Journal of Infrastructure, Policy and Development 2024, 8(16), 8844.
https://doi.org/10.24294/jipd8844

Article

Estimating power generation of a combined cycle power plant using artificial intelligence technique

Vasileois Xezonakis, Olusegun D. Samuel, Christopher C. Enweremadu*

Department of Mechanical, Bioresources and Biomedical Engineering, University of South Africa, Science Campus, Florida, Johannesburg 1710, South Africa

*** Corresponding author:** Christopher C. Enweremadu, enwercc@unisa.ac.za

CITATION

Xezonakis V, Samuel OD, Enweremadu CC. (2024). Estimating power generation of a combined cycle power plant using artificial intelligence technique. Journal of Infrastructure, Policy and Development. 8(16): 8844. https://doi.org/10.24294/jipd8844

ARTICLE INFO

Received: 28 August 2024
Accepted: 30 September 2024
Available online: 24 December 2024

Abstract: Among contemporary computational techniques, Artificial Neural Network (ANN) and Adaptive Neuro-Fuzzy Inference System (ANFIS) are favoured because of their capacity to tackle non-linear modelling and complex stochastic datasets. Nondeterministic models involve some computational intricacies when deciphering real-life problems but always yield better outcomes. For the first time, this study utilized the ANN and ANFIS models for modelling power generation/electric power output (EPO) from databases generated in a combined cycle power plant (CCPP). The study presents a comparative study between ANNs and ANFIS to estimate the power output generation of a combined cycle power plant in Turkey. The inputs of the ANN and ANFIS models are ambient temperature (AT), ambient pressure (AP), relative humidity (RH), and exhaust vacuum (V), correlated with electric power output. Several models were developed to achieve the best architecture as the number of hidden neurons varied for the ANNs, while the training process was conducted for the ANFIS model. A comparison of the developed hybrid models was completed using statistical criteria such as the coefficient of determination (R^2), mean average error (MAE), and average absolute deviation (AAD). The R^2 of 0.945, MAE of 3.001%, and AAD of 3.722% for the ANN model were compared to those of R^2 of 0.9499, MAE of 2.843% and AAD of 2.842% for the ANFIS model. Even though both ANN and ANFIS are relevant in estimating and predicting power production, the ANFIS model exhibits higher superiority compared to the ANN model in accurately estimating the EPO of the CCPP located in Turkey and its environment.

Keywords: adaptive neuro-fuzzy inference system; combined cycle power plant; electric output power; machine learning; neural networks

1. Introduction

The generation of electricity is crucial for improving the quality of societal life. Thus, the continuous need to meet the increasingly emerging global energy demands has led to the development of efficient energy sources such as thermal power stations, wind, and solar energy (Botsaris et al., 2014). Thermal power plants have significantly contributed to satisfying energy needs, although they present several shortcomings regarding electricity performance. This gap is filled with the introduction of various types of combined cycle power plants (CCPP), offering a remarkable improvement in performance of up to 60%, along with environmental benefits (Ersayin and Ozgener, 2015; Hoang and Pawluskie, 2016). Technological evolution has recently brought a new wave of diverse approaches to predicting and diagnosing the performance of energy systems, including the power sector (Ntantis and Li, 2009; Ntantis and Li, 2013; Ntantis and Botsaris, 2016).

Therefore, the incorporation of modern methodologies such as Artificial Intelligence (AI) in terms of artificial neural networks (ANNs) (Ntantis and Botsaris, 2015), alternative soft computing techniques such as Fuzzy Logic (FL) (Ntantis, 2009), and the ANFIS (Panella and Gallo, 2005) has contributed to this direction. These methodologies present exemplary and distinguished characteristics due to their robustness and advanced computational competencies. The incorporation of the multi-layer perceptron neural network (MLPNN) configuration for forecasting the output electric power of a traditional power plant is acknowledged (Pa and Kazemi, 2022). Among these techniques, ANFIS has gained popularity and has been adapted by academics for several real-world applications, contributing to accurate solutions in the primary domain of interest, the energy sector. Interesting and novel applications can be found in energy prediction in housing (Ekki and Aksov, 2011), the stepping motors drive control (Melin and Castillo, 2005), prediction of time series systems (Melin et al., 2021), speed estimation of wind energy (Mohandes et al., 2016), and procedures of membrane separation (Rezakazemi et al., 2016).

An adaptive configuration network is used by ANFIS, a machine learning (ML) technique, coupling AI rules with respective FL models. The integration demonstrates a strong relationship between the independent and dependent thermodynamic parameters, as demonstrated by the present study's optimum electric power (EP) outcomes. Using this state-of-the-art toolbox, the membership functions (MFs) are adjusted in terms of the least squares using a back-propagation method (Mathworks, 2022; Revalthy et al., 2022). The development of ANFIS is demonstrated through the trained and learned input data structure, and its methodology is presented in the methodology section of the current study.

2. Literature review

The ANFIS has gained substantial attention in the energy sector, leading to reliable solutions. Therefore, its adaptation in the solar energy sector (photovoltaic) has contributed to optimum results for photovoltaic (PV) systems using maximum power point tracking. ANFIS has been successfully implemented, leading to the detection of faults in rear-ranged PV systems with efficient solutions (Bendary et al., 2021) and accurate prediction of PV power by achieving a lower error of 6.14%. According to Ibrahim et al. (2024), this is different from the 16% error associated with related regression analysis.

Ibrahim et al. (2017) conducted a comprehensive literature review of the adaptive neuro-fuzzy system's architecture through the PV in different categories, such as solar irradiance, output power forecasting, parameter identification of the photovoltaic system sizing, inverter control, and fault diagnosis. The study by Kaur, Kaur and Khanna (2021) on a novel hybrid model coupling ANFIS with a sub-clustering optimization and a grid partitioning algorithm contributes to optimum solar radiation predictions with superior solutions using performance and accuracy metrics. In the electricity sector for power forecasting, ANFIS depicted an improvement of the mean absolute percentage error by 0.4002%, defeating a time series model (11.4%), a first-order fuzzy time series model (5.74%) and a multi-linear regression model (10.62%) (Lorencin et al., 2019).

Ntantis (2009) has shown that the integration of ANFIS and GA in traditional thermal power plants accurately predicted performance assessment, making it a reliable approach. The synergy between ANFIS and GA as a hybrid model optimized the power extraction of a fuel-cell connected system with an outstanding performance rate of 98% (Ntantis and Li, 2009). Another study by Ntantis and Li (2013) compared the hybrid Particle Swarm Optimization-ANFIS (PSO-ANFIS) model with ANFIS and multiple linear regression techniques for predicting energy demand and highlighted the effectiveness of the PSO-ANFIS method. Furthermore, Ntantis and Li (2015) conducted a study on an integrated approach of the adaptive neuro-fuzzy inference system with an equilibrium optimizer proposed for the efficiency forecasting of a solar parabolic dish collector. An interesting hybrid technique coupling ANFIS with a wavelet transform by these authors (Ntantis and Botsaris, 2016) identified robust solutions for solar energy applications.

In electricity demand, ambient temperature is one of the most well-established predictors of electricity consumption, due to its influence on heating and cooling for residential applications. Therefore, the impact of urban heat islands and increased temperatures on energy consumption for cooling purposes is significant, highlighting the impact of temperature on peak load considering that the hottest days correspond to the highest electricity demand (Santamouris, 2015).

The impact of atmospheric pressure on power generation is vital for the performance of gas turbines, as it improves of the combustion process and contributes to the overall the plant performance. Therefore, an exploration of the correlation between pressure fluctuations and power output changes for plant's operating at high altitudes is also proposed (El Hadik, 1990).

Research has shown that relative humidity affects power generation in the combustion process of a gas turbine. A 15% increase in RH can lead to a 0.7% increase in output power (Shukla and Singh, 2014).

Adopting the adaptive neuro-fuzzy inference methodology based on various meteorological data including monthly mean, minimum, and maximum temperatures, has contributed to estimating solar radiation (Petkovic et al., 2013). Several studies have been conducted in the wind power field, leading to significant outcomes. Therefore, optimal wind power efficiency outcomes via ANFIS are presented by Nabipour et al. (2019). Accurate matching of the dependent (output) data between a climate model and a lower wind potential is also achieved (Gill and Singh, 2017). An additional case study is also performed by Shanbedi et al. (2014) for the statistical estimation of the coefficient of performance of a refrigeration system through the adaptive neuro-fuzzy inference system. The application of ANFIS in forecasting the performance of a two-phase closed thermosiphon, leading to optimum solutions has been studied by Kaur et al. (2021). In their study, Vimala et al. (2024) utilized an integrated tool by combining an ANFIS-based controller and a Taguchi optimization technique to forecast the output power of a grid-connected PV system.

In the CCPP, the output electric power is forecasted via ANFIS and ANNs, and their comparison depicts the effectiveness of ANFIS, achieving performance metrics of mean square error 16.6887 with a strong correlation (R^2) of 0.9711

(Nugraha et al., 2024). An interesting study by Abuayyash et al. (2024) in the renewable resources field in Northern Sumatra (Indonesia) using ANFIS has produced significant outcomes. The study focused on forecasting renewable resources development for various independent factors, achieving a minimal error of 0.000201092% and a forecasted developmental value of 160.44 MW. The selection criteria of both the respective techniques (ANN and ANFIS) over the automated machine learning algorithms (AutoML) are detailed below. Furthermore, ANNs require more computational resources by means of a hyperparameter tuning approach to prevent overfitting. A drawback of ANFIS is the requirement of careful tuning and that its performance is dependent on how the fuzzy rules are defined (Zamani et al., 2015).

Gaps and aims of the study

Table 1 highlights the implementation of AI and ML techniques in traditional and hybrid combined power plants. Recently, many tools have been adopted to model and forecast the EP output of CCPP plants. However, there is inadequate research on selecting pioneering tools or methods, such as the ANFIS-based extrapolative models, to simulate nonlinear patterns in the CCPP.

Table 1. An overview of soft computing tools in CCPP plants.

S/N	Types of plants	Plants ‘choice constraints	Types of Soft computing (SC)	Responses	Remarks	Technical gaps	Refs.
1	Combined cycle power plants	Ambient temperature, exhaust vacuum, atmospheric pressure, and relative humidity	ANN and Electrostatic discharge algorithm (ESDA)	Electric power output forecasting.	The ESDA-ANN models outperformed ASO studies in literature	ANFIS was recommended to determine the efficacy ESDA for power plants	Zhao and Foong (2022)
2	Cycle power plant	fuel gas heat input, CO_2 percentage, and power output	ANN	Heat rate prediction	The best prediction of heat rate data with a regression R^2 value of 0.995 reported	Need to enhance the performance using other SC tools	Zaaoumi et al. (2021)
3	CCPP	Ambient temperature, atmospheric pressure, re5lative humidity, exhaust steam pressure	ANFIS and RF	vacuum (V) and power output of the CCPP	Best performance realized from RF	ANN and ANFIS have not comprehensive studied and correlated with the Plants ‘choice constraints	Bandić et al. (2020)
4	CCPP	Temperature, ambient pressure, relative humidity, exhaust vacuum	ANN	Electric power	ANN reported to be trustable in managing CCPP	The need to adopt ANN as the alternate SC tool in predicting CCPP emphasized.	Akdemi (2016)
5	Solar chimney power plants	Time, ambient temperature, solar radiation	ANN and ANFIS	Air velocity inside chimney	ANFIS model exhibited better performance than ANN.	The study cannot be adaptable to CCPP because it is solar based plant	Amirkhani et al. (2015)

Table 1. (*Continued*).

S/N	Types of plants	Plants 'choice constraints	Types of Soft computing (SC)	Responses	Remarks	Technical gaps	Refs.
6	Combined Cycle Power plant (CCPP)	Ambient Temperature, Exhaust Vacuum, Ambient Pressure, Relative Humidity	Machine Learning Methods (MLAs)	CCPP hourly Electric power prediction	Various MLAs, such accurate prediction achieved through KNN, GBRT, LR, ANN, and DNN for the electric power output	Results show that the state-of-the-art surpasses GBRT by predicting the optimum electric power output.	Siddiqui et al. (2021)
7	Combined Cycle Power Plant (CCPP)	Ambient temperature, Exhaust Vacuum, Ambient pressure, Relative humidity	Hybrid Machine Learning approaches	Power plant's output power with the minimum waste	BOA, PPE, and SVM models forecasted the output power of CCPP during during the power outage to avoid technical issues	BOAPPE methodology improved the convergence speed, avoiding the trapping into local optimum solutions	Wang et al. (2023)
8	Combined Cycle Power Plant boiler	Ambient Temperature, Exhaust Vacuum, Ambient Pressure, Relative Humidity	Hybrid Machine Learning Technique	CCPP hourly output power estimation	Accurate prediction of electric power output achieved through an integrated MLP, ANN, and GA techniques studies	Other heuristic algorithms for MLPs design of power plant proposed	Lorencin et al. (2019)
9	Combined Cycle Power Plant	Ambient Temperature, Vacuum Exhaust, Ambient pressure, Relative Humidity	Multi-model ensemble and a traditional machine learning method such as the RF	An efficient and reliable CCPP power output under full conditions prediction model	Accurate power evaluation of A CCPP plant with the conventional Machine Learning algorithm (RF) for robust and electricity generation and utilization	A more in-depth prediction of electricity using artificial intelligence technologies is aimed for the future.	Qu et al. (2021)
10	Combined Cycle Power Plant	Ambient Temperature, Vacuum Exhaust, Ambient pressure, Relative Humidity	Machine Learning Approaches (MLA's)	CCPP output power complete load forecasting and anomaly detection.	Reduction in shortage in operation in CCPP plant observed with LR, SVM, RF, and ANNs	A data science approach is recommended to uncover hidden information from sensors that is unintelligible to humans.	Hundi and Shahsavari (2020)

LR = Linear Regression; SVM = Support Vector Machines; RF = Random Forests; KNN = the kth-nearest neighbours; GBRT = Gradient-Boosted Regression Rate; DNN = Deep Neural Networks; BOA = Butterfly Optimization Algorithm; PPE = Phasmatodea population evolution algorithm; PPE = Phasmatodea population evolution algorithm; BOAPPE = BOA + PPE; GA = Genetic Algorithm.

In the present study, ANFIS is used to forecast the EP output and consider the heat conversion from the exhaust gas of the gas turbine, contributing to CCPP efficiency. It is centred around the estimation of power generation based on ambient parameters such as Ambient Temperature (AT), Exhaust Vacuum (V), Atmospheric Pressure (AP), and Relative Humidity (RH). The selection of these parameters is crucial for the model's simplicity, accuracy, interpretability, and efficiency for output power prediction However, other parameters that might impact energy demand or generation were deliberately excluded due to data availability and quality, model simplicity and interpretability, relevance of the present and computational efficiency.

The practical implications on power plant operations include adjusting generation capacity to meet electricity demands more efficiently, minimizing energy waste. Applications of these predictions can be found in allocating resources (fuel, labour) to ensure optimal use of the plant's capacity. Another application is better scheduling of maintenance activities during periods of lower demand, reducing the risk of service interruptions. In grid management, balancing supply and demand is crucial to ensure proper electricity supply and reduce the possibility of blackouts or energy shortages. Power grid operators can use time predictions to dynamically control load in different regions, preventing overloading during peak demand periods. Predicting electricity demand based on ambient parameters, can assist grid operators in handling extreme weather conditions that could lead to supply shortages (Saleel, 2021).

The ANFIS structured configuration is implemented using MATLAB's GUI capabilities and command lines. The study considers the entire dataset of 9568 samples, unlike previous studies which used reduced versions. A comparison with existing ANFIS studies is also conducted. Another objective is to compare ANFIS with neural networks (ANN) for different test cases in the energy sector, using key performance indicators such as regression coefficient (R^2), root mean square error (RMSE), mean average error (MAE), and absolute average deviation (AAD). The study aims to achieve robust EP prediction with computational cost benefits in terms of simulation time.

3. Methodology

3.1. Operational diagram and description of the CCPP dataset

The CCPP layout is depicted in **Figure 1** and is designed for electricity production. Initially, the fuel is combined with compressed ambient air and then burned within the combustor. The gas turbine powers a generator to create energy, propelled by the resulting hot gases. Afterwards the exhaust gases are fed into the heat recovery steam generator (HRSG) equipped with a large amount of thermal energy. The steam from the steam turbine is condensed to feedwater into the HRSG, transforming it into high-pressure steam. The energy produced by the steam provides power to the steam turbine attached to another generator. CCPP outperforms standalone systems by exploiting the energy from the gas and steam cycles.

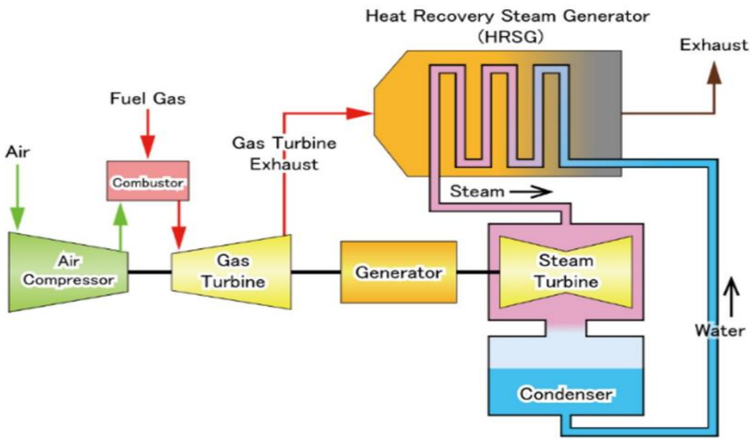

Figure 1. Schematics of the functional (Gas Turbine Combined Cycle Generation, 2020).

In the present test case, the thermodynamic design variables that impact the combined CCPP cycle's plant performance under full working conditions are the ambient temperature (AT), exhaust vacuum (V), ambient pressure (AP), and relative humidity (RH). The main concern is that developing a trustworthy mathematical model for the CCPP has become challenging, as reported recently by Xezonakis et al. (2024). Access to the massive CCPP dataset comprising time-series data from a 420 MW gas-fired power plant for six years (2006–2011), a well-known dataset used in machine learning energy field research.

This data collection includes 9568 hourly measurement samples of the independent variables (AT, V, AP, and RH) and the dependent feature (EPO) including, 674 daily datasets in (.*xls*) format from the original dataset. The data collection was made via sensors deployed in various locations capturing the independent and dependent features. The processing analysis involves the integration of the 674 datasets to form a continuous set. The handling of noisy data improved the signal-to-noise ratio, ensuring that the model was not trained on faulty data points. Additionally, any related disturbances caused by power fluctuations and sensor malfunctions were filtered out. The handling of anomalies or unexpected patterns in the dataset was managed by filtering out values beyond defined thresholds, or eliminating externally affected data. Hence, this structured and clean dataset provides accurate modelling, and its quality improvements ensure its reliability and applicability in real world applications (Faahmi et al., 2022).

The maintenance of the data's quality took place through different steps. The conflicted data points outside the acceptable range were filtered out, while the noisy data resulting from electrical disturbances were eliminated. The complex structure of a CCPP consists of various components, and the performance is highly dependent on factors affecting safety, reliability and availability. Any related anomalies detected in the data are handled through real-time health monitoring of the constituents. Real-time health monitoring contributes to the integrity of a CCPP. The status of these online components can be accessed via the incorporation of sensors, installed on critical parts of equipment in the CCPP. According to this actual data, condition monitoring is considered in each component to assist the operator in detecting and diagnosing the anomalies accurately, taking all the maintenance work needed (Faahmi et al., 2022).

However, there are different challenges in proposing an effective scheme for the CCPP. The respective operational data present inherent features in terms of cross-variable association. A few variables are related to others according to their working mechanism, making the elimination of the interpreted conditions results inconvenient (Zhang et al., 2020). Another reason is the accurate anomaly detection, and the orientation of the monitoring systems sounds difficult due to the unavoidable noises and errors resulting from different working environments and sensor performance. Another parameter to be considered is the complexity of each CCPP system, which obstructs the operator from identifying the causes of faults due to their random nature, causes of securing sufficient representation for all these errors (Hundi and Shahsavari, 2020).

Facing these problems in combined cycle power plants is also done through the adoption of various investigations to reduce the uncertainties in these anomalies. Hence, a novel performance model for CCPP using reliability block diagrams to illustrate illustration of the relationships among the subsystems is proposed (Sabouhi et al., 2016). Furthermore, a coupled generalized regression neural network (GRNN) with a B-spline contributed to reducting collinear issues between sensors (Chen et al., 2015). Therefore, the final version of the dataset presents a clear structure and organization with 9568 rows (samples), where each row represents a single hourly measurement and 5 columns (features). The four columns represent the independent variables (AT, V, AP and RH) and the 5th column represents the dependent feature to be predicted (EPO). This clear and well-structured dataset ensures accurate predictive modelling of the target variable.

A sample description of the input and output variables of the massive dataset is listed in **Table 2**.

Table 2. Actual data received from a CCPP.

Feature	Type	Minimum %	Maximum %	Unit
Ambient Temperature	Input	1.81	37.11	°C
Ambient Pressure	Input	992.89	1033.30	Mbar
Relative Humidity	Input	25.56	100.16	%
Exhaust Vacuum	Input	25.36	81.56	cmHg
Electric Power Output	Output	420.26	495.75	MW

Source: Xezonakis et al. (2024).

3.2. Design and research methodology

3.2.1. ANN configuration

Developing the Multilayer Perceptron (MLP) technique for the current test case (CCPP) using the Neural Network Toolbox in the MATLAB environment, the entire dataset (9568) is split into training and testing sets. **Figure 2** shows the structure of a neural network (ANN) composed of several adaptable units (neurons) and the Multilayer Perceptron (MLP) configuration is visible when one or more neurons exist. The respective values of the input layer are processed by an activation function (transfer function), while the MLP learning process occurs to predict the electric power (EP) output value in terms of the outputs sent to the neurons of the output layer. This section may be divided by subheadings. It should provide a concise and precise description of the experimental results, their interpretation, as well as the experimental conclusions that can be drawn.

Two different types of activation functions such as the logsigmoid (*logsig*) and the hyperbolic tangent function (tansig), are employed with assigned values within the range from 0 to 1. Additional information can be found in Xezonakis et al. (2024). The dependent forecasted variable is the electric power output, and four independent parameters are received by the neural network (ANN) as part of the modelling process: (i) ambient temperature, (ii) exhaust vacuum, (iii) ambient pressure, and (iv) relative humidity. A dataset of 9568 samples for a period of six years (2006–2011) is adopted and the network's performance and reliability are expressed by means of MSE, MAE, AAD, and R^2 metrics. For novelty reasons, 70% of the entire dataset is trained, 15% is tested, and 15% is validated. **Figure 2** depicts a sample neural network (ANN) with the four independent variables (AT, V, AP, and RH), nine neurons in the hidden layer, and an output layer forecasting the electric output parameter. In the current case study, a multilayer perceptron feed forward (MLP) network is applied, based on MATLAB neural networks toolbox capabilities as illustrated and explained in section 3.2.2.

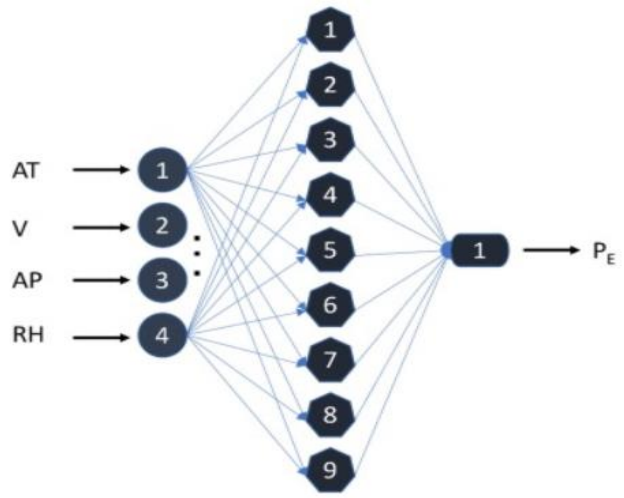

Figure 2. Sample ANN configuration with 9 neurons of the current work.

Figure 3 illustrates the respective architectural diagram of the entire procedure. The proposed testing, training, and validation process takes place through 1000 validation steps and three different training algorithms; Levenberg-Marquardt, Bayesian Regularization, and Scaled Conjugate Gradient are employed for 10,000 iterations via the graphical user interface (GUI) capabilities of the well-established MATLAB code. In the present work, amendments of the dataset percentages of the training/testing/validation provided less accurate and reliable outcomes. Both input and target data are provided, and the modified weight biases match the targets with the real data.

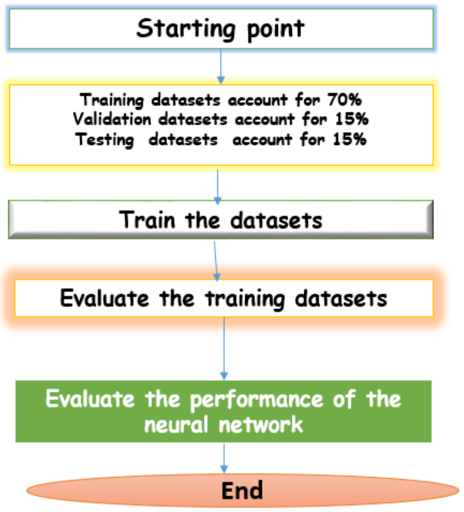

Figure 3. Architecture of the block diagram (Xezonakis et al., 2024).

3.2.2. Neural networks (ANN) metrics

In the current test case for novel outcomes the dataset is split into-training (70%), validation (15%), and testing (15%) sets. The reason for this splitting is to monitor the model's performance during training and to perform early stopping or hyperparameter tuning. The validation set (15%) assists in avoiding overfitting, by providing feedback on the model's performance before it is tested on the test data. The validation set (15%) is also checked during the training process. Different splits might be challenging since ANN is focused on refining the model through validation, and the primary concern is to obtain more accurate results. For the ANFIS model this is justified as: ‘‘ ANFIS is a hybrid model combining fuzzy logic with ANN, and it is less prone to overfitting by means of its inherent structure. Thus, a 70:30 split (training and checking) might be adequate since it does not require a separate validation set for tuning purposes. Furthermore, it relies on rules derived from fuzzy logic, and its simpler parameter space does not necessitate a validation set, compared to ANN, allowing for a direct 70:30 split without sacrificing the model's generalization or performance. The main target is to find the best model for power generation, and this split also contributes to the identification of the model's performance on the unseen data. The lack of validation might reduce any opportunities for the model's optimization during training. Although ANFIS had more data for testing compared to ANN, accurate outcomes are still provided, thus each method is studied individually. The respective dataset settings are summarized in **Table 3**. The network's database testing process evaluates performance and accuracy using Mean Square Error (MSE), MAE, AAD, and correlation coefficient (R^2), excluding the MBE for long-term multidimensional forecasted data. **Figure 4** illustrates the current multilayer perceptron (MLP, 4-20-1-1) structure, showing the four input parameters with the hidden layers (20) and the output layer, predicting the output variable.

Table 3. Dataset setting for features and types (Xezonakis et al., 2014).

Feature	Type
Data size	9568
Adopted Variables	AT, V, AP and RH
Number of hidden layers (neurons)	20
Training set (%)	70
Testing set (%)	15
Validation set (%)	15
Training Algorithms	LM, BR and SCG

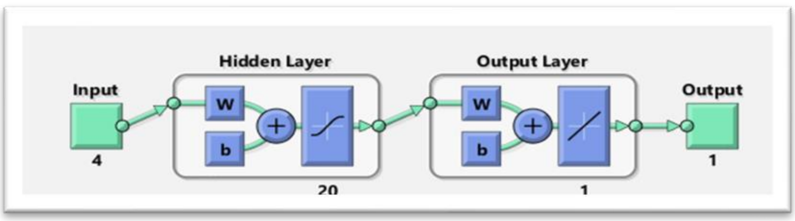

Figure 4. ANN structure with 4 input parameters and 20 neurons (Xezonakis et al., 2024).

3.2.3. Assessment of the performance metrics

Statistical variables such as regression coefficient, root mean square error, mean average error, and absolute average deviation were adopted to determine the predictive superiority of the model techniques. Equations (1)–(4) were used to evaluate the statistical metrics of the ANN and ANFIS models. The results were utilised to assess the superiority and effectiveness of the model techniques.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - \ddot{y}_i)^2}{\sum_{i=1}^{n}(\bar{y}_i - \ddot{y}_i)^2} \tag{1}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n} |\hat{y}_i - y_i|^2 \tag{2}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |\hat{y}_i - y_i|, \tag{3}$$

$$AAD = \frac{1}{n}\sum_{i=1}^{n} |y_i - \overline{y_i}| \tag{4}$$

where y_i, $\hat{y}_i$, and $\overline{y_i}$ are the actual (experimental) data, predicted data, and mean value of the data set. The main reason for selecting these performance indicators is that these values will be compared with identical studies (Ani and Agu, 2022; Samuel et al., 2021; Tiwari et al., 2012).

3.2.4. Correlation analysis of data

The correlation matrix~~,~~ quantifies the linear relationship between each ambient parameter (AT, V, AP, and RH) and the target variable (EPO), as shown in **Table 4**.

The negative value in the correlation (R = −0.9481) for AT implies that the EPO) shows that as the ambient temperature increases the electricity demand decreases and in higher temperatures this also contributes to decreased power generation. A strong negative correlation (R = −0.6898) for V indicates the influence of the output power (EPO) as part of the power generation. A moderately strong correlation (R = 0.5184) for AP indicates that the target variable is due to the impact on the turbine performance in power generation systems or the weather conditions affecting any energy consumption patterns. The weak positive correlation (R = 0.3898) shows that a higher value of RH is related to higher electricity demand or power output. Humidity also affects the cooling systems, which might increase energy consumption.

Table 4. Correlation matrix between the independent and the dependent parameters.

Variables	AT	V	AP	RH	EPO
AT	1.000	0.8441	−0.5075	−0.5425	−0.9481
V	0.8441	1.000	−0.4135	−0.3122	−0.8698
AP	−0.5075	−0.4135	1.000	0.0996	0.5184
RH	−0.5425	−0.3122	0.0996	1.000	0.3898
EPO	−0.9481	−0.8698	0.5184	0.3898	1.000

Figure 5 shows a visual interpretation of the correlation matrix via heatmaps, depicting the strength between these variables. The stronger the correlation (closer to +1, or –1), the intensity of the colour increases. **Figure 6a–d** also illustrates a comprehensive view of how the four input parameters are correlated with EPO, and more accurate models can be built for future energy predictions.

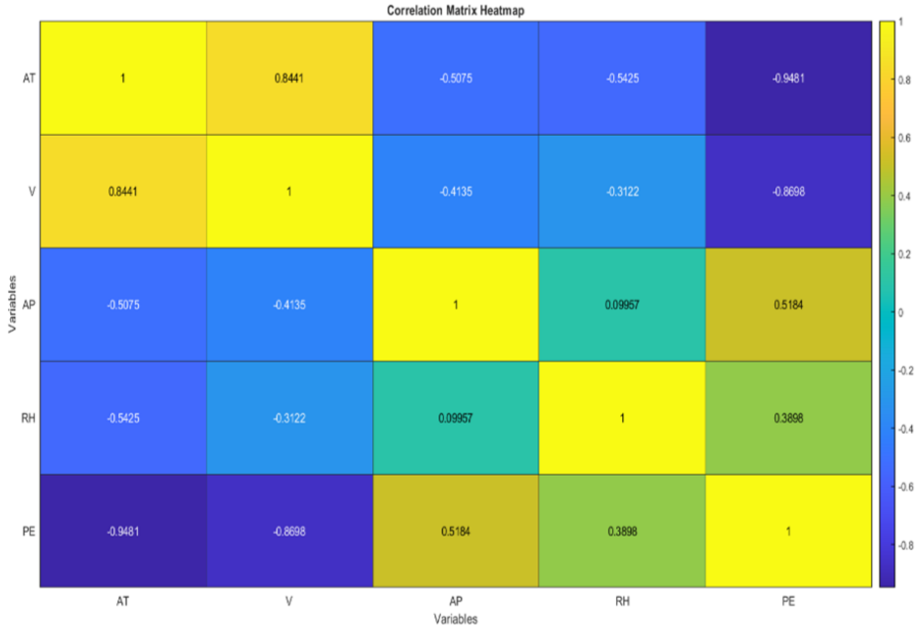

Figure 5. Visualization of the independent and the dependent parameters of the correlation matrix via heatmaps.

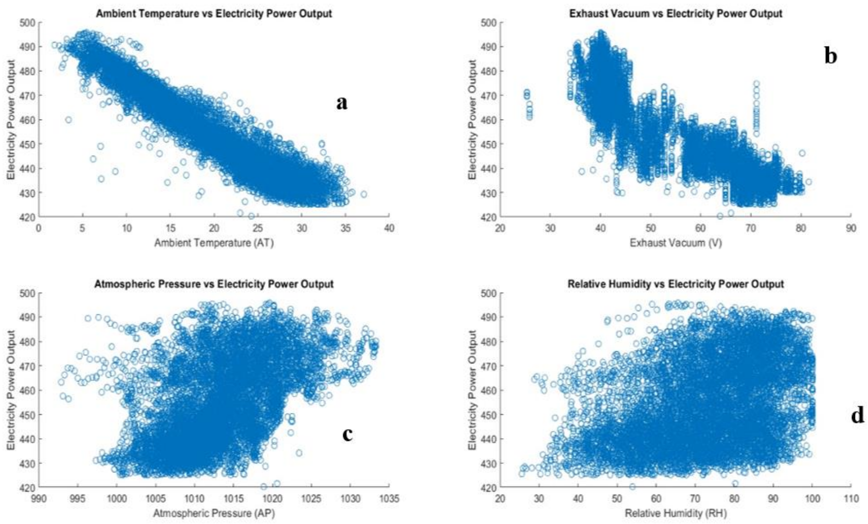

Figure 6. Correlation attitude of each input parameter with the dependent feature (EPO): **(a)** electricity power output (EPO) vs. ambient temperature; **(b)** EPO vs. exhaust vacuum; **(c)** EPO vs. atmospheric pressure; **(d)** EPO vs. relative humidity.

3.2.5. Sensitivity analysis of the input parameters

Figure 7 illustrates the sensitivity and impact of each input parameter (AT, V, AP and AP) on the prediction of both models (ANN, ANFIS). Ambient pressure presents the most fluctuating curve, indicating a significant effect on the predicted output power and consequently, the model is more sensitive to this feature. The sensitivity analysis for both models (ANN, ANFIS) reveals that ambient pressure is the most important feature, while V is of least importance. Robustness depends on which input parameter shows less sensitivity to minor perturbations in the prediction of the output power. Therefore, the sensitivity curves of the two models indicate the superiority of the lowest EPO changes (deviations) fitting with uncertainties in the input variables.

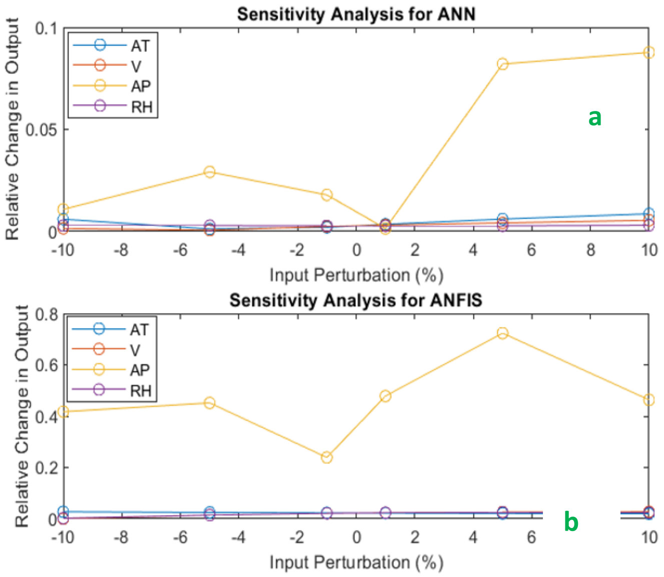

Figure 7. Sensitivity analysis of soft computing tools (ANN, ANFIS) to predict changes in the output parameter: **(a)** ANN; **(b)** ANFIS.

3.2.6. ANFIS methodology

Four input variables constitute the ANFIS architecture: AT, V, AP, and RH, and each parameter relates to three Gaussian MFs. After the introduction of the Fuzzy Inference System (FIS), the first-order Sugeno model implements IF-THEN rules and then uses these input variables. Although any adaptation of a higher-order Sugeno model is conceivable, it is shown that its increasing complexity does not result in reliable solutions (Mathworks, 2022; Revalthy, et al., 2022). Therefore, the four design variables have 48 rules altogether because the four variables are multiplied by the total number (12) of MFs. ANFIS is a hybrid model combining fuzzy logic with ANN and it is less prone to overfitting by means of its inherent structure. Thus, a 70:30 split (training and checking) might be adequate since it does not require a separate validation set for tuning purposes. Furthermore, it relies on rules derived from fuzzy logic, and its simpler parameter space does not necessitate a validation set compared to ANN, allowing for a direct 70:30 split without sacrificing the model's generalization or performance. The main target is to find the best model for power generation, thus this split also contributes to the identification of the model's performance on the unseen data. The lack of validation might reduce any opportunities for the model's optimization during training. Although ANFIS had more data for testing compared to ANN still accurate outcomes are provided, thus each method is studied individually. A sample architecture showing the four input parameters within the five-layer configuration is depicted in **Figure 8**. Further details on the training, checking, validating the FIS model for optimal EPO solutions can be found in the flow chart of the entire algorithm ANFIS (see **Figure 9**). **Figure 10** shows the structured fuzzy inference system (FIS) of the four independent features (AT, V, AP, RH) and the dependent parameter (EP), based on ANFIS (GUI) attributes. **Figures 11–14**, depict the Sugeno MF fuzzy model with Gaussian shaped MF of each input parameter (AT, V, AP, and RH) for the fuzzification process with optimal outcomes incorporated. These membership functions are divided into low, medium and high fuzzy sets described by Gaussian functions. **Table 5** summarizes the attributes of these features.

Table 5. Summary of the membership attributes of the input features.

Input feature	Range	MF1 (low), Gaussian range	MF2 (medium), Gaussian range	MF3 (high), Gaussian range
AT	[1.81, 37.11]	[7.3312, 1.6389]	[7.3328, 19.386]	[7.6892, 370632]
V	[25.36, 81.56]	[11.6570, 25.2809]	[11.5163, 53.4434]	[12.2653, 81.4055]
AP	[992.89, 1033.33]	[8.5480, 993.1911]	[8.2323, 1013.1619]	[8.6337, 1032.955]
RH	[25.56, 100.16]	[16.1217, 26.0420]	[15.7726, 62.9738]	[15.8058, 100.1423]

Layer 1: Three different membership functions corresponding to each input parameter (AT, V, AP and RH) are transformed into linguistic labels and the adapted Gaussian membership function (Gaussmf) is defined by Equation (5). The ANFIS technique involves layers 1–4 explained as follows:

$$F(x, \sigma, \quad c) = e^{-\frac{(x-c)^2}{2\sigma^2}} \tag{5}$$

where as x: the input value,σ: the standard deviation, and c: mean value.

Layer 2: A combined product-of the 12 membership functions using the four input variables through weighted functions ($w_i,\ i = 1\ ,2, 3$) is considered.

Layer 3: Calculation of the ratio between the *i-th* rule's firing strengths ($\overline{w_i}, i = 1\ ,2\ and\ 3$) and the sum of all the individual rule's firing strengths ($w_i, i = 1\ ,2, 3$) is assumed.

Layer 4: An emulsion between the previous layer (3) and the first-order Sugeno fuzzy model denoted, as ($\overline{w_i}\, f_i,\ i = 1, 2\ and\ 3$), is depicted.

Layer 5: Incorporation of one node, presents the network's output as the sum of the weighted output from the previous layer (4) for optimum forecasting of the output parameter (EP).

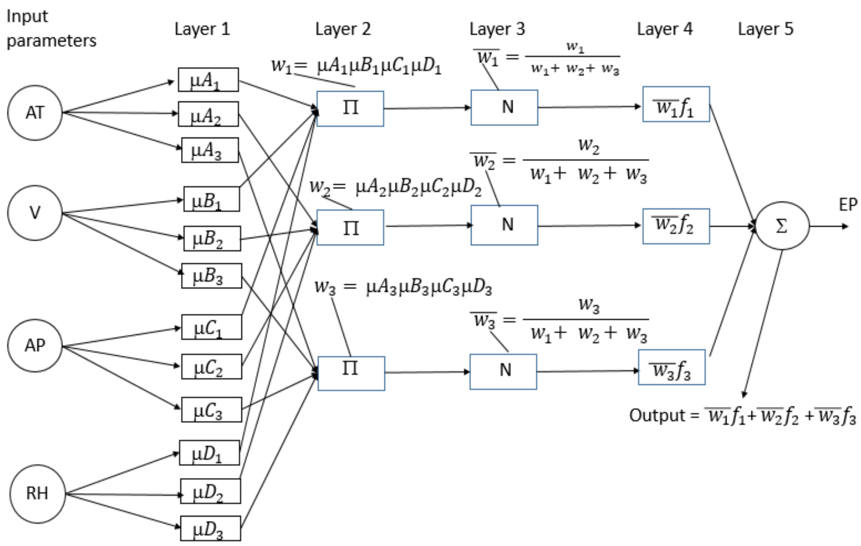

Figure 8. ANFIS network with four input variables and five layers.

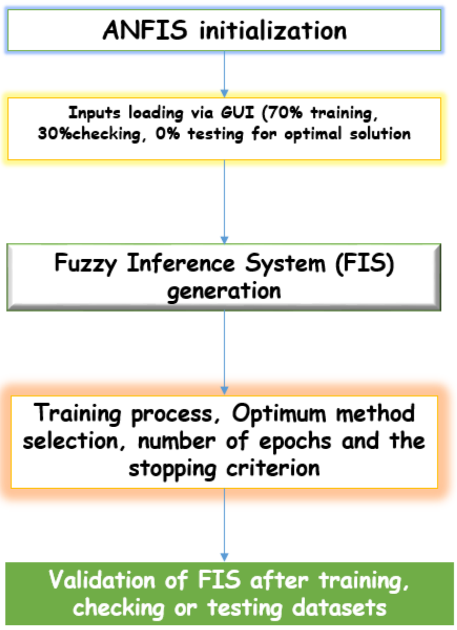

Figure 9. Flowchart of ANFIS toolbox.

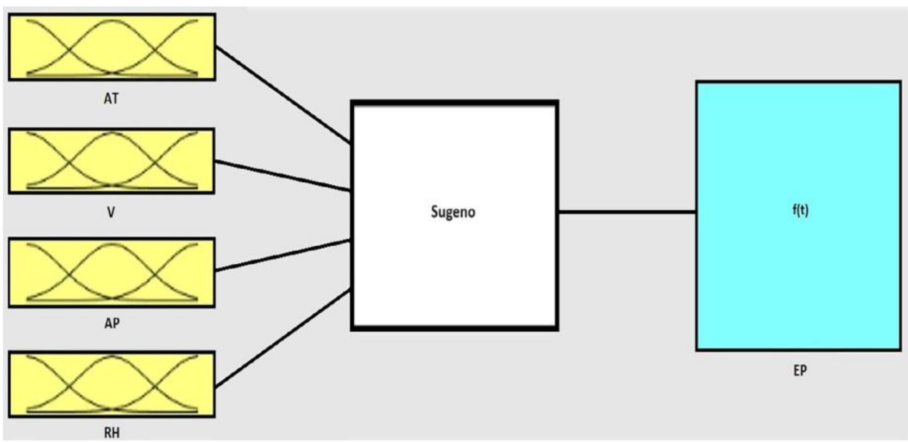

Figure 10. FIS model of the four independent variables.

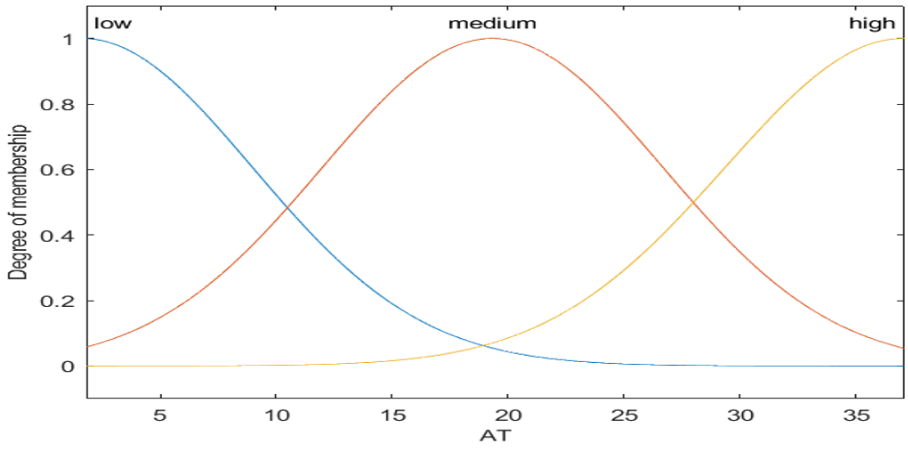

Figure 11. Gaussian membership function of the ambient temperature features.

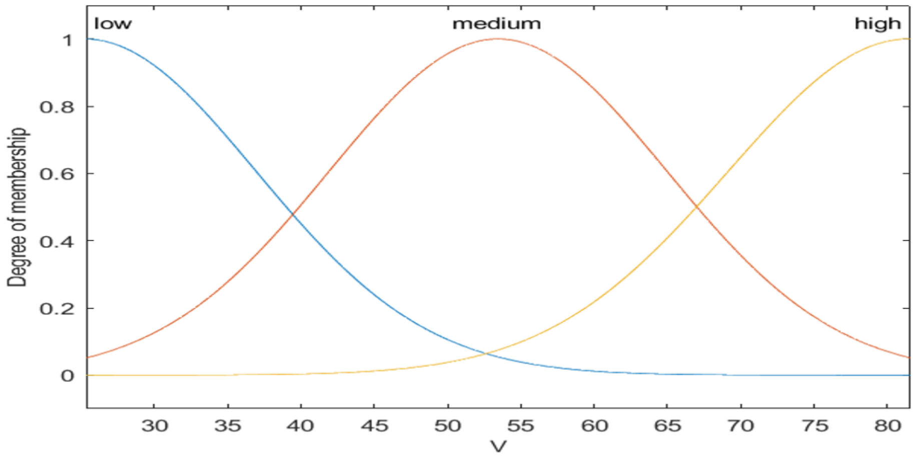

Figure 12. Gaussian membership function of the exhaust vacuum feature.

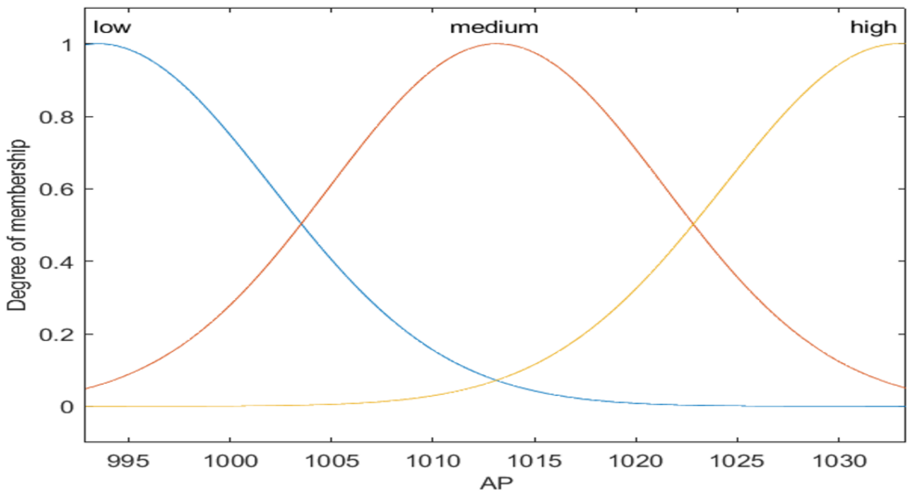

Figure 13. Gaussian membership function of the ambient pressure feature.

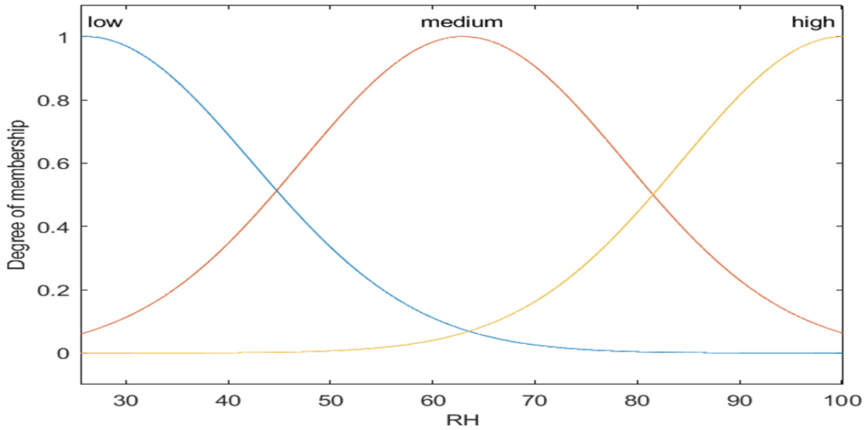

Figure 14. Gaussian membership function of the relative humidity feature.

4. Results and discussion

4.1. ANN model analysis

This section highlights the key findings and their analysis as well as a few interesting conclusions regarding their relevance. Each simulation yields a different outcome after the completion of the training procedure with the parameters for the entire dataset due to the network's bias and randomly initialized weights.

After comparising the three training algorithms (LM, BR and SCG) for the four independent features, it is evident that the BR algorithm stands out with outstanding outcomes based on the MSE performance metric and the correlation coefficient R^2, as illustrated in **Table 6**.

Table 6. Comparison between the training algorithms (LM, BR and SCG, (Xezonakis, et al., 2024).

Training Algorithm	Mean Square Error (MSE %)	Correlation Coefficient (R^2)
LM	16.9874	0.9426
BR	16.3068	0.9441
SCG	16.6200	0.9395

Figures 15 and **16** show the lowest value of MSE = 16.3068% for the given number of epochs for both the training and testing datasets as well as for the entire dataset and robust outcomes with a very strong correlation R^2 = 0.944.

The respective values of the other performance metrics of the ANN test case for all the datasets are: MSE = 16.3068, MAE = 13.853% and AAD = 3.722%.

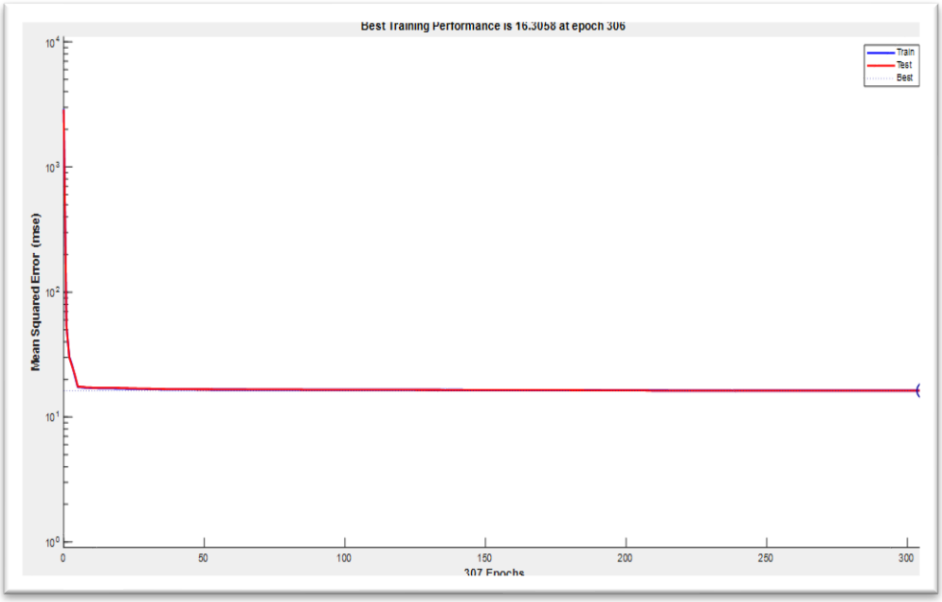

Figure 15. Optimum training performance of the BR network for 20 neurons (Xezonakis et al., 2024).

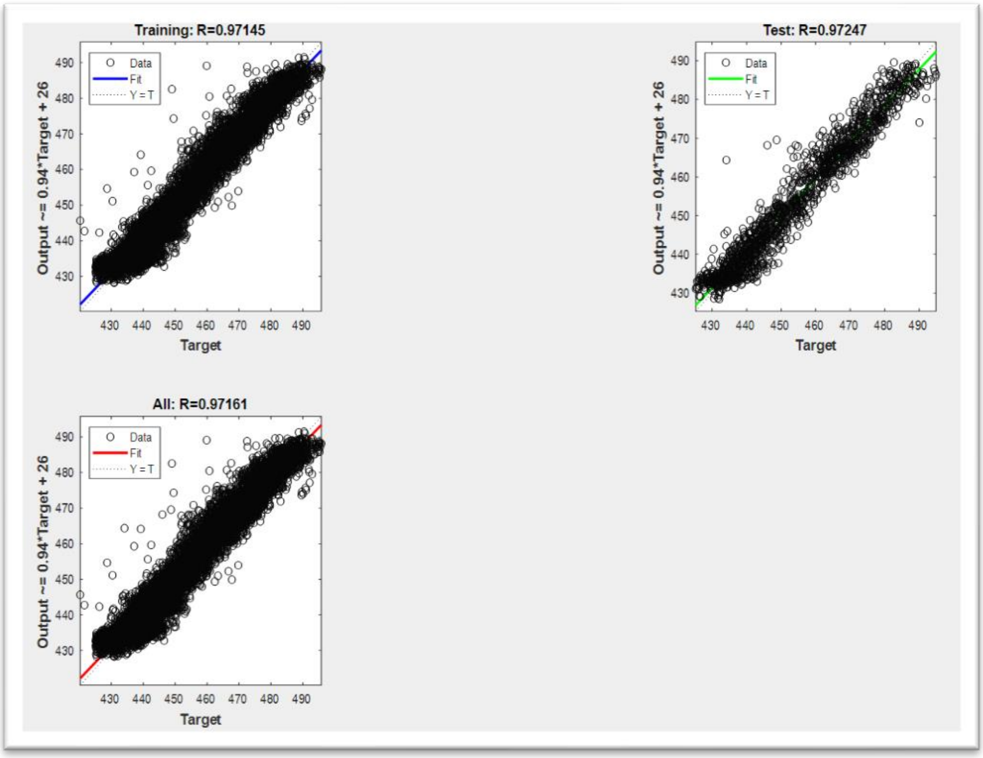

Figure 16. Regression analysis of the sample BR configuration (Xezonakis et al., 2024).

4.2. ANFIS model analysis

The next step after setting up the tools in the methodology section such as the FIS model with the novel Gaussian membership function (Gaussmf), is the optimization process presented in the flow chart diagram (**Figure 8**). **Figure 17**, after the initialization of the training process, depicts the convergence history of the combined use of the least square error (LSE) and steepest gradient (SG) algorithms for 100 epochs, resulting in RMSE converged values of 3.875% (training error) and 3.193% (checking error). **Figure 18** illustrates a sample of the stages associated with the FIS model after implementing the IF-THEN rules (81), including the optimal impact on the dependent characteristic (EP) intended for the generated FIS model with the assigned value of 1.510^3 W. The ANFIS model is designed to map the inputs (AT, V, AP and RH) to the EPO, using 81 fuzzy rules, which is illustrated in the Appendix. The use of Gaussian membership functions, combined with the linear Sugeno output makes this system effective for complex modelling, as well as nonlinear relationships between the input parameters and EP. The respective defuzzification process implements a weighted average, ensuring that the output is a crisp value by means of the fuzzy rule-based evaluation.

Figures 19–21 depict the three-dimensional responses showing the effect of the EP output parameter on the four input characteristics (AT, V AP, and RH), with optimal outcomes highlighted in yellow for the three different parameter combinations (AT, V, EP), (AT, RH, EP) and (AT, AP, EP). Therefore, the distinguished characteristics in terms of the robustness and validity of the ANFIS toolbox are confirmed. **Figure 22** visualizes the ANFIS performance outcomes of the determination coefficient R^2 for the entire dataset (0.9499), indicating a very strong fit between the actual and predicted data.

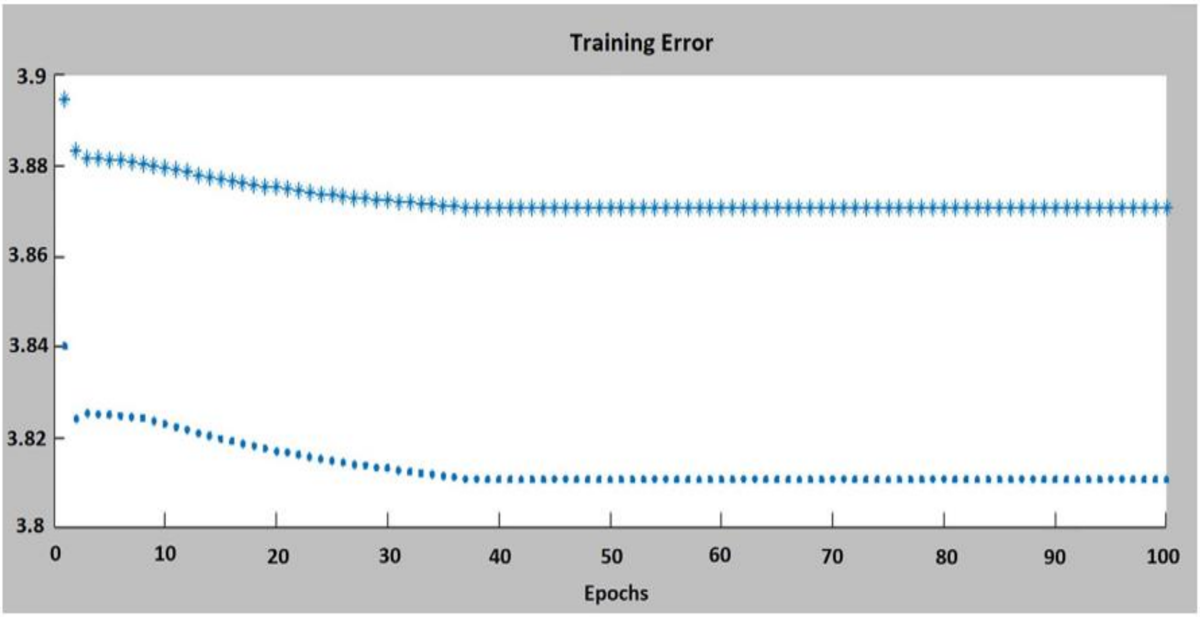

Figure 17. Convergence process of the RMSE of ANFIS for training error (*, upper curve) and the checking error (*, lower curve) outcome.

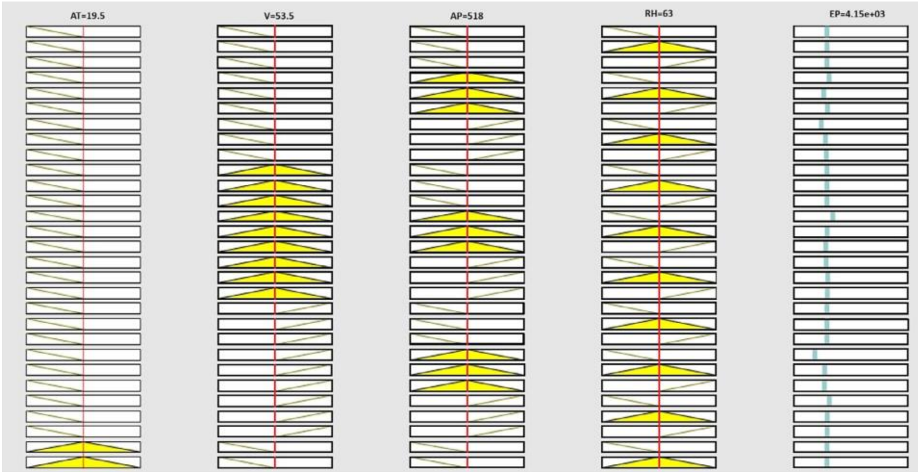

Figure 18. A sample example of the ANFIS editor's IF AND THEN rules.

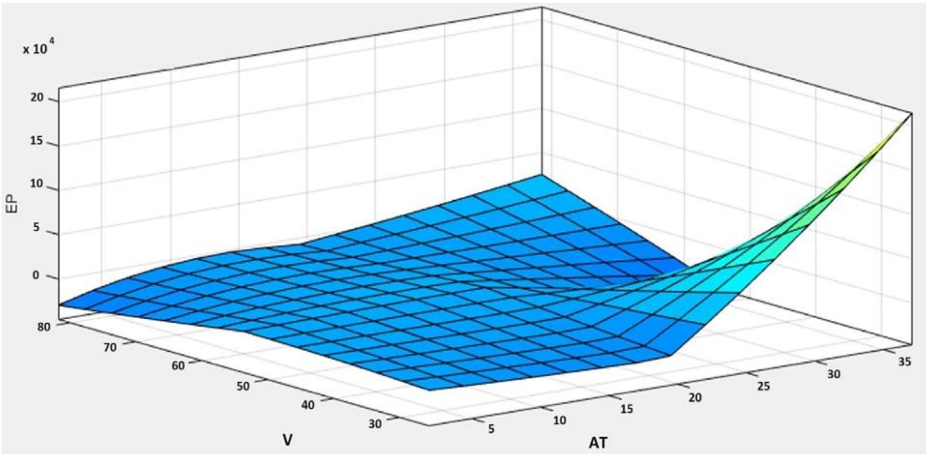

Figure 19. Predicted surface response of EP versus V and AT features.

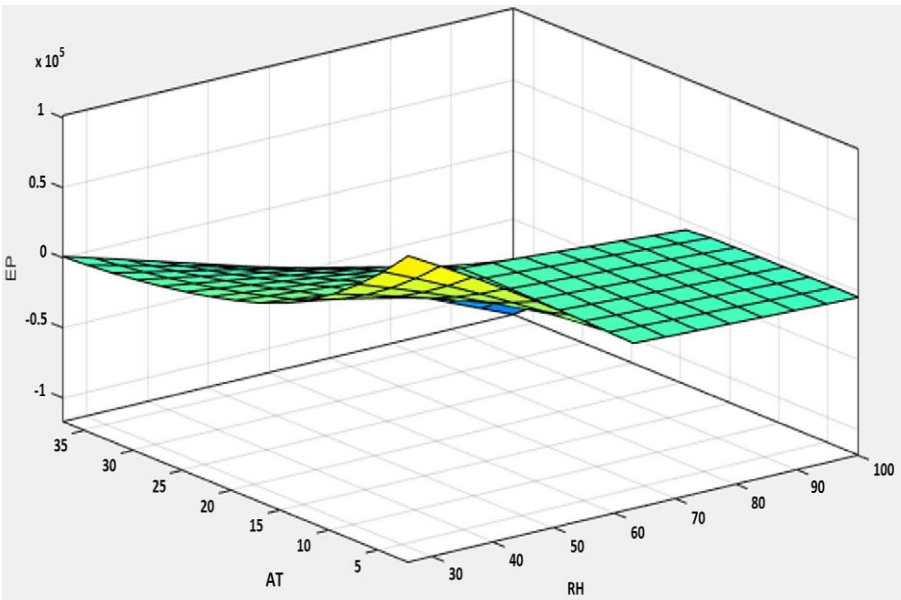

Figure 20. Predicted surface response of EP versus V and AT feature.

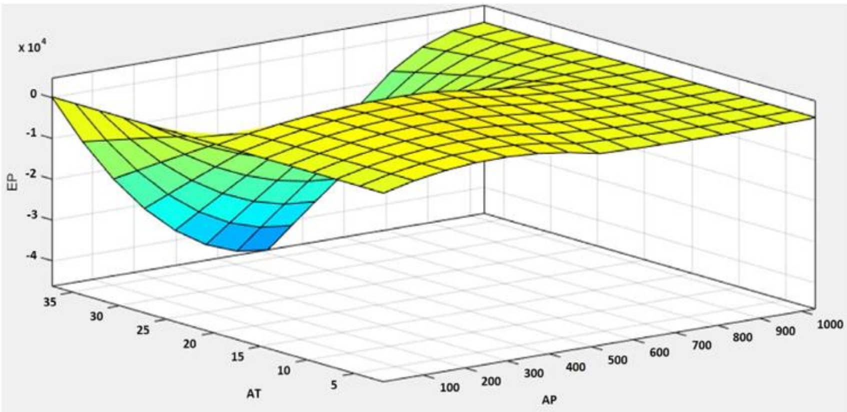

Figure 21. Predicted surface response of EP versus AT and AP features.

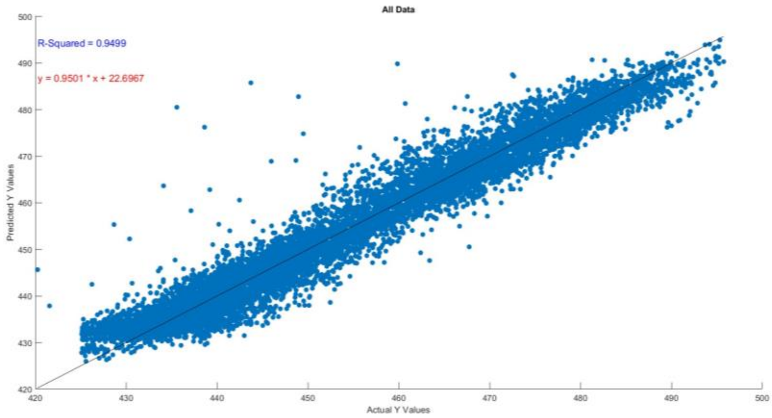

Figure 22. Regression analysis of the entire dataset.

Table 7 presents the efficacy prediction of the ANFIS structured network, via the training and checking error values of eight different membership functions (MF) according to the software's capabilities implemented for the four independent parameters. The adapted Gaussian membership functions (*Gaussmf*) for both the training and checking data are more efficient. In terms of computational cost benefits of the simulations, it took 2 minutes for ANFIS compared to the ANN simulation which took 6 minutes to complete.

Table 7. Training and checking errors for the eight different MF.

MFs	**Types of MFs**	**Training Error %**	**Checking Error %**
3-3-3-3	Gaussmf	3.875	3.193
3-3-3-3	Trapmf	3.965	3.212
3-3-3-3	Gbellmf	3.886	3.196
3-3-3-3	Trimf	3.893	3.246
3-3-3-3	Gauss2mf	3.879	3.254
3-3-3-3	Pimf	3.982	3.284
3-3-3-3	Dsigmf	3.883	3.295
3-3-3-3	Psigmf	3.878	3.310

This computational cost gain illustrates the advantages of ANFIS and refers to the training process of fitting the model to the entire dataset, optimizing the parameters and validating the model's performance. This process can be computationally intensive for large datasets. When it comes to inference it refers the time taken by the model to predict unseen data after training. Inference time is usually faster as it involves applying of the pre-trained model to predict the target feature (EPO) based on input parameters.

The computational environment impacts various factors affecting computational cost such as:

1) Hardware specifications (CPU/GPU): The type of processor affects computational time, with standard CPUs being sufficient for less complex models but taking longer.
2) Memory (RAM): The amount of available RAM is crucial for handling large datasets. Higher RAM allows for faster data loading and more efficient computations.

Regarding software specifications, MATLAB is the main programming language used, with libraries optimized for speed but requiring more computational power. The practical significance of time differences is related to the one-time or periodic cost in the training process.

Inference time comparisons relate to the time taken to generate predictions based on new input data. In this present study, on electric output power forecasting using four independent parameters (AT, V, AP and RH), the model's practical utility in real-time applications is affected. Scalability is crucial for large datasets and predictive tasks in power generation, especially when deployed in power grid platforms.

Table 8 compares the two main RMSE metrics scores (training and validation errors) between this study and relevant studies in energy field, showing the superiority of the current investigation. The RMSE values in study align with the literature.

Table 8. Comparison of various RMSE performance metrics from the energy sector.

RMSE Training Error %	RMSE Checking Error %	References
3.875	3.193	Current study
3.891	3.246	Pa and Kazemi (2022)
3.905	3.256	Rezakazemi, et al. (2016)
3.882	3.261	Bendary et al. (2021)

4.2.1. Critical/statistical analysis of ANN and ANFIS models for CCPP plant

Figure 23 describes the evaluated statistical indices adopted for the critical analysis to demonstrate the effectiveness and superiority of ANN and ANFIS models including R^2, RMSE, SEP, MAE, and AADE. As shown, the ANFIS exhibited better prediction fitting capabilities compared to the ANN model in predicting the CCPP located in Turkey and its environment.

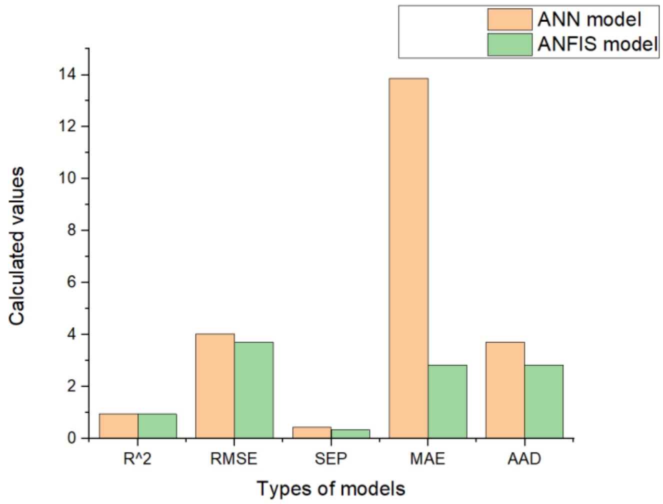

Figure 23. Relative statistical indices of ANN and ANFIS models.

The effective demonstration of the value of the adopted models (ANN, ANFIS) is crucial comparing their performance with simpler baseline models (Linear Regression). Therefore, a baseline model contributes in this direction to handle improvements brought by more complex models in terms of accuracy, efficiency and generalization. The proposed model (LR) assumes a linear relationship between the independent and the target variable, that the independent variables are highly not correlated with each other, as well as constant variance of the error terms. **Table 9** illustrates a comparison between the four metrics (MSE, MAE, AAD and R^2) with the linear regression analysis (LR) measures. Linear Regression performs exceptionally well compared to the main metrics. The MSE and MAE are extremely close to zero, and the R^2 is perfect (1.00), which suggests that the model explains 100% of the variance in the data. This also indicates that the model fits the data almost perfectly, affirming the superiority of LR compared to the metrics of both models (ANN, ANFIS).

From **Table 9**, the MAE for ANFIS (2.843) is significantly lower than for ANN (13.853), indicating that ANFIS predictions are on average, much closer to the actual values. The AAD for ANFIS (2.842) is also lower than for ANN (3.722), suggesting that the deviations in ANFIS predictions are more tightly clustered, reflecting a more consistent model performance. The impact on real world performance, shown by the large reduction in MAE (from 13.853 to 2.843), indicates that ANFIS predictions are nearly 80% more accurate than those of ANN.

The 0.5% improvement in the R^2 value between the ANN and ANFIS models may seem small at first glance, but in real-world applications, even minor increases in predictive accuracy translates to significant benefits, particularly in high-impact domains like power plant operations, grid management, and energy planning.

In practical terms this improvement means that the model makes better predictions of future power demand or output, reducing the error in forecasts. In grid management, accurate predictions can prevent over- or underestimating energy needs, thus improving resource allocation. However, accurate demand forecasting assists in energy production and waste reduction, leading to substantial cost savings for utilities, and better management of fuel usage and other operational costs. Furthermore, it might assist more in the performance gains in power plant management with financial impact attributes, translating into substantial financial savings in large-scale operations. The risk of under or over supplying energy is reduced accurately with direct cost implications.

Table 9. Comparison between (LR) with the key performance indicators.

Model	MSE	MAE	AAD	R^2
Linear Regression (LR)	0.1637×10^{-12}	0.2713×10^{-25}	0.01203×10^{-12}	1.0000
ANN	16.309	13.853	3.722	0.9451
ANFIS	13.853	2.843	2.842	0.9499

4.2.2. Error analysis

Error analysis is a critical part of evaluating the performance of machine learning models, allowing us to understand where and how models perform well and where they fail to capture underlying data patterns. In this analysis, we will focus on three models: Linear Regression, Artificial Neural Network (ANN), and Adaptive Neuro-Fuzzy Inference System (ANFIS), using metrics like residuals, residual vs. fitted values plots and distribution of residuals. The plot of residuals vs. fitted (predicted) values is a key diagnostic tool and ideally, residuals should be scattered randomly around zero, indicating that the model captures the relationship between the variables without systematic errors. Systematic patterns in this plot (e.g., curvature, funnel shapes) indicate non-linear problems. **Figure 24** illustrates, the residuals versus the fitted plots for the three models (LR, ANN, and ANFIS). Linear regression assumes a linear relationship of the residuals that should be scattered randomly around zero, thus the model is appropriate. ANN presents a large spread in residuals versus the fitted values, indicating that the model does not predict certain data points well. In the case of the network architecture, the number of training data is insufficient. ANFIS residuals illustrate systematic patterns, suggesting that the fuzzy rules or the membership functions should be adjusted, or because the complex patterns of the data cannot be captured. The distribution of residuals is typically visualized via histograms, giving insight into the error distribution. For a good quality model, the residuals should follow a normal distribution centred around zero. **Figure 25a–c** illustrates the distribution of residuals for each performance model. The residual distribution of the LR model is normal predicting the output appropriately without the presence of any outliers. In the ANN model, the residuals show a heavy skew, suggesting that neural networks struggle to capture the true data, indicating potential overfitting or improper training. The ANFIS model's residuals are not normally distributed, indicating the need for better-defined membership functions to capture the relationship between the input and output parameters. Therefore, the error analysis reveals that LR is more efficient for simple linear relationships but struggles with non-linear data. ANN is more powerful in capturing complex patterns but is sensitive to overfitting, and ANFIS requires careful configuration although it provides a balance between interpretability and performance. The following section will compare the comparison between the current study with identical ones using ANN and ANFIS based on using performance metrics (MSE, MAE, AAD and R^2).

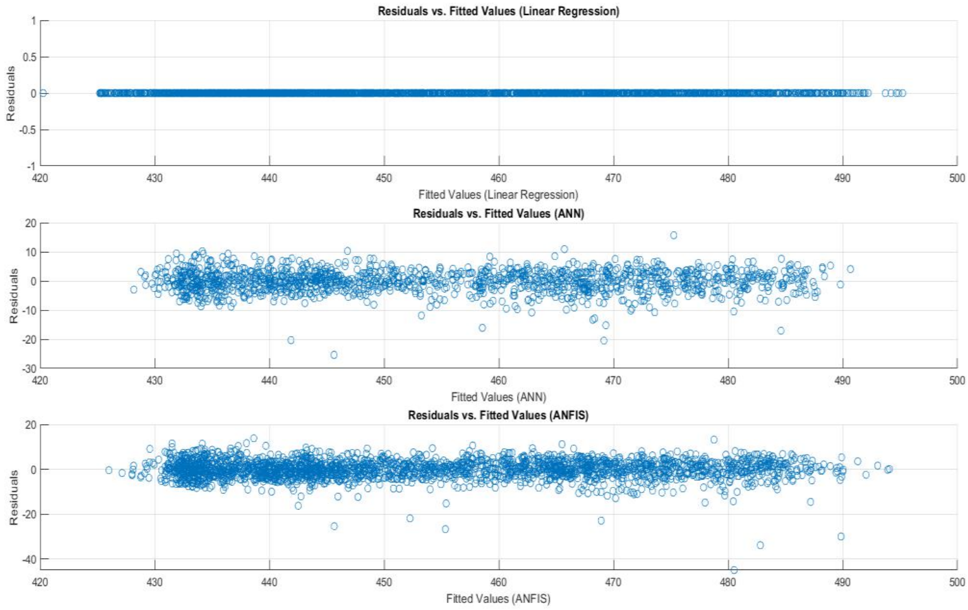

Figure 24. Residuals versus fitted values of LR, ANN and ANFIS.

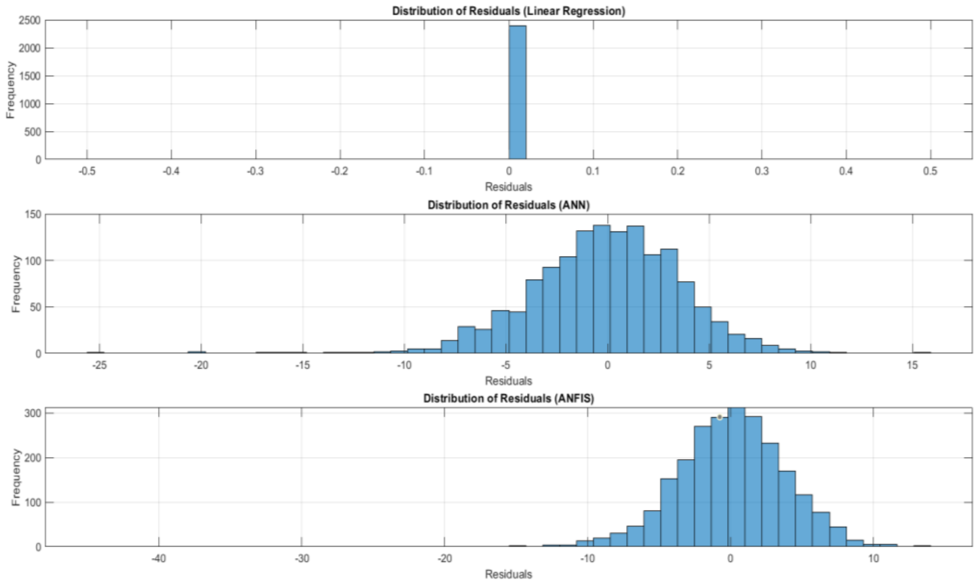

Figure 25. Distribution of residuals of LR, ANN and ANFIS models: **(a)** frequency vs. residuals for linear regression; **(b)** frequency vs. residuals for ANN; **(c)** frequency vs. residual.

4.2.3. Hybrid of ANN and ANFIS models for CCPP plant

Table 10 summarizes the hybrid models for CCPP plants and compares them with those in the literature (Ani and Agu 2022; Samuel et al., 2021; Tiwari et al., 2012). The differences between the present hybrid models and those in the literature can be attributed to plant conditions, and the topologies associated with the soft computing tool used.

Table 10. Statistical comparison ANN and ANFIS models for CCPP plant.

Refs.	Model tools	MSE	MAE	AAD	R^2
Present study	ANN	16.3068	13.853	3.722	0.9444
	ANFIS	13.853	2.843	2.842	0.9499
Samuel et al. (2021)	ANN	50.447	4.344	6.0529	0.8979
	ANFIS	9.3850	1.5711	1.9124	0.9786
Tiwari et al. (2012)	ANN	27.98	3.81	5.87	0.33
	ANFIS	30.96	3.99	5.87	0.41
Ani and Agu, 2022	ANN	15.133	8.8718	6.234	0.8862
	ANFIS	5.1203	5.2737	4.581	0.9349

5. Conclusion

In this paper, the use of neural networks (ANN) and adaptive neuro-fuzzy inference system (ANFIS) in a combined cycle power plant has been investigated using the actual dataset (9568) from an authenticated test case in Turkey. The independent variables are ambient temperature, exhaust vacuum, ambient pressure and relative humidity, while the dependent parameter is electric output power. The performance of ANN and ANFIS models is assessed using the entire dataset for testing both models. A comparison between ANN and ANFIS methodologies specifies better validation and performance attributes of the ANFIS model.

The conclusions of the current study are summarized as follows:

1) The computational cost of using the soft computing technique (ANFIS) is improved by 4 min compared to neural networks (ANN). ANFIS capabilities ensure robust and optimal prediction of electric output power considering the entire actual dataset, which is beneficial for controlling CCPP operation.
2) ANFIS is more efficient compared to similar energy studies, as verified by the improved RMSE performance metric in **Table 5**.
3) The preference for ANFIS over ANN in the present and similar studies, based on key performance indicators (MSE, R^2, AAD, and MAE) in **Table 6**, validates its optimal performance with an exemption of the opposite outcome, presented in (Tiwari, Bajpai and Dewangan (2012). This study could serve as a starting point for leveraging various machine learning techniques such as support vector machine (SVM) and combined hybrid models for alternative and optimized predictions of CCPP and other types of combined power plants worldwide.
4) Error analysis revealed the superiority of the linear regression (LR) model, due to its distinguished features compared to ANN and ANFIS.
5) Although a massive dataset is adopted the network's overall performance resulted in accurate and reliable outcomes with a substantial significant impact on the operation of the CCPP.

Author contributions: Conceptualization, VX and CCE; methodology, VX, ODS and CCE; software, VX; validation, VX, ODS and CCE; formal analysis, VX; investigation, VX, ODS and CCE; resources, VX; data curation, VX; writing—original draft preparation, VX; writing—review and editing, ODS and CCE;

visualization, VX, ODS and CCE; supervision, ODS and CCE; project administration, CCE. All authors have read and agreed to the published version of the manuscript.

Acknowledgments: The authors would like to thank Ahmad Saad for copyediting and proofreading and Svetla Meijer for advice on the editing process of the figures.

Conflict of interest: The authors declare no conflict of interest.

References

Abuayyash K, Alsamamra H, Teir MA, et al. (2024). Wind speed prediction in Jerusalem using machine learning algorithms: A case study of using ANFIS and KNNR. American Journal of Modern Energy, 10(2), 25-37. https://doi.org/10.11648/j.asme.20241002.12

Akdemir B. (2016). Prediction of hourly generated electric power using artificial neural network for combined cycle power plant. International Journal of Electrical Energy, 4(2), 91-95. doi: 10.18178/ijoee.4.2.91-95

Amirkhani S, Nasirivatan S, Kasaeian, et al. (2015). ANN and ANFIS models to predict the performance of solar chimney power plants. Renewable Energy, 83, 597-607. https://doi.org/10.1016/j.renene.2015.04.072

Ani KA, Agu CM. (2022). Predictive comarison and assessment of ANFIS and ANN, as effective efficient tools in modeling degradation of total petroleum hydrocarbon (TPH). Cleaner Waste Systems, 6, 100107. https://doi.org/10.1016/j.clwas.2022.100052

Bandić L, Hasičić M, Kevrić J. (2020). Prediction of power output for combined cycle power plant using random decision tree algorithms and ANFIS. In: Avdaković, S., Mujčić, A., Mujezinović, A., Uzunović, T., Volić, I. (eds). Advanced Technologies, Systems, and Applications IV, Proceedings of the International Symposium on Innovative and Interdisciplinary Applications of Advanced Technologies (IAT 2019). Lecture Notes in Networks and Systems, 83, 406-416. Springer, Cham. https://doi.org/10.1007/978-3-030-24986-1_32

Bendary AF, Abdelaziz AV, Ismail MM, et al. (2021). Proposed ANFIS-based approach for fault tracking detection, clearing, and rearrangement of photovoltaic systems. Sensors, 21(7), 2-16. https://doi.org/10.3390/s21072269

Botsaris PN, Ntantis EL, Topalis C. (2014). Case study of solar energy use for the thermal needs of ferrous casting industrial unit. In: Proceedings of the 10th National Conference on Renewable Energy Resources, November 26-28, 2014, Thessaloniki, Greece

Chen J, Li H, Sheng D, Li W. (2015). A hybrid data-driven modeling method on sensor condition monitoring and fault diagnosis for power plant. International Journal of Electrical Power & Energy Systems, 71, 274–284. https://doi.org/10.1016/j.ijepes.2015.03.012

Ekki BB, Aksov UT. (2011). Prediction of building energy needs in early stage of design by using ANFIS. Journal of Engineering Systems Applications, 38(5), 5352-5358. https://doi.org/10.1016/j.eswa.2010.10.021

Elfaki E, Hassan A. (2018). Prediction of the electrical output power of combined cycle power plant using regression ANN model. International Journal of Computer Science and Control Engineering, 6(2), 9–2. doi: 0.5281/zenodo.1285164

El-Hadjik AA. (1990). The impact of atmospheric conditions on gas turbine performance. Journal of Engineering Gas Turbines Power. 112(4), 590-596. https://doi.org/10.1115/1.2906210

Ersayin E, Ozgener L. (2015). Performance analysis of combined cycle power plants: a case study. Renewable and Sustainable Energy Reviews, 43, 823- 842. https://doi.org/10.1016/j.rser.2014.11.082

Faahmi ATWK, Kashyzadeh KR, Ghorbani S. (2022). A comprehensive review on mechanical failures cause vibration in the gas turbine of combined cycle power plants. Engineering Failure Analysis, 134, 106094. https://doi.org/10.1016/j.engfailanal.2022.106094

Gas Turbine Combined Cycle Generation. (2020). Available online: File:Gas Turbine Combined Cycle Generation 01.svg - Wikimedia Commons (assessed 15th July, 2023).

Gill J, Singh J. (2017). Performance analysis of vapour compression refrigeration system Using an adaptive neuro-fuzzy inference system. International Journal of Refrigeration, 82, 436-446. https://doi.org/10.1016/j.ijrefrig.2017.06.019

Guerra MIS, De Araujo FMU, De Carvalho Neto JT, et.al. (2024). Survey on adaptive neural fuzzy inference system (ANFIS) architecture applied to photovoltaic systems. Energy Systems, 15, 505-541. https://doi.org/10.1007/s12667-022-00513-8.

Guleryuz D. (2021). Determination of industrial energy demand in Turkey using MLR, ANFIS and PSO-ANFIS. Journal of Artificial Intelligence Systems, 3, 16-34. https://doi.org/10.33969/AIS.2021.31002

Hoang T, Pawluskiewicz DK. (2016). The efficiency analysis of different combined cycle power plants based on the impact of selected parameters. International Journal of Smart Grid Clean Energy, 5 (2), 77-85. doi: 10.12720/sgce.5.2.77-85

Hundi P, Shahsavari R. (2020). Comparative studies among machine learning methods for performance estimation and health monitoring of thermal power plants. Applied Energy, 265,114775. https://doi.org/10.1016/j.apenergy.2020.114775

Ibrahim AM, IA, Lawan SM, Abdukabir R, et.al. (2024). Solar radiation prediction using an improved adaptive neuro-fuzzy inference system (ANFIS) optimization ensemble. In: Constantin Voloşencu, C (ed). Adaptive Neuro-Fuzzy Inference System as a Universal Estimator. IntechOpen, 1-22. http://dx.doi.org/10.5772/interchopen.1003891

Ibrahim TK, Mohammed MK, Awad OI, et.al. (2017). The optimum performance of the combined cycle power plant: A comprehensive review. Renewable and Sustainable Energy Reviews, 79, 459-474. https://doi.org/10.1016/j.rser.2017.05.060

Kaur S, Kaur T, Khanna R.. (2021). Adaptive neuro-fuzzy inference system-based output power controller in grid-connected photovoltaic systems. Energy Sources ,Part A: Recovery, Utilization, and Environmental Effects, 1890860. https://doi.org/10.1080/15567036.2021.1890860.

Krishnan S, Wu E. (2019). AlphaClean: Automatic generation of data cleaning pipelines. arXIv preprint arXiv: 1904.11827. https://doi.org/10.48550/arXiv.1904.11827

Lorencin I, Mrzljak V. Car Z. (2019). Genetic algorithm approach to the design of multi-layer perceptron for combined cycle power plant electrical power output estimation. Energies, 12(22),4352. https://doi.org/10.3390/en12224352.

Mathworks, ANFIS NeuroFuzzy. MATLAB software. 2022. https://uk.mathworks.com/help/fuzzly/anfis.html

Melin P, Castillo O, (2005). Intelligent control of a stepping motor drive using an adaptive neuro-fuzzy logic inference system, Information Sciences, 170 (2-4), 133-151. https://doi.org/10.1016/j.ins.2004.02.015

Melin P, Soto J, Castillo OC, et al. (2012). A new approach for time series prediction using ensembles of ANFIS models. Applied Energy, 88(11), 3494 – 3506. https://doi.org/10.1016/j.eswa.2011.09.040

Mohandes M, Rehman S, Rahman SM. (2011). Estimation of wind speed profile using adaptive neuro-fuzzy inference system (ANFIS). Applied Energy, 88 (11), 4024 – 4032. https://doi.org/10.1016/j.apenergy.2011.04.015.

Nabipour N, Mosavi A, Hajnal E, et al. (2019). Modelling Climate change impact on wind power resources adaptive neuro-fuzzy inference system. Engineering Applications of Computational Fluid Mechanics, 14(1), 491-506. https://doi.org/10.1080/19942060.2020.1722241

Ntantis EL. (2009). Capability expansion of non-linear gas path analysis. PhD thesis, Cranfield University, UK.

Ntantis EL, Li YG. (2009). The impact of measurement noise on gas turbine GPA diagnostics. In: Proceedings of the Sixth International Conference on Condition Monitoring & Machinery Failure Prevention Technologies; June 2009, Dublin, Ireland.

Ntantis EL, Li YG. (2013) The impact of measurement noise in GPA diagnostic analysis of a gas turbine engine. International Journal of Turbo Jet-Engines, 30(4), 401-408. https://doi.org/10.1515/tjj-2013-0024

Ntantis EL, Botsaris PN. (2015). Diagnostic methods for an aircraft engine performance. Journal of Engineering Science and Technology Review, 8(4), 64-72. doi: 10.25103/jestr.084.10

Ntantis EL, Botsaris PN. (2016). The impact of gas turbine component leakage fault on GPA performance diagnostics. Journal of Engineering Science and Technology Review, 9(1), 116-123. doi: 10.25103/jestr.091.18

Nugraha YT, Simnjuntak PM, Irwanto M. (2024). Analysis of forecast of renewable development in North Sumatra using ANFIS. Jurnal Media Elektro, XIII(1), 27-36. doi: 10.35508/jme.v13i1.15310

Olatomiwa L, Mekhilef S, Shamsirband S, et al. (2015). Potential of support vector regression for solar radiation prediction in Nigeria. Natural Hazards, 77, 1055-1068. https://doi.org/10.1007/s11069-015-1641-x

Pa M, Kazemi A. (2022). ANFIS–based prediction of power generation for combined cycle power plant. ArXIV: 2210.09011. https://doi.org/10.48550/arXiv.2210.09011

Panella M, Gallo AS. (2005). An input-output clustering approach synthesis of ANFIS Networks. IEEE Transactions of Fuzzy Systems. 13(1), 69-81. doi: 10.1109/TFUZZ.2004.839659

Pawar N. (2022). Development of photovoltaic power generation model using ANFIS. International Journal of Engineering Management Resources, 12(4), 199-208. https://doi.org/10.31033/ijemr.12.4.26

Petkovic D, Cojbasic A, Nikolic V. (2013). Adaptive neuro-fuzzy approach for wind turbine power coefficient. Renewable and Sustainable Energy Reviews, 28, 191-195. https://doi.org/10.1016/j.rser.2013.07.049.

Quamming Y., Menngshuo W., Hugo J.E., Isabelle G., and Yang Y., (2018)., ‘‘ Talking human out of learning applications: A survey of automated learning, ‘‘.

Qu Z, Xu J, Chi R, et al. (2021). Prediction of electricity generation from a combined cycle plant based on a stacking ensemble and its hyperparameter optimization with a grid-search method. Energy, 227, 120309. https://doi.org/10.1016/j.energy.2021.120309

Revalthy SR, Kirubarakan V, Rajeshwaran,M, et.al. (2022). Design and analysis of ANFIS-based MPPT method for solar photovoltaic applications. International Journal of Photoenergy, 2022:9625564. https://doi.org/10.1155/2022/9625564

Rezakazemi M, Mosavi A, Shirizian,S. (2016). ANFIS pattern for molecular membranes Separation optimization. Journal of Molecular Liquids, 274, 470-476. https://doi.org/10.1016/j.molliq.2018.11.017

Sabouhi H, Abbaspour A, Fotuhi-Firuzabad M, Dehghanian P. (2016). Reliability modeling and availability analysis of combined cycle power plants. International Journal of Electrical Power & Energy Systems, 79, 108–119. https://doi.org/10.1016/j.ijepes.2016.01.007

Saleel CA .(2021). Forecasting the enerhu output from a combined cycle power plant using deep learning models. Case Studies in Thermal Engineering, 28, 101693. https://doi.org/10.1016/j.csite.2021.101693.

Salisu S, Mustafa MW, Mustapha M. (2018). A wavelet-based solar radiation prediction in Nigeria using the adaptive neuro-fuzzy approach. Indonesian Journal of Electric Engineering Computer Science, 12(3), 907-915. http://doi.org/10.11591/ijeecs.v12.i3.pp907-915

Samuel OD, Okwu MO, Tartibu LK, et al. (2021). Modelling of Nicotiana Tabacum L. oil. Biodiesel production: Comparison of ANN and ANFIS. Frontier in Energy Research, 8, 612165. https://doi.org/10.3389/fenrg.2020.612165

Savrum MM, Inci M. (2021). Adaptive neuro-fuzzy inference system combined with genetic algorithm to improve power extraction capability in fuel cell applications. Journal of Cleaner Production, 299,1269944. https://doi.org/10.1016/j.jclepro.2021.126944

Santamouris M., (2015). On the energy impact of urban heat island and global warmimg on buildings. Joirnal of Energy and Buildings, Vol.82, SP. 100, https://doi.org/10.1016/j.enbuild.2014.07.022.

Shanbedi M., Amiri A, Rashidi S, et al. (2014). Thermal performance Prediction of two-phase closed thermosiphon using adaptive neuro-fuzzy inference system. Heat Transfer Engineering, 36(3), 315-324. https://doi.org/10.1080/01457632.2014.916161

Shukla AK, Singh O. (2014). Effect of compressor inlet temperature and relative humidity on gas tutbine cycle performance. International Journal of Scientific & Engineering Research, 5 (5), 664-670.

Siddiqui R, Anwar H, Ullah F, et al. (2021). Power prediction of combined cycle power plant (CCPP) using machine learning algorithm-based aradigm. Wireless Communications and Mobile Computing, 2021, 9966395. https://doi.or(2015g/10.1155/2021/9966395

Tay KG, Muwafaq H, Choy CC. (2019). Electricity consumption forecasting using adaptive neuro-fuzzy inference system (ANFIS). Universal Journal of Electrical Electronic Engineering, 6(58), 37-48. Doi:10.13189/ujeee.2019.061606

Tiwari MK., Bajpai S, Dewangan LK. (2012). Prediction of industrial solid waste with ANFIS model and its comparison with ANN model – a case study of Durg- Bhilai Twin City India. International Journal of Engineering and Innovative Technology, 2(6)

UCI Machine Learning Repository: Combined Cycle Power Plant Data Set. (2012). Available online: https://archive.ics.uci.edu/ml/datasets/combined+cycle+power+plant (accessed on 20th July, 2023).

Vimala C, Jayanthi N, Vijayalakshmi P, et al. (2024). Enhancing predictive modelling in power plants: A comparative analysis of ANFIS and ANN for electrical energy output estimation. In: Proceedings of the International Conference on Inventive Computation Technologies (ICITCT), 2024, Nepal. 10.1109/ICICT60155.2024.10544753

Wang X, Sun X, Chu S, et al. (2023). Improved butterfly algorithm applied to prediction of combined cycle power plant. Mathematics and Computers in Simulation, 24,337-353. https://doi.org/10.1016/j.matcom.2022.08.009

Xezonakis V. (2024). Enhancing electric energy production in Combined Gas Plant (COGAS) and Combined Cycle Power Plant (CCPP) through artificial neural networks and adaptive neuro-fuzzy logic inference system. PhD thesis, University of South Africa. Deparment of Mechanical, Bioresources and Biomedical Engineering.

Xezonakis V, Samuel OD, Enweremadu CC. (2024). Modelling and Output Power Estimation of a Combined Gas Plant and a combined cycle plant using an Artificial Neural Network approach. Journal of Engineering, 2024, 5540010. https://doi.org/10.1155/2024/5540010

Zaaoumi A, Bah A, Ciocan M, et al. (2021). Estimation of tbe energy production of a parabolic trough solar power plant using analytical and artificial neural networks models. Renewable Energy, 170, 620-628. https://doi.org/10.1016/j.renene.2021.01.129

Zamani HA, Taghanaki SR, Karimi M, Anabloo M, Dandashi A. (2015). Implementing ANFIS for prediction of reservoir oil gas. Journal of Natural Gas Science and Engineering, 25, 325-334. https://doi.org/10.1016/j.jngse.2015.04.008

Zayed ME, Zhao J, Li W, et al. (2021). A hybrid adaptive neuro-fuzzy inference system integrated with equilibrium optimizer for predicting the energetic performance of solar dish collector. Energy, 235,121289. https://doi.org/10.1016/j.energy.2021.121289

Zhao Y, Foong LK. (2022). Predicting electrical power output for a combined cycle power plant using a novel artificial neural network optimized by electrostatic discharge algorithm. Measurement, 198, 11405. https://doi.org/10.1016/j.measurement.2022.111405

Zhang Y, Dong ZY, Kong W, Meng K. (2020). A composite anomaly detection system for data-driven power plant condition monitoring. IEEE Transactions on Industrial Informatics, 16 (7), 4390–4402. doi: 10.1109/TII.2019.2945366

Appendix

A sample rule of the model is the following:

1) Rule 1: If AT is low, V is low, AP is low and RH is low, then the output power (EP) is calculated by 81 MF1 linear functions defined as:

$$\mathrm{MF1} = \mathrm{linear}, [a_1, a_2, a_3, a_4, a_5] \tag{6}$$

where: a_1, a_2, a_3, a_4 are coefficients applied to the input parameters AT, V, AP and RH and a_5 is a constant. These functions use the Sugeno model to generate crisp output on a fuzzy input combination.

2) Rule 2: If AT is low, V is low, AP is low and RH is medium, then the respective power output is determined identically by MF2, analogous to MF1.

Therefore, this pattern continues for all combinations of the input membership functions across the four variables. ANFIS model is designed to map the inputs (AT, V, AP and RH) to the (EP), using 81 fuzzy rules. The use of Gaussian membership functions, combined with the linear Sugeno output makes this system effective for complex modelling, as well as nonlinear relationships between the input parameters and EP. The respective defuzzification process implements a weighted average, ensuring that the output is a crisp value by means of the fuzzy rule-based evaluation.