Machine Learning to support prospective Life Cycle Assessment of emerging chemical technologies

Blanco, C.F.; Pauliks, N.; Donati, F.; Engberg, N.; Weber, J.

Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in
Green and Sustainable Chemistry

Machine learning to support prospective life cycle assessment of emerging chemical technologies

C. F. Blanco[1,2], N. Pauliks[1], F. Donati[1], N. Engberg[3] and J. Weber[4]

Increasing calls for safer and more sustainable approaches to innovation in the chemical sector necessitate adapted methods for the environmental assessment of emerging chemical technologies. While these technologies are still in the research and development phase, gaining an early understanding of their potential implications is crucial for their eventual introduction into markets worldwide. Life Cycle Assessment (LCA) is a core tool which has been recently adapted for such purpose. Prospective LCA approaches aim to develop plausible future-oriented models which account for the evolution of factors both intrinsic and extrinsic to the technologies assessed. Such future-oriented models introduce many indeterminacies, which could, to some extent, be addressed by Machine Learning techniques. Recent demonstrations of such techniques in the context of prospective LCA, as well as promising avenues for further research, are critically discussed.

Addresses
[1] Institute of Environmental Sciences (CML), Leiden University, the Netherlands
[2] TNO, Circularity and Sustainability Impacts Department, Utrecht, the Netherlands
[3] Delft University of Technology, Industrial Design Engineering, Delft, the Netherlands
[4] Delft University of Technology, Department of Intelligent Systems, Delft, the Netherlands

Corresponding author: Blanco, C.F. (c.f.blanco@cml.leidenuniv.nl)

Current Opinion in Green and Sustainable Chemistry 2024, **50:**100979

This review comes from a themed issue on **AI and Sustainability**

Edited by **Jana Weber** and **Kevin Leonard**

Available online 18 October 2024

For complete overview of the section, please refer the article collection - AI and Sustainability

https://doi.org/10.1016/j.cogsc.2024.100979

Introduction

The call for a more responsible approach to innovation has spurred methodological advancements for the environmental assessment of novel materials and products. This is of particular importance for emerging chemical technologies, which are largely driving global oil consumption [1], belong to the hard-to-abate sectors [2] and whose products preponderate in most aspects of our lives [3]. Since the 1980's, Life Cycle Assessment (LCA) has prevailed as the preferred tool for assessing product and service performance across diverse environmental criteria [4]. In assessing multiple criteria across entire product life cycles, LCA is unique in its ability to reveal undesirable environmental trade-offs associated with the consumption of products and services.

LCA was originally conceived as an *ex-post* assessment, wherein the material/energy consumptions of product systems and their environmental emissions could — at least in principle — be measured based on past data. However, with the shift to responsible innovation, the LCA community is now proposing future-oriented approaches for early environmental appraisal of emerging technologies that are still under development and can only partly be empirically measured [5—9]. Henceforth, we will refer to these approaches collectively as *prospective* LCA [10]. In the context of the chemical industry, a prospective LCA might be required to redesign a current chemical production system anticipating future scenarios, or for assessment during novel chemical discoveries and process development [11]. The ultimate goal is to incorporate sustainability criteria during the initial stages of chemical (re)discovery and process (re)development, maximising the potential for significant changes while minimising resource expenditures [12].

In order for LCA results to be meaningful, they must represent a marketable technology and foreseeable conditions in which it would operate when released. This may require future projections of 10 or more years, depending on technological maturity. However, future projections are characterised by high levels of uncertainty. While *ex-post* LCA methods already grappled with challenges regarding uncertainty, data availability and variability [13—15], prospective LCA introduces a novel challenge (i.e., the need to anticipate how both internal and external factors will evolve during the research and development stages of emerging technologies) [16—19]. Some examples of such indeterminacies in the context of the chemical industry are: characteristics of novel chemistry itself (yields, reaction conditions and

Given the role as Guest Editor, Jana Weber had no involvement in the peer review of the article and has no access to information regarding its peer-review. Full responsibility for the editorial process of this article was delegated to Kevin Leonard.

further), process scale-up behaviour, anticipated future chemical supply chains and process electrification scenarios.

To meet the challenge, prospective LCA has employed process modelling, scenario analysis and various forms of uncertainty analysis and sensitivity analysis [16,20,21]. In addition to these, first demonstrations of the use of Machine Learning (ML) that can help address key knowledge gaps have been presented. In the following sections we will discuss selected examples of the latter, distinguished according to three main applications (for a detailed explanation of the LCA model and its components and definitions, we refer the reader to Ref. [22]):

1. Surrogate models for early-stage screening: ML-derived models that contribute to early-stage screening and ranking without LCA model construction. This includes algorithms that predict LCA scores, related early-stage sustainability metrics, or fill in data gaps that allow for early-stage assessment (Figure 1: A).
2. Informing prospective LCA models: ML algorithms to aid LCA model construction, assessment and interpretation by estimating model parameters within specific phases of LCA:
   - *Goal and scope definition*: parameters needed to define the function, alternatives, functional units, reference flows and system boundaries of future product systems (Figure 1: B)
   - *Life Cycle Inventory Analysis*: economic (Figure 1: C) and environmental (Figure 1: D) inputs and outputs of unit processes in future product systems
   - *Life Cycle Impact Assessment*: characterisation factors that translate environmental inputs/outputs to impacts (Figure 1: E, F, G, H), as well as normalisation factors (Figure 1: I)
   - *Interpretation*: parameters to quantify the magnitude of uncertainty (Figure 1: J) and sensitivity (Figure 1: K) in prospective LCA models. Novel techniques have also been presented to identify scenarios of interest (Figure 1: L)

Here we also consider ML algorithms developed for models outside the LCA domain (e.g. models of functional performance and reliability, landscape parameters that are used in impact models) which can serve to inform and complement the above elements of the LCA model.

3. Decision support: ML techniques that can enhance decision-making following the results of prospective LCA (e.g., in the determination of pareto-optimal technology configurations via multi-objective optimisation) (Figure 1: M)

In Figure 1 we also differentiate between elements that have been targeted in published case studies where ML has been clearly integrated in prospective LCA (green), potential elements where we believe ML can already support prospective LCA (yellow), and elements unexplored in the context of prospective LCA but with potential integration in the future (red).

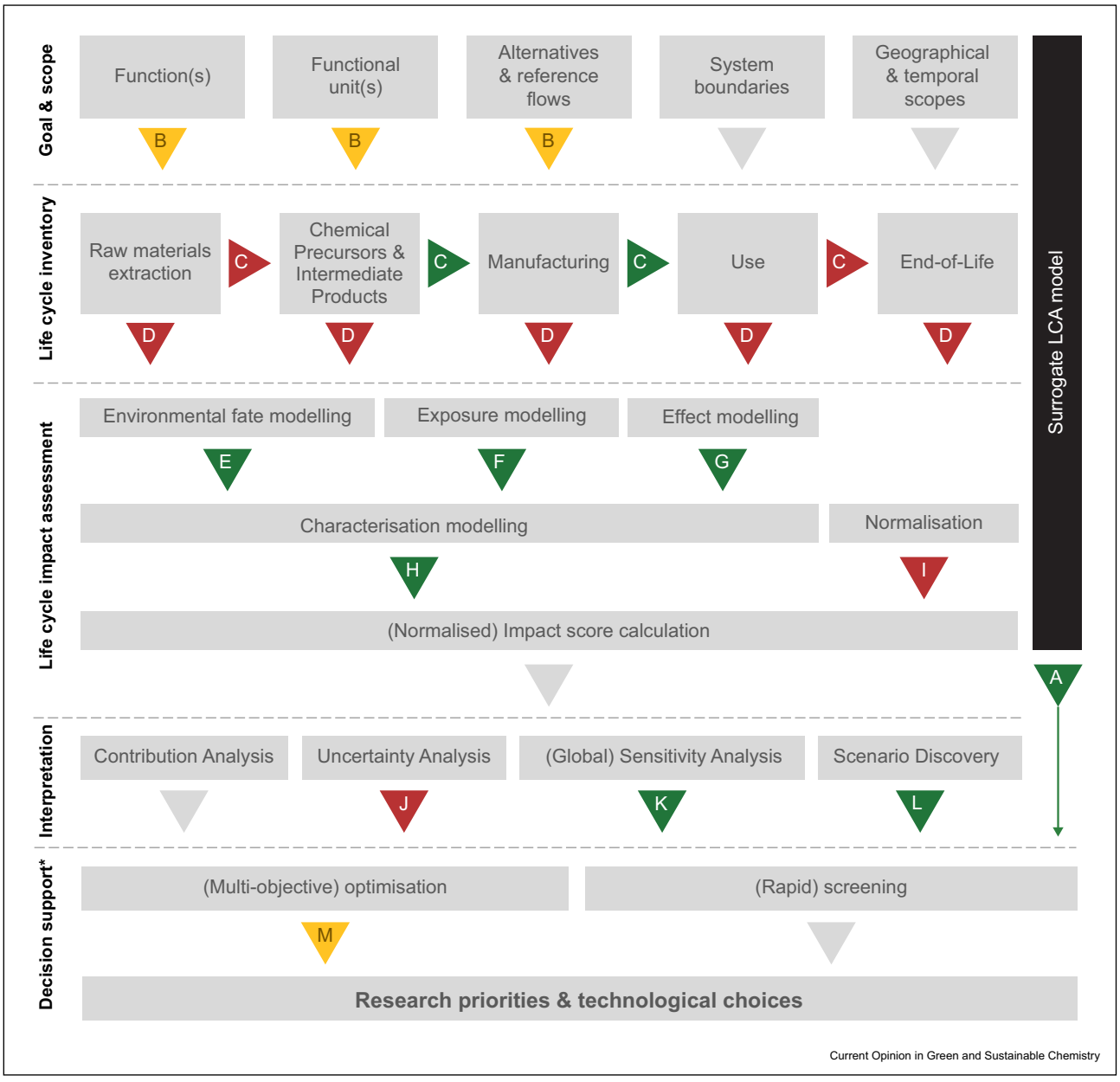
Surrogate models for early-stage screening

One important application for ML in LCA of emerging chemical technologies is to enable the early-stage screening of thousands of novel products, molecule alternatives, or reaction/production pathways. The known chemical design space is estimated to contain about 100 million molecules, whereas the unknown chemical space can be approximated to go up to the order of 10^{60} molecules [12]. Considering the laborious nature of modelling new products and compiling LCA databases, a sensible initial application for ML is to approximate selected sustainability criteria or the results for the entire LCA model, using predictions based on existing LCA databases. The general intention is not to entirely circumvent LCA models of novel products or novel production routes, but to screen for most promising configurations in large and early-stage decision-making space and then explore more limited configurations through full prospective LCA models [23].

We note two distinct approaches here. First, different reaction pathways towards specified products or product categories can be assessed through optimisation approaches [24–26]. Common applications are biorefineries, bioplastics, or further emerging process or product classes. We expect ML to play a supporting role in filling data gaps for such large-scale and early-stage screening schemes. For instance, models like the Molecular Transformer may complete known or predict novel reactions along the pathways [27–29] or may be used to predict reaction yields [30]. Artificial Neural Networks can be used to inform about the reaction context [31] and QSPR models (quantitative structure-property relationships) or Graph Neural Networks can make predictions for a range of molecular properties needed to evaluate the reactions (e.g. toxicity or thermodynamic properties) [32–34].

Second, life-cycle impacts have been estimated directly for specific chemicals based on molecular and thermodynamic properties and process-level parameters. Such an approach was developed in the FineChem [35,36] and recently updated FineChem 2 [37] tools, which predict product carbon footprints, critical raw materials and the endpoint Ecoindicator99 from existing LCA databases and molecular structures. In a similar fashion, the CLiCC tool [38], and the implementation by Sun et al. [39] directly calculate additional impact scores such as cumulative energy demand, acidification, human

Figure 1



Future-oriented elements in prospective LCA of emerging chemical technologies that might be informed by ML. Triangles represent output variables with the information obtained from different modelling steps that have (green) or may be (yellow) predicted or informed by ML techniques. In red we indicate variables for which ML support is largely unexplored to date but may be an interesting research topic. *Decision support is not a phase of the standard LCA model, but is included here as a key aspect of early-stage environmental assessment of emerging technologies. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

health, ecosystem quality, metal depletion, freshwater ecotoxicity, and particulate matter formation. Additional demonstrations of this approach are found in Kleinekorte et al. [40] and Karka et al. [41].

Notably, the literature on surrogate models for LCAs does not differentiate explicitly between *ex-post* and prospective LCA predictions. The current state of research in screening applications is on improving model accuracy, generating relevant datasets, and integrating tools for multiple data gaps. As of now, this has been irrespective of the stage of technological development. In the following section, we discuss ML applications to help fill knowledge gaps within the full prospective LCA models and aid in model construction, assessment and interpretation.

Informing prospective LCA models

Future-oriented goal and scope

Indeterminacies in the first phase of the LCA method, *Goal and Scope*, are contextual and relate to the adequacy of high-level definitions and modelling choices regarding (e.g. function, functional unit, reference flow and system boundaries) [42]. From a prospective LCA standpoint, the additional question is how future developments could influence any of such modelling choices. For example, Hischier et al. [19] rightly highlight the challenge in defining a functional unit for nanomaterials. So-called multifunctional nanomaterials may deliver more than one function and in different quantities [43]. This will largely depend on the properties of the nanomaterials (e.g. electronic, magnetic, optical, catalytic) which may not be fully understood until later R&D stages. This often-overlooked aspect is hard to capture in prospective LCA models, yet it is very relevant as LCA models are highly sensitive to functional performance variations [44]. There are already several publications demonstrating the use of ML to predict such performance-related properties of nanomaterials [45]. ML has also been used to predict the functional performance of other technologies often assessed in prospective LCA (e.g. long-term energy yield of photovoltaic installations [46,47] and battery storage) [48]. Here we observe that such technology-specific models would likely be produced by technology experts (rather than LCA specialists) and would only inform functional unit definitions in the LCA model. However, the scopes, assumptions and limitations of these models would be inherited by the prospective LCA model and so must be properly understood and communicated.

Prediction of prospective inventory data

The second phase of LCA includes the collection of data, namely the material and energy inputs and outputs associated with the unit processes or activities that comprise the product system or value chain, from raw materials extraction to end of life. Naturally, this aspect has been the focus of most prospective LCA strategies to date, where proposed approaches include, e.g. foreseeable roadmaps/scenarios in raw materials and energy sourcing and manufacturing [8].

A central question for prospective LCA of emerging chemical technologies is how the synthesis of novel chemicals and materials will be upscaled for mass production [20,49]. Evidently, this will have a large influence on the materials and energy requirements of a commercialised product. Several authors have advocated for the use of process simulation to develop upscaled models of the manufacturing stage [50–52]. ML has already been used in process simulation optimisation in, e.g. polymer composites [53] and semiconductors [54]. As an example, Lockner & Hopmann [55] train an ANN to optimise parameters for an injection moulding process; injection flow rate, holding pressure time, holding pressure, cooling time, melt temperature, and cavity wall temperature. All of these parameters could determine energy inputs of foreground processes in a prospective LCA model of novel polymers. The combination of ML-supported process simulation for LCA shows large potential to address data gaps in the inventories of upscaled chemical technologies [56].

One often overlooked but highly sensitive aspect of the inventory phase is the choice of allocation factors. They inform the model on how to allocate environmental burdens when unit processes perform more than one economic function (multifunctional), including chemical co-products and recycling of waste. Allocation factors can be based on the economic values of each product/service, on physical principles of the products (mass/energy content) or others. As for economic factors, ML-based forecasting has been widely applied to commodity prices, and thus we can imagine the use of such techniques for multifunctional systems in the LCA framework for finding allocation factors that depict future developments. Examples are co-produced metals such as copper and molybdenum [57]. We want to emphasise that the interpretation of LCA results is highly sensitive to allocation factors. Thus, while this approach allows to consider future economic scenarios, it also introduces novel model uncertainties which may or may not be desirable. If allocation factors are based on ML commodity price predictions, we recommend carefully executed *exploratory* scenario approaches complemented by other types of scenario analysis.

Prediction of prospective characterisation factors

Perhaps the most active area of work on integrating ML with LCA of chemicals is on the prediction of (human and eco-) toxicity characterisation factors (CFs) for the life cycle impact assessment phase of LCA [58–61]. These approaches typically rely on molecular descriptors as model inputs and molecule-specific CFs as outputs. Subsequently, the CFs are then multiplied by the emitted chemical masses in each compartment (obtained during the inventory phase) to produce an impact score. In a more nuanced approach, von Borries et al. [62] propose ML to predict the most influential parameters required to calculate the CF's, including fate, exposure and effect factors such as degradation half-life and bioaccumulation.

Other impact categories next to toxicity are currently underexplored. As emerging chemical technologies advance through R&D, impact characterisation factors may be contingent on evolving extrinsic factors. LCA impact categories such as abiotic resource depletion and water scarcity are temporally dependent and highly sensitive to future trends. An emerging chemical

technology may be better or worse placed to benefit from such trends depending on the quantity and location of the materials and water demand of its value chain [63]. Through forecasting environmental variables and coupling them with different characterisation methods, insights into the future evolution of impact mechanisms can be obtained. This is especially interesting for CFs whose underlying environmental data is not covered in Integrated Assessment Models, as is the case with the water scarcity method AWARE [64].

Interpretation

The ubiquitous nature of uncertainty and complexity in prospective LCA calls for a more nuanced and sophisticated interpretation of results. Various authors have argued for the application of Global Sensitivity Analysis (GSA) [65], as it can add robustness in the presence of numerous, large and heterogeneous uncertainties, and it can also aid in model simplification and dimensionality reduction by identifying poorly known parameters of low or high relevance [16,44,66,67]. ML-based techniques have been recently proposed to conduct GSA [68—70], which could streamline its implementation in prospective LCA. For example, random forest algorithms were used to produce surrogate (meta) models. Here, the random forest variable importance measure can be used to define sensitivity measures in a computationally efficient way, while accounting for correlations and interactions in high-dimensional LCA models (numerous uncertain variables) [69].

Other authors have recognised the presence of unquantifiable and so-called *deep* uncertainties as well as more general forms of indeterminacy in prospective LCA which fall in the realm of post—normal science [71]. The *scenario discovery* approach employs ML and other algorithms to systematically identify combinations of uncertain parameters that produce scenarios that fail to meet performance criteria [72—74]. Such an approach was demonstrated for a case study of geothermal heating networks, where diverse design alternatives with uncertain parameters were screened to find combinations of parameter ranges that would produce scenarios of interest (e.g. exceeding a water consumption threshold as calculated by LCA) [68]. A more recent implementation of scenario discovery in the context of prospective LCA applied to the development of microalgal compounds for aquaculture can be found in Jouannais et al. [71].

Decision support systems

Given prospective LCA's mission to guide technological development, we believe the analysis in this review must extend beyond discussion of LCA results and step well into decision-making. This means translating LCA model outcomes — including uncertainty and sensitivity analysis — into actionable decisions. ML has been increasingly used within Decision Support Systems [75] (e.g. in multi-objective optimisation) [76,77]. A noteworthy example is the SUSCAPE tool [78], which optimises chemical processes while taking into consideration LCA indicators. SUSCAPE uses neural networks to derive surrogate models from full process models (flow sheet simulators or large differential equation systems). The surrogate models are then passed to multi-objective optimisation algorithms to produce a set of solutions for optimal process design. In the final step, the solution set is filtered and ranked using Data Envelope Analysis techniques. Multi-criteria decision analysis (MCDA) tools have also been proposed as a final step to rank the optimal solutions based on stakeholder preferences [79].

Discussion and conclusions

It is well-reckoned within the scientific modelling community that the future is unpredictable. Thus, prospective LCA of emerging chemical technologies supported by ML should focus on ensuring that the decisions made during R&D are the best ones given the information available to us at present. ML no doubt has a promising potential to enhance the broader decision-making process within which prospective LCA is embedded, guiding it towards safer and more sustainable innovation. At the same time, there is a risk of ML obscuring the analyses and/or creating a false sense of confidence in future scenarios. This is a risk, as modellers must acknowledge that ML algorithms bring their own sources of error, even if best practices such as rigorous data preparation and model validation are applied. The irreducible errors should be analysed and clearly disclosed.

ML will also tend to drive results towards expected values or ranges which are for the most part predicated on past data. The future does not necessarily behave like the past, so it is imperative to keep a close eye on potential *black swans*, i.e. low-likelihood but high-consequence scenarios and emergent behaviour resulting from complex systems in which LCA product systems are embedded. The other side of this coin is that we might also want to allow room for creativity and out-of-the-box thinking, which is not only based on statistical predictions.

We conclude by advocating for the continued and energised pursuit of research in the integration of ML into prospective LCA, embracing ML not as a silver bullet but as an important decision-support tool in a broader catalogue of prospective LCA strategies.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the author(s) used ChatGPT in order to improve language and readability. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank Bartolomeus Häussling Löwgren for his valuable insights and comments on part of this work.

Data availability

No data was used for the research described in the article.

References

Papers of particular interest, published within the period of review, have been highlighted as:

* of special interest
** of outstanding interest

1. Kätelhön A, Meys R, Deutz S, Suh S, Bardow A: **Climate change mitigation potential of carbon capture and utilization in the chemical industry**. In *Proceedings of the national academy of sciences*; 2019:11187–11194. https://doi.org/10.1073/pnas.1821029116. vol. 116.
2. Paltsev S, Morris J, Kheshgi H, Herzog H: **Hard-to-Abate Sectors: the role of industrial carbon capture and storage (CCS) in emission mitigation**. *Appl Energy* 2021, **300**, 117322. https://doi.org/10.1016/j.apenergy.2021.117322.
3. Santos A, Barbosa-Póvoa A, Carvalho A: **Life cycle assessment in chemical industry – a review**. *Curr Opin Chem Eng* 2019, **26**: 139–147. https://doi.org/10.1016/j.coche.2019.09.009.
4. Guinée JB, Heijungs R, Huppes G, Zamagni A, Masoni P, Buonamici R, *et al.*: **Life cycle assessment: past, present, and future**. *Environ Sci Technol* 2011, **45**:90–96. https://doi.org/10.1021/es101316v.
5. Cucurachi S, Van Der Giesen C, Guinée J: **Ex-ante LCA of emerging technologies**. *Procedia CIRP* 2018, **69**:463–468. https://doi.org/10.1016/j.procir.2017.11.005.
6. Arvidsson R, Tillman A-M, Sandén BA, Janssen M, Nordelöf A, Kushnir D, *et al.*: **Environmental assessment of emerging technologies: recommendations for prospective LCA**. *J Ind Ecol* 2017, **22**:1286–1294. https://doi.org/10.1111/jiec.12690.
7. Thonemann N, Schulte A, Maga D: **How to conduct prospective life cycle assessment for emerging technologies? A systematic review and methodological guidance**. *Sustainability* 2020, **12**:1192. https://doi.org/10.3390/su12031192.
8. Adrianto LR, van der Hulst MK, Tokaya JP, Arvidsson R, Blanco CF, Caldeira C, *et al.*: **How can LCA include prospective elements to assess emerging technologies and system transitions? The 76th LCA Discussion Forum on Life Cycle Assessment, 19 November 2020**. *Int J Life Cycle Assess* 2021, **26**:1541–1544. https://doi.org/10.1007/S11367-021-01934-W. 8 2021.
9. Wender BA, Foley RW, Prado-Lopez V, Ravikumar D, Eisenberg DA, Hottle TA, *et al.*: **Illustrating anticipatory life cycle assessment for emerging photovoltaic technologies**. *Environ Sci Technol* 2014, **48**:10531–10538. https://doi.org/10.1021/es5016923.
10. Arvidsson R, Svanström M, Sandén BA, Thonemann N, Steubing B, Cucurachi S: **Terminology for future-oriented life cycle assessment: review and recommendations**. *Int J Life Cycle Assess* 2023. https://doi.org/10.1007/s11367-023-02265-8.
11. Kralisch D, Ott D, Gericke D: **Rules and benefits of Life Cycle Assessment in green chemical process and synthesis design: a tutorial review**. *Green Chem* 2015, **17**:123–145. https://doi.org/10.1039/C4GC01153H.
12. Reymond J-L, Awale M: **Exploring chemical space for drug discovery using the chemical universe database**. *ACS Chem Neurosci* 2012, **3**:649–657. https://doi.org/10.1021/cn3000422.
13. Huijbregts MAJ: **Application of uncertainty and variability in LCA**. *Int J Life Cycle Assess* 1998, **3**:273–280.
14. Ross S, Evans D, Webber M: **How LCA studies deal with uncertainty**. *Int J Life Cycle Assess* 2002, **7**:47. https://doi.org/10.1007/BF02978909.
15. Lloyd SM, Ries R: **Characterizing, propagating, and analyzing uncertainty in life-cycle assessment**. *J Ind Ecol* 2007, **11**: 161–179. https://doi.org/10.1162/jiec.2007.1136.
16. Cucurachi S, Blanco CF: **Practical solutions for ex-ante LCA illustrated by emerging PV technologies**. In *Ensuring the environmental sustainability of emerging technologies*. Edited by Florin M-V, Ed, Lausanne: EPFL International Risk Governance Center; 2023:149–167.
17. Gavankar S, Anderson S, Keller AA: **Critical components of uncertainty communication in life cycle assessments of emerging technologies**. *J Ind Ecol* 2015, **19**:468–479. https://doi.org/10.1111/jiec.12183.
18. Lacirignola M, Blanc P, Girard R, Pérez-López P, Blanc I: **LCA of emerging technologies: addressing high uncertainty on inputs' variability when performing global sensitivity analysis**. *Sci Total Environ* 2017, **578**:268–280. https://doi.org/10.1016/j.scitotenv.2016.10.066.
19. Hischier R, Salieri B, Pini M: **Most important factors of variability and uncertainty in an LCA study of nanomaterials – findings from a case study with nano titanium dioxide**. *NanoImpact* 2017, **7**:17–26. https://doi.org/10.1016/j.impact.2017.05.001.
20. Tsoy N, Steubing B, van der Giesen C, Guinée J: **Upscaling methods used in ex ante life cycle assessment of emerging technologies: a review**. *Int J Life Cycle Assess* 2020, **25**. https://doi.org/10.1007/s11367-020-01796-8.
21. van der Giesen C, Cucurachi S, Guinée J, Kramer GJ, Tukker A: **A critical view on the current application of LCA for new technologies and recommendations for improved practice**. *J Clean Prod* 2020, **259**, 120904. https://doi.org/10.1016/j.jclepro.2020.120904.
22. European Commission - Joint Research Centre: *International reference life cycle data system (ILCD) handbook – general guide for life cycle assessment – detailed guidance*. 2010.
23. ** Weber JM, Guo Z, Zhang C, Schweidtmann AM, Lapkin AA: **Chemical data intelligence for sustainable chemistry**. *Chem Soc Rev* 2021, **50**:12013–12036. https://doi.org/10.1039/D1CS00477H. *Shows how different ML parts can play together for early-stage screening.*
24. Voll A, Marquardt W: **Reaction network flux analysis: optimization-based evaluation of reaction pathways for biorenewables processing**. *AIChE J* 2012, **58**:1788–1801. https://doi.org/10.1002/aic.12704.
25. Ulonska K, Skiborowski M, Mitsos A, Viell J: **Early-stage evaluation of biorefinery processing pathways using process network flux analysis**. *AIChE J* 2016, **62**:3096–3108. https://doi.org/10.1002/aic.15305.
26. Zhang D, del Rio-Chanona EA, Shah N: **Screening synthesis pathways for biomass-derived sustainable polymer production**. *ACS Sustainable Chem Eng* 2017, **5**:4388–4398. https://doi.org/10.1021/acssuschemeng.7b00429.
27. Schwaller P, Laino T, Gaudin T, Bolgar P, Hunter CA, Bekas C, *et al.*: **Molecular transformer: a model for uncertainty-calibrated chemical reaction prediction**. *ACS Cent Sci* 2019, **5**: 1572–1583. https://doi.org/10.1021/acscentsci.9b00576.
28. Zhang C, Arun A, Lapkin A: **Completing and balancing database excerpted chemical reactions with a hybrid mechanistic - machine learning approach**. *ChemRxiv* 2023.
29. Wijngaarden VM, Vogel G, Weber JM (in press). Completing partial reaction equations with rule and language model-based methods (in press). Computer aided chemical engineering n.d.

30. Schwaller P, Vaucher AC, Laino T, Reymond J-L: **Prediction of chemical reaction yields using deep learning**. *Mach Learn Sci Technol* 2021, **2**, 015016. https://doi.org/10.1088/2632-2153/abc81d.

31. Gao H, Struble TJ, Coley CW, Wang Y, Green WH, Jensen KF: **Using machine learning to predict suitable conditions for organic reactions**. *ACS Cent Sci* 2018, **4**:1465–1476. https://doi.org/10.1021/acscentsci.8b00357.

32. Wieder O, Kohlbacher S, Kuenemann M, Garon A, Ducrot P, Seidel T, *et al.*: **A compact review of molecular property prediction with graph neural networks**. *Drug Discov Today Technol* 2020, **37**:1–12. https://doi.org/10.1016/j.ddtec.2020.11.009.

33. Gharagheizi F, Mehrpooya M: **Prediction of standard chemical exergy by a three descriptors QSPR model**. *Energy Convers Manag* 2007, **48**:2453–2460. https://doi.org/10.1016/j.enconman.2007.04.005.

34. Weber JM, Guo Z, Lapkin AA: **Discovering circular process solutions through automated reaction network optimization**. *ACS Engineering Au* 2022, **2**:333–349. https://doi.org/10.1021/acsengineeringau.2c00002.

35. Wernet G, Hellweg S, Fischer U, Papadokonstantakis S, Hungerbühler K: **Molecular-structure-based models of chemical inventories using neural networks**. *Environ Sci Technol* 2008, **42**:6717–6722. https://doi.org/10.1021/es7022362.

36. Wernet G, Papadokonstantakis S, Hellweg S, Hungerbühler K: **Bridging data gaps in environmental assessments: modeling impacts of fine and basic chemical production**. *Green Chem* 2009, **11**:1826. https://doi.org/10.1039/b905558d.

37. ** Zhang D, Wang Z, Oberschelp C, Bradford E, Hellweg S: **Enhanced deep-learning model for carbon footprints of chemicals**. *ACS Sustain Chem Eng* 2024, **12**:2700–2708. https://doi.org/10.1021/acssuschemeng.3c07038.

A very recent and relevant demonstration of deep learning in life cycle impact assessment.

38. ** Song R, Keller AA, Suh S: **Rapid life-cycle impact screening using artificial neural networks**. *Environ Sci Technol* 2017, **51**:10777–10785. https://doi.org/10.1021/acs.est.7b02862.

Presents methods to improve interpretability and accuracy of surrogate models for LCA with a focus on mutual information-permutation importance (MI-PI) feature selection method.

39. Sun Y, Wang X, Ren N, Liu Y, You S: **Improved machine learning models by data processing for predicting life-cycle environmental impacts of chemicals**. *Environ Sci Technol* 2023, **57**:3434–3444. https://doi.org/10.1021/acs.est.2c04945.

40. ** Kleinekorte J, Kleppich J, Fleitmann L, Beckert V, Blodau L, Bardow A: **APPROPRIATE life cycle assessment: a PROcess-specific, PRedictive impact AssessmenT method for emerging chemical processes**. *ACS Sustain Chem Eng* 2023, **11**:9303–9319. https://doi.org/10.1021/acssuschemeng.2c07682.

Important work that integrates the process level descriptors into the prediction of life cycle impacts.

41. ** Karka P, Papadokonstantakis S, Kokossis A: **Digitizing sustainable process development: from ex-post to ex-ante LCA using machine-learning to evaluate bio-based process technologies ahead of detailed design**. *Chem Eng Sci* 2022, **250**, 117339. https://doi.org/10.1016/j.ces.2021.117339.

Develops a modular methodological framework for streamlining LCA calculations using machine learning tools to estimate life cycle impact assessment scores at early stages of biorefinery process development. Refers to the prospective problem and illustrates the larger data pipeline and prediction challenge.

42. Igos E, Benetto E, Meyer R, Baustert P, Othoniel B: **How to treat uncertainties in life cycle assessment studies?** *Int J Life Cycle Assess* 2019, **24**:794–807. https://doi.org/10.1007/s11367-018-1477-1.

43. Nam NH, Luong NH: *Nanoparticles: synthesis and applications. Materials for biomedical engineering*. Elsevier; 2019:211–240. https://doi.org/10.1016/B978-0-08-102814-8.00008-1.

44. Cucurachi S, Blanco CF, Steubing B, Heijungs R: **Implementation of uncertainty analysis and moment-independent global sensitivity analysis for full-scale life cycle assessment models**. *J Ind Ecol* 2021. https://doi.org/10.1111/jiec.13194.

45. * Anuoluwa Bamidele E, Olanrewaju Ijaola A, Bodunrin M, Ajiteru O, Martha Oyibo A, Makhatha E, *et al.*: **Discovery and prediction capabilities in metal-based nanomaterials: an overview of the application of machine learning techniques and some recent advances**. *Adv Eng Inf* 2022, **52**, 101593. https://doi.org/10.1016/j.aei.2022.101593.

Provides a broad overview of ML algorithms to predict nanomaterial properties, some of which are determinant of functional performance.

46. Mishra S, Palanisamy P: **Multi-time-horizon solar forecasting using recurrent neural network**. In *2018 IEEE energy conversion congress and exposition (ECCE)*. IEEE; 2018:18–24. https://doi.org/10.1109/ECCE.2018.8558187.

47. Sreekumar S, Bhakar R: **Solar power prediction models: classification based on time horizon, input, output and application**. In *2018 international conference on inventive research in computing applications (ICIRCA)*. IEEE; 2018:67–71. https://doi.org/10.1109/ICIRCA.2018.8597288.

48. Ling C: **A review of the recent progress in battery informatics**. *npj Comput Mater* 2022, **8**:33. https://doi.org/10.1038/s41524-022-00713-x.

49. Piccinno F, Hischier R, Seeger S, Som C: **From laboratory to industrial scale: a scale-up framework for chemical processes in life cycle assessment studies**. *J Clean Prod* 2016, **135**:1085–1097. https://doi.org/10.1016/j.jclepro.2016.06.164.

50. Reuter MA: **The simulation of industrial ecosystems**. *Miner Eng* 1998, **11**:891–918. https://doi.org/10.1016/S0892-6875(98)00078-8.

51. Liu Y, Syberfeldt A, Strand M: **Review of simulation-based life cycle assessment in manufacturing industry**. *Prod Manuf Res* 2019, **7**:490–502. https://doi.org/10.1080/21693277.2019.1669505.

52. Ferdous J, Bensebaa F, Pelletier N: **Integration of LCA, TEA, Process Simulation and Optimization: a systematic review of current practices and scope to propose a framework for pulse processing pathways**. *J Clean Prod* 2023, **402**, 136804. https://doi.org/10.1016/j.jclepro.2023.136804.

53. Cassola S, Duhovic M, Schmidt T, May D: **Machine learning for polymer composites process simulation – a review**. *Compos B Eng* 2022, **246**, 110208. https://doi.org/10.1016/j.compositesb.2022.110208.

54. Sato R, Kutsukake K, Harada S, Tagawa M, Ujihara T: **Machine learning for semiconductor process simulation described by coupled partial differential equations**. *Adv Theory Simul* 2023, **6**. https://doi.org/10.1002/adts.202300218.

55. Lockner Y, Hopmann C: **Induced network-based transfer learning in injection molding for process modelling and optimization with artificial neural networks**. *Int J Adv Des Manuf Technol* 2021, **112**:3501–3513. https://doi.org/10.1007/s00170-020-06511-3.

56. Pylianidis C, Snow V, Overweg H, Osinga S, Kean J, Athanasiadis IN: **Simulation-assisted machine learning for operational digital twins**. *Environ Model Software* 2022, **148**, 105274. https://doi.org/10.1016/j.envsoft.2021.105274.

57. Pauliks N: *A machine learning approach to reduce uncertainty in ex-ante LCA*. Master's thesis. Institute of Environmental Sciences (CML), Leiden University; 2023.

58. Hou P, Jolliet O, Zhu J, Xu M: **Estimate ecotoxicity characterization factors for chemicals in life cycle assessment using machine learning models**. *Environ Int* 2020, **135**, 105393. https://doi.org/10.1016/j.envint.2019.105393.

59. Marvuglia A, Leuenberger M, Kanevski M, Benetto E: **Random forest for toxicity of chemical emissions: features selection and uncertainty quantification**. *J Environ Account Manag* 2015, **3**:229–241. https://doi.org/10.5890/JEAM.2015.09.003.

60. Servien R, Latrille E, Patureau D, Hélias A: **Machine learning models based on molecular descriptors to predict human and environmental toxicological factors in continental freshwater**. *Peer Commun J* 2022, **2**:e15. https://doi.org/10.24072/pcjournal.90.

61. Hou P, Zhao B, Jolliet O, Zhu J, Wang P, Xu M: **Rapid prediction of chemical ecotoxicity through genetic algorithm optimized neural network models**. *ACS Sustainable Chem Eng* 2020, **8**:12168–12176. https://doi.org/10.1021/acssuschemeng.0c03660.

62. ** von Borries K, Holmquist H, Kosnik M, Beckwith KV, Jolliet O, Goodman JM, *et al.*: **Potential for machine learning to address data gaps in human toxicity and ecotoxicity characterization**. *Environ Sci Technol* 2023, **57**:18259–18270. https://doi.org/10.1021/acs.est.3c05300.

Proposes and demonstrates a framework to prioritize toxicity impact characterisation parameters for ML model development based parameter sensitivity and data availability.

63. Engberg N: *Feasibility of machine learning to forecast time-sensitive and spatially explicit characterisation factors*. Institute of Environmental Sciences (CML), Leiden University; 2023.

64. Boulay A-M, Bare J, Benini L, Berger M, Lathuillière MJ, Manzardo A, *et al.*: **The WULCA consensus characterization model for water scarcity footprints: assessing impacts of water consumption based on available water remaining (AWARE)**. *Int J Life Cycle Assess* 2018, **23**:368–378. https://doi.org/10.1007/s11367-017-1333-8.

65. Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, *et al.*: *Global sensitivity analysis*. The Primer. John Wiley and Sons; 2008. https://doi.org/10.1002/9780470725184.

66. Lacirignola M, Blanc P, Girard R, Pérez-López P, Blanc I: **LCA of emerging technologies: addressing high uncertainty on inputs' variability when performing global sensitivity analysis**. *Sci Total Environ* 2017, **578**:268–280. https://doi.org/10.1016/j.scitotenv.2016.10.066.

67. Ravikumar D, Seager TP, Cucurachi S, Prado V, Mutel C: **Novel method of sensitivity analysis improves the prioritization of research in anticipatory life cycle assessment of emerging technologies**. *Environ Sci Technol* 2018, **52**:6534–6543. https://doi.org/10.1021/acs.est.7b04517.

68. Jaxa-Rozen M, Pratiwi AS, Trutnevyte E: **Variance-based global sensitivity analysis and beyond in life cycle assessment: an application to geothermal heating networks**. *Int J Life Cycle Assess* 2021, **26**:1008–1026. https://doi.org/10.1007/s11367-021-01921-1.

69. Antoniadis A, Lambert-Lacroix S, Poggi J-M: **Random forests for global sensitivity analysis: a selective review**. *Reliab Eng Syst Saf* 2021, **206**, 107312. https://doi.org/10.1016/j.ress.2020.107312.

70. * Aouichaoui ARN, Al R, Sin G: *DeepGSA: plant data-driven global sensitivity analysis using deep learning*. 2022:1759–1764. https://doi.org/10.1016/B978-0-323-85159-6.50293-1.

Presents a tool which aims at enabling non-specialist practitioners to leverage Deep Learning-based models for global sensitivity analysis.

71. Jouannais P, Blanco CF, Pizzol M: **ENvironmental Success under Uncertainty and Risk (ENSURe): a procedure for probability evaluation in ex-ante LCA**. *Technol Forecast Soc Change* 2024, **201**, 123265. https://doi.org/10.1016/j.techfore.2024.123265.

72. Kwakkel JH, Jaxa-Rozen M: **Improving scenario discovery for handling heterogeneous uncertainties and multinomial classified outcomes**. *Environ Model Software* 2016, **79**:311–321. https://doi.org/10.1016/j.envsoft.2015.11.020.

73. Kwakkel JH: **The Exploratory Modeling Workbench: an open source toolkit for exploratory modeling, scenario discovery, and (multi-objective) robust decision making**. *Environ Model Software* 2017, **96**:239–250. https://doi.org/10.1016/J.ENVSOFT.2017.06.054.

74. Bryant BP, Lempert RJ: **Thinking inside the box: a participatory, computer-assisted approach to scenario discovery**. *Technol Forecast Soc Change* 2010, **77**:34–49. https://doi.org/10.1016/j.techfore.2009.08.002.

75. Ali R, Hussain A, Nazir S, Khan S, Khan HU: **Intelligent decision support systems—an analysis of machine learning and multicriteria decision-making methods**. *Appl Sci* 2023, **13**, 12426. https://doi.org/10.3390/app132212426.

76. Gaspar-Cunha A, Costa P, Monaco F, Delbem A: **Many-objectives optimization: a machine learning approach for reducing the number of objectives**. *Math Comput Appl* 2023, **28**:17. https://doi.org/10.3390/mca28010017.

77. Häussling Löwgren B, Weigert J, Esche E, Repke J-U: **Uncertainty analysis for data-driven chance-constrained optimization**. *Sustainability* 2020, **12**:2450. https://doi.org/10.3390/su12062450.

78. Gonzalez-Garay A, Guillen-Gosalbez G: **SUSCAPE: a framework for the optimal design of SUStainable ChemicAl ProcEsses incorporating data envelopment analysis**. *Chem Eng Res Des* 2018, **137**:246–264. https://doi.org/10.1016/j.cherd.2018.07.009.

79. Azapagic A, Stamford L, Youds L, Barteczko-Hibbert C: **Towards sustainable production and consumption: a novel DEcision-support framework IntegRating economic, environmental and social sustainability (DESIRES)**. *Comput Chem Eng* 2016, **91**:93–103. https://doi.org/10.1016/j.compchemeng.2016.03.017.