu
Email: chris.hoyle@oregonstate.edu
Email: Bryony.DuPont@oregonstate.edu

ABSTRACT

Global concerns about climate change and resource management have escalated the need for sustainable consumer products. In light of this, sustainable design methodologies that supplement the product design process are needed. Current research focuses on developing sustainable design curricula, adapting classical design methods to accommodate environmental sustainability, and sustainability tools that are applicable during the early design phase. However, concurrent work suggests that sustainability-marketed and innovative products still lack a reduction of environmental impact compared to conventional products. Life cycle assessment (LCA) has proven to be an exceptional tool used to assess the environmental impact of a realized product. However, LCA is a reactive tool that does not proactively reduce the environmental impact of novel product concepts. Here we develop a novel methodology, the PeeP method, using historical product LCA data with kernel density estimation to provide an estimated environmental impact range for a given product design. The PeeP method is tested using a series of case studies exploring four different products. Results suggest that probability density estimations developed through this method reflect the environmental impact of the product at both the product and component level. In the context of sustainable design research, the PeeP method is a viable methodology for assessing product design environmental impact prior to product realization. Our methodology can allow designers to identify high-impact components and reduce the cost of product redesign in practice.

1 INTRODUCTION

Sustainable product design is defined as the consideration of the three tenets of sustainability (ecological, economic, and social) impact during product development. Sustainable product design, specifically *Design for the Environment* (*DfE*), continues to be an emergent topic in engineering design and a growing concern in novel product development. The Sustainable Market Share Index, published by the Center for Sustainable Business, suggests that the sustainable product market has grown from 88.2 billion USD to 113.9 billion USD from 2013 to 2018 [1]. Furthermore, sustainability-marketed products make up 54.7 percent of consumer goods market growth, while only making up 16.1 percent of the consumer goods market share [2]. These trends suggest that today's consumer is becoming more sustainability-aware and shifting purchasing habits toward sustainability-marketed products. Unfortunately, many adopted DfE methods rely on reactive tools such as life cycles assessment in iterative approaches after a product has been fully realized.

Tools like life cycle assessment require a finalized product design where manufacturing methods, transportation logistics, and material selection are defined. In practice, engineering designers lack robust proactive tools for the consideration of environmental impact during the early design phase.

In engineering design research, several methods have been introduced to promote sustainable product design earlier in concept development. The popular Theory of Inventive Problem Solving (TRIZ) method for creative problem solving has been adapted to enabling sustainable product design [3]. Eco-design tools, such as the GREEn Quiz (Guidelines and Regulations for Early design for the Environment), allow designers to explore sustainable design guidelines early in the design phase [4–6]. Recently, there has been interest in improving environmental responsibility education in engineering design curricula [7]. However, designers are still struggling to make meaningful shifts toward environmentally-friendly product design. This is evident in recent research which suggests that eco-labeled products and innovative products can fail to be more sustainable than the conventional alternatives they are replacing [8, 9].

Here we aim to contribute to the growing research area of sustainable design methods by introducing a **P**robabilistic approach for **E**stimating the **E**nvironmental impact of **P**roduct concepts, termed the *PeeP* method. The PeeP method uses kernel density estimation with a product repository and historical product LCA data to estimate the environmental impact range of a product, based on the designer-supplied bill of materials. Using Monte Carlo sampling, the PeeP method generates sample data points representative of each component's environmental impact on the supplied complete or incomplete bill of materials. Finally, through kernel density estimation, probability density functions are presented for each component and the overall product concept, denoting the ranges of the probable environmental impact of the design and the contribution by each component. In this paper, we demonstrate and validate the PeeP methodology through the use of sourced design repository data and four example product concepts.

2 BACKGROUND

The PeeP method described in this paper leverages repository data used to drive sustainable product design. However, the data requirements for the PeeP method are approachable and can be satisfied with other data sources that a design may have access to. The foundations of the PeeP method rely on background knowledge of data-driven product design, design repositories, component naming standardization, sustainable product design, research in life cycle assessment, and probability estimation.

2.1 Data-Driven Product Design

Data-driven product design has been at the forefront of product development research for the last decade. 'Big Data,' in tandem with data-mining, allows designers to make design decisions based on intuition supported by evidence parsed from datasets. [10–12]. Wang et al. mined product review data using sentiment analysis to capture emotive response gaps between design intention and consumer preference [13].Zhang et al. also used sentiment analysis on reviews to understand potential pain points of cellular phone design, indicating areas of improvement [14]. The current literature in data-driven design most commonly uses text mining, sentiment analysis, and available customer data [15–19]. However, other data sources have been used in data-driven product design.

Kong, Li & Zheng characterized areas of mass product customization using expert knowledge and patent data [20]. Sangelkar 2012, leveraged association rule mining against expert knowledge and product data to identify changes in user activity between universal and conventional products [16]. Yan & Xu 2007, used support vector machines with design-time data to predict design times for plastic mold [21]. Generated data has also been used to create decision trees for cellular phone configuration design [22].

In previous literature, there is an abundance of generated (i.e., the researchers use algorithmically generated data that represents real data) or collected user data (i.e., parsed from online websites or surveys) used in data-driven product design. However, there is an under-representation of 'real' prior product data used in data-driven product design methodologies. The proposed PeeP method moves the data-driven design state-of-the-art toward employing product data from evolving product databases. By developing a methodology considering supplied product data flexibility, the PeeP method is repeatable and useful given a variety of source data.

2.2 Design Repositories and Knowledge Discovery

Design repositories house design data to provide knowledge to the end-user [23, 24]. Given concerns about intellectual property, there is a lack of available and robust design repositories available to researchers and industry alike. Efforts have been made to encourage the adoption of such databases by publishing research demonstrating repository standardized schemes and methodologies that streamline repository utilization. [25–27]. Despite weak adoption, novel design repositories are still being introduced in research [28]. For demonstrating the PeeP method, we use the Oregon State University Design Repository (OSDR) and the Oregon State Sustainable Design Repository(SDR), both of which have been extensively employed in academic research and are consistently curated and maintained [29–31].

Design repository research primarily focuses on using repository data to inform and expand current function-based design theories. One such effort has been to use design repository data to aid in and automating concept generation [32–34] from functional modeling. Functional modeling is a design tool used in the early design phase to describe the function relationships required to meet the design requirements of a product concept. Functional modeling is often completed without the need of explicit design decisions. Leveraging functional-based design as a bridge to the early design phase, Soria et al. introduced a methodology for identifying human error by using functional modeling. This method was tested with repository data [35,36]. In related work, automate functional modeling has been explored through the use of the Oregon State University design repository [37]. The primary goal of this work is to utilized automated functional assignment to aid in *Design-for-X (DfX)* methodologies [38]. *Design-for-X* is defined as the specialize approach toward meeting a specific design objective (X). Though these methods look to apply *Design-for-X (DfX)* objectives to the early design phase through function, these methods often only used design repository data in validation. There is a lack of leveraging the repository product data in primary knowledge discovery, as opposed to secondary uses such as case study facilitation.

Design repositories specifically are potential candidates for applying the theory of Knowledge Discovery in Databases (KDD) [39–41]. As in, repositories can provide the necessary 'big data' required in modern data-mining approaches. However, modern data-mining methods are reliant on extensive and complete knowledge data sources. To explore this space, Williams et al. discuss design repository effectiveness as a data source for neural networks [42]. For advanced knowledge discovery methods, modern repositories may not provide accurate results. In contrast, there is still a need to explore novel methods that do not directly rely on extensive data sets. The PeeP method provides a stepping stone in research by using repository data with probability-based knowledge discovery to draw conclusions based on data available.

2.3 Component Naming Standardization

In addition to answering the need for standardized design repository schema, recent research has defined how to standardize the data within the repository structure. Research has introduced standardization for naming functions within a functional decomposition and when describing component functions [43–45]. This standard terminology is called *functional-basis terms*. Future work expanded the functional basis terms to include a new basis terms list to describe components [46]. Our proposed methodology capitalizes on this research by renaming common-named components with component basis terms. The component basis standardization reduces data noise introduced by user-defined common component names.

2.4 Life Cycle Assessment (LCA)

Life Cycle Assessment is a method of determining the total life-cycle sustainability impacts of a process, system, or design [47,48]. Though methods exist to estimate social and fiscal sustainability, LCA is primarily used to explore environmental sustainability [49,50]. LCA works by taking life cycle inventory data, processing the data through an LCA methodology, and displaying numerical results that can be used to interpret the sustainability of a product. In practice, LCA is used to compare products, systems, and processes. There are several LCA methodologies, most notably ReCiPe, TRACI, and CML [51–53]. The use of standardized LCA methods is facilitated through LCA programs such as SIMAPRO or GaBi [54, 55]. Life Cycle Assessment remains a foundation tenant in sustainability-based research. However, LCA introduces research, practical, and utility challenges.

A primary challenge for bringing Life Cycle Assessment to the early design stage of product design is that LCA is a reactive tool. LCA is often used to measure the impact of the product after the completion of the design process. This limits LCA as an iteration tool, rather than a proactive tool to mitigate impact prior to product realization. Other challenges include lack of geometry data, time investment, the ability to translate LCA indicators intuitively, handling uncertainty in LCA, and how to address bias introduced through assumptions [56–58].

In the areas of material and building development research, LCA has been successfully used during design to mitigate impact [59–61]. However, the use scenarios for buildings and materials are easily modeled using abundant existing historical data with low uncertainty. For data-driven product design research, proactive LCA-based methodologies are sparse likely due to the complex assumptions needed for product LCAs and lack of available historical product-level LCA data. However, successful research examples do exist.

In 2010, Bohm et al. demonstrated that design repositories can be used to integrate LCA data into the early design phase of product design. However, this method still relies on the manual use of LCA methods [62]. In 2017, Arlitt et al. introduced the Function-based Design for sustainability method (FDS). This work provides introductory steps toward leveraging component-level LCA, function, and component data to enhance sustainable design knowledge during early product design. The FDS method allows users to identify the most environmentally impactful functions and clusters components that solve that function by environmental impact [63]. Marwah et al. also used clustering with component level LCA data to estimate the impact of introduced components based on likeness to components within clusters [64]. Our proposed methodology aims to follow the state-of-the-art by adding novel methodologies that bring LCA into the early design phase while addressing some of the highlight issues found in LCA driven research.

2.5 Kernel Density Estimation

Kernel Density Estimation (KDE), realized by Emanuel Parzen in the 1960s, is a non-parametric method for determining the probability density function of random variables [65]. The KDE method uses kernel functions to interpolate a density function across a range of random variables. The kernel function determines the contribution of neighboring observations - within the bandwidth or 'smoothing factor'- based on the distance, and determines the aggregated probability of the origin data point. Today, kernel density estimation research has focused on expanding the method to include multivariate capabilities and optimizing bandwidth selection [66–70].

In related research, KDE is used as a powerful statistical tool used in predictive and forecasting analysis, particularly in the socio-geography space. Kernel density estimation is commonly used to predict crime and create crime 'hot spot' maps using weather, temporal, social media, demographics, and geolocation data [71–74]. A conditional variant of KDE was used to forecast electricity use through smart meter data and energy-cost tariff models [75].

In the engineering design domain, KDE has been used to track shape correspondence between two non-rigid 3D models [76]. Reliability-Based Robust Design Optimization was combined with KDE to developed probability density functions for manufacturing uncertainty in electronic power steering motors [77]. Kernel density estimation has also been leveraged to determine the probability of environmental conditions surrounding offshore wind turbines [78]. The results of this work are useful in making structural decisions of wind turbine designs. Recently, KDE has proven useful in modeling tolerances between components [79]. Though the use of KDE is limited in the engineering design space, our proposed methodology explores the implications of KDE in such a field. PIn particular to the data-driven product design, KDE is beneficial over classical probability methods such as Bayesian statistics. Product data is often disparate, unbalanced, and sparse; thus, assuming a prior distribution is challenging.

3 METHODS

Here we describe the novel probabilistic method for estimating the environmental impact of product concepts (PeeP). This section includes data selection and processing, probability distribution, meta probability distributions, and the assumptions and limitations. Along with the description of the methodology, an example case is discussed throughout this section. An overview of the PeeP method is shown in Figure 1.

3.1 Data Selection and Processing

The data need for the PeeP method requires that the practitioner have multiple consumer product product bills of material

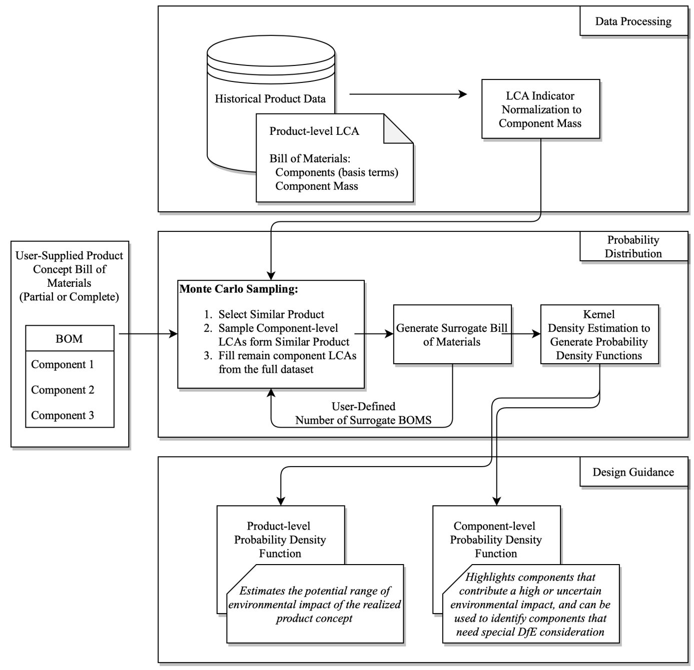


FIGURE 1: PeeP method overview

(including component weight), along with Life Cycle Assessments completed for each product. In industry practice, user-supplied product data is ideally from the same product family. This data is then subject to further processing as described below.

3.1.1 Data Selection

To facilitate the demonstration of the methods presented here, we need to source specific product information. The required product information includes product name, bill of materials (in component basis terms), component weight, and LCA impact data. Component basis terms are selected to remove the variability caused by the common component name. In this regard, component basis terms allow for standardization of the components used in the PeeP method, as presented in section 4.

For method validation, the prescribed data is sourced from the Oregon State University Design Repository (OSDR) and Sustainable Design Repository (SDR) [24, 30]. Using historical product data from both repositories, a total of 44 products are compiled into a data set. The sourced data set includes 487 components as defined by 70 component basis terms. For each product, the LCA software GaBi is used to determine the environmental impact of each product [55]. The sourcing of LCA data for each product is subject to the assumptions made during the creation of the SDR. Succinctly put, the ReCiPe indicators were determined under the assumption all products are landfilled, transported the same distance, and manufactured in the same locations. In the demonstration, we chose to use the aggregated normalized ReCiPe end-point indicator for environmental impact, Species.yr [51]. Species.yr is the measure damage to ecosystem diversity quantified by the loss of species during a year time increment with consideration of terrestrial, freshwater, and marine species loss. In practice, the approach described here can utilize any number of LCA indicators from a variety of LCA tools. Table 1 shows an example of the data sourced for demonstration.

3.1.2 Processing

The sourced data needs to be processed to create the needed and exhaustive probabilistic distributions. The data processing procedure begins by normalizing the LCA indicators to the mass, m_p, of the product, as shown in eq. 1.

$$LCA_{norm} = LCA_{product}/m_p \tag{1}$$

Classically LCA is represented though functional units. However, LCAs of different products and sources are likely not to share the same functional unit. The normalization of LCA, LCA_{norm}, indicators by mass allows for the direct comparison of the sourced historical product data without worry of disparity in product size or functional unit definitions. The product LCA indicator is reduced to impact per gram of product mass. Within the normalized indicator, the components have a ratio of contribution to the overall product LCA,LCA_{norm}. To develop this ratio r_c, the component mass m_c is divided by the product mass.The perceived component impact is calculated, shown in eq. 2, by multiplying the normalized LCA indicator, LCA_{norm}, and component mass ratio, r_c.

$$LCA_{comp} = LCA_{norm} * m_c/m_p \tag{2}$$

For each product, components are combined in tuples that represent every combination of components found within a product and the summation of their adjusted component impact, eq. 3 .

$$LCA_{tuple} = \sum_{i=1}^{n} LCA_{comp_n} \tag{3}$$

This tuple combination allows for an increased number of probability distributions. Multiple probability distributions allow our approach to anticipate and relate the user-supplied bill of materials to similar combinations of components. The creation of more data points allows for the account of percent similarity between a supplied bill of material and the created distributions.

A final example of completed data is shown in table 2. Component and component tuples, compared to the overall impact of the product, are analogous to the impact of a product the component is generally found. This data used to create the environmental impact meta-probability distributions of an incomplete bill of materials. Component and component tuples as compared to the normalized and adjusted component impact are used in probability summation to create the impact meta probability distributions of the known bill of materials.

3.2 Probability Distribution

The processed data points are grouped into clusters defined by component basis name. The individual component clusters are used as inputs for the KDE algorithm to provide probability distributions for specific components and component clusters. These distributions add an exhaustive approach to displaying knowledge for the end-user who may be only interested in the impact of specific components and configurations of components. Furthermore, the calculations of probability distributions for component configurations can reduce the computation time when compiling the meta-distributions outlined in section 3.3.

For the PeeP method, probability distributions are estimated using the Kernel Density Estimation (KDE) method [65]. The KDE method allows for estimation of the probability density function given a set of independent and random variables that have no known prior density distribution. The Kernel Density Estimator is shown in, eq. 4,

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} k\left(\frac{x-x_i}{h}\right) \tag{4}$$

where k is the kernel function and h is the bandwidth. For this work, the kernel function is assumed to be Gaussian, shown in eq. 5. The Gaussian function is chosen for convenience and generally does not play a key role in KDE [80].

$$k(x) = \frac{1}{\sqrt{2\pi}}\exp(-\tfrac{1}{2}x^2) \tag{5}$$

In contrast, the bandwidth parameter greatly influences KDE. The bandwidth parameter, h, provides smoothing to the generated probability density function. If the bandwidth is too small, the PDF is noisy, and the variables' randomness is highlighted. If the bandwidth is too large, the resultant PDf can be 'over smoothed,' and probabilistic features of the data can be lost. Research is ongoing for optimizing bandwidth selection in the KDE process [66]. For this work, the rule-of-thumb approach, Scott's Rule, calculates the optimum bandwidth factor [81]. The Scott's Rule bandwidth equation is shown in eq. 6, where σ is the standard deviation of the the observations and n is the number of observations. The Scott's Rule is selected for its popularity and

TABLE 1: EXAMPLE PRODUCT DATA REQUIREMENTS FOR THE PeeP METHOD

Product ID	Product Name	Component Name	Component Basis	Weight (g)	Product LCA (Species.yr)
830	Mech. Pencil	Top	Cap	1.20	3.92×10^{-10}
830	Mech. Pencil	Tip	Insert	0.33	3.92×10^{-10}
830	Mech. Pencil	Spring	Spring	0.11	3.92×10^{-10}
830	Mech. Pencil	Feeding Assembly	Assembly	0.02	3.92×10^{-10}
830	Mech. Pencil	Tube	Tube	5.87	3.92×10^{-10}
830	Mech. Pencil	Eraser	Unclassified	0.54	3.92×10^{-10}
830	Mech. Pencil	Lead	Unclassified	0.01	3.92×10^{-10}

TABLE 2: EXAMPLE PRODUCT DATA LCA NORMALIZATION

Mechanical Pencil				
ID	Tuple Name	Weight (g)	Product LCA (Species.yr)	Normalized LCA (Species.yr / gram)
1	Mech. Pencil	8.08	3.92×10^{-10}	2.19×10^{-11}
2	Top	1.20	3.92×10^{-10}	3.25×10^{-12}
3	Tube	5.87	3.92×10^{-10}	1.59×10^{-11}
...	....	......	....	......
9	Top + Tube	7.07	3.92×10^{-10}	1.92×10^{-11}
...	....	......	....	......
N	Top + Tip ... + Lead	8.08	3.92×10^{-10}	2.19×10^{-11}

validation in research.

$$h = \left[\frac{24\sqrt{\pi}\sigma^3}{n}\right]^{\frac{1}{3}} \tag{6}$$

Overall, The KDE method is chosen due to the uniqueness of the densities found within the data set. Within the example data set used for method validation, it is clear that PDF for each component varies greatly. One component probability distribution may exhibit normal behavior, while other components can exhibit exponential or logarithmic distribution behavior. For certain, it can not be determined if any given standard probability distribution model can define the supplied data set. KDE represents a robust method of estimating PDF regardless of the data supplied and limits assumptions drawn by fitting distributions to incomplete data sets.

3.3 Meta Probability Distributions

Probabilistic meta-distributions are generated to estimate the potential impact of a product design given a completed or partial bill of materials. The data used for the meta-distributions is generated through the summation of component cluster-related LCA impact data representative of each component within the product concept BOM. Using KDE, meta-probability density functions are generated based on multiple components found within the product concept BOM. In short, the meta-data used as input to the KDE is a number of surrogate LCA-rich BOMs generated from the user-supplied product concept BOM. The creation of meta-data follows two different approaches depending on the completeness of the supplies bill of material. The generated meta-data is used in KDE following the same equations found in section 3.2.

For the completed bill of materials, surrogate BOMs are generated by taking the Cartesian product of possible LCA impact indicators per component and cluster found in the product concept BOM. Computationally, taking the Cartesian product of a large bill of materials is taxing. A Monte Carlo sampling approach is used to limit the computation time needed to generate representative data. For each Monte Carlo iteration, several surrogate BOMs are generated. For the initial surrogate BOM, each product concept BOM component is randomly assigned a normalized impact indicator as grabbed from the set of LCA indicators representative of that component. To generated the remaining data points in the iteration, each similar product in the data set is explored. A similar product is defined by sharing at least two components with the provided complete bill of materials. The identified similar component cluster and related LCA impact are added to a surrogate BOM. The remaining non-similar component LCA impacts are randomly assigned akin to the first data point generated. The meta-data is generated over many iterations to create a surrogate BOM data set with a size governed by eq. 7; where ep is the number of iterations and k is the number of similar products. Finally, within each surrogate BOM, the component cluster LCA impacts are summed to estimate the overall weight normalized environmental impact of the bill of materials.

$$N_{metadata} = ep \times \left(x_i + \sum_{i=1}^{k} x_k \right) \tag{7}$$

During metadata generation, we can not assume that the historical product data provided for the PeeP method is complete. However, we do assert that clusters that make up multiple components within the supplied BOM are more representative of the potential impact of the product design, primarily if the supplied data set is closely related to the supplied BOM product design. As such, the LCA impacts related to similar products within the data set are favored over using independent components to make estimate environmental impact. However, in large data sets, this may not always be helpful, especially in unique configurations of components in novel product concepts, such is why a singular component data point is generated.

For the partial bill of materials, Monte Carlo sampling is used similarly, as mentioned before. However, here there is no summation within the data point LCA impact indicators. The partial bill of materials LCA impacts is calculated by the product level non-normalized impact data related to each component or component cluster, as in the representation of the impact of the total product that a specific component or configuration of components is likely to be found. Each defined metadata point is internally averaged to determine the aggregated product's environmental impact of the product design given the limited bill of materials.

3.4 Assumptions and Limitation

Due to the wide variety of components that fall into standard component basis terms, the final output for the normalized data probability density distributions needs to be in terms of impact over grams of product weight. The normalization is done to remove the uncertainty related to using data sets that have a bias toward defining classically ubiquitous components as high or low impact. For example, the environmental impact of a housing component generally scales with the mass of the housing. Bias can be introduced to the PeeP method by defining the PDF for housings using only smaller products with less significant housings. If a user then estimates the impact of a large product with significant housing, the housing impact estimation will be invalid. Unfortunately, the removal of this bias requires some designer cognition to fully utilized the PeeP method with a complete BOM. Meaning that if a designer has a completed BOM of materials but does not know an estimation of the realized product mass, the utility of the PeeP approach can be limited to being only comparative. Though this may be the case, the knowledge gained from using the PeeP method still carries utility. In this regard, the PeeP method can still be used to scale product mass to fit within defined environmental constraints.

The PeeP method does not account for material choice. The perceived component impact is calculated under the assumption that any component with the same weight is subject to the same ratio of impact regardless of material choice. This assumption is faulty as we know that metal-based manufacturing methods are often more impact than polymer manufacturing. As such, the PDFs are subject to data bias depending on the distributions of materials within the source data. If the source data only has one material type represented for a component, and the user intends to make the component from a different material, the pdf can be skewed. With a robust data set, this limitation is minimized with the assumption that the PeeP method takes into account the component's general material distribution. The pdfs created from a robust data set should have a bandwidth that accounts for many material related impacts per component and represent the most probable impact caused by the most probable material the component is manufactured with.

The nature of estimating environmental impact given a partial bill of materials introduces reasonable uncertainty to the PeeP method. The component level PDFs are calculated under the probability that the component is found in products with a specific range of impact. The uncertainty of the probability of impact given a set of known components leads to a significant range of realized product impact. However, depending on the completeness of the partial BOM, a user can use the normalized meta PDFs as a sanity check to realize the representative impact range that will apply to a given design.

4 Method Validation and Knowledge Navigation

This section highlights the demonstration of the PeeP method and the exploration of knowledge that this method can provide. A set of bill of materials are selected from the dataset to represents a variety of consumer products. These consumer products, Stapler, Mechanical Pencil, Apple Peeler Corer, and Game Controller, represent BOMs of differing complexities. Table 3 shows the complete set of bill of materials for each product. Partial bills of material are also tested to demonstrate the estimation of environmental impact given incomplete design information. We create partial BOMs from the previous complete BOMs by removing any component that is not directly related to the product's primary function. For example, also shown in Table 3, an incomplete bill of materials where the designer is sure of the mechanical pencil's internal functional components but undecided on any external components such as housing, grip, or clip.

The results of these examples are leveraged against a known LCA indicator data of the validation products as sourced from the OSDR and SDR. The following sections explore the accuracy of the calculated KDEs for the completed and partial BOMs. Furthermore, the sections will demonstrate how to interpret the PeeP method's results and explore the implications of the method.

4.1 Completed Bill of Materials

Figures 2-5, show the complete bill of materials KDE probability distribution functions for each example product shown in table 3. Overall, we can observe that the generated probability functions are robust enough to capture the tested products' actual environmental impacts. Given this validation test case, the ranges of environmental impact estimations appear to be wide enough to capture products with uncertain embedded impact due to various use phases, material choice, and geometry. A user can explore the probable impact of a given product through the PeeP method and prior knowledge of product BOM.

4.1.1 Mechanical Pencil We calculated the mechanical pencil PDF by using 550 sample BOMs generated during the data processing and sampling stage. The PDF provides environmental impact estimation within a range of $[0 \times 10^{-11}$ - $7 \times 10^{-11}]$ Species.yr per gram of product, Figure 2. The actual impact of the mechanical pencil is 4.85×10^{-11} Species.yr/gram. The multi-modal nature of the probability density function can be attributed to the inclusion of two 'unclassified' components in the complete bill of materials. The 'unclassified' components, in this case, are the pencil lead and eraser. In general, due to the wide variation of the environmental impact of 'unclassified' components, the bill of materials that include such components appear to have wider ranges of impact distribution.

4.1.2 Apple Peeler Corer Shown in Figure 3, the Apple Peeler Corer (APC) KDE, generated with 1250 sample BOMs, shows that the complete BOM should have an impact estimation range of $[0.0 \times 10^{-10}$ - $9.0 \times 10^{-10}]$ Species.yr / gram. The actual impact of the APC is 4.92×10^{-11} Species.yr/gram.

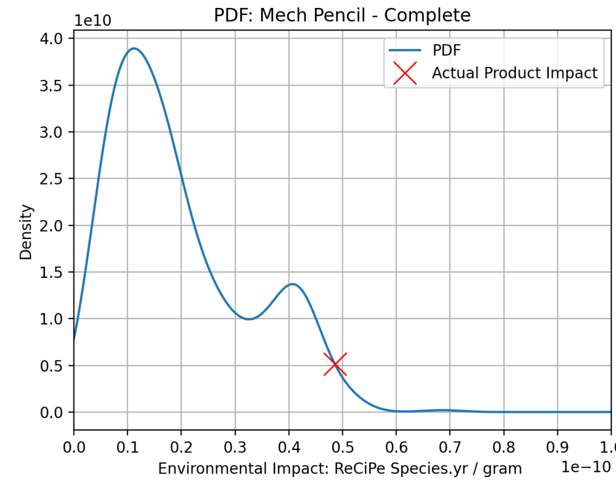


FIGURE 2: Probability density function for the complete Mechanical Pencil bill of materials

While the APC PDF is smoother than the Mechanical Pencil PDF, the range is of impact estimation is substantially larger. Here we can observe the effects of a larger BOM when using the PeeP method. The larger the bill of materials, the less variation is caused by individual components. However, larger BOMs lead to larger ranges of estimated impact.

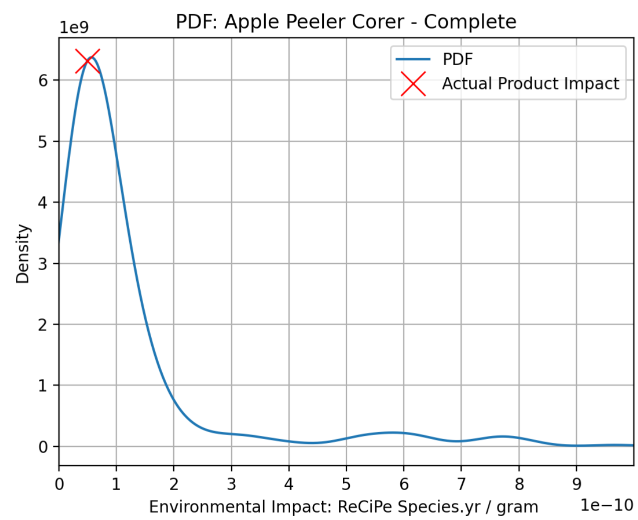


FIGURE 3: Probability density function for the complete Apple Peeler Corer bill of materials

TABLE 3: BILL OF MATERIALS FOR EACH TEST PRODUCT

	Products							
	Mechanical Pencil		Apple Peeler Corer		Stapler		Hand Dryer	
ID	CBOM	PBOM	CBOM	PBOM	CBOM	PBOM	CBOM	PBOM
1	Assembly	Tube	Support	Support	Support	Support	Housing	Housing
2	Tube	Spring	Shaft	Shaft	Support	Guider	Electric Motor	Electric Motor
3	Cap	Cap	Link	Crank	Guider	Container	Fan	Fan
4	Insert	Assembly	Latch Release	Blade	Container	Spring	Housing	Housing
5	Spring		Crank	Nut-bolt	Spring		Seal	Circuit Board
6	Unclassified		Blade	Blade	Cover		Circuit Board	Heating Element
7	Unclassified		Nut-bolt	Support	Cover		Heating Element	
8			Spring	Handle	Link		Unclassified	
9			Screw		Spring		Support	
10			Link				Fastener	
11			Blade					
12			Support					
13			Handle					
14			Screw					
15			Handle					

4.1.3 Stapler Shown in Figure 4, the Stapler KDE, generated with 1400 sample BOMs, shows that the complete BOM should have an impact estimation range of [0.0×10^{-10} - 9.5×10^{-10}] Species.yr / gram. The actual impact of the stapler is 3.98×10^{-11} Species.yr/gram. The PDF for the stapler shows that there is no convergence of low probability till much after the most probable peak between [0.0×10^{-10} - 2.5×10^{-10}] Species.yr / gram. We hypothesize that this is caused by the BOM containing a large ratio of high occurrence, high variation components, such as support and cover. This result is particularly interesting as the PeeP method does not suffer from overfitting the data to the major distribution. In this case, the generated PDF suggests that the impact can safely be assumed to be 2.5×10^{-10} Species.yr/gram or under. However, the results highlight that the user should carry special consideration for products that have features that can lead to a higher impact. In the case of our source data, we anticipate that the trail off is caused by products with high use phases or large geometry.

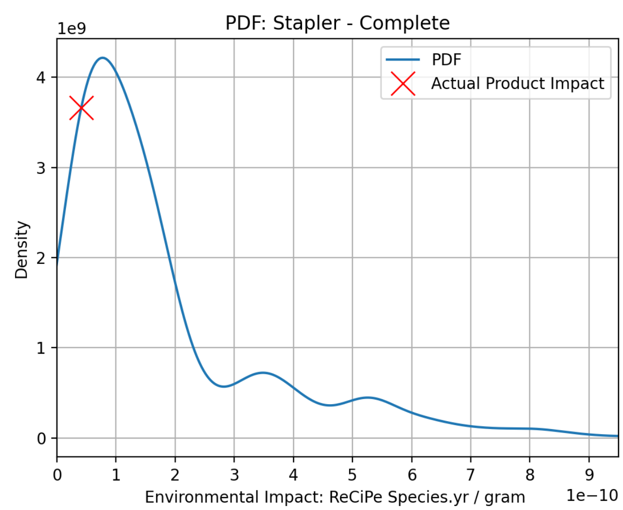


FIGURE 4: Probability density function for the complete Stapler bill of materials

4.1.4 Hand Dryer

Shown in Figure 5, the Hand Dryer KDE, generated with 1450 sample BOMs, shows that the complete BOM should have an impact estimation range of $[0.0 \times 10^{-10}$ - $9.75 \times 10^{-10}]$ Species.yr / gram. The actual impact of the Hand Dryer is 1.19×10^{-10} Species.yr/gram. The effective environmental estimation impact range for the Hand Dryer is $[0.0 \times 10^{-10}$ - $5.5 \times 10^{-10}]$ Species.yr / gram. The hand dryer used in this validation happens to be an Eco-labeled hand dryer. Thus the actual impact of the product is lower than might be expected. However, found in a separate study, the actual impact of a conventional dryer is 3.7×10^{-10} Species.yr / gram [9]. This dryer was not included in the source data for validation. Assuming the same bill of materials is representative of the conventional dryer, the generated Hand Dryer PDF appears to capture that probability of a more environmentally impactful hand dryer. This is an important demonstration of the utility of the PeeP method. That is, the PeeP method is able to ascertain a range that is representative of a wide variety of use cases of the given BOM, regardless of the perceived bias in the data.

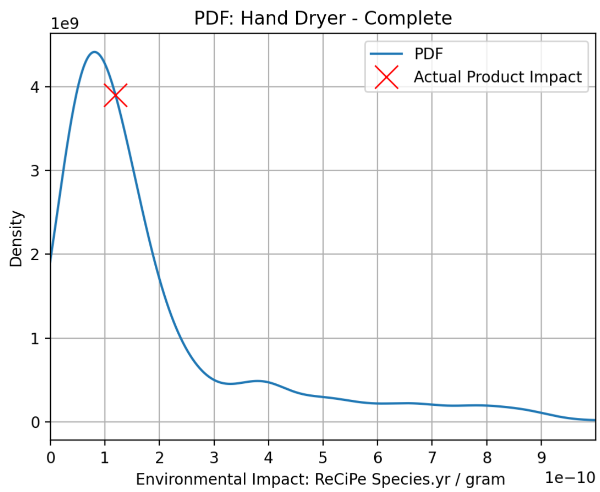


FIGURE 5: Probability density function for the complete Stapler bill of materials

4.2 Partial Bill of Materials

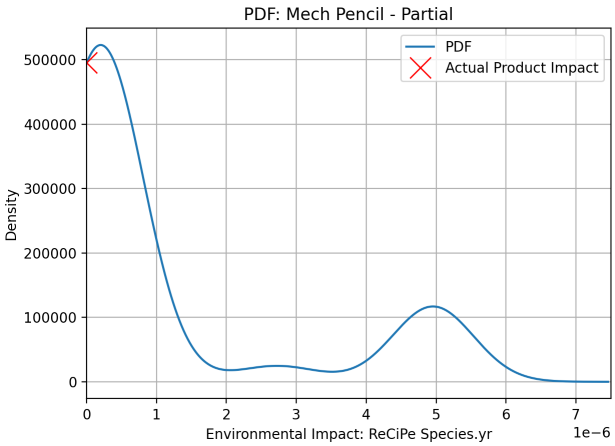
The results of the Partial bill of materials validation are shown in figure 6. These PDFs are useful in the case of concurrent and ongoing design, where a user can determine components that are required during design. These PDFs are useful in scenarios where specific designs must contain certain components either to complete the desired function or as prescribed by the design constraints of the project. The results shown provide a 'best guess' estimation based on similarities between the partial BOM and environmental impact of the source products in the data. As such, the PDFs generated with a partial BOM are much more of a fuzzy look at potential environmental impact than using a complete BOM with the PeeP methodology.

As shown in figure 6, all PDFs captured the true impact of the complete product. However, there are substantial impact estimation ranges due to the large uncertainty of the bill of materials. The true impacts for each product are as follows: Mechanical pencil 3.92×10^{-10} Species.yr, Apple Peeler Corer 3.25×10^{-8} Stapler Species.yr, 9.94×10^{-9} Species.yr, and Hand Dryer 1.05×10^{-6} Species.yr. Besides the hand dryer PDF, the large environmental impact ranges appear to dwarf the actual impacts of the product. For example, the impact of the mechanical pencil's actual impact is 3.92×10^{-10} Species.yr, while the suggested impact range is $[0.0 \times 10^{-6}$ - $7.0 \times 10^{-6}]$ Species.yr. While the true impact does lie in the most probable range, it is nearly impossible for the user to consider environmental impacts lower. We hypothesis that in this case, a user is unlikely to assume an impact more than two orders of magnitude lower than the x-axis range. More than likely, the user will assume an impact range of $[0.0 \times 10^{-6}$ - $1.5 \times 10^{-6}]$ Species.yr, thus likely overestimating the realized environmental impact of the partial BOM product. However, these results are likely suffering from the wide view of the source data.

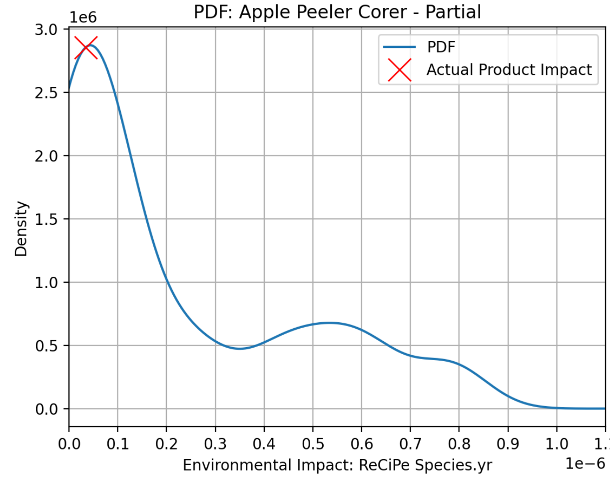
The data used to validate the PeeP method features products that purposely represent a full breadth of consumer products. The products in the source data range from an impact of $[3.92 \times 10^{-10}$ - $9.91 \times 10^{-6}]$ Species.yr with an average product impact of 4.85×10^{-7} Species.yr. All of the generated PDFs in figure 6 capture the average impact of the products within the test data set. These findings suggest that if a more poignant data set were selected, such as product data found in private company repositories, the PeeP method would still help estimate product impact based on partial BOM. Concisely put, the use of partial BOM is subject to the bias of the data set used. In the test case, the average product within the source data has a reasonably impactful use phase. The Mechanical Pencil, Stapler, and Apple Peeler Corer do not feature a significant use phase impact than the data set's bias. The Hand dryer does have a comparable use phase impact to the source data set products.

4.3 Component Level Kernel Density Estimation

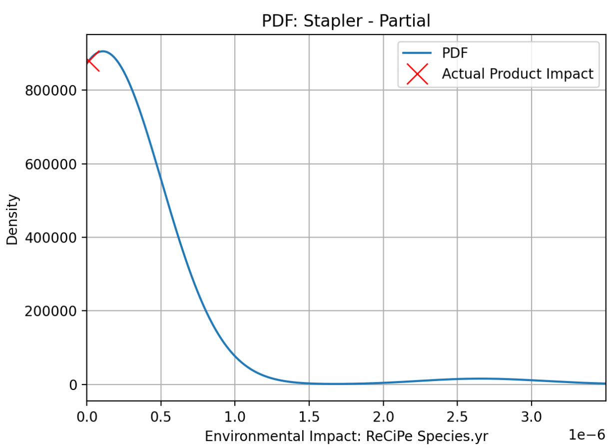
As mentioned previously, the single component KDEs can be used to highlight the estimation of environmental impact contribution caused by individual components. In figure 7, the individual component KDE probability distribution functions for the component 'Cap' and 'Unclassified' is shown. These components are selected to support the results generated from the complete BOM PDF for the mechanical pencil, figure 2. In figure 7, we can observe that the unclassified components contribute to the uncertainly of impact estimation of the mechanical pen-



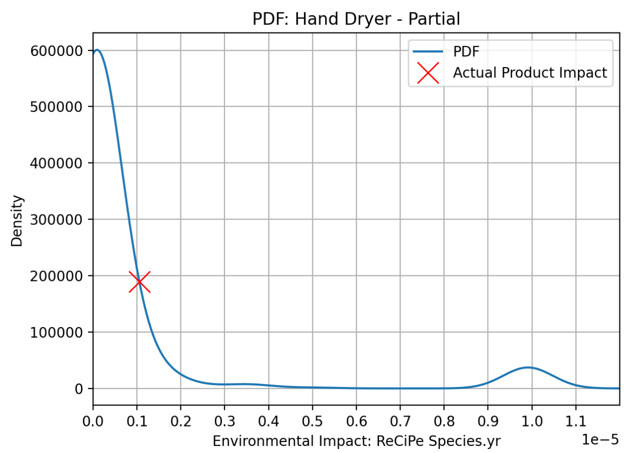
(a) Mechanical Pencil partial BOM probability density function



(b) Apple Peeler Corer partial BOM probability density function



(c) Stapler partial BOM probability density function



(d) Hand Dryer partial BOM probability density function

FIGURE 6: Probability density functions for partial bill of materials of (a) Mechanical Pencil (b) Apple Peeler Corer (c) Stapler (d) Hand Dryer

cil, as shown by the combined component PDF. The environmental impact contribution for the cap ranges from [0.0×10^{-11} - 0.9×10^{-11}] Species.yr./gram; whilst the unclassified component ranges from [0.0×10^{-6} - 4.0×10^{-11}] Species.yr/gram. Using the PeeP method, users can look at contributions caused by individual components or component clusters against a complete BOM probability density function to identify components that contribute to a wide range of impacts.

5 DISCUSSION

The PeeP method is useful by designers who need to estimate the environmental impact of a product design given some information regarding the components used. Particularly, suppose the design phase is nearing completion, and a designer knows the likely complete bill of materials. In that case, the PeeP method can offer robust and precise predictions of the product's potential environmental impact. Using the single level component PDF generated from the PeeP methodology, the user can identify the 'pain' point within the design. Components that relay a large uncertainly to the impact of the product design can highlight and be more carefully considered by the designer. Finally, if the design is iterative-based, a designer can still utilize the partial bill of materials to understand the potential downstream impacts of a weakly realized design. Ultimately, the PeeP method offers insight into product design's environmental sustainability prior to high-cost design activities such as beta-prototyping, user case studies, and product deployment. The PeeP method is a step in making sustainable design proactive as opposed to reactive. However, the PeeP method is subject to a

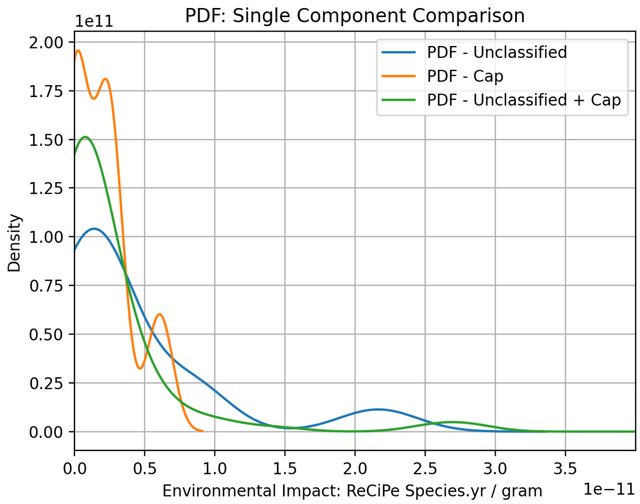


FIGURE 7: Probability density function for Cap, Unclassified and Cap + Unclassified component(s)

number of realized drawbacks.

The PeeP method requires a reasonable expectation of designer cognition to identify where among the probability distributions their products genuinely lie. Factors including, intended product use, exotic material choice, and purposeful sustainable design, require the designer to empirically identify if those fact ores positively or negative skew impact estimations. Designers also need to be aware of how well their design is represented in the supplied repository.

The PeeP methods accuracy is mainly dependent on the repository data and the percent representation of the supplied BOM. In the highlighted method validation, this was not a problem for the completed bill of materials environmental impact estimations. We hypothesis that this was due to a sufficiently diverse representation of consumer products. However, when introducing uncertainly and fuzzy approximations of the partial bill of materials, the PeeP method could not convey an environmental impact estimation that was reasonably representative of the related product's full BOM. In this regard, it is essential that the designer supply historical product data closely related to the product design as possible.

6 CONCLUSION

In this paper, we introduced a novel method for estimating the environmental impact of a given bill of materials. This method helps address growing concerns of product sustainability, and greater climate change concerns felt around the globe. As such, designers can employ the PeeP method to help inform environmentally sustainable component selection during the later stages of product design. In short, this method allows the designer to understand how components contribute to the environmental impact of the design and determine a range of estimated environmental impact of the potential realized product.

The PeeP method is leveraged against user-supplied historical product data. This data needs to include product BOMs, LCA impact metrics per product, product, and component weight. The input data is normalized to estimate the component impact ranges based upon weight. Once supplied with a design's complete or partial bill of materials, the PeeP method employs kernel density estimation to generate probability density functions for the bill of material, each BOM component, and all unique component combinations. These generated PDFs allow visualization of probable potential environmental impact caused by components, the complete bill of materials, or partial bill of materials. As explored through method validation, the PeeP method demonstrates promise in becoming a useful tool for global sustainability-minded designers.

The PeeP method presented in our work aims to address some of the presented issues in current LCA focused product design research. First, the probabilistic approach of the methodology helps account for embedded uncertainty by relying on the user intuition to define the “reasonable range” of impact based on their supplied product and product found in the data set. Furthermore, the PeeP method displays probability density functions that show the uncertainty within each component LCA. Second, the PeeP method uses complete product LCAs. Early product design research that used LCAs are often reliant on component level LCAs. Component level LCAs do not take into account impact as related to the product the component is found within. It is disingenuous to estimate the impact of a product design solely on LCAs based on manufacturing and disposal. Product use does accounts for a significant portion of the impact. Furthermore, using complete LCAs is generally what is more likely found in design or product repositories given the comparative and iterative nature of LCA as found in industry. Third and finally, the PeeP method contributes to the ongoing state-of-the-art in moving questionable design into a proactive approach used during design. However, a number of drawbacks within PeeP methodology can be improved upon and expanded on in future work.

The PeeP method can be improved by implementing adjusters to account for the material type of the components. This should improve provided LCA impact ranges per component. As of now, large data standard deviations can be caused by the implicit nature of material selection within the PeeP method. Furthermore, penalty constraints can be developed to depreciate the effects of component weight in generating the PDFs. These constraints can help address scaling issues where component environmental impact does not linearly scale with weight. Penalty constraints can also account for exotic materials that do not meet the standard assumptions. For example, titanium is less dense than steel. Given the PeeP method's current iteration, titanium will be assumed to be less impactful even though titanium processing can be dramatically more impactful than steel processing.

The PeeP method can be expanded by introducing functional definitions as attached to components. Through the development of novel methodologies, the PeeP method can be leveraged in functional modeling to determine environmental impact ranges of functional chains given all possible component combinations that solve the functional chain. Furthermore, with the development of grammar rules and component compatibility, future methods can provide component combinations that are realistic and environmentally sustainable. The realization of this methodology expansion can help move the PeeP method earlier in the design phase, where designers are designing toward desired functions.

ACKNOWLEDGMENT

This material is based upon work supported by the National Science Foundation under Grant No. CMMI-1826469. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

[1] Kronthal-Sacco, R., and Whelan, T., 2019. Sustainable Share Index: Research on IRI Purchasing Data (2013-2018). Tech. Rep. 9, NYU STERN Center for Sustainable Business.

[2] Kronthal-Sacco, R., and Whelan, T. Sustainable Market Share IndexTM Research on 2015-2020 IRI Purchasing Data Reveals Sustainability Drives Growth, Survives the Pandemic. Tech. rep., NYU STERN Center for Sustainable Business.

[3] Russo, D., Regazzoni, D., and Montecchi, T., 2011. "Eco-design with TRIZ laws of evolution". *Procedia Engineering,* ***9***, pp. 311–322.

[4] DuPont, B., and Wisthoff, A., 2015. "Exploring the Retention of Sustainable Design Principles in Engineering Practice Through Design Education". In Proceedings of the ASME 2015 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference IDETC/CIE 2015, Vol. 8, American Society of Mechanical Engineers, pp. 1–4.

[5] Wisthoff, A., and DuPont, B., 2016. "A Method for Understanding Sustainable Design Trade-Offs During the Early Design Phase". In Sustainable Design and Manufacturing 2016, R. Setchi, R. J. Howlett, Y. Liu, and P. Theobald, eds., Springer International Publishing, pp. 271–280.

[6] Faludi, J., 2015. "A sustainable design method acting as an innovation tool". *Smart Innovation, Systems and Technologies,* ***35***, pp. 201–212.

[7] Faludi, J., and Gilbert, C., 2019. "Best practices for teaching green invention: Interviews on design, engineering, and business education". *Journal of Cleaner Production,* ***234***, pp. 1246–1261.

[8] Gilchrist, B. P., Tumer, I. Y., Stone, R. B., Gao, Q., and Haapala, K. R., 2012. "Comparison of Environmental Impacts of Innovative and Common Products". In 2012 International Design Engineering Technical Conferences &Computers and Information in Engineering Conference, ASME, pp. 1–10.

[9] Ferrero, V., Raman, A. S., Haapala, K. R., and DuPont, B., 2019. "Validating the Sustainability of Eco-Labeled Products Using a Triple-Bottom-Line Analysis". *Smart and Sustainable Manufacturing Systems,* ***3***(1), nov.

[10] Li, J., Tao, F., Cheng, Y., and Zhao, L., 2015. "Big Data in product lifecycle management". *International Journal of Advanced Manufacturing Technology,* ***81***(1-4), pp. 667–684.

[11] Tao, F., Cheng, J., Qi, Q., Zhang, M., Zhang, H., and Sui, F., 2018. "Digital twin-driven product design, manufacturing and service with big data". *International Journal of Advanced Manufacturing Technology,* ***94***(9-12), pp. 3563–3576.

[12] Brynjolfsson, E., and McAfee, A., 2012. "Big Data : The Management Review". *Harvard Business Review*(October), pp. 1–12.

[13] Wang, W. M., Li, Z., Liu, L., Tian, Z. G., and Tsui, E., 2018. "Mining of affective responses and affective intentions of products from unstructured text". *Journal of Engineering Design,* ***29***(7), jul, pp. 404–429.

[14] Zhang, H., Rao, H., and Feng, J., 2018. "Product innovation based on online review data mining: a case study of Huawei phones". *Electronic Commerce Research,* ***18***(1), mar, pp. 3–22.

[15] Tucker, C., and Kim, H., 2011. "Predicting Emerging Product Design Trend By Mining Publicly Available Customer Review Data". *DS 68-6: Proceedings of the 18th International Conference on Engineering Design (ICED 11), Impacting Society through Engineering Design, Vol. 6: Design Information and Knowledge, Lyngby/Copenhagen, Denmark, 15.-19.08.2011*(August), pp. 43–52.

[16] Sangelkar, S., Cowen, N., and McAdams, D., 2012. "User activity-product function association based design rules for universal products". *Design Studies,* ***33***(1), pp. 85–110.

[17] Liao, S. H., Hsieh, C. L., and Huang, S. P., 2008. "Mining product maps for new product development". *Expert Systems with Applications,* ***34***(1), pp. 50–62.

[18] Qi, J., Zhang, Z., Jeon, S., and Zhou, Y., 2016. "Mining customer requirements from online reviews: A product improvement perspective". *Information and Management,* ***53***(8), pp. 951–963.

[19] Alkahtani, M., Choudhary, A., De, A., and Harding, J. A.,

2019. "A decision support system based on ontology and data mining to improve design using warranty data". *Computers and Industrial Engineering,* ***128***(April 2018), pp. 1027–1039.

[20] Kong, D., Li, M., and Zheng, W., 2018. "To Identify Technology Frontier for Mass-Customized Production Service Converged with Artificial Intelligence Based on Patent Data Mining". *2018 15th International Conference on Service Systems and Service Management, ICSSSM 2018*(9184028), pp. 1–6.

[21] Yan, H. S., and Xu, D., 2007. "An approach to estimating product design time based on fuzzy ν-support vector machine". *IEEE Transactions on Neural Networks,* ***18***(3), pp. 721–731.

[22] Tucker, C. S., and Kim, H. M., 2009. "Data-driven decision tree classification for product portfolio design optimization". *Journal of Computing and Information Science in Engineering,* ***9***(4), pp. 1–14.

[23] Bohm, M. R., and Stone, R. B., 2004. "Product design support: Exploring a design repository system". *American Society of Mechanical Engineers, Computers and Information in Engineering Division, CED*, pp. 55–65.

[24] Szykman, S., Sriram, R. D., Bochenek, C., Racz, J. W., and Senfaute, J., 2000. "Design Repositories: Engineering Design's New Knowledge Base". *IEEE Intelligent Systems and Their Applications,* ***15***(3), pp. 48–55.

[25] Bohm, M. R., Stone, R. B., Simpson, T. W., and Steva, E. D., 2008. "Introduction of a data schema to support a design repository". *CAD Computer Aided Design,* ***40***(7), pp. 801–811.

[26] Phelan, K., Wilson, C., and Summers, J. D., 2014. "Development of a design for manufacturing rules database for use in instruction of dfm practices". In Proceedings of the ASME Design Engineering Technical Conference, Vol. 1A, pp. 1–7.

[27] Bharadwaj, A., Xu, Y., Angrish, A., Chen, Y., and Starly, B., 2019. "Development of a pilot manufacturing cyberinfrastructure with an information rich mechanical cad 3D model repository". *ASME 2019 14th International Manufacturing Science and Engineering Conference, MSEC 2019,* ***1***, pp. 1–8.

[28] Rossi, M., Marconi, M., Menghi, R., and Papetti, A., 2020. "A Knowledge Repository to Support Ecodesign Implementation in Manufacturing Companies". In Design Tools and Methods in Industrial Engineering, C. Rizzi, A. O. Andrisano, F. Leali, F. Gherardini, F. Pini, and A. Vergnano, eds., Springer International Publishing, pp. 651–661.

[29] Szykman, S., Sriram, R. D., Bochenek, C., and Racz, J. W., 1999. "The NIST Design Repository Project". *Advances in Soft Computing - Engineering Design and Manufacturing*, pp. 5–19.

[30] Ferrero, V., Wisthoff, A., Huynh, T., Ross, D., and DuPont, B., 2018. "A Sustainable Design Repository for Influencing the Eco-Design of New Consumer Products". *EngrXIV*, p. Under Review.

[31] Oman, S., Gilchrist, B., Tumer, I. Y., and Stone, R., 2014. "The development of a repository of innovative products (RIP) for inspiration in engineering design". *International Journal of Design Creativity and Innovation,* ***2***(4), oct, pp. 186–202.

[32] Rajagopalan, V., Bryant, C. R., Johnson, J., McAdams, D. A., Stone, R. B., Kurtoglu, T., and Campbell, M. I., 2005. "Creation of Assembly Models to Support Automated Concept Generation". In Volume 5a: 17th International Conference on Design Theory and Methodology, Vol. 2005, ASME, pp. 259–266.

[33] Bryant, C. R., McAdams, D. A., Stone, R. B., Kurtoglu, T., and Campbell, M. I., 2005. "A computational technique for concept generation". *Proceedings of the ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference - DETC2005,* ***5***, pp. 267–276.

[34] Bryant, C. R., McAdams, D. A., Stone, R. B., Kurtoglu, T., and Campbell, M. I., 2006. "A validation study of an automated concept generator design tool". *Proceedings of the ASME Design Engineering Technical Conference,* ***2006***, pp. 1–12.

[35] Soria Zurita, N. F., Stone, R. B., Demirel, O., and Tumer, I. Y., 2018. "The Function-Human Error Design Method (FHEDM)". In Volume 7: 30th International Conference on Design Theory and Methodology, American Society of Mechanical Engineers.

[36] Soria Zurita, N. F., Stone, R. B., Onan Demirel, H., and Tumer, I. Y., 2020. "Identification of Human–System Interaction Errors During Early Design Stages Using a Functional Basis Framework". *ASCE-ASME J Risk and Uncert in Engrg Sys Part B Mech Engrg,* ***6***(1), mar.

[37] Mikes, A., Edmonds, K., DuPont, B., and Stone, R., 2020. "Optimizing an algorithm for data mining a design repository to automate functional modeling". No. IDETC2020-22346, American Society of Mechanical Engineers.

[38] Tensa, M., Edmonds, K., Ferrero, V., Mikes, A., Soria Zurita, N., Stone, R., and DuPont, B., 2019. "Toward Automated Functional Modeling: An Association Rules Approach for Mining the Relationship between Product Components and Function". *Proceedings of the Design Society: International Conference on Engineering Design,* ***1***(1), pp. 1713–1722.

[39] Fayyad, U., Piatetsky-Shapiro, G., and Smyth, P., 1996. "The KDD Process for Extracting Useful Knowledge from Volumes of Data". *COMMUNICATIONS OF THE ACM,* ***39***(11), pp. 27–34.

[40] Fayyad, U., Piatetsky-Shapiro, G., and Smyth, P., 1996. "Knowledge Discovery and Data Mining: Towards a Unify-

ing Framework". In KDD-96 Proceedings, Vol. 14, pp. 82–88.
[41] Fayyad, U., Piatetsky-Shapiro, G., and Smyth, P., 1996. "From Data Mining to Knowledge Discovery in Databases". *AI Magazine,* ***17***(2), pp. 37–54.
[42] Williams, G., Meisel, N. A., Simpson, T. W., and McComb, C., 2019. "Design repository effectiveness for 3D convolutional neural networks: Application to additive manufacturing". *Journal of Mechanical Design, Transactions of the ASME,* ***141***(11), pp. 1–12.
[43] Stone, R. B., and Wood, K. L., 2000. "Development of a Functional Basis for Design". *Asme,* ***122***(December 2000).
[44] Hirtz, J., Stone, R. B., McAdams, D. A., Szykman, S., and Wood, K. L., 2002. "A functional basis for engineering design: Reconciling and evolving previous efforts". *Research in Engineering Design - Theory, Applications, and Concurrent Engineering,* ***13***(2), pp. 65–82.
[45] Cheong, H., Chiu, I., Shu, L. H., Stone, R. B., and McAdams, D. A., 2011. "Biologically meaningful keywords for functional terms of the functional basis". *Journal of Mechanical Design, Transactions of the ASME,* ***133***(2), pp. 1–11.
[46] Kurtoglu, T., Campbell, M. I., Bryant, C. R., Stone, R. B., and Mcadams, D. A., 2005. "Deriving a Component Basis for Computational Functional Synthesis". In ICED 05: 15th International Conference on Engineering Design: Engineering Design and the Global Economy.
[47] ISO, E., 2006. "ISO 14044: 2006". *Environmental management-Life cycle assessment-Requirements and guidelines. European Committee for Standardization*.
[48] Guinée, J. B., Heijungs, R., Huppes, G., Zamagni, A., Masoni, P., Buonamici, R., Ekvall, T., and Rydberg, T., 2011. "Life cycle assessment: Past, present, and future". *Environmental Science and Technology,* ***45***(1), pp. 90–96.
[49] Jørgensen, A., 2013. "Social LCA - A way ahead?". *International Journal of Life Cycle Assessment,* ***18***(2), pp. 296–299.
[50] Stanford University, 2005. GUIDELINES For LCCA in Buildings. Tech. Rep. October.
[51] Huijbregts, M., Steinmann, Z., P.M.F. Elshout, R., Stam, G., Verones, F., Vieira, M., Hollander, A., Zijp, M., and van Zelm, R., 2016. *ReCiPe 2016: A harmonized life cycle impact assessment method at midpoint and endpoint level*. National Institute for Public Health and the Environment.
[52] Bare, J. C., Norris, G. a., and Pennington, D. W., 2003. "The Tool for the Reduction and Assessment Impacts". *Journal of Industrial Ecology,* ***6***(3), pp. 49–78.
[53] Guinée, J., Gorrée, M., Heijungs, R., Huppes, G., Kleijn, R., de Koning, A., van Oers, L., Wegener Sleeswijk, A., Suh, S., Udo de Haes, H., de Bruijn, H., van Duin, R., and Huijbregts, M., 2002. *CML 2001*. Dordrecht.
[54] Goedkoop, M., De Schryver, A., Oele, M., Durksz, S., and de Roest, D., 2008. "Introduction to LCA with SimaPro 7". *PRe Consultants, The Netherlands*.
[55] Sphera. GaBi Software.
[56] Igos, E., Benetto, E., Meyer, R., Baustert, P., and Othoniel, B., 2019. "How to treat uncertainties in life cycle assessment studies?". *International Journal of Life Cycle Assessment,* ***24***(4), pp. 794–807.
[57] Brundage, M. P., Bernstein, W. Z., Hoffenson, S., Chang, Q., Nishi, H., Kliks, T., and Morris, K. C., 2018. "Analyzing environmental sustainability methods for use earlier in the product lifecycle". *Journal of Cleaner Production,* ***187***, pp. 877–892.
[58] Tensa, M., Ross, D., Ferrero, V., and DuPont, B., 2018. "Assessing the impact of product use variation on environmental sustainability". *Proceedings of the ASME Design Engineering Technical Conference,* ***4***.
[59] Rezaei, F., Bulle, C., and Lesage, P., 2019. "Integrating building information modeling and life cycle assessment in the early and detailed building design stages". *Building and Environment,* ***153***(September 2018), pp. 158–167.
[60] Bartolozzi, I., Daddi, T., Punta, C., Fiorati, A., and Iraldo, F., 2020. "Life cycle assessment of emerging environmental technologies in the early stage of development: A case study on nanostructured materials". *Journal of Industrial Ecology,* ***24***(1), pp. 101–115.
[61] Göswein, V., Rodrigues, C., Silvestre, J. D., Freire, F., Habert, G., and König, J., 2020. "Using anticipatory life cycle assessment to enable future sustainable construction". *Journal of Industrial Ecology,* ***24***(1), pp. 178–192.
[62] Bohm, M. R., Haapala, K. R., Poppa, K., Stone, R. B., and Tumer, I. Y., 2010. "Integrating life cycle assessment into the conceptual phase of design using a design repository". *Journal of Mechanical Design, Transactions of the ASME,* ***132***(9).
[63] Arlitt, R., Van Bossuyt, D. L., Stone, R. B., and Tumer, I. Y., 2017. "The function-based design for sustainability method". *Journal of Mechanical Design, Transactions of the ASME,* ***139***(4), pp. 1–12.
[64] Marwah, M., Shah, A., Bash, C., Patel, C., and Ramakrishnan, N., 2011. "Using Data Mining to Help Design Sustainable Products". *Computer,* ***44***(8), pp. 103–106.
[65] Parzen Emanuel, 1962. "On Estimation of a Probability Density Function and Mode". *The Annals of Mathematical Statistics,* ***33***(3), pp. 1065–1076.
[66] Sheather, S. J., and Jones, M. C., 2008. "A Reliable Data-Based Bandwidth Selection Method for Kernel Density Estimation A". *Journal of the Royal Statistical Society. Series B (Methodological),* ***53***(3), pp. 683–690.
[67] Turlach, B. A. "Bandwidth selection in kernel density estimation: A review". In CORE and Institut de Statistique.
[68] Kristan, M., Leonardis, A., and Skočaj, D., 2011. "Multivariate online kernel density estimation with Gaussian ker-

nels". *Pattern Recognition,* ***44***(10-11), pp. 2630–2642.
[69] Chen, Y. C., 2017. "A tutorial on kernel density estimation and recent advances". *Biostatistics and Epidemiology,* ***1***(1), pp. 161–187.
[70] Gramacki, A., 2018. *Nonparametric kernel density estimation and its computational aspects*. Springer International Publishing.
[71] Hu, Y., Wang, F., Guin, C., and Zhu, H., 2018. "A spatio-temporal kernel density estimation framework for predictive crime hotspot mapping and evaluation". *Applied Geography,* ***99***(January), pp. 89–97.
[72] Wang, X., Gerber, M. S., and Brown, D. E., 2012. "Automatic crime prediction using events extracted from twitter posts". *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics),* ***7227 LNCS***, pp. 231–238.
[73] Gerber, M. S., 2014. "Predicting crime using Twitter and kernel density estimation". *Decision Support Systems,* ***61***(1), pp. 115–125.
[74] Flaxman, S., Chirico, M., Pereira, P. A., and Loeffler, C., 2019. "Scalable high-resolution forecasting of sparse spatiotemporal events with kernel methods: A winning solution to the NIJ "real-time crime forecasting challenge"". *Annals of Applied Statistics,* ***13***(4), pp. 2564–2585.
[75] Arora, S., and Taylor, J. W., 2016. "Forecasting electricity smart meter data using conditional kernel density estimation". *Omega (United Kingdom),* ***59***, pp. 47–59.
[76] Vestner, M., Litman, R., Rodolà, E., Bronstein, A., and Cremers, D., 2017. "Product manifold filter: Non-rigid shape correspondence via kernel density estimation in the product space". *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017,* ***2017-January***, pp. 6681–6690.
[77] Jang, J., Cho, S. G., Lee, S. J., Kim, K. S., Kim, J. M., Hong, J. P., and Lee, T. H., 2015. "Reliability-based robust design optimization with kernel density estimation for electric power steering motor considering manufacturing uncertainties". *IEEE Transactions on Magnetics,* ***51***(3), pp. 15–18.
[78] Haselsteiner, A. F., Ohlendorf, J.-H., and Thoben, K.-D., 2017. "Environmental contours based on kernel density estimation". *arXiv preprint arXiv:1710.07156*(October), pp. 1–4.
[79] Goka, E., Beaurepaire, P., Homri, L., and Dantan, J. Y., 2019. "Probabilistic-based approach using Kernel Density Estimation for gap modeling in a statistical tolerance analysis". *Mechanism and Machine Theory,* ***139***, pp. 294–309.
[80] Yen-Chi, C., 2018. "Lecture 6: Density Estimation: Histogram and Kernel Density Estimator". pp. 1–11.
[81] Scott, D. W., 2015. *Multivariate density estimation: theory, practice, and visualization*. John Wiley \& Sons.