



Advances in application of machine learning to life cycle assessment: a literature review

Ali Ghoroghi¹ · Yacine Rezgui¹ · Ioan Petri¹ · Thomas Beach¹

Received: 7 September 2021 / Accepted: 15 February 2022
© The Author(s) 2022

Abstract

Purpose Life Cycle Assessment (LCA) is the process of systematically assessing impacts when there is an interaction between the environment and human activity. Machine learning (ML) with LCA methods can help contribute greatly to reducing impacts. The sheer number of input parameters and their uncertainties that contribute to the full life cycle make a broader application of ML complex and difficult to achieve. Hence a systems engineering approach should be taken to apply ML in isolation to aspects of the LCA. This study addresses the challenge of leveraging ML methods to deliver LCA solutions. The overarching hypothesis is that: LCA underpinned by ML methods and informed by dynamic data paves the way to more accurate LCA while supporting life cycle decision making.

Methods In this study, previous research on ML for LCA were considered, and a literature review was undertaken.

Results The results showed that ML can be a useful tool in certain aspects of the LCA. ML methods were shown to be applied efficiently in optimization scenarios in LCA. Finally, ML methods were integrated as part of existing inventory databases to streamline the LCA across many use cases.

Conclusions The conclusions of this article summarise the characteristics of existing literature and provide suggestions for future work in limitations and gaps which were found in the literature.

Keywords Life cycle assessment · Machine learning · Optimization · Prediction · Decision-making · ANN · SVM · Random forest · Hybrid methods

1 Introduction

Life Cycle Assessment (LCA) is a series of procedures set for the collection and assessment of the inputs and outputs of materials or energy, as well as the subsequent impacts on the environment incurred due to the running of a system or product throughout that entity's life cycle (ISO 14040.2 Draft). The LCA provides a framework for the definition of the scope, and the goal of the assessment, analysis of the inventory (LCI, life cycle inventory), assessment of the impact (LCIA, life cycle impact assessment), and finally, the interpretation from these procedures (Guinee 2002). The purpose, entities (systems, products) and the degree of sophistication

are defined in the LCA framework's goal and scope definition step. The life cycle inventory (LCI) is the step in which the system boundaries are defined. The key outcome from the LCI is the inventory which collates inputs and outputs to the environment. The life cycle impact assessment (LCIA) is how its relevance expresses the inventory to the impact categories. This step quantifies the impact through weighting and normalization. The interpretation is the final step in which the results from the LCIA are evaluated and used to make recommendations (Guinee 2002). LCA is a vital instrument to help reduce the overall environmental burden and provide insights into upstream and downstream trade-offs associated with environmental pressures, health & wellbeing, and the consumption of natural resources. As such, LCA can inform policy-making by providing valuable information on environmental performance, and thus contributing to performance targets within the Environmental Technology Action Plan (ETAP) and for Energy-using Products within the EuP Directive, in green public procurement (GPP), and in Environmental Product Declarations (EPDs).

Communicated by Matthias Finkbeiner.

✉ Ali Ghoroghi
ghoroghi@cardiff.ac.uk

¹ School of Engineering - Cardiff University, Queen's Buildings, 14-17 The Parade, Cardiff, Wales CF24 3AA, UK

In addition, the recent special report on the impacts of global warming of 1.5°C was yet another call to implement measures to mitigate GHG emissions and to devise new adaptation scenarios (IPCC 2021; Sala et al. 2021). In this context, LCA helps quantify the environmental pressures, the trade-offs, and areas for achieving improvements considering the entire life cycle of built assets from design to recycling. However, current approaches to LCA do not consistently factor in (both in the foreground and background inventory systems) life cycle variations in: (a) building usage, (b) energy supply (including from renewable sources), and (c) building and environmental regulations; as well as other changes over the building/district lifetime (Anand and Amor 2017; Bueno et al. 2016; Skaar and Jørgensen 2013). These include (a) change in the energy mix of a building/district or upgrading/retrofitting the energy system(s) in place; and (b) time-increase of energy demand during the lifetime of a building due to a wide range of reasons, including changes in occupancy patterns.

As such, LCA is an important instrument to help reduce the overall environmental burden of buildings and provide insights into the upstream and downstream trade-offs that are associated with environmental pressures, health and well-being, and the consumption of natural resources. As such, LCA can inform policymaking by providing valuable information on the environmental performance of built assets. However, the current LCA methods and tools face several limitations and challenges, including: (a) site-specific considerations (Bueno et al. 2016), several local impacts need to be considered in building assessments, such as the microclimate; (b) model complexity (Anand and Amor 2017), buildings involve a wide range of material/products, interacting as part of a complex assembly or system; (c) scenario uncertainty (Anand and Amor 2017; Bueno et al. 2016), the long use phase of buildings, including the potential for future renovation, poses uncertainty problems in LCA that are not currently addressed; (d) health and wellbeing (Bueno et al. 2016; Skaar and Jørgensen 2013), traditional LCA methodologies do not address indoor and outdoor environmental impacts on health and well-being; (e) recycled material data

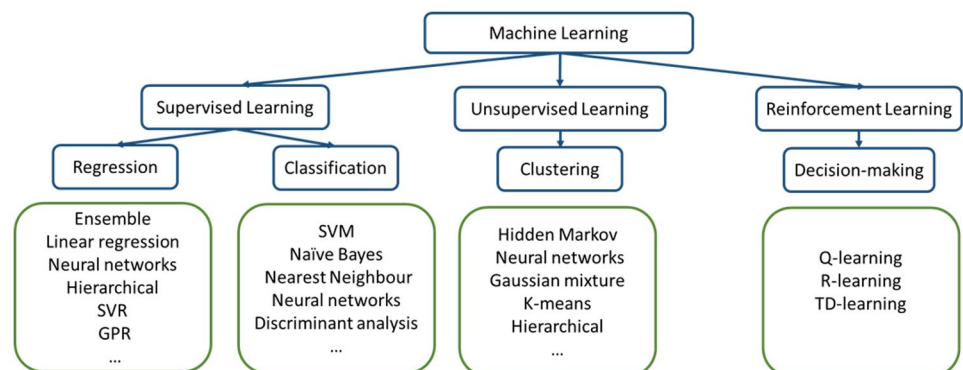
(Anand and Amor 2017; Negishi et al. 2018), lack of data on using waste and recycled materials as new building materials; and (f) lack of consideration for social and economic aspects (Anand and Amor 2017; Negishi et al. 2018).

The sheer number of input parameters and their uncertainties that contribute to the full life cycle make a broader application of ML complex and difficult to achieve. Hence the need to adopt a cartesian, i.e., “Divide and Conquer”, or systems engineering approach, whereby the strategy to reduce and mitigate the environmental impact of a complex artefact, in our case a built asset, should be divided into an ensemble of discrete and manageable scenarios, such as optimizing the energy mix of an energy system. By addressing these discrete scenarios in isolation using ML, a broader reduction of environmental impacts via LCA is feasible.

ML methods are from a subtype of Artificial Intelligence (AI) methods that learn from data to improve their accuracy without the need to be programmed again. ML is creating such a model that can find patterns by studying a set of training data and developing an algorithm without human involvement (Mitchell 1997). ML algorithms are typically categorized into four groups: Supervised learning, Unsupervised learning, Semi-supervised learning, and Reinforcement learning as shown in Fig. 1.

Areas of data science, including ML, are presently used to fill gaps in data for LCA. Furthermore, they have been used to develop accessory tools for LCA that can model and predict a product’s environmental impact based on information from the design phase. ML has the capability of being integrated as a real-time algorithm, assessing production or changes in processes and responding with potential alternatives for better or less environmentally impactful production. ML approaches have been applied to different disciplines of LCA. These include the prediction of missing data, forecasting impact parameters both directly and indirectly, and optimization algorithms in LCA. ML methods have also been used to overcome incompleteness or uncertainty in data to deliver actionable recommendations for the LCA (Algren et al. 2021). One potential advantage of ML in LCA is that it can reduce the cost of data collection. In other words, with ML,

Fig. 1 Various types of ML techniques



the most informative attributes can be identified and focused on collecting them while ignoring other attributes that may not contribute significantly to the model's accuracy. Sometimes missing data in LCA is encountered, and ML can help predict those missing values, improving the data available for LCA. Based on the type of ML, various statistics and visualizations can be used to evaluate the predicted data, including accuracy, confusion matrix, receiver operating characteristic curve, cluster distortion, and means squared error.

ML thrives in applications where there is a requirement to solve mathematical models accurately and efficiently. Consequently, it can be adapted to provide ideas or methods for an optimization process. It can be implemented as part of a real-time decision-making process where potential improvements in the performance of a system throughout its life cycle are identified. Optimization methods can then be applied to the process. This makes it particularly useful in the design process instead of the entire LCA. This study shows that the ML can be coupled with standard optimization methods to increase their capability of quickly exploring promising regions. Figure 2 provides the standard ML and LCA deployment processes which should be considered in the investigation of ML methods in LCA.

The paper reviews the application of ML to LCA with a focus on Buildings, Districts and Cities, while also including several useful related applications described under a "Miscellaneous" heading. A plethora of studies has explored the application of LCA in buildings with most studies focusing on energy use and GHG emissions (Asif 2019; Elkhayat et al. 2020; Lyu and Chow 2020). However, the literature thus far lacks a comprehensive review of the different applications of ML in LCA, the trends in current practices as well as some of the gaps in research. This review aims to address this by proposing and answering research questions which will establish the current practices and future works which are required in this field.

1.1 Goal and scope

This paper aims to investigate the role of ML methods in LCA across three levels;

- Buildings,
- Districts and cities
- Miscellaneous

As such, our review focuses on built assets considered (a) in isolation or (b) within a District or wider city level. Built assets can be of any type, including residential, public, or industrial. At the districts and cities level, the role of ML in human structures like roads, pavements, bridges, parks, railways is investigated, informed by the literature. ML in other related studies like chemical, agriculture, and products is considered and reported at the "Miscellaneous" level. As to the investigated LCA requirements, this review adopts an exploratory approach in that it reports use cases involving the application of ML in Buildings and wider Districts. As such, a bottom-up approach, driven by applications of ML in LCA, has helped identify the most common requirements addressed in the literature. It is worth noting that environmental certification schemes, such as BREEAM, are not considered in this paper. When ML is used at the districts level, the building as an attribute (categorical attribute) can be considered in the model, which will help capture the differences between different building types within the districts. This attribute will have to be tested to see the level of information (accuracy, for example) that it brings to the model.

In this study, fundamental limitations and challenges faced by current ML methods in LCA, applications, motivations, constraints and their role in predictions and optimizations are considered (Fig. 3).

Fig. 2 ML and LCA deployment processes

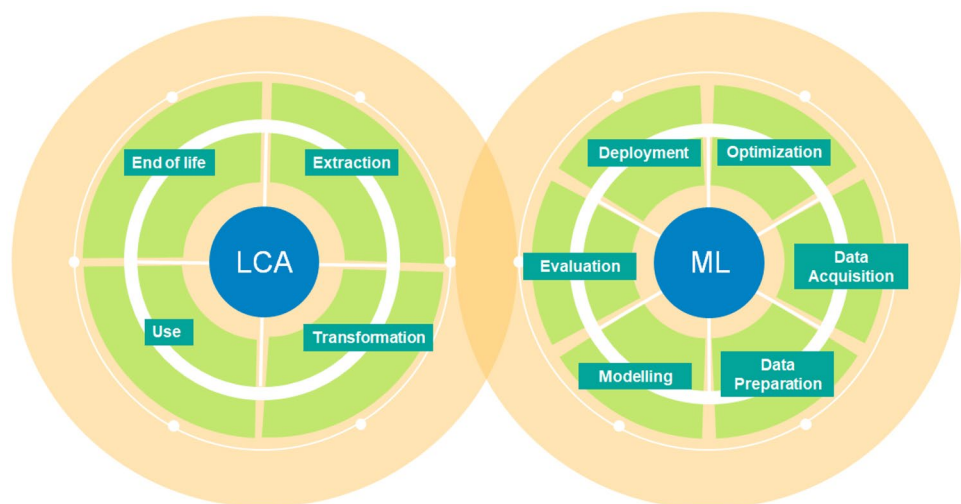
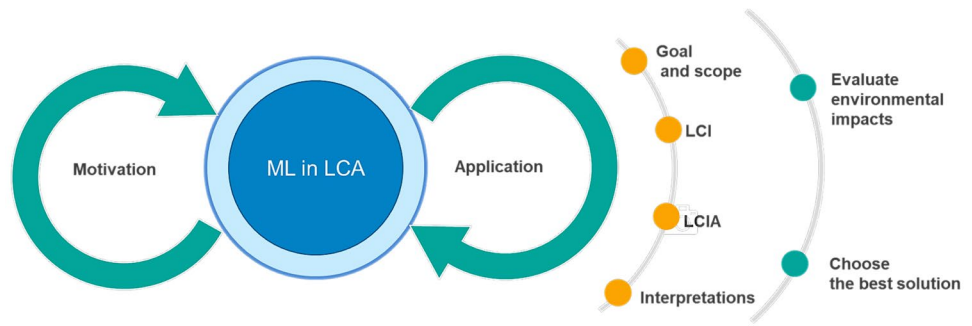


Fig. 3 The applications, motivations, constraints and ML methods in LCA that are considered in this review



The significant contributions of this paper are collating a literature survey to determine use of ML techniques for LCA by answering the following research questions:

1. How has ML been used in LCA?
2. What is the role and efficacy of ML methods in optimization in LCA?
3. Can ML methods integrate and contextualize existing inventory databases to provide a sound basis to streamline the LCA?
4. What are the gaps in research in order to guide future research for ML in LCA?

LCA is explored, and the current state of the art reported in the literature is identified to answer these questions. ML techniques tailored to LCA and specific AI techniques that can advance LCA's establishment and delivery of the smart technology are investigated. Gaps in research will then be identified in order to guide future research for ML in LCA.

The contents of this paper are organized as follows: Sect. 2 lays out the methodology for identifying and including studies for the review. Section 3 discusses research and provides an overview of ML methods in LCA. Section 4 talks about ML and optimization in LCA. The results and discussion are described in Sect. 5. Finally, the findings are evaluated and concluded in Sect. 6.

2 Methodology

A literature review in applying ML in LCA was performed, and 81 relevant studies were analysed according to the research questions. The review presented here aims to identify, evaluate and interpret all available research relevant to LCA using ML models. This section outlines the process for selecting included papers. This methodology was based on five phases.

Planning phase In this phase, scope, literature research questions and databases were determined. Google scholar was chosen for the search database as well as Scopus and Web of Science. Citavi (SWISS ACADEMIC SOFTWARE GMBH, 2021) was used for managing the collected references because of its broad functionality. The publication years of studies were determined to be between the years 2000 and 2021.

Search phase In this step, the search process was developed to select appropriate studies. After defining the research questions in the planning phase, the main terms were defined. Similar terms or interchangeable terms were identified and connected using Boolean OR and AND operators. Table 1 shows the search terms used.

Table 1 Search terms

Main term	Search terms
Life cycle assessment	Life cycle assessment OR Lifecycle assessment OR LCA OR Life cycle impact assessment OR LCIA OR Life cycle analysis OR Life cycle inventory OR LCI
Prediction	Predict OR Predicting OR Prediction
Machine learning	Machine learning OR Deep learning OR ML OR ANN OR NN OR BPNN OR MLP OR SVM OR RF OR KNN OR hybrid. While these search terms included the key ML methods, many more were considered for inclusion from results yielded from these searches.
Optimization	Optimization OR Optimize OR Optimizing OR Optimized OR Optimum OR Optimal OR Decision making

Filtering phase At first, the contents of the papers were assessed through screening of titles and abstracts and the following of inclusion criteria were applied.

1. Language: English
2. Document types: Only full-text, conference or journal papers or books
3. Time interval: The publication years of selected primary studies are between the years 2000 and 2021 to narrow to more relevant results based on current practices in the field of LCA.

After removing the duplicates, the papers selected through their abstract screening were reviewed in full, and those that did not consider ML techniques in LCA or provide primary research findings in this topic were excluded. In the next step, the relevance of a paper based on its introduction and the conclusion/discussion was determined. In total, this yielded 81 primary research papers. These references were imported into our reference manager Citavi.

Evaluation phase In this phase, the articles were assessed for their quality and impact. Three main points were considered for this phase:

1. Is the methodology clear?
2. Are results provided in full?
3. Is the paper relevant to the research questions of this review?

Finally, a decision is made regarding the inclusion of the paper in a full review for this paper. Some papers may have been included for context or interest despite a lack of methodology.

Extraction phase The collected references were managed using Citavi. For each selected paper, relevant information was collected and a justification for each inclusion was noted. Each paper was then analysed and the following information was extracted and recorded: the model used, the optimal model found by the authors, the type of application that ML was targeting in the paper, and finally, the scale at which LCA was applied in this paper.

3 ML methods in LCA

In this section, related works about ML methods and motivation are presented. For ML methods, each studied zones are made bold.

Luque et al. presented a conceptual framework for the integration of AI and LCA. Throughout their study, the relevance of using sensing when addressing an objective of intelligent sustainability in engineering projects has emerged (Luque et al. 2020). Adedeji et al. present a roadmap to

using AI techniques in LCI. The data chain for efficient resident data availability for LCA studies was considered to focus on AI integration. Also, a framework for using AI in LCI was developed (Adedeji et al. 2020).

At the **buildings** level, through the combined use of ML in LCA, it may be possible to significantly reduce environmental impacts (Barros and Ruschel 2021). D'Amico et al. employed ML methods in civil and structural engineering in order to reduce building impacts (D'Amico et al. 2019b). Barros and Ruschel performed a systematic literature review of the scientific research conducted for architecture, engineering and construction industries in the context of LCA and ML (Barros and Ruschel 2021). They show that the most investigated environmental indicators were energy consumption and Global Warming Potential (GWP). Significantly, they found that ML was predominantly used for prediction. In the case of a regionalized bottom-up model created using ML techniques, environmental profiles for individual households were assessed by (Frö melt et al. 2020). At the **districts and cities** level, Manfren et al. presented a review of modelling tools for identifying optimal solutions for district-wide energy systems. They introduced a framework for the key concepts of a local energy management system in an urban area. This framework has a multicriteria perspective and uses ML to find optimal solutions for providing energy services through distributed generation (Manfren et al. 2011). Furthermore, DeRousseau et al. examined the various problem formulations which are commonly seen in the field of concrete mixture design optimization that can necessitate models based on the linear combination, statistics, ML, and physics (DeRousseau et al. 2018). Also, in LCA at the **miscellaneous** level of production, ML algorithms can have an impact in reducing GHG emissions in LCA for geographically differentiated and contextualized design measures; however, they are still underutilized for such applications (Milojevic-Dupont and Creutzig 2021). Kurdi et al. reviewed methods for simulation in tribology to model tribo-contact scenarios and investigated LCA with simulation combined with ML (Kurdi et al. 2020). Wu and Wang reviewed ML methods applied to toxicity prediction and discussed the ML algorithm's input parameter to enhance prediction accuracy (Wu and Wang 2018). Gust et al. demonstrated that in toxicological and regulatory assessment for novel materials where fewer characterization data are available, probabilistic adverse quantitative outcome pathway can leverage using supervised ML models (Gust et al. 2015). In later sections, we discuss the most commonly used ML techniques in LCA.

3.1 Neural networks

Artificial Neural Networks (ANNs), also known as Neural Networks (NNs) or simulated neural networks (SNNs), are

a subset of ML and are at the heart of Deep-Learning algorithms. Their name and structure are inspired by the human brain, mimicking how biological neurons signal to one another (Livingstone 2008). ANNs are favourable as they overcome some limitations commonly seen with traditional software, such as collecting environmental and energy data, physical problem and software language, long computational time, and the need to calibrate a model. Consequently, ANN models provide a superior and more reliable decision support tool for engineers and architects, reducing uncertainties in the LCA field. Furthermore, the implementation of ANN in software can accommodate the development of an appropriate decision support tool. Thus, ML algorithms and techniques may be capable of increasing accuracy in LCA and reducing the simulation time (Sharif and Hammad 2019; Barros and Ruschel 2021; D'Amico et al. 2019a). However, the validity of the NN solution is directly and powerfully proportional to the reliability of the database, which tends to be the most difficult to implement. Ziyadi et al. implemented quantitative uncertainty analysis methods to characterize and quantify uncertainties in a Life Cycle Inventory Analysis (LCIA) model. An ANN model was trained and tested to propagate input variability through a system using interval analysis. Monte Carlo sampling was then used to propagate input uncertainty directly and was compared to an indirect nonlinear optimization method that tries to maximize output range (Ziyadi and Al-Qadi 2019; Barros and Ruschel 2021). At the **buildings** level, ANN mainly was used for optimizing building performance and for impact prediction of energy consumption and GWP. It was suggested that advances in LCA and ML could help calculate and analyze building environmental indicators and develop and improve LCA methods. Shi and Xu presented a systematic LCA method to analyze the environmental performance of construction materials. Furthermore, BPNN and the hybrid algorithm GA-BP were introduced to evaluate building materials. Compared with BPNN, the hybrid GA-BP algorithm was shown to be of better value for selecting construction materials environmentally and has greater precision (Shi and Xu 2009). D'Amico et al. used ANN to simultaneously solve the energy and environmental balance along the building life cycle. The authors developed a decision support tool that quickly and reliably determines buildings' performance with minimum effort. The reliable data and ML combination significantly contribute to the increase in speed and accuracy of LCA (Barros and Ruschel 2021; D'Amico et al. 2019a). The results showed that ANN helps predict energy demand and building LCA (D'Amico et al. 2019a). Considering that the importance of the design phase to carbon emissions during a building's life cycle, Xikai et al. presented a regression model of carbon emissions using designing factors. Also, to determine the designing factors for a predictive model; Multilayer Perceptron (MLP) was

used to develop regression models (Xikai et al. 2019). Sharif and Hammad proposed an ANN model to obtain complex data generated from the simulation-based multiobjective optimization model. This model tried to predict energy consumption to improve buildings' energy performance-critical element of building energy conservation. The outcome of this study showed that the proposed ANN models could efficiently predict the LCA for the whole building renovation scenarios considering the building envelope, HVAC, and lighting systems (Sharif and Hammad 2019). Also, Sharif proposed a simulation-based multiobjective optimization model for optimizing the selection of renovation scenarios for existing buildings by minimizing total energy consumption (TEC) considering LCA. He developed a surrogate ANN for selecting near-optimal building energy renovation methods; and developed deep ML Models to generate renovation scenarios considering TEC (Arani 2020). In the building sector's construction, the material with their embodied energy of all the materials that fall under the main category like wood, cement, plastic and the material that release less energy is provided as input data to the NN (Mukherjee et al. 2019). Płoszaj-Mazurek et al. showed the relationships between the parameters of buildings and the possibility of introducing Carbon Footprint estimation and implementing building optimization at the initial design stage. They used Convolutional Neural Networks (CNN) to analyze an image of the urban layout and consider its influence on the building's Total Carbon Footprint (Płoszaj-Mazurek et al. 2020). Azari et al. investigated the ideal building envelope design using a multiobjective optimization algorithm. This was based on the office building's energy use and life cycle environmental impacts. The input variables for design were insulation material, window type, window frame material, wall thermal resistance and south and north window-to-wall ratios. The optimal iteration of these variables was found to design the building with the smallest possible operational energy and environmental impact. The eQuest 3.65 simulation tool was used to calculate active energy. LCA and Athena IE was used to find an estimated LCA. In addition, an ANN and genetic algorithm (GA) approach were implemented to generate further combinations and find the ideal design iteration. The environmental impact categories included global warming, acidification, eutrophication, formation of air pollution, and ozone depletion (Azari et al. 2016; Barros and Ruschel 2021). Xia et al. introduced a green building assessment index, developed using the life cycle theory and a back-propagation neural network (BPNN), through a Chinese and international building classification system. The assessment index was intended for scientific assessment as the basis for choosing the best plan for green building systems (Xia and Liu 2013; Barros and Ruschel 2021). Oduyemi et al. produced an ANN model for estimating operation and maintenance costs

of buildings (Oduyemi et al. 2015). Life cycle cost analysis (LCCA) compares different design elements, specifications, and materials based on the installation, operation, maintenance and residual costs to evaluate the total life cost of construction. Alqahtani et al. used ANNs to develop a framework for LCCA of construction projects. This was used to estimate the entire cost of construction and uses cost significant items to find the main cost contributions affecting the accuracy of estimation (Alqahtani and Whyte 2013). Wang and Shen created a stochastic Markov model to increase the accuracy of life cycle energy consumption forecasting. This was done by involving longitudinal uncertainties in building conditions, degree days, and valuable life. The Markov building deterioration model was developed using historical data of similar situations and was used to predict the building's useful life and expected condition at any given time. Deterioration of building and temperature changes were used to simulate yearly variation in energy consumption. Energy consumption was estimated with the available data set to calculate annual energy consumption using NN. The proposed stochastic model results in a more restricted distribution, being similar to measured data. It may be implied that the longitudinal uncertainty in the thermal condition of the building and the temperature can account for some uncertainty in the variation of the energy performance (Wang and Shen 2013; Barros and Ruschel 2021). Duprez et al. developed a technique using ML for predicting GWP of building design alternatives with a high coefficient of determination. The original model was compared to three metamodels, Multiple Linear Regression (MLR), Support Vector Regression (SVR) and ANN, to compare their ability to estimate GWP accurately. The authors concluded that ANN offered better results than MLR and SVR (Duprez et al. 2019).

At the **city** level, Perrotta et al. used the application of Boruta Algorithm (BA) and NN to evaluate and calculate a fleet of trucks' fuel consumption to estimate the emissions for pavement roads. The authors showed that NN is appropriate for analyzing data from fleet and road asset management databases. The resulting NN model was used to estimate the impact of rolling resistance parameters (pavement roughness and macrotexture) on fuel consumption (Perrotta et al. 2018). Furthermore, Perrotta et al. used truck telematics, road geometry and condition data to investigate the fuel consumption prediction of fleets of trucks. Three ML techniques, Support Vector Machine (SVM), Random Forest (RF) and ANN, were developed and compared in performance (Perrotta 2017).

In the **miscellaneous** level Wisthoff et al. studied the relationship between product design decisions and eventual LCA. Their study developed a search tree of sustainable design knowledge in the early design phase, and to assist in quantifying the impact of these design decisions;

the study used an MLP method to relate the LCA of 37 case study products to product attributes to help the designer to redesign the product to reduce the impact (Wisthoff et al. 2016). Smetana et al. focused on analyzing evolutionary similarities and differences between two complex modular systems, NN and blockchain technologies, on evaluating their potential for application to material flow analysis (MFA) and LCA. The authors concluded that the combination of NN and blockchain could form a more efficient system for MFA and LCA (Smetana et al. 2018). Chiang et al. introduced a design for environment methodology to evaluate derivative consumer electronic product development using a BPNN model and a technique for order preference by similarity to ideal solution (TOPSIS) method (Chiang et al. 2011). Zhu et al. presented a research framework for greening the continuous sitagliptin manufacturing process with LCA and NN's aid. Deep learning NN models were developed to predict LCA according to the chemicals in a database with known LCA values and corresponding molecular descriptors (Luque et al. 2020). Li et al. developed an ANN approach to estimate unknown eco-indicators for missing environment impact information for several vital materials used in electronic products and integrate recycling scenarios in LCA (Li et al. 2008). The result showed that the ANN-based approach was accurate enough in forecasting the missing materials. Kaab et al. employed two ANNs and an adaptive neuro-fuzzy inference system (ANFIS) model for predicting LCA and output energy of sugar cane production (Kaab et al. 2019). Romeiko et al. presented a model for estimating LCA spatially at the county scale, with corn production developed by applying ANN (Romeiko et al. 2020a). For the cost estimation of a product's life cycle in the product design process, Leszczyński and Jasiński used ANNs and compared them with a parametric estimation (Leszczynski and Jasinski 2020). Marvuglia et al. developed an automatic selection strategy using combinations of a General Regression Neural Network (GRNN) and a set of linear models, based on partial least squares (PLS) regression for USEtox factor. The authors found that linear models have lower predictive power (prediction of toxicity factors) compared to GRNN nonlinear model (Marvuglia et al. 2015; Barros and Ruschel 2021). Song et al. developed ANN models to estimate the LCA of chemicals in the market. Using molecular structure information, they trained multilayer ANNs for life cycle impacts of chemicals using six impact categories. The application domain (AD) of the model was estimated for each impact category within which the model exhibits higher reliability (Song et al. 2017). Also, Song continued an attempt to harness the power of ML techniques to address the data deficiencies in LCA and an ANN, and Random Forest predictive models were

developed to estimate approximate life cycle impacts of chemicals (Song 2019). Li et al. used nine molecular fingerprints to describe pesticides, binary and ternary classification models constructed to predict aquatic toxicity of pesticides via six machine learning methods: ANN, Naïve Bayes (NB), K-Nearest Neighbours (KNN), Classification Tree (CT), RF and SVM (Li et al. 2017). Amini Toosi et al. explored the possibility of an ANN-based LCA model for the conceptual design phase by classifying products according to their environmental and product characteristics. The product classification ultimately identified was used to create classification schemes with the C4.5 decision tree algorithm. An ANN-based approach with product attributes as inputs and environmental impact drivers as outputs were developed to predict the approximate LCA of grouping members. The predicted results seemed to be satisfactory (Seo et al. 2005). Cornago et al. introduced a model which resembles the deep neural network (DNN) to forecast the hourly day-ahead electricity consumption in an LCA aware scheduling system. This information allows to schedule the production to minimize the LCA impacts relative to the electricity consumption. (Cornago et al. 2020). Understanding and developing the LCA of activated carbon produced from diverse biomass feedstocks is critical and time-consuming for biomass screening and process optimization for sustainability. Liao et al. addressed this problem by developing a high accuracy ANN model and kinetic-based process simulation to estimate primary energy consumption and GHG emissions across various woody biomass (Liao et al. 2020). Nabavi-Pelesaraei et al. used historical data to predict future agricultural energy, and they showed that agricultural energy output and its LCA could be readily predicted by ANN (Nabavi-Pelesaraei et al. 2018). Sousa et al. proposed an ANN model using product attributes, which are characteristics of product concepts, and environmental inventory data from pre-existing LCAs. The product design team then use the new high-level attributes to obtain LCA for a new quickly product concept (Sousa et al. 2000). Also, Sousa and Wallace developed an ANN-based learning surrogate in approximate LCA of product design concepts (Sousa and Wallace 2006). Kleinekorte et al. proposed a predictive LCA framework of chemicals using ANN networks. The results show that the proposed ANN was able to predict whether a technology change has the potential to reduce climate change impacts (Kleinekorte et al. 2019b). Park and Seo proposed a BPNN model for an approximate LCA for the conceptual design phase by classifying products according to their environmental and product characteristics. For approximate LCA, the product attributes and environmental impact drivers (EID) were identified to predict the environmental impacts of products. The results

showed BPNN is more accurate than multiple regression analysis in the prediction of the results of LCA (Park and Seo 2003). Milczarski et al. applied ANN to validate the production process's quality and parameters in the food processing industry (Milczarski et al. 2020).

3.2 Support vector machines

Support vector machines (SVMs) have been scarcely involved in LCA. SVM is an ML algorithm based on a theory proposed by Vapnik called the statistical learning theory. It has proven to have unique advantages when working with smaller samples, nonlinear and high dimensional pattern recognition and can also be used in conjunction with other ML problems such as function fitting. SVM aims to solve the optimization problem and to find the optimal classification hyperplane in the high-dimensional feature space in order to work with complicated data classification (Cortes and Vapnik 1995).

At the **buildings** level, Shan et al. explored ML-based electroencephalogram (EEG) methods in the human-computer interaction domain for a potentially more accurate and objective human-building interaction. The machine learning-based EEG methods can be the primary feedback mechanism of wellbeing and performance to the building life cycle platform. Linear discriminant analysis (LDA) and SVM machine learning classifiers were demonstrated. Together with EEG indices, these two ML-based EEG methods can be the primary feedback mechanism of wellbeing and performance to the building (Shan et al. 2017). Liu et al. proposed a methodology that couples multiobjective optimization and SVM and decision tree classifiers to extract design heuristics (Comfort temperatures, etc.). The methodology has been demonstrated on sustainable residential system design via Techno-Ecological Synergy in LCA (TES-LCA) methodology (Liu and Bakshi 2018).

At the **districts and cities** level, Perrotta et al. presented the application of SVM to fuel consumption modelling of articulated trucks for a large dataset. Again, SVM demonstrated a good level of accuracy (Perrotta 2017).

At the level of **miscellaneous**, Hou et al. compared the performance of SVM beside other ML models with the performance of the Ecological Structure-Activity Relationships (ECOSAR) model. This is proven to be the best model among several existing aquatic ecotoxicity QSAR tools and linear regression models for estimating HC50 values of chemicals based on their physical-chemical properties and their classification of the mode of action (Hou et al. 2020). Pradeep Kumar et al. developed an SVM model to delineate vanadium-derived strengthening effects in HSLA steels in the field of production. In addition, they created a ML model to predict the yield strength of V-HSLA steels. Materials savings are translated to embodied energy and

carbon savings using LCA databases in a life cycle inventory process, subtracting the costs incurred in the production of vanadium feedstock (Pradeep Kumar et al. 2021). Li et al. used SVM to predict the aquatic toxicity of pesticides and develop a tool for an early evaluation of aquatic pesticide toxicity in environmental risk assessment. They found that SVM exhibited high accuracy (Li et al. 2017). Romeiko et al. compared the SVM and Gradient Boosting Regressor (GBR) model for estimating spatially explicit life cycle global warming and eutrophication, with corn production. The results indicated that the GBR model built with monthly weather, features yielded higher predictive accuracy for life cycle, global warming impact, and life cycle EU (Romeiko et al. 2019). Milczarski et al. applied SVM, ANN, RF, KNN and C4.5 to validate the production process's quality and its parameters in the food processing industry. The results showed that using the RF algorithm had the best results of processes classification (Milczarski et al. 2020).

3.3 Random forest

Random forest is a type of supervised learning algorithm. It is a collection of decision trees, each trained with the “bagging” method. The principle of the bagging method is that combining learning models can improve the outcome (Breiman 2001). This ML algorithm has been relatively well-used in the LCA due to its high predictive accuracy and its built-in variable importance measures (Hou et al. 2020; Hou 2019).

At the **buildings** level, Xikai et al. applied RF beside three regression techniques to develop regression models of carbon emissions to predict designing factor during the building's life cycle (Xikai et al. 2019). Frömel used RF, KNN and LASSO-Regression to attribute missing water supply, electricity, and heating information. The predicted data were then converted to quantities using price data. Household budget survey finds the existence of similar socio-economic household archetypes in consumption. These archetypes diverging from general macro-trends suggest that the proposed approach may be beneficial in improving understanding of consumption and informing policymakers' future decisions for impactful environmental measures targeting specific consumer groups (Frömel et al. 2018; Frömel 2018). DeRousseau et al. applied RF as the best method between various ML methods like regression models and for predicting concrete compressive strength for field concrete mixtures given the model performance metrics in the field of concrete mixture design optimization (DeRousseau 2020).

At the **districts and cities** level, Perrotta et al. developed an RF model beside other ML algorithms to investigate the fuel consumption prediction of large fleets of trucks based on truck telematics and road geometry and condition data.

The study also shows that although all three methods make it possible to develop models with good precision, the RF slightly outperforms SVM and ANN (Perrotta 2017).

At the **miscellaneous** level, Cheng et al. assessed the impacts of different combinations of feedstocks and pyrolysis conditions on climate change, energy, and economic performance. First, they built an RF model to predict the yields and characteristics of biochar for selected feedstocks at varied pyrolysis conditions. Then, they applied LCA and financial analysis to RF model outputs to determine GWP, energy return on investment (EROI), and minimum product selling price (MPSP) of biochar (Cheng et al. 2020a). Also, Cheng et al. evaluated the energy, climate change, and economic performance of slow pyrolysis of multiple feedstocks under various processing conditions via the integration of RF, LCA, and financial analysis. The results showed this integration is helpful for efficiently evaluating many possible pyrolysis systems producing biochar to sequester atmospheric CO₂ (Cheng et al. 2020b). Also, Cheng et al. evaluates the feasibility of hydrothermal treatment (HTT) with carbon capture and storage (CCS) as energy-producing negative emissions technology (NET) and compares such system with traditional bioenergy with carbon capture and sequestration (BECCS) system. RF was developed to predict product yields and characteristics from HTT of various feedstocks. The model results were then integrated into an LCA model to compute two metrics EROI and GWP. Results showed that RF models had better prediction accuracy than regression tree and multiple linear regression models for HTT of feedstocks and predicted the mass yields of various products and the energy and carbon contents of biocrude and hydrochar (Cheng et al. 2020a). Rojek and Dostatni used RF beside some ML methods as modelling tools supporting selecting materials in ecodesign (Rojek and Dostatni 2020). Gu developed an LCA model to reduce the life cycle environmental impacts of metal-organic frameworks; he combined a conventional LCA with RF and yielded some preliminary heuristics for sustainable design of metal-organic frameworks with some life cycle impact (Gu 2018). Beyond LCA, Hou developed an RF model in chemical risk management to predict the ecotoxicity of new chemicals or as a screening process to identify chemicals with high predicted ecotoxicity potential to further test in priority (Hou 2019). Milczarski et al. applied RF to validate the production process's quality and its parameters in the food processing industry. The results showed that using the RF algorithm had the best results of processes classification (Milczarski et al. 2020).

3.4 Hybrid and ensemble ML techniques

The use of ML methods, including singles, ensembles, and hybrids, have been dramatically increasing. Hybrid methods

combine at least two ML and soft computing methods to achieve superior outcomes. Ensemble methods use a series of ML classification trees as opposed to one. By doing so, the accuracy of the model is significantly increased. Ensemble methods are categorized as supervised learning algorithms. Ensemble methods increase the training. The ensemble method allows for different training algorithms, making training more flexible. Kishk et al. proposed an integrated life cycle costing (LCC) that utilizes statistics, fuzzy set theory, and ANNs to deal with incomplete information, human judgment, and uncertainty. The authors claim that these models should also provide estimates from different levels of data, and information availability (Kishk and Al-Hajj 1999).

At the **buildings** level, Feng et al. developed a quantitative method using fuzzy C-means clustering and an extreme learning machine (FCM-ELM) for assessing buildings' environmental performance in early decision stages, considering uncertainty associated with complex design decisions. The results show that the model is at least as reliable and accurate as the Monte Carlo methodology (Feng et al. 2019). Also, Feng developed an LCA method that integrated discrete-event simulation and process-based LCA using the Bayesian regularization back-propagation neural networks (BRBNN), RT, ensemble learning (EL) and ELM algorithms to extract knowledge about the relationships between construction planning and project performance (Feng 2020). Azari et al. proposed a hybrid ANN and GA approach as the optimization technique to explore optimum building envelope design concerning energy use and LCA in a low-rise office building. The categories within the LCA were global warming, acidification, eutrophication, smog formation, and ozone depletion (Azari et al. 2016). In the context of building material properties, Shi et al. considered a systematic method derived from LCA theory to analyze the green performance of construction materials. The authors proposed a BPNN and GA-BP hybrid algorithm to evaluate green building materials. They showed that with BPNN, the GA-BP hybrid algorithm is favourable for selecting green building materials and achieves higher accuracy (Shi and Xu 2009). Wang et al. introduced a Markov chain based stochastic approach and an ANN model to project periodic energy consumption distribution for each joint energy state of building condition and temperature. Comparing the traditional deterministic model and the developed model shows that the proposed model improved the result (Wang and Shen 2013). Duprez et al. combined Sobol Sensitivity Analysis (SA) and an ANN to building LCA. The Sobol method displayed satisfactory results with the computation of quantitative indices. SA was used in the ANN training, and the subsequent model predicted the GWP of new design alternatives. It was able to do this in a time-efficient manner and with a coefficient of determination higher than 0.9 (Duprez et al. 2019).

For **miscellaneous** uses, Kleinekorte et al. proposed a fully automated framework, including selecting suitable subsets of descriptors, called feature selection and optimization of the network architecture. They used a GA to determine the optimal network architecture and an ANN to predict the environmental impact for a given chemical. The results show that the environmental impact is expected correctly, and the framework can serve as an initial screening tool for identifying environmentally beneficial process alternatives (Kleinekorte et al. 2019a). Lysenko et al. proposed a method that the gradient-boosted classifier tree ensemble model (GBM) is chosen for the small number of positive (toxic) drugs in a training dataset with missing values. The ML leverages the identity of drug targets and off-targets, functional impact score computed from Gene Ontology annotations, and biological network data to predict drug toxicity (Lysenko et al. 2018). Li et al. introduced a modular Scorecard-based LCA architecture with a Bayesian Network (BN). The energy consumption is assessed by an overall modular Scorecard-based LCA architecture embedded with a BN energy prediction model. Seo and Kim proposed a hybrid GA and NN model for an approximate LCA. The GA was employed as an optimization method of relevant feature selection, determining the number of hidden layers and processing elements. For approximate LCA, the product attributes and environmental impact drivers (EID) were identified to predict the environmental impacts of products (Seo and Kim 2007). The results show that the hybrid model improves the prediction accuracy of the BN model, and the BN is suitable for small data sets (Li et al. 2017). Zhou et al. proposed integration of ANN with GA to optimize the multiobjective function of material selection in product design considering LCA (Zhou et al. 2009).

3.5 Other types of ML

Slapni et al. presented a framework that used Weka 3.6.10, a JAVA program package for machine learning algorithms to predict the missing characterization factors (CFs) in environmental interventions to reduce deviation from the European Union normalization factors (EU NFs) and a nominated regional NFs to calculate LCA (Slapnik et al. 2014).

At the **buildings** level, Duprez et al. proposed a method for predicting GWP of building design alternatives with a high coefficient of determination. The authors used MLR, SVR and ANN. MLR and SVR performed poorly when predicting new values as they could not cope with complexity as for MLR or were prolonged as for SVR models. (Duprez et al. 2019). Xikai et al. presented a study on the regression model of carbon emissions in residential buildings using designing factors. Four regression techniques, Principal Component Analysis (PCR), RF, MLP and SVR, were used to develop regression models, and the results show that SVR had the optimal predictive power (Xikai et al. 2019). Shan used LDA and SVM machine learning classifiers to

established EEG based methods to improve human-building interaction in the indoor environment and use them in a building LCA platform (Shan et al. 2017). Płoszaj-Mazurek et al. introduced a study of regenerative design guidelines for parametric modelling of building designs with calculated total Carbon Footprint. They used the GBR model to predict optimal building features and the CNN to predict the total carbon footprint of a building design based on fundamental building features and the urban layout. The results of multicriteria analyses showed relationships between the parameters of buildings and the possibility of introducing carbon footprint estimation and implementing building optimization at the initial design stage (Płoszaj-Mazurek et al. 2020). Østergaard used an MLR model to estimate more accurate lifespans, which can help to reduce the uncertainty of sustainability assessments of buildings in LCA. The regression model proved to estimate the lifespan with lower errors than the general approach relying on a single fixed value for all building locations, uses and building materials (Østergaard et al. 2018). Feng developed an LCA method that integrated discrete-event simulation and process-based LCA. The BRBNN, RT, EL and ELM algorithms were used to extract knowledge about the relationships between construction planning and project performance (Feng 2020).

At the **districts and cities** level, Alam used multiple linear regression, polynomial regression, decision tree regression and support vector regression models using calculated CO₂ emission as a response variable for the LCA model for different phases of the pavement life cycle. The models determined the significant contributors and quantified the CO₂ emission in pavement material production, initial construction, maintenance and use phase; they found that SVM and ANN performed better than other methods (Alam 2020). Renard et al. developed a reinforcement learning (RL) decision support tool that minimizes the global warming impacts of a pavement system over its life cycle. Renard et al. presented an approach to LCA modelling that implements a reinforcement learning algorithm called Q-learning, which helps decision-makers account for several sources of uncertainty in pavement infrastructure (Renard et al. 2021a).

In the context of **miscellaneous** applications, Romeiko et al. used the boosted regression tree (BRT) model to identify the leading contributors among soil, weather, and farming practice parameters affecting the life cycle impacts in Soybean Production. The authors used a combination of Environmental Policy Integrated Climate and process-based LCA models to quantify life cycle GWP, EU and acidification (AD) impacts. BRT has been used in discovering the driving factors for spatial and temporal trends in transportation, public health, and other disciplines (Romeiko et al. 2020b). Also, Romeiko et al. compared the predictive accuracies of SVR, linear regression (LR), ANN, gradient boosted regression tree (GBRT), and extreme gradient boosting (XGBoost) for estimating spatially

explicit LCA at the county scale, with corn production in a case study. The results indicated that the GBRT model yielded the highest predictive accuracy with cross-validation (CV) values of 0.8 for the life cycle GW impacts (Romeiko et al. 2020a). Bui and Perera proposed a decision support framework comprising the life cycle cost analysis and advanced data analytics based on Gaussian Mixture Models (GMM) with the expectation-maximization (EM) algorithm for data clustering. GMM is a case of an unsupervised learning algorithm in which GMM is a probabilistic-model technique for distributing data into different clusters by Gaussian distributions. This framework prepared an intelligent decision support tool for ship owners to achieve optimized vessel performance and comply with stringent environmental regulations (Bui and Perera 2020). Hamrol et al. presented an integrated eco-design of products and technological processes, ensuring the appropriate selection of materials and connections from the point of view of recyclability. The method was implemented in an expert system using the classification method decision tree induction as the classification method. The expert system offers a practical solution that makes it possible to change material or connection without consulting the product designer. Moreover, it is consistent with concurrent engineering design (Dostatni et al. 2018). Hou et al. used KNN, SVM, ANN, RF, Adaptive boosting (AdaBoost) and Gradient boosting machine (GBM) for estimating HC50 values of chemicals based on their physical-chemical properties and their classification of the mode of action. Among the machine learning models, RF had the best predictive performance (Hou et al. 2020). Cheng et al. used the MLR, regression tree (RT), and RF to predict product yields and characteristics from HTT of various feedstocks. The model results were then integrated into an LCA model to compute EROI and net GWP. Results showed random forest models had better prediction accuracy than regression tree and multiple linear regression to model HTT of feedstocks (Cheng et al. 2020a). Rojek and Dostatni compared the effectiveness of RBF networks, Kohonen networks, and RF as modelling tools supporting selecting materials in ecodesign and showed that ML methods effectively supported selecting materials in ecodesign. The study has proven ML methods to be highly useful and effective in selecting materials in designed products (Rojek and Dostatni 2020). Gu used a built-in decision tree model (ID3) package coupled with conventional LCA to speed up understanding metal-organic frameworks based via connecting the LCA results with ML technique (Gu 2018). Lee et al. developed a rapid predictive model to quantify life cycle GW and eutrophication (EU) impacts of corn production using the BRT model to estimate future life cycle environmental impacts of corn production (Lee et al. 2020). Nabavi-Pelesaraei et al. conduct energy output and environmental impact prediction of paddy production on ANN and adaptive neuro-fuzzy inference system (ANFIS). According to the results, multi-level ANFIS

is chosen as a better model than ANN models due to higher computation speed, and higher accuracy (Nabavi-Pelesaraei et al. 2018). Ma and Kim (Ma and Kim 2015) presented an algorithm, predictive usage mining for life-cycle assessment (PUMLCA). This displayed a higher forecasting accuracy when data had complexity. Through modelling usage patterns, trend, seasonality and level, predictive LCA was performed for agricultural machinery in real-time. This showed an accurate estimate of environmental impact (Barros and Ruschel 2021). SAAB presents a proposed LCA calculator for implementing an efficient LCA computation; they used Spark MLlib, a library built on Apache Spark, MPI and OpenMP for LCA algorithms. The results showed that the combination of MPI/OpenMP provided much better performance for computing algorithms than Spark MLIB in LCA (Saab 2019). Abdella et al. presented a framework integrating the economic input-output LCA with logic regression and k-means clustering to deal with multiple decision-making units in food consumption categories and sustainability indicators (Abdella et al. 2020). Olafasakin et al. developed a Kriging-based reduced order model (ROM) to predict pyrolysis yields of feedstock samples based on the output of a detailed chemical kinetic pyrolysis mechanism for assessing the costs and emissions of a pyrolysis biorefinery (Olafasakin et al. 2021).

KNN classification is one of the most fundamental and straightforward classification models in traditional supervised learning. Consequently, it is often one of the first choices for a classification study when it is tiny or no prior knowledge about the data distribution (Peterson 2009; Frömelt et al. 2018; Hou et al. 2020). Hou et al. proposed three data-driven frameworks to estimate the missing data in LCA. The results show that KNN models have better prediction performance than ECOSAR and linear regression models for estimating some parameters for chemicals in USEtox (Hou et al. 2020). Serajiantehrani used KNN and MLR, decision tree regression, and gradient boosting regression methods for the complete construction and environmental costs of trenchless cementitious spray-applied pipe linings, cured-in-place pipe with polyester resin, and sliplining with high-density polyethylene pipe methods by evaluation and analysis of the construction and environmental costs based on the actual data. The results show that Multi-linear Regression had the optimal predictive (Serajiantehrani 2020). Milczarski et al. applied KNN and C4.5 to validate the production process's quality and parameters in the food processing industry (Milczarski et al. 2020). Romeiko et al. compared the SVM and GBR model for estimating spatially explicit life cycle global warming and eutrophication, with corn production. The results indicated that the GBR model built with monthly weather, features yielded higher predictive accuracy for life cycle, global warming impact, and life cycle EU (Romeiko et al. 2019).

In this section, applying ML in LCA is explored, and the current state of the art reported in the literature is identified to answer the above questions. ML techniques tailored to LCA and specific AI techniques that can advance LCA's establishment and delivery of the smart technology are investigated. Table 2 shows details of the literature survey on ML methods in LCA. The papers are divided into three types of prediction impact (PI), decision making (DM) and literature review (LR). The applied method in each paper is identified, and the scale at which the model was applied is shown. The majority of studies identified in this review were for impact prediction, but many had multiple objectives and often incorporated decision making. Many niche applications were also found in this review and the discussed studies show the adaptability of ML techniques for LCA.

4 ML and optimization in LCA

In its most basic form, LCA does not always include a systematic way of optimizing alternatives for environmental impacts mitigation. Combining LCA with ML methods may be a useful way of generating optimized process alternatives as part of an LCA (Wallace et al. 2014). ML models can perform faster and with lower storage requirements when estimating model outputs than other traditional process-based models. They are also more flexible when being integrated into other processes and simulation platforms. These allow ML models to attempt more runs of a simulation and achieve better outcomes for a range of computationally demanding tasks. These include optimization, prediction, and validation. Also, ML models can be fine-tuned by altering trainable parameters through an optimization procedure. LCA can be used to assess technological solutions from an environmental perspective. In conjunction, ML can be used as an optimizer alone or combined with other optimization algorithms to find the best solution according to constraints in LCA.

Luque et al. developed a conceptual framework for the integration of AI and LCA. The study focused on the sensorization of industrial plants and the treatment of data through ML algorithms in the field of sustainability optimization (Luque et al. 2020). Ziyadi et al. developed an ML surrogate model to perform direct Monte Carlo sampling as well as indirect nonlinear optimization to provide grounds for objective model uncertainty analysis for LCA applications (Ziyadi and Al-Qadi 2019).

At the **buildings** level, Sharif and Hammad developed an ANN model to analyze renovation scenarios to minimize total energy consumption in LCC and LCA. They developed a set of data to represent renewal scenarios from results obtained by Simulation-Based Multi-Objective Optimization (SBMO). ANNs were developed as surrogate models

Table 2 Summary of literature survey on ML in LCA for included papers

Ref-ID	Individual / ensemble models	The best model	Types of Papers			Scale		
			PI	DM	LR	Buildings	Districts & cities	Miscellaneous
(Shi and Xu 2009)	BPNN ^a , GA-BP ^b	GA-BP		✓		✓		
(Romeiko et al. 2020b)	BRT ^c		✓					✓
(Cheng et al. 2020a)	RF ^d		✓					✓
(Xia and Liu 2013)	BPNN		✓			✓		
(Wisthoff et al. 2016)	MLP ^e		✓					✓
(Pradeep Kumar et al. 2021)	SVM ^f		✓			✓	✓	
(Lee et al. 2020)	BRT		✓					✓
(Kurdi et al. 2020)	General ML				✓			✓
(Saab 2019)	Spark MLlib ^g		✓					✓
(Manfren et al. 2011)	General ML			✓			✓	
(Azari et al. 2016)	ANN-GA		✓			✓		
(Ziyadi and Al-Qadi 2019)	ANN		✓				✓	
(Ploszaj-Mazurek et al. 2020)	CNN ^h		✓			✓		
(Gu 2018)	RF, ID3 ⁱ , naive Bayes	ID3	✓					✓
(Rojek and Dostatni 2020)	RF, RBF networks ^j , Kohonen networks	Kohonen networks	✓					✓
(Barros and Ruschel 2021)	General ML				✓	✓		
(Marvuglia et al. 2015)	GRNN ^k , MCUVE-PLS ^l , Random Frog,	GRNN		✓			✓	
(D'Amico et al. 2019b)	General ML				✓	✓		
(Gust et al. 2015)	General ML				✓			✓
(Wang and Shen 2013)	Markov - ANN		✓			✓		
(Cheng et al. 2020b)	RF, MLR ^m , RT ⁿ	RF	✓					✓
(Duprez et al. 2019)	Sobol SA ^o - ANN		✓			✓		
(Shan et al. 2017)	SVM, LDA ^p , NB ^{ai} , KNN	LDA, SVM	✓			✓		
(Hou et al. 2020)	ANN, SVM, RF, KNN ^q , AdaBoost ^{al}	KNN	✓					✓
(Cornago et al. 2020)	DNN ^r		✓				✓	
(Serajiantehrani 2020)	KNN, MLR, DTR ^s , Gradient Boosting	MLR	✓			✓	✓	
(Sharif and Hammad 2019)	ANN			✓		✓		
(Hou 2019)	RF, ANN - GA	RF	✓					✓
(DeRousseau et al. 2018)	General ML		✓			✓	✓	
(Xikai et al. 2019)	MLP, RF, PCR ^t , SVR ^x	SVR	✓			✓		
(Leszczynski and Jasinski 2020)	MLP		✓					✓
(Feng et al. 2019)	FCM-ELM ^y		✓			✓		
(D'Amico et al. 2019a)	ANN		✓	✓		✓	✓	
(Zhu et al. 2020)	ANN		✓					✓
(Adedeji et al. 2020)	General ML			✓				✓
(Dostatni et al. 2018)	General ML		✓	✓				✓
(Luque et al. 2020)	General ML				✓			✓
(Kleinekorte et al. 2019a)	ANN-GA		✓					✓
(Perrotta et al. 2018)	BA ^u - NN		✓				✓	
(Bui and Perera 2020)	GMM ^z			✓				✓
(Li et al. 2017)	BN ^{taa}			✓				✓
(Ma and Kim 2015)	PUMLCA		✓	✓				✓

Table 2 (continued)

Ref-ID	Individual / ensemble models	The best model	Types of Papers			Scale		
			PI	DM	LR	Buildings	Districts & cities	Miscellaneous
(Renard et al. 2021b)	Q-learning		✓				✓	
(Nabavi-Pelesaraei et al. 2018)	ANN, ANFIS ^{ab}	ANFIS	✓				✓	
(Liao et al. 2020)	ANN		✓				✓	
(Slapnik et al. 2014)	Weka ^{ac}		✓				✓	
(Mukherjee et al. 2019)	ANN		✓			✓		
(Oduyemi et al. 2015)	ANN		✓			✓		
(Romeiko et al. 2020a)	ANN, LR ^{ad} , SVR, GBRT ^{ae} , XGBoost ^{af}	GBRT	✓					✓
(Kaab et al. 2019)	ANN, ANFIS	ANN	✓					✓
(Li et al. 2008)	ANN		✓					✓
(Chiang et al. 2011)	BPNN		✓					✓
(Lysenko et al. 2018)	GBM ^{ag}		✓					✓
(Kishk and Al-Hajj 1999)	ANN-FST ^{ah}		✓			✓		
(Smetana et al. 2018)	General ML		✓					✓
(Perrotta 2017)	ANN, SVM, RF		✓				✓	
(Seo et al. 2005)	ANN		✓					✓
(Alqahtani and Whyte 2013)	BPNN		✓			✓		
(Østergaard et al. 2018)	MLR		✓			✓		
(Hou et al. 2018)	Similarity-based link prediction		✓				✓	
(Li et al. 2017)	ANN, SVM, RF, NB, KNN, CT ^{aj}	SVM, ANN	✓					✓
(Frörmelt et al. 2020)	General ML		✓				✓	
(Wu and Wang 2018)	General ML		✓					✓
(Alam 2020)	SVR, MLR, DTR, Polynomial regression	SVR	✓				✓	
(Song et al. 2017)	ANN		✓					✓
(Frörmelt et al. 2018)	RF, KNN, LASSO ^{ak}		✓			✓		
(Liu and Bakshi 2018)	SVM, DTC ^{am}			✓		✓		
(Chen et al. 2021)	Multiagent deep reinforcement learning			✓			✓	
(Algren et al. 2021)	General ML				✓			
(DeRousseau 2020)	RF		✓			✓	✓	
(Seo and Kim 2007)	NN-GA		✓					✓
(Renard et al. 2021a)	Q-learning			✓			✓	
(Feng 2020)	BRBNN ^{an} , RT, EL ^{aq} , ELM	EL		✓		✓		
(Abdella et al. 2020)	k-means clustering, logistics regression		✓					✓
(Romeiko et al. 2019)	SVM, GBR	GBR	✓					✓
(Olafasakin et al. 2021)	Kriging-based ROM ^{ap}		✓					✓
(Zhou et al. 2009)	NN-GA		✓					✓
(Kleinekorte et al. 2019b)	ANN		✓					✓
(Milczarski et al. 2020)	ANN, RF, KNN, C4.5, SVM	RF	✓					✓
(Sousa et al. 2000)	ANN		✓					✓
(Park and Seo 2003)	ANN		✓					✓
(Sousa and Wallace 2006)	ANN		✓					✓
(Song 2019)	ANN, RF		✓					✓

Table 2 (continued)

Ref-ID	Individual / ensemble models	The best model	Types of Papers			Scale		
			PI	DM	LR	Buildings	Districts & cities	Miscellaneous
(Abokersh et al. 2020)	ANN		✓				✓	
(Arani 2020)	ANN		✓	✓		✓		

^aBack-Propagation Neural Networks^bBack-Propagation Neural Networks and Genetic Algorithms^cBoosted Regression Tree^dRandom Forest^eMulti-Layer Perceptron^fSupport Vector Machine^gApache Spark's machine learning library^hConvolutional Neural NetworkⁱIterative Dichotomiser 3 (decision tree learning)^jRadial Basis Function network^kGeneralized Regression Neural Network^lMonte Carlo version of Uninformative Variable Elimination Partial Least Squares^mMultiple Linear RegressionⁿRegression Tree^oSensitivity Analysis^pLinear Discriminant Analysis^qK-Nearest Neighbor^rDeep Neural Network^sDecision Tree Regression^tPrincipal Component Regression^xSupport Vector Regression^yFuzzy C-means clustering-Extreme^uBoruta Algorithm learning machine^zGaussian Mixture Models^{aa}Bayesian Network^{ab}Adaptive Network-based Fuzzy Inference System^{ac}A JAVA program package for machine learning algorithms^{ad}Linear Regression^{ae}Gradient Boosted Regression Tree^{af}Extreme Gradient Boosting^{ag}Gradient Boosting Machine^{ah}Fuzzy Set Theory^{ai}Naïve Bayes^{aj}Classification Tree^{ak}Least Absolute Shrinkage and Selection Operator regression^{al}Adaptive Boosting^{am}Decision Tree Classifier^{an}Bayesian regularization back-propagation neural networks^{aq}Ensemble learning^{ap}Reduced Order Model ROM

of actual computationally complex buildings. The computational time saved with the use of the proposed substitute models was found to be significant (Cornago et al. 2020). Also, Sharif proposed a simulation-based multi-objective optimization model for optimizing the selection of renovation scenarios for existing buildings by minimizing total energy consumption (TEC) considering LCA. Furthermore, he developed a surrogate ANN for selecting near-optimal building energy renovation methods; and developed deep ML Models (MLMs) to generate renovation scenarios considering TEC and LCC (Arani 2020). Azari et al. used a multi-objective optimization algorithm to explore ideal building envelope design by analyzing energy use and LCA of office buildings. Their approach combined an ANN and GA to find the optimal design (Azari et al. 2016). Feng developed an LCA method that integrated discrete event simulation and process-based LCA. The optimization method achieved real-time environmental optimization by introducing ML methods into simulation-based optimization. Płoszaj-Mazurek et al. applied the CNN method to optimize the carbon footprint of buildings in regenerative architectural design. The BRBNN, RT, EL and ELM algorithms were used to extract knowledge about the relationships between construction planning and project performance (Feng 2020). The results show ML methods could be a research tool for exploring vast design spaces in the field of sustainable architectural design (Płoszaj-Mazurek et al. 2020; Płoszaj-Mazurek 2020). Renard et al. implemented a Q-learning to optimize a pavement construction and maintenance plan to minimize the expected global warming impact of a pavement facility (Renard et al. 2021a). Liu et al. proposed a methodology that couples multiobjective optimization and ML to extract design heuristics (comfort temperatures and other related parameters). The methodology has been demonstrated on sustainable residential system design via TES-LCA methodology (Liu and Bakshi 2018).

At the **districts and cities** level, Chen et al. used multi-agent deep reinforcement learning to optimize dissolved oxygen and chemical dosage in water treatment plants. The outcome was designed from an LCA perspective to achieve sustainable optimization. They showed that the optimization based on LCA had results that achieved lower environmental impacts compared to the baseline scenario (Chen et al. 2021). Abokersh et al. developed a multiobjective optimization framework using an ANN model comprising the Bayesian optimization approach; assisted sensitivity analysis. ANN method was used to inherent sustainability principles in the design of solar assisted district heating in different urban sized communities in an optimization framework (Abokersh et al. 2020). DeRousseau et al. examined the various problem formulations commonly seen in concrete mixture design optimization that can necessitate models based on the linear

combination, statistics, ML, and physics. They used ML methods for predicting the compressive field strength of concrete (DeRousseau et al. 2018; DeRousseau 2020).

For **miscellaneous** uses, Zhou et al. proposed integration of ANN with GA to optimize the multiobjective of material selection in product design considering LCA (Zhou et al. 2009). In the context of decision-making support for LCA, Marvuglia et al. presented an evaluation of two different grouping techniques for categorizing materials based on their environmental performance. The agglomerative clustering technique and self-organizing map helped distinguish variables that could be used to establish classes of materials using their environmental performance (Marvuglia et al. 2015). The authors implemented GRNN and a set of linear models based on PLS regression, hoping to develop an automatic selection strategy of the critical variables according to the modelled output (USEtox factor) (Marvuglia et al. 2015; Barros and Ruschel 2021). Cornago et al. proposed an LCA aware scheduling framework, in which a production schedule is optimized with a lower environmental impact using predicted the hourly day-ahead electricity consumption by a DNN model (Cornago et al. 2020). Romeiko et al. presented a model for estimating LCA spatially at the county scale in corn production. This was developed by applying ML methods that could be used for corn supply chain optimization, corn-based biorefinery siting, and feedstock landscape optimization (Romeiko et al. 2020a).

Figure 4 represents the general relation of ML methods in the field of Optimization in LCA in the reviewed researches.

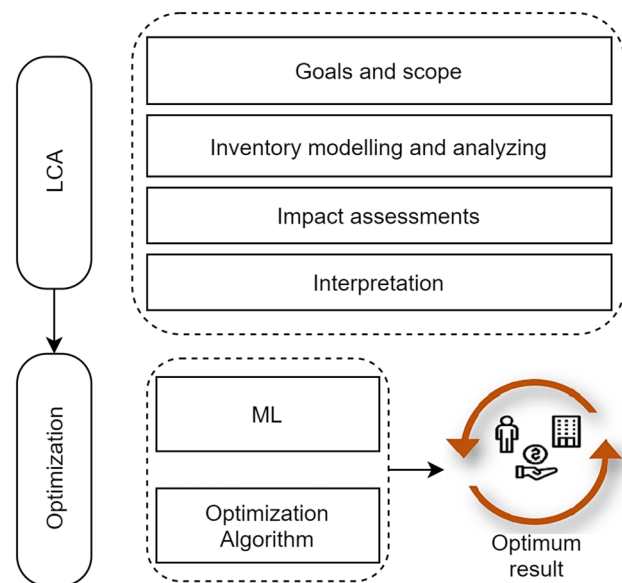
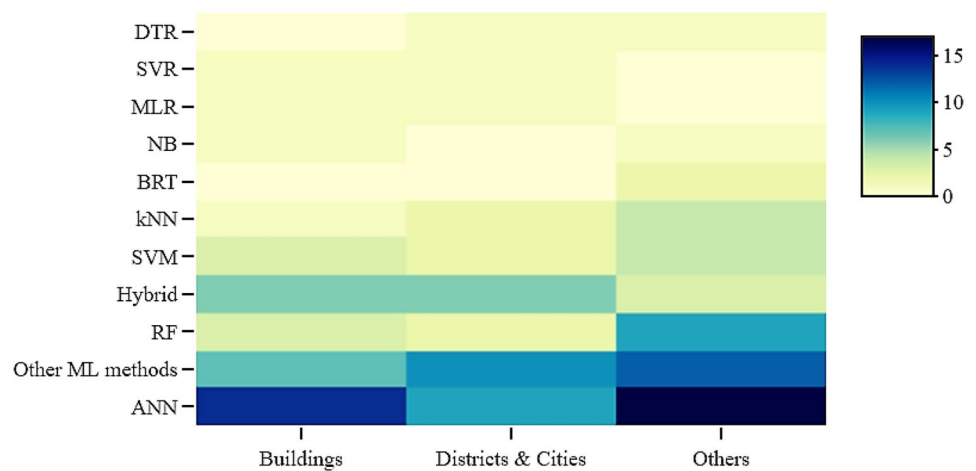


Fig. 4 ML and optimization methodology in LCA

Fig. 5 Heatmap of hit-points for each ML method

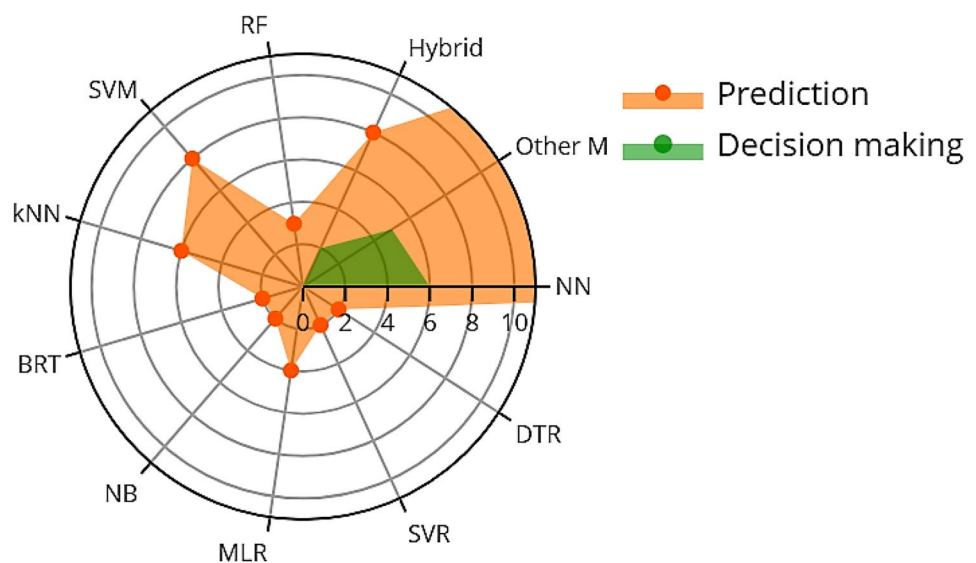
5 Results and discussion

This paper collaborated a literature survey to determine the use of ML techniques for LCA by answering the research questions. Gaps in research for ML in LCA were identified to guide future research. In the following sections, the highlights of reviewed papers and the limitation of using ML methods in LCA will be discussed.

5.1 Limitations of ML methods in LCA

Based on reviewed papers, the limitations of ML methods in LCA are;

1. LCA and training powerful analytical models with ML are expensive and depend on large amounts of hand-crafted, structured training data. Computational cost and
2. Some ML models, known as black-box models such as DNN, RF and SVMs, are exceedingly complex and make it very difficult to predict how they will perform in a specific context. Similarly, their users may not be able to review and understand the recommendations given by these models for intelligent systems.
3. Early design stages often are limited in detailed information, which is typically required for thorough assessments and thus need quick decisions on varying, numerous and loosely-defined concepts. These make the early use of detailed LCA impractical. For predictive modelling and experimental studies to be compatible, standardization of the conditions, experiments and reporting

Fig. 6 Radar graph showing the applications of ML methods in prediction and decision making

are needed in order to achieve consistency and to be reproducible.

4. Data-based approach is a method to fill in data gaps in LCA studies. It depends on the available data and how we choose to use it statistically, so we can recognize a good pattern from the data and make a prediction (Song 2019). Therefore, a large amount of data is required for an LCA while one of the key limitations on the application of ML algorithms side is a lack of high-quality and real-world-collected data sets.

5.2 Highlights of reviewed papers

The significant contributions of this paper are collaborating literature survey to determine use of ML techniques for LCA by answering the following research questions:

- How has ML been used in LCA?
- What is the role and efficacy of ML methods in optimization in LCA?
- Can ML methods integrate and contextualize existing inventory databases to provide a sound basis to streamline the LCA?
- What are the gaps in research in order to guide future research for ML in LCA?

Applying ML in LCA is explored, and the current state of the art reported in the literature is identified to answer the

above questions. ML techniques tailored to LCA and specific AI techniques that can advance LCA's establishment and delivery of the smart technology are investigated. In this section, the results of the research are shown by figures and tables.

How has ML been used in LCA? Table 2 shows details of the literature survey on ML methods in LCA. The papers are divided into three types of prediction impact (PI), decision making (DM) and literature review (LR). The applied method in each paper is identified, and the scale at which the model was applied is shown. The papers included in this paper answer and support the above research question. In the included literature, many different applications at different scales were demonstrated to be beneficial in accurate and efficient LCA.

The associated heatmap, Fig. 5, shows that ANN is the most commonly applied method at all three levels of categorisation in this paper, particularly at the buildings level and then at the miscellaneous level which includes a variety of niche applications. Hybrid techniques were the next most used ML method at the building and districts and cities level.

Figure 6 is a radar graph showing that the most common application of ML methods have been for predictions. Individually, NN was most commonly used, followed by hybrid methods. In response to the first research question, Figs. 5 and 6 display a list of included studies that supported the use of ML-based prediction methods to predict LCA accurately.

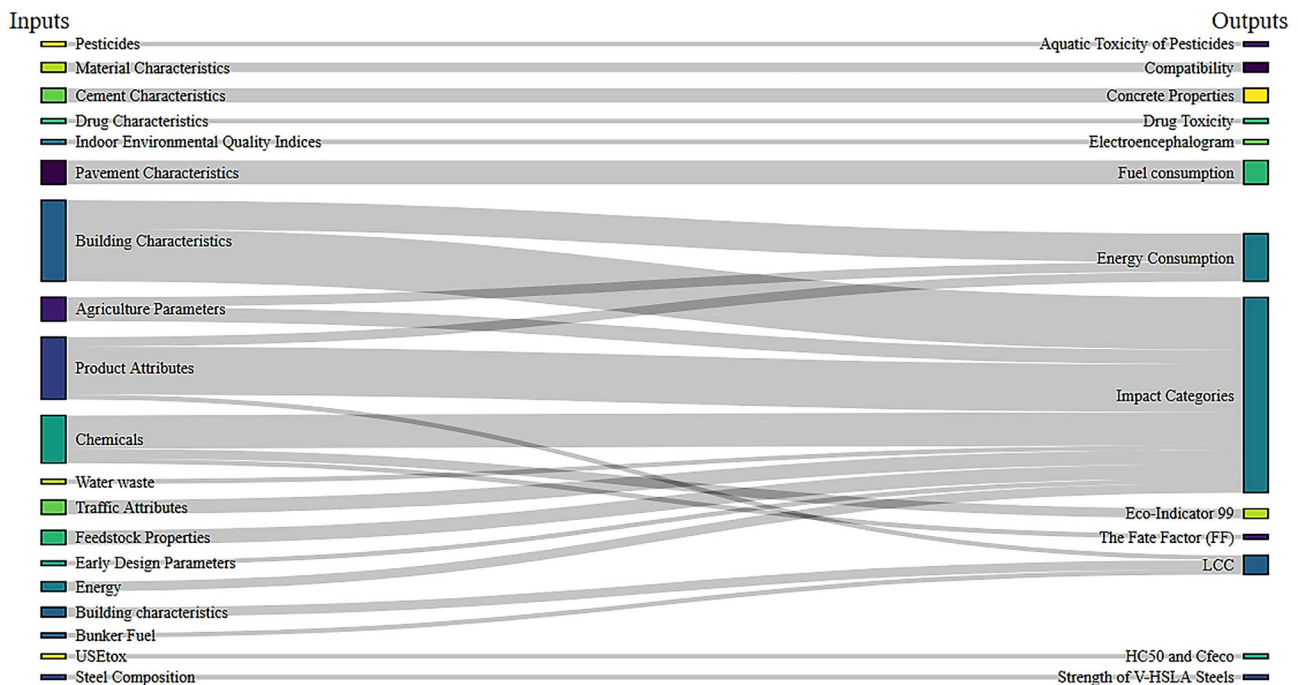


Fig. 7 Sankey diagram to show the relationship between the inputs and outcomes of ML methods in LCA

What is the role and efficacy of ML methods in Optimization in LCA? The results of this study show that the ML methods are capable of matching detailed LCA results and predicting missing data or trends of variables while staying within the accuracy of typical LCA. Furthermore, ML extends outside of LCA in processes such as data cleaning, predicting system output or performance, ecosystem informatics, and optimization. ML algorithms could also be applied in screening or cleaning data for LCI, estimating flow data for unit processes, improving the quality and quantity of data used to determine CFs, and can be used to generate optimized scenarios. They are especially suitable for supporting real-time decisions of construction environmental optimization. This study shows that the ML can be coupled with standard optimization methods to increase their capability of quickly exploring promising regions.

Can ML methods integrate and contextualize existing inventory databases to provide a sound basis to streamline the LCA? Many included studies in this review utilised pre-established databases in order to perform LCA. ML methods are capable of integrating these existing databases, although with all LCA the quality of the data and the nature of the

database may have an impact on the quality of LCA. However, ML methods identified in this paper can be used to fill in gaps if pre-existing databases are partially complete.

Figure 7 shows predictors and outcomes of ML methods that have been used in LCA applications. Characteristics are shown as the most commonly used inputs. Impact categories were the most frequently assessed outcome of these applications.

What are the gaps in research in order to guide future research for ML in LCA? Table 3 shows details of the literature survey on ML methods in LCA for different levels of the built environment. In this paper, the levels are categorized as buildings vs district & cities. For each level of the built environment, different categories of LCA are identified. The applications in which ML methods have been utilized have been marked with an asterisk. This is a roadmap for researchers in LCA who want to apply ML techniques to identify gaps in research. This paper has identified a significant gap in research in the 'End of Life' phase and 'Benefits beyond the system, for buildings. These include demolition, disposal and transport, as well as recycling. At the districts

Table 3 Use cases at different levels of the built environment

Buildings				
Product stage	Constructions	Use stage	End of life	Benefits beyond the system
★ Early design	Construction waste	Retrofit waste	Demolition	Reuse
Raw material supply	★ Construction materials	Retrofit materials	★ Waste processing	Recycling
★ Manufacturing	Transport	Renovation waste	Disposal	Transport
★ Transport		★ Renovation materials	Transport	
		Maintenance waste		
		Maintenance materials		
		★ Operational use of energy		
		Operational use of water		
		Transport		
Districts & cities				
Open spaces ^a	Networks ^b	Mobility ^c	★ Consumption patterns	
Construction waste	Construction waste	Vehicle manufacturing		
★ Construction materials	Construction materials	★ Daily mobility		
Maintenance waste	Maintenance waste			
★ Maintenance materials	Maintenance materials			
Public lighting	Demolition waste			
Public wastes	★ Energy systems			
Demolition wastes	★ Water systems			
Water consumption				

^aRoads, Pavements, Parking, Green spaces, etc.

^bNetworks of Water, Electricity, Water-waste, Gas and District heating, etc.

^cPassenger cars, Public transport, etc.

and cities level, the most significant opportunities for ML in LCA research lie in the ‘Networks’ and ‘Open Spaces’.

The research questions posed in this paper were answered through this literature survey. In the included papers, authors claimed and displayed that ML can be applied to different aspects of the LCA and be a useful tool. ML methods were shown to be applied efficiently in optimization scenarios in LCA. Finally, ML methods were integrated into existing inventory databases to streamline the LCA across many use cases. However, ML-based techniques have been employed less for real-time monitoring and control of real-world LCA.

Future research should focus on using ML technologies in real-time applications to monitor, optimize, and control the built-environment systems. ML models may be more comprehensible than other black-box approaches due to their transparency. Furthermore, hybrid ML applications may expand on the benefits of ML models and overcome limitations to case-specific scenarios for optimizing LCA through their interpolation and extrapolation capabilities. Advanced stochastic metaheuristics should be used in refining ML model training parameters to maximize their accuracy and reliability.

6 Conclusion

LCA, when done successfully, provides a systems view of products systematically and quantitatively and can thus act as a decision support tool. It can then guide the design and give insights on areas for improvement and innovation. However, performing detailed LCA is expensive, time-consuming, and requires a large amount of data.

The ML methods in LCA have received considerable attention as countries are continuing and growing to address the importance and protection of the environment. The climate regulations have encouraged industries to apply LCA using various intelligent technologies. The rapid development of modern technologies, including sensors, information, wireless transmission, network communication, cloud computing, and smart devices have been led to an enormous amount of data accumulation. Therefore, LCA researches have adopted the opportunities made possible by the development of computational techniques and ML methods to improve predictive models. ML methods belong to the category of data-based predictive models and thus aims to use computational methods to allow an algorithm to find a meaningful pattern from an extensive data set.

The contributions of this paper are as follows:

- This study presented a review of ML models utilized for LCA. It presented a thorough review and critical discussion of various ML technologies to solve function approximation, optimization, monitoring, and control

problems in LCA research. Moreover, the advantages and disadvantages of using ML technologies in LCA are highlighted to direct future policymakers efforts in this domain.

- The reviews show that if computational levels in LCA are divided into three categories, inventory, modelling and optimization, ML is most used at the inventory level for prediction and finding the missing data; and optimizing during the model simulation. The fundamental limitations and challenges faced by applying ML methods in LCA are model complexity and scenario uncertainty.
- The review identifies that developing ML techniques, including predictive model control and optimization algorithms, can help the policymakers deliver actionable knowledge to inform various control strategies and corrective measures to reduce the gap between predicted and actual environmental impact. This review finds that ML methods can match the LCA results within the accuracy of typical LCA studies and correctly predict the trends.

This review has identified research gaps and given an overview of the progression in this field to aid researchers’ understanding of key concepts for applying ML in LCA. Future research should focus on using ML technologies in real-time applications to monitor, optimize, and control the built-environment systems. ML models may be more comprehensible than other black-box approaches due to their transparency. Furthermore, hybrid ML applications may expand on the benefits of ML models and overcome limitations to case-specific scenarios for optimizing LCA through their interpolation and extrapolation capabilities. Optimization uses can be particularly impactful in life cycle alternatives where the environmental impact of a process, product or system can be reduced. ML has the capability of being integrated as a real-time algorithm, assessing production or changes in processes and responding with potential alternatives. These can be less environmentally impactful and help decision-makers choose the optimum available options of the design, construction/production, facilities management, and demolition processes. Advanced stochastic metaheuristics should be used in refining ML model training parameters to maximize their accuracy and reliability. Nevertheless, ML may not be appropriate for every application and should be considered alongside the cost, length of time and delays which incur from some ML techniques. In the future, the integration of ML models within LCA may be commonplace following further research into applications such as utilizing access to dynamic data and providing detailed and accurate environmental impacts.

Acknowledgements The authors would like to acknowledge funding from the Engineering and Physical Sciences Research Council (EPSRC) under the contract EP/T019514/1.

Declarations

Conflicts of interest The authors have no competing interests to declare that are relevant to the content of this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abdella GM, Kucukvar M, Onat NC, Al-Yafay HM, Bulak ME (2020) Sustainability assessment and modeling based on supervised machine learning techniques: The case for food consumption. *J Clean Prod* 251:119661. <https://doi.org/10.1016/j.jclepro.2019.119661>. <https://www.sciencedirect.com/science/article/pii/S0959652619345317>
- Abokersh MH, Vallés M, Cabeza LF, Boer D (2020) A framework for the optimal integration of solar assisted district heating in different urban sized communities: A robust machine learning approach incorporating global sensitivity analysis. *Appl Energy* 267:114903. <https://doi.org/10.1016/j.apenergy.2020.114903>. <https://www.sciencedirect.com/science/article/pii/S0306261920304153>
- Adediji PA, Akinlabi SA, Madushele N, Olatunji OO (2020) Potential roles of artificial intelligence in the lci of renewable energy systems. In: Emamian SS, Awang M, Yusof F (eds) *Advances in Manufacturing Engineering*. Springer Singapore, Singapore, pp 275–285
- Alam MR (2020) Pavements—performance; pavements, asphalt—environmental aspects. Master's thesis, Memorial University of Newfoundland, Newfoundland and Labrador, Canada
- Algren M, Fisher W, Landis AE (2021) Chapter 8 - machine learning in life cycle assessment. In: Dunn J, Balaprakash P (eds) *Data Science Applied to Sustainability Analysis*. Elsevier, pp 167–190. <https://doi.org/10.1016/B978-0-12-817976-5.00009-7>. <https://www.sciencedirect.com/science/article/pii/B9780128179765000097>
- Alqahtani A, Whyte A (2013) Artificial neural networks incorporating cost significant items towards enhancing estimation for (life-cycle) costing of construction projects. *AJCEB* 13:51–64. <https://doi.org/10.5130/ajceb.v13i3.3363>
- Anand CK, Amor B (2017) Recent developments, future challenges and new research directions in LCA of buildings: A critical review. *Renew Sust Energ* 67:408–416. <https://doi.org/10.1016/j.rser.2016.09.058>. <https://www.sciencedirect.com/science/article/pii/S1364032116305524>
- Arani SAS (2020) Optimizing energy performance of building renovation using traditional and machine learning approaches. <https://spectrum.library.concordia.ca/986791/>, unpublished
- Asif M (2019) An empirical study on life cycle assessment of double-glazed aluminium-clad timber windows. *Int J Build Pathol* 37(5):547–564. <https://doi.org/10.1108/IJBPA-01-2019-0001>
- Azari R, Garshasbi S, Amini P, Rashed-Ali H, Mohammadi Y (2016) Multi-objective optimization of building envelope design for life cycle environmental performance. *Energy Build* 126:524–534. <https://doi.org/10.1016/j.enbuild.2016.05.054>. <https://www.sciencedirect.com/science/article/pii/S0378778816304236>
- Barros NN, Ruschel RC (2021) Machine learning for whole-building life cycle assessment: A systematic literature review. In: Toledo Santos E, Scheer S (eds) *Proceedings of the 18th International Conference on Computing in Civil and Building Engineering*. Springer International Publishing, Cham, pp 109–122
- Breiman L (2001) Random forests. *Mach Learn* 45(1):5–32. <https://doi.org/10.1023/A:1010933404324>
- Bueno C, Hauschild MZ, Rossignolo JA, Ometto AR, Mendes NC (2016) Sensitivity analysis of the use of life cycle impact assessment methods: a case study on building materials. *J Clean Prod* 112:2208–2220. <https://doi.org/10.1016/j.jclepro.2015.10.006>. <https://www.sciencedirect.com/science/article/pii/S0959652615013876>
- Bui KQ, Perera LP (2020) A decision support framework for cost-effective and energy-efficient shipping. <https://doi.org/10.1115/OMAE2020-18368>
- Chen K, Wang H, Valverde-Pérez B, Zhai S, Vezzaro L, Wang A (2021) Optimal control towards sustainable wastewater treatment plants based on multi-agent reinforcement learning. *Chemosphere* 279:130498. <https://doi.org/10.1016/j.chemosphere.2021.130498>. <http://dx.doi.org/10.1016/j.chemosphere.2021.130498>
- Cheng F, Luo H, Colosi LM (2020a) Slow pyrolysis as a platform for negative emissions technology: An integration of machine learning models, life cycle assessment, and economic analysis. *Energy Convers Manag* 223:113258. <https://doi.org/10.1016/j.enconman.2020.113258>. <https://www.sciencedirect.com/science/article/pii/S0196890420308013>
- Cheng F, Porter MD, Colosi LM (2020b) Is hydrothermal treatment coupled with carbon capture and storage an energy-producing negative emissions technology? *Energy Convers Manag* 203:112252. <https://doi.org/10.1016/j.enconman.2019.112252>. <https://www.sciencedirect.com/science/article/pii/S0196890419312580>
- Chiang TA, Che ZH, Wang TT (2011) A design for environment methodology for evaluation and improvement of derivative consumer electronic product development. *J Syst Sci Syst Eng* 20(3):260. <https://doi.org/10.1007/s11518-011-5168-8>
- Cornago S, Vitali A, Brondi C, Low JSC (2020) Electricity technological mix forecasting for life cycle assessment aware scheduling. *Procedia CIRP* 90:268–273. <https://doi.org/10.1016/j.procir.2020.01.099>. <https://www.sciencedirect.com/science/article/pii/S2212827120302663>. 27th CIRP Life Cycle Engineering Conference (LCE2020) Advancing Life Cycle Engineering : from technological eco-efficiency to technology that supports a world that meets the development goals and the absolute sustainability
- Cortes C, Vapnik V (1995) Support-vector networks. *Mach Learn* 20(3):273–297. <https://doi.org/10.1007/BF00994018>
- D'Amico A, Ciulla G, Traverso M, Lo Brano V, Palumbo E (2019a) Artificial neural networks to assess energy and environmental performance of buildings: an Italian case study. *J Clean Prod* 239:117993. <https://doi.org/10.1016/j.jclepro.2019.117993>. <https://www.sciencedirect.com/science/article/pii/S095965261932863X>
- D'Amico B, Myers R, Sykes J, Voss E, Cousins-Jenvey B, Fawcett W, Richardson S, Kermani A, Pomponi F (2019b) Machine learning for sustainable structures: a call for data. *Structures* 19:1–4. <https://doi.org/10.1016/j.istruc.2018.11.013>. <https://www.sciencedirect.com/science/article/pii/S2352012418301395>
- DeRousseau M, Kasprzyk J, Srubar W (2018) Computational design optimization of concrete mixtures: a review. *Cem Concr Res* 109:42–53. <https://doi.org/10.1016/j.cemconres.2018.04.007>. <https://www.sciencedirect.com/science/article/pii/S0008884617313601>
- DeRousseau MA (2020) Concrete mixture design using machine learning, life cycle assessment, and multi-objective optimization. PhD thesis, Department of Civil, Architectural, and Environmental Engineering, Ann Arbor. <http://abc.cardiff.ac.uk/login?url=https://www.>

- proquest.com/dissertations-theses/concrete-mixture-design-using-machine-learning/docview/2441568304/se-2?accountid=9883
- Dostatni E, Rojek I, Hamrol A (2018) The use of machine learning method in concurrent ecodesign of products and technological processes. In: Hamrol A, Ciszak O, Legutko S, Jurczyk M (eds) *Advances in Manufacturing*. Springer International Publishing, Cham, pp 321–330
- Duprez S, Fouquet M, Herreros Q, Jusselme T (2019) Improving life cycle-based exploration methods by coupling sensitivity analysis and metamodels. *SCS* 44:70–84. <https://doi.org/10.1016/j.scs.2018.09.032>. <https://www.sciencedirect.com/science/article/pii/S2210670718309818>
- Elkhayat YO, Ibrahim MG, Tokimatsu K, Ali AAM (2020) A comparative life cycle assessment of three high-performance glazing systems for office buildings in a hot desert climate zone. *Clean Technol Environ Policy* 22(7):1499–1515. <https://doi.org/10.1007/s10098-020-01891-2>. <https://doi.org/10.1007/s10098-020-01891-2>
- Feng K (2020) Environmentally friendly construction processes under uncertainty: assessment, optimisation and robust decision-making. PhD thesis, Luleå University of Technology, Industrialized and sustainable construction
- Feng K, Lu W, Wang Y (2019) Assessing environmental performance in early building design stage: An integrated parametric design and machine learning method. *SCS* 50:101596. <https://doi.org/10.1016/j.scs.2019.101596>. <https://www.sciencedirect.com/science/article/pii/S2210670718318870>
- Frömlert A (2018) Data mining meets life cycle assessment: Towards understanding and quantifying environmental impacts of individual households. PhD thesis, ETH Zurich, Zurich. <https://doi.org/10.3929/ethz-b-000282689>
- Frömlert A, Dürrenmatt DJ, Hellweg S (2018) Using data mining to assess environmental impacts of household consumption behaviors. *Environ Sci Technol* 52(15):8467–8478. <https://doi.org/10.1021/acs.est.8b01452>
- Frömlert A, Buffat R, Hellweg S (2020) Machine learning based modeling of households: A regionalized bottom-up approach to investigate consumption-induced environmental impacts. *J Ind Ecol* 24:639–652
- Gu X (2018) Metal-organic frameworks for post-combustion carbon capture - a life cycle assessment. PhD thesis, The Ohio State University. http://rave.ohiolink.edu/etdc/view?acc_num=osu152415298686088
- Guinee JB (2002) Handbook on life cycle assessment operational guide to the ISO standards. *Int J Life Cycle Assess* 7(5):311. <https://doi.org/10.1007/BF02978897>
- Gust K, Collier Z, Mayo M, Stanley J, Gong P, Chappell M (2015) Limitations of toxicity characterization in life cycle assessment: Can adverse outcome pathways provide a new foundation? *IEAM* 12. <https://doi.org/10.1002/ieam.1708>
- Hou P (2019) Data-driven environmental system analysis: Addressing data Gaps in life cycle assessment. PhD thesis, The University of Michigan. <https://deepblue.lib.umich.edu/handle/2027.42/151638>
- Hou P, Cai J, Qu S, Xu M (2018) Estimating missing unit process data in life cycle assessment using a similarity-based approach. *Environ Sci Technol* 52(9):5259–5267. <https://doi.org/10.1021/acs.est.7b05366>
- Hou P, Jolliet O, Zhu J, Xu M (2020) Estimate ecotoxicity characterization factors for chemicals in life cycle assessment using machine learning models. *Environ Int* 135:105393. <https://doi.org/10.1016/j.envint.2019.105393>. <https://www.sciencedirect.com/science/article/pii/S0160412019314412>
- IPCC (2021) Summary for Policymakers, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, book section SPM. <https://doi.org/10.1017/CBO9781107415324.004>. www.climatechange2013.org
- Kaab A, Sharifi M, Mobli H, Nabavi-Pelesaraei A, wing Chau K (2019) Combined life cycle assessment and artificial intelligence for prediction of output energy and environmental impacts of sugarcane production. *Sci Total Environ* 664:1005–1019. <https://doi.org/10.1016/j.scitotenv.2019.02.004>. <https://www.sciencedirect.com/science/article/pii/S0048969719304838>
- Kishk M, Al-Hajj A (1999) An integrated framework for life cycle costing in buildings. In: *Proceedings of the COBRA 1999 RICS Construction and Building Research Conference*, vol 2. pp 92–101
- Kleinekorte J, Kröger L, Leonhard K, Bardow A (2019a) A neural network-based framework to predict process-specific environmental impacts. In: Kiss AA, Zondervan E, Lakerveld R, Özkan L (eds) *29th European Symposium on Computer Aided Process Engineering, Computer Aided Chemical Engineering*, vol 46. Elsevier, pp 1447–1452. <https://doi.org/10.1016/B978-0-12-818634-3.50242-3>. <https://www.sciencedirect.com/science/article/pii/B9780128186343502423>
- Kleinekorte J, Welz M, Fleitmann LHJ, Kröger LC, Leonhard K, Bardow A (2019b) Combining process short cuts and artificial neural networks for predictive life cycle assessment of chemicals. <https://publications.rwth-aachen.de/record/764985>, rWTH-2019-07176
- Kurdi A, Alhazmi N, Alhazmi H, Tabbakh T (2020) Practice of simulation and life cycle assessment in tribology—a review. *Materials* 13(16). <https://doi.org/10.3390/ma13163489>. <https://www.mdpi.com/1996-1944/13/16/3489>
- Lee EK, Zhang WJ, Zhang X, Adler PR, Lin S, Feingold BJ, Khwaja HA, Romeiko XX (2020) Projecting life-cycle environmental impacts of corn production in the u.s. midwest under future climate scenarios using a machine learning approach. *Sci Total Environ* 714:136697. <https://doi.org/10.1016/j.scitotenv.2020.136697>. <https://www.sciencedirect.com/science/article/pii/S0048969720302072>
- Leszczynski T, Jasinski T (2020) Comparison of product life cycle cost estimating models based on neural networks and parametric techniques—a case study for induction motors. *Sustainability* 12(20). <https://doi.org/10.3390/su12208353>. <https://www.mdpi.com/2071-1050/12/20/8353>
- Li F, Fan D, Wang H, Yang H, Li W, Tang Y, Liu G (2017) In silico prediction of pesticide aquatic toxicity with chemical category approaches. *Toxicol Res* 6:831–842. <https://doi.org/10.1039/C7TX00144D>. <http://dx.doi.org/10.1039/C7TX00144D>
- Li J, Wu Z, Zhang HC (2008) Application of neural network on environmental impact assessment tools. *Int J Sustain Manuf* 1:100–121. <https://doi.org/10.1504/IJSM.2008.019229>. <https://www.inderscienceonline.com/doi/abs/10.1504/IJSM.2008.019229>
- Li Y, Zhang H, Roy U, Lee YT (2017) A data-driven approach for improving sustainability assessment in advanced manufacturing. In: *2017 IEEE International Conference on Big Data (Big Data)*. pp 1736–1745. <https://doi.org/10.1109/BigData.2017.8258116>
- Liao M, Kelley S, Yao Y (2020) Generating energy and greenhouse gas inventory data of activated carbon production using machine learning and kinetic based process simulation. *ACS Sustain Chem Eng* 8(2):1252–1261. <https://doi.org/10.1021/acssuschemeng.9b06522>. <https://doi.org/10.1021/acssuschemeng.9b06522>
- Liu X, Bakshi BR (2018) Extracting heuristics for designing sustainable built environments by coupling multiobjective evolutionary optimization and machine learning. In: Eden MR, Ierapetritou MG, Towler GP (eds) *13th International Symposium on Process Systems Engineering (PSE 2018)*, *Computer Aided Chemical Engineering*, vol 44. Elsevier, pp 2539–2544. <https://doi.org/10.1016/B978-0-444-64241-7.50418-3>. <https://www.sciencedirect.com/science/article/pii/B9780444642417504183>
- Livingstone DJ (2008) Artificial neural networks: methods and applications (Methods in Molecular Biology), 1st edn. Humana Press
- Luque A, De Las Heras A, Ávila Gutiérrez MJ, Zamora-Polo F (2020) Adapts: An intelligent sustainable conceptual framework for

- engineering projects. *Sensors* 20(6). <https://doi.org/10.3390/s20061553>. <https://www.mdpi.com/1424-8220/20/6/1553>
- Lysenko A, Sharma A, Boroevich KA, Tsunoda T (2018) An integrative machine learning approach for prediction of toxicity-related drug safety. *Life Sci Alliance* 1(6):e201800098–e201800098. <https://pubmed.ncbi.nlm.nih.gov/30515477>
- Lyu Y, Chow Tt (2020) Economic, energy and environmental life cycle assessment of a liquid flow window in different climates. *Build Simul* 13(4):837–848. <https://doi.org/10.1007/s12273-020-0636-z>. <https://doi.org/10.1007/s12273-020-0636-z>
- Ma J, Kim HM (2015) Predictive usage mining for life cycle assessment. *Transp Res D Transp Environ* 38:125–143. <https://doi.org/10.1016/j.trd.2015.04.022>. <https://www.sciencedirect.com/science/article/pii/S1361920915000541>
- Manfred M, Caputo P, Costa G (2011) Paradigm shift in urban energy systems through distributed generation: Methods and models. *Appl Energy* 88(4):1032–1048. <https://doi.org/10.1016/j.apenergy.2010.10.018>. <https://www.sciencedirect.com/science/article/pii/S0306261910004204>
- Marvuglia A, Kanevski M, Benetto E (2015) Machine learning for toxicity characterization of organic chemical emissions using usetox database: Learning the structure of the input space. *Environ Int* 83:72–85. <https://doi.org/10.1016/j.envint.2015.05.011>. <https://www.sciencedirect.com/science/article/pii/S0160412015001269>
- Milczarski P, Zielinski B, Stawska Z, Hlobaz A, Maslanka P, Kosinski P (2020) Machine learning application in energy consumption calculation and assessment in food processing industry. In: Rutkowski L, Scherer R, Korytkowski M, Pedrycz W, Tadeusiewicz R, Zurada JM (eds) *Artificial Intelligence and Soft Computing*. Springer International Publishing, Cham, pp 369–379
- Milojevic-Dupont N, Creutzig F (2021) Machine learning for geographically differentiated climate change mitigation in urban areas. *SCS* 64:102526. <https://doi.org/10.1016/j.scs.2020.102526>. <https://www.sciencedirect.com/science/article/pii/S2210670720307423>
- Mitchell TM (1997) *Machine Learning*, 1st edn. McGraw-Hill Inc, USA
- Mukherjee A, Guptab H, Guptac JK (2019) Embodied energy assessment of construction material in india using artificial neural network. *Int J Tec Innov Mod Eng Sci* 5:64–75
- Nabavi-Pelesaraei A, Rafiee S, Mohtasebi SS, Hosseinzadeh-Bandbafha H, Wing Chau K (2018) Integration of artificial intelligence methods and life cycle assessment to predict energy output and environmental impacts of paddy production. *Sci Total Environ* 631–632:1279–1294. <https://doi.org/10.1016/j.scitotenv.2018.03.088>. <https://www.sciencedirect.com/science/article/pii/S0048969718308428>
- Negishi K, Tiruta-Barna L, Schiopu N, Lebert A, Chevalier J (2018) An operational methodology for applying dynamic life cycle assessment to buildings. *Build Environ* 144:611–621. <https://doi.org/10.1016/j.buildenv.2018.09.005>. <https://www.sciencedirect.com/science/article/pii/S0360132318305535>
- Oduyemi O, Okoroh M, Dean A (2015) Developing an artificial neural network model for life cycle costing in buildings. In: *Proc. of 31st Annual ARCOM Conference*. pp 843–852
- Olafasakin O, Chang Y, Passalacqua A, Subramaniam S, Brown RC, Mba Wright M (2021) Machine learning reduced order model for cost and emission assessment of a pyrolysis system. *Energy Fuels* 35(12):9950–9960. <https://doi.org/10.1021/acs.energyfuels.1c00490>
- Østergaard N, Thorsted L, Miraglia S, Birkved M, Rasmussen FN, Birgisdóttir H, Kalbar P, Georgiadis S (2018) Data driven quantification of the temporal scope of building lcas. *Procedia CIRP* 69:224–229. <https://doi.org/10.1016/j.procir.2017.11.057>. <https://www.sciencedirect.com/science/article/pii/S2212827117308302>, 25th CIRP Life Cycle Engineering (LCE) Conference, 30 April – 2 May 2018, Copenhagen, Denmark
- Park JH, Seo KK (2003) Approximate life cycle assessment of product concepts using multiple regression analysis and artificial neural networks. *KSME INT J* 17(12):1969–1976. <https://doi.org/10.1007/BF02982436>
- Perrotta F, Parry T, Neves LC (2017) Application of machine learning for fuel consumption modelling of trucks. In: *2017 IEEE International Conference on Big Data (Big Data)*. pp 3810–3815. <https://doi.org/10.1109/BigData.2017.8258382>
- Perrotta F, Parry T, Neves LC, Mesgarpour M (2018) A machine learning approach for the estimation of fuel consumption related to road pavement rolling resistance for large fleets of trucks. In: *Proc. of the Sixth International Symposium on Life-Cycle Civil Engineering (IALCCE 2018)*, Ghent, Belgium. <http://eprints.nottingham.ac.uk/id/eprint/51400>
- Peterson LE (2009) K-nearest neighbor. *Scholarpedia* 4(2):1883. <https://doi.org/10.4249/scholarpedia.1883>, revision #137311
- Płoszaj-Mazurek M (2020) Machine learning-aided architectural design for carbon footprint reduction. *Builder* 276:35–39. <https://doi.org/10.5604/01.3001.0014.1615>
- Płoszaj-Mazurek M, Rynska E, Grochulska-Salak M (2020) Methods to optimize carbon footprint of buildings in regenerative architectural design with the use of machine learning, convolutional neural network, and parametric design. *Energies* 13(20). <https://doi.org/10.3390/en13205289>
- Pradeep Kumar P, Santos DA, Braham EJ, Sellers DG, Banerjee S, Dixit MK (2021) Punching above its weight: life cycle energy accounting and environmental assessment of vanadium microalloying in reinforcement bar steel. *Environ Sci: Processes Impacts* 23:275–290. <https://doi.org/10.1039/D0EM00424C>. <http://dx.doi.org/10.1039/D0EM00424C>
- Renard S, Corbett B, Swee O (2021a) Minimizing the global warming impact of pavement infrastructure through reinforcement learning. *Resour Conserv Recycl* 167:105240. <https://doi.org/10.1016/j.resconrec.2020.105240>. <https://www.sciencedirect.com/science/article/pii/S0921344920305553>
- Renard S, Corbett B, Swee O (2021b) Minimizing the global warming impact of pavement infrastructure through reinforcement learning. *Resour Conserv Recycl* 167:105240. <https://doi.org/10.1016/j.resconrec.2020.105240>. <https://www.sciencedirect.com/science/article/pii/S0921344920305553>
- Rojek I, Dostatni E (2020) Machine learning methods for optimal compatibility of materials in ecodesign. *B POL ACAD SCI-TECH* 68(No. 2 (i.a. Special Section on Computational Intelligence in Communications)):199–206. http://journals.pan.pl/Content/115175/PDF/04D_199-206_01290_Bpast.No.68-2_10.04.20_K1A_SS_.pdf
- Romeiko XX, Guo Z, Pang Y (2019) Comparison of support vector machine and gradient boosting regression tree for predicting spatially explicit life cycle global warming and eutrophication impacts: A case study in corn production. In: *2019 IEEE International Conference on Big Data (Big Data)*. pp 3277–3284. <https://doi.org/10.1109/BigData47090.2019.9005581>
- Romeiko XX, Guo Z, Pang Y, Lee EK, Zhang X (2020a) Comparing machine learning approaches for predicting spatially explicit life cycle global warming and eutrophication impacts from corn production. *Sustainability* 12(4). <https://doi.org/10.3390/su12041481>. <https://www.mdpi.com/2071-1050/12/4/1481>
- Romeiko XX, Lee EK, Sorunmu Y, Zhang X (2020b) Spatially and temporally explicit life cycle environmental impacts of soybean production in the U.S. Midwest. *ES&T* 54(8):4758–4768. <https://doi.org/10.1021/acs.est.9b06874>. PMID: 32202767
- Saab F (2019) *Parallelca : a foreground aware parallel calculator for life cycle assessment*. PhD thesis, École de technologie supérieure. <http://oatd.org/oatd/record?record=oai:espace.etsmtl.ca:2405&q=id%3A%22oai%3Aespace.etsmtl.ca%3A2405%22>

- Sala S, Amadei AM, Beylot A, Ardente F (2021) The evolution of life cycle assessment in European policies over three decades. *Int J Life Cycle Assess*. <https://doi.org/10.1007/s11367-021-01893-2>
- Seo KK, Kim WK (2007) Approximate life cycle assessment of product concepts using a hybrid genetic algorithm and neural network approach. In: Szczuka MS, Howard D, Śizak D, Kim HK, Kim TH, Ko IS, Lee G, Sloat PMA (eds) *Advances in Hybrid Information Technology*. Springer, Berlin Heidelberg, Berlin, Heidelberg, pp 258–268
- Seo KK, Min SH, Yoo HW (2005) Artificial neural network based life cycle assessment model for product concepts using product classification method. In: Gervasi O, Gavrilova ML, Kumar V, Laganá A, Lee HP, Mun Y, Taniar D, Tan CJK (eds) *Computational Science and Its Applications - ICCSA 2005*. Springer, Berlin Heidelberg, Berlin, Heidelberg, pp 458–466
- Serajiantehrani R (2020) Development of a machine learning-based prediction model for construction and environmental costs of trenchless spray-applied pipe linings, cured-in-place pipe, and sliplining methods in large diameter culverts. PhD thesis, The University of Texas. <https://rc.library.uta.edu/uta-ir/handle/10106/28867>
- Shan X, Yang EH, Zhou J, Chang V (2017) Human-building interaction under various indoor temperatures through neural-signal electroencephalogram (EEG) methods. *Build Environ* 129. <https://doi.org/10.1016/j.buildenv.2017.12.004>
- Sharif SA, Hammad A (2019) Developing surrogate ANN for selecting near-optimal building energy renovation methods considering energy consumption, LCC and LCA. *J Build Eng* 25:100790. <https://doi.org/10.1016/j.job.2019.100790>. <https://www.sciencedirect.com/science/article/pii/S2352710218311811>
- Shi Q, Xu Y (2009) The selection of green building materials using GA-BP hybrid algorithm. In: 2009 International Conference, AICI, vol 3. pp 40–45. <https://doi.org/10.1109/AICI.2009.74>
- Skaar C, Jørgensen RB (2013) Integrating human health impact from indoor emissions into an LCA: a case study evaluating the significance of the use stage. *Int J Life Cycle* 18(3):636–646. <https://doi.org/10.1007/s11367-012-0506-8>
- Slapnik M, Istenic D, Pintar M, Udovc A (2014) Extending life cycle assessment normalization factors and use of machine learning a Slovenian case study. *Ecol Indic* 50. <https://doi.org/10.1016/j.ecolind.2014.10.028>
- Smetana S, Seebold C, Heinz V (2018) Neural network, blockchain, and modular complex system: The evolution of cyber-physical systems for material flow analysis and life cycle assessment. *Resour Conserv Recycl* 133:229–230. <https://doi.org/10.1016/j.resconrec.2018.02.020>. <https://www.sciencedirect.com/science/article/pii/S092134491830079X>
- Song R (2019) Machine learning for addressing data deficiencies in life cycle assessment. PhD thesis, University of California. <http://abc.cardiff.ac.uk/login?url=https://www.proquest.com/dissertations-theses/machine-learning-addressing-data-deficiencies/docview/2217716309/se-2?accountid=9883>. copyright - Database copyright ProQuest LLC; ProQuest does not claim copyright in the individual underlying works; Last updated - 2021-06-04
- Song R, Keller AA, Suh S (2017) Rapid life-cycle impact screening using artificial neural networks. *Environ Sci Technol* 51(18):10777–10785. <https://doi.org/10.1021/acs.est.7b02862>
- Sousa I, Wallace D (2006) Product classification to support approximate life-cycle assessment of design concepts. *Technol Forecast Soc Change* 73(3):228–249. <https://doi.org/10.1016/j.techfore.2004.03.007>. <https://www.sciencedirect.com/science/article/pii/S0040162505000934>, technological Change and the Environment
- Sousa I, Wallace D, Eisenhard JL (2000) Approximate life-cycle assessment of product concepts using learning systems. *J Ind Ecol* 4(4):61–81. <https://doi.org/10.1162/10881980052541954>. <https://onlinelibrary.wiley.com/doi/abs/10.1162/10881980052541954>
- Wallace RJ, Marvuglia A, Benetto E, Tiruta-Barna L (2014) A new approach to optimization with life cycle assessment: Combining optimization with detailed process simulation. In: Murgante B, Misra S, Rocha AMAC, Torre CM, Rocha JG, Falcão MI, Taniar D, Apduhan BO, Gervasi O (eds) *Computational Science and Its Applications - ICCSA 2014 - 14th International Conference*, Guimarães, Portugal, June 30 - July 3, 2014, Proceedings, Part III, Springer, Lecture Notes in Computer Science, vol 8581. pp 707–720
- Wang E, Shen Z (2013) Lifecycle energy consumption prediction of residential buildings by incorporating longitudinal uncertainties. *J Civ Eng* 19(sup1):S161–S171. <https://doi.org/10.3846/13923730.2013.802744>
- Wisthoff A, Ferrero V, Huynh T, DuPont B (2016). Quantifying the impact of sustainable product design decisions in the early design phase through machine learning. <https://doi.org/10.1115/DETC2016-59586>
- Wu Y, Wang G (2018) Machine learning based toxicity prediction: from chemical structural description to transcriptome analysis. *Int J Mol Sci* 19(8):2358. <https://doi.org/10.3390/ijms19082358>. <https://pubmed.ncbi.nlm.nih.gov/30103448>
- Xia LM, Liu JJ (2013) Research on green building assessment system based on BP neural network and life cycle assessment (LCA). In: *Architecture, Building Materials and Engineering Management*, Trans Tech Publications Ltd, Applied Mechanics and Materials, vol 357. pp 508–514. <https://doi.org/10.4028/www.scientific.net/AMM.357-360.508>
- Xikai M, Lixiong W, Jiwei L, Xiaoli Q, Tongyao W (2019) Comparison of regression models for estimation of carbon emissions during building's lifecycle using designing factors: a case study of residential buildings in Tianjin, China. *Energy Build* 204:109519. <https://doi.org/10.1016/j.enbuild.2019.109519>. <https://www.sciencedirect.com/science/article/pii/S0378778819313672>
- Zhou CC, Yin GF, Hu XB (2009) Multi-objective optimization of material selection for sustainable products: artificial neural networks and genetic algorithm approach. *Mater Des* 30(4):1209–1215. <https://doi.org/10.1016/j.matdes.2008.06.006>. <https://www.sciencedirect.com/science/article/pii/S0261306908002689>
- Zhu X, Ho CH, Wang X (2020) Application of life cycle assessment and machine learning for high-throughput screening of green chemical substitutes. *ACS Sustain Chem Eng* 8(30):11141–11151. <https://doi.org/10.1021/acssuschemeng.0c02211>
- Ziyadi M, Al-Qadi IL (2019) Model uncertainty analysis using data analytics for life-cycle assessment (LCA) applications. *Int J Life Cycle Assess* 24(5):945–959. <https://doi.org/10.1007/s11367-018-1528-7>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.