Republic of Iraq

Ministry of Higher Education and Scientific Research

University of Misan/Collage of Engineering

Department of Electrical Engineering

PREDICTION OF PHOTOVOLTAIC POWER GENERATION BASED ON ARTIFICIAL INTELLIGENCE

By

Shahad Mohammed Radhi

B.Sc. Electrical engineering, 2021

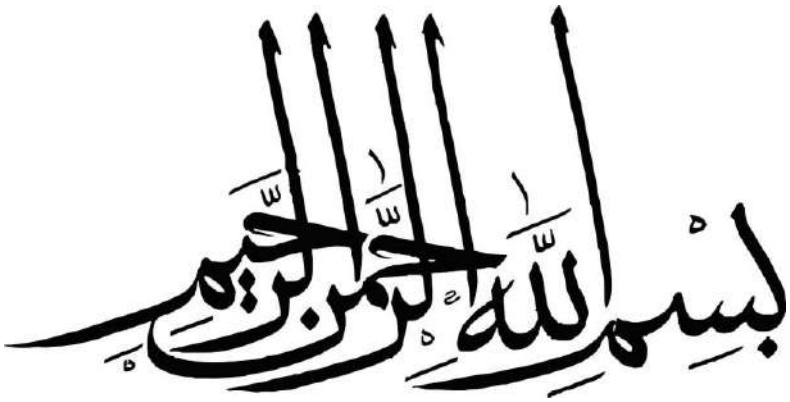
A THESIS

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Electrical Engineering

Supervisor Name: Assist. prof. Dr. **Sadeq Duair Aneed**



"وَقُلْ رَبِّ زِدْنِي عِلْمًا"

سورة طه آية:١١٤

Statement of Authorship

This thesis was completed as part of the MSc. (**Electrical Engineering**) at **College of Engineering -University of Misan-**. This is my own unaided work. Where the work of others has been used or drawn on then it has been fully attributed to the relevant source.

Signature:

Name: Shahad Mohammed Radhi

Date: / /2024

Certification of the Examining Committee

We certify that we have read the thesis entitled" **Prediction Of Photovoltaic Power Generation Based On Artificial Intelligence"** which is being submitted by **Shahad Mohammed Radhi** and as the Examining Committee, examined the student in its content. In our opinion, the thesis is adequate for award of the degree of Master of Science in Electrical Engineering.

Signature:

Name: **Assist. Prof. Dr. Sadeq D. Aneed**

(Supervisor)

Date: / /2024

Signature:

Name: **Assist Prof. Dr. Khalid M. Abdul-Hassan**

(Member)

Date: / /2024

Signature:

Name: **Assist Prof. Dr. Mohammed Kh. AL-NUSSAIRI**

(Member)

Date: / /2024

Signature:

Name: **Prof. Dr. Adel Manaa Dakhil**

(Chairman)

Date: / /2024

Approval of the College of Engineering

Signature:

Name: **Prof. Dr. Abbas Oda Dawood**

Dean of the College of Engineering

Date: / /2024

Supervisor Certification

I certify that this thesis which is entitled" **Prediction of Photovoltaic Power Generation Based On Artificial Intelligence"** which is being submitted by "**Shahad Mohammed Radhi**" was prepared under my supervision at College of Engineering, University of Misan, as a partial fulfilment of the requirements for the degree of Master of Science in Electrical Engineering.

Signature:

Supervisor: **Assist Prof. Dr. Sadeq Duair Aneed**

Date: / / 2024

I forward this thesis for debate by examining committee

Signature:

Name: **Assist Prof. Dr. Mohammed Kh. AL-NUSSAIRI**

Head of Electrical Engineering Department

Date: / / 2024

DEDICATION

Let history write in its pages that I drew the dream on the top of the mountains and walked on the road repeating there is no good in a dream without my actions...

With all faith and humility, I dedicate this thesis to our awaited Imam, the Mahdi (May God hasten his reappearance).

To the one whose name I carry with pride, to the one who harvested the thorns from my path to pave the way for me to knowledge… (my dear father).

To the one who was my source of strength and inspiration, to the owner of the sincere supplication and the compassionate heart to the one who sacrificed her time and effort to secure a better future for me...(to my mother).

To those who supported me with all love and removed all the troubles from my path, paving the way for me, planting confidence and determination within me...to (my brothers).

To the great man who was the source of my strength and encouragement, who was a motivation for me in every step...to (my dear husband).

Finally, to that soul that overcame all difficulties and endured many challenges, but did not give up, to my strength that appeared in moments of weakness, and to my determination that always pushed me to move forward despite all obstacles, I dedicate this well-deserved success to myself.

ACKNOWLEDGEMENTS

Firstly, I would like to express my gratitude to Allah Almighty for His blessings that enabled me to complete my studies. I would also like to thank Assistant Professor Dr. Sadeq Duair Aneed for his time, efforts, and patience in following up on my presentation, and teaching me how to deal with research challenges with a spirit of perseverance and critical thinking. His sound advice and precise guidance motivated me to achieve this accomplishment. I am proud to be one of his students.

I extend my thanks to the Dean of the College of Engineering, the head of the Department of Electrical Engineering, and the faculty members of the Department of Electrical Engineering for their cooperation. I am also appreciative of Engineers Youssef Sami and Saif Al-Din Ali for their assistance in collecting practical data, and Dr. Qudama Al-Yasiri for his support with long-term weather data.

Additionally, I would like to thank Engineer Mustafa Sadiq and Mohammed Saeed for their help, and I express my deep appreciation to my family for their continuous support.

Abstract

A Photovoltaic (PV) energy is considered one of the most inexhaustible renewable energy sources in the world due to its abundance and sustainability, as well as its low operational costs. However, its power production depends on whether factors such as irradiance, temperature, etc. Therefore, a PV power forecasting prediction is a crucial stage to utilize the stability, quality and management of a hybrid power grid. In this thesis, the PV forecasting prediction model based on theoretical data and real data is designed using various machine learning techniques. The theoretical data are obtained from the website of (climate one building), while the real data are collected from the PV experimental prototype installed at the engineering college of Misan University in Iraq. To enhance the PV forecasting prediction model, an artificial neural network (ANN) technique based on the gray wolf optimization (GWO) and genetic algorithm (GA) as learning methods are utilized. Then, the Python approach is used to design this PV forecasting based on five fitness functions, R2, MAE, RMSE, MSE, and RE. Finally, the data are analyzed and tested over short-term and medium-term time horizons to ensure model performance and forecast accuracy, thus improving PV power production across different times and weather conditions. The results indicate that the ANN model based on the GA algorithm captures the PV power generation pattern with higher accuracy across various weather conditions compared to traditional ANN and ANN-GWO prediction models. This is evident from the higher Pearson correlation coefficient (R^2) values achieved during different months in the theoretical data for medium-term photovoltaic power prediction. Additionally, the GA-based ANN model achieved higher R^2 values under sunny, cloudy, and rainy conditions in the experimental data for short-term photovoltaic power prediction.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLES

Symbol	Definition	Unit
B_i	The Bias of The ANN Model	
C	The Capacitor of The Boost Converter	(F)
D	The Duty Cycle	-
Dp	The Change Power	(Watts)
Dv	The Change Voltage	(Volt)
G	The Solar Irradiation	(Watts/M^2)
I_D	The Current of PV Diode	(A)
I_o	The Saturation Current of The PV Diode	(A)
Io	The Output Current	(A)
I_{ph}	The Current Generator from The Solar Cell	(A)
I_{PV}	The Output Current of The PV Cell	(A)
Is	Is the Input Current	(A)
I_{SC}	The Short Circuit Current of PV Cell	(A)
I_{sh}	The Shunt Circuit Current of PV Cell	(A)
Ki	The Temperature Coefficient	-
L_{min}	Inductor of The Boost Converter	(H)
N	The PV Diode Factor	-
Pa	The Actual Power	(Watts)
Pf	The Predicted Power	(Watts)
Po	The Output Power	(Watts)
P_{prev}	Previous Power Values	(Watts)
Ps	The Input Power	(Watts)
R_S	The Series Resistance of PV Cell	(Ohm)
R_{sh}	The Shunt Resistance of PV Cell	(Ohm)
T	The Ambient Temperature	(C)
T	Iterations	-
Ts	The Switching Period	(Second)
V_d	The Voltage Across the PV Diode	(Volt)
V_o	The Output Voltage	(Volt)
V_{prev}	Previous Voltage Values	(Volt)
V_{PV}	The Voltage Output of PV Cell	(Volt)
Vs	The Input Voltage	(Volt)
W_{ij}	The Connection Weights of ANN Model	
X_i	The Input of ANN Model	
Y_i	The Real Output	(Watts)

$\vec{r}_1$ And $\vec{r}_2$	The Random Vectors Within the Range of [0, 1]	-
$\vec{x}_a$	The Best Search Agent	-
$\vec{x}_p$	The Position Vector of The Prey	-
$\vec{x}_\beta$	The Second-Best Search Agent	-
$\vec{x}_\delta$	The Third Best Search Agent	-
$\overrightarrow{A}$ And $\vec{C}$	The Coefficient Vectors	-
$\vec{x}$	The Position Vector of a Grey Wolf	-
$\Delta V_{\mathbf{Out}}$	The Ripple of The Output Voltage	(Volt)

LIST OF ABBREVIATIONS

Full From	Abbreviation
Adaptive Neuro-Fuzzy Inference System	ANFLS
Ambient Temperature	AT
Ant Colony Optimization	ACO
Artificial Neural Networks	ANN
Bipolar Junction Transistor	BJT
Coefficient of Determination	COD
Convolutional Autoencoder	CAE
Convolutional Neural Network	CNN
Direct Horizontal Irradiation	DHI
Direct Normal Irradiation	DNI
Elastic Net	EN
Empirical Modal Analysis	EMA
Empirical Modal Decomposition	EMD
Extreme Learning Machine	ELM
Feed-Forward Neural Network	FFNN
Gated Recurrent Unit	GRU
Gaussian Process Regression	GPR
Generative Adversarial Network	GAN
Genetic Algorithm	GA
Global Horizontal Irradiation	GHI
Gray Wolf Optimization	GWO
Improved Whale Optimization Algorithm	IMWOA
Levenberg Marquardt Back Propagation	LM
Linear Regression Model	LRM
Long Short-Term Memory	LSTM
Mean Absolute Error	MAE
Mean Absolute Percentage Error	MAPE
Mean Absolute Relative Normalized Error	MARNE
Mean Bias Error	MBE
Mean Square Error	MSE

Multilayer Perceptron Neural Network	MLPNN
Multiple Regression	MR
Neural Network Ensemble	NNE
Pearson correlation coefficient	R^2
Practical Swarm Optimization	PSO
Principal Component Analysis	PCA
Radial Basis Function Neural Network	RBFNN
Random Forest	RF
Recurrent Neural Network	RNN
Relative Humidity	RH
Resilient Back Propagation	RBP
Root Mean Square Error	RMSE
Salp Swarm Algorithm	SSA
Scaled Conjugate Gradient	SCG
Sine Cosine Algorithm	SCA
Support Vector Regression	SVR
Symbolic Regression	SR
Virtual Mode Decomposition	VMD
Wavelet Decomposition	WD
Wavelet Packet Decomposition	WPD
Wind Speed	WS

LIST OF PUBLICATIONS

Two scientific papers from this research were published in journals and one paper was accepted in the Proceedings of the 6th International Scientific Conference on Engineering Sciences and Advanced Technologies in the American Institute of Physics (AIP)(ISSN: 0094-243X, 1551-7616) Hilla College- Babylon- Iraq / 28-29 November 2024 Scopus.

1. S. M. Radhi, S. D. Al-majidi, M. F. Abbod, and H. S. Al-raweshidy, “Predicting Solar Power Generation Utilized in Iraq Power Grid Using Neural Network,” ***Misan Journal of Engineering Sciences***, 3(1), 2024:38-62. **(published)**

2. S. M. Radhi, S. D. Al-majidi, M. F. Abbod, and H. S. Al-raweshidy, "Machine Learning Approaches for Short-Term Photovoltaic Power Forecasting." ***Energies,*** 17(17), 2024: 4301. **(published)**

3. S. M. Radhi, S. D. Al-majidi, “Enhancing Solar Energy Predictions in Misan, Iraq with Grey Wolf Algorithm-Driven Neural Networks”, Accepted for publication in the 6th International Scientific Conference on Engineering Sciences and Advanced Technologies at the ***American Institute of Physics (AIP)***, Hilla College, Babylon, Iraq / 28-29 November 2024. **(Scopus)**

CHAPTER ONE

Introduction

1.1. Background

Over the past few years, the global population has caused a substantial rise in the continuous energy consumption derived from coal, oil, natural gas, and other resources[1]. These resources have become increasingly hazardous due to the worsening of global warming caused by the large-scale emission of carbon dioxide[2]. Currently, the primary source of worldwide electrical energy generation is derived from fossil fuels, which account for over 80%[3]. This ratio is projected to result in about 40.4 gigatonnes of carbon dioxide emissions, by 2030[4]. To address the issue of energy scarcity in the future and mitigate the adverse consequences of fossil fuel combustion, several researchers have called for the use of renewable energy sources . Therefore, solar energy is widely recognized as an important type of renewable energy due to its cleanliness, abundance, and ease of access[5] PV power generation is a very efficient method of harnessing solar energy[6]. This energy may be conveniently captured with the use of PV panels, which can be installed on rooftops or in large-scale solar farms[7]. Solar energy is converted into electricity and utilized to supply power to the building or incorporate it into the electrical grid. The PV energy production is very competitive in electricity generation. However, it faces significant limitations due to the instability of the power system resulting from its production changing throughout the day in response to the availability of solar radiation as the intensity of solar radiation fluctuates due to factors such as geographical location, time of day, and season. Therefore, solar radiation reaches its peak during the afternoon. In addition, the strength and spread of solar radiation can also be affected by the sky conditions and wind speed[8]. The

geographical location of Iraq is characterized by a hot desert climate, especially in Misan Governorate, where summer temperatures rise from 40 to 50 Celsius and may increase during severe heat waves. In addition, it receives high levels of solar radiation. Due to these unique climatic conditions, Misan Governorate is considered one of the ideal areas for investing in PV power projects and generating power using PV panels.

Therefore, this thesis is concerned with developing methods for predicting energy production for solar PV systems in Misan. Specifically, this focuses on forecasting solar energy production for 10 to 13 hours each day at 3-minute intervals and also being taken at 1-hour intervals. The repeated use of this forecast-horizon provides ample time for PV power plant operators and the energy market to evaluate the situation and make informed decisions. All case studies included in this thesis examine a particular forecasting problem and employ several approaches and techniques, such as (a) persistence, (b) statistical, (c) machine learning (ML), and (d) hybrid methods. These approaches are classified based on their use of historical data on PV power production and associated climate factors as shown in Figure (1-1). In the persistence model, the forecasted PV power output is equal to the real power output from the previous day at a corresponding hour. This technique relies only on past PV power production data to forecast PV power generation. Statistical approaches are used to forecast PV power generation by conducting a statistical analysis of various input data. Thus, these strategies are based on historical time series data. In contrast, machine learning methods require a large dataset to provide accurate predictions about PV power generation. A machine learning model is an advanced technology capable of processing linear, non-linear, and non-stationary data patterns. A hybrid model is created by combining two or more methodologies to develop a forecast model. The hybrid model demonstrates superior performance

compared to the single model across many forecasting challenges by using the particular strengths of each technique. The time period over which the PV power output is to be forecasted is called the forecast horizon. The forecast accuracy varies with the change of the forecast horizon in the same model used with the same parameters. However, according to most researchers' reports, the PV power generation forecast can be divided into four categories based on the time horizon, as shown in Figure (1-2) and each part will be explained in the literature review. Finally, the thesis outlines the process of collecting and analyzing the precise data in the subsequent chapters.

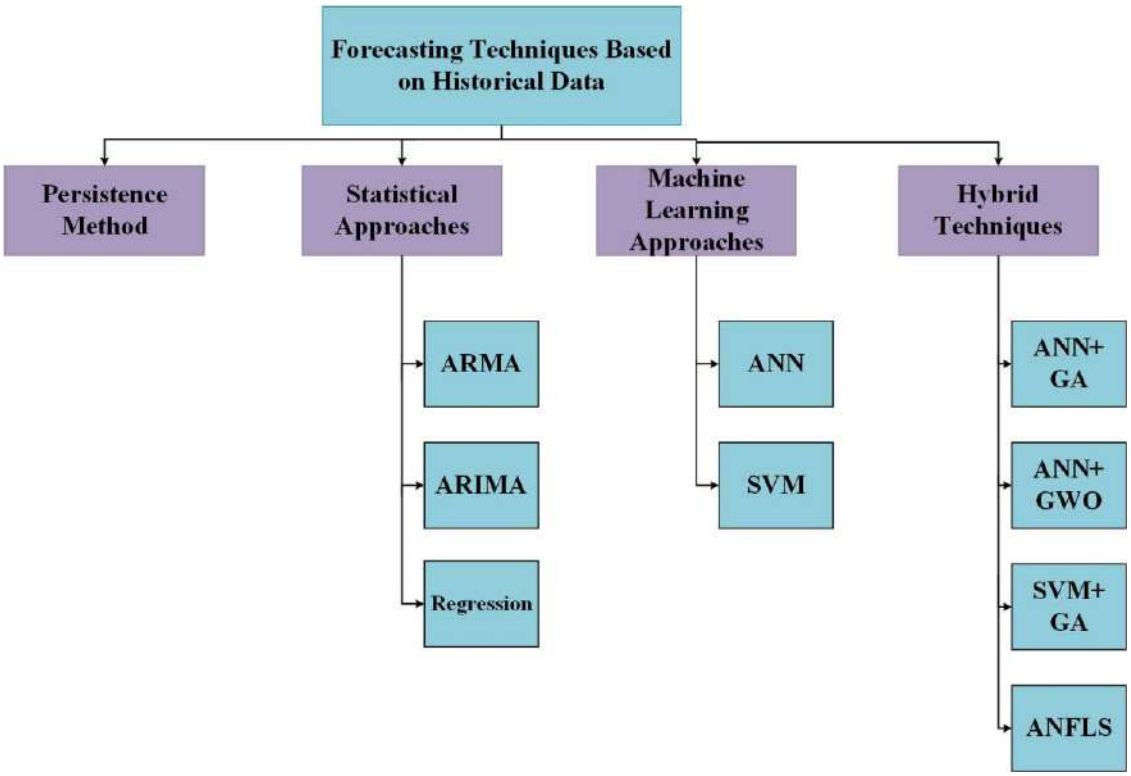


Figure 1-1 Classification of PV power forecasting based on historical data.

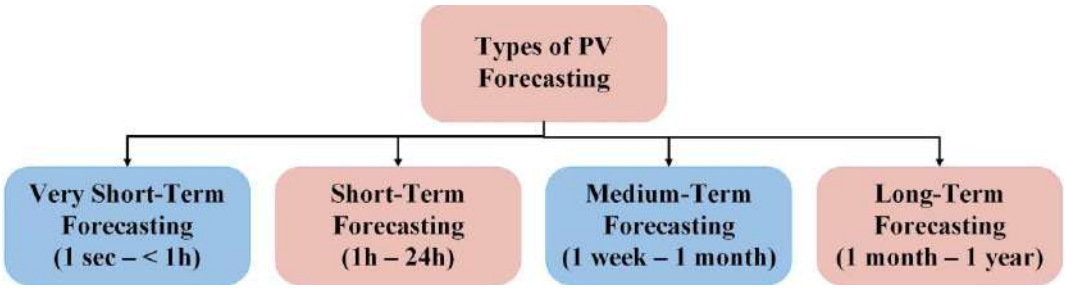


Figure 1-2 Time-based classification of PV power forecasts

1.2. Photovoltaic System

PV systems have achieved worldwide popularity for providing environmentally friendly and sustainable energy[9]. Where the deployment of PV systems has seen substantial growth in recent years [10]. Solar cells, sometimes referred to as PV cells, convert sunlight into energy via the photon-voltage effect phenomenon[11], as seen in Figure (1-3). The cells are organized in a grid-like formation and surrounded by a strong frame, with a protective cover to protect them from external factors. These solar cells have PV technology to catch sunlight and convert it into electrical power. In actuality, without being exposed to solar radiation, the PV cell functions similarly to a diode and does not produce any electrical current. When a PV cell is directly exposed the solar radiation, photons with wavelengths longer than the energy of silicon in layer n induce electrons to migrate from layer n to layer p, creating many holes. By iteratively reproducing this situation, a voltage variation is generated within the cell, which causes the movement of electrons and creates an electric current. The intensity of the light incident on a PV cell directly affects the magnitude of the electrical current produced. Moreover, the temperature and properties of semiconductors have a significant influence on the performance of the solar cell[12]. Each class of solar power system has its own set of advantages and disadvantages, and the choice depends on aspects such as efficiency requirements, space, and cost concerns. However, PV power generation is dependent on fluctuating weather

conditions and is impacted by factors such as time, air temperature, module temperature, wind speed, direction, and humidity. The system largely depends on the solar radiation absorbed by the panels. Nevertheless, this radiation is not uniform [10]. Therefore, the unpredictability and instability of solar energy supply may be ascribed to basic variables that contribute to the challenge of making accurate predictions. These aspects must be tackled in order to guarantee the dependability of the energy system. Accurately the forecasting of energy production in solar power plants is essential for improving management, maximizing efficiency, and guaranteeing the secure and steady operation of the power grid [13]-[14]. Accurate predictions of PV power will help independent power producers or energy authorities improve energy planning and management[15].

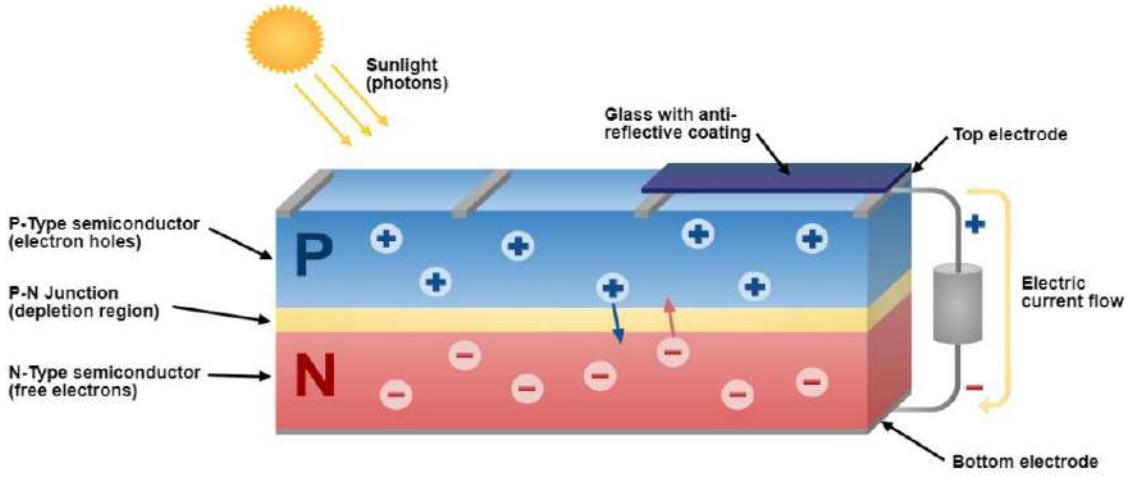


Figure 1-3 The structure of a PV cell

1.3. Literature Review

The expansion of PV has several challenges such as the unpredictable nature of energy production, influenced by meteorological factors, which can cause system disparities and affect the stability of the electrical grid. So recent literature has examined different approaches to produce and optimize PV power generation estimates to enhance efficiency in the face of seasonal or geographical fluctuations.

Researchers classify PV power forecasting based on factors such as forecast duration, weather patterns, and forecasting techniques. Accordingly, the forecasting of PV power generation can be classified into four groups according to the time frame.

1.3.1. Very Short-Term Forecasting

This prediction which has been made over period of time from 1 second to less than 1 hour, it will be clarified in below by research papers which has been suggested it in the last years.

The paper [16], developed a hybrid model combining FFNN, GA, and ANFIS for short-term PV generation forecasting, using data from Greece. The model optimized ANFIS with GA, followed by FFNN for final prediction, achieving lower error rates than separate models. It achieved an MAE of 0.4425% and an NRMSE of 6.3426%.

In the paper [17], a deep learning approach combining CNN and LSTM was presented to predict the electrical energy production of a 451.82 MW PV facility in Limburg, Belgium. The model used data from March 2015 to March 2016, recorded every 15 minutes, with data split by season for training and testing. The proposed model demonstrated superior performance compared to reference methods, achieving MAE of 1.028 and RMSE of 2.095, with better accuracy in summer due to higher solar radiation.

In the paper [18], an improved ACO approach was suggested to optimize SVM model parameters for forecasting energy production at the Desert Knowledge Australia Solar Centre. The model utilized data from 38 locations, including temperature, humidity, and solar radiation, recorded every 5 minutes. The I-ACO-SVM model demonstrated high predictive accuracy, achieving an MSE of 0.0349, RMSE of 0.1868, MAE of 0.1569, and an R2 value of 0.997.

In the paper[19], a hybrid deep learning model combining WPD and LSTM was proposed to predict PV power output one hour ahead with five-minute intervals. The model was trained on data from a PV system in Alice Springs, Australia, collected from June 2014 to May 2015, with testing from June 2015 to June 2016. The WPD-LSTM technique outperformed other models, achieving better MBE, MAPE, and RMSE values, demonstrating its potential for accurate PV power prediction.

The paper [20], suggested a hybrid model combining SR and MLP was developed to forecast PV power one month ahead using meteorological data from an Australian solar farm. The model utilized data from January 2017 to December 2018, with testing in January 2019. The hybrid algorithm outperformed individual models, achieving an RMSE of 5.58 kW, MAE of 3.3 kW, and an R2 value of 0.993, indicating superior forecasting performance.

In the paper[21], two models, ANFIS and ANFIS-PSO, were developed to predict PV system performance at the National Polytechnic Institute in Mexico City. The data, collected from 15 October 2020 to 12 December 2022, included weather and energy variables recorded every 5 minutes. The results showed that incorporating PSO into the ANFIS model significantly improved prediction accuracy, with ANFIS-PSO achieving RMSE = 0.754 kW and MAPE = 0.556%, demonstrating enhanced effectiveness for PV system performance prediction.

In the paper[22], an approach combining GAN and CAE was proposed to improve PV power forecasting accuracy using data from Jiangsu Province, China. The dataset, collected from January to December 2017 at 5-minute intervals, was divided by weather conditions (sunny, cloudy, rainy). The CAE-GAN model achieved superior performance with low errors, including an average MAE of 0.9215 and MAPE of 16.73%, effectively addressing challenges in PV forecasting.

In the paper[23], fifteen machine learning and deep learning algorithms were evaluated for PV power prediction in five California cities using weather data and

PV output from 2011 to 2020. The models were trained on data from one city and tested on others. Among the models, LSTM and GRU showed the highest accuracy, with GRU outperforming LSTM, consistently achieving an R2 score of 0.94, indicating its superior performance for PV forecasting.

1.3.2. Short-Term Forecasting

This type of forecast includes predictions within a time frame of 1 to 24 hours. The subsequent paragraphs provide a literature analysis of past short-term projections.

The paper[24], employed a hybrid GA/PSO/ANFIS approach to predict PV power generation in a microgrid in Beijing, using hourly data from 2015. The model's performance was evaluated with a dataset from 2016, showing superior accuracy with an RMSE of 7.89%, NMAE of 3.98%, and MAE of 5.31% compared to other methods.

This paper[2], proposed a simulation-based energy prediction model for PV systems using ANFIS. The model, tested with data from Thailand, demonstrated superior accuracy, achieving an MAE of 1.1952 and RMSE of 0.1184, outperforming the PSO-ANN hybrid model.

The paper[25], introduced DNN-GA hybrid models for solar irradiation prediction in Morocco. The models showed the LSTM-GA approach outperformed others, with MSE and MAE values of 0.0015 and 0.027 in summer, highlighting its effectiveness in addressing vanishing gradient issues.

The paper[26], developed a hybrid PV power prediction model combining DFFNN and RNN. Data from South Korea indicated that the PV Hybrid Network (PVHybNet) significantly improved prediction accuracy, achieving an R2 value of 92.7%, surpassing individual network performances.

The paper[27], compared ANFIS and MLP with empirical models for PV power forecasting. The ANFIS model achieved an exceptionally low NRMSE of less than 5.35 x 10-4%, outperforming other models, especially the hybrid AI approach.

The paper[28], employed a stacked LSTM model to predict PV power output 1.5 hours in advance in Cyprus. The model outperformed others, with an RMSE of 0.09394 during cross-validation, demonstrating high predictive accuracy.

The paper[29], proposed the IMWOA-SVM model for PV power prediction, optimized for both sunny and cloudy weather conditions. The model showed exceptional accuracy, with an RMSE of 0.263 and R2 of 0.995 in sunlight, demonstrating strong prediction capabilities.

The paper[7], compared ANN and MR models for PV power prediction using data from Hungary. The ANN model exhibited higher accuracy, with a COD of 0.95, MAE of 16.05, and RMSE of 28.90, showing improved performance with hybrid input methods.

The paper[30], introduced the LSTM-GPR hybrid model for short-term PV power prediction, using data from the University of Illinois. The hybrid model showed superior performance, with MAPE of 9.43%, outperforming individual LSTM and GPR models.

The paper[31], compared ANN algorithms (LM, RBP, SCG) using meteorological data from Tamil Nadu, India. The ANN model trained with the LM method achieved high efficiency, with R-values of 0.9376 for training and 0.9340 for testing data.

In the paper[14], enhanced PV power forecasting by combining WNN and GA. The method optimized performance, achieving a 3.5% relative error, though accuracy decreased on overcast and rainy days, with errors of 7.8% and 10.1%, respectively.

This paper[32], combined ANN with VMD and ACO for PV power prediction in Beijing, using hourly data from 2019. The VMD-ACO-2NN model outperformed others, with an RMSE of 0.0232 and R2 of 0.9768, showcasing high prediction accuracy.

In the paper[33], developed an MLPNN model for daily PV power prediction in Nigeria, with data from 2021. The model achieved high performance with an R2 of 93.53% during the rainy season but showed limited suitability for one-day prediction in dry seasons.

1.3.3. Medium-Term Forecasting

In this type, forecasts are made one week to one month in advance. The following research provides a review of the literature on medium-range forecasts in recent years.

The paper[34], used a technique called CSO to adjust the nonlinear and linear parameters of the RBFNN model. First, the data were collected from solar panels installed on buildings in the Netherlands. They were recorded every 15 minutes from 1st May to 31th December 2018. The model was trained by incorporating weather data from the three months of July, September, and December from the original dataset. In conclusion, the CSO-RBFNN model achieved remarkable average RMSE values of $3.843x10^{-3}$ in summer, $4.131x10^{-3}$ in autumn, and $2.846x10^{-3}$ in winter, exceeding all other methods tested.

The paper[35], conducted a comprehensive investigation into how meteorological factors affect performance parameters. They conducted correlation and interdependence studies using a grid-connected solar PV plant in Zawiyat Konta, Adrar Province, as a case study. Meteorological data were collected from February 2017 to January 2018, including ambient temperature, radiation, humidity, atmospheric pressure, and wind speed. After data processing, they were used as training input for a random forest model. The study showed remarkable prediction

accuracy, with a coefficient of determination of over 0.99, surpassing the highest results of previous investigations that achieved 0.98.

1.3.4. Long-Term Forecasting

This type of forecast involves long-term of projections that extend from one year to several years in the future. Some previous research has been conducted on the topic of long-term expectations.

In the paper[36], developed an LSTM model optimized with EMD and SCA to predict solar power output from a PV system in Alice Springs. Data from 2017 were used, with meteorological inputs showing stronger prediction accuracy in August. The EMD-SCA-LSTM model achieved high correlation with actual values, with RMSE of 0.5283 and R2 of 0.9210.

The paper[37], proposed combining CNN with SSA to predict PV power, using data from Taiwan's 500 kW PV plant. The results showed that CNN-SSA outperformed other methods like SVM-SSA and LSTM-SSA, achieving the best prediction accuracy in sunny conditions with a MAPE of 5.34%.

The paper[38], introduced a hybrid model combining LSTM and CNN for long-term solar energy prediction using data from 1990-2013. The model outperformed traditional methods, achieving a low MAPE of 2.83 and an R-value of 0.9, demonstrating high precision in forecasting PV system energy production.

In 2021, used a tree-based ML approach to predict solar PV power using meteorological parameters by **the paper[39].** used a tree-based ML approach for predicting solar PV power in Saudi Arabia using five different models. The GWO-FM and ENBS-FM models showed comparable performance, while the ENBG-FM model performed best with low RMSE and fast prediction time.

The paper[40], proposed an SSA-based model to predict SPV power in Saudi Arabia, comparing it with GWO and Levenberg–Marquardt ANN models. The SSA

model demonstrated superior performance with lower RMSE and MSE values and a high R2 value of 0.99312, showcasing its computational efficiency.

Table 1-1Summary of recent PV power forecasting studies

N	Authors	Location	Testes method	Data length	Result
Short-Term Forecasting					
1	(Panapakidis and Christoforidis, 2017)	Greece	FFNN-GA-ANFIS	2012 to 2014	with MAE of 0.4425% and MARNE of 2.6349%.
2	(G. Li et al., 2020)	Limburg, Belgium,	CNN-LSTM	2015-2016 15 minutes	MAE and RMSE were 1.028 and 2.095 respectively.
3	(M. Pan et al., 2020)	Australian	improved ACO-SVM	2018-2019 5 minutes	RMSE of 0.1868, MAE of 0.1569.
4	(P. Li et al., 2020)	Australian	WPD-LSTM	2014-2016	MBE=0.0067,MAPE=2.4002, and RMSE=0.2357
5	(Trabelsi et al., 2022)	Australian	SR-MLP	2017-2019 5 minutes	RMSE=5.58 kW, MAE = 3.3 kW
6	(Lara-Cerecedo et al., 2023)	Mexico	ANFIS-PSO	2020-2022	RMSE = 0.754 kW, MAE = 0.325 kW,
7	(X. Pan et al., 2023)	China	GAN-CAE	2017, 5 minutes	MAE=0.9215, MAPE of 16.73%.
8	(Sauter et al., 2023)	California	fifteen machine learning	2011-2020	GRU outperforming the LSTM with a consistent R2 score of 0.94.
Short-Term Forecasting					
9	(Semero, Zhang, et al., 2018)	Beijing	GA/PSO/ANFIS	2015-2016	RMSE=7.89%,MAE=5.31 %
10	(System et al., 2020)	Thailand	ANFIS	2018	MAE=1.1952 and RMSE =0.1184
11	(Bendali et al., 2020)	Morocco	DNN-GA	2016-2019	MSE=0.0015, and MAE=0.027 for LSTM-GA
12	(Carrera et al., 2020)	South Korea	DFFNN-RNN PVHybNet	2013-2015	R2 value of 92.7%

13	(Ben Ammar et al.,2021)	University of Exeter.	ANFIS and MLP	three-day, August 11 to 13, 2017	NRMSE below 0.00052%.
14	(Konstantinou et al., 2021)	Nicosia, Cyprus	DRNN	2016-2019	RMSE = 0.11368
15	(Y. W. Liu et al., 2021)	Australia	IMWOA-SVM	_	RMSE=0.263, MAE =0.212
16	(AlShafeey & Csáki, 2021)	Hungary	ANN-MR	2017- 2020	COD = 0.95, MAE = 16.05, MSE = 835.68, and RMSE = 28.90.
17	(Y. Wang et al., 2021	China	LSTM-GPR	2016 to 2017	MAPE for LSTM_GPR of 9.43%
18	(Geetha et al., 2022)	India	ANN algorithms	_	a training data R-value of 0.9376 and a testing data R-value of 0.9340.
19	(Zhang & Zhang, 2022)	China	WNN-GA	30 min	WNN with GA results in a relative error of 3.5 percent.
20	(Netsanet et al., 2022)	Beijing, China	ANN-VMD-ACO	Hourly data from 2019	RMSE = 0.0232 and R2= 0.9768
21	(Adeyemi et al., 2022)	Nigeria	MLPNN and data pre-processing.	2021	R2=93.53%, and the MAPE= 5.93% during the rainy season.
Medium-Term Forecasting					
22	(Yang et al., 2020)	Netherlands	RBFNN- CSO	Monthly weather data	The CSO-RBF model produced summer, fall, and winter average RMSE values of 0.003843, 0.004131, and 0.002846 respectively.
23	(Ziane et al., 2021)	Adrar Province	RF, PCA	2017-2018 15-minute	The study achieved a coefficient of determination of over 0.99.
Long-Term Forecasting					
24	(Zhou et al., 2020)	Alice Springs	EMD-SCA-LSTM	2017	RMSE=0.5283, MAE= 0.3063
25	(Aprillia et al., 2020	Taiwan	CNN-SSA	2017	MAPE values of 5.34%, and 42.55% in the sunny and rain model respectively.

26	(Ray et al., 2020)	Australian	LSTM-CNN	1990-2013	MAPE= 2.83 and R2= 0.9.
27	(Alaraj et al., 2021)	Saudi Arabia	SVR-FM, DT-FM, ENBG-FM, ENBS-FM, and GWO-FM	four years	RMSEs=15.91 throughout training and 19.66 W during testing for ENBG-FM.
28	(Alaraj et al., 2023)	Saudi Arabia	SSA	Ten years	R2= 0.99312

1.4. Motivations

Due to world development, the traditional energy supply is fast depleting, increasing the quantity of energy required for traditional energy and resulting in problems like cost, energy crises, and environmental challenges. Solar energy, despite its advantages, is highly variable due to its dependence on factors such as solar radiation, temperature, cloud cover, and solar hours, and it is only available intermittently during the day [41]. This variability, conditions, and intermittency of solar energy can lead to variations within the PV system, affecting the stability of the interconnected electrical grid. Due to the intermittent and uncontrollable nature of solar energy production, accurate prediction of solar generation is critical for both the grid and operators. The main duty of the network operator is to plan the supply and demand of electricity to maintain a balance between them. Where adverse weather conditions result in a rapid surge in demand. For example, demand typically increases during daylight hours and in the summer or winter seasons when heating or cooling systems are being utilized. An imbalance between supply and demand could lead to blackouts or voltage drops, while inadequate production or excess capacity without storage solutions can strain the grid. Thus, utilizing big data and machine learning techniques to enhance the precision of demand and production predictions might facilitate more effective planning and operation. Furthermore,

enhanced meteorological predictions could provide more precise anticipation of renewable energy generation, where batteries can store surplus energy during times of low demand and discharge it during times of high demand. Hence, it is imperative to precisely forecast solar energy production to avoid any variations or disruptions in the provision and thereby sustain grid stability[42]. This thesis examines the performance of advanced machine learning techniques, specifically artificial neural networks (ANN), as well as optimization approaches like the genetic algorithm (GA) and the gray wolf optimization (GWO) to enhance PV energy production forecasting by accurately modeling non-linear relationships between weather conditions and power output. ANN provides a flexible and powerful framework for modeling complex patterns, while GA and GWO offer robust optimization strategies to refine the ANN structure and parameters. These optimization algorithms help in achieving optimal network performance by adjusting the number of hidden layers, neurons, and other crucial parameters, leading to more reliable and precise forecasting results. This combination of machine learning and optimization techniques contributes to the advancement of renewable energy management and planning, especially in regions where solar energy production is influenced by highly variable meteorological conditions.

1.5. Aim and Objectives

The research aims to study the impact of changes in weather conditions in Iraq on the performance of solar panels and the quantum of power production from PV systems. Since changes in weather conditions lead to fluctuations in power production, accurate forecasting contributes to achieve a balance between supply and demand for power. Thus, the results help develop plans for the future expansion of solar panels in Iraq, which helps achieve sustainability in electricity generation from renewable energy sources and reduces dependence on fossil fuels.

The main objectives of this thesis are as follows:

1. Enhancing forecast accuracy based on medium-term meteorological data for 2021, collected hourly and divided into four seasons in Misan, to explore the impact of seasonal climate changes on PV power production. Then the short-term experimental data for 2024, collected every 3 minutes and divided into weather conditions in Misan, to analyze the impact of each type of weather on PV power production. This data provides an accurate and comprehensive representation of different weather conditions, helping to build accurate forecast models for seasonal production fluctuations and different weather conditions based on solar radiation and temperature.
2. Study the performance of current state-of-the-art machine learning methods for solar energy forecasting and propose improvement methods to find the best accuracy for forecasts of PV energy production based on data used.
3. Comparison of theoretical and experimental results to demonstrate the effectiveness of the models used to predict PV power production under the influence of different data and weather conditions.

1.6. Thesis Contributions

The main contributions of this research are as follows:

1. A MATLAB simulation model was created to gather actual power data by importing solar radiation and temperature values from the website (climate onebuilding). This model represents a solar system consisting of 4 PV modules. The system was enhanced by incorporating a boost converter, which allows the electricity produced by the panels to be used to power loads.
2. A small PV system is installed at the University of Misan College of Engineering in Iraq, to collect a training and real data set from the PV array.

3. The ANN methodology was formulated based on empirical data gathered from the University of Misan in Iraq in 2024, specifically in the periods of January, March, and June. Theoretical data, on the other hand, was sourced from the website (climate onebuilding) of 2021.
4. The ANN was used to predict power production due to its ability to handle nonlinear and complex data. In order to enhance the performance of the network, the GA and GWO were used. The objective of using the two algorithms is to explore the potential of both the old and modern algorithms in improving the performance of NN and increasing the accuracy of predicting PV power production.
5. The data were subjected to training by both algorithms followed by a comparative analysis. The results indicated that the model utilizing the ANN-GA exhibited superior performance compared to the model utilizing the ANN-GWO when tested across different datasets and weather conditions.

1.7. The Thesis Structure

The thesis has been divided into six chapters, which are as follows:

- Chapter One: This chapter provides an overview of the study, including the background of PV energy, the motivation behind the research, and the objectives and contributions of the thesis. In addition, studies on the prediction of PV energy generation are presented, including theoretical data and experimental data in different prediction periods. Finally, a summary of the content in each chapter of this thesis is provided.
- Chapter Two: This chapter introduces a PV cell model. A MATLAB-SIMULINK model was developed to simulate a PV system that captures real energy by incorporating solar irradiance and temperature parameters in the embedded solar panels. Additionally, the model includes a DC-DC boost converter, which is

controlled using traditional Perturb and Observe (P&O) methods. The chapter further explains the process of collecting experimental data to validate the model.

- Chapter Three: This chapter details the technology used, specifically the enhanced ANN model optimized through a combination of GA and GWO. It explains the processing and division of both experimental and theoretical data, followed by a comprehensive outline of the methodology employed to achieve accurate predictions and performance assessment.
- Chapter Four: This chapter presents the experimental and theoretical results, discussing each in detail before comparing them. The analysis highlights differences and similarities between the two data sets, examining how each method performs under various conditions. The comparison aims to validate the model's accuracy and assess its applicability in real-world scenarios.
- Chapter Five: In conclusion, the most important conclusions are presented and highlighted. Furthermore, some critical recommendations are made to extend and enhance the current work with suggested avenues for future research.

CHAPTER TWO

Mathematical Modeling of PV System

2.1. Introduction

This chapter focuses on the modeling and PV data collection for the solar energy system utilized in this thesis. It begins with an overview of developing a solar cell model, followed by the collection of theoretical data. This data is generated through a MATLAB simulation model that includes solar panels connected to a boost converter, with the power output regulated by a Maximum Power Point Tracking (MPPT) system using the P&O algorithm. Details of the data collection methodology are provided. Subsequently, experimental data is collected from a small-scale PV system, comprising four PV panels equipped with sensors to measure solar irradiance, temperature, and current. Finally, the most important performance metrics used in this thesis are explained to evaluate the PV system's performance.

2.2. Modelling of a PV Cell

The primary function of PV cells is to convert light into electrical energy through the phenomena of photo-voltage effect. The most prevalent design is the single diode configuration with both series and shunt resistors, as depicted in Figure (2-1). The PV cell is depicted in the diagram as a current source, denoted as I_{Ph}, which is linked in parallel to a diode. The present source of current, denoted as I_{Ph}, is generated when light interacts with PV cells. The magnitude of I_{Ph} is directly proportional to the intensity of solar radiation. In a typical solar cell, resistances are not explicitly incorporated but are integrated and linked with the PV diode in real-world applications. The equations can quantitatively represent the PV module

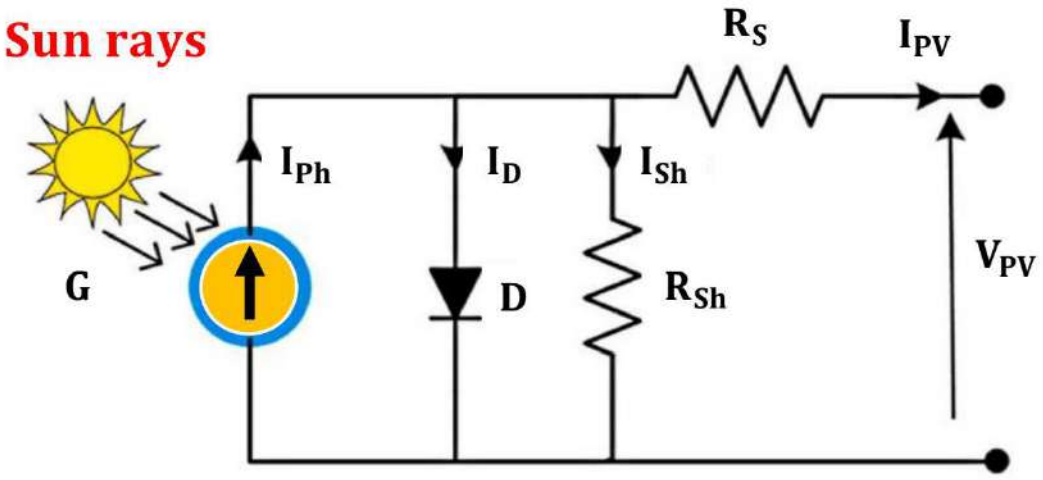


Figure 2-1 Single-diode model of the PV module

The total current (I_{Ph}) generated by the module can be determined by applying Kirchhoff's law, as stated in equations (2-1) [43]:

$$I_{PV} = I_{Ph} - I_D - I_{Sh} \qquad (2\text{-}1)$$

Where the shunt current going through the shunt resistance is denoted as I_{Sh}, the PV diode current is denoted as I_D, and the PV output current is denoted as I_{PV} [43].

Where I_{PV} represents the present generator as stated in equation (2-2):

$$I_{PV} = [I_{SC} + K_i(T - 298)]\frac{G}{1000} \qquad (2\text{-}2)$$

Where G represents the solar irradiation, T represents the ambient temperature of the climate conditions, I_{SC} represents the short circuit current of the PV cell, K_i represents the temperature coefficient, which is determined by Shockley's Equation(2-3)[43].

$$I_D = I_0\left(\exp\left[\frac{qv_D}{nKT}\right] - 1\right) \qquad (2\text{-}3)$$

Where the saturation current of the PV diode is denoted as Io, the voltage across the PV diode is represented by Vd, the electrical charge is symbolized as

q (1.69×10^{-19} C), the Boltzmann constant is denoted as k (1.38×10^{-23} J/K), and the PV diode factor is represented by N[43]. The following equation depicts the leakage current (**I_{sh}**)

$$I_{Sh} = \frac{(V_{PV} + I_{PV} R_s)}{R_{Sh}} \qquad (2\text{-}4)$$

So, the output current of the PV cell resulting from substituting equation number (2-3) and (2-4) in equation number (2-1)

$$I_{PV} = I_{Ph} - I_o \left(\exp\left[\frac{q(V_{PV}+I_{PV} R_s)}{nKNT}\right] - 1\right) - \frac{(V_{PV} + I_{PV} R_s)}{R_{Sh}} \qquad (2\text{-}5)$$

Where I_{PV} represents the current output of the PV system, R_S the Series resistance of PV cell, R_{Sh} the Shunt resistance of PV cell and V_{PV} represents the voltage output of the PV system.

2.3. Theoretical Test Data Components

In this section, the components of the theoretical work will be described. The theoretical work is a crucial aspect of understanding how to design and improve the performance of these systems. The theoretical work components include a MATLAB simulation model for collecting real energy from meteorological data and a set of key components that help build a comprehensive model that can be used to analyze and improve performance. These components include the solar panels, the DC/DC boost converter, and the method used to control it.

2.3.1. Proposed System Description

The main objective of this section is to collect the actual power by using the data collected from meteorological in 2021 to make predictions for PV power generation. More precisely, the data used in this thesis is related to the city of Amara, Iraq, located at coordinates 31.54° N, and 47.2° E see Figure (2-2). The data can be

obtained from the website (climate.onebuilding). In this section, a solar power system scheme will be implemented using simulation in MATLAB/Simulink, MATLAB is one of the main programs for modeling, analyzing, and solving problems related to dynamic systems, through which the voltage, current, and actual power are collected using solar radiation and temperature as inputs to the solar panels. The collected data is used in an Excel file. The simulated model in MATLAB/Simulink is shown in Appendix A.

The main component used in the simulation scheme includes the solar panel array, in this thesis, four solar panels with a power of 1400 W are used for power generation. Solar panels are one of the main components of the PV system, which convert solar power into electrical power and the type of panels is selected based on their efficiency and suitability for the target environment. The specific characteristics of the solar panels utilized are outlined in Table (2-1) [44]. It is important to observe that many PV solar cells are connected in parallel and series to obtain the required current and voltage for the solar panel[11]. In addition, to determine the output power of the PV system, the temperature and solar radiation data must be utilized as inputs for the solar panel and then actual PV power is measured using a sensed voltage and current of a PV Simulink operation. For this reason, with the increase in temperature, the voltage decreases and the current increases slightly as a result, the PV power is reduced. On the other hand, with the increase of solar radiation, the current of the PV system also increases because it is directly proportional to the solar radiation, while the voltage of the PV systems shows little change[45]. Therefore, a boost converter, a crucial element in enhancing the solar panels' output voltage.

Table 2-1 Overview of the Simulink PV module test

Characteristics	Values
Cell Number	90
Open circuit voltage	41.07 V
Maximum power voltage	34.23 V
Short circuit current	11.25 A
Maximum power current	10.23 A
Maximum power point	350 W
Temperature Coefficient (Voc)	-0.272
Temperature Coefficient (Isc)	+0.061%

Figure 2-2 Location at the city of Amara, Iraq(Coordinates: 31.54° N, and 47.2° E)

2.3.2. DC-DC Boost Converter

A DC-DC converter is built and connected to capture energy from PV arrays for integration with the power grid, by utilizing a boost converter, as shown in Figure (2-3). It is possible to make the PV source linear and extract the maximum power by suitably altering the duty cycle. A boost converter is commonly used as a voltage regulator to convert a fluctuating DC voltage into a stable DC value. As a result, the output voltage experiences variations when there are changes in radiation or temperature. The DC-DC boost converter is centered around a transistor that governs

the amplified processing under the control of a controller. The MOSFET, BJT, and IGBT are often used transistors in a DC-DC converter. Nevertheless, the MOSFET transistor is commonly favored for the construction of the DC-DC boost converter because it can efficiently function under demanding loads and higher frequency situations [46][47], while also demonstrating reduced power losses[11]. In the equation below, duty cycle (D) (let D=0.5) represents the ratio of conversion of input voltage (V_s) to output voltage (V_O) [48]:

$$V_O = \frac{V_s}{1-D} \qquad (2\text{-}6)$$

The output voltage of the converter is denoted as (V_0), whereas the input voltage of the converter is denoted as (V_s). The switching period is represented by the symbol T_s. DT_s indicate the duration when the switch is in the ON state, while $(1\text{-}D)T_s$ represents the length when the switch is in the OFF state[48]. In an ideal circuit, the converter's input power is equal to the output power as shown equation.

$$P_o = Ps \;\Rightarrow V_O I_O = V_s I_s \qquad (2\text{-}7)$$

Where P_o is the output power and P_s is the input power of the DC-DC converter, I_o is the output current and I_s is the input current of the DC-DC converter.

A-Inductor of the Boost Converter*:*

Where f =50 KHz is the frequency of switching, R=70 ohm.

$$L_{min} = \frac{D(1-D)^2 R}{2f} \qquad (2\text{-}8)$$

$$= \frac{0.5*(1-0.5)^2*70}{2*50} = 0.0875 \text{ m}H$$

The boost converter intended for continuous-current operation will require an inductor value that exceeds the minimum value, L_{min}. From a design standpoint, it is advantageous to express the inductor value, L, about desired change in current, Δi_L is the ratio between the ripple of the input current to the output current, and the optimal value is within 20% to 40% for this ratio[48].

$$L = \frac{V_S DT}{\Delta i_L} = \frac{V_S D}{\Delta i_L f} \qquad (2\text{-}9)$$

B- Capacitor of the Boost converter:

where ($\Delta \mathbf{V_o}/\mathbf{V_o}$) is the Voltage Ripple Factor, which is the ratio between the ripple of the output voltage and the output voltage. Typically, this ratio is limited to a range of 1% to 5%[48], let ratio 2%.

$$C = \frac{D}{R\left(\frac{\Delta V_o}{V_o}\right) f} \qquad (2\text{-}10)$$

$$= \frac{0.5}{70*(0.02)*50} = 0.00714 \text{ m}F$$

where **C** is the capacitor, which is used to filter the ripple PV voltage, Ir is solar radiation and Te is Temperature.

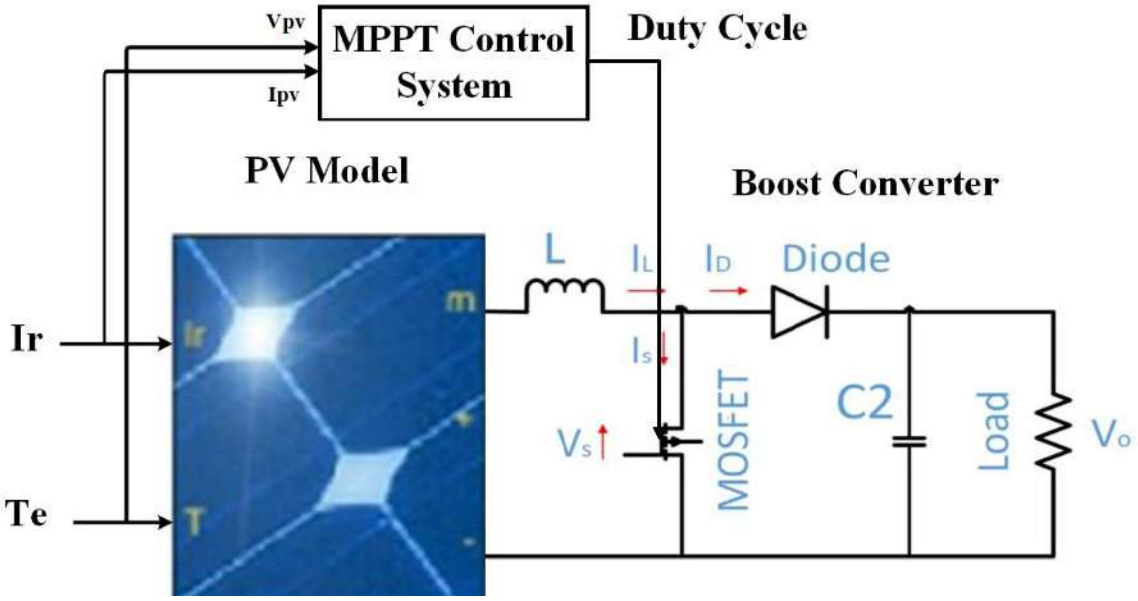


Figure 2-3 Block Diagram of a PV System with MPPT Control and Boost Converter

2.3.3. Perturb and observe (P&O) Algorithm

As mentioned previously, renewable energy sources, depend on environmental factors such as temperature and radiation within solar energy systems. Therefore, the optimization of system performance to operate at maximum power levels within

specific timeframes is achieved through the implementation of Maximum Power Point Tracking (MPPT). The control strategy in solar energy systems plays a crucial role in enhancing the efficiency of solar panels by employing MPPT techniques. These techniques involve using the boost converter mechanism to adjust the duty cycle of power electronics switches. The P&O controller is a commonly used method in MPPT, as depicted in Figure (2-4), which includes monitoring voltage and current fluctuations from the solar panels and comparing them to previous PV power and voltage values. This comparison allows for adjustments in the algorithm to regulate the reference voltage for the power converter by introducing slight disturbances to the duty cycle of the converter or MOSFET. Continual monitoring and adjustment of the duty ratio based on output power variations ensure optimal system performance. The affordability and simplicity of implementation make the P&O algorithm a popular choice for PV-MPPT technologies, adapting effectively to changes in solar irradiance and temperature to maintain consistent performance[49].

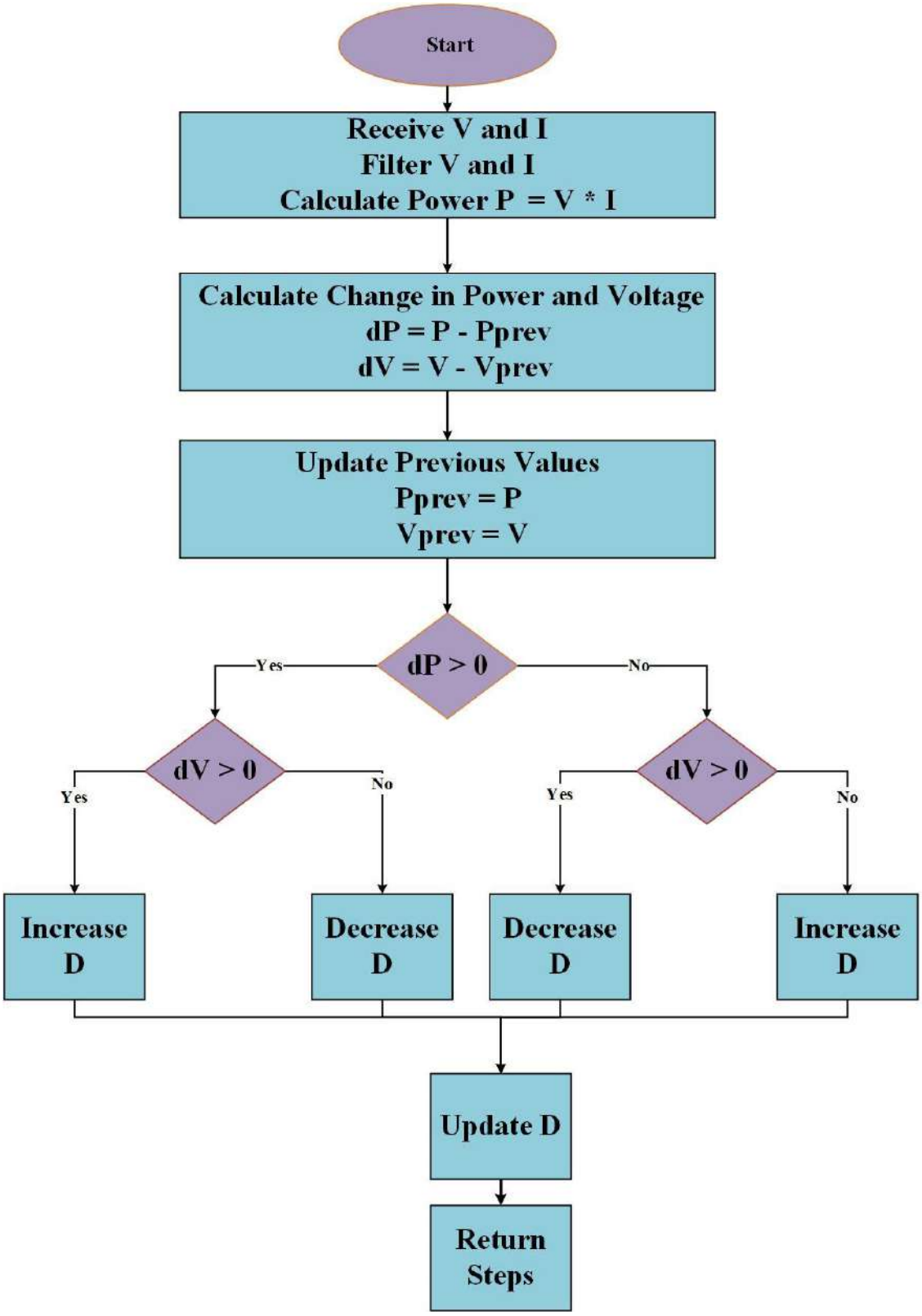


Figure 2-4 Flowchart for MPPT Control System Using P&O Algorithm

2.3.4. Collected Theoretical Data

The primary aim of this study is to create forecasts for PV power generation in order to guarantee the robustness of grid-connected PV systems. The weather data retrieved from the publicly accessible website (climate onebuilding), covering the period from 2007 to 2021, is utilized. The data utilized for the year 2021 comprises measurements taken on an hourly basis. To provide further clarification, the data included in this analysis pertains only to Amara, a city situated in the southern part of Iraq. The overall data collected during 2021 amounts to 8760 readings, encompassing the entirety of each day. This dataset contains hourly measurements of solar radiation, temperature, wind speed, and relative humidity, including data collected during nighttime hours. Nevertheless, this thesis employs solar radiation and temperature as inputs for the model utilized to forecast PV power generation. Therefore, after constructing the simulation model in MATLAB/Simulink, the given data is inputted into the MATLAB model to gather the active power values on an hourly basis, specifically for the months that were included in the simulation. Next, the data is extracted from the Excel file and placed in specified variables. The temperature data is assigned to the variable "Te" and the solar radiation data is assigned to the variable "Ie". Next, the Simulink model, which is the designated name of the model saved in the computer, is invoked. The voltage (V_PV), current (I_PV), and power (P_PV4) values are extracted using the (To Workspace block). The values are then stored in the Ve, Ie, and P matrices, correspondingly. The data is structured in an Excel file for future usage. See the code in Appendix A and the theoretical data for January in Appendix B or use the Google Drive link to view the full data.

2.4. Experimental Test Data Components

This section includes many elements of the experimental work, which involves creating and utilizing a data acquisition system to monitor voltage, current, and radiation in a DC circuit (specifically solar panels), as well as temperature. The main processing unit for this system will be a Siemens S7-1200 PLC. The Programmable Logic Controller (PLC) utilized in the system is equipped with 6 inputs and 6 outputs. The device is programmed using the development software (TIA PORTAL) provided by the vendor. The system architecture is depicted in Figure (2-5). Furthermore, the explanation can be divided into multiple sections covering topics such as solar panels and devices used for measurement and control. The experimental data collection phase occurred between December 2023 and June 2024.

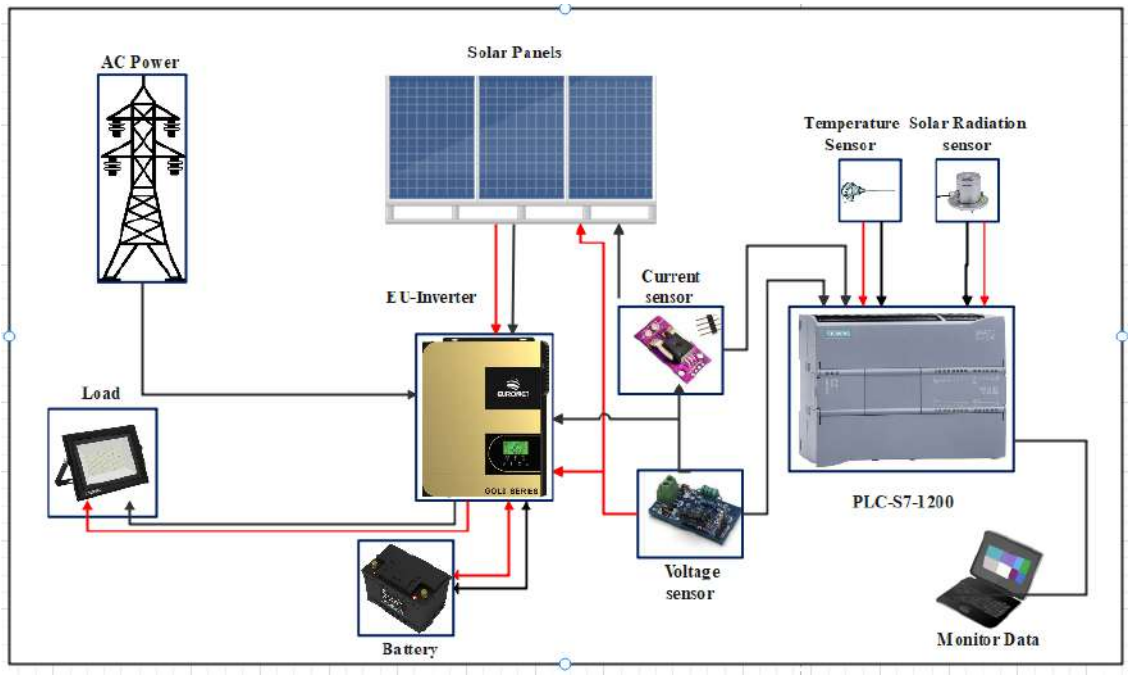


Figure 2-5 System Architecture

Figure 2-6 The outdoor PV system installed at the Engineering campus of University Misan, Iraq

2.4.1. Solar Panels (EU-M350W)

Solar panels convert sunlight into electricity through PV cells, which capture photons and generate electricity via the PV effect. These PV cells are grouped into modules, and several modules form an array or system. A typical PV system includes four solar panels, an inverter to convert DC electricity to AC, and control and monitoring devices. The solar cells are usually connected in series to achieve the desired voltage, and then in parallel to enhance current. Solar panels are a sustainable and eco-friendly energy source, reducing greenhouse gas emissions and lowering energy costs. However, their efficiency is dependent on sunlight availability and strength, and their performance is often evaluated under standard conditions that differ from actual installation environments. In this study, the Euronet EU-M350W solar panels are used, with four panels in total. Each panel has a maximum power output of 350W, operates at 34.23V with a current of 10.23A at peak power. The panels are connected in series to maximize voltage and positioned optimally to

increase sunlight exposure. The installation of the outdoor PV system is located on the engineering campus of the University of Misan in Iraq, as shown in Figure (2-6). Detailed characteristics of the panels are provided in Table (2-2).

Table 2-2 Specifications of the PV module test

Characteristics	Values
Cell Number	90
Open circuit voltage	41.07 V
Maximum power voltage	34.23 V
Short circuit current	11.25 A
Maximum power current	10.23 A
Maximum power point	350 W
Temperature Coefficient (Voc)	-0.272
Temperature Coefficient (Isc)	+0.061%
Dimension (mm)	1755*1038*35mm
PV cell model	EU-M350W
Operating Temperature	-40^{o}C to +80^{o}C

2.4.2. Sensors Used at Work

A solar radiation sensor measures the intensity of solar radiation or sunlight received by the surface of a PV panel and is usually expressed in W/m^2, as shown in Figure (2-7). It helps monitor the amount of solar energy received by PV panels, which can be used to analyze their performance and efficiency[50]. It is commonly used in meteorological, agricultural, and solar energy applications. The pyranometer must be installed in an open area, away from any obstacles that may block sunlight. It must be installed horizontally and properly leveled to ensure accurate measurement. The pyranometer should also be calibrated regularly to maintain accuracy. Data showed that on March 5 (a cloudy day), radiation peaked at 158 W/m², while on March 9 (a sunny day), it reached 976 W/m², reflecting the impact of varying weather on PV output, which rises with morning sunlight and declines by afternoon, As shown in the figure below (2-8) to represent sunny weather and (2-9)

to represent rainy weather. The specifications of the Pyranometer Solar Radiation Sensor are shown in Table (2-3).

Figure 2-7 Solar pyranometer

Table 2-3 The specifications of the RS485 Modbus Pyranometer Solar Radiation Sensor

Attribute	Value
Measuring range	0～1500W/m^2
Sensitivity	7~14μV/wm
Output Type	Digital, Relay
Temperature characteristics	±2% (-10°C～40°C)
Response time	≤35 seconds (99%)
Working temperature	-50 ° C ~ 50 ° C

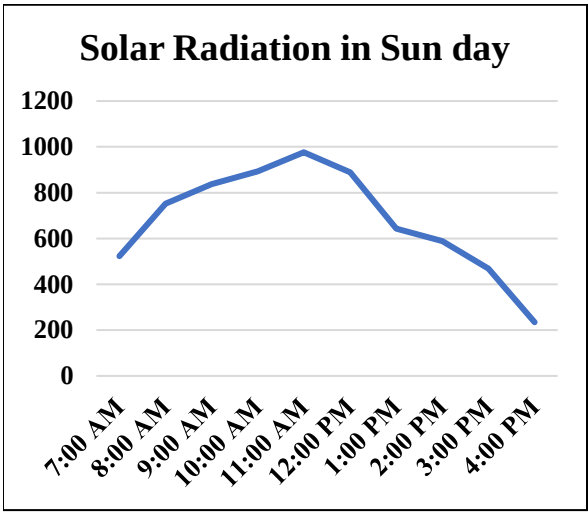


Figure 2-8 Solar radiation on a Sunny day

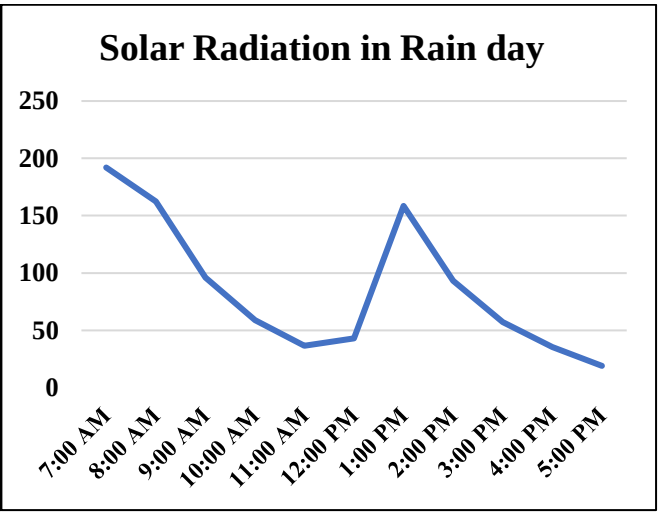


Figure 2-9 Solar radiation on a Rainy day

In addition, a thermocouple sensor records ambient temperature around the PV panels, generating voltage based on the temperature difference between connected metal wires and the current sensor. The temperature sensor utilized is depicted in Figure (2-10) and the current sensor in Figure (2-11). Finally, voltage readings, ranging from 10 to 160 V, were collected from the inverter, as direct voltage sensors were unavailable. This voltage powers the load and charges batteries, which then light the area at night.

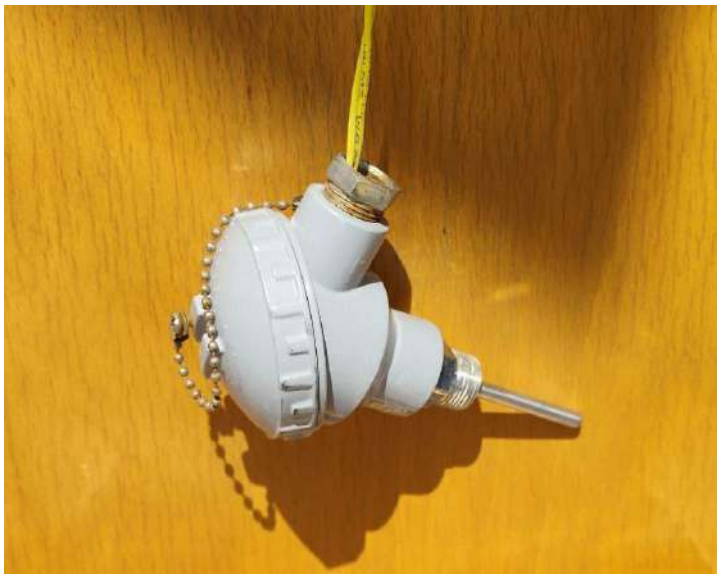


Figure 2-10 Temperature Sensor

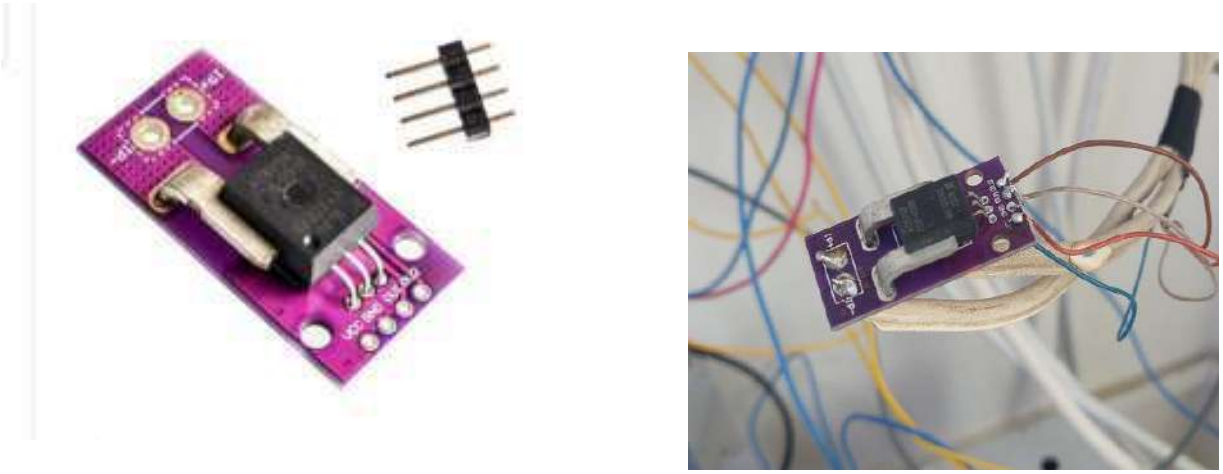


Figure 2-11 current sensors

2.4.3. PLC (S7-1200 Station)

A programmable logic controller (PLC) is a specialized computer designed to control manufacturing processes and machinery, accommodating both AC and DC power sources. Known for its reliability and robust construction, the PLC is durable and well-suited to harsh industrial environments. Equipped with a powerful CPU and high-speed communication features, it supports seamless data exchange and efficient process management. Its programmable memory allows it to handle tasks such as timing, sequencing, arithmetic, and data processing. The PLC S7-1200 AC/DC Relay, programmed using Siemens TIA Portal software, offers versatility and supports complex programming with an intuitive interface, including ladder logic and debugging tools. The PLC setup includes an external analog input card with four inputs, enhancing its compatibility for various applications.[51].

The Siemens TIA Portal software is used to program the PLC S7-1200 AC/DC Relay. Figure (2-12) below shows the TIA Portal software interface. Figure (2-13) shows the PLC with its extension. The specifications of the Siemens S7-1200 PLC are shown in Table (2-4).

Table 2-4 The specifications of the Siemens S7-1200 PLC

Attribute	Value
Brand	Siemens
Memory	4 MB
Output Type	Digital, Relay
Minimum Operating Temperature	-20° C
Maximum Operating Temperature	+60° C
Number of Outputs	10 (Digital Output, Relay Output)
Manufacturer Series	S7-1200
Programming Language Used	FBD, LAD, SCL
Output Current	2A
Input Type Analogue, Digital	Number of Inputs 14 (Digital Input, 2 switches as Analogue Input)
Dimensions	100 x 110 x 75 mm

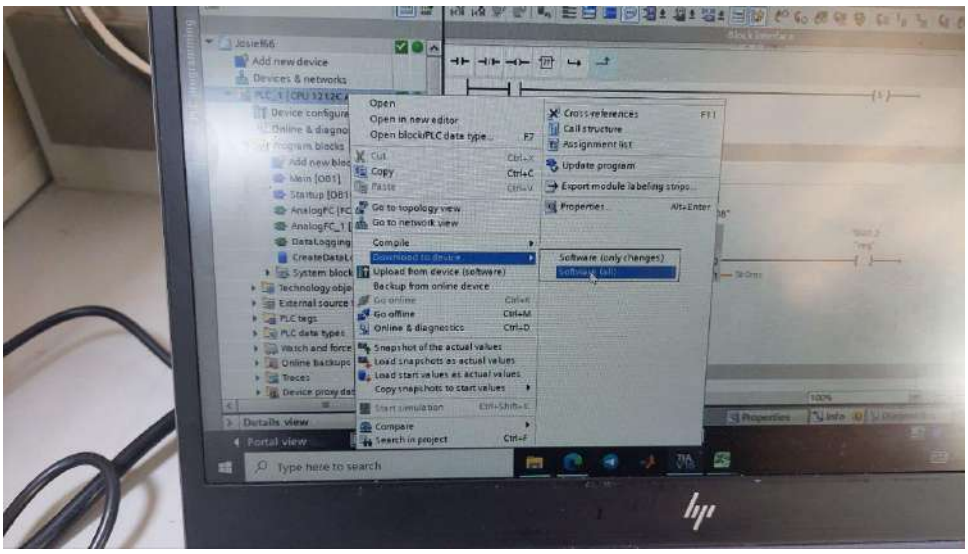


Figure 2-12 Programmer PLC

Figure 2-13 The Programmable Logic Controller (PLC) and the extension it provides

2.4.4. Euronet 5.5kW Inverter Gold

The EuroNet 5.5k Gold inverter acts as a device that converts direct current (DC) generated by PV panels into alternating current (AC), and is suitable for both residential and industrial applications, as shown in Figure (2-14). The Euronet 5.5kW Gold power inverter is accurately designed to achieve an extremely high-efficiency rate, featuring a conversion efficiency of up to 95%. This demonstrates its ability to efficiently convert DC electricity into AC electricity, thus providing a reliable and environmentally friendly power source. In addition, the device is characterized by its ability to seamlessly switch to battery power in the event of a power outage, thus ensuring continuous operation of devices even in the absence of a power grid. Therefore, in this thesis, this inverter has been integrated with a charger that allows direct charging of batteries from the power grid. This function improves the process of maintaining a constant charge for the batteries.

One of the main features of the inverter is that the feature of interface designed for ease of use, allowing real-time monitoring and management of operational parameters. In addition, the device is made of flexible materials and a strong frame, ensuring its durability and flexibility in harsh conditions. Moreover, its compact and lightweight design makes it easy to install across different solar power systems. The battery operates at 12 volts. The system is designed to operate alongside utility power, and in the event of a power outage, it will seamlessly switch to PV electricity.

Figure 2-14 Inverter device for PV power System

2.4.5. Collected Real Data

The experimental data of this section are collected to predict the performance of PV power generation from a small-scale PV system installed on the campus of the College of Electrical Engineering, University of Misan, Iraq. This system includes a meteorological station containing on solar radiation and temperature sensors, which

are used to record data which based on changes in weather conditions. The data is collected for using a central controller, a Siemens Automation S7-1200 programmable logic controller, which is carefully organized to ensure the accuracy of data collection at the meteorological station. The PLC receives data from solar radiation, temperature, and current sensors. The PLC can then process this data, perform calculations, and execute program instructions to control different aspects of the system. In addition, it can perform programming and adjustments remotely via a network interface. Based on the programming instructions, the PLC can record the sensed data over time, and it has a user-friendly interface that allows for easy programming and monitoring. In addition, the main computer controls the movement of data, processes it, and stores it on the server. When data is received from the PLC and network devices, the data blocks are classified based on protocols, separated according to individual sensors, then organized, converted into physical quantities, and stored along with metadata that enables comprehensive identification of sensors and measurements. Within a main database containing temporal information. Each record is systematically arranged and stored in memory daily. To read the electrical power and weather data, a computer was connected to the PLC as Figure (2-15). A supervisory control and data acquisition system was used to monitor and control the system. The PLC reads the data every 3 minutes periodically during the day and shuts down at night. The data was recorded on the computer and stored in a Microsoft Excel sheet. Finally, a data set of approximately 12,613 readings from January, March, and June, at a rate of 3 minutes, was archived within an electronic PLC device. The readings for variables such as temperature and solar radiation are recorded and measured by dedicated temperature and radiation sensors. Calculating the PV power generation parameters requires the use of a current sensor to extract the current which is then accessed by the PLC via the data log generation

directive while the voltage values are taken from the inverter across a range of approximately 10 to 160 volts. See more information in Appendix C. Now, collected the details are completed to the next contribution.

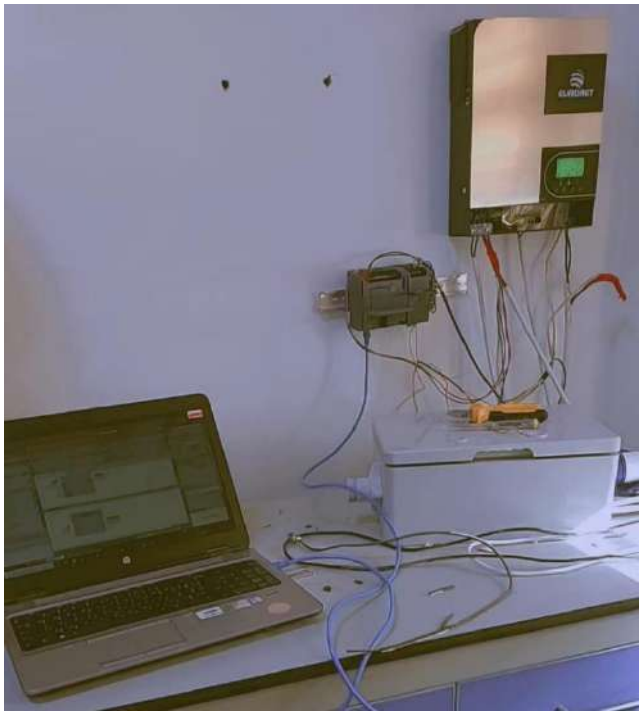


Figure 2-15 The computer connection to PLC to download PV power system data

2.5. Evaluation Indicators

Choosing the appropriate evaluation metrics is essential to examine the performance and effectiveness of the ANN model in predicting PV energy generation. These metrics provide the model's accuracy, reliability, and predictive ability, thus enhancing the credibility of the predictive results and their applicability. This thesis has five commonly used evaluation metrics, R2, MAE, RMSE, MSE, and RE, which can be called using the sklearn library in Python. It is worth noting that the accuracy of the model shows an inverse relationship with the values of these evaluation metrics. At the same time, the level of accuracy of the model shows a direct relationship with R2[52].

- Mean Square Error (MSE) is defined as [52]

$$MSE = \frac{1}{M}\sum_{m=1}^{M}(Pa - Pf)^2 \qquad (2\text{-}11)$$

where m ranges from 1 to M, with M being the total number of data points in the test set, Pa actual power and Pf predicted power values in the test set.

- Root Mean Square Error (RMSE) is defined as[52]

$$RMSE = \sqrt{\frac{\sum_{m=1}^{M}(\mathrm{Pa}-\mathrm{Pf})^2}{M}} \qquad (2\text{-}12)$$

- Mean Absolute Error (MAE) is defined as [52]

$$MAE = \frac{1}{M}\sum_{m=1}^{M}(\mathrm{Pa} - \mathrm{Pf}) \qquad (2\text{-}13)$$

The closer the MAE and RMSE values are to zero, the better the prediction performance.

- Pearson's coefficient of determination (R^2) measures the degree to which a statistical model successfully predicts an outcome. a model's predictive ability is considered greater when its R^2 value approaches 1, as demonstrated in Equation (2-14) [52].

$$R^2 = 1 - \sqrt{\frac{\sum_{m=1}^{M}(\mathrm{Pa}-\mathrm{Pf})^2}{\sum_{m=1}^{M}\left(\mathrm{Pa}-mean(\mathrm{Pa})\right)^2}} \qquad (2\text{-}14)$$

- Relative error is a metric employed to quantify the precision of a measurement or estimation in comparison to the actual value or accepted norm. The error is commonly quantified as a percentage, indicating its magnitude relative to the real number. The equation for computing relative error is[52]:

$$Relative\ Error(\%) = \frac{\mathrm{Pa}-\mathrm{Pf}}{\mathrm{Pa}} \times 100 \qquad (2\text{-}15)$$

2.6. Summary

The first part introduces the process of creating a model for the solar cell whilst the second part describes how to create a simulation model in MATLAB based on site data that contains radiation and temperature. The Four solar panels were used and connected to a boost converter. The main goal of the model was to collect the necessary PV current and PV voltage data through the use of solar radiation and temperature. The model provided almost perfect measurements. Next, the most important Python libraries used in this thesis are described. The fourth part explains that a PV system was previously implemented that contains sensors, solar panels, and a PLC controller. The data obtained from the four input channels of the expansion unit are stored in their transfer registers, which are interrogated by the PLC through data register generation instructions and data register writing instructions. The communication method is based on transferring data from PLC to computer via router interface. Using a web server interface to connect the data acquisition system to the computer terminal provides a stable and fast enough communication method. Therefore, linking the data acquisition system with the computer terminal via the router standard results in a real-time monitoring and control system, suitable for working in different environments. In the last part, the most important evaluation indicators used in this thesis were presented.

CHAPTER THREE

Machine Learning Forecasting Model

3.1. Introduction

In this section, ANN and optimization methods based on GA and GWO are explained in detail, which are used in optimizing artificial neural networks. The GA and GWO provide optimization of the number of layers and neurons in the ANN. In this chapter, the most important Python libraries used to implement this work will be presented. In addition to the mechanism of data collection, processing, and implementation of the proposed methodologies.

3.2. Basic Principles of ANN, GA, and GWO

In this section, ANN and optimization methods based on GA and GWO are explained in detail, which are used in ANN optimization. Where GWO and GA provide optimization of the number of layers and neurons of the ANN based on performance metrics such as R2, MAE, RMSE, and MSE values of the training samples.

3.2.1. Artificial Neural Networks (ANN)

ANN is the model of information processing that imitates the operation of the organic nervous system in the human brain [53]- [54]. The origin of ANN concepts may be traced back to the previous century when they were first suggested as a solution to complex issues [55]-[54]. This period was notable for the introduction of the first McCulloch-Bates neural model in 1943[54]. The ANN has become increasingly popular in various fields, with many institutions using it to tackle challenges in different sectors of the economy and human activities that were traditionally handled by operations research [56]-[53]. ANN is notable for its ability to be applied in scientific and engineering fields[53]. In addition, the flexibility of

ANN makes it highly relevant in the field of data analysis, leading to increased interest in using it for energy prediction. However, this application relies on having reliable data to estimate output functions that fit with real-world situations.

The ANN architectures can be classified into two primary categories: feedforward networks (FFNN) and feedback networks as shown in Figure (3-1). Out of these options, feedforward networks are more commonly used since they are more efficient in terms of memory utilization during operation[11]. They have shown significant resilience in managing nonlinear systems such as solar arrays. The FFNN can be categorized into four types: single-layer, multi-layer, radial function networks, and Bayesian regularized neural networks (BRANN). Each of these networks can utilize distinct learning algorithms. Out of these options, the multilayer ANN is the most frequently used variety since it can modify the weights of the hidden layer [21]. A multilayer feedforward ANN usually consists of three layers: input, hidden, and output, as shown in Figure (3-2). The neurons in each layer are connected through weights and bias terms from the previous layers. Equation (3-1) can be used to formally represent this interrelated processing structure.

$$y_i = \sum_{i=1}^{m} w_{ij} x_i + b_i \qquad (3\text{-}1)$$

Where the variable xi represents the input training node, whereas the connection weights (w_{ij}) are associated with the input nodes, hidden nodes, and layer nodes. The bias(b_i) pertains to the concealed nodes and output layer nodes, whereas m denotes the quantity of input signals.

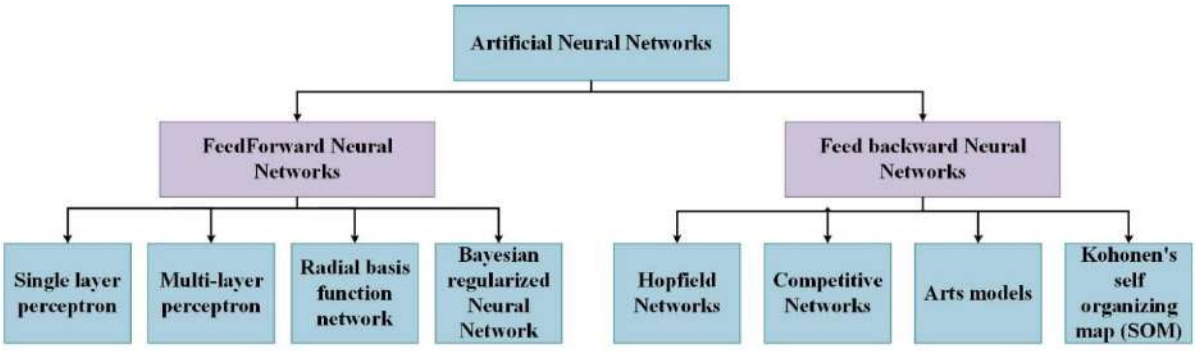


Figure 3-1 Review framework for artificial neural networks classification.

The development of FFNN has two main challenges: The first challenge optimizing the architecture of the NN by determining the optimal number of hidden layers and neurons. The second challenge improving the initial weights of the training nodes. This thesis aimed to improve the best arrangement the number of layers and neurons. In this thesis, the model is trained by inputting the sun's radiation and temperature information using the Keras library in Python. The model is compiled with (Adam) optimizer and mean squared error (MSE) as the loss function. In addition, the interconnections between neurons in each layer are established using weights from other neurons and bias terms (b) from preceding layers. Random weights are assigned to nodes to enhance the model's performance.

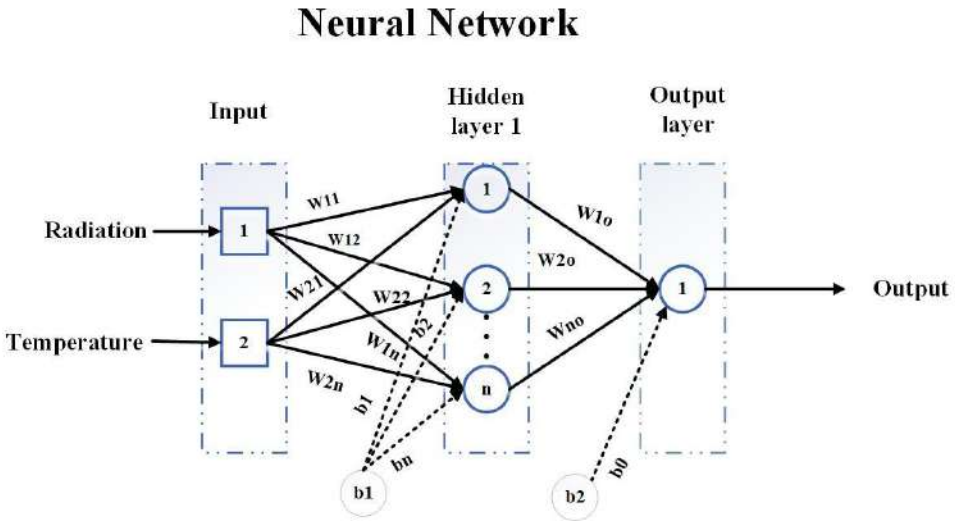


Figure 3-2 Neural Network Structures

3.2.1.1. Activation Function

In ANNs, most neurons conduct nonlinear computations akin to their biological counterparts. These neurons typically exhibit point-like behavior, governed by a single nonlinear activation function, denoted as f(x), which connects the input summation to the output activity. This nonlinearity is pivotal for the efficacy of ANN[57]. A widely adopted activation function is the Rectified Linear Unit (ReLU), represented as $f(x) = \max(x, 0)$ [58]. The derivative of ReLU at $x = 0$ is technically undefined but conventionally set to 0 in practical applications[59]. ReLU and its variations are commonly employed in feedforward networks, whereas the hyperbolic tangent (tanh) function is preferred for recurrent networks[60].

3.2.1.2. Hidden Layer Size

In the design process optimizing the hidden layers and neurons in the feedforward ANN is crucial. Therefore, this improvement aims to find a balance between computational efficiency and the model's ability to accurately represent the data distribution [[61]–[62]]-[63]. If the hidden layers of an ANN contain too many units, this can lead to significant computational problems and possible overfitting or underfitting in the regression model. On the other hand, having too few units in the hidden layers may result in faster computations but a suboptimal linear regression fit. Traditionally, determining the appropriate size of hidden layers has involved a trial-and-error approach. However, this method is considered inadequate due to its long execution time[11].

3.2.2. Genetic Algorithm (GA)

Throughout history, nature has consistently served as a profound wellspring of inspiration for humans. John Holland initially proposed the optimization method known as (GA)[64] and later acquired prominence through the research conducted by David Goldberg in 1989. The concepts of natural selection and genetics influence GA. The method has shown success in resolving optimization problems and has proven to be efficient in intelligently exploring a large and complex search space[65]. The main purpose of a GA is to create and manipulate several people using appropriate genetic operators to discover the solutions. Therefore, GA is categorized as a global search method that relies on the concept of gathering multiple solutions rather than relying on a single solution[66]-[67] The GA commences by initializing a population consisting of a collection of solutions represented as chromosomes. A new population is formed by extracting solutions from an existing population[68]. New solutions (offspring) are chosen based on their fitness, with the more suited solutions having a higher chance of reproducing.

The genetic optimization process follows several stages to generate a new population. Initially, a population size of 10 chromosomes is set, with a crossover probability of 0.8 and a mutation probability of 0.1, allowing for a maximum of 5 iterations. In the selection phase, fitter individuals are given higher chances to reproduce. Crossover involves the exchange of genes between two parents to produce offspring, potentially with higher fitness. Mutation modifies genes within a chromosome to increase diversity and prevent local optima. After mutation, the fitness value of each individual is calculated, and the individual with the lowest fitness value is considered optimal. The process iterates until a stopping criterion is met, such as reaching the maximum number of iterations or achieving an improved solution. The life cycle of GA progresses through various stages, beginning with

population initialization, followed by selection, crossover, mutation, and ultimately the termination condition, as depicted in Figure (3-3).

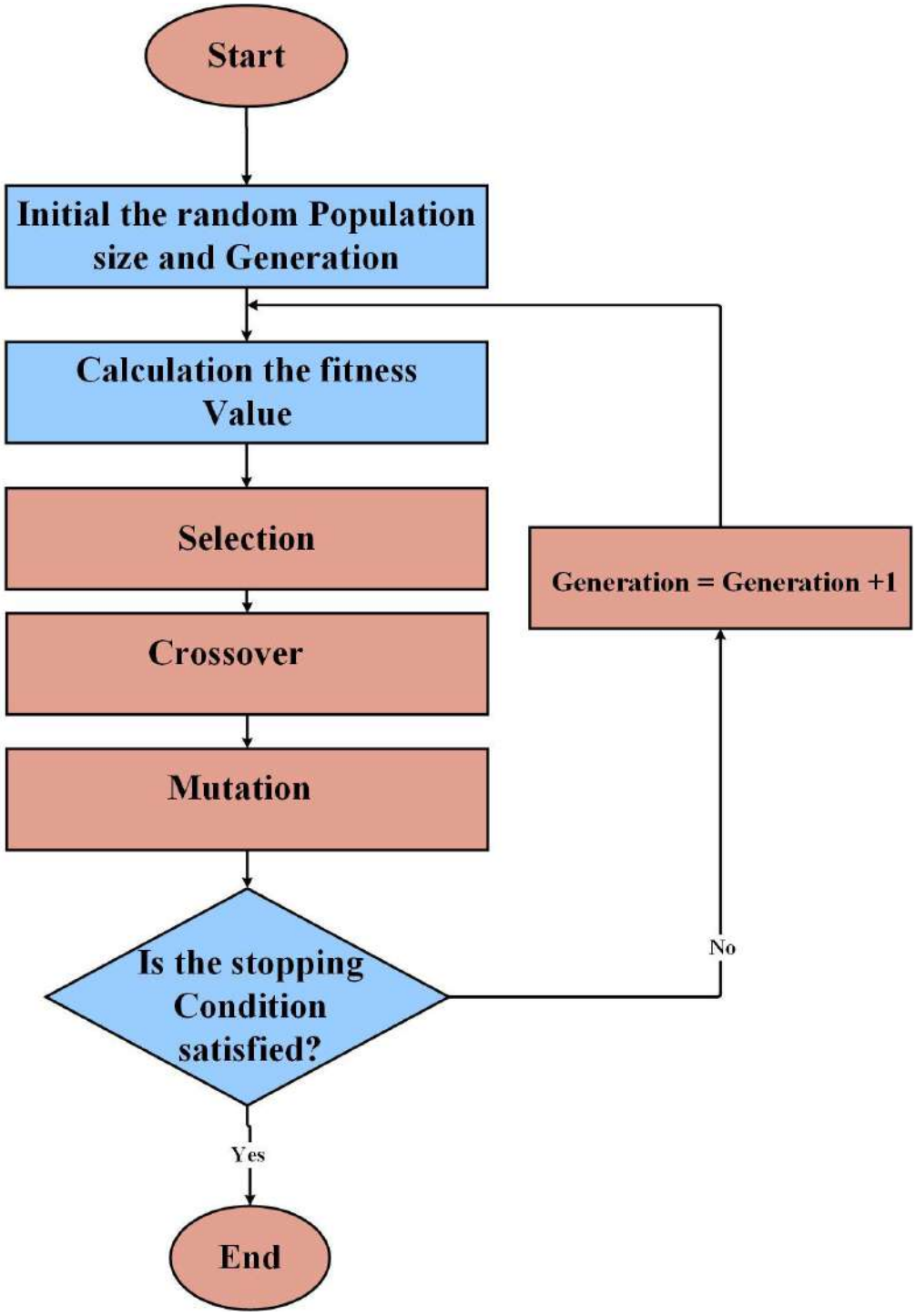


Figure 3-3 The flowchart of GA

3.2.3. Grey Wolf Optimization (GWO)

The GWO algorithm, which was introduced by Mirjalili et al. in 2014, is designed to replicate the social structure and hunting behavior observed [69] and to address optimization problems through the simulation of the hunting behavior of gray wolves. In this model, the position of each wolf is considered as a potential solution to the optimization problem. Grey wolves live primarily in social groupings, with an average group size ranging from 5 to 12 individuals[70]. The social hierarchy assumes a pivotal part in the act of hunting as shown in Figure (3-4)[71]. Thus, the population can be categorized into four different layers, namely alpha (α) (the main and dominant leader who is responsible for decision-making, i.e., the best manager who exerts the most effective control over the group), and beta (β) (alpha's adviser and are subordinate wolves who assist the alphas in group decision-making and other duties), delta wolves (submit to alpha and beta) and omegas (the lowest in the hierarchy)[72]- [73].

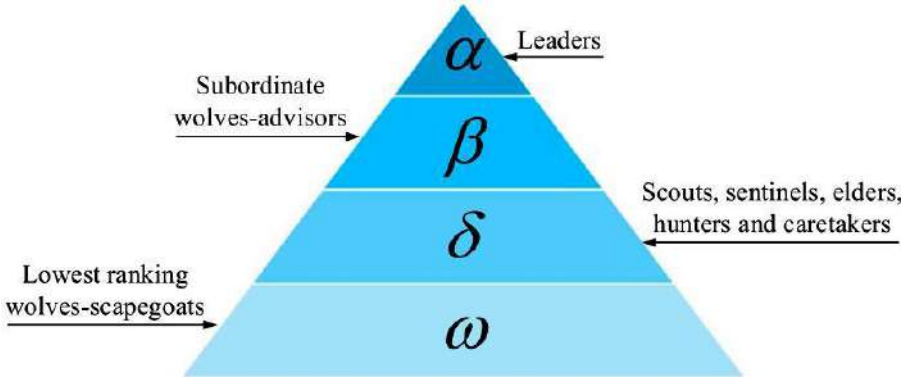


Figure 3-4 Hierarchy of grey wolf population[71]

The primary stages of grey wolf hunting are as follows:1) Tracking, chasing, and moving towards the prey. 2) Pursuing, and persistently bothering the prey until it ceases all movement. 3) Initiate an assault on the prey. The steps are depicted in Figure (3-5), [69]. To simulate the hunting behavior of gray wolves [72]., the three

most optimal solutions, namely alpha, beta, and delta, are selected based on the assumption that these wolves possess superior knowledge regarding the prospective location of their prey. Subsequently, it is guaranteed that additional search agents adjust their placements based on the position of the optimal search agent[73]. Figure (3-6) illustrates the process by which a search agent adjusts its position in a two-dimensional search space based on the values of alpha, beta, and delta. The final position will be randomly located within a circle determined by the positions of alpha, beta, and delta in the search space. Alpha, beta, and delta wolves use their sensory input to approximate the location of the prey, while the remaining wolves randomly adjust their positions in the vicinity of the prey [73]**.** The steps of the GWO Algorithm for addressing optimization issues are depicted in Figure (3-7).

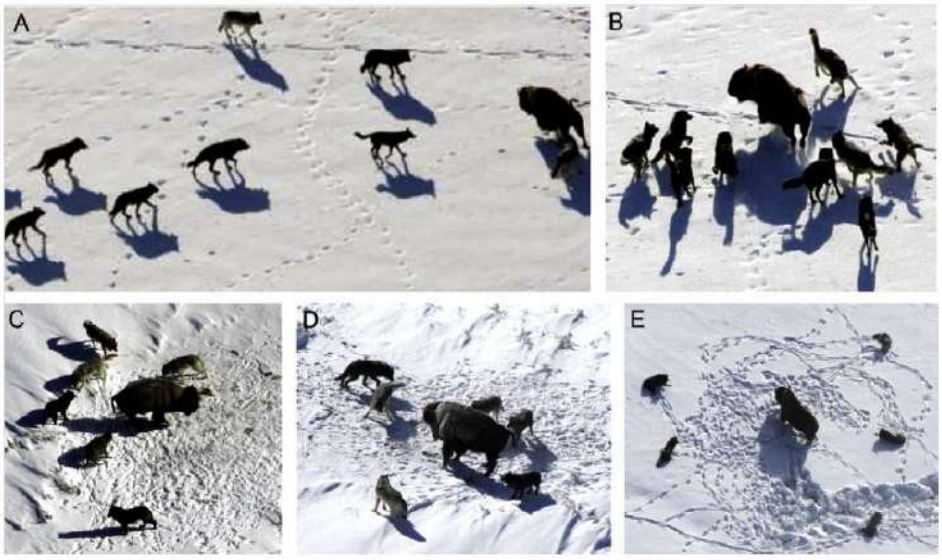


Figure 3-5 The hunting behavior of grey wolves[69]**.**

The GWO algorithm has several different definitions, which are performed using the following equations. Equations (3-2) and (3-3) are utilized to represent the prey surrounding behavior exhibited by gray wolves[73].

$$\vec{D} = \left|\vec{C} \cdot \vec{x}_p(t) - \vec{x}(t)\right| \qquad (3\text{-}2)$$

$$\vec{x}(t+1) = \vec{x}_P(t) - \vec{A} \cdot \vec{D} \qquad (3\text{-}3)$$

In the given equations, the variable t denotes the quantity of current iterations. The vectors $\overrightarrow{A}$, and $\vec{C}$ are, used to represent the coefficients[73]. The position vector of the prey is denoted as $\vec{x}_p$, whereas the position of a grey wolf is represented by $\vec{x}$ [73]. The vectors $\overrightarrow{A}$, and $\vec{C}$ are computed utilizing Equations (3-4) and (3-5). [73] The value of $\vec{a}$ is linearly lowered from 2 to 0 during the iteration, while $\vec{r}_1$ and $\vec{r}_2$ represent random vectors within the range of [0, 1].

$$\vec{A} = 2\vec{a} \cdot \vec{r}_1 - \vec{a} \qquad (3\text{-}4)$$

$$\vec{c} = 2 \cdot \overrightarrow{r_2} \qquad (3\text{-}5)$$

In order to simulate the hunting behavior of gray wolves, the three most optimal solutions, namely alpha, beta, and delta, are selected based on the assumption that these wolves possess superior knowledge regarding the prospective location of their prey. Subsequently, it is guaranteed that additional search agents adjust their placements based on the position of the optimal search agent. These processes are performed using the following equations[72].

$$\vec{D}_a = \left|\vec{C}_1 \cdot \vec{x}_a - \vec{x}\right|, \quad \vec{D}_\beta = \left|\vec{C}_2 \cdot \vec{x}_\beta - \vec{x}\right|, \quad \vec{D}_\delta = \left|\vec{C}_3 \cdot \vec{x}_\delta - \vec{x}\right| \qquad (3\text{-}6)$$

$$\vec{x}_1 = \vec{x}_a - \overrightarrow{A_1} \cdot \vec{D}_a, \quad \vec{x}_2 = \vec{x}_\beta - \overrightarrow{A_2} \cdot \vec{D}_\beta \ , \vec{x}_3 = \vec{x}_\delta - \overrightarrow{A_3} \cdot \vec{D}_\delta \qquad (3\text{-}7)$$

The three optimal solutions, $\vec{\boldsymbol{x}}_{\boldsymbol{a}}$ the best search agent, $\vec{\boldsymbol{x}}_{\boldsymbol{\beta}}$ the second-best search agent, and $\vec{\boldsymbol{x}}_{\boldsymbol{\delta}}$ the third best search agent, are chosen to simulate gray wolf hunting behavior, with additional search agents adjusting their positions based on the optimal search agent's position using equations (3-8).

$$\vec{x}(t+1) = \frac{\vec{x}_1(t) + \vec{x}_2(t) + \vec{x}_3(t)}{3} \qquad (3\text{-}8)$$

In the given equations, $\vec{A}$ represents a stochastic value inside the interval [-2a, 2a] [73]. When $|\overrightarrow{A}|$ is less than 1, the grey wolves are compelled to attack their

victim, however, when $|\vec{A}|$ is greater than 1, the grey wolves are compelled to distance themselves from the prey to locate a more suitable prey. After satisfying a termination requirement, the gray wolf optimization method is concluded[73].

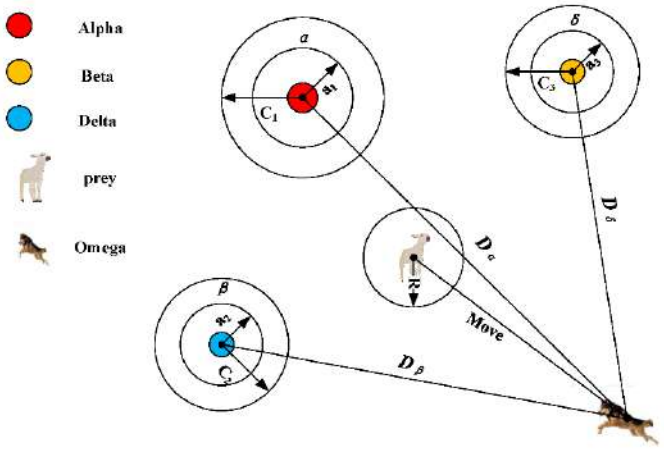


Figure 3-6 Graphical abstract of GWO[73]

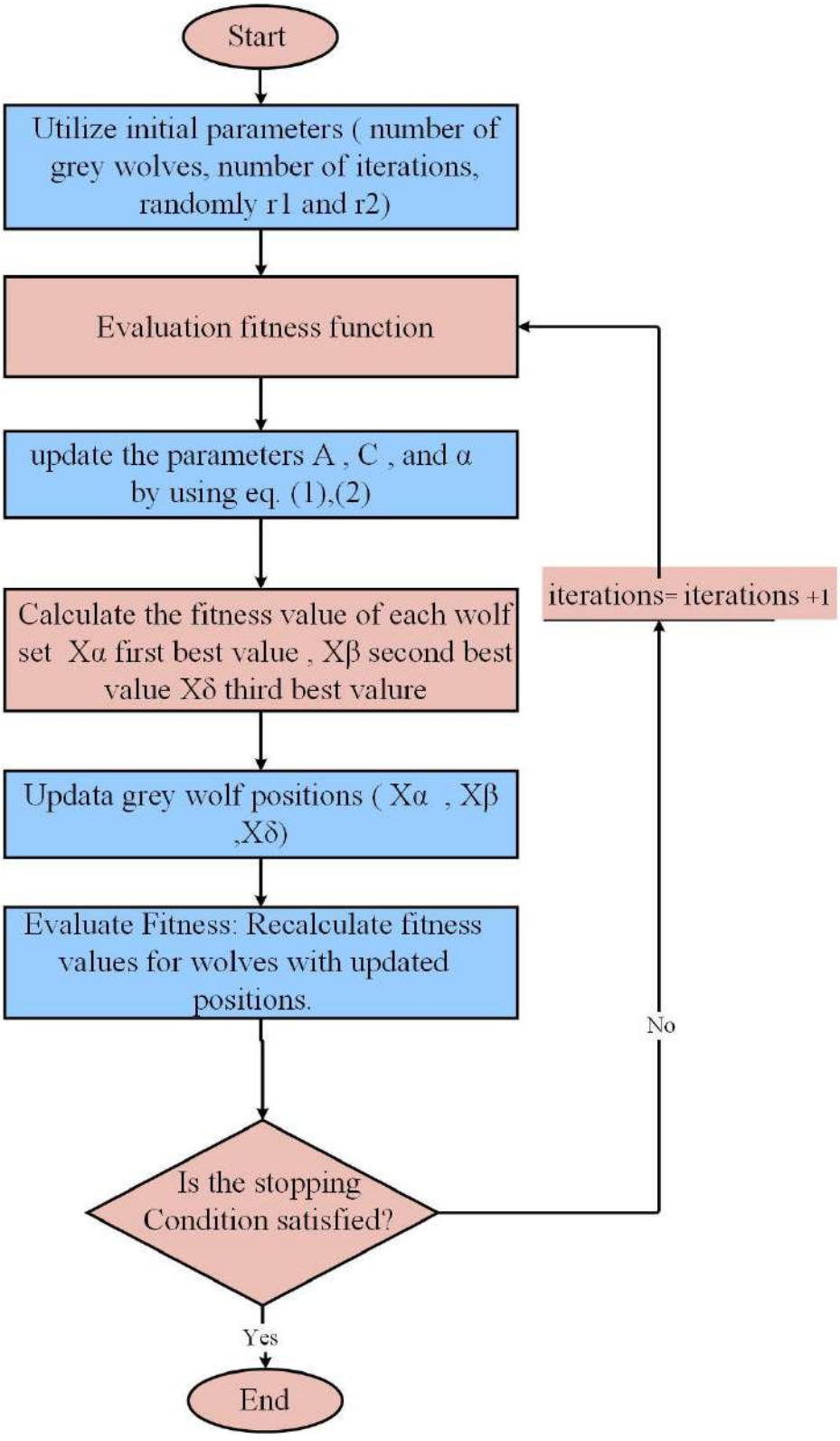


Figure 3-7 The flowchart of grey wolf optimization

3.3. Python

Python, developed by Guido van Rossum in the late 1980s, has become a fundamental programming language in AI due to its broad application, especially in machine learning and neural network development. Key libraries like TensorFlow, Keras, PyTorch, Matplotlib, Seaborn, Pandas, and NumPy support a wide range of tasks from data manipulation and visualization to building complex predictive models. In this thesis, Google Colab was used as the coding environment, leveraging its free access to GPU and TPU resources and pre-installed packages to streamline model creation. Key libraries include Matplotlib for visualizations, Seaborn for statistical graphics, Pandas for data manipulation, and NumPy for numerical processing. TensorFlow is specifically used to build prediction models, leveraging machine learning to handle large datasets effectively. The process of building models in Keras includes defining network layers, compiling the model with loss functions and optimizers, fitting it to data, evaluating accuracy, and making predictions. Together, these tools create a cohesive environment for data analysis, visualization, and the development of AI-driven models.

3.4. Proposed Methodological Methods

Determining the number of hidden layers and neurons is an important step for implementing an ANN model because the low MSE value can be learned quickly. Therefore, in this work, the number of hidden layers and neurons was determined based on the GA and GWO algorithms. Figure (3-10) shows the proposed GA optimization framework to find the optimal structure (number of hidden layers and neurons) of the ANN. Figure (3-11) shows the proposed framework of the GWO model which is used to find the optimal structure (number of hidden layers and neurons) of ANN.

Table 3-1 shows the constant parameter settings for the models used.

Parameters	Description
Activation function	ReLU
Number of inputs	Solar radiation and temperature
Number of outputs	1
Maximum epochs	1000
Number of Iteration	5
Number of populations	10
Optimization method	Adam
Layer Number of NN	1
Hidden Neurons Number of NN	10
Mutation rate	0.1
Crossover rate	0.8
loss function	MSE
Number of wolves	10

3.4.1. Theoretical Data Processing and Analysis

In this section, the data are pre-processed by extracting 7 specific columns from the Excel file, which are date, time, solar radiation, temperature, voltage, current, and actual power. The objective of this pre-processing is to improve the quality and relevance of the data. In the filtering step, any anomalous data is eliminated, data collected at nighttime are discarded, and only data collected during daytime is considered, as the solar panels function based on the sunlight they get. Effective data preparation is crucial for constructing precise prediction models as historical data may contain erroneous information, which can negatively impact the performance of the model. As a result, the data is processed and missing data is removed. Additionally, if any errors are detected, the procedure is repeated to ensure the

accuracy of the data, as shown in Figure (3-8).

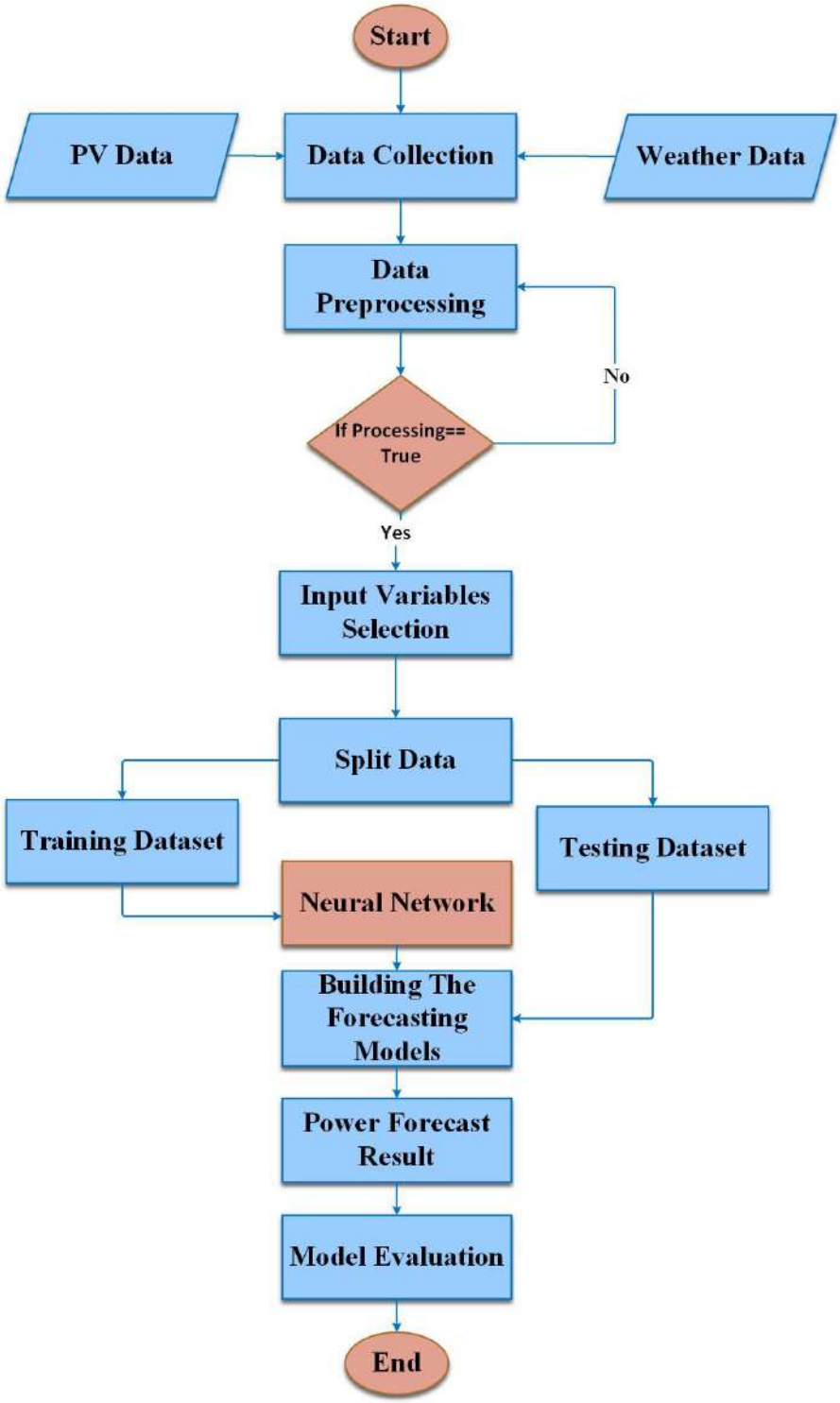


Figure 3-8 The framework for the suggested approach of predicting PV power accurately using Python language.

The Pandas package in Python is utilized to parse Excel files and provide the data in a tabular format, including the columns and rows. When completion of the processing, the data is divided into a training set and a test set, which is an essential initial stage in constructing a machine-learning model. The objective of this stage is to divide a part of the data for training the model while reserving another part for testing and assessing its performance on unseen data. The data is divided into separate sections using manually defined indexes. Where the function range (A1, A2) in Python is utilized to generate a list that includes numbers, such as those ranging from A1 to A2. These numbers indicate the indices of the rows that will be allocated to the test set. The train_data includes rows that are outside the specified test_indices. After partitioning the data into a training set and a test set, feature variables, and a goal variable are generated for both sets. The solar radiation and temperature data (third and fourth columns) are saved in the training and testing sets as variables named X_train and X_test. The variables y_train and y_test store the last column, which indicates the actual power.

In the data analysis phase of this thesis, four specific months are chosen, with each month reflecting one of the seasons in Iraq. January was selected as the symbol of the winter season, March as the symbol of spring, July as the symbol of summer in Iraq, and September as the symbol of autumn. The training of each class was conducted independently, and the results were extracted and will be addressed in Section 4.5. Subsequently, the monthly dataset was divided into training and testing subsets. The training subset consisted of 60% of the data, representing samples from the initial portion of the month. Conversely, the testing subset comprised 40% of the data, representing the data from the latter part of the month.

After the divided data then trained using an FFNN that includes various layers, namely the input layer, hidden layers, and output layer. The model is trained using

a dataset of sun radiation and temperature as input. The process includes developing and instructing an ANN model using the TensorFlow and Keras libraries. The neural model is defined as a series of layers using the Sequential () function. Following this, an initial hidden layer with 10 units is included. The Rectified Linear Unit (ReLU) activation function, commonly employed in FFNN, incorporates a single unit into the output layer. Moreover, the connectivity between neurons in each layer is established by the weights of other neurons and bias terms (b) in the preceding layers. The weights of the nodes are randomly given to enhance the performance of the model. Subsequently, the loss function MSE is selected, followed by the utilization of the Adam optimizer. The optimizer is highly regarded for its ability to efficiently and effectively train deep learning models due to its combination of rapid adaptation and stability. "Adaptive Moment Estimation" is abbreviated as "Adam". The penultimate step is to train the model using training data for 1000 epochs. Subsequently, the R2, MAE, RMSE, and MSE metrics are computed, followed by a comparison between the actual and predicted values.

3.4.2. Experimental Data Processing and Analysis

Forecasting power generation in solar power plants is crucial for enhancing the control and distribution efficiency of the plants and distributed the secure and steady operation of the power grid. The data were collected at the University of Misan College of Engineering building, and the data collecting methodology and tools employed were elucidated in Chapter Three. Once the data collection is finished, it is examined and processed. In this section, the data is pre-processed by extracting seven specific columns (date, time, solar radiation, temperature, current, voltage, and actual power) from the Excel file. Abnormal data is removed through pre-processing, collected data at nighttime is discarded, and only collected data during

daytime is considered, as previously mentioned, using Excel. The objective of data processing is to acquire datasets of superior quality, which is essential for constructing a precise predictive model. Conversely, the purpose of training is to provide the system with the power to utilize its knowledge of novel data, enabling it to make predictions.

After processing, the data samples are categorized into three groups according to weather patterns: cloudy, rainy, and sunny. These categories are spread across three seasons, in 2024. It should be noted that the autumn season has not yet occurred in Iraq and is expected to arrive in late September. Weather types were classified by sampling from two days in each season. The Python Pandas library is utilized for the analysis Excel files and for presenting data in a structured fashion, comprising columns and rows. Next, the data is divided into a training set and a test set. The function alternating_indices in Python is defined to divided the data. This function accepts two parameters: the number of samples, represented by the variable n, and the training ratio, represented by the variable train_ratio. A training percentage of 60% was selected in this thesis. Subsequently, a collection of indicators is generated, with indicators that span from 0 to n-1. Afterward, two lists, train_indices, and test_indices, are initialized to store indices for training and testing. Ultimately, the data is traversed in increments of 20 steps, and pointers are added to the two lists based on the computed blocks.

Figure (3-9) displays the most recent PV and meteorological data utilized throughout the testing process. The acquired dataset is examined and categorized to determine the precise meteorological conditions that are pertinent to a specific day. The sampling occurs at regular intervals of exactly 3 minutes, lasting for approximately 12 hours every day. The frequency of sampling is adjusted based on the quantity of radiation detected or the amount of sunshine that reaches the solar

panels. There is a noticeable reduction in the length of daylight hours on some days, particularly in January. Consequently, there is some unmeasured data, especially during periods of heavy rainfall or when there are maintenance works inside the Electrical Engineering College building, which hinders the correct collection process. In addition, during the heavy rainfall case, the data readings became unavailable due the radiation is very low and the radiation sensor was unable to read it. Therefore, these measurements are removed from the data set in the absence of solar radiation data. Subsequently, the input variables are introduced into the respective ANN model, contingent upon the specific meteorological conditions. The evaluation of the prediction model is conducted using metrics such as MAE, RMSE, R2, and MSE.

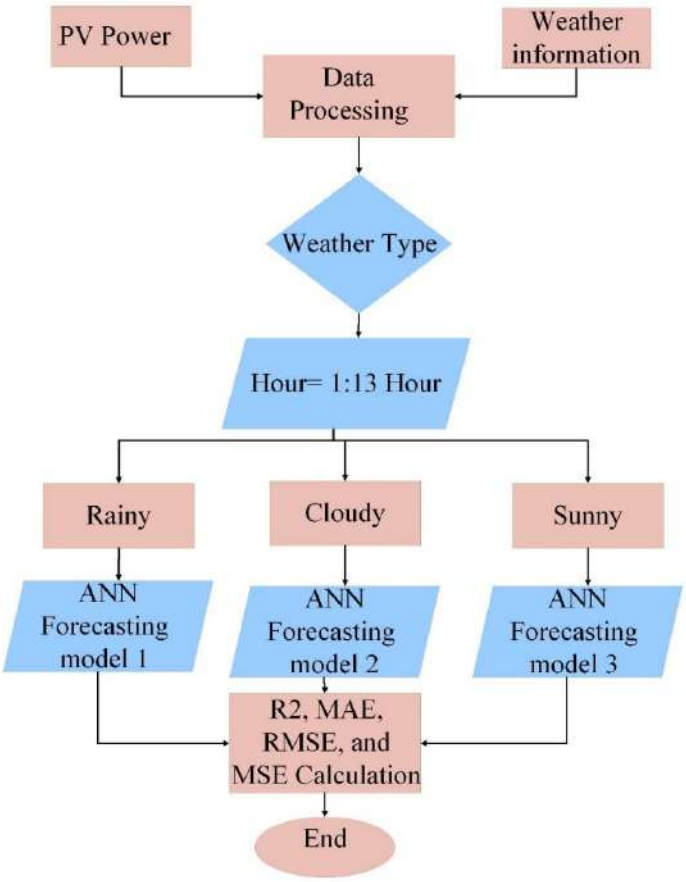


Figure 3-9 Testing stage

3.4.3. A Framework Of NN-GA

A GA is a worldwide multi-agent optimization algorithm that relies on biological evolution, and due to its robust exploration capabilities and high performance, GA has been employed to identify the optimal topology for ANN. In this section, the ANN architecture will be fine-tuned by utilizing a GA to determine the best number of hidden layers and neurons, as explained in Figure (3-10). This procedure will be carried out after acquiring the data, performing preprocessing tasks, and splitting the data, as detailed in the current chapter. The method commences by defining parameters, including the population size and the number of models to be generated in each generation. Initially, the population size is set to 10. The mutation rate is determined by selecting the value of 10% for the occurrence of a mutation in the model. Subsequently, the crossover rate was established at 80%, representing the likelihood of hybridization taking place between the two models. Finally, the number of iterations is defined, and a total of five iterations are selected. The permissible range for the number of hidden layers in the ANN is between 1 and 5 layers, while the range for the number of neurons in each layer is between 10 and 128 neurons to avoid overfitting. Then, this thesis utilized the performance measures R2, MAE, RMSE, MSE, and RE as the fitness function.

The mutation function generates a new model by inserting random values into the original model. These mutations occur as random alterations in the number of neurons within the dense layers of the NN. The crossover function generates a progeny model by randomly picking neurons from two-parent models. A few individuals with exceptional physical fitness are chosen for direct genetic transmission to the following generation. Subsequent, hybridization and mutations are carried out using the existing population in order to produce the next generation. Every model undergoes training using the training data and is then assessed using a

fitness evaluation function. If the new model surpasses the current optimal model, the optimal model is updated. This process continues until the specified number of iterations is reached, and the final findings are then displayed. The code can be found in Appendix D.

3.4.4. A Framework Of ANN-GWO

To enhance the accuracy of the ANN model, a hybrid approach is employed by combining the GWO algorithm with ANN. Once the data processing and division procedure is finished, the Python libraries are specified to carry out the training process and extract the results. Subsequently, a function is established to construct the ANN model, taking into account the specified quantities of hidden layers and neurons, as explained in Figure (3-11). The ANN-GWO approach requires the specification of two parameters: the number of wolves and the number of iterations, as indicated in Table (3-1). The remaining parameters are assigned randomly. During the initial stage, the ANN undergoes training utilizing the GWO algorithm. The GWO algorithm utilizes particular functions called "Initialize Wolf" and "Update Position" to represent and adjust the positions of wolves. The algorithm stochastically assigns the positions of grey wolves within the defined intervals of the hidden layer count and neuron count. The positions of the wolves are updated using the GWO equations, which consider the three most optimal positions of the wolves (alpha, beta, delta). The GWO algorithm is utilized to determine the optimal configuration of hidden layers and the number of neurons. The process is iterated a fixed number of times, and in each iteration, each wolf is assessed and their places are adjusted according to the superior wolves. Ultimately, the ultimate model is trained using the optimized parameters. Subsequently, the model is assessed by employing the test suite and quantifying its performance using metrics such as MSE, MAE, RMSE, R^2, and RE. The code can be found in Appendix E.

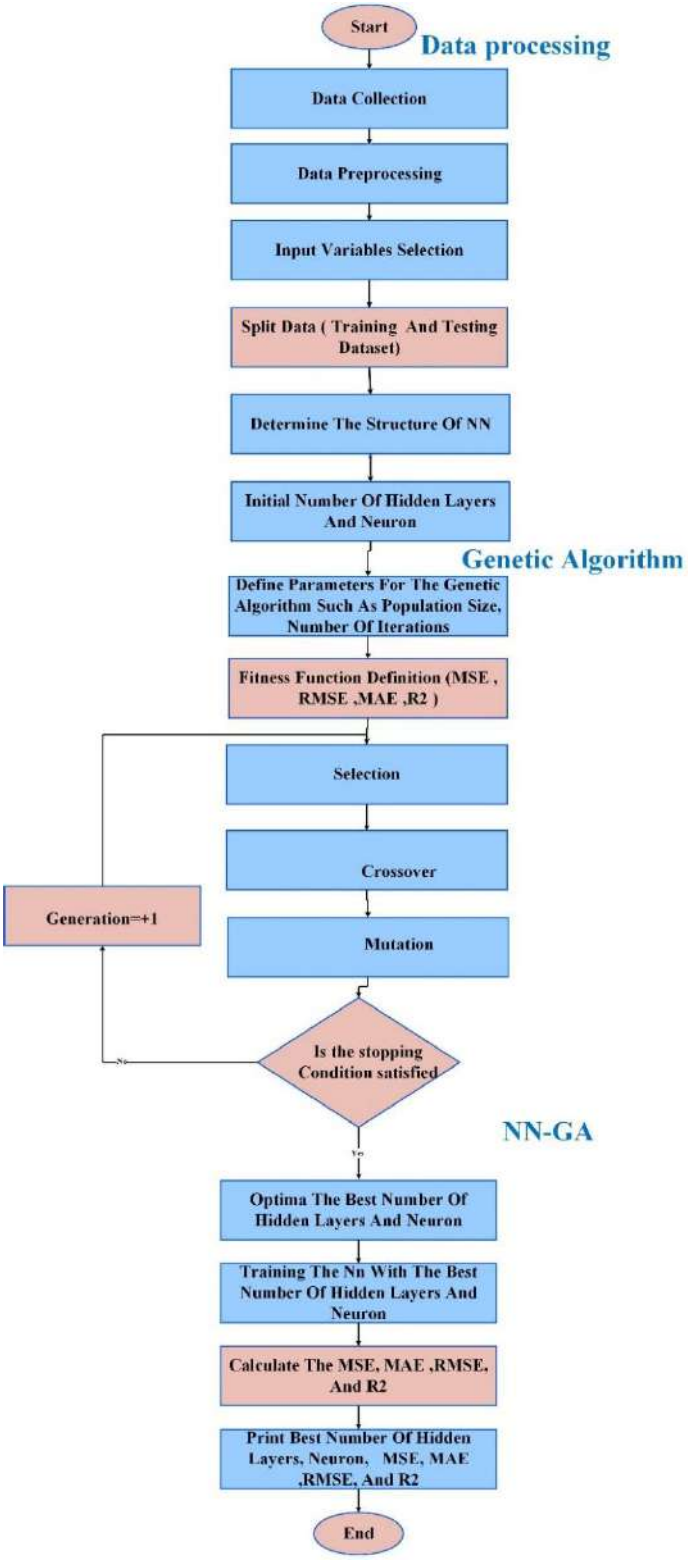


Figure 3-10 A Framework Of GA-NN For PV Power Forecasting

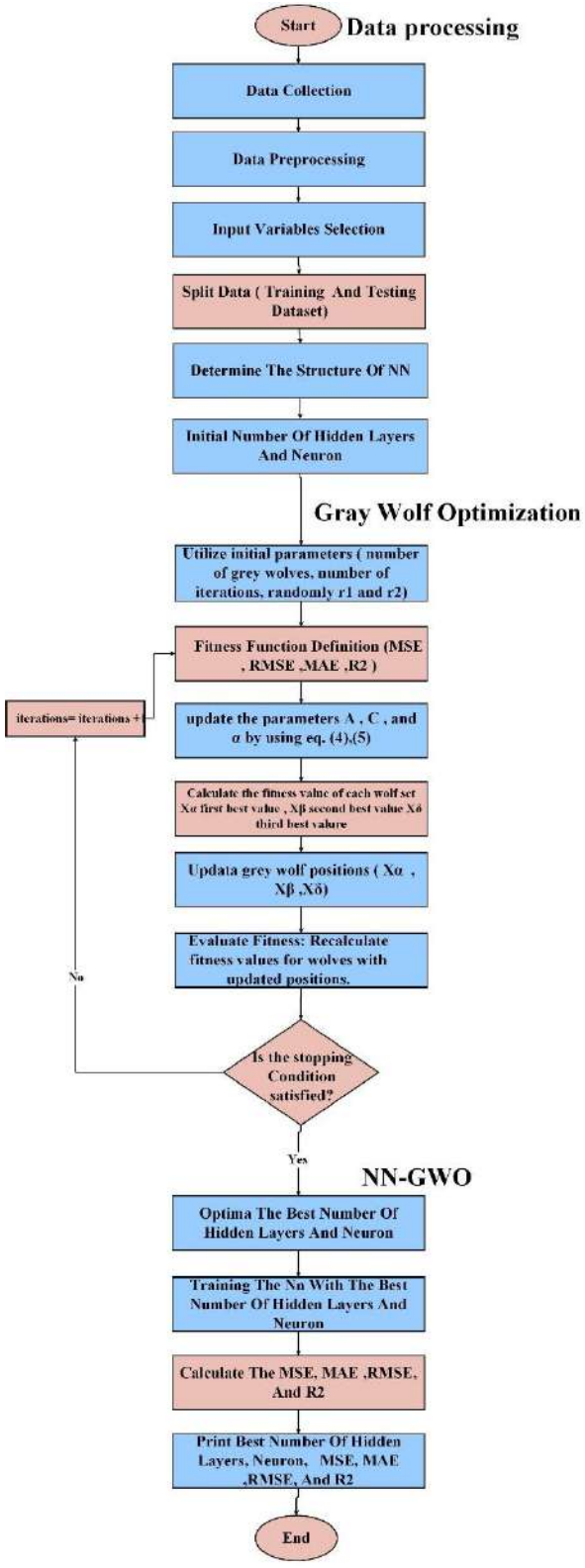


Figure 3-11 A Framework Of GWO-NN For PV Power Forecasting

3.5. Summary

The thesis initially outlines the structure and fundamental concepts behind three primary techniques used: the ANN, the GA, and the GWO algorithm. Each of these techniques plays a specific role in improving the accuracy and adaptability of the predictive model for PV energy generation. The ANN is introduced as a machine learning method that mimics the human brain's structure to process complex patterns. The GA is described as an optimization method inspired by natural selection, aimed at enhancing the ANN by identifying the optimal configuration for hidden layers and neurons. Similarly, the GWO algorithm, inspired by the social hierarchy and hunting behavior of gray wolves, is presented as another approach to optimize the ANN, improving its performance in handling diverse weather conditions. The thesis then provides an overview of the Python libraries and tools essential to this study. Libraries such as keras, Pandas, and TensorFlow are employed for data manipulation, model training, and evaluation, ensuring efficient handling of both theoretical and experimental datasets. In the methodology section, a detailed explanation of the data preprocessing steps is provided. This includes dividing the theoretical and experimental data into distinct categories based on weather conditions and seasons. The thesis then describes the process of integrating ANN with GA and GWO. Both optimization techniques are used to refine the ANN by adjusting the number of hidden layers and neurons, thereby improving the model's accuracy in predicting energy production across different meteorological conditions. This optimization process enables the ANN to adapt to fluctuations in weather, resulting in a more reliable forecast model.

CHAPTER FOUR

Theoretical and Experimental Results

4.1. Introduction

This chapter presents the theoretical results during January, March, July, and September within one hour for 2021. The presented results cover the effect of using monthly data on the accuracy of solar panel performance and the experimental results during different weather conditions rainy, cloudy, and sunny within 3 minutes for 2024. The presented results cover the effect of using different weather conditions on the accuracy of solar panel performance by using the methods described above. A comparison of the theoretical and experimental results from the current research will be covered in the last part.

4.2. Results of The Theoretical Part

This section forecasts PV power generation using weather data from January 1 to December 31, 2021, with hourly intervals. Since solar energy generation in Iraq is influenced by seasonal and climatic variations, data from four months January (winter), March (spring), July (summer), and September (autumn) was analyzed to predict PV energy output. The data was initially divided into separate training and testing sets for each month, then trained using a NN with randomly assigned weights via Python tools like TensorFlow and Keras. The models were evaluated against actual PV energy outputs. GA and GWO techniques were applied to fine-tune the number of hidden layers and neurons further to optimize the NN architecture.

Below are the results for those specific months:

4.2.1. January Forecasting Term

The purpose of this assessment is to evaluate the effectiveness of the suggested prediction model throughout the winter season. To provide a more accurate representation, January was chosen to represent the winter season. A total of 310 sample data tests were collected in January to develop the prediction model for the PV grid-connected system. The data were split into two sets: the training data, spanning from 8:00 am on 1 January 2021, to 11:00 am on 19 January, and the test data, covering the period from 12:00 pm on 19 January to 5:00 pm on 31 January.

Table (4-1) presents the MAE, RMSE, and MSE values from the ANN model, which are approximately 114.8937W, 150.6157W, and 22685.0952W, respectively. These metrics indicate a significant level of forecasting precision. Figure (4-1)(a) shows the MAE across training epochs, where a substantial decrease is observed during the first 600 epochs, from about 125W to 118W, with a slower decline from 118W to 114W as epochs increase from 600 to 1000, indicating the model is nearing optimal performance. Figure (4-1)(b) depicts the MSE, which drops notably during the first 400 epochs, continuing to decrease at a slower rate until stabilizing around 22685W from 600 to 1000 epochs. This suggests that the model is stabilizing and the rate of improvement is slowing. Figure (4-1)(c) shows the R-squared value, which increases gradually from 600 to 1000 epochs, reaching about 0.8246, signaling that the model is approaching its ideal performance. There were also slight declines in R-squared values for March (3.6%), July (0.77%), and September (10.79%) during the ANN model training. Figure (4-1)(d) shows the RMSE, which starts at approximately 163W and decreases to 150W as the epochs progress, suggesting that further training could enhance performance but must be monitored to avoid overfitting.

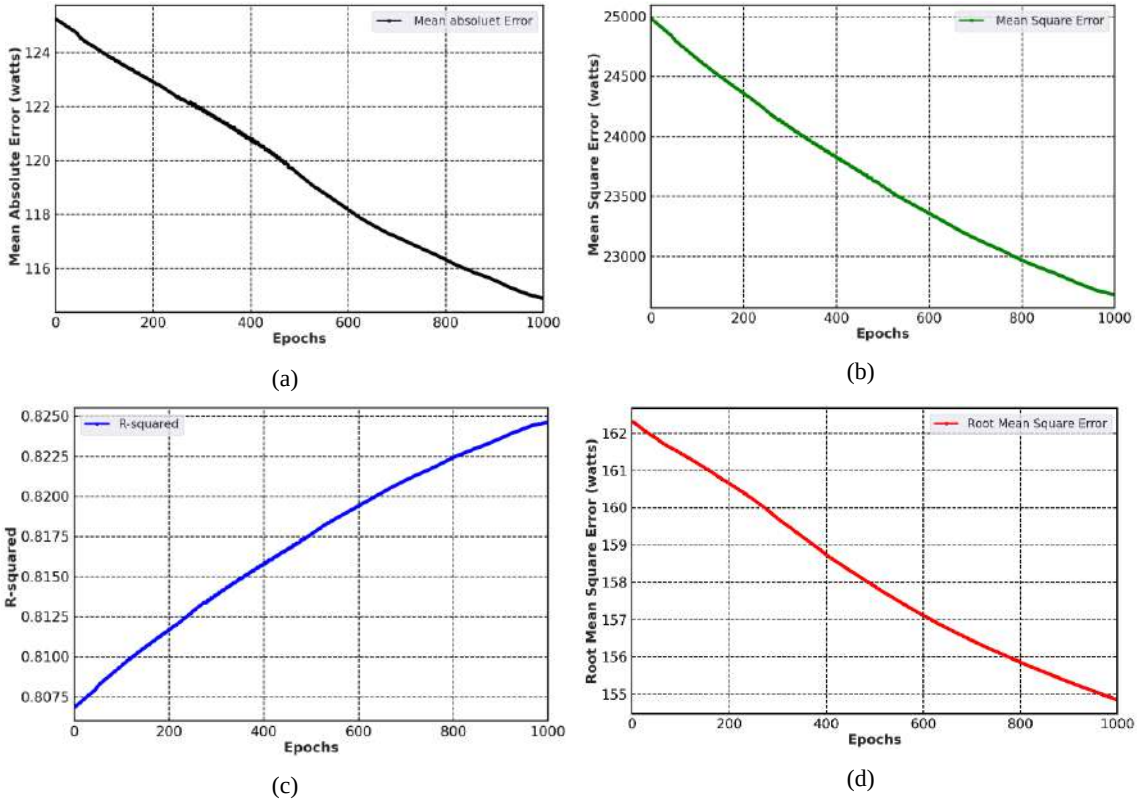


Figure 4-1 The results obtained only by employing neural networks in January (a) MAE, (b) MSE, (c)R2, (d) RMSE

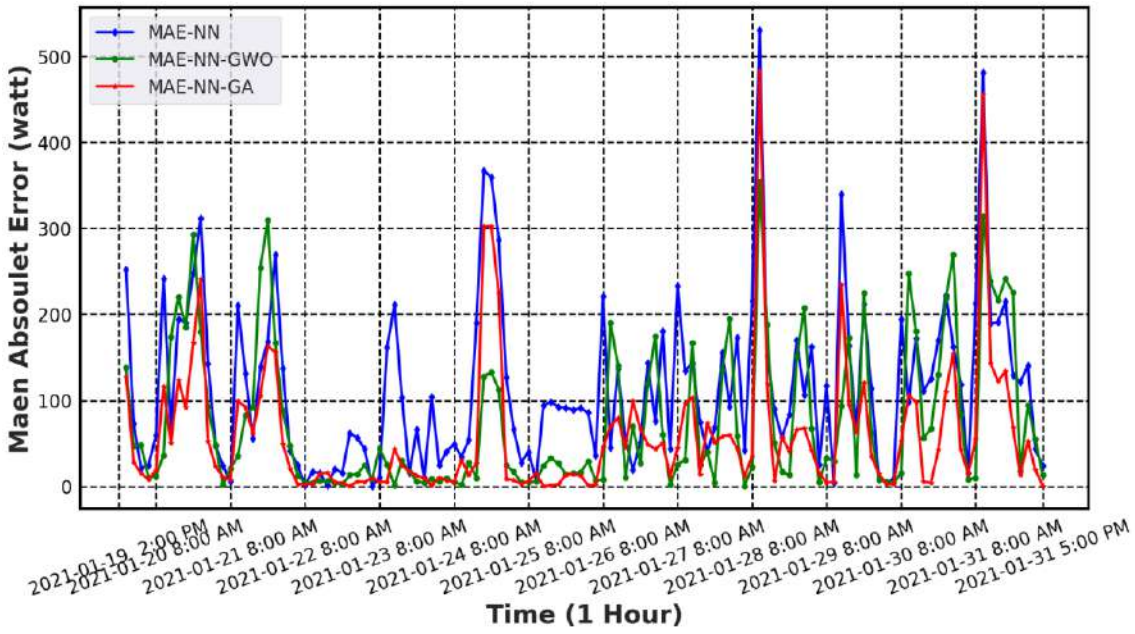
The data shown in Table (4-1) demonstrates that the suggested ANN-GA model surpasses both the ANN-GWO and ANN models in terms of performance, where the MAE values recorded by ANN-GA, ANN, and ANN-GWO are 59.4175 W, 114.8937 W, and 75.6923 W, respectively. Figure (4-2) (a) illustrates the MAE observed when comparing the predicted and actual power generation. The horizontal axis (X) represents time in hourly increments, ranging from 3:00 PM on 19 January 2021, to 5:00 PM on 31 January 2021. The Y-axis indicates the MAE in watts, ranging from 0 to 500 watts. The MAE-ANN-GA technique demonstrates superior accuracy in forecasting PV power generation, exhibiting the lowest MAE values and minimal volatility during the specified period. While the MAE-ANN-GWO approach has a favorable level of accuracy, which is higher than that of MAE-ANN

but lower than that of MAE-ANN-GA. This demonstrates that ANN-GA has a strong correlation with the actual power, hence indicating the efficacy of the GA in enhancing the accuracy of predictions. According to this analysis, the utilization of enhanced algorithms such as the GWO and GA yields greater precision and improved consistency in forecasting PV power generation, as opposed to employing an unimproved ANN.

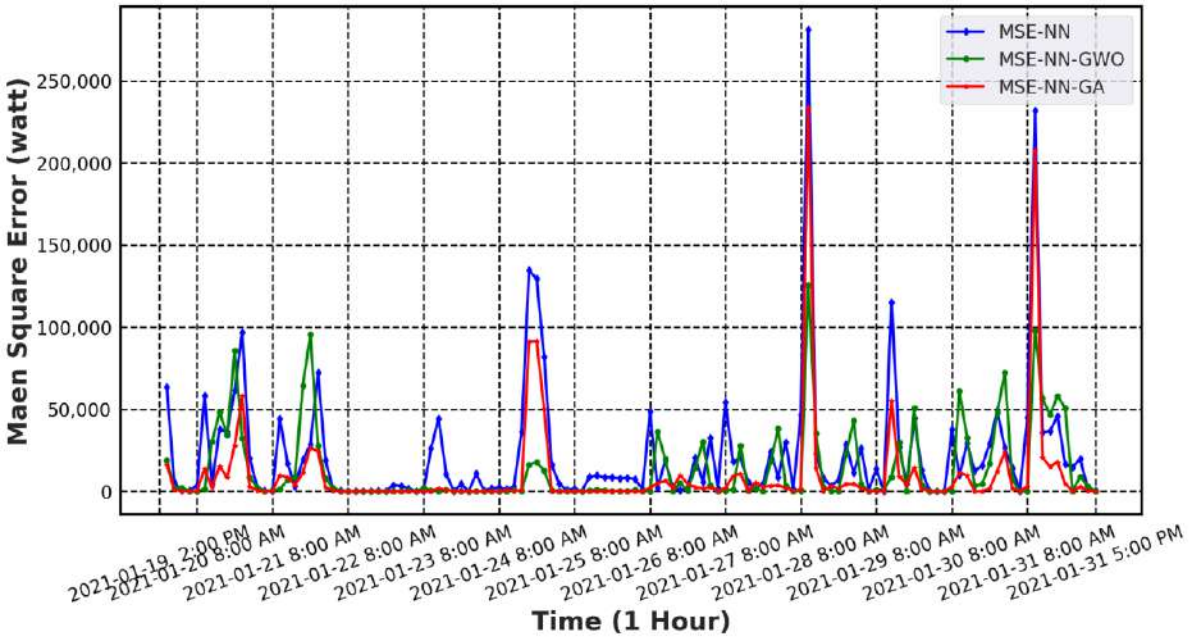
Figure (4-2) (b) displays the passage of time in hourly intervals on the horizontal axis (X), ranging from 2:00 PM on 19 January 2021, to 5:00 PM on 31 January 2021. The Y-axis displays the MSE in watts, with a range of 0 to 250,000 watts. Between the dates 2021-01-19 and 2021-01-21, the MSE-NN model exhibits significant variations, suggesting a lack of accuracy throughout certain periods. The MSE-NN-GWO model demonstrates a high level of stability, as evidenced by its low error levels. while the model MSE-NN-GA has better performance in comparison to other methods. Overall, the MSE-NN-GA method demonstrates superior accuracy in forecasting PV power generation. This is evident from Table (4-1). It also has the lowest MSE values and exhibits less variability over time. Specifically, the RMSE values for the NN, NN-GA, and NN-GWO methods are 150.6157, 99.8946, and 115.0495 watts, respectively. In addition, the R2 values for the ANN, ANN-GA, and ANN-GWO in Table (4-1) are 0.8246, 0.9228, and 0.8976, respectively, as indicated in Scheme (4-13) (a).

Table 4-1 Summarized forecasting results in January days for the PV prediction model.

Days	Method	RMSE	MSE	MAE	R^2
January	ANN	150.6157	22685.0952	114.8937	0.8246
	ANN-GA	99.8946	9978.9501	59.4175	0.9228
	ANN-GWO	115.0495	13236.4071	75.6923	0.8976



(a)

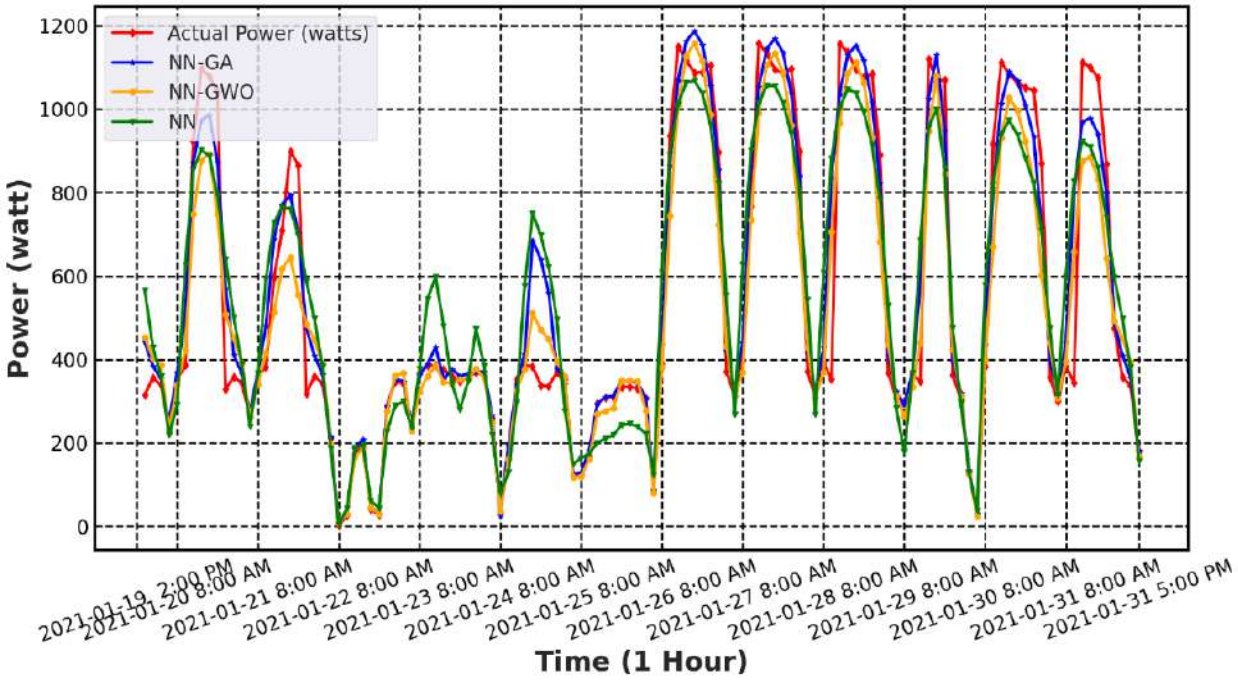


(b)

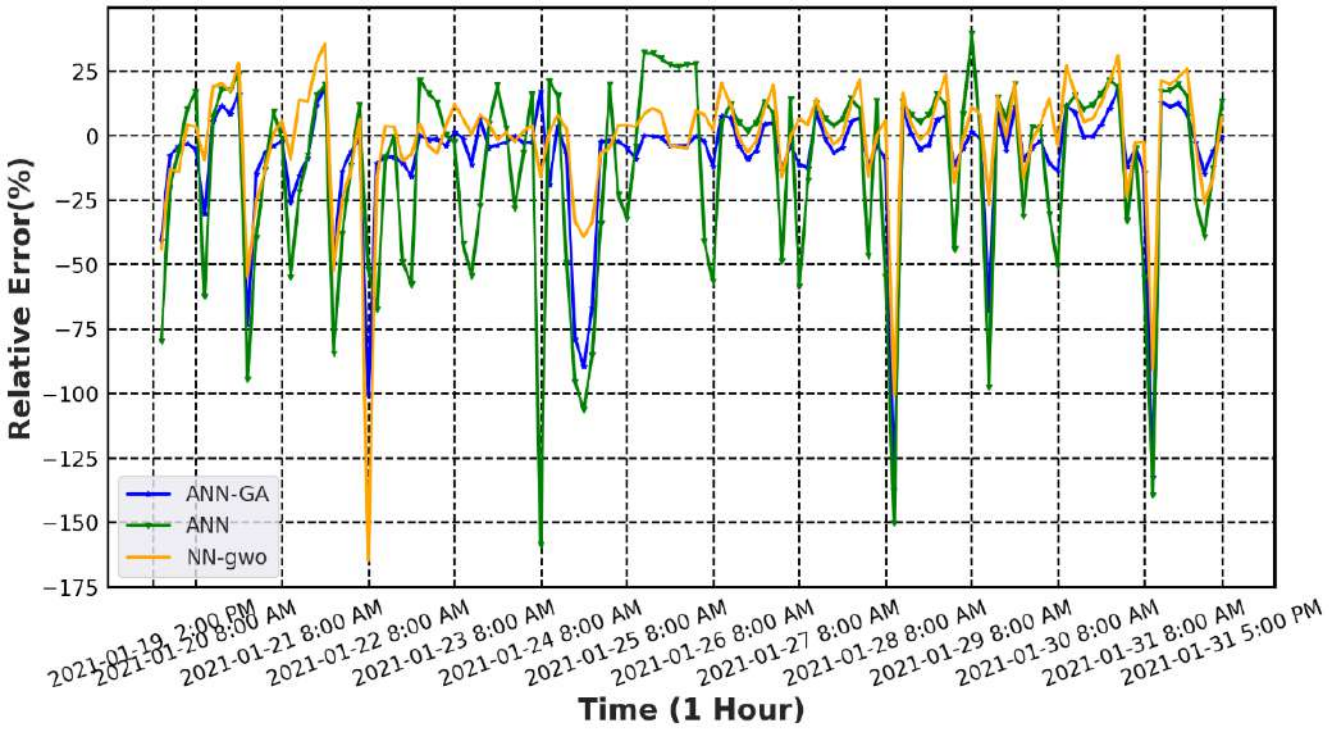
Figure 4-2 (a) The Result MAE in January (b) The Result MSE in January

Figure (4-3) (a) illustrates the temporal progression of electrical power (watts) during a specific period. The power values range from 0 to 1200 watts at one-hour intervals. The graph illustrates the Actual Power and forecasts generated by various

neural network models (ANN-GA, ANN-GWO, ANN) throughout a specified timeframe. Peaks in real power can be noted at specific periods of the day, such as the morning, where it gradually increases until it reaches its highest value at noon. This shows an increase in energy consumption during these times. Then, the power output starts to gradually decrease around 1 pm due to the decrease in sunlight. The data presented in this figure demonstrates that the enhanced models ANN-GA and ANN-GWO exhibit higher levels of forecast stability in comparison to the basic model but with minor discrepancies. These models offer a potent tool for enhancing energy management by the provision of precise predictions of daily energy usage. This can assist in optimizing resource allocation and minimizing waste. There was a decline in output from 8:00 AM to 5:00 PM on 01-22-2021, with power production predicted to reach approximately 390 watts during that period. This phenomenon can be ascribed to the decrease in solar radiation resulting from the presence of thick cloud cover due to rain. In addition, there are instances when the projected electricity production in Iraq for January may be lower than the actual output due to constraints such as reduced daylight hours and insufficient solar radiation. Figure (4-3) (b) shows the relative error between the predicted PV output and the actual PV generation for January. As depicted in Figure (4-3) (b), the ANN-GA prediction model exhibits superior stability compared to the ANN-GWO and ANN prediction models. Furthermore, the ANN technique exhibits prominent spikes, suggesting the presence of significant flaws in the forecast. Hence, the ANN-GA prediction model attains the minimum relative error of approximately 6.7%, whereas the ANN-GWO and ANN prediction models obtain 7.5% and 14.29% respectively.



(a)



(b)

Figure 4-3 (a) The Comparison between actual power and prediction power (b) Relative error with ANN, ANN-GA, and ANN-GWO in January

4.2.2. March Forecasting Term

The objective of this assessment is to assess the efficacy of the proposed prediction model throughout the spring season. To be more precise, March was selected to symbolize the arrival of the spring season. A total of 371 samples were taken hourly throughout the day. As stated in Section 4.3, the data is partitioned into separate sets for training and testing purposes. The training data is collected from 3 March at 7:00 am until 19 March at 10:00 am. On the other hand, the test data extends from 19 March at 11:00 am to the end of the month at 6:00 pm.

The results of training the neural network (NN) are summarized in Table (4-2), showing the MAE, RMSE, and MSE values as 88.7082W, 139.6889W, and 19512.9818W, respectively. These values reflect the model's predictive accuracy, with a reduction indicating better performance. Figure (4-4)(a) illustrates the MAE across training epochs, where a significant decrease in error is observed during the first 400 epochs, from about 102W to 94W. After 400 epochs, the improvement slows down, reaching around 90W by 800 epochs, suggesting the model is nearing its optimal performance. Figure (4-4)(c) shows the R-squared value, which increases significantly during the first 400 epochs, indicating an initial rapid improvement. Figure (4-4)(d) displays the RMSE, which decreases from 147 to 139.6889 as the epochs increase. Additionally, the R-squared values increased by 3.6% for January and 2.86% for July, reflecting the positive impact of longer daylight and sufficient solar radiation on PV system efficiency. The correlation coefficient of 0.8555 in Table (4-2) indicates a strong positive relationship between R-squared and forecast accuracy.

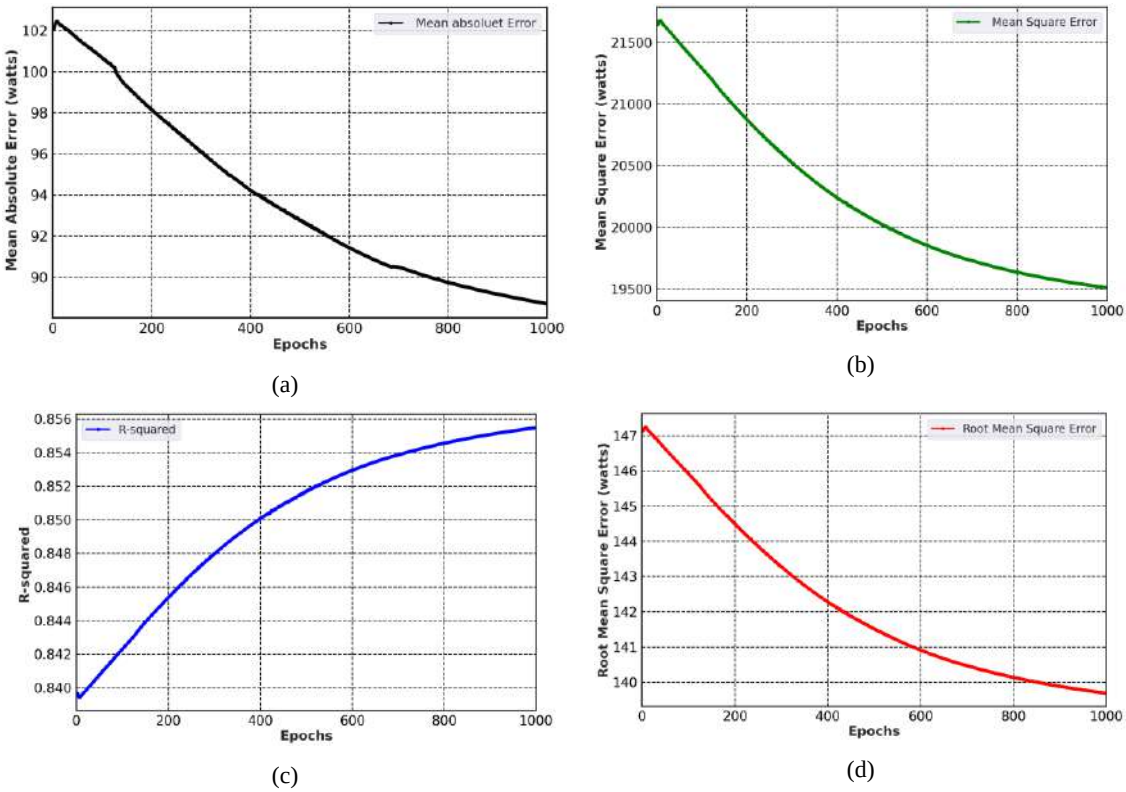


Figure 4-4 The results obtained only by employing neural networks in March (a) MAE, (b) MSE, (c)R2, (d) RMSE

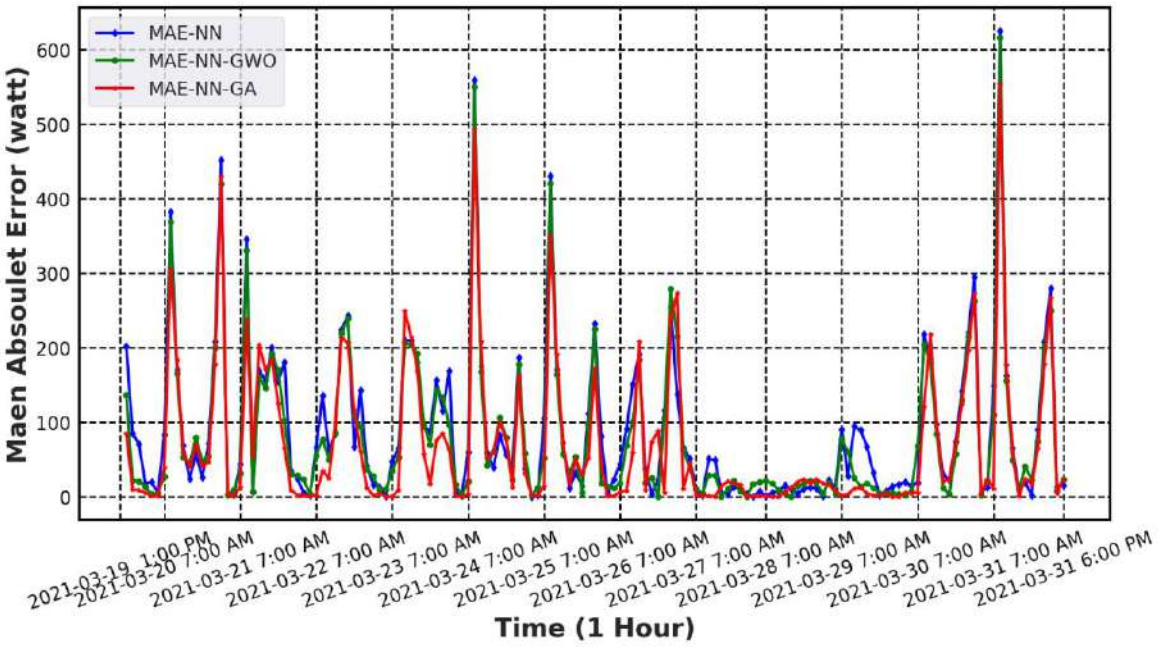
The data presented in Table (4-2) demonstrates that the proposed ANN-GA model surpasses the ANN-GWO and ANN models in terms of performance. This is evident from the MAE values reported by the ANN-GA, ANN, and ANN-GWO models, which are 69.7706 W, 88.7082W, and 81.2287 W, respectively. Figure (4-5) (a) displays the MAE obtained from the comparison between the predicted and actual power generation. The X-axis depicts the time in hourly intervals, ranging from 1:00 PM on 19 March 2021 to 6:00 PM on 31 March 2021. The Y axis represents the MAE measured in watts. The chart illustrates significant oscillations in the period from 2021-03-19 to 2021-03-23 model MAE-ANN, reaching up to 400 watts. This suggests low accuracy. Both MAE-ANN-GWO and MAE-ANN-GA exhibit lesser fluctuations compared to MAE-NN, showing higher prediction

accuracy. From 2021-03-24 to 2021-03-28, the model MAE-ANN exhibited significant fluctuations, suggesting instability and low accuracy. In contrast, the model MAE-ANN-GA performed well, displaying smaller fluctuations, the average, and fluctuation have decreased over the period.

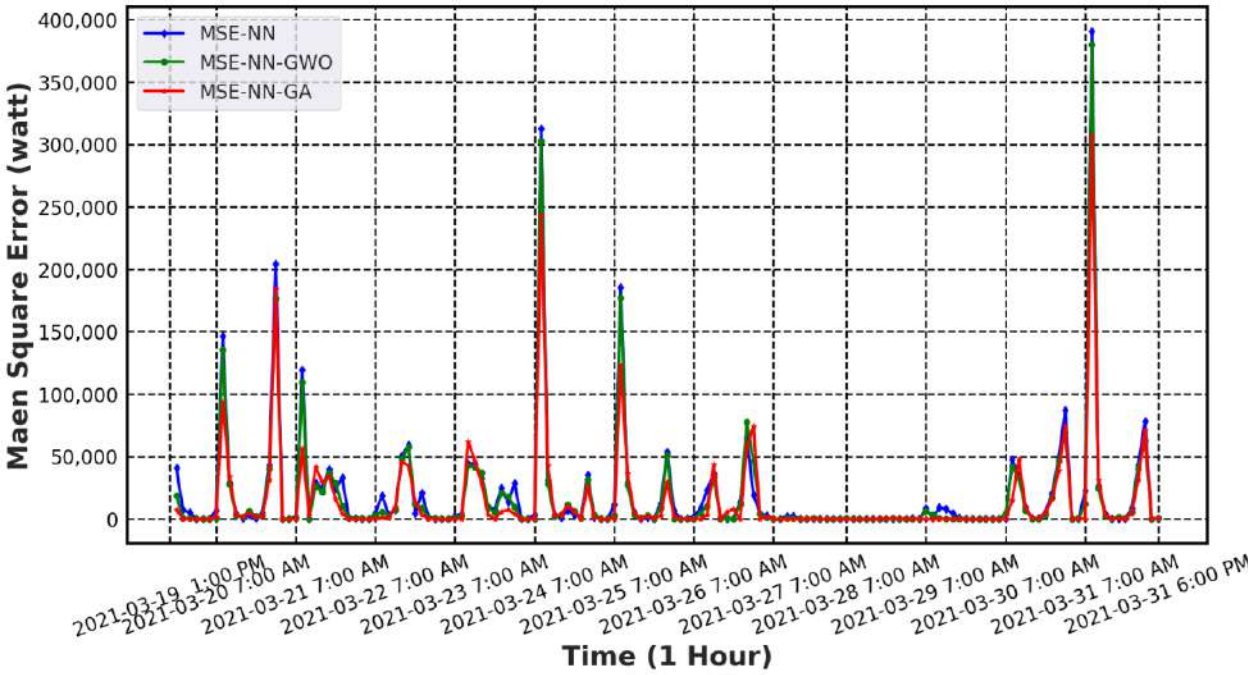
Figure (4-5) (b) represents the MSE in watts, ranging from 0 to 400,000 watts. Table (4-2) demonstrates that the suggested ANN-GA model surpasses the ANN-GWO and ANN models in terms of performance. The MSE values obtained from the ANN-GA, ANN, and ANN-GWO models are 14926.9957 W, 19512.9818 W, and 17554.4754 W, respectively, indicating clear differences amongst the models. The stability of the model in MSE-ANN-GA (red line) is greater than that of the other models. The consistent stability seen in the model suggests that the GA exhibits a higher level of resilience when confronted with abrupt fluctuations in data since it maintains a value near zero for the majority of the time. This characteristic signifies its greater precision in forecasting solar energy production. The RMSE values for the ANN, ANN-GA, and ANN-GWO approaches are 139.6889, 122.1761, and 132.4933, respectively, as shown in Table (4-2). In addition, the R2 values for the ANN, ANN-GA, and ANN-GWO are 0.8555, 0.8894, and 0.8699, respectively, as indicated in Scheme (4-13) (b).

Table 4-2 Summarized forecasting results in March days for the PV prediction model.

Days	Method	RMSE	MSE	MAE	R^2
March	ANN	139.6889	19512.9818	88.7082	0.8555
	ANN-GA	122.1761	14926.9957	69.7706	0.8894
	ANN-GWO	132.4933	17554.4754	81.2287	0.8699



(a)

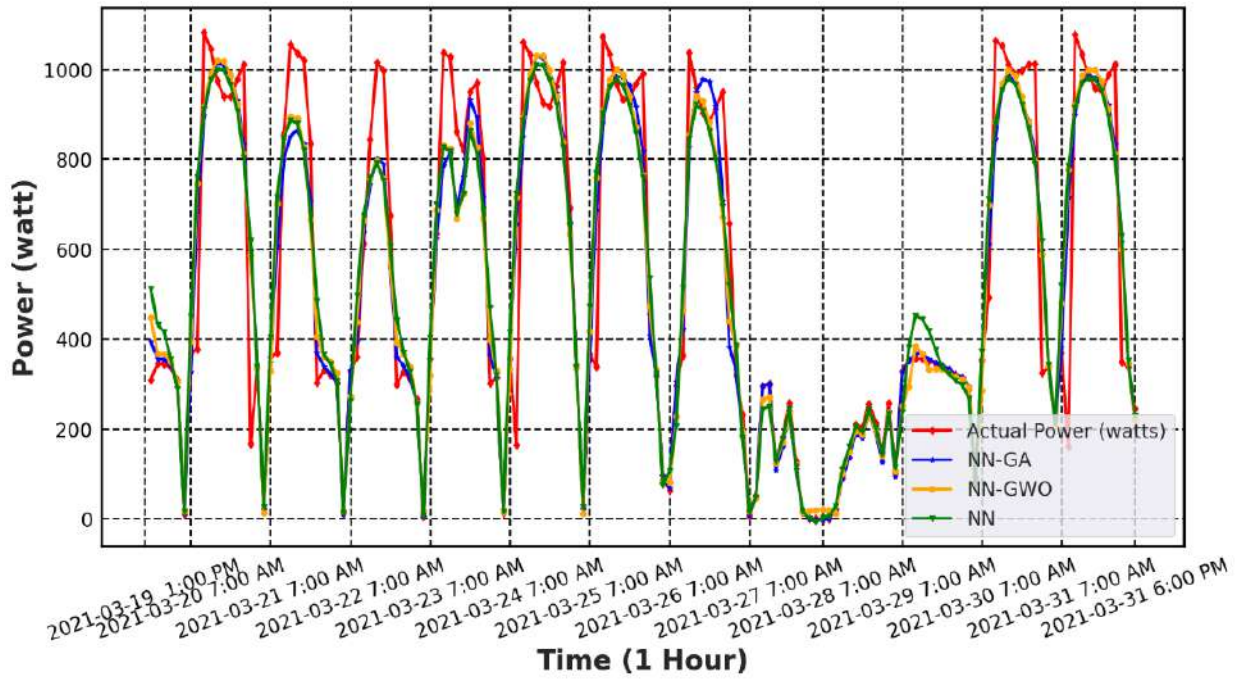


(b)

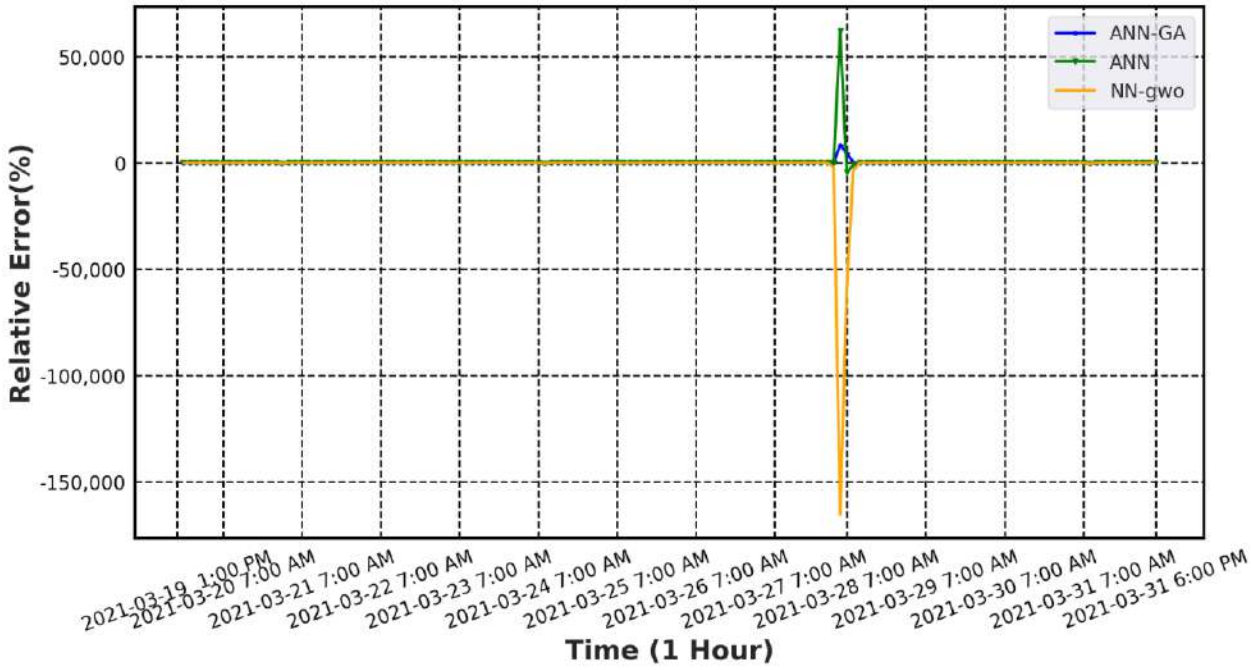
Figure 4-5 (a) The Result MAE in March (b) The Result MSE in March

Figure (4-6) (a) shows a comparison between the real power and the prediction power (in watts) for the three techniques for a specified time frame. The horizontal

axis (X) reflects time in hourly intervals. All approaches exhibit strong performance in identifying high and low points, demonstrating a high capability to forecast daily patterns of PV energy. In addition, Figure (4-6) (a) exhibits a positive trajectory in power generation between 7:00 a.m. and 12:00 p.m., followed by a progressive decrease. Between 7:00 AM on the 27th and 5:00 PM on the 28th, there was a decline in power generation. Throughout this time frame, the intensity of solar radiation varied between 7.8 and 190 w/m^2 as a result of the dense cloud cover and precipitation, as depicted in Figure (4-6) (a). Nevertheless, the correlation between predicted and real energy demonstrated enhancement in comparison to January. Based on this visual examination, it can be concluded that NN-GA and NN-GWO exhibit a modest advantage over NN in some periods, mostly because they are closer to the actual power. Figure (4-6) (b) shows the relative error between the predicted PV power output and the real PV power generation for March. Figure (4-6) (b) demonstrates that all approaches (ANN-GA, ANN, NN-GWO) consistently exhibit values close to zero over the majority of the periods. This suggests that the predictions acquired using these methods were extremely precise throughout this time frame. A distinct peak is observed in the blue line (ANN-GA) at approximately 8:00 AM on 27 March 2021. Following the peak, there is a rapid decline in the orange line (NN-GWO), reaching highly negative values (-150,000 W), due to an abrupt shift in weather conditions and an escalation in solar radiation. This suggests that NN-GWO might exhibit more sensitivity to certain variations in the data as compared to ANN-GA. While the green line (ANN) remains consistently close to zero. The ANN prediction model achieves a low relative error of approximately 54.9%, whereas the ANN-GWO and ANN-GA prediction models achieve 76.5% and 40.8% respectively.



(a)



(b)

Figure 4-6 (a) The Comparison between actual power and prediction power (b) Relative error with ANN, ANN-GA, and ANN-GWO in March

4.2.3. July Forecasting Term

The primary objective of this evaluation is to assess the accuracy of the proposed forecast model during the summer, with July chosen as the representative month for the season. Due to the longer daylight hours in Iraq during summer, a larger sample size of 435 was collected, exceeding previous months. The data was divided into two sets: the first dataset, used for training the model, spans from 07/01 05:00:00 to 07/19 11:00:00, while the second dataset, for testing, covers 07/19 12:00:00 to 07/31 19:00:00.

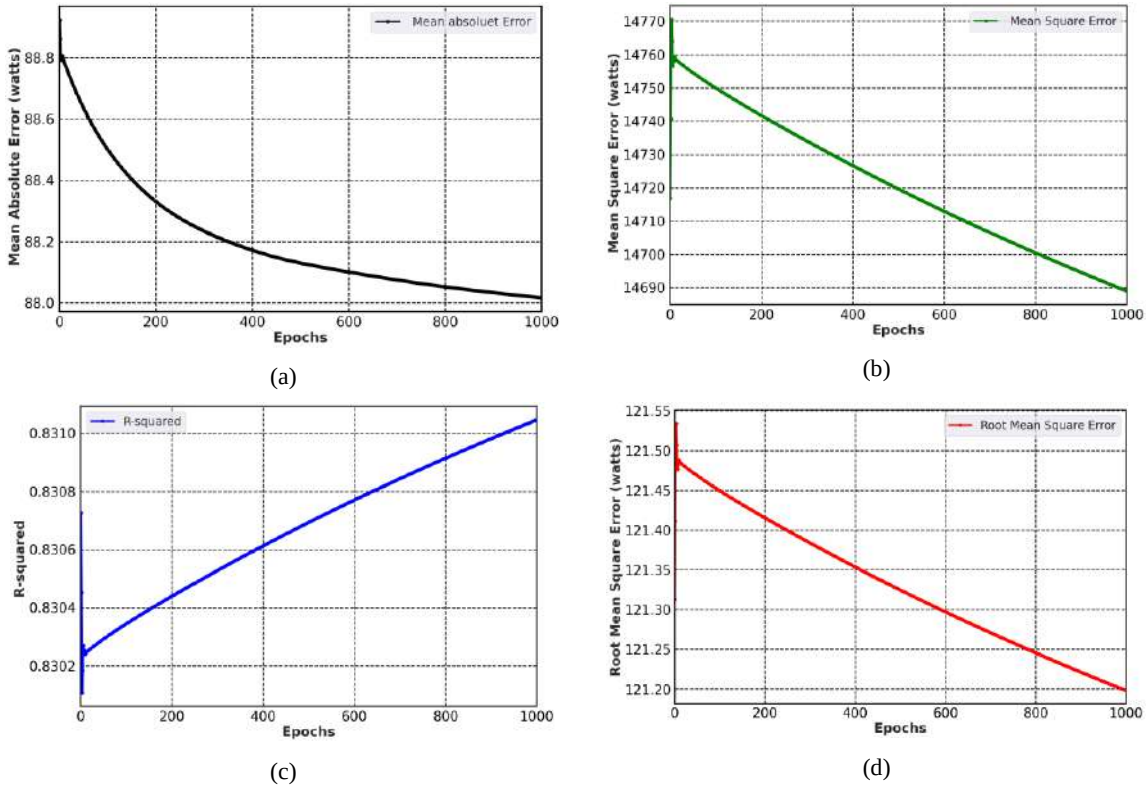
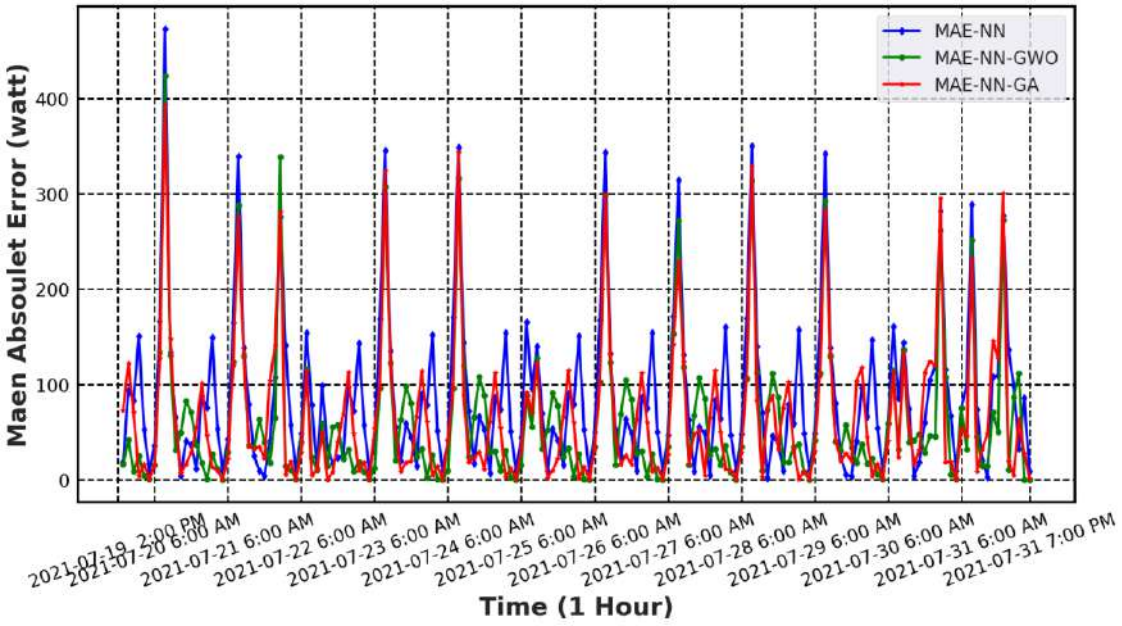


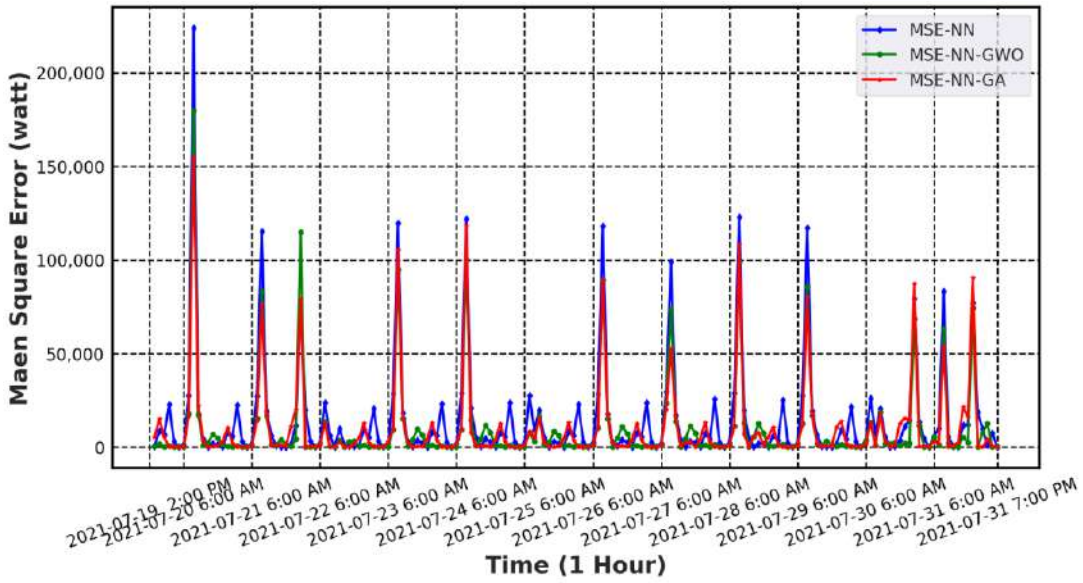
Figure 4-7 The results obtained only by employing neural networks in July(a) MAE, (b) MSE, (c)R2, (d) RMSE

Table (4-3) and Figure (4-7) (a, b, d) display the respective values of MAE, MSE, and RMSE. The MAE value is 88.0168 W, the MSE value is 14688.9915 W, and the RMSE value is 121.1981 W. Therefore, a reduction in the specified evaluation

measures signifies an improvement in the precision of the forecast. The decline commences about 800 epochs and subsequently reaches a state of stabilization. Furthermore, Figure (4-7) (c) presents the R2 value of the suggested model as 0.8310, suggesting that a higher R2 value equates to a higher level of prediction accuracy.



(a)

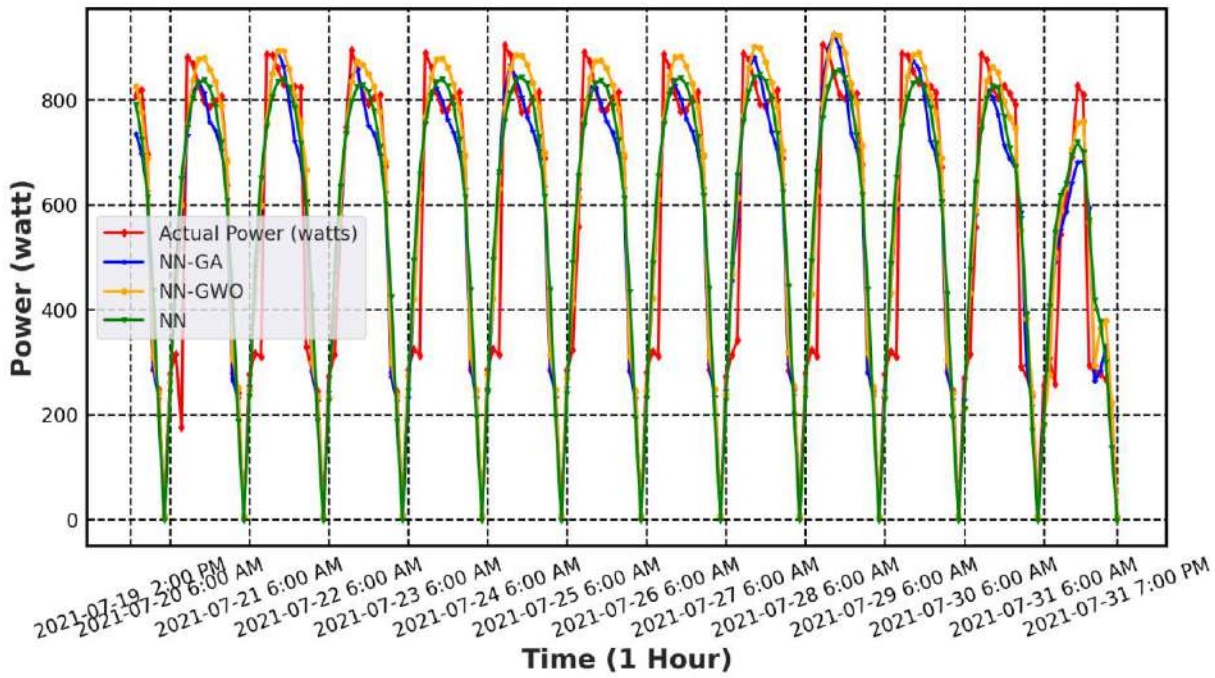


(b)

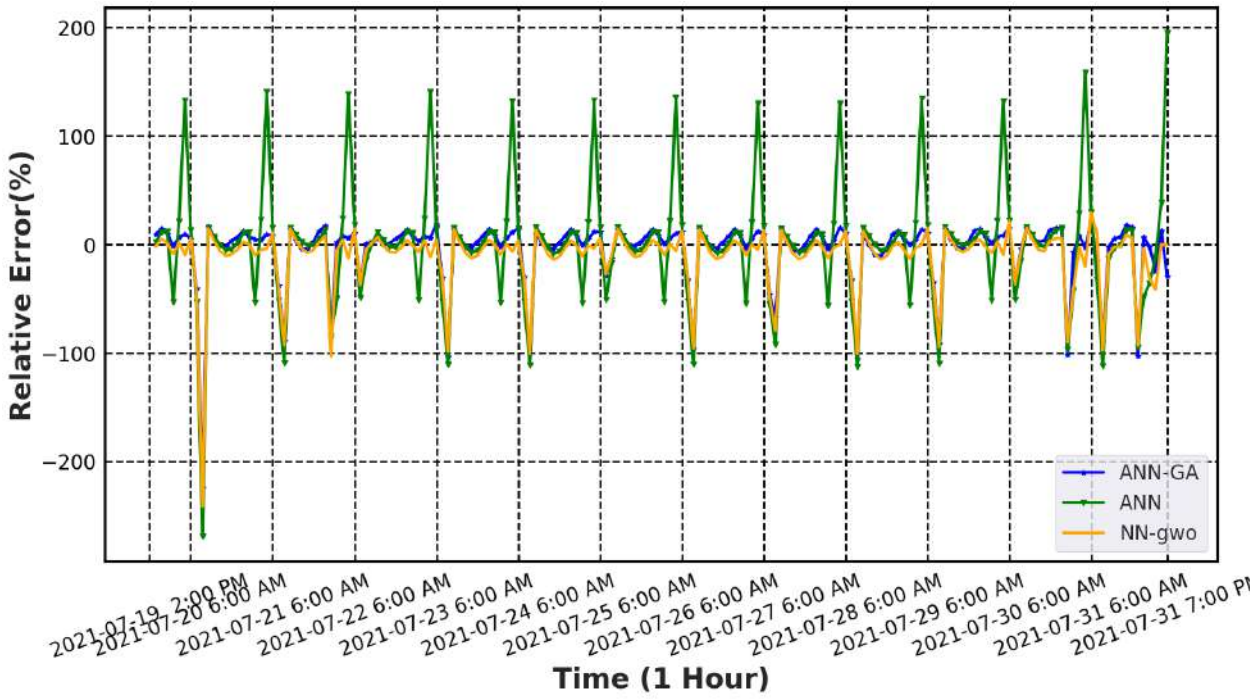
Figure 4-8 (a) The Result MAE in July (b) The Result MSE in July

The data presented in Table (4-3) demonstrates that the suggested ANN-GWO model exhibits superior performance compared to the ANN-GA and ANN models. The MAE values reported by the ANN-GA, ANN, and ANN-GWO models are 66.3552 W, 88.0168 W, and 62.9410 W, respectively, indicating an evident difference in performance. The ANN-GWO model shows better performance during this July, as shown in the results in Table (4-3). Figure (4-8) (a) illustrates the MAE resulting from the comparison of predicted and actual power generation. The Y axis displays MAE measured in watts, with a range from 0 watts to 400 watts. The drawing exhibits a consistent pattern among the three lines, with minor variations distinguishing them. The recurring pattern in the data may be attributed to variations in solar radiation throughout the day, with solar energy intensifying from 7 a.m. and peaking at 12 p.m., followed by a slow decline. This change has a direct impact on the precision of the models, resulting in an elevation of the MAE during specific time intervals, such as particular days. Figure (4-8) (b) displays the power output of MSE in watts, with a range from 0 to 200,000 watts in hourly. Table (4-3)

demonstrates that the suggested ANN-GWO model surpasses the ANN-GA and ANN models in terms of performance. The MSE values derived from the ANN-GA, ANN, and ANN-GWO models are 10309.3373W, 14688.9915W, and 9834.1092W, respectively. These results indicate significant discrepancies between the models. The stability of the MSE-NN-GWO model, represented by the green line, surpasses that of the other models. The RMSE values for the ANN, ANN-GA, and ANN-GWO techniques are 101.5349, 121.1981, and 99.1671, respectively, as indicated in Table (4-3). In addition, the R2 values for the ANN, ANN-GA, and ANN-GWO are 0.8310, 0.8814, and 0.8869, respectively, as indicated in Scheme (4-13) (c).



(a)



(b)

Figure 4-9 (a) The Comparison between actual power and prediction power (b) Relative error with ANN, ANN-GA, and ANN-GWO in July

Table 4-3 Summarized forecasting results July days for PV prediction model.

Days	Method	RMSE	MSE	MAE	R^2
	ANN	121.1981	14688.9915	88.0168	0.8310
July	ANN-GA	101.5349	10309.3373	66.3552	0.8814
	ANN-GWO	99.1671	9834.1092	62.9410	0.8869

Figure (4-9) (a) compares the actual and predicted power for July. The X-axis indicates the time in hour intervals, ranging from 2:00 PM on 19 July to 7:00 PM on 31 July. The Y-axis displays electrical power in watts, with a range from 0 to 900 watts. The graph exhibits a diurnal recurring pattern that mirrors the alternation of daylight, with power growing during daylight hours and diminishing after 12:00 PM.

The predicted lines (blue, orange, green) adhere closely to this pattern. During peak periods, often around midday, all approaches demonstrate strong concurrence with the real capacity. The blue and orange lines accurately track the peak, demonstrating strong predictive accuracy during these periods. During periods of low power, such as after 1 pm or early morning, there are minor variances between the predicted power and the actual power. Where the green line exhibits fluctuations during these time intervals, suggesting that enhanced techniques may provide more precise predictions for periods characterized by quick shifts, and the enhanced techniques (NN-GA and NN-GWO) exhibit a remarkable level of precision in forecasting solar power, as it nearly aligns with the real power. Despite the higher quantity of samples, energy production in this month did not surpass that of September. The decrease in productivity can be ascribed to the elevated temperatures experienced in Iraq last July. Figure (4-9) (a) demonstrates a consistent and balanced correlation between actual and expected power, indicating that it is an opportune moment to generate electricity utilizing PV systems. Figure (4-9) (b) displays the discrepancy between the estimated PV power production and the actual PV power generation for July. Figure (4-9) (b) demonstrates that the three models (ANN-GA, ANN, NN-GWO) exhibit a recurring pattern in the predicted energy, which aligns with the daily fluctuations in solar radiation. Nevertheless, the fundamental ANN model exhibits marginally different performance, potentially attributable to the absence of supplementary optimization approaches employed in the other two models. Finally, the ANN-GWO prediction model demonstrates a significantly lower relative error of approximately 8%, whereas the ANN and ANN-GA prediction models exhibit relative errors of 16.77% and 8.1%, respectively.

4.2.4. September Forecasting Term

In September, a substantial quantity of data, amounting to 390 samples, was gathered to assess the efficacy of the suggested prediction model over the autumn season. The data for this month was split into two categories: training data, which was collected from 1 September at 06:00 to 18 September at 18:00, and testing data, which was collected from 19 September at 6:00 a.m. to 30 September at 18:00 to evaluate the model's effectiveness.

The optimal MAE, MSE, and RMSE values are around 71.6435W, 8574.6566W, and 92.5994W, respectively when using ANN, as indicated in Table (4-4). The value represents the spectrum of training episodes that have been finished, spanning from 0 to 1000 epochs. The MAE, MSE, and RMSE axes are quantified in watts. During the first 600 epochs, the average deviation for MAE reduced dramatically from approximately 80 to approximately 73 W. This signifies a discernible enhancement in the model's performance during this stage. The level of imprecision gradually diminishes, albeit at a reduced pace, as the number of epochs rises from 600 to 1000. Similarly, the RMSE and MSE drop with an increase in the number of epochs as Figure (4-10). The data also indicated that the MAE and RMSE values were significantly lower compared to the values recorded in January, with reductions of 37.64% W and 38.52%, respectively. In addition, the ANN model demonstrates an outstanding R2 value of 0.9244, surpassing other findings.

The data in Table (4-4) demonstrates that the suggested ANN-GA model surpasses the performance of both the ANN-GWO and ANN models. The MAE values reported by the ANN-GA, ANN, and ANN-GWO models are 57.9227 W, 71.6435 W, and 63.6414 W, respectively, clearly indicating the differences in performance amongst the models. Figure (4-11) (a) shows the MAE resulting from the comparison between the predicted and actual power generation. The range is

from 0 to 350 watts. The chart displays significant variations during the early hours of 27 September, with the MAE-ANN model reaching a peak of 338.34 W, the MAE-ANN-GA model reaching 304.06 W, and the MAE-ANN-GWO model reaching 323.47 W due to these changes the increase in solar radiation during one hour leads to large differences in the accuracy of weather forecasts. Thus, employing the enhanced ANN-GA method yields superior precision and enhanced reliability in forecasting PV power generation, as compared to utilizing the unimproved ANN and the GWO algorithm.

Figure (4-11) (b) displays the MSE of three distinct approaches throughout an hourly period on the horizontal (X) axis, covering the dates from 19 September 2021 to 30 September 2021. The Y axis represents the MSE in watts, ranging from 0 to 120,000 watts. From 19 September to 21 September, the three models exhibit low values, suggesting a stable performance. However, on 27 September, there is a significant surge in power consumption. The usage reaches a peak of over 92452.68 W while employing MSE-ANN-GA, around 104636.92 W when utilizing MSE-ANN-GWO, and 114472.08 W when utilizing MSE-ANN. This surge in power consumption corresponds to a substantial rise in the MSE caused by the rapid change in solar radiation ،and varies between 63 and 409 w/m^2 throughout one hour. In general, the MSE-ANN-GA technique has a higher level of accuracy when it comes to predicting PV power generation. This information is evident from Table (4-4). Additionally, it exhibits the lowest MSE values and demonstrates less volatility over time. The RMSE values for the ANN, ANN-GA, and ANN-GWO approaches are 92.5994 W, 82.0618 W, and 87.1780 W, respectively. The MSE values for the ANN, ANN-GA, and ANN-GWO approaches are 8574.6566, 6734.1462, and 7600.0110 W, respectively. In addition, the R2 values for the ANN, ANN-GA, and ANN-GWO are 0.9244, 0.9406, and 0.9329, respectively, as indicated in Scheme (4-13) (d).

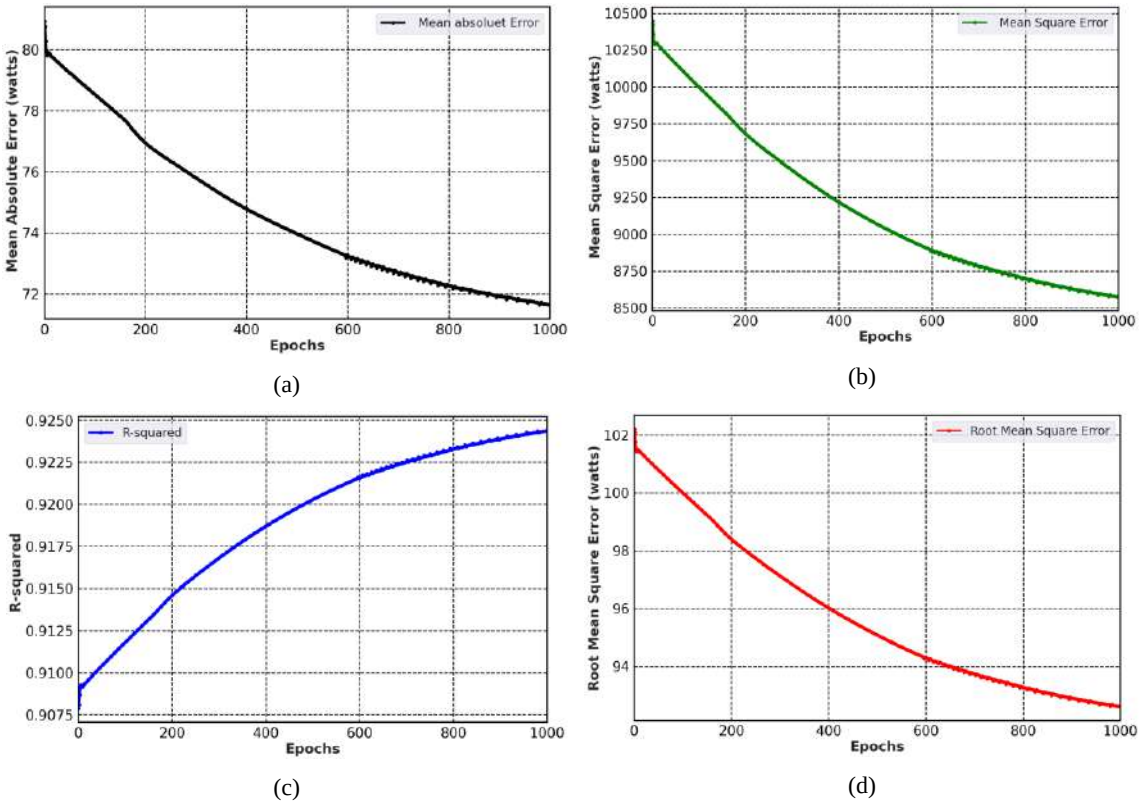


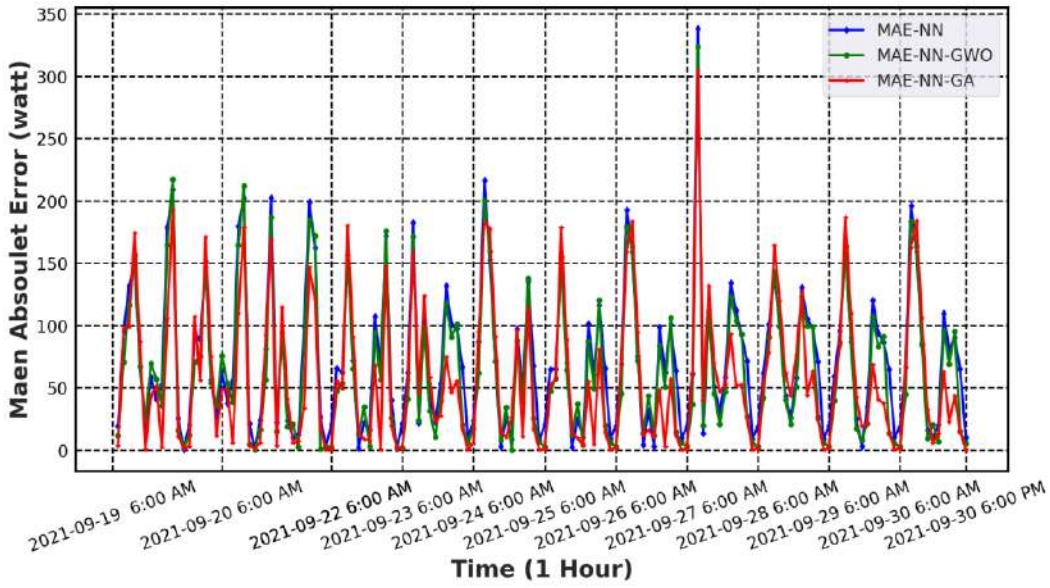
Figure 4-10 The results obtained only by employing neural networks in September (a) MAE, (b) MSE, (c)R2, (d) RMSE

Table 4-4 Summarized forecasting results in September days for the PV prediction model.

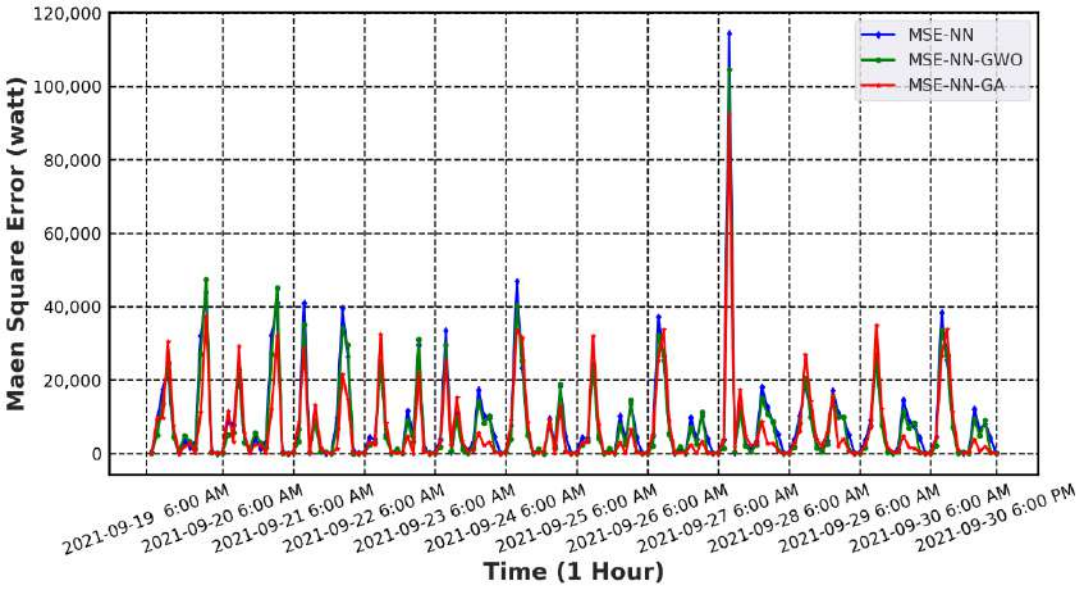
Days	Method	RMSE	MSE	MAE	R^2
	ANN	92.5994	8574.6566	71.6435	0.9244
September	ANN-GA	82.0618	6734.1462	57.9227	0.9406
	ANN-GWO	87.1780	7600.0110	63.6414	0.9329

Figure (4-12) (a) presents a comparison between the actual power and prediction power in September using three different approaches. In September, there is a consistent daily pattern that mirrors the cycle of daylight. This is because the power

improves during the daytime when there is ample sunlight and mild temperatures. The NN-GA and NN-GWO models exhibit higher concordance with the actual power, particularly during peak and off-peak periods. The NN, enhanced by the utilization of the GA, exhibits superior performance in comparison to other results. This can be attributed to the presence of optimal meteorological circumstances, such as clear skies and ample sunlight, which are conducive to the compatibility of this method. Figure (4-12) (b), displays the relative error over time during September, ranging from -100 to 200 percent. The graph displays variations in the relative inaccuracy throughout September. The ANN-GA prediction model has superior stability when compared to the ANN-GWO and ANN prediction models. Furthermore, the ANN technique demonstrates more pronounced oscillations, indicating a significant variability in the forecasts, particularly in the troughs where the power approaches substantial negative values. The NN-GWO model, shown by the orange color, exhibits a comparable pattern to the ANN-GA model but displays more significant oscillations, particularly near the peaks. Hence, the ANN-GA prediction model attains the minimum relative error of approximately 6%, whereas the ANN-GWO and ANN prediction models achieve 10% and 13.9% respectively.



(a)



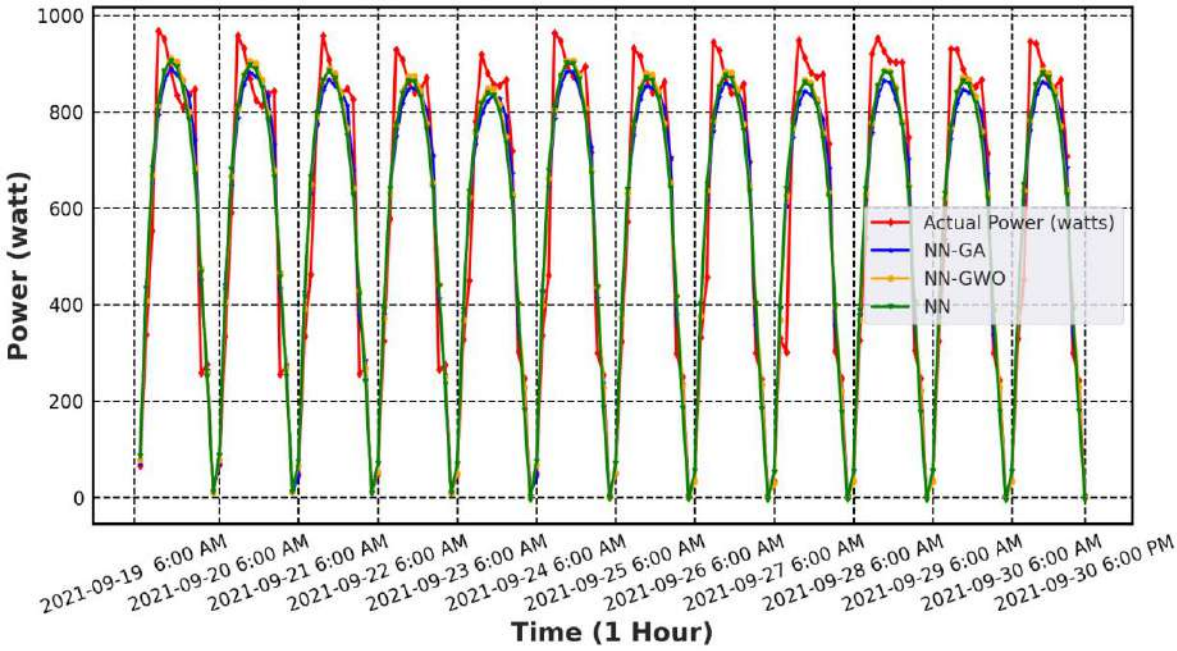
(b)

Figure 4-11 (a) The Result MAE in September (b) The Result MSE in September

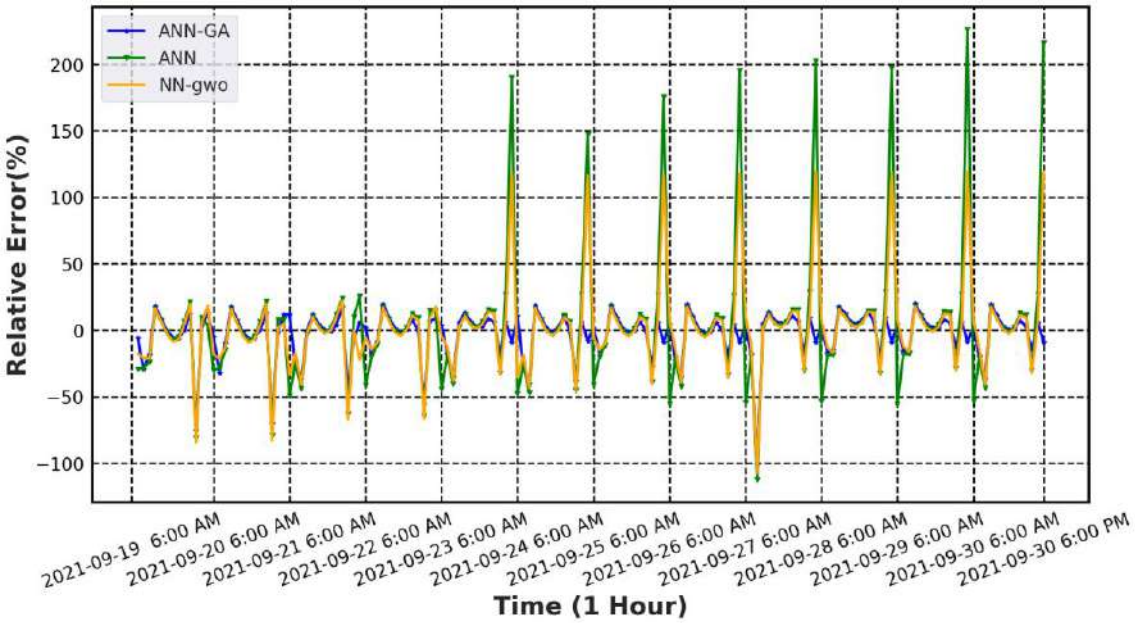
Table 4-5 Relative percentage error for January, March, July, and September

Method	January	March	July	September
ANN-GA	6.7%	40.8%	8.1%	6%
ANN-GWO	7.5%	76.5%	8%	10%

ANN	14.29%	54.9%	16.77%	13.9%



(a)



(b)

Figure 4-12 (a) The Comparison between actual power and prediction power with ANN, ANN-GA, and ANN-GWO in September (b) Relative error with ANN, ANN-GA, and ANN-GWO in September

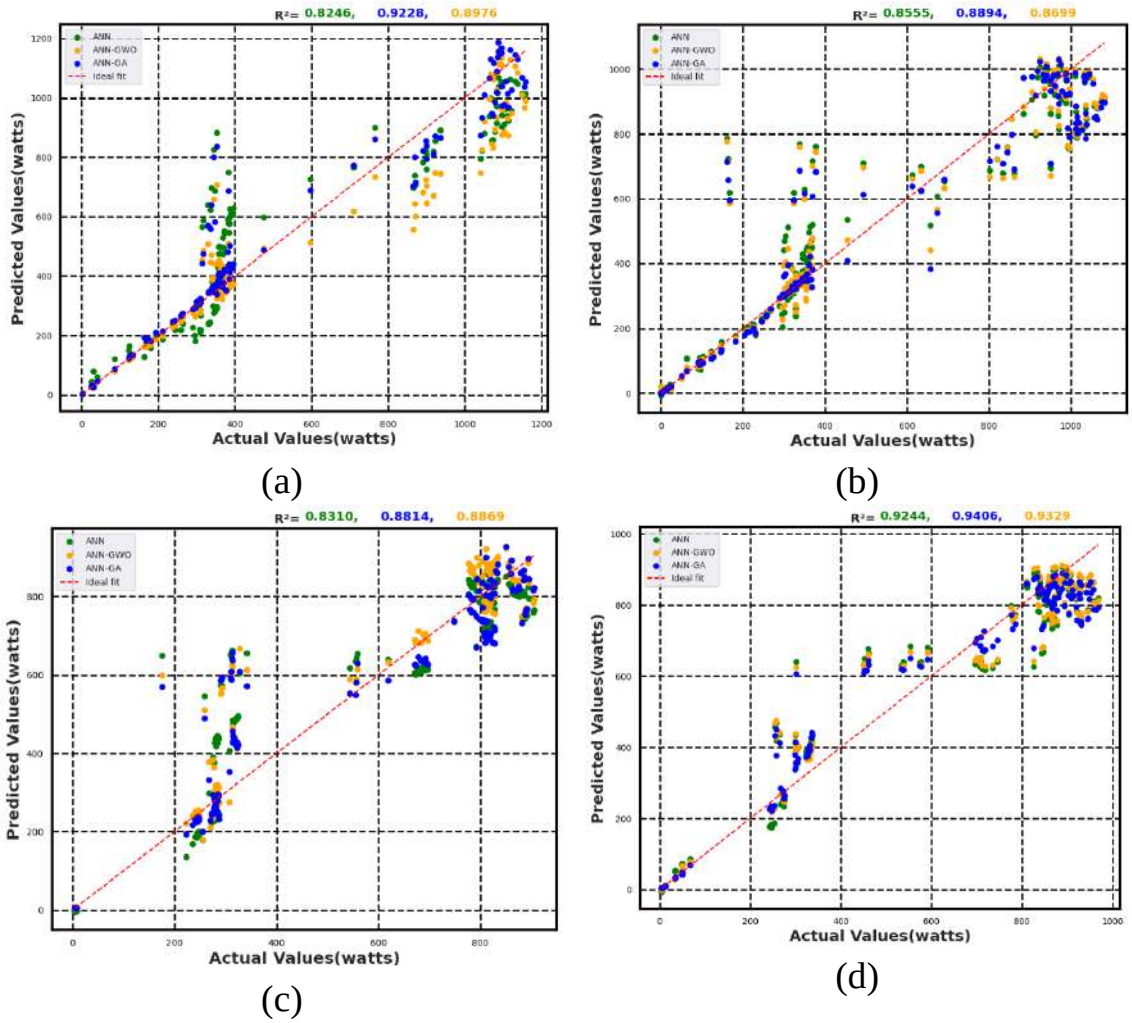


Figure 4-13 Scatter diagrams of ANN, ANN-GA, and ANN-GWO (a)January, (b)March, (c) July, (d)September

4.2.5. Discussion

After reviewing the results for each month individually, it is evident that September exhibited the most suitable values according to the defined criteria. Consequently, September is forecasted to have the maximum potential for PV generation, because there is a large amount of sun radiation caused by the absence of clouds and mild temperatures. In contrast, solar generation in January decreased due to the limited duration of daylight and the presence of unexpected weather

patterns, which included weak sunlight, frequent showers, and gloomy skies. As a result, it is regarded as the least favored month for forecasting. Nevertheless, solar energy that is now accessible continues to contribute to electricity generation. Conversely, July is characterized by high temperatures and abundant sunlight, making it a favorable time frame for energy production in the solar system. Although the ANN demonstrated satisfactory performance, it fell short of achieving the intended level when compared to the other methods discussed in Chapter 2 of this thesis. The reason for this is that certain techniques have undergone enhancements and have been integrated with other technology, leading to increased abilities in predicting future events. Hence, this thesis enhanced its performance by employing two techniques: GA and GWO. Post-enhancement, the results demonstrated that the enhanced neural network surpassed the unimproved neural network in all chosen months. Upon comparing the GWO with the GA, we observe that Table 6 presents the optimized selection of hidden layers and neurons, as well as the corresponding time taken for each technique. It is evident that the GA exhibits faster performance and higher accuracy compared to the GWO. On the contrary, in July, the GWO outperformed the GA. This suggests that the GWO is more adept at handling high temperatures, particularly in the hotter climate of southern Iraq during July. Ultimately, the enhanced and integrated techniques proven to be superior in accurately forecasting PV power generation across various circumstances.

Table 4-6 shows a comparison of results for the models used

Modeling	**parameter**	**January**	**March**	**July**	**September**
ANN-GA	Layer Number	4	4	3	4
	Hidden Neurons Number	[34,101,44, 121]	[30,102,1 03,51]	[13,115,86]	[11,120,82,86]
	Time	14 mint	20 mint	17 mint	16 mint

ANN-GWO	Layer Number	5	4	3	4
	Hidden Neurons Number	11	18	23	21
	Time	30 mint	22 mint	53 mint	37 mint

4.3. Results of The Experimental Part

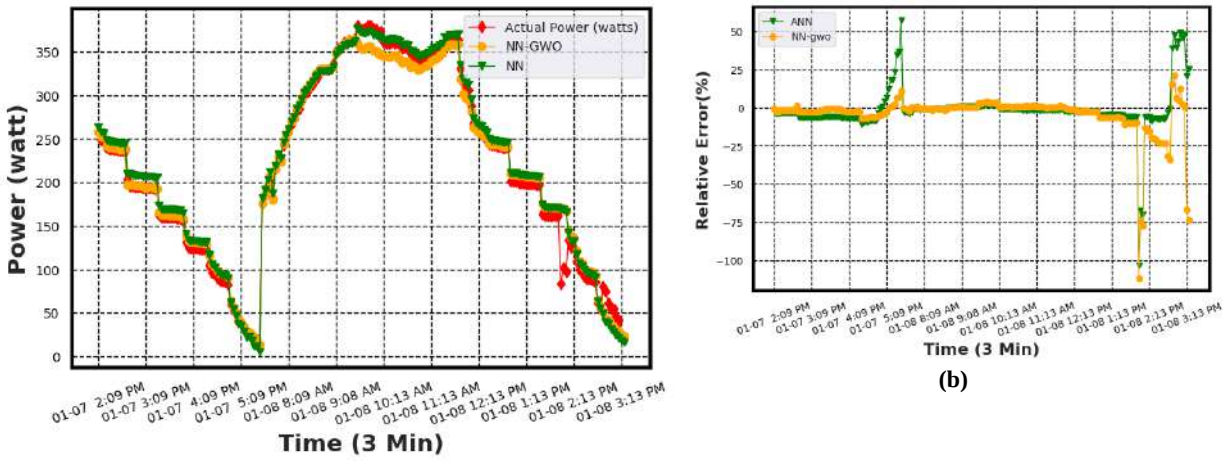
The data used in this study was collected in 2024 during experimental testing of PV cells at the University of Misan College of Engineering, Iraq (31.8907° N, 47.1078° E), covering January, March, and June. Due to the variable nature of weather conditions, the experimental data may contain errors or missing intervals. To address this, ANN and optimization techniques were applied to adjust the number of hidden layers and neurons. An experimental study was conducted to evaluate PV power generation and the effectiveness of the proposed technique under various weather conditions: six days of cloudy, rainy, and sunny weather. The fall season was excluded due to the recency of the 2024 data, as the fall season in Iraq begins in late September. Python, using Matplotlib and Seaborn libraries, was utilized to create visual representations of the results.

4.3.1. Results for January

This analysis utilizes real data obtained in January 2024, the data sampling is conducted at three-minute intervals, from 7 a.m. until 5 p.m., due to the absence of energy generation by solar cells during nighttime hours. The data was classified into three weather patterns (cloudy, rainy, and sunny), with three days chosen for each pattern. The data samples were first partitioned, allocating 60% for training and the remaining 40% for testing. Next, the number of hidden layers and neurons is determined based on GWO because it converges strongly to find the correct solutions and is a modern algorithm with good performance. Finally, the results of the ANN-GWO model show high performance in all weather conditions (sunny, rainy, and cloudy) compared to ANN through the Table and Figure below.

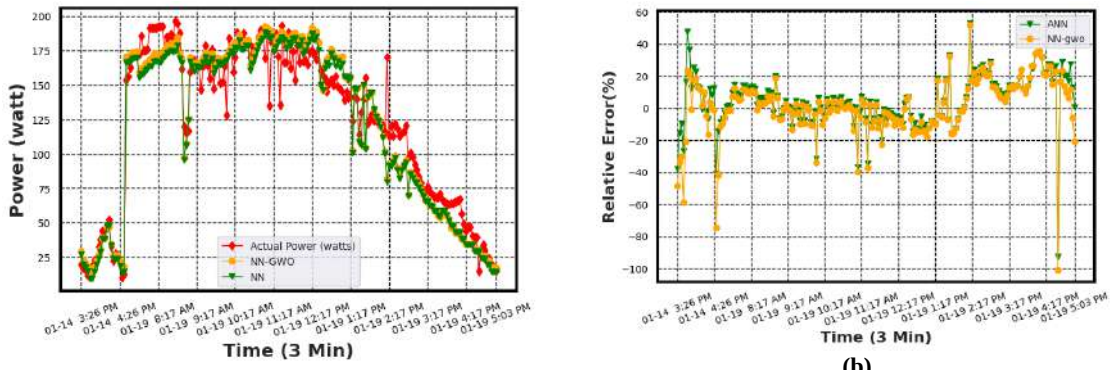
Table 4-7 Summarized forecasting results Sunny days for PV prediction model.

Days	Method	RMSE	MSE	MAE	RE(%)	R^2
cloudy	ANN	16.2648	264.5423	12.2325	13%	0.9215
	ANN-GWO	14.8086	219.2948	11.2481	12.3%	0.9349
rainy	ANN	20.6886	428.0175	13.1810	16.8%	0.6754
	ANN-GWO	20.0763	403.0617	11.6983	14.5%	0.6943
Sunny	ANN	12.3783	153.2229	7.8102	7.2%	0.9870
	ANN-GWO	11.8840	141.2314	7.1700	5.1%	0.98799



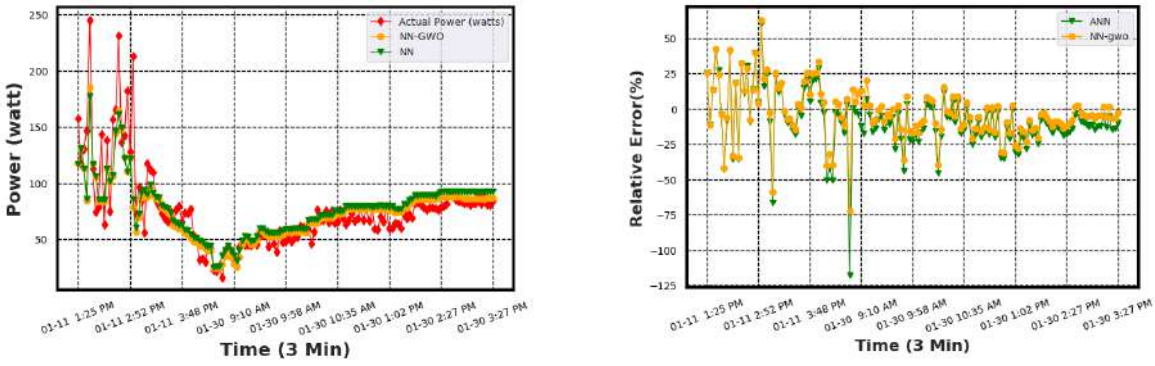
(a) (b)

Figure 4-14 (a) The Comparison between actual power and prediction power (b) Relative error with ANN, and ANN-GWO in Sunny days



(a) (b)

Figure 4-15 (a) The Comparison between actual power and prediction power (b) Relative error with ANN, and ANN-GWO in Cloudy days



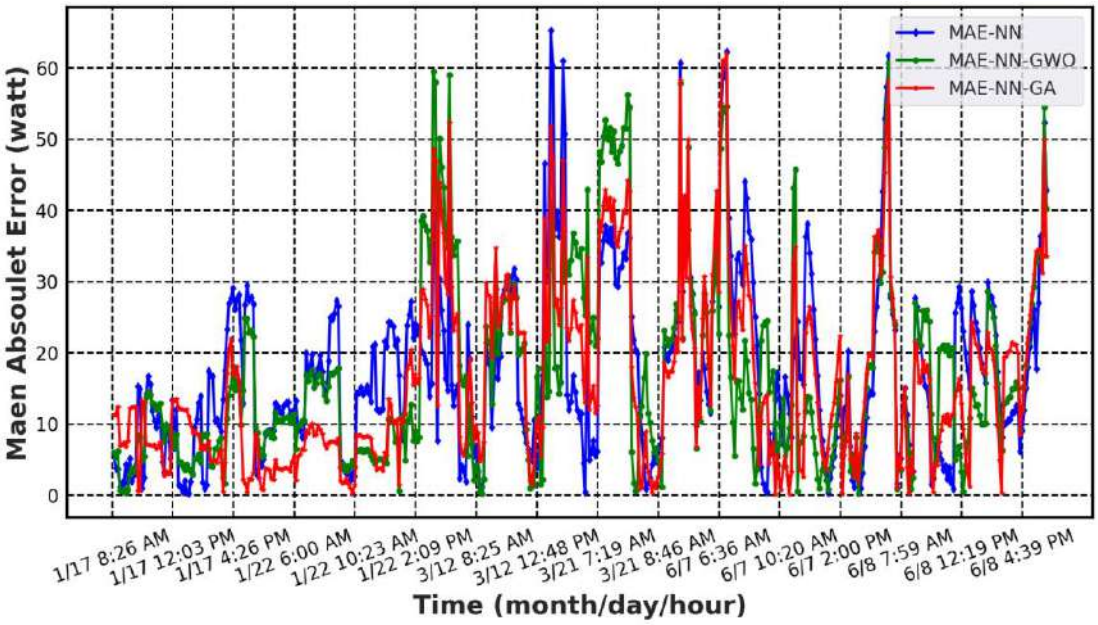
(a) (b)

Figure 4-16(a)The Comparison between actual power and prediction power (b) Relative error with ANN, and ANN-GWO in Rainy days

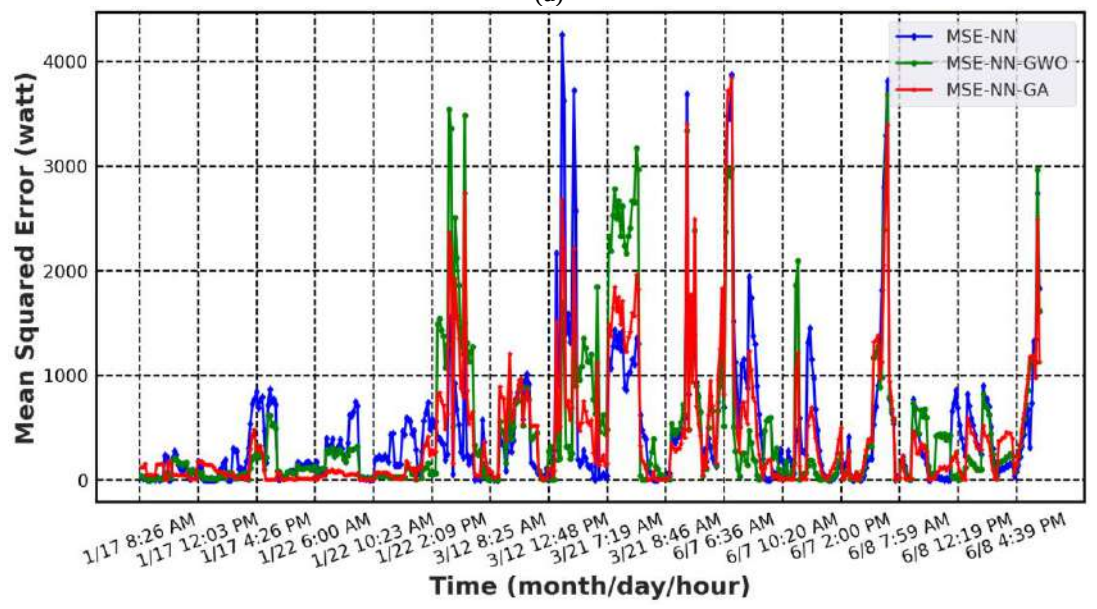
4.3.2. Sunny Days

This classification aims to evaluate the efficacy of the proposed forecasting model under sunny weather conditions. To be more specific, a total of 1,158 samples were chosen across a span of 6 days characterized by unobstructed skies. The samples were evenly distributed throughout three seasons to represent sunny weather. 361 samples were collected on 17 and 22 January to represent the winter season. 12 March and 21 March were designated to symbolize the arrival of spring, with a combined total of 384 samples. Similarly, 7 June and 8 June were picked to

represent the summer season, with a total of 413 samples. The number of samples varies due to the impact of sunshine on the solar panels. Moreover, the length of daylight varies over the different seasons. The data was partitioned with 60% of the data allocated for training the model and the remaining 40% for testing purposes.



(a)



(b)

Figure 4-17 (a)The Result MAE of the Sunny Days, (b) The Result MSE of the Sunny Days

Figure (4-17) (a), displays the MAE in watts on the y-axis, ranging from 0 to 60 W for all proposed approaches. The x-axis shows the monitoring hours, with data collected every 3 minutes. Figure (4-17) (a), displays the difference in MAE between the predicted and observed power generation during sunny weather circumstances. The GA-ANN, ANN, and ANN-GWO approaches, which are introduced in this thesis, are used for this analysis. The GA-ANN method occasionally yielded reduced MSE and MAE values in comparison to alternative methods. On days characterized by clear skies, the MAE values were 17.3919 for the ANN, 16.0403 for ANN-GA, and 17.0240 for ANN-GWO. As indicated in Table (4-8). Furthermore, the graph illustrates that the three models exhibit convergence in the MAE when it is sunny. This convergence is attributed to the consistent stability of solar radiation and temperature, resulting in a gradual decrease in error. The enhancements achieved through the utilization of the genetic algorithm demonstrate efficacy in diminishing errors and enhancing the precision of forecasts. Therefore, stable atmospheric conditions contribute to minimizing mistakes in energy predictions, hence enhancing the performance of the models.

Table 4-8 Summarized forecasting results Sunny days for PV prediction model.

Days	Method	RMSE	MSE	MAE	R^2
	ANN	21.5097	462.6680	17.3919	0.9532
Sunny	ANN-GA	20.5089	420.6154	16.0403	0.9574
	ANN-GWO	21.8529	477.5520	17.0240	0.9516

Figure (4-17) (b), displays the MSE in watts on the y-axis, ranging from 0 to 4000 W for all proposed approaches. The x-axis depicts the monitoring hours, with measurements taken every 3 minutes. Figure (4-17) (b), illustrates the difference in MSE between the predicted and observed power generation during sunny weather

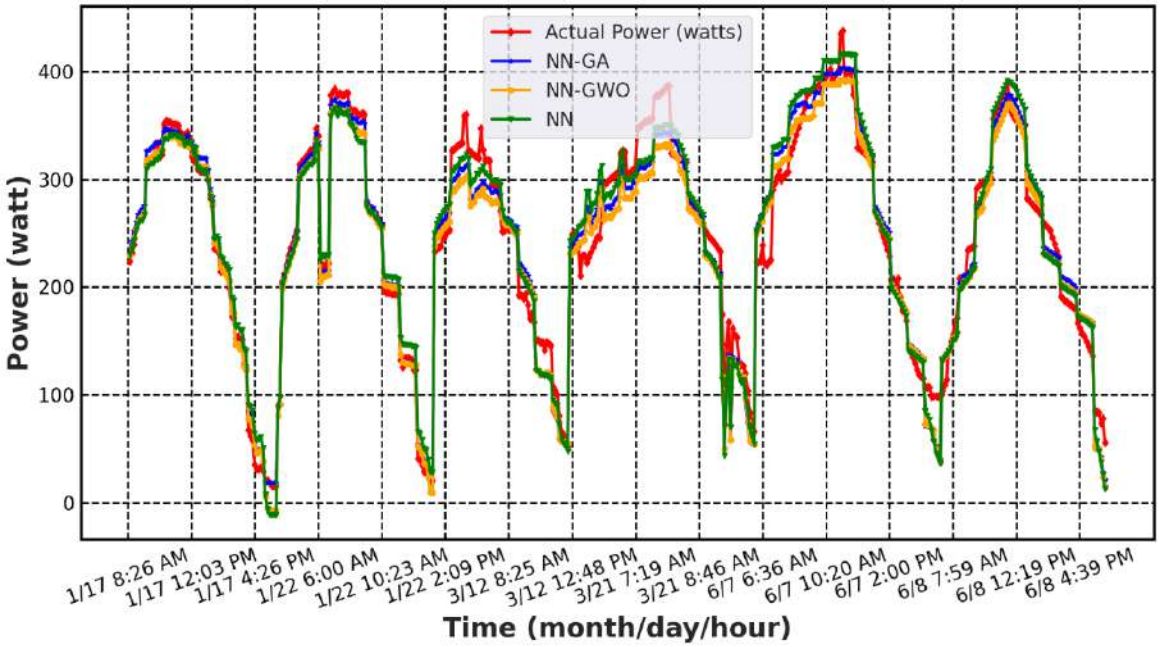
conditions. The ANN-GA, ANN, and ANN-GWO approaches, which are introduced in this thesis, are used for this analysis. The ANN-GA method occasionally yielded reduced MSE and MAE values in comparison to alternative methods. On days when the sky was clear, the MSE values were 462.6680 for the ANN, 420.6154 for ANN-GA, and 477.5520 for ANN-GWO. As indicated in Table (4-8). Furthermore, the diagram illustrates that the three models exhibit convergence in the MSE when it is sunny, albeit with variations in performance across different time intervals. The utilization of the GWO with the GA has demonstrated efficacy in minimizing errors and enhancing the precision of predictions as compared to the rudimentary MSE-ANN model. Moreover, Table (4-8) displays the R2 of the ANN-GA, ANN, and ANN-GWO, with values of 0.9574, 0.9532, and 0.9516, as indicated in Scheme (4-23) (b).

Table 4-9 Model Performance Comparison in Sunny Weather

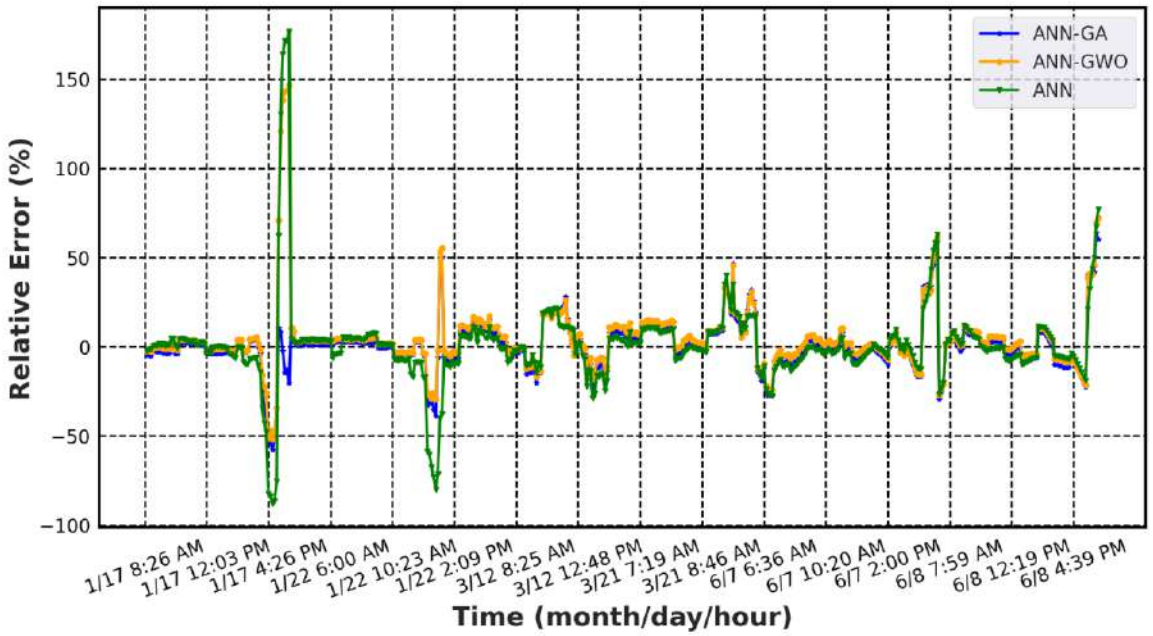
Indicator	ANN	ANN-GA	Ratio (%) for ANN-GA	ANN-GWO	Ratio (%) for ANN-GWO
MAE	17.3919	16.0403	7.77%	17.0240	2.11%
RMSE	21.5097	20.5089	4.65%	21.8529	1.57%
MSE	462.6680	420.6154	9.08%	477.5520	3.12%
R^2	0.9532	0.9574	0.44%	0.9516	0.17%

The suggested ANN-GA technique has the lowest error rate in comparison to ANN-GWO and ANN alone. The ANN-GA strategy effectively minimizes errors and enhances forecast accuracy across various weather circumstances, making it the most suitable option for predicting PV power generation. It consistently outperforms other proposed methods in all performance indicators. The MAE of the ANN-GA is lower than 7.77% compared to the regular ANN. On the other hand, the ANN-GWO demonstrates an improvement of 2.11% compared to the regular ANN. Similarly, the MSE shows a 9.08% improvement compared to ANN, however, the performance

of ANN-GWO decreases by 3.12% compared to ANN. When considering the RMSE, the ANN-GA model demonstrates a 4.65% enhancement over the ANN model. Conversely, the performance of the ANN-GWO model declines by 1.57% compared to the ANN model. Furthermore, about R2, ANN-GA showed a 0.44% enhancement compared to ANN, whereas the performance of ANN-GWO declined by 0.17% compared to ANN, as indicated in Table (4-9).



(a)



(b)

Figure 4-18(a) The Comparison between actual power and prediction power (b) Relative error with ANN, ANN-GA, and ANN-GWO in Sunny days

Furthermore, Figure (4-18) (a) displays the power generation of the PV plant during bright weather. It also compares the expected and real power output using the proposed models. The horizontal axis, known as the X-axis, depicts time in 3-minute intervals. The vertical axis, known as the Y-axis, indicates electrical power measured in watts, ranging from 0 to 450 watts. The true capacity can vary significantly throughout different periods, reaching its highest point on 7 June at 11:27 AM and its lowest point on 17 January at 5:03 PM. It is evident that all the suggested models consistently generate energy on a sunny day. The energy output exhibits an upward trend from 9:00 to 12:00, followed by a slow decline. These findings suggest that the precision of all models is comparatively greater on days with clear weather. In addition, Figure (4-18) (a) demonstrates that all of the proposed prediction models align well with the actual power curve. Overall, the enhanced models utilizing GA exhibit greater performance and accuracy by closely aligning with the real power levels in both high and low situations, unlike the basic ANN model. This underscores

the necessity of employing optimization techniques in ANN to improve the precision of forecasts in energy-related applications. Enhanced models effectively track both major and minor changes, making them efficient instruments for managing electricity.

Figure (4-18) (b) depicts the discrepancy between the predicted PV output and the actual PV generation on sunny days, expressed as the relative error. The X-axis displays the testing period separated into 3-minute intervals, while the Y-axis reflects the relative error percentage for three different methodological tests. Figure (4-18) (b) indicates that the ANN-GA forecasting model exhibits superior stability when compared to the ANN-GWO and ANN prediction models. Furthermore, the ANN technique demonstrates more significant and noticeable fluctuations, especially within the time intervals of 1/22 and 6/8, where the relative error surpasses 100% and drops below -100%. Therefore, the ANN-GA forecasting model obtains the lowest relative percentage error of approximately 4.5%, whereas the ANN-GWO and ANN prediction models have error rates of around 5.5% and 6.5%, respectively.

Based on the comparison, the ANN demonstrates higher accuracy in sunny weather conditions than the ANN-GWO model. Upon comparing the ANN-GWO model with the ANN model, it becomes apparent that the ANN-GWO model exhibits greater performance in sunny weather conditions when assessed using the MAE metric. Nevertheless, the analysis conducted using R2, MSE, and RMSE did not meet the desired threshold. It is worth noting that as periods shift, specific patterns may arise, which might have varying effects on the performance of each model. However, it has been confirmed that the ANN-GWO model performs better than ANN in situations marked by cloudiness and wetness.

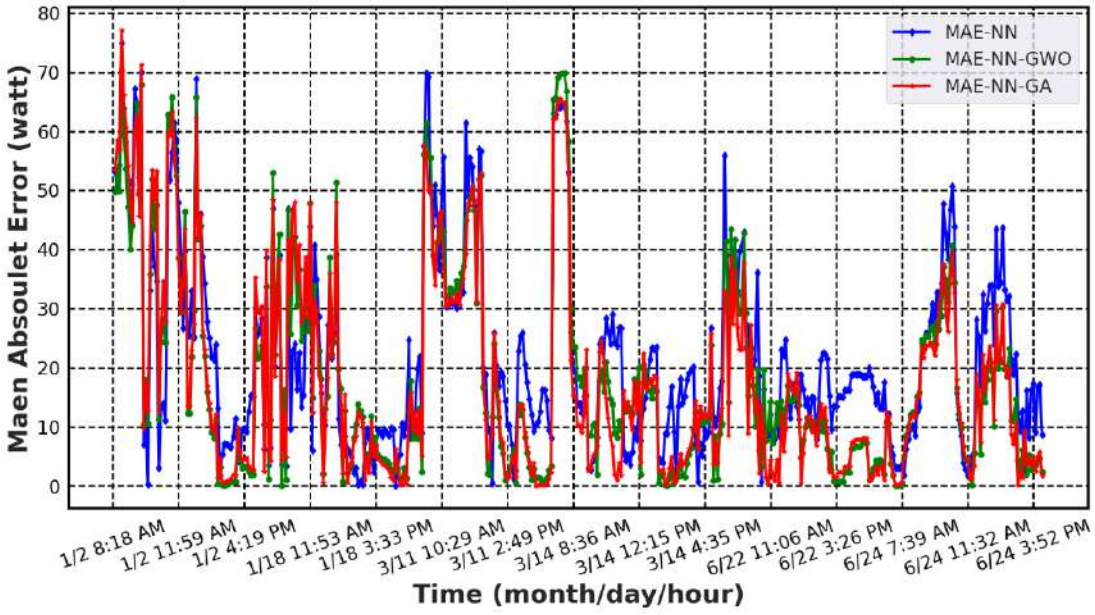
4.3.3. Cloudy Days

The purpose of this classification is to evaluate the effectiveness of the proposed prediction model in forecasting outcomes during cloudy weather conditions. A total of 1063 samples were gathered from 6 uniformly dispersed overcast days across three seasons. A total of 330 samples were collected on the 2nd and 18th of January to accurately reflect the winter season. Similarly, 356 samples were gathered on the 11th and 14th of March to represent the spring season, and 377 samples were obtained on the 22nd and 24th of June to accurately represent the summer season. During gloomy weather, the overall number of samples decreases compared to bright weather because clouds obscure sunlight, preventing solar panels from obtaining adequate light. Moreover, the sensor may lack sufficient sensitivity to precisely assess the extent of radiation it detects leading to some readings not being recorded.

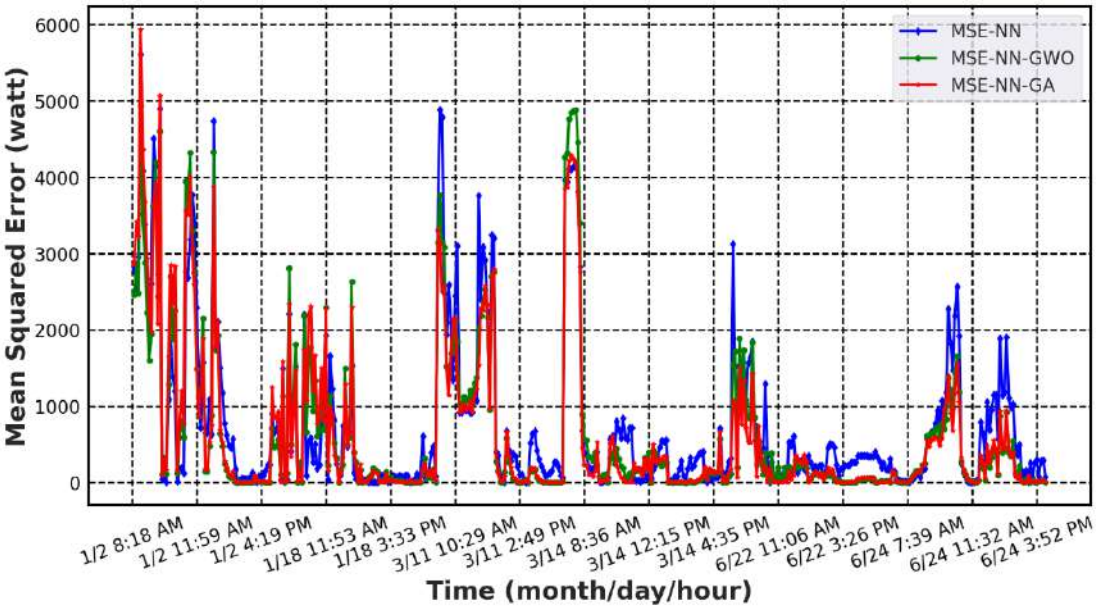
Figure (4-19) (a), the x-axis indicates time in 3-minute intervals. The y-axis represents the average absolute error in watts, ranging from 0 to 80 watts. On January 2, the MAE-NN error varied between 0 and 70 W. The MAE-NN-GA error spans a range of 0 to 60 watts. At the same time, the MAE-NN-GWO error spans a range of 0 to 70 watts. And between 18 January and 11 March, the MAE-NN error fluctuates between 0 and 60 watts. The MAE-NN-GA error spans a range of 0 to 50 watts. The MAE-NN-GWO model's inaccuracy range is between 0 and 55 watts. On cloudy days, solar radiation is variable and irregular, making the energy output more difficult to predict. Consequently, any abrupt fluctuations in solar radiation result in heightened inaccuracies in forecasts. The MAE-NN model exhibits inconsistent performance in the majority of periods, with error fluctuations ranging from 0 to 70 W. The MAE-NN-GA model outperforms MAE-NN, exhibiting an error range of 0 to 50 watts, this model has greater stability during periods characterized by overcast

weather conditions. MAE-NN-GWO exhibits a moderate level of performance, falling between MAE-NN and MAE-NN-GA, the error in this case varies from 0 to 55 watts. To summaries, the three models exhibit worse accuracy in overcast weather as compared to sunny days, mostly due to significant variations in solar energy and temperatures. Optimization strategies enhance the performance of models during cloudy conditions, resulting in improved accuracy and stability of forecasts.

Figure (4-19) (b) displays the time-dependent mean square error values (in watts) ranging from 0 to 6000 watts, with a frequency of 3 minutes. As stated before, on 2 January, the MSE-NN and MSE-NN-GWO models achieved an output of approximately 5000 watts, but the MSE-NN-GA model achieved an output of around 4000 watts. From 14 March to 22 June, all models consistently maintain a stability of approximately 1000 watts. It is worth mentioning that ANN-GA exhibited the highest inaccuracy during the beginning of period 2 and specifically on 18 January. Nevertheless, the model's performance subsequently enhanced, surpassing both ANN and ANN-GWO. The MAE is obtained by computing the mean value of the ANN-GA, ANN, and ANN-GWO models, which yield 17.8099 W, 21.8080 W, and 18.1780 W, respectively. The MSE values for the models are 615.0131 W, 744.8360 W, and 629.6826 W, respectively, as Table (4-10).



(a)



(b)

Figure 4-19 (a)The Result MAE of the cloudy Days, (b) The Result MSE of the cloudy Days

Table 4-10 Summarized forecasting results in Cloudy days for PV prediction model.

Days	Method	RMSE	MSE	MAE	R^2

Cloudy	ANN	27.2917	744.8360	21.8080	0.9209
	ANN-GA	24.7995	615.0131	17.8099	0.9347
	ANN-GWO	25.0934	629.6826	18.1780	0.9332

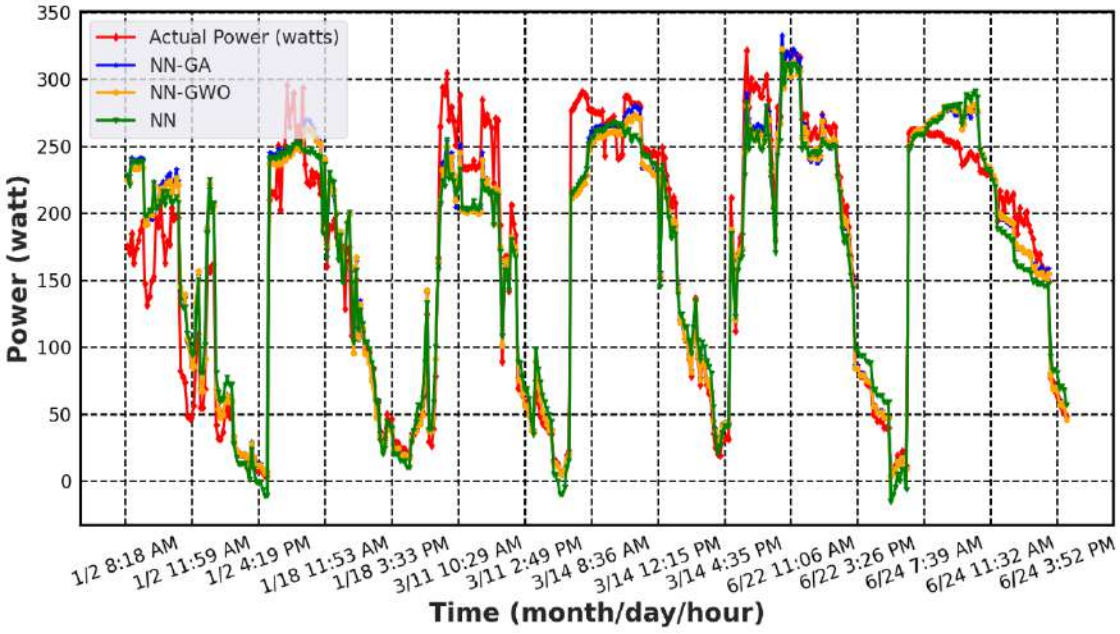
Table (4-10) presents a concise overview of the forecast findings specifically for cloudy weather conditions. The prediction outcomes are fully evaluated using MSE, RMSE, MAE, and R2. The MSE of the ANN-GA is 615.0131 Watts. The RMSE is 24.7995 Watts, and the MAE is 17.8099 Watts. The R2 values for the ANN-GA, ANN, and ANN-GWO models are presented in Table (5-4) as 0.9347, 0.9209, and 0.9332, respectively. The prediction result of ANN-GA is the most optimal, as demonstrated in Scheme (4-23) (f).

Furthermore, the ANN-GA approach consistently outperformed both the ANN and ANN-GWO methods across all performance indicators. The MAE value for ANN-GA has decreased by 18.34% compared to ANN. Similarly, the ANN-GWO method showed a notable improvement of 16.65% compared to the ANN method, as shown in Table (4-11). Regarding the R2 index, both ANN-GA and ANN-GWO exhibited enhancements of 1.48% and 1.32%, respectively, in comparison to ANN. Both ANN-GA and ANN-GWO exhibited a 17.43% and 15.46% enhancement in MSE as compared to ANN. In terms of the RMSE index, the ANN-GA approach showed a 9.13% enhancement over the ANN method, whilst the ANN-GWO method demonstrated an 8.05% improvement over the ANN method.

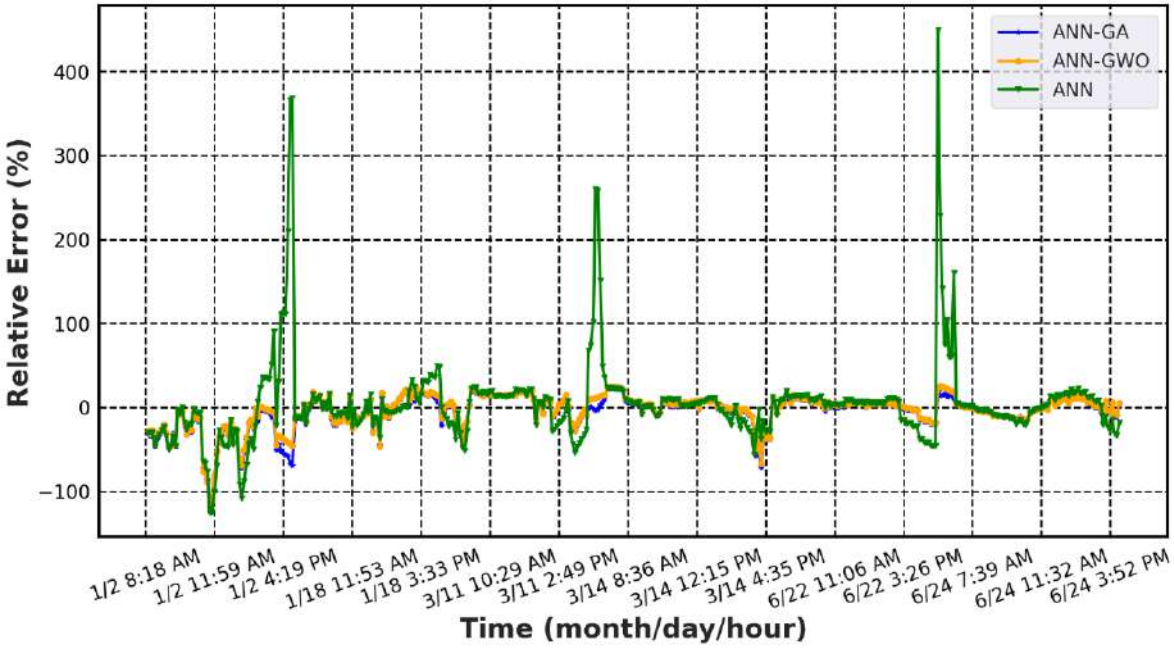
Table 4-11 Model Performance Comparison in Cloudy Weather

Indicator	ANN	ANN-GA	Ratio (%) for ANN-GA	ANN-GWO	Ratio (%) for ANN-GWO
MAE	21.8080	17.8099	18.34%	18.1780	16.65%
RMSE	27.2917	24.7995	9.13%	25.0934	8.05%
MSE	744.8360	615.0131	17.43%	629.6826	15.46%
R^2	0.9209	0.9347	1.48%	0.9332	1.32%

Figure (4-20) (a) depicts a comparative analysis of power generation under overcast weather conditions throughout time, divided into 3-minute intervals. The analysis includes the use of short-term ANN, ANN-GA, and ANN-GWO models to predict and compare the actual power generation. Noticeable surges in energy usage can be observed during specific periods, such as on January 18th at 12:19 p.m. These peaks are caused by increased energy demand. At some time intervals, such as 3/14 at 8:36 AM, discernible decreases in energy levels may be noticed. These periods arise due to reduced activity produced by cloud cover that obstructs sun energy. Furthermore, solar panels encounter reduced solar radiation under overcast conditions in contrast to clear and sunny days. Consequently, the generation of electricity from these solar panels declines during these certain time intervals. It can be concluded that the improved models NN-GA and NN-GWO exhibit marginally superior performance in comparison to the basic model ANN, hence emphasizing the significance of employing optimization techniques in neural networks to enhance prediction accuracy. These models can effectively monitor substantial fluctuations in real capacity, demonstrating their capability to generate precise predictions of future energy usage. Finally, during calm times, the models provide good predictions. However, when there are quick fluctuations, the ANN-GA and ANN-GWO algorithms exhibit superior performance with higher accuracy and greater adaptability to changes in weather conditions.



(a)



(b)

Figure 4-20 (a) The Comparison between actual power and prediction power (b) Relative error with ANN, ANN-GA, and ANN-GWO in Cloudy days

Figure (4-20) (b), displays several surprising changes, marked by significant increments and decrements in relative error when comparing the actual power and the forecasted power during cloudy days. The green line (ANN) exhibits significant volatility, characterized by substantial spikes and dips in relative error reaching up to 400% and -100% respectively, suggesting substantial inaccuracies in the forecast and revealing significant deficiencies in the projection. The blue line, representing (NN-GA), and the orange line, representing (NN-GWO), exhibit comparatively stable patterns with fewer significant changes when compared to the green line. The relative error of these methods ranges around zero, and most values fall between -100% and 100%. Therefore, it can be inferred that ANN-GA methods exhibit superior accuracy and stability in generating PV predictions when compared to ANN-GWO and ANN approaches. The ANN-GA prediction model achieves the lowest relative error of approximately 6.5%, whilst the ANN-GWO and ANN prediction models achieve 7% and 12.5% respectively. Although solar panels do not operate at full efficiency on cloudy days, they still generate some energy, but at a lower rate compared to clear days.

4.3.4. Rainy Days

The purpose of this classification is to evaluate the efficacy of the proposed forecasting model under rainy weather conditions. Specifically, 809 samples were taken on specific rainy days, with the collection period spanning 6 days and the samples uniformly dispersed over three seasons. To depict the winter season, deliberately selected the days with the most intense precipitation, namely 11th and 30th January, and gathered a total of 212 samples from these days. In addition, a total of 210 samples were gathered on 19th and 24th March to accurately represent the spring season. Furthermore, an additional 387 samples were collected on 10th

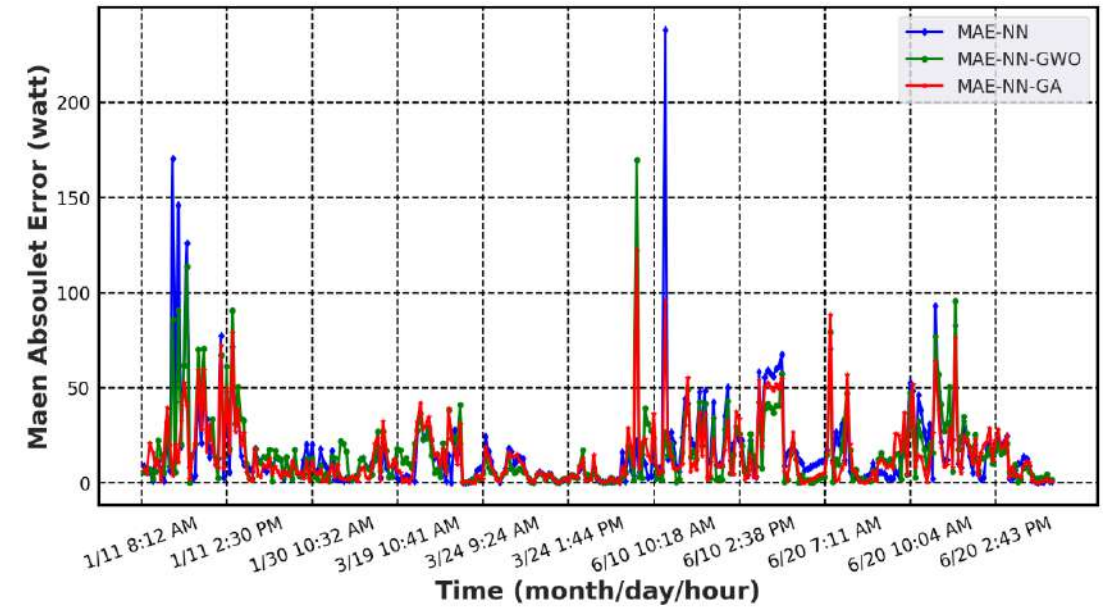
and 20th June to effectively depict the summer season. In January, the length of daylight hours fell due to rainy weather, and there were periods of time during the day where no measurements were taken. The lack of certain measurements in January and March resulted from intense precipitation, which led to the entire shutdown of the system for many hours due to rain-saturated clouds obstructing radiation. On 10 June, there was a slight precipitation caused by the increasing temperature. Furthermore, a day with overcast skies was also incorporated. Today marks the twentieth day of compensating for the lack of days with a mix of bright and cloudy weather. This is because June usually encounters strong sunlight and elevated temperatures in southern Iraq.

Table 4-12 Summarized forecasting results Rainy days for PV prediction model.

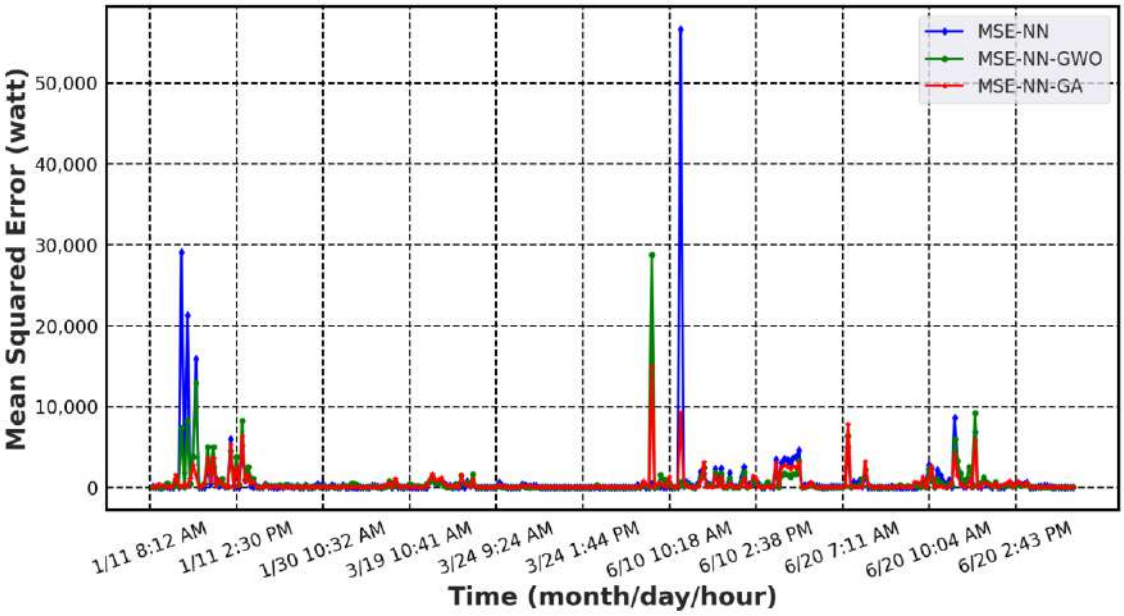
Days	Method	RMSE	MSE	MAE	R^2
Rainy	ANN	29.0796	845.6225	16.4599	0.8250
	ANN-GA	22.3597	499.9573	14.3067	0.8965
	ANN-GWO	25.4140	645.8738	15.8127	0.8663

Figure (4-21) (a) displays the MAE between the expected and actual power generation specifically in rainy weather circumstances. Figure (4-12) (a) depicts the MAE of power, measured in watts, which spans a range of 0 to 200 watts. The three models exhibit variability in their performance, occasionally converging during certain periods and diverging during others. The two methodologies, MAE-NN-GWO and MAE-NN-GA, exhibit comparable efficacy in the majority of time intervals, although MAE-NN displays certain inconsistencies, suggesting its lower adaptability to abrupt fluctuations in rainy conditions. According to the data presented in Figure (4-21) (a), it is evident from Table (4-12) that the MAE associated with ANN-GA is occasionally lower than both ANN and ANN-GWO.

On days with rain, the ANN-GA, ANN, and ANN-GWO had MAE values of 14.3067 W, 16.4599 W, and 15.8127 W, respectively.



(a)



(b)

Figure 4-21 (a)The Result MAE of the Rainy Days, (b) The Result MSE of the Rainy Days

Figure (4-21) (b), displays the difference in MSE between the projected and observed power generation specifically during rainy weather conditions. The horizontal axis, also known as the X-axis, shows time in 3-minute intervals. On the other hand, the vertical axis, or the Y-axis, represents the MSE of power measured in watts. The range of values on this axis goes from 0 to 50,000 watts. From the figure, certain periods show a clear increase in the MSE. The data in Figure (4-21) (b), show that the MSE for all models is nearly zero. However, there are two specific periods where the MSE deviates significantly. From 8 AM to 12 PM on 11 January, the MSE value for the ANN-GWO model reaches 30,000 W. Similarly, during the period from 10 to 11 AM on the morning of 10 June, there is a spike in the ANN model, with the MSE increasing up to 50,000 W, due to the excessive energy use during these periods, where the peak solar radiation level was recorded at 467.5926 w/m2 on 10 June at 10:32 a.m. On the other side, the solar irradiance level was recorded at 193.8658 w/m2, and 133.9699 w/m^2, between this point, in 3-minute intervalley. In this case, the fluctuation in the ANN prediction is due to this rapid change in solar irradiance. This sudden fluctuation in the input irradiance indicates a limitation in its ability to handle rapid variations in the input data. Furthermore, changes in solar radiation and temperatures caused by rainy weather contribute to higher absolute error and thus lead to inaccurate forecasts. Thus, GA-ANN is comparatively more effective than both ANN and ANN-GWO, as demonstrated in Table (4-12). It is important to highlight that GA regularly demonstrates higher performance in many settings. The MSE values for the ANN-GA, ANN, and ANN-GWO are 499.9573 W, 845.6225 W, and 645.8738 W, respectively. Table (4-12) presents the R2 values of the ANN-GA, ANN, and ANN-GWO, which are presented as 0.8965, 0.8250, and 0.8663, respectively. These values indicate a significantly weaker correlation compared to clear-sky conditions, as seen in Figure (4-23) (d)

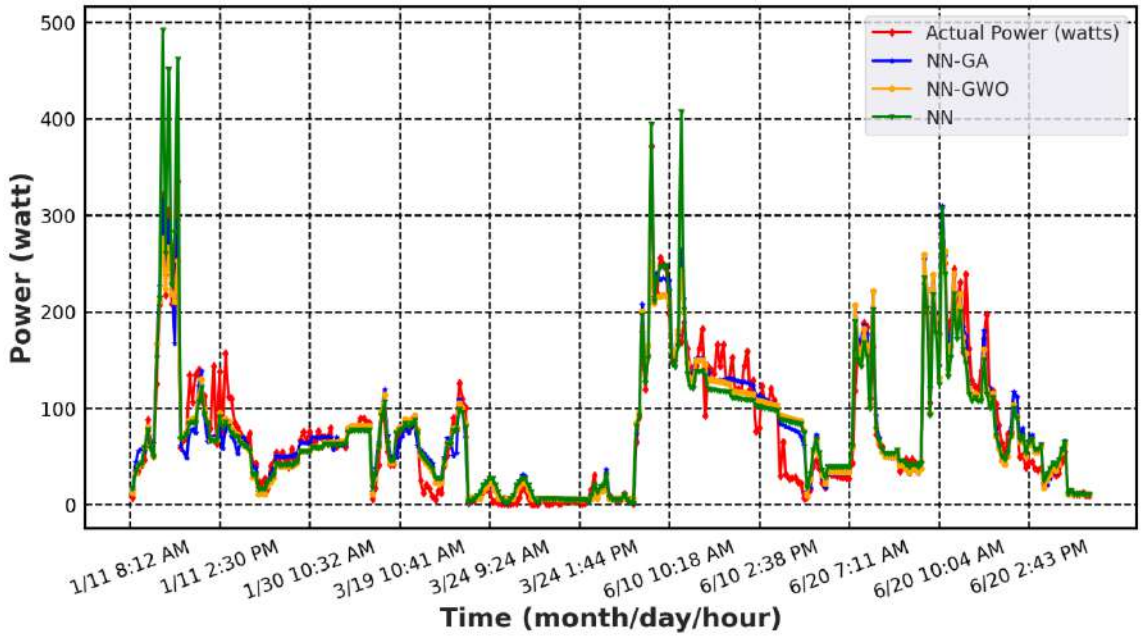
Figure (4-22) (a), displays the comparison between the real capacity and the produced capacity of the three models over a specified testing period. The horizontal axis, referred to in the figure, the time is split into intervals of 3 minutes. The vertical axis, sometimes known as the Power axis, represents the measurement of electrical power in watts. Figure (4-22) (a), illustrates that the forecasts generated by the ANN-GA, ANN-GWO, and ANN models aim to track variations in real power. These many models exhibit similarities in certain areas and display minor variations in other areas. Moreover, there are substantial enhancements in energy levels in particular instances, particularly on 11 January at 10:18 p.m. and 10 June at 10:18 a.m., due to an abrupt alteration in solar radiation. Consequently, there is a rise in energy use throughout these periods. In contrast, the 3/24 dates exhibit a reduced number of low regions due to the occurrence of continuous heavy rainfall throughout the day. Power generation was significantly affected on 24 March at 4:04 p.m. The solar radiation reached 19.8206 w/m^2. This decrease was a sunset result, which reduced the amount of solar energy reaching the solar panels, leading to a significant decrease in power generation. In addition, energy output is reduced in comparison to sunny and cloudy days as a result of the presence of rain and clouds, which might impede the solar radiation from reaching the solar panels. Furthermore, the presence of water droplets on solar panels can generate reflecting regions, leading to an increase in light reflection and a decrease in light absorption. This decrease can result in reduced solar energy efficiency on days with rainfall. Nevertheless, the projected values generated by the suggested model persistently align with actual energy measurements, and the prediction values with the ANN optimization models still provide a good prediction of the actual power measurement compared to the independent ANN. While Figure (4-22) (b), illustrates the correlation between time (x-axis) and relative error (y-axis) for three distinct models: ANN-GA (blue), ANN-

GWO (orange), and ANN (green). Time encompasses certain intervals inside the calendar months. Most of the time, modest relative errors close to zero are commonly found. However, there are occasional significant departures in the relative error, particularly for the ANN-GWO (in orange) and the ANN (in green). There is a distinct period, specifically on 03/24 at 9:24 am, when significant and abrupt downward changes occur. These oscillations result in a substantial reduction in the relative inaccuracy, with values as low as -4820.6% recorded on 24 March at 11:34 a.m. or the model ANN-GWO. During the same time frame, the ANN model (shown by the color green) has variances of approximately -4628.5%. Due to sudden changes in the environment in which the data was collected, leading to difficulty in forecasting accurately. As a result, the ANN-GA forecasting model acquires the lowest relative percentage error about 40%, while, the ANN and ANN-GWO prediction model reaches 44% and 56.5%, respectively.

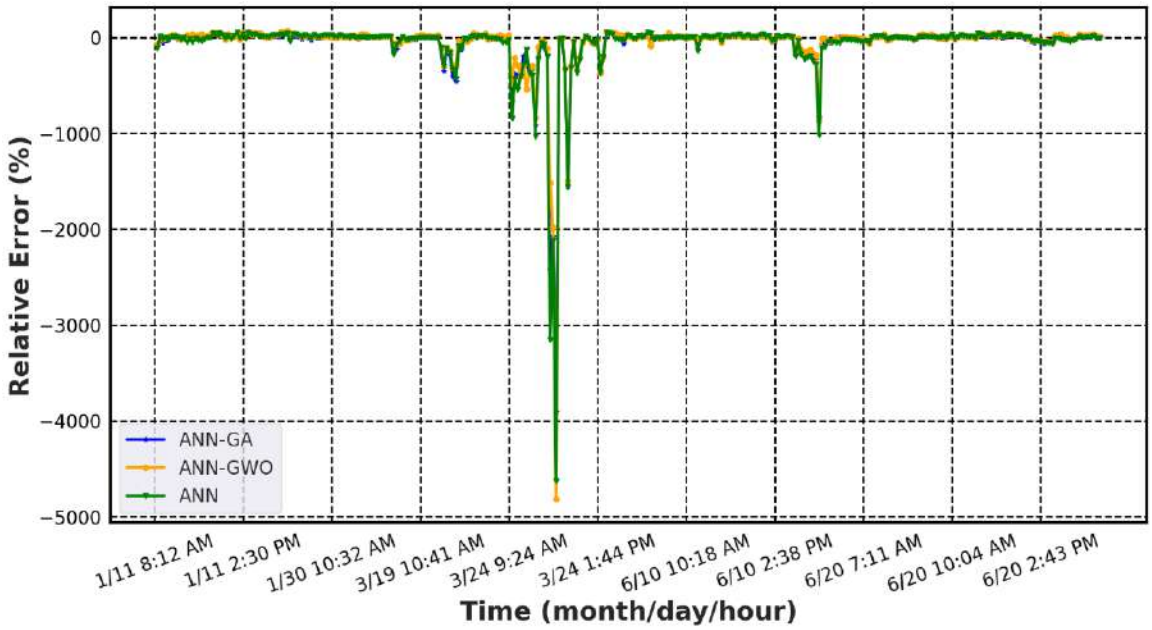
It was expected that the performance of the rainy and heavy cloud models would be lower than the rest of the models due to the greater variation in PV energy. However, the GA-ANN method maintained a superior MAE value of 13.08% of ANN while the ANN-GWO method, improved by 3.93% compared to ANN, and ANN-GA achieved 9.52% of ANN-GWO and performed better when finding the MAE value in cloudy weather. For R2, ANN-GA and ANN-GWO show 8.68% and 5.02% improvement over ANN respectively. Similarly, for MSE, ANN-GA, and ANN-GWO show 40.88% and 23.61% improvement over ANN respectively. The RMSE value of the ANN-GA method shows a 23.12% improvement compared to ANN while the ANN-GWO method shows a 12.61% improvement compared to ANN. Thus, ANN-GA is the best performer in all performance indicators (MAE, RMSE, R2, MSE) ,as shown in the model performance comparison Table(4-13).

Table 4-13 Model Performance Comparison in Rainy Weather

Indicator	ANN	ANN-GA	Ratio (%) for ANN-GA	ANN-GWO	Ratio (%) for ANN-GWO
MAE	16.4599	14.3067	13.09%	15.8127	3.93%
RMSE	29.0796	22.3597	23.12%	25.4140	12.61%
MSE	845.6225	499.9573	40.88%	645.8738	23.61%
R^2	0.8250	0.8965	8.68%	0.8663	5.02%



(a)



(b)

Figure 4-22 (a) The Comparison between actual power and prediction power (b) Relative error with ANN, ANN-GA, and ANN-GWO in Rainy days

Finally, Table (4-15) provides a summary of the comparison of PV energy in terms of relative percentage error for three classification tests. Improving PV prediction is essential for increasing the accuracy of weather forecasts on rainy and cloudy days. Consequently, the ANN has been created by the utilization of two distinct algorithms. Therefore, the ANN-GA prediction model demonstrated superior performance compared to other approaches in all weather conditions. This was evident from the much lower measured values, suggesting the impressive performance of this model. In addition, Table (4-14) contains the best number of hidden layers and neurons extracted from each algorithm along with the execution time.

Table 4-14 shows a comparison of results for the models used

Modeling	parameter	Sunny	Cloudy	Rainy
ANN-GA	Layer Number	1	3	3

	Hidden Neurons Number	49	[32,52,48]	[17,67,31]
	Time	40 Min	33 Min	32 Min
	Layer Number	1	1	4
ANN-GWO	Hidden Neurons Number	49	56	69
	Time	54 Min	56 Min	57 Min

Table 4-15 Relative percentage error for sunny, cloudy, and rainy days

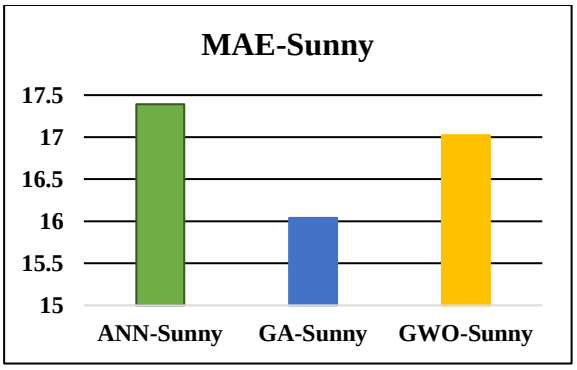
Method	Sunny days	Cloudy days	Rainy days
ANN-GA	4.5%	6.5%	40%
ANN-GWO	5.5%	7%	56.5%
ANN	6.5%	12.5%	44%

4.3.5. Discussion

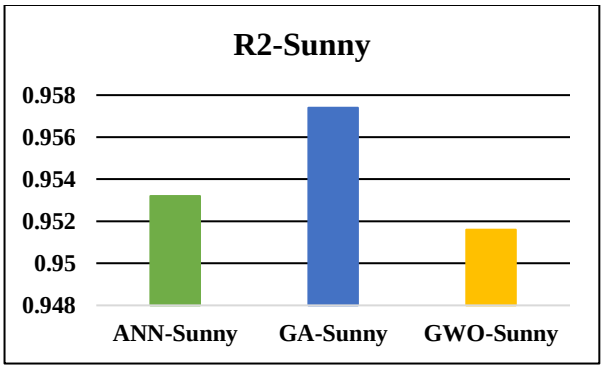
In general, a PV system is affected by production under variable weather conditions, as it highly depends on the amount of radiation. Therefore, on rainy days, the solar radiation was feeble, while on cloudy days, there were obvious fluctuations in solar radiation, which presented great challenges for the stable prediction of PV energy. Moreover, the performance accuracy on sunny days is relatively better than in other weather conditions. This is due to the nature of the weather conditions in Iraq. As a result, two machine training algorithms were implemented to evaluate the overall performance of the proposed neural network for PV energy prediction. Specifically, it identifies improvements in NN achieved through GA and GWO

across three distinct weather categories. Then the two algorithms were compared to ensure their performance in different weather conditions.

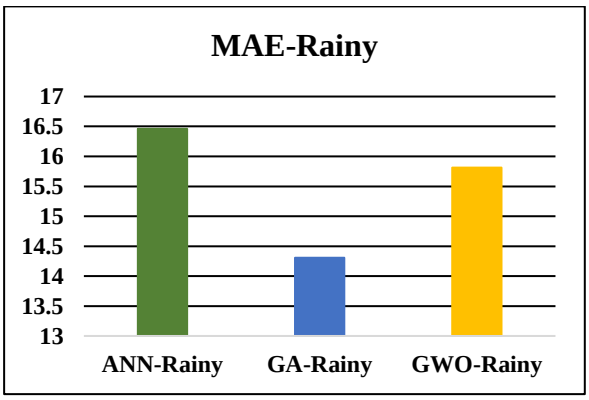
Based on the comparison condition, the traditional artificial neural network performs well and accurately during sunny weather conditions. However, the accuracy decreases in rainy and cloudy weather conditions. On the other hand, the ANN-GWO model shows high performance during all weather conditions (rainy, cloudy) compared to ANN. Therefore, ANN-GA shows the best performance in the weather conditions in Iraq, especially since Iraq weather has different conditions in different seasons of the year, this model can handle high performance during the weather conditions. In addition, ANN-GA has a good running speed compared to other models, for example, the running time for the ANN-GA model ranges from 32 to 40 minutes, whereas the ANN-GWO model takes from 54 to 57 minutes.



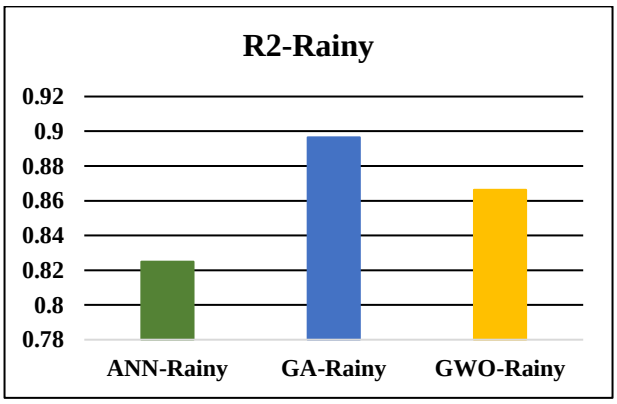
(a)



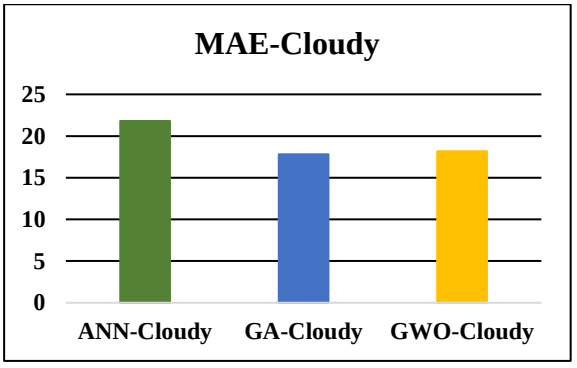
(b)



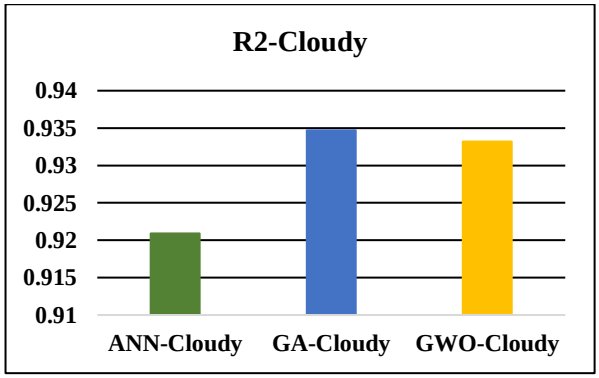
(c)



(d)



(e)



(f)

Figure 4-23 Comparison of MAE and R2 performance measurements across different meteorological seasons.

4.4. Comparison between Theoretical and Experimental Results

The thesis aims to predict PV power generation in the southern Iraq region, particularly in Misan Governorate, by analyzing weather conditions and enhancing ANN. The effectiveness of prediction models was evaluated using hourly weather data from 2021 and 3-minute actual data from 2024. For the theoretical results during (January, March, and September), the GA achieved the best performance among other methods, while in July the GWO achieved the best performance due to the nature of the climate in Misan province, which enjoys high solar radiation and temperature, so the GWO may be able to adapt better to high temperatures. On the other hand, September is considered one of the most accurate and performing months because the values of MAE, RMSE, MSE were the lowest and R2 was the highest for the ANN-GA method, and the improvement rate ranged between 2.51% to 32.83% for improving RMSE. Additionally, it ranged from 32.49% to 54.87% concerning the amelioration of MSE. Moreover, the improvement varied from 2.51% to 16.97% for the enhancement of MAE. Lastly, a range of 1.89% to 6.29% was observed for the enhancement of R2.

For the experimental results, the genetic algorithm was also the best in accuracy and performance compared to ANN, and ANN-GWO in all types of weather (sunny, rainy, and cloudy). In addition, sunny weather was the best in terms of overall model performance, achieving the highest value of R2 and the lowest value of MSE, RMSE, while rainy weather achieved the best performance in terms of MAE. In summary, the improvement percentage ranged between 1.17% to 23.12% for improving RMSE. Additionally, the range for the enhancement of MSE was observed to be between 2.33% and 40.89%. Furthermore, the improvement for MAE was quantified between 2.02% and 18.35%. Lastly, the improvement percentages

for R2 were found to span from 0.16% to 7.98%. Furthermore, Table (4-16) compares the methods used. The GA algorithm performs better in all weather conditions because it has proven effective in avoiding falling into local maximum solutions due to its random mechanisms, such as mutations. This makes it suitable for dealing with PV power production forecasting data.

In both cases, the experimental results (short-term) taken based on weather conditions (sunny, cloudy, and rainy) were generally better than the theoretical results (medium-term) measured by season (January, March, July, and September) based on the values of RMSE, MSE, MAE was lower in the experimental results and these values ranged from 72.3% to 94.6%, while the R2 value was close, but sometimes the experimental results were better. The difference was small, equivalent to about 1.75%, as seen in Table (4-17) and Figure (4-24), the comparison between experimental and theoretical results can be observed when comparing September with Sunny weather, this comparison was made based on the best result obtained from theoretical and experimental results. Therefore, the table reveals a large difference between theoretical and experimental results due to the difference in the data collection mechanism and time intervals, as well as due to the difference in the load used when collecting theoretical and experimental data. However, the performance was good in both cases, but the theoretical results do not take the rapid fluctuations in solar radiation resulting from cloud movement. In addition, the results appear ideal and more accurate and convergence than the experimental results, which are affected by the change in solar radiation as it is captured every 3 minutes. This shows that relying on experimental data, especially in fluctuating weather conditions, provides a more accurate and better representation of reality than theoretical models. Table (4-18) shows a comparison between experimental and theoretical results.

Table 4-16 The characteristics as a comparison between ANN, ANN-GA, and ANN-GWO

Characteristics	ANN Result	ANN-GA Result	ANN-GWO Result
Performance	Good (Satisfactory)	Best Performance	Very Good (but lower than ANN-GA)
Speed	Fastest	Medium Speed	Slowest
Accuracy	Good Accuracy	Highest Accuracy	High Accuracy (lower than ANN-GA)
Training Time	Shortest	Moderate	Longest
Model Complexity	Complexity increases with more layers and neurons	Complexity increases with more layers and neurons	Complexity increases with more layers and neurons
Python Implementation	Used	Used	Used
Flexibility	Less flexible	Highly flexible	Flexible but complex

Table 4-17 The comparison between the experimental and theoretical data in September and Sunny days

Methods	Parameter	Theoretical Result	Experimental Result	Difference	Percentage Error (%)
ANN	MAE	71.6435	17.3919	54.2516	76.04%
	RMSE	92.5994	21.5097	71.0897	76.8%
	R2	0.9244	0.9532	0.0288	3.05 %
	MSE	8574.6566	462.6680	8111.9886	94.6%
ANN-GA	MAE	57.9227	16.0403	41.8824	72.3%
	RMSE	82.0618	20.5089	61.5529	75.00%
	R2	0.9406	0.9574	0.0168	1.75%
	MSE	6734.1462	420.6154	6313.5308	93.75%
ANN-GWO	MAE	63.6414	17.0240	46.6174	73.25%
	RMSE	87.1780	21.8529	65.3251	74.93%
	R2	0.9329	0.9516	0.0187	1.97%
	MSE	7600.0110	477.5520	7122.459	93.72%

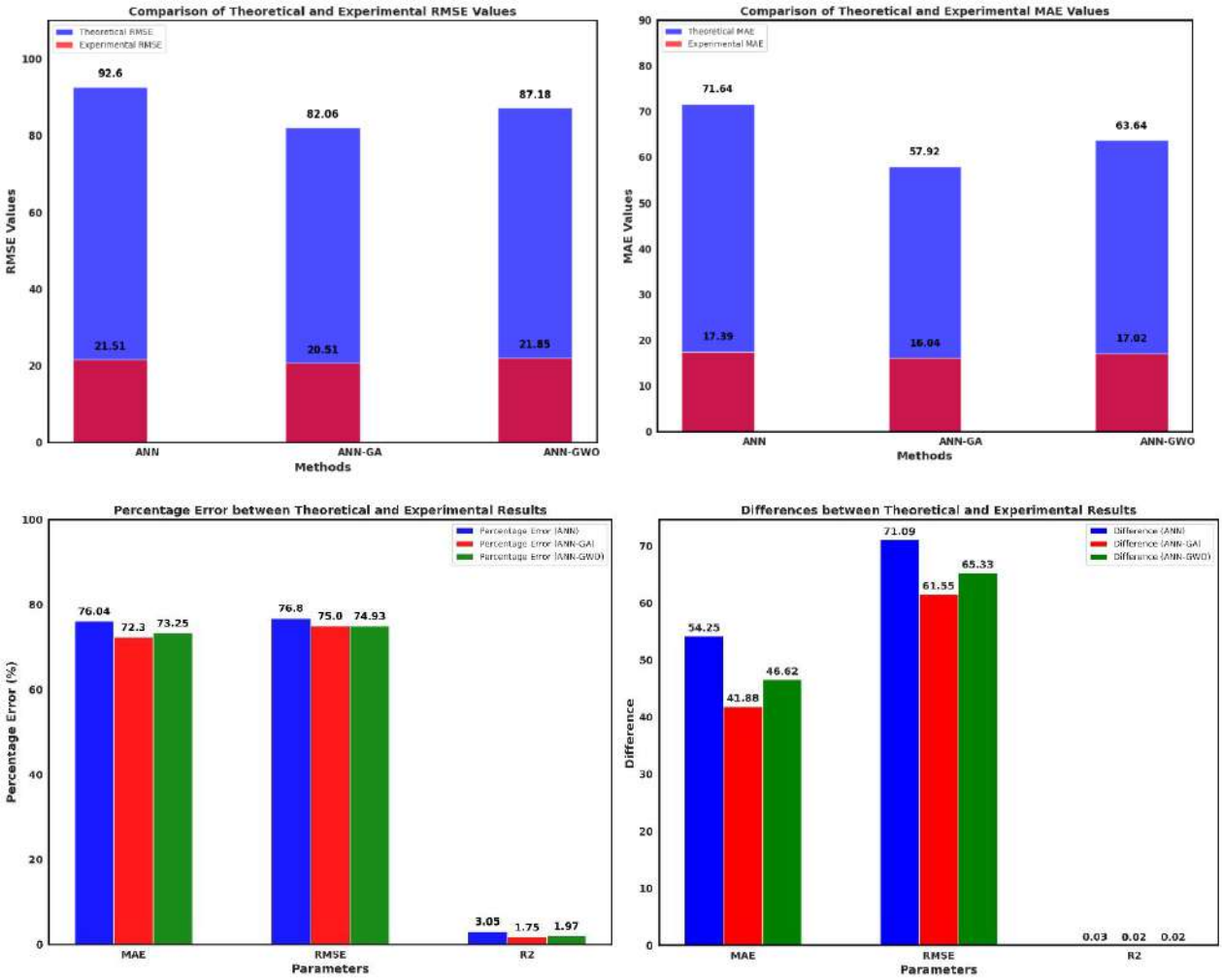


Figure 4-24 The comparison between the experimental and theoretical data in September and Sunny days

4.5. Summary

The enhanced ANN technique, which utilized the hidden layer and neuron optimization algorithm, was employed. The optimization was performed using two methods: the GA and the GWO. Consequently, the artificial neural network model experienced a decrease in the mean square error, MAE, and RMSE. Therefore, the results indicate that September has the highest performance, providing MAE

improvement from 2.51% to 16.97% compared to other month for theoretical data while the proposed MSE of ANN-GA model, which resulted in 40.88%, 9.08% and 18.34% improvement compared to ANN model for rainy, sunny and permanent days respectively for experimental data. Furthermore, experimental evidence has demonstrated that the utilization of ANN-GA for optimization yields superior performance. This approach effectively minimizes errors across all performance metrics and substantially enhances forecast accuracy across various weather circumstances, surpassing both ANN and ANN-GWO.

Table 4-18 The comparison between theoretical and experimental results

Characteristics	Theoretical Results	Experimental Results
Forecasting Time Interval	Hourly (Seasonal: January, March, July, September)	Every 3 minutes (Weather-based: Rainy, Cloudy, Sunny)
Production range Power(W)	0 to 1200	0 to 400
Forecasting Methods	ANN, ANN-GA, ANN-GWO	ANN, ANN-GA, ANN-GWO
Best Performing Method	ANN-GA	ANN-GA
MAE, MSE, RMSE, R^2 Values	Large (Not close to zero, but a good fit overall)	Small (Close to zero, better fit)
Accuracy	There is a good match between predicted and actual values, as data does not capture rapid changes in solar irradiance	Some slight fluctuations between predicted and actual values due to rapid changes in solar radiation.
Impact of Time Resolution	Longer time resolution obscure variations and rapid fluctuations in solar radiation	Shorter time resolution takes rapid fluctuations in solar radiation more effectively
Complexity of Prediction	Lower	Higher
Performance (Overall)	Good performance	Good performance

CHAPTER FIVE

Conclusion and Suggestions for Future Work

5.1. Conclusions

In this thesis, machine learning techniques were used to design the PV power forecasting model based on theoretical data and experimental data for the geographic areas of Iraq (Misan) for the first time. To sum up, the PV data is collected from multiple sources, including meteorological theoretical data and PV experimental data. Then, the effectiveness of artificial neural network technology in predicting PV power generation was investigated. Next, two optimization techniques, Genetic Algorithm (GA) and Grey Wolf Optimizer (GWO), were applied to enhance the ANN model's structure by optimizing the layers and neuron counts for improved prediction accuracy. The PV forecasting model is focused on the impact of solar radiation and temperature prediction at 3-minute and 1-hour timescales. Python is also used to obtain experimental and theoretical results in the process of minimizing MAE, MSE, RMSE, and R2. Data was processed seasonally, focusing on selected months (January, March, June, and September) to analyze theoretical results, and segmented by weather conditions (rainy, cloudy, sunny) to analyze experimental results. Results revealed that the ANN-GA model outperformed ANN and ANN-GWO in forecasting PV power under varying weather, demonstrating a superior ability to minimize errors across performance metrics. The experimental data, more accurate due to its sensitivity to weather changes, showed that ANN-GA achieved notable error reductions (MSE) under all weather conditions compared to ANN (9.08% in sunny, 17.40% in cloudy, and 40.88% in rainy conditions). Similarly, theoretical data showed MSE improvements across seasons for ANN-GA over

ANN, with notable gains in January (58.38%) and March (27.19%). Overall, the findings underscore that using GA for ANN optimization yielded more substantial performance improvements than GWO, underscoring the effectiveness of intelligent optimization techniques for achieving precise PV power predictions.

Current PV energy prediction models mainly consider solar radiation and temperature, overlooking other environmental factors like dust, humidity, and atmospheric pressure that impact solar panel performance. Future work could enhance prediction accuracy by integrating these factors, along with effects like partial shading and dust accumulation. Additionally, neural network performance plateaued after five training iterations, suggesting potential saturation. To address this, hybrid models combining optimization algorithms like ALO-MLP or WOA-MLP can improve weight tuning and predictive accuracy, offering a promising direction for model enhancement.

5.2. Future Directions

Future work will include several key directions:

1. Urban PV Production Prediction: Investigate energy production predictions for PV panels in urban areas, addressing specific challenges such as shading from buildings and pollution.
2. PV Expansion on Grid Efficiency: Analyze the impact of expanding PV systems on grid stability and efficiency, focusing on the effects of series and parallel configurations on voltage, current management, and load performance.
3. Long-Term Climate Change and Solar Prediction: Develop models to assess how long-term climate change forecasts affect the accuracy of solar radiation predictions, utilizing advanced forecasting methods.

4. Hybrid Prediction Models: Improve prediction accuracy by combining statistical and machine learning models, with real data from various locations to validate and compare model effectiveness.

REFERENCES

[1] P. Lin, Z. Peng, Y. Lai, S. Cheng, Z. Chen, and L. Wu, "Short-term power prediction for photovoltaic power plants using a hybrid improved Kmeans-GRA-Elman model based on multivariate meteorological factors and historical power datasets," *Energy Convers. Manag.*, vol. 177, no. July, pp. 704–717, 2018, doi: 10.1016/j.enconman.2018.10.015.

[2] P. System, U. Adaptive, P. Dawan, K. Sriprapha, S. Kittisontirak, and T. Boonraksa, "Comparison of Power Output Forecasting on the Inference Systems and Particle Swarm," *energies MDPI*, 2020.

[3] R. Luna-Rubio, M. Trejo-Perea, D. Vargas-Vázquez, and G. J. Ríos-Moreno, "Optimal sizing of renewable hybrids energy systems: A review of methodologies," *Sol. Energy*, vol. 86, no. 4, pp. 1077–1088, 2012, doi: 10.1016/j.solener.2011.10.016.

[4] R. Quadrelli and S. Peterson, "The energy-climate challenge: Recent trends in CO2 emissions from fuel combustion," *Energy Policy*, vol. 35, no. 11, pp. 5938–5952, 2007, doi: 10.1016/j.enpol.2007.07.001.

[5] Y. Li, Y. Su, and L. Shu, "An ARMAX model for forecasting the power output of a grid connected photovoltaic system," *Renew. Energy*, vol. 66, pp. 78–89, 2014, doi: 10.1016/j.renene.2013.11.067.

[6] A. Bracale, G. Carpinelli, and P. De Falco, "A Probabilistic Competitive Ensemble Method for Short-Term Photovoltaic Power Forecasting," *IEEE Trans. Sustain. Energy*, vol. 8, no. 2, pp. 551–560, 2017, doi: 10.1109/TSTE.2016.2610523.

[7] M. AlShafeey and C. Csáki, "Evaluating neural network and linear regression photovoltaic power forecasting models based on different input methods," *Energy Reports*, vol. 7, pp. 7601–7614, 2021, doi: 10.1016/j.egyr.2021.10.125.

[8] G. Codato, A. P. Oliveira, J. Soares, J. F. Escobedo, E. N. Gomes, and A. D. Pai, "Global and diffuse solar irradiances in urban and rural areas in southeast Brazil," *Theor. Appl. Climatol.*, vol. 93, no. 1–2, pp. 57–73, 2008, doi: 10.1007/s00704-007-0326-0.

[9] M. Z. Jacobson and M. A. Delucchi, "Providing all global energy with wind,

water, and solar power, Part I: Technologies, energy resources, quantities and areas of infrastructure, and materials," *Energy Policy*, vol. 39, no. 3, pp. 1154–1169, 2011, doi: 10.1016/j.enpol.2010.11.040.

[10] U. K. Das *et al.*, "Forecasting of photovoltaic power generation and model optimization: A review," *Renew. Sustain. Energy Rev.*, vol. 81, no. January, pp. 912–928, 2018, doi: 10.1016/j.rser.2017.08.017.

[11] S. D. Al-Majidi, M. F. Abbod, and H. S. Al-Raweshidy, "A particle swarm optimisation-trained feedforward neural network for predicting the maximum power point of a photovoltaic array," *Eng. Appl. Artif. Intell.*, vol. 92, no. September 2019, p. 103688, 2020, doi: 10.1016/j.engappai.2020.103688.

[12] A. Thesis, S. To, and T. H. E. Council, "Multilevel Inverters with Reduced Switches for PV System in AC Microgrid," no. October, 2023.

[13] J. C. Mojumder, H. C. Ong, W. T. Chong, N. Izadyar, and S. Shamshirband, "The intelligent forecasting of the performances in PV/T collectors based on soft computing method," *Renew. Sustain. Energy Rev.*, vol. 72, no. August 2015, pp. 1366–1378, 2017, doi: 10.1016/j.rser.2016.11.225.

[14] C. Zhang and M. Zhang, "Wavelet-based neural network with genetic algorithm optimization for generation prediction of PV plants," *Energy Reports*, vol. 8, pp. 10976–10990, 2022, doi: 10.1016/j.egyr.2022.08.176.

[15] U. K. Das *et al.*, "Forecasting of photovoltaic power generation and model optimization: A review," *Renew. Sustain. Energy Rev.*, vol. 81, no. August 2017, pp. 912–928, 2018, doi: 10.1016/j.rser.2017.08.017.

[16] I. P. Panapakidis and G. C. Christoforidis, "A hybrid ANN/GA/ANFIS model for very short-term PV power forecasting," *2017 11th IEEE Int. Conf. Compat. Power Electron. Power Eng. CPE-POWERENG 2017*, no. August, pp. 412–417, 2017, doi: 10.1109/CPE.2017.7915207.

[17] G. Li, S. Xie, B. Wang, J. Xin, Y. Li, and S. Du, "Photovoltaic Power Forecasting with a Hybrid Deep Learning Approach," *IEEE Access*, vol. 8, pp. 175871–175880, 2020, doi: 10.1109/ACCESS.2020.3025860.

[18] M. Pan *et al.*, "Photovoltaic power forecasting based on a support vector machine with improved ant colony optimization," *J. Clean. Prod.*, vol. 277, p. 123948, 2020, doi: 10.1016/j.jclepro.2020.123948.

[19] P. Li, K. Zhou, X. Lu, and S. Yang, "A hybrid deep learning model for short-

term PV power forecasting," *Appl. Energy*, vol. 259, no. July, p. 114216, 2020, doi: 10.1016/j.apenergy.2019.114216.

[20] M. Trabelsi *et al.*, "An Effective Hybrid Symbolic Regression–Deep Multilayer Perceptron Technique for PV Power Forecasting," *Energies*, vol. 15, no. 23, pp. 1–14, 2022, doi: 10.3390/en15239008.

[21] L. O. Lara-Cerecedo, J. F. Hinojosa, N. Pitalúa-Díaz, Y. Matsumoto, and A. González-Angeles, "Prediction of the Electricity Generation of a 60-kW Photovoltaic System with Intelligent Models ANFIS and Optimized ANFIS-PSO," *Energies*, vol. 16, no. 16, 2023, doi: 10.3390/en16166050.

[22] X. Pan, J. Zhou, X. Sun, Y. Cao, X. Cheng, and H. Farahmand, "A hybrid method for day-ahead photovoltaic power forecasting based on generative adversarial network combined with convolutional autoencoder," *IET Renew. Power Gener.*, vol. 17, no. 3, pp. 644–658, 2023, doi: 10.1049/rpg2.12619.

[23] E. Sauter, M. Mughal, and Z. Zhang, "Evaluation of Machine Learning Methods on Large-Scale Spatiotemporal Data for Photovoltaic Power Prediction," *Energies*, vol. 16, no. 13, 2023, doi: 10.3390/en16134908.

[24] Y. K. Semero, J. Zhang, and D. Zheng, "PV power forecasting using an integrated GA-PSO-ANFIS approach and Gaussian process regression based feature selection strategy," *CSEE J. Power Energy Syst.*, vol. 4, no. 2, pp. 210–218, 2018, doi: 10.17775/cseejpes.2016.01920.

[25] W. Bendali, I. Saber, B. Bourachdi, M. Boussetta, and Y. Mourad, "Deep Learning Using Genetic Algorithm Optimization for Short Term Solar Irradiance Forecasting," *4th Int. Conf. Intell. Comput. Data Sci. ICDS 2020*, 2020, doi: 10.1109/ICDS50568.2020.9268682.

[26] B. Carrera, M. K. Sim, and J. Y. Jung, "PVHybNet: A hybrid framework for predicting photovoltaic power generation using both weather forecast and observation data," *IET Renew. Power Gener.*, vol. 14, no. 12, pp. 2192–2201, 2020, doi: 10.1049/iet-rpg.2018.6174.

[27] R. Ben Ammar, M. Ben Ammar, and A. Oualha, "Comparative study among physical models and artificial intelligence methods for PV power forecasting," *18th IEEE Int. Multi-Conference Syst. Signals Devices, SSD 2021*, pp. 1038–1046, 2021, doi: 10.1109/SSD52085.2021.9429452.

[28] M. Konstantinou, S. Peratikou, and A. G. Charalambides, "Solar photovoltaic forecasting of power output using lstm networks," *Atmosphere (Basel).*, vol.

12, no. 1, pp. 1–17, 2021, doi: 10.3390/atmos12010124.

[29] Y. W. Liu, H. Feng, H. Y. Li, and L. L. Li, "An improved whale algorithm for support vector machine prediction of photovoltaic power generation," *Symmetry (Basel).*, vol. 13, no. 2, pp. 1–26, 2021, doi: 10.3390/sym13020212.

[30] Y. Wang, B. Feng, Q. S. Hua, and L. Sun, "Short-term solar power forecasting: A combined long short-term memory and gaussian process regression method," *Sustain.*, vol. 13, no. 7, 2021, doi: 10.3390/su13073665.

[31] A. Geetha *et al.*, "Prediction of hourly solar radiation in Tamil Nadu using ANN model with different learning algorithms," *Energy Reports*, vol. 8, pp. 664–671, 2022, doi: 10.1016/j.egyr.2021.11.190.

[32] S. Netsanet, D. Zheng, W. Zhang, and G. Teshager, "Short-term PV power forecasting using variational mode decomposition integrated with Ant colony optimization and neural network," *Energy Reports*, vol. 8, pp. 2022–2035, 2022, doi: 10.1016/j.egyr.2022.01.120.

[33] K. O. Adeyemi, V. Eniola, G. M. Kalu-Uka, M. Zarmai, M. Uthman, and E. Bala, "Forecasting Photovoltaic Energy Generation Using Multilayer Perceptron Neural Network," *Int. J. Renew. Energy Res.*, vol. 12, no. 4, pp. 1742–1753, 2022, doi: 10.20508/ijrer.v12i4.13306.g8599.

[34] Z. Yang, M. Mourshed, K. Liu, X. Xu, and S. Feng, "A novel competitive swarm optimized RBF neural network model for short-term solar power generation forecasting," *Neurocomputing*, vol. 397, pp. 415–421, 2020, doi: 10.1016/j.neucom.2019.09.110.

[35] A. Ziane *et al.*, "Photovoltaic output power performance assessment and forecasting: Impact of meteorological variables," *Sol. Energy*, vol. 220, no. April, pp. 745–757, 2021, doi: 10.1016/j.solener.2021.04.004.

[36] N. R. Zhou, Y. Zhou, L. H. Gong, and M. L. Jiang, "Accurate prediction of photovoltaic power output based on long short-term memory network," *IET Optoelectron.*, vol. 14, no. 6, pp. 399–405, 2020, doi: 10.1049/iet-opt.2020.0021.

[37] H. Aprillia, H. T. Yang, and C. M. Huang, "Short-term photovoltaic power forecasting using a convolutional neural network-salp swarm algorithm," *Energies*, vol. 13, no. 8, 2020, doi: 10.3390/en13081879.

[38] B. Ray, R. Shah, M. R. Islam, and S. Islam, "A New Data Driven Long-Term

Solar Yield Analysis Model of Photovoltaic Power Plants," *IEEE Access*, vol. 8, pp. 136223–136233, 2020, doi: 10.1109/ACCESS.2020.3011982.

[39] M. Alaraj, A. Kumar, I. Alsaidan, M. Rizwan, and M. Jamil, "Energy Production Forecasting from Solar Photovoltaic Plants Based on Meteorological Parameters for Qassim Region, Saudi Arabia," *IEEE Access*, vol. 9, pp. 83241–83251, 2021, doi: 10.1109/ACCESS.2021.3087345.

[40] M. Alaraj, I. Alsaidan, A. Kumar, M. Rizwan, and M. Jamil, "Advanced Intelligent Approach for Solar PV Power Forecasting Using Meteorological Parameters for Qassim Region, Saudi Arabia," *Sustain.*, vol. 15, no. 12, 2023, doi: 10.3390/su15129234.

[41] Z. Wang, "SOLAR POWER FORECASTING A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy in the School of Computer Science at The University of Sydney," *Univ. Sydney*, no. September, 2019.

[42] J. G. Kim, D. H. Kim, W. S. Yoo, J. Y. Lee, and Y. B. Kim, "Daily prediction of solar power generation based on weather forecast information in Korea," *IET Renew. Power Gener.*, vol. 11, no. 10, pp. 1268–1273, 2017, doi: 10.1049/iet-rpg.2016.0698.

[43] A. I. M. Ali, M. A. Sayed, and E. E. M. Mohamed, "Modified efficient perturb and observe maximum power point tracking technique for grid-tied PV system," *Int. J. Electr. Power Energy Syst.*, vol. 99, no. December 2017, pp. 192–202, 2018, doi: 10.1016/j.ijepes.2017.12.029.

[44] S. M. Radhi, S. D. Al-majidi, M. F. Abbod, and H. S. Al-raweshidy, "Predicting Solar Power Generation Utilized in Iraq Power Grid Using Neural Network," vol. 3, no. 1, 2024.

[45] F. A. Abbas, A. A. Obed, M. A. Qasim, S. J. Yaqoob, and S. Ferahtia, "An efficient energy-management strategy for a DC microgrid powered by a photovoltaic/fuel cell/battery/supercapacitor," *Clean Energy*, vol. 6, no. 6, pp. 827–839, 2022, doi: 10.1093/ce/zkac063.

[46] Z. Waradzyn, R. Stala, A. Mondzik, A. Penczek, A. Skala, and S. Pirog, "Efficiency Analysis of MOSFET-Based Air-Choke Resonant DC-DC Step-Up Switched-Capacitor Voltage Multipliers," *IEEE Trans. Ind. Electron.*, vol. 64, no. 11, pp. 8728–8738, 2017, doi: 10.1109/TIE.2017.2698368.

[47] M. A. Alqarni, "A High Efficiency Photovoltaic Inverter System

Configuration with Maximum Power Point Tracking," no. March, 2016.

[48] W. Hart Danial, *Commonly used Power and Converter Equations*. 2010.

[49] S. D. Al-Majidi, M. F. Abbod, and H. S. Al-Raweshidy, "Design of an intelligent MPPT based on ANN using a real photovoltaic system data," *2019 54th Int. Univ. Power Eng. Conf. UPEC 2019 - Proc.*, pp. 1–6, 2019, doi: 10.1109/UPEC.2019.8893638.

[50] M. R. Nugraha and A. Adriansyah, "Optimization of sensor model for solar radiation measurement with a pyranometer," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 739, no. 1, 2021, doi: 10.1088/1755-1315/739/1/012080.

[51] V. K. Yadav, S. Tyagi, G. Kumar, N. Kumar, and S. Namekar, "Automatic Car Washing Using PLC," *Int. J. Eng. Sci. Comput.*, vol. 3, no. 9, pp. 40–43, 2016, [Online]. Available: http://ijesc.org/

[52] P. Forecasting, "Power Forecasting," pp. 1–23, 2024.

[53] O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, N. A. E. Mohamed, and H. Arshad, "State-of-the-art in artificial neural network applications: A survey," *Heliyon*, vol. 4, no. 11, p. e00938, 2018, doi: 10.1016/j.heliyon.2018.e00938.

[54] Z. Pang, F. Niu, and Z. O'Neill, "Solar radiation prediction using recurrent neural network and artificial neural network: A case study with comparisons," *Renew. Energy*, vol. 156, pp. 279–289, 2020, doi: 10.1016/j.renene.2020.04.042.

[55] A. K. Jain, J. Mao, and K. M. Mohiuddin, "Artificial neural networks: A tutorial," *Computer (Long. Beach. Calif).*, vol. 29, no. 3, pp. 31–44, 1996, doi: 10.1109/2.485891.

[56] M. A. Boyacioglu, Y. Kara, and Ö. K. Baykan, "Predicting bank financial failures using neural networks, support vector machines and multivariate statistical methods: A comparative analysis in the sample of savings deposit insurance fund (SDIF) transferred banks in Turkey," *Expert Syst. Appl.*, vol. 36, no. 2 PART 2, pp. 3355–3366, 2009, doi: 10.1016/j.eswa.2008.01.003.

[57] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Networks*, vol. 2, no. 5, pp. 359–366, 1989, doi: 10.1016/0893-6080(89)90020-8.

[58] H. Li, J. Wang, M. Tang, and X. Li, "Polarization-dependent effects of an Airy beam due to the spin-orbit coupling," *J. Opt. Soc. Am. A Opt. Image Sci. Vis.*, vol. 34, no. 7, pp. 1114–1118, 2017, doi: 10.1002/ecs2.1832.

[59] G. R. Yang and X. J. Wang, "Artificial Neural Networks for Neuroscientists: A Primer," *Neuron*, vol. 107, no. 6, pp. 1048–1070, 2020, doi: 10.1016/j.neuron.2020.09.005.

[60] A. Aksoy, Y. E. Ertürk, S. Erdoğan, E. Eyduran, and M. M. Tariq, "Estimation of honey production in beekeeping enterprises from eastern part of Turkey through some data mining algorithms," *Pak. J. Zool.*, vol. 50, no. 6, pp. 2199–2207, 2018, doi: 10.17582/journal.pjz/2018.50.6.2199.2207.

[61] R. Akkaya, A. A. Kulaksiz, and Ö. Aydoğdu, "DSP implementation of a PV system with GA-MLP-NN based MPPT controller supplying BLDC motor drive," *Energy Convers. Manag.*, vol. 48, no. 1, pp. 210–218, 2007, doi: 10.1016/j.enconman.2006.04.022.

[62] H. Hamdi, C. Ben Regaya, and A. Zaafouri, "Real-time study of a photovoltaic system with boost converter using the PSO-RBF neural network algorithms in a MyRio controller," *Sol. Energy*, vol. 183, no. February, pp. 1–16, 2019, doi: 10.1016/j.solener.2019.02.064.

[63] M. Mohammadhassani, H. Nezamabadi-Pour, M. Z. Jumaat, M. Jameel, and A. M. S. Arumugam, "Application of artificial neural networks (ANNs) and linear regressions (LR) to predict the deflection of concrete deep beams," *Comput. Concr.*, vol. 11, no. 3, pp. 237–252, 2013, doi: 10.12989/cac.2013.11.3.237.

[64] {Holland, John H.}, "No Title Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence}," no. MIT Press, p. {159-170},.

[65] D. Saadaoui, M. Elyaqouti, K. Assalaou, D. Ben hmamou, and S. Lidaighbi, "Parameters optimization of solar PV cell/module using genetic algorithm based on non-uniform mutation," *Energy Convers. Manag. X*, vol. 12, no. July, p. 100129, 2021, doi: 10.1016/j.ecmx.2021.100129.

[66] S. N. Sivanandam and S. N. Deepa, *Introduction to genetic algorithms*. 2008. doi: 10.1007/978-3-540-73190-0.

[67] A. B. Hassanat, V. B. S. Prasath, M. A. Abbadi, S. A. Abu-Qdari, and H. Faris, "An improved Genetic Algorithm with a new initialization mechanism based

on Regression techniques," *Inf.*, vol. 9, no. 7, 2018, doi: 10.3390/info9070167.

[68] D. Gupta and S. Ghafir, "An Overview of methods maintaining Diversity in Genetic Algorithms," *Int. J. Emerg. Technol. Adv. Eng.*, vol. 2, no. 5, pp. 56–60, 2012.

[69] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey Wolf Optimizer," *Adv. Eng. Softw.*, vol. 69, pp. 46–61, 2014, doi: 10.1016/j.advengsoft.2013.12.007.

[70] N. Mittal, U. Singh, and B. S. Sohi, "Modified Grey Wolf Optimizer for Global Engineering Optimization," *Appl. Comput. Intell. Soft Comput.*, vol. 2016, 2016, doi: 10.1155/2016/7950348.

[71] S. K. Mosavi, E. Jalalian, and F. S. Gharahchopog, "a Comprehensive Survey of Grey Wolf Optimizer Algorithm and Its Application," *Int. J. Adv. Robot. Expert Syst.*, vol. 1, no. 6, pp. 23–45, 2021.

[72] A. I. Lawah, A. A. Ibrahim, S. Q. Salih, H. S. Alhadawi, and P. S. Josephng, "Grey Wolf Optimizer and Discrete Chaotic Map for Substitution Boxes Design and Optimization," *IEEE Access*, vol. 11, pp. 42416–42430, 2023, doi: 10.1109/ACCESS.2023.3266290.

[73] M. Colak, M. Yesilbudak, and R. Bayindir, "Daily photovoltaic power prediction enhanced by hybrid GWO-MLP, ALO-MLP and WOA-MLP models using meteorological information," *Energies*, vol. 13, no. 4, 2020, doi: 10.3390/en13040901.

Appendix A

The Simulink model of the PV system is linked to the electrical grid.

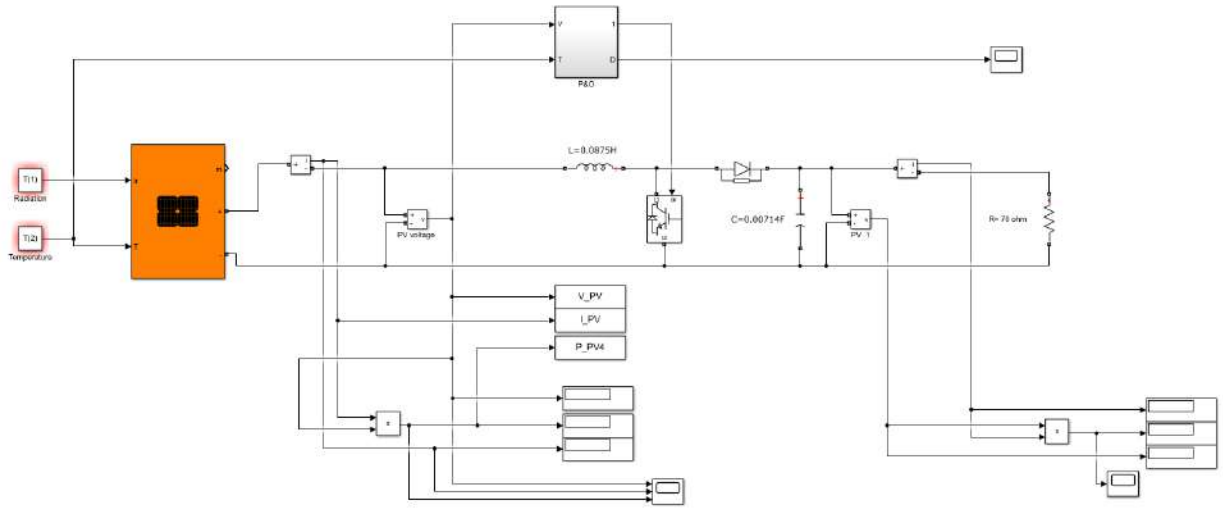


Figure B.1. The Simulink model of the PV system is linked to the electrical grid based on MATLAB simulation

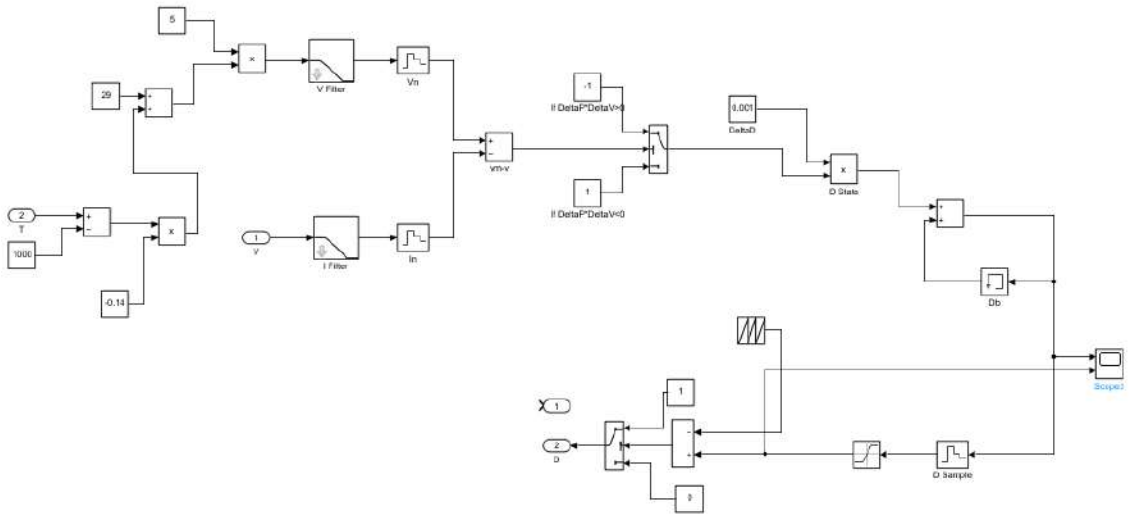


Figure B.2. P&O algorithm

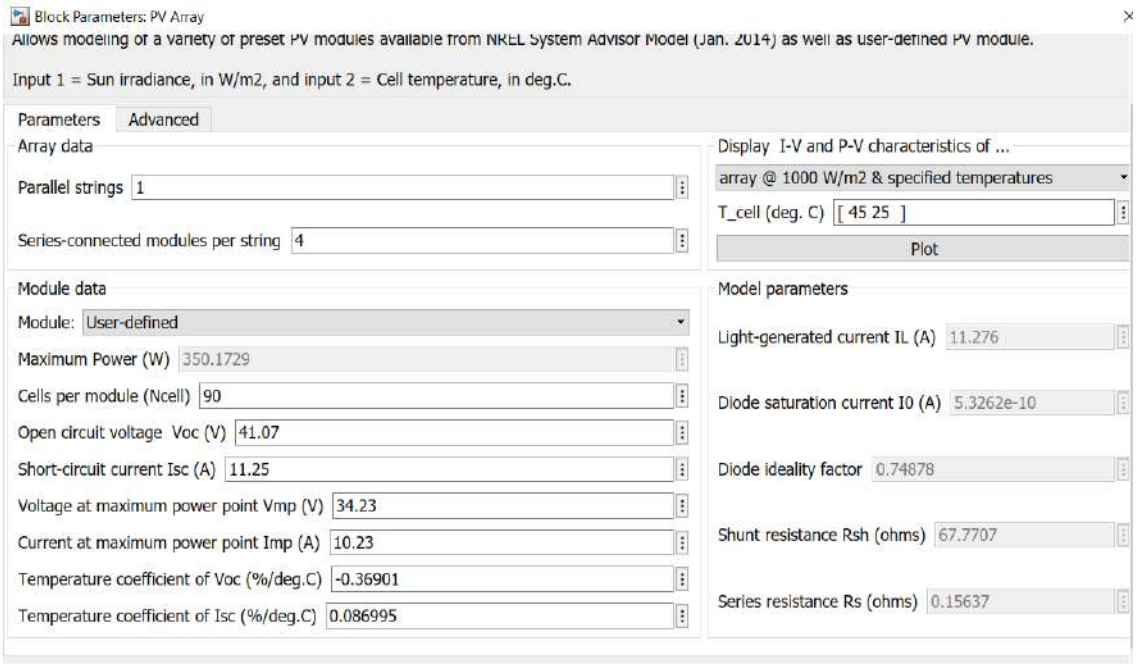


Figure B.3. Block parameters PV array

Code Used to Collected Data in Workspace

```
clc
clear all
Te=xlsread('name file','sheet1','C2:D1159');
for i=1:n
    T=Te(i,:);
    assignin('base','Ir',T);
    sim('PV.slx');
    %V(i)= V_PV4(length(V_PV4));
    %I(i)= I_PV4(length(I_PV4));
    Ve(i)= V_PV(length(V_PV));
    Ie(i)= I_PV(length(I_PV));
    P(i)= P_PV4(length(P_PV4));
end
```

Appendix B

Theoretical Data for January

Date/Time	Radiation	Temperature	voltage	current	power
01/01 08:00:00	428.5125	9.6125	161.744583	2.30079293	372.1407929
01/01 09:00:00	643.6	9.74	166.6279923	2.11671074	352.7032608
01/01 10:00:00	751.3125	10.6575	158.1940481	5.889781171	931.728326
01/01 11:00:00	783.2125	11.01	159.7334388	5.660780009	904.2158572
01/01 12:00:00	782.9	12.5125	151.6820686	7.244240612	1098.821401
01/01 13:00:00	744.075	15.485	156.0272031	5.625860998	877.7873568
01/01 14:00:00	621.35	17.835	161.6426773	1.996946122	322.7917177
01/01 15:00:00	438.1625	18.605	156.7409952	2.229264088	349.4170716
01/01 16:00:00	227.175	17.755	144.6611191	2.058459455	297.7790485
01/01 17:00:00	26.55	16.69	20.323617	0.294840682	5.992229097
01/02 08:00:00	247.75	9.2625	151.7551233	2.158767553	327.6040362
01/02 09:00:00	271.6375	9.815	153.9585827	2.189692453	337.1219466
01/02 10:00:00	231.425	10.3625	148.865102	2.11991703	315.5816648
01/02 11:00:00	265.7125	11.65	152.4246909	2.167808152	330.4274875
01/02 12:00:00	353.7375	13.65	156.742829	2.229290199	349.4252524
01/02 13:00:00	345.575	15.65	155.2246832	2.207587541	342.6720768
01/02 14:00:00	233.4625	16.8625	146.0292647	2.077268179	303.3419449
01/02 15:00:00	133.8875	17.3625	97.36946673	1.448788346	141.0677487
01/02 16:00:00	70.125	17.8625	52.07132004	0.771313786	40.16332697
01/02 17:00:00	8.325	17.575	6.968836313	0.092952866	0.647773309
01/03 08:00:00	460.5625	10.2725	162.2389145	2.30785965	374.4246444
01/03 09:00:00	648.3	11.12	165.9576939	2.069964489	343.526533
01/03 10:00:00	733.475	11.9725	156.282231	6.003002009	938.1625466
01/03 11:00:00	770.9875	13.7125	158.6856871	5.369897188	852.1258249
01/03 12:00:00	785.5625	16.16	150.440588	7.132284902	1072.985134
01/03 13:00:00	778.125	18.6125	149.092749	7.045288998	1050.401504
01/03 14:00:00	710.9625	20.115	151.791163	5.702947106	865.6569738
01/03 15:00:00	604.3375	20.715	159.6262505	1.99711807	318.7924693
01/03 16:00:00	409.6875	21.315	154.3052702	2.194444158	338.6142988
01/03 17:00:00	51.075	20.8125	38.38710863	0.566138818	21.73243231
01/04 08:00:00	187.5	11.92	132.8278538	1.960435684	260.4004644
01/04 09:00:00	299.05	12.7725	154.3814766	2.195556974	338.9533275
01/04 10:00:00	370.9375	13.62	157.4615749	2.239565103	352.6454483
01/04 11:00:00	383.5	14.9975	157.1275321	2.234789823	351.1470097
01/04 12:00:00	381.325	16.845	155.9689602	2.218227445	345.9746281

Appendix C

Experimental Data for Sunny Days

Date	LOC_T	radiation	Temperature	current	voltage	Power
01/17/2024	7:46 AM	472.51	10.4	1.1	133	146.3
01/17/2024	7:49 AM	530.53	13	1.23	133	163.59
01/17/2024	7:53 AM	546.88	13.5	1.27	133	168.91
01/17/2024	7:56 AM	382.81	14	0.89	133	118.37
01/17/2024	7:59 AM	574.8	14	1.34	133	178.22
01/17/2024	8:03 AM	603.44	15.6	1.4	133	186.2
01/17/2024	8:06 AM	631.8	16.5	1.47	133	195.51
01/17/2024	8:09 AM	643.95	17.2	1.5	131	196.5
01/17/2024	8:13 AM	656.11	17.4	1.53	131	200.43
01/17/2024	8:16 AM	672.31	17.5	1.56	131	204.36
01/17/2024	8:19 AM	676.36	18	1.57	131	205.67
01/17/2024	8:23 AM	720.92	18.8	1.68	131	220.08
01/17/2024	8:26 AM	736.83	18.9	1.71	131	224.01
01/17/2024	8:29 AM	759.26	18.8	1.77	131	231.87
01/17/2024	8:33 AM	785.59	19	1.83	131	239.73
01/17/2024	8:36 AM	809.9	20.1	1.88	135	253.8
01/17/2024	8:39 AM	830.01	20.4	1.93	135	260.55
01/17/2024	8:43 AM	838.4	20.6	1.95	135	263.25
01/17/2024	8:46 AM	845.05	19.8	1.97	135	265.95
01/17/2024	8:49 AM	854.46	20.9	1.99	135	268.65
01/17/2024	8:53 AM	878.62	19.8	2.05	135	276.75
01/17/2024	8:56 AM	890.91	20	2.07	135	279.45
01/17/2024	8:59 AM	899.16	20.5	2.09	135	282.15
01/17/2024	9:03 AM	899.31	20.7	2.09	135	282.15
01/17/2024	9:06 AM	911.02	20.3	2.12	135	286.2
01/17/2024	9:09 AM	935.33	21.3	2.18	140	305.2
01/17/2024	9:13 AM	947.77	21.7	2.21	140	309.4
01/17/2024	9:16 AM	955.73	21.8	2.23	140	312.2

To view the full data used, please click on the link below.

- **((https://drive.google.com/drive/folders/1pfdXrwwMT47bTcFr_ljWUtI0EKYhmQ4-?usp=drive_link))**
- **(https://climate.onebuilding.org/)**

Appendix D

ANN-GA Algorithm (python code)

#Install all necessary Libraries resources

```
import numpy as np
from sklearn.model_selection import train_test_split
from keras.models import Sequential
from keras.layers import Dense
from keras.optimizers import Adam
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.metrics import r2_score, mean_squared_error, mean_absolute_error
import time
# Load your dataset
data = pd.read_excel('Name File')  # Replace 'your_dataset.csv' with your dataset
file
col_names=['column1', 'column2', 'column3', 'column4', 'column5', 'column6',
'column7']
data.columns = col_names
data.head()
# Specify the indexes by which you want to partition the data for theoretical
data.
test_indices = list(range(234, 390))
# Splitting data into training and test set based on indexes
train_data = data[~data.index.isin(test_indices)]
test_data = data[data.index.isin(test_indices)]
# Create the required variables
t_train = train_data.iloc[:, 0]
X_train = train_data.iloc[:,1:3]
y_train = train_data.iloc[:, -1]
t_test = test_data.iloc[:, 0]
X_test = test_data.iloc[:, 1:3]
y_test = test_data.iloc[:, -1]
# Specify the indexes by which you want to partition the data for Experimental
data.
def alternating_indices(n, train_ratio=0.6):
    indices = np.arange(n)
    train_indices = []
    test_indices = []
    block_size = int(train_ratio * 20)
    test_block_size = 20 - block_size
    for i in range(0, n, 20):
        train_indices.extend(indices[i:i+block_size])
```

```
        test_indices.extend(indices[i+block_size:i+block_size+test_block_size])
    return train_indices, test_indices
train_indices, test_indices = alternating_indices(len(data))
train_data = data.iloc[train_indices]
test_data = data.iloc[test_indices]
# Create the required variables
t_train = train_data.iloc[:, 0:1]
X_train = train_data.iloc[:, 2:4]
y_train = train_data.iloc[:, -1]
t_test = test_data.iloc[:, 0:1]
X_test = test_data.iloc[:, 2:4]
y_test = test_data.iloc[:, -1]


# Define parameters
population_size = 10
mutation_rate = 0.1
crossover_rate = 0.8
num_iterations = 5
num_hidden_layers_range = (1, 5)  # Range for number of hidden layers
num_neurons_range = (10, 128)  # Range for number of neurons in each layer
# Fitness function (Example: R2, MAE, RMSE)
def evaluate_fitness(model, X_test, y_test):
    y_preds_GA = model.predict(X_test)
    r2 = r2_score(y_test, y_preds_GA)
    mae = mean_absolute_error(y_test, y_preds_GA)
    rmse = np.sqrt(mean_squared_error(y_test, y_preds_GA))
    mse = (mean_squared_error(y_test, y_preds_GA))
    return r2, mae,  rmse,mse, y_preds_GA
# Mutation function
def mutate(model):
    mutated_model = Sequential.from_config(model.get_config())
    for layer in mutated_model.layers:
        if isinstance(layer, Dense):
            if np.random.rand() < mutation_rate:
                num_neurons = np.random.randint(*num_neurons_range)
                layer.units = num_neurons
    return mutated_model
# Crossover function
def crossover(parent1, parent2):
    child = Sequential()
    for layer1, layer2 in zip(parent1.layers, parent2.layers):
        if isinstance(layer1, Dense) and isinstance(layer2, Dense):
            units = min(layer1.units, layer2.units)  # Choose the minimum units
```

```
            child.add(Dense(units=units, input_dim=layer1.input_shape[1],
activation='relu'))  # Input layer or hidden layer
    child.add(Dense(units=1))  # Output layer
    return child
# Genetic Algorithm
best_model = None
best_r2 = -float('inf')
best_mae = float('inf')
best_rmse = float('inf')
best_mse = float('inf')
r2_history = []
mae_history = []
rmse_history = []
mse_history = []
best_r2_history = []
for iteration in range(num_iterations):
    # Generate population
    population = []
    for _ in range(population_size):
        num_hidden_layers = np.random.randint(*num_hidden_layers_range)
        model = Sequential()
        model.add(Dense(np.random.randint(*num_neurons_range),
input_dim=X_train.shape[1], activation='relu'))  # Input layer
        for _ in range(num_hidden_layers):
            model.add(Dense(np.random.randint(*num_neurons_range),
activation='relu'))  # Hidden layers
        model.add(Dense(1))  # Output layer
        model.compile(optimizer='adam', loss='mse')  # Using mean squared error
as loss for regression task
        population.append(model)
# Evaluate fitness
    for model in population:
        model.fit(X_train, y_train, epochs=1000, batch_size=32, verbose=0)
        r2, mae, rmse, mse, y_preds_GA = evaluate_fitness(model, X_test, y_test)
        r2_history.append(r2)
        mae_history.append(mae)
        rmse_history.append(rmse)
        mse_history.append(mse)
        if r2 > best_r2:
            best_r2 = r2
            best_model = model
            best_mae = mae
            best_rmse = rmse
            best_mse = mse
```

```
            best_y_preds_GA = y_preds_GA
            best_r2_history.append(best_r2)
    # Crossover
    for i in range(1, len(population)):
        if np.random.rand() < crossover_rate:
            parent1 = population[i-1]
            parent2 = population[i]
            population[i] = crossover(parent1, parent2)
    # Mutation
    for i in range(1, len(population)):
        population[i] = mutate(population[i])
# Elitism: Keep the best individual
    population[0] = best_model
    print("Iteration:", iteration + 1)
    print("Best R2:", best_r2)
    print("Best MAE:", best_mae)
    print("Best RMSE:", best_rmse)
    print("Best MSE:", best_mse)
    print()
# Print the optimized number of hidden layers and neurons
print("Optimized number of hidden layers:", len(best_model.layers) - 2)
for i, layer in enumerate(best_model.layers[1:-1]):
    print(f"Number of neurons in hidden layer {i+1}: {layer.units}")
_, best_mae, best_rmse, best_mse, _ = evaluate_fitness(best_model, X_test,
y_test)
print("Best MAE:", best_mae)
print("Best RMSE:", best_rmse)
print("Best MSE:", best_mse)
# Record start time
start_time = time.time()
# Record end time and calculate execution time
end_time = time.time()
execution_time = end_time - start_time
print("Execution time:", execution_time, "seconds")
```

Appendix E

ANN-GWO (Python code)

```
# Define the neural network model
def create_model(num_hidden_layers, num_neurons):
    model = Sequential()
    model.add(Dense(num_neurons, input_dim=X_train.shape[1],
activation='relu'))  # Input layer
    for _ in range(num_hidden_layers):
        model.add(Dense(num_neurons, activation='relu'))  # Hidden layers
    model.add(Dense(1))  # Output layer
    model.compile(optimizer='adam', loss='mse', metrics=['mae'])
    return model
# Define the objective function to minimize (Mean Squared Error)
def objective_function(num_hidden_layers, num_neurons):
    model = create_model(num_hidden_layers, num_neurons)
    model.fit(X_train, y_train, epochs=1000, verbose=0)
    y_pred = model.predict(X_test)
    mse = mean_squared_error(y_test, y_pred)
    return mse
# Initialize gray wolf positions
def initialize_wolves(num_wolves, num_dimensions):
    wolves = np.random.rand(num_wolves, num_dimensions)
    wolves[:, 0] = wolves[:, 0] * 4 + 1  # Hidden layers between 1 and 5
    wolves[:, 1] = wolves[:, 1] * 127 + 10  # Neurons between 10 and 128
    return wolves
# Update wolf positions using GWO equations
def update_positions(alpha, beta, delta, wolves, a=2):
    for wolf in wolves:
        for i in range(len(wolf)):
            r1 = np.random.random()
            r2 = np.random.random()
            A1 = 2 * a * r1 - a
            C1 = 2 * r2
            D_alpha = abs(C1 * alpha[i] - wolf[i])
            X1 = alpha[i] - A1 * D_alpha
            r1 = np.random.random()
            r2 = np.random.random()
            A2 = 2 * a * r1 - a
            C2 = 2 * r2
            D_beta = abs(C2 * beta[i] - wolf[i])
            X2 = beta[i] - A2 * D_beta
            r1 = np.random.random()
```

```
            r2 = np.random.random()
            A3 = 2 * a * r1 - a
            C3 = 2 * r2
            D_delta = abs(C3 * delta[i] - wolf[i])
            X3 = delta[i] - A3 * D_delta
            wolf[i] = (X1 + X2 + X3) / 3
        wolf[0] = np.clip(wolf[0], 1, 5)  # Hidden layers between 1 and 5
        wolf[1] = np.clip(wolf[1], 10, 128)  # Neurons between 10 and 128
    return wolves
# Perform GWO optimization
def gray_wolf_optimization(max_iter, num_wolves):
    num_dimensions = 2
    wolves = initialize_wolves(num_wolves, num_dimensions)
    best_position = None
    best_error = float('inf')
    errors = []
    for _ in range(max_iter):
        for wolf in wolves:
            num_hidden_layers = int(wolf[0])
            num_neurons = int(wolf[1])
            error = objective_function(num_hidden_layers, num_neurons)
            if error < best_error:
                best_error = error
                best_position = wolf.copy()
        errors.append(best_error)
        alpha, beta, delta = wolves[np.argsort([objective_function(int(wolf[0]),
int(wolf[1])) for wolf in wolves])[:3]]
        wolves = update_positions(alpha, beta, delta, wolves)
    best_num_hidden_layers, best_num_neurons = int(best_position[0]),
int(best_position[1])
    return best_num_hidden_layers, best_num_neurons, best_error, errors
# Run GWO optimization
best_hidden_layers, best_neurons, best_error, errors =
gray_wolf_optimization(max_iter=5, num_wolves=10)
# Train the neural network with the optimized parameters
best_model = create_model(best_hidden_layers, best_neurons)
history = best_model.fit(X_train, y_train, epochs=1000, verbose=0)
# Evaluate the model
best_y_pred = best_model.predict(X_test)
best_mae = mean_absolute_error(y_test, best_y_pred)
best_mse = mean_squared_error(y_test, best_y_pred)
best_rmse = np.sqrt(best_mse)
best_r2 = r2_score(y_test, best_y_pred)
print("Best MSE MAE RMSE R2:",best_mse, best_mae, best_rmse, best_
```

جمهورية العراق

وزارة التعليم العالي والبحث العلمي

جامعة ميسان

كلية الهندسة

تخمين انتاج القدرة الكهربائية من الألواح الشمسية باستخدام الذكاء الاصطناعي

رسالة مقدمة الى مجلس كلية الهندسة في جامعة ميسان كجزء من متطلبات الحصول على شهادة الماجستير في علوم الهندسة الكهربائية (قدرة)

اعداد

شهد محمد راضي

بكالوريوس هندسة كهربائية

باشراف

الأستاذ المساعد الدكتور صادق دعير عنيد

تشرين الثاني 2024

الخلاصة

تعتبر الطاقة الكهروضوئية من أكثر مصادر الطاقة المتجددة التي لا تنضب في العالم بسبب وفرتها واستدامتها، فضلاً عن انخفاض تكاليف تشغيلها. ومع ذلك، يعتمد إنتاجها على عوامل مثل الإشعاع ودرجة الحرارة وما إلى ذلك. لذلك، فإن التنبؤ بالطاقة الكهروضوئية هو مرحلة حاسمة للاستفادة من استقرار وجودة وإدارة شبكة الطاقة الهجينة. في هذه الأطروحة، تم تصميم نموذج التنبؤ بالطاقة الكهروضوئية بناءً على البيانات النظرية والبيانات الفعلية باستخدام تقنيات التعلم الآلي المختلفة. يتم الحصول على البيانات النظرية من موقع الويب (climate one building)، بينما يتم جمع البيانات الفعلية من النموذج الأولي التجريبي للطاقة الكهروضوئية المثبت في كلية الهندسة بجامعة ميسان في العراق. لتعزيز نموذج التنبؤ بالطاقة الكهروضوئية، يتم استخدام تقنية الشبكة العصبية الاصطناعية (ANN) القائمة على تحسين الذئب الرمادي (GWO) والخوارزمية الوراثية (GA) كطرق للتعلم. يتم استخدام نهج Python لتصميم هذا التنبؤ بالطاقة الكهروضوئية بناءً على أربع وظائف لياقة، R2 وMAE وRMSE وMSE. في هذه الدراسة، تم تحليل البيانات واختبارها على مدى فترات زمنية قصيرة ومتوسطة المدى لضمان أداء النموذج ودقة التنبؤ، وبالتالي تحسين إنتاج الطاقة الكهروضوئية عبر أوقات وظروف جوية مختلفة. أخيرًا، تشير النتائج إلى أن نموذج ANN القائم على خوارزمية GA يلتقط نمط توليد الطاقة الكهروضوئية بدقة أعلى عبر ظروف جوية مختلفة مقارنة بنماذج التنبؤ ANN وANN-GWO التقليدية. يتضح هذا من قيم معامل ارتباط بيرسون (R^2) الأعلى التي تم تحقيقها خلال أشهر مختلفة في البيانات النظرية للتنبؤ بالطاقة الكهروضوئية متوسطة المدى. بالإضافة إلى ذلك، حقق نموذج ANN القائم على GA قيم R^2 أعلى في ظل ظروف مشمسة وغائمة وممطرة في البيانات التجريبية للتنبؤ بالطاقة الكهروضوئية قصيرة المدى.