Sarah Gül, Michael Spielmann, Annekatrin Lehmann, Diana Eggers, Vanessa Bach, Matthias Finkbeiner

Benchmarking and environmental performance classes in life cycle assessment—development of a procedure for non-leather shoes in the context of the Product Environmental Footprint

Journal article | **Accepted manuscript (Postprint)**
This version is available at https://doi.org/10.14279/depositonce-9127

Benchmarking and environmental performance classes in life cycle assessment—development of a procedure for non-leather shoes in the context of the Product Environmental Footprint

Sarah Gül[1*], Michael Spielmann[2], Annekatrin Lehmann[4], Diana Eggers[3], Vanessa Bach[4], Matthias Finkbeiner[4]

Abstract
Purpose In the process of developing Product Environmental Footprint Category Rules (PEFCR)—currently tested in various pilots in the Single Market for Green Products initiative of the European Commission—the definition of product category benchmarks and environmental performance classes is a crucial element of each PEFCR. Whilst life cycle assessment (LCA) methodology developed over the last 20 years can be used for many other topics to be tackled in the pilots, there is a clear lack of methodology for the determination of benchmarks and environmental performance classes. In this article, hence, we address this gap and develop a procedure for benchmarking and environmental performance classes in LCA.
Methods To do this, given requirements and definitions of the PEF guidelines on both subjects are taken as a basis and are refined by using common LCA techniques like hot spot and sensitivity analyses. The specific steps of the procedure are applied systematically in a case study using sports shoes as an example.
Results and discussion The resulting procedure involves the definition of a scenario vector, which is composed of relevant life cycle phases as well as the lifetime of the product (i.e. sports shoes) as variables. On the basis of the hot spot and sensitivity analyses, these variables are quantified, first, to generate the benchmark and, second, to determine the environmental performance classes around the benchmark for each considered impact category individually. In addition, the influence of data uncertainty on the class distribution is assessed with the help of the Monte Carlo simulation.
Conclusions The results of the application in the case study demonstrate the high impact of the product's lifetime on the final environmental performance classes, and the importance of data quality. Limitations are identified regarding data availability and the harmonisation of the classes to potentially create a PEF label. A debate is induced on the validity of such a label when considering the fact that the characterisation methods and factors proposed in the PEF guidelines may not be complete or accurate enough.

Keywords Benchmarking · Consumer goods · Data uncertainty · Environmental performance classes · Hot spot analysis · LCA · Life cycle assessment · PEF · PEFCR · Product Environmental Footprint · Product Environmental Footprint Category Rules · Sensitivity analysis · Sports shoes

* Sarah Gül
sarahguel@yahoo.de

[1] Ecofys Germany, Berlin 10117, Germany
[2] QUANTIS Germany, Berlin 10122, Germany
[3] thinkstep (former PE INTERNATIONAL), Leinfelden-Echterdingen 70771, Germany
[4] Department of Environmental Technology, Technical University Berlin, Berlin 10623, Germany

1 Introduction

The Product Environmental Footprint (PEF) Guide (European Commission 2013) was published in April 2013 by the European Commission as part of the initiative to create a Single Market for Green Products (European Commission 2014a). It aims at establishing a life cycle-based method harmonising existing life cycle assessment (LCA) methods

EU-wide. The PEF Guide requires the development and testing of PEF Category Rules (PEFCR), which complement the PEF Guide in further specifying requirements and, where appropriate, intend to enable comparisons and comparative assertions of PEF studies (European Commission 2013; European Commission 2014b). Besides the PEF Guide, the PEF Pilot Guidance (European Commission 2014b) is the second major PEF document prepared and regularly updated by the European Commission in order to guide a 3-year pilot phase for selected product categories. The pilot's first wave started in November of 2013 and was followed by a second wave in June 2014, which included food, feed and beverages.

To succeed in the PEF's aim of harmonising existing standards, several challenges are currently faced and addressed in the pilot phase. This concerns choices of internationally agreed and tested impact categories, their indicators and assessment methods (Finkbeiner 2014), and specific definitions of PEF elements such as the representative product. Moreover, it is yet unclear how relevant impact categories should be determined or how to deal with existing ISO-based environmental labelling schemes, such as type III environmental product declarations (EPDs) or type I schemes (e.g. Blue Angel, EU Ecolabel, EU Energy label) (Lehmann et al. 2015). These issues are, however, not further addressed within the scope of this paper.

This paper focuses on one of the remaining open questions, which is the identification of one or more benchmark(s) for the selected pilot product categories and the development of five environmental performance classes A to E. The introduction of both, the benchmark(s) and the environmental performance classes, shall support the communication of PEF results and is an essential element for the definition of a possible PEF label (Lehmann et al. 2014).

There is a clear research need for the identification of benchmark(s) and determination of environmental performance classes as there is no concrete procedure provided in the PEF Guide and PEF Pilot Guidance. In this paper, we develop a procedure for benchmarking and environmental performance classes by taking into account state-of-the-art LCA methodology and applying it to the example of non-leather shoes, more specifically, sports shoes. The procedure is developed under the presumption that all essential PEF elements are given when applying the procedure, such as the definition of the representative product or the choice of applicable impact categories for a considered product category. The challenges and potential shortcomings related to these elements are, however, not discussed in this paper. In the process of developing the procedure, those factors that have the highest impact on the environmental performance classes for sports shoes and are thus relevant for the determination of the benchmark are identified. Finally, the results are assessed for their sensitivity towards data uncertainty, which represents one of the main constraints in LCA analyses and is also acknowledged in the PEF Guide (European Commission 2013).

2 State of the art of defining benchmarks and environmental performance classes

A review of existing literature research reveals very little on how to define benchmarks and environmental performance classes. With regard to benchmarking, Nissinen et al. (2007) developed the "eco-benchmark", which is composed of a set of five common and familiar products typically used by the average Finnish household: rye bread and cheese for food, a car drive, laundry wash and an apartment. This approach does not create a benchmark that is representative for an entire product category; instead, it determines the benchmarks based on selected accuracy and quality criteria (Nissinen et al. 2007). Regarding the definition of environmental performance classes, i.e. the classification of products, the most widely applied approach is the EU Energy label. This label, however, is limited in its scope, as it focuses exclusively on the energy efficiency of household appliances during the use phase (European Commission 2010a), i.e. it is primarily directed at energy-intensive products such as refrigerators and washing machines. The process of determining the various classes of the EU Energy label is not further explained in available literature, and it, hence, does not provide useful input for the development of environmental performance classes as intended by the PEF Pilot Guidance.

3 Methodology

The methodology used in this paper to develop a procedure for benchmarking and environmental performance classes is composed of the development of the concept (Section 3.1) and its application in a case study of sports shoes (Section 3.2), which supported the development of the procedure.

3.1 Concept

To address the mentioned research gap, a stepwise procedure to quantify a benchmark and to determine environmental performance classes is developed starting from the requirements and suggestions given in the PEF Guide and PEF Pilot Guidance on this matter. In the PEF Pilot Guidance, the benchmark is defined as the environmental profile score of the representative product of a product category (European Commission 2014b). The representative product can be a real or virtual product depending on whether full technical and market information is available and accessible or not. The environmental profile score can mainly be defined as the results of a life cycle impact assessment (LCIA) or as a set of life cycle inventory (LCI) data. Since the PEF Pilot Guidance requires the consideration of 14 impact categories (status July 2014), the benchmark can theoretically comprise up to 14 individual scores. According to the PEF Pilot Guidance, a minimum of three individual scores will exist as at least three relevant impact categories shall be included, yielding three individual scores (European Commission 2014b).

The PEF Pilot Guidance requires the identification of five environmental performance classes A to E, with class A representing the best environmentally performing class and with the centre class C being defined by the benchmark. It further defines that the environmental performance classes are to be determined by estimating the spreads around the benchmark (European Commission 2014b). This implies that the benchmark needs to be quantified first. Then, based on the defined benchmark, the environmental performance classes can be developed. This paper proposes a comprehensive procedure for defining benchmarks and environmental performance classes, which can be useful for the PEF pilot phase. In order to develop this procedure, the suggestions given in the PEF Pilot Guide are taken into account and are further refined by using state-of-the-art LCA methodology, i.e. hot spot analysis to identify the most relevant life cycle stages and processes and sensitivity analysis to create the ranges of the classes.

3.2 Case study

The specific steps are developed whilst applying them incrementally to the case of sports shoes. This example is taken from the ongoing PEF pilot "non-leather shoes" representing the representative product of the sub-category "sport shoes" (status December 2014) (SAC 2014). The unit of analysis in the case study of this paper is set to "one pair of sports shoes" and the reference flow to "wear in good condition for one year". These two elements are defined based on discussions within the PEF pilot non-leather shoes, which derived the unit of analysis and reference flow from existing Product Category Rules and the EU Ecolabel, a type I environmental label, on footwear. A cradle-to-grave approach is applied. To create the life cycle model of the sports shoes, data on all life cycle stages is sourced from a study by Cheah et al. (2013) of a typical and existent pair of sports shoes. As the publically available market and technical information for sports footwear is sparse, the life cycle model was refined with information on (i) the energy source by using current trade statistics for the European sports footwear market (CBI 2010) and (ii) the average energy and water consumption during footwear production from the EU Ecolabel (Kowalska et al. 2013). The resulting pair of shoes is chosen as the representative product for the case study of this paper and can be backed up by existing market information on annual top sales of products with similar characteristics to the pair of sports shoes of Cheah et al. (2013). In this case therefore, a real product is used as the representative product, while a virtual product would have involved an "averaging" of the sports shoes on the market. We would like to emphasise that sports shoes is only an example for a potential representative product in order to illustrate how benchmarks and scenarios could be defined. The discussion of whether sports shoes are indeed representative of the product category non-leather shoes is out of the scope of this paper.

4 Results and discussion

On the basis of the given definitions and implications for the benchmark and environmental performance classes in the PEF Pilot Guidance, as described in the previous section, the benchmark is defined as the mean of class C in the proposed procedure. Accordingly, an overall "best case" of the product category in question is derived from the profile of the benchmark (i.e. the representative product) representing the mean of class A and an overall "worst case" representing the mean of the lowest class E (Fig. 1). For the proposed procedure, these three means are treated as scenarios.

As mentioned above, the benchmark (i.e. the mean of class C), which is derived from the representative product, is the starting point, and the other two scenarios (i.e. the mean of classes A and E) are developed from here. Hence, a conceptual basis is defined for the proposed procedure. It accounts for the fact that these three main scenarios are embedded in a common life cycle model—the life cycle model of the representative product—which is then implemented in LCA software (e.g. GaBi Software, PE INTERNATIONAL 2014) to quantify the environmental performance for selected impact categories. The scenarios, however, differ in certain aspects of this life cycle model (e.g., mass of components, use of adhesives, different production technologies, use of recycled materials), which characterise them for being the benchmark and best and worst case scenarios of the analysed product category. These aspects need to be adaptable to the three scenarios and are thus considered "variable". In the case of sports shoes, the variables are defined as the four life cycle phases most relevant to the environmental performance, i.e. manufacturing, transportation, use, and end of life, as well as the lifetime of the product. These five variables are aggregated and presented in the form of a vector $\overrightarrow{S_i}$. This vector, therefore, represents the respective scenario S_i with i being the benchmark, best or worst case. The variable manufacturing is further subdivided into energy and materials, which themselves are broken down into consumption and source, and type and weight, respectively. In our example of sports shoes, the use phase is restricted to maintenance. However, when the developed procedure is applied to other product categories, like energy-intensive

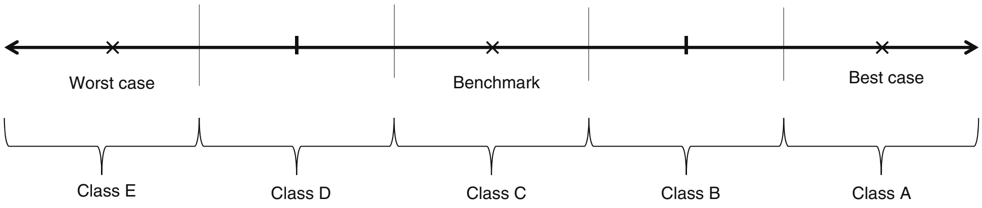


Fig. 1 Role of the benchmark and best and worst case scenarios in the developed procedure as the respective means of classes *C*, *A* and *E*

products, the use phase variable needs to be further specified. Hence, depending on the product category in question, the variables of the vector might be chosen differently.

In the following, the vector—the conceptual basis of the developed procedure—is illustrated for the example of sports shoes (Eq. 1).

$$\vec{S_i} = \begin{cases} \text{manufacturing (energy (consumption; source), materials (type; weight))} \\ \text{transportation} \\ \text{use phase (maintenance)} \\ \text{end of life} \\ \text{lifetime} \end{cases} \tag{1}$$

The vector serves as the framework for all the main steps of the proposed procedure, which are composed of the quantification of the benchmark for a certain product category (Section 4.1) and the development of the environmental performance classes around the benchmark (Section 4.2). The proposed procedure includes an additional last step (Section 4.3), in which the effect of data uncertainty on the accuracy of the environmental performance classes is analysed. These three main steps are explained in the following in more detail.

4.1 Preliminary choice of impact categories and quantification of benchmark(s)

In order to quantify the benchmark, the life cycle model of the representative product, as explained above, needs to be determined. The determination follows the requirements defined in the PEF Guide (European Commission 2013) and comprises the definition of the unit of analysis, reference flow, and system boundaries. It further includes the specification of all life cycle phases of the representative product, which covers the life cycle phases included in the scenario vector. Also, any further assumptions that are needed to complete the life cycle model are made in order to be able to carry out an LCIA.

The benchmark is finally quantified by carrying out an LCIA in LCA software using the life cycle model of the representative product. According to the PEF Guide, all 14 recommended impact categories shall be applied in the first phase of developing the PEFCRs ("screening study"), and benchmarks and performance classes shall be determined for the relevant categories. To define which impact categories are relevant, the PEF guidelines recommended initially using normalisation. As first results from the screening studies (status February 2015) showed that this is not feasible, each PEF pilot can now decide individually which impact categories are relevant for the respective product category and which are not.

However, at the time of developing this procedure for determining benchmarks and environmental performance classes, the screening study for non-leather shoes had not yet started. Consequently, relevant impact categories were not yet defined. Thus, for the case study of sports shoes, we selected five impact categories (climate change, terrestrial eutrophication, resource depletion, particulate matter, aquatic eutrophication). For this selection, we considered both the evaluation of the PEF-recommended LCIA methods provided in the ILCD Handbook (European Commission—Joint Research Centre 2011) and current LCA practice: first, we chose only those impact categories with LCIA methods assigned levels I and II, i.e. LCIA methods which are "satisfactory" or "need some improvements". Categories with LCIA methods rated level 3 (e.g. water depletion) were not chosen to be included in the case study on sports shoes. Second, we chose impact categories that are commonly considered in current LCA practice. Hence, this selection results in a benchmark composed of five scores as displayed in Table 1.

4.2 Determination of environmental performance classes

To develop the environmental performance classes around the benchmark, several sub-steps have to be undertaken with the proposed procedure starting with the hot spot analysis, followed by the determination of best and worst case scenarios, and lastly with respect to the sensitivity analysis:

Table 1 Benchmark for the example of one pair of sports shoes quantified through life cycle impact assessment for selected EF default impact categories (European Commission 2013)

Environmental profile of benchmark for sports shoes (LCIA results)		
Climate change (IPCC 2007)	kg CO_2-eq.	9.86
Terrestrial eutrophication (Seppälä et al. 2006)	mol N-eq.	0.108
Resource depletion (CML 2002)	kg Sb-eq.	2.9×10^{-5}
Particulate matter (RiskPoll 2009)	kg $PM_{2.5}$-eq.	4.46×10^{-3}
Aquatic eutrophication (Struijs et al. 2009)	kg P-eq.	3.69×10^{-4}

First, a hot spot analysis is carried out on the LCIA results of the life cycle model of the representative product. Hot spots refer to all life cycle phases in general, but also to specific input and output flows contributing a considerable percentage to the overall impact category result. The term "considerable" is not precisely defined in literature and depends, for example, on the complexity of the product and the resulting range of hot spots. Based on common practice and experiences of the authors of this study, the threshold for the contribution proposed in the context of this study ranges from 5 to 10 %. These values are chosen independently of the PEF guidelines, which recently defined hot spots as "either (i) life cycle stages, processes and elementary flows cumulatively contributing at least 50% to any impact category, or (ii) at least the two most relevant impact categories, life cycle stages, processes and at least two elementary flows (minimum 6)" (James, Galatola 2015). The threshold value chosen for this study could be adapted according to this recent PEF proposal as the procedure itself works the same way for any threshold. The hot spot analysis for the life cycle model is carried out individually for each impact category as the hot spots are expected to differ among them. It consists of two steps:

1. Using the LCIA results, hot spot life cycle phases of the scenario vector for each impact category are identified. This is depicted in Fig. 2, which illustrates the results of the case study for sports shoes: for climate change, all life cycle phases but the use phase represent hot spots when assuming a 5 % threshold.
2. Since the "manufacturing" variable is composed of the "energy" and "materials" sub-variables, each is analysed separately to identify their respective hot spots and, thus, determine their specific contributions to the overall results. As shown in Fig. 2 for the example of sports shoes, electricity for the sub-variable energy and polyurethane and norbornen copolymer and nylon for the sub-variable materials were the specific hot spots in the case of climate change.

For the "lifetime" variable of the product, the above steps cannot be applied as the lifetime is not a life cycle phase as such, which can be analysed using the LCIA results directly. It is suggested here that the lifetime is only relevant for the environmental impact in cases where the use phase is of low environmental importance: If the lifetime dominates the environmental impact of the product (e.g. due to more durable material input), the use phase usually is of little relevance and vice versa. This implies the assumption that whenever a product is characterised by high durability and, hence, potentially higher environmental impact in the manufacturing phase, the use phase results to be of little significance. This, however, is valid only for products that are not energy-intensive. To define a threshold for the procedure, it is assumed that the product's lifetime represents a hot spot if the use phase is not dominant, i.e. contributing less than 50 % to the overall result of the respective impact category. This threshold is chosen as an illustrative example based on expert opinions and could, therefore, be adapted as soon as further insight to this topic is available. In the case study, "lifetime" is found to be a hot spot for the impact category climate change (Fig. 2).

Second, the best and worst case scenarios in terms of industry practices have to be determined as they lead to the quantification of the means of classes A and E. For the procedure, the approaches to specify both extreme scenarios are proposed as follows:

Best case: The best case reflects the goal of "good" industry practices, which are then reflected by class A in a potential label. The variables for the best case scenarios can obviously be defined based on *existing* best practices, which are typically aggregated in the industry-specific Best Available Technique Reference (BREF) documents[1] as suggested by the PEF Pilot Guidance. They could also take into account existing eco-innovations of the market if available. Following this strategy, the possibility of directing the respective sector in a certain way with minimised environmental impact is not fully exploited because the best case scenario would only reflect the state of the art. Therefore, it is important to leave room for future improvement of the environmental performance of products covered by the product category. This can be done by including ideal conditions: "Best case" would then rather mean "ideal best case".

Worst case: Contrary to the *ideal* best case, the intention here is not to define the worst case possible, but to actually reflect the state of the art in terms of bad practices. Examples are the consideration of maximum energy consumption for production or the maximum product weight (i.e. material input) observed in the sector, but also the lowest share of recycling as well as the lowest known lifetime of the product in question. However, if no specific data or literature

[1] For the case of sports shoes: There is no BREF document for apparel products available, but only for the materials used (see also Joint Research Centre 2015).

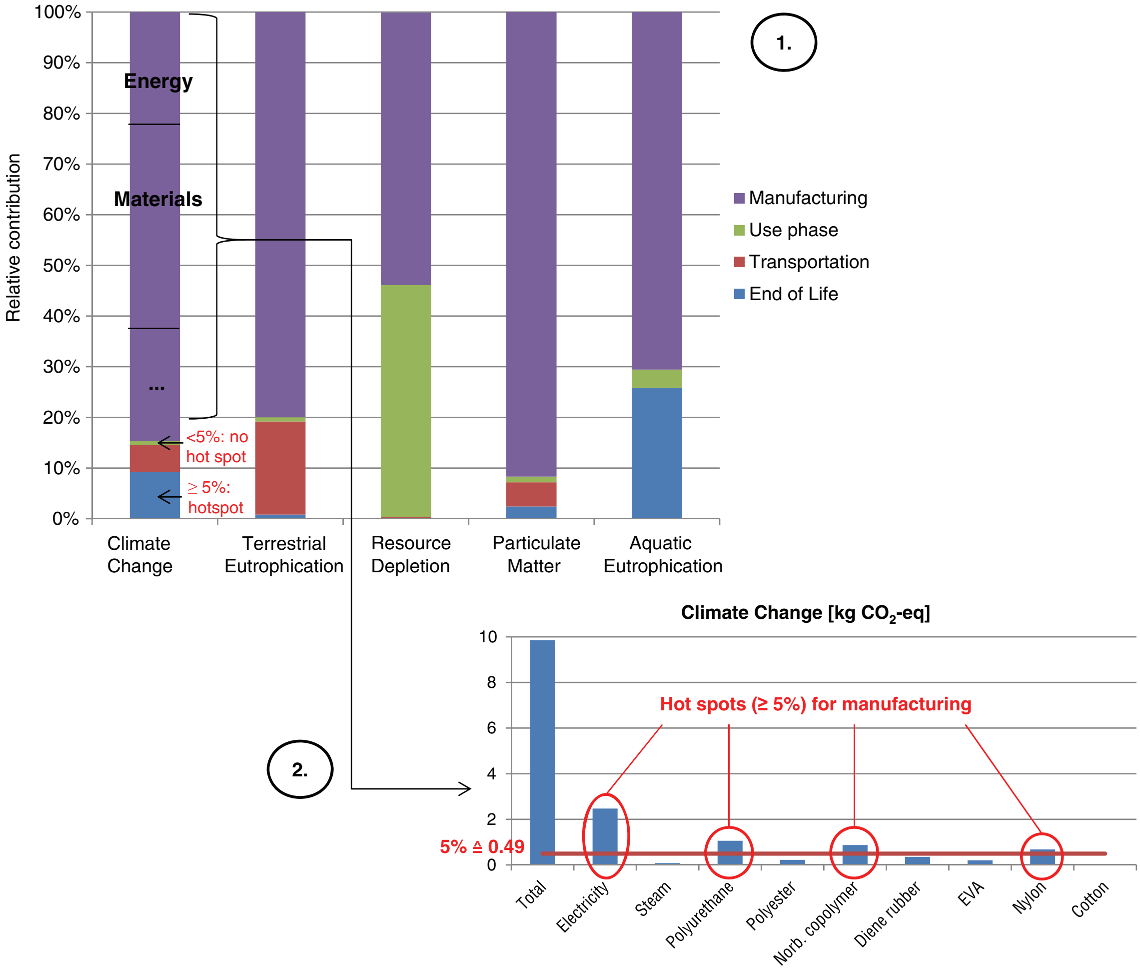


Fig. 2 Illustrative example for the two steps of the hot spot analysis carried out on (1) the life cycle phases of the scenario vector and (2) the sub-variables of manufacturing, contributing 5 % to the overall LCIA result

information on worst practices is available, worst *possible* practices can be approximated instead.

An example for the specification of a best case scenario in the case of sports shoes could be as follows (Eq. 2)):

$$\vec{S_i} = \begin{cases} energy\ (electricity,\ heat) : \text{minimum consumption and "green" energy sources} \\ materials : \text{minimum waste rate and choice of materials with low environmental impact} \\ transportation : \text{low-emitting modes of transport and minimum distances} \\ use\ phase : \text{minimum necessary maintenance} \\ end\ of\ life : \text{no landfill and maximum rate of recycling} \\ lifetime : \text{maximum lifetime indicated through testing} \end{cases} \quad (2)$$

Third, sensitivity analyses are carried out to identify the ranges of classes A, C and E around their determined means. The analyses again focus on the hot spots identified for the respective impact category. Hence, those variables representing hot spots—e.g. energy or transportation—are modified by ±10 % for each mean of the three scenarios “benchmark”, “best” case and “worst” case. As with the hot spot threshold, this range was defined based on common practices and serves as an illustrative example. The thresholds could be adapted when applying the procedure in practice. As depicted in Fig. 3, this can result in a minimum number of three classes if the resulting classes A, C and E cover the entire scale (see orange ovals in Fig. 3). It can also result in a maximum of five classes (see red circles in Fig. 3) as it is intended by the PEF Pilot Guidance.

Based on the sensitivity analyses, the final environmental performance classes can be quantified by carrying out further LCIA on the class borders. Figure 4 gives an example of

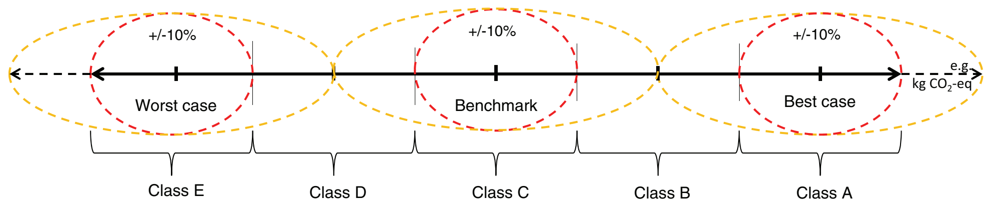


Fig. 3 Determination of the ranges of the environmental performance classes based on the benchmark and best and worst case scenarios through sensitivity analysis (±10 %) resulting in three (*orange ovals*) to five classes (*red circles*)

environmental performance classes for climate change that were quantified for the case study of sports shoes.

The example demonstrates that the environmental performance classes do not necessarily result in a uniform distribution using the proposed procedure. The ranges of the individual classes depend mainly on how extreme the best and worst case scenarios turn out to be and how high the effect is of changing the hot spots for the respective impact category. For the case of sports shoes, the inclusion of lifetime in the proposed procedure has been found to have a high impact as it clearly has an amplifying effect on the class ranges, resulting, for example, in an absolute class range for climate change of 1 to 56 kg CO_2-eq.

4.3 Data uncertainty analysis on environmental performance class borders

The effect of data uncertainty on the borders of the environmental performance classes can be assessed with the help of statistical tools, such as the Monte Carlo simulation. Data uncertainty in LCA is generally related to the inaccuracy or lack of data input (Lo et al. 2005).

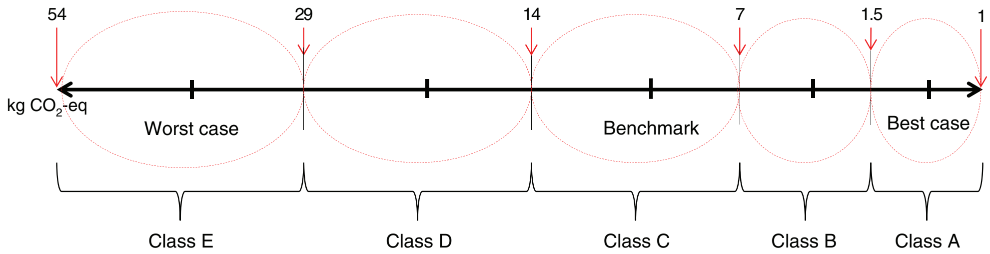
It is usually challenging to identify the concrete uncertain data. Therefore, here, we propose assessing the determined hot spots for data uncertainty. Thereby, it is accounted for the fact that they are the main contributors to the overall environmental impact, and their uncertainties, thus, have a high impact. The main goal of assessing the effect of data uncertainty for environmental performance classes is to look at the resulting distribution of all class borders and to find any potential overlapping of classes to be able to interpret the generated environmental performance classes in a comprehensive way.

For the case study on sports shoes, a data quality assessment of the hot spots was conducted to get an idea of the importance of data quality for the definition of benchmarking and environmental performance classes. This was done by assuming that the quality of the data used to create the sports shoes' life cycle model corresponds to the PEF minimum data quality requirements (European Commission 2013). Based on this data quality assumption, the Monte Carlo simulation generated probability distributions around the class borders. As depicted in Fig. 5 for the example of the impact category climate change, the probability distributions around the left and right class borders of class D show slight overlap. Hence, the minimum data requirements currently defined in the PEF Guide may be too low to allow for a performance class system with five classes, and more stringent data quality requirements may need to be applied to the development process of the benchmark and environmental performance classes to generate robust results and avoid the overlap of adjacent class borders. For this case study, the PEF minimum data quality requirements were chosen as a first step. Future studies should, therefore, also investigate how different, more stringent requirements would influence the resulting distribution of class borders.

5 Conclusions and outlook

This paper presents a procedure developed in the context of PEF to identify a benchmark for a product category and develop environmental performance classes around this benchmark. The gradual application in the case study for sports shoes generated reasonable results. There are, however, constraints that were identified during the development of the proposed procedure, which are summarised in the following.

Fig. 4 Quantified final environmental performance classes of one pair of sports shoes for the impact category climate change



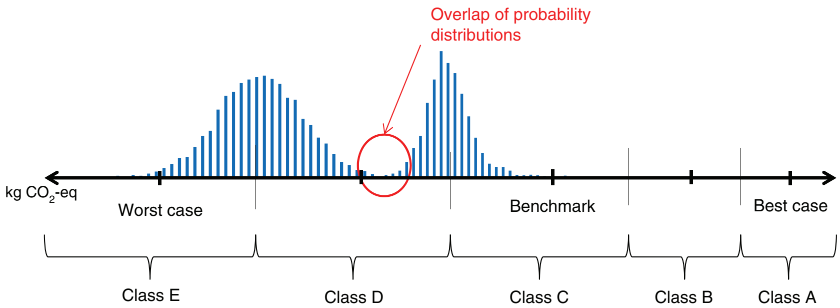


Fig. 5 Example for data uncertainty distributions around the left and right class borders of class *D* overlapping in the case of sports shoes

The quantification of benchmark(s) and environmental performance classes can generally only be regarded as valid if the representative product is actually representative of the sector. Furthermore, the representative product based on PEF requirements does not necessarily represent the average of environmental impacts (Lehmann et al. 2015). According to the PEF Pilot Guidance, the determination of the representative product includes the selection of relevant processes and impact categories. Hence, this could result in excluding environmental impacts based on the analysis of the representative product, while they may be relevant for other products of the same product category. This implies the need to align the definition of the representative product and of the benchmark with the current state of the art of the respective product development as it was done in the case of the Energy Label for household appliances (European Commission 2010a). Moreover, the choice of impact categories and characterisation methods of the European Commission for the PEF Guide is also faced with restrictions as some of them are still not extensively used yet or bear high uncertainty (Finkbeiner 2014).

A general constraint of LCA studies lies in the difficulty to acquire specific and accurate data. The lack of reporting and documentation of factory-specific emissions limits the scope of interpretation and conclusion on the actual environmental performance of the product. For the proposed procedure, this is particularly important as the benchmark's robustness depends highly on the amount and quality of market and technical information available. It is therefore of critical importance to determine the required quantity and accuracy of data to create an appropriate benchmark, since the benchmark is used as the starting point for developing the environmental performance classes. The uncertainty of the "representativeness" of the benchmark, thus, decreases with a higher level of data availability and quality. The importance of data quality is supported by the findings of the case study on sports shoes, which implies a need for more investigation on defining appropriate PEF data quality requirements for the specific case of the benchmark development.

One of the main challenges with regard to communicating PEF results based on benchmarks and performance classes, for example in terms of a PEF label, will be the harmonisation of the environmental performance classes across the impact categories. To avoid manipulation, it has to be ensured that a product does not reach class A in one impact category by sacrificing other impacts that are only reflected in one of the other considered impact categories. This is particularly relevant for those impact categories that differ almost entirely in their hot spots, which was found to be true for climate change and aquatic eutrophication in the case study example of sports shoes. According to the PEF Pilot Guidance, applying normalisation and weighting is foreseen to generate an overall (single) result showing the environmental performance of a product. This, for example, would allow for communicating the benchmark and environmental performance classes to the consumer. However, both normalisation and weighting are related to numerous challenges, such as the availability of suitable normalisation factors or (potentially subjective) weighting factors. The difficulties of normalisation were recently acknowledged by the European Commission based on findings and discussions within the PEF pilot phase (see Section 4.1), and should, hence, be subject to further studies. To refine the developed procedure, further approaches to determine the best and worst case scenarios could be investigated. Cost consideration, particularly for the best case scenario, could be a particular focus in this analysis.

The definite inclusion of lifetime as a vector variable was found to be controversial, especially for non-energy-intensive products. This discussion on including the lifetime needs to be led by a reflection of the consumer-producer relationship. The consumer behaviour plays a crucial role for the lifetime of a product and therefore represents a high uncertainty factor. Fashion trends and innovations cycles (e.g. in the case of electronic devices) are often enforced by the producers themselves, triggering higher consumption and a shorter use phase. This is clearly a contradiction to producing increasingly durable products and, thereby, potentially assigning a high environmental performance class due to a high lifetime. Improving the communication between the producer and the consumer on the importance of the product's lifetime might, thus, be essential to rule out this contradiction. For a final adaptation of a procedure to quantify benchmark(s) and environmental performance classes, a decision on the inclusion of lifetime as a variable needs to be made.

Another influential characteristic of the proposed procedure is the use of hot spot analysis to create the environmental performance classes. On the one hand, this allows identifying critical points of improvement. On the other hand, however, it bears the risk that other minor critical points of the product's life cycle are ignored. This is accompanied by the fact that even though 14 impact categories are considered by the PEF Guide, there is a chance of missing environmental impacts that are not covered by these impact categories or are not fully assessed due to inaccurate LCIA methods and characterisation factors, for example in the case of loss of biodiversity. This could lead to two potential consequences: (1) the benchmark and the environmental performance classes might be distorted causing misinterpretation, and (2) in the case of introducing a PEF label, its scope and validity could be questioned. One intermediate option could be to introduce further requirements (i.e. thresholds) for the environmental performance classes on the use of toxic chemicals similar to the EU Ecolabel (European Commission 2010b), for example, which might have a relatively low absolute impact, but contribute higher in relative terms. One positive effect could be that this provides a certain guideline or framework, in which the best case, i.e. class A, can be reached, which is not the case in the proposed procedure. This approach would be in line with the findings by Lehmann et al. (2014) who suggest that for a product to reach class A, it also needs to comply with certain specific criteria that are adapted from existing environmental labelling schemes. However, before introducing such information in the form of additional environmental impacts, the determination of missing characterisation factors should be addressed.

Generally, the proposed procedure was developed alongside the case study of sports shoes and might, therefore, integrate aspects specific to the product category, particularly for the case of the vector variables. As the procedure is intended to be applicable to other product categories that are consumer goods, it would need to be tested in follow-up studies to identify those specific aspects and generalise them if necessary. Further research could also assess the applicability to intermediate goods, for example in the context of the current PEF pilot phase.

References

Centre for the Promotion of Imports from developing countries – CBI (2010) The EU market for sports footwear. CBI Market Information Database

Cheah L, Duque Ciceri N, Olivetti E et al (2013) Manufacturing-focused emissions reductions in footwear production. J Clean Prod 44:18–29

European Commission (2010a) Commission Delegated Regulation (EU) No 1060/2010 of 28 September 2010 supplementing Directive 2010/30/EU of the European Parliament and of the Council with regard to energy labelling of household refrigerating appliances. Off J Eur Union

European Commission (2010b) Regulation (EC) No 66/2010 of the European Parliament and of the Council of 25 November 2009 on the EU Ecolabel. Off J Eur Union

European Commission (2013) Commission Recommendation (2013/179/EU) of 9 April 2013 on the use of common methods to measure and communicate the life cycle environmental performance of products and organisations—annex II: product environmental footprint (PEF) guide. Off J Eur Union 54

European Commission (2014a) Single market for green products initiative. http://ec.europa.eu/environment/eussd/smgp/index.htm. Accessed 02 June 2014

European Commission (2014b) Environmental Footprint Pilot Guidance document. Guidance for the implementation of the EU Product Environmental Footprint (PEF) during the Environmental Footprint (EF) pilot phase, v. 4.0, May 2014

European Commission – Joint Research Centre (2011) International reference life cycle data system (ILCD) handbook—recommendations for life cycle impact assessment in the European context, 1st edn. Publications Office of the European Union, Luxemburg

Finkbeiner M (2014) Product environmental footprint—breakthrough or breakdown for policy implementation of life cycle assessment? Int J Life Cycle Assess 19(2):266–271

James K, Galatola M (2015) Screening and hotspot analysis: procedure to identify the hotspots and the most relevant contributions (in terms of, impact categories, life cycle stages, processes and flows). Version 4.0. European Commission

Joint Research Centre (2015) Reference documents. http://eippcb.jrc.ec.europa.eu/reference. Accessed June 02 2014

Kowalska M, Cordella M, Wolf O et al (2013) Background report—revision of Ecolabel for the product group "footwear". Joint Research Centre, European Commission, Brussels

Lehmann A, Bach V, Berger M et al (2014) Applying PEF in practice—challenges related to the development of PEFCRs and benchmarks. SETAC Europe 24th Annual Meeting, 15 May 2014, Basel, Switzerland

Lehmann A, Bach V, Finkbeiner M (2015) Product environmental footprint in policy and market decisions: applicability and impact assessment. Integr Environ Assess Manag 11(3):417–424

Lo SC, Ma H, Lo SL (2005) Quantifying and reducing uncertainty in life cycle assessment using the Bayesian Monte Carlo method. Sci Total Environ 340:23–33

Nissinen A, Grönroos J, Heiskanen E et al (2007) Developing benchmarks for consumer-oriented life cycle assessment-based environmental information on products, services and consumption patterns. J Clean Prod 15:539–549

Sustainable Apparel Coalition – SAC (2014) Non-leather shoes: input to steering committee approval. Prepared by members of the SAC and PE INTERNATIONAL. Accessible from the Environmental Footprint wiki. Accessed 15 June 2014