Environmental and Climate Technologies
2025, vol. 29, no. 1, pp. 243–258
https://doi.org/10.2478/rtuect-2025-0017
https://content.sciendo.com

Machine Learning Approach for Predicting Environmental Impact: A Neuro-Fuzzy Model for Life Cycle Impact Assessment of Strawberry Production

Maksims FEOFILOVS[1*], Majid ZAEEMI[2], Andrea CAPPELLI[3], Francesco ROMAGNOLI[4]

[1,4]*Institute of Energy Systems and Environment, Riga Technical University, Azenes iela 12/1, Riga, LV-1048, Latvia*

[2]*Bioeconomy in Transition Research Group, IDEA, Unitelma Sapienza - University of Rome, Viale Regina Elena, 295, 00161 Rome, Italy*

[3]*Department of Chemical Engineering Materials Environment (DICMA), Faculty of Civil and Industrial Engineering, Sapienza University of Rome, Via Eudossiana 18, Rome, 00184, Italy*

Received 04.04.2025; accepted 02.06.2025

Abstract **– Artificial Intelligence (AI) is transforming traditional methods reliant on human knowledge by introducing machine learning techniques, which offer effective solutions to complex challenges. An example of such a case is the evaluation of the environmental impacts of products throughout their life cycle. This study bridges the gap in life cycle assessment (LCA) by leveraging AI to predict environmental impacts in agriculture, specifically by using LCA data from one cultivation system to model another. We employed Adaptive Neuro-Fuzzy Inference Systems (ANFIS) to predict CO_2 equivalent emissions for open-field strawberry production, utilizing greenhouse strawberry data. The novelty lies in combining machine learning with LCA to address data scarcity and improve predictive accuracy in agricultural impact assessments. The model was trained with data generated in MATLAB and validated against emissions computed using the Ecoinvent 3.10 database and SimaPro software. Among the three fuzzy inference system (FIS) generation approaches - Fuzzy C-Means (FCM), Subtractive Clustering (SC), and Grid Partitioning (GP) FCM exhibited the highest the accuracy. This methodology showcases AI's potential to transform LCA, enabling more efficient, data-driven sustainability assessments.**

Keywords **– ANFIS; artificial intelligence; carbon footprint; global warming potential; sustainability.**

1. Introduction

Environmental challenges posed by industrial and agricultural activities necessitate robust tools to quantify and mitigate impacts. Life Cycle Assessment (LCA) is widely adopted for this purpose, offering comprehensive evaluations of environmental impacts across a product's life cycle. Despite its strengths, LCA's reliance on extensive, high-quality data creates

* Corresponding author.
E-mail address: maksims.feofilovs@rtu.lv

significant barriers, particularly for sectors like agriculture, due to global, regional, and local variations in management practices [1].

Machine learning techniques can provide new opportunities to deal with data access problems. Machine learning methods can be used to predict the best approximation for the missing data or to assist in the decision-making process by analyzing large data sets [2]. Machine learning applications could be used to define a set of benchmarks within green building rating systems based on LCA [3, 4]. Furthermore, machine learning techniques can be used to analyze large datasets derived from Environmental Product Declarations (EPDs) to determine consistent benchmarks for environmental performance across product categories [5], which is often an important step in setting criteria for green procurement schemes. Similarly, in the agricultural sector benchmarking of emission reduction credits from organically managed crop production can be mentioned as another case of potential machine learning application within the LCA context. Distinguishing regional variability in data will add complexity to benchmarking procedures. Thus, machine learning could support green public procurement initiatives by enabling more precise benchmarks for goods and services based on environmental performance [6]. By automating the analysis of EPD data, machine learning applications can help stakeholders identify products with lower environmental impacts [7].

This study introduces an innovative application of ML in LCA by employing Adaptive Neuro-Fuzzy Inference Systems (ANFIS) to predict Carbon dioxide equivalent (CO_2 eq) emissions for open-field strawberry production based on greenhouse data. This approach demonstrates a novel use of AI to address data scarcity in agriculture, advancing LCA methodology. By comparing three FIS generation methods – Fuzzy C-Means (FCM), Subtractive Clustering (SC), and Grid Partitioning (GP) – we identify optimal modeling strategies. Our findings underline AI's transformative potential in making LCA more accessible and precise, supporting sustainable decision-making in agriculture and beyond.

2. MACHINE LEARNING APPLICATIONS IN LCA

Machine learning is a methodology for analysing and drawing inferences from data based on statistics and it includes various algorithms for different types of questions [8]. Three primary categories of machine learning are distinguished: supervised, unsupervised, and reinforcement learning [9]. Supervised learning refers to when algorithms are used to model the relationships between one or more independent variables and a dependent variable. Unsupervised learning involves algorithms that identify patterns within datasets or categorize data into groups. Reinforcement learning algorithms develop optimized strategies or actions based on a defined set of constraints or a reward system [10].

Supervised learning algorithms require the modeler to provide a dataset of independent variables and a dependent variable. Independent variables are called features, and dependent variables are called targets. With enough input data, a supervised learning algorithm can produce a model of the relationship between features and labels that is precise enough to be used in real life [10]. On the other hand, unsupervised learning algorithms operate without data labels. The algorithm sorts data points into groups or finds structure within a dataset. In the reinforcement learning approach, there is no training dataset. The algorithm tries different actions or pathways within a system of rewards, goals, or constraints that provide feedback on the quality of answers [9].

The search results in Scopus scientific database (see Fig. 1) shows that number of publications including "machine learning" AND "life cycle assessment" keywords has steadily increased from 2010 to 2023. Notably, there has been a significant surge since 2020, with number of publications rapidly increasing from around 500 in 2020 to over 3000 in 2023. The database statistics show that approximately 50 per cent of publications are related to the sectors of Energy, Engineering, and Environmental Science.

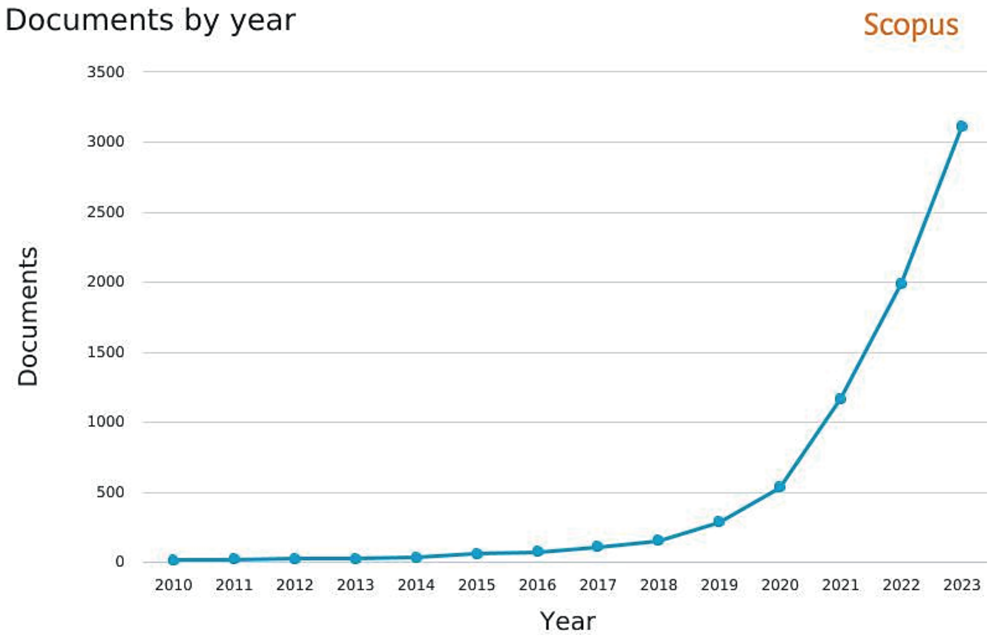

Fig. 1. Number of publications including "machine learning" AND "life cycle assessment" keywords.

Furthermore, Ghoroghi *et al.*, 2022 [2], Xue *et al.*, 2024 [11] and Barros & Ruschel, 2021 [12] provide comprehensive reviews of publications on machine learning applications in LCA, highlighting the depth and diversity of research in this rapidly evolving field. Studies have demonstrated machine learning effectiveness in predicting environmental impacts, identifying key impact contributors, and applying it to topics related to agriculture, building design, and clean energy systems. Machine models have improved the application of LCA, facilitating more informed decision-making on topics related to sustainability metrics.

2.1. Machine Learning Applications for Life Cycle Inventory

Machine learning has been applied to improve the prediction of data for LCI across many disciplines and demonstrated usefulness for estimating environmental emissions. *OrganoRelease* model [13] was developed to estimate the release factors of organic chemicals into various environmental compartments using an Artificial Neural Network (ANN) model. This model is significant for its ability to predict chemical emissions from the use and post-use of consumer products. Similarly, another study [14] focused on the building sector, developing a surrogate ANN to optimize building renovation scenarios. Their model estimates energy consumption, life cycle cost, and environmental impacts, providing a comprehensive tool for life cycle inventory optimization.

A slightly different approach reported in the literature [15] involves a linear regression model to predict emissions from dual fuel technology. In this case, data gaps in the LCI are filled for dual fuel diesel engines in oilfield operations by estimating Greenhouse Gas (GHG) emissions, non-methane hydrocarbons, and carbon monoxide. In the field of activated carbon production, [16] an ANN integrated with *Aspen Plus Simulation* is reported to estimate the energy use and carbon footprint of activated carbon derived from biomass. The model enhances the LCA framework by accurately predicting the yield of the final product from various biomass feedstocks.

A surrogate-based optimization framework with ANN models in the study of Nguyen *et. al*., 2019 [17] was developed to estimate greenhouse gas emissions and nutrient leaching in agricultural landscapes. The work focused on irrigated corn production systems in eastern Colorado, USA, providing high spatial resolution estimates that support resource management and environmental impact assessments.

Thilakarathna *et al*., 2020 [18], explored the optimization of concrete mixes to minimize embodied carbon values. By comparing five machine learning models, they were able to predict the compressive strength of high-strength concrete and calculate the associated carbon footprints. Cheng, Porter *et al*., 2020 [19], further contributed to the field by developing machine learning models to predict product yields and characteristics from the hydrothermal treatment of various feedstocks, enabling the calculation of energy consumption and GHG emissions.

The study by Cornago *et al*., 2020 [20], focused on the energy sector by forecasting the hourly power generation mix for different energy sources. Their model feeds this data into an LCA framework to estimate the carbon emissions of the energy mix, thus supporting more sustainable production scheduling. Naseri *et al*., 2020 [21] aimed to design sustainable concrete mixtures by developing an ANN model to predict the compressive strength of concrete. This model also considers cost, energy consumption, and life cycle carbon emissions, providing a holistic approach to sustainable concrete mix design.

2.2. Machine Learning Applications for Life Cycle Impact Assessment

2.2.1. Building sector

Several studies have demonstrated the potential of machine learning to optimize design and minimize environmental footprints within the building sector. For instance, Płoszaj-Mazurek *et al*., 2020 [22] applied convolutional neural networks and parametric design to optimize the carbon footprint of buildings, developing a supplementary tool that enables rapid estimates of environmental impacts during the early design phase. Similarly, Xikai *et al*., 2019 [23] focused on predicting life cycle carbon emissions for residential buildings in Tianjin, China, using regression models, emphasizing the importance of early-stage impact assessments. D'Amico *et al*., 2019 [24] went a step further by creating a decision support tool with ANN to estimate the energy and environmental performance of buildings quickly and reliably, streamlining the design process while ensuring sustainability.

Duprez *et al*., 2019 [25] advanced life cycle-based exploration methods by combining sensitivity analysis with metamodels, introducing ANN models to predict the global warming potential of new building designs. The work highlights the role of machine learning in facilitating informed, sustainable choices early in the design process, contributing to greener building practices.

2.2.2. Agriculture sectors

The agriculture sector is one of the main areas of application directions of machine learning techniques to predict LCA outcomes, particularly in estimating environmental impacts like GHG emissions and nutrient leaching. For example, a study by [26] explored the environmental impact of strawberry production, comparing open field and greenhouse cultivation in Iran using ANN and adaptive neuro-fuzzy inference systems (ANFIS). Their study demonstrated how machine learning can help quantify the environmental benefits of different agricultural practices. Similarly, [27] applied ANN models to predict the environmental impacts of potato production, while another study [28] focused on rice farms.

Machine learning techniques have also been applied to other crops. Khanali *et al*., 2017 [29] used ANN models to predict the yield and life cycle impacts of black tea, green tea, and oolong tea production in Iran, emphasizing energy inputs and environmental consequences. Similarly, study by Mousavi-Avval *et al*., 2017 [30] combined LCA with ANFIS to model energy use and environmental emissions for canola production, demonstrating the broad applicability of these methods to various crops.

In an international context, Vlontzos and Pardalos, 2017 [31] used ANN models combined with Data Envelopment Analysis (DEA) to predict GHG emissions from agricultural production in EU countries, extending historical data to future projections. This approach showcased the ability of machine learning to provide forward-looking environmental assessments. Nabavi-Pelesaraei *et al*., 2018 [32] further demonstrated this by integrating AI methods with LCA to predict energy output and environmental impacts for paddy production.

More recent studies continue to explore the intersection of machine learning and agriculture under changing environmental conditions. For example, Lee *et al*., 2020 [33] used boosted regression tree models to predict the life cycle impacts of corn production in the US Midwest under future climate scenarios, while Romeiko *et al*., 2020 [34] compared six different machine learning techniques to assess the GWP and eutrophication impacts from corn production. Lastly, Kaab *et al*., 2019 [35] applied LCA with artificial intelligence to predict the energy output and environmental impacts of sugarcane production, using both ANN and ANFIS models to evaluate different sugarcane farming systems in Iran.

2.2.3. Other sectors

In other sectors, machine learning has demonstrated successful application results when applied to LCIA. For example, Ozbilen *et al*., 2013 [36] developed an ANN model for nuclear-based hydrogen production, predicting GWP and acidification potential. Similarly, Marvuglia *et al*., 2015 [37] applied machine learning to predict the toxicity of organic chemical emissions, enhancing the accuracy of chemical fate models and streamlining the process for rapid LCA predictions.

In a related effort, Slapnik *et al*., 2015 [38] provided normalization factors for traditional impact categories and additional categories specific to Slovenia. By computing and comparing chemical characterization factors (CFs) within Slovenia and across Europe, the work highlights regional variations and provides a basis for more localized LCA predictions. Both studies show the importance of machine learning in reducing complexity and improving the accuracy of chemical impact predictions. Hou *et al*., 2020 [39] further advanced the field by developing machine learning models to estimate ecotoxicity characterization factors for 552 chemicals in the USEtox database. By predicting hazardous concentrations, their study fills critical data gaps, enhancing the accuracy of chemical CFs and supporting more comprehensive ecotoxicity assessments.

For the mining industry Asif *et al*., 2019 [40] created an integrative model combining life cycle inventory and ANN to manage air pollution from metal mining systems. The model estimates the carbon footprint by considering GHG emissions, fuel consumption, and operational data, offering a comprehensive tool for mitigating the environmental impacts of mining activities. Machine learning has also been used to refine chemical impact assessments. For instance, Song *et al*., 2017 [41] introduced a novel ANN-based approach for rapid screening of chemical life-cycle impacts, which is crucial for timely decision-making in chemical production. Zhu *et al.*, 2020 [42] took this further by applying ANN models to identify green chemical substitutes with minimal environmental impacts, particularly in the pharmaceutical industry.

2.3. Machine Learning Applications for Life Cycle Interpretation

The application of machine learning techniques in life cycle interpretation is helping to provide significant insights into results of LCA studies, identifying key drivers of impacts, and managing uncertainties. Azari *et al.*, 2016 [43] explored the optimization of building designs to achieve superior energy and environmental performance of buildings. The study employed a hybrid approach combining ANN and Genetic Algorithms (GA) to facilitate multi-objective optimization. In a similar way, Zhao *et al.*, 2019 [44] used a random forest model to identify the primary drivers of household carbon footprints among herdsmen in Inner Mongolia, China.

Supervised machine learning techniques to assess sustainability in food consumption sectors are reported in the study of Abdella *et al.*, 2020 [45]. By applying centroid-based clustering and logistic regression, different food industries are classified and analysed. The approach helped to identify key aspects influencing sustainability, enhancing the ability to manage food consumption impacts.

Ziyadi and Al-Qadi, 2019 [46] focused on model uncertainty analysis for life cycle assessments. The work combined an ANN surrogate model with interval, Bayesian, and model correction analysis methods to estimate uncertainties in the life cycle assessment of a pavement section in Chicago. The methodology highlighted the importance of addressing uncertainties to improve the reliability of LCA results.

The study of Romeiko *et al.*, 2020 [47] examined the spatially and temporally explicit life cycle impacts of soybean production in the US Midwest. Using boosted regression trees, the study identified key contributors to the life cycle impacts, such as soil, weather, and farming practices. This allowed the study to explain the significance of spatial and temporal variations in environmental impacts to enhance supply chain designs.

A framework for the optimal integration of solar-assisted district heating systems in urban communities was developed by Abokersh *et al.*, 2020 [48]. The authors applied global sensitivity analysis with machine learning and created a decision-making tool that supports the transition to clean energy in Europe. The use of ANN as a surrogate model for TRNSYS in this study enabled computationally intensive optimizations, demonstrating the potential of machine learning in complex system integrations.

This literature review illustrates the diverse applications of machine learning in life cycle interpretation and how machine learning models contribute to more effective life cycle assessment, thus supporting decision-making related to sustainability questions across various disciplines.

3. CASE STUDY OF NEURO-FUZZY MODEL APPLICATION FOR LIFE CYCLE IMPACT ASSESSMENT

Using MATLAB, data for 33 greenhouse producers was generated to train the adaptive neuro-fuzzy inference systems ANFIS model. The *Ecoinvent* 3.10 database and SimaPro software were utilized to calculate GHG emissions as CO_2 eq emissions, which serve as the model's target data. This section also covers fuzzy logic, fuzzy inference systems (FIS), and adaptive neuro-fuzzy inference systems (ANFIS), highlighting the methodology and performance evaluation of different ANFIS models in predicting CO_2 eq emissions for both greenhouse and open-field strawberry production systems.

3.1. Machine Learning Approach

Lotfi Zadeh [49] created the concept of fuzzy logic during the 1960s. Fuzzy logic is a computing approach based on "degrees of truth" rather than the binary "true or false" (0 or 1) Boolean logic. His research primarily focused on the problem of computer understanding of natural language, which is not easy to translate into the absolute terms of 0 and 1. Fig. 2 shows the difference between Boolean and Fuzzy logic.

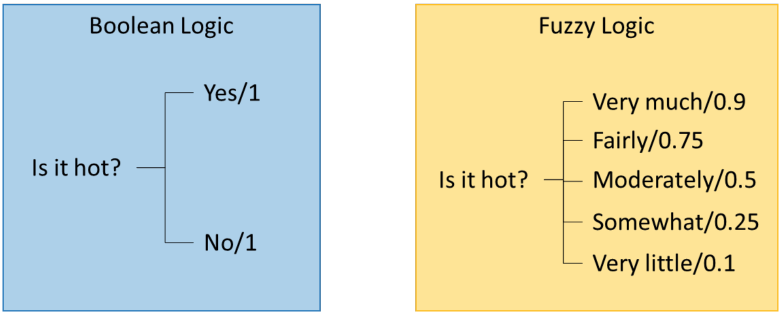

Fig. 2. Boolean logic vs. Fuzzy logic.

System modelling using conventional mathematical tools, such as differential equations, is often inadequate for handling uncertain systems. The FIS captures the qualitative aspects of human knowledge and reasoning processes, eliminating the need for precise quantitative analysis. This fuzzy modelling or fuzzy identification, was first explored systematically by Takagi and Sugeno [50]. FIS, presented in Fig. 3, involves interpreting values of input vector based on some sets of rules, assigning values to the output vector. The truth of a statement is expressed in degrees rather than as a binary true or false value.

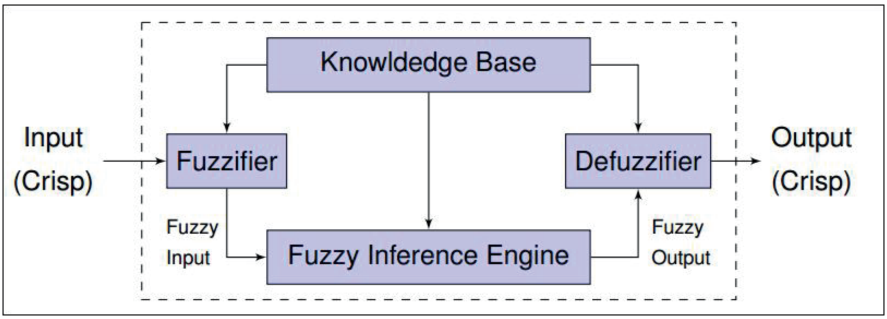

Fig. 3. A diagram of fuzzy inference system.

According to [52], FIS structure is explained through four functional blocks:

- Rule base containing several fuzzy *if-then* rules and a database that defines the membership functions of the fuzzy sets used in the fuzzy rules that are embedded in the Knowledge Base;

- Decision-making unit which performs the inference operations on the rules (Fuzzy Inference Engine);
- Fuzzification interface which transforms the inputs into degrees of match with linguistic values (Fuzzifier);
- Defuzzification interface which transforms the fuzzy results of the inference into an output (Defuzzifier).

Three approaches are used to generate the FIS in the model, which are Fuzzy C-Means (FCM), Subtractive Clustering (SC), and Grid Partitioning (GP). These approaches are defined in MATLAB as genfis1 (GP), genfis2 (SC), and genfis3 (FCM).

Fuzzy clustering is an unsupervised method for analysing data and constructing models [52]. Objects on the boundaries between several classes are not forced to fully belong to one of the classes. Instead, membership degrees are assigned between 0 and 1. The FCM employs fuzzy partitioning such that a data point can belong to all groups with different grades between 0 and 1.

Clustering algorithms require the user to specify the number of cluster centres and their locations. Thus, the quality of the solution depends strongly on the choice of initial values. Yager and Filev [53] proposed a simple and effective algorithm, called the mountain method, for estimating the number and initial location of cluster centers. Their method is based on gridding the data space and computing a potential value for each grid point based on its distances to the actual data points [54]. A grid point with many data points nearby will have a high potential value. The grid point with the highest potential value is chosen as the first cluster center. When the first cluster center is chosen, the potential of all grid points is reduced according to their distance from the cluster center. Grid points near the first cluster center will have greatly reduced potential. The next cluster center is then placed at the grid point with the highest remaining potential value. This procedure of acquiring a new cluster center and reducing the potential of surrounding grid points repeats until the potential of all grid points falls below a threshold.

In GP method, data are divided into grids, based on the type and number of member functions in each dimension [55]. ANFIS-GP begins with zero output and progressively learns the different fuzzy set rules and functions through the training procedure.

3.2. Development of Adaptive Neuro-Fuzzy Inference System (ANFIS)

An adaptive network is a network structure consisting of nodes and directional links [51]. Part or all of the nodes are adaptive, which means their outputs depend on the parameter(s) of the nodes, and the learning rule specifies how these parameters should be changed to minimize a prescribed error measure.

According to [56], ANFIS is a rule-based system with three components: membership functions of input–output variables, fuzzy rules, and output characteristics and system results. The system has a feedforward neural network structure and each layer is a neuro-fuzzy system component capable of learning and generalizing from the training data. This approach allows to find non-linear relationships between inputs and outputs.

In this study fuzzy logic toolbox of MATLAB version: 9.12.0.2039608 (R2022a-academic use) was used with the following inputs: diesel fuel, electricity, irrigation water, machinery, nitrogen, phosphate, potassium, farmyard manure, pesticides and plastic (nylon). The kg of CO_2 eq emissions is considered as the output of the model. Input clustering is used to divide the inputs into four groups. Results are obtained by aggregating the contribution of each input to the total kg of CO_2 eq emissions as seen in Fig. 4.

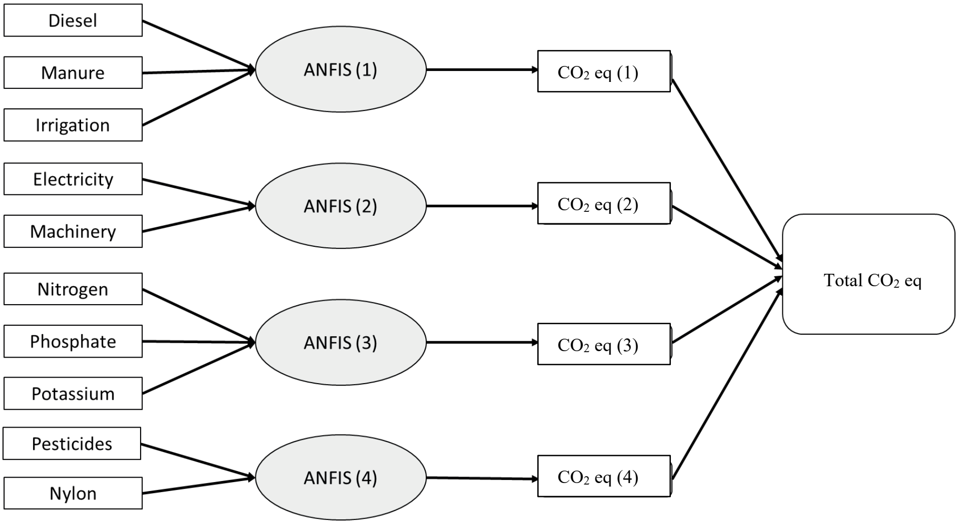

Fig. 4. ANFIS structure to predict kg of CO_2 eq of greenhouse strawberry production.

Five adjustments were made in the structure of ANFIS network to define configuration would produce the most accurate model with the lowest errors. The adjustments were made on the number of membership functions, types of membership functions (triangular, trapezoidal, bell-shaped, Gaussian and sigmoid), types of output membership function (constant or linear), optimization methods (hybrid or back propagation) and the number of epochs. After this, the model performance was validated using a statistical parameter namely root mean squared error (*RMSE*) Eq. (1):

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(predicted - actual\right)^2} \quad (1)$$

The *RMSE* is a measure of how close the ANFIS predicted values is to the observed one.

3.3. Input Data Generation

The input data for this study were obtained from a study that analysed the environmental impacts of open-field and greenhouse strawberry production in Guilan Province in the north of Iran. Among strawberry producers in the region, 33 greenhouse holders and 77 open field growers were randomly selected and surveyed. The data for open-field and greenhouse production systems are summarized in Table 1.

In the next step, 10 out of 12 inputs were selected and the data for 33 greenhouse producers generated using average and standard deviation of the sample. This data is further used as training data in the ANFIS model. Data generation was accomplished using MATLAB version: 9.12.0.2039608 (R2022a-academic use). Labor and microelements were excluded due to the lack of clarity in the details provided in the source literature.

In order to calculate the target data to be used in the model, *Ecoinvent* 3.10 database and SimaPro software were employed. The generated data were used as input for SimaPro analysis and then the amount of CO_2 eq was calculated using IPCC 2021 GWP100 V1.02.

TABLE 1. LIFE CYCLE INVENTORY DATA FOR STRAWBERRY PRODUCTION (PER 1 TON OF PRODUCED STRAWBERRIES)

Inputs	**Greenhouse**			**Open field**	
	Units	**Average**	**SD**	**Average**	**SD**
Machinery	kg	94.8	40.8	1059.3	578.2
Labor	h	1183	356.4	570.8	219.6
Diesel fuel	L	3.9	2.2	5.56	3.9
Electricity	kWh	450.8	227.9	24.7	5.2
Chemical fertilizers					
Nitrogen (N)	kg	9.8	3.9	40.7	13.8
Phosphate (P_2O_5)	kg	29.4	11.8	38.7	15.1
Potassium (K_2O)	kg	15.2	2.7	48.1	24.6
Microelements	kg	2.7	1.2	–	–
Farmyard manure	kg	955.6	405.7	1018.7	345.3
Pesticides	kg	1	0.3	6.4	2.5
Water for irrigation	m^3	358.8	185.5	189.1	39.5
Plastic	kg	27.58	23.5	–	–

4. RESULTS OF STRAWBERRY PRODUCTION CASE STUDY

To build the model and predict the CO_2 eq emissions for the production of one ton of strawberries in a greenhouse system, four Adaptive Neuro-Fuzzy Inference Systems (ANFIS) were developed utilizing the grouped inputs. Each FIS component within ANFIS has its own inputs and outputs. There are three inputs for ANFIS (1), namely, Diesel, Machinery, and Irrigation. The output is CO_2 eq related to the inputs which has been calculated using IPCC GWP100 v1.02. The rule view showing the mechanism of membership functions related to each amount of inputs, and the surface which shows the changes of each input on the CO_2 eq emission.

There are five layers in the ANFIS structure which are input, input membership functions, rule, output membership functions, and output.

The model initiates a dialogue box in which to input values for each emission-contributing factor in strawberry production, computes the CO_2 eq emissions, and displays the calculated amount.

To assess the accuracy of the model's predictions, the Root Mean Square Error (*RMSE*) for each ANFIS is illustrated in Fig. 5. A comparison between the output and predicted (target) data indicates that the Fuzzy C-Means (FCM) approach is more effective for ANFIS (1) and ANFIS (2), whereas the Grid Partitioning approach demonstrates better performance for ANFIS (3) and ANFIS (4). Further details regarding the network structures and their respective *RMSE* values can be found in Table 2.

In order to predict the equivalent CO_2 eq emissions associated with open-field strawberry production, the average values specific to this production system, as outlined in Table 1, were input into the model using various FIS generation approaches. Table 3 illustrates that the equivalent CO_2 eq emissions related to open-field production, predicted by the FCM approach, closely align with the calculated CO_2 eq values obtained using the IPCC GWP100 v1.02 method via SimaPro software.

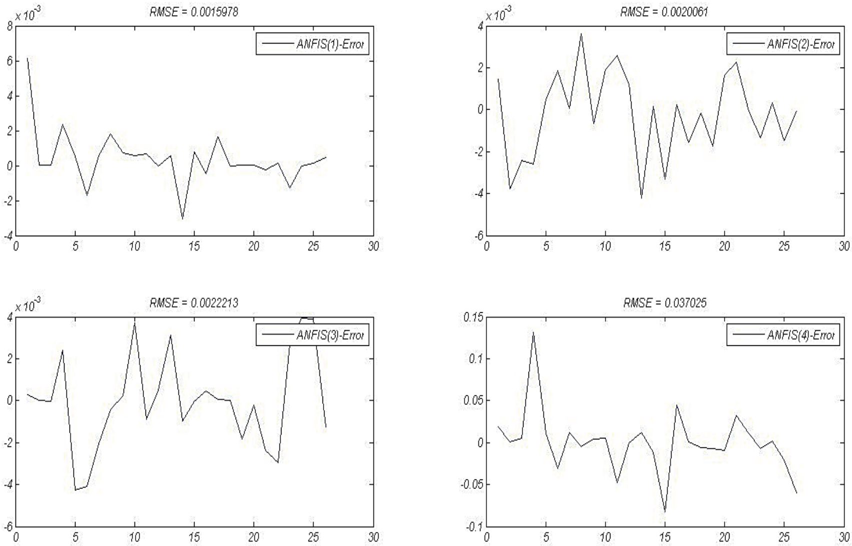

Fig. 5. RMSE for ANFIS (1) to ANFIS (4) using FCM approach.

TABLE 2. CHARACTERISTICS OF THE STRUCTURE OF EACH NETWORK AND THEIR RMSE

Model	FIS generation approach	Type of mf		Number of mf		Learning Method	RMSE
		Input	Output	Input	Epochs		
ANFIS (1)	FCM	gaussmf	linear	5	100	Hybrid	$1.597e^{-3}$
ANFIS (2)	FCM	gaussmf	linear	5	100	Hybrid	$2.006e^{-3}$
ANFIS (3)	FCM	gaussmf	linear	5	100	Hybrid	$2.221e^{-3}$
ANFIS (4)	FCM	gaussmf	linear	5	100	Hybrid	$3.702e^{-2}$
ANFIS (1)	Sub. Clust.	gaussmf	linear	5	100	Hybrid	$1.635e^{-2}$
ANFIS (2)	Sub. Clust.	gaussmf	linear	5	100	Hybrid	$2.945e^{-2}$
ANFIS (3)	Sub. Clust.	gaussmf	linear	5	100	Hybrid	$2.279e^{-3}$
ANFIS (4)	Sub. Clust.	gaussmf	linear	5	100	Hybrid	$7.819e^{-3}$
ANFIS (1)	Grid Part.	gaussmf	linear	5	100	Hybrid	$5.268e^{-3}$
ANFIS (2)	Grid Part.	gaussmf	linear	5	100	Hybrid	$7.836e^{-3}$
ANFIS (3)	Grid Part.	gaussmf	linear	5	100	Hybrid	$1.560e^{-3}$
ANFIS (4)	Grid Part.	gaussmf	linear	5	100	Hybrid	$1.688e^{-3}$

TABLE 3. PREDICTED (MATLAB) AND CALCULATED (SIMAPRO) VALUES FOR GREENHOUSE AND OPEN FIELD STRAWBERRY PRODUCTION BY DIFFERENT APPROACHES

Equivalent CO_2 eq emissions (kg)	MATLAB			SimaPro
	FCM	SC	GP	IPCC GWP100 v1.02
Open field	7790.31	1444.85	1584.53	7784.30
Greenhouse	1604.43	1604.52	1596.29	1604.38

However, the Subtractive Clustering and Grid Partitioning approaches do not demonstrate efficiency in this production system. Conversely, for greenhouse production systems, all three approaches prove proficient, yielding predicted values that closely match those calculated using the IPCC GWP100 v1.02 method via SimaPro.

As can be seen in Table 4, agricultural machinery accounts for approximately 40 % of CO_2 eq emissions in the greenhouse strawberry production system, whereas it plays a more dominant role in open field system contributing to around 90 %. In the greenhouse system, electricity and nylon contribute 22 % and 16 % respectively to total CO_2 eq emissions, while their impacts are relatively insignificant in the open field system.

TABLE 4. CO_2 EQ EMISSION RELATED TO EACH INPUT ITEM

Input Items	CO_2 eq emission, kg	
	Greenhouse	**Open field**
Diesel {GLO}| market group for | APOS, S	1.61	2.29
Electricity, low voltage {GLO}| market group for | APOS, S	323.68	17.74
Agricultural machinery, tillage {GLO}| market for | APOS, S	584.91	6535.84
Nitrogen fertiliser, as N {GLO}| market for | APOS, S	111.90	464.75
Phosphate fertiliser, as P_2O_5 {GLO}| market for | APOS, S	64.15	84.44
Potassium chloride, as K_2O {GLO}| market for | APOS, S	6.93	21.94
Manure, solid, cattle {GLO}| market for | APOS, S	20.20	21.53
Pesticide, unspecified {GLO}| market for | APOS, S	10.25	65.58
Irrigation {GLO}| market group for | APOS, S	33.61	17.71
Nylon 6 {GLO}| market for | APOS, S	258.32	0.01
Calculated CO_2 eq emissions (IPCC GWP100 v1.02 method)	1415.57	7231.82

5. CONCLUSION

The study investigated the integration of a ML technique known as ANFIS into LCA for assessing environmental impacts. Recent literature on the application of machine learning in LCA showed a wide range of ML applications across the phases of Life Cycle Inventory, Life Cycle Impact Assessment, and Life Cycle Interpretation. The information from the Scopus database showed that machine learning techniques in LCA have emerged recently and can serve as support in managing uncertainties in LCA across many sectors, from agriculture to construction and energy systems. This also underlines a growing recognition of ML techniques for sustainability analysis.

A case study on the ability of ML to predict LCA outcomes for open-field strawberry production systems using data derived from LCA assessments was performed. The created ANFIS model demonstrated strong predictive performance for estimating GHG emissions in both greenhouse and open-field production systems. The FCM approach showed the highest accuracy, proving the efficiency of combining fuzzy logic with neural networks for the defined study context.

The study advances knowledge by demonstrating the feasibility of cross-system data application, paving the way for broader ML-based innovations in sustainability assessments. This integration of AI with LCA not only enhances predictive accuracy but also reduces the time and resources required for data collection and modeling. The practical implications extend beyond strawberries, offering a scalable methodology for other crops and systems. Future research should explore expanding this approach to incorporate additional environmental parameters and apply it across diverse agricultural and industrial contexts. By bridging the gap between AI and LCA, this study sets a foundation for more efficient, data-driven sustainability practices. Further investigations should consider extending model validation across a broader range of empirical datasets from open-field cultivation systems. This would help reflect real-world complexity more accurately and reinforce the reliability of cross-system predictions.

ACKNOWLEDGEMENT

The research has been done within Fundamental and Applied Research Project "Carbon farming Certification system: The transition towards result-based agriculture sector (CarbFarmS)", project No. lzp-2023/1-0055, funded by the Latvian Council of Science

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this work the author(s) used Curie service in order to meet the highest standards of academic communication. After using this service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

REFERENCES

[1] Fan J., Liu C., Xie J., Han L., Zhang C., Guo D., Niu J., Jin H., McConkey B. G. Life Cycle Assessment on Agricultural Production: A Mini Review on Methodology, Application, and Challenges. *International Journal of Environmental Research and Public Health* 2022:19(16):9817. https://doi.org/10.3390/ijerph19169817

[2] Ghoroghi A., Rezgui Y., Petri I. Advances in application of machine learning to life cycle assessment: A literature review. *The International Journal of Life Cycle Assessment* 2022:27:433–456. https://doi.org/10.1007/s11367-022-02030-3

[3] Schlegl F., Gantner J., Traunspurger R., Albrecht S., Leistner P. LCA of buildings in Germany: Proposal for a future benchmark based on existing databases. *Energy and Buildings* 2019:194:342–350. https://doi.org/10.1016/j.enbuild.2019.04.038

[4] Dong Y., Ng. S.T., Liu P. A comprehensive analysis towards benchmarking of life cycle assessment of buildings based on systematic review. *Building and Environment* 2021:204:108–162. https://doi.org/10.1016/j.buildenv.2021.108162

[5] Koyamparambath A., Adibi N., Szablewski C., Adibi S. A., Sonnemann G. Implementing Artificial Intelligence Techniques to Predict Environmental Impacts: Case of Construction Products. *Sustainability* 2022:14(6):3699. https://doi.org/10.3390/su14063699

[6] Harvey J., *et al*. (Eds.). (2020). Pavement, Roadway, and Bridge Life Cycle Assessment. Taylor & Francis Group. Proceedings of the International Symposium on Pavement. Roadway, and Bridge Life Cycle Assessment 2020. LCA 2020, Sacramento, CA, 3–6 June 2020.

[7] Tiwari M. K., *et al*. Role of Artificial Intelligence in Green Public Procurement. *Applications of Emerging Technologies and AI/ML Algorithms, Asset Analytics.* Springer 2023:361–371. https://doi.org/10.1007/978-981-99-1019-9_32

[8] Dunn J. B., Balaprakash P. Chapter 1 – Overview of Data Science and Sustainability Analysis. *Data Science Applied to Sustainability Analysis* 2021:1–14. https://doi.org/10.1016/B978-0-12-817976-5.00001-2

[9] Burkov A. The Hundred-Page Machine Learning Book. Quebec City: 2019.

[10] Cerulli G. *Fundamentals of Supervised Machine Learning: With Applications in Python, R, and Stata.* Springer, 2023. https://doi.org/10.1007/978-3-031-41337-7

[11] Xue R. X., Zhang X., Pang Y., Gao F., Xu M., Lin S., et al. (2024). A review of machine learning applications in life cycle assessment studies. *Science of The Total Environment* 2024:912:168969. https://doi.org/10.1016/j.scitotenv.2023.168969

[12] Barros N. N., Ruschel R. C. Machine Learning for Whole-Building Life Cycle Assessment: A Systematic Literature Review. *Proceedings of the 18th International Conference on Computing in Civil and Building Engineering* 2021:98:109–122. https://doi.org/10.1007/978-3-030-51295-8_10
[13] Tao M., Li D., Song R., Suh S., Keller A. A. OrganoRelease – A framework for modeling the release of organic chemicals from the use and post-use of consumer products. *Environmental Pollution* 2018:234:751–761. https://doi.org/10.1016/j.envpol.2017.11.058
[14] Sharif S. A., Hammad A. Simulation-Based Multi-Objective Optimization of institutional building renovation considering energy consumption, Life-Cycle Cost, and Life-Cycle Assessment. *Journal of Building Engineering* 2019:21:429–445. https://doi.org/10.1016/j.jobe.2018.11.006
[15] Meng F., LaFleur C., Wijesinghe A., Colvin, J. Data-driven approach to fill in data gaps for life cycle inventory of dual fuel technology. *Fuel* 2019:246:187–195. https://doi.org/10.1016/j.fuel.2019.02.124
[16] Liao M., Kelley S., Yao Y. Generating Energy and Greenhouse Gas Inventory Data of Activated Carbon Production Using Machine Learning and Kinetic Based Process Simulation. *ACS Sustainable Chemistry & Engineering* 2020:8(2):1252–1261. https://doi.org/10.1021/acssuschemeng.9b06522
[17] Nguyen T. H., Nong D., Paustian, K. Surrogate-based multi-objective optimization of management options for agricultural landscapes using artificial neural networks. *Ecological Modelling* 2019:400:1–13. https://doi.org/10.1016/j.ecolmodel.2019.02.018
[18] Thilakarathna P. S. M., Seo S., Baduge K. S. K., Lee H., Mendis P., Foliente G. Embodied carbon analysis and benchmarking emissions of high and ultra-high strength concrete using machine learning algorithms. *Journal of Cleaner Production* 2020:262:121281. https://doi.org/10.1016/j.jclepro.2020.121281
[19] Cheng F., Porter M. D., Colosi L. M. Is hydrothermal treatment coupled with carbon capture and storage an energy-producing negative emissions technology? *Energy Conversion and Management* 2020:203:112252. https://doi.org/10.1016/j.enconman.2019.112252
[20] Cornago S., Vitali A., Brondi C., Low J. S. C. Electricity Technological Mix Forecasting for Life Cycle Assessment Aware Scheduling. *Procedia CIRP* 2020:90:268–273. https://doi.org/10.1016/j.procir.2020.01.099
[21] Naseri H., Jahanbakhsh H., Hosseini P., Moghadas Nejad F. Designing sustainable concrete mixture by developing a new machine learning technique. *Journal of Cleaner Production* 2020:258:120578. https://doi.org/10.1016/j.jclepro.2020.120578
[22] Płoszaj-Mazurek M., Ryńska E., Grochulska-Salak M. Methods to Optimize Carbon Footprint of Buildings in Regenerative Architectural Design with the Use of Machine Learning, Convolutional Neural Network, and Parametric Design. *Energies* 2020:13(20):5289. https://doi.org/10.3390/en13205289
[23] Xikai M., Lixiong W., Jiwei L., Xiaoli Q., Tongyao W. 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 and Buildings* 2019:204:109519. https://doi.org/10.1016/j.enbuild.2019.109519
[24] D'Amico A., Ciulla G., Traverso M., Lo Brano V., Palumbo E. Artificial Neural Networks to assess energy and environmental performance of buildings: An Italian case study. *Journal of Cleaner Production* 2019:239:117993. https://doi.org/10.1016/j.jclepro.2019.117993
[25] Duprez S., Fouquet M., Herreros Q., Jusselme T. Improving life cycle-based exploration methods by coupling sensitivity analysis and metamodels. *Sustainable Cities and Society* 2019:44:70–84. https://doi.org/10.1016/j.scs.2018.09.032
[26] Khoshnevisan B., Rafiee S., Mousazadeh H. Environmental impact assessment of open field and greenhouse strawberry production. *European Journal of Agronomy* 2013:50:29–37. https://doi.org/10.1016/j.eja.2013.05.003
[27] Khoshnevisan B., Rafiee S., Omid M., Mousazadeh H., Sefeedpari, P. Prognostication of environmental indices in potato production using artificial neural networks. *Journal of Cleaner Production* 2013:52:402–409. https://doi.org/10.1016/j.jclepro.2013.03.028
[28] Khoshnevisan B., Rajaeifar M. A., Clark S., Shamahirband S., Anuar N. B., Mohd Shuib N. L., Gani, A. Evaluation of traditional and consolidated rice farms in Guilan Province, Iran, using life cycle assessment and fuzzy modeling. *Science of The Total Environment* 2014:481:242–251. https://doi.org/10.1016/j.scitotenv.2014.02.052
[29] Khanali M., Mobli H., Hosseinzadeh-Bandbafha H. Modeling of yield and environmental impact categories in tea processing units based on artificial neural networks. *Environmental Science and Pollution Research* 2017:24:26324–26340. https://doi.org/10.1007/s11356-017-0234-5
[30] Mousavi-Avval S. H., Rafiee S., Sharifi M., Hosseinpour S., Shah A. Combined application of Life Cycle Assessment and Adaptive Neuro-Fuzzy Inference System for modeling energy and environmental emissions of oilseed production. *Renewable and Sustainable Energy Reviews* 2017:78:807–820. https://doi.org/10.1016/j.rser.2017.05.002
[31] Vlontzos G., Pardalos P. M. Assess and prognosticate green house gas emissions from agricultural production of EU countries, by implementing DEA Window analysis and artificial neural networks. *Renewable and Sustainable Energy Reviews* 2017:76:155–162. https://doi.org/10.1016/j.rser.2017.03.054

[32] Nabavi-Pelesaraei A., Rafiee S., Mohtasebi S. S., Hosseinzadeh-Bandbafha H., Chau K. Integration of artificial intelligence methods and life cycle assessment to predict energy output and environmental impacts of paddy production. *Science of The Total Environment* 2018:631–632:1279–1294. https://doi.org/10.1016/j.scitotenv.2018.03.088

[33] Lee E. K., Zhang W.-J., Zhang X., Adler P. R., Lin S., Feingold B. J., et al. Projecting life-cycle environmental impacts of corn production in the U.S. Midwest under future climate scenarios using a machine learning approach. *Science of The Total Environment* 2020:714:136697. https://doi.org/10.1016/j.scitotenv.2020.136697

[34] Romeiko X. X., Guo Z., Pang Y., Lee E. K., Zhang X. Comparing Machine Learning Approaches for Predicting Spatially Explicit Life Cycle Global Warming and Eutrophication Impacts from Corn Production. *Sustainability* 2020:12(4):1481. https://doi.org/10.3390/su12041481

[35] Kaab A., Sharifi M., Mobli H., Nabavi-Pelesaraei A., Chau K. Combined life cycle assessment and artificial intelligence for prediction of output energy and environmental impacts of sugarcane production. *Science of The Total Environment* 2019:664:1005–1019. https://doi.org/10.1016/j.scitotenv.2019.02.004

[36] Ozbilen A., Aydin M., Dincer I., Rosen M. A. Life cycle assessment of nuclear-based hydrogen production via a copper–chlorine cycle: A neural network approach. *International Journal of Hydrogen Energy* 2013:38(15):6314–6322. https://doi.org/10.1016/j.ijhydene.2013.03.071

[37] Marvuglia A., Kanevski M., Benetto E. Machine learning for toxicity characterization of organic chemical emissions using USEtox database: Learning the structure of the input space. *Environment International* 2015:83:72–85. https://doi.org/10.1016/j.envint.2015.05.011

[38] Slapnik M., Istenič D., Pintar M., Udovč A. Extending life cycle assessment normalization factors and use of machine learning – A Slovenian case study. *Ecological Indicators* 2016:50:161–172. https://doi.org/10.1016/j.ecolind.2014.10.028

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

[40] Asif Z., Chen Z., Zhu Z. H. An integrated life cycle inventory and artificial neural network model for mining air pollution management. *International Journal of Environmental Science and Technology* 2019:16(4):1847–1856. https://doi.org/10.1007/s13762-018-1813-9

[41] Song R., Keller A. A., Suh S. Rapid Life-Cycle Impact Screening Using Artificial Neural Networks. *Environmental Science & Technology* 2017:51(18):10777–10785. https://doi.org/10.1021/acs.est.7b02862

[42] Zhu X., Ho C.-H., Wang X. Application of Life Cycle Assessment and Machine Learning for High-Throughput Screening of Green Chemical Substitutes. *ACS Sustainable Chemistry & Engineering* 2020:8(30):11141–11151. https://doi.org/10.1021/acssuschemeng.0c02211

[43] Azari R., Garshasbi S., Amini P., Rashed-Ali H., Mohammadi Y. Multi-objective optimization of building envelope design for life cycle environmental performance. *Energy and Buildings* 2016:126:524–534. https://doi.org/10.1016/j.enbuild.2016.05.054

[44] Zhao Y., Zhang Q., Li F. Y. Patterns and drivers of household carbon footprint of the herdsmen in the typical steppe region of Inner Mongolia, China: A case study in Xilinhot City. *Journal of Cleaner Production* 2019:232:408–416. https://doi.org/10.1016/j.jclepro.2019.05.351

[45] Abdella G. M., Kucukvar M., Onat N. C., Al-Yafay H. M., Bulak M. E. Sustainability assessment and modeling based on supervised machine learning techniques: The case for food consumption. *Journal of Cleaner Production* 2020:251:119661. https://doi.org/10.1016/j.jclepro.2019.119661

[46] Ziyadi M., Al-Qadi I. L. Model uncertainty analysis using data analytics for life-cycle assessment (LCA) applications. *The International Journal of Life Cycle Assessment* 2019:24(8):945–959. https://doi.org/10.1007/s11367-018-1528-7

[47] Romeiko X. X., Lee E. K., Sorunmu Y., Zhang X. Spatially and Temporally Explicit Life Cycle Environmental Impacts of Soybean Production in the U.S. Midwest. *Environmental Science & Technology* 2020:54(8):4758–4768. https://doi.org/10.1021/acs.est.9b06874

[48] Abokersh M. H., Vallès M., Cabeza L. F., Boer D. 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. *Applied Energy* 2020:267:114903. https://doi.org/10.1016/j.apenergy.2020.114903

[49] Zadeh L. A. Fuzzy sets. *Information and Control* 1965:8(3):338–353. https://doi.org/10.1016/S0019-9958(65)90241-X

[50] Takagi T., Sugeno M. Fuzzy identification of systems and its applications to modeling and control. *IEEE Transactions on Systems, Man, and Cybernetics* 1980:15(1):116–132. https://doi.org/10.1109/TSMC.1985.6313399

[51] Jang J.-S. R. ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics* 1993:23(3):665–685. https://doi.org/10.1109/21.256541

[52] Suganya R., Shanthi R. Fuzzy C-Means Algorithm: A review. *Semantics Scholar* 2012.

[53] Yager R. R., Filev D. P. Generation of fuzzy rules by mountain clustering. *Journal of Intelligent and Fuzzy Systems* 1994:2(3):209–219. https://doi.org/10.3233/IFS-1994-2301

[54] Bataineh K., Najia M., Saqera M. A comparison study between various fuzzy clustering algorithms. *Semantics Scholar* 2011.

[55] Mola M., Amiri-Ahouee R. ANFIS model based on fuzzy C-mean, grid partitioning, and subtractive clustering for detection of stator winding inter-turn fault in PM synchronous motor. *International Transactions on Electrical Energy Systems* 2021:31(3):e12770. https://doi.org/10.1002/2050-7038.12770

[56] Naderloo L., Alimardani R., Omid M., Sarmadian F., Javadikia P., Torabi M. Y., Alimardani F. Application of ANFIS to predict crop yield based on different energy inputs. *Measurement* 2012:45(6):1406–1413. https://doi.org/10.1016/j.measurement.2012.03.025