Detecting environmental hotspots in extensive portfolios through LCA and data science: a use-case perspective

Tobias Manuel Prenzel[1, 2*], *Alexander* Shevelov[3], and *Daniel* Wehner[1]

[1] Fraunhofer Institute for Building Physics IBP, Department Life Cycle Engineering GaBi, Stuttgart, Germany
[2] University of Stuttgart, Institute for Acoustics and Building Physics IABP, Department Life Cycle Engineering GaBi, Stuttgart, Germany
[3] Adolf Würth GmbH & Co KG, Künzelsau, Germany

Abstract. Today, businesses need to reduce environmental impacts significantly along the entire value chain. Yet, full organisational product stewardship seems tough for extensive portfolios of several thousands of individual products varying in material and functionality, as well as production processes and locations. In addition, identifying relevant levers for improvement is more challenging with an increasing amount of influencing parameters. Moreover, while a quantification of environmental sustainability performance is required to derive sound management decisions, life cycle assessment (LCA) approaches particularly for large portfolios traditionally fail to provide effective, time efficient means of assessing more than a couple of scenarios per study. In this context, Fraunhofer IBP determined the CO_2-footprint of around 24,000 individual screws in the portfolio of Würth, market leader for assembly and fastening materials, to demonstrate a data science framework for efficient scale-up of environmental sustainability assessments. Hereby, the identification of key hotspots in the portfolio along the value chain was focussed, as well as transparently displaying results and levers for improvement. This contribution builds upon proven methods and tools from LCA and data science and a modularly built approach to achieve a high degree of workflow automation. It offers practical insights into CO_2-footprinting and further environmental sustainability analyses for portfolios with large amounts of individual products.

1 Introduction

With rising pressure from legislation, society and business partners, enterprises experience growing need to increase sustainability performance. Their stakeholders require sound and quantitative information on environmental impacts of all products [1]. Yet, realising environmental sustainability on portfolio level still proves challenging [2]. As a global wholesaler of a large and diverse product portfolio with a considerable number of customers

---

* Corresponding author: tobias.manuel.prenzel@ibp.fraunhofer.de

from various industries and service branches, Würth is facing a similar situation. Efficiently managing and measuring environmental and social aspects of each product along the entire value chain and displaying results in a transparent way is required. In order to comply with this demand and contribute more significantly to the protection of livelihoods of present and future generations, Würth has set the goal to integrate the principles of sustainability and circular economy more intensively in their key strategy and core business operations [3]. Hereby, one decisive factor is continuous transformation of the entire product portfolio towards more eco-friendly products with decreased contribution to climate change. While environmental sustainability is a key to product portfolio development also in other industries [2], the specific focus at Würth is data quality and completeness as well as the calculation of performance indicators to find key hotspots along the value chain and point out levers for improvement.

The method of choice for insights into environmental sustainability is life cycle assessment (LCA) according to the international standard series ISO 14040f. Often, LCA studies focus on singled out products or processes with the intention to gain insights that are transferrable to other portfolio elements. However, this is prone to error. Consequently, Wehner et al. [1] proposed to start with the identification of hotspots at portfolio level instead to avoid three key risks: (1) attribution of large initial effort to gain limited insights on only one product, (2) randomly selecting a product to start with that is irrelevant on portfolio level and (3) the use of specific approaches and tools that are not transferable.

Besides, widely deployed methods and tools to conduct LCA studies often rely on manual data collection and processing devoted to one specific goal, which is laborious and inflexible for both LCA experts and data-owners [1]. In addition, although advantageous for informed decision-making, complete system knowledge over a portfolio requires large efforts [2] and LCA experts can become bottlenecks in widespread consulting constellations [4]. Even in digitised manufacturing environments with data accessible in high volume, variety, and velocity, traditional calculation of sustainability metrics is facing major limitations [5]. Overall, the effort for traditional LCA approaches in terms of time and cost is too high for economic portfolio analyses. With this in mind, promising capabilities of tools and methods from data science in environmental sustainability analysis have already been demonstrated [5, 6, 7].

In a multi-purpose portfolio LCA, one of the key elements is handling data in a well-defined, structured manner. Building on international LCA standards, the cross industry standard for data management (CRISP-DM), and the data science life cycle (DSLC), Wehner et al. therefore recently proposed the sustainability data science life cycle (S-DSLC) for operationalizing data science enhanced product stewardship [8]. This concept offers great potential for an overview over a portfolio's sustainability performance and insights into drivers of environmental impacts with a focus on continued usability to drive sustainable business impact. However, the S-DSLC so far has only been demonstrated with artificially generated data. This paper looks into the adaptation, implementation, and evaluation of the S-DSLC on a business use-case with several stakeholders and shares key learnings for an extensive product portfolio LCA.

2 Case study

2.1 Goal and Scope

The goal of the conducted project is the practical application of the S-DSLC framework to enable an environmental hotspot analysis regarding selected elements in the Würth product portfolio. Hereby, the focus is set on the determination of CO_2-footprints (EF3.0 contribution to climate change in kg CO_2 eq) for cradle-to-gate production. While there may be trade-offs when optimising only one specific key performance indicator (KPI) for portfolio-level supply chains [9], the selected environmental impact serves to validate the S-DSLC framework in practise before applying it for other KPIs, i.e. further LCA impact categories. Equally, the selection of a specific part of the product portfolio at Würth enables testing the concept. Therefore, screws following three international norms and their German equivalents function as pilot products due to their high relevance for Würth: hexagon head bolts (ISO 4014 / DIN 931), hexagon head bolts with thread up to head (ISO 4017 / DIN 933), and hexagon socket head cap screw (ISO 4762 / DIN 912). In total, the selected portfolio elements include 34,617 screws.

There are a number of different production processes allocated to ten different production routes at 31 different production facilities, including four different material groups. The most relevant material is steel of different qualities, with aluminium, brass, and polyamide making up minor shares of the selected portfolio. Regarding size and shape of the products, variation ranges from one to 500 mm length and M2 to M39 thread diameter. Processing for all metal screws generally consists of the manufacturing of coils, cutting, forming the screw head, producing the thread, and final surface treatment, as well as several intermediate cleaning and preparation steps. Polymer screw production is based on injection moulding. Particularly for the metal screws, there are several different techniques for some processing steps, such as cold forming and hot forming that need to be considered.

2.2 Applied approach

For evaluating the described portfolio, the S-DSLC framework introduced by Wehner et al. [8] is first prepared in order to transfer it to practical use cases. Hereby, the adaptation particularly reflects cooperation between different roles formerly not represented in the S-DSLC framework. The roles considered for the adaptation are environmental sustainability experts responsible for generating LCA insights, as well as clients providing product data and deploying findings into business practise.

Second, workflows and data pipelines are set up for the automated calculation of LCA results on portfolio level in line with the modified S-DSLC framework. Readily available ERP data serve as input that includes products and their specifications as well as annual sales data. Relying on ERP data “as is” avoids complex manual data preparation, since all steps would be required for each data or modelling update in the future (e.g. additional products, new LCA background datasets, etc.). To be able to reflect product individuality accordingly, LCA data arecalculated for each specific product. Hereby, the selected data science enhanced LCA approach shows similarities to the concept for a “fast product carbon footprinting methodology” proposed by Meinrenken et al. [9]. Yet, dedicated LCA models are used instead of emission factor estimation to increase modelling accuracy for the automated determination of CO_2-footprints per product. These dedicated LCA models are prepared in standard LCA software for frequently recurring production operations for screws (“modules”) and subsequently processed in KNIME Analytics Platform.

Lastly, the S-DSLC and the automated workflows are tested and validated on the described Würth portfolio. In doing so, the interpretation of ERP data relies on supportive documentation regarding company-internal abbreviations, product hierarchies, and material codes. This, in combination with product-specific standards, as well as some unstructured data from product owners form the product-related data. Information on manufacturing processes in different formats and LCA background data for materials, energy supply, and transportation complete the input data. Whenever necessary, generic database processes substitute primary data, e.g. when no supplier data are available. A modular approach allows swapping these easily later on, if required. Additionally, different implemented scenarios facilitate geographical, supplier-specific, and material-specific comparisons. A comprehensible, business-user oriented presentation of results finally extracts insights from the large number of individual results calculated and helps drawing meaningful conclusions.

3 Results

The conducted project proved a concept for automation of LCA modelling suitable for extensive product portfolios. For the transfer of the S-DSLC to the described use case modifications to the theoretical concept play a key role. The adapted S-DSLC depicted in Fig. 1 reflects the cooperation between LCA experts for creation on the one side, and business users for deployment on the other. This accounts for different roles and responsibilities of both and is typical for a large number of companies making use of external knowhow for environmental sustainability assessments. Würth acts as the data owner and supplies input on portfolio, products, and production. Fraunhofer IBP then works on the creation of data science enhanced LCA. In doing so, Fraunhofer IBP provides background data and knowhow to calculate environmental impacts from obtained inputs. Lastly, Würth is responsible for deploying the insights into business practice.

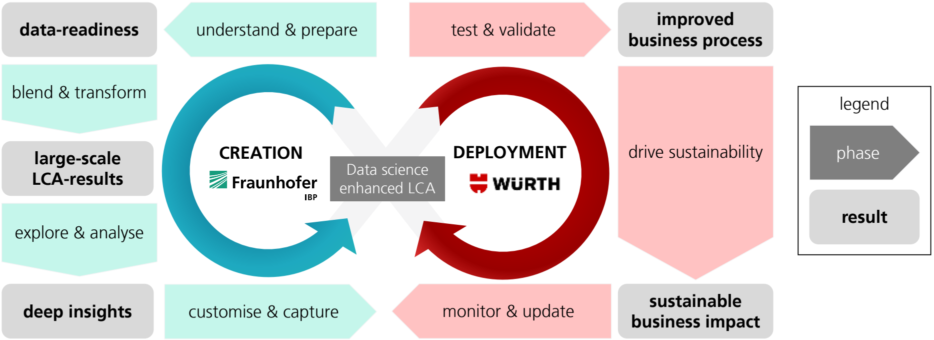


Figure 1. S-DSLC according to Wehner et al [8] adapted for the realised cooperation between LCA experts and business users to conduct large-scale portfolio analyses.

The environmental sustainability assessment for the portfolio conducted at Fraunhofer IBP (cf. “creation”) features elements well known to LCA practitioners. First, a comprehensive product list results from several iterative data preparation and processing steps (pull ERP data, merge, clean, and process). This list contains every single product with an unequivocal identifier as well as all relevant product specifications required for modelling. Second, modules representing parts of the production route result from different data sources. They depict a very fundamental LCA concept in the approach: they break down the entire value chain into aggregated processes and their environmental sustainability metrics, which can be combined for modelling complex production routes after scaling according to product specifications.

A dedicated mapping algorithm that unequivocally determines the production route of a portfolio element from its specifications addresses the large number of combinatory possibilities of production routes, facilities, and material groups. The algorithm is designed as a decision tree, returning a list of modules assigned to each portfolio element. The implementation of the modified S-DSLC also features an automated workflow for scaling and adding the environmental impacts for these modules to yield product specific sustainability metrics. Hereby, scaling is based on product specifications such as material, mass, surface area, etc. Both mapping algorithm and workflow significantly facilitate automation, as manual attribution and scaling of unit processes to a specific product is no longer required. Fig. 2 gives an overview over the standard LCA elements (product list, modules) and approaches from data science (mapping algorithm, automated workflow) used for conducting large-scale portfolio LCAs “with one click”.

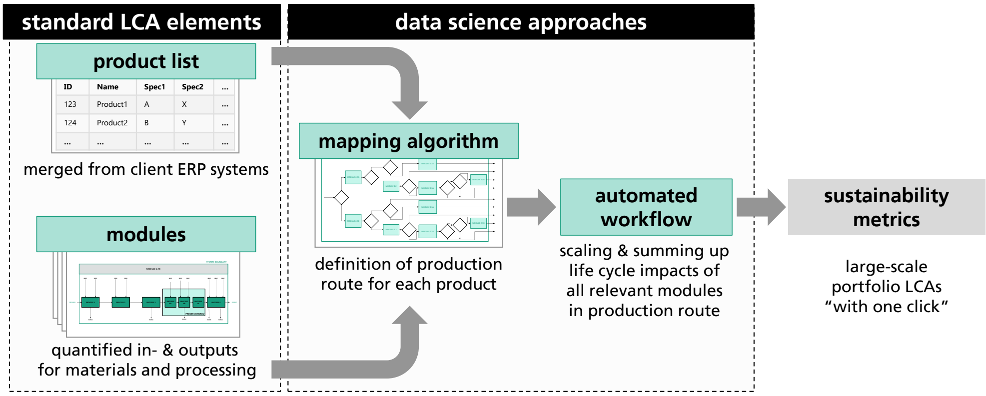
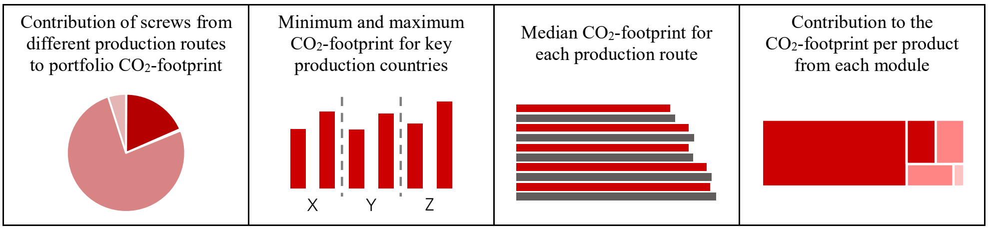


Figure 2. Overview of the modular and scalable approach building on elements from LCA and data science for the automated calculation of large-scale portfolio metrics on environmental sustainability.

After adapting the S-DSLC framework and setting up the described workflows and modules for the considered materials and processing steps (c.f. section 2.1), the calculation of environmental sustainability metrics only requires “one click” and produces results instantly. In the scope of the project, the conducted assessment yielded roughly 24.000 individually calculated CO_2-footprints (EF3.0 climate change in kg CO_2 eq.) for distinct portfolio elements. Already a small deviation in specifications sometimes determines a completely different set of required modules, resulting in significantly different environmental impacts for the product. Additionally, different implemented scenarios facilitate geographical, supplier-specific, and material-specific comparisons. A short report features insights from an exploratory analysis of the calculated CO_2-footprints, such as key drivers in production and scenario comparison. Fig. 3 depicts exemplary result illustrations. The project outcomes also include a list of all calculated CO_2-footprints with an indicator of data confidence for each product and production module depending on underlying modelling data.


Contribution of screws from different production routes to portfolio CO_2-footprint

Minimum and maximum CO_2-footprint for key production countries

Median CO_2-footprint for each production route

Contribution to the CO_2-footprint per product from each module

Figure 3. Exemplary results in the short report derived from an exploratory analysis of the calculated CO_2-footprint results for roughly 24,000 specific products.

Additionally, the study helps to identify relevant data gaps for environmental sustainability modelling of the portfolio, which forms an important project goal at Würth. Due to missing values for the necessary product specifications, the algorithm cannot process some products. A shortlist of all mandatory product specifications missing in the current data set and suitable mitigation strategies is currently used to review products in question and helps increasing data quality within the ERP system.

4 Conclusion

This contribution demonstrated the practical application of the sustainability data science life cycle (S-DSLC) for the evaluation of an extensive product portfolio of roughly 24,000 individual screws varying in a wide set of different parameters. The presented S-DSCL application offers a novel approach that can be integrated into existing LCA practise, without introducing yet another software solution or tool. The integration of readily available ERP data on products and their specifications with data on production processes yielded a quick and cost-effective LCA screening, initially limited to the CO_2-footprint (EF3.0 climate change in kg CO_2 eq). In the future, other impact categories can be evaluated analogously, following the applied approach. The created results, as for any conventional LCA, can feed into dashboards, BI tool, reports, presentations etc. and ultimately provide insights regarding hotspots on portfolio and product level, data gaps, and strategies to proceed. In contrast to standard LCA procedure, automated workflows mostly eliminate the need for manual data processing steps and allow significant increase of scale and efficiency at which portfolio LCAs can be conducted. Moreover, the modular concept ensures simple updateability of portfolio elements, manufacturing steps, and LCA background data individually. Lastly, easy transferability to other product systems, portfolios, and industries is expected due to reusable workflows and data pipelines. To this end, further applications are trialled at Fraunhofer IBP.

At Würth, obtained insights help to gain better understanding of the company's environmental footprint and simultaneously enable easier identification and analysis of environmental weak points. Based on the conducted portfolio LCA, future data collection will play a crucial role in order to identify optimisation potentials. This will serve as a starting point for improvement and strategic decisions that lead to cost and risk reduction in business practise. Furthermore, the data analysis improves quantity and quality of the sustainability-relevant product database at Würth. In addition, it constitutes a data management framework for further evaluations. With this in mind, Würth is able to become more transparent and reliable in internal and external communication through data provision, particularly with regard to fulfilling environmental sustainability requirements of customers and government.

The presented results were obtained by Fraunhofer IBP for Adolf Würth GmbH & Co. KG. In particular, the authors would like to acknowledge the work and dedication from all project participants.

References

1. D. Wehner et al.: *Rethinking LCA und Product Stewardship in Industrieunternehmen*, Abteilung Ganzheitliche Bilanzierung des Fraunhofer-Instituts für Bauphysik IBP, Stuttgart (2021)
2. C. Villamil and S. Hallstedt: *Sustainability integration in product portfolio for sustainable development: Findings from the industry*. Bus Strat Env **30**. pp 388-403. DOI: 10.1002/bse.2627 (2021)
3. Adolf Würth GmbH & Co. KG: *#HelloCircle. Nachhaltigkeitsbericht 2018/2019*. https://www.wuerth.de/web/media/downloads/pdf/nachhaltigkeit/wuerth_nachhaltigkeitsbericht_2019.pdf, accessed 18.10.2021 (2020)
4. N. Otte et al.: *Decentralized LCA in innovation and reporting processes of a large enterprise*. The 10th Int. Conf. on Life Cycle Management, virtual. 08.09.2021. (2021)
5. D. Wehner et al.: *Towards industry 4.0 ready advanced sustainability analytics*. Chapter 30. In: Challenges for Technology Innovation: An Agenda for the Future. Proceedings of the International Conference on Sustainable Smart Manufacturing (S2M 2016), 20-22 October 2016, Lisbon, Portugal (2017)

6. T. Betten et al.: *Integration of Big Data Analytics into Life Cycle Assessment*. LCA XVIII. Fort Collins, CO, USA. 25-27.09.2018 (2018)
7. D. Wehner: *Ordnungsrahmen und Ansätze für das ökologische Risikomanagement bei der Produktentstehung*. Forschungsergebnisse aus der Bauphysik **40**. Fraunhofer Verlag. ISBN: 978-3-8396-1627-7 (2020)
8. D. Wehner et al.: *Workflow automation for multi-purpose LCA of large product portfolios*. The 10th Int. Conf. on Life Cycle Management, virtual. 08.09.2021. (2021)
9. C.J. Meinrenken et al.: *Combining Life Cycle Assessment with Data Science to Inform Portfolio-Level Value-Chain Engineering*. JIEC **18**, 5. DOI: 10.1111/jiec.12182 (2014)