*Corresponding Author's Email: lisa.gobio-thomas@brunel.ac.uk
Proceedings of the World Conference on Climate Change and Global Warming
Vol. 1, Issue. 1, 2024, pp. 44-58
DOI: https://doi.org/10.33422/ccgconf.v1i1.346

ISSN: 3030-0703 online

The Use of Machine Learning in Predicting Future Environmental Impacts of the ASTEP Solar Thermal System: A Case Study

Lisa Baindu Gobio-Thomas[1*], Moschos Papananias[1,2], Mohamed Darwish[1,3] and Valentina Stojceska[1,3]
[1] Brunel University London, College of Engineering, Design and Physical Sciences, Uxbridge UB8 3PH, UK
[2] Brunel University London, Institute of Materials and Manufacturing, Uxbridge UB8 3PH, UK
[3] Brunel University London, Institute of Energy Futures, Centre for Sustainable Energy Use in Food Chains, Uxbridge UB8 3PH, UK

Abstract

An artificial neural network (ANN) model developed in MATLAB was used to predict the environmental performance of an innovative solar thermal system, ASTEP (Application of Solar Energy to Industrial Processes) over a 30-year period. The system was applied to the industrial processes of two end-users, Mandrekas (MAND) and Arcelor Mittal (AMTP). The ASTEP system was designed to supply thermal energy up to 400°C and consist of three main components: a novel rotary Fresnel Sundial, thermal energy storage (TES) and a control system. The actual GHG emissions of the ASTEP system and a solar thermal plant as presented in the literature were used to evaluate the ability of the ANN model to predict GHG emissions. The actual and predicted emissions were compared to assess the accuracy of the model. Validation results showed a difference of 2.13 kgCO2eq/kWh for AMTP's ASTEP system, 2.43 kgCO2eq/kWh for MAND's ASTEP system and 0.32 kgCO2eq/kWh for a third solar thermal plant. These findings indicate that the ANN model could be considered as an effective tool in predicting GHG emissions for solar thermal plants allowing the industry to evaluate their environmental performance and adopt measures to reduce their impact.

Keywords: GHG emissions, environmental impact, machine learning, solar thermal plants, artificial neural network (ANN) regression

1. Introduction

Energy demand and consumption has increased substantially over the years due to the growth in population, industrial activities and socio-economic development worldwide. Fossil fuels are still predominantly used to provide energy in different sectors across the world, releasing greenhouse gas emissions into the atmosphere causing air pollution, global warming and negative health impacts. To mitigate these negative impacts, governments across the world are under growing pressure to reduce GHG emissions and limit global warming to about 1.5ºC as declared in the United Nations Framework Convention on Climate Change (UNFCCC) and 2015 Paris Agreement (UNFCCC, 2024). Consequently, the EU Parliament and its member states have agreed to reduce carbon emissions by at least 55% by 2030, compared to 1990 levels, and to be climate neutral by 2050 (EU Commission, 2024a). To achieve this goal, the EU's 2030 climate and energy framework has a set a target to increase the share of renewable energy to 32% (EU Commission, 2024b). This has led to the provision of a number of incentives such as the "Horizon Europe", "Just Transition Mechanism" and "Innovation Fund" to EU member states to encourage the adoption of renewable energy technologies and the reduction of GHG emissions in industries and other sectors. The industrial sector accounts for the third largest share of energy consumption in the EU (EuroStat, 2024). From 1995 - 2019, the use of energy from fossil-fuels in the industrial sector decreased, while the use of energy from renewable energy systems (RES) increased by more than 60% within the same period (Brodny & Tutak, 2022). This indicates that industries are keen to reduce their environmental impact by using renewable energy for their processes. Solar thermal technologies are renewable energy systems which can be used to provide heating and cooling for industrial processes. This can result in the reduction of GHG emissions and contributes to the decarbonization of EU industries.

A number of studies have been conducted on applying machine learning (ML) techniques in predicting the environmental impact of renewable energy systems. Mujeeb & Javid (2022) assessed the impact of renewable energy generation on carbon emissions in the USA. The authors developed a novel forecasting model which first used Spearman Correlation Analysis (SCA) to select the best inputs for the training of the model, then an Improved Shallow Denoising Autoencoder (ISDAE) for the feature extractor and an Improved Particle Swarm Optimization (IPSO) based Deep Neural Network (DNN) to forecast the carbon emissions of renewable energy source (RES) generation. The results showed a strong correlation between non-renewable energy generation sources and carbon emissions reveling that higher level of RES penetration significantly reduces the carbon emissions. Ahmed et al. (2022) investigated the impact of renewable energy (RE), energy consumption, financial development, gross domestic product and population on CO_2 emissions in China and India. The authors used long short-term memory (LSTM) model and found that energy consumption had the greatest impact, while RE had the lowest impact on CO_2 emissions in both countries. Chukwunoso et al. (2024) predicted the impact of RE and non-renewable energy sources on the CO_2 emissions in USA using different ML algorithms; layered recurrent neural network (L-RNN), feed-forward neural network (FFNN) and convolutional neural network (CNN). The results showed that the L-RNN model outperformed the other models by achieving more accurate CO_2 emission predictions. The predictions indicated that the CO_2 emissions in the USA will increase by about twice the rate if the current trends continue based on their energy generation sources. Alshafeey & Rashdan (2023) predicted the impact of renewable & non-renewable energy sources GHG emissions for the USA, China and the EU. Gradient boosting was used to identify the major factors contributing to GHG emissions and artificial neural network was used to predict their GHG emissions. The predictions showed that increasing nuclear consumption by 25% in China, would lead to a 11% decrease in GHG emissions in China. The model predicted that increasing wind energy consumption by 25% would result in a 3% decrease in GHG emissions. In the EU, increasing coal consumption by 25% would lead to an 11% rise in their GHG emissions. Fang et al. (2024) used LSTM-recurrent neural network with Monte Carlo approach to predict the carbon savings and electricity generation of a concentrated solar thermal gasification of biomass (CSTGB) using a solar tower system and integrated with carbon capture storage (CCS). The model predicted the CSTGB system would achieve a savings of 415,960 tons of CO_2-eq over a 30 years lifespan, when carbon tax is included as revenue and a savings of 132,615 tons of CO_2-eq when carbon tax is excluded.

Each of the studies used different ML techniques to predict the environmental impact of the RE systems. Most of the RE systems assessed in these studies were nuclear, solar PV energy, wind, biomass, geothermal and hydro-energy, with only one study predicting environmental performance of a solar thermal plant.

A review of the studies in the literature that used ML to predict the performance of solar thermal plants showed that these studies used ML techniques to mainly predict the energy generation of the solar thermal plants. None of the studies predicted the GHG emissions of the solar thermal plants. Based on extensive literature review and to the best of the authors' knowledge, there were limited and scarce studies on the prediction of GHG emissions generation of solar thermal plants using ML. This highlights the need for more studies in this area. Therefore, the main objectives of this work are to predict the environmental performance of the newly developed solar thermal system over the next 30 years using ML and to identify the factors that will influence its performance.

1.1 Description of ASTEP system

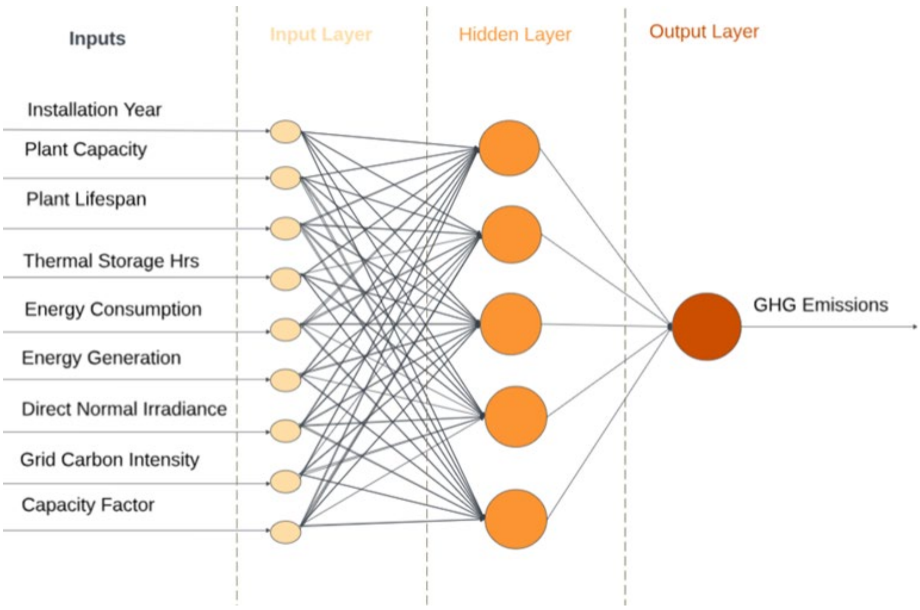
The ASTEP system as an innovative solar thermal technology has been assessed in this study. It consists of three main subsystems; the novel rotary Fresnel Sundial solar collector, the thermal energy storage and the control system. The Fresnel Sundial solar collectors simultaneously rotate around their longitudinal axis and reflect the solar radiation to two elevated receivers as the sun latitude changes, thereby heating up the thermal oil in its receiver tubes (Abbas et al., 2023). The heated thermal oil then flows to a thermal energy storage (TES) tank where the thermal energy is stored and released to the industrial processes when needed. The control system is based on a programmable logic control (PLC) unit and will ensure that the heat supply remains within the process specifications for the temperature, pressure and flow rates. The ASTEP system has a great potential to reduce GHG emissions in various industries by suppling renewable energy for high-temperature processes up to 400ºC to companies located at both low and high latitudes. Increasing the ASTEP system's current capacity of 27.2 MWh to a future prospective capacity of 1828 MWh for AMTP, results in higher annual GHG emissions savings of 559.4 tonnes of CO_2 emissions. Likewise, increasing its current capacity of 27.4 MWh for MAND to a future prospective capacity of 2158 MWh, results in higher annual GHG emissions savings of 909.6 tonnes of CO_2 emissions. This demonstrates the potential GHG emissions reduction of the ASTEP system when used to provide thermal energy to industries at large capacities (Gobio-Thomas et al., 2022).

The ASTEP system has been applied to two end users; Mandrekas (MAND) and ArcelorMittal (AMTP). MAND is a family owned dairy company in Corinth, Greece and is located in a region at a low latitude of 37.93N. The company produces different types of yogurt, yogurt-based dressings and milk desserts. Their production process requires temperatures of up to 175ºC for milk pasteurization and 5ºC for the refrigeration of the dairy products (ASTEP, 2024). The ASTEP system will be used to provide thermal energy for the pasteurization and cooling processes. AMTP is the world's leading steel company and its metal processing plant is located in Iasi, Romania, at a high latitude of 47.1N. The company manufactures welded steel tubes for a wide range of applications. The steel tubes are colour coated and in order to apply this coating on the tubes, the steel tubes need to be pre-heated to a temperature of 220ºC. Therefore, the ASTEP system will be used at AMTP to provide thermal energy to preheat the manufactured tubes before their colour coating process (ASTEP, 2024).

1.2 Artificial Neural Network (ANN)

An Artificial Neural Network (ANN) consists of an input layer, one or more hidden layers and an output layer. In Figure 1, the input layer includes the nine-inputs of the ANN model. The output layer produces the predicted GHG emissions. Between the input and output layers is the hidden layer which consists of a number of neurons and performs computations on the input data. Apart from the nodes in the input layer, all other nodes in the hidden and output layers represent a neuron. The number of neurons in the hidden layer is determined by trial and error during simulations. The size of input and output layers is determined by the dimensions of the input and output data, respectively. Each neuron in the hidden layer receives signals from all inputs, applies a weighted sum, adds a bias term, then passes the result through an activation function and provides a feedforward path to the output layer (Maind & Wankar, 2014).

Figure 1: Artificial Neural Network Architecture



2. Methodology

In this study, supervised learning has been used, where the ANN processes the input data and compares the resulting and actual outputs of GHG emissions. A backpropagation algorithm is then used to reduce the error by adjusting the connection weights of the network and the algorithm is stopped when the predicted GHG emissions matches the actual GHG emission result or reaches the desired accuracy (Thorat, Pandit & Balote, 2023). This study uses neural network regression model to predict the GHG emissions of the solar thermal plants. This is because this type of regression model can analyze complex nonlinear relationships, has a high fault tolerance and a strong ability for information synthesis (Zhou, 2022).

2.1 Data Collection Process

Studies on the environmental assessment of solar thermal plants published between 2010 – 2024 were selected from scientific databases such as Scopus, Science Direct & Web of Science. The studies were carefully reviewed on a case-by-case basis and the relevant studies were selected for detailed analyses and extraction of the specific details of the solar thermal plants were recorded, as presented in Table 1. This included the plant's year of installation, plant capacity, direct normal irradiation (DNI) of the plant's location, its annual energy consumption, annual energy generation, thermal energy storage hours, capacity factor, plant's life span and its greenhouse gas (GHG) emissions.

2.2 Database Compilation

The variables selected in Table 1 have been reported in the literature that they have an impact on GHG emissions and are related to the solar thermal plants. Bosnajakovic & Tadjanovic (2019) stated that the size or capacity of a solar thermal plant is one of the various factors that can influence its environmental impact. Studies have found that the DNI affects the environmental profile of solar thermal power plants with plants located in places with high DNI level achieving lower environmental impact than plants in lower DNI levels (Corona et al., 2016; Guillen-Lambea & Carvalho, 2021). Whitaker et al. (2013) found that the energy consumption of solar thermal plants mainly through its operational and maintenance (O&M) phase was the largest contributor to the GHG emissions of the solar thermal plant. The amount of annual energy generated by a solar thermal plant can impact on its environmental profile as reported by Guillen-Lambea & Carvalho (2021) who states that higher energy production of the plant, results in lower GHG emissions. Gasa et al. (2021) also identified the use of TES as one of the variables that can influence the environmental performance of a solar thermal plant. The authors found that plants with TES have lower environmental impact due to their reduced operational impact. Guillen-Lambea & Carvalho (2021) and Lamnatou & Chemisana (2017) reported that the lifespan of a solar thermal power plant affects its GHG emissions. Klein & Rubin (2013) and Gasa et al. (2021) stated that the capacity factor influences the environmental impact of a solar thermal plant. Scarlat et al. (2022) found that the carbon intensity of grid electricity can impact on the GHG emissions of a solar thermal plant as high carbon intensity results in higher plant emissions. Carbon intensity data for grid electricity for Europe and America was used in the GHG prediction database (Enerdata, 2024). Table 1 presents the database of the variables used in the GHG prediction model. The current technical and environmental performance data of MAND and AMTP's ASTEP system was also included in the database, resulting in a total of 32 solar thermal power plants used in the database (Table 1).

Table 1: Database of variables used in the GHG prediction model

References	Year of Installation	Capacity Factor (%)	Plant Capacity (MW)	DNI (kwh/m^{2}yr)	Energy Consumption (GJ)	Energy Generation (MWh)	TES (hrs)	Plant Life Span (years)	Carbon Intensity (gCO_2eq/kWh)	GHG Emissions (kgCO_2eq/kWh)
Burkhardt et al. (2011)	2008	0.47	110	2920	340000	426700	6.3	30	493.8	72
Burkhardt et al. (2011)	2008	0.49	110	2920	27000	438800	6.3	30	493.8	35
Burkhardt et al. (2011)	2010	0.47	103	2700	170680	438800	6.3	30	477	26
Burkhardt et al. (2011)	2010	0.47	103	2700	188684	426700	6.3	30	477	28
Klein & Ruben (2013)	2010	0.51	103	2700	213350	443000	6	30	477	39
Klein & Ruben (2013)	2010	0.33	103	2700	157879	288000	6	30	477	24
Whitaker et al. (2013)	2012	0.42	106	2600	185447	378463	6	30	447.4	37
Whitaker et al. (2013)	2012	0.42	103	2700	183481	378463	6	30	447.4	34
Whitaker et al. (2013)	2012	0.42	103	2700	179214	378463	6	30	447.4	33
Whitaker et al. (2013)	2012	0.42	103	2700	209083	378463	6	30	447.4	36
Whitaker et al. (2013)	2012	0.42	103	2700	213350	378463	6	30	447.4	38
Whitaker et al. (2013)	2012	0.42	103	2700	179214	378463	6	30	447.4	32
Asdrubali et al. (2013)	2012	0.63	20	2100	25300	110000	15	25	307	14.2
Asdrubali et al. (2013)	2012	0.36	50	2100	57600	160000	7.5	25	307	20.6
Corona et al. (2014)	2012	0.38	50	2030	189946.8	165687	7.5	25	307	26.6
Corona et al (2014)	2012	0.4	50	2030	319165	174407	7.5	25	307	67.2

References	**Year of Installation**	**Capacity Factor (%)**	**Plant Capacity (MW)**	**DNI (kwh/m^2yr)**	**Energy Consumption (GJ)**	**Energy Generation (MWh)**	**TES (hrs)**	**Plant Life Span (years)**	**Carbon Intensity (gCO_2eq/kWh)**	**GHG Emissions ($kgCO_2eq/kWh$)**
Corona & San Miguel (2015)	2012	0.38	50	2030	191865.5	165687	7.5	25	307	26.9
Corona & San Miguel (2015)	2012	0.38	50	2030	191865.5	165687	7.5	25	307	34
Corona (2016a)	2013	0.38	50	2030	196662.1	165687	7.5	25	301	27.6
Corona et al. (2016b)	2013	0.91	100	2086	620782	797423	14	25	301	45.9
Rodriguez-Serrano et al. (2017)	2015	0.45	100	2600	197850	395700	5	30	289	24
Rodriguez-Serrano et al. (2017)	2017	0.5	100	2600	197850	395700	7	30	295	22.4
Gasa et al. (2021)	2016	0.81	110	3332	175000	776240	17.5	30	272.8	31
Corona et al. (2016b)	2016	0.8	100	2687	556949	735732	14	25	301	28.8
Corona et al. (2016b)	2016	0.8	100	2026	1043243	800647	14	25	301	45.3
Corona et al. (2016b)	2016	0.8	100	2686	908722	863804	14	25	301	37.4
Ko et al. (2018)	2018	0.6	101	2900	288800	656000	12	30	290	24.3
Li et al. (2019)	2019	0.57	10	2150	25700	50000	15	25	651.7	35
Banacloche et al. (2020)	2020	0.4	1	1922	611.5	2052	0	25	559.2	22
Backes et al. (2021)	2021	0.4	0.033	1933	5644	46	0	25	339.3	34.8
MAND ASTEP system (2024)	2024	0.5	0.025	1783	7632	27.8	4	30	336.6	26
AMTP ASTEP system (2024)	2024	0.43	0.025	1497	15227	27.8	4	30	291.8	34.2

2.3 Multiple Linear Regression & ANN Model

Multiple linear regression analysis and the development of the ANN model was performed using MATLAB R2022a (MathWorks, Portola Valley, California). Multiple linear regression analysis of the input variables was conducted to identify the tatistically significant variables that influences the GHG emissions of the solar thermal plants. Neural network regression model was developed to predict the GHG emissions of the ASTEP solar thermal systems for MAND and AMTP industrial processes for a period of 30 years and the methodology is shown in Figure 2. The compiled dataset shown in Table 1 was uploaded into MATLAB and hyperparameters such as the number of hidden layers, number of neurons in the hidden layer, number of epochs and training algorithm were selected. The ANN structure used in this study comprised of 9 inputs with 1 hidden layer as illustrated in Figure 1. The dataset was split into a training set and a test set using a random nonstratified partition with a 90% - 10% split. Feedforward ANN architectures, such as the Multi-Layer Perceptron (MLP) network, have been extensively applied in machine learning applications to model complex non-linear data. Therefore, an MLP network with hyperbolic tangent sigmoid transfer functions and automated regularisation was used for the ANN model. The number of hidden layers and neurons selected for the ANN model was determined by trial and error during simulations, which showed that 1 hidden layer with 5 neurons prevented overfitting and achieved higher accuracy level of the prediction results. In this modelling problem, a large number of hidden neurons will increase the risk of overfitting. The maximum number of training epochs was set to 1000, but the model was trained until it reached convergence. The root means square error (RMSE) value for the training data was 3.2 and the RMSE value of the test data was 1.87. To assess the performance of the model, Figures 3a and 3b show the collected-predicted plots. Overfitting occurs when the model performs well on the training set, but poorly on the test set, i.e., the training accuracy is significantly higher than the testing accuracy. It can be seen, in Figures 3a and 3b, that there is no overfitting on the training data and the model is able to generalise well to unseen data.

Figure 2: Prediction methodology

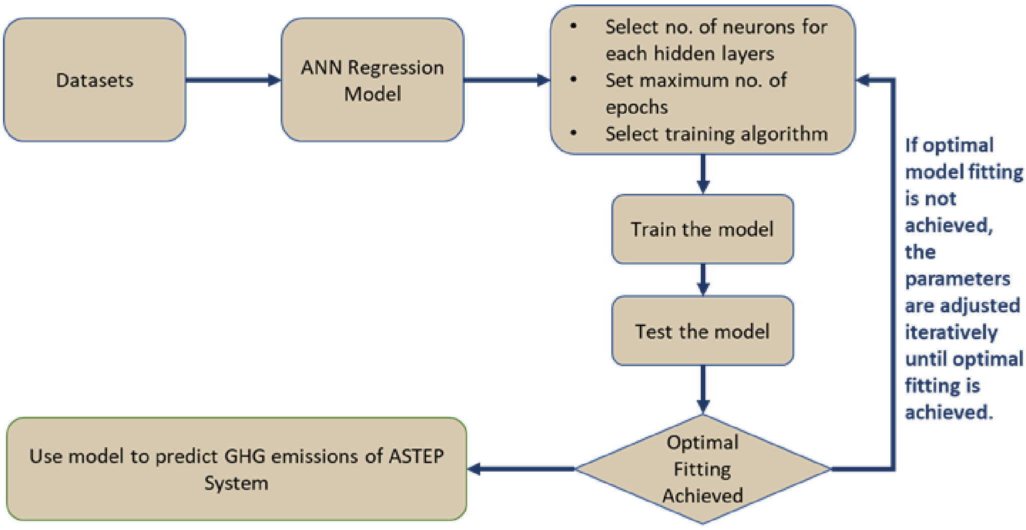


Figure 3a: Training Graph of the ANN model

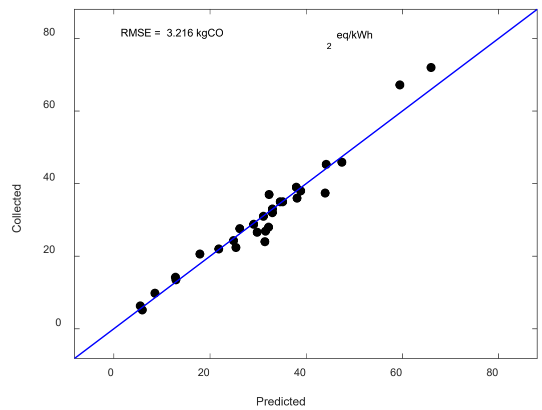
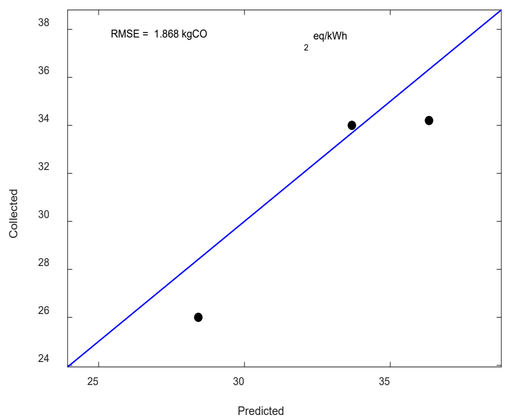


Figure 3b: Testing Graph of the ANN model



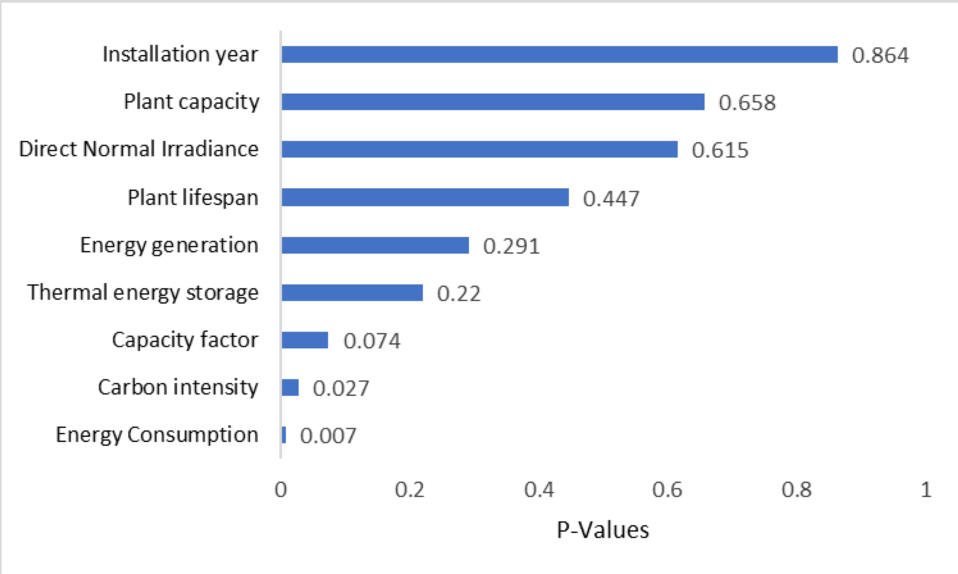
3. Results & Discussion

This section shows the results and discussion of the multiple linear regression analysis of the input variables of the ANN model as welll as the ANN model prediction results for the training and test set.

3.1 Multiple Linear Regression Analysis

The results from multiple linear regression analysis demonstrated that the coefficient of determination (R^2) is 69.09% and the adjusted one is 56.45% for the whole dataset. This suggests that the model fits the data moderately well. Multiple linear regression analysis was also conducted on the input variables of the model. The results showed that the statistically significant variables for 95% confidence level are the energy consumption (p-value = 0.007) and the carbon intensity (p-value = 0.027) as presented in Figure 4. This indicates that energy consumption and carbon intensity have a stronger influence on the GHG emissions of the plant compared to the other input variables of the ANN model. Studies have found that the energy consumption from the manufacturing and operational phases of solar thermal plants is responsible for most of the GHG emissions of the plants (San Miguel & Corona, 2014; Li et al., 2019; Mihoub et al., 2019). The GHG emissions produced during the manufacturing of the plant is a singular event that doesn't occur after the plant has been constructed. However, the emissions from the operational phase is continuous throughout the life span of the plant, typically around 30 years. Therefore, more attention should be given to the GHG emissions generated during the operation of the solar thermal plant from the electricity used to operate its pumps, controls and other auxiliary equipment. This is supported by Mihoub et al. (2019) who found that the energy consumption of the solar thermal power plant's fuel backup system contributed over 90% of the total GHG emissions of the plant. Lamnatou & Chemisana (2017) reported that energy consumption from the operation and maintenance needs of the solar thermal plant is one of the factors that influences its environmental profile. Most solar thermal plants use grid electricity or natural gas as the auxiliary fuels to operate pumps and controls and to prevent freezing of the heat transfer fluid and thermal energy storage fluid of the plants. The broader impact of the findings in this study is for the solar thermal industry to use renewable energy options instead of electricity and natural gas as their auxiliary fuel for the operation of solar thermal plants.

Figure 4. Multiple linear regression results of model's input variables



3.2 ANN model prediction for training & test set

Figures 5a and 5b show the collected (actual) responses along with the predicted responses for the training and test set. The model was tested using 3 samples as presented in Table 2; Sample 1 is AMTP ASTEP system, Sample 2 is another solar thermal plant and Sample 3 is MAND's ASTEP system. It can be seen from Figures 5a and 5b that the predicted GHG emission values of the model are close to the actual emission values in both the training and test sets. Table 2 shows the actual and predicted GHG emissions of the three test samples. The actual GHG emissions of the solar thermal system for Sample 2 in Table 2, was collected from the environment assessment results of a solar thermal plant conducted by Corona & San Miguel (2015). The GHG emissions of AMTP's ASTEP system (Sample 1) and MAND's ASTEP system (Sample 3) were obtained from environmental life cycle assessment of these plants using SimaPro software. The ANN model predicted GHG emissions of 36.33 kgCO_2eq/kWh and 28.42 kgCO_2eq/kWh for AMTP and MAND's ASTEP system for a period of 30 years, respectively. The actual GHG emissions is 34.2 kgCO_2eq/kWh for AMTP and 26 kgCO_2eq/kWh for MAND's ASTEP system. Therefore, the difference in the actual and predicted values is 2.13 kgCO_2eq/kWh for AMTP's ASTEP system and 2.42 kgCO_2eq/kWh for MAND's ASTEP system. The model predicted GHG emissions of 33.68 kgCO_2eq/kWh for the other solar thermal plant in the test sample. The actual GHG emissions of this plant is 34 kgCO_2eq/kWh. Therefore, the difference in the actual and predicted values by the model for this solar thermal plant is 0.32 kgCO_2eq/kWh. This suggests that the model is able to generalise well to unseen data and could be used to predict the GHG emissions of other solar thermal plants.

There were some studies in the literature that predicted the GHG emissions of renewable energy systems. However, these were mainly for solar PV, wind, biomass, geothermal and hydro-energy systems and not for solar thermal plants. There were limited studies in the literature on the application of ML to predict the GHG emissions of solar thermal plants. Most of the studies in the literature used ML to predict the energy generation output of the solar thermal plants and not their GHG emissions. Therefore, the development of an ANN model to predict the GHG emissions of the ASTEP solar thermal system could not be directly compared with other studies in the literature. The closest study found in the literature was by

Fang et al. (2024) who used LSTM-recurrent neural network to predict the carbon savings and electricity generation of a solar thermal plant integrated with carbon capture storage. However, Fang et al. (2024) predicted the carbon savings and not the GHG emissions of the plant. This study provides a unique contribution to the limited body of literature on the application of ML to predict the GHG emissions of solar thermal plants, in particular rotary Fresnel solar thermal plants. The broader implication of the findings of this study on environmental impact assessment practices is the need to use predictive models such as this newly developed ANN model to forecast and assess the environmental performance of solar thermal plants and adopt suitable strategies and measures to minimize their environmental impact.

Figure 5a: ANN model predictions for the training set

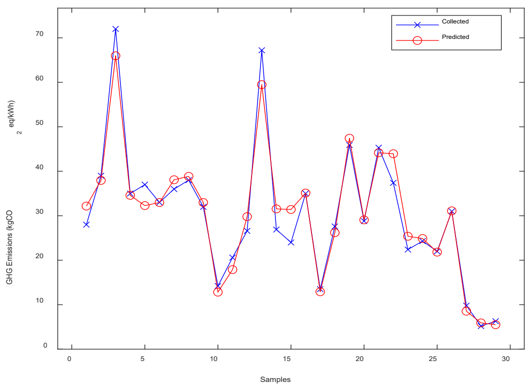


Figure 5b: ANN model predictions for the test set

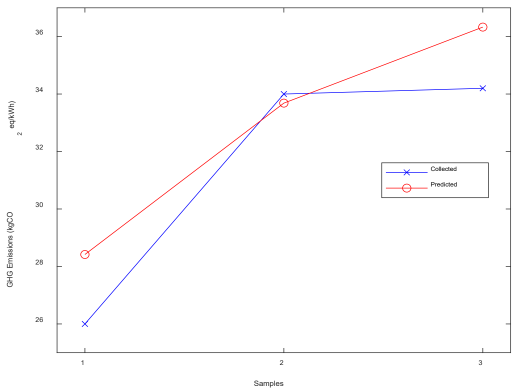


Table 2. Actual & Predicted GHG emissions of test set

Test Samples	Actual GHG emissions ($kgCO_2eq/kWh$)	Predicted GHG emissions ($kgCO_2eq/kWh$)	Residuals ($kgCO_2eq/kWh$)
Sample 1 (MAND)	26	28.42	2.42
Sample 2	34	33.68	0.32
Sample 3 (AMTP)	34.2	36.33	2.13

4. Conclusion

This paper used a newly developed ANN model to predict the GHG emissions of a newly developed rotary Fresnel ASTEP solar thermal system over 30 years. The results of the multiple linear regression analysis showed that the energy consumption and carbon intensity were the highest statistically ($P<0.05$) significant variable influencing the GHG emissions. Energy consumption from the operation of solar thermal plants have been found as one of the factors that significantly reduces ($P<0.05$) on their GHG emissions. Renewable energy should be used to replace electricity and natural gas which are usually the auxiliary fuels used for these plants. These measures will help to reduce the environmental impact of solar thermal plants.

The practical implications of the findings in this study is for the solar thermal industry to use predictive models such as this newly developed ANN model to predict the GHG emissions of solar thermal plants. This will enable the industry to predict the environmental performance of the newely developed solar plants and implement suitable measures to reduce their environmental impact. The practical implication for policy makers is the need to provide subsidies for renewable energy options to make them more environmentally friendly and cost-effective than fossil fuels. This could encourage the solar thermal industry to use renewable energy as their auxiliary fuel instead of fossil fuels, thereby reducing the GHG emissions of the operational phases of the solar thermal plants. The practical implications for researchers in the field of solar thermal energy is the need for more studies to be conducted on the development of machine learning models to forecast the environmental impact of solar thermal plants due to the lack of research in this area. This study contributes to the very limited studies in the literature on the application of machine learning to predict the GHG emissions of solar thermal plants. It also identifies the variables that strongly influence the GHG emissions of the plants. This can facilitate the development of effective strategies to improve the environmental performance of solar thermal plants by targeting the key factors influencing its GHG emissions. The main contributions of this work are the assessment of factors that influence the environmental impact of solar thermal plants and development of an ANN model to predict the GHG emissions of the newly developed ASTEP solar thermal system.

5. Recommendation

The current database should be expanded to include more data on the technical and environmental performance of small capacity solar thermal plants ranging from 25kW to 100kW. This will help to increase the accuracy of the ANN model to predict the GHG emissions of the ASTEP system and other small capacity plants. Future research includes training the ANN model from datasets based only on small capacity plants of 25kW to 100kW to achieve a more accurate prediction of the GHG emissions of ASTEP system. Another recommendation for future research is to use a variety of ML techniques and compare their prediction accuracy results. The ML algorithm with the highest prediction accuracy should then be used to forecast the GHG emissions of the solar thermal plant. This will help to improve the accuracy of the model in predicting the GHG emissions of the rotary Fresnel solar thermal plant.

Acknowledgment

This research is funded by the EU Horizon 2020 research and innovation programme, Application of Solar Energy in Industrial processes (ASTEP), grant number 884411.

References

Abbas, R., Barbera, R., Rovira, A. and Barnetche, M. (2023). Sundial, a new collector for solar heat for industrial processes: Optical and thermal design. *Thermal Science & Engineering Progress,* 44, 102025. https://doi.org/10.1016/j.tsep.2023.102025

Ahmed, Mansoor., Shuai, Chuanmin., & Ahmed, Maqsood. (2022). Influencing factors of carbon emissions and their trends in China and India: a machine learning method. *Environmental Science and Pollution Research,* 29, 48424–48437. https://doi.org/10.1007/s11356-022-18711-3.

AlShafeey, M., & Rashdan, O. (2023). Quantifying the impact of energy consumption sources on GHG emissions in major economies: A machine learning approach. *Energy Strategy Reviews*, 49, 101159. https://doi.org/10.1016/j.esr.2023.101159.

ASTEP. (2024, June 2). *Case studies*. Retrieved June 8, 2024, from https://astepproject.eu/case-studies/

Backes, J. G., D'Amico, A., Pauliks, N., Guarino, S., Traverso, M., lo Brano, V. (2021). Life Cycle Sustainability Assessment of a Dish-Stirling Concentrating Solar Power Plant in the Mediterranean Area. Sustainable Energy Technologies and Assessments, 47. https://doi.org/10.1016/j.seta.2021.101444.

Banacloche, S., Herrera, I. and Lechón, Y. (2020). Towards Energy Transition in Tunisia: Sustainability Assessment of a Hybrid Concentrated Solar Power and Biomass Plant. Science of the Total Environment, 744. https://doi.org/10.1016/j.scitotenv.2020.140729.

Bosnjakovic, M., & Tadijanovic, V. (2019). Environment Impact of a Concentrated Solar Power Plant. *Technical Journal,* 13, 68-74.

Britannica. (2024, June 7). *Mean square error*. Retrieved June 7, 2024, from https://www.britannica.com/science/mean-squared-error

Brodny, J., & Tutak, M. (2022). Analysis of the efficiency and structure of energy consumption in the industrial sector in the European Union countries between 1995 and 2019. *Science of the Total Environment*, 808, 152052. https://doi.org/10.1016/j.scitotenv.2021.152052.

Chukwunonso, B, P., AL-Wesabi, I., Shixiang, L., Khalil AlSharabi, K., Al-Shamma, A, A., Hassan M., Farh, H, M, H., Saeed, F., Kandil, T., & Al-Shaalan, A, M. (2024). Predicting carbon dioxide emissions in the United States of America using machine learning algorithms. Environmental Science and Pollution Research, 31, 33685–33707. https://doi.org/10.1007/s11356-024-33460-1.

Corona, B.; Cerrajero, E., López, D. and San Miguel, G. (2016a). Full Environmental Life Cycle Cost Analysis of Concentrating Solar Power Technology: Contribution of Externalities to Overall Energy Costs. *Solar Energy*, *135*, pp. 758–768. https://doi.org/10.1016/j.solener.2016.06.059.

Corona, B., Ruiz, D. and San Miguel, G. (2016b). Life Cycle Assessment of a HYSOL Concentrated Solar Power Plant: Analyzing the Effect of Geographic Location. Energies, 9, pp.413. https://doi:10.3390/en9060413.

Enerdata. (2024, May 21). *Energy & emissions projections 2050 – EnerOutlook, 2023 Edition*. Retrieved May 25, 2024, from https://eneroutlook.enerdata.net/forecast-world-co2-intensity-of-electricity-generation.html

European Commission. (2024a, May 25). *2030 Climate Target Plan*. Retrieved May 28, 2024, from https://ec.europa.eu/clima/eu-action/european-green-deal/2030-climate-target-plan_en.

European Commission. (2024b, May 27). *The Just Transition Mechanism: making sure no one is left behind.* Retrieved May 29, 2024, from https://ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal/finance-and-green-deal/just-transition-mechanism_en.

EuroStat. (2024, May 10). *Energy Statistics- an overview*. Retrieved May 15, 2024, from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Energy_statistics_-_an_overview#Final_energy_consumption.

Fang, Y., Li, X., Wang, X., Dai, L., Ruan, R., You, S. (2024). Machine learning-based multi-objective optimization of concentrated solar thermal gasification of biomass incorporating life cycle assessment and techno-economic analysis. *Energy Conversion and Management*, 302, 118137. https://doi.org/10.1016/j.enconman.2024.118137

Gasa, G., Lopez-Roman, A., Prieto, C., & Cabeza, L. F. (2021). Life Cycle Assessment (LCA) of a Concentrating Solar Power (CSP) Plant in Tower Configuration with and without Thermal Energy Storage (TES). *Sustainability*, 13, 3672. https://doi.org/10.3390/su13073672

Gobio-Thomas, L.B., Abbas, R., Darwish, M., & Stojceska, V. (2022). Decarbonization of industrial processes using novel Sundial solar thermal units. In 2nd World Conference on Sustainability, Energy & Environment. https://www.doi.org/10.33422/2nd.wscee.2022.12.130

Gobio-Thomas, L.B., Darwish, M., & Stojceska, V. (2023). Environmental impacts on solar thermal power plants used in industrial supply chains. *Thermal Science and Engineering Progress*, 38, 101670. https://doi.org/10.1016/j.tsep.2023.101670

Guillén-Lambea, S., & Carvalho, M. (2021). A critical review of the greenhouse gas emissions associated with parabolic trough concentrating solar power plants. *Journal of Cleaner Production*, 289, 125774. https://doi.org/10.1016/j.jclepro.2020.125774

Klein, S. J. W. & Rubin, E. S. (2013). Life cycle assessment of greenhouse gas emissions, water and land use for concentrated solar power plants with different energy backup systems. *Energy Policy*, 63, 935–950. https://doi.org/10.1016/j.enpol.2013.08.057

Ko, N., Lorenz, M., Horn, R., Krieg, H. and Baumann, M. (2018). Sustainability Assessment of Concentrated Solar Power (CSP) Tower Plants - Integrating LCA, LCC and LCWE in One Framework. Procedia CIRP, 69, pp. 395–400. https://doi.org/10.1016/j.procir.2017.11.049.

Lamnatou, C. & Chemisana, D. (2017). Concentrating solar systems: Life Cycle Assessment (LCA) and environmental issues. *Renewable and Sustainable Energy Reviews,* 78, pp.916 – 932. https://doi.org/10.1016/j.rser.2017.04.065

Li, R., Zhang, H., Wang, H., Tu, Q. and Wang, X. (2019). Integrated hybrid life cycle assessment and contribution analysis for CO2 emission and energy consumption of a concentrated solar power plant in China. *Energy,* 174, pp.310-322. https://doi:10.1016/j.energy.2019.02.066.

Maind, S. B. & Wankar, P. (2014). Research Paper on Basic of Artificial Neural Network. *International Journal on Recent and Innovation Trends in Computing and Communication*, 2 (1). https://ijritcc.org/index.php/ijritcc/article/view/2920

Mihoub, S. (2019). Design, economic, and environmental assessments of linear Fresnel solar power plants. *Environmental Progress & Sustainable Energy,* pp.13350. https://doi.org/10.1002/ep.13350

Mujeeb, S., & Javai, N. (2021). Deep learning based carbon emissions forecasting and renewable energy's impact quantification. *IET Renewable Power Generation*, 17, pp.873–884. https://doi.org/10.1049/rpg2.12641

San Miguel, G. & Corona, B. (2014). Hybridizing Concentrated Solar Power (CSP) with Biogas and Biomethane as an Alternative to Natural Gas: Analysis of Environmental Performance Using LCA. *Renewable Energy*, *66*, pp. 580–587. https://doi.org/10.1016/j.renene.2013.12.023.

Scarlat, N., Prussi, M., & Padella, M. (2022). Quantification of the carbon intensity of electricity produced and used in Europe. *Applied Energy,* 305, 117901. https://doi.org/10.1016/j.apenergy.2021.117901

Thorat, M., Pandit, S.V., & Balote, S. (2023). Artificial Neural Network: A brief study. *International Research Journal of Engineering and Technology (IRJET),* 10, 2. http://dx.doi.org/10.33130/AJCT.2022v08i03.003

UNFCCC. (2024, June 5). *About the Secretariat*. Retrieved June 9, 2024, from https://unfccc.int/about-us/about-the-secretariat?gad_source=1&gclid=Cj0KCQjwsaqzBhDdARIsAK2gqndk30AAng1K6MuQyObc-naslUq0PXNu--krdUyQosHXq5ldy_dAJagaArGDEALw_wcB

Whitaker, M.B., Heath, G.A., Burkhardt, J.J., & Turchi, C.S. (2013). Life Cycle Assessment of a Power Tower Concentrating Solar Plant and the Impacts of Key Design Alternatives. *Environmental Science & Technology*, 47, 5896-5903. https://doi.org/10.1021/es400821x

Zhou, S. (2022). An analysis of the small sample datasets based on machine learning. Proceedings of the 6th International Conference on Electronic Information Technology & Computer Engineering. Xiamen, China, 21-23 October. https://doi.org/10.1145/3573428.3573720