Process Safety and Environmental Protection 184 (2024) 1119–1130

Contents lists available at ScienceDirect

Process Safety and Environmental Protection

journal homepage: www.journals.elsevier.com/process-safety-and-environmental-protection

Machine learning assisted prediction of solar to liquid fuel production: a case study

Muhammad Wakil Shahzad [a,*], Viet Hung Nguyen [b], Ben Bin Xu [a], Rasikh Tariq [c], Muhammad Imran [d], Waqar Muhammad Ashraf [e], Kim Choon Ng [f], Muhammad Ahmad Jamil [a], Amna Ijaz [a], Nadeem Ahmed Sheikh [g]

[a] Mechanical & Construction Engineering Department, Northumbria University, Newcastle Upon Tyne NE1 8ST, UK
[b] Utility Department, Heineken Vietnam Brewery, Hà Nội, Viet Nam
[c] Tecnologico de Monterrey, Institute for the Future of Education, Ave. Eugenio Garza Sada 2501, Monterrey, N.L., 64849, Mexico
[d] Department of Mechanical, Biomedical and Design Engineering, Aston University, Birmingham B4 7ET, UK
[e] The Sargent Centre for Process Systems Engineering, Department of Chemical Engineering, University College London, Torrington Place, London WC1E 7JE, UK
[f] Water Desalination and Reuse Center, King Abdullah University of Science and Technology, Saudi Arabia
[g] Department of Mechanical Engineering, International Islamic University, Islamabad, Pakistan

ARTICLE INFO

Keywords:
Solar energy
Liquid fuel
Formic acid
Two-axis solar tracking system

ABSTRACT

In this era of heightened environmental awareness, the global community faces the critical challenge of climate change. Renewable energy (RE) emerges as a vital contender to mitigate global warming and meet increasing energy needs. Nonetheless, the fluctuating nature of renewable energy sources underscores the necessity for efficient conversion and storage strategies. This pioneering research focuses on the transformation of solar energy (SE) into liquid fuels, with a specific emphasis on formic acid (FA) as a case study, done in Binh Thuan, Vietnam. The paper unveils a technology designed to convert solar energy into formic acid, ensuring its stability and storage at ambient conditions. It involves detailed simulations to quantify the daily and monthly electricity output from photovoltaic (PV) systems and the corresponding mass of formic acid producible through solar energy. The simulation of a dual-axis solar tracking system for the PV panels, intended to maximize solar energy capture, is one of the project's illustrations. The elevation and azimuth angles, which are two essential tracking system parameters, are extensively studied in the present research. The project makes use of machine learning algorithms in the field of predictive modeling, specifically Artificial Neural Networks (ANN) and Support Vector Machines (SVM). These tools play a crucial role in modeling PV power output and formic acid production while accounting for a variety of influencing factors. A comparative study shows that SVM outperforms ANN in accurately predicting the production of FA and PV power generation, both of which are the major goals. This model is a predictive tool that can be used to forecast these goals based on certain causal variables. Overall, it is observed that the maximum power produced with 2-axis solar tracker was achieved in February as 2355 kW resulting in the highest formic acid production of 2.25×10^6 grams. The study's broad ramifications demonstrate solar liquid fuel technology's potential as a long-term fix in the field of renewable energy. In addition to advancing the field of renewable energy storage, the study represents a major step toward tackling the global challenge of climate change.

1. Introduction

The year 2020 witnessed a strong impact of climate change with many unusual natural disasters occurring worldwide, such as grasshopper accidents and terrible forest fires in Australia (Deb et al., 2020; Henry Fountain). In recent years, the area of ice in the two poles has continuously decreased, and the average global temperature has constantly increased due to climate change (Zhang et al., 2015; Post et al., 2019). World Meteorological Organization (WMO) reports that the average global temperatures reached a high record between 2010 and 2019, and the last five years had the highest temperatures in the past 140 years by 58% and 78%, respectively (ANON, 2023). Renewable

* Corresponding author.
E-mail address: muhammad.w.shahzad@northumbria.ac.uk (M.W. Shahzad).

https://doi.org/10.1016/j.psep.2024.02.060
Received 25 November 2023; Received in revised form 4 February 2024; Accepted 20 February 2024
Available online 23 February 2024

energy is considered one of the most effective ways to adapt to climate change (Fräss-Ehrfeld, 2009; Olabi and Abdelkareem, 2022). Solar and wind power are outperforming other renewable energy sources (ANON, 2024a). Experts estimate that the cost of constructing wind and solar power systems has decreased significantly compared to 2009 (IRENA, 2019). They have substantial cost improvement in investment, operation, and maintenance (O&M) over the past ten years with no sign of slowing down. Equipment costs of renewable energy projects tend to decline rapidly (IRNEA, 2022). Therefore, these renewable energy sources are increasingly competing with fossil fuels. As evident from the new report shared by the International Renewable Energy Agency (IRENA), 2020, With 62% of the year's total power capacity increase, renewable energy achieved substantial progress that signaled a trend toward sustainable energy alternatives (Irena, 2021).

Solar energy is a preferred choice for replacing fossil fuels and contributing clean energy due to its cheapness and high availability among accessible renewable resources (Mehrpooya et al., 2020). However, solar energy is intermittent and depends on variable parameters. For that reason, integrating solar energy into the national grid is a big challenge for the power supply system. Integrating the energy storage systems into the solar power plant is the solution to that problem and reduces solar energy instability (Chong et al., 2016). The storage system of solar energy plays an important role in buffering the periodic power supply of day-night and seasonal cycles of buffering the intersperse power grid usage (Kubicek et al., 2017). Energy storage can be categorized into different forms, such as chemical, thermal, kinetic, pumped, magnetic, electrical, thermochemical, and electrochemical (Koohi--Fayegh and Rosen, 2020). One of the potential considerable solar storage solutions is converting excess solar energy (when solar power produced is greater than the demand) into hydrogen, with the challenge of storing and transporting it (ANON, 2019a; Dutta, 2021). Hydrogen, the most abundant element, is considered a key player in the quest for sustainable energy. With an atomic mass of 1 g/mol, it forms water when combined with oxygen and is involved in creating various organic compounds. Hydrogen is gaining attention as a clean and renewable energy source, primarily because its combustion only produces water, eliminating toxic waste and CO_2 emissions (Singh et al., 2020; Tarhan and Çil, 2021). Its production methods include thermochemical gasification of natural gas, heavy hydrocarbon gasification, and biomass gasification and pyrolysis, each with unique processes and outputs (Megia et al., 2021). Additionally, water electrolysis, splitting water into hydrogen and oxygen using electricity has a certain potential for scaling up but is not truly commercial yet (Zheng et al., 2023). These technologies, while promising, still face challenges in scaling up for widespread industrial use. Since hydrogen can't be stored at ambient temperatures and pressure (Barthelemy et al., 2017). In recent years, the technologies and equipment for hydrogen storage are still limited in capacity and can only be met on a small scale. Therefore, the produced activities of hydrogen have not been expected. Secondly, although the raw material source for hydrogen production is endless, producing hydrogen from the electrolysis process is expected to reduce (Badgett et al., 2021).

Several hydrogen storage techniques have been developed to overcome these challenges, such as chemical Hydrides, Methanol(CH_4), Ammonia(NH_3), and Formic Acid (HCOOH) (Müller et al., 2017). Table 1 summarizes the critical variables that show the benefits of Formic Acid (FA) in comparison to other viable solutions such as liquid H_2, Methylcyclohexane (MCH), and (NH_3). FA is a promising candidate for Low Organic Hydrogen Carriers (LOHC) due to its unique properties, including low flammability and low toxicity, and its accessibility for large-scale production (Chatterjee et al., 2021; Dutta et al., 2022). Since most systems associated with FA dehydrogenation (decarboxylation) have a low-reaction enthalpy, H_2 can be generated at a mild temperature (below 100 °C), well within the US DOE's (Department of Energy) target (Trimm, 2005). Among these techniques, formic acid (HCOOH) has been considered an ideal hydrogen storage method, due to its high volumetric hydrogen density, at 53 g of H_2 per liter (Mellmann et al., 2016). It can be easily stored under ambient temperature with a high energy density liquid and its low toxicity characteristic. The formic acid market data demonstrate the high demand for this material with steady growth (5% growth from 2014 to 2019) and increasing price (ANON, 2019b). FA is the elementary carboxylic acid that can be effortless to bio-degradable, and therefore it is environmentally friendly (Mardini and Bicer, 2021a). The formic acid in fuel cell applications can produce 7000 kWh yearly. FA has a variety of applications in, pharmaceuticals, Direct Formic acid fuel cells (DFAC), and food chemicals (Eppinger and Huang, 2017) as shown in Fig. 1. FA can also be used in generating electricity purposes by transforming chemical energy into direct current electricity. In 1996, formic acid was also explored as a fuel (Qingfeng et al., 2001).

Since the formic acid is produced from CO_2 and H_2 using the electricity generated from the PV system and can be stored in ambient conditions. Therefore, it can be consumed as liquid fuel and applied in fuel cell applications to generate electricity. Solar power capacity is the maximum electrical output that a solar energy system can produce. The solar power capacity of Vietnam is very low compared to the leading countries like China, the United States, and India as shown in Fig. 2. Vietnam is witnessing a time of significant expansion in solar energy capacity, which signals a strong trend toward renewable energy sources inside the country (ANON, 2024a). Consequently, it is imperative to carry out additional studies in this field since it has the potential to significantly improve Vietnam's solar energy capacity and bring it into line with the developments of other top nations in solar energy utilization. This research and development could be a crucial step towards improving Vietnam's standing in the world of renewable energy. Numerous research on the viability evaluation and electrolyze optimization of photovoltaic farms and rooftop solar power plants to produce hydrogen in Vietnam have been conducted (Phan-Van et al., 2023; Phap et al., 2022). Depending on the solar energy potential of each province, the hybrid power system in the industrial units in Quang Nam Province, Binh Thuan Province, and Can Tho Province can produce 17, 386 kg/year to 17,422 kg/year of hydrogen to supply fuel cells (Phap et al., 2022). Vu Minh Phap and the group also compared various provinces in Vietnam namely Hai Duong, Quang Nam, Binh Thuan, and

Table 1
Comparison of various energy vectors (Dutta et al., 2022).

Properties	Hydrogen	Methylcyclohexane	Ammonia	Formic Acid
Phase	liquid	Liquid	Liquid	Liquid
Density (kg/m^3)	70.8	770	610	1220
Boiling point (°C)	–253	101	–33	100
Volumetric H_2 content (kgH_2/m^3)	70.8	47.1	107.7/120	53
Volumetric energy density (MJ/L)	8.49	5.66	12.92/14.4	6.36
Gravimetric H_2 content (wt%)	100	6.1	17.65	4.38
The gravimetric energy density (MJ/kg)	120	7.35	21.18	5.22
Dehydrogenation Enthalpy ($kJ/molH_2$)	0.907	69.8	30.6	31.2
Enthalpy of Vaporization ($kJ/molH_2$)	–	10.8	15.1	NA
H_2 release	Evaporation	Dehydrogenation (350 °C)	Cracking (>425 °C)	Dehydrogenation (<100°C)
An explosive limit in the air (vol%)	4–75	1.2–6.7	15–28	
Combustibility / Harmfulness	Extremely combustible	Harmful	Harmful	Less Combustibility / less harmfulness

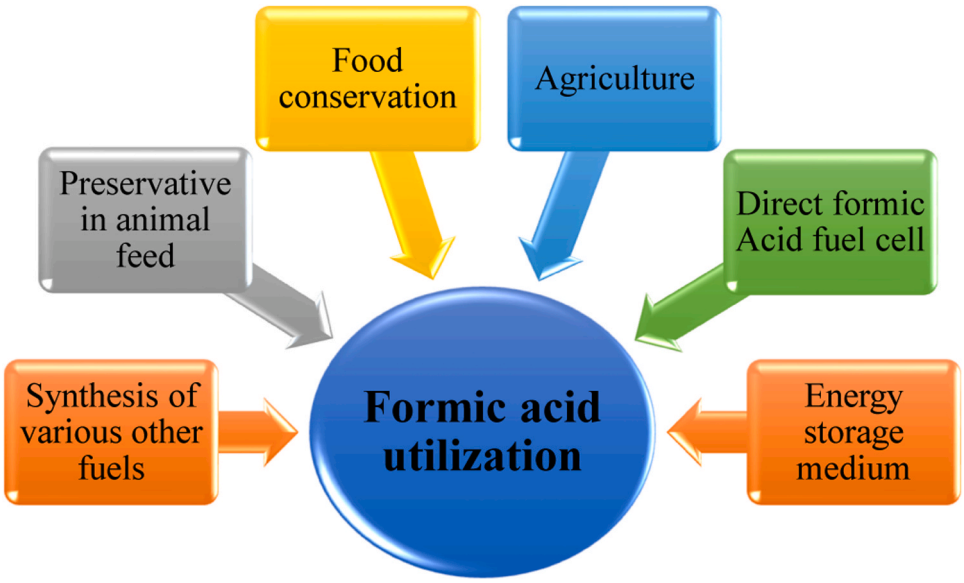

Fig. 1. Formic acid applications.

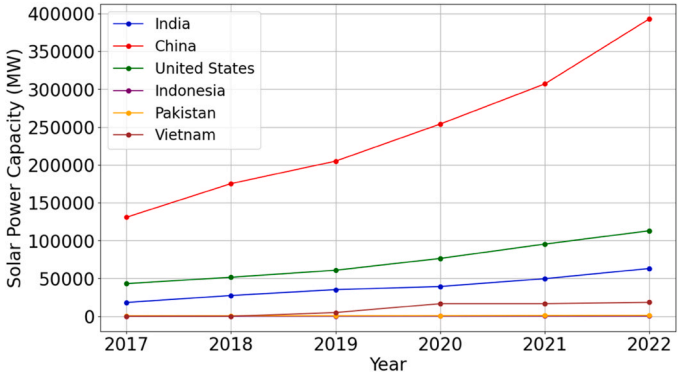

Fig. 2. Solar power capacity comparison of Vietnam with other countries [33].

Can Duo among which Binh Thuan has excellent solar energy potential which indicates it can produce the highest amount of electricity and solar power.

The above review suggests that there is a dire need for the development and implementation of solar-based energy production and storage systems. Among different available energy carriers, formic acid has the highest energy density which shows that it can store energy efficiently and improve the performance of energy systems (refer to Table 1). Furthermore, other energy carriers have challenges with transport and storage issues like ammonia needing to be stored at cryogenic temperature, and hydrogen needs 700 bar pressure for storage. Formic acid can be stored at ambient temperature and pressure hence it is a potential energy carrier. Therefore, it is considered an energy storage medium integrated with 2-axis solar tracking. The 2-axis solar tracking PV offers several benefits like up to 45% higher power production compared to fixed tilt systems by tracking sun orientation throughout the day. The current study provides a detailed investigation of Formic Acid production potential using a 2-axis solar tracking system in Binh Thuan, Vietnam using Machine Learning. This is because the conventional modeling techniques require extensive technical equation development and simulation parameters which are not well established for formic acid production in the considered area. On the other hand, with the development and advancement of data-driven techniques Artificial intelligence-based Machine Learning Models can accurately predict the system performance at assorted operating conditions. Therefore, the current study employs two ML models Artificial Neural Networks (ANN) and Support Vector Machine (SVM) to predict the Formic Acid production trends using a 2-axis solar tracker system. The study is aimed to contribute to advancing the field of renewable energy storage and will serve as a major step toward tackling the global challenge of climate change.

2. Materials and methods

2.1. Vietnam's significance and machine learning using dual axis solar tracker

In Vietnam, there have been studies on various types of modeling by machine learning, deep learning, and Artificial Intelligence, to predict the electricity generation capacity, short-term solar \radiation forecast, and forecasting output of industrial Solar Power Plants respectively (Pham and Tran, 2023; Huynh, 2020; Quang et al., 2021). Hence, based upon the literature review, it is understood that there is still a requirement to present a case study of solar to fuel production as no research has been conducted for Binh Thuan, Vietnam, and there is no research on the usage of machine learning (ML) tools on its predictive performance, as ML can simplify the multidimensional chemical and thermal models associated with it. Following the gaps identified in the literature, this research systematically addresses the solar energy potential (corresponding PV production) for the region of Binh Thuan, Vietnam, which is eventually used to evaluate the formic acid production potential. Owing to their ability to generalize, artificial neural networks (ANN) are the most recommended methods above other machine learning algorithms. Since there is a growing demand for academic programs linked to artificial intelligence in science, math, and engineering, there will be a significant increase in demand for machine learning techniques in the energy sectors in the upcoming years. For ML algorithms to be successfully implemented and shared by significant players in the energy sector, data generation, management, and safety are anticipated to be critical, which will encourage the development of ambitious energy management projects (Rangel-Martinez et al., 2021).

To produce power for the liquid fuel conversion process, the PV production system is taken into consideration using a two-axis solar tracking system. To keep track of the sun's orientation, utilize a solar tracker. When utilizing two-axis trackers, the screen is oriented to monitor in the direction of the highest amount of sunlight throughout the day by modifying the angles of the tracker (both elevation and

azimuth angles) (Muthukumar et al., 2023). By dynamically adapting to the sun's position, a dual-axis solar tracking system maximizes solar panel efficiency by optimizing sunlight exposure throughout the day and the seasons. This method is especially useful for optimizing solar energy use because it tracks changes in the sun's elevation and azimuth. It continuously works better than fixed-tilt systems, with the potential to increase energy production by 30–45 percent, enhancing return on investment. Its capacity to deliver a consistent power output increases power supply reliability and lessens reliance on the grid, making it a feasible option for locations with grid power constraints. Furthermore, because of its adaptability to a variety of terrains, it can be deployed in a variety of situations, such as roofs and uneven landscapes, making it an excellent choice for a range of geographic areas (ANON, 2024b). MATLAB software is used to implement simulation. There are three steps of simulation.

1. The first step will be solar tracking system simulation and relative estimation of PV production.
2. The second step involves the potential generation of formic acid through the evaluated PV production.
3. In the last step, machine learning is applied to estimate the PV production and the relative solar to fuel production.

The location chosen to implement simulation in this research is Binh Thuan Province, Vietnam. Binh Thuan province is in the time zone UTC+07, at the latitude of 11° 06' 05" and the longitude of 107° 56' 30" (see Fig. 3). The data of the solar irradiance on that position is collected on the Global Solar Atlas platform by the World Bank Group which is displayed in Table 2. The data of the monthly averages of direct normal irradiance (DNI) is collected and shown in Table 3 which shows yearly 1355kWh/m^2 DNI is collected and Table 4 shows the daily averages of direct normal irradiance.

The visualizations of solar irradiance daily and monthly are implemented in MATLAB simulations. It is necessary to examine which time of the day or month has a high density of solar irradiation. Then, the data is used to simulate the motion of the sun for the two-axis solar tracking system. Two main parameters of the two-axis solar tracking system are the Elevation angle and the Azimuth angle. These two parameters are computed to respond to three specific days in a year. The 1st of January, 21st of March, and 21st of June are chosen for simulation. The computations of the Elevation angle and Azimuth angle are based on these three typical days and the solar irradiance data in Binh Thuan. The power produced by the PV system will supply to the compression and reactor system to form formic acid. In this research, the total area of photovoltaic panels is assumed as 100 m^2. The PV array's efficiency is considered 15%, corresponding to the two-axis solar tracking system.

Table 2
Data of the solar irradiance positioned at the latitude. of 11° 06' 05" and the longitude of 107° 56' 30".

Description	Value
Specific photovoltaic power output (kWh/kWp)	1461
Direct normal irradiation (kWh/m^2)	1333
Global horizontal irradiation (kWh/m^2)	1812
Diffuse horizontal irradiation (kWh/m^2)	850
Global tilted irradiation at the optimum angle(kWh/m^2)	1861
Air temperature (°C)	22.5
Optimum tilt of PV modules(°)	13
Terrain elevation (m)	791

Table 3
The monthly averages direct normal irradiance (DNI).

Month	DNI (kWh/m^3)	Month	DNI (kWh/m^3)
January	128	July	80
February	157	August	88
March	154	September	82
April	130	October	99
May	117	November	112
Jun	94	December	113
Yearly	**1355**		

2.2. The two-axis solar tracking system

The elevation angle and the azimuth angle are two critical parameters used to track the sun's position. The azimuth angle and the elevation angle should be computed throughout the day to simulate the position of the sun, which varies continuously. The PV array has the potential to increase power generated in the range of 30–60% (Ibrahim et al., 2018). Considering that equipping the two-axis solar tracker could significantly impact the PV system efficiency. The two-axis tracker allows the PV to follow the sun in any orientation. The PV system with a two-axis tracker could absorb maximum solar irradiance to produce electricity. The

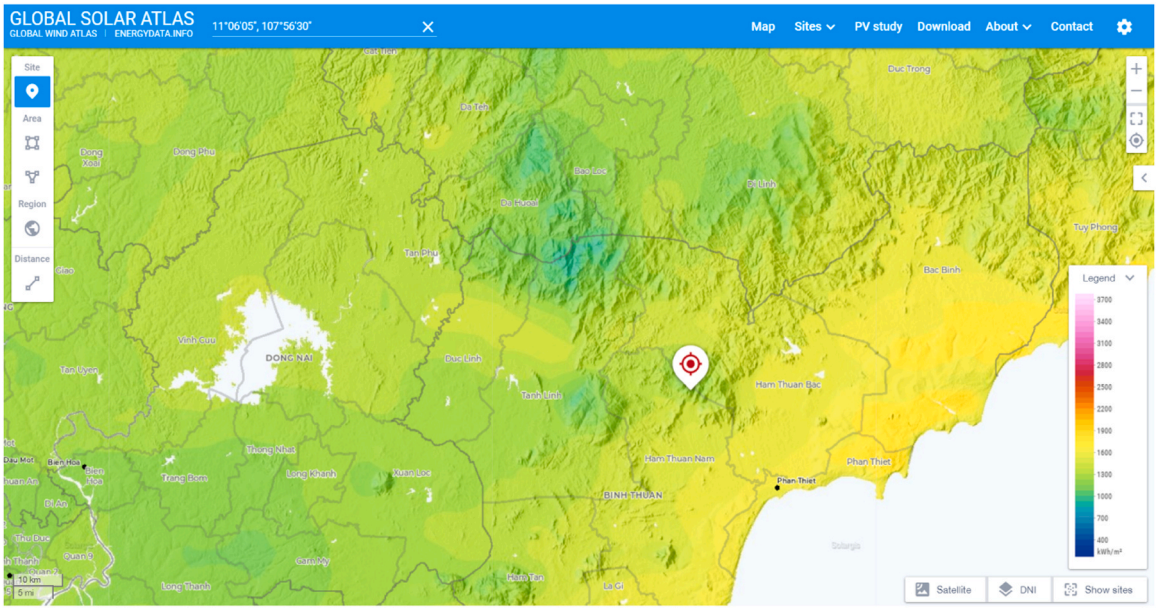

Fig. 3. Direct normal irradiance map of Binh Thuan province.

Table 4
The daily averages direct normal irradiance.

Daily solar irradiance (Wh/m^2)												
Hour	**Jan**	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**
0 - 1	0	0	0	0	0	0	0	0	0	0	0	0
1 - 2	0	0	0	0	0	0	0	0	0	0	0	0
2 - 3	0	0	0	0	0	0	0	0	0	0	0	0
3 - 4	0	0	0	0	0	0	0	0	0	0	0	0
4 - 5	0	0	0	0	0	0	0	0	0	0	0	0
5 - 6	0	0	0	6	24	17	7	0	7	10	0	0
6 - 7	54	107	155	195	231	156	116	116	138	202	197	117
7 - 8	315	409	388	376	341	249	188	197	239	321	367	313
8 - 9	433	544	511	478	424	325	252	273	311	383	433	410
9 - 10	515	630	572	499	453	356	283	301	323	404	472	464
10 - 11	542	672	591	492	451	365	308	336	324	390	465	469
11 - 12	522	653	538	418	396	365	313	334	324	362	421	441
12 - 13	478	615	497	402	390	365	318	350	327	341	390	414
13 - 14	412	562	483	428	360	324	285	319	274	278	331	350
14 - 15	343	521	465	439	317	276	242	283	214	228	280	291
15 - 16	289	440	397	351	240	203	158	191	140	172	207	229
16 - 17	204	342	282	236	161	144	102	125	107	121	147	147
17 - 18	29	114	80	0	0	0	0	23	0	0	8	11
18 - 19	0	0	0	0	0	0	0	0	0	0	0	0
19 - 20	0	0	0	0	0	0	0	0	0	0	0	0
20 - 21	0	0	0	0	0	0	0	0	0	0	0	0
21 - 22	0	0	0	0	0	0	0	0	0	0	0	0
22 - 23	0	0	0	0	0	0	0	0	0	0	0	0
23 - 24	0	0	0	0	0	0	0	0	0	0	0	0

two-axis solar tracking system has great potential to enhance the PV's power output, with a drawback of significant investment cost for this system. As for two-axis solar tracking, in terms of engineering, the two-axis solar tracker is an advanced technology. Yet in terms of financing, the investor is expected to make the decision depending on the economic benefit each solar tracking system brings to their PV plant.

2.3. Estimation of photovoltaic power production

The power generated from PV is calculated by the following equation:

$$P_{PV} = G_T * A_{PV} * \eta_{PV} \tag{1}$$

Where, P_{PV} is the PV power [watts, W], G_T is the solar irradiance in a tilted module [W/m^2], A_{PV} is the PV array area [m^2], and η_{PV} is the PV converter efficiency.

2.4. Solar to formic acid conversion

The reaction between CO_2 and H_2 occurs in the conversion reactor to generate the formic acid. In the last stage, the formic acid goes into the distillation column to reduce the pressure and temperature. The reaction between CO_2 and H_2 occurs in the conversion reactor to generate the formic acid. In the last stage, the formic acid goes into the distillation column to reduce the pressure and temperature. The reaction is demonstrated in the equation below.

$$H_2 + CO_2 \rightarrow HCOOH \tag{2}$$

The efficiency of the reaction to form formic acid is around 19%, based on previous research (Müller et al., 2017). The equations shown below are used to calculate the mass of formic acid produced.

$$moles_{FA} = 3600 \times \frac{P_{PV}}{\Delta G} \times eff_{reaction} \tag{3}$$

$$mass_{FA} = moles_{FA} \times molarmass_{FA} \tag{4}$$

Where, $moles_{FA}$ are the moles of formic acid, P_{PV} is the power generated by the PV system [kW], $mass_{FA}$ is the mass of formic acid[g], ΔG = 32.9 kJ/mol is the energy needed to form formic acid, $eff_{reaction}$ = 19% is the efficiency of the reaction to form formic acid, and $molarmass_{FA}$ = 46 g/mol is the molecular mass of formic acid.

2.5. Machine learning based process modelling of solar to fuel system

Machine learning-based modelling algorithms have demonstrated their usefulness in approximating the physical behavior of the systems with good accuracy (Zhang et al., 2023; Ashraf and Dua, 2023). The algorithms can build the functional mapping between the input-output variables of the system by harnessing the data associated with the variables (Krzywanski et al., 2024). Machine learning-based algorithms can mine the data to extract the functional relationships that are difficult to develop through conventional analytical techniques (Enke and Thawornwong, 2005). Moreover, machine learning algorithms offer computational efficacy to expedite the process analysis which is sometimes computationally prohibitive in mathematical model-based analytics (Afzal et al., 2021).

Machine learning algorithms are deployed for modeling the PV power production and formic acid production under the influence of the operating parameters as discussed in the previous sections. Artificial neural networks (ANN) and support vector machines (SVM) are two well-established and reliable data-driven machine learning algorithms that can effectively simulate the complex and high-dimensional output

space of the objective function (Ashraf et al., 2021; Li et al., 2020). Furthermore, the modeling comparison of the two tools is presented to select the well-performing algorithm suited for modeling the PV power and formic acid production. The details associated with the working of the ANN and SVM algorithms, hyperparameters tuning, and model development are provided in the following.

2.5.1. Artificial neural network algorithm

Artificial neural networks (ANN) and support vector machine (SVM) are the two versatile classes of algorithms that are used quite often for the modelling tasks. The working mechanism of building the functional map is significantly different by design, however, the two algorithms have demonstrated particular utility in various domains of the applications including chemical process industries (Zayed et al., 2023), energy systems, signal processing (ANON, 2002), and smart manufacturing (Ishfaq et al., 2023).

ANN is one of the powerful modelling algorithms of ML and mimics the information processing taking place in the human brain to approximate the function space against the causal input variables. Three processing layers called input layer, hidden layer, and output layer are embedded in the architecture of the ANN for the information processing. The observations associated with the set of input variables, i.e., X_i where $i = 1,2,3,\ldots,N$ are the total observations are passed on from the input to the hidden layer of ANN. Hidden layer is the main processing unit of ANN and extensive computations take place there to generate a signal that is transmitted to the output layer of the ANN. Further information processing takes place at the output layer of ANN to simulate a response. The mathematical representation of the working of the ANN algorithm is described as follows:

$$\hat{y}_i = f_2\left(\sum_{i=1}^{N} W_2\left[f_1\left(\sum_{i=1}^{N} X_i W_1 + b_1\right)\right] + b_2\right) \tag{5}$$

here, W_1, W_2 are the matrices containing the weigt connections from the input to the hidden layer and hidden to the output layer of ANN respectively. b_1, b_2 and f_1, f_2 are the bias matrices and activation function applied on the hidden and output layer of ANN respectively. Whereas, $\hat{y}_i$ is the simulated response produced as the result of information processing at the processing layers of ANN.

2.5.2. Support vector machine algorithm

Support vector machine is another excellent modelling algorithm of ML and uses the structural risk minimization principle to build the data-driven model for the application. The algorithm projects the high-dimensional and hard-to-classify data onto the higher dimensions by the kernel and subsequently, solves the problem linearly. The hyperplanes are introduced to separate the data points. SVM incorporated Vapnik's ε-intensive loss function and approximates the problem as an inequality-constrained optimization problem to maximize the boundary around the hyperplane to tolerate the modelling error. The data points lying on the margin lines are treated as support vectors and play a pivotal role in the predictive mechanism of the SVM model. Karush-Kuhn-Tucker conditions are incorporated to ensure the global optimality of the nonlinear optimization problem for building an effective SVM-based predictive model. The standard form of support vector regression with $C > 0$ (box constraint) and $\varepsilon > 0$(epsilon) is written as:

$$(w, \beta, \beta^{*\min} \frac{1}{2} w^T w + C \sum_{i}^{N} \beta_i + C \quad \sum_{i}^{N} \beta^*) \tag{6}$$

such that

$$w^T \varphi(x_i) + b - z_i \quad \leq \varepsilon + \beta_i \tag{7}$$

$$z_i \quad - w^T \varphi(x_i) - b \leq \varepsilon + \beta_i^* \tag{8}$$

$$\beta_i, \beta_i^* \geq 0, i = 1, 2, 3, \ldots, n$$

The dual problem is

$$\alpha, \alpha^{*\min} \frac{1}{2}(\alpha - \alpha^*)^T Q(\alpha - \alpha^*) + \varepsilon \quad \sum_{i}^{N} (\alpha + \alpha^*) + \sum_{i}^{N} z_i(\alpha_i - \alpha_i^*) \tag{9}$$

subject to

$$e^T(\alpha_i - \alpha_i^*) = 0 \tag{10}$$

$$0 \leq \alpha_i, \alpha_i^* \leq C, i = 1, 2, 3, \ldots, N \tag{11}$$

$$Q_{i,j} = K(x_i, x_j) \equiv \varphi(x_i)^T \varphi(x_j) \tag{12}$$

After solving the dual problem, the approximate function is:

$$\hat{y}(x) = \sum_{i=1}^{N} (\alpha_i + \alpha_i^*) K(x_i, x) + b \tag{13}$$

here, w is the weight vector; β, β^* are the slack variables; $\varphi(.)$ is the mapping function; Q is the regression function; α_i, α_i^* are the function parameters; $K(,.)$ is the kernel function; and b is the bias introduced in the model and $i = 1, 2, 3, \ldots, N$ equal to a total number of observations. Whereas $\hat{y}(x)$ is the output variable mapped on the list of input variables.

2.5.3. Performance evaluation metrics of ML algorithms

Three performance indicators namely coefficient of determination (R^2), mean absolute error (MAE), and root mean square error (RMSE) are introduced to gauge the modeling performance of ANN and SVM models (Krzywanski and Nowak, 2016; Tariq et al., 2023). The value of R^2 varies from zero (no correlation) to one (perfect correlation) which indicates the strength of association between the actual and model simulated responses. Whereas MAE and RMSE are the two error terms that illustrate the deviation that may exist in the model responses compared with the actual values. Therefore, the smaller the error terms are, the more accurate predictions are made by the model. The mathematical expressions of the performance indicators are given:

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{N} (y_i - \overline{y_i})^2} \tag{14}$$

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |y_i - \hat{y}_i|, and \tag{15}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2} \tag{16}$$

Where y_i and $\hat{y}_i$ are the actual and model simulated values respectively and $i = 1, 2, 3, \ldots, N$ equal to the total number of observations. Similarly, $\overline{y_i}$ is the mean of the actual values.

3. Results and discussion

3.1. Solar characteristics of the subject location

The average daily solar irradiance is visualized in Fig. 4(a) and Fig. 4 (b) illustrates the monthly solar irradiance which provides information on the months that solar irradiance has a high density in a year. The results depicted in Fig. 4a, provide insights into the temporal variation of solar irradiance throughout the day. Notably, sunrise at 6 and sunset at 19 delineate the period during which the daily solar irradiance peaks, particularly between 9 and 14, with its zenith at 11 with 450 Wh/m^2

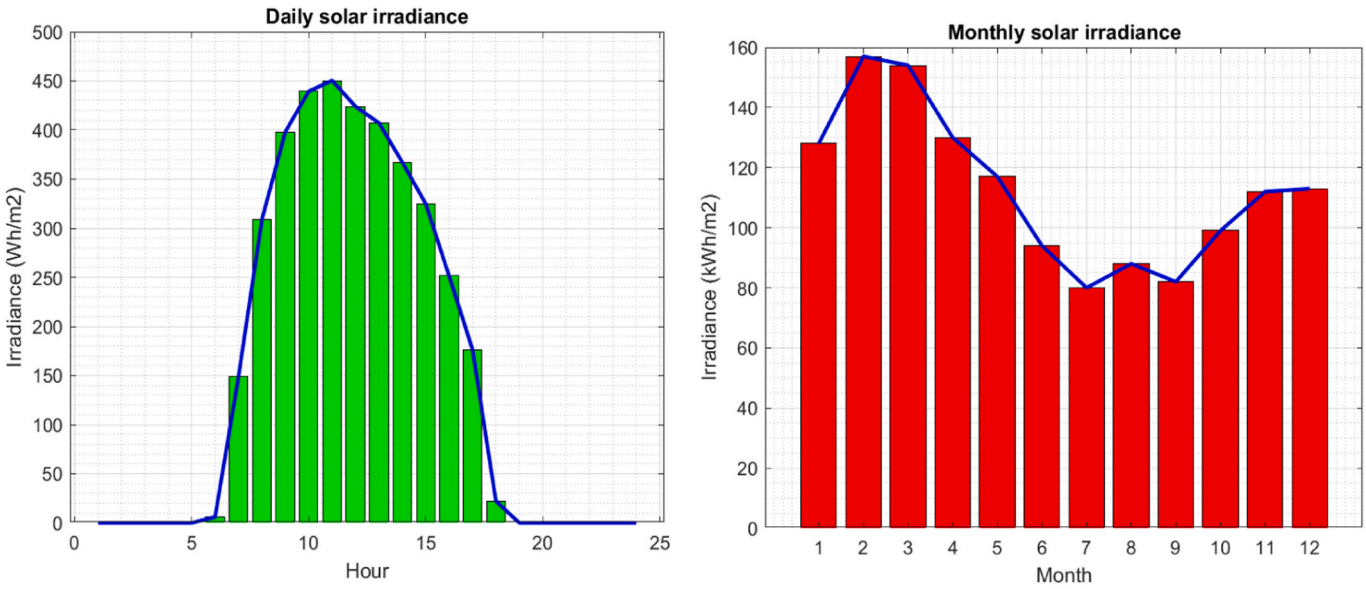

Fig. 4. Daily (a) and monthly(b) solar irradiance.

solar irradiance. This temporal distribution signifies that the primary power output from the PV system occurs between 9 and 14, with lower irradiance levels observed during the morning and night. Examining Fig. 4b, which showcases monthly solar irradiance patterns, reveals that February and March exhibit the highest irradiance levels. Conversely, a significant decline in total solar irradiance occurs from April to July. Specifically, the highest recorded irradiance is 157 kWh/m^2 in February, contrasting with the lowest observed in July at 80 kWh/m^2.

Table 4 represents daily solar irradiance (Wh/m^2), with the colors that change from yellow to deep orange and back again corresponding to the sun's position in the sky, with yellow signifying lower light levels at sunrise and sunset and deep orange representing the sun's highest point. The following conclusions can be drawn from Table 4:

- No solar irradiance was captured from midnight to just before dawn (0–1 AM to 4–5 AM) in the year, as indicated by consistent zeroes across the time slots in respective months.
- There is a slight commencement of solar irradiance observed at 5–6 AM, marked in yellow, indicating the first light of dawn, with irradiance present in some months, particularly April, May, June, July, September, and October.
- There is a spike in irradiance levels between 6 and 7 AM, which is shown by a change in colour from yellow to orange. This indicates the start of solar energy availability and the dawn.
- The sun's ascent towards its zenith is represented by the maximum solar irradiance levels of the day, which occur between 7 and 8 AM and 10–11 AM, with colors ranging from orange to deep orange.
- The irradiance gradually decreases in the post-noon hours, and colors eventually return to deep orange by 3–4 PM, then orange, and finally yellow, as the sun sets.
- The mild orange and yellow hues of the late afternoon (4–5 PM to 6–7 PM) show a decrease in irradiance as the sun approaches the horizon.

Figs. 5a, 5b and Table 5 represent two graphs and a table that relate the elevation and azimuth angles of the sun to specific dates, alongside the equation of time (EoT) and the time correction factor for Binh Thuan. Fig. 5a displays the sun's elevation angle above the horizon on

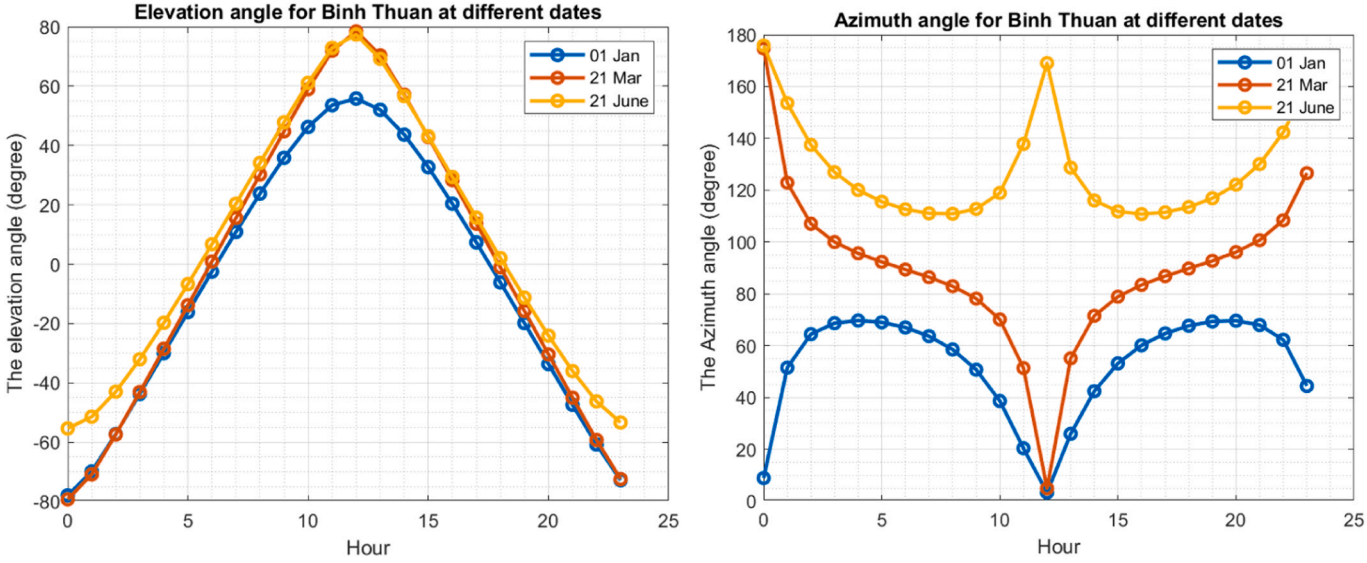

Fig. 5. (a)Elevation angle and (b)Azimuth angle.

Table 5
The equation of time and the time correction factor.

Description	1st January	21st Mar	21st Jun
The equation of time	-3.7052	-7.8428	-1.4474
The time correction factor	8.0612	3.9236	10.3189

three separate dates: January 1st, March 21st, and June 21st, throughout the day. Every day at solar noon, the elevation reaches its peak; the variation in elevation reflects the seasonal shift in the sun's path across the sky. The sun reaches its highest point in the sky on June 21, which is usually the summer solstice. This results in the largest elevation angles. On the same three dates, Fig. 5b displays the azimuth angle graph, which depicts the sun's position as a function of compass direction over a day. Around 90 degrees azimuth is where the sun rises in the east, 270 degrees azimuth is where it sets in the west, and 180 degrees azimuth is when it reaches its lowest point due south at solar noon. As we progress from January 1st to June 21st, the curve's shape reflects the changing of the seasons, with earlier sunrises and later sunsets.

The values of the time correction factor and the equation of time (Eot) for the three previously mentioned dates are listed in Table 5. The elliptical shape of the Earth's orbit and its axial tilt are the reasons behind the EoT, which explains the discrepancy between solar time and clock time. The time correction factor accounts for both the observer's specific longitude concerning the time zone meridian and the EoT. The EoT is negative on January 1st, suggesting that solar noon happens earlier than clock noon. The EoT becomes less detrimental as the year goes on until June 21st. The fact that the time correction factor is positive indicates that forwarding the clock is necessary to bring the clock in line with solar time. To accurately calculate the solar noon and, consequently, the elevation and azimuth angles, it is essential to interpret the relationship between the EoT and the time correction factor. The sun's trajectory varies on different dates as a result of the tilt and orbit of the Earth, and this is reflected in the angles. Solar panels are properly oriented to collect the most sunlight when the EoT and time correction correct these angles for the difference between solar and clock time.

Table 6, for Bin Thaun, shows the hour angle, azimuth angle, elevation angle, and local solar time in 24 hours, for three separate dates: January 1, March 21, and June 21. The local solar time represents time as an angle, with 15 degrees representing one hour. The values increase incrementally and consistently throughout the hours. The hour angle is computed from the local solar time. The sun is west of the meridian (before solar noon) when the values are negative, and east of the meridian (after solar noon) when the values are positive. The hour angle rises (negative to positive) as the day goes from 0 to 24 hours. The sun's height in the sky is depicted from the elevation angle, which can be calculated by the latitude, the hour angle, and the declination angle. Early in the day, negative elevation angles mean that sunrise has not yet occurred, and the sun is behind the horizon. The elevation angle rises and eventually reaches positive values as the time approaches solar noon, signifying that the sun emerged and is rising in the sky. The azimuth angle in Table 4 indicates the direction of the sun according to the compass. 0 degrees denotes north, 90 degrees east, 180 degrees south, and 270 degrees west on the azimuth angle. The azimuth angles begin in the northeast on January 1st and March 21st, pass through the south (positive values), and then shift towards the northwest as the sun sets.

3.2. Power production from the PV system and estimation of the production of formic acid

The daily power output in kilowatts (kW) from a photovoltaic (PV) system during day is shown in Table 7, broken down by month. From Table 7 it is evident that since PV systems depend on solar energy, it is to be expected that the power generation values are primarily zeros during the night (from "0–1" to "5–6" and from "18–19" to "23–24"). Across all months, the peak power generation happens in the middle of the day, approximately between "8–9" and "16–17" hours, with a peak around "10–11" and "11–12" hours. Power generation varies with the seasons:

- Higher electricity generation during peak hours is observed in May through August, which corresponds with longer summer daylight hours.
- Lower power generation is observed in January and December, which corresponds with the winter's shortened daylight hours.

With the use of data displayed in Table 7, one can plan for energy storage or grid distribution, analyze a PV system's annual performance, and spot trends in power generation. The power produced by the PV

Table 6
Local solar time, hour angle, elevation angle, and azimuth angle.

Hour	The local solar time (degree)			The hour angle (degree)			The elevation angle (degree)			The Azimuth angle (degree)		
	1st Jan	21st Mar	21st Jun	1st Jan	21st Mar	21st Jun	1st Jan	21st Mar	21st Jun	1st Jan	21st Mar	21st Jun
0–1	0.134	0.065	0.172	-177.985	-179.019	-177.420	-77.936	-79.258	-55.357	8.909	174.730	175.834
1–2	1.134	1.065	1.172	-162.985	-164.019	-162.420	-69.874	-70.856	-51.405	51.514	122.912	153.629
2–3	2.134	2.065	2.172	-147.985	-149.019	-147.420	-57.257	-57.419	-42.997	64.444	107.080	137.513
3–4	3.134	3.065	3.172	-132.985	-134.019	-132.420	-43.715	-43.097	-32.049	68.681	100.007	126.963
4–5	4.134	4.065	4.172	-117.985	-119.019	-117.420	-29.939	-28.514	-19.752	69.714	95.649	120.089
5–6	5.134	5.065	5.172	-102.985	-104.019	-102.420	-16.153	-13.831	-6.719	69.028	92.345	115.560
6–7	6.134	6.065	6.172	-87.985	-89.019	-87.420	-2.494	0.885	6.727	67.033	89.415	112.656
7–8	7.134	7.065	7.172	-72.985	-74.019	-72.420	10.896	15.593	20.397	63.675	86.433	111.082
8–9	8.134	8.065	8.172	-57.985	-59.019	-57.420	23.801	30.248	34.151	58.537	82.950	110.912
9–10	9.134	9.065	9.172	-42.985	-44.019	-42.420	35.835	44.771	47.833	50.722	78.184	112.797
10–11	10.134	10.065	10.172	-27.985	-29.019	-27.420	46.247	58.950	61.131	38.648	70.133	118.950
11–12	11.134	11.065	11.172	-12.985	-14.019	-12.420	53.595	71.930	72.901	20.393	51.350	137.849
12–13	12.134	12.065	12.172	2.015	0.981	2.580	55.830	78.454	77.410	3.304	4.906	169.080
13–14	13.134	13.065	13.172	17.015	15.981	17.580	52.036	70.388	69.199	25.966	55.106	128.714
14–15	14.134	14.065	14.172	32.015	30.981	32.580	43.675	57.132	56.637	42.427	71.524	116.067
15–16	15.134	15.065	15.172	47.015	45.981	47.580	32.720	42.884	43.149	53.160	78.928	111.834
16–17	16.134	16.065	16.172	62.015	60.981	62.580	20.398	28.337	29.419	60.133	83.456	110.788
17–18	17.134	17.065	17.172	77.015	75.981	77.580	7.335	13.671	15.679	64.728	86.840	111.477
18–19	18.134	18.065	18.172	92.015	90.981	92.580	-6.144	-1.040	2.069	67.694	89.793	113.498
19–20	19.134	19.065	19.172	107.015	105.981	107.580	-19.851	-15.754	-11.261	69.347	92.745	116.908
20–21	20.134	20.065	20.172	122.015	120.981	122.580	-33.649	-30.429	-24.088	69.638	96.138	122.136
21–22	21.134	21.065	21.172	137.015	135.981	137.580	-47.391	-44.991	-36.012	67.968	100.720	130.088
22–23	22.134	22.065	22.172	152.015	150.981	152.580	-60.793	-59.253	-46.260	62.264	108.409	142.334
23–24	23.134	23.065	23.172	167.015	165.981	167.580	-72.819	-72.438	-53.374	44.438	126.599	160.686

Table 7
The daily PV power (kW).

Hour	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep	Oct	Nov	Dec
0–1	0	0	0	0	0	0	0	0	0	0	0	0
1–2	0	0	0	0	0	0	0	0	0	0	0	0
2–3	0	0	0	0	0	0	0	0	0	0	0	0
3–4	0	0	0	0	0	0	0	0	0	0	0	0
4–5	0	0	0	0	0	0	0	0	0	0	0	0
5–6	0	0	0	0.09	0.36	0.26	0.11	0.00	0.11	0.15	0	0
6–7	0.81	1.61	2.33	2.93	3.47	2.34	1.74	1.74	2.07	3.03	2.96	1.76
7–8	4.73	6.14	5.82	5.64	5.12	3.74	2.82	2.96	3.59	4.82	5.51	4.70
8–9	6.50	8.16	7.67	7.17	6.36	4.88	3.78	4.10	4.67	5.75	6.50	6.15
9–10	7.73	9.45	8.58	7.49	6.80	5.34	4.25	4.52	4.85	6.06	7.08	6.96
10–11	8.13	10.08	8.87	7.38	6.77	5.48	4.62	5.04	4.86	5.85	6.98	7.04
11–12	7.83	9.80	8.07	6.27	5.94	5.48	4.70	5.01	4.86	5.43	6.32	6.62
12–13	7.17	9.23	7.46	6.03	5.85	5.48	4.77	5.25	4.91	5.12	5.85	6.21
13–14	6.18	8.43	7.25	6.42	5.40	4.86	4.28	4.79	4.11	4.17	4.97	5.25
14–15	5.15	7.82	6.98	6.59	4.76	4.14	3.63	4.25	3.21	3.42	4.20	4.37
15–16	4.34	6.60	5.96	5.27	3.60	3.05	2.37	2.87	2.10	2.58	3.11	3.44
16–17	3.06	5.13	4.23	3.54	2.42	2.16	1.53	1.88	1.61	1.82	2.21	2.21
17–18	0.44	1.71	1.20	0	0	0	0	0.35	0	0	0.12	0.17
18–19	0	0	0	0	0	0	0	0	0	0	0	0
19–20	0	0	0	0	0	0	0	0	0	0	0	0
20–21	0	0	0	0	0	0	0	0	0	0	0	0
21–22	0	0	0	0	0	0	0	0	0	0	0	0
22–23	0	0	0	0	0	0	0	0	0	0	0	0
23–24	0	0	0	0	0	0	0	0	0	0	0	0

system is used to supply the compression and reaction system to form formic acid.

The monthly patterns for PV power generation in KW are depicted in Table 8. From Table 8, it is observed that February has the highest monthly power generation at 2355 kW, closely followed by 2310 kW in March and 1920 kW in January. An apparent decline from May (1950 kW) to August (1200 kW), while June had the lowest output at 1410 kW. Gradual increase in power generation is observed after August, peaking at 1230 kW in September 1485 kW in October 1680 kW in November, and 1695 kW in December. The yearly performance of the PV system is summarized in these monthly statistics, which can be helpful for long-term planning and analyzing patterns in energy output.

The outcomes to produce formic acid, evaluated on a daily in grams (g), is presented in Table 9. Each cell represents the amount of formic acid produced in grams during every hour for each month. The mass of produced formic acid can be calculated by subsequent Eqs. (3) and (4). The continuous 0.00 results from "0–1" to "5–6" and from "18–19" to "23–24" show that no formic acid synthesis is detected during the night for any month as there is no solar energy. Depending on the month, production runs from the early morning hours of "6–7" until the "17–18" hours. If the manufacturing process is solar-powered or temperature-influenced, the midday production peak, or "7–8" to "16–17" hours, may coincide with the peak hours of sunlight. It is evident from Table 9 that throughout numerous months, the hours of "7–8," "8–9," "9–10," and "10–11" were shown to have the highest values of formic acid production; this suggests a high degree of activity or efficiency during these times frames. Seasonally, higher values are observed during peak production hours in April through July, which may indicate improved production conditions or increased demand. August is usually when the values start to decline and stay that way until December, except for a tiny uptick in October and November during specific hours.

Table 8
The monthly PV power (kW).

Month	PV Power (kW)	Month	PV Power (kW)
January	1920	July	1200
February	2355	August	1320
March	2310	September	1230
April	1950	October	1485
May	1755	November	1680
Jun	1410	December	1695

Table 10 presents the monthly formic acid production in grams and it is observed that February witnessed a peak in the monthly production of formic acid, with almost 2.25 million grams produced. The production starts to fall in May and reaches its lowest point in July (1.15 million grams). Finally, the production shows progressive growth starting in August and reaching 1.6 million grams in both November and December.

4. Machine learning based modeling of PV power and formic acid production

The data split ratio of 0.8, 0.1, and 0.1 respectively is implemented as a training, testing, and validation dataset for the data allocation for the development of ANN. The number of hidden layer neurons is the key hyperparameter to optimize the predictive performance of the ANN. We have varied the hidden layer neurons from 4 to 10 which is the 1× to 2.5× of the input variables – a common practice used by the researchers (Tariq et al., 2023; Ashraf et al., 2023). Tangent hyperbolic and linear functions are applied on the hidden and output layer of ANN respectively as an activation function. The parameters' weight and biases are optimized by the Levenberg Marquardt algorithm which does not require computing an expensive Hessian matrix for the parametric optimization. The modeling performance comparison allows the selection of a better predictive ANN model for the considered two output variables.

Similarly, rigorous hyperparameter tuning is carried out to train a well-predictive SVM model. Box constraint, kernel scale, and epsilon are varied in the range of 0.001–1000, 0.001–1000, and 0.016814–1681.4152 respectively. Whereas, different kernel functions including linear, quadratic, cubic, Gaussian functions etc., are tried to find the best combination and set-values of the hyperparameters by Bayesian optimization in conjunction with expected improvement per second plus algorithm. The two algorithms are implemented in MATLAB 2019 version b.

The modeling performance of ANN and SVM algorithms to predict the PV power (referred to as power) and Formic Acid production (called formic acid) is presented in Fig. 4(a-d). The prediction interval is also presented around the model-based predictions to indicate the range of variation in the simulated responses. Moreover, the distribution of true and model-predicted responses are also shown along the edges of plots to show how closely the model-simulated responses are districted with

Table 9
The daily produced formic acid (gram).

Hour	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep	Oct	Nov	Dec
0–1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1–2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2–3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3–4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4–5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5–6	0.00	0.00	0.00	86.07	344.29	243.87	100.42	0.00	100.42	143.45	0.00	0.00
6–7	774.65	1534.95	2223.52	2797.33	3313.76	2237.87	1664.05	1664.05	1979.65	2897.75	2826.02	1678.40
7–8	4518.77	5867.22	5565.97	5393.83	4891.74	3571.98	2696.91	2826.02	3428.52	4604.84	5264.72	4490.08
8–9	6211.51	7803.84	7330.44	6857.05	6082.40	4662.22	3615.01	3916.26	4461.38	5494.25	6211.51	5881.57
9–10	7387.82	9037.53	8205.51	7158.30	6498.42	5106.92	4059.72	4317.93	4633.53	5795.50	6770.98	6656.21
10–11	7775.15	9640.03	8478.07	7057.88	6469.73	5236.03	4418.35	4820.02	4647.87	5594.66	6670.56	6727.94
11–12	7488.24	9367.47	7717.77	5996.33	5680.73	5236.03	4490.08	4791.33	4647.87	5192.99	6039.37	6326.27
12–13	6857.05	8822.35	7129.61	5766.81	5594.66	5236.03	4561.80	5020.85	4690.91	4891.74	5594.66	5938.95
13–14	5910.26	8062.05	6928.77	6139.78	5164.30	4647.87	4088.41	4576.15	3930.61	3987.99	4748.29	5020.85
14–15	4920.43	7473.90	6670.56	6297.58	4547.46	3959.30	3471.56	4059.72	3069.89	3270.73	4016.68	4174.48
15–16	4145.79	6311.93	5695.08	5035.20	3442.87	2912.09	2266.56	2739.95	2008.34	2467.39	2969.47	3285.07
16–17	2926.44	4906.09	4045.37	3385.49	2309.59	2065.72	1463.22	1793.16	1534.95	1735.78	2108.76	2108.76
17–18	416.01	1635.36	1147.62	0.00	0.00	0.00	0.00	329.94	0.00	0.00	114.76	157.80
18–19	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
19–20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
20–21	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
21–22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
22–23	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
23–24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 10
The monthly produced formic acid (gram).

Month	Formic Acid (gram)	Month	Formic Acid (gram)
January	1.84E+06	July	1.15E+06
February	2.25E+06	August	1.26E+06
March	2.21E+06	September	1.18E+06
April	1.86E+06	October	1.42E+06
May	1.68E+06	November	1.61E+06
Jun	1.35E+06	December	1.62E+06

respect to true distribution which is helpful to visualize the model-based data representation.

The two algorithms have demonstrated good performance in modeling the power and formic acid production. Referring to Fig. 6(a-d), R^2, MAE, and RMSE in modeling the power and formic acid production by SVM algorithm are 1.0, 0.23 kW, 0.29 kW, and 1.0, 10.47 g, 18.98 g respectively which are better that of the ANN-based predictions, i.e., 0.88, 4.22 kW, 5.43 kW and 1.0, 17.11 g and 26.36 g. Thus, it is noted that the modeling efficacy of the SVM to predict the two objectives is superior as measured in terms of higher R^2 value and the lower MAE and RMSE. Therefore, the SVM model is selected to be deployed for predicting the power and formic acid production under the causal effects of the input variables. In future studies, the model would be embedded in conducting optimization analytics which would estimate the optimal values of the causal variables, and the power and formic acid production would be maximized.

5. Conclusion

Employing a two-axis solar tracking system results in up to 60% increase in efficiency compared to a static PV system. However, different factor like high initial investment cost and relatively insignificant economic benefits often discourage the integration of this tracking system with PV systems equipped with battery storage. Moreover, the storage limitations hinder the accumulation of all power produced by the PV system. The current study proposed a novel solution to address these challenges by combining the PV system with the conversion of excess power to formic acid which is lucrative to efficiently utilize the whole generated power. This strategic integration mitigates the economic constraints associated with the high-cost investment of a two-axis solar tracking system, thereby shortening the payback period, and delivering substantial economic benefits to both investors and society at large. For this purpose, a case study is conducted based on Vietnam's Binh Thuan Province. The data cis collected for solar irradiance and simulated using MATLAB. Then a Machine Learning based model is developed to predict Formic Acid production under diverse operating conditions. The findings of the study can be used for establishing operational management strategies, evaluating the effectiveness and capacity of formic acid production facilities, and comprehending the impact of outside variables like temperature and sunshine on production. The major findings of the study are summarised as follows:

- Efficiency Improvement: It is observed that installing a two-axis solar tracking system in PV plants can significantly boost solar energy capture and formic acid production. For instance in the current case the maximum power production was observed to 2355 kW in February. It shows that by efficiently using excess electricity, this integration solves the problem of solar power's inherent variability.
- Economic Viability: By maximizing the return on investment in two-axis solar trackers, the conversion of solar energy into formic acid presents an economically appealing model. For instance the highest formic acid production was achieved in February 2.25 $\times 10^6$ grams because of the highest solar irradiance. This system is an appropriate choice for widespread adoption because it reduces energy waste and can shorten the payback period.
- Overall, the optimal solar irradiance and sun angles that maximize solar energy for the photovoltaic cells used in the production process are probably the cause of the high formic acid production in February.
- Machine Learning Insights: By comparing the two algorithms, the Support Vector Machine (SVM) algorithm proved to be a more precise predictor of PV system power output and the subsequent production of formic acid than the Artificial Neural Network (ANN). The results directly contribute to the larger aim of boosting the stability and uptake of renewable energy sources by supporting the study's objective of creating a dependable and commercially viable solar energy storage system.
- Future Implications: The study's findings emphasize the strategic importance of solar liquid fuel technology, in particular formic acid, as an alternative for the renewable energy industry. This could

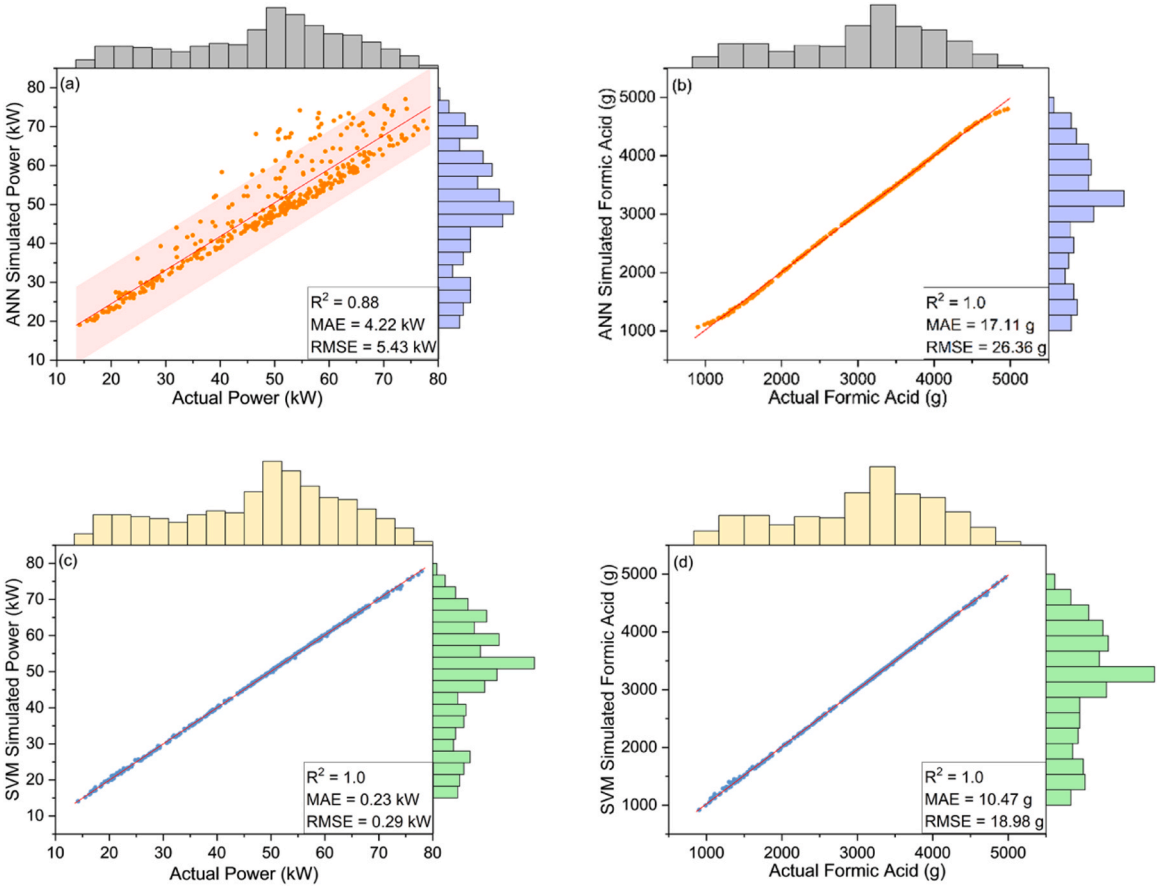

Fig. 6. Modeling performance of the ANN and SVM algorithms in simulating the PV power and formic acid production. (a) and (b) present the actual and ANN simulated responses for Power and Formic Acid production respectively. Whereas (c) and (d) show the SVM model-based modeling performance for Power and Formic Acid production respectively. SVM appears to be well suited to modeling the Power and Formic Acid production and the performance indicators like R^2, MAE, and RMSE are better than that of ANN.

accelerate the shift to a global energy framework that is more resilient to climate change and more sustainable.

This study bridges the gap between technological potential and practical implementation by addressing the mismatch between peak solar production and energy demand, highlighting the role of innovative energy storage in driving the renewable sector forward.

Recommendations

The formic acid's application can be examined in a direct formic acid fuel cell system. The fuel cell system is a system that converts thermochemical energy into electricity. The direct formic acid fuel cell is a net-zero carbon dioxide system. The products of formic acid in the electricity process in the fuel cell are only CO_2 and H_2O. Carbon dioxide and water, then can be recycled to produce formic acid and hydrogen again. The solar-driven Chlor-alkali process and the direct fuel cell are the most systems used recently. As, the Chlor-alkali process consumes a low total of energy, and has low investment, and operation cost. The operation and maintenance (O&M) cost of the Chlor-alkali process includes electricity utilization and reinstatement of the membrane. The direct formic acid fuel cell has electromotive force at 1.45 V compared to other types of polymer electrolyte membrane fuel cells and direct methanol fuel cells, at 1.18 V (Mardini and Bicer, 2021b). With these advantages, the direct formic acid fuel cell has great potential to develop on a commercial scale and could be applied widely in daily life.

Declaration of Competing Interest

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

Acknowledgment

The authors would like to thank Northumbria University and the British Council grant for the SAFECONOMY - Solar to Alternative Fuel Economy Research Consortia and Capacity Building Platform project. The authors also would like to thank RAEng / Leverhulme Trust Research Fellowships Tranche 19 for the FAM project (LTRF2223–19–103) awarded to Dr. Shahzad Northumbria University UK.

References

Afzal, A., Saleel, C.A., Bhattacharyya, S., Satish, N., Samuel, O.D., Badruddin, I.A., 2021. Merits and limitations of mathematical modeling and computational simulations in mitigation of COVID-19 pandemic: a comprehensive review, 2021 29:2 Arch. Comput. Methods Eng. 29, 1311–1337. https://doi.org/10.1007/S11831-021-09634-2.

ANON State of the Global Climate | 1 | World Meteorological Organization 2023. ⟨https://wmo.int/publication-series/state-of-global-climate⟩ (accessed January 15, 2024).

ANON Solar Energy Solutions vs. Other Renewable Sources: Which is the Best Choice? | CleanMax ⟨https://www.cleanmax.com/solar-update/how-solar-energy-compares-to-other-renewable-sources-of-energy.php⟩ (accessed January 25, 2024a).

ANON International Renewable Energy Agency T. RENEWABLE POWER-TO-HYDROGEN INNOVATION LANDSCAPE BRIEF 2019a.

ANON Formic Acid Market worth $618,808.7 Thousand by 2019b ⟨https://www.marketsandmarkets.com/PressReleases/formic-acid.asp⟩ (accessed January 15, 2024).
ANON Solar Power by Country 2024. ⟨https://worldpopulationreview.com/country-rankings/solar-power-by-country⟩ (accessed January 17, 2024a).
ANON Advantages Dual-Axis Solar Tracking | Zenithund New Energy ⟨https://en.zenithund.com/news/ns_3276.html⟩ (accessed January 17, 2024b).
ANON Handbook of NEURAL NETWORK SIGNAL PROCESSING 2002.
Ashraf, W.M., Dua, V., 2023. Artificial intelligence driven smart operation of large industrial complexes supporting the net-zero goal: coal power plants. Digit. Chem. Eng. 8, 100119 https://doi.org/10.1016/J.DCHE.2023.100119.
Ashraf, W.M., Uddin, G.M., Arafat, S.M., Krzywanski, J., Xiaonan, W., 2021. Strategic-level performance enhancement of a 660 MWe supercritical power plant and emissions reduction by AI approach. Energy Convers. Manag 250. https://doi.org/10.1016/J.ENCONMAN.2021.114913.
Ashraf, W.M., Uddin, G.M., Tariq, R., Ahmed, A., Farhan, M., Nazeer, M.A., et al., 2023. Artificial intelligence modeling-based optimization of an industrial-scale steam turbine for moving toward net-zero in the energy sector. ACS Omega 8, 21709–21725. https://doi.org/10.1021/ACSOMEGA.3C01227/ASSET/IMAGES/LARGE/AO3C01227_0012.JPEG.
Badgett, A., Ruth, M., James, B., Pivovar, B., 2021. Methods identifying cost reduction potential for water electrolysis systems. Curr. Opin. Chem. Eng. 33, 100714 https://doi.org/10.1016/J.COCHE.2021.100714.
Barthelemy, H., Weber, M., Barbier, F., 2017. Hydrogen storage: recent improvements and industrial perspectives. Int. J. Hydrog. Energy 42, 7254–7262. https://doi.org/10.1016/J.IJHYDENE.2016.03.178.
Chatterjee, S., Parsapur, R.K., Huang, K.W., 2021. Limitations of ammonia as a hydrogen energy carrier for the transportation sector. ACS Energy Lett. 6, 4390–4394. https://doi.org/10.1021/acsenergylett.1c02189.
Chong, L.W., Wong, Y.W., Rajkumar, R.K., Rajkumar, R.K., Isa, D., 2016. Hybrid energy storage systems and control strategies for stand-alone renewable energy power systems. Renew. Sustain. Energy Rev. 66, 174–189. https://doi.org/10.1016/J.RSER.2016.07.059.
Deb, P., Moradkhani, H., Abbaszadeh, P., Kiem, A.S., Engström, J., Keellings, D., et al., 2020. Causes of the widespread 2019–2020 Australian bushfire season. Earths Future 8. https://doi.org/10.1029/2020EF001671.
Dutta, S., 2021. Review on solar hydrogen: its prospects and limitations. Energy Fuels 35, 11613–11639. https://doi.org/10.1021/ACS.ENERGYFUELS.1C00823/ASSET/IMAGES/MEDIUM/EF1C00823_M031.GIF.
Dutta, I., Chatterjee, S., Cheng, H., Parsapur, R.K., Liu, Z., Li, Z., et al., 2022. Formic acid to power towards low-carbon economy. Adv. Energy Mater. 12 https://doi.org/10.1002/aenm.202103799.
Enke, D., Thawornwong, S., 2005. The use of data mining and neural networks for forecasting stock market returns. Expert Syst. Appl. 29, 927–940. https://doi.org/10.1016/J.ESWA.2005.06.024.
Eppinger, J., Huang, K.W., 2017. Formic acid as a hydrogen energy carrier. ACS Energy Lett. 2, 188–195. https://doi.org/10.1021/acsenergylett.6b00574.
Fräss-Ehrfeld C. Renewable energy sources: a chance to combat climate change. 2009.
Henry Fountain. Climate Change Affected Australia's Wildfires, Scientists Confirm - The New York Times ⟨https://www.nytimes.com/2020/03/04/climate/australia-wildfires-climate-change.html%20⟩⟨(accessed%20January%2025,%202024)⟩. (accessed January 27, 2024).
Huynh A. Development of deep learning predictive models for short-term solar radiation forecasting: case study in Vietnam 2020.
Ibrahim, A.I.Bin, Rahman, F.D.B.A., Rohaizat, M.Bin, 2018. Dual axes solar tracker. Int. J. Electr. Comput. Eng. (IJECE) 8, 1887–1892. https://doi.org/10.11591/IJECE.V8I3.PP1887-1892.
IRENA (International Renewable Energy Agency). Renewable Power Generation Costs in 2018. International Renewable Energy Agency 2019:88.
Irena. Renewable Power Generation Costs 2020. 2021.
IRNEA. IRENA (2022), Renewable Power Generation Costs in 2021, International Renewable Energy Agency, Abu Dhabi. ISBN 978-92-9260-452-3. International Renewable Energy Agency 2022:160.
Ishfaq, K., Sana, M., Ashraf, W.M., Dua, V., 2023. Sustainable EDM of Inconel 600 in Cu-mixed biodegradable dielectrics: modelling and optimizing the process by artificial neural network for supporting net-zero from industry. J. Clean. Prod. 421, 138388 https://doi.org/10.1016/J.JCLEPRO.2023.138388.
Koohi-Fayegh, S., Rosen, M.A., 2020. A review of energy storage types, applications and recent developments. J. Energy Storage 27, 101047. https://doi.org/10.1016/J.EST.2019.101047.
Krzywanski, J., Nowak, W., 2016. Artificial intelligence treatment of SO2 emissions from CFBC in air and oxygen-enriched conditions. J. Energy Eng. 142 https://doi.org/10.1061/(ASCE)EY.1943-7897.0000280.
Krzywanski, J., Skrobek, D., Sosnowski, M., Ashraf, W.M., Grabowska, K., Zylka, A., et al., 2024. Towards enhanced heat and mass exchange in adsorption systems: the role of AutoML and fluidized bed innovations. Int. Commun. Heat. Mass Transf. 152, 107262 https://doi.org/10.1016/J.ICHEATMASSTRANSFER.2024.107262.
Kubicek, M., Bork, A.H., Rupp, J.L.M., 2017. Perovskite oxides – a review on a versatile material class for solar-to-fuel conversion processes. J. Mater. Chem. A Mater. 5, 11983–12000. https://doi.org/10.1039/C7TA00987A.
Li, J., Pan, L., Suvarna, M., Tong, Y.W., Wang, X., 2020. Fuel properties of hydrochar and pyrochar: prediction and exploration with machine learning. Appl. Energy 269, 115166. https://doi.org/10.1016/J.APENERGY.2020.115166.
Mardini, N., Bicer, Y., 2021a. Formic acid synthesis and utilization for solar energy storage through solar-driven chloralkali process and fuel cells. Energy Storage 3, e235. https://doi.org/10.1002/EST2.235.
Mardini, N., Bicer, Y., 2021b. Direct synthesis of formic acid as hydrogen carrier from CO2 for cleaner power generation through direct formic acid fuel cell. Int. J. Hydrog. Energy 46, 13050–13060. https://doi.org/10.1016/J.IJHYDENE.2021.01.124.
Megia, P.J., Vizcaino, A.J., Calles, J.A., Carrero, A., 2021. Hydrogen production technologies: from fossil fuels toward renewable sources. a mini review. Energy Fuels 35, 16403–16415. https://doi.org/10.1021/ACS.ENERGYFUELS.1C02501.
Mehrpooya, M., Ghorbani, B., Shahsavari, A., Zaitsev, A., 2020. Conversion and storage of solar energy in the forms of liquid fuel and electricity in a hybrid energy storage system using methanol and phase change materials. Energy Convers. Manag 209, 112669. https://doi.org/10.1016/J.ENCONMAN.2020.112669.
Mellmann, D., Sponholz, P., Junge, H., Beller, M., 2016. Formic acid as a hydrogen storage material – development of homogeneous catalysts for selective hydrogen release. Chem. Soc. Rev. 45, 3954–3988. https://doi.org/10.1039/C5CS00618J.
Müller, K., Brooks, K., Autrey, T., 2017. Hydrogen storage in formic acid: a comparison of process options. Energy Fuels 31, 12603–12611. https://doi.org/10.1021/ACS.ENERGYFUELS.7B02997/SUPPL_FILE/EF7B02997_SI_001.PDF.
Muthukumar, P., Manikandan, S., Muniraj, R., Jarin, T., Sebi, A., 2023. Energy efficient dual axis solar tracking system using IOT. Meas.: Sens. 28, 100825 https://doi.org/10.1016/J.MEASEN.2023.100825.
Olabi, A.G., Abdelkareem, M.A., 2022. Renewable energy and climate change. Renew. Sustain. Energy Rev. 158 https://doi.org/10.1016/j.rser.2022.112111.
Pham, V.H.S., Tran, H.D., 2023. Research on applying machine learning models to predict the electricity generation capacity of rooftop solar energy systems on buildings. Asian J. Civ. Eng. 24, 3413–3423. https://doi.org/10.1007/S42107-023-00722-1.
Phan-Van, L., Dinh, V.N., Felici, R., Duc, T.N., 2023. New models for feasibility assessment and electrolyser optimal sizing of hydrogen production from dedicated wind farms and solar photovoltaic farms, and case studies for Scotland and Vietnam. Energy Convers. Manag 295, 117597. https://doi.org/10.1016/J.ENCONMAN.2023.117597.
Phap, V.M., Sang, L.Q., Ninh, N.Q., Binh, D.Van, Hung, B.B., Huyen, C.T.T., et al., 2022. Feasibility analysis of hydrogen production potential from rooftop solar power plant for industrial zones in Vietnam. Energy Rep. 8, 14089–14101. https://doi.org/10.1016/J.EGYR.2022.10.337.
Post, E., Alley, R.B., Christensen, T.R., Macias-Fauria, M., Forbes, B.C., Gooseff, M.N., et al., 2019. The polar regions in a 2°C warmer world. Sci. Adv. 5 https://doi.org/10.1126/SCIADV.AAW9883.
Qingfeng, L., Hjuler, H.A., Bjerrum, N.J., 2001. Phosphoric acid doped polybenzimidazole membranes: Physiochemical characterization and fuel cell applications. J. Appl. Electrochem 31, 773–779. https://doi.org/10.1023/A:1017558523354.
Quang, N., Duy, L., Van, B., 2021. Energy QD-EET on, 2021 undefined. Applying Artificial Intelligence in Forecasting the Output of Industrial Solar Power Plant in Vietnam. EudlEu 8, 1–9. https://doi.org/10.4108/eai.29-3-2021.169166.
Rangel-Martinez, D., Nigam, K.D.P., Ricardez-Sandoval, L.A., 2021. Machine learning on sustainable energy: A review and outlook on renewable energy systems, catalysis, smart grid and energy storage. Chem. Eng. Res. Des. 174, 414–441. https://doi.org/10.1016/J.CHERD.2021.08.013.
Singh, R., Altaee, A., Gautam, S., 2020. Nanomaterials in the advancement of hydrogen energy storage. Heliyon 6, e04487. https://doi.org/10.1016/J.HELIYON.2020.E04487.
Tarhan, C., Çil, M.A., 2021. A study on hydrogen, the clean energy of the future: Hydrogen storage methods. J. Energy Storage 40, 102676. https://doi.org/10.1016/J.EST.2021.102676.
Tariq, R., Ali, M., Sheikh, N.A., Shahzad, M.W., Xu, B.Bin, 2023. Deep learning artificial intelligence framework for sustainable desiccant air conditioning system: Optimization towards reduction in water footprints. Int. Commun. Heat. Mass Transf. 140, 106538 https://doi.org/10.1016/J.ICHEATMASSTRANSFER.2022.106538.
Trimm, D.L., 2005. Minimisation of carbon monoxide in a hydrogen stream for fuel cell application. Appl. Catal. A Gen. 296, 1–11. https://doi.org/10.1016/J.APCATA.2005.07.011.
Zayed, M.E., Kabeel, A.E., Shboul, B., Ashraf, W.M., Ghazy, M., Irshad, K., et al., 2023. Performance augmentation and machine learning-based modeling of wavy corrugated solar air collector embedded with thermal energy storage: support vector machine combined with Monte Carlo simulation. J. Energy Storage 74, 109533. https://doi.org/10.1016/J.EST.2023.109533.
Zhang, W., Ashraf, W.M., Senadheera, S.S., Alessi, D.S., Tack, F.M.G., Ok, Y.S., 2023. Machine learning based prediction and experimental validation of arsenite and arsenate sorption on biochars. Sci. Total Environ. 904, 166678 https://doi.org/10.1016/J.SCITOTENV.2023.166678.
Zhang, G., Yao, T., Xie, H., Wang, W., Yang, W., 2015. An inventory of glacial lakes in the Third Pole region and their changes in response to global warming. Glob. Planet Change 131, 148–157. https://doi.org/10.1016/J.GLOPLACHA.2015.05.013.
Zheng, Y., Ma, M., Shao, H., 2023. Recent advances in efficient and scalable solar hydrogen production through water splitting. Carbon Neutrality 2, 1–15. https://doi.org/10.1007/S43979-023-00064-6/FIGURES/9.