International Journal of Photoenergy

Modeling and Forecasting for Energy Production of Photovoltaic (PV) Systems

Lead Guest Editor: Laurentiu Fara
Guest Editors: Bogdan-Gabriel Burduhos, James Connolly, Alexandros G. Charalambides, and Anne Migan Dubois

Modeling and Forecasting for Energy Production of Photovoltaic (PV) Systems

International Journal of Photoenergy

Modeling and Forecasting for Energy Production of Photovoltaic (PV) Systems

Lead Guest Editor: Laurentiu Fara
Guest Editors: Bogdan-Gabriel Burduhos, James Connolly, Alexandros G. Charalambides, and Anne Migan Dubois

Contents

Forecasting of Energy Production for Photovoltaic Systems Based on ARIMA and ANN Advanced Models
Laurentiu Fara, Alexandru Diaconu, Dan Craciunescu, and Silvian Fara
Research Article (19 pages), Article ID 6777488, Volume 2021 (2021)

Sky Image-Based Localized, Short-Term Solar Irradiance Forecasting for Multiple PV Sites via Cloud Motion Tracking
Lasanthika H. Dissawa, Roshan I. Godaliyadda, Parakrama B. Ekanayake, Ashish P. Agalgaonkar, Duane Robinson, Janaka B. Ekanayake, and Sarath Perera
Research Article (27 pages), Article ID 9973010, Volume 2021 (2021)

Hindawi
International Journal of Photoenergy
Volume 2022, Article ID 1898132, 12 pages
https://doi.org/10.1155/2022/1898132

Research Article

Forecasting Tennis Match Results Using the Bradley-Terry Model

Aisha Fayomi,[1] **Rizwana Majeed**,[2] **Ali Algarni**,[1] **Sohail Akhtar**,[3] **Farrukh Jamal**,[4] **and Jamal Abdul Nasir**[2]

[1]*Faculty of Science Department of Statistics, King Abdulaziz University, Jeddah, Saudi Arabia*
[2]*Department of Statistics, Government College University, Lahore, Pakistan*
[3]*Department of Mathematics and Statistics, University of Haripur, Haripur, Pakistan*
[4]*Department of Statistics, The Islamia University of Bahawalpur, Pakistan*

Correspondence should be addressed to Farrukh Jamal; farrukh.jamal@iub.edu.pk

Received 17 December 2021; Revised 15 February 2022; Accepted 16 February 2022; Published 8 March 2022

Academic Editor: Laurentiu Fara

Forecasting has been playing an important role in different fields of life, i.e., in decision-making activities of management, to predict uncertain events within an organization, in weather forecasting, in flood forecasting, etc. Stakeholders involved in betting market take advantage of tennis forecasting directly or indirectly. Winning probability calculated using forecasting models helps the bettors in deciding whether to place a bet or not. Keeping in view the importance of tennis forecasting, the Bradley-Terry model is used to model men's professional tennis for predicting match outcomes in tennis matches of men's singles. Model coefficients are estimated using data from January 2019 to September 2020 of 3439 matches. Ratings for each player are calculated using model coefficients. Player rankings are then calculated using these ratings. Comparison of model rankings with ATP rankings has shown satisfactory results. Winning probability for each player is calculated using model coefficients and ratings. These probability predictions are evaluated against four measures of performance. The results reveal that surface on which a game is played on contributes significantly towards a player's performance. Due to this impact of the surface, our model has produced superior player rankings for certain players who had been ranked very low in official ATP rankings. According to most of the performance measures, the model has shown good results for clay court data which are closely followed by hard court data. To calculate return on investment, model results are compared with the bookmakers' average odds and best available odds. It has been found that return on investment for a fitted model is highest in the case of clay court data in comparison to bookmaker's average odds and best odds.

1. Introduction

Forecasting has been playing an important role in different fields of life. In recent years, it has also got significant importance in sports, i.e., golf, cricket, soccer, and tennis. The ability to predict the outcome of a sporting event accurately is something which fascinates the sporting world [1–5]. Application of statistical analysis for forecasting sports has evolved rapidly to meet demand from gambling, coaches, and media. For example, media channels provide more insightful coverage of a sporting tournament due to predictions [6]. Modeling and predicting the results of tennis matches in particular have gained significant attention in recent years. The prediction models reveal the interesting characteristics of various playing styles, making them suitable for coaching purposes [7, 8]. The increasing trend of online betting markets, such as Betfair, offers an additional and extremely competitive motivator for research in prediction models [9]. Tennis sport is an attractive research candidate due to the large amount of freely available historical data.

Most of the research done in modeling tennis matches focuses on prematch prediction, in which the aim is to predict the winning probability of either player before starting the match. Prematch prediction is in opposition to the prediction of in-play matches, in which winning probabilities during a match are predicted [10, 11]. Our focus in this research is solely on the prediction of winning probability before starting the match.

In this study, paired comparisons for forecasting tennis match results will be done. Several methods are available to forecast match outcomes in tennis. For example, Bradley-Terry-type model was employed to forecast match results for the top tier of men's professional tennis, the ATP tour [12]. Similarly, a high-dimensional dynamic model was used to forecast tennis match results [13]. Furthermore, tennis match outcomes were forecasted using within-match betting markets [14]. Kovalchik compared predictive performance of different forecasting models for predicting the tennis outcomes [15].

Paired comparison method is a handy decision-making tool, which describes and compares the values to each other. For making decision in the absence of any objective data, the paired comparison method is a very handy tool. This method is also known as the pairwise comparison [16]. In pairwise comparison, the items which we have to compare are judged in pairs to check which one is preferred and to rank the items being compared [17, 18]. The main reason for applying pairwise comparison methods is simplicity in judging two items instead of several items at once. This method can be used in various situations, i.e., in subjective evaluation criteria or when important priorities are unclear, for example, modeling competitive ability in sports and choice behavior, i.e., preference of the democratic presidential candidate to the Republican candidate or the preference of one soft drink to another.

1.1. Rating. A rating is a numerical value assigned to each team or player created on their contributions, while a ranking is the ordinal position based on the ratings [19]. The rating plays an important role in forecasting and prediction. It is used to indicate the strength of items or objects relative to each other and is used in almost every field. The player's rating is a system used to rate players individually. The rating can be done domestically or internationally, and each of the teams/players is ranked according to their rating or individual performance. The focus of this research is on the player's ranking to predict match outcomes [20, 21].

1.2. Betting Odds and Betting Strategies. Tennis bets are mainly placed in two kinds of markets: bookmakers or betting exchanges. In the former type, odds are offered by a bookmaker and customers who accept odds and directly place money at these odds against the bookmaker [22]. In betting exchange, odds are offered by customers and are placed against each other; for each paired bet, exchange simply takes a small commission. Generally, odds are more favorable in exchange markets, but due to the limitation of historical data on such odds, the models are mostly compared against traditional bookmaker's odds in research. The problem in making bets is to find successful bets, where the considered probability of occurrence is greater than the corresponding probability calculated by the bookmakers' odds, to get a positive expected return. A statistical model capable to predict probabilities of the results of tennis matches accurately can form a profitable betting strategy. Bets can be placed on different events relevant to various aspects of tennis match, both before starting the match and during the match. However, bets placed before starting the match will be considered here [23, 24].

The purpose of betting strategies is to take advantage from those cases in which odds are undervalued. In these cases, the actual probability of the event in an outcome occurring is greater than the implied probability calculated using odds. If the predictive model is available, a simple betting strategy is placing a unit bet when the model shows that the odds are undervalued.

2. Data and Model

For the current study, data on tennis matches of men's singles have been obtained from tennis-data.co.uk. The dataset consists of 3439 men's single matches from January 2019 to September 2020. The dataset contains information about tournament name; date of match; location; ATP rankings; name of the winning and losing player; surface (carpet, clay, hard court, or grass); the results in sets, games, and matches; and the series. "Series" relates to the importance of tournament in terms of the prize money available and ranking points; it is divided into ATP250, ATP500, Grand Slam, Masters 1000, and Masters Cup. Results of previous 52 weeks of tournament competition have been used to derive ATP rankings. Awarded points are based on prestige of the tournament and on the progress of player. Records for the odds of bookmakers are also available at the website of tennis data for each game, which are utilized to measure the forecasting performance.

The purpose of the study is to forecast the result of the tennis matches between two competing players. The available data has been analyzed under Bradley-Terry modeling framework using the BradleyTerry2 package in R to obtain ranking of tennis players for the future.

2.1. Model. The Bradley-Terry model is a simple, popular, and widely used method for handling data on paired comparisons which is used to find the probabilities of the potential outcomes when individuals or objects are judged against each other in the form of pairs. This model is used to predict matches and tournaments and for testing the efficiency of betting markets. The default modeling approach to the statistical analysis of tennis matches is also based on the Bradley-Terry model. Forecasting tennis is further complicated by the surface effect because some players play better on some surfaces than on others, and the influence of the surface on match outcomes is also assessed by the Bradley-Terry model [25, 26].

Suppose a set of K elements are compared with each other in the form of pairs. To find the probability that "i" is preferred to "j," for two elements "i" and "j" taken from this set, the following model was suggested by Bradley and Terry in 1952.

$$P(i \text{ is preffered to } j) = \frac{\lambda_i}{\lambda_i + \lambda_j}, \tag{1}$$

where the parameter $\lambda_l > 0$ belongs to element $l \in \{1, 2, \cdots, K\}$ which represents rating and is denoted by $\lambda = \{\lambda_i\}_{i=1}^{K}$.

This model has many applications in different fields as mentioned by Hunter; an extensive bibliography published by Davidson and Farquhar (1976) on the method of paired comparisons also includes a considerable amount of work [27, 28]. For example, it has been adopted to rank players by the European Go Federation and the World Chess Federation and Hastie and Tibshirani suggested that this model is a standard approach for constructing multiclass classifiers using the results of binary classifiers [29]. Different extensions of this model have been suggested to handle draws, home advantage, team comparisons, and multiple comparisons [30, 31]. In particular, the Plackett-Luce model was developed which is the popular extension for multiple comparisons, has been used for ranking of multiple objects, defined a prior distribution of permutations, and has also been used for choice models [32]. In the monographs of David [33] and Diaconis [34], detailed discussion about the statistical foundations of these models has also been provided.

3. Ranking of Tennis Players on All Surfaces

Keeping in view the scope of the current study, it is of great interest to utilize the abilities of players $\{\alpha i\}$ to model a new ranking system as an alternative to official rankings published by ATP and compare both the ranking systems. Model parameters have been estimated using data for the period from January 2019 to September 2020, to obtain ranking of tennis players, assigning equal weights to all matches played on any surface.

Rankings produced by the model at the end of September 2020 for the top fifteen players have been shown in Table 1 along with ATP rankings. The rankings produced by the model and the ATP rankings are more or less the same for the top 7 players, except a few discrepancies such as Thiem D. and Federer R., whose ATP ranking is 3 and 4, respectively, whereas ranking of these players according to our model is 4 and 3, respectively. This change in ranking of Thiem D. and Federer R. is due to the difference in the performance of these players on different surfaces; i.e., Thiem D. performs well on clay whereas clay is Federer's least favored surface and he has managed to win only one title at the Roland Garros (French Open). Moreover, Federer R. has played on all surfaces, i.e., hard court, clay court, and grass court whereas Theim D. never won a game on grass court in our dataset. Theim D. has won more games, i.e., 63 in our dataset, than Federer R., who managed to win 58 games. Federer R. has won less games than Theim D., but he has played on all surfaces that contributed in his better overall performance than Theim D. It is also a historical fact that Federer R. has won more Wimbledon (grass court game) titles than any other player. Therefore, it is quite evident that grass is his most preferred playing surface.

Similarly, Berrettini M. has been ranked 7 in our model whereas he is at number 8 according to official ATP rankings. Zverev A., who is at number 7 in official ATP ranking, has been ranked 19 in our model. This is due to the fact that Zverev A. has managed to win only two games on grass court. Although Zverev A. has won more games, i.e., 53 in total than Berrettini M., who managed to win 49 games, Berrettini M. played equally well on all surfaces which accounted for his better ranking position. For Dimitrov G., our model has amazingly predicted the same ranking as awarded by ATP (Figure 1).

Table 1: Model ranking and ATP ranking of top 20 players on all surfaces.

Model rank	Players	Estimates	Rating	Std. error	ATP ranking
1	Djokovic N.	0.2725	1.3132	0.5181	1
2	Nadal R.	0.0000	1.0000	0.0000	2
3	Federer R.	-0.1672	0.8460	0.5154	4
4	Thiem D.	-0.231	0.7937	0.4977	3
5	Medvedev D.	-0.9094	0.4028	0.4928	5
6	Tsitsipas S.	-1.0045	0.3662	0.4585	6
7	Berrettini M.	-1.3522	0.2587	0.5062	8
8	Rublev A.	-1.3591	0.2569	0.4869	12
9	De Minaur A.	-1.4134	0.2433	0.5524	28
10	Raonic M.	-1.4193	0.2419	0.5577	21
11	Wawrinka S.	-1.4465	0.2354	0.5167	17
12	Schwartzman D.	-1.474	0.2290	0.503	14
13	Struff J. L.	-1.5372	0.2150	0.4939	32
14	Monfils G.	-1.5478	0.2127	0.5247	9
15	Opelka R.	-1.5483	0.2126	0.5134	36
16	Bautista Agut R.	-1.5949	0.2029	0.495	10
17	Kyrgios N.	-1.6348	0.1950	0.5648	42
18	Krajinovic F.	-1.7045	0.1819	0.5476	29
19	Zverev A.	-1.7391	0.1757	0.5187	7
20	Dimitrov G.	-1.7691	0.1705	0.5243	20

After 7 players in our model ranking, there are some discrepancies, due to the fact that ATP official ranking system ranks players based on how frequently a player has participated in events, instead of considering absolute performance of players, whereas the model being used in this study ranks players merely on the basis of past performance drawing no distinction such as how frequently a player has participated in events.

The ability of players can also be estimated using comparison intervals calculated using quasistandard errors as depicted in Figure 2. This method not only provides comparison with the reference player, but readily made comparisons between any pair of competing players are also provided.

3.1. Winning Probabilities on All Surfaces. Using values of estimates from model fit given in Table 2 and Figure 3, without considering the surface game played on, at the end of 2020, the estimated winning probabilities of competing players can be calculated using the following equation:

$$\hat{\pi}_{ij} = \frac{\exp\left(\hat{\alpha}_i - \hat{\alpha}_j\right)}{1 + \exp\left(\hat{\alpha}_i - \hat{\alpha}_j\right)}. \quad (2)$$

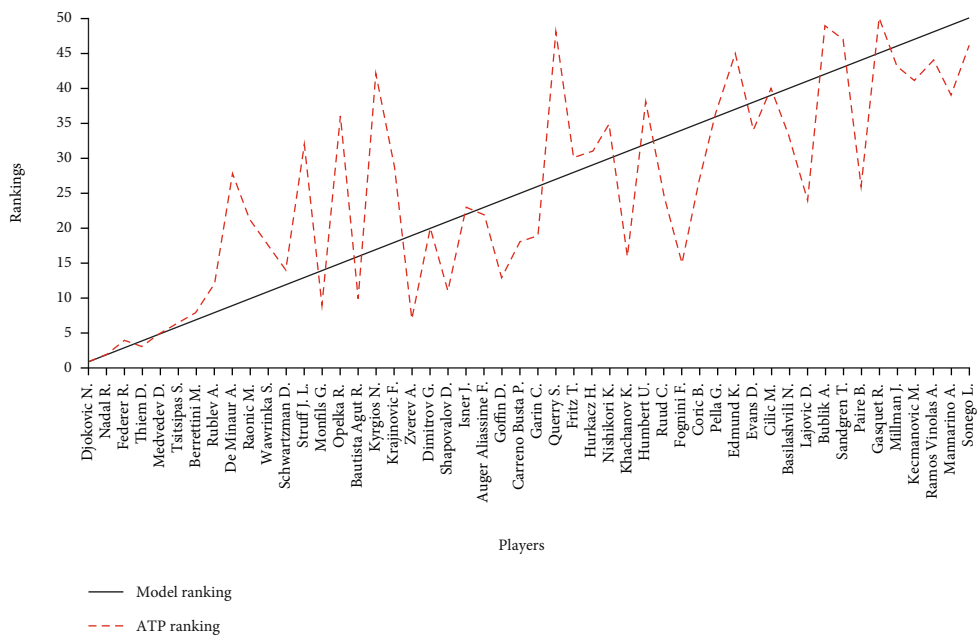


Figure 1: Comparison of model ranking with ATP ranking for top 50 tennis players on all surfaces.

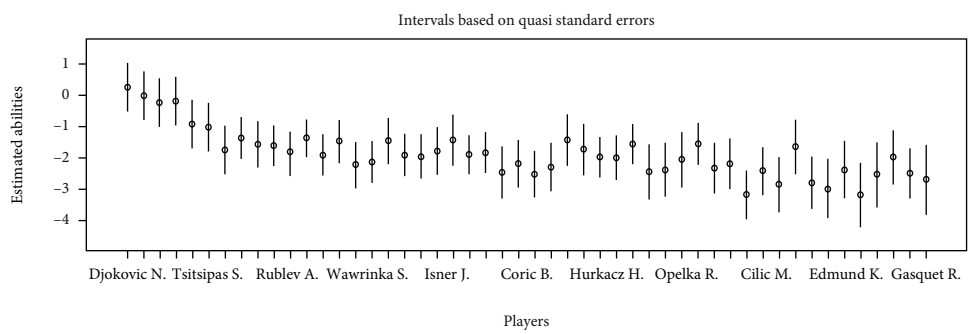


Figure 2: Estimated relative abilities of tennis players on all surfaces.

Table 2: Wining probabilities of players on all surfaces. DN: Djokovic N.; NR: Nadal R.; FR: Federer R.; TD: Thiem D.; MD: Medvedev D.; TS: Tsitsipas S.; ZA: Zverev A.; BM: Berrettini M.; DMA: De Minaur A.; RM: Raonic M.

	DN	NR	FR	TD	MD	TS	BM	RA	DMA	RM
DN		0.568	0.608	0.623	0.765	0.782	0.835	0.836	0.844	0.844
NR	0.432		0.542	0.557	0.713	0.732	0.794	0.796	0.804	0.805
FR	0.392	0.458		0.516	0.677	0.698	0.766	0.767	0.777	0.778
TD	0.377	0.443	0.484		0.663	0.684	0.754	0.755	0.765	0.766
MD	0.235	0.287	0.323	0.337		0.524	0.609	0.611	0.623	0.625
TS	0.218	0.268	0.302	0.316	0.476		0.586	0.588	0.601	0.602
BM	0.165	0.206	0.234	0.246	0.391	0.414		0.502	0.515	0.517
RA	0.164	0.204	0.233	0.245	0.389	0.412	0.498		0.514	0.515
DMA	0.156	0.196	0.223	0.235	0.377	0.399	0.485	0.486		0.501
RM	0.156	0.195	0.222	0.234	0.375	0.398	0.483	0.485	0.499	

Winning probabilities of players in the left-hand column over the players in the top row, on all surfaces.

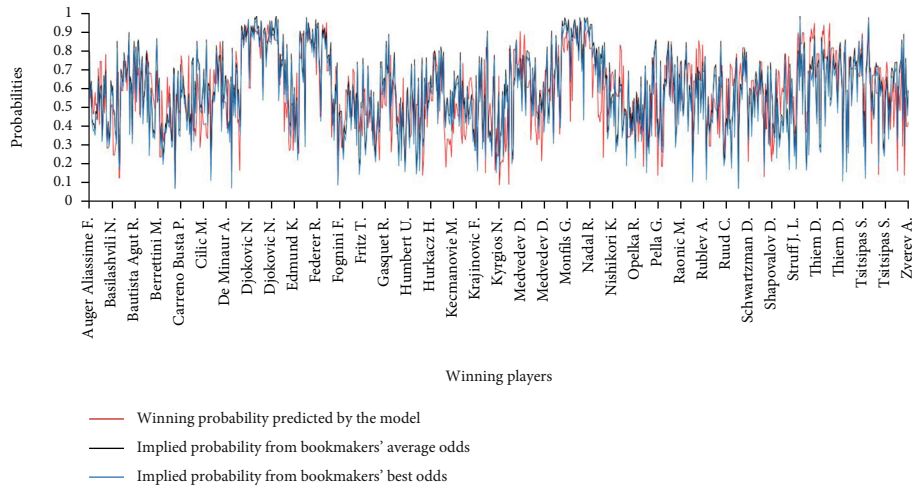


Figure 3: Match by match comparison of predicted probability with implied probability from bookmakers' average and best odds on all surfaces.

Table 3: Model ranking and ATP ranking of top 20 players on hard court.

Model rank	Player	Estimate	Rating	Std. error	ATP ranking
1	Djokovic N.	0.1416	1.1521	0.6918	1
2	Nadal R.	0.0000	1.0000	0.0000	2
3	Thiem D.	-0.0556	0.9459	0.6985	3
4	Federer R.	-0.2224	0.8006	0.7063	4
5	Medvedev D.	-0.6465	0.5239	0.6387	5
6	Tsitsipas S.	-1.0456	0.3515	0.6267	6
7	Raonic M.	-1.0508	0.3497	0.7172	21
8	De Minaur A.	-1.2714	0.2804	0.6871	28
9	Opelka R.	-1.3611	0.2564	0.6866	36
10	Kyrgios N.	-1.3994	0.2467	0.7363	42
11	Struff J. L.	-1.4251	0.2405	0.6964	32
12	Monfils G.	-1.4462	0.2355	0.6782	9
13	Berrettini M.	-1.4814	0.2273	0.7028	8
14	Rublev A.	-1.4879	0.2259	0.6456	12
15	Wawrinka S.	-1.5668	0.2087	0.6779	17
16	Bautista Agut R.	-1.6233	0.1973	0.6697	10
17	Dimitrov G.	-1.6310	0.1957	0.6989	20
18	Krajinovic F.	-1.6960	0.1834	0.6986	29
19	Carreno Busta P.	-1.7105	0.1808	0.6790	18
20	Zverev A.	-1.7151	0.1800	0.7016	7

The estimated winning probability of Djokovic N. vs. Nadal R. is

$$\hat{\pi}_{12} = \frac{\exp\left(\hat{\alpha}_1 - \hat{\alpha}_2\right)}{1 + \exp\left(\hat{\alpha}_1 - \hat{\alpha}_2\right)} = \frac{\exp\left(0.2725 - 0\right)}{1 + \exp\left(0.2725 - 0\right)} = 0.568. \quad (3)$$

To check the statistical significance of the difference between Djokovic N. and Nadal R., difference of the estimates $(\alpha_1{}^{\wedge} - \alpha_2{}^{\wedge} = 0.2725)$ is compared with its standard error (SE = 0.5181) having p value = 0.5990 indicating that the difference between the performance of Djokovic N. and Nadal R. is insignificant.

A 95% confidence interval for $\alpha_1 - \alpha_2$ is

$$0.2725 \pm 1.96(0.5181) \text{ or } (-0.743, 1.288), \quad (4)$$

and the 95% confidence interval for probability π_{12} of Djokovic to win obtained using the above interval is (0.32, 0.78).

4. Ranking of Tennis Players on Hard Court

In our model ranking for all surfaces, we saw some discrepancies which were actually due to the difference of performance on different surfaces. Some players perform well on hard court, some perform well on clay court, and some give better results on grass court. Keeping in view this fact, a separate analysis has been carried out for hard court and clay court competitions. This separate analysis has proved that surface has a very strong impact on a player's performance. Especially, we found that true ability of a player is judged by his performance on hard court. In the following, results from our model for hard court are presented.

Ranking produced by the model at the end of September 2020 for the top fifteen players on hard court has been shown in Table 3 along with ATP rankings (Figure 4). The model ranking and ATP ranking for the 1st six players are the same which shows the model has produced good results when surface effect has been taken into consideration. Raonic M. whose ranking is 21 according to ATP ranking has been ranked 7 by the model. This is due to the reason that Raonic M. has played and won more matches on hard court than any other surface. On clay court, he has played and won only 1 match, and on grass court, he managed to win 8 matches. Similarly, De Minaur A. has won 35 matches on hard court out of a total of 38 he has won. He has won

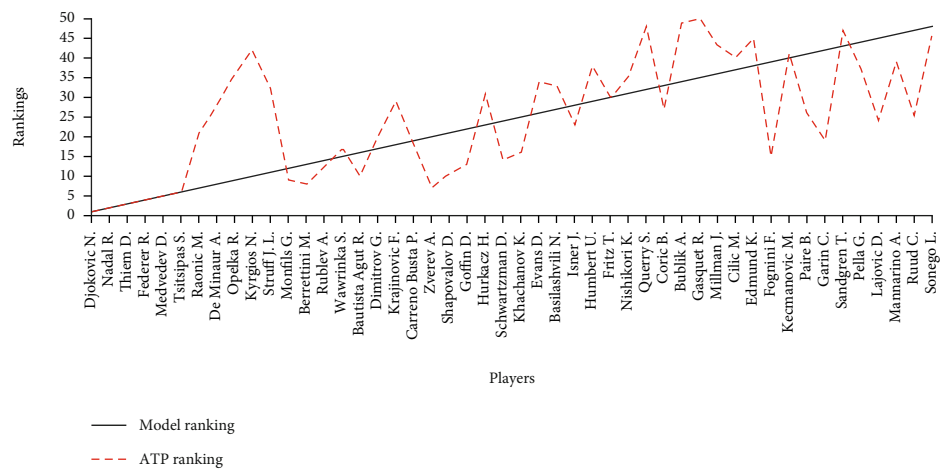


Figure 4: Comparison of model ranking with ATP ranking for top 50 tennis players on hard court.

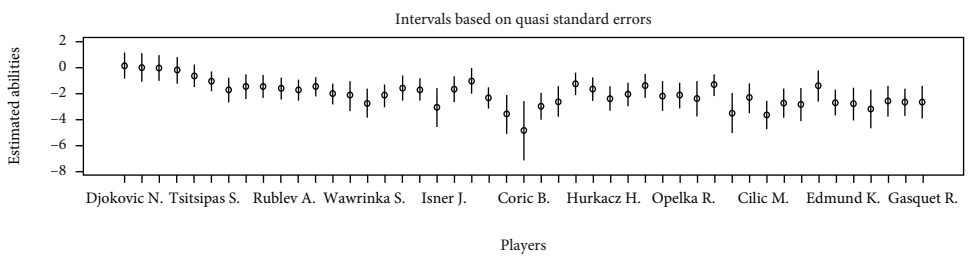


Figure 5: Estimated relative abilities of tennis players on hard court only.

Table 4: Wining probabilities of players on hard court. DN: Djokovic N.; NR: Nadal R.; TD: Thiem D.; FR: Federer R.; MD: Medvedev D.; TS: Tsitsipas S.; RM: Raonic M.; DMA: De Minaur A.; OR: Opelka R.; KN: Kyrgios N.

	DN	NR	TD	FR	MD	TS	RM	DMA	OR	KN
DN		0.535	0.549	0.590	0.687	0.766	0.767	0.804	0.818	0.824
NR	0.465		0.514	0.555	0.656	0.740	0.741	0.781	0.796	0.802
TD	0.451	0.486		0.542	0.644	0.729	0.730	0.771	0.787	0.793
FR	0.410	0.445	0.458		0.604	0.695	0.696	0.741	0.757	0.764
MD	0.313	0.344	0.356	0.396		0.598	0.600	0.651	0.671	0.680
TS	0.234	0.260	0.271	0.305	0.402		0.501	0.556	0.578	0.588
RM	0.233	0.259	0.270	0.304	0.400	0.499		0.555	0.577	0.586
DMA	0.196	0.219	0.229	0.259	0.349	0.444	0.445		0.522	0.532
OR	0.182	0.204	0.213	0.243	0.329	0.422	0.423	0.478		0.510
KN	0.176	0.198	0.207	0.236	0.320	0.412	0.414	0.468	0.490	

Winning probabilities of players in the left-hand column over the players in the top row, on hard court.

only 2 matches on grass court and 1 match on clay court. Because of this reason, he has been ranked 8 in our model. Likewise, Opelka R. whose ATP rank is 36 has been ranked 9 in our model due to the reason that he has played and won more matches on hard court than any other surface. Opelka R. has won 32 matches on hard court out of a total of 36 wins. He has won only two matches on clay court and two on grass court.

The ability of players can also be estimated using comparison intervals calculated using quasistandard errors as depicted in Figure 5. This method not only provides comparison with the reference player, but readily made comparisons between any pair of competing players are also provided.

4.1. Winning Probabilities on Hard Court. Using values of estimates from model fit given in Table 4 and Figure 6, estimated winning probabilities of competing players can be calculated using the following equation:

$$\hat{\pi}_{ij} = \frac{\exp\left(\hat{\alpha}_i - \hat{\alpha}_j\right)}{1 + \exp\left(\hat{\alpha}_i - \hat{\alpha}_j\right)}. \quad (5)$$

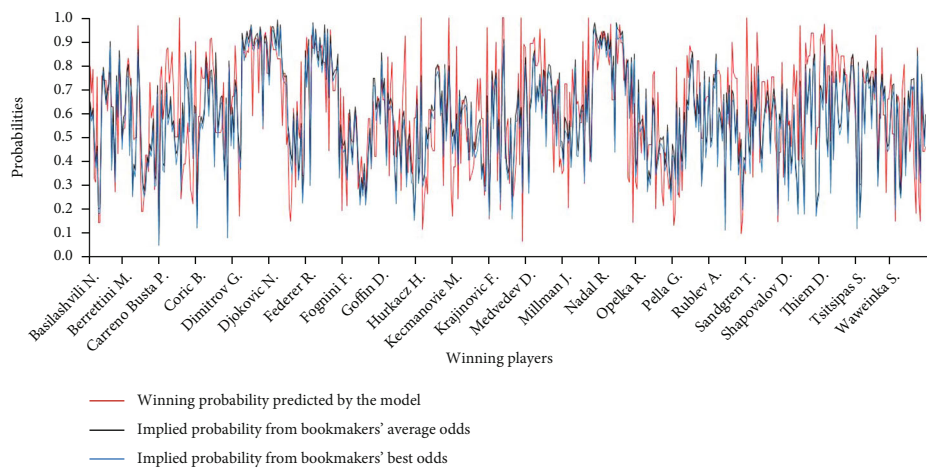


Figure 6: Match by match comparison of predicted probability with implied probability from bookmakers' average and best odds on hard court.

Table 5: Model ranking and ATP ranking of top 15 players on clay court.

Model rank	Player	Estimate	Rating	Std. error	ATP ranking
1	Djokovic N.	0.1119	1.1184	0.8803	1
2	Nadal R.	0.0000	1.0000	0.0000	2
3	Thiem D.	-0.3216	0.7250	0.7930	3
4	Federer R.	-0.6513	0.5214	0.9713	4
5	Wawrinka S.	-0.7988	0.4499	0.9701	17
6	Garin C.	-0.8865	0.4121	0.8665	19
7	Rublev A.	-0.9208	0.3982	0.9329	12
8	Tsitsipas S.	-0.9294	0.3948	0.8067	6
9	Ruud C.	-0.9939	0.3701	0.8548	25
10	Schwartzman D.	-1.1248	0.3247	0.8390	14
11	Fognini F.	-1.4655	0.2310	0.8729	15
12	Lajovic D.	-1.5521	0.2118	0.8869	24
13	Pella G.	-1.561	0.2099	0.8798	37
14	Struff J. L.	-1.9137	0.1475	0.8762	32
15	Paire B.	-1.9992	0.1354	1.0323	26

The estimated winning probability of Djokovic N. vs. Nadal R. is

$$\hat{\pi}_{12} = \frac{\exp\left(\hat{\alpha}_1 - \hat{\alpha}_2\right)}{1 + \exp\left(\hat{\alpha}_1 - \hat{\alpha}_2\right)} = \frac{\exp\left(0.1416 - 0\right)}{1 + \exp\left(0.1416 - 0\right)} = 0.535. \tag{6}$$

To check the statistical significance of the difference between Djokovic N. and Nadal R., difference of the estimates $(\alpha_1{}^{\wedge} - \alpha_2{}^{\wedge} = 0.1416)$ is compared with its standard error (SE = 0.6918) having p value = 0.8378 indicating that the difference between performance of Djokovic N. and Nadal R. is insignificant.

A 95% confidence interval for $\alpha_1 - \alpha_2$ is

$$0.1416 \pm 1.96(0.6918) \text{ or } (-1.2143, 1.4975), \tag{7}$$

and the 95% confidence interval for probability π_{12} of Djokovic to win obtained using the above interval is (0.23, 0.82).

5. Ranking of Tennis Players on Clay Court

Rankings produced by the model at the end of September 2020 for the top fifteen players on clay court have been shown in Table 5 along with ATP rankings (Figure 7). The model ranking and ATP ranking for first four players are the same which shows the model has produced good results when surface effect has been taken into consideration. It is pertinent to note that Wawrinka S., who was ranked 11 and 15 in all surface model and hard-court model, respectively, is now ranked 5 in the clay court model, whereas Garin C., who was ranked 26 and 42 in all surface models and hard-court model, respectively, is now ranked 6 in the clay court model. This significant upward change in ranking is due to the fact that Wawrinka S. and Garin C. have performed well on clay court as compared to hard court. Wawrinka S. has won 8 matches out of 14 on clay court while Garin C. has won 32 matches out of 43 on clay court.

The ability of players can also be estimated using comparison intervals calculated using quasistandard errors as depicted in Figure 8. This method not only provides comparison with the reference player, but readily made comparisons between any pair of competing players are also provided.

5.1. Winning Probabilities on Clay Court. Using values of estimates from model fit given in Table 6 and Figure 9, estimated winning probabilities of competing players on clay

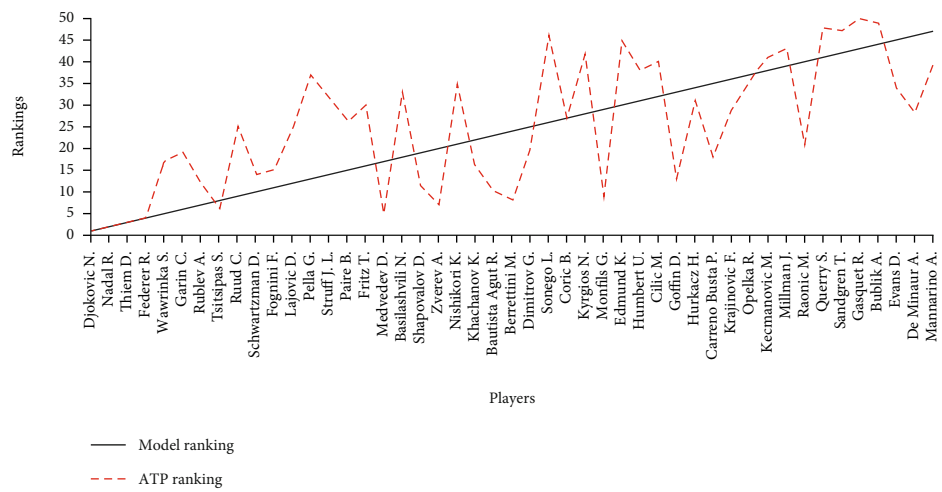


Figure 7: Comparison of model ranking with ATP ranking for top 50 tennis players on clay court.

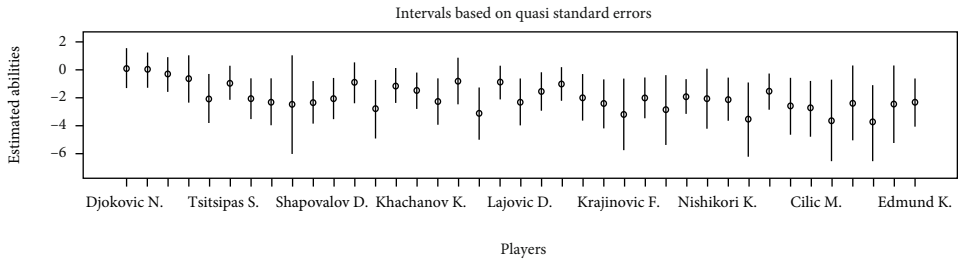


Figure 8: Estimated relative abilities of tennis players on clay court.

Table 6: Wining probabilities of players on clay court. DN: Djokovic N.; NR: Nadal R.; TD: Thiem D.; FR: Federer R.; WS: Wawrinka S.; GC: Garin C.; RA: Rublev A.; TS: Tsitsipas S.; RC: Ruud C.; SD: Schwartzman D.

	DN	NR	TD	FR	WS	GC	RA	TS	RC	SD
DN		0.528	0.607	0.682	0.713	0.731	0.737	0.739	0.751	0.775
NR	0.472		0.580	0.657	0.690	0.708	0.715	0.717	0.730	0.755
TD	0.393	0.420		0.582	0.617	0.638	0.645	0.647	0.662	0.691
FR	0.318	0.343	0.418		0.537	0.559	0.567	0.569	0.585	0.616
WS	0.287	0.310	0.383	0.463		0.522	0.530	0.533	0.549	0.581
GC	0.269	0.292	0.362	0.441	0.478		0.509	0.511	0.527	0.559
RA	0.263	0.285	0.355	0.433	0.470	0.491		0.502	0.518	0.551
TS	0.261	0.283	0.353	0.431	0.467	0.489	0.498		0.516	0.549
RC	0.249	0.270	0.338	0.415	0.451	0.473	0.482	0.484		0.533
SD	0.225	0.245	0.309	0.384	0.419	0.441	0.449	0.451	0.467	

Wining probabilities of players in the left-hand column over the players in the top row, on clay court.

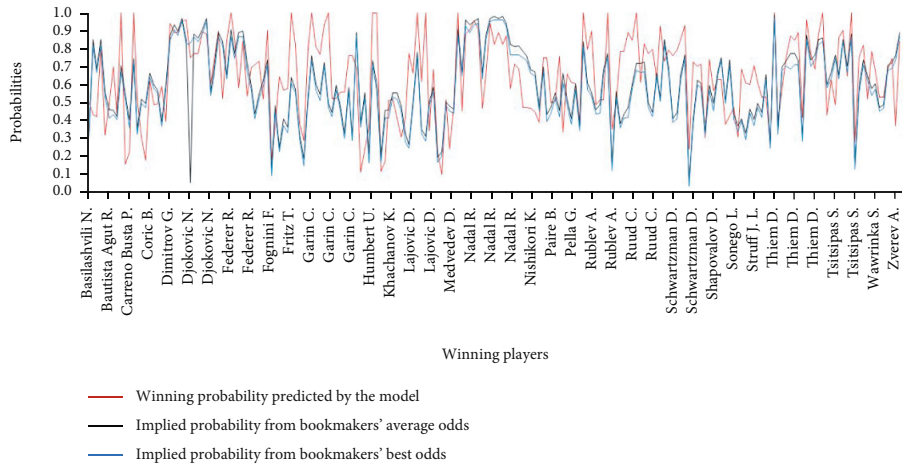


Figure 9: Match by match comparison of predicted probability with implied probability from bookmakers' average and best odds on clay court.

Table 7: Measures of predictive performance for fitted model.

	Fitted model			
	All surfaces	Hard court	Clay court	
Classification accuracy m_1	66.49%	67.89%	73.27%	
Average probability m_2	0.5871	0.6134	0.6499	
Average log probability m_3	-0.5993	-0.5744	-0.5142	
Return on investment m_4	-8.60%	8.20%	19.80%	(Avg. odds)
	8.81%	21.82%	22.77%	(Max. odds)

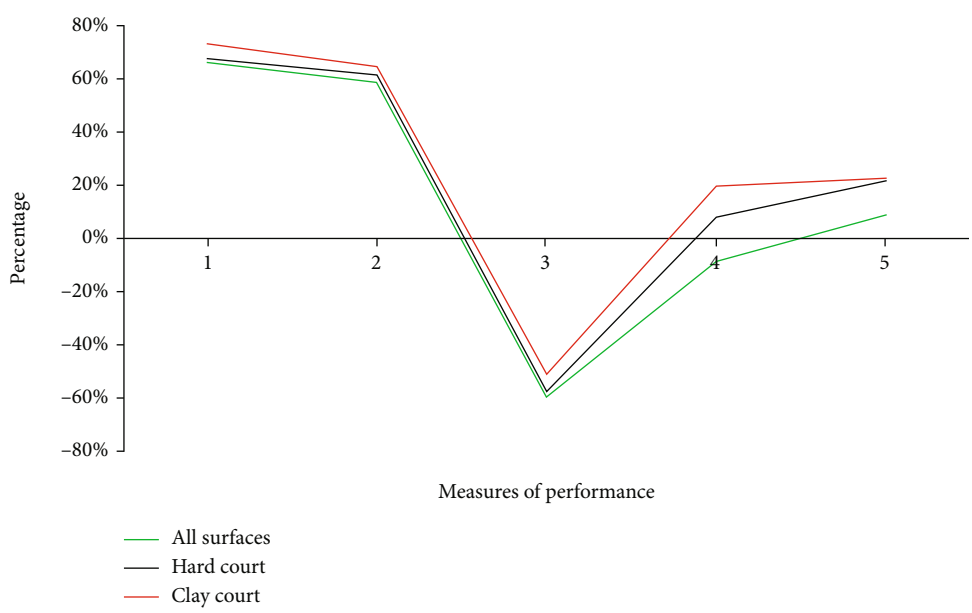


Figure 10: Measures of predictive performance for the fitted model.

Table 8: Measures of predictive performance for bookmakers' average odds.

	Bookmakers' average odds		
	All surfaces	Hard court	Clay court
Classification accuracy m_1	66.81%	66.14%	65.35%
Average probability m_2	0.6032	0.5979	0.6064
Average log probability m_3	-0.57	-0.579	-0.58
Return on investment m_4	26.88%	26.00%	23.76%

Table 9: Measures of predictive performance for bookmakers' best odds.

	Bookmakers' best odds		
	All surfaces	Hard court	Clay court
Classification accuracy m_1	62.46%	61.78%	62.38%
Average probability m_2	0.575	0.5694	0.5826
Average log probability m_3	-0.627	-0.636	-0.624
Return on investment m_4	26.44%	24.96%	25.74%

can be calculated using the following equation:

$$\hat{\pi}_{ij} = \frac{\exp\left(\hat{\alpha}_i - \hat{\alpha}_j\right)}{1 + \exp\left(\hat{\alpha}_i - \hat{\alpha}_j\right)}. \tag{8}$$

The estimated winning probability of Djokovic N. vs. Nadal R. is

$$\hat{\pi}_{12} = \frac{\exp\left(\hat{\alpha}_1 - \hat{\alpha}_2\right)}{1 + \exp\left(\hat{\alpha}_1 - \hat{\alpha}_2\right)} = \frac{\exp\left(0.1119 - 0\right)}{1 + \exp\left(0.1119 - 0\right)} = 0.528. \tag{9}$$

To check the statistical significance of the difference between Djokovic N. and Nadal R., difference of the estimates $(\alpha_1{}^{\wedge} - \alpha_2{}^{\wedge} = 0.1119)$ is compared with its standard error (SE = 0.8803) having p value = 0.8988 indicating that the difference between performance of Djokovic N. and Nadal R. is insignificant.

A 95% confidence interval for $\alpha_1 - \alpha_2$ is

$$0.1119 \pm 1.96(0.8803) \text{ or } (-1.6135, 1.8373), \tag{10}$$

and the 95% confidence interval for probability π_{12} of Djokovic to win obtained using the above interval is (0.17, 0.86).

6. Measures of Performance

In Table 7, four measures of predictive performance are presented. According to all performance measures, the model is overall well fitted for all data, but for clay court matches, it shows best results for which model accuracy is 73.27%, higher than any other model in our data. Moreover, average probability and average log probability which are two important performance measuring criteria also show that the model is well fitted for clay court data. Similarly, return on investment is highest for clay court matches on average bookmakers' available odds.

From Table 7, it is obvious that when the effect of surface is considered, the model provides good return on investment as model predictions have been compared against bookmaker's average odds and best available odds.

From Figure 10, it is quite clear that for all the measures of predictive performance, the model is overall well fitted. It shows best results for clay court in the fitted model, closely followed by results for hard court. So, it can be concluded that surface has significant impact on the performance of the players and ultimately on the outcome of the match.

Table 8 presents the measures of predictive performance for bookmakers' average odds. In this case, classification accuracy for "all surface" data is higher as compared to the fitted model case and is closely followed by classification accuracy for hard court data.

Average probability and average log probability are highest for clay court data. As far as return on investment is concerned, it is highest when the model is applied on data of all surface types. The same is the case when the bookmakers' best available odds are utilized for comparison purpose as shown in Table 9.

7. Closing Remarks

The objective of this study was to introduce a model for predicting match outcomes in men's professional tennis using ratings obtained from the fitted model to rank players keeping in view the effect of different surfaces. In studies conducted previously on this topic, the modeling approach was restricted to the models in which information on official rankings was used. In the current study, the Bradley-Terry approach has been applied on historical match results to obtain match forecasts which are more accurate according to several criteria because the evidence from soccer, tennis, and golf suggests that the official rankings of teams and players although useful predictors do not contain the entire information required for forecasting results. To achieve the objectives of the study and implement the Bradley-Terry approach, historical data on men's singles was obtained from January 2019 to September 2020 of 3439 matches. The dataset was categorized into three major categories with respect to surface type; all surfaces (consisting of matches' data on hard court, clay court, and grass court), hard court, and clay court. The Bradley-Terry model was applied on the dataset for each surface category. In our analysis on dataset of all surface types, equal weightage has been given to hard court, clay court, and grass court. Analysis for hard court and clay court has been done separately for each surface.

The analysis revealed that surface of the match has significant impact on the performance of the players. Due to this impact of the surface, our model produced superior player rankings for certain players who had been ranked very low in official ATP rankings. Using coefficients of model, ratings of players were found for each category of surface. These ratings were then used to calculate wining probabilities of players. To check the adequacy of the model,

four performance criteria were used to measure the performance of the model predictions. According to most of the performance measures, the model has shown good results for clay court data which are closely followed by hard court data. To calculate return on investment, model results were compared with the bookmakers' average odds and best available odds. It was found that return on investment was highest for clay court data when compared with both bookmaker's average odds and best odds.

Data Availability

The data used in the article is given therein.

Conflicts of Interest

No conflict of interest was declared by the authors.

Acknowledgments

The Deanship of Scientific Research (DSR), King Abdul-Aziz University, Jeddah, supported this work, under grant no. G: 358–247–1442. The authors, therefore, gratefully acknowledge the DSR technical and financial support.

References

[1] B. L. Boulier and H. O. Stekler, "Are sports seedings good predictors?: an evaluation," *International Journal of Forecasting*, vol. 15, no. 1, pp. 83–91, 1999.

[2] P. K. Newton and J. B. Keller, "Probability of winning at tennis I. Theory and data," *Studies in Applied Mathematics*, vol. 114, no. 3, pp. 241–269, 2005.

[3] W. J. Knottenbelt, D. Spanias, and A. M. Madurska, "A common-opponent stochastic model for predicting the outcome of professional tennis matches," *Computers & Mathematics with Applications*, vol. 64, no. 12, pp. 3820–3827, 2012.

[4] S. Akhtar, P. Scarf, and Z. Rasool, "Rating players in test match cricket," *Journal of the Operational Research Society*, vol. 66, no. 4, pp. 684–695, 2015.

[5] P. Scarf, X. Shi, and S. Akhtar, "Modelling batting strategy in test cricket," in *In Progress in Industrial Mathematics at ECMI 2008*, pp. 481–489, Springer, Berlin, Heidelberg, 2010.

[6] F. J. Klaassen and J. R. Magnus, "Forecasting the winner of a tennis match," *European Journal of Operational Research*, vol. 148, no. 2, pp. 257–267, 2003.

[7] M. De Araujo Fernandes, "Using soft computing techniques for prediction of winners in tennis matches," *Machine Learning Research*, vol. 2, no. 3, pp. 86–98, 2017.

[8] P. Scarf and S. Akhtar, "An analysis of strategy in the first three innings in test cricket: declaration and the follow-on," *Journal of the Operational Research Society*, vol. 62, no. 11, pp. 1931–1940, 2011.

[9] F. Lisi and G. Zanella, "Tennis betting: can statistics beat bookmakers?," *Electronic Journal of Applied Statistical Analysis*, vol. 10, no. 3, 2017.

[10] S. Akhtar and P. Scarf, "Forecasting test cricket match outcomes in play," *International Journal of Forecasting*, vol. 28, no. 3, pp. 632–643, 2012.

[11] P. Scarf, X. Shi, and S. Akhtar, "On the distribution of runs scored and batting strategy in test cricket," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, vol. 174, no. 2, pp. 471–497, 2011.

[12] I. McHale and A. Morton, "A Bradley-Terry type model for forecasting tennis match results," *International Journal of Forecasting*, vol. 27, no. 2, pp. 619–630, 2011.

[13] S. J. Koopman and R. Lit, "The analysis and forecasting of tennis matches by using a high dimensional dynamic model," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, vol. 182, no. 4, pp. 1393–1409, 2019.

[14] S. Easton and K. Uylangco, "Forecasting outcomes in tennis matches using within-match betting markets," *International Journal of Forecasting*, vol. 26, no. 3, pp. 564–575, 2010.

[15] S. A. Kovalchik, "Searching for the GOAT of tennis win prediction," *Journal of Quantitative Analysis in Sports*, vol. 12, no. 3, pp. 127–138, 2016.

[16] R. A. Bradley, "14 paired comparisons: some basic procedures and examples," *Handbook of Statistics*, vol. 4, pp. 299–326, 1984.

[17] N. Dingle, W. Knottenbelt, and D. Spanias, "On the (page) ranking of professional tennis players," in *Computer Performance Engineering*, pp. 237–247, Springer, Berlin, Heidelberg, 2012.

[18] G. H. Kuper, G. Sierksma, and F. C. Spieksma, *Using tennis rankings to predict performance in upcoming tournaments (No. 14034-EEF)*, University of Groningen, Research Institute SOM (Systems, Organisations and Management), 2014.

[19] S. R. Clarke and D. Dyte, "Using official ratings to simulate major tennis tournaments," *International Transactions in Operational Research*, vol. 7, no. 6, pp. 585–594, 2000.

[20] L. Knorr-Held, "Dynamic rating of sports teams," *Journal of the Royal Statistical Society: Series D (The Statistician)*, vol. 49, no. 2, pp. 261–276, 2000.

[21] Š. Lyócsa and T. Výrost, "To bet or not to bet: a reality check for tennis betting market efficiency," *Applied Economics*, vol. 50, no. 20, pp. 2251–2272, 2018.

[22] M. J. Dixon and S. G. Coles, "Modelling association football scores and inefficiencies in the football betting market," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 46, no. 2, pp. 265–280, 1997.

[23] T. Barnett and S. R. Clarke, "Combining player statistics to predict outcomes of tennis matches," *IMA Journal of Management Mathematics*, vol. 16, no. 2, pp. 113–120, 2005.

[24] N. Dewart and J. Gillard, "Corrigendum to: Using BradleyTerry models to analyse test match cricket," *IMA Journal of Management Mathematics*, vol. 30, no. 2, 2019.

[25] D. Firth, "Bradley-Terry models inR," *Journal of Statistical Software*, vol. 12, no. 1, pp. 1–12, 2005.

[26] D. R. Hunter, "MM algorithms for generalized Bradley-Terry models," *The Annals of Statistics*, vol. 32, no. 1, pp. 384–406, 2004.

[27] R. R. Davidson and P. H. Farquhar, "A bibliography on the method of paired comparisons," *Biometrics*, pp. 241–252, 1976.

[28] T. Hastie and R. Tibshirani, *Classification by pairwise coupling*, In Advances in neural information processing systems, 1998.

[29] P. V. Rao and L. L. Kupper, "Ties in paired-comparison experiments: a generalization of the Bradley-Terry model," *Journal of the American Statistical Association*, vol. 62, no. 317, pp. 194–204, 1967.

[30] A. Agresti and M. Kateri, *Categorical Data Analysis*, Springer, 2011.

[31] R. L. Plackett, "The analysis of permutations," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 24, no. 2, pp. 193–202, 1975.

[32] P. Diaconis, "Group representations in probability and statistics," *Lecture notes-monograph series*, vol. 11, 1988.

[33] H. A. David, "Note on "an extreme null distribytion approach to the problem of paired comparisons" by Particia S. Costello and Douclas A. Wolfe," *Communications in Statistics-Theory and Methods*, vol. 17, no. 11, pp. 4005–4009, 1988.

[34] P. Diaconis, "Group representations in probability and statistics," *Lecture notes-monograph series*, vol. 11, pp. 1–192, 1988.

Hindawi
International Journal of Photoenergy
Volume 2022, Article ID 9194537, 21 pages
https://doi.org/10.1155/2022/9194537

Research Article

A Practical Approach for Predicting Power in a Small-Scale Off-Grid Photovoltaic System using Machine Learning Algorithms

Aadyasha Patel,[1] **O. V. Gnana Swathika**,[1] **Umashankar Subramaniam**,[2,3] **T. Sudhakar Babu**,[4] **Alok Tripathi**,[1] **Samriddha Nag**,[1] **Alagar Karthick**,[5] **and M. Muhibbullah**[6]

[1]*School of Electrical Engineering, VIT Chennai, Chennai 600127, India*
[2]*Department of Communications and Networks, Renewable Energy Laboratory, College of Engineering, Prince Sultan University, Riyadh 11586, Saudi Arabia*
[3]*Department of Energy and Environmental Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Saveetha Nagar, Thandalam, Chennai, 602105 Tamilnadu, India*
[4]*Department of Electrical and Electronics Engineering, Chaitanya Bharathi Institute of Technology (CBIT), Hyderabad 500075, India*
[5]*Renewable Energy lab, Department of Electrical and Electronics Engineering, KPR Institute of Engineering and Technology, 641407, Coimbatore, Tamilnadu, India*
[6]*Department of Electrical and Electronic Engineering, Bangladesh University, Dhaka 1207, Bangladesh*

Correspondence should be addressed to M. Muhibbullah; m.muhibbullah@bu.edu.bd

Received 28 October 2021; Accepted 18 January 2022; Published 12 February 2022

Academic Editor: Laurentiu Fara

Climate change and the energy crisis substantially motivated the use and development of renewable energy resources. Solar power generation is being identified as the most promising and abundant source for bulk power generation. However, solar photovoltaic panel is heavily dependent on meteorological data of the installation site and weather fluctuations. To overcome these issues, collecting performance data at the remotely installed photovoltaic panel and predicting future power generation is important. The key objective of this paper is to develop a scaled-down prototype of an IoT-enabled datalogger for photovoltaic system that is installed in a remote location where human intervention is not possible due to harsh weather conditions or other circumstances. An Internet of Things platform is used to store and visualize the captured data from a standalone photovoltaic system. The collected data from the datalogger is used as a training set for machine learning algorithms. The estimation of power generation is done by a linear regression algorithm. The results are been compared with results obtained by another machine learning algorithm such as polynomial regression and case-based reasoning. Further, a website is developed wherein the user can key in the date and time. The output of that transaction is predicted temperature, humidity, and forecasted power generation of the specific standalone photovoltaic system. The presented results and obtained characteristics confirm the superiority of the proposed techniques in predicting power generation.

1. Introduction

Renewable or nonconventional sources of energy are something that replenishes itself at the speed of its consumption. Some examples of renewable energy are solar, wind, tidal, waves, and geothermal. Renewable energy ventures are being undertaken in developed as well as developing countries. The majority of the nonconventional energy is harnessed to get electricity, which is more efficient, clean, does not pollute the environment, and is costeffective in the long run. Among all the nonconventional forms of available resources, solar energy is most abundantly found and the amount of

solar energy that hits the surface of the earth in an hour is enough to fulfill global needs for an entire year. This power from the Sun is used in a variety of ways, such as solar heating, solar thermal energy, photovoltaic, and photosynthesis. For large scale utility solar installations, solar thermal is employed.

SAPV system or off-grid system is one that is not connected to any electricity distribution system. They are classified into two types: direct-coupled system and stand-alone system. A direct-coupled system does not have a battery to store energy and can be used only when there is enough sunlight available. Stand-alone systems, on the other hand, employ battery backups to be used at any time of the day. They may also be equipped with a sensor and datalogger to sense and record the different parameters that need to be observed, respectively, such as pyranometer, anemometer, temperature sensor, humidity sensor, and system current and voltages. For further analysis of the obtained data, ML tools could be used.

Data collection to digitize the industry and society has become a priority nowadays. A summary of the up-to-date ML with its practical applicability is discussed in [1, 2]. The demand for new research concerned with AutoML (automated machine learning) is discussed in [3] so that dependency on domain experts and time-consuming data manipulation is reduced. To estimate the power production of an established PV plant, a comparative analysis of three different methods is performed in [4] between a Sandia National Laboratories model, a multilayer perceptron neural network, and a regression approach. Statistical ML approaches demonstrated more accurate power predictions. Authors in [5] propose a new feature selection-based distributed ML approach to sense the active signatures of diverse power system events. The methodology presented in [5] is demonstrated in an interrelated two-area-based microgrid with numerous kinds of energy generation arrangements. Residential energy management system is proposed in [6] such that possible loads are effectively switched on to local energy storage based on its charge-discharge cycles and grid availability. The automation of switchover is achieved using artificial neural networks and support vector machine, machine learning algorithms. The work presented in [7] gives a summary of predicting means of solar irradiation using ML styles. In [8], the application of deep learning artificial intelligence procedures is reported to predict the energy ingestion and power generation combined with the weather forecasting mathematical simulation. Ensemble-based models are proposed in [9] to fulfil a long-term forecasting for territorial PV generation. Various predictive models based on ML style are trained and confirmed in [10] to estimate the real PV output power in orientation with a satisfactory time period accurately. In [11], an in-depth analysis of the current methods used in the prediction of solar irradiance is presented to enable the selection of a suitable forecast method for the proposed system.

Authors in [12, 13] elaborated the design of an economic datalogger capable of measuring electrical and meteorological data, which also meets the IEC requirements for SAPV systems to operate in remote areas. The data from the logger is collected via a SD card since telecommunications is a concern. All the required parameters are monitored with high accuracy and low power consumption. The datalogger consists of Arduino Mega 2560 board with ATmega2560 chip and DS1307 real-time clock chip for date stamping in the SD card every time data is logged. The results and findings are recorded and compared with the data that was taken by one commercial datalogger DataTaker DT80 during the testing stage in [14]. In [15], a datalogger, to monitor isolated PV systems in developing nations, was developed using Arduino. The system met all the relevant requirements of the IEC standard. It is tested in harsh weather conditions and exhibited comparable performance to commercial systems, hence is found to be reliable. The microcontroller-based data acquisition systems are designed and developed for feasible operation in [16–20]. In [21], the author developed a wireless data acquisition system for weather station monitoring consisting of sensors to measure various atmospheric parameters, whose data that is collected and conditioned using precision electronic circuits and interfaced to a PC using RS232 linking via a wireless unit. The processing and display of the collected data are done using the LabVIEW program. Also, the data is available over the internet to any user. The key task of observing temperature and humidity along with the transmission of this data in the short message service form to users' cellular phone is done by the system which also offers a data-logging facility. For further analysis by [22], the logged data can be transferred to a PC having a graphical user interface program.

[23] discuss a software program that could be used to acquire any Unidalog (universal datalogger) remotely as remote terminal unit by using the GSM network as an intermediate for data communication. Processing software program design models using transformative development models to get the system iteratively to make segments are part of the telemetry monitoring. Further, in [24, 25], a lossless algorithm is presented based on the statistical information with a compression ratio of up to 14: 1 to significantly reduce the storage and telecommunication costs of large volumes of measured data. The development of an economical field datalogger model using Raspberry Pi and industrial sensors is described, and an online rain flow count algorithm is implemented in [26]. The software and hardware style for constructing a wireless datalogger for a thermal authentication system to measure and operate in a stated temperature range and also agreeing with International and European protocols for validation of pharmaceutical, biotechnology, and the medical device is described in [27], where the temperature ranges from -60°C to 150°C with total system measurement accuracy of 0.1°C. CBR algorithm is employed to discover the past cases which are the most identical to the present case in [28]. The exactness of the model is confirmed using past 10 years data.

Few features of photovoltaics, such as loss of load probability and peak solar hour, were demonstrated and simulated by authors in [29, 30]. Better battery management systems are proposed in [31–34] to extend the life cycle of the batteries and smooth transition between charging and discharging cycles. Several methodologies are studied in

[35–38] to minimize the cost and size of the SAPV system with DC-DC converters. The parallel connection of the MPPT system decreases the undesirable effect of power converter losses in the total efficiency since only a portion of the generated power is processed by the MPPT system as studied in [39]. For this MPPT to operate with functions of the step-up converter and battery charger, a simple bidirectional DC-DC power converter is proposed. The operational characteristics of the proposed circuit are being investigated with the implementation of a model in a real-time application. In [40], authors discussed the scheme that comprises three DC-DC converters and that operate in accordance with the main bus load demand and the PV panel power. Execution of control elements through the usage of One-Cycle Control permits great accurateness of convergence, fast response to transient conditions, and low-cost analogue application. Simulations of transition conditions are achieved through actual states such as the absence of sunshine, cloud-edge effects, sluggishly passing clouds. The design and the factors that affect the performance of the SAPV consist of PV sizing, the consequence of shadowing, temperature, and dust, leading to reduced SAPV output power and progressive degradation of the PV generator and cells permanently. Special attention is drawn by authors in [41] to the need for remote monitoring and examination and standardized upkeep measures, and a cautious design segment is encouraged for SAPV systems to remain a favored power substitute for telecommunication applications. An appropriate and rational technique for selecting the optimum constraints of the particle swarm optimization algorithm takes into account the topology, constraints of the DC-DC converter, and the configuration of solar panels. The ideal value of the sampling period for the digital MPP controllers delivering their peak performance is determined in [42] based on an innovative method.

Compared to the direct duty cycle method, the voltage reference control method has the attraction of fast convergence and minor oscillations in steady state, also analytical calculations of controller gains are possible in [43]. In [44], the design and cost estimation of SAPV along with MPPT, inverter, charge controller, and lead-acid batteries are studied. The sizing of SAPV is achieved by taking into account the radiation and electrical data for a typical household in remote areas or villages using MATLAB/Simulink. The Life Cycle Cost analysis is used to evaluate the economic feasibility of the system. The results are encouraging, in the sense that, Yemeni remote areas could be electrified. The errors made during designing, costing, sizing, installing, and maintenance of solar PV system are discussed in [45]. The design was done in 2 dissimilar ways. The total price came out to be between US$422.5 and US$ 107.5. An alternative method to MPPT was discussed in [46] grounded on the transformation of the energy parameters into new charge-related parameters specified by the Joint Research Center and the IEC Standard 61724. In [47], authors proposed a scheme which when compared with traditional schemes needs a reduced number of switching elements in the battery path, hence significantly reducing the cost and size of inductors involved with battery charging circuits by employing simple power management topology. The efficiency of the power conversion stage varies with levels of irradiance. Hence, to maximize the system efficiency over a broader array of real-time working environments, a novel adaptive control scheme was proposed in [48].

For the fitted electrical devices in an independent house, the battery storage necessities are to address on the kind, size, and functioning sequence. The relation between the accessible stored energy and the size/operational order of the house's electrical equipment are highlighted in [49]. The rules for the selection of a suitable site/location together with the technique for the assessment of solar energy supply at the preferred site are provided in [50]. Sizing of various solar system components such as panels, batteries, inverter, charge controllers, and other accessories with daily load energy demand considerations was studied in [50] to design and install a solar PV system. The energy management of a SAPV structure is designed taking into consideration the drastic changes in weather conditions and rapid changes in the energy necessities of the client. A storage unit such as a super-capacitor is used to bring under control the mismatch of energy among the load and the SAPV system. The effectiveness and performance of the system in [51] is evaluated using MATLAB/Simulink software. The case studies of several SAPV systems in various parts of the world from [52–55] present information on guidelines, components required, the energy demand and consumed portfolio, and the process of calculation.

IoT-based solutions are proposed by researchers in [56–58] to make available appropriate and suitable source management, load shedding, data acquisition, and control of the SAPV systems and monitor and evaluate the electrification projects. [59] show that IoT provides the capability to bring into use the MATLAB, ThingSpeak, and other tools/functions by granting the authority to one person to operate the forecasting system. Overall, the entire effort lessens. MAPE does not depend on the scale, there is possibility of intuitive interpretation and it is also easy to compute is presented in [60]. The authors in [61–63] discussed that the simulation and design of a hybrid system in an islanded location is actually feasible ecologically and economically. There is a drop in emission of CO_2, lesser energy, and net cost. The work done in [64–68] evaluates the numerous strategies of load dispatch of islanded microgrid systems. The outcome attained gives a proper parameter to compare and approximate sizes of modules and its related costs.

This paper was aimed at developing a novel scaled-down prototype of an IoT-enabled datalogger for PV panels that are installed in a remote location where human intervention is not possible due to harsh weather conditions or other circumstances. An IoT platform is used to store and visualize the captured data. The collected data from the datalogger is used as a training set for ML algorithms. The estimation of power generation is done by a LR algorithm. Further, a website is developed wherein the user can input the date and time. The output of that transaction is predicted temperature, humidity, and forecasted power generation of the specific standalone PV system. Hence, the proposed IoT-enabled datalogger extensively contributes in realizing a

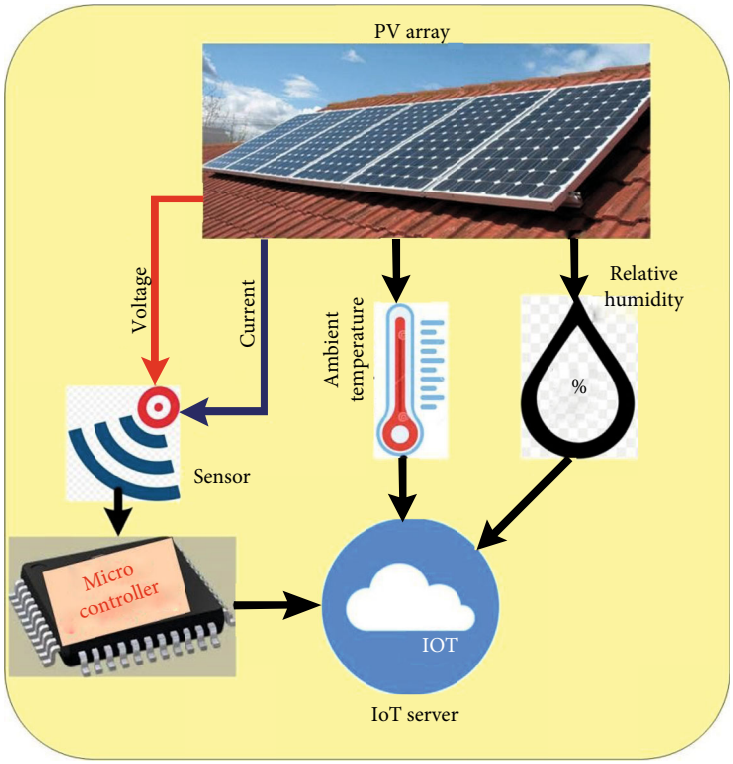


Figure 1: Data from PV panel to an IoT server.

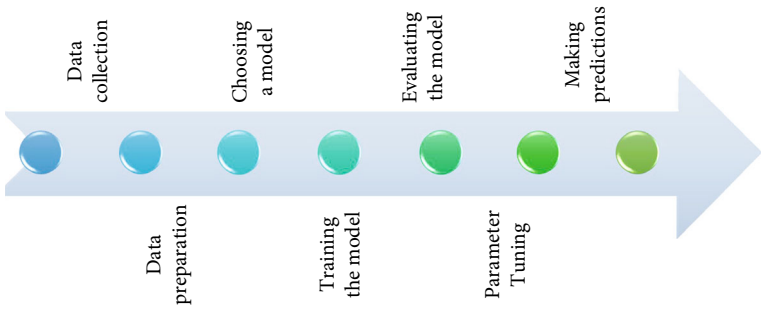


Figure 2: Block diagram of the basic machine learning stages.

self-sufficient small-scale off-grid photovoltaic system since it does not require involvement of any individual in providing and predicting power. The proposed small-scale off-grid photovoltaic system has applications in electrification of secluded, rural, isolated, and remote areas/homes. It can also be utilized by regions facing frequent power cuts. This system is also capable of supplying electricity under natural disaster circumstances.

The organization of the article is as follows: Section 2 explains the methodology to be followed for realizing the IoT-enabled datalogger for SAPV networks. It also discusses how ML algorithms are employed to forecast power generation. The purpose of creating the website is also discussed in this section. Section 3 discusses the system designing in detail for the hardware prototype of the proposed datalogger. It also explains how the logged data serves as the training data to the ML algorithms. The output of the algorithm is the forecasted power generation. Section 4 elaborates how the datalogger is implemented on the 40 W PV panel and also how the data is logged. LR algorithm is effectively used on the training data to predict the power generation in the website as well. Section 5 concludes working model of the datalogger for SAPV networks.

2. Methodology

2.1. Datalogger for SAPV Network. The main objective of this work is to develop a low-cost IoT-enabled datalogger for a remotely deployed standalone solar PV system. The datalogger contains calibrated digital sensors and a microcontroller. The sensors monitor and logs performance parameters of the PV system and the microcontroller which has a connection to an IoT server through a Wi-Fi connection send the measured data to the server. While collecting data, the sensors may tend to log erroneous values or miss logging a few values. To avoid the above-mentioned scenarios, two

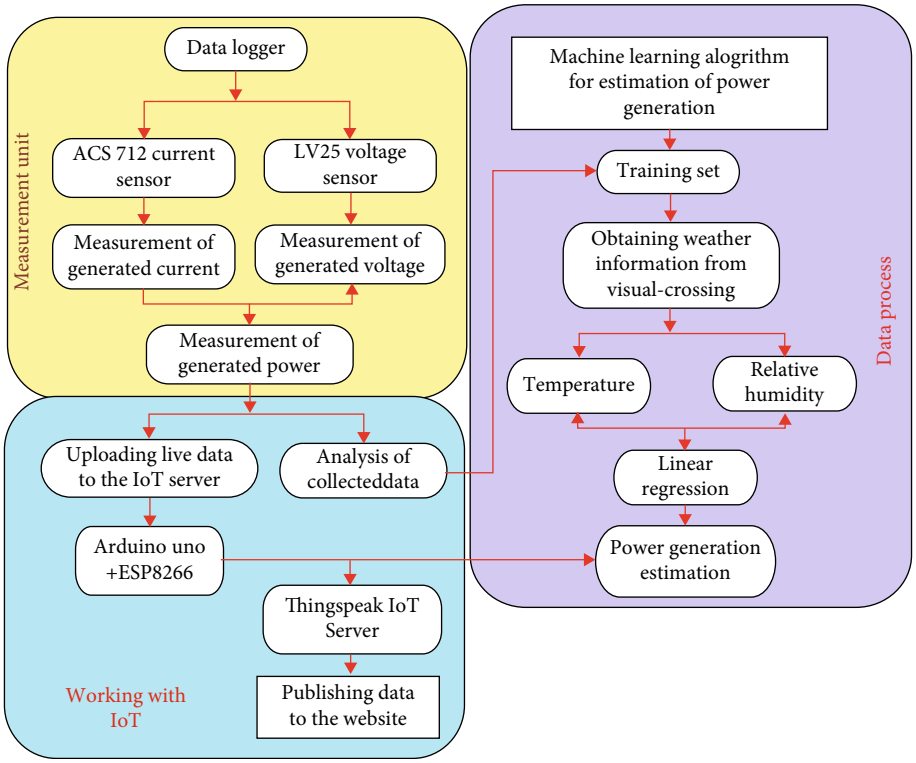


Figure 3: The methodology of proposed machine learning-based power forecasting in SAPV networks.

data predicting algorithms are employed. The performance of the two algorithms is compared, and based on the application, one of the algorithms is deployed. On the server-side, users can view live generation performance and can export data to an excel sheet for further data analysis.

A datalogger is designed for the PV system as demonstrated in Figure 1. The datalogger consists of a sensor to measure photovoltaic parameters such as voltage and current and a microcontroller that is connected to an IoT server. The sensors collect information and transmit it to the computer via direct serial connection or directly being sent to an IoT server, where it is further viewed and analyzed. Appropriate software and weather data providing platforms are used to store, display, and process the collected data. The collected data then proceeds to the ML algorithm for estimating power generation for the next 30 days.

2.2. Machine Learning-based Prediction Algorithms. The machine learning stages are shown in Figure 2. The datalogger initially collects performance parameters of PV panel, voltage, and current and sends it to the ThingSpeak server. This server provides storing and graphical representation facility. The collected data from the datalogger is used as a training set for the ML program where the algorithm analyzes the collected data and attempts to relate measured data with weather parameters such as temperature and relative humidity.

The algorithms employed here are the following:

2.2.1. Linear Regression. LR is a mathematical approach to derive a linear relationship between two variables. It does so by finding a linear line that best fits the data points. The equation of a line is used to predict output values that are not present in the data set. LR is fast and used extensively in analysing and predicting data sets. Upon executing the calculation of LR, it creates a LR line equation. The equation has two variables: temperature and relative humidity, and the substitution of these two values will predict the power generation. This process also minimizes errors.

2.2.2. Polynomial Regression. PR has an independent variable x and a dependent variable y. The relationship between the variables is a polynomial of nth degree. Hence, the plot forms a curve. This algorithm can adjust to multiple variety of curves but requires additional effort to realize the appropriate fit and interprets the role of its independent variables.

2.3. GUI for Viewing Forecasted Power. A webpage, which is created, acts as the GUI to view the generated parameters. The user needs to feed the date and time so that the webpage fetches its corresponding temperature and relative humidity from weather forecasting servers to display generated wattage. The constant record of performance and failure data is enabled by IoT so that it can used for analytics of predicting and forecasting the impending power generation potentials, revenue generation, etc. Photovoltaic systems fitted at isolated or faraway places from the control center can be accessed using IoT which also helps in improving the efficiency of the system, reducing human involvement and supervision time, and facilitating network management.

Table 1: Specifications of voltage sensing device.

Component	Model no.	Detection range	Accuracy	Analog resolution
Voltage sensor	LV25	0.02445 V to 25 V	±0.01 V	0.00489 V

3. System Designing

Figure 3 gives a detailed overview of the methodology incorporated in this article. The design is categorized into two parts: hardware and software. The hardware platform involves realizing a low-cost data monitoring system that gathers data and stores them in the IoT server. To predict power generation, the machine learning algorithm requires at least 4 months of performance data of PV panel for best results. Due to time constraints, the collection of so much data was implausible. Hence, appropriate approximate values were used for training sets. The software part consists of MATLAB for ML algorithm, an IDE for Arduino, and a webpage for viewing the real-time performance of solar PV panel and estimation of generation. ThingSpeak platform is programmed to show the voltage and current of a standalone PV panel.

3.1. IoT Based Datalogger

3.1.1. Selection of Microprocessor. The selection of suitable hardware was based on the use of open-source software and hardware that will allow us to attain the economical objective of the final system. Among the plentiful microprocessors based upon open-source hardware available in the market, Arduino Uno is chosen for its low-cost, flexibility, and widespread popular developer community. The Arduino DIY board is based on the ATmega328 microcontroller and has 14 digital I/O pins, 6 analogue inputs, a power jack, USB connection, and a reset button. The board comprises of everything essential to support the microcontroller and it can work using an external supply or any standard USB port using USB type A to USB type B connector cable.

It offers advantages over other development kits such as the following:

(i) Easy programming tool: Arduino programming is done by its free IDE software being beginner-friendly yet flexible enough for advanced users. Arduino IDE is a very small, lightweight program that requires very basic system requirements:

(a) Microsoft Windows XP with SP2, Windows 7, Windows 8/8.1, and Windows 10 operating system

(b) Microsoft .NET Framework 3.5 or higher

(c) Intel Pentium/AMD Athlon processor or equivalent running at 1 GHz or more

(d) 512 MB RAM (1 GB RAM recommended)

(e) 10 MB free hard drive space or more (only for PROGRAMINO IDE for Arduino)

(f) It can run on any modern low-end PC, and unlike Raspberry Pi, it does not require any external development environment or secondary operating system to work with

(g) It uses a modified version of basic C/C++ programming language and has a built-in compiler with an error notifier

(ii) Low-cost: Arduino board especially Nano and Uno variants are relatively cheaper than their counterparts by a high margin. The least expensive Arduino Uno board costs around ₹250, and it can even be assembled by hand which will reduce its costs even further. Compatible sensors are also cheap and easily available in the market

(iii) Cross-platform: Arduino IDE is available on most of the commonly used operating systems. Whereas its rivals like Raspberry Pi require their own OS and are difficult to set up especially in Macbooks. Most of the DIY development board does not support Linux and ChromeOS but Arduino does

(iv) Expandable software: The Arduino IDE supports C/C++ language, and it receives OTA update over time. It also has a board manager to download and installs an external development board. It also has a library manager which lists available external library from its GitHub page. Many extensions are also available for the Arduino IDE

(v) Expandable Hardware: Arduino supports UART, I2C, and SPI protocol through which it can connect to a plethora of modems. Arduino is an open-source project, so an experienced circuit designer can add more features to it and publish it for others to use. There are many modified versions of Arduino board available to choose from as per one's need and requirement

3.1.2. Selection of Sensors and Internet Module. The sensors used to collect the data required to carry out this work are described in the following subsections.

(1) Voltage sensor. The voltage is measured and recorded by the new low-cost voltage sensor LV25, which works on the principle of resistive voltage divider design. The voltage detection sensor is very cheap and easily available in the market. Its interfacing with Arduino Uno board is also very simple. It uses a potential divider to decrease any input voltage by a factor of 5 permitting the user to use the analog pin of a microcontroller to monitor voltage greater than its capacity. The specifications of the voltage-sensing device are given in Table 1. Some of the features of LV25 are as follows:

(i) Small in size and compact

(ii) High personnel safety

(iii) High-degree of accuracy

Table 2: Specifications of current sensing device.

Component	Model no.	Range	Internal conductor resistance	Output error	Bandwidth	Output sensitivity
Current sensor	ACS712	0 – 5 A	1.2 mΩ	1.5% at TA = 25°C	80 kHz	66 to 185 mV/A

Table 3: Specifications of Wi-Fi module.

Component	Manufacturer	Model no.	Operating values		Recommended values	
			Voltage	Current	Voltage	Current
Wi-Fi module	Espressif systems	ESP8266	3.3 V	400 mA	3.3 V	300 – 800 mA

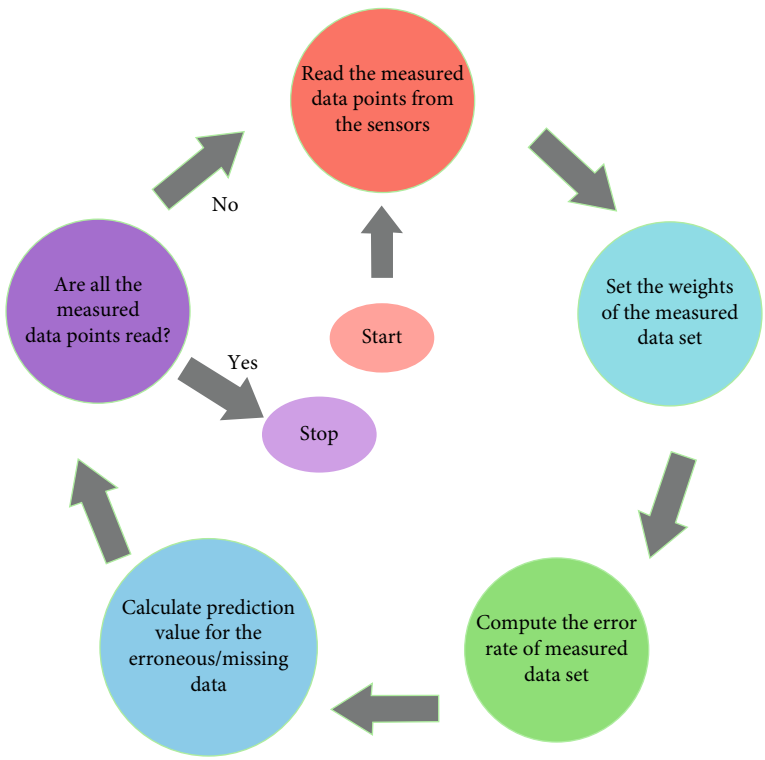


Figure 4: Prediction process of erroneous/missing data.

(iv) Nonsaturable

(v) Wide dynamic range

(vi) No external components required

(vii) Easy to use with microcontrollers

(2) Current sensor. The PV current is measured by ACS712 hall-effect sensor. It is a low-cost sensor and is compatible with Arduino Uno. The device consists of a precise, low-offset, and linear hall sensor circuit with a copper conduction path located near the surface of the die. The hall-effect sensor detects current by magnetic field generation. Upon detection, a proportional voltage is generated which becomes the raw data measurement for the microcontroller. Arduino Uno uses this raw data to calculate proportional current for the corresponding voltage. The specifications of the current sensing device are given in Table 2. Some of the features of ACS712 are as follows:

(i) Better accuracy

(ii) Minimum power loss

(iii) Low-noise analogue signal path

(iv) Device bandwidth is set via the new filter pin

(v) Stable output offset voltage

(vi) Near zero magnetic hysteresis

(3) Wi-Fi module. The Arduino Uno board cannot connect to the internet on its own. Hence, a small and cost-efficient Wi-Fi module ESP8266 is being employed here. It provides internet access using a built-in micro-Wi-Fi modem and

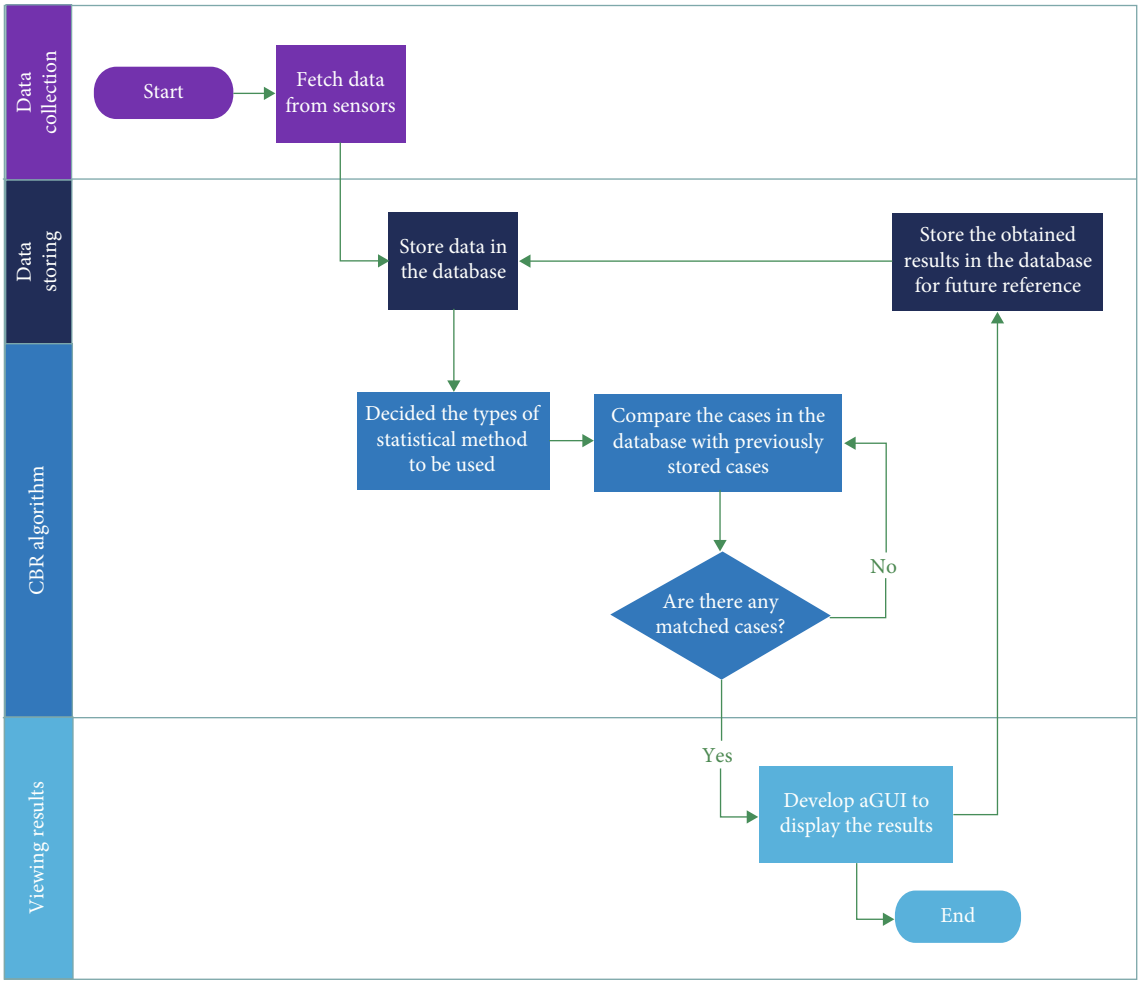


Figure 5: CBR algorithm.

Table 4: Specifications of Arduino Uno.

Parameter	Voltage		Operating current	Microcontroller
	Operating	Recommended		
Rating	5 V	5 – 12 V	0.04 A	Built-in ATmega328

Table 5: Rating of the PV panel.

Parameter	Maximum voltage	Maximum current	Maximum power generation
Rating	16.4 Vmp	2.44 Amp	40 W

supports Wi-Fi 802.11 b/g/n protocol offering a complete and self-contained Wi-Fi networking solution. It allows to either host the application or offload all Wi-Fi networking functions from another application processor.

ESP8266 can boot up directly from an external flash drive. It has its memory and integrated cache storage. It remembers data like previously connected Wi-Fi SSID (service set Iidentifier) and password and automatically gets connected to the remembered Wi-Fi network if it is available. It can provide internet accessibility basically to any device which supports a UART connection. The specifications of the Wi-Fi module are given in Table 3. Some features of ESP8266 are as follows:

(i) Low-cost: ESP8266 is very cheap and can be found in the market for less than ₹400. It is fast and easy to use and compatible with most of the development boards

(ii) Plug and play: It requires very basic wiring with the Arduino. It has two inbuilt LED (light-emitting diode) for showing transmission and receiving operation

(iii) Integrated Circuit: ESP8266 does not require any external resistor or capacitor, it has an integrated TR switch, balun, low-noise amplifier, power amplifier and matching network; integrated PLL (phase-locked loop), regulators, power management units,

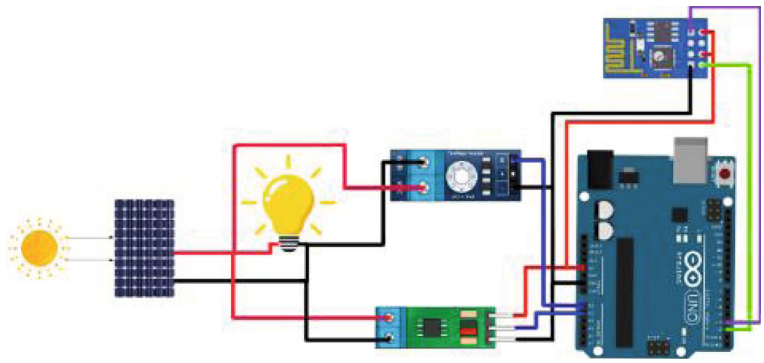


Figure 6: Circuit diagram of datalogger.

Figure 7: Testing datalogger with solar PV emulator.

and +19.5 dBm output power in 802.11b mode; integrated temperature sensor; and integrated MIMO antenna

(iv) Fast boot and recovery: ESP8266's latest firmware allows it to boot up quickly and recover cache memory in no time. It wakes up and transmits packets in less than 2 ms and has very low power consumption

3.2. Forecasting Based on Linear and Polynomial Regression Algorithm. The collected data from the datalogger is used as a training set for the machine learning program, whereas the machine learning algorithm analyzes the collected data and attempts to relate measured data with weather parameters such as temperature and relative humidity. Linear and PRs are a mathematical approach to derive a linear relationship between two variables. It does so by finding a linear line that best fits the data points. The equation of a line is used to predict output values that are not present in the data set. LR is fast and is used extensively in analysing and predicting data sets. The prediction process of erroneous/missing data is shown in Figure 4.

The estimation of power generation is executed by linear and PR algorithms between generated wattage and temperature, humidity for estimating power generation. Irradiance, wind direction, wind speed, and dew point are also the key elements that affect PV panels. Irradiance measurement instrument (pyranometer) is an expensive tool, and in monitoring wind direction, wind speed is also an arduous thing to do. But if the cost is not a concern, then one can further enhance the prediction by adding these factors too.

Figure 8: Solar PV emulator.

3.3. Challenges. During the development of the prototype, there were many issues for which alternative solutions were brainstormed and implemented. The first constraint faced was the implementation of ESP8266 in the circuit since programming it as per need was a challenge. The initial plan was to use a GSM module, but it was expensive and can only send a text message to a GSM number, and the proposed work required something which could store data in the cloud.

Secondly, IoT is often thought to be the future of the internet. Ever since the world has got fast personal internet, people are working towards an interconnected world where

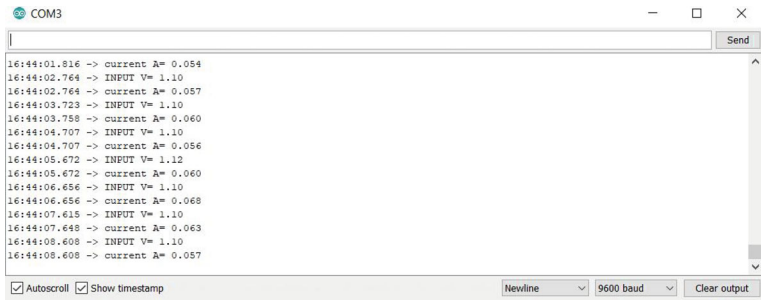


Figure 9: Monitoring power generation of Solar PV emulator.

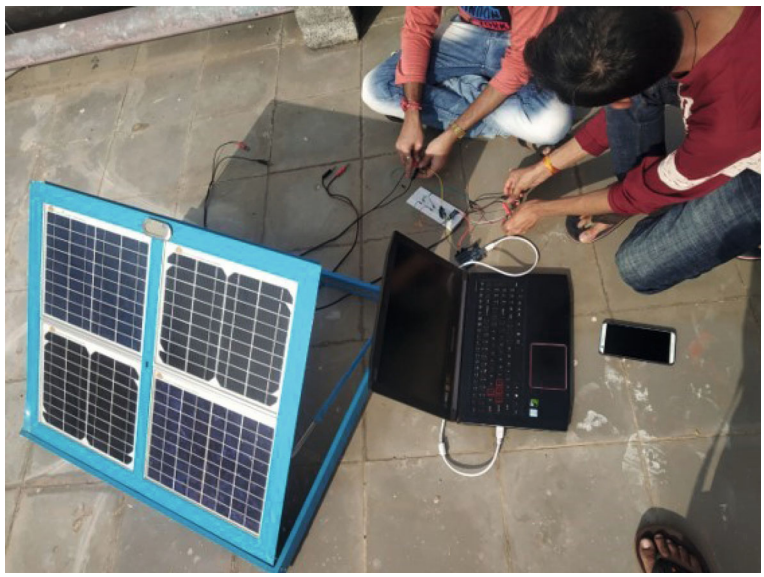


Figure 10: Setting up datalogger with roof-top panel to monitor performance data.

	A	B	C
1	Date	Time	Power Generation
2	01-02-2020	09:00:31	35.38568841
3	01-02-2020	09:30:31	35.84256994
4	01-02-2020	10:00:31	36.29945147
5	01-02-2020	10:30:31	35.38568841
6	01-02-2020	11:00:31	36.39082778
7	01-02-2020	11:30:31	36.48220409
8	01-02-2020	12:00:31	36.75633301
9	01-02-2020	12:30:31	36.93908562
10	01-02-2020	01:00:31	36.75633301
11	01-02-2020	01:30:31	35.84256994
12	01-02-2020	02:00:31	36.6649567
13	01-02-2020	02:30:31	36.61926855
14	01-02-2020	03:00:31	36.75633301
15	01-02-2020	03:30:31	35.38568841
16	01-02-2020	04:00:31	35.84256994
17	01-02-2020	04:30:31	36.29945147
18	02-02-2020	10:30:22	36.48220409
19	02-02-2020	11:00:22	36.75633301
20	02-02-2020	11:30:22	37.07615008
21	02-02-2020	12:00:22	36.39082778
22	02-02-2020	12:30:22	35.38568841
23	02-02-2020	01:00:22	36.75633301
24	02-02-2020	01:30:22	35.38568841
25	02-02-2020	02:00:22	36.61926855
26	02-02-2020	02:30:22	36.75633301
27	02-02-2020	03:00:22	32.18751769
28	02-02-2020	03:30:22	33.37540968
29	02-02-2020	04:00:22	33.55816229

	A	B	C
30	04-02-2020	08:30:52	34.01504382
31	04-02-2020	09:00:52	34.47192535
32	04-02-2020	09:30:52	35.38568841
33	04-02-2020	10:00:52	35.84256994
34	04-02-2020	10:30:52	36.29945147
35	04-02-2020	11:00:52	36.57358039
36	04-02-2020	11:30:52	36.29945147
37	04-02-2020	12:00:52	35.84256994
38	04-02-2020	12:30:52	36.29945147
39	04-02-2020	01:00:52	36.29945147
40	04-02-2020	01:30:52	36.29945147
41	04-02-2020	02:00:52	35.84256994
42	04-02-2020	02:30:52	36.29945147
43	04-02-2020	03:00:52	36.48220409
44	04-02-2020	03:30:52	36.75633301
45	04-02-2020	04:00:52	37.07615008
46	04-02-2020	04:30:52	36.39082778
47	05-02-2020	09:00:47	35.38568841
48	05-02-2020	09:30:47	36.75633301
49	05-02-2020	10:00:47	35.38568841
50	05-02-2020	10:30:47	35.84256994
51	05-02-2020	11:00:47	36.29945147
52	05-02-2020	11:30:47	35.38568841
53	05-02-2020	12:00:47	36.39082778
54	05-02-2020	12:30:47	36.48220409
55	05-02-2020	01:00:47	36.75633301
56	05-02-2020	01:30:47	36.93908562
57	05-02-2020	02:00:47	36.75633301
58	05-02-2020	02:30:47	35.84256994

	A	B	C
59	05-02-2020	03:00:47	34.01504382
60	05-02-2020	03:30:47	34.47192535
61	05-02-2020	04:00:47	35.38568841
62	06-02-2020	10:00:32	35.84256994
63	06-02-2020	10:30:32	36.29945147
64	06-02-2020	11:00:32	36.57358039
65	06-02-2020	11:30:32	36.29945147
66	06-02-2020	12:00:32	35.84256994
67	06-02-2020	12:30:32	36.29945147
68	06-02-2020	01:00:32	36.29945147
69	06-02-2020	01:30:32	36.29945147
70	06-02-2020	02:00:32	35.84256994
71	06-02-2020	02:30:32	36.29945147
72	06-02-2020	03:00:32	36.48220409
73	06-02-2020	03:30:32	36.75633301
74	06-02-2020	04:00:32	37.07615008
75	07-02-2020	08:30:52	36.39082778
76	07-02-2020	09:00:52	35.38568841
77	07-02-2020	09:30:52	36.75633301
78	07-02-2020	10:00:52	36.93908562
79	07-02-2020	10:30:52	36.75633301
80	07-02-2020	11:00:52	35.84256994
81	07-02-2020	11:30:52	36.6649567
82	07-02-2020	12:00:52	36.61926855
83	07-02-2020	12:30:52	36.75633301
84	07-02-2020	01:00:52	35.38568841
85	07-02-2020	01:30:52	35.84256994
86	07-02-2020	02:00:52	36.29945147
87	07-02-2020	02:30:52	36.48220409

Figure 11: Power generation of 40 W PV panel.

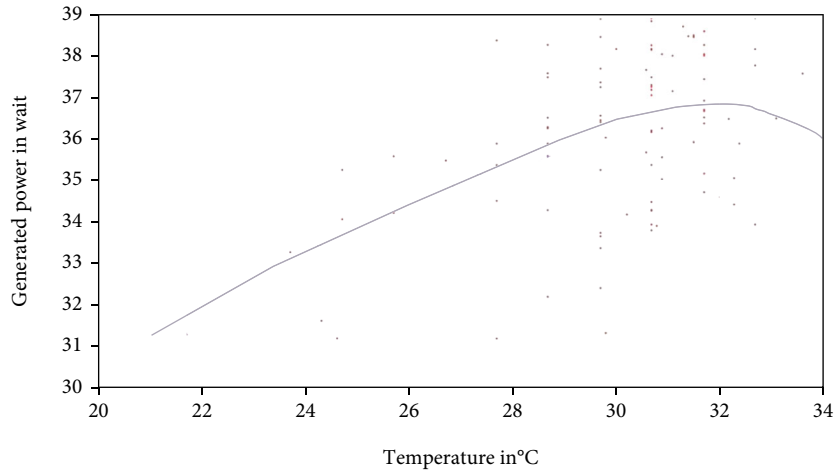


Figure 12: PR between power and temperature.

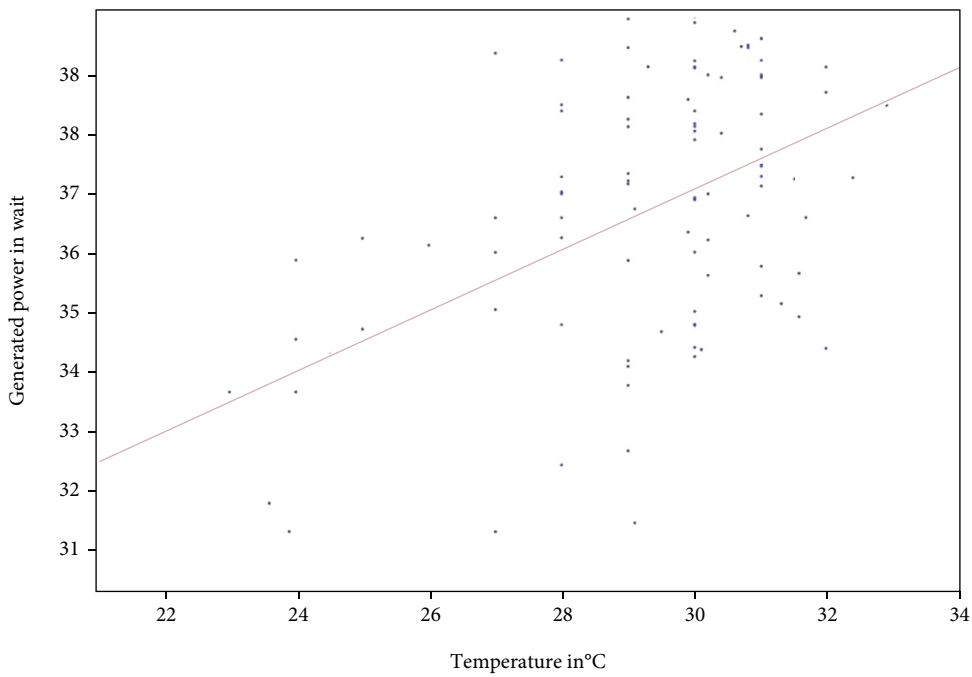


Figure 13: LR between power and temperature.

Table 6: Essential cloud services.

Cloud service	Visual-crossing	ThingSpeak
Description	Weather data acquisition platform	IoT cloud server to save measured data

every machine, be it small or huge, communicates with each other and makes life easier than it already is. The internet covers almost the entire world, and the world has both good and bad sides. On the bright side, we have IoT as an emerging technology that aims to play a crucial role in saving money, conserving energy, better monitoring, and power management, and now we are on the verge of the upcoming superfast fifth generation of internet, the 5G. On the other hand, it also raises questions regarding privacy, data breach, social, and technical challenges. Some researchers have found that making life this much easier may make humankind lazy and inactive.

3.4. Forecasting Based on CBR. The CBR is an artificial intelligence technique that solves new problems by comparing it to the old cases or by fetching solution to the new problems from its past experiences. This algorithm is predominantly applied when the available data is incomplete, erroneous, missing,

Table 7: Methods of analysing meteorological data.

Ref.	Paper title	Type of meteorological data	Method	Inference
[28]	Transformer failure diagnosis using fuzzy association rule mining combined with case based reasoning	—	Hybrid	CBR algorithm is employed to discover the past cases which are the most identical to the present case. The exactness of the model is confirmed using past 10 years data.
[69]	A hybrid algorithm for short-term solar power prediction - sunshine state case study	Hourly	Hybrid	Results obtained from the hybrid algorithm are more accurate with fast convergence compared to the classic algorithm.
[70]	A hybrid ensemble model for interval prediction of solar power output in ship onboard power systems	Hourly	Hybrid	The hybrid algorithm gives outcomes with high efficiency considering the meteorological data along with the ship's swinging as the input parameters.
[71]	A lightweight short-term photovoltaic power prediction for edge computing	Data sampled every 30 minutes	Hybrid	Compared to other standard ML algorithms, the technique employed here is remarkable and is capable of making short-term power predictions.
[72]	A local training strategy-based artificial neural network for predicting the power production of solar photovoltaic systems	Hourly	Intuitive	Various tests were conducted that showed the superiority of the proposed ANN over the benchmark ANN training strategies.
[73]	A practicable copula-based approach for power forecasting of small-scale photovoltaic systems	Daily	Numerical	From the results, it is clear that the mathematical model used here gives satisfactory prediction for cloudy days.
[74]	A solar time based analog ensemble method for regional solar power forecasting	Hourly	Hybrid	The proposed model adapts easily to the changing weather conditions irrespective of the location with high forecasting accuracy, few parameter requirements, data management, etc.,
[75]	Ensemble approach of optimized artificial neural networks for solar photovoltaic power prediction	Daily	Intuitive	Accurate day-ahead power prediction is obtained and is verified against a real case study. The number of hidden neurons in the hidden layer of ANN is optimized using trial and error method of the proposed model.
[76]	Photovoltaic power forecasting with a hybrid deep learning approach	Daily	Hybrid	The proposed hybrid method is compared with three other benchmark methods and is shown to have very small prediction errors.
[77]	Power generation forecast of hybrid PV-wind system	4 hours daily	Numerical	The duration for which data samples are incomplete or missing can be predicted by using the proposed method.
[78]	Prediction of photovoltaic power generation based on general regression and Back propagation neural network	Daily	Numerical	Temperature and irradiance were found to be the key parameters. Back propagation neural network predicted accurate results, but general regression technique was more appropriate for big data sets.
[79]	Probabilistic forecasting of photovoltaic generation: An efficient statistical approach	Daily	Probabilistic	The technique employed here exhibits very high computational efficiency and proves to be remarkably effective.
[80]	Real-time anomaly detection for very short-term load forecasting	Daily	Numerical	A way to detect and replace anomalies/corrupted data is proposed here whose performance surpasses state-of-the-art methods.
[81]	Day-ahead hierarchical probabilistic load forecasting with linear quantile regression and empirical copulas	Daily	Hybrid	A simple linear regression is adopted here for of accurate prediction improvement.
[82]	Direct quantile regression for nonparametric probabilistic forecasting of wind power generation	Hourly	Probabilistic	The proposed linear programming gives a simple solution with high computational efficiency and flexible framework.
[83]	Solar power probabilistic forecasting by using multiple linear regression analysis	Hourly	Numerical	The forecasting result was satisfactory using linear regression

Table 7: Continued.

Ref.	Paper title	Type of meteorological data	Method	Inference
[84]	IOT based online load forecasting using machine learning	Hourly	Numerical	Compared to all the ML algorithms, LR ML algorithm is found to be better as per the parameters is considered in this paper. Forecasting consumption of power for the next-hour is done using the online IoT platform.

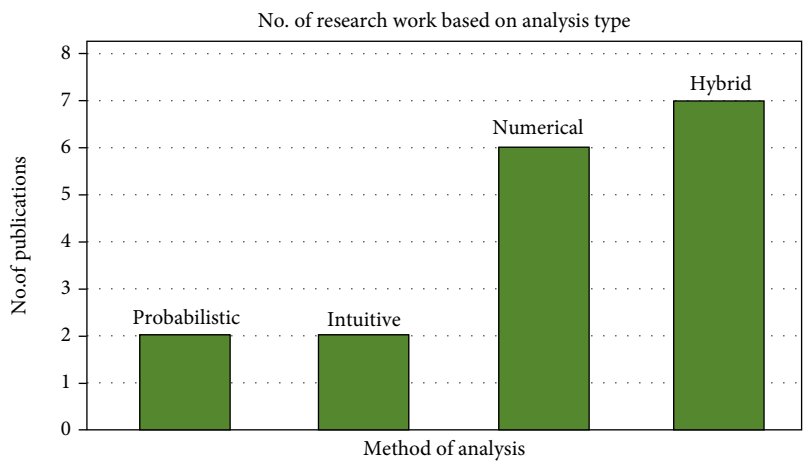


Figure 14: Number of research work based on method of analysis.

	A	B	C	D	E	F	G
1	Date	Time	Temperature	Humidity	Wattage	Predicted wattage with only temperature	Predicted wattage with humidity correction
2	02-11-2020	09:00:23	26.2	78.25	30.214	34.56330166	29.6471
3	02-11-2020	09:30:23	27	69.23	31.121	34.92880688	31.7688
4	02-11-2020	10:00:23	29	65.97	33.684	35.84256994	33.3173
5	02-11-2020	10:30:23	29.5	63.27	34.121	36.07101071	34.0714
6	02-11-2020	11:00:23	30	62.28	34.424	36.29945147	34.4926
7	02-11-2020	11:30:23	30.2	61.5	34.987	36.39082778	34.7359
8	02-11-2020	12:00:23	31	60.89	34.66	36.75633301	35.2202
9	02-11-2020	12:30:23	31.6	59.74	34.36	37.03046192	35.7182
10	02-11-2020	01:00:23	31	58.8	35.124	36.75633301	35.6271
11	02-11-2020	01:30:23	31.6	56.7	35.014	37.03046192	36.3101
12	02-11-2020	02:00:23	32.9	52.5	37.545	37.62440791	37.7218
13	02-11-2020	02:30:23	32.4	54.23	36.454	37.39596715	37.1565
14	02-11-2020	03:00:23	31.3	59.62	34.541	36.89339746	35.6045
15	02-11-2020	03:30:23	30	60.01	34.231	36.29945147	34.9346
16	02-11-2020	04:00:23	32	59.03	33.885	37.21321454	36.0392
17	02-11-2020	04:30:23	31	52.3	36.478	36.75633301	36.8926
18	02-12-2020	10:30:37	29	62.05	33.315	35.84256994	34.0805
19	02-12-2020	11:00:37	29	61.08	33.601	35.84256994	34.2694
20	02-12-2020	11:30:37	30	62.05	34.214	36.29945147	34.5374
21	02-12-2020	12:00:37	29.1	68.85	31.254	35.8882581	32.8023
22	02-12-2020	12:30:37	28	67.45	32.124	35.38568841	32.5723
23	02-12-2020	01:00:37	29	65.97	32.336	35.84256994	33.3173
24	02-12-2020	01:30:37	28	63.47	34.225	35.38568841	33.3472
25	02-12-2020	02:00:37	30	62.27	33.885	36.29945147	34.4946
26	02-12-2020	02:30:37	27	63.33	34.452	34.92880688	32.9176
27	02-12-2020	03:00:37	30.1	65.12	33.852	36.34513963	33.9854
28	02-12-2020	03:30:37	30	66.78	33.745	36.29945147	33.6165
29	02-12-2020	04:00:37	30	58.57	36.124	36.29945147	35.215

Figure 15: Comparison of measured and predicted results.

	A	B	C	D	E	F	G
1	Date	Time	Temperature	Humidity	Predicted wattage with only temperature	Predicted wattage with humidity correction	Wattage
2	02-19-2020	10:30:14	28	61.83	35.38568841	33.6665	34.258
3	02-19-2020	11:00:14	30	48.6	36.29945147	37.1561	36.936
4	02-19-2020	11:30:14	32	43.32	37.21321454	39.0979	38.885
5	02-19-2020	12:00:14	29.9	52.3	36.25376332	36.3901	36.996
6	02-19-2020	12:30:14	30	42.24	36.29945147	38.3944	37.958
7	02-19-2020	01:00:14	31	45	36.75633301	38.3139	37.963
8	02-19-2020	01:30:14	29	49.98	35.84256994	36.4306	36.52
9	02-19-2020	02:00:14	31	48.87	36.75633301	37.5604	36.332
10	02-19-2020	02:30:14	34	41.79	38.1269776	40.3096	39.101
11	02-19-2020	03:00:14	29.9	55.89	36.25376332	35.6911	35.639
12	02-19-2020	03:30:14	30	49.89	36.29945147	36.905	37.269
13	02-19-2020	04:00:14	30	51.5	36.29945147	36.5915	37.025
14	02-19-2020	04:30:14	28	68.88	35.38568841	32.2939	32.854

Figure 16: Testing predicted wattage with measured wattage.

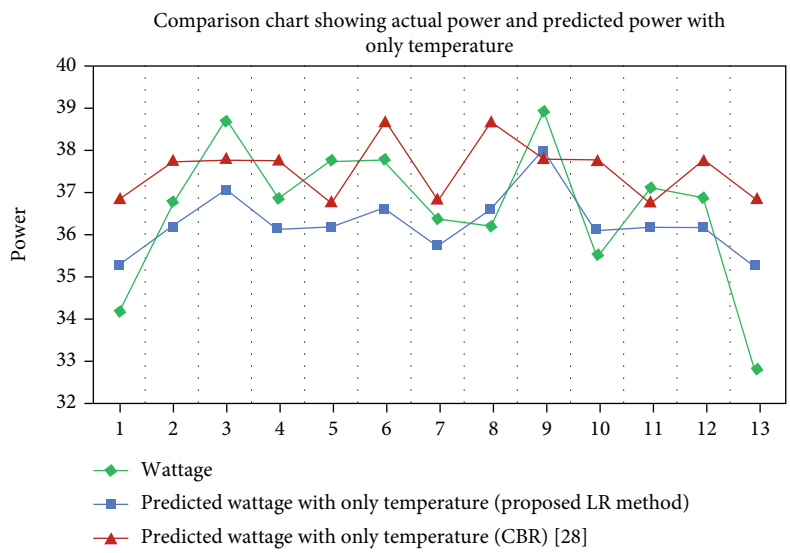


Figure 17: Comparison chart showing actual power and predicted power with only temperature.

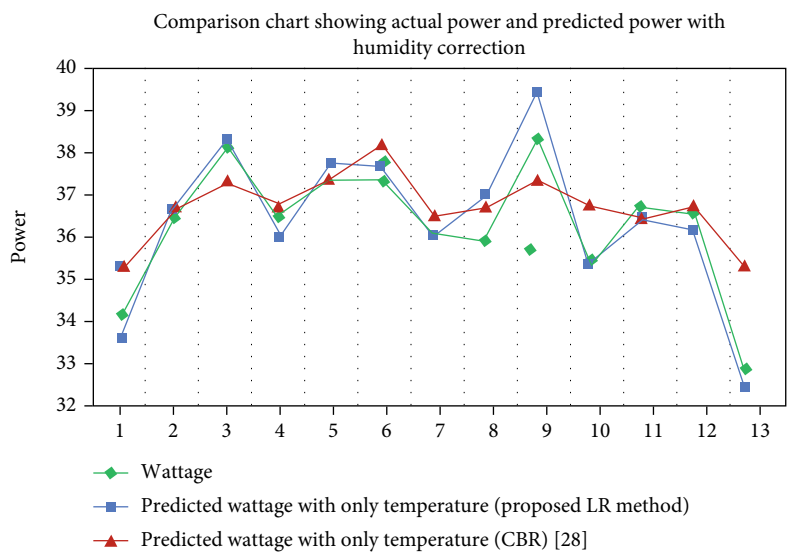


Figure 18: Comparison chart showing actual power and predicted power with humidity correction.

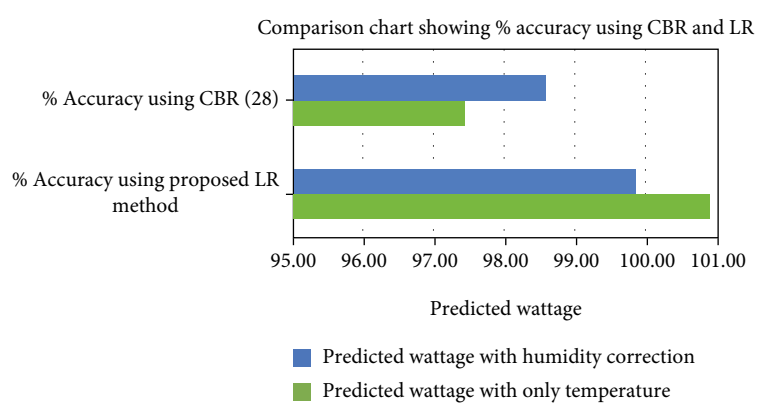


Figure 19: Comparison chart showing % accuracy obtained using CBR and LR.

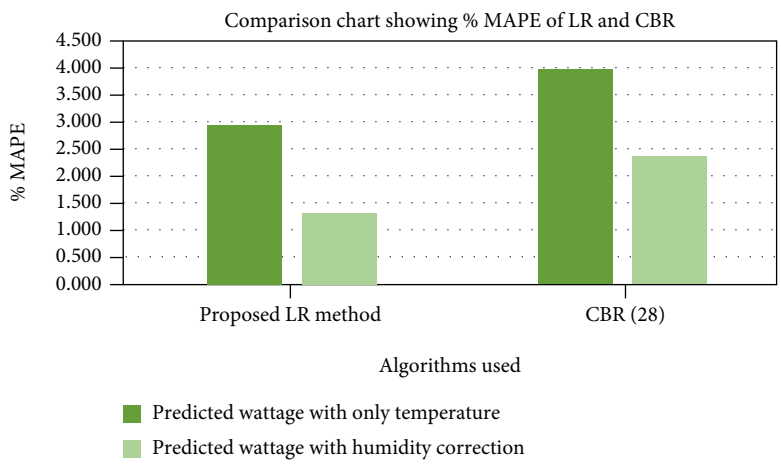


Figure 20: Comparison chart showing % MAPE.

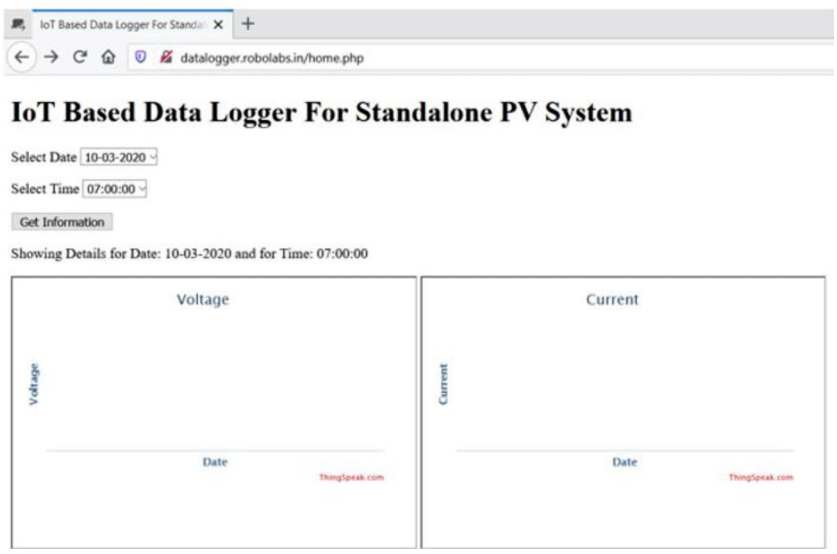


Figure 21: Homepage of developed website.

and corrupted, and the rule-based reasoning does not give satisfactory solution. The implementation of CBR algorithm is depicted in Figure 5.

3.5. Other Component Specifications. The specifications of various components and the essential cloud services used to develop the proposed work are tabulated. Table 4 shows the specifications of Arduino Uno. The ratings of the PV panel used are given in Table 5, and Table 6 lists the essential cloud services used.

4. Experimental Results and Discussion

The proposed model of forecasting of generated power is built and verified under different scenarios to assess its superiority and performance. The assessment and obtained results of the proposed model is discussed in the following subsections.

4.1. Datalogger Realization. The voltage and current sensor are connected to the pins of Arduino Uno which is programmed using an Arduino IDE. ESP8266 Wi-Fi modem is paired with the Arduino Uno to enable wireless capability on the Arduino Uno development board. ESP8266 is flashed with custom firmware to be able to receive and transmit data over the internet. The sensor is manually calibrated and tested with a high precision multimeter. The sensors collect data in real-time and transfer it to the Arduino Uno development board using its analog pins. Upon receiving raw data from the sensors, Arduino Uno is programmed to calculate the performance parameters. This data is sent to be stored in the cloud via the Wi-Fi module.

The voltage and current sensors are connected to digital I/O pins of Arduino, whereas ESP8266 Wi-Fi modem is connected to digital pin 2 and 3 which are acting as TX and RX as shown in Figure 6. The reason behind not using the default TX and RX pin is that Arduino board resets itself each time it starts an operation, and while resetting, there should not be any connection on the default TX and RX pin otherwise it will come up with an error code and will not initialize. The Wi-Fi modem has an inbuilt antenna and it works as either a Wi-Fi hotspot or Wi-Fi receiver with

Figure 22: Website displaying estimated power.

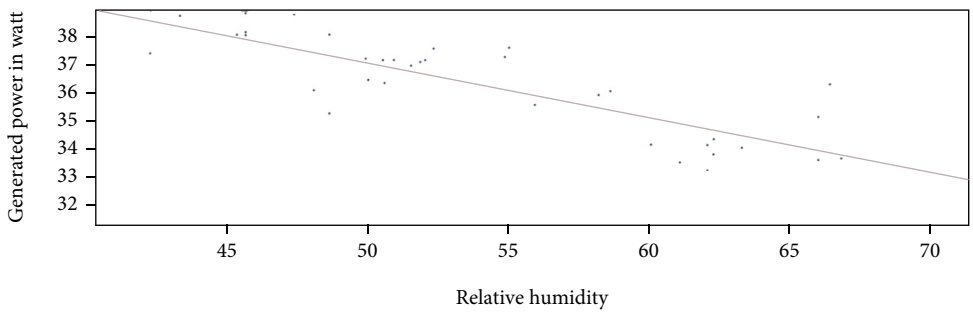


Figure 23: LR between generated power and relative humidity.

inbuilt microprocessor and memory. It does not come with any firmware installed in its memory. The users have to flash the latest firmware by themselves which is provided by the manufacturer. In this project, the latest Non-Os AT firmware version, 1.6.2 is being used. The ESP8266 modem serves as a communicator between the datalogger and IoT server.

IoT capability is incorporated by the ThingSpeak IoT server. It is a free and open-source platform to store and retrieve data. It is one-of-a-kind, easy, and simple authentication system that makes it the most popular IoT server for research and project purposes. ThingSpeak provides reading and writing capabilities to its channels by giving a unique API key for each channel.

An IoT-based datalogger for a standalone PV panel is developed with generation estimation capability. A hardware prototype is realized for a scaled-down version of the datalogger circuit. IoT is incorporated by Arduino Uno and ESP8266 Wi-Fi module to collect the performance parameter of solar PV panels such as current and voltage and to store and visualize these on the ThingSpeak server. The temperature and relative humidity are obtained by Visual-Crossing weather API. Visual-Crossing is a free weather forecasting service. It provides hourly or daily based weather data for any given date and location. First up, a datalogger circuit was assembled to monitor the data of the solar panel. The developed datalogger circuit was tested with a solar PV emulator as shown in Figure 7.

The solar PV emulator emulates power generation through the photovoltaic panels in a closed box. A bright electric bulb emulates solar irradiance, and two small sets of PV arrays work as a PV panel. These PV panels are connected to an inbuilt measurement instrument that displays voltage and current reading in an LCD mounted in front of the emulator system as displayed in Figure 8. The solar PV emulator has two knobs through which temperature or voltage and current generation can be adjusted. It has a switch that switches between external or internal measurement instruments. It has a power outlet pin to attach external measurement monitoring instruments. The datalogger system is attached and tested, and sensors were calibrated.

4.2. Data Collection. The output of the datalogger is monitored in the serial monitor window of Arduino IDE. A screenshot of the same is attached in Figure 9. The PV panel performance data for the next ten days is collected from a 40 W PV panel. Interfacing the datalogger to the PV panel is demonstrated in Figure 10. The power generated is measured and is shown in Figure 11 along with its respective date and time stamp. The monitored data is analyzed and stored on a local computer.

4.3. Forecasting Based on Linear and Polynomial Regression Algorithms. In this section, linear and polynomial regression algorithms are employed to forecast the missing/erroneous values of voltage and current obtained from the datalogger.

4.3.1. Polynomial Regression Algorithm. The data is used as a training set to train the ML algorithm namely PR. The output is a curve. It is considered to be better than LR if the variables in the data set have little to no linear relations between them. The degree of the polynomial equation can also be adjusted for best fitting. Figure 12 shows the PR between power and temperature. The red dots are the measured points and the blue curve is its linearity.

4.3.2. Linear Regression Algorithm. The captured data as in Figure 11 is also used to train another ML algorithm such as LR, and its output is a straight line. Figure 13 shows the LR between power and temperature. The blue dots are the measured points, and the red line is its linearity. Table 7 summarizes the methods of analysing meteorological data for power prediction. From the research papers studied, we find that regression method is simpler to implement. There are better algorithms available for power forecasting, but due to limited data availability and usage of few parametric variables, the regression technique is more suitable. Hence, the proposed algorithm for power forecasting in SAPV systems for the hourly data accessible is LR. Figure 14 shows the number of research articles deliberate based on method of analysis.

4.4. Comparison Summary of Linear and Polynomial Regression Algorithms. The results of both the algorithms are compared. PR does the same thing as LR but by fitting a polynomial equation to the given data sets. Hence, the result is a curve instead of a straight line. While doing internal testing, the results of PR, Figure 12, were not satisfactory when compared to LR as displayed in Figure 13. It is evident that LR fits more accurately than PR. Also, the prediction error and error percentage are noticeably high with the latter. And going higher than 6 degrees of PR resulted in even more inferior curve fitting. Also, going lower than 6 degrees was not satisfactory for the considered system. Hence, LR is chosen as the ML algorithm for the analysis of data from the datalogger of the SAPV network.

4.5. Data Analysis. The ML algorithm namely LR fetched the weather data (temperature and relative humidity) from an API namely Visual-Crossing. Using the fetched data, the ML algorithm is able to estimate the power generation of the PV panel for the next 30 days. Figures 15 and 16 show the comparison between measured and estimated power generation. As seen in Figure 15, column E contains the measured power of the PV panel and column F and G are the estimated values. Column F lists estimated power generation when only the temperature is being considered in the algorithm whereas column G lists estimated power generation when both temperature and relative humidity are being used for estimation. From Figure 15, it is evident that the obtained wattage with only temperature has an accuracy of 94%, and similarly, the obtained wattage with both temperature and humidity correction is 98.955%. The estimated power is found to be more accurate while taking both the parameters, temperature, and relative humidity into account as in Figure 16.

The data recorded in Figure 16 is assessed, and the corresponding results are depicted. Figure 17 compares the wattage with the predicted wattage obtained using both LR and CBR algorithms when only temperature is considered. Figure 18 compares the wattage with the predicted wattage obtained using both LR and CBR algorithms when humidity correction is considered. MAPE is the measure of the forecasting system accuracy. The lower the error, the better will be the prediction. Figures 19 and 20 illustrate the % accuracy and % MAPE attained using CBR and LR algorithms, hence proving that LR is a better power prediction option than CBR. Figure 19 shows that there is a 3.51% increase in the percentage accuracy of predicting wattage with only temperature using the proposed LR method than the CBR method. Also, there is a 1.24% increase in the prediction of wattage with humidity correction between the two methods. Figure 20 depicts the comparison of percentage MAPE between LR and CBR methods. The % MAPE is 2.958 and 3.978 for LR and CBR methods, respectively, for predicted wattage with only temperature. For the prediction of wattage with humidity correction, the % MAPE is 1.321 and 2.374 for LR and CBR methods, respectively. On comparing both the set of values, it is obvious that LR method has lower % MAPE than CBR method.

The measured performance data of the PV panel is sent to the ThingSpeak IoT server by the microcontroller. ThingSpeak provides an easy and hassle-free setup to import data from various IoT-enabled sensors. The developed webpage is shown in Figure 21 containing a graphical representation of current and voltage in real-time. It also has a drop-down menu bar to select date and time, and clicking on Get Information navigates the page to the next page where it shows the estimated power generated for the entered date and time as seen in Figure 22. ThingSpeak then plots a graphical visualization of received data concerning time for each set of parameters individually. These graphs are updated in real-time and can be used to attach anywhere on the web through an embedded link. The website also has a back-end connection to the MATLAB server for the estimation of power generation. MATLAB server is programmed in such a way that upon entering date and time, it obtains weather information such as relative humidity and temperature from Visual-Crossing API for the entered date and time and substitutes these two in the LR algorithm. The algorithm calculates estimated power generation and displays it on the website as can be seen in Figure 22. Figure 23 shows the linear relation between generated power and relative humidity. The blue dots are the measured points and the red line demonstrates their linearity.

5. Conclusion

In this paper, a new low-cost portable datalogger to monitor PV systems is designed, tested, and analyzed. The system design attributed easy-to-obtain hardware and free software, making it available to any investigator or user for the progress of schemes of their design and usage. This adjustability and adaptability makes the system more appropriate for intended applications like monitoring of PV plants and the collection of data from PV panels at isolated sites in developing nations.

LR machine learning algorithm is successfully used. The data from the datalogger serves as the training data set for the algorithm. The forecasted power generation is the output of the machine learning algorithm, which is tested with real-time data and found out to be accurate with less than 10% error. To store the measured data of voltage and current, sensor two channels of Thing Speak are used. Once date

and time are keyed in, forecasted power generation for the entered date and time is computed. The website also has a back-end connection to the MATLAB server for the estimation of power generation.

The price of the proposed system is significantly less than commercially available devices, with little loss of accuracy and precision. Essential data for energy management systems are provided by the power generation estimation system. For use by researchers and in developing nations, this datalogger permits additional study. The subsequent phases include achieving lower power consumption, develop a small SAPV system realized with Arduino, using wireless technologies for communication aspects, scrutinizing the pricing meticulously. The proposed work has a very wide scope and can be conveniently used for standalone PV plants which are usually installed in a remote location where human intervention is not possible due to harsh weather or other circumstances.

Abbreviations

A: Ampere
AC: Alternating current
Amp: Current at maximum power
API: Application programming interface
AT: Attention
CBR: Case based reasoning
dBm: Decibel-milliwatts
DC: Direct current
DIY: Do it yourself
GB: Giga byte
GHz: Giga hertz
GSM: Global system for mobile
I/O: Input and output
I^2C: Inter-integrated circuit
IC: Integrated circuit
IDE: Integrated development environment
IEC: International electrotechnical commission
IoT: Internet of things
kHz: Kilo hertz
LCD: Liquid crystal display
LR: Linear regression
mΩ: Milli ohm
mA: Milli ampere
MAPE: Mean absolute percentage error
MB: Mega byte
MIMO: Multiple-in multiple-out
ML: Machine learning
MPP: Maximum power point
MPPT: Maximum power point tracking
ms: Milli second
mV/A: Milli volt per ampere
OS: Operating system
OTA: Over the air
PC: Personal computer
PR: Polynomial regression
PV: Photovoltaic
PWM: Pulse width modulation
RAM: Random access memory
RX: Receive
SAPV: Stand-alone photovoltaic
SD: Secure digital
SPI: Serial peripheral interface
STBC: Space-time block coding
TA: Ambient temperature
THD: Total harmonic distortion
TR: Transmit/receive
TX: Transmit
UART: Universal asynchronous receiver-transmitter
USB: Universal serial bus
V: Voltage
Vmp: Voltage at maximum power
W: Watt
Wi-Fi: Wireless fidelity.

Data Availability

The data used to support the findings of this study are included in the article.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

References

[1] L. Tuggener, M. Amirian, K. Rombach et al., "Automated machine learning in practice: state of the art and recent results," in *2019 6th Swiss Conference on Data Science (SDS)*, pp. 31–36, Bern, Switzerland, 2019.

[2] Y. Roh, G. Heo, and S. E. Whang, "A survey on data collection for machine learning: a big data - AI integration perspective," *IEEE Transactions on Knowledge and Data Engineering*, vol. 33, p. 1, 2021.

[3] K. Chauhan, S. Jani, D. Thakkar et al., "Automated machine learning: the new wave of machine learning," in *2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)*, pp. 205–212, Bangalore, India, 2020.

[4] G. Graditi, S. Ferlito, and G. Adinolfi, "Comparison of photovoltaic plant power production prediction methods using a large measured dataset," *Renewable Energy*, vol. 90, pp. 513–519, 2016.

[5] M. Al Karim, J. Currie, and T. Lie, "Dynamic event detection using a distributed feature selection based machine learning approach in a self-healing microgrid," *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 4706–4718, 2018.

[6] N. Krishna Prakash and D. Prasanna Vadana, "Machine learning based residential energy management system," in *2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, pp. 1–4, Coimbatore, India, 2017.

[7] C. Voyant, G. Notton, S. Kalogirou et al., "Machine learning methods for solar radiation forecasting: a review," *Renewable Energy*, vol. 105, pp. 569–582, 2017.

[8] T. Sogabe, H. Ichikawa, K. Sakamoto et al., "Optimization of decentralized renewable energy system by weather forecasting and deep machine learning techniques," in *2016 IEEE Innovative Smart Grid Technologies - Asia (ISGT-Asia)*, pp. 1014–1018, Melbourne, VIC, Australia, 2016.

[9] H. Eom, Y. Son, and S. Choi, "Feature-Selective ensemble learning-based long-term regional PV generation forecasting," *IEEE Access*, vol. 8, pp. 54620–54630, 2020.

[10] K. Benhmed, F. Touati, M. Al-Hitmi et al., "PV power prediction in Qatar based on machine learning approach," in *2018 6th International Renewable and Sustainable Energy Conference (IRSEC)*, pp. 1–4, Rabat, Morocco, 2018.

[11] M. Diagne, M. David, P. Lauret, J. Boland, and N. Schmutz, "Review of solar irradiance forecasting methods and a proposition for small- scale insular grids," *Renewable and Sustainable Energy Reviews*, vol. 27, pp. 65–76, 2013.

[12] A. López-Vargas, M. Fuentes, M. V. García, and F. J. Muñoz-Rodríguez, "Low-Cost datalogger intended for remote monitoring of solar photovoltaic standalone systems based on Arduino™," *IEEE Sensors Journal*, vol. 19, no. 11, pp. 4308–4320, 2019.

[13] A. López-Vargas, M. Fuentes, and M. Vivar, "Evaluation of long-term performance of a solar home system (SHS) monitoring system on harsh environments," *Sensors*, vol. 19, no. 24, p. 5462, 2019.

[14] N. Mahzan, A. Omar, L. Rimon, S. M. Noor, and M. Z. Rosselan, "Design and Development of an Arduino based data logger for photovoltaic monitoring system," *International Journal of Simulation: Systems, Science and Technology*, vol. 17, pp. 15.1–15.5, 2017.

[15] M. Fuentes, M. Vivar, J. M. Burgos, J. Aguilera, and J. A. Vacas, "Design of an accurate, low-cost autonomous data logger for PV system monitoring using Arduino™ that complies with IEC standards," *Solar Energy Materials and Solar Cells*, vol. 130, pp. 529–543, 2014.

[16] A. Mahjoubi, R. Mechlouch, and A. Brahim, "A low cost wireless data acquisition system for a remote photovoltaic (PV) water pumping system," *Energies*, vol. 4, no. 1, pp. 68–89, 2011.

[17] T. A. Jesha and M. T. Iqbal, "Data logging and energy consumption analysis of two houses in St. John's, Newfoundland," in *8th International Conference on Electrical and Computer Engineering*, pp. 816–819, Dhaka, Bangladesh, 2014.

[18] B. Nkom and H. Musa, "Development of a novel microcontroller-based data logger," in *2009 2nd International Conference on Adaptive Science & Technology (ICAST)*, pp. 314–324, Accra, Ghana, 2009.

[19] A. Suzdalenko, "Guidelines for autonomous data logger design," in *2011 IEEE International Symposium on Industrial Electronics*, pp. 1426–1429, Gdansk, Poland, 2011.

[20] M. Engin, "Open source embedded data logger design for PV system monitoring," in *2017 6th Mediterranean Conference on Embedded Computing (MECO)*, pp. 1–5, Bar, Montenegro, 2017.

[21] M. Benghanem, "Measurement of meteorological data based on wireless data acquisition system monitoring," *Applied Energy*, vol. 86, no. 12, pp. 2651–2660, 2009.

[22] G. Nhivekar and R. Mudholkar, "Data logger and remote monitoring system for multiple parameter measurement applications," *Journal of Electrical and Electronics Engineering*, vol. 4, 2011.

[23] D. Munandar and D. Syamsi, "Data logger management software design for maintenance and utility in remote," in *2014 The 1st International Conference on Information Technology, Computer, and Electrical Engineering*, pp. 74–78, Semarang, Indonesia, 2014.

[24] S. Badhiye, B. Chatur, and B. V. Wakode, "Data Logger System: A Survey," *International Journal of Computer Technology and Electronics Engineering (IJCTEE)*, pp. 24–26, 2011.

[25] M. Kovac, "Efficient data compression algorithm for data loggers, measurement equipment and remote data analysis applications," in *Proceedings IEEE Southeastcon '98 'Engineering for a New Era'*, pp. 47-48, Orlando, FL, USA, 1998.

[26] S. Chaudhary, P. Ghimire, P. B. Thogersen, and P. de Place Rimmen, "Field data logger prototype for power converters," in *2014 IEEE International Conference on Power Electronics, Drives and Energy Systems (PEDES)*, pp. 1–4, Mumbai, India, 2014.

[27] C. M. Vancea and L. Viman, "Wireless data logger for thermal validation systems," *2011 IEEE 17th International Symposium for Design and Technology in Electronic Packaging (SIITME)*, 2011, pp. 295–298, Timisoara, Romania, 2011.

[28] C. Guo, B. Wang, Z. Wu et al., "Transformer failure diagnosis using fuzzy association rule mining combined with case-based reasoning," *IET Generation, Transmission and Distribution*, vol. 14, no. 11, pp. 2202–2208, 2020.

[29] R. Uthirasamy, V. K. Chinnaiyan, S. Vishnukumar et al., "Design of boosted multilevel DC-DC converter for solar photovoltaic system," *International Journal of Photoenergy*, vol. 2022, Article ID 1648474, 23 pages, 2022.

[30] S. Kaliappan, R. Saravanakumar, A. Karthick et al., "Hourly and day ahead power prediction of building integrated semi-transparent photovoltaic system," *International Journal of Photoenergy*, vol. 2021, Article ID 7894849, 8 pages, 2021.

[31] A. K. Podder, A. K. Das, E. Hossain et al., "Integrated modeling and feasibility analysis of a rooftop photovoltaic systems for an academic building in Bangladesh," *International Journal of Low Carbon Technologies*, vol. 16, no. 4, pp. 1317–1327, 2021.

[32] U. Subramaniam, M. M. Subashini, D. Almakhles, A. Karthick, and S. Manoharan, "An expert system for COVID-19 infection tracking in lungs using image processing and deep learning techniques," *BioMed Research International*, vol. 2021, Article ID 1896762, 17 pages, 2021.

[33] S. S. Ganesh, G. Kannayeram, A. Karthick, and M. Muhibbullah, "A novel context aware joint segmentation and classification framework for glaucoma detection," *Computational and Mathematical Methods in Medicine*, vol. 2021, Article ID 2921737, 19 pages, 2021.

[34] A. Hmidet, U. Subramaniam, R. M. Elavarasan et al., "Design of efficient off-grid solar photovoltaic water pumping system based on improved fractional open circuit voltage MPPT technique," *International Journal of Photoenergy*, vol. 2021, Article ID 4925433, 18 pages, 2021.

[35] P. M. Kumar, R. Saravanakumar, A. Karthick, and V. Mohanavel, "Artificial neural network-based output power prediction of grid-connected semitransparent photovoltaic system," *Environmental Science and Pollution Research*, vol. 29, pp. 10173–10182, 2022.

[36] D. Singh, R. Chaudhary, and A. Karthick, "Review on the progress of building-applied/integrated photovoltaic system," *Environmental Science and Pollution Research*, vol. 28, no. 35, pp. 47689–47724, 2021.

[37] N. M. Kumar, M. Samykano, and A. Karthick, "Energy loss analysis of a large scale BIPV system for university buildings in tropical weather conditions: a partial and cumulative performance ratio approach," *Case Studies in Thermal Engineering*, vol. 25, p. 100916, 2021.

[38] A. R. Prasad, R. Shankar, C. K. Patil, A. Karthick, A. Kumar, and R. Rahim, "Performance enhancement of solar photovoltaic system for roof top garden," *Environmental Science and Pollution Research*, vol. 28, no. 36, pp. 50017–50027, 2021.

[39] V. Chandran, M. G. Sumithra, A. Karthick et al., "Diagnosis of cervical cancer based on ensemble deep learning network using colposcopy images," *BioMed Research International*, vol. 2021, Article ID 5584004, 15 pages, 2021.

[40] R. Kabilan, V. Chandran, J. Yogapriya et al., "Short-term power prediction of building integrated photovoltaic (BIPV) system based on machine learning algorithms," *International Journal of Photoenergy*, vol. 2021, Article ID 5582418, 11 pages, 2021.

[41] N. Y. Jayalakshmi, R. Shankar, U. Subramaniam et al., "Novel multi-time scale deep learning algorithm for solar irradiance forecasting," *Energies*, vol. 14, no. 9, p. 2404, 2021.

[42] S. Obukhov, A. Ibrahim, A. A. Z. Diab, A. S. Al-Sumaiti, and R. Aboelsaud, "Optimal performance of dynamic particle swarm optimization based maximum power trackers for stand-alone PV system under partial shading conditions," *IEEE Access*, vol. 8, pp. 20770–20785, 2020.

[43] M. Killi and S. Samanta, "Voltage-sensor-based MPPT for stand-alone PV systems through voltage reference control," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 7, no. 2, pp. 1399–1407, 2019.

[44] A. Al-Shetwi, M. Sujod, A. Tarabsheh, and I. Altawail, "Design and economic evaluation of electrification of small villages in rural area in Yemen using stand-alone PV system," *International Journal of Renewable Energy Research (IJRER)*, vol. 6, pp. 289–298, 2016.

[45] T. Kulworawanichpong and J. Mwambeleko, "Design and costing of a stand-alone solar photovoltaic system for a Tanzanian rural household," *Sustainable Energy Technologies and Assessments*, vol. 12, pp. 53–59, 2015.

[46] F. J. Muñoz, G. Almonacid, G. Nofuentes, and F. Almonacid, "A new method based on charge parameters to analyse the performance of stand- alone photovoltaic systems," *Solar Energy Materials and Solar Cells*, vol. 90, no. 12, pp. 1750–1763, 2006.

[47] S. Parsekar and K. Chatterjee, "A novel strategy for battery placement in standalone solar photovoltaic converter system," in *2014 IEEE 40th Photovoltaic Specialist Conference (PVSC)*, pp. 2751–2756, Denver, CO, USA, 2014.

[48] S. Babaa, M. Armstrong, and V. Pickert, "High efficiency standalone photovoltaic system using adaptive switching of an interleaved boost converter," in *6th IET International Conference on Power Electronics, Machines and Drives (PEMD 2012)*, pp. 1–7, Bristol, UK, 2012.

[49] M. Nassereddine, J. Rizk, M. Nagrial, and A. Hellany, "Battery sustainable PV solar house: storage consideration for off grid," in *2018 Third International Conference on Electrical and Biomedical Engineering, Clean Energy and Green Computing (EBECEGC)*, pp. 34–38, Beirut, Lebanon, 2018.

[50] W. Ali, H. Farooq, A. U. Rehman, Q. Awais, M. Jamil, and A. Noman, "Design considerations of stand-alone solar photovoltaic systems," in *2018 International Conference on Computing, Electronic and Electrical Engineering (ICE Cube)*, pp. 1–6, Quetta, Pakistan, 2018.

[51] K. Sehil and M. Darwish, "Effective power management in a stand-alone PV system," in *2018 53rd International Universities Power Engineering Conference (UPEC)*, pp. 1–5, Glasgow, UK, 2018.

[52] C. Egbon, A. Oyekola, and T. Lie, "Design of stand alone photovoltaic system in developing countries: a case study of Kano, Nigeria," in *2018 Australasian Universities Power Engineering Conference (AUPEC*, pp. 1–6, Auckland, New Zealand, 2018.

[53] S. R. Spea and H. A. Khattab, "Design sizing and performance analysis of stand-alone PV system using PVSyst software for a location in Egypt," in *2019 21st International Middle East Power Systems Conference (MEPCON)*, pp. 927–932, Cairo, Egypt, 2019.

[54] G. Yang and M. Chen, "Methodology for precisely estimating the generation of standalone photovoltaic system," in *2009 International Conference on Environmental Science and Information Application Technology*, vol. 1, pp. 48–51, Wuhan, China, 2009.

[55] L. Fara and D. Craciunescu, "Output analysis of stand-alone PV systems: modeling, simulation and control," *Energy Procedia*, vol. 112, pp. 595–605, 2017.

[56] O. V. G. Swathika and K. T. M. U. Hemapala, "IOT based energy management system for standalone PV systems," *Journal of Electrical Engineering and Technology*, vol. 14, no. 5, pp. 1811–1821, 2019.

[57] M. S. Shaik, D. Shah, R. Chetty, and R. R. Marathe, "A LoRaWAN based open source IOT solution for monitoring rural electrification policy," in *2020 International Conference on COMmunication Systems & NETworkS (COMSNETS)*, pp. 888–890, Bengaluru, India, 2020.

[58] A. López-Vargas, M. Fuentes, and M. Vivar, "IoT application for real-time monitoring of solar home systems based on Arduino™ with 3G connectivity," *IEEE Sensors Journal*, vol. 19, no. 2, pp. 679–691, 2019.

[59] J. R. K. K. Dabbakuti, A. Jacob, V. R. Veeravalli, and R. K. Kallakunta, "Implementation of IoT analytics ionospheric forecasting system based on machine learning and ThingSpeak," *IET Radar, Sonar and Navigation*, vol. 14, no. 2, pp. 341–347, 2020.

[60] S. Kim and H. Kim, "A new metric of absolute percentage error for intermittent demand forecasts," *International Journal of Forecasting*, vol. 32, no. 3, pp. 669–679, 2016.

[61] S. S. Arefin, "Optimization techniques of islanded hybrid microgrid system," in *Renewable Energy - Resources, Challenges and Applications*, p. Ch. 23, IntechOpen, 2020.

[62] S. K. S. Arefin, "Optimization of islanded microgrid considering system stability and reliability according to the dispatch strategies," 2021, https://researchrepository.rmit.edu.au/discovery/delivery/61RMIT_INST:RMITU/12272691400001341.

[63] S. K. A. Shezan and C. Y. Lai, "Optimization of hybrid wind-diesel-battery energy system for remote areas of Malaysia," in *2017 Australasian Universities Power Engineering Conference (AUPEC)*, pp. 1–6, Melbourne, VIC, Australia, 2017.

[64] M. F. Ishraque, S. A. Shezan, M. M. Rashid et al., "Techno-economic and power system optimization of a renewable rich islanded microgrid considering different dispatch strategies," *IEEE Access*, vol. 9, pp. 77325–77340, 2021.

[65] M. Fatin Ishraque, S. A. Shezan, M. M. Ali, and M. M. Rashid, "Optimization of load dispatch strategies for an islanded microgrid connected with renewable energy sources," *Applied Energy*, vol. 292, p. 116879, 2021.

[66] S. K. A. Shezan, "Feasibility analysis of an islanded hybrid wind-diesel-battery microgrid with voltage and power

response for offshore Islands," *Journal of Cleaner Production*, vol. 288, p. 125568, 2021.

[67] S. A. Shezan, K. N. Hasan, A. Rahman, M. Datta, and U. Datta, "Selection of appropriate dispatch strategies for effective planning and operation of a microgrid," *Energies*, vol. 14, no. 21, p. 7217, 2021.

[68] S. A. Shezan, A. S. S. Rawdah, and Z. Rahman, "Design and implementation of an islanded hybrid microgrid system for a large resort center forPenang Islandwith the proper application of excess energy," *Environmental Progress & Sustainable Energy*, vol. 40, no. 4, pp. 1–15, 2021.

[69] A. Asrari, T. X. Wu, and B. Ramos, "A hybrid algorithm for short-term solar power prediction—sunshine state case study," *IEEE Transactions on Sustainable Energy*, vol. 8, no. 2, pp. 582–591, 2017.

[70] S. Wen, C. Zhang, H. Lan, Y. Xu, Y. Tang, and Y. Huang, "A hybrid ensemble model for interval prediction of solar power output in ship onboard power systems," *IEEE Transactions on Sustainable Energy*, vol. 12, no. 1, pp. 14–24, 2021.

[71] X. Chang, W. Li, and A. Y. Zomaya, "A lightweight short-term photovoltaic power prediction for edge computing," *IEEE Transactions on Green Communications and Networking*, vol. 4, no. 4, pp. 946–955, 2020.

[72] S. Al-Dahidi, M. Louzazni, and N. Omran, "A local training strategy-based artificial neural network for predicting the power production of solar photovoltaic systems," *IEEE Access*, vol. 8, pp. 150262–150281, 2020.

[73] S. Rajabalizadeh and S. M. M. Tafreshi, "A practicable copula-based approach for power forecasting of small-scale photovoltaic systems," *IEEE Systems Journal*, vol. 14, no. 4, pp. 4911–4918, 2020.

[74] X. Zhang, Y. Li, S. Lu, H. F. Hamann, B. Hodge, and B. Lehman, "A solar time based analog ensemble method for regional solar power forecasting," *IEEE Transactions on Sustainable Energy*, vol. 10, no. 1, pp. 268–279, 2019.

[75] S. Al-Dahidi, O. Ayadi, M. Alrbai, and J. Adeeb, "Ensemble approach of optimized artificial neural networks for solar photovoltaic power prediction," *IEEE Access*, vol. 7, pp. 81741–81758, 2019.

[76] 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.

[77] M. J. Sanjari, H. B. Gooi, and N. C. Nair, "Power generation forecast of hybrid PV–wind system," *IEEE Transactions on Sustainable Energy*, vol. 11, no. 2, pp. 703–712, 2020.

[78] J. Zhong, L. Liu, Q. Sun, and X. Wang, "Prediction of photovoltaic power generation based on general regression and back propagation neural network," *Energy Procedia*, vol. 152, pp. 1224–1229, 2018.

[79] C. Wan, J. Lin, Y. Song, Z. Xu, and G. Yang, "Probabilistic forecasting of photovoltaic generation: an efficient statistical approach," *IEEE Transactions on Power Systems*, vol. 32, no. 3, pp. 2471-2472, 2017.

[80] J. Luo, T. Hong, and M. Yue, "Real-time anomaly detection for very short-term load forecasting," *Journal of Modern Power Systems and Clean Energy*, vol. 6, no. 2, pp. 235–243, 2018.

[81] T. Zhao, J. Wang, and Y. Zhang, "Day-ahead hierarchical probabilistic load forecasting with linear quantile regression and empirical copulas," *IEEE Access*, vol. 7, pp. 80969–80979, 2019.

[82] C. Wan, J. Lin, J. Wang, Y. Song, and Z. Y. Dong, "Direct quantile regression for nonparametric probabilistic forecasting of wind power generation," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 2767–2778, 2017.

[83] M. Abuella and B. Chowdhury, "Solar power probabilistic forecasting by using multiple linear regression analysis," in *SoutheastCon 2015*, pp. 1–5, Lauderdale, FL, USA, 2015.

[84] M. P. Raju and A. J. Laxmi, "IOT based online load forecasting using machine learning algorithms," *Procedia Computer Science*, vol. 171, pp. 551–560, 2020.

Hindawi
International Journal of Photoenergy
Volume 2022, Article ID 7015818, 11 pages
https://doi.org/10.1155/2022/7015818

Research Article

Hourly Forecasting of Solar Photovoltaic Power in Pakistan Using Recurrent Neural Networks

Sohrab Khan,[1] **Faheemullah Shaikh**,[1] **Mokhi Maan Siddiqui**,[1] **Tanweer Hussain**,[2] **Laveet Kumar**,[2] **and Afroza Nahar**[3]

[1]*Department of Electrical Engineering, Mehran University of Engineering & Technology, Jamshoro 76062, Pakistan*
[2]*Department of Mechanical Engineering, Mehran University of Engineering & Technology, Jamshoro 76062, Pakistan*
[3]*Department of Computer Science, Faculty of Science and Technology, American International University – Bangladesh, Dhaka 1229, Bangladesh*

Correspondence should be addressed to Afroza Nahar; afroza@aiub.edu

Received 28 October 2021; Revised 10 December 2021; Accepted 15 December 2021; Published 5 January 2022

Academic Editor: Laurentiu Fara

The solar photovoltaic (PV) power forecast is crucial for steady grid operation, scheduling, and grid electricity management. In this work, numerous time series forecast methodologies, including the statistical and artificial intelligence-based methods, are studied and compared fastidiously to forecast PV electricity. Moreover, the impact of different environmental conditions for all of the algorithms is investigated. Hourly solar PV power forecasting is done to confirm the effectiveness of various models. Data used in this paper is of one entire year and is acquired from a 100 MW solar power plant, namely, Quaid-e-Azam Solar Park, Bahawalpur, Pakistan. This paper suggests recurrent neural networks (RNNs) as the best-performing forecasting model for PV power output. Furthermore, the bidirectional long-short-term memory RNN framework delivered high accuracy results in all weather conditions, especially under cloudy weather conditions where root mean square error (RMSE) was found lowest 0.0025, R square stands at 0.99, and coefficient of variation of root mean square error (RMSE) Cv was observed 0.0095%.

1. Introduction

Electricity is a fundamental instrument to continue socio-economic activities. Pakistan has been facing an electricity shortage for many years due to heavy reliance on expensive imported fuel, suboptimal transmission and distribution systems, and poor revenue collection [1]. Pakistan's current power generation mix is heavily skewed towards imported and high-carbon fuels. Pakistan produced 67.5% of its total power consumption through thermal (furnace oil, coal, regasified liquified gas, and nuclear), 32% from renewables (hydel, wind, solar, and bagasse), and the rest is being imported from Iran. Among other renewable energy sources, solar PV power only contributes less than 1% of country's total power consumption [2].

Solar PV has emerged as a reliable technology and competitive power source globally among its renewable and non-renewable counterparts. The solar PV power installed capacity increased by 22% in 2019, with the second-largest renewable generation growth slightly behind wind [3]. Pakistan is also blessed with a vast potential to generate solar PV power and could be an essential and clean source of energy for country's future energy needs. World Bank (WB) reported that Pakistan could meet its electricity demand by only utilizing as small as 0.071% of Pakistan's land area for solar PV power generation [4]. Since the output of solar PV power is intermittent, it depends entirely on the availability of sunshine hours throughout the day, solar irradiance, angle of incident, the circuit of a cell, and metrological conditions [5, 6]. Hence, integrating solar PV-generated power with a grid creates challenges in grid planning and operation. Forecasting solar PV power is an important planning activity, and a robust solar power forecast is essential to mitigate solar-induced variability and facilitate solar PV grid integration [7]. There has been little study done indigenously in Pakistan on solar power forecasting. Most of the

studies focus on solar system performance and solar radiations [8–11]. Pakistan has taken steps towards the adoption of renewable energy. In a recent meeting for climate change, leaders of the world met in Washington, where Pakistan assured that it would shift to 60% and 30% on renewable energy and use of electric vehicles, respectively, by 2030 [12]. To achieve this huge target, it is essential to accurately forecast the power output of available renewable energy sources, especially PV systems.

In reference [13], a statistical model autoregressive integrated moving average (ARIMA) for weather forecast was proposed, the model with the lowest mean square error value was found accurate with MSE value 0.00029, this work also suggests that by reducing training dataset, accuracy decays (mean square error value increases). In reference [14], a study for the monthly mean temperature forecast for 36 months was completed. The original temperature historical data was obtained from automatic weather station Nanjing. Seasonal autoregressive integrated moving average (SARIMA) model is suggested best fit model for temperature forecast, the mean square error (MSE) for the last three years of validation data is 0.84, 0.89, and 0.94, respectively. MSE values are relatively low, with a slight increase of 0.05 every year. Since the increase in error is not of more importance, the model can be safely utilized for temperature forecast.

In reference [15], by utilizing data from National Solar Radiance DataBase (NSRDB), a multisite (i.e., Atlanta, Newyork, Huawei)-based study was conducted for accurate forecasting of short-term solar irradiance. This study compared different models, i.e., ARIMA, support vector regression (SVR), back propagation neural network (BPNN), and RNN. In conclusion to this study, LSTM results were more accurate, especially on mixed days and cloudy days. Moreover, root mean square error (RMSE), R squared, and mean absolute percentage error (MAPE) values for LSTM in other complicated weather situations were also found competitive than models as mentioned above. Hence, this study suggests LSTM as an accurate model for short-term solar irradiance forecast. In reference [16], a comparative analysis of general regression (GR) and back propagation (BP) models was conducted. First, temperature and irradiance were found most important factors for input data by using Pearson's correlation coefficient. Then, learning vector quantization neural network (LVQ NN) was implemented to observe three different weather conditions, sunny, cloudy, and rainy. Finally, a comparative study of the backpropagation neural network and general regression neural network was conducted. The conclusion to this study suggests the backpropagation model as a more accurate model than the general regression model for PV power generation forecast.

In reference [17], the case study of South Korea was investigated to predict the amount of PV power generation at new sites. Dataset of 164 different sites that contained weather information, estimated solar irradiance, plant capacity, and electricity trading was studied on LSTM model for prediction. It was observed that LSTM could learn complex and nonlinear patterns between power output and factors affecting it at different sites. It is concluded that the proposed LSTM model can be beneficial in accurately predicting PV power output, in any region, with known historical weather data. In reference [18], the backpropagation method is proposed to forecast 24-hour PV power accurately. Before model implementation, a correlation analysis was conducted to investigate the relationship between power output and ambient temperature. Based on this correlation analysis, hourly solar radiation intensity, highest daily temperature, lowest daily temperature, average daily temperature, and hourly PV output were given as inputs to forecast PV power. According to results obtained, a model with settings of 28 neurons in the input layer, 11 neurons in the output layer, and 20 hidden nodes was found to perform the BP model best for PV power output forecast.

A study in reference [19] proposed a novel model to forecast the one-day power output of a single 20 MW PV power plant. Support vector machines (SVM) with weather classification was studied in this work. The process divided weather conditions into four types, clear sky, cloudy, foggy, and rainy. The presented model showcased promising results with low forecasting errors. RMSE was observed at 2.10, and mean relative error (MRE) was found at 8.64% for the chosen site. In reference [20], a hybrid approach is proposed for three different neural networks (FFNN, GRNN, and MLP) to forecast 24 hours PV output for 16 different rooftop solar panels of capacity 250 W. Firstly, stepwise regression was used to select meteorological parameters that are strongly correlated with PV power generation. Then, these parameters were used as input to three different single-stage models (feedforward neural network, general regression neural network, and multilayer regression) and their corresponding hybrid models. Hybrid models prediction results were found very close to measured values. Furthermore, accuracy metrics also pointed out that hybrid models are slightly better than single-staged counterparts.

In reference [21], a case study compared artificial neural networks (ANNs) and RNNs to predict solar irradiance. The study proposed deep learning RNN as a better performing model for forecasting solar radiations. In the results, compared to ANNs, for RNN, a significant improvement of 47% was observed in normalized mean bias error (NMBE), and a 26% improvement in RMSE was also observed. It was observed that with an increase in sampling frequency from 1 hour to 10 minutes, coefficient of variation of RMSE (Cv(RMSE)) of ANN dropped by approximately 30%, and CV (RMSE) of RNN dropped by about 2.19%. This study also suggested that adding a moving average algorithm to predicting model accuracy can improve RNN. In reference [22], the study proposes two PV output prediction models using LSTM and GRU (gate recurrent unit) without knowledge of future meteorological information. This study utilized meteorological information of morning hours to estimate the PV power output around noon. The results found that the proposed GRU-based model could capture the seasonal trend between PV power output in peak zone and its preceding zones more effectively than that of LSTM-based model. Furthermore, even in increased difficulty levels, the GRU-based model performs more accurately than other models.

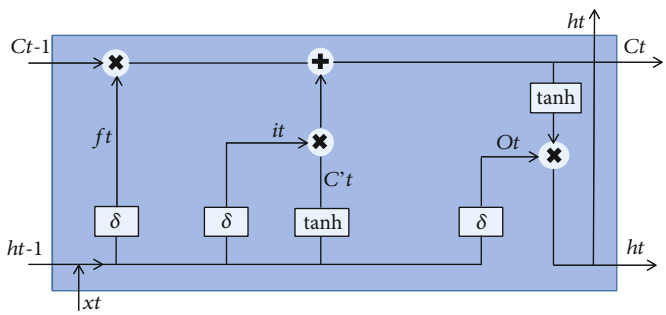


Figure 1: Cell structure of LSTM.

In reference [23], Bi-LSTM for accurate forecasting of solar irradiance hourly and daily was proposed in the study based on two different sites. Multiple models, i.e., vanilla LSTM, attention-based LSTM, Bi-LSTM, and convolutional neural networks, were invested gated for this study, and the models were developed based on single location univariate and multiple location data. Performance and evaluation of model were investigated based on rolling window evaluation. Results indicated that Bi-LSTM and attention-based LSTMs could be used for daily solar irradiance forecasts. In reference [24], a study of single-layer and multilayer LSTM models was conducted for the accurate forecast of PV power generation. Cv (RMSE) was used as a precision method, the results observed for single-layer and multilayer model were 13.8% and 13.2%, respectively. A very little difference in error was observed between the single- and multilayer LSTM model's forecast. Multilayer LSTM showed reduced error. Hence, the accurate forecast is achievable through multilayer LSTM.

Different authors have tested numerous studies and different methodologies. The goal of every author was to propose an accurate model for forecasting. Much work has been already done, and a lot more is completed for precise forecasting. It is considered that deep learning methods have proved themselves in discussed studies for accurate short term time series forecasting. In few studies [13, 14, 25, 26], statistical models were also highlighted as better performing methods to forecast accurately.

In the current study, real-world data has been utilized for accurate PV power output forecast using deep learning LSTM and Bi-LSTM for the first time in the scenario of Pakistan. In addition, a dropout mechanism has been incorporated to prevent overfitting and ensure model's accuracy. Moreover, this study will test different hidden layers for LSTM and consider the best fit of these models. Conclusively, LSTM and Bi-LSTM will be compared, and a final accurate model will be suggested. This study comprises different time series forecasting algorithms for PV power forecasts. Statistical and artificial intelligence-based methodologies are both included in the procedure. The seasonal autoregressive integrated moving average (SARIMA) is a statistical model utilized in this study. The goal of SARIMA is to secure seasonality in data, while our dataset comprises seasonality. long-short-term memory (LSTM) and bidirectional LSTM (Bi-LSTM) are two forms of recurrent neural networks (RNNs) that are studied. The following are the key novelties of this paper:

(i) A deep learning Bi-LSTM is proposed as an accurate power forecasting model for grid-connected PV systems in the study

(ii) Evaluation and comparison of various forecasting models, including statistical and neural network techniques, for time series forecasting of large-scale PV systems

(iii) For accuracy concerns, the study examined over multilayers of LSTM

(iv) The paper includes the time frames for which the forecasting models under consideration are effective

2. Methodology

2.1. Data Description. Data utilized in current work is provided by Quaid-e-Azam Solar Park, Bahawalpur, which is 100 MW power plant. It is a collective 1000 MW project under the China Pakistan Economic Corridor (CPEC) energy section. In the first phase, 100 MW was completed and has been in operation since August 2016 [27]. Data were recorded at intervals of 15 minutes at the power plant. Using Equation (1) [28], the data was averaged on an hourly basis to make signal smoother and improve algorithms' operation.

Furthermore, during operation, it was found that time series forecasts work more accurately on hourly averaged data than that of 15 minutes interval data [29]. The dataset is of one-year time period from 01-January-2019 to 31-December-2019. During the given time, it was observed that power output was constantly zero between 7 PM to 7 AM. Therefore, we only considered the data between 7 AM to 7 PM. The dataset was further separated according to weather conditions such as sunny days, cloudy days, rainy days, partially cloudy days, dusty days, and foggy days.

$$\text{Average} = \frac{t_1 + t_2 + t_3 + t_4}{n}, \tag{1}$$

where n is the number of observations, and $t1$, $t2$, $t3$, and $t4$ are time intervals of an hour.

2.2. Forecasting Models. Forecasting models used for current work are long-short-term memory (LSTM) and bidirectional long-short-term memory (Bi-LSTM), both special types of recurrent neural networks. A statistical approach, seasonal autoregressive integrated moving average (SARIMA), was also studied and compared with RNNs. A brief yet necessary introduction to these models is given, respectively.

2.2.1. Recurrent Neural Networks. Artificial neural networks (ANNs) are a set of algorithms that mimic the human brain. RNNs are ANN types that contain a loop that helps information pass from one step to another. RNNs have a memory based on the previous information, and they look at the previous state to predict the next state [30]. The special kind of RNN, LSTM, was introduced in 1997 to overcome long-term dependency issues [31]. LSTM contains three types of layers, namely input, hidden, and output layer. Unlike other neural networks, LSTMs include memory blocks that are connected through layers. Block's state and output are handled by gates present in each block [32]. Three gates (forget gate, input gate, and output gate) and a cell state make a single LSTM Block. Figure 1 shows the cell structure of LSTM, where $ct-1$ is the previous cell state, $ht-1$ is the hidden layer at time $t-1$, ht is the hidden layer at time t, ct is preliminary input, and xt is input at time t [33].

(1) *Forget gate:* forget gate decides what information to keep and discard from the cell state. The sigmoid layer of forget gate generates either 0 or 1 as a value (ft in Figure 1), 0 means discard, and 1 means keep [34]

(2) *Input gate:* input gate updates the values to cell state; simply, it decides what to store in cell state. A sigmoid layer generates what values to update in cell state (*it* in Figure 1), and tanh (hyperbolic tangent) layer generates a new vector candidate value to be added to cell state ($C^{'}t$ in Figure 1); information from the sigmoid layer and tanh layer is then combined and updated to cell state [30]

(3) *Output gate:* the output gate generates output based on the input and memory of the block. The sigmoid layer has information about the required output (*Ot* in Figure 1), then values are pushed between -1 to 1 using tanh layer. Now, the output from the sigmoid layer and tanh layer is multiplied to get desired output [32]

Besides LSTM, another recurrent neural network algorithm Bi-LSTM is also being employed in this work, a special type of RNNs. The core idea of Bi-LSTM is to present two different neural networks, forward and backward connections to the same output [33]. Therefore, Bi-LSTM takes information from previous contexts and gets information from the future [35]. Bi-LSTM is the LSTMs operating forward and also backward. The forward generates output using information from history, while the backward uses information from the future, which helps in accurate forecasting [33]. Figure 2 depicts Bi-LSTM, the $xt-1$, xt, and $xt+1$ is a set of inputs from past information, the current input and future information, respectively. The mentioned *LSTM* in Figure 2 is the cell structure of LSTM described in Figure 1 [33].

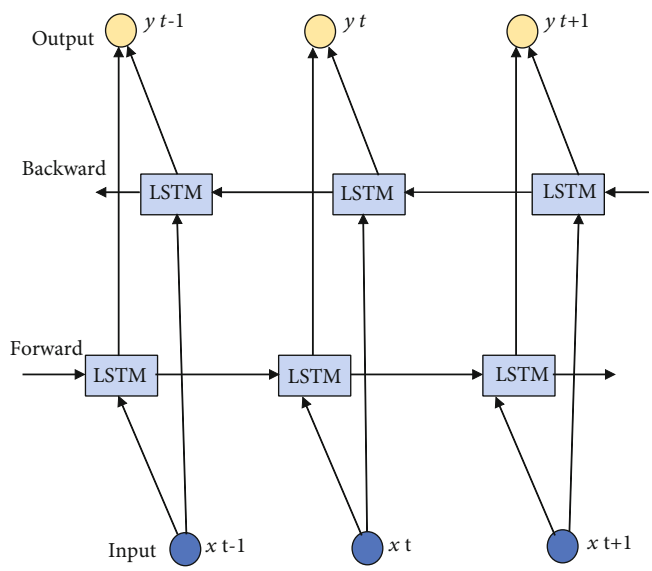


Figure 2: Bidirectional LSTM.

Table 1: Configuration of layers of RNNs.

Model	Optimizer	Neurons	Hidden layers	Neurons division	Epochs
Bi-LSTM	Adam	32	1	32	100
LSTM	Adam	64	2	32 + 32	100

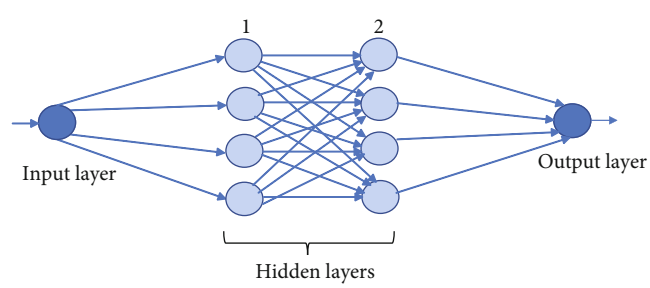


Figure 3: Block diagram of layers configuration of RNN.

2.2.2. Statistical Model. For time series forecasting, the widely used statistical approach is autoregressive integrated moving average, which is acronymically called ARIMA. An extension to ARIMA for seasonal data is seasonal ARIMA (SARIMA) model. With autoregression, integration, and moving average, additional seasonality parameters are added to form SARIMA, as mentioned below in Equation (2) [36].

$$\text{SARIMA} = (p, d, q) * (P, D, Q)^{n}. \quad (2)$$

Table 2: Cv (RMSE) results of hourly forecast of PV power output using LSTM with four different hidden layers.

No. of layers	Sunny	Cloudy	Rainy	Partial cloudy	Dusty	Foggy
1	0.08%	0.14%	1.51%	0.73%	0.65%	0.85%
2	0.15%	0.21%	0.60%	0.51%	0.50%	0.84%
3	0.47%	1.32%	2.12%	0.95%	0.86%	1.45%
4	0.52%	1.47%	0.88%	0.92%	1.08%	1.31%

Table 3: Hourly forecast of PV power output using LSTM and Bi-LSTM, RMSE, R square, and Cv (RMSE) of the respective model.

Weather	Model	RMSE	R square	Cv (RMSE) %
Sunny	LSTM (2 layers)	0.06	0.99	0.15
	Bi LSTM	0.06	0.99	0.15
Cloudy	LSTM (2 layers)	0.058	0.99	0.21
	Bi LSTM	0.0025	0.99	0.0095
Rainy	LSTM (2 layers)	0.157	0.91	0.60
	Bi LSTM	0.12	0.95	0.54
Partial cloudy	LSTM (2 layers)	0.18	0.81	0.51
	Bi LSTM	0.06	0.99	0.17
Dusty	LSTM (2 layers)	0.18	0.80	0.50
	Bi LSTM	0.08	0.99	0.22
Fog	LSTM (2 layers)	0.17	0.85	0.84
	Bi LSTM	0.072	0.98	0.33

Here, p, d, q are nonseasonal parameters and P, D, Q are seasonal parameters with n being order of seasonality, which could be 4 for the quarter and 12 for annual [37]. Seasonal and nonseasonal parts are quite similar, except backshifts are involved in seasonal time [36]. SARIMA in mathematical representation is given below in Equation (3) [38].

$$\varnothing_p(B)\Phi_P(B^n)W_t = \theta_q(B)\Phi_Q(B^n)\omega_t. \tag{3}$$

As described in Equation (1), p, q notations are nonseasonal parameters. They are the same in Equation (2) also. In Equation (2) the P, Q represent order of seasonal autoregressive (AR) and moving average (MA), respectively, and n is the length of period. AR and MA are mathematically represented in Equations (4) and (5), respectively where Φ is respective weight of lagged values and ε is error at respective lagged values [39]. In Equation (3), the B represents backshifts and ω value for noise at time t [38]. W_t is stationary variable, which can further be mathematically explained as Equation (6) [40]. Equation (7) is obtained by merging the value of W_t (given in Equation (6)) in Equation (3).

$$Y_t = \beta_1 + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \cdots + \Phi_P Y_{t-P}, \tag{4}$$

$$Y_t = \beta_2 + \omega_1 \varepsilon_{t-1} + \omega_2 \varepsilon_{t-2} + \cdots + \omega_Q \varepsilon_{t-Q} + \varepsilon_t, \tag{5}$$

$$W_t = (1 - B^n)^D (1 - B)^d X'_t, \tag{6}$$

$$\varnothing_p(B)\Phi_P(B^n)(1 - B^n)^D(1 - B)^d X'_t = \theta_q(B)\Phi_Q(B^n)\omega_t, \tag{7}$$

where n is seasonality length and X'_t is appropriate predifferencing of series X_t (the input value) to ensure constant variance of transformed series over time.

2.3. Implementation of Forecasting Models

2.3.1. Recurrent Neural Networks. This study applied recurrent neural networks to datasets of different weather conditions, sunny days, cloudy days, rainy days, partially cloudy days, dusty days, and foggy days to forecast hourly PV power. The standard LSTM and Bi-LSTM model are available under the *Keras* package that runs on Tensorflow in R programming (R Studio). For PV power forecasting, real-time data in the current study is utilized. For an accurate forecast of PV power output, the real-time power output variable of the dataset is taken into account in the modeling. Then, split the data of PV output into training and validation datasets, with 80% of the data for training and 20% for validation. The models LSTM and Bi-LSTM were trained on the training data and tested on validation data. Min-max normalization method was applied to datasets, and normalized the data between 0 and 1 values, making it easier for models to understand [41]. For LSTM, in this study, four different hidden layers were tested, with the help of different precision methods, the best fit out of these four different models was considered. In this work, number of neurons has been fixed for all datasets based on mean square error [42]. The maximum number of chosen epochs was 100 for current work. Forecasting performance of models is validated by RMSE, R square, and coefficient of variation of RMSE acronymically known Cv(RMSE). The configuration of the layers of RNNs is shown in Table 1 and Figure 3.

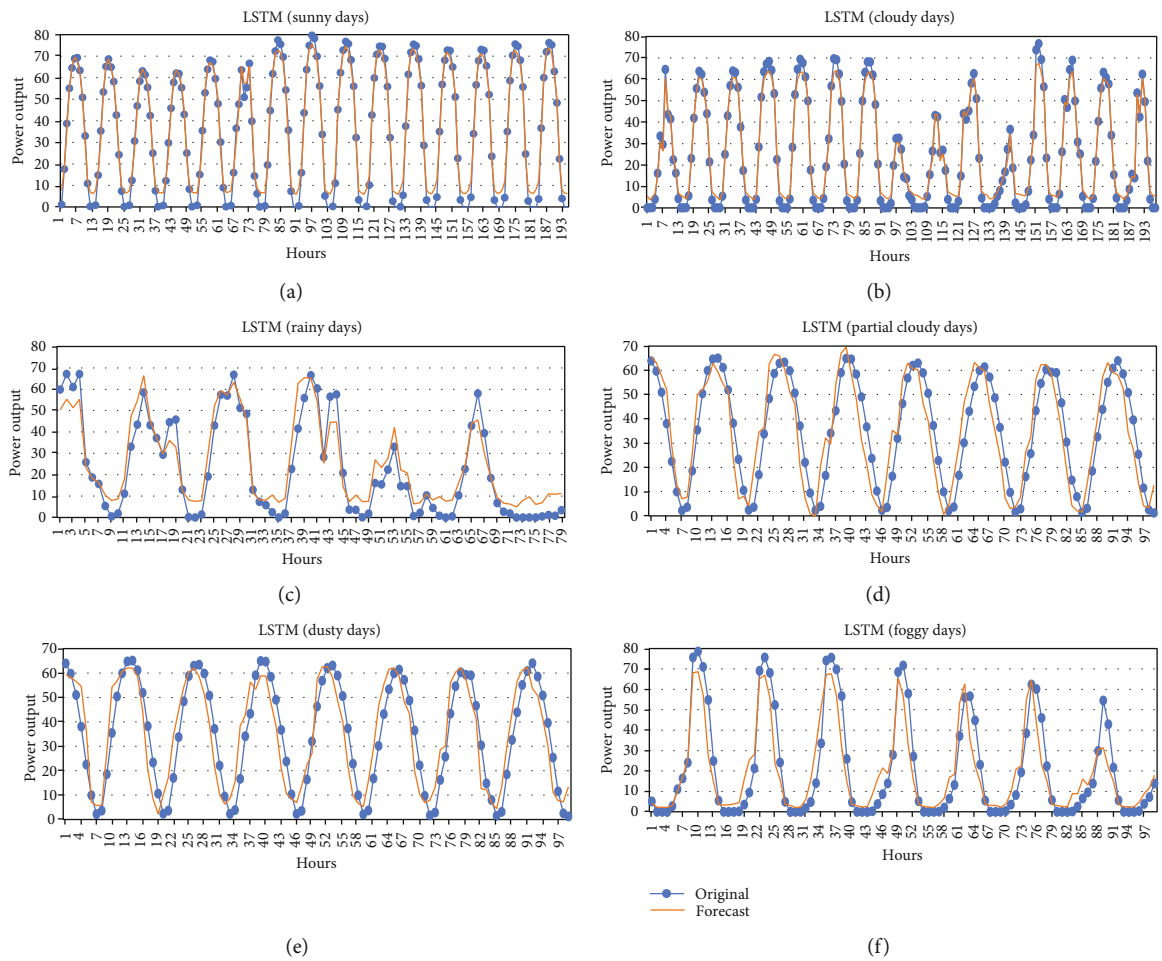


Figure 4: LSTM results for (a) sunny days, (b) cloudy days, (c) rainy days, (d) partial cloudy days, (e) dusty days, and (f) foggy days.

2.3.2. Statistical Model. Akaike's information criterion (AIC) is used to make model parameters effective in ARIMA, ARMA, and SARIMA [36]. The seasonality in the dataset is inherent. Therefore, the SARIMA model was adopted in this study. For the PV power output forecast, the PV power output variable from a real-world dataset has been used. Furthermore, the dataset was partitioned into two sets, one for training and one for testing, to achieve accuracy. Training data makes up around 80% of the dataset, whereas testing data makes up about 20% of real-world datasets. AIC estimates model fit, but AIC does not have any sign of unquestionable quality [43]. SARIMA model for this work has been chosen with a low AIC value (9.387). AIC is determined in [29].

$$\mathrm{AIC}\,(p+q) = N\ \log\left(r^2\right) + 2(p+q), \tag{8}$$

where N represents observations and p and q represents autoregressive and moving average order, respectively.

2.4. Precision Methods. This study adopted the following three precision measures to check the performance of the LSTM and Bi-LSTM in hourly solar PV power forecasts.

$$\mathrm{RMSE} = \sqrt{\sum_{i=1}^{M} \frac{(P_i - O_i)^2}{M}}, \tag{9}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{M}(P_i - O_i)^2}{\sum_{i=1}^{M}(P_i - O_i)^2}, \tag{10}$$

where M is the number of samples, P is the forecasted value, and O is the original value.

$$\mathrm{Cv(RMSE)} = \frac{\mathrm{RMSE}_{\mathrm{forecast}}}{\mathrm{Average}_{\mathrm{forecast}}} \times 100. \tag{11}$$

The idea is that the lower the Cv(RMSE) and RMSE values, the better is accuracy, and higher the R square value, the better is accuracy.

3. Results and Discussion

Table 2 represents the experimental results of different hidden layers tested for LSTM. Table 3 represents the RMSE, R square, and Cv (RMSE) for LSTM (2 layers) and Bi-LSTM with different weather conditions. LSTM 2-layer model was

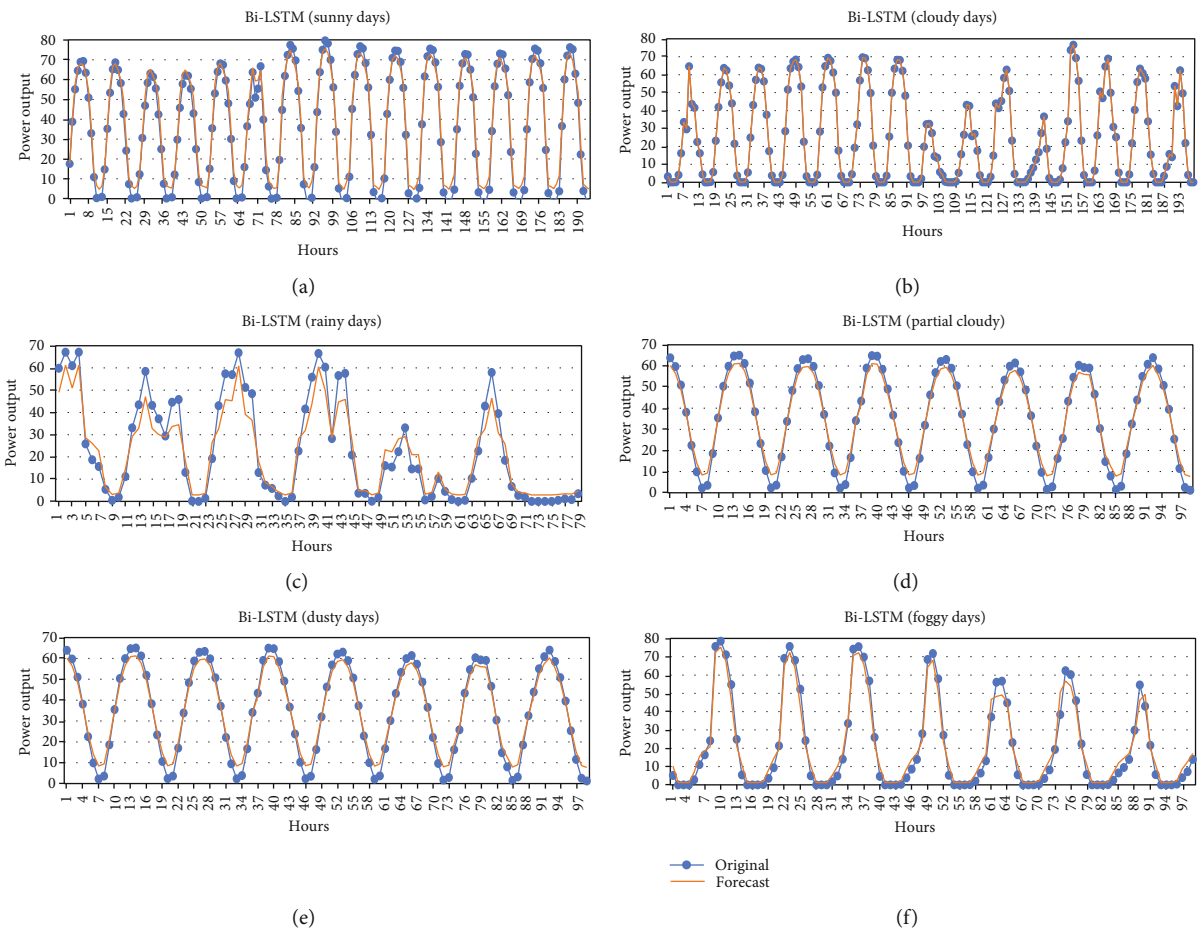


Figure 5: Bi-LSTM results for (a) sunny days, (b) cloudy days, (c) rainy days, (d) partial cloudy days, (e) dusty days, and (f) foggy days.

chosen best fit after conducting a multilayer experiment of LSTM (refer to Table 1). It is important to recall that the dataset of all the weather conditions is certainly not the same. All of the weather conditions have a different number of hours.

3.1. Recurrent Neural Networks

3.1.1. Long-Short-Term Memory RNN. Figure 4 shows a graphical illustration of the LSTM model with 2 hidden layers. The overall performance of LSTM is remarkable in graphical representations.

Figure 4(a)shows that sunny day results of the LSTM model are shown. RMSE, R square, and Cv (RMSE) values for sunny days have been found as 0.06, 0.99, and 0.15%, respectively, which justifies model's accuracy. Figure 4(b) shows that cloudy day results of the LSTM model are shown. RMSE, R square, and Cv (RMSE) values for cloudy weather found are 0.058, 0.99, and 0.21%, respectively. Figure 4(c) shows that rainy day results of the LSTM model are shown; some notable deviations in original and forecasted values can be seen in rainy data results. RMSE, R square, and Cv (RMSE) values for rainy data is 0.157, 0.91, and 0.60%, respectively. Figure 4(d) shows that partial cloudy day results of the LSTM model are shown. It defines a few deviations in forecasted and original values. The RMSE, R square, and Cv (RMSE) 0.18, 0.81, and 0.51%, respectively, were found for partial cloudy data. Figure 4(e) shows that dusty day results of the LSTM model are shown. Dusty days observed nearly identical outcomes as partial cloudy days. RMSE, R square, and Cv (RMSE) 0.18, 0.80, and 0.50%, respectively. Figure 4(f) shows that foggy day results of the LSTM model are shown; few notable deviations can be seen in its graph, RMSE, R square, and Cv (RMSE) value for foggy weather stand at 0.17, 0.85, and 0.84%, respectively. Few authors [32, 44] have also used LSTM for forecasting of photovoltaic power, and the RMSE values achieved are within the similar range as obtained in this research study.

3.1.2. Bidirectional LSTM. Figure 5 shows graphical outcomes for the Bi-LSTM model. Figure 5(a) shows that sunny day results of the Bi-LSTM model are shown. It is found that there are minor deviations in original and forecasted values. RMSE, R square, and Cv (RMSE) values of the Bi-LSTM model for sunny day situations are 0.06, 0.99, and 0.15%, respectively. Figure 5(b) shows that cloudy day results of the Bi-LSTM model are shown, in which results of forecasts are pretty just like the original values. RMSE, R square, and Cv (RMSE) values found are 0.0025, 0.99, and 0.0095%, respectively. Bi-LSTM executed greater accuracy than LSTM

Table 4: RMSE and R square values for respective weather parameters of SARIMA model.

Weather	Model	RMSE	R square	Cv (RMSE) %
Sunny	SARIMA	1.43	0.26	4.57
Cloudy	SARIMA	1.76	0.40	7.03
Rainy	SARIMA	2.545	0.48	10.79
Partial cloudy	SARIMA	1.654	0.229	6.55
Dusty	SARIMA	2.53	0.274	7.75
Fog	SARIMA	3.81	0.19	25.4

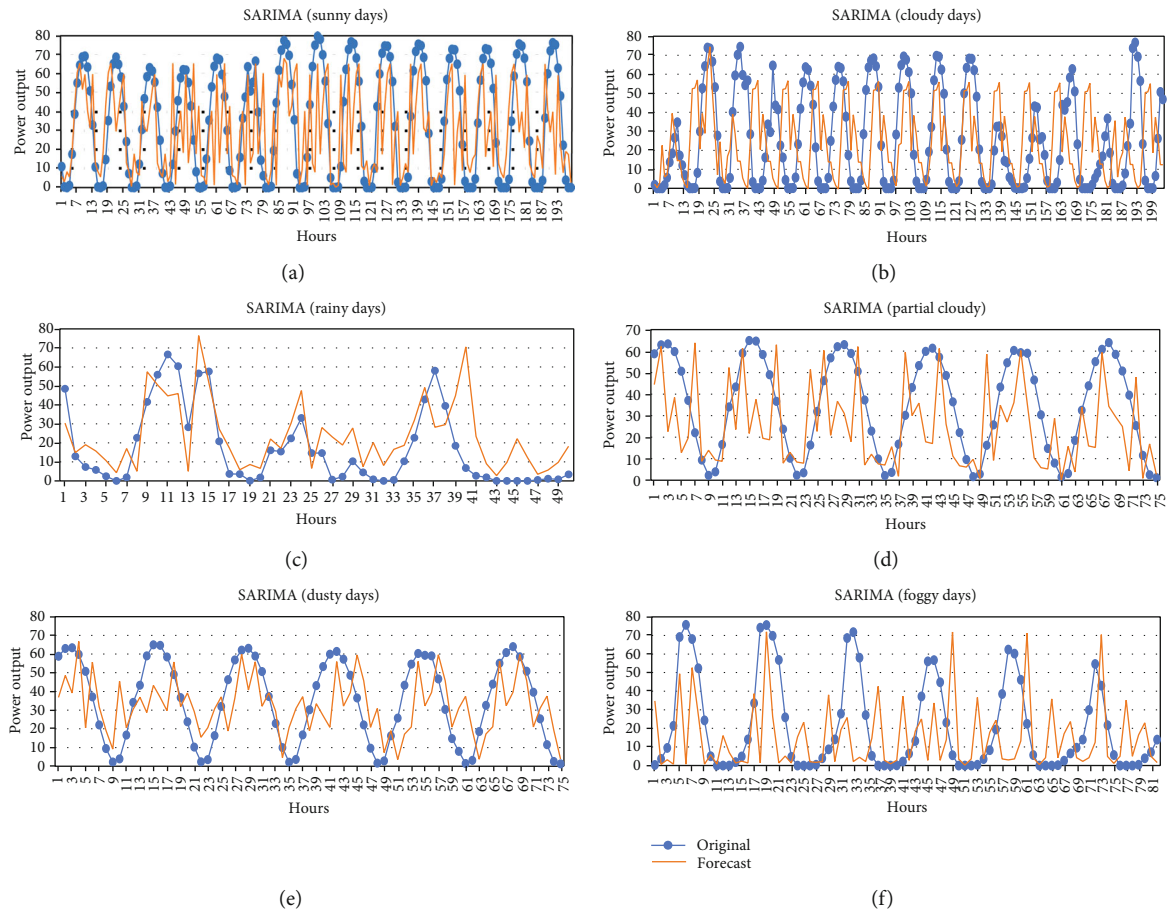


Figure 6: SARIMA results for (a) sunny days, (b) cloudy days, (c) rainy days, (d) partial cloudy days, (e) dusty days, and (f) foggy days.

on cloudy weather situations. Figure 5(c) shows that rainy day outcomes of Bi-LSTM are shown, a few substantial deviations are found at a few factors of original and forecasted data. RMSE, R square, Cv (RMSE) values found are 0.12, 0.95, and 054%, respectively. Figure 5(d) shows that partial cloudy weather outcomes are shown, the consequences are exceptional, and deviations are negligible; furthermore, RMSE, R square, and Cv (RMSE) values stand at 0.06, 0.99, and 0.17%, respectively. This defines the accuracy of the Bi-LSTM model. Figure 5(e) shows that dusty day results are shown. Bi-LSTM results are excellent, the graph appears pleasant with low deviations. RMSE, R square, and Cv (RMSE) results are 0.08, 0.99, and 0.22%, respectively, which also indicates the accuracy of the model. Figure 5(f) shows

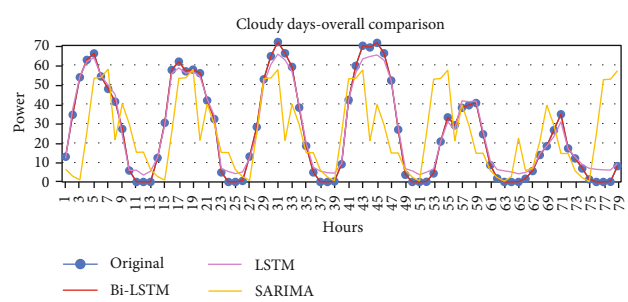


Figure 7: Forecasting results of LSTM, Bi-LSTM, and SARIMA.

Table 5: Results of LSTM, Bi-LSTM, and SARIMA.

Weather	Model	RMSE	R square	CV (RMSE) %
Sunny	LSTM (2 layers)	0.06	0.99	0.15
	Bi-LSTM	0.06	0.99	0.15
	SARIMA	1.43	0.26	4.57
Cloudy	LSTM (2 layers)	0.058	0.99	0.21
	Bi-LSTM	0.0025	0.99	0.0095
	SARIMA	1.76	0.40	7.03
Rainy	LSTM (2 layers)	0.157	0.91	0.60
	Bi-LSTM	0.12	0.95	0.54
	SARIMA	25.45	0.48	10.79
Partial cloudy	LSTM (2 layers)	0.18	0.81	0.51
	Bi-LSTM	0.06	0.99	0.17
	SARIMA	16.54	0.229	6.55
Dusty	LSTM (2 layers)	0.18	0.80	0.50
	Bi-LSTM	0.08	0.99	0.22
	SARIMA	2.53	0.274	7.75
Fog	LSTM (2 layers)	0.17	0.85	0.84
	Bi-LSTM	0.072	0.98	0.33
	SARIMA	3.81	0.19	25.4

that foggy day results of Bi-LSTM are shown. Very minor deviations too on few of the points are found. RMSE, R square, and Cv (RMSE) value found are 0.072, 0.98, and 0.33%, respectively. The results of those precision techniques endorse accuracy for the Bi-LSTM model.

3.2. Statistical Model. In this study, a statistical forecast method was also taken into consideration for the time series forecast of PV power. The overall performance of RNNs was superior to SARIMA. Table 4 represents the precision outcomes for the SARIMA model, and graphical results also are shown in Figure 6.

3.3. Comparative Analysis. The findings of the deep learning RNN model were found to be the most accurate on a cloudy weather dataset. Figure 7 and Table 5 show a thorough comparison of all the models studied over a dataset of cloudy days.

4. Conclusion

The hourly PV electricity output forecast is vital for operation, maintenance, and overcoming the demanding situations faced by the grid-linked PV plants. A simple statistical model for time series forecasting of hourly solar PV electricity, SARIMA, and the performances of LSTM and Bi-LSTM recurrent neural networks have been examined in this paper. The models have been trained and examined on absolutely distinctive weather parameters.

The study recommends that the RNNs overall performance is superior to that of the SARIMA model. The RNNs have deep systems to remedy intense troubles which include vanishing gradients; for this reason, they have carried out forecasting with such excessive accuracy. Furthermore, in the assessment of 2 distinctive RNNs, findings recommend bidirectional long-short-term memory (Bi-LSTM) carried out greater accuracy than long-short-term memory (LSTM). It is found that Bi-LSTM scored higher R squares in each weather condition, whilst it remained lowest 0.95 on rainy days. The RMSE and Cv (RMSE) values for Bi-LSTM have been also recorded lowest in all weathers; RMSE was observed maximum at rainy weather with 0.12 value, and maximum Cv (RMSE) for Bi-LSTM was found at rainy data with 0.54%. The graphical representations of Bi-LSTM (shown in Figure 4) also propose that actual values are very near the forecasted values under distinctive weather conditions. Hence, for short time forecasting of the power output of grid-linked PV power plant, this paper suggests bidirectional LSTM recurrent neural network as exceptional model with high accuracy.

Nomenclature

PV: Photovoltaic
ANN: Artificial neural network
RNN: Recurrent neural network
WB: World Bank
LSTM: Long-short-term memory
Bi-LSTM: Bidirectional long-short-term memory
RMSE: Root mean square error
Cv (RMSE): Coefficient of variation of room mean square error
ARIMA: Autoregressive integrated moving average
SARIMA: Seasonal autoregressive integrated moving average
AIC: Akaike's information criterion.

Data Availability

Data are available upon request to submitting or the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] NEPRA, *Annual report annual report 2019-20*, Islamabad, Pakistan, 2019https://nepra.org.pk/publications/Annual.Reports.ph.

[2] NEPRA, "State of industry report 2020," *Regulation*, vol. 53, no. 9, pp. 1689–1699, 2020.

[3] "Solar - fuels & technologies - IEA," June 2021, https://www.iea.org/fuels-and-technologies/solar.

[4] World Bank, "Expanding renewable energy in Pakistan's electricity mix," 2020, September 2021, https://www.worldbank.org/en/news/feature/2020/11/09/a-renewable-energy-future-for-pakistans-power-system.

[5] "How environmental factors affect solar panel performance - naked solar," September 2021, https://nakedsolar.co.uk/solar-pv/environmental-factors/.

[6] M. A. Al-Nimr, S. Kiwan, and H. Sharadga, "Simulation of a novel hybrid solar photovoltaic/wind system to maintain the cell surface temperature and to generate electricity," *International Journal of Energy Research*, vol. 42, no. 3, pp. 985–998, 2018.

[7] K. N. Nwaigwe, P. Mutabilwa, and E. Dintwa, "An overview of solar power (PV systems) integration into electricity grids," *Materials Science for Energy Technologies*, vol. 2, no. 3, pp. 629–633, 2019.

[8] Z. Abbas, K. Harijan, P. H. Shaikh, G. D. Walasai, and F. Ali, "Effect of ambient temperature and relative humidity on solar PV system performance: a case study of Quaid-e-Azam Solar Park, Pakistan," *Sindh University Research Journal-SURJ (Science Series)*, vol. 49, no. 4, pp. 721–726, 2017.

[9] Z. Abbas, K. Harijan, P. Hameed, and F. Bhayo, "Effect of dust on the performance of photovoltaic system (a case study of Quaid-E-Azam Solar Park Bahawalpur)," *Sindh University Research Journal-SURJ (Science Series)*, vol. 1, no. 6, pp. 73–79, 2017.

[10] A. A. Khosa, T. U. Rashid, N. U. H. Shah, M. Usman, and M. S. Khalil, "Performance analysis based on probabilistic modelling of Quaid-e-Azam Solar Park (QASP) Pakistan," *Energy Strategy Reviews*, vol. 29, article 100479, 2020.

[11] M. H. Aziz, "Performance evaluation of an on grid 100 MW Quaid-E-Azam Solar Power Bahawalpur," 2018, https://www.sciencedirect.com/science/article/pii/S2211467X2030033X.

[12] A. Iqbal, *Pakistan will shift to 60pc clean energy by 2030, world assured - newspaper - DAWN.COM*, Dawn.Com, 2021.

[13] A. G. Salman and B. Kanigoro, "Visibility forecasting using autoregressive integrated moving average (ARIMA) models," *Procedia Computer Science*, vol. 179, no. 2019, pp. 252–259, 2021.

[14] P. Chen, A. Niu, D. Liu, W. Jiang, and B. Ma, "Time series forecasting of temperatures using SARIMA: an example from Nanjing time series forecasting of temperatures using SARIMA: an example from Nanjing," *IOP Conference Series: Materials Science and Engineering*, vol. 394, no. 5, article 052024, 2018.

[15] Y. Yu, J. Cao, and J. Zhu, "An LSTM short-term solar irradiance forecasting under complicated weather conditions," *IEEE Access*, vol. 7, pp. 145651–145666, 2019.

[16] J. Zhong, L. Liu, Q. Sun, and X. Wang, "Prediction of photovoltaic power generation based on general regression and back propagation neural network," *Energy Procedia*, vol. 152, pp. 1224–1229, 2018.

[17] Y. Jung, J. Jung, B. Kim, and S. U. Han, "Long short-term memory recurrent neural network for modeling temporal patterns in long-term power forecasting for solar PV facilities: case study of South Korea," *Journal of Cleaner Production*, vol. 250, article 119476, 2020.

[18] L. Liu, D. Liu, Q. Sun, H. Li, and R. Wennersten, "Forecasting power output of photovoltaic system using a BP network method," *Energy Procedia*, vol. 142, pp. 780–786, 2017.

[19] J. Shi, W. J. Lee, Y. Liu, Y. Yang, and P. Wang, "Forecasting power output of photovoltaic systems based on weather classification and support vector machines," *IEEE Transactions on Industry Applications*, vol. 48, no. 3, pp. 1064–1069, 2012.

[20] P. Ramsami and V. Oree, "A hybrid method for forecasting the energy output of photovoltaic systems," *Energy Conversion and Management*, vol. 95, pp. 406–413, 2015.

[21] Z. Pang, F. Niu, and Z. O'Neill, "Solar radiation prediction using recurrent neural network and artificial neural network: a case study with comparisons," *Renewable Energy*, vol. 156, pp. 279–289, 2020.

[22] D. Lee and K. Kim, "PV power prediction in a peak zone using recurrent neural networks in the absence of future meteorological information," *Renewable Energy*, vol. 173, pp. 1098–1110, 2021.

[23] B. Brahma and R. Wadhvani, "Solar irradiance forecasting based on deep learning methodologies and multi-site data," *Symmetry*, vol. 12, no. 11, article 1830, 2020.

[24] M. K. Park, J. M. Lee, W. H. Kang, J. M. Choi, and K. H. Lee, "Predictive model for PV power generation using RNN (LSTM)," *Journal of Mechanical Science and Technology*, vol. 35, no. 2, pp. 795–803, 2021.

[25] F. Shaikh and Q. Ji, "Forecasting natural gas demand in China: logistic modelling analysis," *International Journal of Electrical Power & Energy Systems*, vol. 77, pp. 25–32, 2016.

[26] F. Shaikh, Q. Ji, P. H. Shaikh, N. H. Mirjat, and M. A. Uqaili, "Forecasting China's natural gas demand based on optimised nonlinear grey models," *Energy*, vol. 140, pp. 941–951, 2017.

[27] *Quaid-e-Azam 1000MW Solar Park (Bahawalpur) Quaid-e-Azam*China-Pakistan Economic Corridor (CPEC) Official WebsiteJune 2021, http://cpec.gov.pk/project-details/10.

[28] Cuemath.com, "Average formula - what is the average formula? Examples," September 2021, https://www.cuemath.com/average-formula/.

[29] H. Sharadga, S. Hajimirza, and R. S. Balog, "Time series forecasting of solar power generation for large-scale photovoltaic plants," *Renewable Energy*, vol. 150, pp. 797–807, 2020.

[30] C. Olah, "Understanding LSTM networks – colah's blog," 2015, September 2021, https://colah.github.io/posts/2015-08-Understanding-LSTMs/.

[31] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[32] M. Abdel-Nasser and K. Mahmoud, "Accurate photovoltaic power forecasting models using deep LSTM-RNN," *Neural*

Computing and Applications, vol. 31, no. 7, pp. 2727–2740, 2019.

[33] B. Iung, "Cœur et grossesse," *EMC – Trait médecine AKOS*, vol. 8, no. 2, pp. 1–4, 2013.

[34] M. AlKandari and I. Ahmad, "Solar power generation forecasting using ensemble approach based on deep learning and statistical methods," *Applied Computing and Informatic*, 2020.

[35] A. Graves, A. R. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 3, pp. 6645–6649, Vancouver, BC, Canada, 2013.

[36] P. Chen, A. Niu, D. Liu, W. Jiang, and B. Ma, "Time series forecasting of temperatures using SARIMA: an example from Nanjing," *IOP Conference Series: Materials Science and Engineering*, vol. 394, no. 5, 2018.

[37] R. J. Hyndman and G. Athanasopoulos, *8.9 seasonal ARIMA models | forecasting: principles and practice*, O Texts, 2nd edition, 2018.

[38] M. Farsi, D. Hosahalli, B. R. Manjunatha et al., "Parallel genetic algorithms for optimizing the SARIMA model for better forecasting of the NCDC weather data," *Alexandria Engineering Journal*, vol. 60, no. 1, pp. 1299–1316, 2021.

[39] G. Shmueli and K. C. Lichtendahl Jr., "Practical time series forecasting with R: a hands-on guide," 2015, http://www.amazon.com/Practical-Time-Series-Forecasting-Hands-On/dp/0991576632.

[40] J. Flaherty and R. Lombardo, "Modelling private new housing starts in Australia," in *A paper presented in the Pacific-Rim Real Estate Society conference*, pp. 24–27, University of Technology Sydney, Australia, 2000.

[41] C. Martínez, "Neural networks for business analytics with R | Udemy," 2021, https://www.udemy.com/course/neural-networks-for-business-analytics-with-r/learn/practice/1302928?start=start-page#questions.

[42] A. Mellit, S. Sağlam, and S. A. Kalogirou, "Artificial neural network-based model for estimating the produced power of a photovoltaic module," *Renewable Energy*, vol. 60, pp. 71–78, 2013.

[43] H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp. 716–723, 1974.

[44] D. Lee and K. Kim, "Recurrent neural network-based hourly prediction of photovoltaic power output using meteorological information," *Energies*, vol. 12, no. 2, 2019.

Hindawi
International Journal of Photoenergy
Volume 2021, Article ID 4228658, 14 pages
https://doi.org/10.1155/2021/4228658

Research Article

Estimation of the Uncertainty due to Each Step of Simulating the Photovoltaic Conversion under Real Operating Conditions

Anne Migan Dubois,[1] **Jordi Badosa**,[2] **Vincent Bourdin**,[3] **Moira I. Torres Aguilar**,[2] **and Yvan Bonnassieux**[4]

[1]*Université Paris-Saclay, CentraleSupélec, CNRS, Laboratoire Génie Électrique et Électronique de Paris (GeePs), 91192 Gif-Sur-Yvette, France*
[2]*École Polytechnique, IPSL, Université Paris-Saclay, ENS, Sorbonne Université, CNRS, Laboratoire de Météorologie Dynamique (LMD), 91128 Palaiseau, France*
[3]*CNRS, Laboratoire Interdisciplinaire Des Sciences du Numérique (LISN), 91405 Orsay Cedex, France*
[4]*CNRS, École Polytechnique, Laboratoire de Physique et Des Couches Minces (LPICM), Route de Saclay, 91128 Palaiseau, France*

Correspondence should be addressed to Anne Migan Dubois; anne.migan-dubois@geeps.centralesupelec.fr

Received 6 July 2021; Accepted 23 August 2021; Published 15 October 2021

Academic Editor: Alberto Álvarez-Gallegos

The simulation of photovoltaic installations is a major issue for their sizing, their smart grid operation, and their fault detection and diagnosis. In this article, we study in detail every step of the simulation chain, either from the global horizontal irradiance and the ambient temperature (i.e., 4 steps of simulation) or considering the global in-plane irradiance and the module operating temperature (i.e., 1 step of simulation). The average quality estimation of the models is made through the calculations of average annual error between estimations and measurements, from 2016 to 2020. We have shown that the most uncertain step is the conversion of the global irradiance in its diffuse and direct components (17.2%, 2 models tested). If the model goes up to the in-plane irradiance, the average annual error decreases to 5.3% (6 models tested). The photovoltaic module temperature calculation induces an error of less than 2Â°C (4 models tested with 2 configurations). Meanwhile, the photoelectrical conversion shows a 3.5% error, similar to the measurement uncertainties, considering as input, the modules temperature, and the in-plane irradiance. If the simulation goes from the global irradiance and the ambient temperature measured locally, the estimation leads to a 6.7% average annual error. If the local measurements are not available, we can use the closest meteorological station's records (13 for our study), and the error becomes 12.1%. Finally, we can also use the satellite images that lead to a 15.2% error, for average per year. The impact of available input shows that modeling the DC photovoltaic production, using global horizontal irradiance and ambient temperature, gives rise to an error of 6.6% for local measurements, 12.1% for weather station measurements, and 15.2% for satellite images estimations. This article thus draws up a review of the existing models, allowing to calculate the DC production of a photovoltaic module, depending on the atmospheric conditions, and highlights the most precise or most critical steps, considering in situ and weather station ground-based measurements, and also estimation from satellite images.

1. Introduction

Photovoltaic (PV) production mainly depends on the solar radiation incident on PV modules. Solar resource variability and uncertainty associated with the modeling of PV energy production are one of the most important factors that influence the grid stability (with wind turbines), regardless of the size of the power grid [1]. The ability to precisely predict the energy produced by PV systems is of great importance and has been identified as one of the key challenges for massive

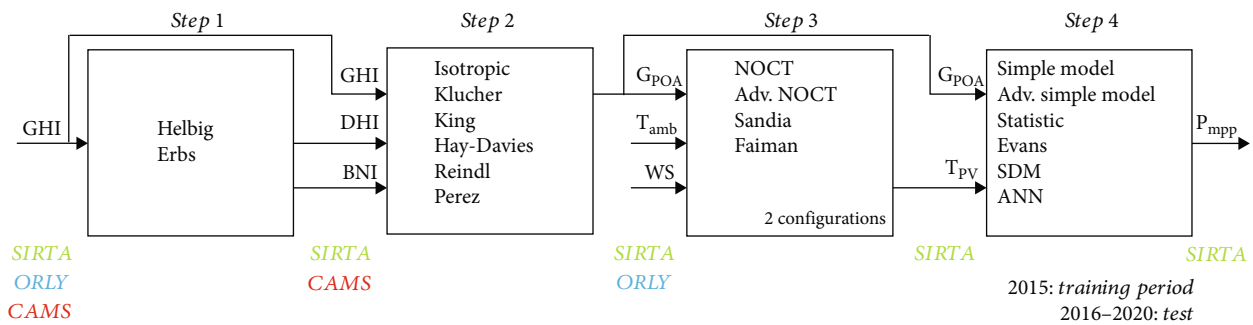


Figure 1: Schematic of the process for the simulation of PV production, studied models, and data sources available for each step.

PV integration [2, 3]. It is also a milestone in the sizing step of PV installations. This can also be used for PV fault detection and diagnosis.

This study focuses on evaluating the uncertainty on PV production estimation at each step of the modeling process. Each step is studied independently, and also its impact on the whole simulation chain is evaluated.

Few articles study the impact of the uncertainty of the modeling process. We can cite [4] which examines the uncertainty in long-term PV system yield predictions by statistical modeling (using Solar Advisor Model software) of a hypothetical 10 MW crystalline silicon PV system in Toronto (Canada). In this case study, uncertainties were estimated to be about 3.9% for year-to-year climate variability, 5% for long-term average horizontal irradiation, 3% for estimation of the in-plane radiation, 3% for power rating of modules, 2% for losses due to dirt and soiling, 1.5% for losses due to snow, and 5% for other sources of error. By performing statistical simulations, it was found that the combined uncertainty is approximately 8.7% for the first year of operation and 7.9% for the average yield over the PV system lifetime.

The study led by Sandia Laboratory [5] explains that the solar resource uncertainty (due to measurements, variability, spectrum...) is between 5 to 17%, the transposition of the horizontal irradiances to the in-plane irradiance is between 0.5 to 2%, the energy simulation and power plant losses induce uncertainty of 3 to 5%, and annual degradation uncertainty is about 0.5 to 1%. Using in situ ground-based measurements can reduce the uncertainty by up to 3.5%.

In reference [6], the authors compare yield predictions and monitoring data for 26 PV power plants located in southern Germany and Spain. If rather old radiation databases are used, the model systematically underestimates PV production by about 5% due to increased irradiance in recent years. Using recent satellite-derived irradiance avoids this underestimation. According to them, the main factor for the uncertainty of yield predictions is the aging, that is to say, the observed decrease of performance ratio. In this study, it decreases by 0.5% per year on average with a relatively high spread between systems. This decrease is attributed to nonreversible degradations and reversible effects, like soiling. The conclusion is that the uncertainty of the state-of-the-art yearly yield predictions using recent solar irradiance data is estimated to about 8%.

Only two publications compare the uncertainty of different photoelectric conversion models [7, 8]. Those two papers, written by the same authors, compare the annual PV yield prediction errors of four models: single-point efficiency, single-point efficiency with temperature correction, PVUSA (Photovoltaics for Utility Scale Applications), and single-diode model (SDM), against outdoor measurements for different grid-connected PV systems in Cyprus over a 4-year evaluation period. The best agreement between the modeled results and outdoor measurements for crystalline silicon PV technologies was obtained using SDM. The energy yield for thin-film technologies was more accurately predicted using the PVUSA model.

Our approach in this paper is similar to indirect forecasts: first, we calculate (or measure) the solar global in-plane irradiance (G_{POA}) and PV module operating temperature (T_{PV}), and then, using a PV performance model, we calculate the power produced at the maximum power point (P_{MPP}) [9]. Only the DC side of the photoelectric conversion is considered (no inverter). The different steps of the PV simulation are summarized in Figure 1, together with the considered models and the data sources.

This study focuses on the evaluation of the uncertainty on PV production estimation, step by step using different models, based on various data sources: in situ (SIRTA) and (ORLY) weather station ground-based measurements and geostationary satellite-based estimations (CAMS) during five years of operation.

For this purpose, we firstly present the data sources in Section 1. Section 2 describes the models that are used for each step of the modeling process with their accuracy. The results are presented in Section 3: it deals with modeling error estimation for the whole PV system simulation focusing on the uncertainty weight of each step and considering different data sources. These results are summarized in a conclusion and some perspectives are drawn.

2. Experimental Data Sources

In this study, three sources of data are explored. The highest level of precision corresponds to in situ ground-based measurements realized in our experimental research platform, where we do PV and atmospheric measurements, namely, SIRTA. Then, we consider the nearest weather station at ORLY airport where we use irradiances and ambient temperature measurements. Lastly, we use satellite-based estimations to proceed estimation of irradiance components

Figure 2: Outdoor characterization PV platform located at SIRTA.

using CAMS. All of these data sources are described in the following subsections.

The considered data are hourly averages and only with solar elevation angles higher than 10°.

2.1. Outdoor Photovoltaic Characterization Platform. Our experimental research platform is installed at the Atmospheric Research Observatory SIRTA [10] located in Palaiseau (France, 48.7 N, 2.2 E), 18.3 South-West of Notre-Dame de Paris.

A PV test bench was installed there in 2014 and comprises six commercial PV modules issued from different technologies (Figure 2). In this paper, we only consider the crystalline-silicon PV module, the second one from the left in Figure 2. All the equipment and sensor are listed in Table 4 of Appendix A, with their accuracy.

The current-voltage characteristics are measured with Agilent DC electronic loads (6060B), each minute from sunrise to sunset. The maximum power point (P_{MPP}) is derived from this characteristic. Under normal operating conditions, the voltage drop along the junction line is typically 0.5 at P_{MPP}, which results in a systematic error of less than 1.6% on the measured power. In addition, due to the transient mode of measurement, an uncertainty of ±5.6% of the measured power should be considered.

The PV module operating temperature (T_{PV}) is measured with 4-wired class A platinum sensors (Pt100) glued on the back-sheet. Their resolution is 0.01°C, and the accuracy is between ±0.15°C and ±0.30°C when the temperature is between -20°C and +100°C. In the transient mode of measurement with short pulses of 1 mA, there is no problem of self-heating of the probes. The overall uncertainty on the measured temperature is about ±0.40°C. The difference between the cell temperature and the probe glued on its back even with the highest irradiance is always less than 2°C, as we simulated it by common thermal laws [11]. These sensors are read by a digital multimeter Tektronix (DMM 3700A).

The global in-plane irradiance (G_{POA}) is measured with a second class solar radiometer (Hukseflux SR01) installed in the same plane as the PV modules. The SR01 is compliant with the ISO 9060 second class specifications. It was compared to a secondary standard for 2 weeks, and an overestimation of less than 0.3% on instantaneous measurements was observed. Thus, we can estimate that the SR01 is accurately calibrated.

2.2. In Situ Atmospheric Ground-Based Measurements SIRTA. Ground-based measurements are realized at SIRTA. It is a reference meteorological and climate observatory with more than 150 remote sensing and in situ instruments. In terms of radiometric measurements, the site is part of the Baseline Surface Radiation Network (BSRN) since 2003 [12]. Global horizontal irradiance (GHI), diffuse horizontal irradiance (DHI), and ground albedo (Albedo) measurements are realized following BSRN standards with Kipp & Zonen CMP22. The direct normal irradiance (BNI) is measured by a Kipp & Zonen CHP1 pyrheliometer. Both CMP22 and CHP1 are compliant with ISO 9060 secondary standard specifications, and they are expected to have a daily uncertainty lower than ±1%. The ambient temperature (T_{amb}) is measured at 1.5 m above the ground by a class A Guilcor platinum sensor (Pt100) with a precision of ± 0.15°C. A Vector A100R anemometer attached to the PV platform measures local wind speed (WS) with an accuracy of 0.1 ms^{-1} and a threshold of 0.3 ms^{-1}.

2.3. Weather Station Atmospheric Ground-Based Measurement ORLY. The weather station which is the nearest to our PV test bench is ORLY airport, located at Athis-Mons (France, 48.4 N, 2.2 E), about 13 km East from SIRTA, opened in 1921. T_{amb} is measured with class B sensors with a precision of ±0.5°C. WS measurements are made by class B anemometers whose accuracy is 10% or 0.5 ms^{-1}. GHI is measured with a pyranometer Kipp & Zonen CM6B compliant with ISO 9060 first-class specifications.

2.4. Satellite Irradiance Estimation CAMS. This study also considers estimations of irradiances from Meteosat geostationary satellite observations computed by Copernicus Atmosphere Monitoring Service (CAMS) [13, 14].

The temporal resolution is one minute, hourly averaged, and the spatial resolution results from interpolation to the point of interest. Meteosat pixels have an elliptic shape, and their average diameter ranges from 3 km to 7 km depending on the viewing geometry of the satellite. The estimations are GHI, DHI, and BNI in clear sky conditions; verbose mode with all atmospheric input parameters used for clouds, aerosols, ozone, water vapor, and the surface reflective properties. T_{amb} and WS are not estimated with satellite images. We use those that are measured in ORLY weather station, if needed.

Table 1: Available data versus data sources.

Data source	GHI	DHI	BNI	G_{POA}	WS and T_{amb}	T_{PV}
SIRTA	✓	✓	✓	✓	✓	✓
ORLY	✓				✓	
CAMS	✓	✓	✓			

Table 2: SDM validation in STC.

Parameter	STC	Model	Relative difference
P_{MPP}	245.7 W	246.5 W	0.32%
V_{MPP}	30.7 V	30.8 V	0.26%
I_{MPP}	8.0 A	8.0 A	0.58%

All the installed sensors and data-logging devices for the three data sources, with their accuracy, are listed in Table 4 of Appendix A.

The available data used for this study is detailed in Table 1:

3. Theoretical Modeling and Model Accuracy

G_{POA} and T_{PV} are needed to estimate P_{MPP}. Either they are available through in situ measurement or they should be estimated, adding up to three modeling steps (as presented in Figure 1).

All presented models are evaluated by yearly averages of data (2016, 2017, 2018, 2019, and 2020). Complete models can be found in the literature listed. All those theoretical models have been programmed using Python, an object-oriented, multiparadigm, and multiplatform programming language. All computed data are summarized in Figure 9 of Appendix B.

3.1. Modeling Performance Indicators. In order to compare estimations to measurements, we compute relative mean bias error (rMBE) and relative mean absolute error (rMAE), as defined in the following equations.

$$\mathrm{rMBE} = \frac{\sum_{i=1}^{N}\left[X_{\mathrm{calc}}(i) - X_{\mathrm{meas}}(i)\right]}{\sum_{i=1}^{N} X_{\mathrm{meas}}(i)} \times 100, \quad (1)$$

$$\mathrm{rMAE} = \frac{\sum_{i=1}^{N}\left|X_{\mathrm{calc}}(i) - X_{\mathrm{meas}}(i)\right|}{\sum_{i=1}^{N} X_{\mathrm{meas}}(i)} \times 100, \quad (2)$$

where X can be DHI, BNI, G_{POA}, or P_{MPP}, X_{calc} correspond to the result of X modeling, X_{meas} is the measured value of X, and N is the number of measured (or calculated) values of X, during one year. The indicators for T_{PV} are expressed in and not normalized which would have been nonsense.

Boxplot is drawn for each rMBE and rMAE calculated. The central red mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually using the "+" symbol. It should be noted that the use of a boxplot to represent 5 values (5 years from 2016 to 2020) is not necessarily very relevant, but it is visually very meaningful.

In this part, the performance of each step of Figure 1 is estimated individually. Input data are in situ ground measurements from SIRTA, for all presented results in this part, as the objective is to evaluate the uncertainty of each modeling step.

3.2. Step 1: Estimation of the Diffuse Fraction. In this study, we consider two models to compute DHI and BNI components of the irradiance from measured GHI. This input is taken from SIRTA in situ ground measurements as the objective is to evaluate only step 1.

For Helbig model, the fraction of the diffuse irradiance is calculated from GHI using an empirical relationship, giving an estimated value of DHI and BNI [15]. The empirical coefficients taken for this study are those found in [15] because modeling performance is not enhanced by using coefficients that have been computed with real data.

Assuming that the diffuse fraction of the hourly total radiation (DHI/GHI) is strongly correlated with the hourly clearness index (k_T), Erbs proposes an empirical relation using the combined data of four American locations [16]. k_T is the ratio of the hourly global radiation to the hourly extraterrestrial radiation.

rMAE and rMBE are calculated using SIRTA in situ ground measurements of DHI and BNI compared to modeled ones using GHI also measured at SIRTA.

As can be seen in Figure 3, the dispersion is wider for the estimation of BNI. DHI is more often underestimated, whereas BNI is overestimated. Helbig models seem to perform better than Erbs, with rMBE = −3.1 ± 1.1% and rMAE = 17 ± 0.9% for the estimation of DHI and rMBE = 5.4 ± 1.6% and rMAE = 16 ± 1.9% for BNI.

3.3. Step 2: Transposition of the Horizontal Irradiances in the Plane of Array. To calculate G_{POA}, needed input data are GHI, DHI, and BNI solar irradiances as well as the tilt angle and the ground albedo. These three measured input are taken from SIRTA in situ ground measurements as the objective is to evaluate only step 2.

G_{POA} is calculated by the sum of the beam, diffuse, and albedo irradiances in the plane of array. The beam irradiance is calculated with geometric relation between the BNI and the angle of incidence AOI (angle between the sun's rays and the PV array). AOI can easily be computed by knowing the sun's position (solar azimuth and zenith angles) and the array's geometry (tilt and azimuth angles). The irradiance due to the ground albedo is calculated thanks to the view factor in front of the array, the celestial vault is assumed to be uniform, and the ground is supposed to have a

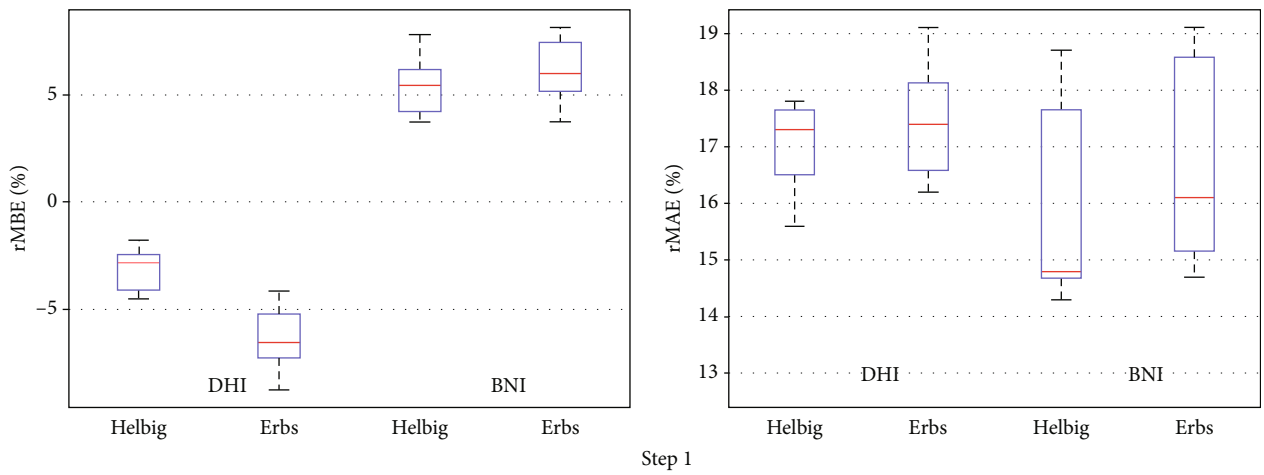


Figure 3: rMBE (left) and rMAE (right) of Helbig and Erbs models to calculate DHI and BNI from GHI measured at SIRTA. Computed data are summarized in Figure 9 of Appendix B.

Lambertian behavior. The coefficient of reflection of the ground is 0.2 in our case, as representative of grass.

Six models have been considered representing different ways to estimate the diffuse irradiance arriving on PV modules, as shown in Figure 1.

The isotropic sky diffuse model assumes that the solar radiosity of the celestial vault is uniformly distributed over the complete skydome [17]. It is the simplest of the tilt surface models. Under completely cloudy skies, this model is quite a good calculation. The five other models do not consider diffuse radiation as isotropic.

Klucher found that the isotropic model gave good results for overcast skies but underestimates irradiance under clear and partly overcast conditions, when there is increased intensity near the horizon and in the circumsolar region of the sky [18].

King model provides the portion of the total radiation incident on the surface that will be used by the module for energy production. It accounts for variation in the solar spectrum as a function of absolute air mass through an empirical function as well as for optical losses incurred by the angle of incidence. These two empirical functions were proposed by the Sandia Laboratory [19] and developed through outdoor testing of relevant modules [20]. Klucher modified this clear sky model by imposing a modulating factor which forces the anisotropic correction factor to approach unity under cloudy sky conditions so that the model reduces to the isotropic sky model.

Hay and Davies developed a model to predict the tilted surface diffuse radiation accounting for both circumsolar and isotropic diffuse radiation [21]. Realizing that the isotropic hypothesis becomes less accurate under clear sky conditions, they defined an anisotropy index which corresponds to a portion of the diffuse radiation to be treated as circumsolar with the remaining portion considered isotropic. The circumsolar diffuse is projected onto the tilted surface, and the remaining diffuse radiation is treated as isotropic diffuse. Under clear skies, the anisotropy index is high, and the circumsolar diffuse is more critical than the isotropic diffuse. Under cloudy skies, the anisotropy index goes to zero, and all diffuse is treated as isotropic. This model does not account for horizon brightening diffuse radiation.

In the model of Perez et al. [22], the contribution of diffuse radiation from circumsolar, isotropic, and horizontal regions is determined by two empirically derived coefficients [23]. The empirical coefficients are based on two years of data from Carpentras and Trappes, both in France. The empirical coefficient for the diffuse radiation from circumsolar is completed by measurements realized in five U.S. locations [24]. Perez et al. approximated the horizon brightening effects by applying a correction factor to the isotropic diffuse radiation.

In the reference [25], Reindl et al. investigate the performance of five models which estimate the hourly tilted surface radiation by comparing with measured energy on six experimental sites. He proposes to apply the horizon brightening correction factor used by Perez to the isotropic term in the Hay-Davies model.

rMAE and rMBE are calculated using SIRTA in situ ground measurements of G_{POA} compared to modeled one, using GHI, DHI, and BNI also measured at SIRTA.

As can be seen in Figure 4, Klucher model best performs regarding error (rMSE = 4, 7 ± 0, 6%) and also bias (rMSE = − 0, 1 ± 0, 4%).

3.4. Step 3: Estimation of PV Module Temperature. In this part, we evaluate four models to calculate T_{PV} from T_{amb}, G_{POA}, and WS. These three inputs are taken from SIRTA in situ ground measurements as the objective is to evaluate only step 3.

The nominal operating cell temperature (NOCT) is defined as the temperature reached by open-circuited cells in a module under a G_{POA} of 800, a T_{amb} of 20 ($T_{amb,NOCT}$), and a WS of 1. An approximate expression for calculating the module operating temperature is given in [26]. In our first calculations, we have considered a typical value of NOCT (48). In a second step, we fit the value of NOCT to one year of measurements (2015) using the Levenberg-Marquardt method. The obtained NOCT is equal to 45.

This model does not consider the impact of wind speed on module temperature, and we can imagine a nonproportional relation between them. This effect is included in the following

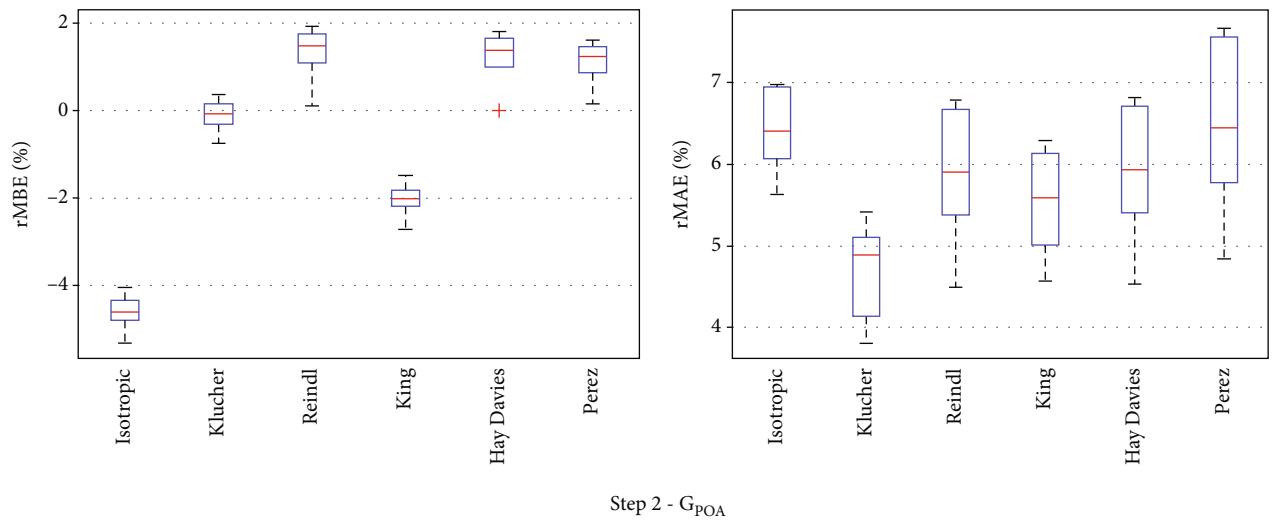


Figure 4: rMBE (left) and rMAE (right) of the presented models to calculate G_{POA} from GHI, DHI, and BNI measured at SIRTA. Computed data are summarized in Figure 9 of Appendix B.

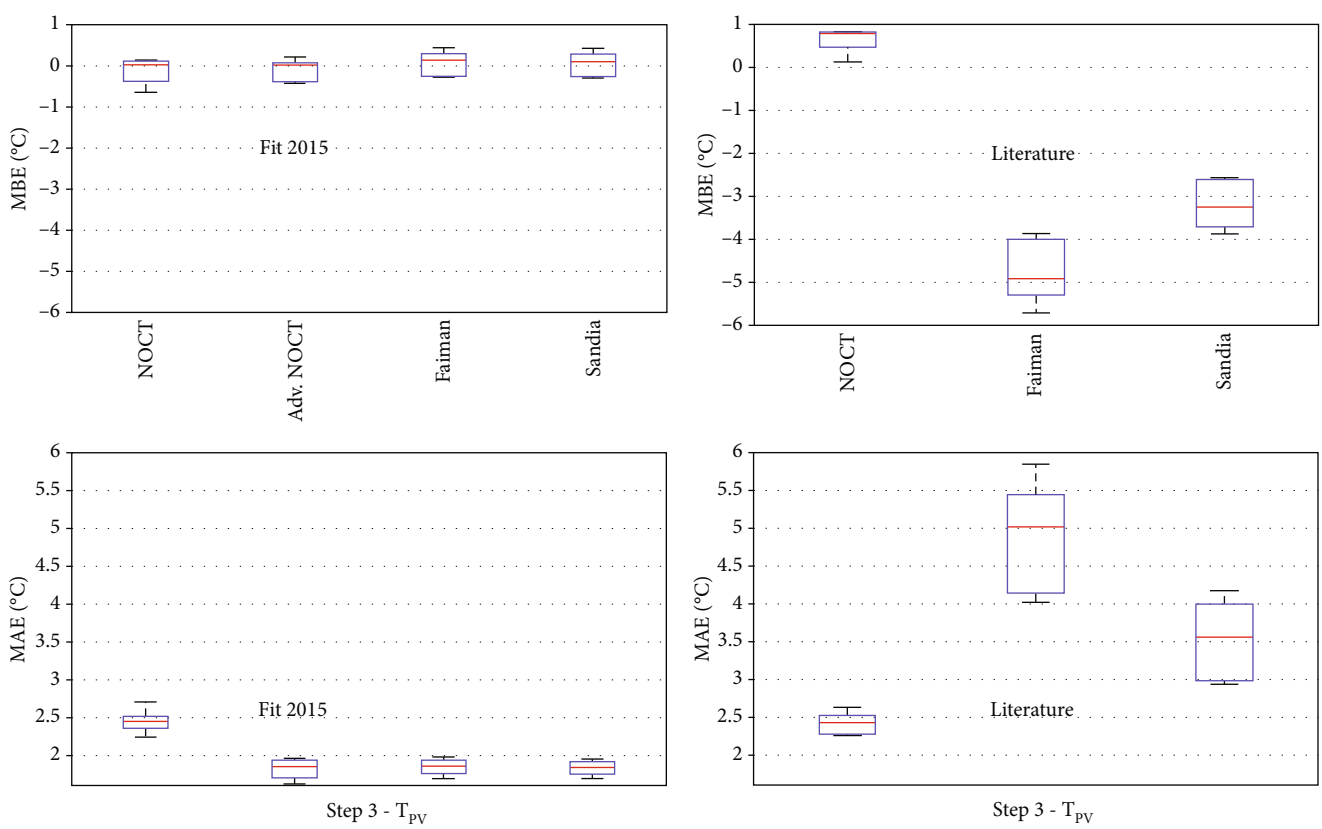


Figure 5: MBE (upper) and MAE (lower) of the presented models to calculate T_{PV} from GHI, T_{amb}, and WS, fitted coefficients (left) and literature coefficients (right). Computed data are summarized in Figure 9 of Appendix B.

improved NOCT formula that has been developed and named Advanced NOCT and is presented here [27]:

$$T_{PV} = T_{amb} + \frac{G_{POA,th}}{G_{POA,NOCT}} \times (a \cdot G_{POA,th} + b) \times \frac{\text{NOCT} - T_{amb,NOCT}}{1 + (\text{WS}/\text{WS}_{NOCT})^{c}}. \quad (3)$$

For module producing electricity, $G_{POA,th}$ is the fraction of energy converted into heat and is defined by the following equation:

$$G_{POA,th} = G_{POA} \times (1 - \text{CE}). \quad (4)$$

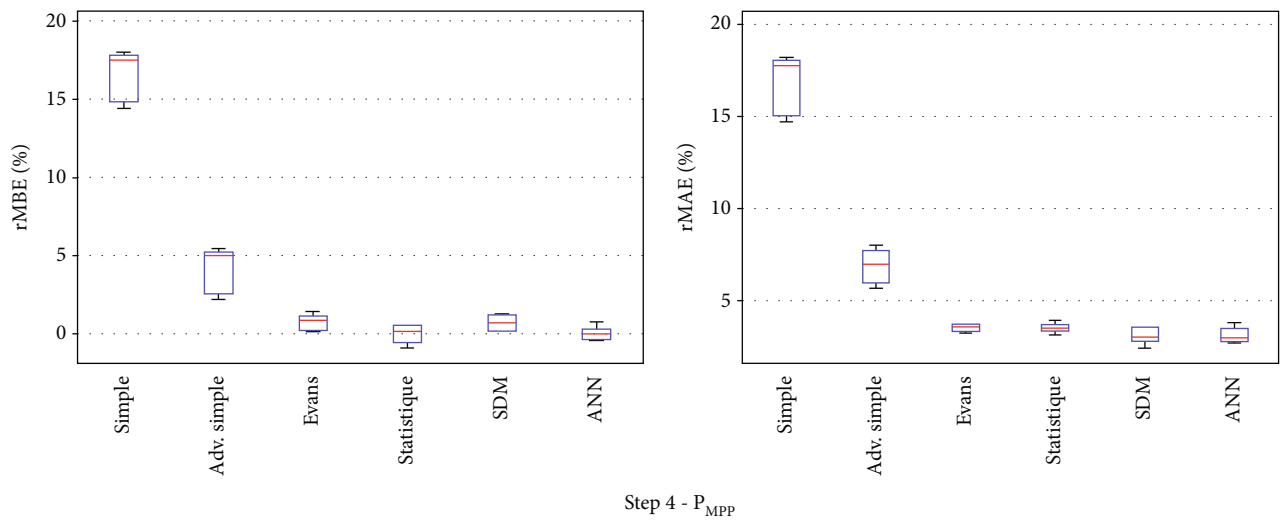


Figure 6: rMAE and rMBE of the presented models to calculate P_{MPP} from GHI and T_{PV}. Computed data are summarized in Figure 9 of Appendix B.

Table 3: Best results for each modeling step considering SIRTA in situ ground measurements. All computed data can be found in Figure 9 of Appendix B.

Step	Input	Model	Output	rMBE/MBE mean ± std	rMBE/MAE mean ± std
1	GHI	Helbig	DHI	−3.1 ± 1.1%	17.0 ± 0.9%
			BNI	5.4 ± 1.6%	16.0 ± 1.9%
2	GHI DHI BNI	Klucher	G_{POA}	−0.1 ± 0.4%	4.7 ± 0.6%
3	GHI T_{amb} WS	Adv. NOCT	T_{PV}	−0.1 ± 0.3	1.8 ± 0.1
4	G_{POA} T_{PV}	Evans	P_{MPP}	0.7 ± 0.5%	3.5 ± 0.2%

The coefficients a, b, and c are obtained by fitting with the Levenberg-Marquardt method, the value of T_{PV} measured during one year (2015), with the formula given in equation (3). We found that $a = 3.78e - 4$, $b = 2.29$, and $c = 0.474$. This study is in progress, in particular, to validate the location and technological universality of the proposed coefficients.

In the mid-1980s, a thermal model was developed at Sandia Laboratory for system engineering and performance modeling purposes [28]. Although rigorous, this early model has proven to be unnecessarily complex, not applicable to all PV technologies, and not easily adaptable to site-dependent influences. A simpler empirically-based thermal model was then developed at Sandia Laboratory [19]. This simple model has proven to be very adaptable and entirely adequate for system engineering and design purposes by providing the expected module operating temperature with an accuracy of about ±5. The empirical coefficients establish the upper limit for module temperature at low wind speeds and high solar irradiance, and the rate at which T_{PV} drops as WS increases, respectively. To evaluate this model, we firstly consider parameter values proposed by the authors ($a = -3.47$, $b = -0.0594$). In a second step, we fit these coefficients to one year of measurements (2015) using the Levenberg-Marquardt method. The obtained coefficients are $a = -3.1398$, $b = -0.305$.

Faiman presented a module temperature model [29] based on simple heat transfer concepts. To evaluate this model, we firstly consider empirical coefficient values proposed by the authors ($U_0 = 25.0$, $U_1 = 6.84$). In a second step, we fit these coefficients to one year of measurements (2015) using the Levenberg-Marquardt method. The obtained coefficients are $U_0 = 21.78$, $U_1 = 9.86$. Given the coefficients obtained for these last three models, Adv. NOCT is more sensitive to WS than the others.

MAE and MBE are calculated using SIRTA in situ ground measurements of T_{PV} compared to modeled one using G_{POA}, T_{amb}, and WS also measured at SIRTA.

As shown in Figure 5, the best performance is not obtained for coefficients taken directly from literature. Moreover, if you can do a data fitting during one year, all the four proposed methods have almost the same performance. For the rest of the study, we will choose the Adv

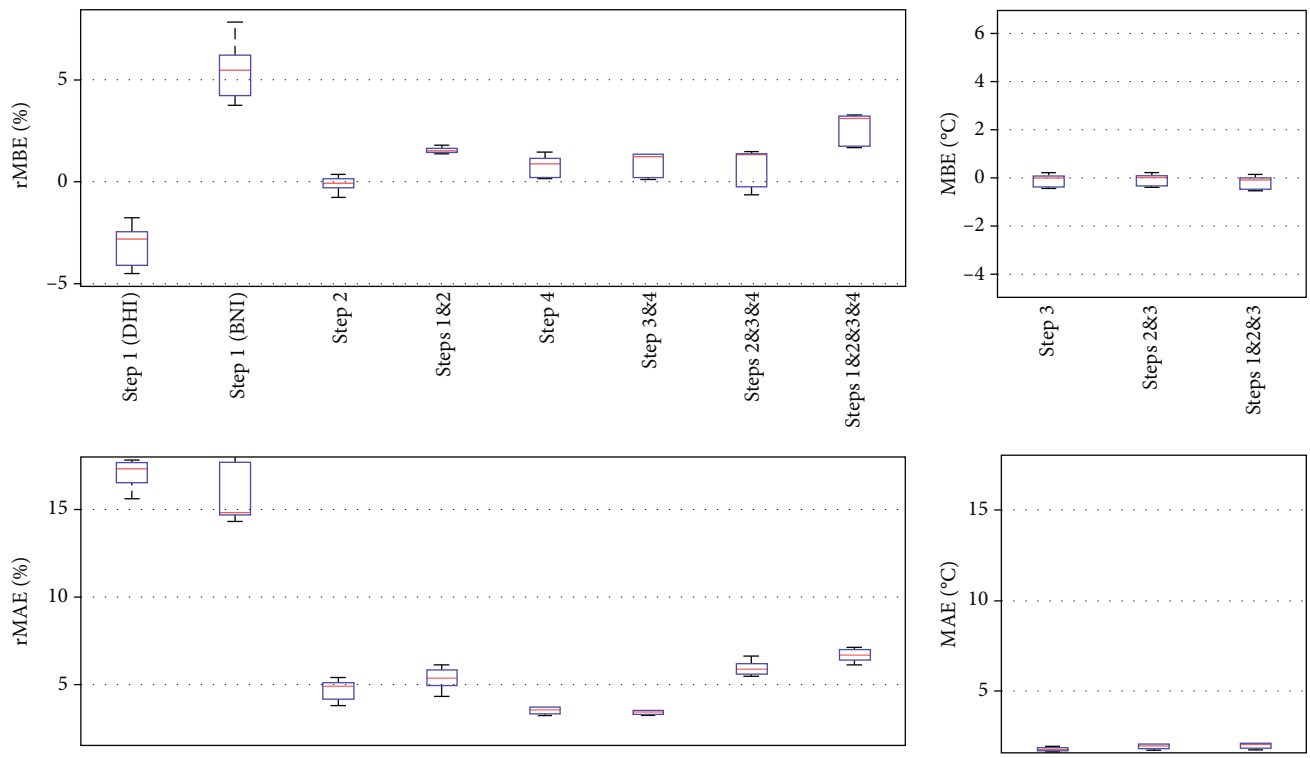


Figure 7: Modeling error estimation in each step of the simulation process for SIRTA measurements. The models are Helbig (step 1), Klucher (step 2), Adv. NOCT (step 3), and Evans (step 4). Computed data can be found in Figure 10 of Appendix C.

NOCT model because according to the coefficients found, it is a little more sensitive to WS. Its accuracy is MBE = −0.1 ± 0.3 and MAE = 1.8 ± 0.1.

3.5. Step 4: Photoelectric Conversion Models. In this study, we consider six different photoelectric conversion models [30] to calculate P_{MPP} from G_{POA} and T_{PV}. These two inputs are taken from SIRTA in situ measurements as the objective is to evaluate only step 4. Soiling effect and aging had not been taken into account.

The single point efficiency model, named here simple model, considers a constant conversion efficiency (CE) measured by the manufacturer during a flash test in standard test conditions (STC) [31]. For the studied PV module, $\mathrm{CE}_{\mathrm{STC}} = 15.28\%$.

We have improved this model by considering the conversion efficiency equal to the average measurements during 2015: $\mathrm{CE}_{2015} = 13.65\%$.

Evans model is a physics-based model that considers the linear variation of $\mathrm{CE}_{\mathrm{STC}}$ with module temperature and its exponential decreasing with the low light effect [32]. The temperature coefficient of the power is taken equal to -0.48% from the manufacturer datasheet and the low light effect coefficient is $\gamma = 0.12$, as advised by Evans in his article [32], for a crystalline silicon PV module.

The statistical model does not need internal information from the system to describe its performance. It is a data-driven approach that can extract relations on past historical data of couples (G_{POA}, T_{PV}) to predict the future behavior of the PV module (P_{MPP}) [33]. For a given hour h, in the year N, the statistical model looks for similar atmospheric conditions in 2015, that is to say, the ensemble of points $h_i \cap h_j$ respecting the following two conditions: $G_{\mathrm{POA}}(2015, h_i) = G_{\mathrm{POA}}(N, h) \pm I_G$ and $T_{\mathrm{PV}}(2015, h_j) = T_{\mathrm{PV}}(N, h) \pm I_T$. The algorithm starts with $I_G = 0$ and $I_T = 0$. If the ensemble of common points in illumination and temperature $(h_i \cap h_j)$ is empty, I_G and I_T intervals increase gradually by 5 compared to the target conditions, with a limit of 50. If several points in this ensemble are found, the given P_{MPP} at the output of the model is the average. Thus, the quality of the historical data is essential for an accurate simulation of this statistical model.

The single-diode electrical model (SDM) with 4 parameters is based on the Shockley diode equation [34], with a current source to model the photo-current, a single-diode junction, and a series resistance to model the contact losses. The photocurrent depends on T_{PV} and G_{POA}, and the diode saturation current and the series resistance are temperature dependent. The ideality factor of the diode is constant. All the parameters used in the above equation are determined by fitting the manufacturer flash test, and the STC modeling is compared in Table 2.

This 4-parameter SDM is more straightforward than the 5-parameter one and gives quite a good simulation accuracy.

The artificial neural network (ANN) was built using the feed-forward neural network structure with a weighted linear combination and sigmoid function [35, 36]. The architecture chosen is one output (P_{MPP}), two inputs (G_{POA} and T_{PV}), and one hidden layer. The training period was one year (2015), of which 2/3 of the data were used for the

Table 4: Equipment and sensors used for all the presented measurements.

Param.	Manufact.	Model	Std	Accuracy	Meas. uncert.
		PV module level measurements			
P_{MPP}	Agilent	6060B	IEC 60891	Electronic load: ±0.1% Voltmeter: 0.1% ± 300 mV Ammeter: 0.1% ± 350 mA	±7%
T_{PV}	TC	578-093 Class A	IEC 60751	[-20 °C; +100 C]: ±0.15°C 100.00 ± 0.06Ω at 0 C	±0.4 C $\lvert T_{cell} - T_{PV} \rvert \leq 1°C$
G_{POA}	Hukseflux Tektronix	SR01 2nd class 3706A	ISO 9060	±20% in hourly totals DC voltmeter: ±0.0025%	
		Atmospheric in situ measurements (SIRTA)			
GHI	Kipp & Zonen	CM22 2ry std	ISO 9060	Daily uncertainty ≤1%	
BNI	Kipp & Zonen	CHP1 2ry std	ISO 9060	Daily uncertainty ≤1%	
T_{amb}	Guilcor	Pt100 Class B	IEC 60751	[-100 °C; +100 C]: ±0.15 C 100.00 ± 0.06Ω at 0 C	±0.4 C
WS	Vector	A100R		[0.3 m/s; 10 m/s]: 0.1 m/s	
		Atmospheric measurements from nearby weather station (ORLY)			
GHI	Kipp & Zonen	CM6B 1 nd class	ISO 9060	Daily uncertainty ≤5%	
T_{amb}		Pt100 Class B	IEC 60751	[-100 °C; +100 C]: ±0.15°C	±0.4°C
WS		Class B		100.00 ± 0.06Ω at 0 °C 10%or 0.3 m/s	
		Geostationary satellite observations (CAMS)			
GHI, DHI, BNI	CAMS			Pixel diameter ≈4.5 km	

training and 1/3 for the validation. Dataset for the test is each year from 2016 to 2020.

rMBE and rMAE are calculated using SIRTA in situ measurements of P_{MPP} compared to modeled one using G_{POA} and T_{PV} also measured at SIRTA.

As can be seen in Figure 6, except for simple model, all of them have almost the same performance. SDM can be improved by either implement the 5-parameter SDM or better taking into account the dependence of the parameters with the temperature. The double-diode model could also be used to be more accurate at low irradiances. Moreover, SDM is tricky to tune because parameters are not given in the datasheet, and flash tests are not always available. Statistical and ANN needs a historical dataset to train the model. For these reasons, we decided to select the Evans model as the best for the following study, with rMBE = 0.8 ± 0.5% and rMAE = 3.5 ± 0.2%.

To summarize, this theoretical study shows that best performances are obtained, for each step, by using the Helbig model for step 1, the Klucher model for step 2, the Adv. NOCT model for step 3, and the Evans model for step 4 (see Figure 1). The modeling step with less accuracy is step 1. All those results are gathered in Table 3.

4. Results: Error in the Simulation of the P_{MPP} from Step to Step

In this section, we firstly study the uncertainty of each successive step, just for SIRTA in situ ground-based measurements, in order to evaluate the propagation of the error and the most uncertain simulation stem.

In the second part, we compare different ways to reach an estimation of P_{MPP}, with decreasing the difficulty of accessing data (that is to say, from SIRTA to ORLY and CAMS input data) and increasing the number of modeling steps (from only step 4 to all steps 1 to 4), as we can see in Figure 1.

4.1. Error Propagation in the Simulation of P_{MPP} from Step to Step. First, let us study the uncertainty of each step, just for SIRTA in situ ground-based measurements, in order to evaluate the propagation of the error and the most uncertain one. Figure 7 summarizes and completes what has been obtained in previous sections. All computed data can be found in Figure 10 of Appendix C.

First of all, this study confirms that the most uncertain step is the first one, to estimate the diffuse fraction (17.0%). Nevertheless, if step 2 (4.7%) is associated to step 1, having in mind

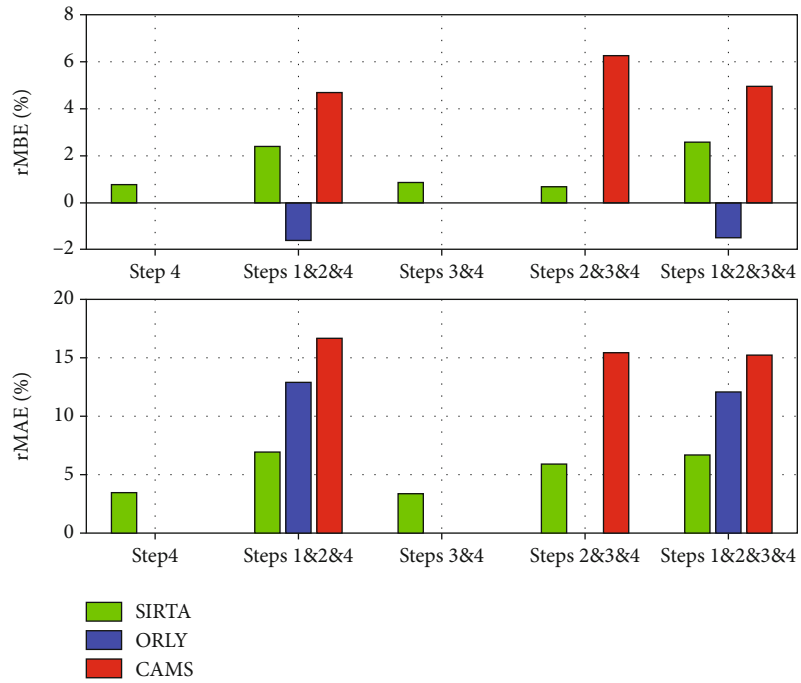


Figure 8: Modeling mean annual rMBE (top) and mean annual rMAE (bottom) of the whole simulation process as a function of the accessibility of the data sources (SIRTA, ORLY, and CAMS). The models are Helbig (step 1), Klucher (step 2), Adv. NOCT (step 3), and Evans (step 4).

that the critical variable is G_{POA} and not DHI, the error decreases down to 5.3%. This error is of the same order of magnitude as the precision of measuring instruments.

The estimation of T_{PV} is not a critical step. The associated uncertainty is lower than 2 (right part of Figure 7). This error must be separated from that of the other steps because they are not of the same type. Moreover, the impact of the T_{PV} modeling error does not linearly impact the modeling error of P_{MPP}.

In addition, the estimation of T_{PV} from environmental measurements is not critical in the simulation of P_{MPP} because the error whatever T_{PV} is measured (step 4 or steps 1 and 2 and 4) or estimated (steps 3 and 4 or steps 1 and 2 and 3 and 4) remains around 3.5% or 6.9%. On the other hand, the computation of G_{POA} has a very strong sensitivity on the simulation of P_{MPP}, confirming what has already been said before. The error goes from 3.5% with G_{POA} measured (step 4 or steps 3 and 4) to 6.9% when it is calculated (steps 1 and 2 and 4 or steps 1 and 2 and 3 and 4).

Moreover, we can see with this study that the errors almost accumulate, if steps 1 and 2 are considered together: the total error in P_{MPP} modeling (steps 1 to 4) is 6.7%, and the mean squared error of steps 1 and 2 and step 4 is 6.4%.

4.2. Error in the Simulation of P_{MPP} according to Data Source. In this part, different ways to reach an estimation of P_{MPP}, with decreasing the difficulty of accessing data (that is to say, from SIRTA to ORLY and CAMS input data) and increasing the number of modeling steps (from only step 4 to all steps 1 to 4) are studied, as can be seen in Figure 1.

Figure 8 presents the results of the global simulation, for all the data sources. The chosen models are the same as for the previous study. The differences between all the simulations are the data sources (in situ ground-based measurements SIRTA, ORLY weather station ground-based measurements, and CAMS satellite-based estimations).

This figure shows whatever the data sources, and the trends are the same: (i) modeling P_{MPP} with GHI is enough to have good performance compared to using GHI, DHI, and BNI; (ii) T_{PV} modeling hardly generates any error in the estimation of P_{MPP}.

Moreover, the less accessible the data are, the more accurate the simulation is. The best results are obtained for SIRTA in situ ground-based measurements with a rMAE of 6.7% for the whole simulation chain. Then, the nearest weather station ORLY (12.1%) and the worth are given by CAMS satellite images estimations (15.2%). This is mainly due to the fact that ground-based measurements are not biased but satellite-based estimations are. It can be explained by, on the one hand, the pixels size of the satellite images, and also, on the other hand, the uncertainties related to the estimation of GHI, DHI and BNI, mostly in overcast conditions.

5. Conclusion

In this paper, we have studied, step by step, the simulation process of the PV DC energy production, with a special focus on the calculation of the uncertainty.

Four steps compose the simulation chain, and basic and well-known models were considered.

(1) Step 1 ⟶ from GHI to GHI, DHI, and BNI: 2 models

Variable	Source	Step 1	Step 2	Year	rMBE	rMAE	Mean rMBE	Std rMBE	Mean rMAE	Std rMAE
DHI	SIRTA	helbig		2016	−2,82	16,80	−3,148	1,089	17,020	0,879
DHI	SIRTA	helbig		2017	−3,96	15,60				
DHI	SIRTA	helbig		2018	−2,68	17,30				
DHI	SIRTA	helbig		2019	−4,51	17,80				
DHI	SIRTA	helbig		2020	−1,77	17,60				
DHI	SIRTA	erbs		2016	−5,59	16,70	−6,368	1,696	17,440	1,115
DHI	SIRTA	erbs		2017	−6,79	16,20				
DHI	SIRTA	erbs		2018	−6,56	17,40				
DHI	SIRTA	erbs		2019	−8,76	19,10				
DHI	SIRTA	erbs		2020	−4,14	17,80				
BNI	SIRTA	helbig		2016	5,66	18,70	5,408	1,557	15,980	1,920
BNI	SIRTA	helbig		2017	7,81	17,30				
BNI	SIRTA	helbig		2018	4,37	14,80				
BNI	SIRTA	helbig		2019	5,46	14,80				
BNI	SIRTA	helbig		2020	3,74	14,30				
BNI	SIRTA	erbs		2016	6,00	19,10	6,148	1,678	16,720	1,934
BNI	SIRTA	erbs		2017	8,14	18,40				
BNI	SIRTA	erbs		2018	5,64	15,30				
BNI	SIRTA	erbs		2019	7,23	16,10				
BNI	SIRTA	erbs		2020	3,73	14,70				
GPOA	SIRTA		Isotrophic	2016	−4,62	6,93	−4,606	0,460	6,428	0,555
GPOA	SIRTA		Isotrophic	2017	−5,32	6,97				
GPOA	SIRTA		Isotrophic	2018	−4,65	6,4				
GPOA	SIRTA		Isotrophic	2019	−4,44	5,63				
GPOA	SIRTA		Isotrophic	2020	−4,6	6,21				
GPOA	SIRTA		Klucher	2016	−0,164	5,41	−0,108	0,413	4,672	0,637
GPOA	SIRTA		Klucher	2017	−0,754	5				
GPOA	SIRTA		Klucher	2018	0,363	4,89				
GPOA	SIRTA		Klucher	2019	0,0933	3,81				
GPOA	SIRTA		Klucher	2020	−0,0774	4,25				
GPOA	SIRTA		Reindl	2016	1,7	6,78	1,325	0,709	5,894	0,915
GPOA	SIRTA		Reindl	2017	1,92	6,63				
GPOA	SIRTA		Reindl	2018	1,48	5,9				
GPOA	SIRTA		Reindl	2019	0,107	4,49				
GPOA	SIRTA		Reindl	2020	1,42	5,67				
GPOA	SIRTA		King	2016	−2,01	6,29	−2,034	0,441	5,538	0,697
GPOA	SIRTA		King	2017	−2,72	6,08				
GPOA	SIRTA		King	2018	−1,49	5,59				
GPOA	SIRTA		King	2019	−1,93	4,57				
GPOA	SIRTA		King	2020	−2,02	5,16				
GPOA	SIRTA		Haydavies	2016	1,59	6,81	1,221	0,711	5,926	0,914
GPOA	SIRTA		Haydavies	2017	1,81	6,67				
GPOA	SIRTA		Haydavies	2018	1,38	5,93				
GPOA	SIRTA		Haydavies	2019	−0,00338	4,53				
GPOA	SIRTA		Haydavies	2020	1,33	5,69				
GPOA	SIRTA		Perez	2016	1,6	7,52	1,102	0,564	6,508	1,153
GPOA	SIRTA		Perez	2017	1,42	7,66				
GPOA	SIRTA		Perez	2018	1,24	6,44				
GPOA	SIRTA		Perez	2019	0,15	4,84				
GPOA	SIRTA		Perez	2020	1,1	6,08				
GPOA	SIRTA	helbig	Isotropic	2016	−3,26	6,6	−3,132	0,189	5,924	0,590
GPOA	SIRTA	helbig	Isotropic	2017	−3,24	6,05				
GPOA	SIRTA	helbig	Isotropic	2018	−2,83	5,9				
GPOA	SIRTA	helbig	Isotropic	2019	−3,06	4,98				
GPOA	SIRTA	helbig	Isotropic	2020	−3,27	6,09				
GPOA	SIRTA	helbig	Klucher	2016	1,44	6,12	1,534	0,155	5,326	0,677
GPOA	SIRTA	helbig	Klucher	2017	1,39	5,71				
GPOA	SIRTA	helbig	Klucher	2018	1,79	5,36				
GPOA	SIRTA	helbig	Klucher	2019	1,55	4,32				
GPOA	SIRTA	helbig	Klucher	2020	1,5	5,12				
GPOA	SIRTA	helbig	Reindl	2016	3,57	7,9	3,222	0,925	6,756	1,242
GPOA	SIRTA	helbig	Reindl	2017	4,35	7,83				
GPOA	SIRTA	helbig	Reindl	2018	3,15	6,48				
GPOA	SIRTA	helbig	Reindl	2019	1,8	4,85				
GPOA	SIRTA	helbig	Reindl	2020	3,24	6,72				
GPOA	SIRTA	helbig	King	2016	−0,648	6,35	−0,558	0,168	5,490	0,732
GPOA	SIRTA	helbig	King	2017	−0,632	5,81				
GPOA	SIRTA	helbig	King	2018	−0,271	5,47				
GPOA	SIRTA	helbig	King	2019	−0,55	4,35				
GPOA	SIRTA	helbig	King	2020	−0,687	5,47				
GPOA	SIRTA	helbig	Haydavies	2016	3,47	7,9	3,128	0,928	6,756	1,242
GPOA	SIRTA	helbig	Haydavies	2017	4,26	7,83				
GPOA	SIRTA	helbig	Haydavies	2018	3,06	6,48				
GPOA	SIRTA	helbig	Haydavies	2019	1,7	4,85				
GPOA	SIRTA	helbig	Haydavies	2020	3,15	6,72				
GPOA	SIRTA	helbig	Perez	2016	2,19	7,8	1,996	0,573	6,858	1,110
GPOA	SIRTA	helbig	Perez	2017	2,76	7,97				
GPOA	SIRTA	helbig	Perez	2018	1,97	6,79				
GPOA	SIRTA	helbig	Perez	2019	1,17	5,22				
GPOA	SIRTA	helbig	Perez	2020	1,89	6,51				

Variable	Source	Coef source	Step 3	Year	MBE	MAE	Mean MBE	Std MBE	Mean MAE	Std MAE
TPV	SIRTA	Fit	noct	2016	0,10	2,24	−0,143	0,331	2,448	0,165
TPV	SIRTA	Fit	noct	2017	0,12	2,45				
TPV	SIRTA	Fit	noct	2018	−0,65	2,40				
TPV	SIRTA	Fit	noct	2019	−0,30	2,70				
TPV	SIRTA	Fit	noct	2020	0,02	2,45				
TPV	SIRTA	Fit	advnoct	2016	0,20	1,63	−0,118	0,279	1,820	0,139
TPV	SIRTA	Fit	advnoct	2017	0,02	1,73				
TPV	SIRTA	Fit	advnoct	2018	−0,44	1,93				
TPV	SIRTA	Fit	advnoct	2019	−0,39	1,96				
TPV	SIRTA	Fit	advnoct	2020	0,01	1,85				
TPV	SIRTA	Fit	Faiman	2016	0,43	1,70	0,054	0,310	1,848	0,111
TPV	SIRTA	Fit	Faiman	2017	0,25	1,78				
TPV	SIRTA	Fit	Faiman	2018	−0,26	1,92				
TPV	SIRTA	Fit	Faiman	2019	−0,27	1,98				
TPV	SIRTA	Fit	Faiman	2020	0,13	1,86				
TPV	SIRTA	Fit	Sandia	2016	0,41	1,70	0,035	0,310	1,832	0,100
TPV	SIRTA	Fit	Sandia	2017	0,23	1,77				
TPV	SIRTA	Fit	Sandia	2018	−0,26	1,90				
TPV	SIRTA	Fit	Sandia	2019	−0,30	1,95				
TPV	SIRTA	Fit	Sandia	2020	0,10	1,84				
TPV	SIRTA	Literature	noct	2016	0,79	2,29	0,626	0,296	2,420	0,151
TPV	SIRTA	Literature	noct	2017	0,81	2,49				
TPV	SIRTA	Literature	noct	2018	0,13	2,26				
TPV	SIRTA	Literature	noct	2019	0,57	2,63				
TPV	SIRTA	Literature	noct	2020	0,83	2,43				
TPV	SIRTA	Literature	Faiman	2016	−3,87	4,03	−4,734	0,769	4,874	0,763
TPV	SIRTA	Literature	Faiman	2017	−4,04	4,18				
TPV	SIRTA	Literature	Faiman	2018	−5,15	5,30				
TPV	SIRTA	Literature	Faiman	2019	−5,70	5,84				
TPV	SIRTA	Literature	Faiman	2020	−4,91	5,02				
TPV	SIRTA	Literature	Sandia	2016	−2,56	2,94	−3,188	0,591	3,522	0,549
TPV	SIRTA	Literature	Sandia	2017	−2,62	3,00				
TPV	SIRTA	Literature	Sandia	2018	−3,65	3,94				
TPV	SIRTA	Literature	Sandia	2019	−3,87	4,17				
TPV	SIRTA	Literature	Sandia	2020	−3,24	3,56				

Variable	Source	Step 4	Year	rMBE	rMAE	Mean rMBE	Std rMBE	Mean rMAE	Std rMAE
PMPP	SIRTA	ANN	2016	−0,02	2,77	0,031	0,473	3,127	0,456
PMPP	SIRTA	ANN	2017	−0,39	3,37				
PMPP	SIRTA	ANN	2018	−0,35	3,81				
PMPP	SIRTA	ANN	2019	0,14	2,74				
PMPP	SIRTA	ANN	2020	0,77	2,95				
PMPP	SIRTA	Evans	2016	0,24	3,34	0,751	0,545	3,509	0,214
PMPP	SIRTA	Evans	2017	0,16	3,71				
PMPP	SIRTA	Evans	2018	0,87	3,55				
PMPP	SIRTA	Evans	2019	1,04	3,23				
PMPP	SIRTA	Evans	2020	1,45	3,71				
PMPP	SIRTA	Simple	2016	15,00	15,15	16,565	1,718	16,757	1,687
PMPP	SIRTA	Simple	2017	14,42	14,70				
PMPP	SIRTA	Simple	2018	18,06	18,18				
PMPP	SIRTA	Simple	2019	17,57	17,75				
PMPP	SIRTA	Simple	2020	17,79	18,00				
PMPP	SIRTA	Adv. simple	2016	2,71	6,04	4,115	1,534	6,850	0,997
PMPP	SIRTA	Adv. simple	2017	2,20	5,65				
PMPP	SIRTA	Adv. simple	2018	5,45	7,61				
PMPP	SIRTA	Adv. simple	2019	5,01	6,96				
PMPP	SIRTA	Adv. simple	2020	5,21	7,99				
PMPP	SIRTA	Statistque	2016	−0,43	3,48	−0,024	0,634	3,513	0,284
PMPP	SIRTA	Statistque	2017	−0,91	3,63				
PMPP	SIRTA	Statistque	2018	0,54	3,90				
PMPP	SIRTA	Statistque	2019	0,15	3,13				
PMPP	SIRTA	Statistque	2020	0,54	2,43				
PMPP	SIRTA	SDM	2016	0,16	3,02	0,704	0,540	3,084	0,470
PMPP	SIRTA	SDM	2017	1,22	3,54				
PMPP	SIRTA	SDM	2018	1,27	3,53				
PMPP	SIRTA	SDM	2019	0,70	2,91				
PMPP	SIRTA	SDM	2020	0,17	2,42				

Figure 9: Computed rMBE and rMAE for SIRTA in situ ground-based measurements, for each modeling step. Colors are performance indicators (from green for the best to red for the worse).

(2) Step 2 ⟶ from GHI, DHI, and BNI to G_{POA}: 6 models

(3) Step 3 ⟶ from G_{POA}, T_{amb}, and WS to T_{PV}: 4 models in 2 configurations

(4) Step 4 ⟶ from G_{POA} and T_{PV} to P_{MPP}: 6 models

Step 1 is the one that leads to the greatest uncertainty (17%). As the output variables of this step are only intermediate variables, if we combine steps 1 and 2, then, the error drops to 5.3%. Moreover, it is the estimation of G_{POA}, the step which induces the greatest error (1.8 for step 3 and 3.5% for step 4).

This remark remains valid if we are interested in the estimation of P_{MPP}. Indeed, the modifications from measured G_{POA} (step 4) to estimated G_{POA} (steps 1 and 2 and 4) induces an increase of the error from 3.5% to 6.9%.

The temperature does not have a great influence in the estimation of P_{MPP}. Whether estimated (steps 3 and 4) or measured (step 4), the error remains stable around 3.5%.

For the global simulation, our study was interested in the performances obtained according to 3 sources of environmental data (GHI, T_{amb}, and WS). As might be expected, the closer this data is to the PV installation, the better the performance is. For example, for local measurements, the error is 6.6%. If data from the nearest weather station is used, the error increases to 12.1%, and for data estimated from satellite images, then, the error becomes 15.2%.

In conclusion, the proposed study makes it possible to evaluate the simulation quality of the P_{MPP} of a PV installation. Each step was evaluated independently and successively, according to 3 data sources. All the computed data are presented in Appendix B, and the raw data are available upon request.

The next steps will be to show the link between all the models and the error propagation, consider different meteorological forecast models, consider different time horizons, go through PV plants instead of a unique PV module, and compare with commercial tools, for instance.

Input	Source	Steps				Year	rMBE	rMAE	Mean rMBE	std rMBE	Mean rMAE	std rMAE
GHI, Tamb, WS	SIRTA	✓	✓	✓	✓	2016	1,74	6,49	2,602	0,804	6,663	0,394
GHI, Tamb, WS	SIRTA	✓	✓	✓	✓	2017	1,71	6,67				
GHI, Tamb, WS	SIRTA	✓	✓	✓	✓	2018	3,22	7,12				
GHI, Tamb, WS	SIRTA	✓	✓	✓	✓	2019	3,08	6,10				
GHI, Tamb, WS	SIRTA	✓	✓	✓	✓	2020	3,26	6,92				
GHI, Tamb, WS	SIRTA	✓	✓	✓	✓	2016	−1,71	16,32	−1,465	1,352	12,089	6,984
GHI, Tamb, WS	ORLY	✓	✓	✓	✓	2017	−2,87	15,80				
GHI, Tamb, WS	ORLY	✓	✓	✓	✓	2018	0,38	14,57				
GHI, Tamb, WS	ORLY	✓	✓	✓	✓	2019	−1,67	1,67				
GHI, Tamb, WS	ORLY	✓	✓	✓	✓	2020						
GHI, Tamb, WS	CAMS	✓	✓	✓	✓	2016	9,35	18,40	4,959	2,554	15,164	6,984
GHI, Tamb, WS	CAMS	✓	✓	✓	✓	2017	4,21	16,42				
GHI, Tamb, WS	CAMS	✓	✓	✓	✓	2018	4,51	14,78				
GHI, Tamb, WS	CAMS	✓	✓	✓	✓	2019	4,06	12,73				
GHI, Tamb, WS	CAMS	✓	✓	✓	✓	2020	2,67	13,49				
GHI, TPV	SIRTA	✓	✓		✓	2016	1,79	6,94	2,407	0,670	6,947	0,447
GHI, TPV	SIRTA	✓	✓		✓	2017	1,61	7,00				
GHI, TPV	SIRTA	✓	✓		✓	2018	2,81	7,37				
GHI, TPV	SIRTA	✓	✓		✓	2019	2,69	6,21				
GHI, TPV	SIRTA	✓	✓		✓	2020	3,13	7,22				
GHI, TPV	ORLY	✓	✓		✓	2016	−1,52	17,52	−1,588	1,117	12,894	7,538
GHI, TPV	ORLY	✓	✓		✓	2017	−2,95	16,84				
GHI, TPV	ORLY	✓	✓		✓	2018	−0,22	15,57				
GHI, TPV	ORLY	✓	✓		✓	2019	−1,65	1,65				
GHI, TPV	ORLY	✓	✓		✓	2020	NaN	NaN				
GHI, TPV	CAMS	✓	✓		✓	2016	9,76	19,88	4,655	2,937	16,558	7,538
GHI, TPV	CAMS	✓	✓		✓	2017	3,90	17,89				
GHI, TPV	CAMS	✓	✓		✓	2018	3,82	16,15				
GHI, TPV	CAMS	✓	✓		✓	2019	3,62	14,20				
GHI, TPV	CAMS	✓	✓		✓	2020	2,18	14,67				
GHI, DHI, BNI, Tamb, WS	SIRTA		✓	✓	✓	2016	−0,13	5,65	0,676	0,992	5,922	0,443
GHI, DHI, BNI, Tamb, WS	SIRTA		✓	✓	✓	2017	−0,65	6,01				
GHI, DHI, BNI, Tamb, WS	SIRTA		✓	✓	✓	2018	1,48	6,62				
GHI, DHI, BNI, Tamb, WS	SIRTA		✓	✓	✓	2019	1,34	5,46				
GHI, DHI, BNI, Tamb, WS	SIRTA		✓	✓	✓	2020	1,34	5,87				
GHI, DHI, BNI, Tamb, WS	ORLY		✓	✓	✓	2016						
GHI, DHI, BNI, Tamb, WS	ORLY		✓	✓	✓	2017						
GHI, DHI, BNI, Tamb, WS	ORLY		✓	✓	✓	2018						
GHI, DHI, BNI, Tamb, WS	ORLY		✓	✓	✓	2019						
GHI, DHI, BNI, Tamb, WS	ORLY		✓	✓	✓	2020						
GHI, DHI, BNI, Tamb, WS	CAMS		✓	✓	✓	2016	10,80	18,80	6,256	2,651	15,380	NaN
GHI, DHI, BNI, Tamb, WS	CAMS		✓	✓	✓	2017	5,90	16,80				
GHI, DHI, BNI, Tamb, WS	CAMS		✓	✓	✓	2018	5,72	15,00				
GHI, DHI, BNI, Tamb, WS	CAMS		✓	✓	✓	2019	4,87	12,90				
GHI, DHI, BNI, Tamb, WS	CAMS		✓	✓	✓	2020	3,99	13,40				
GPOA, Tamb, WS	SIRTA			✓	✓	2016	0,11	3,36	0,849	0,630	3,420	0,135
GPOA, Tamb, WS	SIRTA			✓	✓	2017	0,21	3,52				
GPOA, Tamb, WS	SIRTA			✓	✓	2018	1,23	3,53				
GPOA, Tamb, WS	SIRTA			✓	✓	2019	1,35	3,21				
GPOA, Tamb, WS	SIRTA			✓	✓	2020	1,34	3,48				
GPOA, Tamb, WS	ORLY			✓	✓	2016						
GPOA, Tamb, WS	ORLY			✓	✓	2017						
GPOA, Tamb, WS	ORLY			✓	✓	2018						
GPOA, Tamb, WS	ORLY			✓	✓	2019						
GPOA, Tamb, WS	ORLY			✓	✓	2020						
GPOA, Tamb, WS	CAMS			✓	✓	2016						
GPOA, Tamb, WS	CAMS			✓	✓	2017						
GPOA, Tamb, WS	CAMS			✓	✓	2018						
GPOA, Tamb, WS	CAMS			✓	✓	2019						
GPOA, Tamb, WS	CAMS			✓	✓	2020						
GPOA, Tpv	SIRTA				✓	2016	0,24	3,34	0,751	0,545	3,509	0,214
GPOA, Tpv	SIRTA				✓	2017	0,16	3,71				
GPOA, Tpv	SIRTA				✓	2018	0,87	3,55				
GPOA, Tpv	SIRTA				✓	2019	1,04	3,22				
GPOA, Tpv	SIRTA				✓	2020	1,34	3,71				
GPOA, Tpv	SIRTA				✓	2016						
GPOA, Tpv	ORLY				✓	2017						
GPOA, Tpv	ORLY				✓	2018						
GPOA, Tpv	ORLY				✓	2019						
GPOA, Tpv	ORLY				✓	2020						
GPOA, Tpv	CAMS				✓	2016						
GPOA, Tpv	CAMS				✓	2017						
GPOA, Tpv	CAMS				✓	2018						
GPOA, Tpv	CAMS				✓	2019						
GPOA, Tpv	CAMS				✓	2020						

Figure 10: Computed rMBE and rMAE for SIRTA in situ ground-based measurements, ORLY weather station ground-based measurements, and CAMS satellites images, for each modeling step. Colors are performance indicators (from green for the best to red for the worse).

Appendix

A. Annex 1 Data acquisition equipment and sensors (Table 4).

B. Annex 2 Computed data used in Section 2 (Figure 9).

C. Annex 3 Computed data used in Section 3 (Figure 10).

Abbreviations

AOI: Angle of incidence
BNI: Hourly beam (direct) normal irradiance (Wm^{-2})
CE: Conversion efficiency
DHI: Hourly diffuse horizontal irradiance (Wm^{-2})
GHI: Hourly global horizontal irradiance (Wm^{-2})
G_{POA}: Hourly global plane of array irradiance (Wm^{-2})
MAE: Mean absolute error
MBE: Mean bias error
NOCT: Nominal operating cell temperature (°C)
P_{MPP}: Hourly power at the maximum power point (W)
T_{PV}: Hourly PV module temperature (°C)
T_{amb}: Hourly ambient temperature (°C)
WS: Hourly wind speed (ms^{-1})
k_T: Hourly clearness index
rMAE: Relative mean absolute error
rMBE: Relative mean bias error
ANN: Artificial neurons network
BSRN: Baseline surface radiation network
CAMS: Copernicus atmosphere monitoring service
ORLY: Ground-based measurements weather station
PV: Photovoltaic
PVUSA: Photovoltaics for utility scale applications
SDM: Single-diode electrical model
SIRTA: Site Instrumental de Recherche par Télédétection Atmosphérique (Atmospheric Research Observatory, in English)
STC: Standard test conditions.

Data Availability

Atmospherical data are available with free access: -from satellite images: http://www.soda-pro.com/web-services/radiation/cams-radiation-service-from meteorological weather stations: https://donneespubliques.meteofrance.fr/-from in situ measurements: https://bsrn.awi.de. Photovoltaic data used to support the results of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work benefited from the support of the Energy4Climate Interdisciplinary Center (E4C) of IP Paris and École des Ponts ParisTech. It was supported by 3rd Programme d'Investissements d'Avenir (ANR-18-EUR-0006-02) and by the Foundation of École polytechnique (Chaire "Defis Technologiques pour une Energie Responsable" financed by total).

References

[1] E. Ela, V. Diakov, E. Ibanez, and M. Heaney, *Impacts of variability and un certainty in solar photovoltaic generation at multiple timescales, tech. Rep. NREL/TP-5500-58274*, National Renewable Energy Laboratory, 2013.

[2] M. Rekinger, I.-T. Theologitis, G. Masson et al., "Connecting the sun," in *solar photovoltaic on the road to largescale grid integrationt Tech. rep*, EPIA, 2012.

[3] B. Barth, G. Concas, E. B. Zane et al., *PV Grid : Final Project Report*, BSW-Solar, 2014.

[4] D. Thevenard and S. Pelland, "Estimating the uncertainty in long-term photovoltaic yield predictions," *Solar Energy*, vol. 91, pp. 432–445, 2013.

[5] M. Schnitzer, *Reducing uncertainty in solar energy estimates: a case study*, Sandia PV Performance Modeling Workshop, Santa Clara, Canada, 2013.

[6] B. Müller, L. Hardt, A. Armbruster, K. Kiefer, and C. Reise, "Yield predictions for photovoltaic power plants: empirical validation, recent advances and remaining uncertainties," *Progress in Photovoltaics: Research and Applications*, vol. 24, no. 4, pp. 570–583, 2016.

[7] G. Makrides, B. Zinsser, M. Schubert, and G. E. Georghiou, "Energy yield prediction errors and uncertainties of different photovoltaic models," *Progress in Photovoltaics : Research and Applications*, vol. 21, no. 4, pp. 500–516, 2013.

[8] G. Makrides, B. Zinsser, A. Phinikarides et al., "Photovoltaic model uncertainties based on field measurements," in *2011 37th IEEE Photovoltaic Specialists Conference*, pp. 002386–002390, Seattle, WA, USA, 2011.

[9] A. Migan, F. Calderon Obaldia, J. Badosa, V. Bourdin, Y. Bonnassieux, and L. Dumbia, "Step-by-step evaluation of photovoltaic module performance related to outdoor parameters : evaluation of the uncertainty," in *2017 IEEE 44th Photovoltaic Specialist Conference (PVSC)*, pp. 626–631, Washington, USA, 2017.

[10] M. Haeffelin, L. Barthès, O. Bock et al., "SIRTA, a ground-based atmospheric observatory for cloud and aerosol research," *Annali di geofisica (Bologna, Italy : 1993)*, vol. 23, no. 2, pp. 253–275, 2005.

[11] D. Chigara, *Etudes de l'échange thermique des modules photovoltaïques: mise en place d'un protocole de calibration de sondes platine pour la mesure de température et thermographie infrarouge, Tech. rep.*, LGEP, 2014.

[12] "Baseline Surface Radiation Network (BSRN)," 2018, http://bsrn.awi.de.

[13] Z. Qu, A. Oumbe, P. Blanc et al., "Fast radiative transfer parameterisation for assessing the surface solar irradiance: the heliosat-4 method," *Energy Meteorology*, vol. 26, no. 1, pp. 33–57, 2017.

[14] M. Schroedter-Homscheidt, "The Copernicus atmosphere monitoring service (CAMS) radiation service in a nutshell," in *22nd SolarPACES Conference 2016*, UAE, Abu Dhabi, United Arab Emirates, 2016.

[15] N. Helbig, *Application of the radiosity approach to the radiation balance in complex terrain*, Thesis at University of Zurich, 2009.

[16] D. G. Erbs, S. A. Klein, and J. A. Duffie, "Estimation of the diffuse radiation fraction for hourly, daily and monthly- average global radiation," *Solar Energy*, vol. 28, no. 4, pp. 293–302, 1982.

[17] H. Hottel and B. B. Woertz, "Performance of Flat Plate Solar Heat Collectors," *Transactions of the American Society of Mechanical Engineers*, vol. 64, pp. 91–104, 1955.

[18] T. Klucher, "Evaluation of models to predict insolation on tilted surfaces," *Solar Energy*, vol. 23, no. 2, pp. 111–114, 1979.

[19] D. L. King, W. E. Boyson, and J. A. Kratochvil, *Photovoltaic array performance model, tech. Rep.*, Sandia National Laboratories, 2004.

[20] D. L. King, J. A. Kratochvil, and W. E. Boyson, "Measuring solar spectral and angle-of-incidence e_ects on pv modules and solar irradiance sensors," in *Proceedings of the 26th IEEE PV Specialists Conference*, pp. 1113–1116, Anaheim, CA, USA, 1997.

[21] J. Hay and J. Davies, "Calculation of the solar radiation incident on an inclined surface," in *Proceedings of the first Canadian solar radiation data workshop*, pp. 373–380, Toronto, Canada, 1980.

[22] R. Perez, P. Ineichen, R. Seals, J. Michalsky, and R. Stewart, "Modeling daylight availability and irradiance components from direct and global irradiance," *Solar Energy*, vol. 44, no. 5, pp. 271–289, 1990.

[23] R. Perez, R. Seals, P. Ineichen, R. Stewart, and D. Menicucci, "A new simplified version of the perez diffuse irradiance model for tilted surfaces," *solar energy*, vol. 39, no. 3, pp. 221–231, 1987.

[24] R. Perez, R. Stewart, R. Seals, and T. Guertin, *The development and veri_-cation of the perez di_use radiation model, Tech. rep.*, Sandia National Laboratories, 1988.

[25] D. Reindl, W. Beckman, and J. A. Duffie, "Evaluation of hourly tilted surface radiation models," *Solar Energy*, vol. 45, no. 1, pp. 9–17, 1990.

[26] R. Ross, "Flat-plate photovoltaic array design optimization," in *14th IEEE Photovoltaic Specialists Conference*, San Diego, CA, 1980.

[27] V. Bourdin, J. Badosa, F. Calderon-Obaldia, C. Abdel-Nous, A. Migan, and Y. Du Bois, *Bonnassieux, Comparaison de modèles de complexité croissante pour la simulation de la température de modules photovoltaïques*, Journ ées Nationales du PhotoVoltaïque, 2017.

[28] M. K. Fuentes, *A simpli_ed thermal model of photovoltaic modules, tech. Rep.*, Sandia National Laboratories, 1985.

[29] D. Faiman, "Assessing the outdoor operating temperature of photovoltaic modules," *Progress in Photovoltaics : Research and Applications*, vol. 16, p. 307, 2018.

[30] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F. J. Martinez-de-Pison, and F. Antonanzas-Torres, "Review of photovoltaic power forecasting," *Solar Energy*, vol. 136, pp. 78–111, 2016.

[31] A. Labouret, P. Cumunel, J.-P. Braun, and B. Faraggi, *Cellules solaires : Lesbases de l'énergie photovoltaïque*, ETSF, Dunod, dunod Edition edition, 2010.

[32] D. L. Evans, "Simplified method for predicting photovoltaic array output," *Solar Energy*, vol. 27, no. 6, pp. 555–560, 1981.

[33] G. Graditi, S. Ferlito, and G. Adinolfi, "Comparison of photovoltaic plant power production prediction methods using a large measured dataset," *Renewable Energy*, vol. 90, pp. 513–519, 2016.

[34] G. R. Walker, "Evaluating mppt converter topologies using a matlab pv model," *Australian Journal of Electrical ans Electronics Engineering*, vol. 21, no. 1, 2001.

[35] A. Mellit and S. A. Kalogirou, "Artificial intelligence techniques for photovoltaic applications: A review," *Progress in Energy and Combustion Science*, vol. 34, no. 5, pp. 574–632, 2008.

[36] M. Karamirad, M. Omid, R. Alimardani, H. Mousazadeh, and S. N. Heidari, "ANN based simulation and experimental verification of analytical four- and five-parameters models of PV modules," *Simulation Modelling Practice and Theory*, vol. 34, pp. 86–98, 2013.

Hindawi
International Journal of Photoenergy
Volume 2021, Article ID 4925433, 18 pages
https://doi.org/10.1155/2021/4925433

Research Article

Design of Efficient Off-Grid Solar Photovoltaic Water Pumping System Based on Improved Fractional Open Circuit Voltage MPPT Technique

Ali Hmidet,[1] Umashankar Subramaniam,[2] Rajvikram Madurai Elavarasan,[3] Kannadasan Raju,[4] Matias Diaz,[5] Narottam Das,[6,7] Kashif Mehmood,[8] Alagar Karthick,[9] M. Muhibbullah,[10] and Olfa Boubaker[11]

[1]*University of Tunis El Manar, Higher Institute of Medical Technologies of Tunis (ISTMT), Tunisia*
[2]*Renewable Energy Lab, Department of Communications and Networks, Prince Sultan University, Saudi Arabia*
[3]*Department of Electrical and Electronics Engineering, Thiagarajar College of Engineering, Madurai 625015, Tamilnadu, India*
[4]*Department of Electrical and Electronics Engineering, Sri Venkateswara College of Engineering, Sriperumbudur, Tamilnadu, India*
[5]*University of Santiago of Chile, Electrical Engineering Department, Santiago, Chile*
[6]*School of Engineering and Technology, Central Queensland University, Melbourne, VIC 3000, Australia*
[7]*Centre for Intelligent Systems, School of Engineering and Technology, Central Queensland University, Brisbane, QLD 4000, Australia*
[8]*School of Electrical Engineering, Southeast University, Nanjing, China*
[9]*Renewable Energy Lab, Department of Electrical and Electronics Engineering, KPR Institute of Engineering and Technology, Arasur Coimbatore, 641407 Tamilnadu, India*
[10]*Department of Electrical and Electronic Engineering, Bangladesh University, Dhaka 1207, Bangladesh*
[11]*University of Carthage, National Institute of Applied Sciences and Technology, Tunis, Tunisia*

Correspondence should be addressed to M. Muhibbullah; m.muhibbullah@bu.edu.bd

Received 26 July 2021; Revised 13 September 2021; Accepted 14 September 2021; Published 14 October 2021

Academic Editor: Laurentiu Fara

The main application of off-grid solar photovoltaic (SPV) systems is water extraction in rural areas where access to the grid is restricted. In this application, photovoltaic (PV) and pump system regulation are crucial to increase its overall efficiency. In this context, this work presents a simple and efficient off-grid SPV water pumping system (SPVWPS). The designed system is based on a DC-DC boost converter, a three-phase DC-AC inverter, and a three-phase induction motor (IM) coupled to the centrifugal pump. The proposed solution is operated using a control strategy that associates an improved fractional open-circuit voltage (FOCV) method for maximum power point tracking (MPPT) and closed-loop scalar control. This association avoids the use of a speed sensor/encoder and a current sensor for the IM. Finally, the effectiveness of the proposed off-grid SPVWPS and its control system for both steady-state and dynamic conditions of insolation change is verified using a 1KVA rated prototype. The relevance of the drive is also checked in various operating conditions and is found to be adequate for pumping water. Moreover, the proposed method guarantees a fast response, less oscillations around the MPP, a system efficiency of 99%, and a high flow rate due to the extraction of maximum power.

1. Introduction

The generation based on fossil fuels from coal and oil threatens climatic conditions, accelerating carbon emissions [1]. In developing countries, the capacity to mount solar photovoltaic (PV) panels has increased significantly for more than a decade due to volatility in oil prices [2]. However, the applications of energy using solar photovoltaic (SPV) generators can be varied because of their enormous benefits. The SPV isolated systems are low price, secure,

and straightforward solutions for decentralised energy supply. They allow secure and decentralised electrical sources to be installed in areas far away from power plants. The isolated PV systems are used in telecommunications, rural electrification, agricultural, street lighting, signage, control, and rural development. One of the most critical applications of PV systems in agriculture is the water pumping system. Such an application makes it possible to extract water in rural areas, where the cost of installing a conventional line is too high [3–6]. In [7], the authors proposed a fuzzy precompensated hybrid proportional integral controller for permanent magnet motor driven solar water pumping system that overcomes conventional controller shortcomings. In [8], the authors proposed indirect field-oriented control for IM without an energy storage system to improve the performance of the PV water pumping system. The grid-connected motor-driven solar-powered water pumping system with efficient control is proposed in [9], which provides MPP tracking along with bidirectional power flow between the grid and PV.

Different kinds of IM control theories have been tested and validated in recent years to boost performance. Among these control methods, scalar control, known as constant stator flux V/f control, field-oriented control (FOC), and direct torque control (DTC), is significant [10–12]. Scalar control is the most popular form of variable speed drive for its low cost, ease of use, and applications that do not require high regulation at low speeds [13, 14]. The purpose of this technique is achieved by the control of the stator voltage and frequency to conserve the flux at a constant value. Unlike FOC and DTC, this technique does not require extra current sensors.

Since the SPV-based water pumping system is well affected by variable climatic conditions, there is always a need to optimize the energy utilisation. The PV source should always be exploited at its maximum efficiency point, i.e., maximum power point (MPP). Therefore, a convenient MPP Tracking (MPPT) scheme is required. A complete review of MPPT for off-grid SPV systems can be discovered in the literature, which comprises a constant voltage (CV), the fractional open-circuit voltage (FOCV), the fractional short circuit current (FSCC), the perturb and observe (P&O), the incremental conductance (IC), artificial neural network (ANN), fuzzy logic control (FLC), genetic algorithm (GA), and upgraded P&O based techniques [15–28].

In general, these techniques are classified into two categories: direct and indirect techniques [17, 25]. Direct MPPT techniques incorporate methods based on the PV measurement of current and voltage, i.e., P&O, IC, ANN, and FLC methods. On the other hand, the indirect methods use a database including empirical data or mathematical relationships resulting in corrections and approximations to estimate the MPP, i.e., CV, FSCC, and FOCV. The FOCV control is based on a proportional relation between the optimal voltage and the characteristic parameters of the PV module that is the open-circuit voltage [15, 16, 19, 21, 26]. This MPPT technique requires only one sensor for validation. Therefore, it is less expensive than P&O and IC methods. Nevertheless, from the point of view of precision, it is less efficient compared to P&O control due to the estimation of the open-circuit voltage parameter (V_{OC}) and the difficulty of interrupting the conversion to take measurements too often of V_{OC}. So, the algorithm of this command is slower than of the P&O method [17, 27].

Generally, the control algorithms for PV and other applications are implemented using microcontrollers to reduce the size and cost. However, this requires a long development time. Rapid-Control-Prototyping (RCP) is a complete solution to develop, control, optimize, and test control strategies quickly and efficiently. Also, this technique provides a straightforward identification tool for research and development in the power electronics industry [28–33]. To provide an RCP for Real-Time Simulation (RTS), sophisticated control platforms are used in the research and industrial application development, such as dSPACE controllers, OPAL-RT, RT-Box controller, and Imperix. Despite their prices remaining high, those platforms' main advantages are based on a fast, simple, and efficient environment thanks to high-level programming. dSPACE remains the most popular on the market for its ease of use and integrates a software and hardware environment more suited to models designed with MATLAB/Simulink.

In this context, an RCP of an SPV water pumping system is proposed to save time and effort in implementing the proposed solution. It is reposed on the use of a DS1104 card from dSpace. The design of an overall circuit is composed of a DC-DC boost converter, a three-phase inverter, and an IM coupled to the centrifugal pump. From the control point of view, the system is based on a modified fractional open-circuit voltage (FOCV) MPPT method and V/f control to reduce the requirement of measurements. Only the PV voltage and the IM voltages are required. An experimental relationship based on temperature measurement is developed to avoid using the static switch for open-circuit voltage detection in the FOCV MPPT technique. Thus, ensures better power extracted at the output of the PV generator. The RCP technique is used to control, test, and validate the effectiveness of the complete system. Therefore, the proposed off-grid SPV water pumping system is built on a dSPACE DS1104 environment for a real-time command, visualization, and data acquisition. Experimental results verify the effectiveness of the proposed strategy. The SPV water pumping has been tested under different operating regimes exhibiting proper behavior.

The rest of the paper is organized as follows: the modelling of SPVWPS is presented in Section 2 that contains detailed designs and calculations. The proposed control strategy is explained in Section 3, while the experimental results are discussed in Section 4; the conclusions are drawn in Section 5.

2. Modelling of SPVWPS

In this fragment, the design and configuration of the SPVWPS are described in detail. It also demonstrated the mathematical models of all functional blocks related to this SVPWPS which can be realized for the experimental validation.

2.1. System Design and Configuration. The proposed installation of high flow SVPWPS comprises SPV panels that act as a DC source, electronic conversion device, motor pump, and command and control system. The complete architecture of the proposed off-grid scheme is depicted in Figure 1.

The expanded description of the functional blocks is as follows.

(i) A PV generator was built based on ten panels connected in series. Their main specifications are given in Table 1

(ii) A SEMIKRON DC-AC inverter of rated power 20 KVA. This device is designed for both DC/DC converter and three-phase VSI. The control signals are generated using a dSPACE 1104 Digital Signal Processor "DSP" board. The PWM command of the boost converter is generated using the MPPT controller, where the inverter is controlled using the SVPWM technique

(iii) A dSPACE DS1104 is used as an RCP platform. The advantage of this controller is the relatively simple implementation of control strategies using the MATLAB/Simulink environment

(iv) A Lowara CEAM 70/3/A motor pump set consisting of a three-phase squirrel-cage motor connected in star and coupled to a single-impeller centrifugal pump. The scalar control, well known V/f control, is proposed to deliver reference speed to the IM. The characteristics of the motor pump are given in Table 2

(v) A water tank is used to simulate a mechanical load

(vi) A reduced number of voltage and current sensors for verification are used. Temperature and irradiation sensors are used in the PV array

2.2. PV Source Model and Simulation. The electrical equivalent circuit of the PV cell is equivalent to a current source coupled with a diode, a shunt resistor, and a series of resistance, as presented by Figure 2 [34].

The current I_{ph} is the current emitted by the cell under solar radiation; it is the current of photons. It is strongly related to the intensity of the insolation and depends only very slightly on the temperature variation. The current I_d the diode is strongly related to the temperature and the energy of the gap of the junction. It also depends on the voltage of the diode. The shunt current I_{sh} corresponds to the current passing in the parallel resistor R_{sh}.

The mathematical model describing the IV characteristics of a PV cell is given in Equation (1) [16, 21, 22, 34].

$$I_{\mathrm{pv}} = I_{\mathrm{ph}} - I_s\left[\exp\frac{q\left(V_{\mathrm{pv}} + R_s I_{\mathrm{pv}}\right)}{\mathrm{NKT}} - 1\right] - \frac{V_{\mathrm{pv}} + R_s I_{\mathrm{pv}}}{R_{\mathrm{sh}}}, \quad (1)$$

where I_{pv}, I_{ph}, and I_s represent the PV array currents, the photocurrent, and the diode reverse saturation current, respectively. q is the electron charge $(1.60217\,10^{-19}\,\mathrm{C})$, V_{pv} is the global voltage of PV array, K is the Boltzmann constant equal to $1.3806503\,10^{-23}$ j/k, T is the temperature of a P-N junction in Kelvin, N is the diode ideality constant, and R_s and R_{sh} are the series and shunt resistors of the cell, respectively.

The PV panel characteristics are shown in Figure 3 that exhibit nonlinear behaviors related to temperature and insolation conditions. These characteristics are loaded with a resistor under standard test conditions (STC). This figure proves that there is only one operating point that corresponds to the maximum power, called maximum power point "MPP".

2.3. DC/DC Boost Converter Model and Simulation. A boost chopper is used to optimize the power supplied by the PV generator. Figure 4 illustrates a DC/DC boost converter [35, 36]. The capacitor is modelled by its capacitance C and series resistance R_C while the inductor is modeled by its inductance L and series resistance R_L. R_C and R_L are assumed insignificant in this design. Likewise, let us admit the transistor ideal and the effect of the forward voltage across the diode is neglected during conduction (where R is the load resistor).

For the last proposed scheme of a DC/DC boost converter, a dynamic system can be described by the following equations:

$$\begin{cases} \dfrac{dv_O}{dt} = \dfrac{(1-D)}{C} i_L - \dfrac{1}{RC} v_O, \\ \dfrac{di_L}{dt} = \dfrac{1}{L} v_{\mathrm{in}} - \dfrac{(1-D)}{L} v_O. \end{cases} \quad (2)$$

The steady-state regime assumes the following:

(i) The average current in the capacitor is equal to zero, and therefore, the average current in the load is equal to the average current in the diode

(ii) The average voltage across the inductor is reduced to the voltage drop in its resistance; it is generally considered null

(iii) The average voltage across the capacitor is equal to the voltage of the load

(iv) The average voltage across the diode is equal to the difference between the input voltage and the load

$$\begin{cases} I_C = 0 = I_D - I_O = I_D - \dfrac{V_O}{R}, \\ V_L = R_L I_L \simeq 0, \\ V_C = V_O = RI_O = RI_D, \\ V_D = V_{\mathrm{in}} - V_L - V_O = V_{\mathrm{in}} - V_O. \end{cases} \quad (3)$$

Let us also note that there are two modes of waveforms of the inductance current. These modes are, respectively, called discontinuous conduction mode (DCM) and

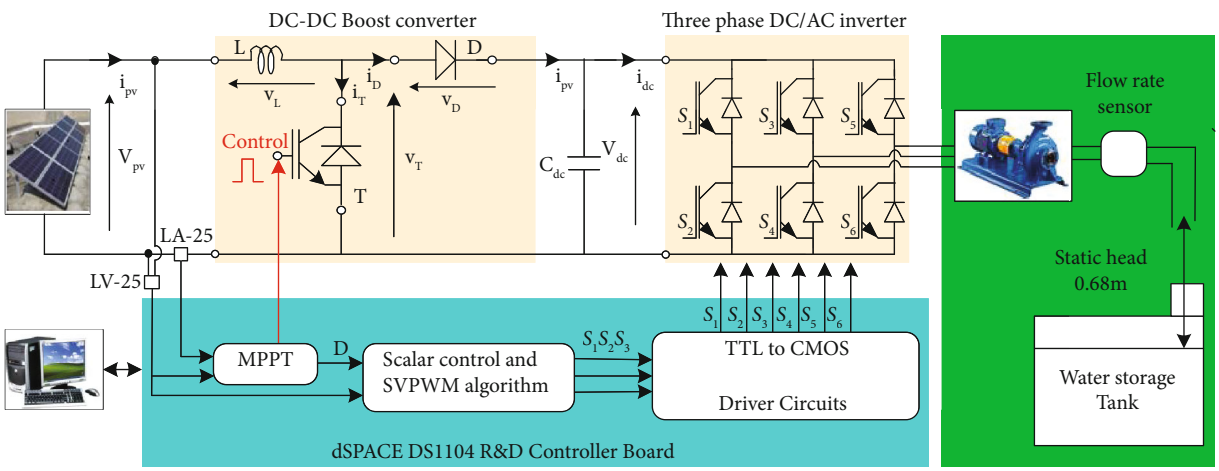


Figure 1: The system architecture of the proposed off-grid SPV water pumping system.

Table 1: PV module specifications.

Characteristic	Value
Maximum power (P_{MPP})	50 Wp
Optimal voltage (V_{MPP})	17.2 V
Optimal current (I_{MPP})	2.9 A
The voltage at the open circuit V_{OC}	20 V
Photocurrent (I_{SC})	3.4 A

Table 2: Motor-pump characteristics.

Characteristic	Value
Pump power rating	0.37 KW
IM power rating	0.61 KW
Voltage rating	220/380 V
Current rating	2.6 A
Speed rating	2800 rpm

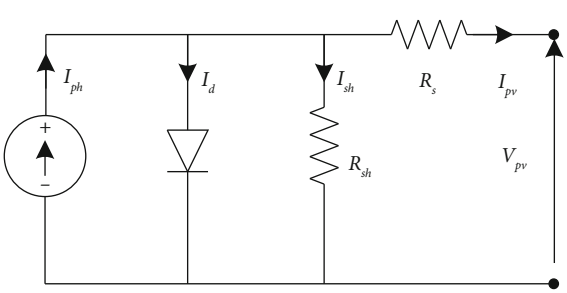


Figure 2: Solar cell equivalent circuit.

continuous conduction mode (CCM). In the case of DCM, the current $i_L(t)$ is cancelled before the cancellation of the commutation period. On the contrary, in the case of CCM, the minimum value of $i_L(t)$ is more significant or at the limit equal to zero. In this work, we assume that our boost converter is operating in the CCM regime. The output voltage of the converter is then governed as follows:

$$V_O = \frac{1}{1-D} V_{\text{in}}. \tag{4}$$

The value of the inductor $L_{\min}$ selected based on estimated inductor ripple current ΔI_{I_L} at a maximum input voltage V_{in} is given in Equation (5), with quantity f_s designates the switching frequency.

$$L_{\min} = \frac{D}{f_s \Delta I_{I_L}} V_{\text{in}}. \tag{5}$$

The minimum value of capacitance $C_{\min}$ that results in the output voltage ripple ΔV_O is given by:

$$C_{\min} = \frac{D}{f_s R \Delta V_o} V_O. \tag{6}$$

The optimum load corresponding to the 100% efficiency of the converter is expressed as follows:

$$R_{\text{opt}} = \frac{1}{(1-D)^2} \frac{V_{\text{in}}}{I_L}. \tag{7}$$

Figure 5 provides the waveforms of the inductance current, the current of the diode, and the inductance voltage, respectively.

2.4. DC-Link Voltage Calculation/DC-Bus Voltage Calculation. For a three-phase voltage inverter whose output voltage V_s can reach 230 V RMS value, the DC bus voltage V_{dc} must obey the following relationship:

$$V_{\text{dc}} = \sqrt{2}\, V_s = \sqrt{2} x 330 = 325\ \text{V}. \tag{8}$$

The DC bus voltage must be greater than 325 V; therefore, a voltage V_{dc} of 400 V can operate the inverter in the best conditions.

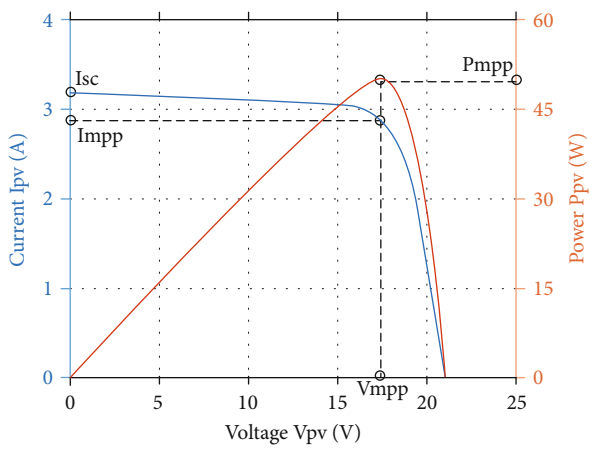


Figure 3: STC characteristics.

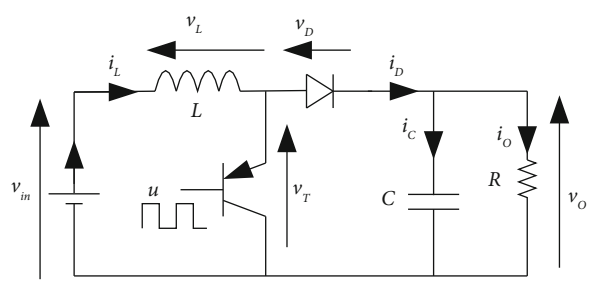


Figure 4: Circuit diagram of DC/DC boost converter.

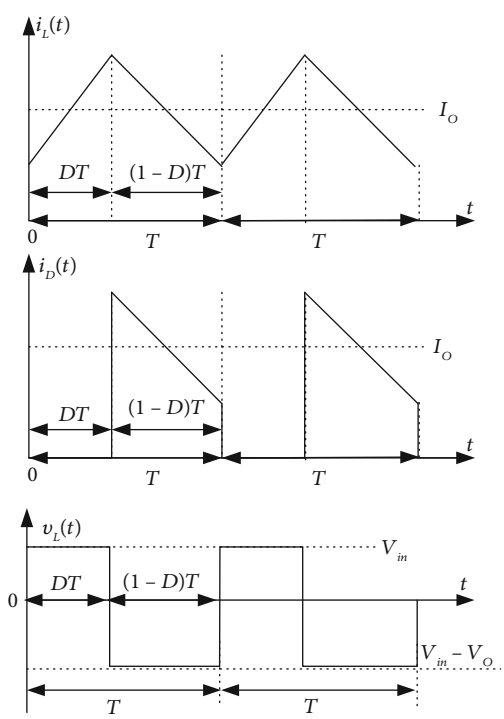


Figure 5: Boost converter waveforms at CCM.

2.5. Design of Boost Converter Parameters. The duty ratio (D) for the boost converter is calculated as follows [37]:

$$D = \frac{V_O - V_{\text{in}}}{V_O} = 1 - \frac{325}{400} = 0.38. \tag{9}$$

Let us assume that the boost converter operates at the MPP and STC conditions; $V_{\text{pv}} = V_{\text{MPP}}, I_{\text{pv}} = I_{\text{MPP}}$, the optimum load resistor R_{opt} is

$$R_{\text{opt}} = \frac{1}{(1-D)^2}\frac{V_{\text{MPP}}}{I_{\text{MPP}}} = \frac{1}{(1-0.38)^2}\frac{172}{2.9} = 150\Omega. \tag{10}$$

For a current ripple ΔI_{I_L} designed of 10%, the minimum value of the inductor is

$$L_{\min} = \frac{D}{f_s \Delta I_{I_L}} V_{\text{MPP}} = \frac{0.38}{10000\,(2.9\,x\,0.1)}\,172 = 22.5\,\text{mH}. \tag{11}$$

The minimum value of the capacitance corresponds to 1% of the voltage ripple ΔV_O and load resistor with $150\,\Omega$ is

$$C_{\min} = \frac{D}{f_s R_{\text{opt}} \Delta V_o} V_O = \frac{0.38x325}{10000x150x(325x0.01)} = 25.33\,\mu\text{F}. \tag{12}$$

Hence, a capacitance of $50\,\mu$F is sufficient. In summary, an inductor of 22.5 mH and a capacitance of $50\,\mu$F can be chosen.

2.6. Design of DC-Link Capacitor. The treated system is designed for water pumping; therefore, the DC-link capacitor C_{dc} connected between the boost converter and the VSI is estimated according to the fundamental frequency as follows [37]:

$$\omega_{\text{rated}} = 2\pi f_{\text{rated}} = 100\pi = 314\,\text{rad/s},$$
$$C_{\text{dc}} = \frac{6aVIt}{\left(V_{\text{dc}}^2 - V_{\text{dc1}}^2\right)} = \frac{6x1.2x132.8x2.9x0.005}{\left(400^2 - 325^2\right)} = 254.976\,\mu\text{F}. \tag{13}$$

Therefore, a capacitor of $250\,\mu$F and 400 V is selected. In the previous relationship, V_{dc} is the estimated dc voltage, V_{dc1} is the selected dc voltage during transients, a is the overloading factor, and t refers to the time by which selected voltage reach at estimated voltage range. Furthermore, V and I designate the phase voltage and current of the motor, respectively.

2.7. Three-Phase DC/AC Inverter Model. The IM is driven by a three-phase VSI [13, 33, 37], whose objective is to furnish in the output a variable voltage and frequency through a PWM controller (S_1–S_6 is the power switch). The switches of any leg of the inverter are complementary. Various PWM techniques can be exploited to produce control signals for a voltage inverter. This work uses the space vector PWM

technique. The SVPWM technique is commonly used for the application of scalar speed control methods. The comparison with other PWM, SVPWM offers better performance because it offers less total harmonic distortion (THD), lower switching losses at the high switching frequency, and enables efficient use of DC-link voltage [4, 11, 18].

The principle of PWM is explained by using the VSI block diagram shown in Figure 1. For a three-phase VSI admitted without loss, the output voltage is strictly defined by the voltage of the DC bus V_{dc} supplying the inverter and by the logic state of the three highest IGBT's (S_1, S_2, S_3). With the three Boolean variables S_1, S_2, and S_3, there are only eight possible combinations. From these combinations, we determine eight voltage vectors, six active vectors denoted from $\bar{v}_1$ to $\bar{v}_6$, and two zero vectors $\bar{v}_0$ and $\bar{v}_7$. It is easily observed that the active vectors form a balanced six-phase system; therefore, the same modulus and chronic phases equal to 60°. Indeed, it is possible to pose the following system of equations where k is an integer indicating the switching combinations.

$$v_k = \begin{cases} \sqrt{\frac{2}{3}} V_{dc} e^{j(k-1)\pi/3} & \text{for } k = 1, \cdots, 6, \\ 0 & \text{for } \quad k = 0, 7. \end{cases} \tag{14}$$

Figure 6 below shows the general principle of the SVPWM technique. Figure 6(a) gives a spatial representation of these vectors while indicating each vector's associated combination of commands. For example, the combination 010 corresponds to $S_1 = 0$, $S_2 = 1$, and $S_3 = 0$. The SVPWM technique exploits the vector voltages diagram and synthesizes a requested reference voltage vector $\bar{v}_{ref}$ from the two neighboring vectors $\bar{v}_k$ and $\bar{v}_{k+1}$ and the null vectors $\bar{v}_0$ or $\bar{v}_7$ as indicated in Figure 6(b). This synthesis must be done in average value over a time interval T_s. The vectors $\bar{v}_k$ and $\bar{v}_{k+1}$ are applied during the time intervals τ_k and τ_{k+1}, respectively, and the null vector is applied during the remaining time τ_0, where $\tau_0 = T_s - \tau_k - \tau_{k+1}$.

In other words, for an average value equivalent to $\bar{v}_{ref}$ over the period T_s, the following relation must be verified.

$$\frac{\tau_k . \bar{v}_k + \tau_{k+1} . \bar{v}_{k+1}}{T_s} = \bar{v}_{ref}. \tag{15}$$

To respect the constraint $\tau_k + \tau_{k+1} \leq T_s$, the V_{ref} the module of the required voltage $\bar{v}_{ref}$ must verify the equation below.

$$V_{ref} \leq \frac{V_{dc}}{\sqrt{2}}. \tag{16}$$

In this context, the synthesis solution is given in Equation (17), where the coefficient ρ nominates a voltage ratio.

$$\begin{cases} \tau_k = T_s \rho \sin\left(\frac{\pi}{3} - \varsigma\right), \\ \tau_{k+1} = T_s \rho \sin(\varsigma), \\ \rho = \frac{\sqrt{2}\, V_{ref}}{V_{dc}}. \end{cases} \tag{17}$$

2.8. Association Motor-Pump Model. A squirrel-cage IM is used in this application. In a stationary reference frame, it is described by five main equations as follows [38]:

$$\frac{d\bar{\phi}_s}{dt} = \bar{v}_s - R_s \bar{i}_s, \tag{18}$$

$$\frac{d\bar{\phi}_r}{dt} = j\omega_r \bar{\phi}_r - R_r \bar{i}_r, \tag{19}$$

$$\bar{\phi}_s = \ell_s \bar{i}_s + m\bar{\phi}_r, \tag{20}$$

$$T_e = -p\Im_m\left(\bar{\phi}_s \bar{i}_s^*\right) = p\left(\phi_{ds} i_{qs} - \phi_{qs} i_{ds}\right), \tag{21}$$

$$T_e - K_f \omega_r - T_r = \frac{j}{p}\frac{d\omega_r}{dt}. \tag{22}$$

In the collection of equations above, variables and parameters with index "s" are at the stator while those with "r" are at the rotor. $\bar{v}$, $\bar{i}$, and $\bar{\varphi}$ designate the vectors of voltage, current, and flux, respectively. R is the resistance, ℓ_s is the leakage inductance, and m denotes the ratio M/L_r, where M is the mutual inductance and L_r is the rotor inductance.

Electrical speed and pole pair number are ω_r and p. T_e and T_r are electromagnetic and mechanical torques, respectively. Various equations can determine the electromagnetic torque, the best known of which is that of Equation (21) where $\Im_m$ designates the imaginary part; φ_{ds}, φ_{qs}, i_{ds}, and i_{qs} are the direct and quadrature components of stator flux and current, respectively. This equation is common for transient stationary and steady-state [14, 38]. The IM mechanical equation is expressed by Equation (22). Here, K_f is the viscous coefficient, and j is the moment of inertia (kg m^2). The centrifugal pump driven by the IM develops a mechanical torque whose dynamics are presented by the following equation [12]:

$$T_r = J\frac{d\omega_r}{dt} + K_f \omega_r + K\omega_r^2. \tag{23}$$

Let it be noted that the steady-state torque produced by the IM is equivalent to the load torque of the pump T_r, the rotor speed is ω_r in rad/s, and K is the centrifugal pump constant. For the water pump used in this test, the proportionality constant K is given according to Equation (24).

$$K = \frac{T_r}{\omega_r^2}. \tag{24}$$

For the IM, the rated torque is 1.2619 Nm, and the rated

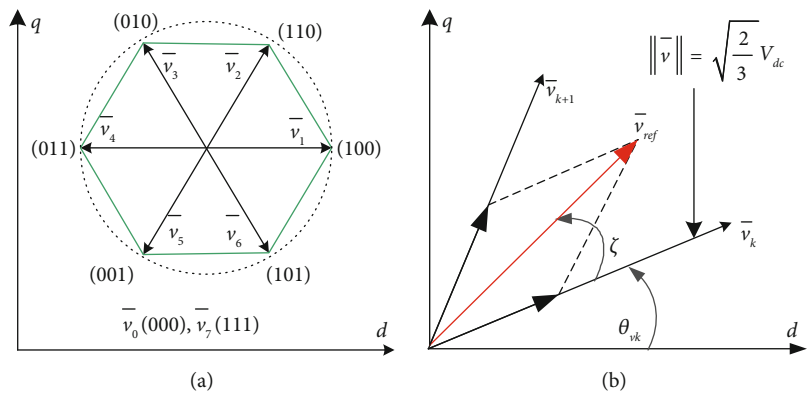


Figure 6: Principle of the SVPWM technique: (a) the voltage space vectors of the three-phase inverter; (b) synthesis of the SVPWM.

speed is 2800 rpm; thus, the proportionality constant K evaluated based on (24) is equal to $14.677\,10^{-6}$ Nm/(rad/s)^2. A KOBOLD C34P flow meter installed on the pump is used to measure the flow rate (Q). The sensor is equipped with a display that delivers a flow rate varying from 0 to 100 l/min. As an indication, a tank with 0.68 m of the head was used in the test bench as a well. The parameters are identified by the variation of the motor supply voltage, which drives the pump at a fixed frequency. However, the IM is connected to the grid via an autotransformer, i.e., a variable supply voltage and a fixed frequency of 50 Hz. Table 3 illustrates the results obtained during the experiment.

The pumping performance of a centrifugal pump is expressed in the form of an $(H-Q)$ curve, depicting the flow Q (e.g., in m^3/h) and the head H (e.g., in m) of the pump. The variation of the pump's speed can give us numerous characteristics $(H-Q)$ as described by the following relation [25, 39]:

$$H = B_1{\omega_r}^2 - B_2\omega_r Q - B_3 Q^2. \tag{25}$$

The optimal values of B_1, B_2, and AB_3 seek for $H = 0.68$ m is $B_1 = 0.039$, $B_2 = -0.3079$, and $B_3 = -0.0024$.

Indeed, the results are used to plot the flow characteristic as a function of speed, Figure 7, and using the "polyfit" function of MATLAB library, the coefficients a_0, a_1, a_2, and a_3 of the equivalent polynomial of flow rate-speed relationship corresponding to Equation (26) are calculated.

$$Q = a_3N^3 + a_2N^2 + a_1N + a_0, \tag{26}$$

$$a_3 = 1.7x10^{-9}, \tag{27}$$

$$a_2 = 9.3x10^{-7}, \tag{28}$$

$$a_1 = 24x10^{-3}, \tag{29}$$

$$a_0 = -64x10^{-3}. \tag{30}$$

A first impression to exploit these results would be to control the flow to the speed. Since the behavior of the system is not linear, it would be interesting to program a speed control routine to ensure equivalent flow control.

2.9. Proposed Control Strategy. The model developed above of the SPVWPS needs a better control scheme to optimize its performance. The proposed strategy embodies the MPPT method for maximum power extraction from the PV generator by the real-time control of the duty ratio of a boost converter, a constant V/f control to give the reference speed of the IM drive, and an SVPWM algorithm to control the power semiconductor devices for better use of dc-bus voltage and a minimum THD.

2.10. Proposed MPPT Structure. The MPPT structure proposed and carried out in this work is the FOCV method. This choice is based on several factors, such as this method is of a simple structure. Low cost, remarkably achievable analogically, and the ability of its easy integration with a DC/DC controller in various applications like the SPVWPS. FOCV method [16, 17, 19, 21, 25–27], based on the use of reference voltage to adjust the duty ratio of the MPPT controller in a feedback control loop, supposes that irradiation and temperature variations on the array have not a significant impact on the MPP voltage value and that MPP voltage can be estimated with a predefined constant voltage V_{REF}. Thus, the optimum voltage is always close to a proportion of the open circuit voltage V_{OC}. Consequently, V_{REF} can be calculated from the empirical relationship shown in Equation (31). It is found that the value of K varies between 72% and 78% [17, 26].

$$V_{\mathrm{REF}} = KV_{\mathrm{OC}}. \tag{31}$$

In this method, the error arising from the comparison of the array voltage with a constant reference voltage is used to continuously adjust the converter duty cycle so that the array operates at a predetermined operating point near the MPP, as seen in Figure 8.

This figure shows that after V_{OC} is sampled by a sampler; V_{REF} which is calculated using Equation (31) is kept constant during one sampling period by hold circuit; now duty ratio D is adjusted to make $V_{\mathrm{pv}} = V_{\mathrm{REF}}$. For the following sample, again V_{OC}, it is sampled, and the same procedure is repeated for each sample.

The implementation of the FOCV technique needs the measurement of open-circuit voltage V_{OC} for all climatic

Table 3: Measurement of flow under supply voltage variation.

Voltage (V)	Speed N (tr/min)	Flow rate Q (l/min)
118	1242	27
136	1482	33.2
180	1904	38.5
230	2167	38.7
280	2310	39.3
330	2385	38.2
380	2738	38

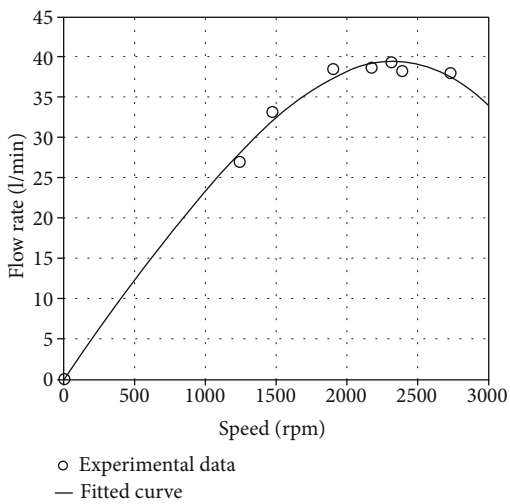


Figure 7: Identification of the flow-speed characteristic.

change. Therefore, it is mandatory to use a static switch in series with the PV generator to establish the open-circuit state. This produces a large oscillation of the output power of the PV array.

Because the voltage V_{OC} varies essentially with temperature (V_{OC} decreases in a slight way, unlike the increase of the temperature), this drawback can be eliminated by measuring only the temperature according to giving relation. To overcome this inconvenience, a series of measurements is carried out on the variation of the V_{OC} and the optimal voltage V_{MPP} as a function of temperature, see Table 4.

Figure 9 confirms that V_{OC} decreases slightly when the temperature increasing and can be approximated by the following linear function:

$$V_{OC} = aT + b, \tag{32}$$

where $a = -0.1754$ and $b = 23.6443$.

On the other hand, Table 4 shows that the constant K of Equation (31) is 0.77, and the average value of the optimal voltage is around 14.72 V in this practical study.

The flowchart of the modified FOCV method is presented in Figure 10. In this figure, the temperature is measured in parallel with the measurement of the voltage of the PV generator to calculate the voltage V_{OC} then the reference voltage. Thus, the maximum power is tracked by comparing the current PV voltage to the already calculated reference voltage.

2.11. Scalar (V/f) Control for IM. The scalar control of an IM is the most prevalent and easiest to date. This approach, also called V/f control, is easy to implement and cost-effective. With the exception of V/f control, FOC and DTC algorithms are difficult and involve additional current sensors for their implementation [9, 13, 38, 40, 41]. The V/f control focuses only on the steady-state model of the IM and assumes that the magnitude of stator flux remains constant in this operating mode. The speed regulation of the motor can be achieved by adjusting the stator voltage magnitude and frequency in such a way that the air gap flux is always conserved at the requisite value. According to Equation (18) at a steady-state and within the low frequency, the voltage drop at the stator resistance can be neglected. Therefore, by keeping the ratio between stator voltage V_s and stator field frequency f_s constant, the magnitude of stator flux Φ_s is controlled as shown in Equation (33) [37, 41], where K_v is the constant of proportionality between the nominal voltage V_n of phase and frequency f_n of the voltage source.

$$\Phi_s \approx \frac{V_s}{2\Pi f_s} = K_v. \tag{33}$$

At low operating frequency under a few Hertz, the voltage drop across to the stator resistance is no longer negligible compared to the leakage reactance. This structure of control should consider the voltage drops to keep a constant stator flux. Inversely, to protect the motor against overvoltage if the rotor speed is more significant than that which corresponds to the rated frequency f_n, the stator voltage must be adjusted accordingly according to the following relation:

$$V_s = \begin{cases} (V_n - V_0)\dfrac{f_s}{f_n} + V_0 & \text{for } f_s < f_n, \\ V_n & \text{for } f_s \geq f_n. \end{cases} \tag{34}$$

Here, V_0 denotes the stator voltage at zero frequency. The presented system tracks the MPP by the FOCV MPPT structure proposed and described above so that the motopump can extract the maximum power available from the SPV array for all solar insolation changes. The MPP tracker is designed here based on the boost converter. So, the output voltage average value of the DC-DC boost converter V_{dc} in CCM mode is expressed in terms of the average value of the input voltage V_{pv} by Equation (4). The following equation gives the voltage magnitude of the IM. Here, ρ denotes the voltage ratio and D the duty ratio adjusted by the controller of the MPP.

$$V_s = \rho V_{dc} = \frac{\rho}{1-D} V_{pv} = \delta V_{pv}. \tag{35}$$

Thus, the stator reference pulsation ω_{sref} is obtained based on the V_s/f_s technique. The resulting references of stator pulsation ω_{sref} and the voltage ratio ρ are applied to the

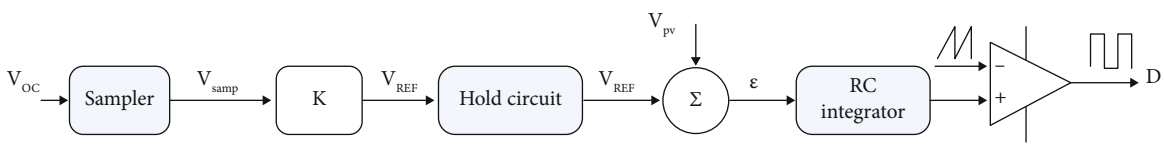


Figure 8: MPPT using the constant voltage method.

Table 4: V_{OC} and V_{MPP} voltage responses versus temperature.

T (°C)	V_{OC} (V)	V_{MPP} (V)	K
22	19.86	15.344	0.7726
23	19.72	15.202	0.7709
26	18.92	14.584	0.7708
28	18.54	14.260	0.7691
32	18.2	14.201	0.7802

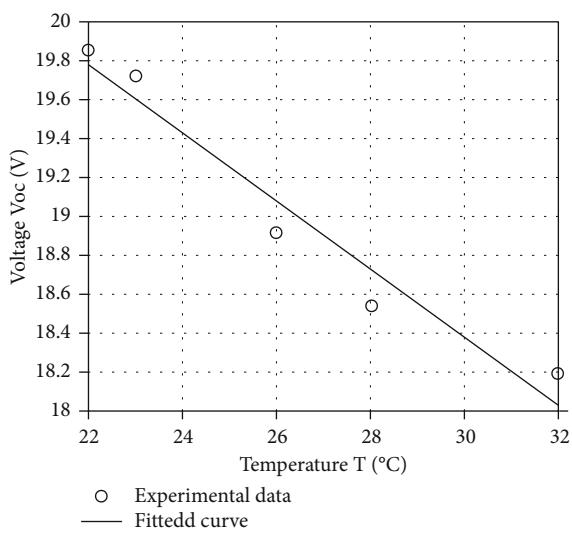


Figure 9: V_{OC} response versus temperature variation.

motor using an SVPWM technique driving three-phase two-level VSI. The schematic diagram of the proposed control is exposed in Figure 11.

3. Experimental Results and Discussions

To check the functionality and analyse the performance of the designed system, an experimental investigation is performed for different insolation conditions. The whole algorithm is designed using the MATLAB/Simulink model and implemented in the dSPACE DS1104 environment. It comprises both the algorithms, likely MPPT algorithm for the power maximisation and the V/f control along with the SVPWM algorithm. The prototype hardware setup was developed in the laboratory for experimental validation (Figure 12) using various elements such as DC-AC inverter, dSPACE DS1104, TTL-CMOS modules, sensors, a series connection 10 panels, motor-pump with a flow meter, and a tank.

The main program of the proposed application based on an RCP system is modelled using Simulink/MATLAB and implemented on the DS1104 R&D Controller Board, as shown in Figure 13. The measured temperature is converted into a reference voltage and compared to the actual voltage of the solar panel to generate an error ΔV. The sum of this error and its integral constitutes the PI regulator. The output of the PI regulator is compared to a triangular signal to give the command to the chopper via a galvanic isolation circuit. A vital building SVPWM block allows determining the time duration and the sector for each reference vector. The inputs of the SVPWM are generated based on the constant V/f control principle. The stator voltage applied to the stator windings of the IM is determined according to the reference vector generated and the actual voltage of the DC bus V_{dc}.

Indeed, other Simulink blocks are used for the acquisition of verification and validation data (insolation, flow rate, electromagnetic torque, currents, power, efficiency, etc.) and are not shown in Figure 13.

4. Test Results for MPPT Solution

The effectiveness of the FOCV technique is based on the use of a voltage sensor (LEM LV-25) for the acquisition of the PV array voltage V_{pv}. Moreover, the simplest form of operation is to maintain the PV generator at a constant voltage identical to the optimal voltage at STC furnished by the manufacturer (neglecting the effects of solar radiation and temperature changes on the MPP voltage). Equations (31) and (32) are applied to obtain the adequate reference voltage for the system. Further, the required information on the temperature is perceived using an electronic circuit (LM35) installed in the roof adjacent to PV panels. The sensed voltage of the PV and the temperature are applied to two ADC's of the DS1104. The computer system internally calculates the reference voltage and compares it to the measured value to generate an error, and it can be compared to a triangular wave. Further, a PWM signal is generated for an insulated gate bipolar transistor (IGBT) using a boost converter. Then, the output voltage of the boost converter is used to determine the stator voltage V_s that needs to be applied to the IM according to Equation (35). The centrifugal pump coupled with IM is maintained under a nominal constant stator flux of 0.7 Wb. Therefore, stator pulsation ω_s is obtained in real-time by dividing the stator voltage V_s by 0.7; it is keeping the rule V/f constant.

The practical results obtained on the three-phase pumping system with an improved FOCV MPPT algorithm for a temperature of 24°C, 15 minutes, and under insolation variation are summarized in the figures below. Figure 14(a) shows the arbitrary scenario captured for the insolation

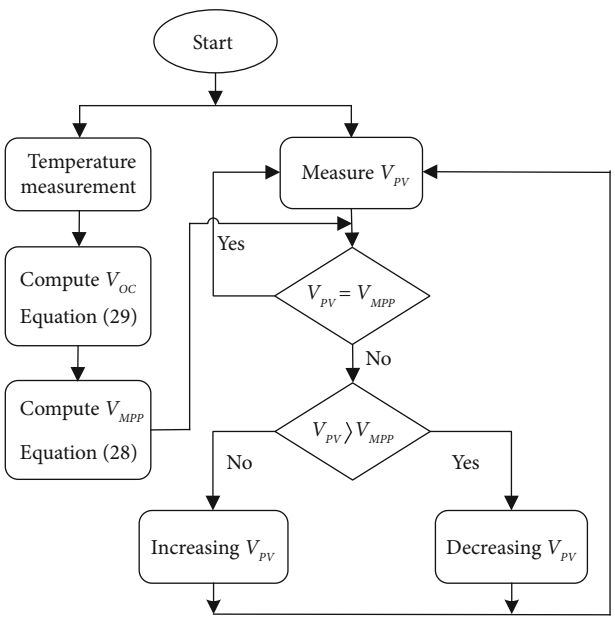


Figure 10: Flowchart of the modified FOCV MPPT method.

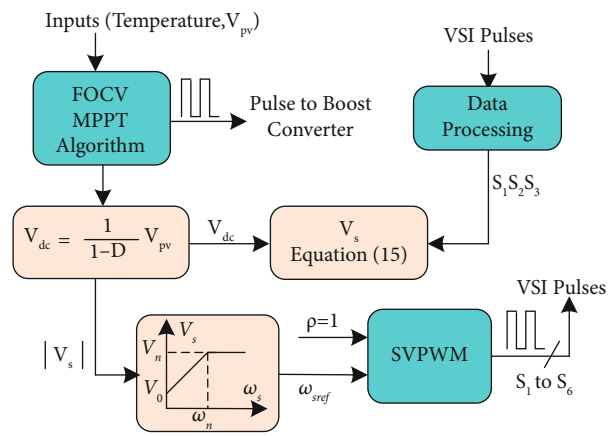


Figure 11: Schematic diagram of V/f control with FOCV MPPT structure.

(Ins), whereas Figure 14(b) denotes the evolution of the duty cycle D for the boost converter to reach the MPP in real-time. The insolation ranges from 500 to 850 W/m^2 generating a variable D of 20 to 35% that is instantly adjusted to extract the maximum power according to the improved FOCV MPPT technique proposed. The instantaneous evolution of the input power (P_{pv}) and the optimal power (P_{pv-opt}) that can be delivered by the SPV generator are presented in Figure 14(c). It is observed that they are confused that confirms that the system extract the maximum power available, and the efficiency is optimal and reaches 100%. In this test situation, the maximum insolation is 850 W/m^2 correspond to an optimal current of 2.48 A (verified according to STC delivered by the manufacturer, Table 1), and according to Equations (31) and (32) for a temperature of 24°C, the average MPP voltage is 150 V; then, the maximum available power for the PV generator is 372 W. This verification calculation is confirmed by Figure 14(d) that presents the movement of MPP against the optimum voltage. Zoom on this figure shows a better movement of the MPP with a maximum power of 375 W (Figure 14(e)).

Figure 14(f) displays that the PV voltage V_{pv} is maintained constant around their optimum reference P_{pv-opt} under insolation variation. This figure proves that the variation of insolation does not have a significant impact on the MPP voltage value and that the boost converter provides a variable output voltage (V_{dc}) According to the duty cycle to reach the MPP. Figures 14(b) and 14(f) prove that for a value of D equivalent to 35% as calculated previously by Equation (9), the voltage delivered V_{dc} is 231 V.

It is a crucial task to study the effectiveness of the proposed algorithm on the complete pumping system. Therefore, the characteristics of stator voltage, pulsation response, stator current, and flow rate are studied, and the observed results are presented. Figure 14(g) proves that the rule of V/f constant is respected because the stator voltage V_s and the pulsation ω_s vary proportionally with a constant confirmed of 0.7 Wb. It is noted in Figure 14(h) that stator current (I_s) of the IM varies proportionally versus the power extracted from the PV array. Figure 14(i) proves that the flow rate (Q) varies in proportion to the available power, and its maximum value accelerates and decelerates depending on the extracted maximum power.

A second experiment is saved for the same temperature but for insolation that considered constant at about 800 W/m^2, over 15 minutes of recording (Figure 15). We observe, respectively, the scenario of insolation for these test conditions, the duty cycle of the converter, the PV voltages and the DC bus voltage, the available PV power and the extracted one, the movement of the MPP point, the stator voltage and pulsation and the current magnitude, and the flow rate.

It is easy to see that the duty cycle is kept constant at about 0.325 which raises the average voltage at the boost converter output to 240 V while keeping the PV voltage equal to an optimal reference voltage around 150 V. The power extracted is close to the optimal power available, which results in a maximum power operation and consequently an efficiency that exceeds 99%. The system then operates on the MPP point as shown in Figure 15(e). The stator quantities are kept constant, and the voltage is proportional to the pulsation with a slope of 0.7 Wb in steady-state operation. Figure 15(h) shows that the water flow rate remains constant at its maximum value for this constant insolation operating condition, and therefore, the motor supply voltage is well controlled.

While comparing the performance of the conventional FOCV method where the reference voltage (V_{REF}) is imposed whatever the variation of the weather conditions or precalculated by measuring the open-circuit voltage (V_{OC}), the proposed structure makes it possible to directly find the reference voltage, which results in a better power extraction and reduced oscillations. Indeed, a third experiment is recorded by imposing a reference voltage V_{REF} of 142 V. The results of this test are shown in Figure 16.

Figure 12: The hardware setup of the SPVWPS: (a) overview of the test bench; (b) motor pump with flow meter and water tank; (c) a series connection of 10 panels.

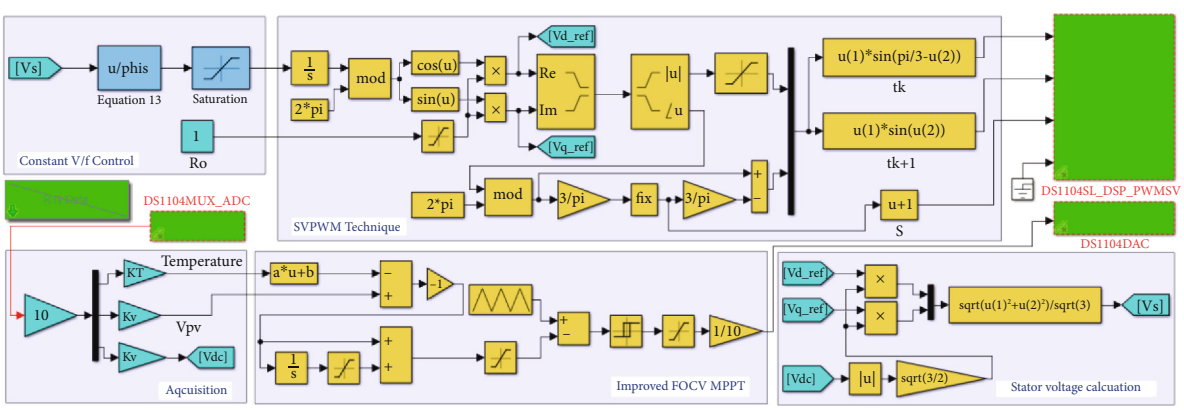


Figure 13: Simulink model of the implemented main RCP program.

Figure 16(a) shows that the efficiency of the boost converter E_{ff} exceeds 91% and reaches 97% for a solar irradiance varying from 500 to 1000 W/m^2. This is confirmed by Figure 15(b) that presents the output power P_{dc} of the boost converter and the available one in the PV generator P_{pv}. The reference voltage is kept constant around 142 V while the voltage at the output of the inverter V_{dc} varies according to the insolation to seek the MPP as shown in Figure 15(c). This objective is reached by the action on the duty cycle D. Figure 15(d) shows that although the system can reach MPP, it shows large oscillations around the reference voltage set at 142 V.

Based on the conventional FOCV MPPT technique, it is observed that the measured voltage varies from 134 to 150 V depending on the variation of the sunlight, leading to an oscillation of 11% compared to the imposed reference voltage. This causes heating of the power components and additional losses, and consequently, some power is lost. On the

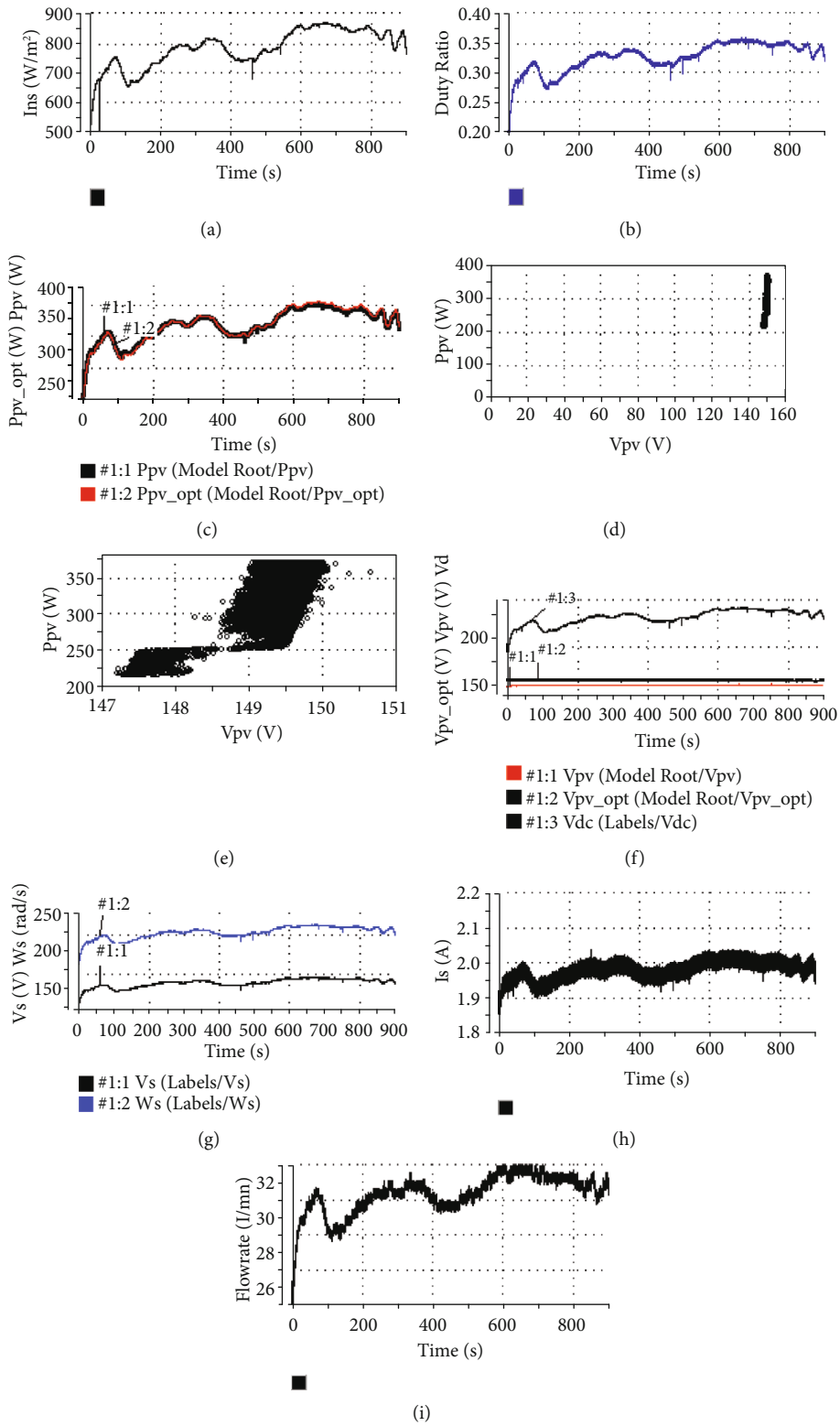


Figure 14: Experimental results of SVPWPS under constant temperature and arbitrary scenario for insolation: (a) evolution of the insolation variation; (b) evolution of the duty cycle; (c) evolution of input and optimal powers; (d) movement of MPP versus the optimum voltage; (e) a zoom of the movement of MPP versus the optimum voltage; (f) evolution of the input and output voltages; (g) stator voltage and pulsation response; (h) stator current response; (i) flow rate evolution.

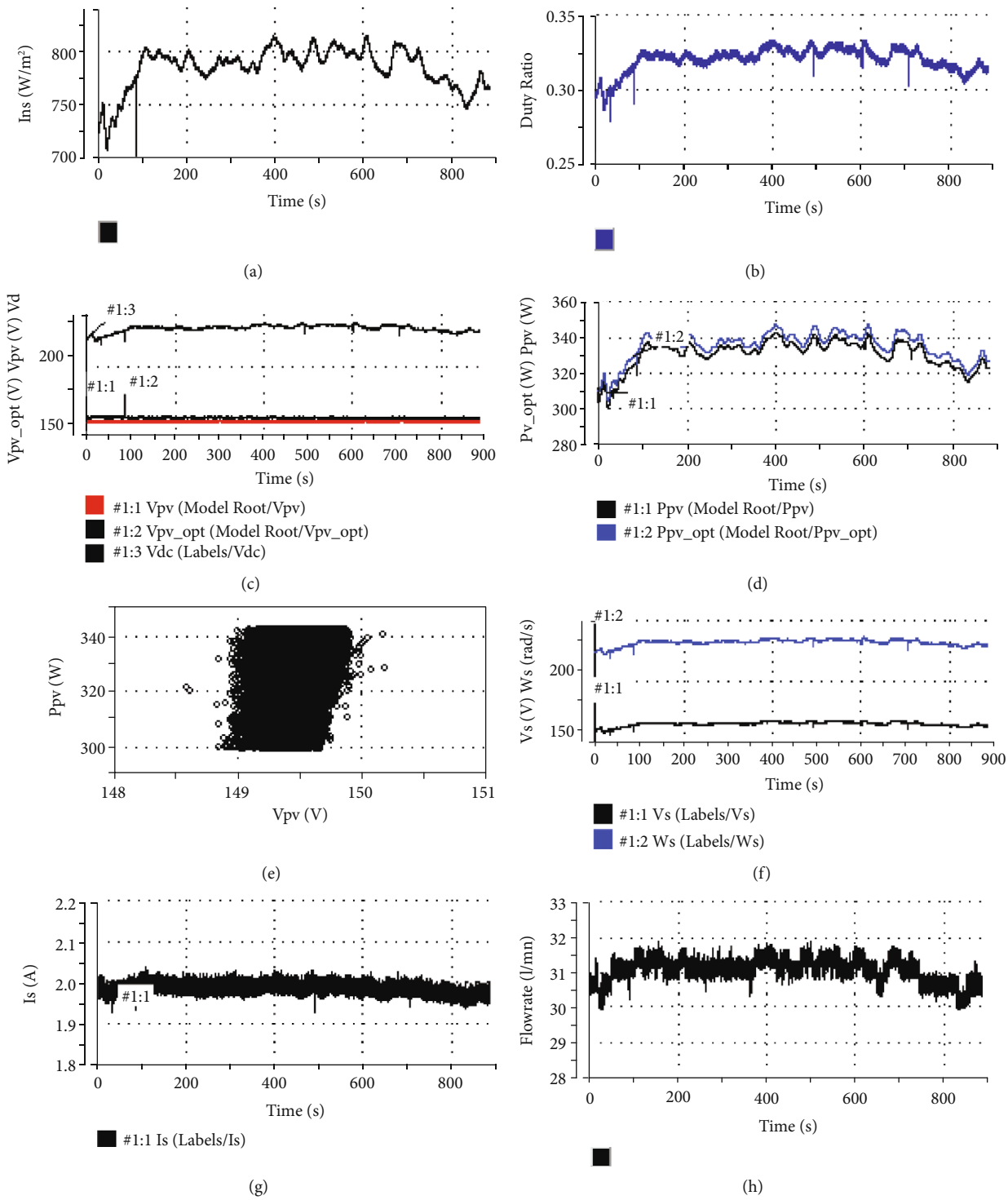


Figure 15: Experimental results of SVPWPS under constant temperature and insolation: (a) evolution of the insolation variation; (b) evolution of the duty cycle; (c) evolution of the input and output voltages; (d) evolution of PV powers; (e) a movement of MPP versus the optimum voltage; (f) stator voltage and pulsation response; (g) stator current response; (h) flow rate evolution.

contrary, by the proposed solution, case of the first and second test, the oscillation of the voltage does not exceed 2.5 V corresponding to 1.6% with respect to the requested optimal voltage. This improvement causes a better flow rate because it increases significantly with the maximum power extracted. Table 5 shows the comparison results of the proposed technique with other classical MPPT techniques. The comparison is done based on previous work [23, 42–44]. And the evaluation based on experimental results shows the superior performance of the proposed algorithm over other techniques in terms of efficiency, tracking speed and accuracy, complexity of algorithm, steady-state oscillation around the

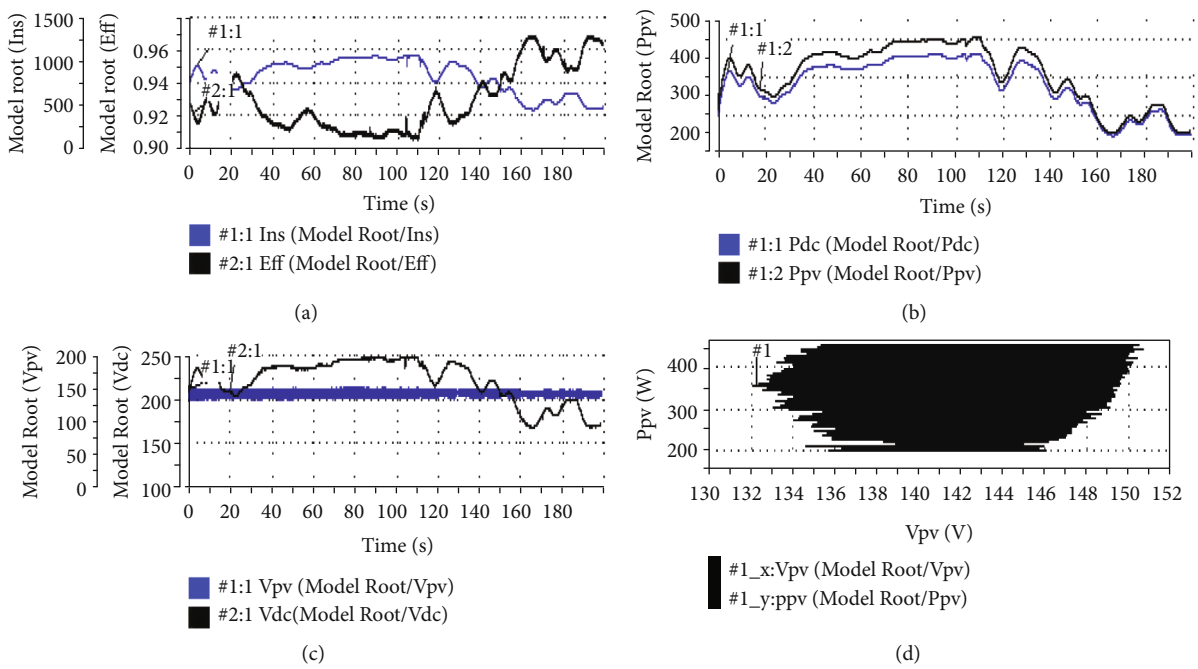


Figure 16: Experimental results of SVPWPS for new reference voltage and arbitrary scenario for insolation: (a) evolution of the insolation and the efficiency of the DC/DC inverter; (b) variation of the input and output powers; (c) evolution of the input and output voltage of the boost converter; (d) movement of MPP versus the PV array voltage.

Table 5: Experimental comparison between the proposed improved-FOCV and MPPT techniques based on classical algorithms.

Technique	Proposed	P&O	IC	FOCV	CV	SCC
Sensed parameters	Temperature	Voltage and current	Voltage and current	Voltage	Voltage	Current
Tracking speed	Fast	Slow	Slow	Slow	Slow	Slow
Tracking accuracy	High	Medium	Medium	Slow	Slow	Medium
Control strategy	Indirect control	Direct control	Direct control	Indirect control	Indirect control	Indirect control
Complexity level	Simple	Medium	Complex	Simple	Simple	Simple
Cost	Inexpensive	Affordable	Expensive	Inexpensive	Inexpensive	Inexpensive
Response to load variation	Faster	Fast	Fast	Slow	Slow	Slow
Steady-state oscillation	Small	Large	Small	Large	Large	Medium
Efficiency	99%	97.8%	98.5%	92.4%	72.8%	93.4%

MPP, cost, sensors needed for implementation, and response to the load variation.

5. Performances during Start-Stop Acquisition

To test the performance of the whole system, a series of measurements are saved at the start and stop of the drive under constant insolation and temperature of 825 W/m^2 and 24°C, respectively. The MPP is tracked immediately after the starting motor. Subsequently, the measurement of PV voltage (V_{pv}), DC bus voltage(V_{dc}), insolation, and global efficiency of the PV system is made and presented in Figure 17. This figure demonstrates that the boost converter operates appropriately and quickly to extract the maximum power according to the improved FOCV MPPT technique adopted and that the global efficiency of the system depending on the available and extracted powers in this condition is 97%. Furthermore, electromagnetic torque (C_{em}), flow rate (Q), and stator current module (I_s) are observed for IM coupled with a centrifugal pump (Figure 18). The electrical variables of the motor (C_{em}) and (I_s) are well controlled by the scalar control, and therefore, the flow rate is also controlled.

The actions on the PV voltage are carried out for two instants, specifically start at 1.5 seconds and stop at 13.5 seconds. The performance of the drive system on starting is adequate, and all parameters reach their equilibrium values directly after starting. Thus, establishing the steady state takes a response time of about one second. It is perceived

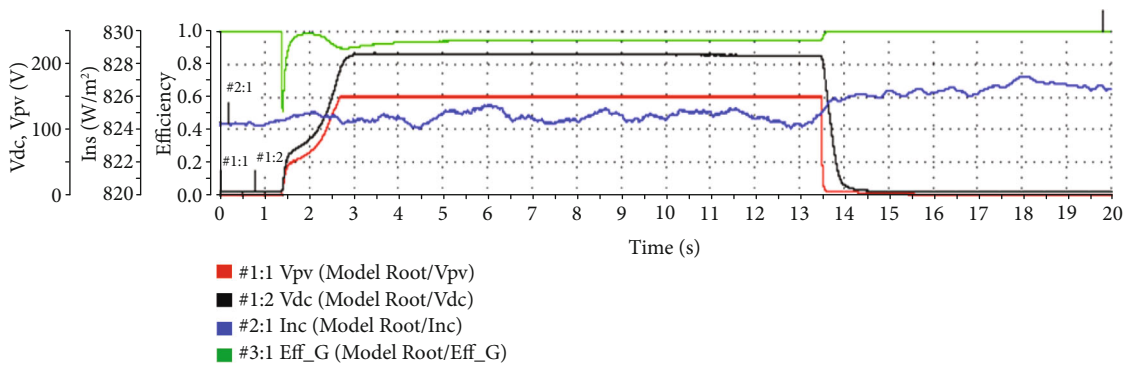


Figure 17: Voltage characteristics.

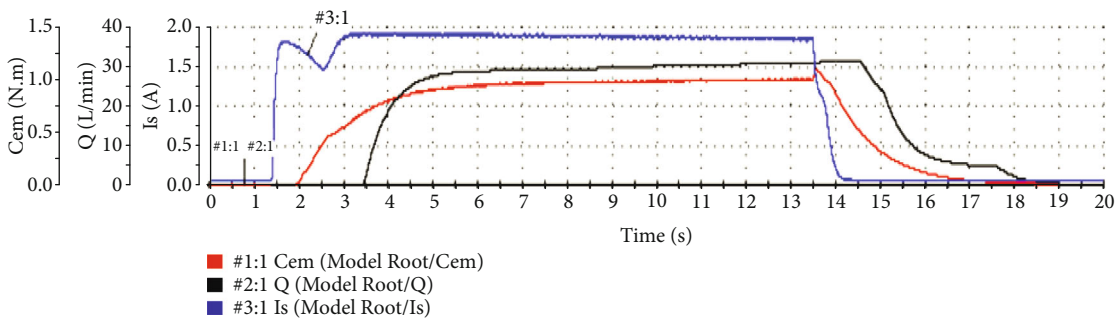


Figure 18: Pump characteristics.

that the source voltage (V_{pv}) is well adjusted to its optimal value of 150 V. The available and extracted powers are practical closes that elucidate an efficiency of about 97%.

For the direct starting of the IM without control, the motor's slip is unity at the time of starting, and the starting stator current is very high, which causes heating problems and a risk of damaging the system. Using the proposed solution, the three-phase inverter is switched off initially because the MPPT algorithm is not running. After starting, the DC-bus voltage is controlled, and the safe starting of IM is obtained by eliminating the possibility of peak current, and the system operates at the MPP. By stopping the simulation of the control MPPT algorithm, the system returns to its initial condition. The flow rate follows the actions of electromagnetic torque and stator current of the IM.

5.1. Steady-State Test. To study the system's performance in a steady-state, a recording is made for 20 seconds at constant insulation and temperature. The characterisation of insolations, V_{pv}, V_{dc}, and efficiency is presented in Figure 19. The boost converter steps up the output voltage to 210 V by action on D's duty cycle to guarantee operation at the MPP. Also, power response to the system specifically P_{pv} and P_{dc} is presented in Figure 20. These powers are at maximum values, and the difference is due to semiconductor losses and the use of measurement sensors. Further, Figure 21 displays the response of C_{em}, Q, and I_s. These variables are constant in this case because the insolation, the temperature, and the load are constant. From the waveforms, it is observed that the PV voltage is settled at MPP voltage, and the variation between input and output power is minimal, which is confirmed by the efficiency of 97%. Moreover, the electromagnetic torque of 1 Nm, the flow rate of 34 l/mn, and the stator current module of 1.8 A are obtained.

Furthermore, the direct and quadratic components I_{ds} and I_{qs} of the stator current I_s and PV current I_{pv} are extracted from the above images and demonstrated in Figure 22. These results are performed in the steady-state condition and ensure that the system is functioning correctly. A Concordia transformation is applied to obtain the current components I_{ds} and I_{qs}. The current I_{pv} shows that the boost converter is operated in CCM operation. From the image, it is observed that the PV current is around 2 A, and the direct and quadratic current components are sinusoidal with 90°C apart from each other.

The scalar control technique controls the steady-state performance of the proposed system. The obtained results prove that this state attained high performance with a flat torque profile, tiny ripples in inductor current, and pure sinusoidal motor current. It shows that the SVPWM technique used to drive the VSI is more convenient than other PWM techniques. Also, scalar control avoids using current sensors for its implementation, unlike FOC and DTC techniques, and therefore, it acts as a simple and inexpensive structure. On the other hand, with the proposed MPPT technique in steady state, the maximum power is extracted, and the flow rate is at the maximum level.

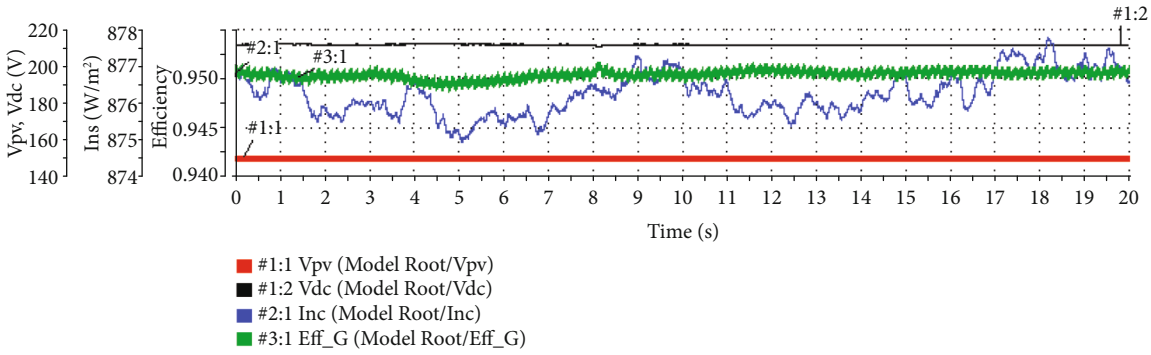


Figure 19: Steady-state voltage characteristics.

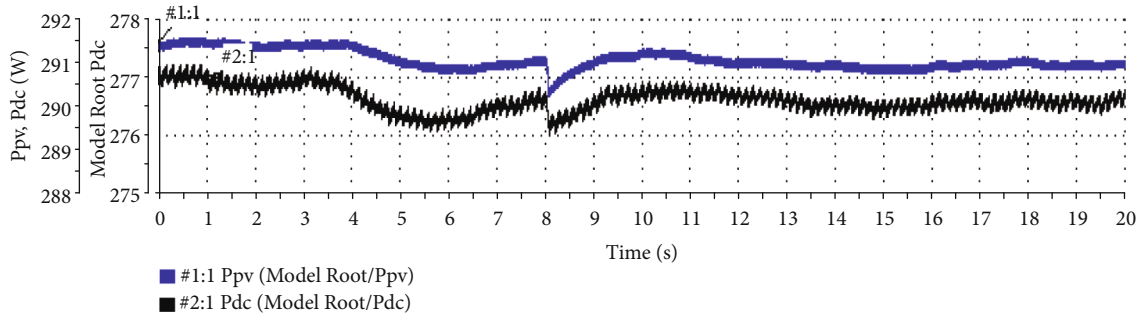


Figure 20: Steady-state power characteristics.

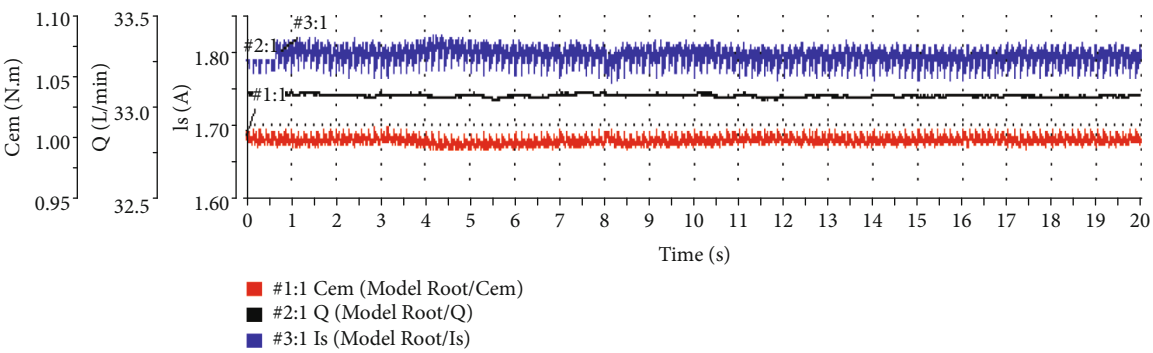


Figure 21: Steady-state pump characteristics.

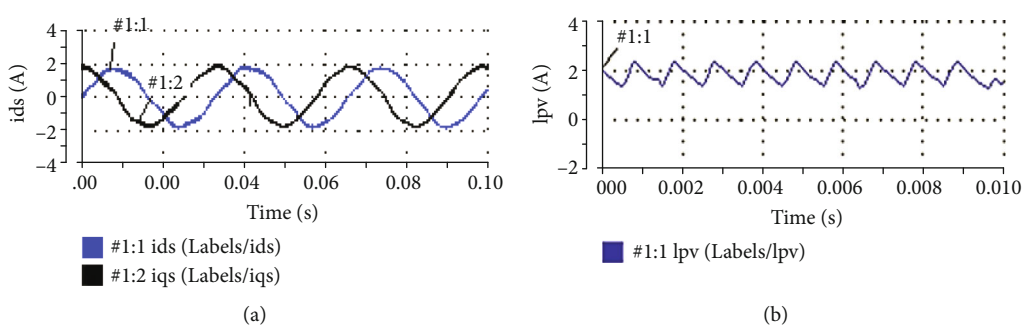


Figure 22: Currents response: (a) zoom on direct and quadratic components I_{ds} and I_{qs}; (b) zoom on PV array current I_{pv}.

Consolidating these results, the main objective of this application is to investigate an efficient, easy-to-use, high-performance SPVWPS to allow the use of electricity in rural areas disconnected from the grid or in the case of a significant energy deficit that negatively affects agricultural production. To accelerate the design process and prove the ability of the proposed algorithm, an RCP methodology is applied. A set of electronic circuits are used in the laboratory to complete this solution: sensors (temperature, insolation, voltages, and currents), TTL-CMOS circuit driver for boost DC/DC, VSI, and flow meter circuit. The key benefits of RCP are increasing rapidly because manual programming has more disadvantages; the cost is reduced compared to the amount of work that would be required to perform a similar test setup from scratch. The control function can be managed in an easy and fast way, and the evaluation is instantaneous. Lastly, the efficiency of RCP is proven, which clearly shows its effectiveness from the test above. Specifically, the withdrawn power from the PV array is at its maximum value for all variations in climatic conditions.

6. Conclusions

The design and implementation of an off-grid SPVWPS are proposed in this work. This system is operated using an RCP platform that integrates an improved FOCV MPPT method and closed-loop scalar control. The mathematical models of the PV array, DC-DC boost converter, three-phase VSI, and IM coupled to centrifugal pump are described in detail. The following conclusions are made:

(i) The proposed scalar control allows controlling the induction motor according to the maximum power available in the PV panel and eliminates extra sensors for speed and current measurement

(ii) An improved fractional open circuit voltage MPPT method enhances the efficiency of the conventional algorithm

(iii) The experimental results displayed the accurate performance of the proposed solution in terms of global efficiency optimization of the SPVWPS, better stability for the exact parameters of the system under start and stop acquisition, and water discharge rates in the steady-state operation

(iv) The effectiveness of the chosen MPPT method for the optimization of the photovoltaic pumping system is tested, and correct steady-state and dynamic operation, including severe solar insolation variations, have been obtained for the whole system

Therefore, the proposed system can be potentially used to design an off-grid solar photovoltaic water pumping system.

Data Availability

The data used to support the findings of this study are included in the article.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

References

[1] F. Johnsson, J. Kjärstad, and J. Rootzén, "The threat to climate change mitigation posed by the abundance of fossil fuels," *Climate Policy*, vol. 19, no. 2, pp. 258–274, 2019.

[2] R. Madurai Elavarasan, L. Selvamanohar, K. Raju et al., "A holistic review of the present and future drivers of the renewable energy mix in Maharashtra, state of India," *Sustainability*, vol. 12, no. 16, p. 6596, 2020.

[3] A. K. Mishra and B. Singh, "Design of solar-powered agriculture pump using new configuration of dual-output buck–boost converter," *IET Renewable Power Generation*, vol. 12, no. 14, pp. 1640–1650, 2018.

[4] C. Ramulu, P. Sanjeevikumar, R. Karampuri, S. Jain, A. H. Ertas, and V. Fedak, "A solar PV water pumping solution using a three-level cascaded inverter connected induction motor drive," *Engineering Science and Technology, an International Journal*, vol. 19, no. 4, pp. 1731–1741, 2016.

[5] R. Antonello, M. Carraro, A. Costabeber, F. Tinazzi, and M. Zigliotto, "Energy-efficient autonomous solar water-pumping system for permanent-magnet synchronous motors," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 1, pp. 43–51, 2017.

[6] U. Sharma, B. Singh, and S. Kumar, "Intelligent grid interfaced solar water pumping system," *IET Renewable Power Generation*, vol. 11, no. 5, pp. 614–624, 2017.

[7] S. Murshid and B. Singh, "Implementation of PMSM drive for a solar water pumping system," *IEEE Transactions on Industry Applications*, vol. 55, no. 5, pp. 4956–4964, 2019.

[8] M. Errouha, A. Derouich, B. Nahid-Mobarakeh, S. Motahhir, and A. El Ghzizal, "Improvement control of photovoltaic based water pumping system without energy storage," *Solar Energy*, vol. 190, pp. 319–328, 2019.

[9] A. K. Mishra and B. Singh, "Grid interactive single-stage solar powered water pumping system utilizing improved control technique," *IEEE Transactions on Sustainable Energy*, vol. 11, no. 1, pp. 304–314, 2020.

[10] C. Liu and Y. Luo, "Overview of advanced control strategies for electric machines," *Chinese Journal of Electrical Engineering*, vol. 3, 2017.

[11] O. Ellabban, J. Van Mierlo, and P. Lataire, "A comparative study of different control techniques for an induction motor fed by a Z-source inverter for electric vehicles," in *2011 International Conference on Power Engineering, Energy and Electrical Drives*, pp. 1–7, Malaga, Spain, 2011.

[12] I. Boldea, "Control issues in adjustable speed drives," *IEEE Industrial Electronics Magaziner*, vol. 2, no. 3, pp. 32–50, 2008.

[13] T. H. Dos Santos, A. Goedtel, S. A. O. Da Silva, and M. Suetake, "Scalar control of an induction motor using a neural sensorless technique," *Electric Power Systems Research*, vol. 108, pp. 322–330, 2014.

[14] Chun-Chieh Wang and Chih-Hsing Fang, "Sensorless scalar-controlled induction motor drives with modified flux observer," *IEEE Transactions on Energy Conversion*, vol. 18, no. 2, pp. 181–186, 2003.

[15] A. R. Jordehi, "Maximum power point tracking in photovoltaic (PV) systems: a review of different approaches," *Renewable and Sustainable Energy Reviews*, vol. 65, pp. 1127–1138, 2016.

[16] A. K. Podder, A. K. Das, E. Hossain et al., "Integrated Modeling and Feasibility Analysis of a Rooftop Photovoltaic Systems for an Academic Building in Bangladesh," *International Journal of Low-Carbon Technologies*, 2021.

[17] D. Singh, R. Chaudhary, and A. Karthick, "Review on the progress of building-applied/integrated photovoltaic system," *Environmental Science and Pollution Research*, vol. 28, no. 35, pp. 47689–47724, 2021.

[18] C. Pazhanimuthu, I. Baranilingesan, and A. Karthick, "An improved control algorithm for series hybrid active power filter based on SOGI-PLL under dynamic load conditions," *Solid State Communications*, vol. 333, 2021.

[19] N. M. Kumar, M. Samykano, and A. Karthick, "Energy loss analysis of a large scale BIPV system for university buildings in tropical weather conditions: a partial and cumulative performance ratio approach," *Case Studies in Thermal Engineering*, vol. 25, 2021.

[20] O. Singh and S. K. Gupta, "A review on recent MPPT techniques for photovoltaic system," in *2018 IEEMA Engineer Infinite Conference (eTechNxT)*, pp. 1–6, New Delhi, India, 2018.

[21] V. Salas, E. Olías, A. Barrado, and A. Lázaro, "Review of the maximum power point tracking algorithms for stand-alone photovoltaic systems," *Solar Energy Materials and Solar Cells*, vol. 90, no. 11, pp. 1555–1578, 2006.

[22] A. R. Prasad, R. Shankar, C. K. Patil, A. Karthick, A. Kumar, and R. Rahim, "Performance enhancement of solar photovoltaic system for roof top garden," *Environmental Science and Pollution Research*, vol. 28, no. 36, pp. 50017–50027, 2021.

[23] R. Kabilan, V. Chandran, J. Yogapriya et al., "Short-term power prediction of building integrated photovoltaic (BIPV) system based on machine learning algorithms," *International Journal of Photoenergy*, vol. 2021, 11 pages, 2021.

[24] N. Y. Jayalakshmi, R. Shankar, U. Subramaniam et al., "Novel multi-time scale deep learning algorithm for solar irradiance forecasting," *Energies*, vol. 14, no. 9, p. 2404, 2021.

[25] A. Hmidet, N. Rebei, and O. Hasnaoui, "Experimental studies and performance evaluation of MPPT control strategies for solar-powered water pumps," in *2015 Tenth International Conference on Ecological Vehicles and Renewable Energies (EVER),*, pp. 1–12, Monte Carlo, Monaco, 2015.

[26] V. S. Chandrika, M. M. Thalib, A. Karthick et al., "Performance assessment of free standing and building integrated grid connected photovoltaic system for southern part of India," *Building Services Engineering Research and Technology*, vol. 42, no. 2, pp. 237–248, 2021.

[27] M. A. Eltawil and Z. Zhao, "MPPT techniques for photovoltaic applications," *Renewable and Sustainable Energy Reviews*, vol. 25, pp. 793–813, 2013.

[28] S. Kıvrak, T. Özer, and Y. Oğuz, "Design and Implementation of dspic33fj32mc204 Microcontroller–Based Asynchronous Motor Voltage/Frequency Speed Control Circuit for the Ventilation Systems of Vehicles," *Measurement and Control*, vol. 52, no. 7-8, pp. 1039–1047, 2019.

[29] L. Jun, G. Ying-Qing, and W. Hai-Quan, "Rapid prototyping real-time simulation platform for digital electronic engine control," in *2008 2nd International Symposium on Systems and Control in Aerospace and Astronautics*, pp. 1–5, Shenzhen, China, 2008.

[30] A. Hmidet and O. Boubaker, "Real-Time Low-Cost Speed Monitoring and Control of Three-Phase Induction Motor via a Voltage/Frequency Control Approach," *Mathematical Problems in Engineering*, vol. 2020, Article ID 6913813, 14 pages, 2020.

[31] J. Aravena, D. Carrasco, M. Diaz et al., "Design and implementation of a low-cost real-time control platform for power electronics applications," *Energies*, vol. 13, no. 6, 2020.

[32] S. Wendel, A. Geiger, E. Liegmann et al., "UltraZohm - a powerful real-time computation platform for MPC and multi-level inverters," in *2019 IEEE International Symposium on Predictive Control of Electrical Drives and Power Electronics (PRECEDE)*, pp. 1–6, Quanzhou, China, 2019.

[33] A. Hmidet, R. Dhifaoui, and O. Hasnaoui, "Development, implementation and experimentation on a dSPACE DS1104 of a direct voltage control scheme," *Journal of Power Electronics*, vol. 10, no. 5, pp. 468–476, 2010.

[34] A. R. Jordehi, "Parameter estimation of solar photovoltaic (PV) cells: a review," *Renewable and Sustainable Energy Reviews*, vol. 61, pp. 354–371, 2016.

[35] P. Sahu, D. Verma, and S. Nema, "Physical design and modelling of boost converter for maximum power point tracking in solar PV systems," in *2016 International Conference on Electrical Power and Energy Systems (ICEPES)*, pp. 10–15, Bhopal, India, 2016.

[36] M. Lasheen, A. K. Abdel Rahman, M. Abdel-Salam, and S. Ookawara, "Adaptive reference voltage-based MPPT technique for PV applications," *IET Renewable Power Generation*, vol. 11, no. 5, pp. 715–722, 2017.

[37] B. Singh, U. Sharma, and S. Kumar, "Standalone photovoltaic water pumping system using induction motor drive with reduced sensors," *IEEE Transactions on Industry Applications*, vol. 54, no. 4, pp. 3645–3655, 2018.

[38] A. Hmidet, R. Dhifaoui, and O. Hasnaoui, "A new direct speed estimation and control of the induction machine benchmark: design and experimental validation," *Mathematical Problems in Engineering*, vol. 2018, Article ID 9215459, 10 pages, 2018.

[39] V. K. Arun Shankar, S. Umashankar, S. Paramasivam, and N. Hanigovszki, "A comprehensive review on energy efficiency enhancement initiatives in centrifugal pumping system," *Applied Energy*, vol. 181, pp. 495–513, 2016.

[40] B. Singh and S. Shukla, "Induction Motor Drive for PV Water Pumping with Reduced Sensors," *IET Power Electronics*, vol. 11, no. 12, pp. 1903–1913, 2018.

[41] A. Smith, S. Gadoue, M. Armstrong, and J. Finch, "Improved method for the scalar control of induction motor drives," *IET Electric Power Applications*, vol. 7, no. 6, pp. 487–498, 2013.

[42] L. Xu, R. Cheng, and J. Yang, "A new MPPT technique for fast and efficient tracking under fast varying solar irradiation and load resistance," *International Journal of Photoenergy*, vol. 2020, Article ID 6535372, 18 pages, 2020.

[43] A. Nadeem, H. A. Sher, and A. F. Murtaza, "Online fractional open-circuit voltage maximum output power algorithm for photovoltaic modules," *IET Renewable Power Generation*, vol. 14, no. 2, pp. 188–198, 2020.

[44] R. B. Bollipo, S. Mikkili, and P. K. Bonthagorla, "Hybrid, optimization, intelligent and classical PV MPPT techniques: review," *CSEE Journal of Power and Energy Systems*, vol. 7, 2020.

Hindawi
International Journal of Photoenergy
Volume 2021, Article ID 1211953, 22 pages
https://doi.org/10.1155/2021/1211953

Research Article

Standalone and Minigrid-Connected Solar Energy Systems for Rural Application in Rwanda: An In Situ Study

Kuo-Chi Chang,[1,2,3] **Noel Hagumimana**,[4] **Jishi Zheng**,[4] **Godwin Norense Osarumwense Asemota**,[5] **Jean De Dieu Niyonteze**,[6] **Walter Nsengiyumva**,[7] **Aphrodis Nduwamungu**,[5] **and Samuel Bimenyimana**[8,9]

[1]*Department of Applied Intelligent Mechanical and Electrical Engineering, Yu Da University of Science and Technology, Miaoli County, Taiwan*
[2]*Department of Business Administration, North Borneo University College, Sabah, Malaysia*
[3]*School of Electronic, Electrical and Physics, Fujian University of Technology, Fuzhou, China*
[4]*Fujian Province Key Laboratory of Automotive Electronics and Electric Drive, Fujian University of Technology, Fuzhou 350118, China*
[5]*University of Rwanda, African Centre of Excellence in Energy for Sustainable Development, Kigali 4285, Rwanda*
[6]*Carnegie Mellon University Africa, Kigali, Rwanda*
[7]*Laboratory of Optics, Terahertz and Non-destructive Testing, School of Mechanical Engineering and Automation, Fuzhou University, Fuzhou 350108, China*
[8]*Huaqiao University, Intelligence and Automation in Construction Provincial Higher-Educational Engineering Research Centre, 361021 Xiamen, China*
[9]*Hello Renewables Ltd., Kigali, Rwanda*

Correspondence should be addressed to Kuo-Chi Chang; albertchangxuite@gmail.com and Noel Hagumimana; hagumanoel@gmail.com

Received 3 May 2021; Revised 26 July 2021; Accepted 6 September 2021; Published 5 October 2021

Academic Editor: James Connolly

In recent years, several factors such as environmental pollution, declining fossil fuel supplies, and product price volatility have led to most countries investing in renewable energy sources. In particular, the development of photovoltaic (PV) microgrids, which can be standalone, off-grid connected or grid-connected, is seen as one of the most viable solutions that could help developing countries such as Rwanda to minimize problems related to energy shortage. The country's current electrification rate is estimated to be 59.7%, and hydropower remains Rwanda's primary source of energy (with over 43.8% of its total energy supplies) despite advances in solar technology. In order to provide affordable electricity to low-income households, the government of Rwanda has pledged to achieve 48% of its overal electrification goals from off-grid solar systems by 2024. In this paper, we develop a cost-effective power generation model for a solar PV system to power households in rural areas in Rwanda at a reduced cost. A performance comparison between a single household and a microgrid PV system is conducted by developing efficient and low-cost off-grid PV systems. The battery model for these two systems is 1.6 kWh daily load with 0.30 kW peak load for a single household and 193.05 kWh/day with 20.64 kW peak load for an off-grid PV microgrid. The hybrid optimization model for electric renewable (HOMER) software is used to determine the system size and its life cycle cost including the levelized cost of energy (LCOE) and net present cost (NPC) for each of these power generation models. The analysis shows that the optimal system's NPC, LCOE, electricity production, and operating cost are estimated to 1,166,898.0 USD, 1.28 (USD/kWh), 221, and 715.0 (kWh per year, 37,965.91 (USD per year), respectively, for microgrid and 9284.4(USD), 1.23 (USD/kWh), and 2426.0 (kWh per year, 428.08 (USD per year), respectively, for a single household (standalone). The LCOE of a standalone PV system of an independent household was found to be cost-effective compared with a microgrid PV system that supplies electricity to a rural community in Rwanda.

1. Introduction

Small electricity systems that can run independently, known as off-grid microgrids, could play a pivotal role in the development of electricity systems based on decentralized renewable energy (RE) technologies. These networks are more cost-effective than stretching transmission lines to rural places [1, 2], thereby providing the possibility to produce sufficient electricity in countries where the national demand surpasses the regular production. In East Africa, for example, the energy deficiency is a significant impediment to social and economic growth. The capital expenses of wide grids can be incredibly expensive for developing countries leading to a shortage of roads and utilities [3–5]. To this end, community-based microgrids are considered the best option that would help rural areas in developing countries to reap the benefits of geospecific renewable energy sources.

The people who are not connected to nor served by the public or private power grid are referred to as "off-grid users." According to the authors' definition in Ref. [6], the term "off-grid" refers to a system and way of life that allows people to function without the assistance of remote infrastructure, including an electrical grid. It is a method of gaining access to electricity that is used in countries and areas where there is limited access to electricity due to a dispersed or remote population. It corresponds to living without relying on one or more public services, commonly referred to as electrical grids. Off-grid users are people who live off the grid, and those systems can be categorized as standalone power systems, minigrids, and microgrids, which should typically provide energy to a smaller community. In this research, the HOMER software (HOMER Pro, version 3.13.1) had been used to model, simulate, and optimize potential renewable energy sources, as well as solutions to ensure universal access to energy Rwandan for off-grid users. HOMER has a built-in optimizer through the proprietary derivative-free method that was adopted in Section 2. The simulation models took place in Rwanda's Western province (Rutsiro, Rwanda, 1°56.3′S, 29°19.5′E). To optimize standalone solar systems, several site visits were conducted in the Rutsiro district of Rwanda's Western province, precisely to the location of the typical sample residential house used in this section. The owner of the residential house listed his electric household items, along with their rated power and daily usage hours. For this study, a sample size of 121 residential houses was chosen.

2. Literature Review: A Comparative Analysis of Standalone and Minigrid-Connected Solar Energy in a Rural Area

With the mounting consequences of global warming, pollution, scarcity of fuel, and energy use, renewable energy sources (RES) is constantly getting more attention around the world. Therefore, the need for renewable energy to plan and build a grid-connected or standalone, microgrid, the minigrid system has risen and will continue to increase. When the price of conventional energy is compared to the price of renewable energy, renewable energy is much less costly [7]. Given that many of Africa's rural areas are plagued by an unsustainable energy system, building standalone, minigrid can solve energy problems for scattered people [8]. In developing nations like Rwanda, where power outages are common, implementing supportable energy development and clean energy needs extensive preparation, particularly given the financial impacts. As a result, the HOMER (hybrid optimization model of electric renewable) Pro software can design, prepare, and simulate the model in a variety of environments, including restrained and unrestrained systems, standalone, grid systems, and/or storage. The microgrid system design has advantages that lead to efficient source loading for microgrids and facilitate power systems operators. HOMER's benefits lie in the features used in the design, planning, and simulation of the microgrid model discussed in [9, 10]. In contrast to African countries, developed countries like the United States, China, and Japan have increased their investment in renewable energy by billions of dollars. Researchers are attempting to produce more electricity from cost-effective resources that are not detrimentally impacting the environment [11].

HOMER was used to examining selected rural places in Nigeria based on the availability of wind and solar energies so that healthcare centers or clinics in isolated regions can provide quick delivery of medical services to the people who need them. It uses the best technical and economic design and sizing of hybrid electric power system components like wind, PV, battery, and inverter systems, where PV/wind/diesel/battery hybrid setup is best for rural health centers, while PV/diesel/battery hybrid systems are best for Port Harcourt considering the quality of renewable energy potential [12].

Tourist destinations in the South China Sea, Malaysia were at risk due to the widespread use of diesel generators and pollutants from diesel-based power plants. HOMER software was used for economic and technical analysis of the system. The best optimized standalone hybrid energy system consists of PV, wind, diesel generator, converter, and battery. The output has proved the diesel-only system has a higher net present cost, cost of energy, and CO_2 emission compared to the optimized hybrid renewable energy system [13].

The study on decentralized power stations in Sabah, Malaysia [14], with a diverse combination of photovoltaic (PV), diesel generators, system converters, and storage batteries. The impact of PV integration using HOMER was properly quantified by analyzing the practical behaviors of different PV penetration levels. The analysis based on technical, economic, and environmental constraints has resulted in satisfying the load demand with the minimum total net present cost (NPC) and the levelized cost of energy (LCOE). The sensitivity analysis and the impact of different PV penetration levels on the system performance and the generation of harmful emissions has been carried out. The findings reveal an increase in the use of renewable energy (RE) sources in energy generation, as well as a decrease in the reliance on standalone diesel generators.

The ability to supply power for rural health clinics (RHC) in six geopolitical regions of Nigeria has also been

Table 1: Summary of comparative analysis based on standalone, microgrid, on-grid, and off-grid results.

S. no.	Authors & references	Year	Location	Adopted technologies	Load type	Consumption type	Method	Objectives
1.	M.K Deshmukh, Athokpam Bharatbushan Singh [23]	2018	—	Standalone	Street lighting	Electrical	HOMER	The objective is to quantitatively estimate energy losses due to the standalone operation mode
2.	U Subramaniam et al. [24]	2020	—	On-grid and off-grid	Villages, islands, and hilly areas	Electrical	Hybrid PV battery with controller	The current method can work in various operating modes, and during transient and steady-state situations. With both off-grid and on-grid situations, the suggested power management controls were approved.
3.	C Marino et al. [25]	2020	Italy	Standalone photovoltaic	Residential user	Electrical	Comparative analysis of the costs of a standalone and a grid-networked PV system vs. grid distance	The study looked at the economics of an islanded PV project with two configurations that measure diminishing self-sufficiency.
4.	MH Mohamed Hariri et al. [26]	2020	—	Grid-connected	Villages, islands	Electrical	Grid synchronization and islanding detection methods	This review highlights the recent development of systems for generating grid-connected PV (GPV) involving many sub-components, like DC-DC converters, PV modules, maximum power point tracking (MPPT), and inverter technologies
5.	FA Alturki, EM Awwad [27]	2020	Saudi Arabia	Standalone	Remote community	Electrical	Hybrid photovoltaic (PV)/wind turbine (WT)/biomass/pump hydro/storage	The aim was sizing and price reduction of islanded hybrid WT/PV/biomass/pump-hydro storage-energy systems
6.	T Wu et al. [28]	2020	—	Grid-connected	Load serving	Electrical	Salp swarm algorithm (SSA).	This study provides a new approach to maximizing the scale of grid-connected renewable energy sources integrated with the salp swarm algorithm (SSA) pumped storage system. This method enables different energy sources to be explored and their combination to contact the base in the optimum configuration of the hybrid system
7.	BE Türkay, AYTelli renewable energy [29]	2011	Turkey	Standalone and grid-connected	Pilot area	Electrical	HOMER	The research explores the viability of using wind and solar energy. Using hydrogen as storage in combination with traditional grid-based electricity to fulfill the

Table 1: Continued.

S. no.	Authors & references	Year	Location	Adopted technologies	Load type	Consumption type	Method	Objectives
								electricity needs of the pilot area
8.	D Mazzeo et al. [30]	2020	Koppen	Standalone and grid-connected	Office building district.	Electrical	Hybrid renewable system	The goal of this work is to bridge the absence of direct comparisons between the technoeconomic output of islanded and grid-networked investigations in the same operating environment, providing global technoeconomic mapping, and optimizing islanded and grid-networked PV-wind systems
9.	A Chakir et al. [31]	2019	—	Grid-connected	Load serving	Electrical	MATLAB/Simulink	The research focused on the grid-connected development system's management, connection with the grid, and storage hybrid renewable energy system's management
10.	R Srivastava et al. [32]	2020	—	Standalone and grid-connected	Load serving	Electrical	Review	The various cases were examined based on their location, design and year of development, as well as the power, the technology used, and performance that can help design a PV plant considering the achievements of the previously commissioned plant. The material of the PV module and panel tilt angle was found to be crucial for the design of a PV plant
11.	AC Duman, Ö Güler [33]	2020	Turkey	Grid-connected	Load serving	Electrical	HOMER	The study focused on a cost-benefit analysis of grid-affiliated rooftop PV systems for private use. There was the suggestion to increase the number of private PV incentives and cultivate a regional support system, considering solar differences among regions
12.	HM Ridha et al. [34]	2020	—	Standalone photovoltaic	Remote areas	Electrical	Review	The purpose of the research was to provide a thorough analysis of the recent progress in the design of standalone PV systems. Multiobjective optimization (MOO) and multicriteria decision-making (MCDM)

Table 1: Continued.

S. no.	Authors & references	Year	Location	Adopted technologies	Load type	Consumption type	Method	Objectives
								methodologies, including the mathematical models used to measure the PV module power output and storage battery
13.	MJ Mayer et al. [35]	2019	Hungarian region	Grid-connected	Load serving	Electrical use	Mathematical model	Grid-connected, ground-mounted technoeconomic optimization of genetic algorithm-based photovoltaic power plants on a comprehensive mathematical model. The target function is the internal rate of return and a genetic algorithm performs the optimization
14.	HA Kazem et al. [36]	2020	Oman	Grid-connected	Load management	Electrical	X-ray diffraction (XRD) and X-ray fluorescence (XRF)	Dust is one major parameter affecting photovoltaic efficiency, yield, and profitability linked to the grid. The proposed model in the paper took account of the dust on the grid-affiliated photovoltaic output power innovatively
15.	E Aykut, ÜK Terzi [37]	2020	Marmara University, Turkey	Grid-connected hybrid	Load serving	Electrical	HOMER	The research focused on technology, cost-benefit, and environmental analyses of grid-affiliated hybrid wind/PV/biomass systems, Marmara University, Goztepe campus. The performance of the hybrid electricity system was assessed using both the net present cost (NPC) and cost of energy (COE) and found to be cheaper
16.	R Khezri et al. [38]	2020	Australia	Grid-connected	Households serving	Electrical		This research specifies the optimum solar capacity for grid-affiliated households, photovoltaic (PV), and battery energy storage (BES) to minimize the net present cost of electric power networks
17.	BK Das [39]	2020	Bangladesh	Standalone and grid-connected	Load serving	Electrical	HOMER	The study sized an islanded and grid-affiliated solar PV electricity provision to a small neighborhood. The result reflects major cost savings through the incorporation of the PV module into the grid

Table 1: Continued.

S. no.	Authors & references	Year	Location	Adopted technologies	Load type	Consumption type	Method	Objectives
18.	ALM Maher [40]	2019	Palestine	Grid-connected and standalone	Industrial zone	Electrical	Open distribution source simulator (OpenDSS)	This study provided a layout for a grid-affiliated PV system and an islanded PV system. Factors influencing device design and size were also described and analyzed. The results showed a good improvement in the overall energy losses and voltage profile concerning load and capacity production
19.	HA Attia, F delAma Gonzalo [41]	2018	United Arab Emirates	Standalone	Remote building	Electrical	Fuzzy logic control	It provided a precise Buck-Boost DC-DC converter design powered by the fuzzy logic controller (FLC). The research concentrated on suggesting a suitable solar PV panel model configuration and attachment
20.	A Iqbal, MT Iqbal [42]	2019	Pakistan	Standalone PV	Rural area	Electrical	HOMER Pro	Thermal modeling of a typical Pakistani rural house was performed using BEopt throughout this research to assess the hourly load profile. System studies indicated that such a system can primarily reinforce the lighting and loading of appliances in a rural household
21.	Y Chaibi et al. [43]	2019	—	Standalone PV	Load serving	Electrical	Sliding mode MPPT/ MATLAB Simulink	The study tracked how the maximum power point (MPP) transferred the maximum available power to the load. The control fed the AC load by a sinusoidal output current
22.	MA Omar, MM Mahmoud [44]	2018	Palestine	Grid-connected and standalone	Residential sector	Electrical	Unconventional PV system/MATLAB software	The study proposed a viable solution to the problem of energy output in the private dwellings sector using unpredictable PV systems operating in islanded and grid-affiliated modes. The battery storage framework enables private dwellings to secure stable energy operations
23.	PK Bonthagorla, S Mikkili [45]	2020	—	Grid-connected/ standalone	Load serving	Electrical	MATLAB/Simulink	This paper briefly discusses the modeling, simulation, and performance evaluation of hybrid and conventional

Table 1: Continued.

S. no.	Authors & references	Year	Location	Adopted technologies	Load type	Consumption type	Method	Objectives
								array configurations during different PSCs in MATLAB/Simulink environment
24.	M Salimi et al. [46]	2021	—	Hybrid grid-connected	Load serving	Electrical	MATLAB/Simulink	A fresh concept to active modeling and locked-loop control of hybrid grid-affiliated multi-input multioutput (MIMO) renewable energy systems was addressed
25.	M Dali et al. [47]	2010	—	Grid-connected and standalone modes	Load serving	Electrical	Standalone inverter	A hybrid system associated with the grid was described in the study. The experimental findings show that the system can operate parallel to or independent of the grid
26.	MI Hlal et al. [48]	2019	Malaysia	Off-grid or standalone	Load management	Electrical	Non-dominated separating genetic algorithm (NSGA-II) method	A multiobjective optimization design accounted for lossy load likelihood (LLP), energy cost (COE), price of battery life loss, and cost of service, substitution, and repair.
27.	KNB Akshai, R Senthil [49]	2020	—	Standalone	Household electricity	Electrical	Simulation software PVsyst	The study is based on the evaluation of economic expenses of grid-affiliated and islanded photovoltaic systems using PVsyst
28.	S Odeh et al. [50]	2019	Palestinian	Standalone	Load serving	Electrical	PVsyst software	This research proposes a hybrid system consisting of an array of photovoltaic (PV) and rechargeable batteries integrated into the distribution grid to share loads with the grid system
29.	J Kumar et al. [51]	2020	—	PV grid-tied	Residential load serving	Electrical	System advisor model (SAM) software	The research resolved the power grid stability and control problems. The PV grid system consists of an 8.0 kW PV array and battery energy storage unit connected to the power grid over AC or DC links
30.	J Kumar [52]	2020	—	Grid-connected	Island	Electrical	PVsyst software	This study mainly explores the design features of a solar photovoltaic device based on a grid connection. The analysis sheds light on various subjects like creating Sankey energy loss diagrams, efficiency

Table 1: Continued.

S. no.	Authors & references	Year	Location	Adopted technologies	Load type	Consumption type	Method	Objectives
31.	ZB Duranay, H Guldemir [53]	2019	—	Standalone	Load management	Water pumping	MATLAB/Simulink	proportion, and total photovoltaic plant capacity A water-pumping double-deck converter and inverter for a single-phase islanded PV system were investigated. The single-phase islanded PV system was modeled using insolation and temperature values as simulation data
32.	Y Cui et al. [54]	2020	—	Grid-connected	Domestic building	Electricity	@risk software	The study was a technoeconomic evaluation of grid-affiliated residential construction applications photovoltaic (PV) systems. The system met the residential electricity demand from April to October and the 1530.23 kWh excess electricity was supplied to the grid
33.	N Gupta et al. [55]	2017	—	Grid-connected	Load serving	Electrical	Pareto analysis and logic gate representations	The research developed grid-connected PV system sensitivity and reliability models. For PV cell and DC-DC converters, analytical relationships of first-order sensitivity are formed and the developed models can be implemented to any PV system for better performance
34.	MP Bonkile, V Ramadesigan [56]	2019	—	Standalone	Load management	Physics-based battery	Single-particle model (SPM)	For an islanded PV-battery energy storage (BES) hybrid device, a power management control strategy is suggested in the research. The evaluation shows that the power management design was successful and met many islanded PV-BES hybrid systems goals, without overcharging, no output excess power generation, and no power transfer to the dump load
35.	SS Dheeban et al. [57]	2019	—	Standalone	Load management	Electrical	MATLAB Simulink	The research concept of this paper includes the mathematical simulation of the solar panels and a battery backup study of the standalone unit

Table 1: Continued.

S. no.	Authors & references	Year	Location	Adopted technologies	Load type	Consumption type	Method	Objectives
36.	E Roumpakias, A Stamatelos [58]	2019	Greece	Grid-connected	Load serving	Electrical	Performance ratio (PR), yield factor (YF), reference yield (YR), capacity factor (CF), and an array to capture losses (LC)	The emphasis of the research was on the efficiency of a grid-affiliated PV system in Central Greece that was operational for six years. The study indicates a slight efficiency reduction over the years, which declined between 1 and 4 percent
37.	BR VS, GG Devadhas [59]	2019	—	Standalone	Load serving	Electrical	Sub-maximum power point tracking (S-MPPT)	This research recommends a single-phase linear PV default scheduler system. It initializes any device to zero in the shortest period
38.	N Manoj Kumar et al. [60]	2017	Malaysia	Grid connected	Load serving	Electrical	Photovoltaic geographical information systems (PVGIS) and Watts PV software	The primary target of the research was to build a sun-based PV plant at two diverse campuses. The specialized feasibility used the open rack or free stand mounting position crystalline innovation-based PV plant utilizing the PVGIS and PV Watts software. The specialized presentation acquired through PVGIS is very similar to the PV Watts results
39.	N Kumar et al. [61]	2019	—	Grid-connected	Load serving	Electrical	Fifth-order general integrator (FOGI)	In the study, an instinctive control procedure dependent on 'fifth-order general integrator (FOGI)' was proposed for framework-associated sun-powered photovoltaic (PV) energy conversion system (SECS)
40.	YZ Alharthi et al. [62]	2019	—	Grid-connected	Load serving	Electrical	HOMER	The study assessed a hybrid renewable energy system linked to the power grid with 15000.0 kW daily load demand and 2395.3 kW peak load. The net present cost (NPC), levelized energy cost (LCOE), and system environmental effects were examined
41.	E Kurt et al. [63]	2019	—	Grid-connected	Load serving	Electrical	PSCAD	The efficiency of a DC grid-affiliated PV device under insolation and temperature variations was investigated in the study. The DC-DC boost

Table 1: Continued.

S. no.	Authors & references	Year	Location	Adopted technologies	Load type	Consumption type	Method	Objectives
42.	H Rezk et al. [64]	2019	Egypt	Standalone	Irrigation	Electrical	HOMER	converter was constructed to increase the system's performance An optimal islanded irrigation solar PV battery system (BS) for Al Minya, Egypt, was used in the study paper. The energy costs obtained were lower than those previously reported due to the correct selection of PV size and shape, as well as the correct selection of the site

achieved by hybrid optimization model for electric renewable (HOMER). At the selected sites, the technoeconomic feasibility of using hybrid photovoltaic/wind/diesel with battery storage systems to match the load of a typical rural healthcare center was evaluated. The system is based on long-term daily meteorological data ranging between 18 and 39 years in this study. The findings of HOMER simulations show that the hybrid system is the best solution for all of the study's locations. Since the diesel-only system provides the highest COE and emits CO2, the hybrid systems involving PV/diesel/battery are considered ideal for RHC at remote locations within Iseyin and Port-Harcourt, due to the quality of renewable energy potential [15].

Fossil fuels like oil and gas are still playing a role in energy generation though people are now considering an alternative that provides energy demand by reducing it via energy efficiency and environment-friendly use of that energy resources. Since transport consumes a lot of conventional energy and generates greenhouse gases, therefore, the proposed measure to alter this issue is to use electric transportation. The HOMER program was used in this study [16] to develop and optimize a wind-solar hybrid energy charging station that will be beneficial for supplying power from renewable resources effectively and sustainably, managing grid load, and establishing additional charging stations.

PV systems have used distributed microgrids as efficient local electricity sources in regulated environments for energy consumers and inexhaustible energy generation. The global deployment of PV microgrids has expanded while taking the benefit of daily unrestricted solar insolation. In Rwanda, the average daily solar irradiation is between 4.0 and 5.0 $kWh/m^2/day$ [17]. The highest solar radiation for the selected site is seen in July where the value is 5.87 $kWh/m^2/day$. Energy storage has been proposed, with the backup used during peak demand, power shortages, blackouts, or some other power loss in grid-connected systems. Global studies show that the world's total implemented photovoltaic capacity has been steadily increasing [18]. Rwanda is educating private investors on how to implement solar energy projects and narrow the gap between electricity demand and supply [19]. Sustainable power sources to replace fossil fuels have been prioritized throughout the world for both economic and environmental reasons. The authors in [20, 21], confirmed the feasibility of a stable standalone electricity generation system for off-grid users using HOMER to model, evaluate, and optimize sustainable power sources that replace traditional energy sources. Table 1 below summarizes the successfully implemented researches made on a standalone, microgrid, and grid-connected solar systems in different parts of the world and their results prove to be viable.

This study gives a complete comparison of the state-of-the-art deterministic methodology to build a minigrid, including the influence of operating strategies to provide recommendations on conceptual models and operating strategies to researchers, developers, and professionals in the field. The standalone like, home lighting system (HLS) requires no major maintenance, and consumers could use it without being subjected to any influences, whereas minigrids are administered by cooperative societies founded by local governments and beneficiaries. In order to provide minigrid services in underdeveloped countries, it is necessary to establish an appropriate business strategy. The comparison as explained in literature, and Table 1 below was mainly based on consumer characteristics, net present cost, and the cost of energy, where it will depend on the quantity of the consumer as well as the different components available to be used in each household. Therefore, the result of the method adopted in our research compared to the electricity tariff in Rwanda is much more viable.

HOMER is particularly well suited to assessing prospective electricity possibilities in rural areas, as well as investigating the technical and economical effectiveness of hybrids provided a village load and energy resource availability. Therefore, this means different areas with different solar resources will provide different output solar powers where a minigrid supply system is being proposed for rural electrification programs. When the obtained energy costs were compared to the existing, current costs of electricity in Rwanda, it was discovered that this was the most cost-effective option. Previously, the solar home system was a simple choice, which included the

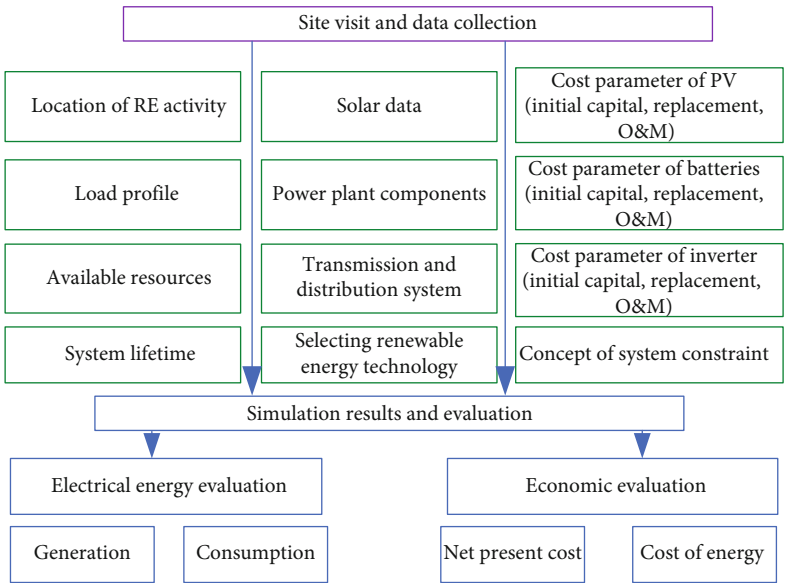


Figure 1: Illustration of the framework for analysis of the study. Abbreviations: RE: renewable energy; O&M: operation and maintenance.

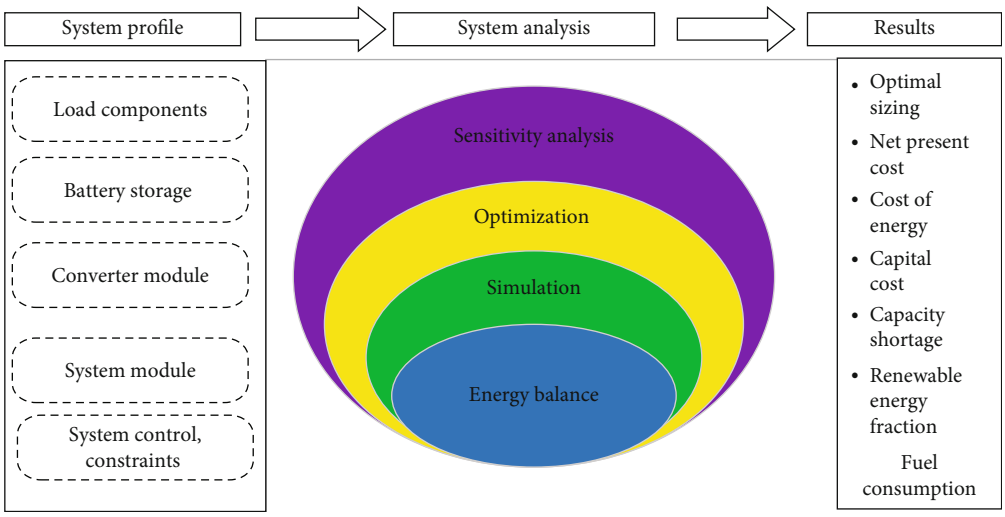


Figure 2: A detailed schematic representation of HOMER software.

implementation of PV panels, batteries, charge controllers, and inverter units for every residence and business structure in the village that used roof areas [22].

3. Methodology

HOMER software analyzed the data gathered from governmental energy organizations considering different photovoltaic systems uses in Rwanda's rural settlements [65]. HOMER software created a variety of models that demonstrate how different natural sustainable energy sources combine to produce green power in this study. In addition, we started working with power plant owners and operators throughout the research to ensure the study's reliability. According to these experts, relevant guidelines for the rural electrification planning process are lacking, posing risks, causing market distortions, and necessitating research projects for new electric power plants.

The methodology of this study is depicted in Figure 1 below. All of the study requirements were conducted based on our team's site visits and data collection. The data collected include electricity load demand profile, available resources, power plant production capacity, solar power plant components, and constraints. HOMER software performed the technoeconomic analyses in this research. The purpose of these technical and economic analyses was to develop a practicable off-grid photovoltaic system that would suit Rwanda's power sector at lower tariffs and maximum availability.

3.1. HOMER Pro Software. The HOMER Pro microgrid software is the world's standard software for optimizing microgrids, from community power and islet utility services to grid-affiliated sites, and army assets [65]. Figure 2 presents

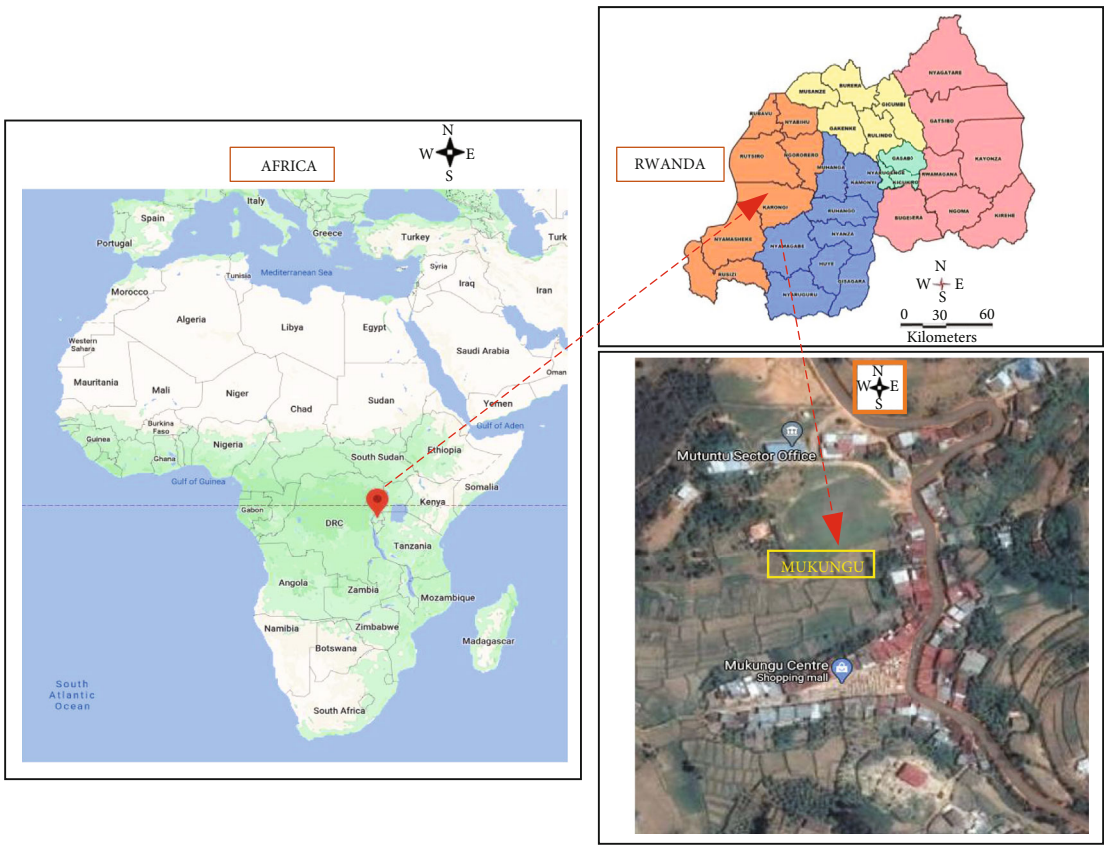


Figure 3: Location map of the selected study area (Mukungu village).

a schematic representation of the HOMER software. Imitation, optimization, and sensitivity analysis are the three primary functions of HOMER. It can strengthen power balance measurements, load profiling, location-specific tools, and system components are all factors taken into consideration by HOMER.

HOMER simulates feasible systems with device configurations in the simulation model. After each simulation, there is an optimization step. To achieve the best possible match, all imitated systems are categorized and refined according to specified parameters. Sensitivity analysis, on the other hand, is an optional function that allows HOMER users to model resource variables that are outside their control, like fuel prices and wind speeds. As a result, researchers can see how the ideal system changes because of these modifications [66]. The optimization ellipsoid encircles the simulation ellipsoid, showing that an imitation consists of several simulations. The sensitivity analysis ellipsoid, encircles the optimization ellipsoid, as in Figure 2.

3.2. Data Collection. In this survey, data were collected from 121 households in four Rwandan provinces, excluding Kigali city, using a specially designed questionnaire. Residents in the area were asked a series of questions as part of this study. The data were then summarized and analyzed using the Statistical Package for the Social Sciences (SPSS version 23.0). Consequently, the total energy

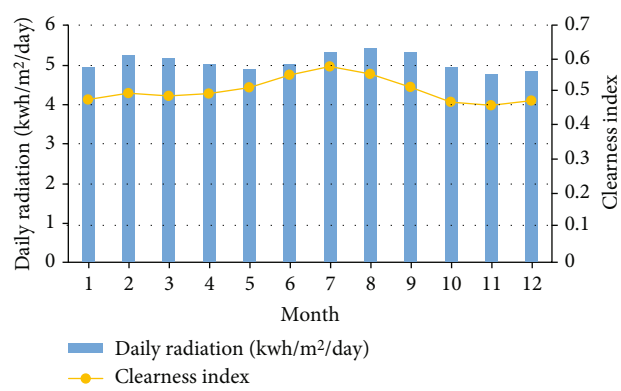


Figure 4: Solar energy profile at the preferred site.

consumption of people living in Rwanda's off-grid areas was calculated, as well as the energy needs of each house. As a result, this investigation is aimed at approximating consumer requests, which is a prerequisite for designing a power plant.

Off-grid solar power deployment necessitates a year's worth of solar irradiation. The National Aeronautics and Space Administration provided the input data for solar resources over a year in this case (NASA). Other data from the Rwanda Meteorology Agency was obtained and

Table 2: The projected everyday electricity demand for 121 houses in the village.

No.	Equipment in 121 households	No. in use	Power consumption (W)	Total power consumption (W)	Hours of use/day (Hrs/day)	Watt-hours/day
1	Lamps	891	10.0	8910.0	5.0	44550.0
2	Mobile phone	335	10.0	3350.0	8.0	26800.0
3	Ceiling fan	0	75.0	0.0	0.0	0.0
4	Radio	113	20.0	2260.0	5.0	11300.0
5	Television	49	120.0	5880.0	5.0	29400.0
6	Computer	4	100.0	400.0	5.0	2000.0
7	Refrigerator	5	500.0	2500.0	24.0	60000.0
8	Iron	19	1000.0	19000.0	1.0	19000.0
Daily total energy consumption in 121 households					193,050.0 Watt-hours/day = 193.05kWh/day	

Abbreviations: W: Watt; kWh: kilowatt-hour; Hrs: hours.

compared to NASA's data to ensure that the solar resource data obtained was correct. All these data were accurate, and they satisfy the requirement to be used in our study. Finally, the obtained data helped us to evaluate and verify the integration of solar power systems into Rwanda's power system.

3.3. Selected Site. Rwanda's government had approved a rural electrification strategy in the termination of 2016, in which the government, private industry, and relevant stakeholders collaborated to significantly boost rural electrification and establish lofty potential targets. Thus, in Rwanda's rural areas, pico/minihydropower, and minigrids from solar energy have been successfully implemented [67]. Mukungu village located in the Karongi District of Rwanda's Western province was chosen for this study, with GPS coordinates of S 02°13.9310′ and E 29°24.590′. In addition, the details of the chosen village location are shown in Figure 3. This village has a picohydropower plant that provides energy to up to 400 households in the off-grid region, promoting economic development. Solar energy, fortunately, can also be used as an alternative energy source in this situation.

3.4. Solar Resource Availability Evaluation. HOMER utilizes four renewable energy sources: biomass, hydro, solar, and wind, as well as other fuels that the system's equipment requires [65, 66]. Rwanda has abundant renewable energy resources, and it is attempting to electrify Rwanda's off-grid villages. The Mukungu village solar resources were extracted from the surface meteorology and solar website of NASA. The solar energy profile at the preferred study site is depicted in Figure 4.

Generally, the PV array's power output is determined by the angle of incidence of the solar radiation on the earth. HOMER calculates the PV array output in Equation (1) below [65]:

$$P_{\mathrm{PV}} = Y_{\mathrm{PV}} f_{\mathrm{PV}} \left(\frac{\bar{G}_T}{\bar{G}_{T,\mathrm{STC}}} \right) [1 + \alpha_P (T_c - T_{c,\mathrm{STC}})]. \quad (1)$$

The temperature power coefficient is zero if the temperature effects on the PV array are not modeled by HOMER software. Equation (1) becomes [65]

$$P_{\mathrm{PV}} = Y_{\mathrm{PV}} f_{\mathrm{PV}} \left(\frac{\bar{G}_T}{\bar{G}_{T,\mathrm{STC}}} \right). \quad (2)$$

On the other hand, HOMER can define the monthly average clearness index K_T using the following [65]:

$$K_T = \frac{H_{\mathrm{ave}}}{H_{\mathrm{o,ave}}}, \quad (3)$$

where Y_{PV} is the PV array rated capacity (kW), f_{PV} is the PV discounting factor (%), $\bar{G}_T$ is the PV array incident solar irradiation (kW/m^2), $\bar{G}_{T,\mathrm{STC}}$ is the emitted radiation under normal assessment conditions (1 kW/m^2), α_P is the temperature power coefficient (%/°C), T_c is the PV cell temperature (°C), $T_{c,\mathrm{STC}}$ is the PV cell temperature standard test conditions (at 25°C); H_{ave} is the normal monthly irradiation from the earth (kWh/m^2/day); and $H_{\mathrm{o,ave}}$ is the extraterrestrial horizontal insolation (kWh/m^2/day).

3.5. Load Details of the Selected Site. All types of electrical appliances at home, as well as the time that the appliances are used by the residents, determine energy consumption [68]. The estimated load was determined in this study based on a survey directed at various communities across the country. Experienced judges tested a series of the developed questionnaires for validity and used them to gather energy consumption data from respondents. Representatives from 121 households completed the questionnaires about their household electrical devices and monthly power consumption pattern in the research. Table 1 shows the summary of results obtained through analysis of other studies and their input data. The daily energy demand of 121 homes, as well as their respective power ratings, are shown in Table 2.

The electricity demand in remote areas is lower than in cities, according to a comprehensive energy consumption

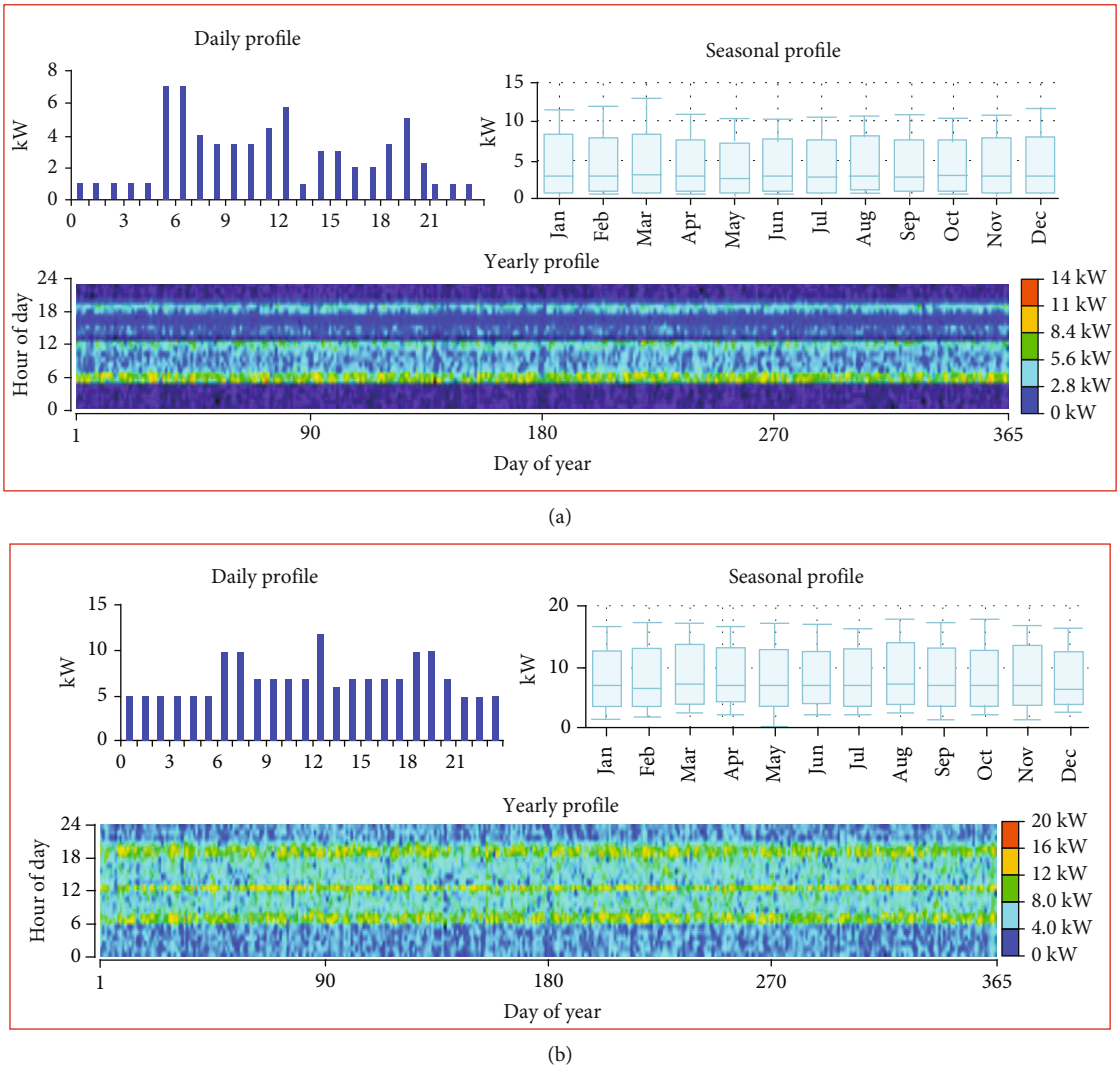


Figure 5: The daily, seasonal, and yearly load profiles: (a) electric load for the distinct off-grid household and (b) electric load for microgrids for a rural community in an off-grid area.

survey conducted in Rwanda. Household appliances that use electricity include radios, light bulbs, mobile phones, ceiling fans, electric irons, refrigerators, and laptops. Within each hour of the year, we must measure the sum of primary load in kilowatts using HOMER, either by importing hourly data from a file or by permitting HOMER to create hourly data from typical everyday load profiles. Consequently, HOMER generates typical load results depending on the consumer's features of everyday load [65]. The photovoltaic systems were designed, and their performance was evaluated in this study, taking into account the following suppositions:

(i) 193.05 kWh per day is the primary load, and 20.64 kW peak load was assumed for an outside-grid PV microgrid for the rural society (121 households)

(ii) The prime load of 1.6 kWh per day and 0.30 kW peak load were assumed for a remote grid solar PV microgrid system in the rural community

(iii) The project's lifespan was decided based on the component warranty, which was estimated to be around 25 years. The load profile used in this survey during imitation is shown in Figure 5

3.6. System Design Components. We used various components in this research depending on the photovoltaic systems we needed to simulate. Photovoltaic solar was the resource in the HOMER analysis. In addition, electricity is stored and converted using batteries and a converter. The efficiency and cost of each of the system's components have a significant impact on the design outcomes. Data from Rwandan generating companies and private sector companies' mini-grid remote grids, as well as existing literature, were used to develop the study's technical and cost parameters. Tables 2 and 3 show the performance and cost of each component for an islanded PV system for a single home and an outside-grid PV microgrid for a remote neighborhood, respectively.

Table 3: System elements and their costs for an islanded PV of an individual home.

No.	Component	Rated capacity	Capital cost (USD)	Replacement cost (USD)	O&M cost/year (USD)	Lifetime (years)
1	Converter (system converter)	1.0 kW	3000.00	2500.00	800.00	15
2	Batteries (generic 1.0 kWh lead acid)	1.0 kWh each	300.00	2,00.00	40.00	10
3	PV (generic flat plate)	1.0 kW	1500.00	1100.00	40.00	25

3.7. Design and Modeling of Selected PV Systems in Rwanda. Rwanda has a large number of untapped renewable energy source sites. Electricity is generated using hydro, solar, methane, peat, geothermal, wind, and waste energy. According to Rwanda's Environmental Management Agency (REMA) Outlook report from 2007, there are approximately 1200 MW of untapped power generation resources in Rwanda [67]. Unfortunately, so many of these resources remain unexplored. Rwanda's gross electricity generation was just 224.6 MW in 2019 [67].

HOMER is a sophisticated numerical modeling framework that offers much more details than traditional statistical modelers. It can perform imitation and responsiveness analysis with modest data [69]. Better designs are produced for likely inputs using net present cost (NPC), which is cost-effective. In addition, it generates power balance equations for every one of the 8760 hours annually, to simulate network operations. Consequently, it helps to determine viable configurations and approximate the installation cost and implementation of the power system over the project's life [69–71].

HOMER computes the average annualized cost for every item utilizing diversified costs and penalties for device pollutants. This value is also used to calculate the overall net present cost and the levelized energy cost (LCOE or COE) [65]. In HOMER software, the NPC can be evaluated using [65, 66]

$$\mathrm{CRF}(i, N) = \frac{i(1+i)^N}{(1+i)^N - 1}, \tag{4}$$

where i is the yearly actual interest rate and N is the duration (years). Also, Equation (5) assesses the levelized energy cost [65, 66]:

$$\mathrm{COE} = \frac{C_{\mathrm{ann,tot}}}{E_{\mathrm{prim}} + E_{\mathrm{def}} + E_{\mathrm{grid,sales}}} \tag{5}$$

where $C_{\mathrm{ann,tot}}$ is the overall annualized cost, E_{prim} and E_{def} are the yearly overall basic and postponed load, respectively, and $E_{\mathrm{grid,sales}}$ is the yearly power grid sales.

3.7.1. Standalone Solar Photovoltaic for a Single Residential House. First and foremost, those who are unable to connect to and be assisted by the publicly or privately owned utility grids are known as "off-grid users" [6].

As a result, standalone PV, minigrids, and microgrids can be used to provide electricity to such users. As previously stated, this part used HOMER Pro software to design, imitate, and process available inexhaustible power generation technologies to ensure everyone in Rwanda has access to sustainable energy. Importantly, several site visits to a delegate of selected residential houses in Rutsiro district, Western province, Rwanda, were made to design standalone solar photovoltaic systems, efficiently. Moreover, the electrical devices, power rating, and hours of use variables were used during the studies.

The sketch of the islanded remote grid photovoltaic system for an individual household is shown in Figure 6(a). The distinct islanded solar home system, comprises the PV panel, batteries storage, converter, DC, and AC buses, and electric load. As shown in Figure 6, the total load profile was 1.6 kWh/day and 0.30 kW daily peak load. The estimated daily energy consumption for 121 residential houses in the village was 193.05 kWh/day, as demonstrated in Table 2 above. So, the daily energy load average for one residential house in an off-grid area can be easily estimated to be 1.6 kWh/day per residential house.

Previous researchers and experts in renewable energy and minigrids or microgrids have provided insights on what hybrid systems are, why we need them, their uses, and applications for sustainable energy development. According to the previous surveys and the outcome of this study, standalone systems for one single household and a community are the first-choice decision to be made while they may not be cost-effective.

3.7.2. Off-Grid PV Microgrid System for Rural Community. The microgrid is important to intelligent power systems for increasing the distribution system's energy supply reliability and resilience. A microgrid is an interconnected collection of distributed energy and demand entities that function in either grid-connected or island mode within the network. Microgrids comprise small cell phone towers (as well as nanogrids), large commercial, industrial, and military facilities with generation capacities ranging from kilowatts to megawatts [72–75].

During our study, Figure 6(b) is the representation of the off-grid solar photovoltaic microgrid system for rural areas. The average daily load for that rural area was 193.05 kWh per day and a daily 20.64 kW peak (for 121 households). In addition, the photovoltaic system provides DC, while the converter transforms DC to AC and vice versa, which is supplied to the battery storage facility. Undeniably, research on various configurations or architectures of microgrid systems is gaining more attention to achieve the goals of carbon emission reduction.

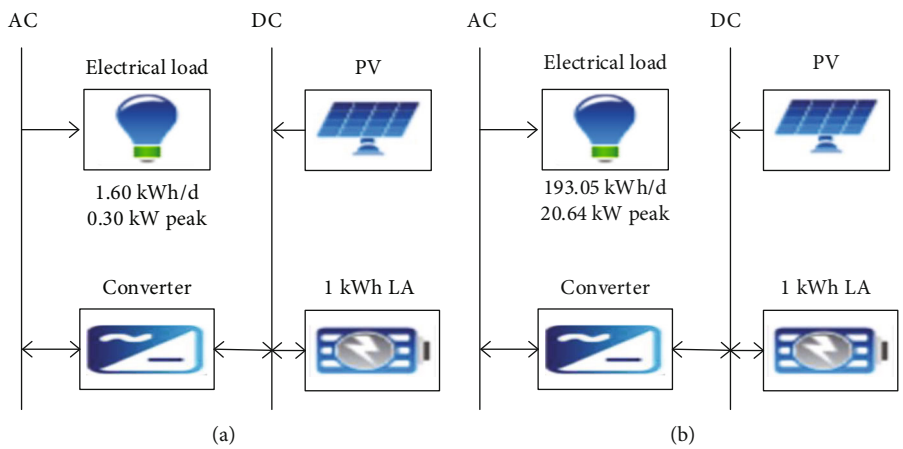


Figure 6: Various types of photovoltaic solar systems: (a) the standalone photovoltaic for a single residential household and (b) the PV microgrid system for the rural community in the off-grid area.

Table 4: Comparative simulation analysis for the proposed PV systems (we did not consider the sensitivity variables).

Resource	System architectures	Electricity production (kWh per year) (portion)	Cost summaries		
			Total NPC (USD)	LCOE (USD/kWh)	Operating cost (USD per year)
Solar	Standalone (1 household)	2426.0	9284.4	1.23	428.08
Solar	Microgrid (community)	221,715.0	1,166,898.0	1.28	37,965.91

4. Simulation and Optimization Results Using Homer Pro Software

HOMER's micropower optimization model was used to process the modeling and simulation results. For each responsiveness case it solves, HOMER imitates each system in the search space and rates all practicable systems in order of decreasing net present cost. HOMER optimizes small power systems by simulating a variety of device options under different restrictions and stimuli. These systems are compared using optimization tables. The optimization table contains information about each system's architecture, such as the number of batteries, converter size, and PV capacity. It also includes information on costs like the levelized energy cost (LCOE), net present cost (NPC), running costs, preliminary capital, and clean energy proportion. Also, two solar energy systems were designed in this research using a large number of hourly parameters in the HOMER software simulation. The simulations and analyses took into account a variety of solar radiation values. Without taking into account the sensitivity variables, Table 4 illustrates the imitation and processing of two dissimilar remote grid solar PV for the selected survey site.

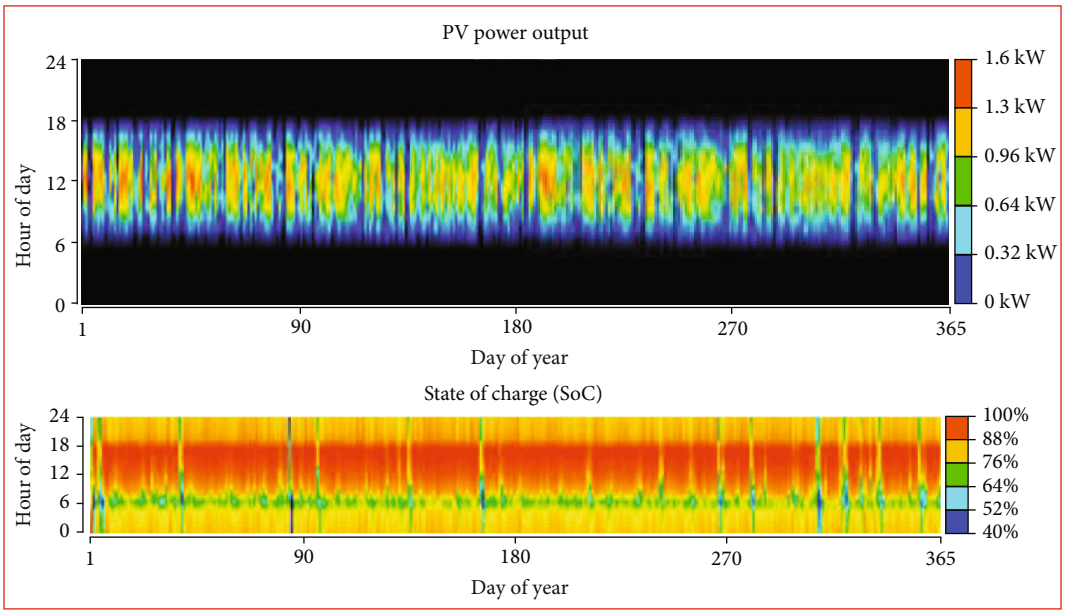
The simulation in this study considered different photovoltaic systems. As illustrated in Table 5 above, the minimum levelized cost of energy (LCOE) found from the simulation results was USD 1.23 per kWh for a standalone photovoltaic for an individual household. This standalone system generates 2426 kWh total yearly production and comprises 1.64 kW PV, 3 strings of batteries, and 0.262 kW of a system converter. The total operating cost and NPC for such photovoltaic systems are USD 428.08 and USD 9284.41 per year, respectively.

In contrast, the off-grid PV microgrid system for rural communities has shown a high LCOE compared to the standalone PV for an individual household. It generates 221,715.0 kWh total yearly production and comprises 150.0 kW PV, 443 strings of batteries, and 20.8 kW of system converter. For this photovoltaic system, the total NPC, LCOE, and operating costs were also USD 1,166,898.00, USD 1.28 per kWh, and USD 37,965.91 per year, respectively. In addition, PV output power and batteries' charge state (SOC) of the simulated photovoltaic systems are graphically illustrated in Figure 7 below, where (a) is an islanded solar PV system for a dwelling house and (b) is a PV microgrid system for a remote neighborhood in the off-grid area.

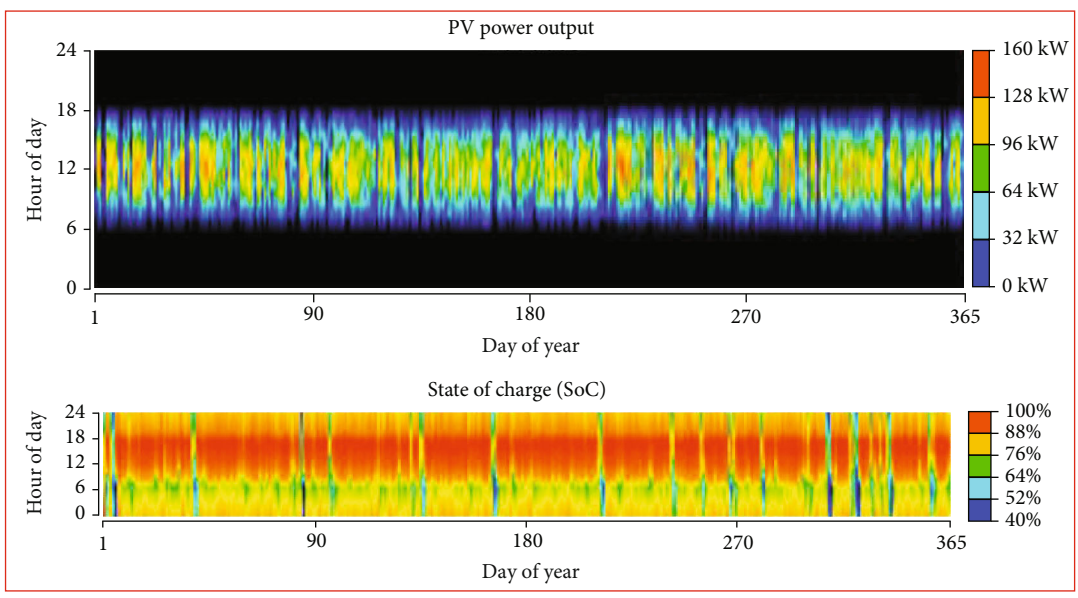
Because access to electricity is a key driver of development and welfare, Rwanda's government has set a goal of providing electricity to 100 percent of all the population by 2024. Rwanda has future prosperity of renewable resources, including wind, solar, geothermal, hydro, and methane gas, all of which should be explored before making any decisions. This will undoubtedly encourage development projects, bringing the total capacity of electricity generation to 556.0 MW by 2024. Unquestionably, the findings of this study show that for off-grid users, small solar standalone systems for individual households are preferable because they can start providing energy more rapidly at a low price.

Table 5: Structure items and their expense for off-grid PV microgrid systems for the rural community.

No.	Component	Rated capacity	Capital cost (USD)	Replacement cost (USD)	O&M cost/year (USD)	Lifetime (years)
1	Converter (system converter)	10.0 kW	21,164.00	16,000.00	8000.00	15
2	Batteries (generic 1 kWh lead acid)	1.0 kWh each	1702.00	1000.00	60.00	10
3	PV (generic flat plate)	10.0 kW	18,500.00	15,000.00	20.00	25



(a)



(b)

Figure 7: PV power output and SOC of the simulated solar photovoltaic systems: (a) the standalone photovoltaic for a single residential house and (b) the off-grid solar photovoltaic microgrid system for the rural community.

5. Discussion

The electricity prices is constantly increasing due to the world's fast growing population that needs access to sustainable electricity to sustain modern life expectancy. In Sub-Saharan Africa (SSA), for example, people living without access energy remain a determining factor that contributes to persistent poverty [5]. In this area, urban communities are still served by inefficient and unstable networks, while rural areas still lack access to electricity, except for power given to fairly wealthy households by small/private generators. Using fossil fuels to produce energy has long been regarded as unappealing due to the release of hothouse gases into the environment that raises the overall carbon trail. The latter encompasses disastrous consequences including increased global warming as well as its related consequences [76, 77].

In the current era of accelerated development and globalization, countries all over the world are looking at the low-cost PV systems to replace their existing power generation mix to ensure the reliability, affordability, and sustainability of potential power systems [78]. In fact, most governments have made renewable energy production a top priority, not only to minimize their overall carbon emissions and achieve international climate targets but also to gain wider socioeconomic benefits. And as per the International Energy Agency, 1.3 billion people everywhere in the world cannot have access to reliable electricity, particularly in the countryside of the developing world where the expansion of the utility grid is exceptionally difficult [79]. With distributed and independent control solutions, the microgrid model has confirmed to be one of the most realistic solutions that could be used to distribute inexhaustible energy sources (DRES) and can mitigate the perceived complications of deployment with increased stability with natural catastrophes, physical/cyberattacks, and cascading power blackouts [80].

To date, conventional energy resources cannot provide enough energy to meet the demand and are generally not environmentally friendly. Solving this problem of the energy gap, solar energy can yield an adequate solution [81]. However, due to every site requirement, they provide unpredictable power generation. Renewable energy presents a challenge of power quality, reliability, power system stability, and reactive power compensation. The intermittent nature of renewable energy like the solar, wind is less predictable and time-variant. The influence of dust on PV panels in the UAE was researched, and it was discovered that after 5 weeks of outside exposure, there was a 10% drop in power production [82]. Due to its stochastic and random character, renewable energy systems pose substantial issues to traditional grids, such as frequency variation, voltage fluctuation, and harmonics.

The low efficiency and unreliability of PV systems [83] are the most serious challenges. This article's technoeconomic model simulates minigrid, microgrid performance utilizing meteorological data, demand profiles, technology capabilities, and pricing data to identify the ideal component sizing of hybrid minigrids for rural electrification. The findings show how system sizing is influenced by location, renewable resource availability, technological cost, and performance which cause output power unstable. HOMER assesses different designs using the levelized cost of electricity (LCOE), but it cannot assess different financial models [84].

Rural areas' big issue is lacking consumer demand density and generally consists of low-income groups; therefore, project rate of return on investment is difficult to achieve as planned. High costs, low energy efficiency, and a lack of suitable rules and information are among issues that PV systems confront [85]. Unlike consumers in developed countries who can afford the high upfront costs of installing solar panels on their roofs to produce electricity, the number of Africans in stricken need of solar power cannot accommodate such an investment, despite the fact that solar power has a positive economic and environmental case. The global solar market is controlled by industrialized countries such as China, Europe, and the United States, making it difficult for industry knowledge and skills to spread to local businesses. The state's taxes and regulations have made solar-powered town electrification prohibitively expensive [86]. Because system functioning necessitates real-time measurements of solar irradiation and ambient temperature data, the data collected is limited due to flaws in the measuring equipment [14]. Because minigrid payback times can easily exceed several years, providing a regulatory environment that includes valid agreements or subsidies is necessary to limit risks for investors [87]. Long-term financing for minigrid projects is frequently difficult to come through due to inflation is either high or uncertain [88].

Challenges regarding policies are as follows: The necessity of policy support for off-grid electricity is critical where mostly there is no long-term electrification strategy [89]. Licensing challenges are as follows: Retail or generation licensing procedures that are complicated, costly, and time-consuming deter investors and businesses from starting minigrid initiatives [90]. Tariff setting challenges are as follows: Tariff design conflict is exacerbated by the fact that, in comparison to cheaper grid-based electricity, off-grid system developers must charge significantly higher tariffs to meet investment and operating costs [91]. The challenges are shown in Table 6.

The construction of a distributed power generation plant with a transmission and distribution systems for the generated power is typically the most cost-effective solution in isolated areas where grid expansion are considered expensive. Solar energy is an especially appealing renewable choice for most of the African countries because it is decentralized, abundant, and cost-effective as technology progresses. It is also resistant to supply and price swings while it remains equally qualified for funding from mutual and multinational organizations aiming to increase the renewable energy outputs in these countries. This is accomplished by inexhaustible energy sources available as well as the introduction of microgrids/standalone systems as ideal solutions to rural electrification problems in developing nations. In particular, microgrids, standalone remote-grid systems are suitable for off-grid lighting because they minimize device costs by combining streetlight storage and using pole-mounted solar PV

Table 6: Challenges of off-grid electrification in a nutshell.

Policies	Tariff	Licensing setting	Financial	Technical	Social
No energy access plan identifies off-grid areas	Minigrids economic feasibility is under doubt	Licenses are expensive	Minor projects are ignored by financing programs	When dimensioning, there is no consideration for future need	There are no community-based educational initiatives
There is no long-term electrification plan in place	Tariffs are too high for rural populations to sustain	Licensing is a time-consuming and complicated process	Programs for short-term funding	Technical standards are inadequate	There are not enough examples of productive use scenarios
Regulations that are only in effect for a short period of time and are subject to change	There is no suitable payment method	There is no differentiation made between project sizes	Fiscal incentives are not available	Components that are mismatched with the environment	
External stakeholders are not involved in the collaboration because it is done in isolation	Tariff criteria are strict, and there is no distinction between comprehensive financial		The rate of return on investment is predicted to below	There is no monitoring mechanism in place, and there are no responsible, certified employees	There was no community input throughout the planning stage
In minigrid initiatives, there is no clear description of stakeholder duties			Initial investment costs are high	There are not enough restrictions in place to assure dependable operation and maintenance	There is not enough technical expertise to hire local workers

for both charging batteries and distributed generation. These technologies ensure a critical position in meeting the global energy demands, and they are more than capable of providing power in a more effective, safe, secure, and updated manner.

The use of standalone solar PV systems can provide significant energy and environmental benefits over grid-connected solar PV systems. In communities with traditional energy and the greatest solar capacity, standalone solar PV systems present the strongest air pollution control benefits. In fact, the solar rooftop provides environmental benefits by replacing traditional (conventional) grid electricity. The standalone and microgrid systems simulated in this paper have provided best results; however, due to financial instability of most of the Sub-Saharan countries, a standalone PV system proves to be more viable to those scattered households.

6. Conclusion

Limited access to energy slows down the development and makes it harder for governments and people to establish growth targets. In this study, we designed and simulated off-grid PV power systems to provide electricity to a Rwandan remote county using HOMER software. Simulation results revealed that an islanded PV system for a dwelling home is the ideal off-grid power generation system for use in rural areas. The system is particularly cost-effective compared with a microgrid PV system that supplies electricity to a rural community in Rwanda. Results indicate that the total NPC, LCOE, and operating costs of a standalone energy system are estimated to USD 9284.40, USD 1.23 per kWh, and USD 428.08 per year, respectively. This is also evidenced by our results in Table 5.

Consistent with the aforementioned, not only could standalone PV power systems be the ideal solution to the electrification of rural areas in Rwanda but also these systems could help the government and environmental agencies in the efforts to minimize weather-related problems and stir up the development of green energy systems as the country strives to provide reliable and sustainable energy to all its citizens. It is also believed that the proposed standalone solar PV system would equally contribute to the development of future renewable energy generation systems in other countries with similar environmental, climate, weather, and meteorological conditions around the world. In particular, neighboring countries such as Burundi, Democratic Republic of Congo (DRC), Tanzania, and Uganda, and all other countries in the region are estimated to be good candidates for such a system.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] N. U. Blum, R. Sryantoro Wakeling, and T. S. Schmidt, "Rural electrification through village grids–Assessing the cost

competitiveness of isolated renewable energy technologies in Indonesia," *Renewable and Sustainable Energy Reviews*, vol. 22, pp. 482–496, 2013.

[2] S. Szabó, K. Bódis, T. Huld, and M. Moner-Girona, "Energy solutions in rural Africa: mapping electrification costs of distributed solar and diesel generation versus grid extension," *Environmental Research Letters*, vol. 6, no. 3, p. 034002, 2011.

[3] J. P. Murenzi and T. S. Ustun, "The case for microgrids in electrifying Sub-Saharan Africa," in *IREC2015 The Sixth International Renewable Energy Congress*, Sousse, Tunisia, March 2015.

[4] P. Buchana and T. S. Ustun, "The role of microgrids & renewable energy in addressing Sub-Saharan Africa's current and future energy needs," in *IREC2015 The Sixth International Renewable Energy Congress*, Sousse, Tunisia, March 2015.

[5] U. Deichmann, C. Meisner, S. Murray, and D. Wheeler, "The economics of renewable energy expansion in rural Sub-Saharan Africa," *Energy Policy*, vol. 39, no. 1, pp. 215–227, 2011.

[6] S. Bimenyimana, G. N. O. Asemota, J. D. D. Niyonteze, C. Nsengimana, P. J. Ihirwe, and L. Li, "Photovoltaic solar technologies: solution to affordable, sustainable, and reliable energy access for all in Rwanda," *International Journal of Photoenergy*, vol. 2019, Article ID 5984206, 29 pages, 2019.

[7] S. Rehman and L. M. al-Hadhrami, "Study of a solar PV-diesel-battery hybrid power system for a remotely located population near Rafha, Saudi Arabia," *Energy*, vol. 35, no. 12, pp. 4986–4995, 2010.

[8] A. O. Adewuyi and O. B. Awodumi, "Renewable and non-renewable energy-growth-emissions linkages: review of emerging trends with policy implications," *Renewable and Sustainable Energy Reviews*, vol. 69, pp. 275–291, 2017.

[9] U. Sureshkumar, P. S. Manoharan, and A. P. S. Ramalakshmi, "Economic cost analysis of hybrid renewable energy system using HOMER," in *IEEE-International Conf. Adv. Eng. Sci. Manag. ICAESM-2012*, vol. 8, pp. 94–99, 2012.

[10] S. Bahramara, "Optimal planning of hybrid renewable energy systems using HOMER: a review," *Renewable and Sustainable Energy Reviews*, vol. 62, pp. 609–620, 2016.

[11] D. Toke, "Renewable financial support systems and cost-effectiveness," *Journal of Cleaner Production*, vol. 15, no. 3, pp. 280–287, 2007.

[12] L. Olatomiwa, R. Blanchard, S. Mekhilef, and D. Akinyele, "Hybrid renewable energy supply for rural healthcare facilities: an approach to quality healthcare delivery," *Sustainable Energy Technologies and Assessments*, vol. 30, pp. 121–138, 2018.

[13] M. Hossain, S. Mekhilef, and L. Olatomiwa, "Performance evaluation of a stand-alone PV-wind-diesel-battery hybrid system feasible for a large resort center in South China Sea, Malaysia," *Sustainable Cities and Society*, vol. 28, pp. 358–366, 2017.

[14] L. M. Halabi, S. Mekhilef, L. Olatomiwa, and J. Hazelton, "Performance analysis of hybrid PV/diesel/battery system using HOMER: a case study Sabah, Malaysia," *Energy Conversion and Management*, vol. 144, pp. 322–339, 2017.

[15] L. Olatomiwa, S. Mekhilef, and O. S. Ohunakin, "Hybrid renewable power supply for rural health clinics (RHC) in six geo- political zones of Nigeria," *Sustainable Energy Technologies and Assessments*, vol. 13, pp. 1–12, 2016.

[16] O. Ekren, C. Hakan Canbaz, and Ç. B. Güvel, "Sizing of a solar-wind hybrid electric vehicle charging station by using HOMER software," *Journal of Cleaner Production*, vol. 279, p. 123615, 2021.

[17] T. I. J. de Dieu Uwisengeyimana and A. Teke, "Current overview of renewable energy resources in Rwanda," *The Journal of Energy and Natural Resources*, vol. 5, no. 6, pp. 92–97, 2016.

[18] J. Dong, M. M. Olama, T. Kuruganti et al., "Novel stochastic methods to predict short-term solar radiation and photovoltaic power," *Renewable Energy*, vol. 145, pp. 333–346, 2020.

[19] J. D. Dieu Niyonteze, F. Zou, G. N. Osarumwense Asemota, and S. Bimenyimana, "Solar-powered mini-grids and smart metering systems , the solution to Rwanda energy crisis," vol. 1311, p. 012002, 2019.

[20] A. R. Gautam, K. Gourav, J. M. Guerrero, and D. M. Fulwani, "Ripple mitigation with improved line-load transients response in a two-stage DC-DC-AC converter: adaptive SMC approach," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 4, pp. 3125–3135, 2018.

[21] D. Kumar Lal, B. Bhusan Dash, and A. K. AkellaOptimization of PV/wind/micro-hydro/diesel hybrid power system in HOMER for the study area," *International Journal on Electrical Engineering and Informatics*, vol. 3, no. 3, pp. 307–325, 2011.

[22] N. Hagumimana, J. Zheng, G. N. O. Asemota et al., "Concentrated solar power and photovoltaic systems: a new approach to boost sustainable energy for all (Se4all) in Rwanda," *International Journal of Photoenergy*, vol. 2021, Article ID 5515513, 32 pages, 2021.

[23] M. K. Deshmukh and A. B. Singh, "Modeling of Energy Performance of Stand-Alone SPV System Using HOMER Pro," *Energy Procedia*, vol. 156, pp. 90–94, 2019.

[24] O. A. Controller-in-loop, *A hybrid PV-battery system for on-grid and off-grid applications—controller-in-loop simulation validation*.

[25] C. Marino, A. Nucara, M. F. Panzera, M. Pietrafesa, and A. Pudano, "Economic comparison between a stand-alone and a grid connected PV system vs. grid distance," *Energies*, vol. 13, no. 15, p. 3846, 2020.

[26] M. Hafeez, M. Hariri, M. Khairunaz, M. Desa, and S. Masri, *Grid-connected PV generation*, vol. 2040, 2020.

[27] F. A. Alturki and E. M. Awwad, "Sizing and cost minimization of standalone hybrid WT/PV/biomass/pump-hydro storage-based energy systems," *Energies*, vol. 14, no. 2, p. 489, 2021.

[28] T. Wu, H. Zhang, and L. Shang, "Optimal sizing of a grid-connected hybrid renewable energy systems considering hydroelectric storage," *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, pp. 1–17, 2020.

[29] B. E. Türkay and A. Y. Telli, "Economic analysis of standalone and grid connected hybrid energy systems," *Renewable Energy*, vol. 36, no. 7, pp. 1931–1943, 2011.

[30] D. Mazzeo, N. Matera, P. de Luca, C. Baglivo, P. Maria Congedo, and G. Oliveti, "Worldwide geographical mapping and optimization of stand-alone and grid- connected hybrid renewable system techno-economic performance across Koppen- Geiger climates," *Applied Energy*, vol. 276, p. 115507, 2020.

[31] A. Chakir, M. Tabaa, F. Moutaouakkil et al., "Optimal energy management for a grid connected PV-battery system," *Energy Reports*, vol. 6, pp. 218–231, 2020.

[32] R. Srivastava, A. N. Tiwari, and V. K. Giri, "An overview on performance of PV plants commissioned at different places in the world," *Energy for Sustainable Development*, vol. 54, pp. 51–59, 2020.

[33] A. C. Duman and Ö. Güler, "Economic analysis of grid-connected residential rooftop PV systems in Turkey," *Renewable Energy*, vol. 148, pp. 697–711, 2020.

[34] H. M. Ridha, C. Gomes, H. Hizam, M. Ahmadipour, A. A. Heidari, and H. Chen, "Multi-objective optimization and multi-criteria decision-making methods for optimal design of standalone photovoltaic system: a comprehensive review," *Renewable and Sustainable Energy Reviews*, vol. 135, p. 110202, 2021.

[35] M. J. Mayer and G. Gróf, "Techno-economic optimization of grid-connected, ground-mounted photovoltaic power plants by genetic algorithm based on a comprehensive mathematical model," *Solar Energy*, vol. 202, pp. 210–226, 2020.

[36] H. A. Kazem, M. T. Chaichan, A. H. A. al-Waeli, and K. Sopian, "A novel model and experimental validation of dust impact on grid-connected photovoltaic system performance in Northern Oman," *Solar Energy*, vol. 206, pp. 564–578, 2020.

[37] E. Aykut and Ü. K. Terzi, "Techno-economic and environmental analysis of grid connected hybrid wind/photovoltaic/-biomass system for Marmara University Goztepe campus," *International Journal of Green Energy*, vol. 17, no. 15, pp. 1036–1043, 2020.

[38] R. Khezri, A. Mahmoudi, and M. H. Haque, "Optimal capacity of solar PV and battery storage for Australian grid-connected households," *IEEE Transactions on Industry Applications*, vol. 56, no. 5, pp. 5319–5329, 2020.

[39] B. K. Das, "Optimal sizing of stand-alone and grid-connected solar PV systems in Bangladesh," *International Journal of Energy for a Clean Environment*, vol. 21, no. 2, pp. 107–124, 2020.

[40] M. Al-Maghalseh, "Generation unit sizing, economic analysis of grid connected and standalone PV power plant," *Int. J. Energy Appl. Technol.*, vol. 6, no. 1, pp. 1–7, 2019.

[41] H. A. Attia and F. D. A. Gonzalo, "Stand-alone PV system with MPPT function based on fuzzy logic control for remote building applications," *International Journal of Power Electronics and Drive Systems*, vol. 10, no. 2, p. 842, 2018.

[42] A. Iqbal and M. T. Iqbal, "Design and analysis of a stand-alone PV system for a rural house in Pakistan," *International Journal of Photoenergy*, vol. 2019, Article ID 4967148, 8 pages, 2019.

[43] Y. Chaibi, M. Salhi, and A. el-jouni, "Sliding mode controllers for standalone PV systems: modeling and approach of control," *International Journal of Photoenergy*, vol. 2019, Article ID 5092078, 12 pages, 2019.

[44] M. A. Omar and M. M. Mahmoud, "Design and Simulation of a PV System Operating in Grid-Connected and Stand- Alone Modes for Areas of Daily Grid Blackouts," *International Journal of Photoenergy*, vol. 2019, Article ID 5216583, 9 pages, 2019.

[45] P. K. Bonthagorla and S. Mikkili, "Performance investigation of hybrid and conventional PV array configurations for grid-connected/standalone PV systems," *CSEE Journal of Power and Energy Systems*, pp. 1–16, 2020.

[46] M. Salimi, F. Radmand, and M. Hosseini Firouz, "Dynamic modeling and closed-loop control of hybrid grid-connected renewable energy system with multi-input multi-output controller," *Journal of Modern Power Systems and Clean Energy*, vol. 9, no. 1, pp. 94–103, 2021.

[47] M. Dali, J. Belhadj, and X. Roboam, "Hybrid solar-wind system with battery storage operating in grid-connected and standalone mode: control and energy management - experimental investigation," *Energy*, vol. 35, no. 6, pp. 2587–2595, 2010.

[48] M. I. Hlal, V. K. Ramachandaramurthy, A. Sarhan, A. Pouryekta, and U. Subramaniam, "Optimum battery depth of discharge for off-grid solar PV/battery system," *J. Energy Storage*, vol. 26, p. 100999, 2019.

[49] K. N. B. Akshai and R. Senthil, "Economic evaluation of grid connected and standalone photovoltaic systems using PVSyst," *IOP Conference Series: Materials Science and Engineering*, vol. 912, no. 4, 2020.

[50] S. Odeh and I. Ibrik, "Performance assessment of standalone PV systems for rural communities," *Australian Journal of Mechanical Engineering*, pp. 1–10, 2019.

[51] J. Kumar, N. R. Parhyar, M. K. Panjwani, and D. Khan, "Design and performance analysis of PV grid-tied system with energy storage system," *International Journal of Electrical and Computer Engineering*, vol. 11, no. 2, pp. 1077–1085, 2021.

[52] B. K. K. Prasad, K. P. Reddy, K. Rajesh, and P. V. Reddy, "Design and simulation analysis of 12.4 kWp grid connected photovoltaic system by using PVSYST software," *The International Journal of Recent Technology and Engineering*, vol. 8, no. 5, pp. 2859–2864, 2020.

[53] Z. B. Duranay and H. Guldemir, "Modelling and simulation of a single phase standalone PV system," in *2019 11th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, Pitesti, Romania, June 2019.

[54] Y. Cui, J. Zhu, F. Meng, S. Zoras, J. McKechnie, and J. Chu, "Energy assessment and economic sensitivity analysis of a grid-connected photovoltaic system," *Renewable Energy*, vol. 150, pp. 101–115, 2020.

[55] N. Gupta, R. Garg, and P. Kumar, "Sensitivity and reliability models of a PV system connected to grid," *Renewable and Sustainable Energy Reviews*, vol. 69, pp. 188–196, 2017.

[56] M. P. Bonkile and V. Ramadesigan, "Power management control strategy using physics-based battery models in standalone PV-battery hybrid systems," *J. Energy Storage*, vol. 23, pp. 258–268, 2019.

[57] S. S. Dheeban, N. B. Muthu Selvan, and C. Senthil Kumar, "Design of standalone PV system," *International Journal of Scientific and Technology Research*, vol. 8, no. 11, pp. 684–688, 2019.

[58] E. Roumpakias and A. Stamatelos, "Performance analysis of a grid-connected photovoltaic park after 6 years of operation," *Renewable Energy*, vol. 141, pp. 368–378, 2019.

[59] V. S. Bibin Raj and G. G. Devadhas, "Design and development of new control technique for standalone PV system," *Microprocessors and Microsystems*, vol. 72, p. 102888, 2020.

[60] N. Manoj Kumar, K. Sudhakar, and M. Samykano, "Techno-economic analysis of 1 MWp grid connected solar PV plant in Malaysia," *International Journal of Ambient Energy*, vol. 40, no. 4, pp. 434–443, 2019.

[61] N. Kumar, V. Saxena, B. Singh, and B. K. Panigrahi, "Intuitive control technique for grid connected partially shaded solar PV-based distributed generating system," *IET Renewable Power Generation*, vol. 14, no. 4, pp. 600–607, 2020.

[62] Y. Z. Alharthi, M. K. Siddiki, and G. M. Chaudhry, "Techno-economic analysis of hybrid PV/wind system connected to utility grid," in *2019 IEEE Texas Power and Energy Conference (TPEC)*, pp. 1–6, College Station, TX, USA, February 2019.

[63] E. Kurt and G. Soykan, "Performance analysis of DC grid connected PV system under irradiation and temperature variations," in *2019 8th International Conference on Renewable Energy Research and Applications (ICRERA)*, pp. 702–707, Brasov, Romania, November 2019.

[64] H. Rezk, M. A. Abdelkareem, and C. Ghenai, "Performance evaluation and optimal design of stand-alone solar PV-battery system for irrigation in isolated regions: a case study

in Al Minya (Egypt)," *Sustainable Energy Technologies and Assessments*, vol. 36, p. 100556, 2019.

[65] HOMER Energy LLC, "HOMER Pro version 3.7 user manual," *HOMER Energy*, p. 416, 2016, http://www.homerenergy.com/pdf/HOMERHelpManual.pdf.

[66] T. Lambert, P. Gilman, and P. Lilienthal, "Micropower system modeling with HOMER," *Integration of Alternative Sources of Energy: Farret/Integration of Alternative Sources of Energy*, pp. 379–418, 2006.

[67] J. D. D. Niyonteze, F. Zou, G. Norense Osarumwense Asemota, S. Bimenyimana, and G. Shyirambere, "Key technology development needs and applicability analysis of renewable energy hybrid technologies in off-grid areas for the Rwanda power sector," *Heliyon*, vol. 6, no. 1, article e03300, 2020.

[68] S. Firth, K. Lomas, A. Wright, and R. Wall, "Identifying trends in the use of domestic appliances from household electricity consumption measurements," *Energy and Buildings*, vol. 40, no. 5, pp. 926–936, 2008.

[69] W. M. Amutha and V. Rajini, "Cost benefit and technical analysis of rural electrification alternatives in southern India using HOMER," *Renewable and Sustainable Energy Reviews*, vol. 62, pp. 236–246, 2016.

[70] A. Jamalaiah, C. P. Raju, and R. Srinivasarao, "Optimization and operation of a renewable energy based pv-fc-micro grid using HOMER," in *2017 International Conference on Inventive Communication and Computational Technologies (ICICCT)*, pp. 450–455, Coimbatore, India, March 2017.

[71] M. Khemariya, A. Mittal, P. Baredar, and A. Singh, "Cost and size optimization of solar photovoltaic and fuel cell based integrated energy system for un-electrified village," *The Journal of Energy Storage*, vol. 14, pp. 62–70, 2017.

[72] H. Louie, E. O'Grady, V. van Acker, S. Szablya, N. P. Kumar, and R. Podmore, "Rural off-grid electricity service in Sub-Saharan Africa [technology leaders]," *IEEE Electrification Magazine*, vol. 3, no. 1, pp. 7–15, 2015.

[73] G. Prinsloo, A. Mammoli, and R. Dobson, "Customer domain supply and load coordination: a case for smart villages and transactive control in rural off-grid microgrids," *Energy*, vol. 135, pp. 430–441, 2017.

[74] M. E. Khodayar, "Rural electrification and expansion planning of off-grid microgrids," *The Electricity Journal*, vol. 30, no. 4, pp. 68–74, 2017.

[75] D. E. Olivares, A. Mehrizi-Sani, A. H. Etemadi et al., "Trends in microgrid control," *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 1905–1919, 2014.

[76] M. Hoel, "Depletion of fossil fuels and the impacts of global warming," *Fuel and Energy Abstracts*, vol. 37, no. 6, p. 460, 1996.

[77] T. S. Ustun, C. Ozansoy, and A. Zayegh, "Recent developments in microgrids and example cases around the world–A review," *Renewable and Sustainable Energy Reviews*, vol. 15, no. 8, pp. 4030–4041, 2011.

[78] IRENA, *Future of solar photovoltaic: deployment, investment, technology, grid integration and socio-economic aspects (A Global Energy Transformation: paper)*, 2019.

[79] R. K. Akikur, R. Saidur, H. W. Ping, and K. R. Ullah, "Comparative study of stand-alone and hybrid solar energy systems suitable for off-grid rural electrification: a review," *Renewable and Sustainable Energy Reviews*, vol. 27, pp. 738–752, 2013.

[80] A. Abdulkarim, N. Faruk, A. O. Oloyode et al., "State of the art in research on optimum design, reliability and control of renewable energy microgrids," *Elektr. J. Electr. Eng.*, vol. 17, no. 3, pp. 23–35, 2018.

[81] T. R. Shah and H. M. Ali, "Applications of hybrid nanofluids in solar energy, practical limitations and challenges: A critical review," *Solar Energy*, vol. 183, pp. 173–203, 2019.

[82] B. M. A. Mohandes, L. el-Chaar, and L. A. Lamont, "Application study of 500 W photovoltaic (PV) system in the UAE," *Applied Solar Energy*, vol. 45, no. 4, pp. 242–247, 2009.

[83] M. A. Green, Y. Hishikawa, E. D. Dunlop et al., "Solar cell efficiency tables (version 53)," *Progress in Photovoltaics: Research and Applications*, vol. 27, no. 1, pp. 3–12, 2019.

[84] A. Niraula, "Scaling up of off-grid solar micro grids: moving towards a 'utility in a box' model for rapid deployment," 2015, https://static1.squarespace.com/static/536b92d8e4b0750dff7e241c/t/55ed3aebe4b049db23e2c0c7/1441610475311/Dissertation_Msc_Anjal.pdf.

[85] M. H. Alaaeddin, S. M. Sapuan, M. Y. M. Zuhri, E. S. Zainudin, and F. M. al- Oqla, "Photovoltaic applications: status and manufacturing prospects," *Renewable and Sustainable Energy Reviews*, vol. 102, pp. 318–332, 2019.

[86] J. Amankwah-amoah, "Solar energy in Sub-Saharan Africa : the challenges and opportunities of technological leapfrogging," *Thunderbird International Business Review*, vol. 57, no. 1, pp. 15–31, 2015.

[87] T. Reber, S. Booth, D. Cutler, X. Li, and J. Salasovich, *Tariff considerations for micro-grids in Sub-Saharan Africa*, US Aid, 2018, https://www.nrel.gov/docs/fy18osti/69044.pdf.

[88] USAID, "Challenges and needs in financing mini-grids Mini-Grids Support Toolkit," Energy U.S. Agency for International Development, 2018, June 2021, https://www.usaid.gov/energy/mini-grids/financing.

[89] A. A. Eras-Almeida, M. Fernández, J. Eisman, J. G. Martín, E. Caamaño, and M. A. Egido-Aguilera, "Lessons learned from rural electrification experiences with third generation solar home systems in latin America: case studies in Peru, Mexico, and Bolivia," *Sustainability*, vol. 11, no. 24, p. 7139, 2019.

[90] E. I. Come Zebra, H. J. van der Windt, G. Nhumaio, and A. P. C. Faaij, "A review of hybrid renewable energy systems in mini-grids for off-grid electrification in developing countries," *Renewable and Sustainable Energy Reviews*, vol. 144, p. 111036, 2021.

[91] M. F. N. Peterschmidt, M. Rohrer, and B. Kondev, "RECP minigrid policy toolkit doublepage," 2014, http://www.minigridpolicytoolkit.euei-pdf.org/policy-toolkit.

Hindawi
International Journal of Photoenergy
Volume 2021, Article ID 7552200, 14 pages
https://doi.org/10.1155/2021/7552200

Research Article

Takagi–Sugeno State Delayed Feedback and Integral Control for PV Systems: Modeling, Simulation, and Control

Hicham El Aiss, Karina A. Barbosa, Hector Chavez, and Carlos Rodriguez

Departamento de Ingenieria Electrica, Universidad de Santiago de Chile, Estación Central, Santiago, Chile

Correspondence should be addressed to Hicham El Aiss; hicham.elaiss@usach.cl

Received 10 June 2021; Revised 10 August 2021; Accepted 11 August 2021; Published 28 September 2021

Academic Editor: Laurentiu Fara

The integration of the large-scale photovoltaic systems has experienced significant growth, which is similarly expected to occur with small-scale photovoltaic systems. Since small-scale systems must be simple in cost-effective components, control strategies must be implemented in low complexity circuits. However, current maximum power point tracking (MPPT) algorithms are generally complex and require electronic components to support variable control gains for different irradiance conditions, preventing simple MPPT implementations suitable for small-scale photovoltaic systems. This paper proposes a new control strategy to tackle the power tracking problem of the power systems. First, a dynamic model of the photovoltaic system is described and converted into a Takagi–Sugeno (T-S) model. Then, an MPPT scheme is proposed in series with a fixed integral and a fuzzy gain state delay feedback controller, which avoids the need for a variable control gain, simplifying the electronic implementation of the control strategy. New delay-dependent stabilization conditions based on the Lyapunov-Krasovskii functional are proposed in terms of a convex optimization problem, where the delayed feedback and integral gains are designed simultaneously. Simulation results using Matlab and Simulink are used to validate the proposed method.

1. Introduction

The increasing demand for electrical energy and climate change has fostered the need for clean energy. Integrating these energies into the power system brings new technological challenges to reevaluate the usual designs, operations, and planning practices. Among clean energy technologies, photovoltaic (PV) systems have experienced significant growth worldwide, due to cost reductions [1] in the last decades. According to [2], the world total PV capacity has increased at an average annual rate of 55% in the period 2009-2013. This way, the development, modeling, and control of PV systems have captured particular attention [3–6].

Although the growth of PV systems is currently associated with large-scale plants, small-scale applications are now gaining interest. Projections of IRENA [1] established small-scale applications like rooftop PV, solar carports, and solar trees as future trends in the industry's development. In particular, small-scale PV systems usually are close to demands (such as a rooftop PV), increasing its cost-effectiveness as pointed out in [7]. Other projections are in line with this premise of future increase in small-scale PV [8, 9].

One limitation of small-scale PV applications is the low performance of control algorithms. In small-scale applications, the control electronic cost is close to that of the power electronic converter, so current small-scale PV devices usually count with very simple circuits, such as Pulse Width Modulation (PWM) [10]. Nevertheless, this simplicity in the electronic implementation has a negative impact on performance [11]. It is then of interest to find better control algorithms that are simple enough to utilize simple electronic implementation.

In general, MPPT control strategies tend to be complex and MPPT systems are typically implemented in a DC-DC power converter to provide power to different applications [12, 5, 6, 13, 14]. In terms of the MPPT algorithms,

currently, one can find the application of diverse techniques (see, for example, [15–17] and [18, 19]). In [18], the authors have presented a comparative analysis for the perturb and observe (P&O) and the incremental conductance (INC) techniques, which are the most widely used. The MPPT algorithms must be implemented in an electronic converter, and the boost converter has shown to be the simplest [20, 21]. Even in the case of the boost converter, it can be noted that the implementation of the control strategies leads to high-level programming needs, and the use of expensive electronic digital signal processor (DSP) is standard [18, 22]. Advanced MPPT algorithms lead to complex implementation frameworks, preventing their use in small-scale applications.

The advanced MPPT proposals are based on variable control gains, which is one of the reasons why they need complex electronic implementations. Various control strategies, such as neural networks [23], genetic algorithms [24, 25], and fuzzy controllers [26–29], have appeared in the boost converter context. Generally speaking, developing a new strategy to control the PV system is a challenge. In the literature, several methods and technologies have been used and investigated. In [30], the H_∞ control approach based on a fuzzy proportional-integral (PI) was used as a state feedback control in the control design and stability conditions of the closed-loop PV system. In [31], the boost converter was controlled using the T-S fuzzy parallel distributed compensation (PDC) to ensure stability and zero tracking error. To achieve MPPT under varying climatic conditions, the authors in [32] propose an H_∞ observer-based fuzzy controller. To improve the efficiency of photovoltaic systems, the authors in [29] have proposed to use a fuzzy controller with adaptive gain. Two distinct rule bases were combined in the proposed controller concept. The first rule base is intended for adjusting the boost converter's duty cycle, while the second rule base is designed for online adjusting of the controller's gain. It is worth noting that the techniques proposed in [29–32] compute their controller gains offline. Theoretically, the above methods ensure good efficiency, fast PV power convergence to the MPPT, and a smoother steady-state response without oscillation around the MPPT. These provide a stable PV power generation. The traditional PI controller usually is sensitive to parameter variations, resulting in a low stability margin [33]. A common aspect of all these approaches is the variable nature of the different control gains to respond to changing irradiance conditions, which requires a complete processing framework. The reason for variable gains relies on the fact that the underlying nonlinear dynamics is addressed by defining different stability regions, leading to constant control policies within each stability region. The resulting control is then an aggregation of various regions to cover the totality of the control space. Hence, the control algorithm selects the control gain associated with the operating stability region, resulting in variable gain algorithms. These variable gain algorithms are difficult to implement with simple electronics, so they are not generally implemented in small-scale PV systems.

It is well known that frequency control using a fuzzy logic controller (FLC) responds effectively to parameter variations. In [34], FLC is carried out to enhance the system's performance. It has three stages: fuzzification, processing, and defuzzification. The robust observed-based MPPT control for PV systems with a DC-DC buck converter is developed by [35], using a Lyapunov approach and LMI formulation. In particular, a T-S model-based method has been used for the MPPT of PV systems, resulting in fast convergence to the maximum power and elimination of the oscillations around the maximum. Significant results for maximizing power point tracking of PV systems were reported by [32, 36].

On the other hand, many LMI stabilization conditions based on T-S fuzzy models have been proposed in the literature using the Lyapunov approach to design controllers for nonlinear systems. The authors in [32, 30] present LMI stabilization conditions of the PV model via fuzzy observer-based and a fuzzy PI state feedback control, respectively. In [37], a robust control problem of fuzzy time-delay systems has been investigated throughout the Wirtinger-type integral inequality and convex techniques to estimate the derivative of Lyapunov-Krasovskii functional, ensuring robust asymptotic stability of the closed-loop systems. A robust H_∞ static output-feedback controller for discrete T-S fuzzy models with input saturation constraint has been developed by [38]. However, the authors of these works did not consider the delay and the delay-control issues, which can destabilize the closed-loop systems. To the best of our knowledge, a state delayed feedback control has never been applied to PV systems with DC-DC boost converter under disturbance effect. The approximation of nonlinear systems to convex structures makes it possible to formulate stability conditions under LMIs [39]. Since LMIs are convex constraints, they are simple to solve using a variety of available solvers, including SeDuMi [40]. As a result, the stability/stabilization criteria of the nonlinear system can be reduced to the feasibility of a set of LMIs. In general, when the feedback gains have been processed as variable parameters in the LMI feasibility issue, automatic stabilizing control is generated from a set of obtained LMIs.

This paper proposes a new control approach for a boost converter in an MPPT PV application based on a state delayed feedback approach and integral control. The contribution is a new state delay feedback algorithm, which allows the determination of a unique, constant controller gains for irradiance's different operating values, including the integral gain. The proposal is based on new sufficient delay-dependent criteria stated as a convex optimization problem to ensure that the closed-loop system is asymptotically stable and minimizes the exogenous impact on the boost converter's output. The theoretical formulation is verified by numerical results using Matlab and Simulink, where a complete simulation of the system is presented.

The organization of the paper is as follows. In Section 2, the Takagi–Sugeno (T-S) fuzzy model for the PV system is presented. The general control strategy proposed in this paper is presented in Section 3, including the control design approach. In Section 4, a complete simulation result is given, considering the PV panel module: 1Soltech 1STH-FRL-4H-250-M60-BLK. The last section presents the conclusion

and general discussion about the proposed control approach for a boost converter in an MPPT PV application.

Notations. The notations used in this paper are quite standard. $\mathbf{R}^n$ and $\mathbf{R}^{n\times n}$ refer to, respectively, the n-dimensional Euclidean space and $n \times n$ real matrices. The superscript "T" means the transpose of a matrix. The notation $P > 0$ means the matrix P is a symmetric and positive definite. The notation $s(M)$ means $M + M^T$.

2. T-S Fuzzy Model for the PV System

A first aspect of the analysis is constructing a functional model of the PV system. A T-S model from the complex nonlinear system that describes the PV system global dynamic behavior allows obtaining an exact representation using local submodels. Then, we will present the equations from the PV panel and the DC-DC boost converter's functioning. The PV panel equations are related to the DC-DC boost converter by a T-S fuzzy model were obtained from [31, 32], which describes the PV system.

2.1. PV System Model. A PV cell is a system that converts light into electricity through the photoelectric effect. A PV cell is generally represented by an equivalent circuit [31, 32] shown in Figure 1.

From Figure 1, the PV current can be defined as follows:

$$I_{\text{pv}} = I_{\text{ph}} - I_d - I_{\text{sh}}, \tag{1}$$

where I_{ph} is the photon-current which is dependent on the solar irradiance and can defined by the following:

$$I_{\text{ph}} = (I_{\text{sc}} + K_{\text{sc}}\Delta T)\frac{S_r}{S_n}, \tag{2}$$

where I_{sc} is the short-circuit current, K_{sc} is the parameter of PV cell short-circuit current, $\Delta T = T - T_{\text{ref}}$ is the variation temperature, T_{ref} is the reference temperature of PV cell, S_r is the solar irradiation, and S_n is the nominal solar irradiation. It is clear that the photon-current increases when the temperature or the solar irradiation increases.

The diode current is defined as follows:

$$I_d = I_s\left(\exp\left\{\frac{V_{\text{pv}}\text{v} + RsI_{\text{pv}}}{V_T}\right\} - 1\right), \tag{3}$$

where V_T is the thermal voltage, V_{pv} is the photovoltaic output voltage, R_s is the shunt resistance, and I_s is the saturation current which is represented by [31, 32]

$$I_s = I_{\text{rs}}\left(\frac{T}{T_{\text{ref}}}\right)^3 \exp\left\{\frac{QE_{\text{bg}}}{nK_b}\left(\frac{1}{T_{\text{ref}}} - \frac{1}{T}\right)\right\}, \tag{4}$$

where Q is the electronic charge, E_{bg} is the semiconductor band-gab energy of the PV cell, K_b is the Boltzman constant, n is the ideal PN-junction characteristic, and I_{rs} is the reverse saturation at T_{ref} and S_n. The value of reverse saturation current I_{rs} may be evaluated through the open circuit voltage V_{oc} and the short circuit current I_{sc} given by

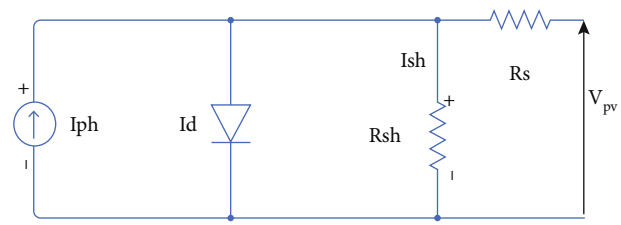


Figure 1: Solar cell equivalent circuit.

$$I_{\text{rs}} = \frac{I_{\text{sc}}}{\exp\{V_{\text{oc}}/V_T\} - 1}. \tag{5}$$

This way, the solar cell can be modeled.

2.2. DC-DC Boost Converter. The DC-DC boost converter (Figure 2) is one of the simplest types of the switch-mode converter [41, 42]. It consists of an inductor, a semiconductor switch, a diode, and a capacitor. The advantage of using the DC-DC converter is that the efficiency is high [43], because all the circuit elements, such as the inductor, capacitor, switch, and diode, present negligible losses. In practice, the efficiency of the DC-DC converter exceeds 90%, which is adequate for an energy converter. Due to the existence of commutations, the boost converter operates in two modes: the inductor stores energy, and the capacitor releases energy when the switch is closed; the inductor releases energy, and the capacitor stores energy when the switch is opened. Additionally, the output voltage can be varied with the duty cycle of the commutator with a Pulse Width Modulation (PWM) strategy. The topology is simple and effective.

Note that the boost converter is normally implemented with a MPPT algorithm, so another converter is necessary to consider a secondary application such as voltage DC source, a battery charger, or a DC/AC inverter [44]. The design of secondary conversion systems is out of the scope of this work.

Regardless of the simplicity of the boost converter, its dynamics are described by nonlinear models [30–32]. A modeling approach based on a T-S fuzzy model is presented below.

2.3. T-S Fuzzy Model. An exact representation of the global dynamic behavior of the PV system can be obtained by a T-S model using local submodels. T-S models are usually defined by if-then fuzzy-rules in the state space. The T-S PV model developed in [31, 32] is described as follows:

$$\dot{x}(t) = \sum_{i=1}^{4}\theta_i(z(t))(Ax(t) + B_i u(t) + Dw(t)), \tag{6}$$

$$y(t) = Ex(t), \tag{7}$$

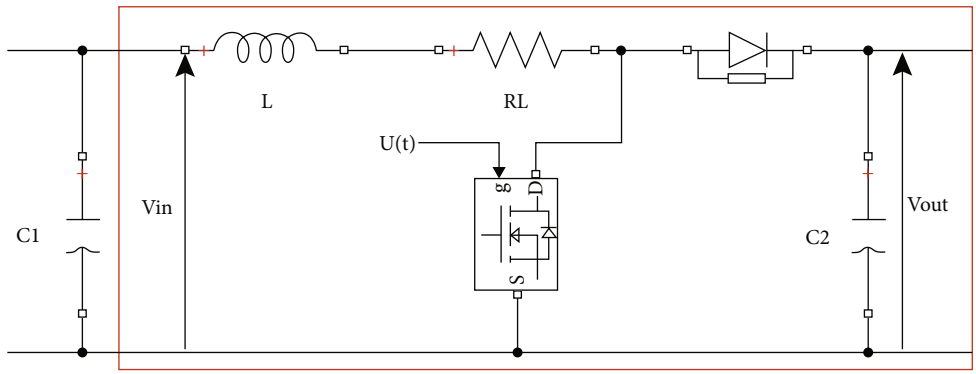


Figure 2: DC/DC boost converter.

where $A \in \mathbf{R}^{n\times n}$, $B_i \in \mathbf{R}^{n\times m}$, $D \in \mathbf{R}^{n\times w}$, and $E \in \mathbf{R}^{ny\times n}$ are constant matrices with

$$A = \begin{bmatrix} -\frac{R_L}{L} & -\frac{1}{L} & \frac{1}{L} \\ \frac{1}{C_2} & -\frac{1}{RC_2)} & 0 \\ -\frac{1}{C_1} & 0 & 0 \end{bmatrix}, B_1 = \begin{bmatrix} \frac{V_{C2_{\min}}}{L} \\ -\frac{I_{L\min}}{C_2} \\ 0 \end{bmatrix}, E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix},$$
$$B_2 = \begin{bmatrix} \frac{V_{C_{2_{\min}}}}{L} \\ -\frac{I_{L\max}}{C_2} \\ 0 \end{bmatrix}, B_3 = \begin{bmatrix} \frac{V_{C2\max}}{L} \\ -\frac{I_{L\min}}{C_2} \\ 0 \end{bmatrix} B_4 = \begin{bmatrix} \frac{V_{C2\max}}{L} \\ -\frac{I_{L\max}}{C_2} \\ 0 \end{bmatrix}, D = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{C_1} \end{bmatrix}, \tag{8}$$

and

$$x(t) = \begin{bmatrix} I_L(t) \\ V_{C2}(t) \\ V_{pv}(t) \end{bmatrix}, w(t) = I_{pv}(t),$$
$$\theta_1(z(t)) = U_{1,\min} * H_{1,\min}, \theta_2(z(t)) = U_{1,\min} * H_{1,\max},$$
$$\theta_3(z(t)) = U_{1,\max} * H_{1,\min}, \theta_4(z(t)) = U_{1,\max} * H_{1,\max}, \tag{9}$$

with $U_{1,\min}$, $U_{1,\max}$, $H_{1,\min}$, and $H_{1,\max}$ are the membership functions.

The choice of the membership function (MF) of the model significantly affects its performance. Since there is no exact procedure in the literature for calculating these membership functions, the best guideline may be the trial and error method. There are various types of membership functions, such as triangle, trapezoid, and Gaussian. The best way to achieve system performance is to use membership functions with a minimum number of changeable parameters. For example, the Gaussian MF was defined by mean and variance, while Sigmoidal MF was represented by a set of two points.

Remark 1. The representation (6) is given by using the sector bounded nonlinearity. This technique is usually applied to obtain linear submodel. Furthermore, the global model (6) is operated in the space $[I_{L,\min} I_{L,\max}] \times [V_{c2,\min} V_{c2,\max}]$.

Note that most physical systems involve time delays in their behavior, and in PV systems, delays can occur from the charging and discharging of electronic components. This delay may be small due to the construction of electronic components, but it is still necessary to take them into account. In general, the delay effect is not taken into consideration when using the boost converter directly. Nevertheless, it is known that delays are sources of instability of the system's performance and a source of uncertainties. Hence, to deal with the delay effect when using the boost converter, we propose to control the boost converter by using state delay feedback and the integral (I) controller. The next section presents the control strategy, including a numerical approach to the control design proposed in this work.

3. Control Strategy

The general control strategy proposed in this work can be seen in Figure 3, where the PV panel inputs are the temperature T and the solar irradiation S_r, while the PV panel outputs are the PV panel current l_{pv} and voltage V_{pv}. The general goal is to design both controllers the integral and the state delay feedback together with the MPPT in order to transfer the maximum power from the PV panel.

In this paper, a reference voltage V_{mppt} is obtained by using the perturb and observe (P&O) algorithm (see Figure 4) in order to get the maximum power point tracking (MPPT).

The proposed control strategy described in Figure 3 shows that the integral control input is the difference between V_{mppt} and V_{pv}. This implies that the input is the slope, as shown in Figure 4. Moreover, it noted that slope as input does not mean that a disturbance supplied to integral control since the slope parameter is part of the MPPT algorithm, and there is a different way to build it.

The strategy used to control the boost converter is shown in Figure 3. The idea behind using the integral controller (I) and the state delay feedback is that the stability region provided does not display any discontinuity, which

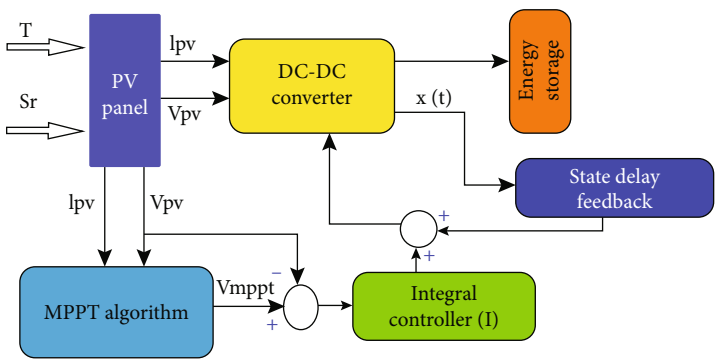


Figure 3: Schema block control method to PV model.

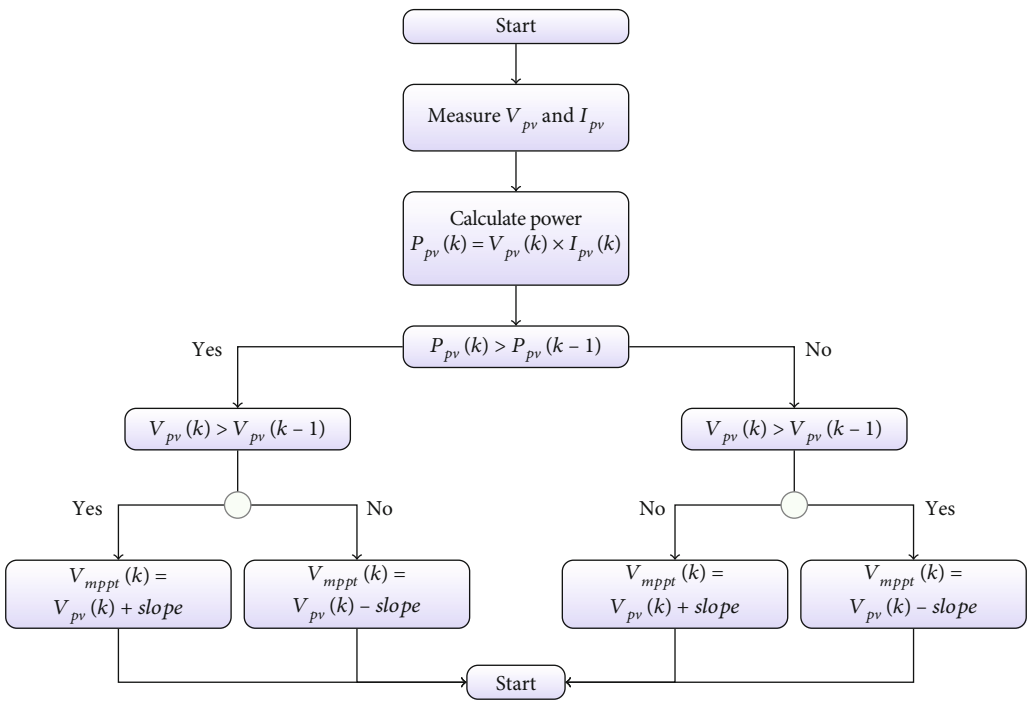


Figure 4: P&O algorithm flowchart.

Table 1: PV panel module: 1Soltech 1STH-FRL-4H-250-M60-BLK.

Specifications	Values
Cells per module (Ncell)	60
Series parallel	1
Maximum power	248.977 W
Voltage at maximum power point	30.7 V
Current at maximum power point	8.11 A
Open circuit voltage	38.4 V
Short-circuit current	8.85 A

allows a more flexible selection of the controller's gains, which can be constant in particular. The integral controller is used to regulate the error between the MPPT output and the PV panel voltage. The state delay feedback is used to command the boost converter and represent the delay's effect on the boost functioning. The control law is defined as

$$u(t) = u_1(t) + u_2(t), \tag{10}$$

where $u_1(t)$ is the overall state delay feedback described in the fuzzy rule if-then as follows:

$$u_1(t) = \sum_{i=1}^{4} \theta_i(z(t)) K_{di} x(t-h), \tag{11}$$

and h is the constant delay, $x(t) \in \mathbf{R}^n$ is the system state, and $K_{di} \in \mathbf{R}^{m \times n}$ are the delay-feedback controller gains. Moreover, $u_2(t)$ is the integral controller given by

$$u_2(t) = K_I \int e(t)dt, \tag{12}$$

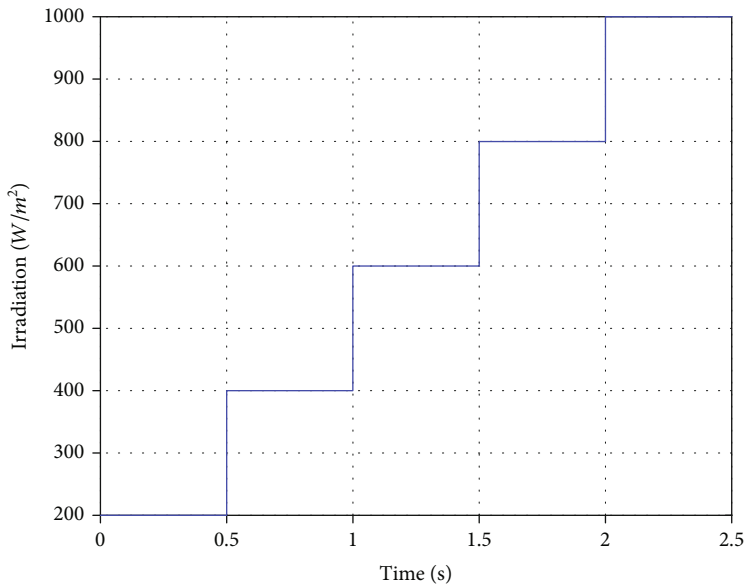


Figure 5: Variation of irradiation.

with K_I the integral controller gain and $e(t) = V_{\text{mppt}} - V_{\text{pv}}$ the error between the MPPT output and the PV output.

The integral controller produces the output signal given by $u_2(t)$, which is proportional to the integral of the input signal $e(t)$. This V_{mppt} is compared with the PV voltage (V_{pv}), and the integral controller receives an error signal ((t)). The desired response can be achieved by designing the integral gain (K_I). Once the boost converter receives energy from the PV panel, the integral controller starts to function, the value of the duty cycle varies, and the input value sensed by the integral controller changes.

The main advantage of the proposed control is its simplicity, mainly because now it is considered a unique constant integral controller for all levels of irradiation. It is well known that the integral controller is used in many applications since it can be implemented practically with basic electronic components [45–48], which significantly simplifies the implementation and costs.

3.1. Control Design. Based on the experimental step response, Ziegler-Nichols and other authors have proposed several rules, for tuning the PI controller [49–51]. Several results associated with the analytical calculation of the gain margin, the phase margin of the delayed systems, and the time-delay ratio have considered [49]. These rules can, of course, be applied to known mathematical models. Such rules suggest a set of values of K_p and K_I that will ensure the stable operation of the system. However, the resulting system may present performance issues, such as excessive overshoots, in which case fine-tuning rules are required until an acceptable result is obtained. The fine-tuning rules of Ziegler-Nichols provide a systematic way of choosing the values of controller gains.

Next, a convex approach is proposed in order to design both the integral and state delay feedback controllers,

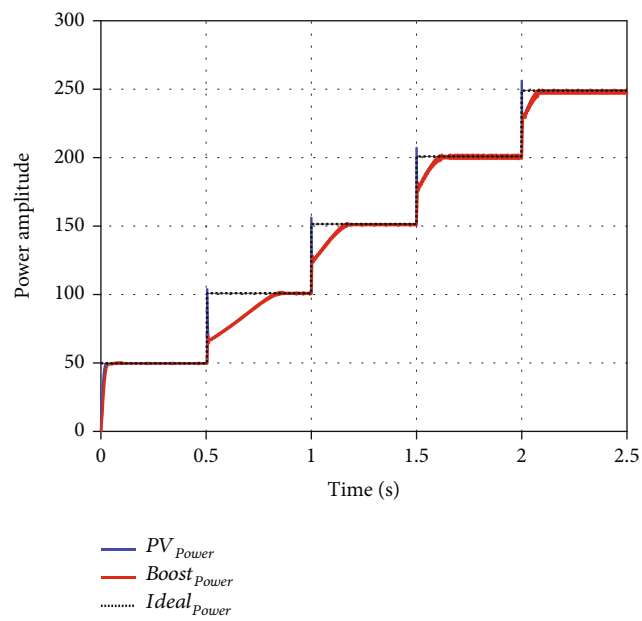


Figure 6: PV power under various levels of irradiation.

obtaining the gains K_I and K_{di} for $i = \{1, 2, 3, 4\}$ that will be used to generate the DC-DC boost converter control signal. In order to obtain the latter, replace (10) and (11) into (6):

$$\dot{x}(t) = A(t)x(t) + A_d(t)x(t-h) + D(t)w(t) + B(t)K_I f(e(t)), \tag{13}$$

$$y(t) = Ex(t), \tag{14}$$

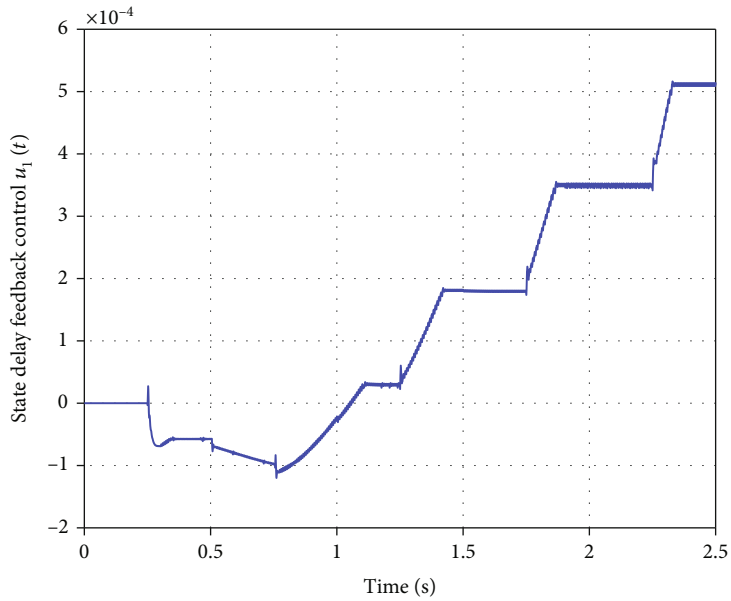


Figure 7: Evolution of control law $u_1(t)$.

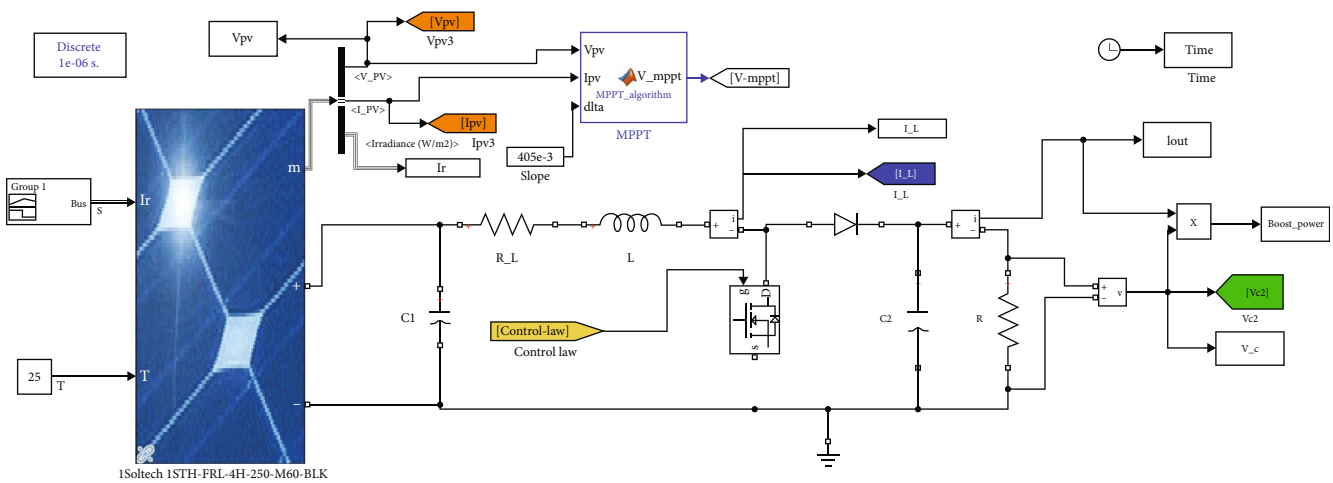


Figure 8: Simulink block model of PV model control.

where

$$A(t) = \sum_{i=1}^{4} \theta_i(z(t))A_i, B(t) = \sum_{i=1}^{4} \theta_i(z(t))B_i,$$

$$A_d(t) = B(t)K_d(t), D(t) = \sum_{i=1}^{4} \theta_i(z(t))D_i, f(e(t)) = \int e(t)dt. \tag{15}$$

It is also assumed that there exists a positive scalar α such that $f^T(e(t))f(e(t)) < \alpha x^T(t)x(t), \forall t > 0$.

The control technique used helps to display the delay in the system formulation and also to show the effect of the delay on the boost converter. Furthermore, a delay-dependent condition of the closed-loop system can be easily provided. Moreover, to minimize the impact of the exogenous disturbances on the output of the boost converter, the H_∞ performance index J is introduced.

$$J = \int_0^t \{y^T(s)y(s) - \gamma^2 \bar{\omega}^T(s)\bar{\omega}(s)\} ds, \tag{16}$$

where γ is a positive scalar.

Our aim is to develop a delay-dependent stabilization method, which provides a controller gain $K(t)$ and integral control parameter K_I such that the closed-loop system (13) is stable for any positive constant delay h and the H_∞ performance index $J < 0$ for all $t > 0$. The next theorem presents a convex-based approach to solve this problem.

Theorem 2. *Consider the photovoltaic system described in (6). Based on the control strategy (10) with (11), the closed-loop system (13) is asymptotically stable with H_∞ performance $J < 0$, if*

there exist positive matrices X, $\tilde{Q}$, and $\tilde{R}$ and matrices $Y_i, i = \{1, 2, 3, 4\}$ and K_I and scalars λ, h, ρ, and γ such that the following inequalities are satisfied.

$$\tilde{\Psi} ii < 0, i = \{1 \leq, 2, 3 \leq, 4\}, \tag{17}$$

$$\Psi ij + \Psi ji < 0, 1i < j4, \tag{18}$$

where

$$\tilde{\Psi} \begin{bmatrix} \begin{bmatrix} \bar{\Psi}^i_{11} & \bar{\Psi}^{ij}_{12} & \tilde{\Psi}^i_{13} & \tilde{\Psi}^i_{14} \\ * & \Psi^{\sim}{}_{22} 00 * * - \gamma I 0 * * * - I] & \lambda h \prod_{ij}^{T} & \sum^{\sim T} X^T \begin{bmatrix} I_n \\ 0 \\ 0 \\ 0 \end{bmatrix} * \tilde{R} - 2\lambda X 00 * * - \gamma I 0 * * * - \rho I_n] \end{bmatrix} \end{bmatrix} \tag{19}$$

with

$$\begin{aligned} \bar{\Pi}\, ij &= \left[A_i X\, B_i Y_j D_i B_i K_I\right], \\ \tilde{\Sigma} &= \Sigma\Omega = [EX\, 0\, 0\, 0], \\ \Omega &= \text{diag}\, \{X, X, I, I\}, \\ \bar{\Psi}^i_{11} &= \tilde{Q} - \tilde{R} + \text{sym}\, \{A_i X\}, \\ \tilde{\Psi}^{ij}_{12} &= \tilde{R} + B_i Y_j, \\ \tilde{\Psi}^i_{13} &= D_i, \\ \tilde{\Psi}^i_{14} &= B_i K_I, \\ \tilde{\Psi}_{22} &= -\tilde{Q} - \tilde{R}. \end{aligned} \tag{20}$$

If the conditions in (17) are satisfied, the controller gains are defined as follows:

$$K_{di} = Y_i X^{-1}, i = \{1, 2, 3, 4\} \text{ and } K_I. \tag{21}$$

Moreover, $f^T(e(t)) f(e(t)) < 1/p x^T(t) x(t), \forall t > 0$.

Proof. The proof can be found in Appendix. □

Theorem 2 proposes sufficient conditions to design the state delay feedback and integral controller in (10), in order to ensure that the T-S PV model is stable and minimizes the impact of exogenous disturbances on the output of the boost converter. Since the proposed stability conditions (17) are Linear Matrix Inequality (LMI) conditions, for a given upper bound delay h and a scalar λ, the controller gains (t) and K_I can easily be determined using effective convex optimization algorithms [39]. On the other hand, note that the traditional state feedback controller (without delay) may lead to more conservative results since they are independent of the delay, which tends to be conservative, especially when the actual delay is small. Although the traditional state feedback controller (without delay) has the advantage of being simple to implement, its performance cannot be better than that of a delayed state-feedback controller, which uses the available information about the size of the delay. The delayed state-feedback controller could be viewed as a compromise between improving performance and implementation simplicity.

Remark 3. Note that the gains of all controllers in the proposed scheme are computed offline. Thus, the need for complex control electronics during the control action is not required. Besides, the simulation test (online control) needed only a couple of gains (K and K_{di}) to get the desired objectives. The considerations set out above show the simplicity and importance of the scheme proposed.

Next, this paper focuses on implementing state delay feedback with an integral controller to transfer the maximum energy from the PV panel to the boost converter's output. Particular attention has been paid to the design of an integral controller parameter K_I and state delay feedback simultaneously. For this reason, the parameters obtained from Theorem 2 will be used to simulate the PV system. Note that, since the conditions on Theorem 2 are a linear function of γ, it is possible to solve an optimization problem in order to minimize the H_∞ performance index J.

4. Simulation Result

In order to evaluate the proposed control strategy in Figure 1, an actual PV module is considered, whose specifications are listed in Table 1. Also, we consider the boost converter parameters in Figure 2 as in [30] that are $L = 10\,\text{mH}$, $R_L = 0.01$, $C_1 = 500\,\mu\text{F}$, $C_2 = 100\,\mu\text{F}$, and $R = 20\,\Omega$.

The first step is to obtain the integral gain K_I and state delay feedback gains in order to ensure that the maximum energy from the PV panel to the boost converter's output. In order to solve this problem numerically using convex optimization algorithms, Theorem 2 has been applied with $h = 0.25$ and $\lambda = 0.005$, and it is also assumed that $K_I < -1.9$, which are values that maintain the convexity of the problem. In this

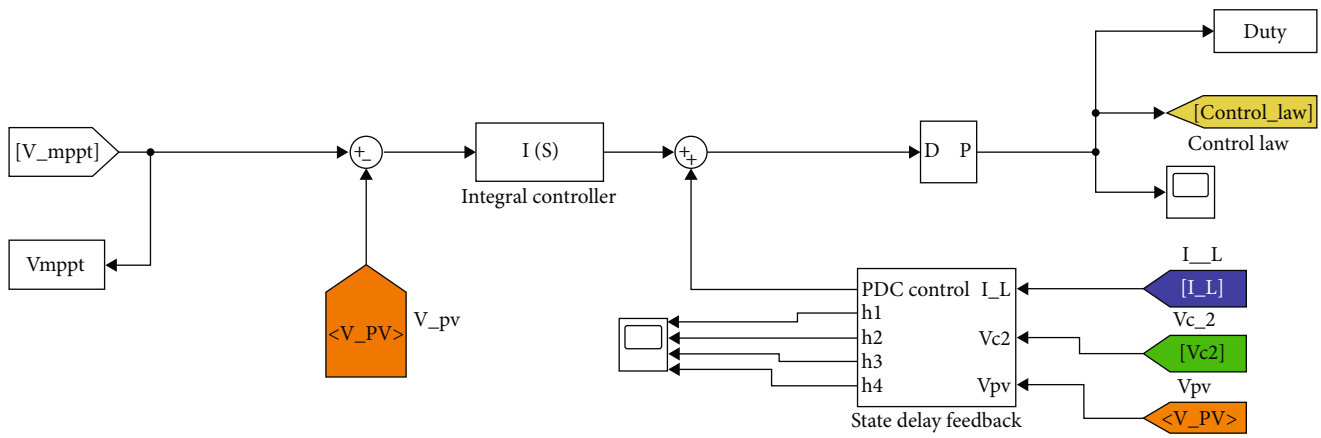


Figure 9: Simulink block model of control strategy.

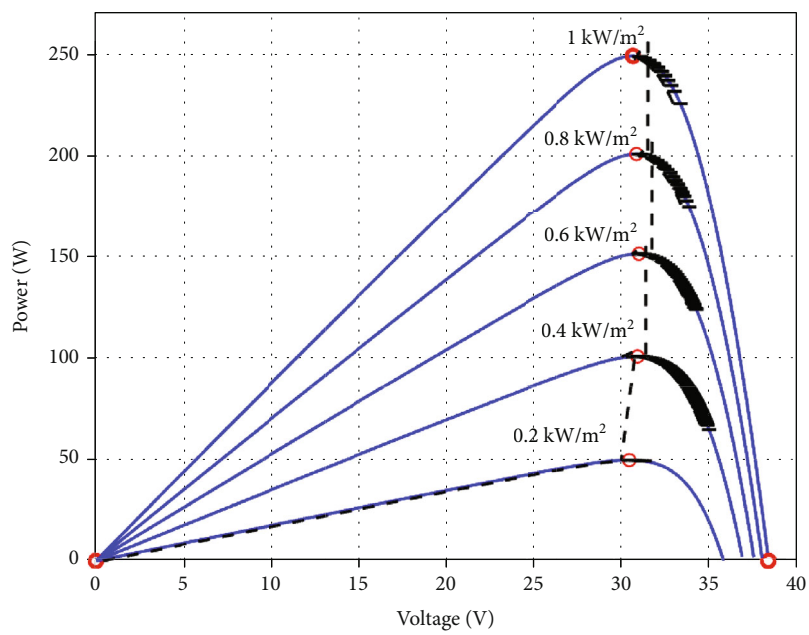


Figure 10: $P - V_{pv}$ curves in steady-state operating and P_{mppt} power from the P&O method.

case, the following gains were simultaneously computed offline: $K_I = -2.0665$ and the T-S fuzzy controller gains.

$$\begin{aligned} K_{d1} &= 10 - 3[0.1204 - 0.0043\ 0.0003], \\ K_{d2} &= 10 - 3[0.1207 - 0.0035 - 0.0067], \\ K_{d3} &= 10 - 4[-0.2964\ 0.0174\ 0.0035], \\ K_{d4} &= 10 - 3[-0.3905\ 0.0112\ 0.0210]. \end{aligned} \tag{22}$$

To verify the performance of the boost converter by using the proposed control strategy in the transient and steady-state period, Equation (23) shows the Sigmoidal membership functions used by the state delay feedback. The irradiation changes are proposed in Figure 5. The irradiation profile has transient mode at 0.5 s, 1 s, 1.5 s, and 2 s. Figure 6 shows the PV and boost converter power. Figure 7 shows the evolution of the control law applied. Figure 8 illustrates the Simulink model of the PV model control, and Figure 9 displays the Simulink model of control strategy implemented.

$$U_{1,\min} = \frac{1}{1 + e^{1.2(I_L - 1)}}, \tag{23}$$

$$H_{1,\min} = \frac{1}{1 + e^{-2.5\left(V_{C_2} - 0.3\right)}}. \tag{24}$$

From Figure 6, the power generated by the boost converter is close to the MPPT value. It can be seen that the MPPT condition is reached for a variety of irradiance conditions while the dynamic behavior occurs in a smooth manner with low picks and oscillations. Also, output power varies proportional to input irradiance, as normally seen in MPPT applications.

As shown in Figure 10, the dynamic trajectory of system follows a MPPT path. Since the algorithm for MPPT is P&O, an oscillatory behavior is observed, as normally occurs in

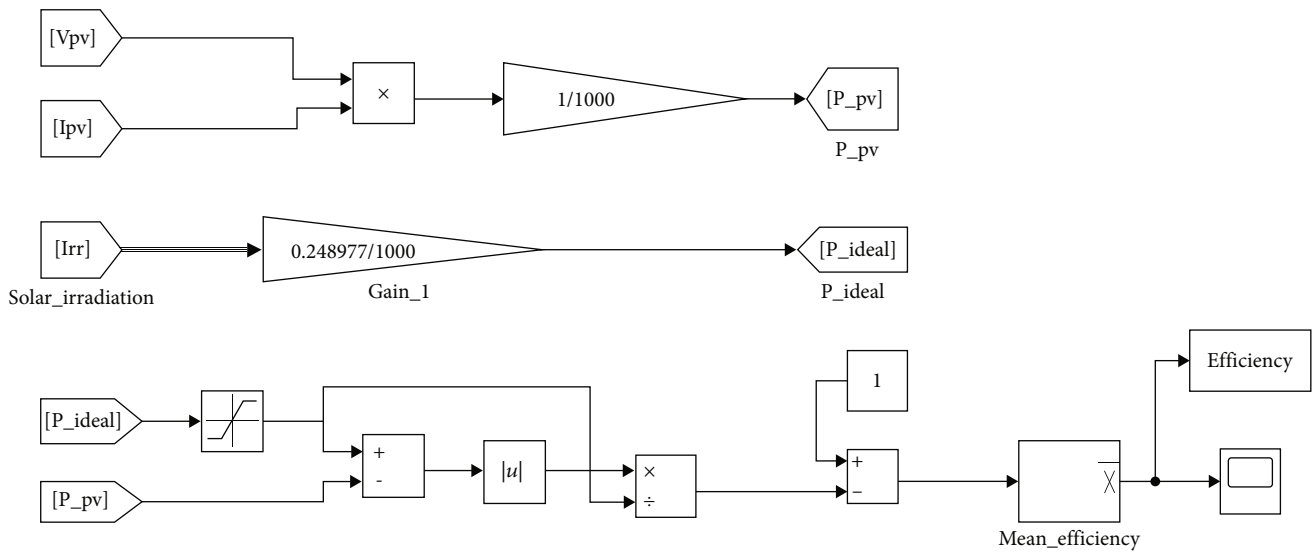


Figure 11: Simulink block model of calculating efficiency.

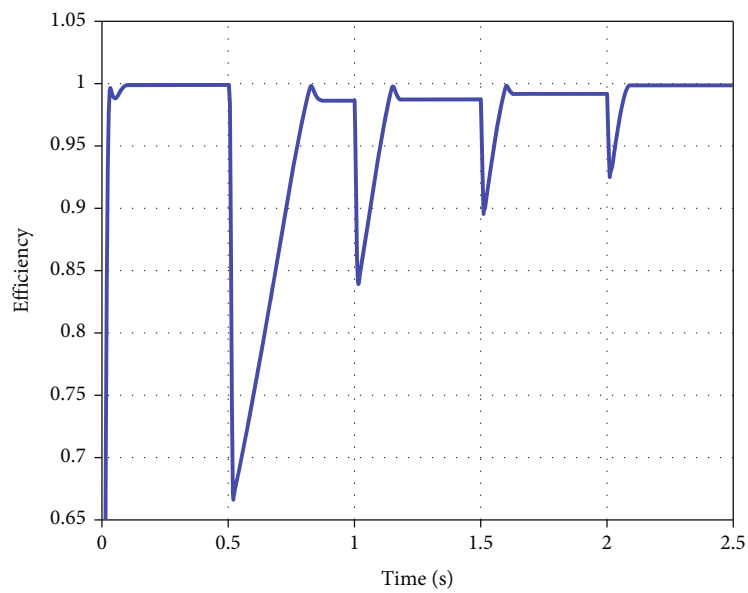


Figure 12: The PV model control efficiency.

these cases. The oscillation is not significant, with a steady-state ripple less than 2%. Furthermore, the proposed MPPT control can detect all the maximum power points, particularly those of small irradiation, while the INC and P&O MPPT control ones have not been detected (see [30, 32] and references therein). This means that the proposed MPPT control can effectively maintain the PV system operating at the MPPT for all the irradiation.

Figure 11 displays the Simulink block used to calculate the efficiency of the PV model based on the control strategy, where "Gain_1" shows the maximum power at 1000 W/m^2. Figure 12 shows the efficiency of the PV model control calculated. It can be seen that the efficiency decreases while the transient behavior occurs, which shows that the operation is not close to MPPT during the transient behavior. In steady state, the efficiency is close to 95%, which is an acceptable efficiency for these applications.

Note that the controller gains were obtained offline, which implies less computation complexity and less time in computation than using the sliding mode control and fuzzy logic control method. The sliding mode control method used the online calculation and required more accurate time to calculate [5, 6]. Besides, the fuzzy logic control method has a complex implementation [52] and requires the active power at each instant k, which takes more accurate time. Moreover, an additional electronic component needs to be connected between the boost converter and the load that introduces additional cost charges [6, 52], which is not

required for the proposed approach in this paper. Furthermore, the introduction of constant time delay allows to provide a delay-dependent condition of the PV system, which is ignored in other existing works [5, 6, 52].

5. Conclusion

This paper proposes a Takagi–Sugeno state delayed feedback and integral control for a boost converter in a PV application. The proposal obtains constant control gains, regardless of the nonlinear nature of the phenomena that usually leads to variable control gains from stratifications of the stability regions. The control framework is based on delay-dependent stabilization conditions using a Lyapunov-Krasovskii functional, leading to a convex optimization problem, where the delayed feedback and integral gains are obtained simultaneously via LMI. By simulation results, one can see that the oscillations and efficiency of the results are adequate for PV applications, showing the applicability of the method. Note also that the proposed control strategy is simple to implement and achieves acceptable performance. In terms of simplicity, all the gains necessary for the control implementation are constant and do not require online processing; instead, the gains are computed offline and then implemented to be proportional to the input signals. In terms of performance, the proposed method shows accurate MPPT tracking, acceptable steady-state efficiency (about 95%), and low steady-state ripple (less than 2%). These aspects show a proposal that simplifies the electronic implementation of MPPT algorithms for small-scale PV developments.

Appendix

Proof of Theorem 2. Let us consider the Lyapunov-Krasovskii functional described as follows:

$$V(x(t)) = x^T(t)Px(t) + \int_{t-h}^{t} xT(s)Qx(s)ds + h\int_{-h}^{0}\int_{t+\lambda}^{t} \dot{x}^T(s)R\dot{x}(s)dsd\lambda. \tag{A.1}$$

Calculating the time derivative of (A.1), we obtain

$$\dot{V}(x(t)) = 2x^T(t)P\dot{x}(t) + x^T(t)Qx(t) - x^T(t-h)Qx(t-h) + \dot{x}^T(t)h^2R\dot{x}(t) + h\int_{t-h}^{t}\dot{x}^T(s)R\dot{x}(s)ds. \tag{A.2}$$

Using the Jensen's inequality, we have

$$\begin{aligned}\dot{V}(x(t)) \le{}& 2x^T(t)P\dot{x}(t) + x^T(t)Qx(t) \\ &- x^T(t-h)Qx(t-h) + \dot{x}^T(t)h^2R\dot{x}(t) \\ &+ \left(x^T(t) - x^T(t-h)\right)R(x(t) - x(t-h)).\end{aligned} \tag{A.3}$$

Assume that the function $f((t))$ is bounded and satisfies the following condition:

$$f^T(e(t))f(e(t)) < \alpha x^T(t)x(t) \Rightarrow 0 < \alpha x^T(t)x(t) - f^T(e(t))f(e(t)), \tag{A.4}$$

where $\alpha = 1/\rho$. Adding the right side of (A.4) to (A.3), we obtain

$$\begin{aligned}\dot{V}(x(t)) \le{}& 2x^T(t)P\dot{x}(t) + x^T(t)Qx(t) - x^T(t-h)Qx(t-h) \\ &+ \dot{x}^T(t)h^2R\dot{x}(t) + \left(x^T(t) - x^T(t-h)\right)R(x(t) - x(t-h)) \\ &+ \alpha x^T(t)x(t) - f^T(e(t))f(e(t)).\end{aligned} \tag{A.5}$$

Consider now the H_∞ performance index J in (16). Under zero initial condition, we have

$$J < \int_0^t \left\{\dot{V}(x(s)) + \gamma^{-1}y^T(s)y(s)\gamma\,\bar{\omega}T(s)\bar{\omega}(s)\right\}ds = \int_0^t \xi_s^T\Psi\xi_s ds, \tag{A.6}$$

where

$$\Psi = \begin{bmatrix} \Psi 11 & \Psi 12 & \Psi 13 & \Psi 14 \\ * & \Psi 22 & 0 & 0 \\ * & * & -\gamma I & 0 \\ * & * & * & -I \end{bmatrix} + h^2\Pi^T R\Pi + \gamma^{-1}\Sigma^T\Sigma. \tag{A.7}$$

with

$$\begin{aligned}\xi_s^T &= \left[x^T(t)x^T(t-h)w^T(t)f^T(e(t))\right], \\ \Pi &= [A(t)A_d(t)D(t)B(t)K_I], \\ \Psi_{11} &= Q - R + \alpha I + \text{sym}\,\{PA(t)\}, \\ \Psi_{12} &= R + PA_d(t), \\ \Psi_{13} &= PD(t), \\ \Psi_{14} &= PB(t)K_I, \\ \Psi_{22} &= -Q - R.\end{aligned} \tag{A.8}$$

To design the controller gains, we define the following matrix $\Omega = \{X, X, I, I\}$ with $X = P^{-1}$. Pre- and post-multiplying Ψ by Ω and its transpose, we obtain

$$\tilde{\Psi} = \Omega T\Psi\Omega = \begin{bmatrix} \tilde{\Psi}11 & \tilde{\Psi}12 & \tilde{\Psi}13 & \tilde{\Psi}14 \\ * & \tilde{\Psi}22 & 0 & 0 \\ * & * & -\gamma I & 0 \\ * & * & * & -I \end{bmatrix} + \Omega^T\left\{h^2\Pi^T\lambda\left(\lambda^2R^{-1}\right)^{-1}\lambda\Pi + \gamma^{-1}\Sigma^T\Sigma\right\}\Omega. \tag{A.9}$$

By using relaxation method, Equation (A.9) can be described as follows:

$$\tilde{\Psi} = \sum_{i=1}^{4} \theta_i^2(z(t))\tilde{\Psi}_{ii} + \sum_{i=1}^{4}\sum_{i<j}^{4} \theta_i(z(t))\theta_i(z(t))\left(\tilde{\Psi}_{ij} + \tilde{\Psi}_{ji}\right). \tag{A.10}$$

By considering the variable changes $\tilde{Q} = XQX$ and $\tilde{R} = XRX$, and the controller design $Y(t) = K(t)$ with Schur Complement Lemma and right-hand side of inequality (A.11), we obtain conditions in (17).

$$\left(\lambda R^{-1} - X\right)R\left(\lambda R^{-1} - X\right) > 0 \Longleftrightarrow -\lambda^2 R^{-1} < \tilde{R} - 2\lambda X. \tag{A.11}$$

If conditions in (17) are satisfied, it means that (A.10) is satisfied. This implies that $J < 0$, and the closed-loop system is asymptotically stable with H_∞ performance. This completes the proof. □

Abbreviations

DSP:	Digital signal processor
INC algorithm:	Incremental conductance algorithm
FLC:	Fuzzy logic controller
LMIs:	Linear matrix inequalities
MF:	Membership function
MPPT:	Maximum power point tracking
PDC:	Parallel distributed compensation
P&O algorithm:	Perturb and observe algorithm
PI:	Proportional-integral
PV systems:	Photovoltaic systems
PWM:	Pulse width modulation
T-S model:	Takagi–Sugeno model.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare no potential conflict of interests.

Acknowledgments

This work has been supported by "Fondo Nacional de Desarrollo Científico y Tecnológico"—Fondecyt, Chile, under grants 3190378 and 1191302.

References

[1] IRENA, "Future of Solar Photovoltaic Deployment, investment, technology, grid integration and socio-economic aspects," in *International Renewable Energy Agency*, Abu Dhabi, 2020.

[2] IRENA, "Remap: a renewable energy roadmap," in *International Renewable Energy Agency*, Abu Dhabi, 2014.

[3] N. Kumar, T. K. Saha, and J. Dey, "Cascaded two level inverter based grid connected photovoltaic system: modelling and control," in *2014 IEEE International Conference on Industrial Technology (ICIT)*, pp. 468–473, Busan, Korea, 2014.

[4] A. H. Ali, H. S. Hamad, and A. A. Abdulrazzaq, "Performance investigation of grid connected photovoltaic system modelling based on MATLAB simulation," *International Journal of Electrical and Computer Engineering*, vol. 8, no. 6, p. 4847, 2018.

[5] D. Xu, G. Wang, W. Yan, and X. Yan, "A novel adaptive command-filtered backstepping sliding mode control for PV grid-connected system with energy storage," *Solar Energy*, vol. 178, pp. 222–230, 2019.

[6] M. R. Kumar, S. Satya Narayana, and G. Vulasala, "Advanced sliding mode control for solar PV array with fast voltage tracking for MPP algorithm," *International Journal of Ambient Energy*, vol. 41, no. 10, pp. 1192–1200, 2020.

[7] G. Masson and I. Kaizuka, *Trends in Photovoltaic Applications 2019*, International Energy Agency Photovoltaic Power Systems Programme, 2019.

[8] Ltd GEMP, *Projections for Distributed Energy Resources–Solar PV and Stationary Energy Battery Systems*, Australian Energy Market Operator, 2020.

[9] A. Jäger-Waldau, *PV Status Report 2019*, European Commission, Joint Research Centre, 2019.

[10] B. Swarnakar and A. Datta, "Design and implementation of PWM charge controller and solar tracking system," *International Journal of Science and Research*, vol. 5, no. 5, pp. 1214–1217, 2016.

[11] S. B. Sepulveda Mora, E. A. Luna Paipa, M. A. Laguado Serrano, and L. F. Bustos Márquez, "Performance comparison between PWM and MPPT charge controllers," *Scientia et Technica*, vol. 24, no. 1, pp. 6–11, 2019.

[12] A. N. A. Ali, M. H. Saied, M. Z. Mostafa, and T. M. Abdel-Moneim, "A survey of maximum PPT techniques of PV systems," in *2012 IEEE Energytech*, pp. 1–17, Cleveland, OH, USA, 2012.

[13] P. Shah and B. Singh, "Adaptive observer based control for roof-top solar PV system," *IEEE Transactions on Power Electronics*, vol. 35, no. 9, pp. 9402–9417, 2019.

[14] A. Belkaid, I. Colak, K. Kayisli, and R. Bayindir, "Improving PV system performance using high efficiency fuzzy logic control," in *2020 8th International Conference on Smart Grid (icSmartGrid)*, pp. 152–156, Paris, France, 2020.

[15] F. Belhachat and C. Larbes, "Comprehensive review on global maximum power point tracking techniques for PV systems subjected to partial shading conditions," *Solar Energy*, vol. 183, pp. 476–500, 2019.

[16] A. F. Mirza, Q. Ling, M. Y. Javed, and M. Mansoor, "Novel MPPT techniques for photovoltaic systems under uniform irradiance and partial shading," *Solar Energy*, vol. 184, pp. 628–648, 2019.

[17] U. Yilmaz, O. Turksoy, and A. Teke, "Improved MPPT method to increase accuracy and speed in photovoltaic systems under variable atmospheric conditions," *International Journal of Electrical Power & Energy Systems*, vol. 113, pp. 634–651, 2019.

[18] S. A. Mohamed and M. Abd El Sattar, "A comparative study of P&O and INC maximum power point tracking techniques for grid-connected PV systems," *SN Applied Sciences*, vol. 1, no. 2, p. 174, 2019.

[19] S. Gupta, O. Singh, and M. Ansari, *Maximum Power Point Tracking Techniques for Photovoltaic System: A Review*, Springer, 2019.

[20] R. Panigrahi, S. K. Mishra, S. C. Srivastava, A. K. Srivastava, and N. N. Schulz, "Grid integration of small-scale photovoltaic systems in secondary distribution network—a review," *IEEE Transactions on Industry Applications*, vol. 56, no. 3, pp. 3178–3195, 2020.

[21] S. Babaa, G. Murr, F. Mohamed, and S. Pamuri, "Overview of boost converters for photovoltaic systems," *Journal of Power and Energy Engineering*, vol. 6, no. 4, pp. 16–31, 2018.

[22] H. S. Lee and J. J. Yun, "Advanced MPPT algorithm for distributed photovoltaic systems," *Energies*, vol. 12, no. 18, p. 3576, 2019.

[23] C. C. Liao, "Genetic k-means algorithm based RBF network for photovoltaic MPP prediction," *Energy*, vol. 35, no. 2, pp. 529–536, 2010.

[24] S. Hadji, F. Krim, and J. P. Gaubert, "Development of an algorithm of maximum power point tracking for photovoltaic systems using genetic algorithms," in *International Workshop on Systems, Signal Processing and their Applications, WOSSPA*, pp. 43–46, Tipaza, Algeria, 2011.

[25] C. Larbes, S. A. Cheikh, T. Obeidi, and A. Zerguerras, "Genetic algorithms optimized fuzzy logic control for the maximum power point tracking in photovoltaic system," *Renewable Energy*, vol. 34, no. 10, pp. 2093–2100, 2009.

[26] A. Messai, A. Mellit, A. M. Pavan, A. Guessoum, and H. Mekki, "FPGA-based implementation of a fuzzy controller (MPPT) for photovoltaic module," *Energy Conversion and Management*, vol. 52, no. 7, pp. 2695–2704, 2011.

[27] A. Al Nabulsi and R. Dhaouadi, "Efficiency optimization of a DSP-based standalone PV system using fuzzy logic and dual-MPPT control," *IEEE Transactions on Industrial Informatics*, vol. 8, no. 3, pp. 573–584, 2012.

[28] T. L. Kottas, Y. S. Boutalis, and A. D. Karlis, "New maximum power point tracker for PV arrays using fuzzy controller in close cooperation with fuzzy cognitive networks," *IEEE Transactions on Energy Conversion*, vol. 21, no. 3, pp. 793–803, 2006.

[29] O. Guenounou, B. Dahhou, and F. Chabour, "Adaptive fuzzy controller based MPPT for photovoltaic systems," *Energy Conversion and Management*, vol. 78, pp. 843–850, 2014.

[30] K. Houda, D. Saifia, M. Chadli, and S. Labiod, "H∞fuzzy proportional integral state feedback controller of photovoltaic systems under asymmetric actuator constraints," *Transactions of the Institute of Measurement and Control*, vol. 43, no. 1, pp. 34–46, 2021.

[31] Z. B. Safia, M. Allouch, N. Harrabi, M. Bahloul, M. Chaabane, and A. Aitouche, "Fuzzy control for maximum power-point tracking of a photovoltaic system," in *2018 7th International Conference on Systems and Control (ICSC)*, pp. 263–269, Valencia, Spain, 2018.

[32] M. Allouche, K. Dahech, M. Chaabane, and D. Mehdi, "Fuzzy observer-based control for maximum power-point tracking of a photovoltaic system," *International Journal of Systems Science*, vol. 49, no. 5, pp. 1061–1073, 2018.

[33] S. Lalouni, D. Rekioua, T. Rekioua, and E. Matagne, "Fuzzy logic control of stand-alone photovoltaic system with battery storage," *Journal of Power Sources*, vol. 193, no. 2, pp. 899–907, 2009.

[34] T. Rajesh, B. Gunapriya, M. Sabarimuthu, S. Karthikkumar, R. Raja, and M. Karthik, "Frequency control of PV-connected micro grid system using fuzzy logic controller," *Materials Today: Proceedings*, vol. 45, pp. 2260–2264, 2021.

[35] C. S. Chiu and Y. L. Ouyang, "Robust maximum power tracking control of uncertain photovoltaic systems: a unified TS fuzzy model-based approach," *IEEE Transactions on Control Systems Technology*, vol. 19, no. 6, pp. 1516–1526, 2011.

[36] H. Khabou, M. Souissi, and A. Aitouche, "MPPT implementation on boost converter by using T–S fuzzy method," *Mathematics and Computers in Simulation*, vol. 167, pp. 119–134, 2020.

[37] X. Wang, S. Ding, X. Zhang, and X. Fan, "Further studies on robust H_∞ control for a class of Takagi–Sugeno fuzzy time-delay systems with application to continuously stirred tank reactor problems," *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, vol. 233, no. 2, pp. 103–117, 2019.

[38] D. Saifia, M. Chadli, S. Labiod, and T. M. Guerra, "Robust H_∞ static output-feedback control for discrete-time fuzzy systems with actuator saturation via fuzzy Lyapunov functions," *Asian Journal of Control*, vol. 22, no. 2, pp. 611–623, 2020.

[39] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, SIAM, 1994.

[40] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimization Methods and Software*, vol. 11, no. 1-4, pp. 625–653, 1999.

[41] F. Asadi and K. Eguchi, "Dynamics and control of DC-DC converters," *Synthesis Lectures on Power Electronics Series*, vol. 6, no. 1, pp. 1–241, 2018.

[42] U. Yilmaz, A. Kircay, and S. Borekci, "PV system fuzzy logic MPPT method and PI control as a charge controller," *Renewable and Sustainable Energy Reviews*, vol. 81, pp. 994–1001, 2018.

[43] Q. Zhao and F. C. Lee, "High-efficiency, high step-up DC-DC converters," *IEEE Transactions on Power Electronics*, vol. 18, no. 1, pp. 65–73, 2003.

[44] S. Serna-Garcés, D. González Montoya, and C. Ramos-Paja, "Control of a charger/discharger DC/DC converter with improved disturbance rejection for bus regulation," *Energies*, vol. 11, no. 3, p. 594, 2018.

[45] J. C. Basilio and S. Matos, "Design of PI and PID controllers with transient performance specification," *IEEE Transactions on Education*, vol. 45, no. 4, pp. 364–370, 2002.

[46] J. A. Morales-Saldaña, R. Loera-Palomo, and E. Palacios-Hernández, "Parameters selection criteria of proportional–integral controller for a quadratic buck converter," *IET Power Electronics*, vol. 7, no. 6, pp. 1527–1535, 2014.

[47] A. R. Al Tahtawi and M. Yusuf, "Low-cost quadrotor hardware design with PID control system as flight controller," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 17, no. 4, p. 1923, 2019.

[48] P. Prommee and K. Angkeaw, "High performance electronically tunable log-domain current-mode PID controller," *Microelectronics Journal*, vol. 72, pp. 126–137, 2018.

[49] A. O'Dwyer, "A summary of PI and PID controller tuning rules for processes with time delay. Part 1: PI controller tuning rules," *IFAC Proceedings Volumes*, vol. 33, no. 4, pp. 159–164, 2000.

[50] A. O'Dwyer, "A summary of PI and PID controller tuning rules for processes with time delay. Part 2: PID controller tuning rules," *IFAC Proceedings Volumes*, vol. 33, no. 4, pp. 211–216, 2000.

[51] D. Valério and J. S. Da Costa, "Tuning of fractional PID controllers with Ziegler–Nichols-type rules," *Signal Processing*, vol. 86, no. 10, pp. 2771–2784, 2006.

[52] T. Hassan, R. Abbassi, H. Jerbi et al., "A novel algorithm for MPPT of an isolated PV system using push pull converter with fuzzy logic controller," *Energies*, vol. 13, no. 15, p. 4007, 2020.

Hindawi
International Journal of Photoenergy
Volume 2021, Article ID 9982187, 17 pages
https://doi.org/10.1155/2021/9982187

Research Article

A Seventeen Multilevel High-Power Application Inverter with Low Total Harmonic Distortion

Ajmal Farooq,[1] Shanshan Tu,[2] Fiaz Ahmad,[3] Muhammad Zeeshan Malik,[4] Obaid U. Rehman,[5] Ghulam Hafeez,[1] and Sadaqat ur Rehman[6]

[1]*Department of Electrical Engineering, University of Engineering & Technology, Mardan, Pakistan*
[2]*Engineering Research Center of Intelligent Perception and Autonomous Control, Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China*
[3]*Department of Electrical & Computer Engineering, Air University, Islamabad, Pakistan*
[4]*School of Electronics and Information Engineering, Taizhou University, Taizhou, 318000, Zhejiang, China*
[5]*Department of Electrical Engineering, Sarhad University of Science & IT, Peshawar, Pakistan*
[6]*Department of Computer Science, National Institute, Mian Wali, Pakistan*

Correspondence should be addressed to Shanshan Tu; sstu@bjut.edu.cn

Received 21 March 2021; Revised 23 July 2021; Accepted 12 August 2021; Published 3 September 2021

Academic Editor: Laurentiu Fara

In this paper, a new topology of multilevel inverter (MLI) is designed with a fewer number of components and low total harmonic distortion (THD) for high-power photovoltaic (PV) systems. The key limitations of conventional MLI topologies are high total harmonic distortion (THD) and the use of a large number of switching components due to which the cost of the overall inverter is high. In conventional MLI, THD can be significantly reduced by the addition of a large value filter element at the input side; however, it will result in increased size and cost. Thus, achieving a pure sinusoidal AC at the output and to maintain a low THD level is a major issue in conventional MLIs. The proposed MLI has the advantage of decreasing the output THD by using a modified form of the cascaded H-Bridge structure and sine pulse width modulation technique. The proposed inverter consists of 6 unidirectional switches and 2 bidirectional switches, and there is no extra requirement for additional voltage balancing capacitors or clinching diodes. The individual switching states and SPWM operation for generating the gate pulses of the proposed MLI are discussed in detail. Relevant waveforms are plotted, equations are derived, and mathematical analysis is carried out. A steady-state analysis of the proposed MLI demonstrates an output voltage with 17 levels while using only four DC sources. Simulation results of the proposed MLI for single-phase and three-phase structures are obtained, and comparison is carried out with existing MLI topologies which shows that the proposed MLI has significantly low THD and better performance. From the results, it is clear that the proposed MLI has a THD of 3.52% in comparison with four conventional MLIs whose THDs are 6.1%, 6.63%, 7.3%, and 9.93%. Moreover, the proposed MLI generates 17 voltage levels by using only 08 switching devices, whereas the conventional MLIs use more than 10 switching devices for the generation of 15 voltage levels.

1. Introduction

Multilevel inverters are a candidate topology for high-voltage and high-power applications in industries nowadays. Multilevel inverters offer various advantages including low THD, simple to deal with, and compact size as compared to conventional inverters [1–4]. Research is ongoing on MLIs, and there are various topologies available in the literature. The selection of MLI for an application is based upon cost, complexity, losses, and THD. Figure 1 shows the shape of the output voltage of a 7-level MLI [5–9].

A diode-clamped MLI is presented in [10, 11]. As clear by its name, diode-clamped MLI requires a propping device. Capacitors are used for the division of DC voltage into

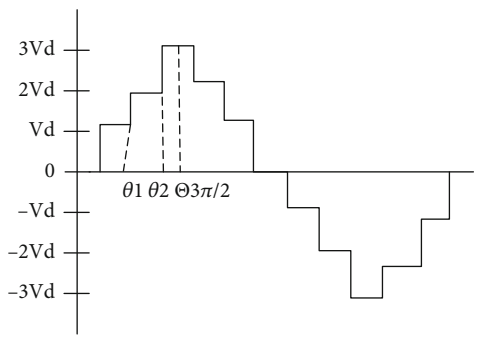


Figure 1: 7-level MLI.

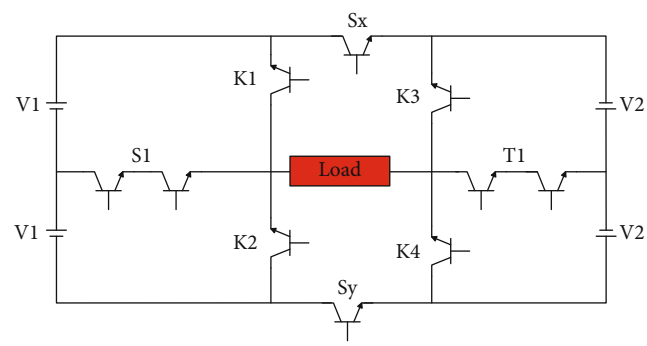


Figure 2: Proposed MLI.

switches. $N-1$ switch sets are necessary for the N number of required levels. The drawbacks of this topology are that in between number of diodes and the count of levels, the quadratic association is very difficult to compute, especially when it ends up upsetting when the count of levels gets higher [12, 13].

A flying capacitor MLI is very much similar to the diode-clamped MLI. The flying capacitor MLI requires that the capacitor must be previously charged for its operation. Besides the benefits of low THD, precharging of capacitors is troublesome [14].

The fundamental and well-understood topology is cascaded H-Bridge MLI. This type of MLI has been utilized for single as well as three-phase transformations. This type of MLI utilizes a certain type of H-Bridge that comprises a diode as well as switches. During an instance, 3 levels of voltage are necessary for MLI. It can be achieved by a solitary or single H-Bridge in cascaded H-Bridge MLI. In cascaded H-Bridge MLI, fewer components like capacitors and switches are required. It requires fewer segments when contrasted with different procedures. But every H-bridge MLI requires an independent DC source [5].

A switch ladder MLI is presented in [6] which is a modified form of cascaded H-Bridge MLI. This MLI utilizes few components and provides an output waveform that is almost sine waveform, which has low THD. This topology also uses more components [15].

A simple circuit (Op-Amp) controlled voltage source MLI is presented that utilized PWM strategy for harmonics decrease and demonstrates the best way to produce SPWM distinctive Op-Amp circuits where the passive type of filters is utilized toward the output for the reduction of harmonics; in this way, the components of the inverter are increased [16].

Another inverter named cascaded H-Bridge MLI with a phase disposition technique is presented in [17]. This work introduced a single phase of cascaded H-bridge MLI, and more elevated levels of voltages were attained with less number of parts utilizing the phase disposition system. The number of voltage levels was enlarged; however, this circuit incited high THD [17].

Reduced switch count multilevel inverter topologies using the cascaded structure and switched capacitor techniques have been presented to lower the THD level and reduce the number of switching devices [18–20].

2. Proposed MLI

There are numerous kinds of MLI available in the market. A lot of research is going upon MLI. The layout circuit of the proposed work is displayed in Figure 2. It is an altered elucidation of the H-bridge type of MLI topology. The fundamental bit of leeway of the presented work as compared to the previously discussed work is that we can achieve more output levels by utilizing a minimum number of components. The proposed MLI has fewer switches and offers low total harmonic distortion.

The proposed MLI uses two bidirectional switches and six unidirectional switches. Each bidirectional switch is comprised of two IGBTS; hence, there are ten IGBTs and the number of independent power supplies is four. So the number of circuit elements is not much more. The more noteworthy thing in the proposed topology is that it can generate 17 levels with only eight switches. The number of levels and fewer switches assumes a significant job in the efficiency of the inverter. The increased level in output voltage demonstrates that this inverter has a low THD.

3. Steady-State Analysis of Proposed MLI

The operation of the proposed MLI is such that these ten switches are turned "on" and "off" at regular intervals by using sinusoidal pulse width modulation (SPWM). This is achieved by turning the different switches "on" assigning them 1 in binary form. The other switches are considered "off," assigning them 0 in binary form. In this way, the switching pattern of all the switches can be determined.

3.1. The 1st Switching Pattern. In the first switching state pattern, the switches K_1, K_3, and S_x are turned "on" and the switches K_2, K_4, S_1, T_1, and S_y are turned "off," and the closed-loop path can be seen as shown in Figure 3. The output voltage is equal to zero as there is no available path for the input voltage to finish the loop.

So V_{out} is

$$V_{\text{out}} = 0. \tag{1}$$

3.2. The 2nd Switching Pattern. In the second switching state pattern, the switches S_1, K_4, and S_y are turned "on" and the

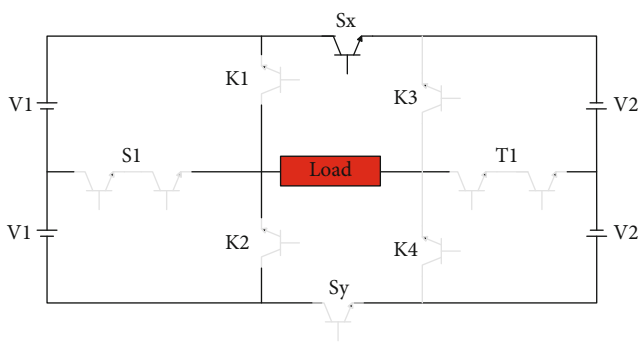


Figure 3: Current flow for the 1st switching pattern.

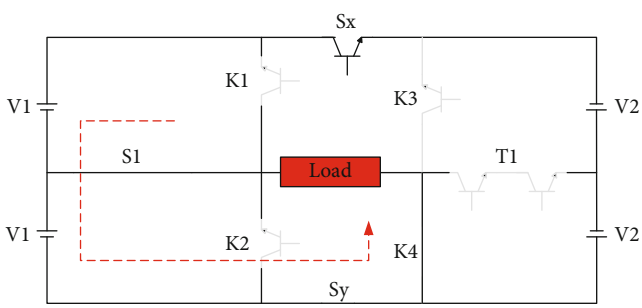


Figure 4: Current flow for the 2nd switching pattern.

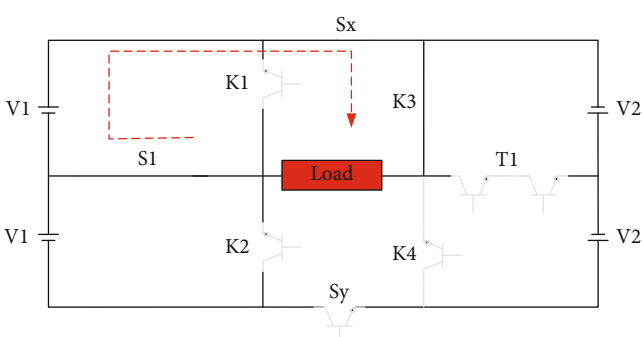


Figure 5: Current flow for the 3rd switching pattern.

switches K_1, K_2, K_3, T_1, and S_x are kept "off," so the circuit diagram can be visualized as in Figure 4.

So V_{out} is

$$V_{out} = V_1. \tag{2}$$

3.3. The 3rd Switching Pattern. In the third switching state pattern, the switches S_1, K_3, and S_x are turned "on" and the switches $K1$, K_2, K_4, T_1, and S_y are kept "off," so the circuit diagram can be visualized as in Figure 5.

So V_{out} is

$$V_{out} = -V_1. \tag{3}$$

3.4. The 4th Switching Pattern. In the fourth switching state pattern, the switches T_1, K_2, and S_y are turned "on" and

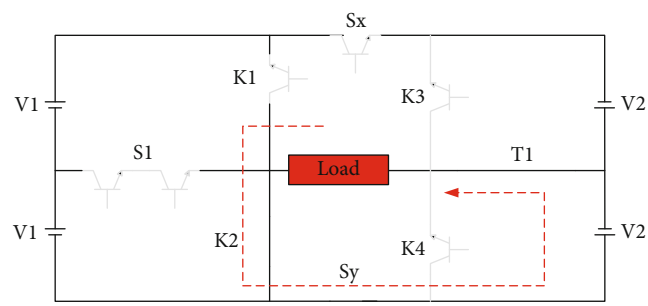


Figure 6: Current flow for the 4th switching pattern.

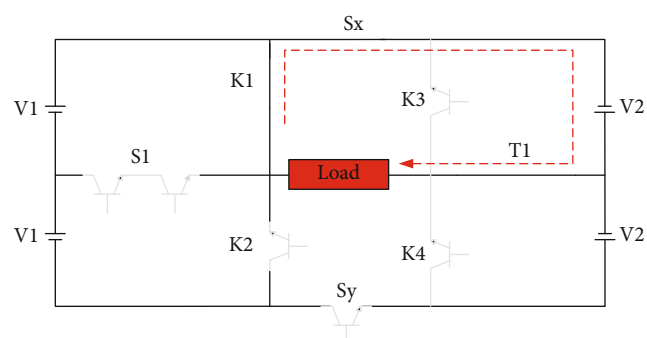


Figure 7: Current flow for the 5th switching pattern.

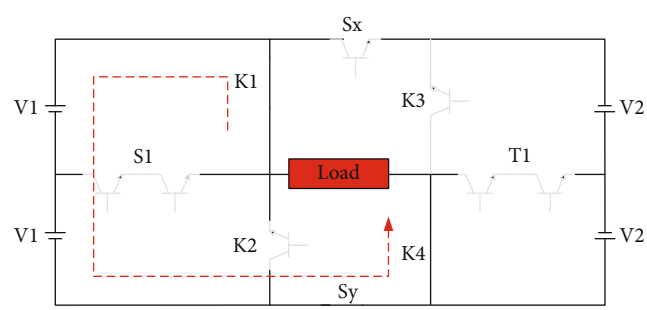


Figure 8: Current flow for the 6th switching pattern.

the switches K_1, K_3, K_4, S_1, and S_x are kept "off," so the circuit path is visualized as in Figure 6.

So V_{out} is

$$V_{out} = V_2. \tag{4}$$

3.5. The 5th Switching Pattern. In the fifth switching state pattern, the switches T_1, K_1, and S_x are turned "on" and the switches K_2, K_3, K_4, S_1, and S_y are kept "off," so the circuit diagram can be visualized as shown in Figure 7.

So V_{out} is

$$V_{out} = -V_2. \tag{5}$$

3.6. The 6th Switching Pattern. In the sixth switching state pattern, the switches K_1, K_4, and S_y are turned "on" and the switches K_2, K_3, S_1, T_1, and S_x are kept "off," so the circuit diagram can be visualized as depicted in Figure 8.

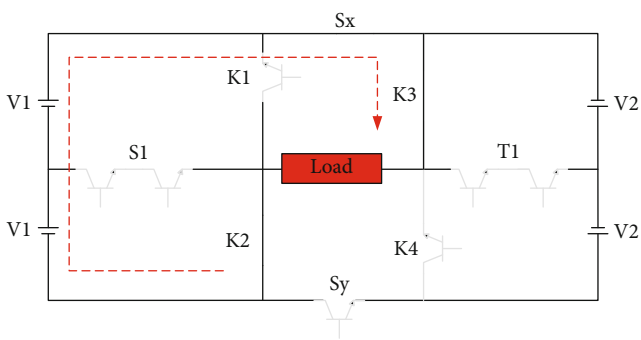


Figure 9: Current flow for the 7th switching pattern.

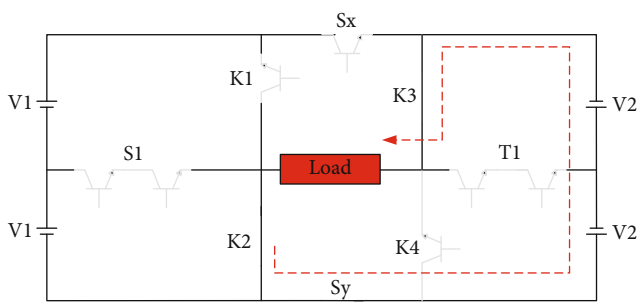


Figure 10: Current flow for the 8th switching pattern.

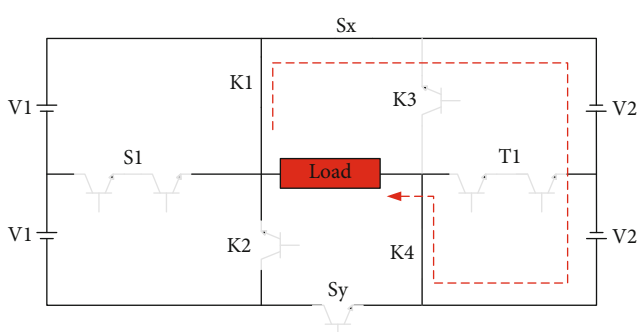


Figure 11: Current flow for the 9th switching pattern.

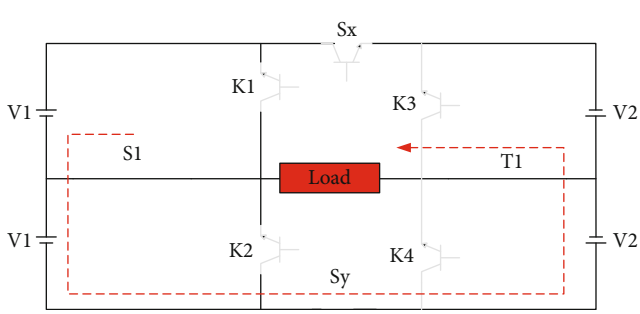


Figure 12: Current flow for the 10th switching pattern.

So V_{out} is

$$V_{\text{out}} = 2V_1. \tag{6}$$

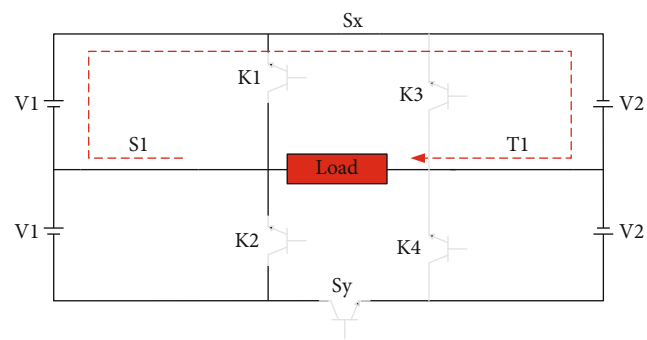


Figure 13: Current flow for the 11th switching pattern.

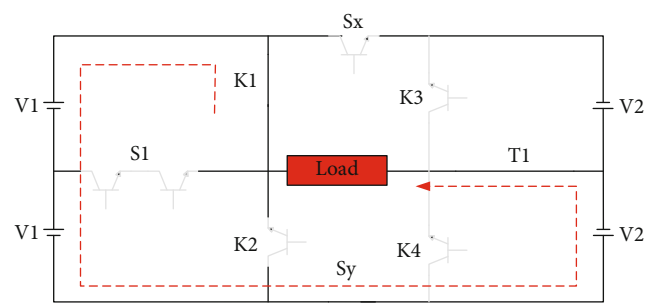


Figure 14: Current flow for the 12th switching pattern.

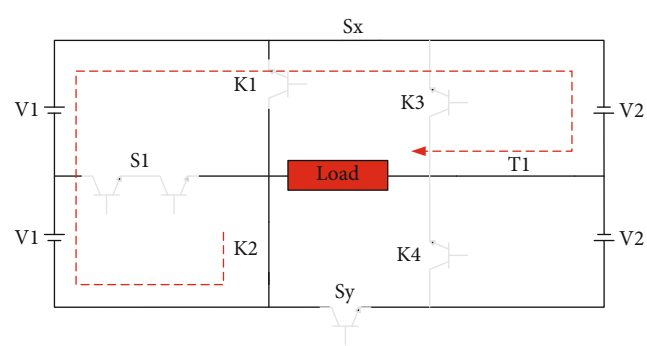


Figure 15: Current flow for the 13th switching pattern.

3.7. The 7th Switching Pattern. In the second switching state pattern, the switches S_x, K_2, and K_3 are turned "on" and the switches K_1, K_4, T_1, S_1, and S_y are kept "off," so the circuit diagram can be visualized as depicted in Figure 9.

So V_{out} is

$$V_{\text{out}} = -2V_1. \tag{7}$$

3.8. The 8th Switching Pattern. In the eighth switching state pattern, the switches K_2, K_3, and S_y are turned "on" and the switches K_1, K_4, S_1, T_1, and S_x are turned "off," so the circuit diagram can be visualized as shown in Figure 10.

So V_{out} is

$$V_{\text{out}} = 2V_2. \tag{8}$$

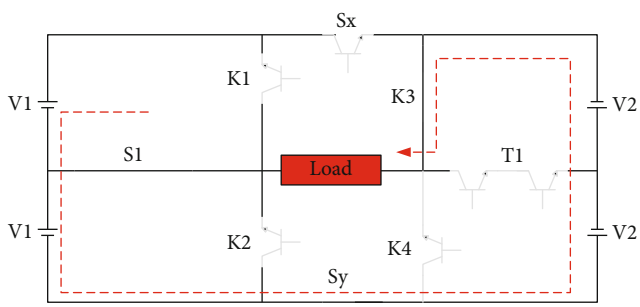


Figure 16: Current flow for the 14th switching pattern.

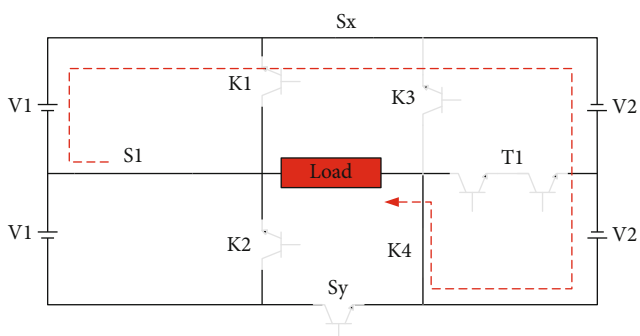


Figure 17: Current flow for the 15th switching pattern.

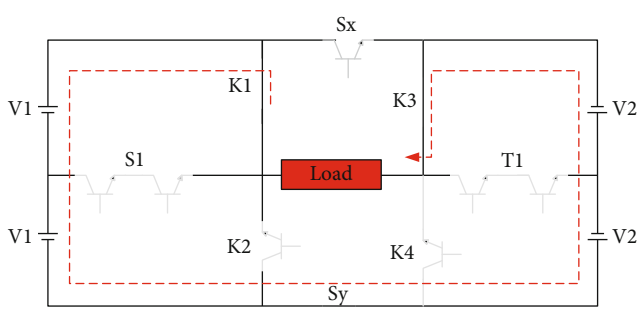


Figure 18: Current flow for the 16th switching pattern.

3.9. The 9th Switching Pattern. In the ninth switching state pattern, the switches K_1, K_4, and S_x are turned "on" and the switches K_2, K_3, S_1, T_1, and S_y are turned "off," so the circuit diagram can be visualized as depicted in Figure 11.

So V_{out} is

$$V_{out} = -2V_2. \tag{9}$$

3.10. The 10th Switching Pattern. In the tenth switching state pattern, the switches S_1, T_1, and S_y are turned "on" and the switches K_1, K_2, K_3, K_4, and S_x are turned "off," so the circuit diagram can be visualized as shown in Figure 12.

So V_{out} is

$$V_{out} = V_1 + V_2. \tag{10}$$

3.11. The 11th Switching Pattern. In the eleventh switching state pattern, the switches S_1, T_1, and S_x are turned "on" and the switches K_1, K_2, K_3, K_4, and S_y are kept "off," and the circuit path can be visualized in Figure 13.

So V_{out} is

$$V_{out} = -(V_1 + V_2). \tag{11}$$

3.12. The 12th Switching Pattern. In the twelfth switching state pattern, the switches K_1, T_1, and S_y are turned "on" and the switches K_2, K_3, K_4, S_1, and S_x are kept "off," and the circuit path can be visualized in Figure 14.

So V_{out} is

$$V_{out} = 2V_1 + V_2. \tag{12}$$

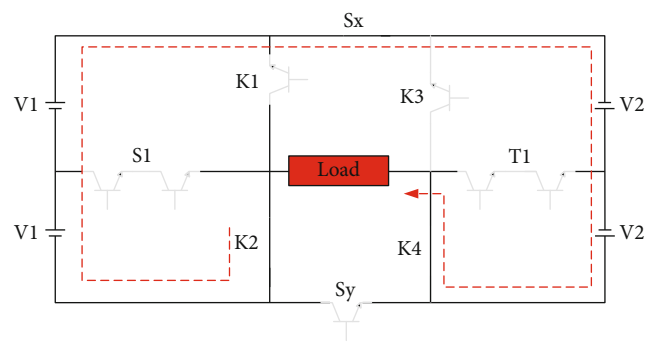


Figure 19: Current flow for the 17th switching pattern.

Table 1: Switching states of a multilevel inverter.

Switch pattern no.	K_1	K_2	K_3	K_4	S_1	T_1	S_x	S_y	V_{out} equation
1	1	0	1	0	0	0	1	0	0
2	0	0	0	1	1	0	0	1	V_1
3	0	0	1	0	1	0	1	0	$-V_1$
4	0	1	0	0	0	1	0	1	V_2
5	1	0	0	0	0	1	1	0	$-V_2$
6	1	0	0	1	0	0	0	1	$2V_1$
7	0	1	1	0	0	0	1	0	$-2V_1$
8	0	1	1	0	0	0	0	1	$2V_2$
9	1	0	0	1	0	0	1	0	$-2V_2$
10	0	0	0	0	1	1	0	1	V_1+V_2
11	0	0	0	0	1	1	1	0	$-V_1-V_2$
12	1	0	0	0	0	1	0	1	$2V_1+V_2$
13	0	1	0	0	0	1	1	0	$-2V_1-V_2$
14	0	0	1	0	1	0	0	1	V_1+2V_2
15	0	0	0	1	1	0	1	0	$-V_1-2V_2$
16	1	0	1	0	0	0	0	1	$2V_1+2V_2$
17	0	1	0	1	0	0	1	0	$-2V_1-V_2$

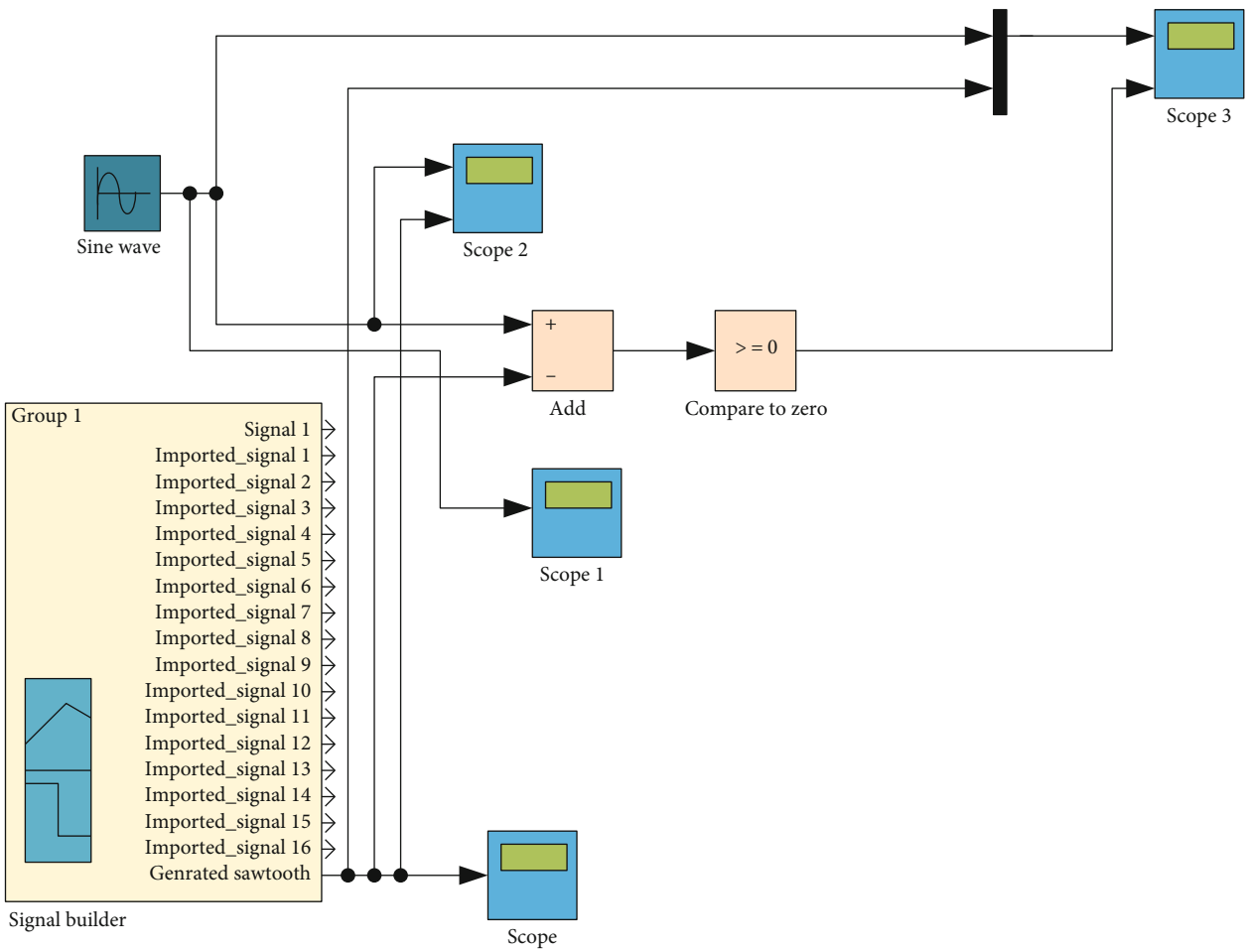


Figure 20: Circuit used for calculating switching sequence for K_1 switch.

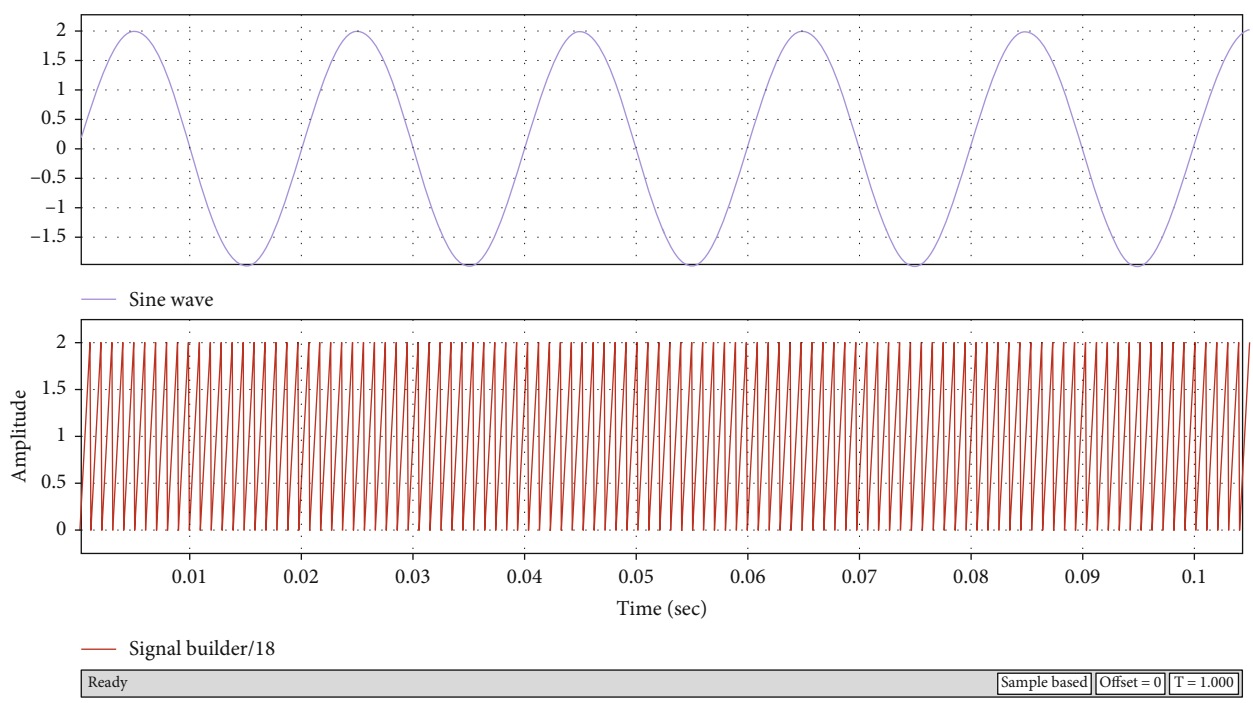


Figure 21: Sine wave and Saw tooth wave generated for calculating K_1 switch sequence.

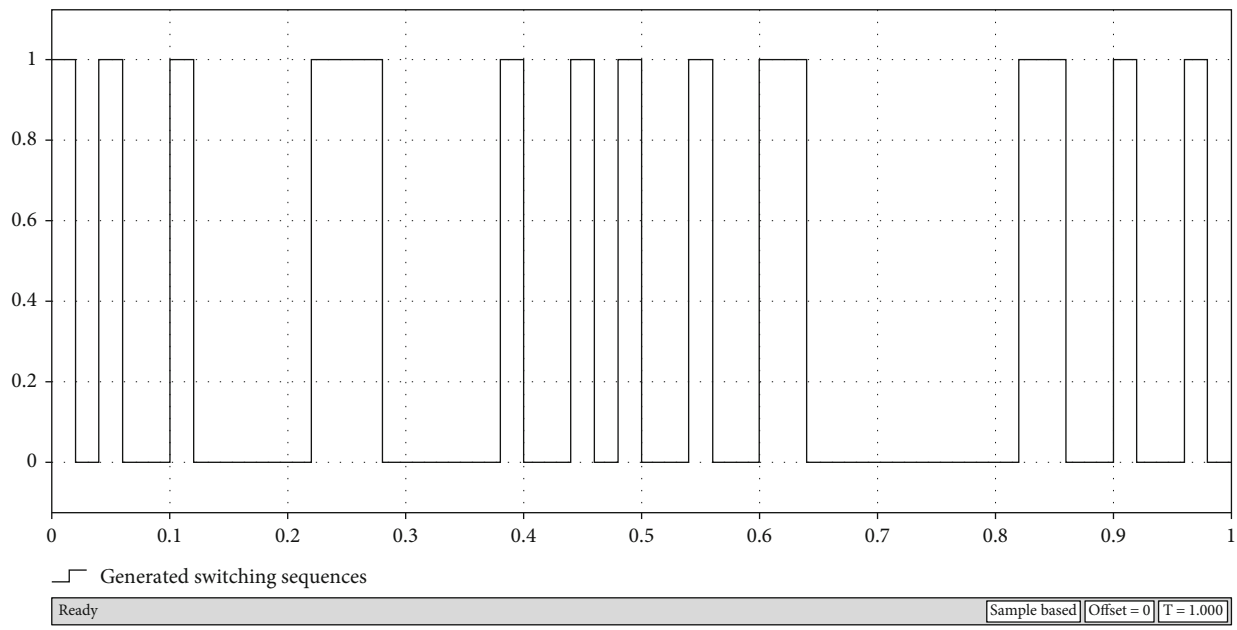


Figure 22: Generated pulses for the K_1 switch.

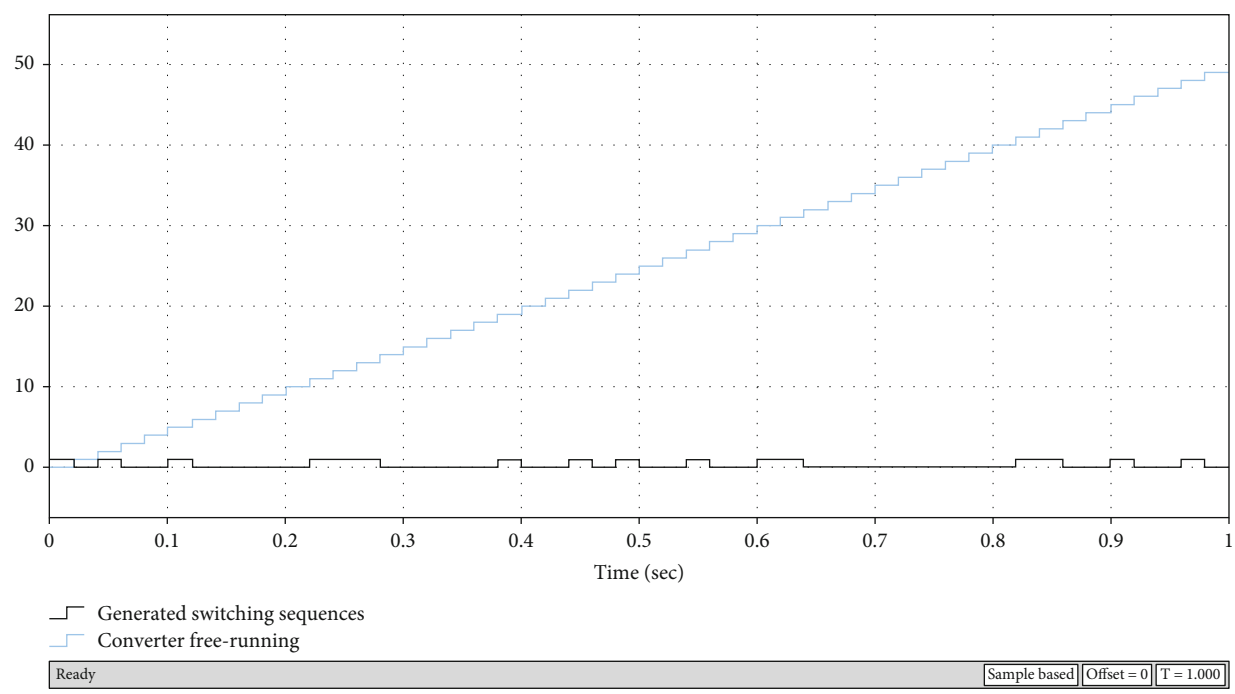


Figure 23: Pulse counter and generated switching sequences for the K_1 switch.

3.13. The 13th Switching Pattern. In the thirteenth switching state pattern, the switches T_1, K_2, and S_x are turned "on" and the switches K_1, K_3, K_4, S_1, and S_y are kept "off," and the circuit path can be visualized in Figure 15.

So V_{out} is

$$V_{out} = -(2V_1 + V_2). \tag{13}$$

3.14. The 14th Switching Pattern. In the fourteenth switching state pattern, the switches S_1, K_3, and S_y are turned "on" and the switches K_1, K_2, K_4, T_1, and S_x are turned "off," and the circuit path can be visualized in Figure 16.

So V_{out} is

$$V_{out} = V_1 + 2V_2. \tag{14}$$

3.15. The 15th Switching Pattern. In the fifteenth switching state pattern, the switches S_1, K_4, and S_x are turned "on" and the switches K_1, K_2, K_3, T_1, and S_y are turned "off," and the circuit path can be visualized in Figure 17.

So V_{out} is

$$V_{\text{out}} = -(V_1 + 2V_2). \tag{15}$$

3.16. The 16th Switching Pattern. In the sixteenth switching state pattern, the switches K_1, K_3, and S_y are turned "on" and the switches K_2, K_4, S_1, T_1, and S_x are kept "off," and the circuit path can be visualized in Figure 18.

So V_{out} is

$$V_{\text{out}} = 2V_1 + 2V_2. \tag{16}$$

3.17. The 17th Switching Pattern. In the seventeenth switching state pattern, the switches K_2, K_4, and S_x are turned "on" and the switches K_1, K_3, S_1, T_1, and S_y are kept "off," the circuit path can be visualized in Figure 19.

So V_{out} is

$$V_{\text{out}} = -2(\mathrm{V}_1 + \mathrm{V}_2). \tag{17}$$

All the switching states and voltage equations calculated for different closed-loop paths of the schematic in Figure 2 are reproduced in Table 1.

4. Voltage Stress and Switching Loss Estimation

To generate the gate pulses based on the switching sequence, a carrier-based adjustment technique is used. As all the switches turns on and off thousand times in a second, there is switching power loss in each switching device due to the on state current and off state voltage. In order to estimate the switching losses, the loss in switch K_1 is calculated theoretically as follows:

$$\begin{aligned} V_1 &= 25\,\text{volts}, \\ V_2 &= 75\,\text{volts}, \\ V_O &= 200\,\text{rms}, \\ F_{\text{SW}} &= 1\,\text{Khz}, \\ I_O &= 5\text{A}. \end{aligned} \tag{18}$$

Voltage stress across switch K_1 is given as

$$V_{K_1(\text{Stress})} = V_1 = 25\,\text{volts}. \tag{19}$$

Current stress through switch K_1 is given as

$$I_{K_1(\text{Stress})} = 5\,\text{amperes}. \tag{20}$$

Switching power loss across switch K_1 is given as

Table 2: Switching sequence of an MLI.

K_1	K_2	K_3	K_4	S_1	T_1	S_X	S_y
1	0	1	0	0	0	1	0
0	0	0	1	1	0	0	1
1	0	0	1	0	0	0	1
0	1	0	0	0	1	0	1
0	0	0	0	1	1	0	1
1	0	0	0	0	1	0	1
0	1	1	0	0	0	0	1
0	1	1	0	0	0	0	1
0	0	1	0	1	0	0	1
0	0	1	0	1	0	0	1
0	0	1	0	1	0	0	1
1	0	1	0	0	0	0	1
1	0	1	0	0	0	0	1
1	0	1	0	0	0	0	1
0	0	1	0	1	0	0	1
0	0	1	0	1	0	0	1
0	0	1	0	1	0	0	1
0	1	1	0	0	0	0	1
0	0	1	0	0	0	0	1
1	0	0	0	0	1	0	1
0	1	0	0	1	1	0	1
0	0	0	0	0	1	0	1
1	0	0	1	0	0	0	1
0	0	0	1	1	0	0	1
1	0	1	0	0	0	1	0
0	1	1	0	1	0	1	0
0	0	1	0	0	0	1	0
1	0	0	0	0	1	1	0
0	1	0	0	1	1	1	0
0	0	0	1	0	1	1	0
1	0	0	1	0	0	1	0
1	0	0	1	0	0	1	0
0	0	0	1	1	0	1	0
0	0	0	1	1	0	1	0
0	1	0	1	1	0	1	0
0	1	0	1	0	0	1	0
0	0	0	1	0	0	1	0
0	0	0	1	0	0	1	0
0	0	0	1	1	0	1	0
0	0	0	1	1	0	1	0
0	0	0	1	1	0	1	0
1	0	0	1	0	0	1	0
1	0	0	1	0	0	1	0
0	1	0	0	0	1	1	0

$$P_{K_1(\text{SW})} = \frac{1}{6}\left[V_{K_1(\text{Stress})} * I_{K_1(\text{Stress})} * F_{\text{SW}}\right] * \left[tr + t_f\right]. \tag{21}$$

tr and t_f have been taken 20 nanoseconds for a typical N-channel Mosfet.

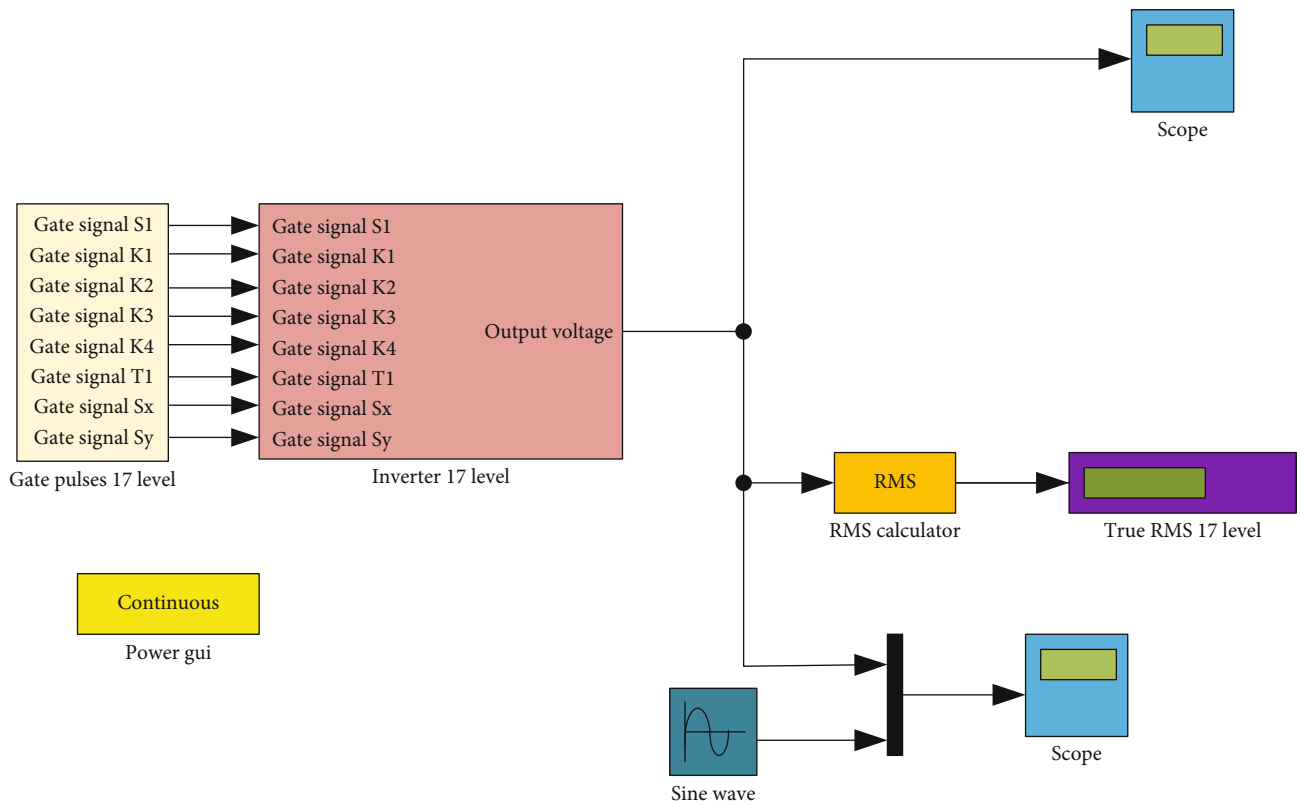


Figure 24: 17-level multilevel inverter.

Table 3: Circuit component values used in simulation and experimental setup.

S. no.	Component	Value
1	Vin	220 volts
2	Lf	5 mH
3	Lb	200 uH
4	C1&C2	330 nF
5	Cbus	10 nF
6	Lm	100 uH
7	Cr	30 nF
8	Lr	20 uH
9	Co	10 nF
10	Switching frequency	10 Hz

$$P_{K_1(\mathrm{SW})} = \frac{1}{6}[25 * 5 * 1000][20 * 10 - 9 + 20 * 10 - 9],$$
$$P_{K_1(\mathrm{SW})} = 0.417 \text{ milliwatts}. \tag{22}$$

Similarly, the switching losses in other switches can be calculated.

Table 4: Output voltage values.

S. no	V_{out} equation	V_{out} value (V)
1	0	0
2	V_1	25
3	$-V_1$	-25
4	V_2	75
5	$-V_2$	-75
6	$2V_1$	50
7	$-2V_1$	-50
8	$2V_2$	150
9	$-2V_2$	-150
10	V_1+V_2	100
11	$-(V_1+V_2)$	-100
12	$2V_1+V_2$	125
13	$-(2V_1+V_2)$	-125
14	V_1+2V_2	175
15	$-(V_1+2V_2)$	-175
16	$2V_1+2V_2$	200

5. Generation of Gate Pulses for Switching Operation

To generate the gate pulses based on the switching sequence, a carrier-based adjustment technique is used since it is useful

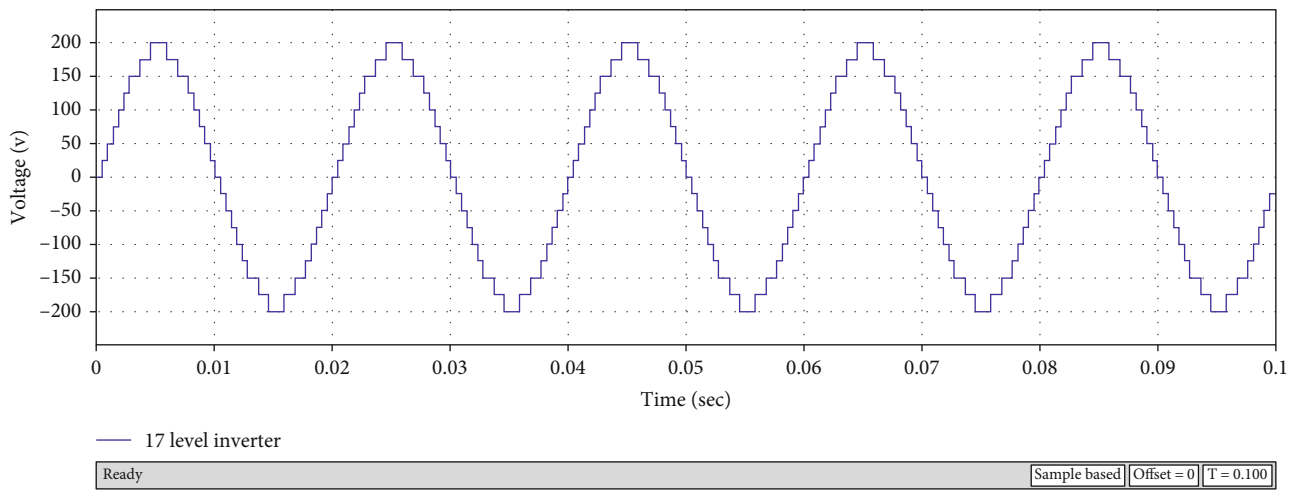


Figure 25: Single-phase 17-level MLI with $t = 0.1$.

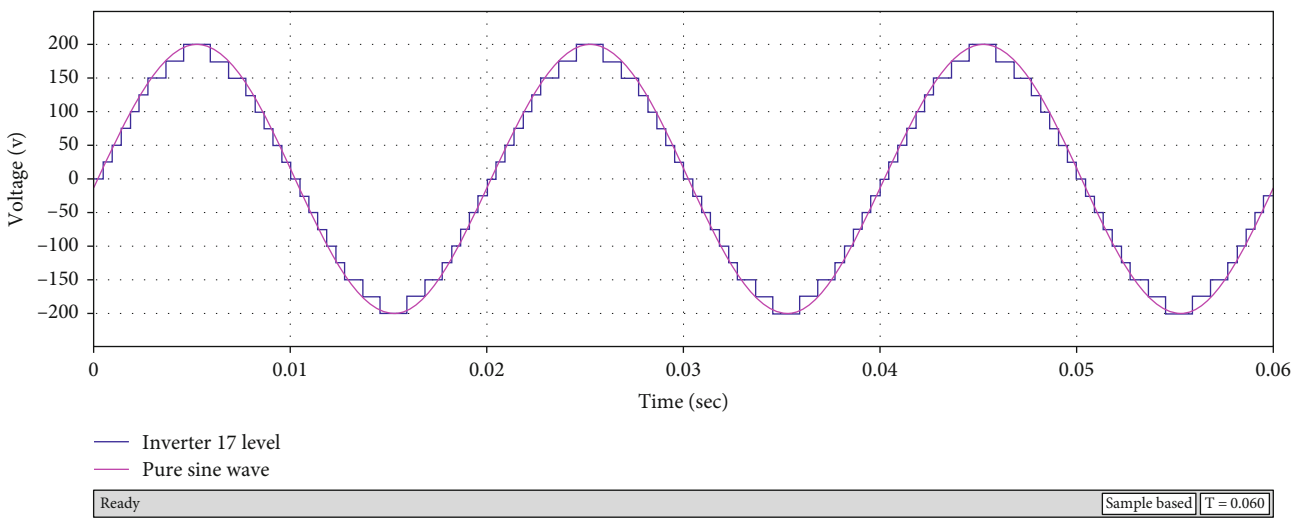


Figure 26: Single-phase 17-level MLI with $t = 0.06$.

in minimizing THD. An example of the implementation of this technique in MATLAB/Simulink is shown in Figure 20, where switching sequence for switch K_1 is generated using a constant SPWM technique. Sine wave and sawtooth waveforms (see Figure 21) are used to generate the switching sequence for switch K_1 (see Figure 22).

Figure 20 shows the circuit, used for calculating the switching sequence of pulses, for the K_1 switch.

Because of playing out the above SPWM method, pulses produced for the K_1 switch are shown in Figure 22.

Figure 22 demonstrates the pulses that are obtained by SPWM activity for the further task of the K_1 switch. To decide on the on and off states of a switch, a counter is used. For example, the switching sequence generation is demonstrated for the switch K_1 as shown in Figure 23. It can be noted that there are 48 states as calculated by the counter that represents the of and off states of the switching sequence for switch K_1. Similarly, switching sequences can be computed for all the switches.

Table 2 demonstrates the total 48 switching sequence for all the eight switches utilized in Figure 2.

6. Results and Discussion

The 17-level multilevel inverter with switches and DC sources was implemented in 203 MATLAB/Simulink as shown in Figure 24. Various values used for generating the simulation results as well as the experimental results are given in Table 3.

6.1. Theoretical Results. For the proposed MLI, a voltage in a ratio of 1 : 3 needs to be selected. There are two DC sources V_1 and V_2 and overall four DC sources. These DC sources can be replaced by photovoltaic panels such that $V_1 = 25$ V and $V_2 = 75$ V. It means that V_2 is three times V_1. Results are obtained for a grid voltage of 200 volts at a frequency of 50 Hz. The switching frequency is 1 kHz. By putting the

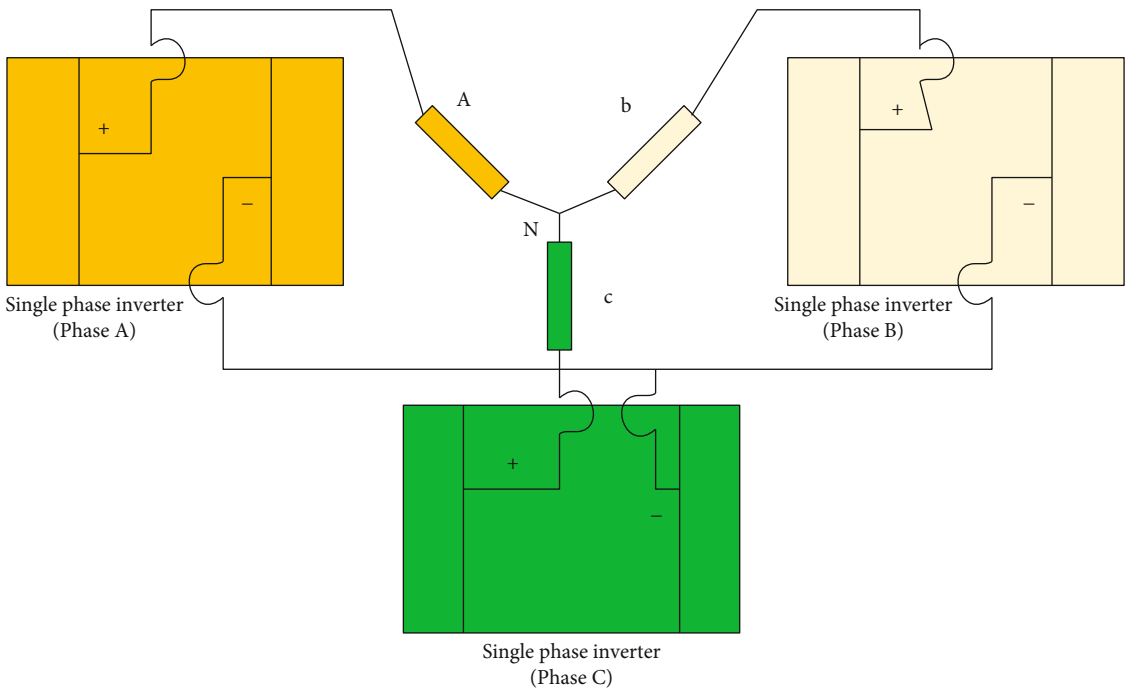


Figure 27: Three-phase inverter connection.

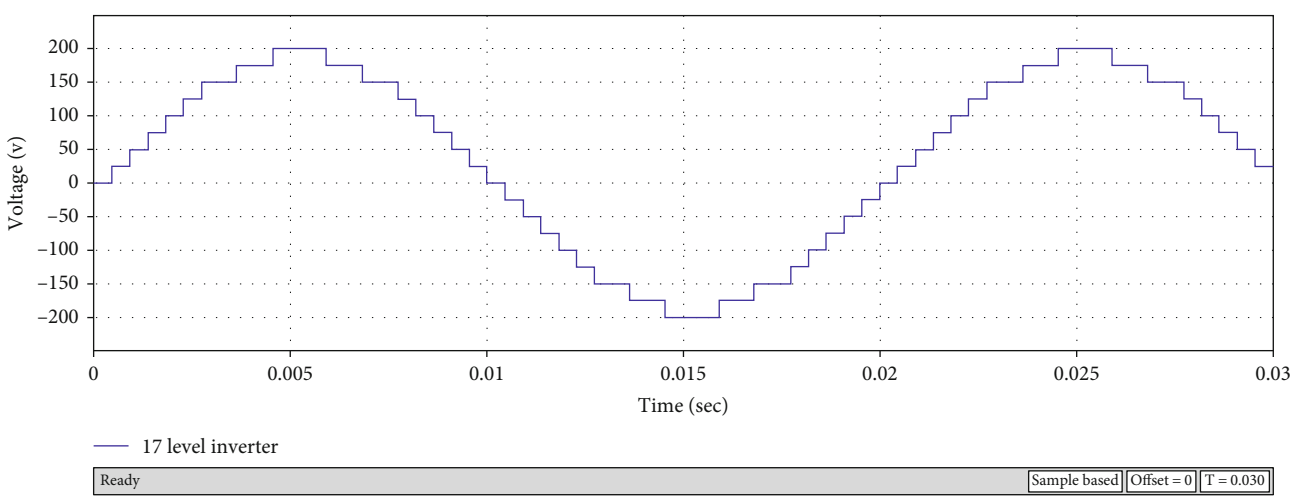


Figure 28: 17-level MLI with $t = 0.03$ for phase a.

estimations of V_1 and V_2 in Table 1, we get the accompanying results in Table 4.

6.2. Simulation Results. The simulation setup is represented in Figure 24. Each switch is modeled by an IGBT in the simulation. The simulated waveforms are in the accompanying figures. The simulation results confirm the expected results as the THD is brought down to 3.52%.

6.2.1. Single-Phase MLI Simulations. The resulting voltage waveform of single-phase MLI is depicted in Figures 25 and 26; it is compared with a clean sinusoidal waveform of the same frequency to get a good visualization of the reduction in THD.

6.2.2. Three-Phase MLI Simulations. In three-phase MLI, the three phases, namely, phase "a," phase "b," and phase "c" are depicted in Figure 27. For phase "b," a 120-degree phase shift is needed w.r.t phase "a." Also for phase "c," another 120-degree phase move w.r.t phase "a" or 120-degree phase shift w.r.t phase "b" is needed. Since there are 48 switching patterns, these need to be accommodated in 360 degrees cycle, so

$$\text{Single Interval Gap} = \frac{360}{48} = 7.5 \text{ degrees}. \tag{23}$$

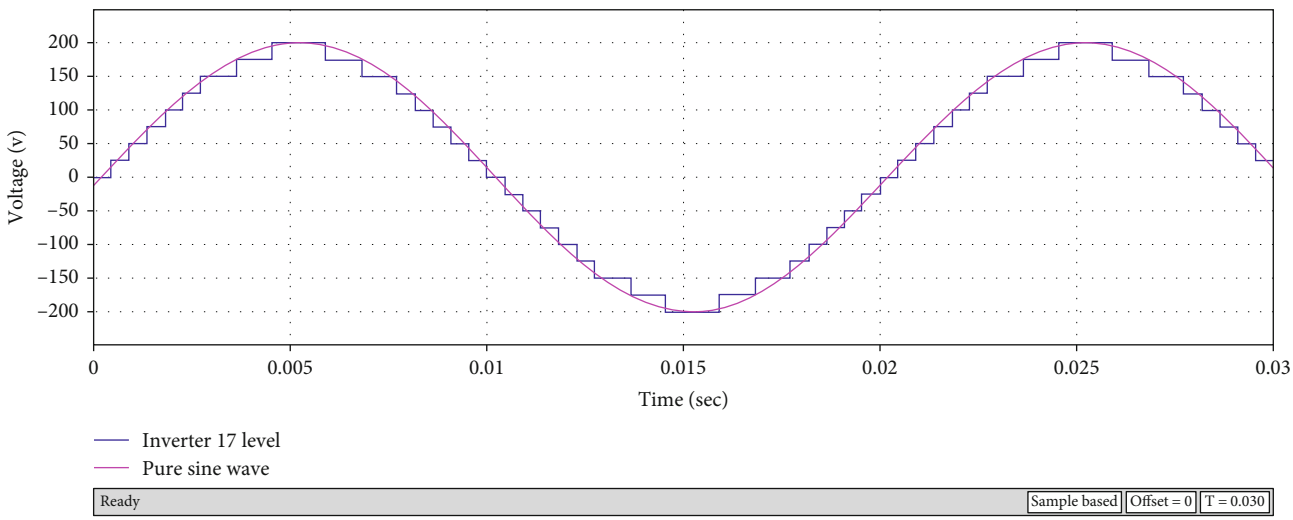


Figure 29: 17-level MLI with $t = 0.03$ for phase a and compared with sine wave.

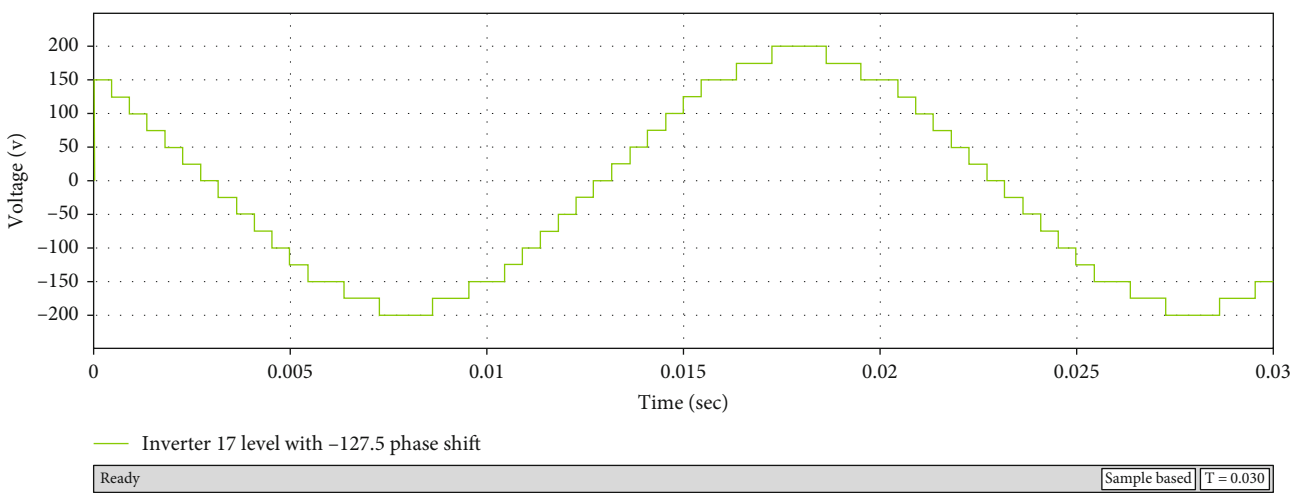


Figure 30: 17-level MLI with $t = 0.03$ for phase b.

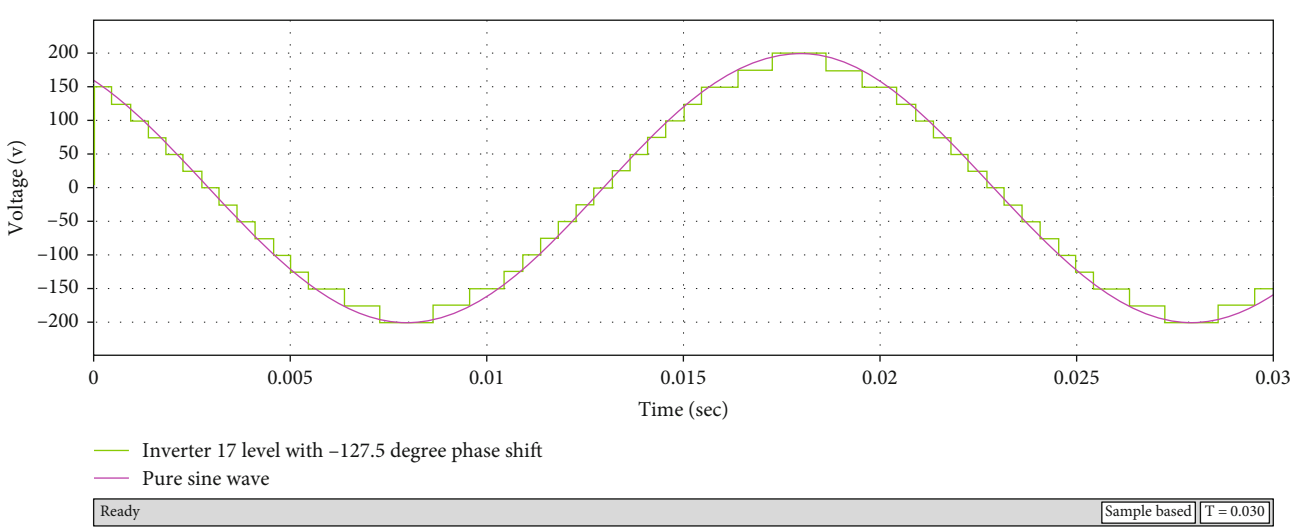


Figure 31: 17-level MLI with $t = 0.03$ for phase b and compared with sine wave.

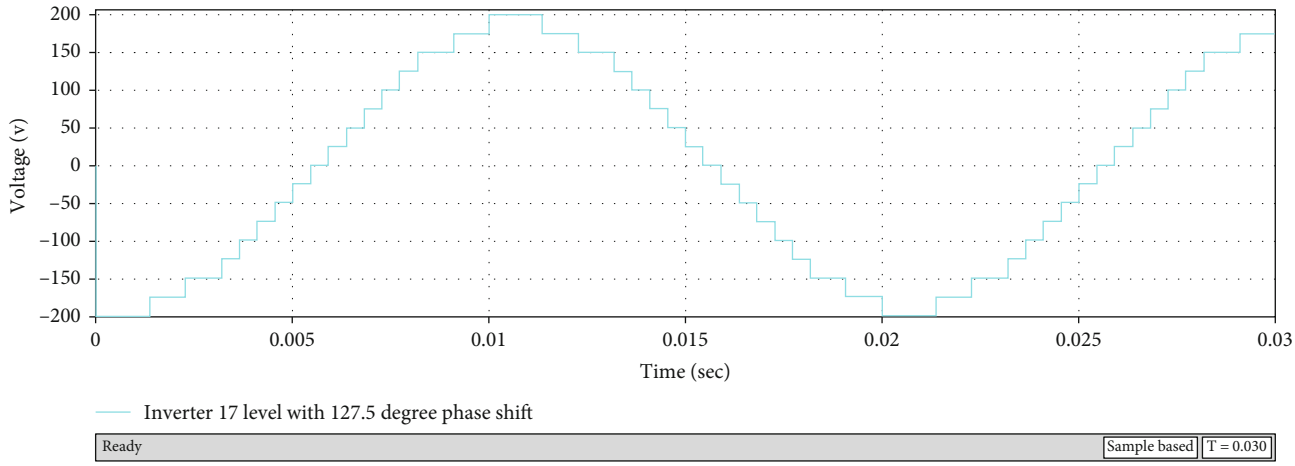


Figure 32: 17-level MLI with $t = 0.03$ for phase c.

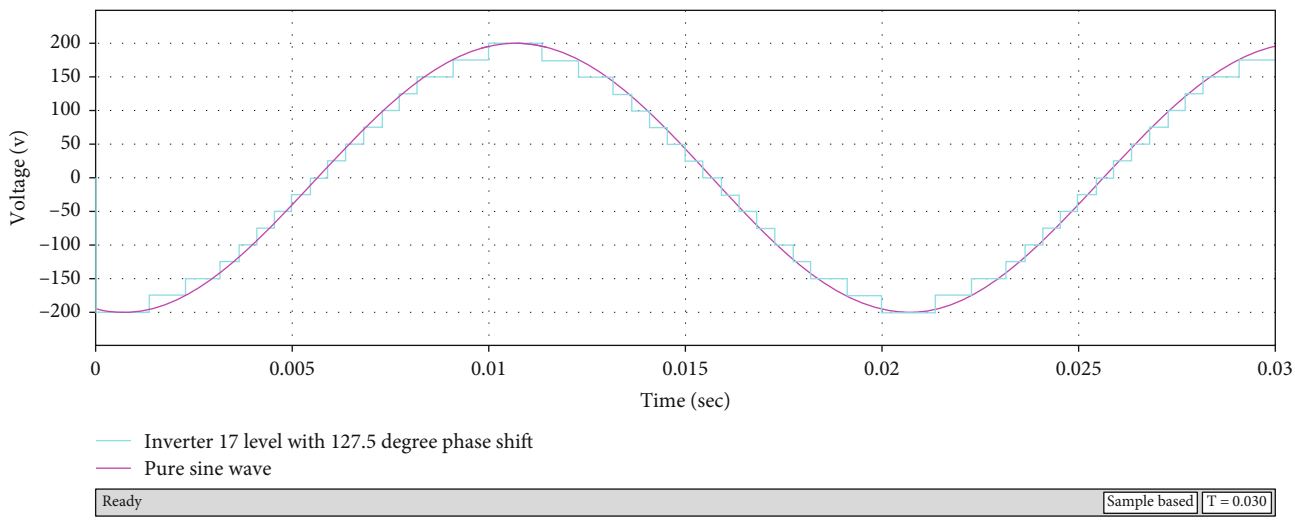


Figure 33: 17-level MLI with $t = 0.03$ for phase c and compared with sine wave.

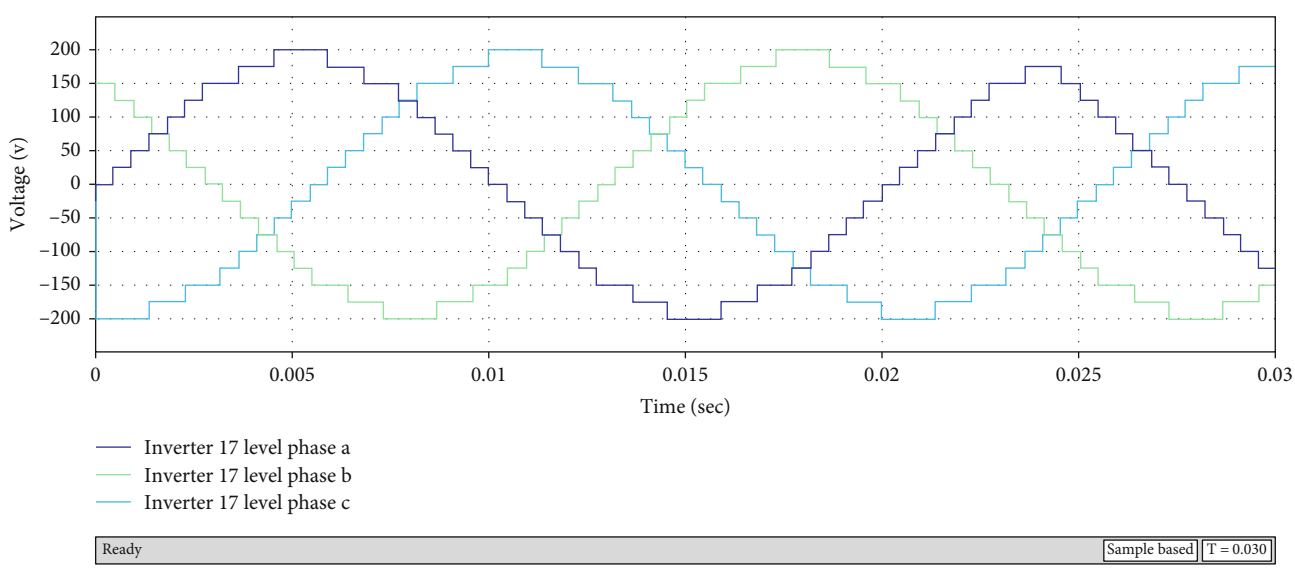


Figure 34: 3-phase 17-level MLI with $t = 0.03$.

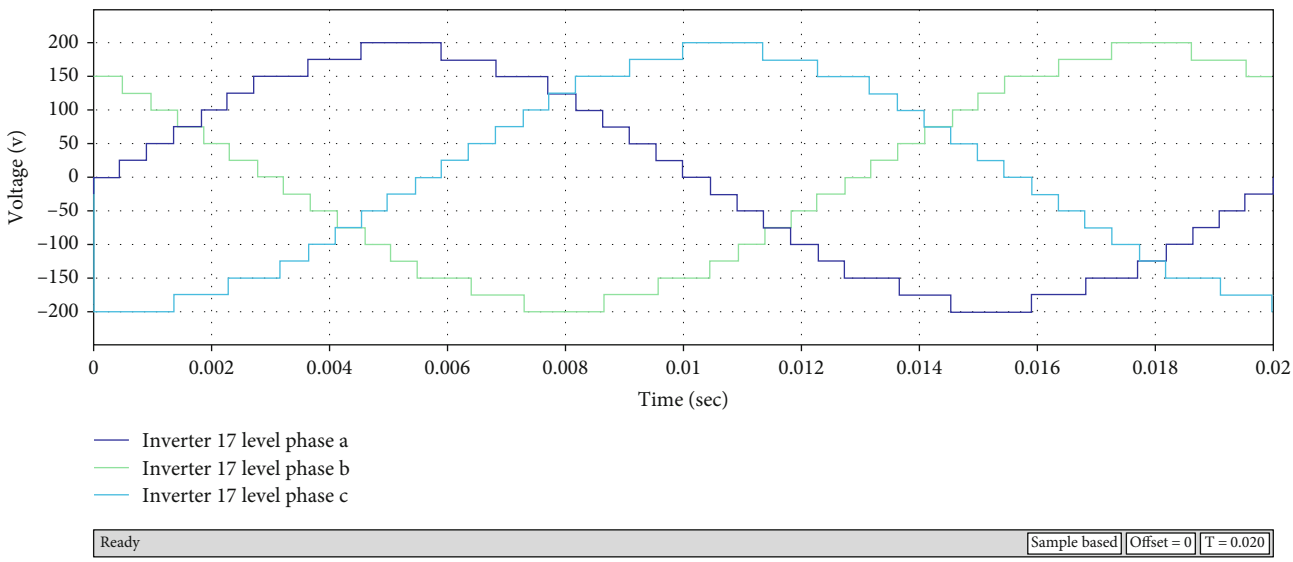


Figure 35: 3-phase MLI output with $t = 0.02$.

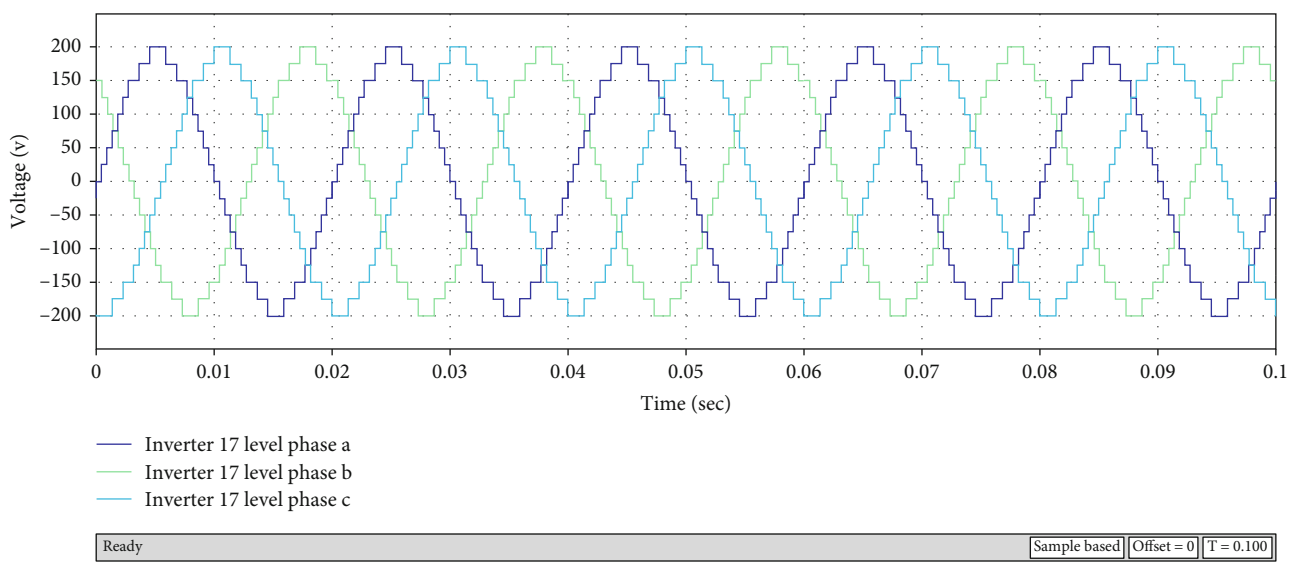


Figure 36: 3-phase MLI output with $t = 0.1$.

For 120-degree phase move,

$$\text{Bits shifts} = 120/7.5 = 16. \tag{24}$$

Therefore, it should be noted that phase "a" will begin at the first switching pattern and phase "b" at the 17th switching pattern and lastly phase "c" will begin at the 33rd switching pattern and then the cycle repeats itself.

(1)Phase a Results. Figures 28 and 29 show simulation waveforms of phase a for a time duration of 30 milliseconds.

(2)Phase b Results. Figures 30 and 31 show simulation waveforms of phase b for a time duration 30 of milliseconds.

(3)Phase c Results. Figures 32 and 33 show simulation waveforms of phase c for a time duration 30 of milliseconds.

(4)3-Phase Results. Figures 34–37 show simulation waveforms for all three phases of proposed MLI.

6.3. Fast Fourier Transform Analysis for THD. FFT investigation is done in Simulink/MATLAB to locate the total harmonic distortion on the frequency spectrum (see Figures 38 and 39). Total harmonic distortion is a helpful procedure to break down any nonlinear conduct of a framework, which is normally done with the help of fast Fourier transform (FFT). The measured signal is changed from the time domain into the frequency domain (see Figure 39). The changed information can be shown in an FFT spectrum in which the response signal's magnitude is plotted versus the frequency. Figure 39 shows the FFT spectrum of a 17-level MLI.

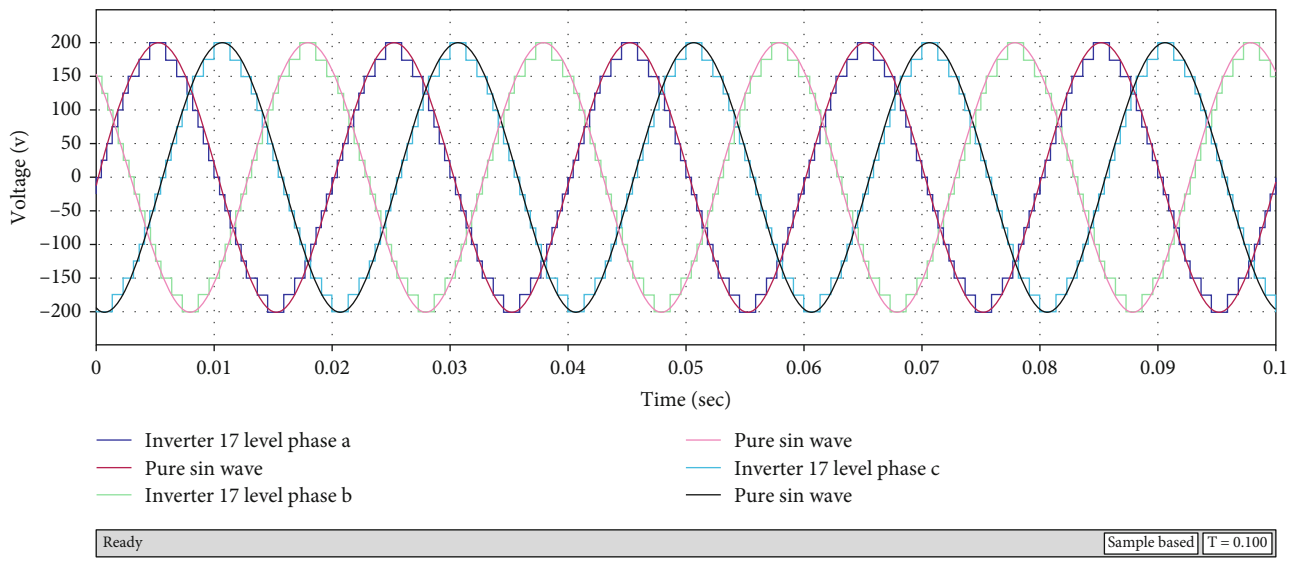


Figure 37: 3-phase MLI output with $t = 0.1$ and compared with sine wave.

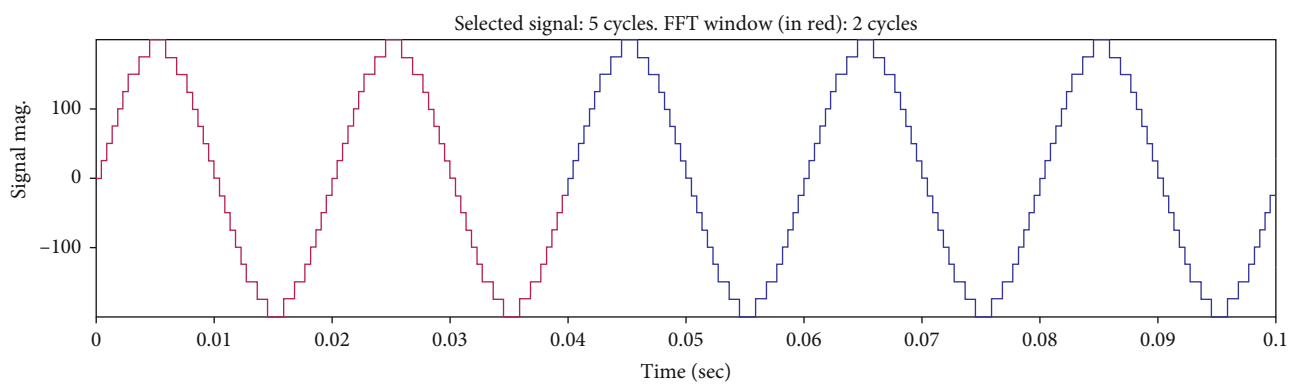


Figure 38: Input MLI signal.

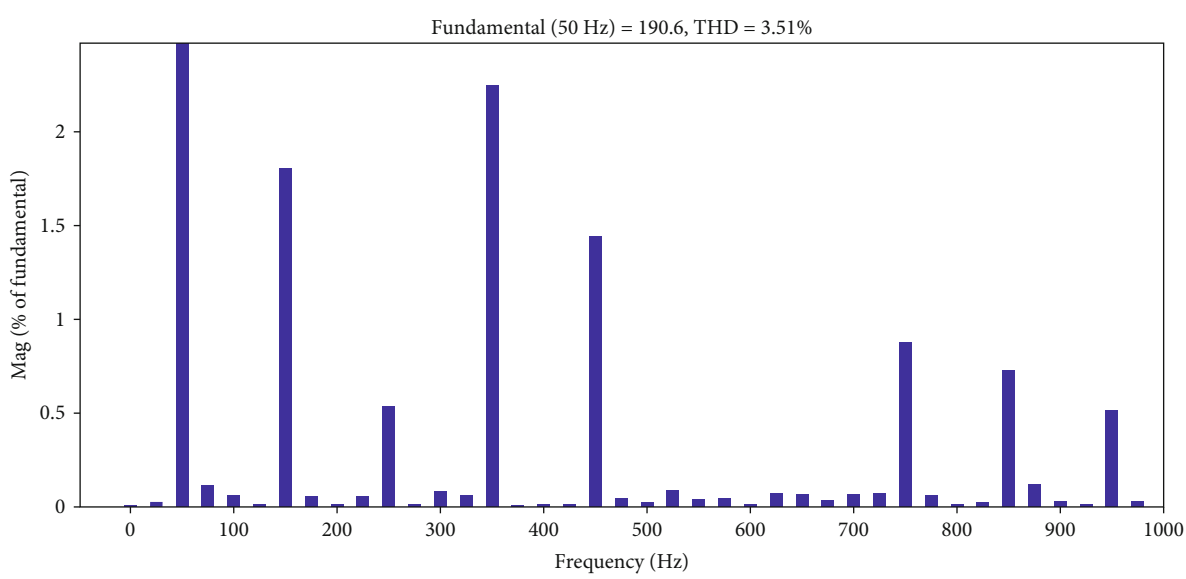


Figure 39: FFT spectrum of a 17-level MLI.

Table 5: Harmonic vs. amplitude.

Harmonic no.	Amplitude (V)
1	100
3	1.7
5	0.5
7	2.3
9	1.4
11	0.1
13	0.01
15	0.7
17	0.7
19	0.5

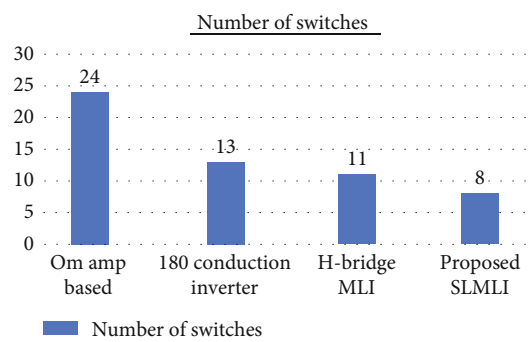


Figure 40: Comparison between numbers of switches.

We record the harmonics up till the 19th harmonics, and we ignore the values after that because those values were too low and almost near to zero. The amplitude of each harmonic is shown in Table 5.

We plugged the values in the THD equation to calculate THD as follows:

$$\begin{aligned}
\mathrm{THD} &= \frac{\sqrt{0.1^2 + 1.7^2 + 0.5^2 + 2.3^2 + 1.4^2 + 0.1^2 + 0.01^2 + 0.7^2 + 0.7^2 + 0.5^2}}{100}, \\
\mathrm{THD} &= \frac{\sqrt{0.01 + 2.89 + 0.25 + 5.29 + 1.96 + 0.01 + 0.0001 + 0.49 + 0.49 + 0.25}}{100}, \\
\mathrm{THD} &= \frac{\sqrt{11.6401}}{100}, \\
\mathrm{THD} &= \frac{3.522 * 100}{100}, \\
\mathrm{THD} &= 3.52\%.
\end{aligned} \tag{25}$$

6.4. Comparison with Other Topologies. A comparison between the number of switches (see Figure 40), number of voltage levels (see Figure 41), and THD (see Figure 42) was conducted between different multilevel inverters and the proposed multilevel inverter. The proposed 17-level MLI uses the SPWM technique for generating the gate signals/pulses. This inverter uses less number of switches, i.e., 8, and generates a maximum number of voltage levels. It also shows a low THD of 3.52%. The THD, where operational amplifier (Op-Amp) circuits were used for pulse generation,

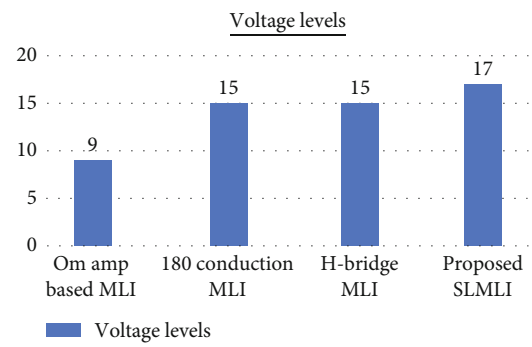


Figure 41: Comparison between numbers of voltage levels.

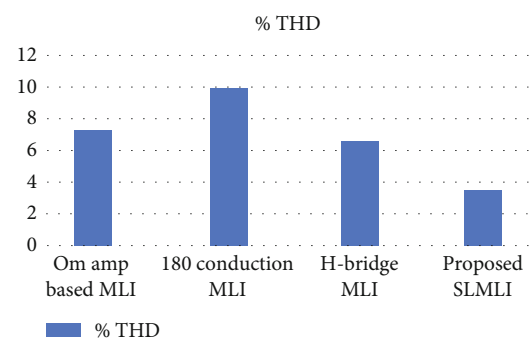


Figure 42: Comparison of %THD.

turns out to be 7.3% whereas 180° conduction MLI had 9.93% THD and H-Bridge MLI had 6.63% THD. Another multilevel inverter that uses the space vector modulation technique for generating pulses using switch ladder topology offers 6.1% THD.

From the above discussions, it is clear that the proposed MLI is superior to conventional MLI's in context of THD, number of switching devices, and voltage levels. The main limitation of proposed MLI is the bad regulation of output voltage and complexity in control of active switches.

7. Conclusions

In this paper, a modified form of a multilevel inverter has been successfully designed for high-power applications. The proposed topology has fewer switches due to which the cost is less. The proposed MLI solution does not require any filter at the output and produces less amount of THD. A thorough inquiry and scrutiny are carried out, and overall switching sequences have been evaluated for the proposed MLI. The THD of the system is 3.5%, and also, there is low voltage stress across the switches and the output is a smooth sinusoidal AC. By setting tailored quality metrics, a theoretical comparison is carried out for the proposed MLI and the conventional MLI's. The results reflect that the proposed MLI has surpassed the conventional MLI in terms of performance, reduced components, and low total harmonic distortion. The theoretical results have been verified by simulation results in MATLAB/Simulink. The results of the simulations align with the theoretical implications. Due to low

THD, the proposed MLI is a well-anticipated candidate for high-voltage/high-power applications. Considering the literature study done upon the MLI, it was seen that the main issue was concerned with the increased components as well as high THD. All of those parameters were improved by using the proposed topology. To reduce more components like DC sources, two capacitors can be used with only one DC source instead of two in the proposed MLI.

Data Availability

All relevant data and its supporting information files are included within the manuscript.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] M. Z. Malik, A. Ali, G. S. Kaloi, A. M. Soomro, M. H. Baloch, and S. T. Chauhdary, "Integration of renewable energy project: a technical proposal for rural electrification to local communities," *IEEE Access*, vol. 8, pp. 91448–91467, 2020.

[2] F. Akar, Y. Tavlasoglu, E. Ugur, B. Vural, and I. Aksoy, "A bidirectional nonisolated multi-input DC–DC converter for hybrid energy storage systems in electric vehicles," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 10, pp. 7944–7955, 2015.

[3] I.-M. Pop-Calimanu, S. Lica, S. Popescu, D. Lascu, I. Lie, and R. Mirsu, "A new hybrid inductor-based boost DC-DC converter suitable for applications in photovoltaic systems," *Energies*, vol. 12, no. 2, p. 252, 2019.

[4] M. Z. Malik, A. Ali, and D. Kumar, "A two cascaded boost converter with high voltage gain module," *International Journal of Computer and Electrical Engineering*, vol. 9, no. 2, pp. 476–483, 2017.

[5] H. Ardi, A. Ajami, F. Kardan, and S. N. Avilagh, "Analysis and implementation of a nonisolated bidirectional DC–DC converter with high voltage gain," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 8, pp. 4878–4888, 2016.

[6] M. Z. Malik, Q. Xu, A. Farooq, and G. Chen, "A new modified quadratic boost converter with high voltage gain," *IEICE Electron. Express*, 2016.

[7] A. Farooq, Z. Malik, Z. Sun, and G. Chen, "A review of non-isolated high step-down Dc-Dc Converters," *International Journal of Smart Home*, vol. 9, no. 8, pp. 133–150, 2015.

[8] M. Z. Malik, A. Ali, Q. Xu, and G. Chen, "A new quadratic boost converter with voltage multiplier cell: an analysis and assessment," *International Journal of Smart Home*, vol. 10, no. 8, pp. 281–294, 2016.

[9] H. Wu, X. Zhan, and Y. Xing, "Interleaved LLC resonant converter with hybrid rectifier and variable-frequency plus phase-shift control for wide output voltage range applications," *IEEE Transactions on Power Electronics*, vol. 32, no. 6, pp. 4246–4257, 2016.

[10] D. Thenathayalan, C. Lee, and J.-H. Park, "High-order resonant converter topology with extremely low-coupling contactless transformers," *IEEE Transactions on Power Electronics*, vol. 31, no. 3, pp. 2347–2361, 2015.

[11] S.-W. Lee and H.-L. Do, "High step-up coupled-inductor cascade boost DC–DC converter with lossless passive snubber," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 10, pp. 7753–7761, 2018.

[12] A. Farooq, Z. Malik, D. Qu, Z. Sun, and G. Chen, "A three-phase interleaved floating output boost converter," *Advances in Materials Science and Engineering*, vol. 2015, Article ID 409674, 8 pages, 2015.

[13] B. Ahmad, J. Kyyra, and W. Martinez, "Efficiency optimisation of an interleaved high step-up converter," *Journal of Engineering*, vol. 2019, no. 17, pp. 4167–4172, 2019.

[14] S. B. Mahajan, P. Sanjeevikumar, P. Wheeler, F. Blaabjerg, M. Rivera, and R. Kulkarni, "X-Y converter family: a new breed of buck boost converter for high step-up renewable energy applications," in *2016 IEEE International Conference on Automatica (ICA-ACCA)*, pp. 1–8, Curico, Chile, 2016.

[15] O. Cornea, G.-D. Andreescu, N. Muntean, and D. Hulea, "Bidirectional power flow control in a DC microgrid through a switched-capacitor cell hybrid DC–DC converter," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 4, pp. 3012–3022, 2016.

[16] A. Tomaszuk and A. Krupa, "Step-up DC/DC converters for photovoltaic applications–theory and performance," *Electrical Review*, vol. 89, pp. 51–57, 2013.

[17] M. Z. Malik, H. Chen, M. S. Nazir et al., "A new efficient step-up boost converter with CLD cell for electric vehicle and new energy systems," *Energies*, vol. 13, no. 7, p. 1791, 2020.

[18] M. D. Siddique, S. Mekhilef, N. M. Shah, and M. A. Memon, "Optimal design of a new cascaded multilevel inverter topology with reduced switch count," *IEEE Access*, vol. 7, pp. 24498–24510, 2019.

[19] M. D. Siddique, S. Mekhilef, N. M. Shah, J. S. M. Ali, and F. Blaabjerg, "A new switched capacitor 7L inverter with triple voltage gain and low voltage stress," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 67, no. 7, pp. 1294–1298, 2020.

[20] M. D. Siddique, J. S. M. Ali, S. Mekhilef, A. Mustafa, N. Sandeep, and D. Almakhles, "Reduced switch count based single source 7L boost inverter topology," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 67, no. 12, pp. 3252–3256, 2020.

Hindawi
International Journal of Photoenergy
Volume 2021, Article ID 6777488, 19 pages
https://doi.org/10.1155/2021/6777488

Research Article

Forecasting of Energy Production for Photovoltaic Systems Based on ARIMA and ANN Advanced Models

Laurentiu Fara,[1,2] **Alexandru Diaconu**,[1] **Dan Craciunescu**,[1] **and Silvian Fara**[1]

[1]*Department of Physics, Faculty of Applied Sciences, Polytechnic University of Bucharest, 060042, Romania*
[2]*Academy of Romanian Scientists, 030167 Bucharest, Romania*

Correspondence should be addressed to Laurentiu Fara; lfara@renerg.pub.ro

Received 21 April 2021; Revised 2 June 2021; Accepted 15 July 2021; Published 9 August 2021

Academic Editor: Alberto Álvarez-Gallegos

Accurate forecasting of solar energy is essential for photovoltaic (PV) plants, to facilitate their participation in the energy market and for efficient resource planning. This article is dedicated to two forecasting models: (1) ARIMA (Autoregressive Integrated Moving Average) statistical approach to time series forecasting, using measured historical data, and (2) ANN (Artificial Neural Network) using machine learning techniques. The main contributions of the authors could be synthetized as follows: (1) analysis and discussion of the experimental and simulated results regarding solar radiation forecast, as well as energy production prediction and forecasting based on ARIMA and ANN models for two case studies: (a) laboratory BIPV system developed at the Polytechnic University of Bucharest and (b) large PV park placed in a specific site of the south of Romania. A variability index of solar radiation was introduced for the model improvement; (2) comparison between the ARIMA and ANN results to highlight the ARIMA model which is more efficient than the ANN one; (3) optimized method defined by the GMDH model (Group Method of Data Handling) proposed to provide a software program for calculation of the PV energy production.

1. Introduction

Solar energy is one of the most promising sources of power generation for residential, commercial, and industrial applications [1–5]. Photovoltaic solar energy, based on PV systems, has increased progressively in recent years due to its advantages of being abundant, inexhaustible, clean, and environmentally friendly [6–8].

Reliable and accurate forecasts play a key role in improving PV solar power plants [1–3, 9–12]. The main challenging problem in the production of solar energy is the intermittent electricity generation using PV systems, due to weather conditions. A variation in temperature and solar irradiance can have a profound impact on the quality of electricity production, leading to a decrease of more than 20% in the PV energy production provided by real PV installations. This limits the integration of PV systems into the grid. Therefore, an accurate short-term forecast of photovoltaic energy is very useful for the efficient daily/hourly management of electricity production and storage in the grid [13].

Accurate forecasting of solar energy is essential for PV plants, in order to facilitate their participation in the energy market and for efficient resource planning [1–3]. Various methods have been reported in the literature for the forecast of PV energy [2, 12]. These methods can be divided into four classes: (i) statistical approaches to time series forecasting, using measured historical data (ARIMA) [14]; (ii) machine learning techniques, in particular Artificial Neural Networks (ANN), based on machine learning approaches [15]; (iii) physical models based on numerical weather prediction and satellite imagery [16]; and (iv) hybrid approaches that represent a combination of the first three methods [17].

The beauty of the ARIMA model lies in its simplicity and can be applied only to stationary time series [14, 18, 19]. Therefore, our data from the time series, which are seasonal and nonstationary, are transformed into a stationary series for the application of the ARIMA model. The model can be developed using sophisticated statistical techniques [20]. The optimal approach is selected and validated using the Akaike Information Criterion (AIC) and the residual sum of squares (SSE).

Another statistical model is based on the seasonal analysis of the ARIMA time series (SARIMA) and is further improved by incorporating short-term solar radiation predictions derived from the NWP (numerical weather prediction) model [19]. Such a model can be used in energy forecast simulation software for buildings, such as Energy Plus, TRNSYS, or eQuest, both for existing buildings and in the design of new buildings [20, 21].

On the other hand, data-based models are determined by mathematical models and measurements. They do not require detailed knowledge of buildings or equipment. Their forecasts are mainly based on historical data, which are available using the control systems implemented in buildings for (1) building automation systems (BAS) and (2) building energy management systems (BEMS). The accuracy of these models depends on the quality of the selected forecast model, as well as the quality and quantity of available data. Such models are easy to adapt to changing conditions and are relatively easy to be used. In most cases, the relationship between the predicted variable and its physical function is not clearly established [17, 22].

In recent decades, researchers have been dedicated to improving efficiency and building energy use through various techniques and strategies. Energy forecasting in an existing building is essential for a variety of applications, such as demand response, fault detection and diagnosis, predictive model control, energy optimization, and management. Energy estimation models are a promising field of research, and this is true given the new advances in artificial intelligence and machine learning. Such models have been widely applied to both building energy systems and HVAC (heating, ventilation, and air conditioning) systems, as they can help with a variety of tasks. Most conventional approaches for solar energy forecasting are limited to discover the data correlation but are not able to analyze them in depth and discover relevant information. With many data in the modern power system, the use of conventional approaches is not appropriate to guarantee an accurate forecast [15, 20, 23].

Recently, Deep Learning (DL) approaches have emerged as powerful machine learning tools that allow shape recognition and regression analysis, as well as prediction applications [24–26]. DL approaches are becoming increasingly popular due to their real ability to describe time series data dependencies. Recently, many Deep Learning models have been proposed, including Boltzmann machines, Deep Belief Networks (DBNs), and Recurrent Neural Networks (RNNs). RNN is a type of neural network that exploits the sequential nature of input data. RNNs are used to model time-dependent data and give good results, which have proven to be successful in several application fields [15, 23, 27]. Short-Term Memory Networks (STMN) are a type of RNNs, which are able to cope with the storage of information for much longer periods of time [16, 28–30]. They are considered to define one of the most used RNN models for predicting data from time series, which perfectly correspond to the problems of PV energy production.

Zhao and Magoulés published a review study in 2012, focusing on the main approaches to energy prediction and forecasting in buildings [20]. Specifically, the authors compared physical models with those based on machine learning, as well as with statistical ones. The authors noted that machine learning-based models had the highest accuracy and flexibility, especially compared to statistical models. Vector support machines have been developed to outperform ANN models; one of the areas recommended for future investigations is that of optimized applications.

In 2017, Wang and Srinivasan [23] explored the use of models based on artificial intelligence (AI) and overall models for predicting and forecasting the energy use of buildings. The authors provided a breakdown of how artificial intelligence as a whole was applied to predict the energy of the buildings. Most AI-based works were applied to an entire building load with hourly data. The authors also explored how overall methods were applied in predicting the energy of buildings. It was noted that such assemblies have been widely applied to fields outside the energy of buildings, and the results showed improved performance compared to unique forecast models.

Akhter et al. [31], Van Deventer et al. [32], Das et al. [33], and Seyedmahmoudian et al. [34] analyzed and compared different forecasting models for PV energy production, including the hybrid ones; their strengths and weaknesses were highlighted.

At the same time, the benefits of model optimization were discussed.

The main objective of this study was to apply the advanced techniques of ARIMA and ANN models for two cases: (1) a lab BIPV system and (2) a large PV park placed in the south of Romania. The obtained simulated and experimental results were compared and allowed to have interesting information regarding the energy production forecast on short term, as well as on long term; it could be very useful for improved performance of energy prediction and forecasting for buildings based on statistical approaches and artificial intelligence tools.

2. Conceptualization and Methodology

2.1. The Advanced Statistical ARIMA Model. The ARIMA (Autoregressive Integrated Moving Average) model establishes a remarkable method used in short-term solar radiation forecasting [14]. The functionality of the ARIMA model is represented in Figure 1 [35, 36]. The software used to make the forecast, respectively, IBM SPSS, is a predictive analysis program, which offers a wide range of statistical procedures: linear regressions, Monte Carlo simulations, geospatial analysis, etc.

The model chosen in this article for the solar radiation forecast is of the ARIMA (1,0,0) or (1,1,0) type. This choice is due to the Akaike Informational Criterion (AIC) [19]. The AIC is used to evaluate the model, depending on the entered data; the aim is to determine the quality of the statistical model on a dataset, establishing whether this model is more efficient than other models. The ARIMA model provides a relative estimate of the information used to determine the processes that could generate the data.

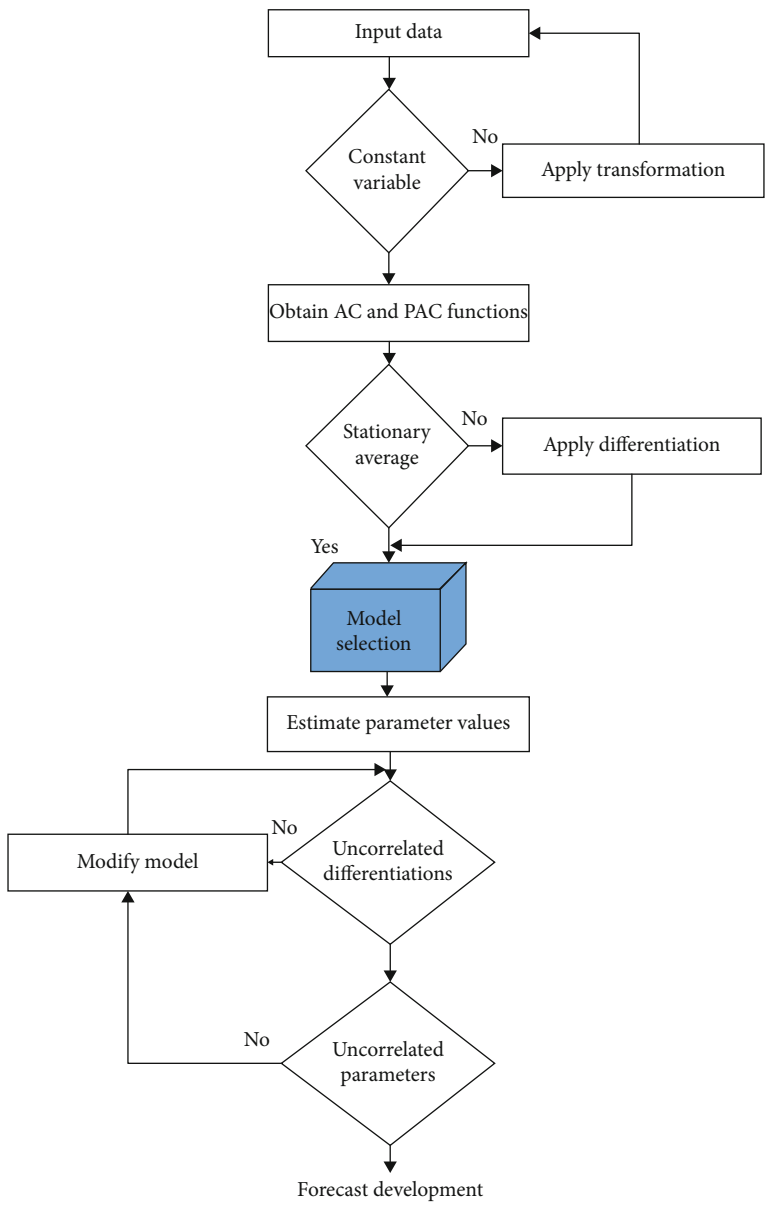


Figure 1: Operation of the ARIMA model for the energy production forecast of a photovoltaic system; AC and PAC autocorrelation functions are used.

The Akaike statistical criterion for the ARIMA model can be calculated based on the following relation [35, 36]:

$$\mathrm{AIC} = 2k - 2\ln(L), \tag{1}$$

where L represents the maximum value given by an estimation function (MLE (Maximum Likelihood Estimation)) and k represents the number of estimated parameters. The minimum value of the AIC result defines that the corresponding model is the most efficient.

The SPSS program [35, 36] automatically calculates the AIC, which is a parameter for model fitting.

If the forecast process contains seasonal fluctuations, as in this case, the process becomes SARIMA $(p, d, q)(P, D, Q)s$, where p is the order of the AR process, d represents the differentiation term, q is the order of the moving average, P represents the order of AR seasonal processes, Q represents the MA order, D is the order of seasonal differentiation, and s represents the length of the seasonal period [19].

The peculiarities of the model used in this forecast are highlighted by defining a backshift operator B of the time series [9]: $B_{Zt} = z_{t-1}$, where z_t and z_{t-1} are two consecutive observations of the time series. Then, B^j can be defined as $B^j{}_{Zt} = z_{t-j}$. Using standard notation, $\phi(B)$ is the autoregression operator, represented as a polynomial in the backshift operator: $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$, and $\theta(B)$ the mobile operator is represented by a polynomial in the operator backshift: $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 \cdots - \theta_q B^q$. Having d—the

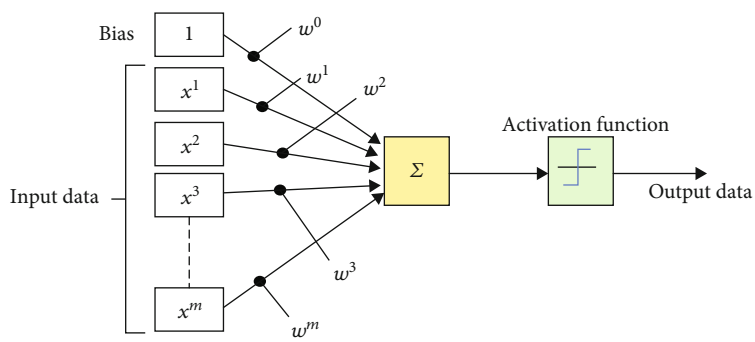


Figure 2: Functionality of the ANN model for the energy production forecast of a photovoltaic system.

degree of nonseasonal differentiation and at the white noise, from a mathematical point of view, the ARIMA model could be written as follows:

$$\phi(B)\,W^d z_t = \theta\,(B)\,a_t, \tag{2}$$

where W is the differentiation operator which is the equivalent of result 1-B.

We call this model ARIMA (p, d, q), where p, d, and q are process orders. In our model, the self-regression delay is determined separately for each month represented by the number of lighting hours, and the hours with zero irradiation (night hours) are neglected. As a consequence, we have used the ARIMA model $(x, 0, 0)$, where x represents the number of hours in a day with solar radiation of those months (for example, for January it is 10, and for July it is 16).

ARIMA $(x, 0.0)$ was used to avoid the SARIMA model (seasonal model) [19, 35, 36]. The seasonality component is taken into account by the autoregression component, where the autoregression is equal to the period. 10 in January represents the fact that the morning value from 8 o'clock correlates with the value of the next day from 8 o'clock, which is in the 10th place from the first value, since we have used only 10 values each day. In other words, in the used experimental data string, there is an autocorrelation of 10 orders.

2.2. The Advanced ANN Model Based on Artificial Intelligence (AI). Neural networks are a set of processing elements that have been developed separately from standard regression techniques. The ability to "mimic" natural intelligence by learning from experience makes this technique very attractive for solar radiation predictions. Functioning like a biological brain, a neural network is made up of a large number of interconnected neurons. Two main classes of neural network architecture can be identified, namely, (1) *the architecture with the propagation of information from the input data to the output data*, also called *feed-forward architecture*, and (2) *the architecture of recurrent networks*. The functionality of the neural network model is represented in Figure 2 (see [15, 25]).

The ANN model is also used to estimate global solar radiation at an hourly interval, based on the following meteorological parameters: air temperature (T), relative humidity (RH), atmospheric pressure (P), wind speed (W), cloudiness (C), and hours of the day (H):

$$R_G = f\,(T, \mathrm{RH}, P, W, C, H). \tag{3}$$

The main parameters of the ANN model used in this study are as follows [35, 36]: the number of neurons in the hidden layer is 1, the initial weight is 0.3, the learning rate is 0.3, and it used the Sigmund logarithmic function. The 10 days of meteorological data are included in the learning process.

3. The PV Energy Production Forecast for a Lab BIPV (Building Integrated Photovoltaic) System

3.1. Results and Discussion: Solar Radiation and Energy Production Forecast Based on the ARIMA and ANN Models for a Lab BIPV System

(a) The BIPV system from the Polytechnic University of Bucharest is a demo system connected to the grid; it was put into operation in July 2008 (see Figure 3) [11]. It consists of six PV panels, an inverter, equipment for monitoring and storing data, and a laboratory station for monitoring meteorological parameters. The power delivered by the PV panels is 615 Wp; they are semitransparent and have a power of 85 Wp each (type 1), respectively, and 120 Wp each (type 2). The system consists of three modules of type 1 and three modules of type 2

Also, to complete the BIPV system and make it functional, there is an inverter that makes the conversion between dc and ac, as well as monitoring and data storage equipment. The inverter is of the Sunny Boy SB700 type and achieves a high efficiency of conversion of direct current into alternating one, which it then delivers to the electricity distribution grid.

With the mentioned technical resources of the studied BIPV system, data regarding the brightness and performance of the PV system were collected, in terms of power (Pac) for a period of five days. Three days were considered for the forecasting process, and the other two days were used to validate the results. Figure 4 illustrates the

Figure 3: BIPV system developed within the laboratory of nanostructures, solar cells, and PV systems from the Department of Physics, Faculty of Applied Sciences, Polytechnic University of Bucharest (inside/outside view).

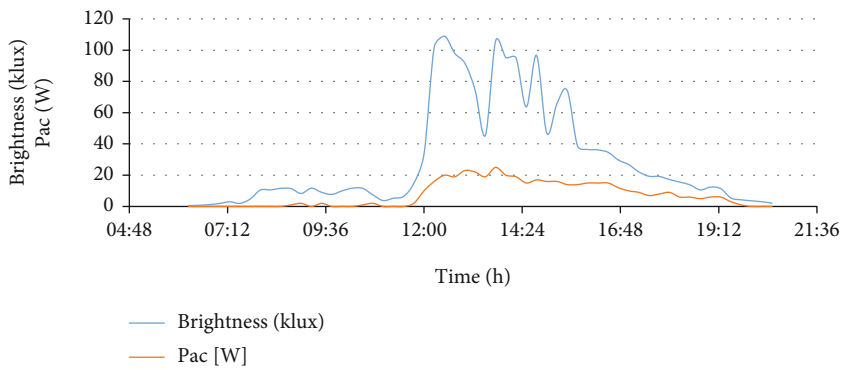


Figure 4: Time evolution of power (W) and brightness (klux) parameters.

evolution of the power and brightness parameters, over a period of one day of the chosen interval.

The detailed comparison, which provides data every 15 minutes, allows us to obtain important information about system performance. An unexpected decrease in power can be seen in the second half of the day. The explanations can be the following:

(i) The photovoltaic window is oriented to the east, so that the incident solar radiation is maximum in the first part of the day

(ii) The system is shaded in the second part of the day, due to the architecture of the building

The solar radiation data, obtained with the help of the lab weather station, were used for the short-term solar radiation forecast, using various methods, in the location where the BIPV system is placed [11].

The results of the ARIMA model are represented by the input dataset for 13.06-15.06 2012, and the model is checked by the data for 16.06-18.06 2012 [11, 35, 36] (see Table 1).

Table 1: The accuracy of the ARIMA model.

Statistical fitting	Average
R^2	0.811
RMSE	18,367.868
MAPE	199.410
MaxAPE	3,838.473
MAE	10,359.011
MaxAE	54,998.512

R^2 represents how much of the variation of the dependent variable is determined by the estimated quantity, RMSE (Root Mean Squared Error) represents the difference between the predicted values and the observed ones, MAPE (Mean Absolute Percentage Error) represents the accuracy of the forecasted model, and MaxAPE (Maximum Absolute Percentage Error) is the largest predicted error, expressed as a percentage. This error is useful to determine the efficiency of the forecast. MAE (Mean Absolute Error) represents the degree of variation of the measured data compared to the forecast. MaxAE (Maximum Absolute Error) is the largest error in the forecast, like MaxAPE, also useful in determining the efficiency of the forecast at the point with maximum error.

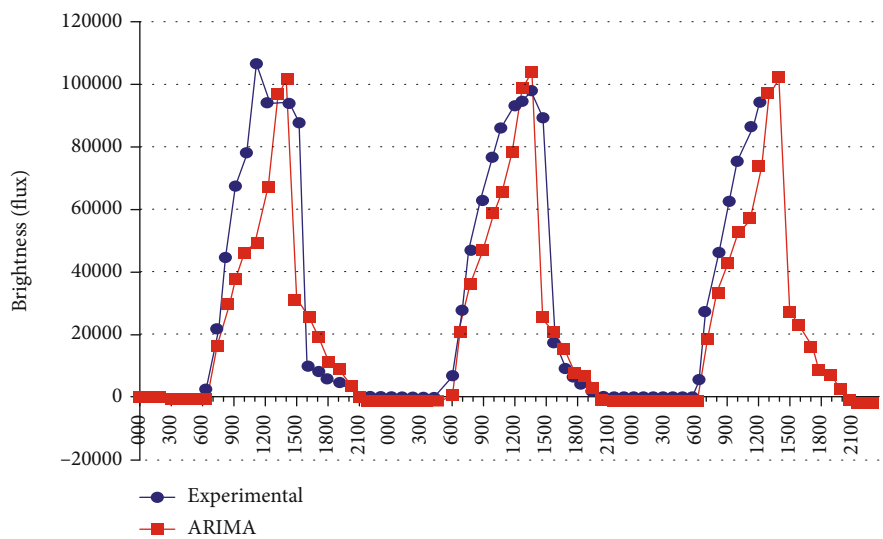


Figure 5: Adjusting the forecasted ARIMA data with the measured ones.

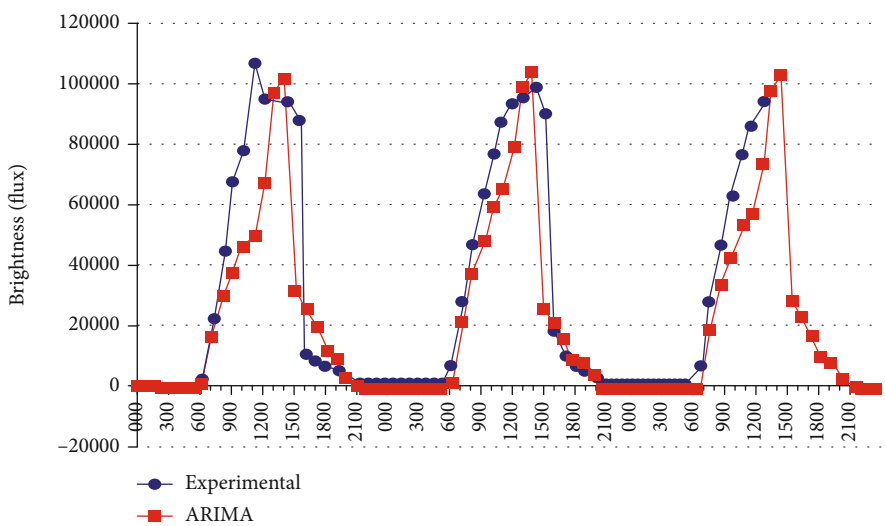


Figure 6: Adjusting the forecasted ANN data with the measured ones.

Figure 5 shows the adjustment of the ARIMA model between the forecasted and measured data.

(b) Taking into account five meteorological variables, namely, air temperature, relative humidity, atmospheric pressure, wind speed, and sunshine duration, the forecast was made for 16.06.2012, using the method of neural networks (ANN) (see Figure 6) [11, 35, 36]

A statistical relationship is defined between global solar radiation and energy production, taking into account the technical parameters of the system. The losses of the system due to its location were also taken into account.

The values of the predicted data for P_{ac} were compared with the measured ones; this situation is represented in Figure 7. The error calculated between the measured and the predicted values of P_{ac} using the ANN model is 8.89%.

3.2. Improving the Forecast Quality by the Variability Index of Solar Radiation. The meteorological data used in the study of the solar radiation forecast for a lab BIPV system come from the Archive of the National Meteorological Authority [9]. The parameters were measured in the Bucharest-Afumati weather station (44° 30′N 26° 13′E, $h = 91$ m, 8 km away from Bucharest, WMO code 15421) and represent data obtained every hour during 2008-2009. The meteorological tools and observations correspond to the WMO instructions (Global Observation System Manual, 2010). The meteorological variables included in this study are global solar irradiance in kJ/m^2, average air temperature in °C at a height of two meters, atmospheric pressure in mbar, relative humidity, wind speed in m/s, and cloudiness in oktas.

At the Bucharest-Afumatzi meteorological station, the global solar irradiance is automatically measured on a horizontal surface, with the pyranometer CM11 Kipp&Zonen-

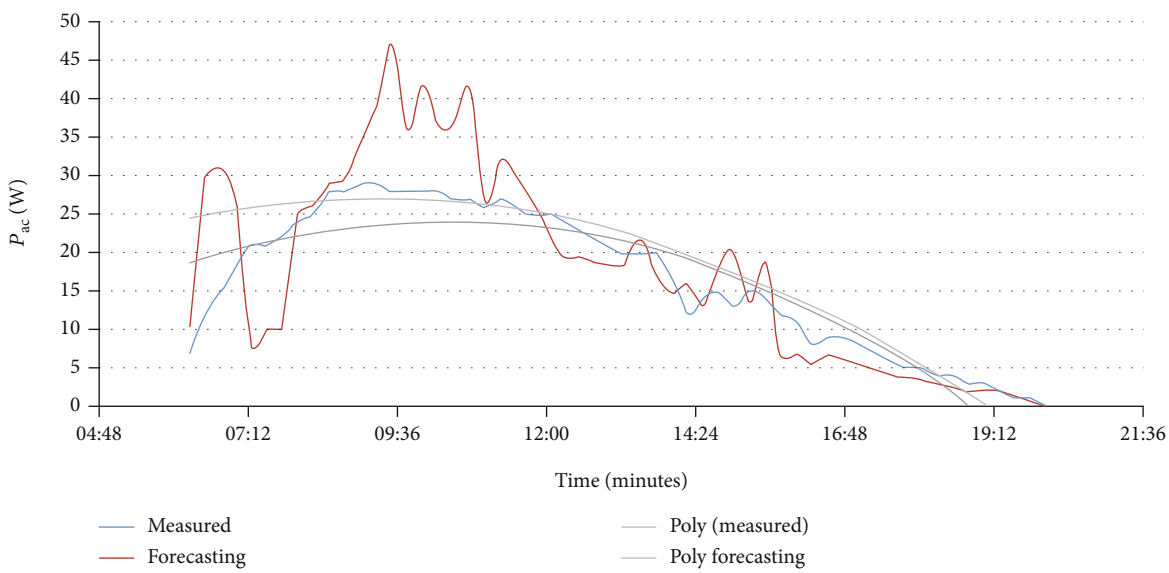


Figure 7: Comparison between measured and forecasted values of P_{ac} for 17.06.2012.

Delft. The pyranometer sensor is a thermopile protected from external effects by a double glass dome. It has the following features: a response time (95%) of 15 seconds, spectral range 305-2800 nm, sensitivity 4-6 μV/W/m^2, nonlinearity $< \pm 0.6\%$ (<1000 W/m^2), this being installed inside the meteorological platform, facing south, at a height of 2 meters.

In the solar radiation forecast of the BIPV lab system from the Polytechnic University of Bucharest, both studied models were used, namely, ARIMA and ANN [9, 11]. The error of the solar radiation forecast depends on its daily variation. At the same time, the daily variation of solar radiation showed an important correlation with clouding; thus, cloud information will be needed to increase the accuracy of the forecast.

However, this type of calculation requires a large amount of input data and a huge computational ability. In most cases, acceptable results cannot be obtained. In order to take into account the influence of the clouding factor but at the same time reduce the number of input data, it is found that the affected situations are determined by an empirical index, as an input parameter, namely [11],

$$V_{DR} = \frac{\sigma_{DR}}{\chi_{MR}}, \tag{4}$$

where V_{DR} is the variability index of the solar irradiance, χ_{MR} is the average of the monthly global solar irradiance, and σ_{DR} is the standard deviation of the daily global solar irradiance.

In fact, this index quantifies the variation of solar irradiance for a given day, compared to the general situation in a given month; the variation of solar irradiance in the previous days is taken into account indirectly by the monthly average. This daily index provides more detailed information, as well as the classification based on synoptic situations, the latter variant including a longer period with various clouding factors. By default, the index provides an important correlation with the clouding factor, $r = 0.66$ in 2008 and $r = 0.61$ in 2009, both being considered at a probability level of 9.99%. The average number of days with cloudiness, which denotes $V_{DR} < 0.5$, is 7.93 in 2008 and 7.69 in 2009, the standard deviation being 0.79 and 0.95, respectively. At the same time, the average number of days with cloudiness, showing a $V_{DR} > 0.5$, is 4.49 in 2008 and 4.57 in 2009, the standard deviation being 2.27 and 2.53, respectively. As a consequence, the value of V_{DR} indicates changes in global solar radiation related to the following situations:

(a) $V_{DR} < 0.5$: fog or cloud cover situations

(b) $V_{DR} > 0.5$: situations with partly cloudy or clear sky

In our study, four different situations are analyzed based on the cloudiness quantified by the V_{DR} index, and then, the forecasts of solar radiation values are developed independently of each other. The main reason for this classification is to reduce the synoptic situations represented by the two classes of cloud cover types.

The four situations include changes in the amounts of cloudiness on the previous day and on the forecast day and are as follows (see Figure 8) [9]:

(1) Cloudy/foggy day on the forecast day and partly cloudy/clear sky on the previous day $V_{DR} < 0.5$ (previous day $V_{DR} > 0.5$)

(2) Cloudy/foggy situation on the forecast day and cloudy/foggy sky situation on the previous day $V_{DR} < 0.5$ (previous day $V_{DR} < 0.5$)

(3) Partly cloudy/clear sky on the forecast day and cloudy/foggy sky on the previous day $V_{DR} > 0.5$ (previous day $V_{DR} < 0.5$)

(4) Partly cloudy/clear sky on the forecast day and partly cloudy/clear sky on the previous day $V_{DR} > 0.5$ (previous day $V_{DR} > 0.5$)

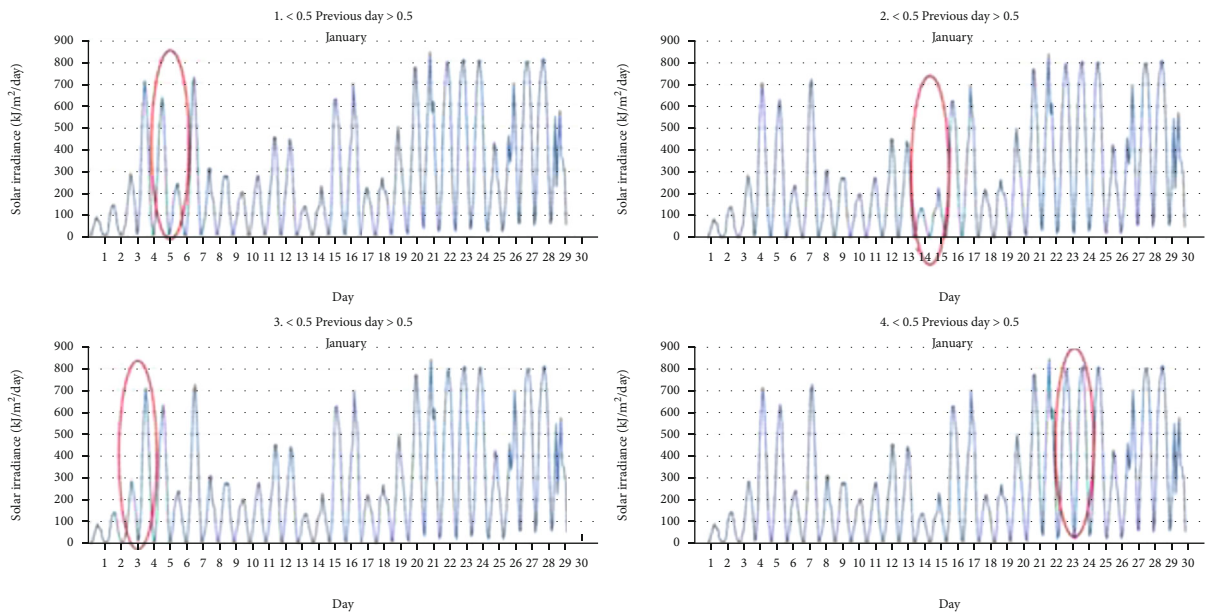


Figure 8: The four cloudy change situations quantified by the V_{DR} index, based on solar irradiance data measured at the Bucharest-Afumatzi Meteorological Station (kJ/m^2/day).

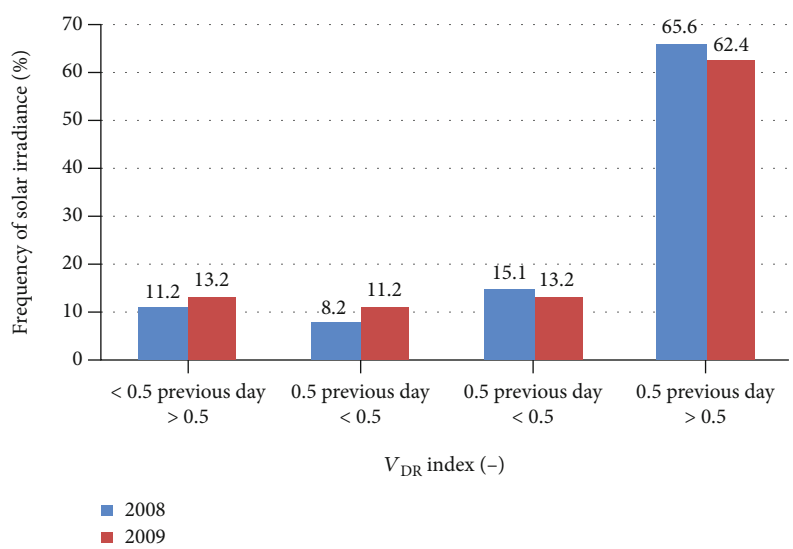


Figure 9: Relative frequency of solar irradiance/clouding situations quantified by the V_{DR} index, using the Bucharest/Afumatzi weather station.

We have analyzed the accuracy of the solar irradiance forecast developed in the four situations using a time series for the ARIMA and ANN models in order to quantify the significance of the forecast error in each case [22, 27]. Considering the frequency of situations during the investigated period, the most common situation is the fourth, which represents 64% of cases. The least common situation is the second, with a frequency of occurrence of only 8-11% of cases. The results of the analysis are represented in Figure 9. By this method, the errors of daily solar radiation forecasts can be quantified even during the forecasting process, separating only three situations, namely, the cloudy situation, the fog situation, and the one with partly overcast sky/clear sky conditions.

3.3. The Solar Radiation Forecast on Short and Long Term

3.3.1. Short-Term Solar Radiation Forecast with Application for a BIPV Lab System. Using the forecast models discussed in Section 3.1, respectively, ARIMA and ANN, we have made a short-term forecast, taking into account the synoptic situations of the days quantified by the V_{DR} index. For each situation, a random example is analyzed for each month, in 2008

Table 2: rRMSE values in the case of the four quantified situations of the V_{DR} index. Processed data from the 2008 year for a forecasting period of 10 days.

Month	<0.5 (the day before) >0.5		<0.5 (the day before) <0.5		>0.5 (the day before) <0.5		>0.5 (the day before) >0.5	
	ARIMA	ANN	ARIMA	ANN	ARIMA	ANN	ARIMA	ANN
1	168.6	194.7	35.6	56.9	79.2	90	9.2	15.9
2	197.9	191.6	58.6	40.5	103.7	115.6	9.9	21.8
3	232.7	209.1	53.3	77.6	47.6	64.4	39.4	42.6
4	191.4	184.4	101.9	77.4	67.4	92.1	22.2	26.3
5	136.6	127	53.5	59.9	51	66.5	12.9	17.2
6	101	95.3	46.8	66.7	51	69.4	16.3	22.9
7	132.3	134.1	85.1	72.5	44.8	69.2	25.6	36.4
8	No data	No data	No data	No data	No data	No data	23.8	20.9
9	No data	No data	No data	No data	No data	No data	6.9	10.9
10	423.3	347.3	No data	No data	100.4	108	12.4	32.6
11	105.9	68.2	70.9	26.6	99.7	105.5	30	24.5
12	136.0	183.2	58.7	58.2	87.4	89.7	35.6	48.2
Average	182.6	173.5	62.7	59.6	73.2	87	20.4	26.7

Table 3: rRMSE values in the case of four quantified situations of the V_{DR} index. Processed data from the 2009 year for a forecasting period of 10 days.

Month	<0.5 (the day before) >0.5		<0.5 (the day before) <0.5		>0.5 (the day before) <0.5		>0.5 (the day before) >0.5	
	ARIMA	ANN	ARIMA	ANN	ARIMA	ANN	ARIMA	ANN
1	291.1	356.1	55.2	65.1	57	76.5	25.7	33.3
2	67.4	111.3	56.2	42.2	86.2	104.8	28.9	18.8
3	299.5	367.3	51	40.6	69.6	100.2	26.8	23.8
4	141.3	108.5	131.8	50.2	103	113	20.8	17
5	237.3	232	60.1	70	55.2	87.5	9.8	14
6	189.1	168.3	No data	No data	69.5	84.6	15	24.2
7	207.6	204.1	No data	No data	87.7	108.5	21.5	28.8
8	327.9	233	No data	No data	61.6	67.7	6.2	19.3
9	469.9	440.4	232.3	66.5	75.8	84.6	15	29
10	No data	No data	131.4	35.4	83.9	94.1	53.4	20.3
11	443.8	551.1	98.7	53	82.1	95.6	16.7	24
12	344.6	334.5	54.5	50.2	87.1	113.8	51.4	33.5
Average	274.5	282.4	96.8	52.6	76.6	94.2	24.3	23.8

and 2009. For both models, the analyzed time series is 10 days. The accuracy of the forecasts is quantified by calculating the mean square relative error, using the relation:

$$\mathrm{rRMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(F_t - A_t)^2} * \frac{100}{G_n}, \tag{5}$$

where F_t represents the predicted value, A_t represents the current value (measured), n is the number of predicted points, and G_n represents the daily average of the current (measured) values.

Tables 2 and 3 contain the results of the forecasts corresponding to the four situations. It is noted that the best results are obtained in the case of partly cloudy/clear sky on the forecast day and partly cloudy/clear sky on the previous day, where the monthly rRMSE varies between 6.2 and 53%, and the annual one is less than 26.7%. In the case of cloudy/foggy sky on the forecast day and cloudy/-foggy sky the day before, as well as partly cloudy/clear sky on the forecast day and partly cloudy/foggy sky on the previous day, the accuracy of the forecasts is relatively similar, ranging between 52.6% and 96.8%. In the case of cloudy/foggy sky on the forecast day with partly cloudy/-clear sky on the previous day, the forecasts show a huge error of over 250%.

Comparing the two models, ARIMA and ANN, it was observed that the overall amplitude of errors is relatively similar in each case, but at the same time, the ARIMA model offers better results by 2.9% in the fourth case. We can

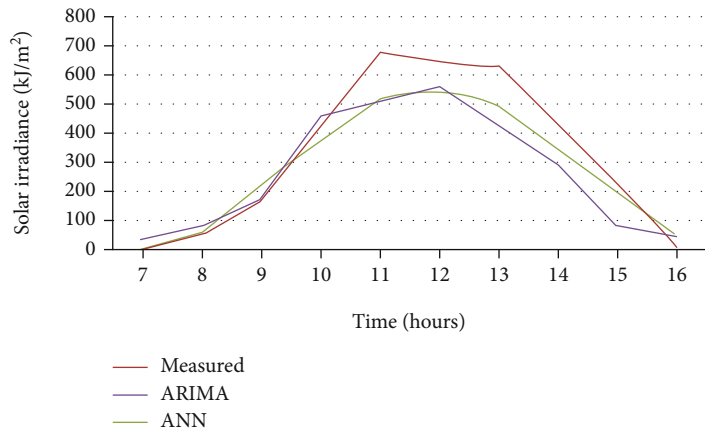


Figure 10: Solar radiation forecast for 20.01.2009 prepared with ARIMA and ANN models.

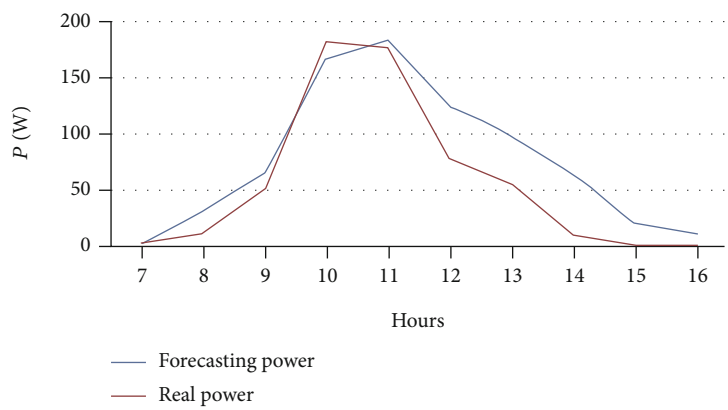


Figure 11: Predicted (forecast) and measured (real) power for 20.01.2009.

conclude that the accuracy of the forecasts depends closely on the daily variation of the solar radiation controlled by the clouding factor, and the weather series forecast provides acceptable results in the case of partly cloudy/clear sky on the forecast day with partly cloudy/clear sky the previous day (see [9, 15]).

Using the models of solar radiation forecast, and establishing the most frequent situation of the days from the meteorological point of view, the solar radiation forecast for January 20, 2009, was made and is represented in Figure 10.

With the help of this forecast, another forecast was elaborated, that of the power delivered by the lab BIPV system, represented in Figure 11.

The error calculated in this case was 26.9% and is consistent with the errors calculated for the solar radiation forecast, namely, 26.7%. This result is due to the improvement of the solar radiation forecast by integrating the data related to different meteorological parameters (temperature, cloud cover, etc.), as well as by separating the synoptic situations according to the variability index.

3.3.2. Long-Term Solar Radiation Forecast with Application for a BIPV Lab System. We have also analyzed the decade variation of solar radiation as long-term changes to be taken into account in solar energy applications. For the analysis of the change trend, we have applied the linear regression model. The annual data come from the World Radiation Data Center database (http://wrdc.mgo.rssi.ru) and represent data corresponding to the period 1975-2006 for the Bucharest weather station. The linear trend is significant at a probability level of 95%, thus identifying an increase in solar radiation. The magnitude of the absolute changes is 36.5 (±2.43) J/cm^2 day^{-1}/decade (Figure 12).

Based on the analysis of multiannual changes in solar radiation (J/cm^2 day^{-1}), the power forecast of the BIPV system for 2013 was made (see Figure 13).

It was observed that the maximum power production of the BIPV system is achieved in May. The total production forecast for 2013 is 163 kW, a result that is in line with the power achieved in the previous year, of 157 kW. We can conclude that the best forecast result can be obtained in situation (4) of partly cloudy/clear sky on the forecasted day and partly cloudy/clear sky on the previous day $V_{DR} > 0.5$, rRMSE—the annual error being 22.3 in the case of ARIMA and 25.3 in the case of ANN. In this case, ARIMA gives the best result in line with Reikard [9, 23] who showed that at 60-minute time resolution, the ARIMA model forecast dominates all time series forecasting methods in four from the six test stations used in the study. In the other synoptic situations, the errors are large. Given the frequency of occurrence of the four synoptic

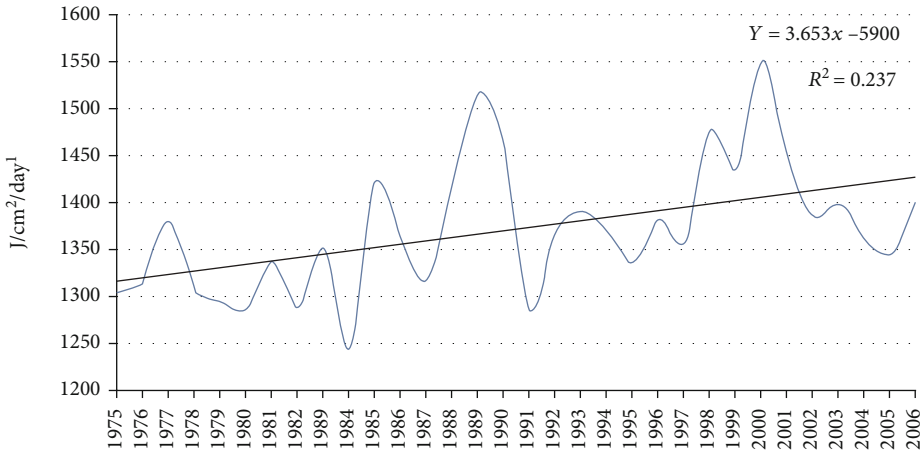


Figure 12: Multiannual variation of solar radiation in Bucharest during 1975-2006.

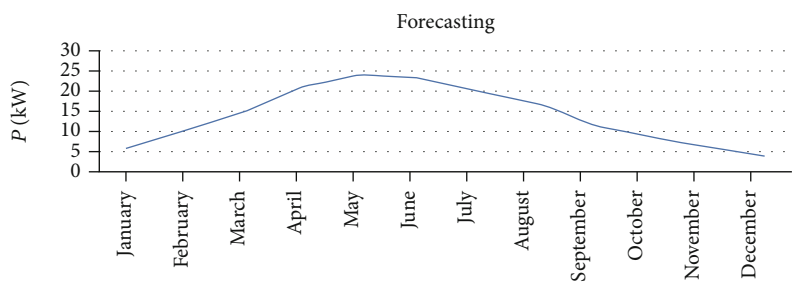


Figure 13: Power forecasted by the BIPV system for 2013, based on the estimation of multiannual solar radiation.

Figure 14: View of the PV park from Grojdibodu.

situations presented above, the situation in which we obtain the best results appears in 65.56% of the days of a year, so acceptable forecasts can be made in 65% of that year. In the case of the other situations, namely, (1), (2), and (3), physical and/or satellite models are required [24].

The results of the study contributed to the improvement of the forecasts of the photovoltaic systems, which will be used for the analysis of an experimental PV park [31, 32].

4. The Energy Production Forecast for a PV Park Using the ARIMA and ANN Models

4.1. Short Presentation of an Experimental PV Park. The experimental PV park from Grojdibodu (located in southern Romania) has in its composition 1931 strings, each string having 21 Suntech PV modules of 245 W. In total, there are 40,551 PV modules totaling an installed power of 9,934 kW. Solar

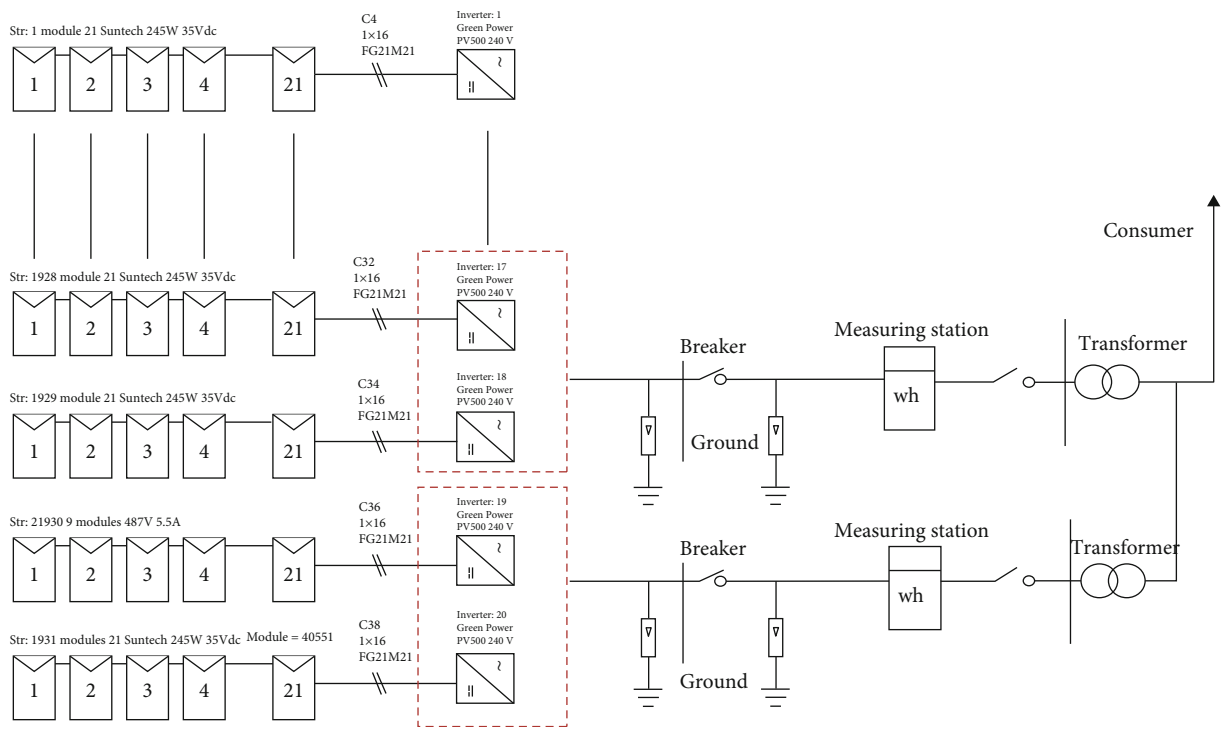


Figure 15: Electrical diagram of the studied PV park.

radiation data are acquired by 2 pyranometers and 10 calibration solar cells. The pyranometers are located in the meteorological station within the PV park, and the solar calibration cells are placed at each transformation center, thus being arranged on the entire surface of 33 ha of the park. The pyranometers and calibration cells are placed in a plane that has the same inclination to the horizontal as PV modules. There are a total of 20 Green Power PV500 inverters. The energy produced is measured at the connection point, and the furthest inverter is located at 400 m from this connection point. Each transformation center contains two inverters. The PV module is made of 72 polycrystalline silicon solar cells.

A view of the PV park from Grojdibodu is presented in Figure 14, and its electrical diagram is shown in Figure 15 [35, 36].

The main meteorological parameters of the Grojdibodu location defining the PV park are indicated in Table 4 [35, 36].

The technical parameters of the studied PV system/PV module, used in the forecast, are presented in Tables 5(a) and 5(b). Figure 16 shows the I-V (current-voltage) characteristics for the Suntech Power PV module used in the componence of the studied PV park.

The technical performance indicators of the Grojdibodu PV system for one year are designated in Table 6.

4.2. Results and Discussion: Energy Production Forecast for the Experimental PV Park. The purpose of the forecast was to size an experimental PV park located in the south of Romania, at Grojdibodu. Thus, a comparison was made for the results obtained by numerical modeling, with those measured in the PV park, as well as a comparative analysis of the programs used in terms of their efficiency. For simulation, we

Table 4: Meteorological parameters for Grojdibodu location.

Solar radiation	Annual average temp.	Annual average wind speed	Albedo
1.78 kWh/m^2/day	11.5°C	3.2 m/s	0.2

Table 5

(a) Technical parameters of the PV park from Grojdibodu

Installed power	9,934 kW
Power of the PV module	245 W Suntech Power 245
Total number of PV modules	40,551
Number of PV modules per string	21
Number of strings	1931
Number of inverters	20—Green Power PV500
Used surface	33 ha

(b) Technical parameters of the PV modules—main components of the PV experimental park

PV module parameters	
Nominal efficiency	13.9098%
Maximum power	269.850 Wcc
Maximum voltage	35 Vcc
Maximum current	7.7 Acc
Open circuit voltage	44.5 Vcc
Short circuit current	8.2 Acc

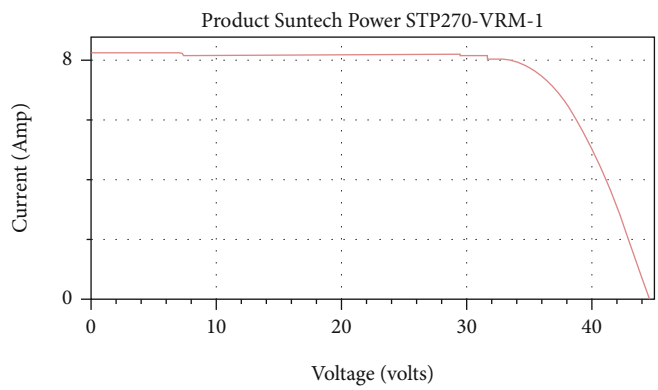


Figure 16: I-V characteristics of Suntech Power PV module.

Table 6: Performance indicators of the PV park from Grojdibodu.

Performance indicators (in the first year)	
Energy produced annually	13,069,573 kW
Capacity factor	15.0%
Energy efficiency	1,316 kWh/kW
Performance ratio	0.83

have used the data measured by the meteorological station within the PV park, considering the optimization of the PV system parameters, the increase of its energy efficiency, and the improvement of its global efficiency [26, 35, 36].

With the help of the ARIMA and ANN models, the forecast of surface solar radiation (SSR) was made, using data obtained from the PV park. Forecasts for 10 days have been developed based on the previous time series. The periods were chosen randomly, and the forecasts were checked for clear or partially clear days, which have the variability index of direct solar radiation $V_{\mathrm{DR}} > 0.5$. The variability index refers to the classification of synoptic situations (clear or cloudy). These situations are not included in the forecasts made. The index helps to validate forecasts; thus, statistical forecasts are more effective in the case of partly cloudy/clear sky on the forecast day and partly cloudy/clear sky on the previous day $V_{DR} > 0.5$. In the case of "cloudy" situations (1), (2), and (3), the rRMSE has large errors; for this reason, the statistical/empirical method for forecasting is not useful, and physical models are needed to improve it.

The correlation between the variability index and nebulosity was calculated in order to validate the index. The index shows the cloudiness on that day, without using cloudy data. The aim was to obtain information about nebulosity from solar radiation data. The validation of the index was performed for a shorter period; for example, in cases where the index had a value > 0.5, we checked whether the nebulosity was low or not, and vice versa. Following the verifications, the correlation was significant; as a result, we have used V_{DR} instead of nebulous data.

Short-term solar radiation forecasts were developed using the ARIMA and ANN models. The efficiency of the forecasted time series is quantified by rRMSE. The model was improved by separating the days with clear sky from the days with a high degree of cloudiness and using the variability index of solar radiation, V_{DR}. The separation of synoptic situations was performed on time series, not on predicted values. It has been established that the most common cases for the use of the V_{DR} index are days with clear sky on the forecast day and the day before. These days represent more than 75.4% of the total number of days of the year.

For the rest of the days when there are situations of cloudy sky or high fog, the forecast provides very large errors. The forecasted results conclude that the ARIMA model is more efficient than the ANN model. The statistics are significant in the case of ARIMA (1,0,14), this being the reason why this variant was chosen for analysis. Also, comparing the results of the forecast with the measured values, we noticed that the ARIMA model (1,0,14) is more efficient than the ARIMA (1,0,7). The ARIMA (1,0,7) and ARIMA (1,0,14) models were selected after performing several iterations. The statistical test used to identify the most significant model is Box Ljung Statistics. This test is applied in the case of ARIMA on the residues of a fitted model, not on the original series, and verifies that these residues do not show autocorrelation.

The results of the measurements and forecasted values for August and September 2013 are presented in Figures 17 and 18. The accuracy of the predictions is quantified using relative absolute mean error (rMAE) and relative mean square error (rRMSE):

$$\begin{aligned} \mathrm{rMAE} &= \frac{1}{n}\sum_{i=1}^{n} |F_t - A_t| \times \frac{100}{G_n}, \\ \mathrm{rRMSE} &= \sqrt{\frac{1}{n}\sum_{i=1}^{n} (F_t - A_t)^2 \times \frac{100}{G_n}}, \end{aligned} \tag{6}$$

where the following notations are introduced: F_t is the predicted value, A_t is the measured value, n is the number of predicted points, and G_n is the daily average of the measured values. Table 7 shows the measured and forecasted values for 10 days in August 2013, while Table 8 shows the same data in September 2013 for 10 days.

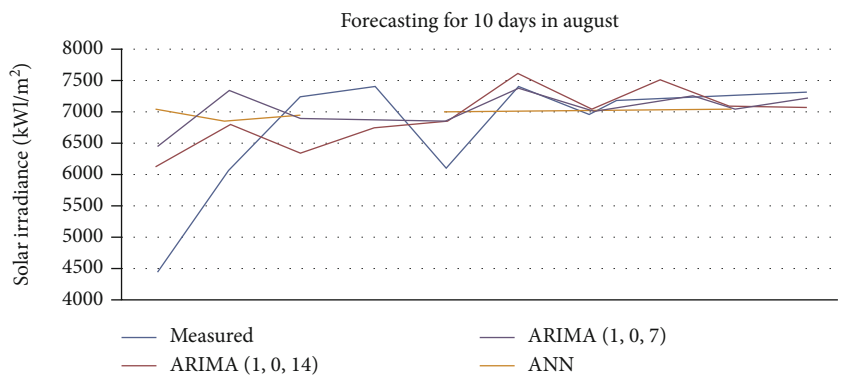


Figure 17: Solar irradiance forecast for Grojdibodu PV park, for 10 days, in August 2013, using the ARIMA and ANN models.

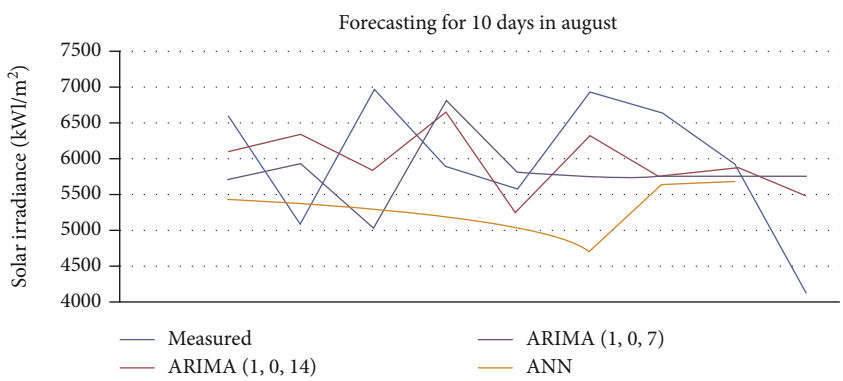


Figure 18: Solar irradiance forecast for Grojdibodu PV park, for 10 days, in September 2013, using the ARIMA and ANN models.

Table 7: Measured and forecasted values for 10 days, August 2013, based on ARIMA and ANN models. This table is reproduced from [35] under the permission of the editor.

Year	Month	Day	ARIMA (1,0,14) (Wh/mp)		ARIMA (1,0,7) (Wh/mp)		ANN (Wh/mp)	
			Measured	Forecasted	Measured	Forecasted	Measured	Forecasted
2013	7	31	4,451.3	6,138.76	4,451.3	6,459.48	4,451.3	7,044
2013	8	1	6,103.6	6,822.83	6,103.6	7,354.92	6,103.6	6,844
2013	8	2	7,240.3	6,353.43	7,240.3	6,916.88	7,240.3	6,963
2013	8	3	741.6	6,754.42	7,416.6	6,878.25	—	—
2013	8	4	6,091	6,878.99	6,091	6,894.17	6,091	7,024
2013	8	5	7,388.08	7,598.88	7,388.08	7,393.5	7,388.08	7,019
2013	8	6	7,101.34	7,006.33	7,101.34	6,974.19	7,101.34	7,002
2013	8	7	7,214.13	7,512.72	7,214.13	7,200.2	7,214.13	7,101
2013	8	8	7,285.4	7,055.09	7,285.4	7,069.23	7,285.4	7,098
2013	8	9	7,325.3	7,109.01	7,325.3	7,210.03	—	—

The two models offer good efficiency in terms of relative errors, but forecasts with errors below 40% are present only in the summer periods, especially in July, August, and September characterized by high values of solar radiation. In the case of low SSR, this methodology produces very large errors.

Based on the predicted values of solar radiation, the electrical power that would be inserted in the national power distribution grid could be calculated taking into account the efficiency of the PV park. Most likely, a global relationship for estimating the energy generated was [35, 36]

$$E = A * H * \mathrm{PR}, \tag{7}$$

where E represents the delivered energy (kWh), A represents the total surface area covered by the PV modules, H represents the average annual solar irradiance on the inclined

Table 8: Measured and forecasted values for 10 days, September 2013, based on ARIMA and ANN models. This table is reproduced from [35] under the permission of the editor.

Year	Month	Day	ARIMA (1,0,14) (Wh/mp)		ARIMA (1,0,7) (Wh/mp)		ANN (Wh/mp)	
			Measured	Forecasted	Measured	Forecasted	Measured	Forecasted
2013	9	19	—	—	—	—	—	—
2013	9	20	6,624.2	6,123.75	6,624.2	5,721.44	6,624.2	5,431
2013	9	21	5,078	6,379.47	5,078	5,973.61	5,078	5,421
2013	9	22	6,997.1	5,857.37	6,997.1	5,038.02	6,997.1	5,315
2013	9	23	5,907	6,697.06	5,907	683.91	5,907	5,210
2013	9	24	5,595.3	5,228.81	5,595.3	5,847.36	5,595.3	5,019
2013	9	25	6,976.2	6,379.57	6,976.2	5,745.71	6,976.2	4,696
2013	9	26	6,661.8	5,745.95	6,661.8	5,755.16	6,661.8	5,668
2013	9	27	5,983.8	5,882.45	5,983.8	5,756.6	5,983.8	5,706
2013	9	28	4,089.2	5,472.23	4,089.2	5,755.73	—	—

Table 9: Main results of the energy production by the experimental PV park from Grojdibodu forecasted for the period 01-09.08.2013.

Day	1	2	3	4	5	6	7	8	9
Hours when the inverter is running (solar irradiance > 10 W/m^2)	13.75	13.25	13	13	13	13.25	13	13	13
Total hours ∗ string	26,565	25,599	25,116	25,116	25,116	25,599	25,116	25,116	25,116
Total solar irradiance (Wh/m^2)	6,103.6	7,240.3	7,416.6	6,091	7,388.08	7,101.34	7,214.13	7,285.4	7,325.3
Exported active energy (MWh)	50.77	59.48	60.74	48.49	60.18	56.62	59	59.16	59.54
PR (%) calculated in SCADA	83.51	82.55	82.22	79.99	81.83	80.07	82.12	81.54	81.66
Real availability (%)	100	100	100	97.5	98.84	98.11	100	100	100
Availability given by the manufacturer (%)	100	100	100	97.5	98.84	100	100	100	100
Events in which energy production is disrupted	—	—	—	Inverter off 6 hours	Inverter: motherboard damaged (replaced)	Inverter off 15 min	—	—	—
∗String hours with energy loss	—	—	—	627.9	291.34	483.82	—	—	—
Estimated energy losses (MWh)	—	—	—	1.8	0.4	1.75	—	—	—

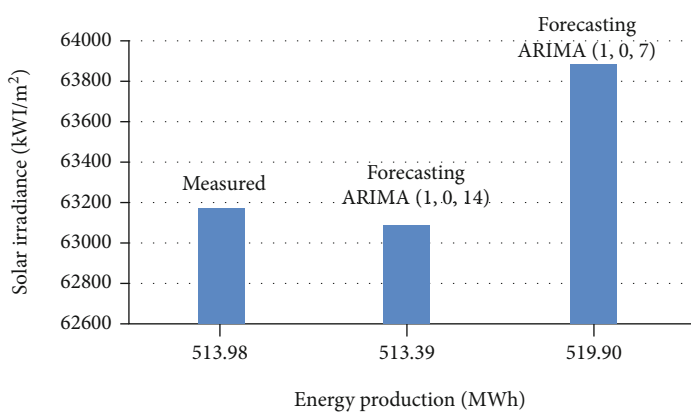


Figure 19: Difference between the measured values of solar irradiance and predicted ones based on the ARIMA model.

PV panels (Wh/m^2), and PR (performance ratio) represents the efficiency of the PV system.

The comparison between the forecasted energy and the energy inserted in the grid is made between 01.08.2013 and 09.08.2013 for the analyzed experimental PV park (see Table 9).

Considering the data collected within the PV park, the monthly exported energy shows that in the period chosen for testing 01-09.08.2013, a quantity of 513.98 MWh was delivered in the distribution grid. The share of the measured solar irradiance, corresponding to this value of the delivered energy, was 63165.75 Wh/m^2.

Table 10: Comparison of different forecasting models.

Ref.	Year of publ.	Forecasting model	Error	Time horizon	Location
[34]	2018	ARIMA, GA, PSO, DE	VAR, WME	1 day	Canada
[33, 34]	2018, 2018	SVM	RMSE, MRE	15 min.	China
[31, 34]	2019, 2018	NWP	MAE, MBE, RMSE,	1 h, 3 h	USA—California state
[32, 34]	2019, 2018	ANN	NRMSE, NMAE, NBE	1 h	Italy
[34]	2018	SVM, NN, MLR, RT	RMSE	2 h	Australia
Present article: BIPV lab system	2021: to be published	ARIMA, ANN	RMSE,MAE, MAPE, MaxAPE, MaxAE	15 min, 10 days (time series)	South of Romania, Bucharest, BIPV lab system
Present article: experimental PV park	2021: to be published	ARIMA, ANN	rMAE, rRMSE	10 days (time series)	South of Romania, experimental PV park

Using the ARIMA model (1,0,14), the solar irradiance for this period was forecasted at 63091 Wh/m^2, respectively, and at 63891 Wh/m^2 using the ARIMA model (1,0,7). Taking into account the technical parameters of the system and the forecast for solar irradiance, the energy forecast with the ARIMA type model will be 513.39 MWh, respectively, and 519.90 MWh (Figure 19).

5. Comparative Analysis of Different Forecasting Models

A comparison of different forecasting models (algorithms) developed in various countries is presented in Table 10.

The results obtained in the present article were compared based on ARIMA and ANN tools with new advanced forecasting models (algorithms), like PSO, GA, SVM, MLR, and RT [31–34]. The accuracy was much better in these cases. However, the authors have studied two interesting PV installations, namely, a BIPV lab system and an experimental PV park; both were placed in the south of Romania. Based on our results, it was established that the ARIMA (1,0.14) model is more efficient than the ANN one and is very close to the experimental case (see Figure 19). Although our approach is dedicated especially to short-term forecast, the long-term solar radiation forecast was discussed too with application for a BIPV system developed at the Polytechnic University of Bucharest. This feature is different from the other more advanced forecasting models presented in Table 10.

6. Optimized Forecasting Method for Energy Production Based on GMDH Model: Case Study—PV Park in Romania

The GMDH (Group Method of Data Handling) represents a family of inductive algorithms for the mathematical modeling of multiparametric datasets that are able to automatically optimize the used models [37]. In the GMDH, continuous or discrete input variables can be introduced, and depending on their type, optimal parametric or nonparametric algorithms can be identified.

In order to optimize the energy production developed by a Romanian PV park, a specific approach was conceived to render efficient solar radiation prediction on the surface of PV panels for a PV park. As the prediction method involves a large amount of data and a great computing power, Romania's territory was divided into 28 regions, considering that each region possesses at least one actinometric station.

For the created regions, important meteorological data will be provided for solar radiation prediction and will be introduced into advanced models for prediction and evaluation (ARIMA, ANN). The goal is to improve prediction efficiency in both effective results and computing time.

Alongside corresponding solar radiation data and associated factors, for an efficient prediction, it will be introducing also the PV park's characteristics [38]. System loss categories are present in all PV systems but differ from one system to another. It is important to mention the fact that the PV park loss is not an aggregate of particular losses. These losses affect the system individually, their impact being calculated for each component. The losses are due to the dust deposits on the module's surfaces, shadows within the emplacement zone, snow deposits, electrical installation's asymmetry, losses between module's connections (wire harness loss), connection loss, etc. Also, the degrading agent must be considered; it represents the production loss during the maintenance period. Considering the PV park's parameters and also the system loss, the energy produced by the PV park was determined both during the days 04, 08, and 14 of July for 2013 and 2014 and during the days 29, 30, and 31 of July 2013 and 2014. The results are presented in Table 11.

Forecasted values show acceptable results correlated with the errors for solar radiation forecasting. Since the forecasts were made using data from previous days without taking into account future forecasts, this method could be optimized by introducing additional parameters that

Table 11: PV park—comparison between measured and forecasted energy using PR of the PV system.

Date	PR (performance ratio) (%)	Measured exported energy (MWh)	Forecasted energy (MWh)	Percentage difference between measured and forecasted values (%)
04 July 2013	81.58	57.47	67.45	17%
04 July 2014	84.98	49.47	66.96	35%
08 July 2013	84.2	57.36	52.79	8%
08 July 2014	84.04	56.83	61.64	8%
14 July 2013	85.72	63.15	67.73	7%
14 July 2014	88.31	30.89	76.71	148%
29 July 2013	80.05	58.11	56.36	3%
29 July 2014	67.02	36.28	44.03	21%
30 July 2013	83.36	49.33	59.87	21%
30 July 2014	83.42	59.75	56.61	5%
31 July 2013	86.31	38.5	62.22	62%
31 July 2014	85.81	35.66	52.35	47%

meteorologically describe the following day. Research continues to identify correlations between meteorological parameters that could accurately be forecasted for a given day and their impact on solar radiation.

7. Conclusions and Prospects

Short-term solar radiation forecasts were developed using the ARIMA and ANN models. The efficiency of the forecasted time series is quantified by rRMSE. The models were improved by separating the days with clear sky from the days with a high degree of cloudiness using the variability index of solar radiation, V_{DR}. The separation of the synoptic situations was performed on time series, not on the predicted values. It has been established that the most common cases for the use of the V_{DR} index are days with clear sky on the forecast day and the day before, where the day "forecast" is represented by the day "tomorrow" and the day "before" is represented by "today" or the day on which the forecast is made. These days represent more than 75.4% of the total number of days of the year, so only these cases were chosen. For the rest of the days when there are situations of cloudy sky or high cloud cover, these forecast solutions offer very large errors. The results of the forecast attest to the fact that the ARIMA model is more efficient than the ANN model. The statistics is significant in the case of ARIMA (1,0,14), this being the reason why this variant was chosen for analysis. Also, comparing the results of the forecast with the measured values, we notice that the ARIMA model (1,0,14) is more efficient than the ARIMA (1,0,7). The ARIMA (1,0,7) and ARIMA (1,0,14) models were selected after performing several iterations. The statistical test used to identify the most significant model is Box Ljung Statistics. This test is applied for ARIMA on the residues of a fitted model, not on the original series, and verifies if these residues do not show autocorrelation. The Box-Ljung model, also known as the statistically modified Box-Pierce model, provides guidance on whether the model is specified correctly. A value significantly lower than 0.05 implies that within the observed data there is a structure that is not relevant for the model. A value greater than 0.05, as determined in the case, indicates that the chosen model is correct (see [9, 35, 36].

The future work will be based on a new optimized forecasting method for PV energy production, presented in Section 6. Taking into account that the optimized prediction and forecasting method would involve a large amount of data and great computing power, Romania's territory could be divided into 28 regions; each region would possess one actinometric (solar radiation) station. A polynomial neural network based on the GMDH (Group Method of Data Handling) approach would be used. It would be represented by a family of inductive algorithms (parametric and nonparametric), which could optimize automatically the energy production of the PV park.

Abbreviations

PV: Photovoltaic
ARIMA: Autoregressive Integrated Moving Average
ANN: Artificial Neural Network
BIPV: Building Integrated Photovoltaics
GMDH: Group Method of Data Handling
AIC: Akaike Information Criterion
SARIMA: Seasonal ARIMA
NWP: Numerical weather prediction
TRNSYS: Transient system simulation tool
eQuest: Quick energy simulation tool
Energy Plus: Open source software
BAS: Building automation systems
BEMS: Building energy management systems
HVAC: Heating, ventilation, and air conditioning
DL: Deep Learning
DBN: Deep Belief Networks
RNN: Recurrent Neural Networks
STMN: Short-Term Memory Networks
AI: Artificial intelligence
IBM SPSS: Predictive analysis program
MLE: Maximum Likelihood Estimation

AR: Autoregressive
MA: Moving average
AC and PAC: Autocorrelation command and partial autocorrelation command
RH: Relative humidity
RMSE: Root Mean Squared Error
MAPE: Mean Absolute Percentage Error
MaxAPE: Maximum Absolute Percentage Error
MAE: Mean Absolute Error
MaxAE: Maximum Absolute Error
WMO: World Meteorological Organization
WRDC: World Radiation Data Center
V_{DR}: Variability index of the solar irradiance
rMAE: Relative mean absolute error
rRMSE: Relative mean squared error
PR: Performance ratio
GA: Genetic algorithm
PSO: Particle swarm optimization
DE: Differential evolution
SVM: Support vector machine
MRE: Mean relative error
VAR: Variance error
WME: Weekly mean error
MBE: Mean bias error
NRMSE: Numerical root mean squared error
NMAE: Numerical mean absolute error
NBE: Numerical bias error
RT: Regression tree
MLR: Multiple linear regression
NN: Neural network.

Data Availability

Previously reported data were used to support this article and were based on prior studies cited at relevant places within the text of the paper as references [9, 11, 35].

Disclosure

A PhD thesis related to this subject was successfully concluded at the Polytechnic University of Bucharest.

Conflicts of Interest

The authors declare no conflict of interest.

Authors' Contributions

Conceptualization was performed by L.F.; methodology was performed by L.F. and A.D.; formal analysis was performed by L.F.; investigation was performed by L.F. and A.D.; software was secured by A.D. and D.C.; validation was performed by L.F., A.D., DC, and S.F.; resources were secured by A.D. and S.F.; data curation was performed by A.D. and S.F.; writing (original draft preparation) was performed by L.F.; writing (review and editing) was performed by L.F.; visualization was performed by L.F., D.C., and S.F.; supervision was performed by L.F.; and project administration was performed by L.F.

Acknowledgments

The authors are grateful to Dr. Cristian Oprea who supplied the solar radiation data from Bucharest-Afumati Meteorological Station that were used in this article. At the same time, the authors appreciate the role of the persons in charge from the PV park from Grojdibodu, south of Romania, who allowed the utilization of some technical information related to this park, as well as specific meteorological data of the site within the present paper.

References

[1] S. Sobri, S. Koohi-Kamali, and N. A. Rahim, "Solar photovoltaic generation forecasting methods: a review," *Energy Conversion and Management*, vol. 156, pp. 459–497, 2018.

[2] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F. J. Martinez-de-Pison, and F. Antonanzas-Torres, "Review of photovoltaic power forecasting," *Solar Energy*, vol. 136, pp. 78–111, 2016.

[3] H. M. Diagne, M. David, P. Lauret, and J. Bolan, "Solar irradiation forecasting: state-of-the-art and proposition for future developments for small-scale insular grids," in *Proceedings of the WREF 2012-World Renewable Energy Forum*, pp. 65–76, Denver, Colorado, 2012.

[4] E. G. Kardakos, M. Alexiadis, S. Vagropoulos, and A. G. Bakirtzis, "Application of time series and artificial neural network models in short-term forecasting of PV power generation," in *2013 48th International Universities' Power Engineering Conference (UPEC)*, pp. 1–6, Dublin, Ireland, 2013.

[5] L. Fara, A. Diaconu, and F. Dragan, "Trends, challenges and opportunities in advanced solar cells technologies and PV market," *Journal of Green Engineering*, vol. 5, no. 4, pp. 157–186, 2016.

[6] L. Fara, C. Cincu, R. Mitroi et al., *Physics and technology of solar cells and photovoltaic systems*, AOSR Publication, Bucharest, 2009.

[7] M. Paulescu, L. Fara, and E. Tulcan-Paulescu, "Models for obtaining daily global solar irradiation from air temperature data," *Atmospheric Research*, vol. 79, no. 3-4, pp. 227–240, 2006.

[8] F. Iacobescu and V. Badescu, "The potential of the local administration as driving force for the implementation of the National PV systems Strategy in Romania," *Renew Energy*, vol. 38, no. 1, pp. 117–125, 2012.

[9] L. Fara, B. Bartok, A. Galbeaza Moraru et al., "New results in forecasting of photovoltaic systems output based on solar radiation forecasting," *Journal of Renewable and Sustainable Energy*, vol. 5, no. 4, 2013.

[10] M. Brabec, V. Badescu, and M. Paulescu, "Nowcasting sunshine number using logistic modeling," *Meteorology and Atmospheric Physics*, vol. 120, no. 1-2, pp. 61–71, 2013.

[11] L. Fara, A. G. Moraru, P. Sterian, A. P. Bobei, A. Diaconu, and S. Fara, "BIPV (building integrated photovoltaic) systems în Romania. Monitoring, modelling and experimental validation," *Journal of Optoelectronics and Advanced Materials*, vol. 5, no. 1-2, pp. 125–130, 2013.

[12] M. Paulescu, O. Mares, E. Paulescu et al., "Nowcasting solar irradiance using the sunshine number," *Energy Conversion and Management*, vol. 79, pp. 690–697, 2014.

[13] L. Fara, D. Craciunescu, L. Fara et al., "Estimation of electricity production for a photovoltaic park using specialized advanced software," in *Proceedings of the EuroSun 2016- International Conference of Solar Energy for Buildings and Industry*, Palma de Mallorca, 2016.

[14] S. Atique, S. Noureen, V. Roy, V. Subburaj, S. Bayne, and J. Macfie, "Forecasting of total daily solar energy generation using ARIMA: a case study," in *2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)*, pp. 0114–0119, Las Vegas, NV, USA, 2019.

[15] J. Runge and R. Zmeureanu, "Forecasting energy use in buildings using artificial neural networks: a review," *Energies*, vol. 12, no. 17, p. 3254, 2019.

[16] X. Qing and Y. Niu, "Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM," *Energy*, vol. 148, pp. 461–468, 2018.

[17] Y.-K. Wu, C.-R. Chen, and H. Abdul Rahman, "A Novel Hybrid Model for Short-Term Forecasting in PV Power Generation," *International Journal of Photoenergy*, vol. 2014, Article ID 569249, 9 pages, 2014.

[18] M. Cococcioni, E. D'Andrea, and B. Lazzerini, "24-hour-ahead forecasting of energy production in solar PV systems," in *2011 11th International Conference on Intelligent Systems Design and Applications*, pp. 1276–1281, Cordoba, Spain, 2011.

[19] S. Vagropoulos, G. Christos Chouliaras, E. G. Kardakos, and A. G. Bakirtzis, "Comparison of SARIMAX, SARIMA, modified SARIMA and ANN-based models for short-term PV generation forecasting," in *2016 IEEE International Energy Conference (ENERGYCON)*, pp. 1–6, Leuven, Belgium, 2016.

[20] H. X. Zhao and F. Magoulès, "A review on the prediction of building energy consumption," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 6, pp. 3586–3592, 2012.

[21] M. Paulescu, V. Badescu, and M. Brabec, "Tools for PV (photovoltaic) plant operators: nowcasting of passing clouds," *Energy*, vol. 54, pp. 104–112, 2013.

[22] A. Mellit, M. Benghanem, A. H. Arab, and A. Guessoum, "A simplified model for generating sequences of global solar radiation data for isolated sites: using artificial neural network and a library of Markov transition matrices approach," *Solar Energy*, vol. 79, no. 5, pp. 469–482, 2005.

[23] Z. Wang and R. S. Srinivasan, "A review of artificial intelligence based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models," *Renewable and Sustainable Energy Reviews*, vol. 75, pp. 796–808, 2017.

[24] A. T. Eseye, J. Zhang, and D. Zheng, "Short-term photovoltaic solar power forecasting using a hybrid Wavelet-PSO-SVM model based on SCADA and Meteorological information," *Renewable Energy*, vol. 118, pp. 357–367, 2018.

[25] K. Wang, X. Qi, and H. Liu, "A comparison of day-ahead photovoltaic power forecasting models based on deep learning neural network," *Applied Energy*, vol. 251, 2019.

[26] F. Harrou, B. Taghezouit, and Y. Sun, "Robust and flexible strategy for fault detection in grid-connected photovoltaic systems," *Energy Conversion and Management*, vol. 180, pp. 1153–1166, 2019.

[27] M. A. Mat Daut, M. Y. Hassan, H. Abdullah, H. A. Rahman, M. P. Abdullah, and F. Hussin, "Building electrical energy consumption forecasting analysis using conventional and artificial intelligence methods: a review," *Renewable and Sustainable Energy Reviews*, vol. 70, pp. 1108–1118, 2017.

[28] M. K. Behera, I. Majumder, and N. Nayak, "Solar photovoltaic power forecasting using optimized modified extreme learning machine technique," *Engineering Science and Technology- an International Journal*, vol. 21, no. 3, pp. 428–438, 2018.

[29] S. Srivastava and S. Lessmann, "A comparative study of LSTM neural networks in forecasting day-ahead global horizontal irradiance with satellite data," *Solar Energy*, vol. 162, pp. 232–247, 2018.

[30] M. C. Mukkamala and M. Hein, "Variants of RMSProp and Adagrad with logarithmic regret bounds," in *Proceedings of the 34th International Conference on Machine Learning*, pp. 2545–2553, Sydney, Australia, 2017.

[31] M. N. Akhter et al., "Review on forecasting of photovoltaic power generation based on machine learning and metaheuristic techniques," *IET Renewable Power Generation*, vol. 13, no. 7, pp. 1009–1023, 2019.

[32] W. VanDeventer, E. Jamei, G. S. Thirunavukkarasu et al., "Short-term PV power forecasting using hybrid GASVM technique," *Renewable Energy*, vol. 140, pp. 367–379, 2019.

[33] U. K. Das, K. S. Tey, M. Seyedmahmoudian et al., "Forecasting of photovoltaic power generation and model optimization: a review," *Renewable and Sustainable Energy Reviews*, vol. 81, Part 1, pp. 912–928, 2018.

[34] M. Seyedmahmoudian, E. Jamei, G. S. Thirunavukkarasu et al., "Short-term forecasting of the output power of a building-integrated photovoltaic system using a metaheuristic approach," *Energies*, vol. 11, no. 5, 2018.

[35] L. Fara, A. Diaconu, and D. Craciunescu, "Forecasting of energy production and operational optmization for photovoltaic systems," *Annals of the Academy of Romanian Scientists, Series on Physics and Chemistry Sciences*, vol. 2, no. 1, pp. 31–60, 2017.

[36] A. Diaconu, *Contributions to the operational optimization of photovoltaic systems in applications, [Ph.D. thesis]*, Polytechnic University of Bucharest, Bucharest, 2017.

[37] https://www.gmdhshell.com/time-series-forecasting.

[38] http://photovoltaic-software.com/PV-solar-energy-calculation.php.

Hindawi
International Journal of Photoenergy
Volume 2021, Article ID 9973010, 27 pages
https://doi.org/10.1155/2021/9973010

Research Article

Sky Image-Based Localized, Short-Term Solar Irradiance Forecasting for Multiple PV Sites via Cloud Motion Tracking

Lasanthika H. Dissawa,[1] Roshan I. Godaliyadda,[1] Parakrama B. Ekanayake,[1] Ashish P. Agalgaonkar,[2] Duane Robinson,[2] Janaka B. Ekanayake,[1] and Sarath Perera[2]

[1]*Department of Electrical and Electronics Engineering, Faculty of Engineering, University of Peradeniya, Peradeniya, Sri Lanka*
[2]*School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, Wollongong, Australia*

Correspondence should be addressed to Lasanthika H. Dissawa; lasanthikadissawa@yahoo.com

Received 15 March 2021; Revised 3 June 2021; Accepted 29 June 2021; Published 13 July 2021

Academic Editor: Bogdan Gabriel BURDUHOS

Power generation through solar photovoltaics has shown significant growth in recent years. However, high penetration of solar PV creates power system operational issues as a result of solar PV variability and uncertainty. Short-term PV variability mainly occurs due to the intermittency of cloud cover. Therefore, to mitigate the effects of PV variability, a sky-image-based, localized, global horizontal irradiance forecasting model was introduced considering the individual cloud motion, cloud thicknesses, and the elevations of clouds above the ground level. The proposed forecasting model works independently of any historical irradiance measurements. Two inexpensive sky camera systems were developed and placed in two different locations to obtain sky images for cloud tracking and cloud-based heights. Then, irradiance values for onsite and for a PV site located with a distance of 2 km from the main camera were forecasted for 1 minute, 5 minutes, and 15 minutes ahead of real-time. Results show that the three-level cloud categorization and the individual cloud movement tracking method introduced in this paper increase the forecasting accuracy. For partially cloudy and sunny days, the forecasting model for 15 min forecasting time interval achieved a positive skill factor concerning the persistent model. The accuracy of determining the correct irradiance state for a 1 min forecasting time interval using the proposed model is 81%. The average measures of RMSE, MAE, and SF obtained using the proposed method for 15 min forecasting time horizon are 101 Wm^{-2}, 64 Wm^{-2}, and 0.26, respectively. These forecasting accuracy levels are much higher than the other benchmarks considered in this paper.

1. Introduction

With the declining prices and promotion of green energy, power generation using solar photovoltaic (PV) progresses to be an alternative variable power generation option in many countries [1, 2]. For example, as of 30th September 2020, there are over 2.56 million PV installations in Australia, with a cumulative capacity of more than 18.5 GW [3]. Furthermore, solar PV was accounted for 5.6% of the total generation in 2019, and it is the fastest-growing generation type in the years 2018 and 2019 [4].

However, due to solar PV variability and intermittency, the increased penetration of solar PV into the power system can cause operational and management issues. The solar irradiance on PV panels varies with date, time, location, and panel orientation relative to the sun [5]. It is well known that the diurnal and annual solar irradiance patterns are highly predictable, and solar variability in longer time intervals can be easily estimated. But, the amount of solar irradiance that reaches the surface of the earth varies by the intermittency of cloud cover, impacting short-term solar energy production. As these variations create significant fluctuations in solar power feed into the grid, the methods that can be used to predict solar irradiance at ground level and thus the corresponding PV power generation are necessary to ensure the effective management of electrical grids [6–8].

Reference [9] categorized solar PV forecasting methods based on forecast time horizon (short-term, medium-term, and long-term), historical data, and forecasting methods. Historical data-based PV forecasting models use PV power

output and related meteorological variables as the inputs. Persistence model, physical models [10], support vector machine-based models [11, 12], and artificial neural network-based models [13] are some of the forecasting methods used to forecast PV power in different forecasting time intervals

Solar power/irradiance forecasting is a powerful tool that can be used to mitigate problems associated with the short-term variations of the generated solar power. Intrahour (from few seconds to few minutes) forecasting is used to identify local ramp up/down events few minutes in advance [14] and can identify pending energy shortfalls, which are helpful for managing PV inverters and energy storage systems effectively. PV fluctuations create issues in maintaining steady-state voltages at the distribution grid within the statutory limits [15]. If a large number of PV plants are connected to the distribution grid, voltages at the PV connection points will rise when the PV generation is high and the captive load is low [16]. Therefore, the voltage rise is considered as one of the dominant issues that limit the ability of the distribution grid to host high PV penetration. As a remedy, or to maintain the voltage within statutory limits, On-Load Tap Changers (OLTC) with smart distribution management systems (S-DMS) [17] are used. Short-term solar forecasting is one of the main building blocks in S-DMS as it is required to predict the network status to control and manage the smart inverters and smart transformers.

Furthermore, short-term forecasting will be beneficial to overcome the partial shading condition that occurred due to passing clouds, which is another major problem in PV systems [18]. In partial shading conditions, reconfigurations of shaded and nonshaded modules in PV arrays enhance the power output by distributing the shading effects evenly [19] without changing the physical location. The short-term cloud shadow forecasts at the location of the PV plant are used as a control signal to the array reconfiguration process to minimize the effect due to the partial shading especially on large-scale PV plants. Further, solar forecasting is used as an input to smart battery management systems to compensate for the PV shortages [20]. Furthermore, short-term solar PV forecasts are used in energy market activities. For example, in Australia, 5-minute PV forecasts are used in the operations in the Australian National Electricity Market (NEM) [21].

Short-term fluctuations in a PV plant mainly occur due to the shadows of clouds and shadows created by fixed objects like buildings, trees, and mountains. The power output of a PV plant due to the shadows created by fixed objects is deterministic as shadows can be mapped onto the PV plant according to the zenith and azimuth angle of the sun at a given time. However, the effects created by the shadows of the clouds vary from time to time depending on cloud velocity, cloud size, cloud position on the sky, cloud thickness, and the texture of the cloud. Hence, it is a random behaviour that requires stochastic prediction as opposed to the deterministic part, which requires extrapolation. Therefore, this paper is focused on a physical forecasting model developed based on cloud shading on PV plant generation.

For short-term power forecasting or real-time power prediction, cloud information from ground-based sky images and time series models based on historical data is widely used [11, 22, 23]. Sky image-based PV forecasting approaches reported in the literature are summarised in Table 1 with the methodology used.

There are some limitations in the existing sky image-based forecasting methods and equipment used to get the data for the forecast. For example, the camera used in [10, 24–26] are expensive to install on a large scale, and some camera systems have proprietary software. Further, the local cloud base height (CBH) information used in [10, 25, 27] is obtained from ceilometers located around 10 km away from the sky imager. This will potentially introduce significant shadow position errors when mapping the cloud shadow onto the ground. Furthermore, the methods presented in [28, 29] need previous irradiance measurements to obtain the forecasts.

The deficiency of the technique introduced in [10, 27] is that the entire cloud area is assumed to be moving at a uniform velocity throughout the image without considering individual cloud movement. Furthermore, in [10], GHI drops due to the shadow of the clouds are assumed to be equal to a constant percentage drop. The use of a single GHI dropping percentage is not robust since the decline of the GHI level may be different according to the thicknesses of the cloud.

The cloud tracking using the Lucas-Kanade optical flow algorithm in [30] needs a higher image capturing frequency to obtain smooth cloud movement. Furthermore, the major weakness of the model presented in [31] is that this model is an onsite forecasting model. Moreover, Reference [32] used the tracking details to find the changes in the features in the sun-blocking window and it did not use cloud motion tracking and shadow casting directly to the forecasting model.

By considering the limitation of the state-of-the-art methods, this paper introduces a multiple-site irradiance forecasting model improved based on the CBH calculation using two cameras located in two different locations. A new cloud pixel identification method was introduced to identify cloud areas in the sky image. Furthermore, a novel approach was introduced to classify the cloudy pixels, where the cloudy pixels were divided into three groups based on the color properties of each cloud pixel. The irradiance dropping factor was defined using the cloud pixel category. Instead of assuming a single dropping factor in [10, 14], an irradiance dropping factor based on image cloud color property was introduced.

Furthermore, individual clouds were tracked separately without assuming the entire cloud area is moving at a uniform velocity throughout the image. A normalized cross-correlation algorithm was utilized to estimate the cloud motion vectors, which is more convenient than the other optical flow techniques. Here, the CCM [34] was applied to each cloud separately, without taking the total cloud area as one segment. This method enables the determination of multiple layer cloud movements. Finally, irradiance forecasts were obtained for onsite PV system (e.g., location 1 in Figure 1) and as well as for PV systems located away from the main camera (e.g., location 3 in Figure 1), by utilizing clear day irradiance profile generated using the ASHRAE clear-sky model [35] together with irradiance drop percentage corresponding to the cloud category and drop occurrence time.

Table 1: Details of the sky image-based forecasting models.

Ref.	Data	Method used	Outcomes
[10]	(i) Sky images Camera: total sky imager 440A Resolution: 640 × 480 pixels Frequency: every 30 s Image format: jpg (ii) CBH	Cloud segmentation: RBR of each image pixel was taken with the help of both clear sky image data and sunshine parameter (SP) Tracking method: cross-correlation method (CCM). The sky images were partitioned into subsets of pixels of equal size of squares Forecasting method: by mapping the cloud shadow onto the ground and considering the average cloud moving velocity (assuming spatial homogeneity of cloud velocity), occlusion time was obtained. The drop occurred in the GHI due to clouds was assumed to be equal to 40% of the clear sky GHI value	Irradiance is forecasted 5 minutes ahead of time
[30]	(i) Sky images Camera: has roughly 60° FOV Resolution: 640 × 480 pixels Frequency: 1 frame/second Image format: jpeg	Cloud segmentation: the difference between the blue color channel and the red color channel of each image pixel was compared with a threshold Tracking method: Lucas-Kanade optical flow algorithm Forecasting method: using linear regression, the pixel moving velocity was obtained According to the velocity, feature point trajectories were developed, and time taken by the feature points to pass a specific location on the image was obtained	Occlusion signals were generated 30 seconds ahead of time
[28]	(i) Sky images Camera: the total sky imager (TSI) Frequency: every 30 s (ii) Pyranometer irradiance measurements	Tracking method: fast cross-correlation algorithm Forecasting method: a linear prediction model was introduced for irradiance forecast based on cloud motion estimations and the previous solar irradiance monitoring data From the motion vectors, the future cloud motion over the location where solar panels reside was estimated. The time series model was defined employing radiation data and the TSI image RBR value change readings concerning the previous step on the selected window (20 × 20)	1 min and 2 min ahead irradiance forecasts were obtained
[27]	(i) Sky images Camera: UCSD sky imager Resolution: 2048 × 2048 pixels Frequency: captured every 30 s (ii) CBH	Cloud segmentation: sky was segmented into three categories: applying the threshold to the RBR channel and comparing the images with a clear sky model Tracking method: CCM to the RBR of two consecutive images as in [10] Forecasting method: the velocity of all clouds was assumed to be homogeneous. Three different values for the clearness index were obtained for three sky conditions to generate irradiance forecasts	5 min, 10 min, and 15 min ahead forecasts were obtained
[24]	(i) Sky images Camera: IP security camera which has a 180° Resolution: 1024 × 1024 pixels Frequency: every 10 s Image format: jpeg	Cloud segmentation: machine learning model developed using pixel color components such as hue, saturation, R, G, and B values of each pixel, RBR, RBD, pixel distance from the sun, and the zenith and azimuth angles of the sun Tracking method: dense optical flow algorithm Forecasting method: according to motion vectors, future sun-occluding paths were constructed. Then, the timing and extent of sun shading events were predicted	The timing and extent of sun shading events were predicted

Table 1: Continued.

Ref.	Data	Method used	Outcomes
[29]	(i) Sky images Camera: wide-angle C-mount camera Resolution: 2592 × 1944 pixels Frequency: every 5 s (ii) Irradiance measurements	Cloud segmentation: RBR method was used Tracking method: Thirion's Demons algorithm Forecasting method: motion velocity was extracted using a dense vector field of cloud displacement vectors. Occlusions were determined using cloud velocities, and the clear sky index was used to improve short-term forecasts, below 3 min with a Kalman filter	Forecasted continuous irradiance for time intervals of up to 10 min
[31]	(i) Sky images Camera: UCSD sky imager Resolution: 1748 × 1748 Frequency: every 30 s	Cloud segmentation: red–blue-ratio (RBR) method Tracking method: variational optical flow (VOF) technique Forecasting method: the VOF forecasts of the binary sky images were transformed to Cartesian coordinates and generated the VOF-based forecast	Cloud trajectory lengths were forecasted for 1 min to 15 min ahead
[32]	(i) Sky images from 3 cameras Camera: the total sky imager Resolution: 640 × 480 pixels Frequency: every 10 s	Cloud segmentation: a supervised classifier was developed to detect clouds at pixel level Tracking method: cloud block-matching method Forecasting method: using three cameras, onsite CBH was obtained. Regression-based forecasting was done using image features of the clouds with cloud block motion vectors and CBHs	15 min ahead irradiance was predicted
[25]	(i) Sky images Camera: UTSA sky imager Resolution: 1024 × 768 pixels (ii) CBH	Cloud segmentation: RBR method Tracking method: optical flow algorithm to the feature points in two consecutive binary images Forecasting method: using CBH measurements and zenith angles of the sun, the cloud shadow was mapped onto the ground. According to the shadow movement and plant location, irradiance drop was forecasted	5 min ahead, irradiance drops were predicted
[33]	(i) Sky images Camera: the total sky imager Resolution: 640 × 480 pixels Frequency: every 30 s (ii) CBH	Cloud segmentation and cloud type classification: RBR method Tracking method: improved Fourier phase correlation method based on affine transform which is corresponding to image-phase-shift-invariance property was utilized Forecasting method: initially, images were undistorted according to the cloud-based height. Then, the blue-sky area was separated, and the clouds were classified. After classifying the clouds, the sky image-irradiance mapping model was developed. Backpropagation neural network (BPNN) and support vector machine (SVM) are adopted for model training to present sky image-irradiance mapping	1 min to 10 min ahead irradiance was predicted

This paper is structured as follows: Section 2 describes the methodology, which provides a detailed description of the new forecasting model covering the cloud segmentation, motion tracking algorithm, and the cross-correlation-based cloud base height calculation method. Section 3 is a case study developed based on the methodology, and it provides the details of the developed hardware setup and the results. The conclusions are presented in Section 4.

2. Methodology

2.1. Data. Visual measurements of the full sky area with a high spatial and temporal resolution are needed to obtain accurate irradiance forecasts from cloud motion tracking. Therefore, the sky images were captured by a camera with a large FOV, which enables to get most of the clouds that make a shadow on the location of the plant and to track the cloud for a longer time duration. Thus, it increases the forecasting time horizon. Obtaining images at a higher spatial and temporal resolution enables us to accurately track the cloud movement. Figure 1 illustrates how two different FOVs capture clouds. Low FOV (FOV 1) lenses capture a small area of the sky. Thus, it might not capture enough details to forecast the irradiance.

In addition to sky images, to forecast the solar irradiance using the movement of the cloud shadows, (a) the direction

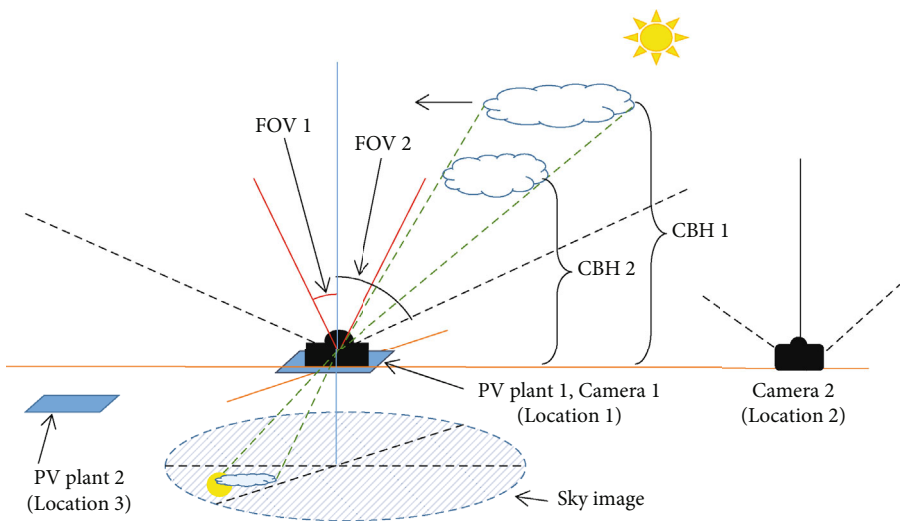


Figure 1: Camera locations, different FOVs of a camera, and how it captures clouds with two CBHs.

of the sky camera with respect to the true north, (b) zenith and azimuth angle of the sun, (c) clear day irradiance profile, and (d) CBH (height above the ground level) are needed. Therefore, the zenith and azimuth angles of the sun and the clear day irradiance profile were calculated based on the information in [19]. For the calculation, the details of time, date, longitude, and latitude of the PV system locations were required.

Typically, the clouds are placed at different layers of the sky. Therefore, the CBH of each cloud is different. The CBH of the clouds above a specific area can be obtained from a ceilometer, but generally, it is expensive. The average CBH values can be obtained from the nearest aviation centers or weather monitoring centers. However, they can introduce significant forecasting errors. Therefore, to forecast the power generation at PV plants located in the neighborhood of the camera location accurately, a local CBH estimation method based on two cameras was introduced. To obtain CBH using the proposed method (to get CBH, clouds captured from camera 1 need to be captured from another location), sky images obtained simultaneously from the second sky camera were considered (camera placed at location 2 as in Figure 1). The irradiance forecasts for several locations were obtained by extracting the image pixels that correspond to the shadow cast on specific PV locations using local CBHs (e.g., location 1 and location 3 in Figure 1).

2.2. Forecasting Methodology. This section provides details of the forecasting methodology developed to forecast irradiance at multiple PV sites. In this method, for one iteration, images taken during one-minute time interval were considered for all 1 min, 5 min, and 15 min irradiance forecasting. Figure 2 shows the flow diagram of the forecasting model. It is divided into nine main sections and is described in detail in this section.

2.2.1. Blue-Sky Area Separation. The correct identification of cloud regions from the sky image is critical as irradiance is forecasted based on the movement of those clouds. A cloud having a large vertical development has a color of a grey shade and creates a substantial drop in the ground level irradiance. Therefore, if a grey cloud is incorrectly identified as a blue-sky area, the result will be significantly erroneous. A blue-sky area separation method was developed to obtain correct cloud regions from the sky images to alleviate this.

In this process, white cloud pixels and pixels of the blue color sky area were separated based on their red and blue component values. However, as discussed in [27], it is not possible to separate pixels related to grey clouds and blue-sky areas only using R, G, or B values. Therefore, to separate only blue-colored pixels from the sky image, $YCbCr$ color space was introduced. The $YCbCr$ color space enables to separate bluish or reddish color components [36] in which Y is the luminance in the $YCbCr$ color plane, Cb is the chrominance dominated by the blue color, and Cr is the chrominance dominated by the red color. Since Cb is strong in places of bluish colors (blue-sky area), it was used with a threshold value to separate the blue-sky area from the sky image. Y, Cb, and Cr components were obtained from RGB pixel values using

$$Y_i = 16 + \frac{65.738}{256} R_i + \frac{129.057}{256} G_i + \frac{25.064}{256} B_i, \quad (1)$$

$$Cb_i = 128 - \frac{37.495}{256} R_i - \frac{74.494}{256} G_i + \frac{112.439}{256} B_i, \quad (2)$$

$$Cr_i = 128 + \frac{112.439}{256} R_i - \frac{94.154}{256} G_i - \frac{18.285}{256} B_i, \quad (3)$$

where Y_i, Cb_i, and Cr_i are Y, Cb, and Cr components of ith pixel and R_i, G_i, B_i are R, G, and B components of the ith pixel.

As the initial step of the blue-sky area separation process, the images captured on sunny, overcast, and partially cloudy sky conditions were manually chosen. From the selected images, pixel indexes related to blue-sky area, white cloud area, and grey cloud areas were extracted manually (equal number of pixels was selected for three categories), and they were labeled. Then, three arrays for three-pixel classes were

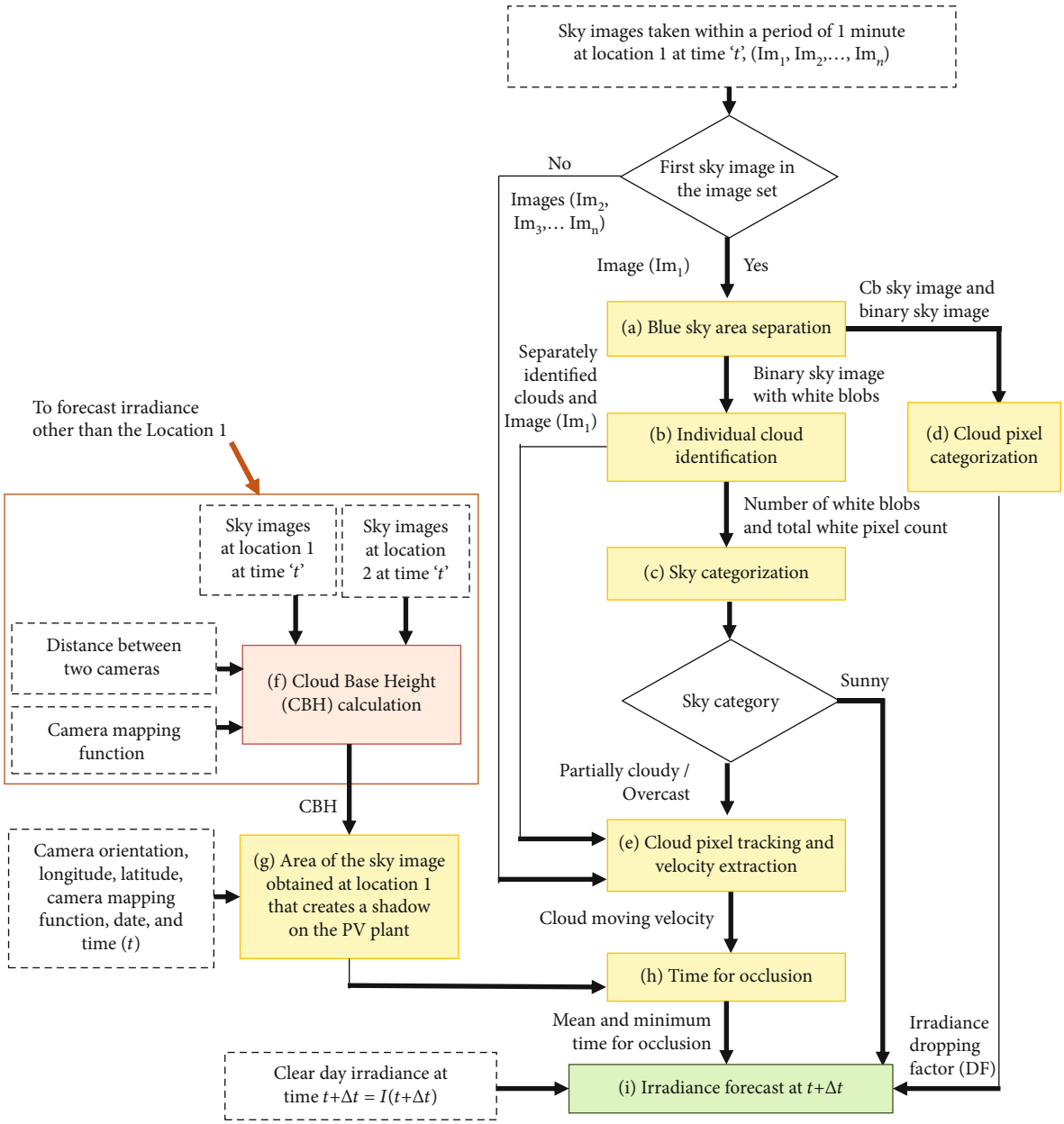


Figure 2: Irradiance forecasting model.

created from the labeled pixels, and the selected images were converted into $YCbCr$ plane images using (1)–(3). After that, Y, Cb, and Cr components of the selected pixel indexes were extracted and placed in the corresponding pixel arrays. Then, each data array was split into two datasets as training dataset and testing dataset. After that, using Y, Cb, and Cr components, a 3D scatter plot was generated from the training dataset. In the scatter plot, three different colors were used to represent the three-pixel categories to understand different clusters. The most dominant property for the classification of pixels related to clouds and the blue area was identified from the scatter plot.

Further, the properties that cannot be used for separation were omitted. Then, considering pixels related to clouds (grey pixels and white pixels) and pixels related to blue-sky area, a histogram was generated for the dominant property (Y or Cb or Cr). After that, the threshold value for the dominant component was found so that the erroneous pixel counts related to both clouds and blue-sky areas are cancelled out. Then, considering the generated scatter plot and the selected threshold value of the dominant pixel separation property, the threshold values for the other parameters were found. Finally, the method was validated using the testing dataset. This cloud pixel identification method was applied to the raw image to generate a binary cloud image where white blobs represent the clouds or the sun and the black area represents the blue-sky region.

The boundaries of the binary image obtained from the thresholding method consist of jagged edges. To smooth out the image boundaries and to remove image noises, image

filters such as median filter, Wiener filter, and statistic filtering functions in Matlab® are widely used [37]. To find out the best filter (less edge distortion filter) for cloud boundary smoothing, three filters were applied to the binary image. The correlation coefficients between the filtered images and the original binary sky image were obtained. Then, the filter corresponding to the highest correlation was chosen for the cloud boundary smoothing.

2.2.2. Individual Cloud Identification. In the blue-sky area separation process, cloud (this may include the sun as well, but it will not be an issue as it does not show any movement) and blue-sky areas were separately identified. In this section, individual clouds were identified separately from the binary sky image generated in the previous section using the connected component algorithm. This is essential to track the individual cloud movement to obtain the cloud moving velocity. Further, using the connected component algorithm, the number of white blobs and the total number of pixel counts in each blob were obtained as they are required under the sky categorization process (explained under Section 2.2.3).

2.2.3. Sky Categorization. The first image in each image group (1 min image set) was categorized into one of the categories: sunny, partially cloudy, or overcast depending on the number of white blobs and on the percentage of white pixels in the binary image. The white pixel percentage was calculated using (4). Then, if the image was classified into partially cloudy or overcast sky conditions, the cloud tracking algorithm was applied to the image set.

$$\text{White Pixel Ratio (WPR)} = \frac{\text{Total number of white pixels}}{\text{Total number of pixels inside the sky area}} \times 100\%. \tag{4}$$

2.2.4. Cloud Pixel Categorization. In the cloud pixel categorization process, pixels in the cloud area were categorized into pixels related to thick clouds, white clouds, and bright white clouds by considering the different grey levels of the cloud pixels in the *Cb* image. For the three cloud pixel categories, irradiance dropping factors were found by comparing the onsite irradiance measurements corresponding to each pixel category with the clear day irradiance measurement. Since the irradiance was forecasted considering the whole cloud, a single irradiance dropping factor was obtained for each cloud. For that, the clouds were categorized as thick clouds, white clouds, and bright white clouds and the irradiance dropping factors of the clouds were assigned based on the pixel category of which the highest number of pixels available in the cloud (mode of the pixel category in the cloud).

2.2.5. Cloud Pixel Tracking and Velocity Extraction. Clouds can be found at different heights in the sky, and depending on the height, they may have different velocities. The cloud velocity provided by weather forecasts usually provides global velocity information. As the accurate prediction of PV drops and shading effects require locally extracted velocities, this section describes a cloud velocity estimation method for individual clouds.

The identified individual clouds (in the first image, Im_1) and the other images in the image set were used as the input data to this process. If the first image was classified into partially cloudy or overcast sky condition, a set of pixel coordinates inside the separately identified cloud regions were selected for tracking using an iterative process such a way that the distances between pixel coordinates in the direction of X or Y have the same pixel difference.

To track the points from one image to the next image, Matlab® normalized cross-correlation function was used [34]. To apply cross-correlation, a template image and a search window were selected. To track a point from the first image to the second image, an $n \times n$ pixel area around the coordinates $[X_1, Y_1]$ on the first image frame was selected as the template image. Following this, an $m \times m$ pixel area ($m > n$) around the coordinates $[X_1, Y_1]$ on the second image frame was selected as the search window. The red component (of the *RGB* image plane) of the template image and search window image segment were considered in the cross-correlation function.

In this process, it was assumed that the image with coordinates $[X_1, Y_1]$ does not move beyond $m \times m$ pixel area over the time interval Δt (time between two images). The maximum correlation points $[X_2, Y_2]$ were selected as the corresponding points for the next image. Since the images were captured at a high rate, the shape of the cloud change is negligible.

The cross-correlation method was applied again to track the points $[X_2, Y_2]$ from the second image frame to the third image frame. Likewise, these steps were repeated for all images and for all clouds. If the image set has six image frames, there are five movement vectors for each selected cloud point from the first image frame to the sixth image frame and they were calculated using the difference between the X and Y coordinates of each point and Δt. Then, the point moving velocity throughout the image set was assigned as the average of the frame-to-frame velocities of that point. After that, histograms of pixel velocity magnitudes and angles were created. Then, the magnitude of the velocity that related to the highest point count in the velocity magnitude histogram and the angle of the velocity that related to the highest point count in the angle histogram were assigned as the velocity vector of the cloud.

After that, the cloud moving velocity was obtained as the median speed and the median direction of the selected points in the cloud.

The above-mentioned cloud motion tracking step is not required if the sky image was categorized as a sunny sky category as in that situation, there will be only one white blob related to the sun.

2.2.6. Cloud Base Height (CBH). Accurate cloud base height (CBH) details are required to forecast irradiance for multiple PV sites located within few kilometers away from the camera location (location 1) because CBH is used to find out the image area that creates a shadow on the PV plant location. When considering the cost and accuracy, the calculation of

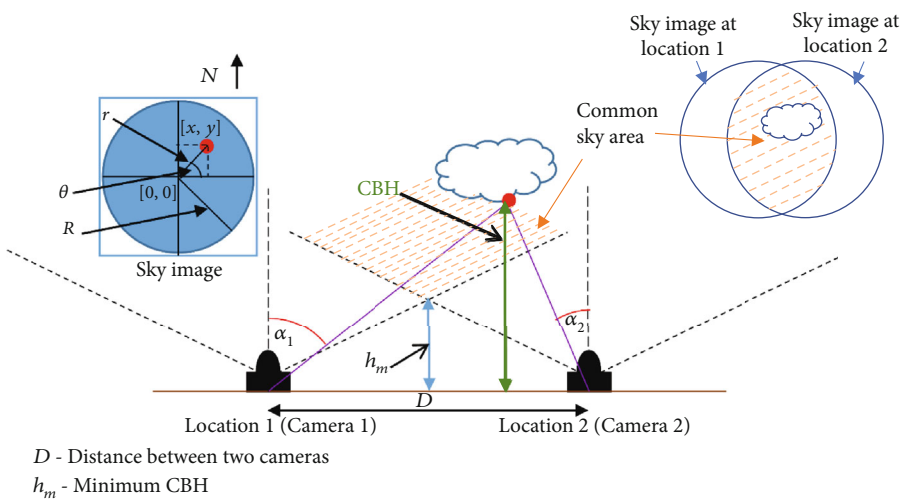


Figure 3: Placement of two cameras for CBH calculation and representation of a sample point in the cloud to the sky image.

the local CBH using sky images [38–40] is the best option; hence, a sky image-based CBH calculation method was used in this paper.

In the sky image-based CBH calculation method, at least two wide-angle cameras are required. The clouds that belong to the overlapping image area (common area as marked in Figure 3) of the two cameras were used to calculate the CBH. The overlapping area of the two images and the minimum value for the CBH (h_m) vary with the FOV and the distance between the two cameras. The clouds which are below h_m cannot be seen in both the images. In other words, the clouds located below h_m cannot be captured by both cameras at the same time. Therefore, it is important to calculate the distance between two cameras considering the average minimum CBH ("h_m") at the PV location and place them accordingly.

The following assumptions were made about the two-camera system when calculating CBH:

(i) Both cameras were placed in a horizontal plane

(ii) The camera height from the base was assumed as zero, and no height difference was considered in the two cameras

(iii) No vertical development of the cloud (the effect of vertical development of the cloud is considered when calculating the irradiance drop percentage)

(iv) Since two identical cameras were used, the equidistant mapping functions of both the cameras were the same

Figure 3 shows the placement of two cameras for CBH calculation and how a sample point in the cloud is represented in the sky image.

To calculate CBH, initially, both images were aligned to the north. Figure 4 shows the flowchart detailing the CBH calculation. Mainly, few cloud boundary points on the binary sky image 1 (from location 1) were selected $[(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)]$. Here, the boundary points of the clouds were found by applying Matlab® "*bwboundaries*" function to the binary cloud image. Considering possible distortions at the boundary of the image obtained from the fisheye lens camera, the cloud boundary points closer to the center of the image were selected (by checking the distance of the boundary point relative to the center point of the image). From the selected points, a number of points were randomly selected via a random function for the calculation of CBH. Then, by assuming different CBH values, the positions of the randomly selected cloud boundary points $[(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)]$ were mapped onto the other sky image captured by the second camera placed at location 2 $[x'_{h,n}, y'_{h,n}]$ using the mapping function of the camera lens. Equations (5)–(9) provide the mathematical equations used to map the cloud boundary points of sky camera 1 to sky camera 2 for different cloud-based heights.

Following the above, image segments around cloud boundary points $[(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)]$ were selected as the template images and the image segments around the points $[(x'_{h,1}, y'_{h,1}), (x'_{2,h}, y'_{h,2}), \cdots, (x'_{h,n}, y'_{h,n})]$ were selected as the search window in CCM. Since the vertical height of the cloud above the base of the cloud was assumed to be relatively small compared to CBH, the appearance of the clouds is assumed to be similar in both images.

$$r_1 = \sqrt{x_n^2 + y_n^2}, \tag{5}$$

$$\alpha_1 = \frac{(\mathrm{FOV}_1/2)}{R_1} \times r_1, \tag{6}$$

$$\alpha_2 = \tan^{-1}\left(\frac{D}{h} - \tan(\alpha_1)\right), \tag{7}$$

$$r_2 = \frac{R_2}{(\mathrm{FOV}_2/2)} \times \alpha_2, \tag{8}$$

$$\left[x'_{h,n}, y'_{h,n}\right] = [r_2 \times \cos(\theta_2), r_2 \times \sin(\theta_2)], \tag{9}$$

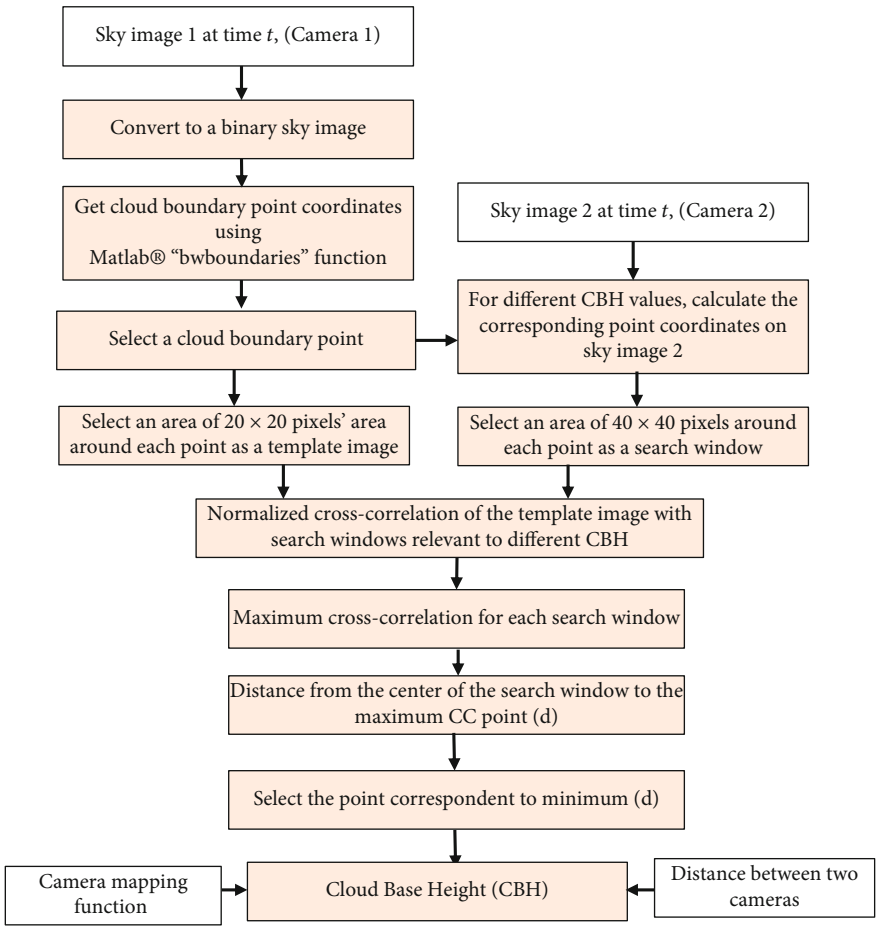


Figure 4: Flow diagram of the CBH calculation method (relevant to Figure 2(f)).

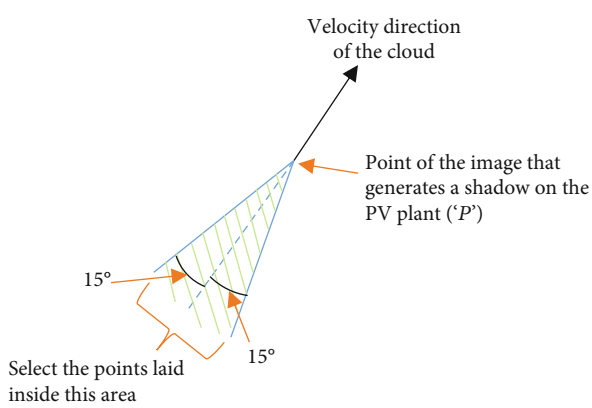


Figure 5: Cloud motion vector.

where "h" is the cloud base heights, $h = [h_m, 400, 600, 800, \cdots, 10000]$; "FOV" is the field of view of the camera; $[x, y]$ are the image coordinates; and "R" is the maximum image radius.

Some boundary points $[x, y]$ have similar maximum cross-correlation values for adjacent CBH values. Due to the fisheye lens distortion, the points corresponding to higher CBHs were mapped closer to each other. Therefore, the corresponding point $[x, y]$ lays inside few searching windows (since the searching windows in the high CBHs overlap) and gives a similar maximum cross-correlation value. The distance between the center points of the search window $[(x'_{hm}, y'_{hm}), (x'_{400}, y'_{400}), (x'_{600}, y'_{600}), \cdots, (x'_{10000}, y'_{10000})]$ and the corresponding maximum cross-correlation points $(c_{hm}, c_{400}, c_{600}, c_{800}, \cdots, c_{10000})$ were compared to obtain an accurate value for CBH. Thus, the CBH corresponding to the correlation point which was located near the center of the search window was selected as the CBH. This method was applied to all other cloud boundary points, and the average value was taken as the CBH.

2.2.7. Area of the Image That Generates a Shadow on the PV Plant. In the onsite irradiance forecasting method, the camera is placed at the PV site. Therefore, if a cloud comes in between the sun and the PV plant, it was identified from the sky image (when the location of the sun on the image was covered by the cloud). The location of the sun on the image was found by using camera orientation, longitude, latitude, camera mapping function, date, and time [41].

In the multiple-site irradiance forecasting, since the PV plants are located few kilometers away from the camera location, the sky image locations which create shadows on the PV

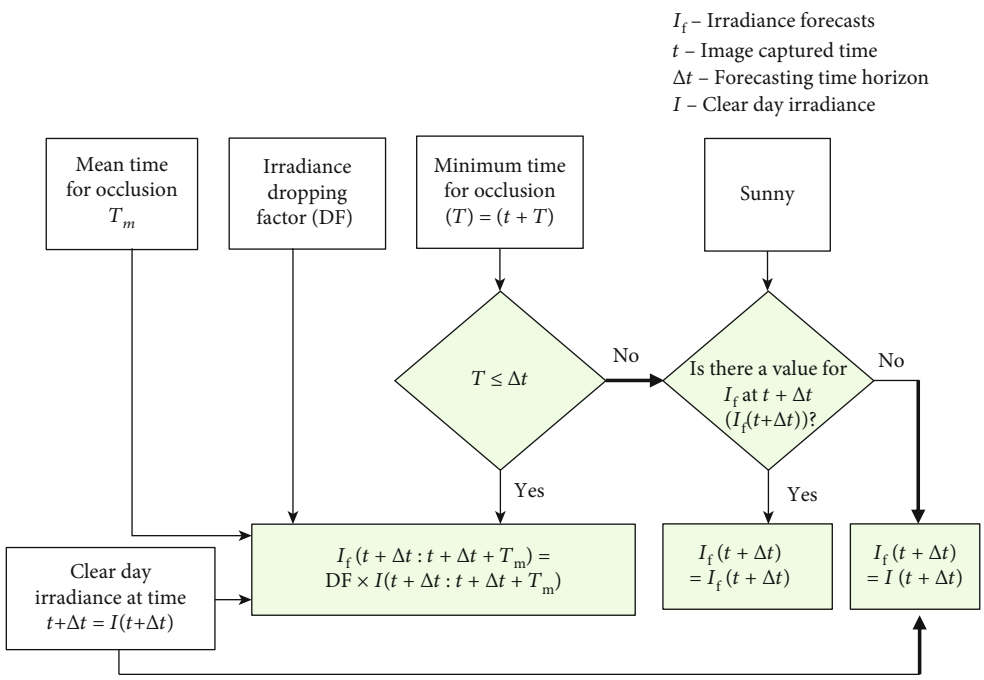


Figure 6: Irradiance forecasting method.

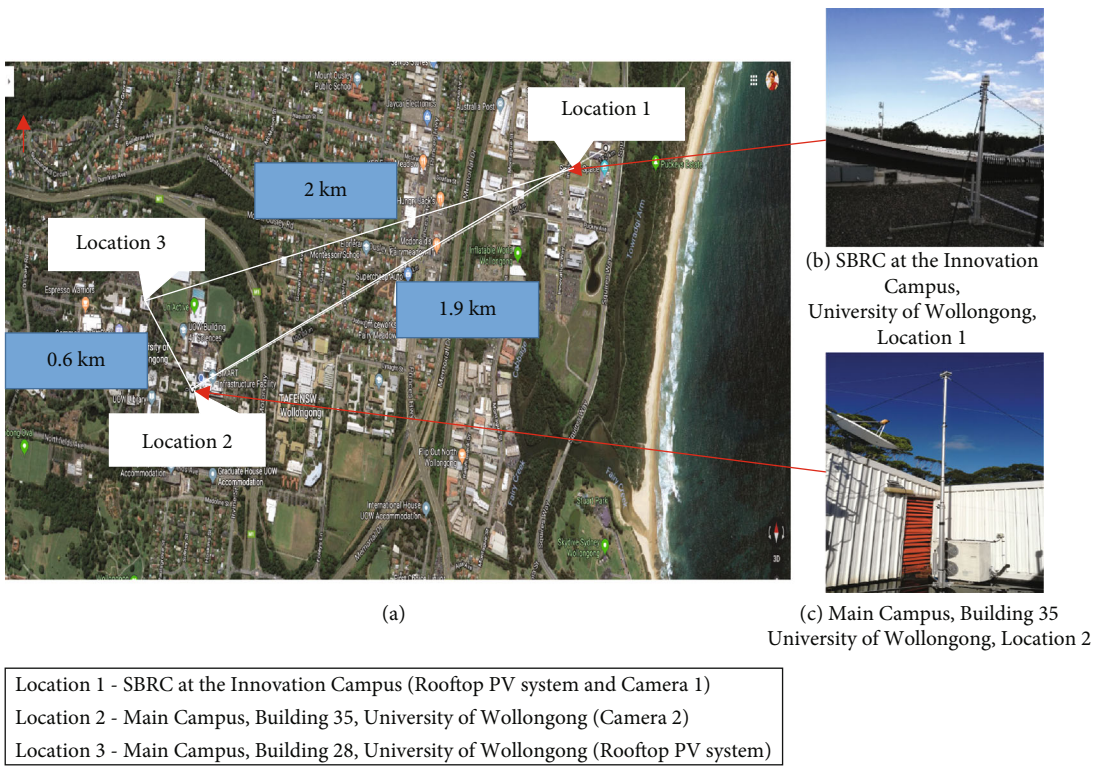


Figure 7: (a) Locations of two cameras and locations of rooftop PV plants, (b) camera 1 at location 1, and (c) camera 2 at location 2.

plants were found according to the calculated CBH, camera orientation, longitude, latitude, camera mapping function, date, and time.

2.2.8. Time for Occlusion. The irradiance drop occurrence time is defined by the speed of the clouds that are directed towards the PV plant. For the onsite irradiance forecasting, it is the time taken by the cloud to reach the point of the sun whereas, in the multiple-site irradiance forecasting, it is the time taken by the cloud to reach the point that makes a shadow on the PV plant. The irradiance drop occurrence time was calculated using the following steps:

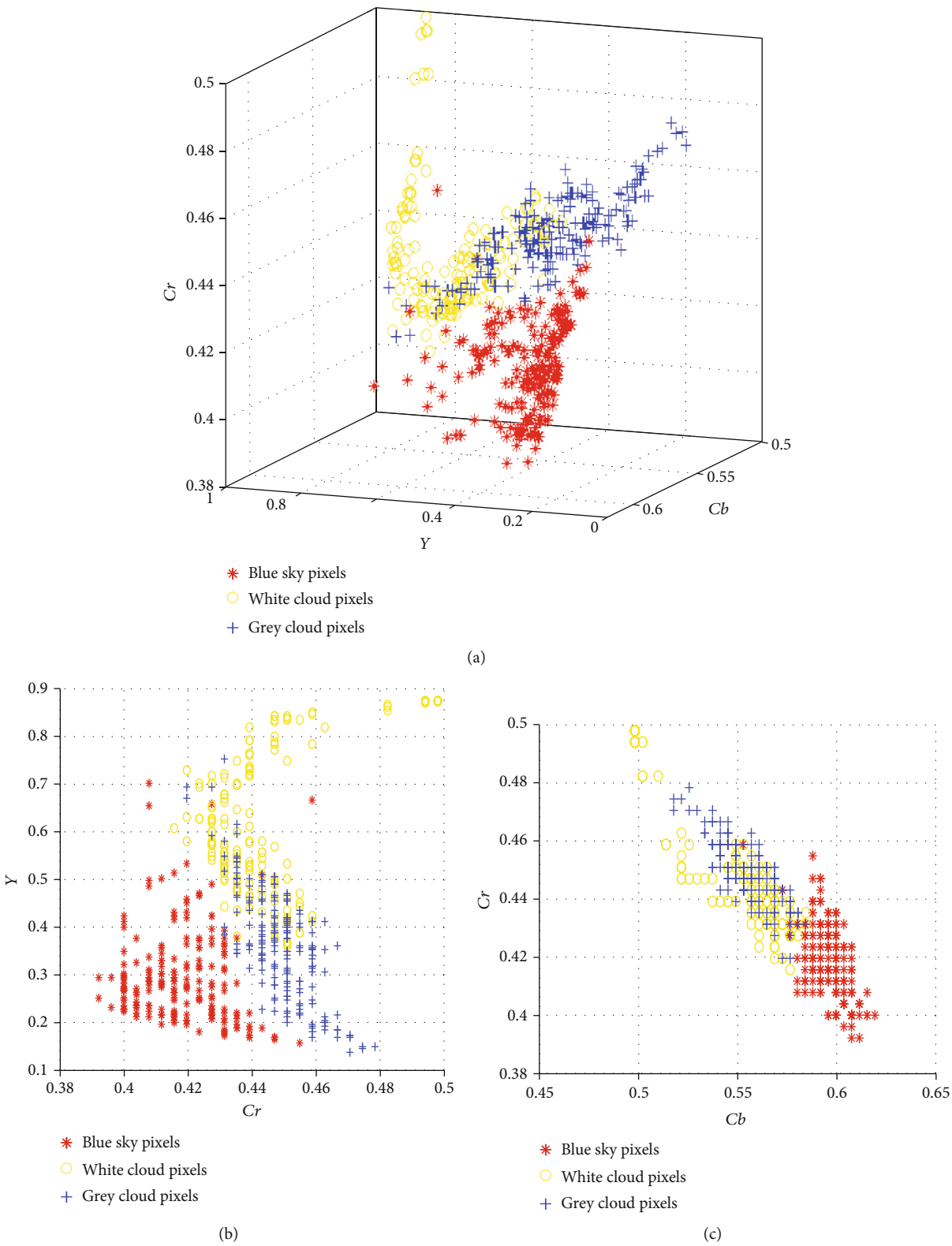
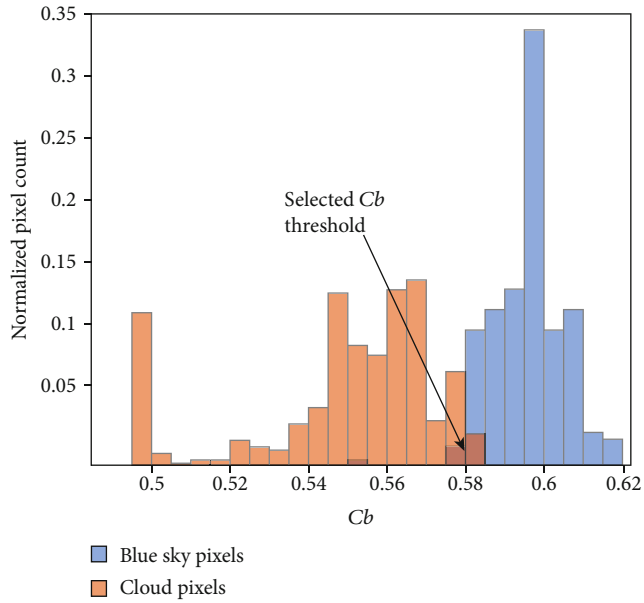
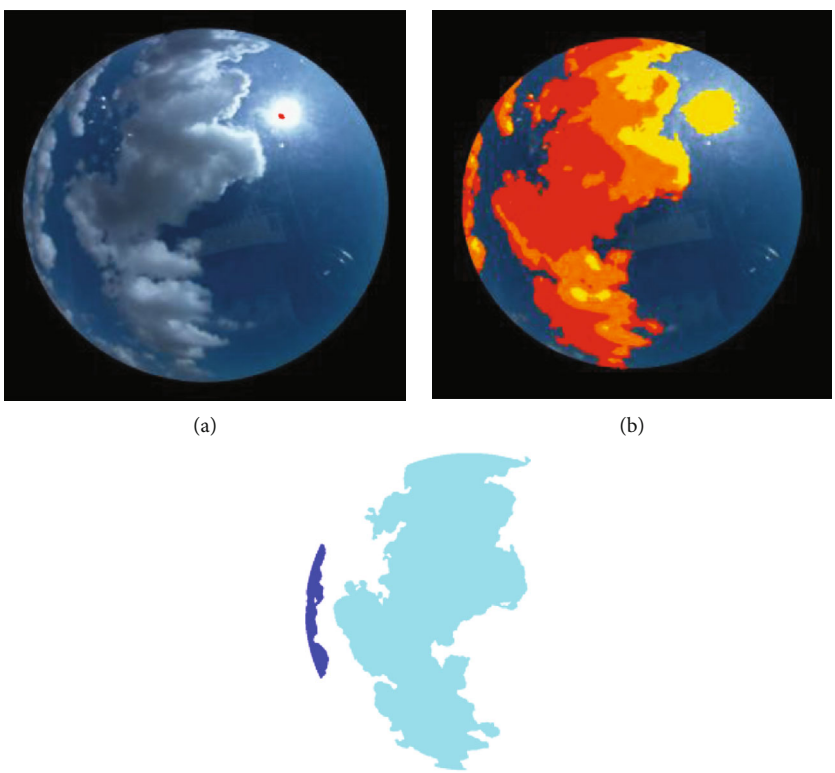


Figure 8: Continued.



(d)

Figure 8: (a) Y, Cb, and Cr scattered plot, (b) Y and Cr scattered plot, (c) Cb and Cr scattered plot for blue-sky pixels and white and grey cloud pixels, and (d) normalized histogram of Cb component of blue-sky pixels and cloud pixels.



(a) (b)

(c)

Figure 9: (a) *RGB* raw image with calculated position of the sun. (b) Identified white patches including both the sun and clouds in the sky image are marked on the *RGB* image. (c) Separately identified cloud regions (area without sun).

Table 2: Sky categorization.

Sunny	Partially cloudy	Overcast
WPR < 15% and only one blob is detected	15% < WPR < 80% and more than one blob are detected	WPR > 80%

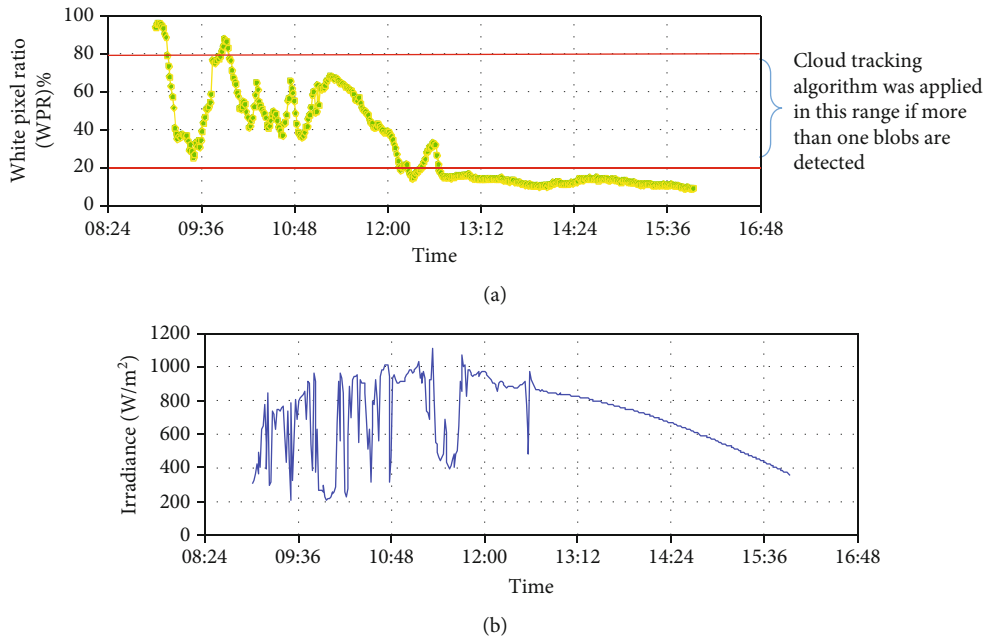


Figure 10: (a) White pixel percentage of the first image in each image set captured on 2018/09/20 and (b) measured irradiance.

(i) The point (in the last image of the image set) that creates a shadow on the PV plant was found (point "P" shown in Figure 5)

(ii) A number of points inside each white blob of the last image frame were selected in such a way that the distances between points in the direction of X or Y are the same

(iii) After that, the backward mapping function of the camera was applied for the points

(iv) Then, the points with a possibility of resulting in a shadow on the PV plant in the expected forecasting time period were found via the motion vector of the cloud and point "P." It is done by identifying the points inside an area covered by two lines drawn with an angle of +15° and -15° relative to the motion vector starting from the point "P" as indicated in Figure 5

(v) According to the calculated velocity of the cloud under Section 2.2.5, the time taken by the selected points to pass the location "P" was calculated

(vi) Then, according to the calculated time for each point to pass the location "P," the minimum (T) and mean (T_m) occlusion values were obtained

If the location "P" is not covered by a cloud at the beginning of the forecasting, the minimum occlusion time is the starting time for occlusion of the cloud. If it is already covered by a cloud at the beginning of the forecasting, the minimum occlusion time is the time taken by the nearest selected point to the point "P."

2.2.9. Irradiance Forecasting. According to the minimum and mean occlusion times (T and T_m) obtained under Section 2.2.8, the irradiance drop was predicted for 1 minute, 5 minutes, and 15 minutes ahead of real-time, as in the flow diagram shown in Figure 6. For the sunny sky condition, since there is no occlusion time, the irradiance is the same as of clear day irradiance (obtained from the ASHRAE model [35]). If the minimum time for occlusion is less than or equal to the forecasting time horizon, there is an irradiance drop, and it was calculated by multiplying the clear day irradiance profile with the corresponding irradiance drop percentage factor obtained for different cloud types (thick grey clouds, white clouds, and bright white clouds). Furthermore, if the minimum time is greater than the considered forecasting time horizon, there will not be an irradiance drop. Hence, the irradiance forecast is equal to the clear day irradiance value. As an example, if the forecasting time horizon is 5 min and the minimum time for occlusion is less than or equal to 5 min, irradiance drop will occur, and if it is greater than 5 min, irradiance drop will not occur.

2.3. Error Metrics and Forecast Performance. Root mean square error (RMSE), mean absolute error (MAE), and true

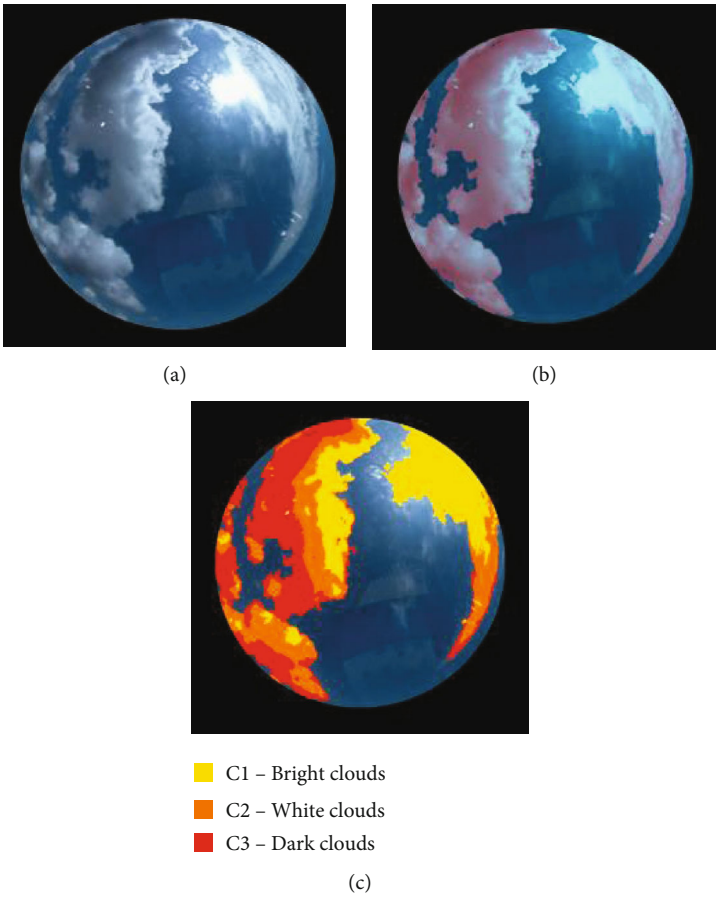


Figure 11: (a) Raw *RGB* image, (b) identified white and grey areas are marked on the raw *RGB* image, and (c) classified cloud pixels from the cloud categorization method.

drop identified percentage were calculated to evaluate the results [29, 42].

The RMSE and MAE were calculated using the predicted irradiance ($I_f(t)$) and measured irradiance ($I_m(t)$) as given by

$$\mathrm{RMSE}_o = \sqrt{\frac{1}{T}\sum_t \left(I_m(t) - I_f(t)\right)^2}, \qquad \mathrm{MAE} = \mathrm{mean}\left(\left|I_m - I_f\right|\right). \tag{10}$$

The percentage of the accurately identified irradiance state (drop or not/1 or 0) from the forecasting model was calculated using (11). The "true state" in (11) is the total number of correctly identified irradiance states, and the "false state" is the total number of incorrectly identified irradiance states (incorrect drops and missed drops) [43].

$$\mathrm{TS} = \frac{\text{True states}}{\text{True states} + \text{Fales states}} \times 100\%. \tag{11}$$

The skill factor indicates the performance of the short-term forecasting models with respect to the persistent model. The forecast accuracy depends on weather conditions and forecasts temporal and spatial resolution. Therefore, forecast accuracies are not comparable site-by-site or hour-by-hour unless normalized by a benchmark. The forecast skill is a way to normalize forecast accuracy [42]. Therefore, the skill factor was calculated for the forecasting results.

The persistence method was defined as the measured irradiance at a time "$t - \delta$" equals to the irradiance at a time "t" (where δ is the forecasting time horizon). The skill factor was calculated using (12), where RMSE_p is the root mean square error of the persistence method and RMSE_c is the root mean square error of the proposed method. SF was calculated for each forecasting time horizon.

$$\mathrm{SF} = 1 - \frac{\mathrm{RMSE}_o}{\mathrm{RMSE}_p}. \tag{12}$$

3. Case Study

3.1. Experimental Setup. Two inexpensive cameras were developed to evaluate the forecasting methodology. To form the sky camera system, a Raspberry Pi single-board computer (3rd Generation, Model B) (US$35), a micro-SD card with a capacity of 32 GB (US$15), and a programmable high-

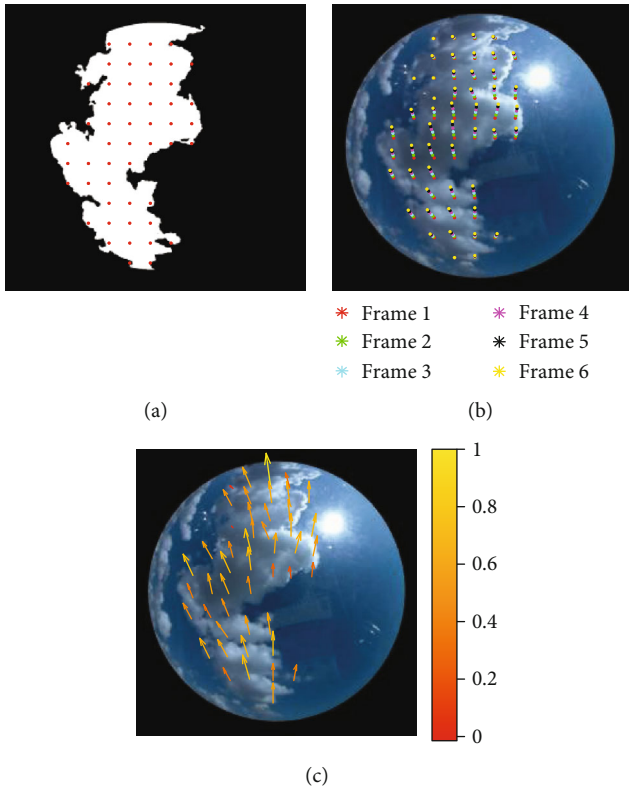


Figure 12: (a) Selected points inside a cloud, (b) tracked points over a 1 minute, and (c) mean of the motion vectors of the selected points.

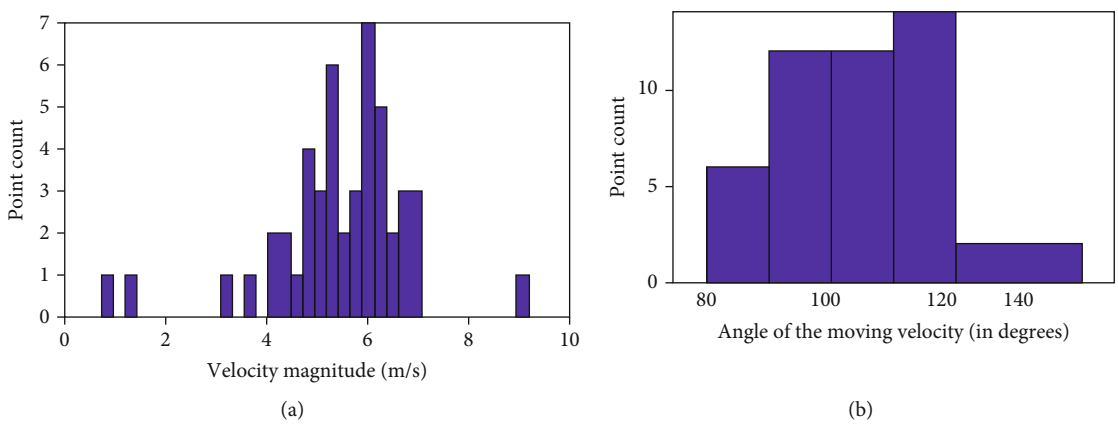


Figure 13: (a) Velocity magnitude histogram and (b) velocity angle histogram.

resolution Pi camera (US$20) with a fisheye lens (US$20) were used.

The installed sky camera systems are shown in Figure 7. The encapsulated Raspberry Pi (RPI) board, together with a programmable high-resolution Raspberry Pi camera module with a wide-angle lens, enables to grab a vast area of the sky onto the image. The Raspberry Pi camera and the RPI board were placed inside a weatherproof enclosure. Both sky camera systems were operated remotely and were programmed to automatically capture images from 8.00 am to 4.45 pm at a rate of 10 seconds. The resolution of the captured images was 1024×768 pixels, and they were stored in jpeg format.

One camera was installed closer to a rooftop PV plant (at location 1), and the second camera was placed 1.9 km away from the first camera (location 2), as shown in Figure 7(a), to collect data. Location 1 is at the Sustainable Buildings Research Centre (SBRC) at the Innovation Campus, University of Wollongong, Australia, while location 2 camera is at the roof of Building 35 at the Main Campus of the University of Wollongong, Australia.

The irradiance levels were forecasted for location 1 and location 3 using the images captured from camera 1 at location 1. The camera fixed at location 2 is only used to obtain the images for CBH. PV data were collected from location 1 to test the onsite forecasting model. To test the multiple-site forecasting model, location 3 (Building 28 rooftop PV system at the main campus, University of Wollongong), which is located with a distance of 2 km from camera 1

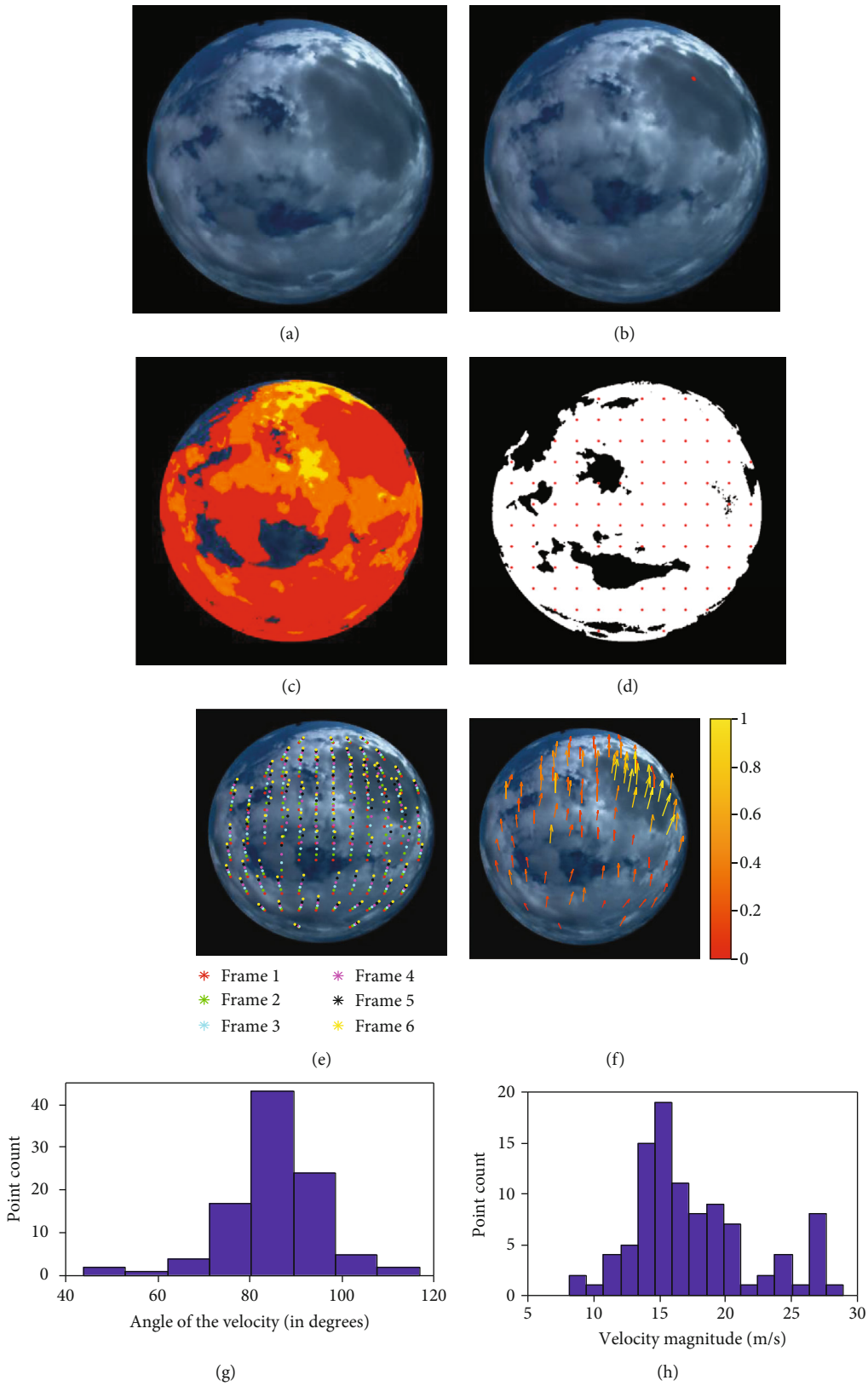


Figure 14: (a) Raw image, (b) location of the sun, (c) separately identified cloud blob related to the sun, (d) selected cloud points inside the selected cloud blob, (e) tracked cloud points over a 1-minute time interval, (f) mean value of the motion vectors of the selected cloud points, (g) velocity angle histogram, and (h) velocity magnitude histogram.

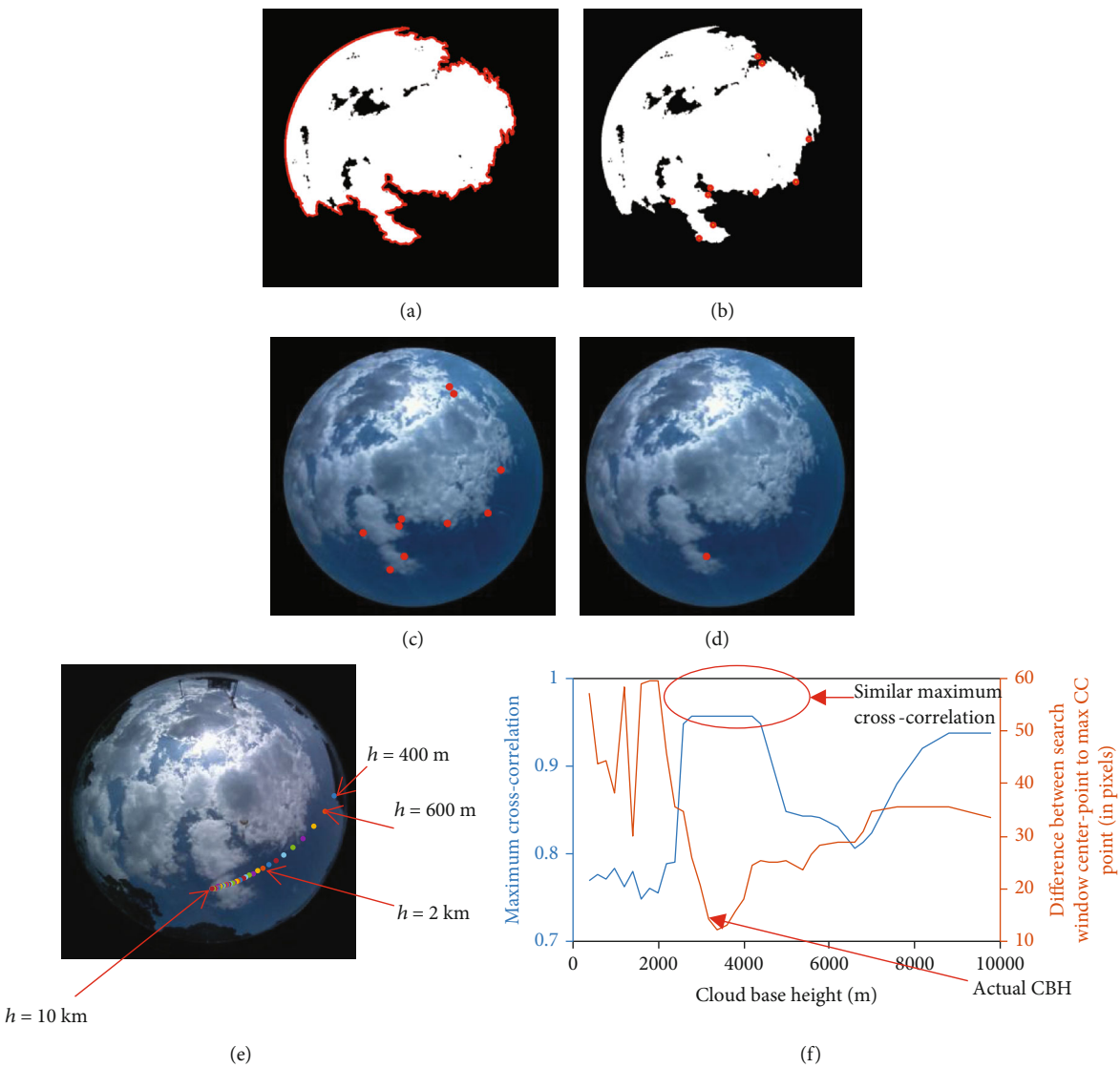


Figure 15: (a) Cloud boundary points extracted from "bwboundaries" function, (b) selected 10 points closer to the center, (c) selected boundary locations on the image captured from the camera placed at location 1, (d) a selected point $[x, y]$, (e) mapped points for the selected point $[x, y]$ for different CBHs on the image taken from the camera placed at location 2, and (f) the maximum CC received for different CBHs (blue) and the pixel difference between the center point of the search window to the maximum CC point for CBHs (brown).

(location 1), was selected as shown in Figure 7(a) (location 1 (34.40°N, 150.90°E) and location 3 (34.41°N, 150.88°E)). Power measurements were acquired from both locations 1 and 3 with a 1-minute resolution to evaluate the forecasting accuracy.

3.2. Irradiance Forecasting Methodology Implementation. This section describes how the irradiance forecasting results were obtained via the experimental setup described in Section 3.1 and the methodology described in Section 2.2.

3.2.1. Blue-Sky Area Separation. As the initial step of irradiance forecasting, the blue-sky area separation process was carried out. The separation of blue pixels (noncloudy pixels) from the sky image was considered in this step and was done using the *YCbCr* threshold method. The training and testing processes were carried out using 50 sky images selected related to the three sky conditions (sunny, overcast, and partially cloudy). For that, a set of six-hundred-pixel indexes related to the blue area, white clouds, and grey clouds were selected from the above-selected sky image set. Then, the corresponding *Y*, *Cb*, and *Cr* components of the selected pixel indexes were extracted, and a scatter plot was generated. Figure 8(a) shows the scatter plot generated via the corresponding *Y*, *Cb* , and *Cr* components of the selected pixel indexes.

As in Figure 8(c), two main clusters were identified, and it is seen that a threshold of *Cb* component separates the pixels related to clouds and pixels related to the blue area of the sky image. Therefore, the *Cb* component was selected as the dominant component that was used to separate bluish pixels from the image. As shown in Figure 8(b), *Y* component cannot be used for separation; therefore, it was omitted. The normalized histogram was generated to determine the *Cb* threshold value for cloud separation, as shown in

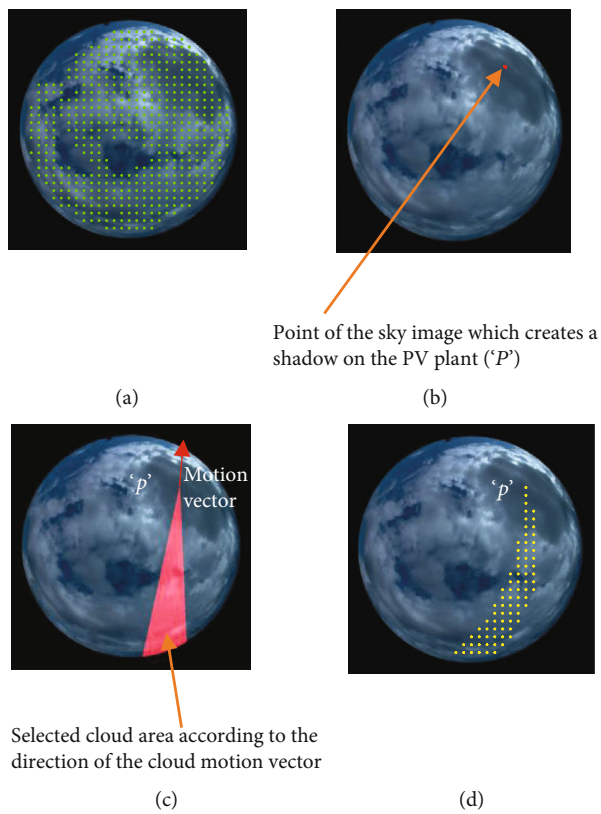


Figure 16: (a) Selected cloud pixels, (b) the image location (*P*) which creates a shadow on PV plant at location 1, (c) selected cloud area according to the direction of the cloud motion vector, and (d) the selected distortion removed points (yellow).

Figure 8(d) (as in Figure 8(c), since *Cb* and *Cr* components related to grey and white cloud pixels are in the same range, both grey and white cloud pixels were considered as the total cloud pixels). The classification decision boundary was adjusted to minimize the classification error, and the optimal location as per histograms illustrated was found to be 0.58. Hence, the thresholding was done based on the actual probabilistic model generated via a manual labeling process. Then, using Figure 8(c) and the selected *Cb* threshold value (0.58), a threshold value for *Cr* was obtained as 0.46 in such a way that all pixels related to the blue area in the sky image are filtered out. For example, if the *Cb* value of a pixel is greater than 0.58 and the *Cr* value is less than 0.46, it was taken as a pixel related to the blue area in the sky image (non-cloud pixel). Then, using threshold values obtained for both *Cb* and *Cr* components, binary cloud images were generated.

A dataset of 100 pixels (25 white cloud pixels, 25 grey cloud pixels, and 50 pixels related to the blue-sky area) was tested to validate the method. The accuracy of identification of the correct pixel category was around 96%.

As discussed in Section 2.2.1, three image filters, median filter, Wiener filter, and statistic filter, were used as image noise filters and as boundary smoothing techniques. After applying three filters to the binary image, the correlation between the filtered image and the binary image was found. The correlation coefficients obtained from the median filtered image, Wiener filtered image, and statistic filtered image were 0.996, 0.882, and 0.993, respectively. Therefore, the median filter was selected as the cloud boundary smoothing technique since it smooths the cloud boundary and maintains the shape of the clouds than the other two filters.

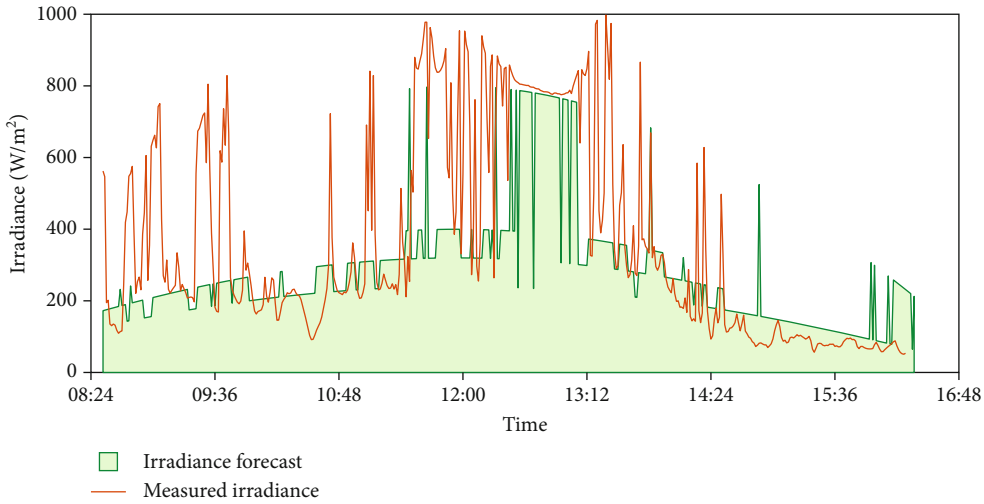
3.2.2. Individual Cloud Identification. The next step of the method is to identify the individual cloud areas on the sky image. Figure 9(a) shows the *RGB* raw image with the calculated position of the sun, Figure 9(b) shows the identified white patches including both the sun and clouds in the sky image, and Figure 9(c) shows the separately identified cloud regions (area without the sun).

3.2.3. Sky Categorization. Then, the sky image was categorized into three sky categories as defined in Table 2 according to the white pixel percentage (WPR) of the binary image.

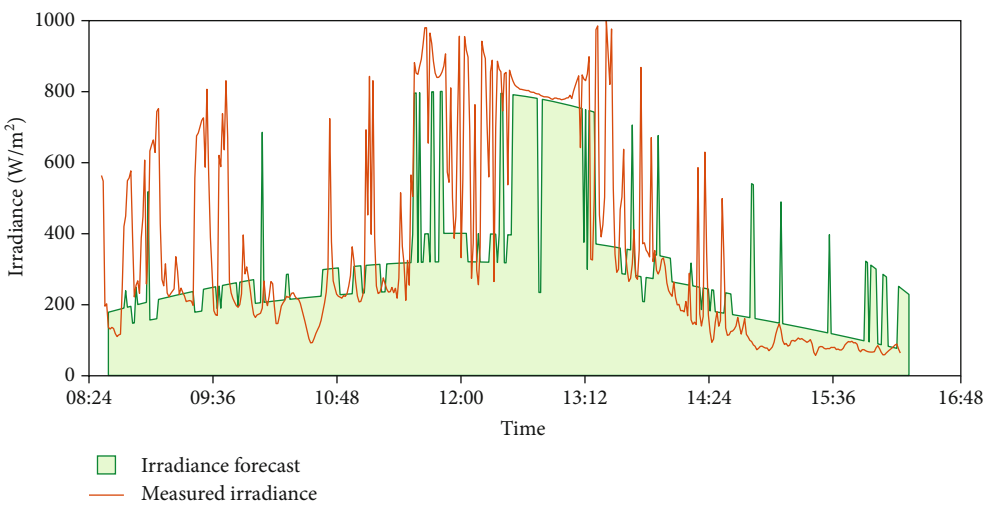
Figure 10 shows the white pixel percentage of the first image in each image set from 8.30 am to 4.45 pm (1 min samples).

3.2.4. Cloud Pixel Categorization. Figures 11(a) and 11(b) show the *RGB* image and the identified cloud areas. By considering the different grey levels of the cloud pixels in the *C b* image, the cloud pixels were categorized into three categories, as shown in Figure 11(c).

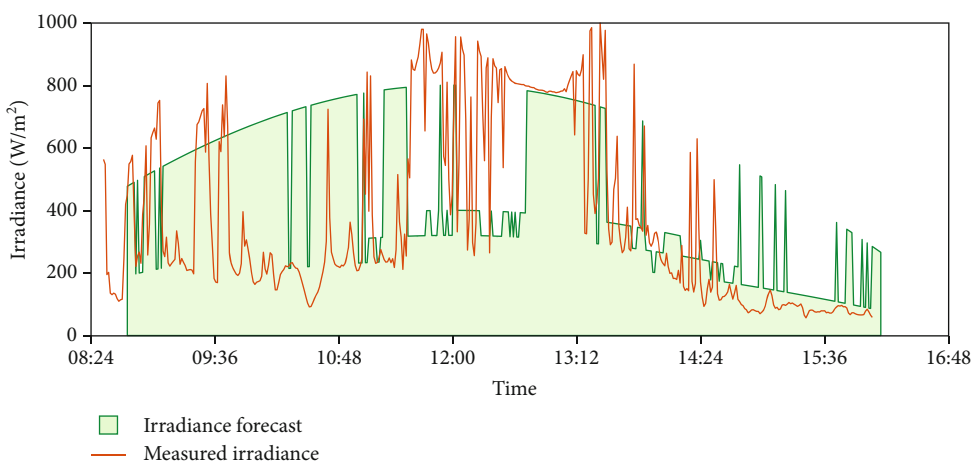
Three irradiance drop levels (C1, C2, and C3) were established for the three cloud groups. For thick clouds (C3), the irradiance was assumed to be 30% of clear day irradiance value, and for white clouds (C2), it was assumed to be 40%.



(a)

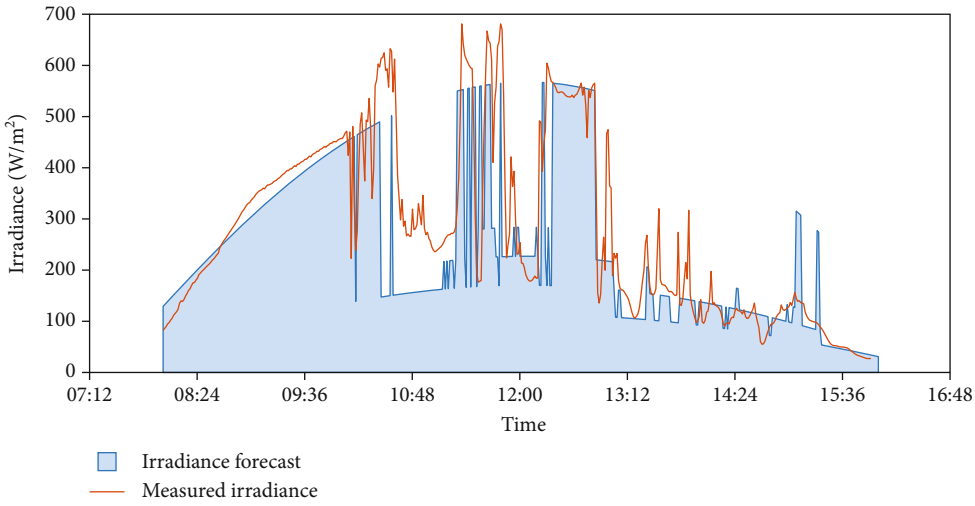


(b)

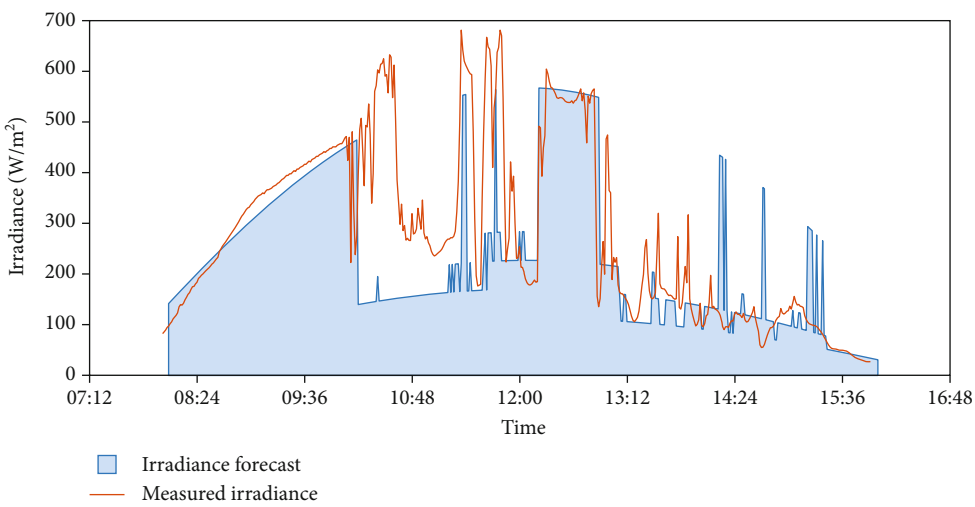


(c)

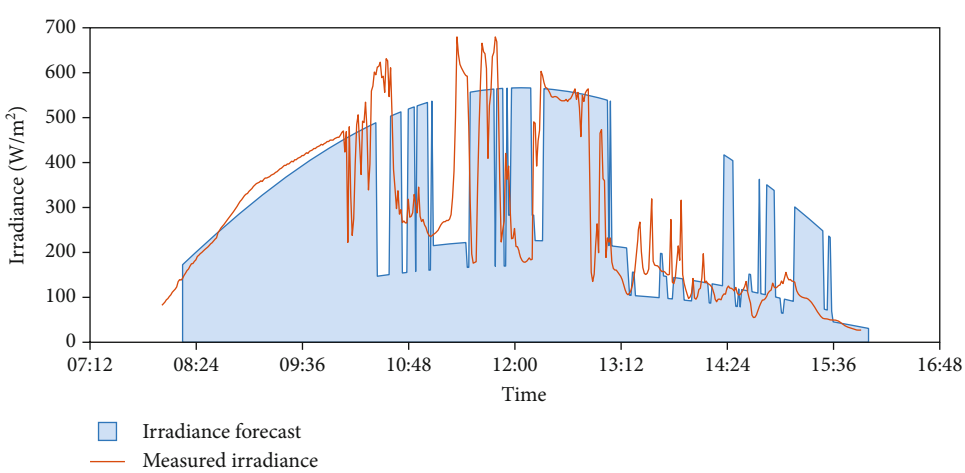
Figure 17: Continued.



(d)



(e)



(f)

Figure 17: (a) 1 minute, (b) 5 minutes, and (c) 15 minutes forecasts at location 1 on 9th April 2019 and (d) 1 minute, (e) 5 minutes, and (f) 15 minutes forecasts at location 3 on 29th July.

Table 3: RMSE, MAE, and TS for irradiance forecasts and irradiance forecasts at location 1 (onsite forecasting) with three-level irradiance drop percentages and with 40% irradiance drop assumption (Sn: sunny; Oc: overcast; Pc: partially cloudy).

Date Sky condition	Time horizon	Forecast with three-level cloud categorization			Forecast with 40% drop assumption [10, 14]			TS/(%)
		RMSE/(Wm^{-2})	MAE/(Wm^{-2})	SF	RMSE/(Wm^{-2})	MAE/(Wm^{-2})	SF	
07/09/2019 (Sn)	1 min	37	29	-0.63	37	29	-0.63	98
	5 min	37	28	-0.02	37	28	-0.02	99
	15 min	36	28	0.41	36	28	0.41	99
08/09/2019 (Sn)	1 min	20	16	-8.6	20	16	-8.6	99
	5 min	20	16	-0.16	20	16	-0.16	100
	15 min	19	15	0.62	19	15	0.62	100
11/09/2019 (Sn)	1 min	13	10	-3.5	13	11	-3.5	98
	5 min	12	10	0.28	12	10	0.28	93
	15 min	12	10	0.76	15	10	0.76	98
13/04/2019 (Pc)	1 min	157	92	-0.20	164	96	-0.25	70
	5 min	213	143	-0.12	229	157	-0.21	62
	15 min	199	120	0.00	202	123	-0.01	61
08/05/2019 (Pc)	1 min	142	92	-0.13	144	92	-0.15	88
	5 min	142	94	0.10	142	94	0.09	87
	15 min	151	101	0.24	151	101	0.23	84
28/07/2019 (Pc)	1 min	136	86	-0.89	144	90	-0.99	79
	5 min	137	88	0.02	148	95	-0.06	78
	15 min	144	89	0.13	146	89	0.12	75
29/07/2019 (Pc)	1 min	113	69	-1.20	118	79	-1.29	80
	5 min	122	73	-0.04	126	81	-0.07	81
	15 min	157	99	0.13	158	104	0.13	62
30/07/2019 (Pc)	1 min	106	51	-1.71	111	50	-1.85	76
	5 min	107	55	-0.53	114	54	-0.63	65
	15 min	80	42	0.19	76	36	0.23	68
10/09/2019 (Pc)	1 min	162	100	-0.87	162	100	0.87	79
	5 min	167	53	-0.27	167	102	-0.27	63
	15 min	130	75	0.11	129	75	0.12	72
09/04/2019 (Oc)	1 min	230	143	-0.64	240	163	-0.71	72
	5 min	224	141	-0.02	228	155	-0.04	74
	15 min	307	230	-0.20	314	243	-0.23	65
05/05/2019 (Oc)	1 min	160	118	-1.01	160	119	-1.01	81
	5 min	206	153	-0.25	206	152	-0.25	66
	15 min	265	220	-0.11	263	218	-0.10	42
12/09/2019 (Oc)	1 min	168	121	-3.6	166	126	-3.55	61
	5 min	188	136	-1.17	183	138	-1.10	63
	15 min	173	119	-0.26	167	122	-0.23	42

On the other hand, for the bright white clouds (C1), the irradiance was assumed as 50% of clear day irradiance.

3.2.5. Cloud Pixel Tracking and Velocity Extraction. After the categorization of the sky condition, the cloud tracking process was performed on each cloud to obtain the cloud velocity. As the first step, few points were selected inside the identified blobs in such a way that the distance between selected points' coordinates is 50 pixels (as in Figure 12(a)). Then, to apply the CCM for each point, the sizes of the template image and the search window were found as follows. The template image size ($n \times n$) was selected to a lower value (n is set to 40 pixels) than that of 50 pixels so that only one point coordinate will lay in the image template. To find the search window size, the below process was considered. According to Reference [44], the average wind speed in the Wollongong area is taken as 19 km/h. The minimum cloud base height considered in our model was 500 m, and the time difference between the captured two image frames is ($\Delta t =$) 10 s. Therefore, the maximum movement of the cloud between two image frames is around 31 pixels. Therefore, the search window size m was set to 70 pixels ($n + 30$).

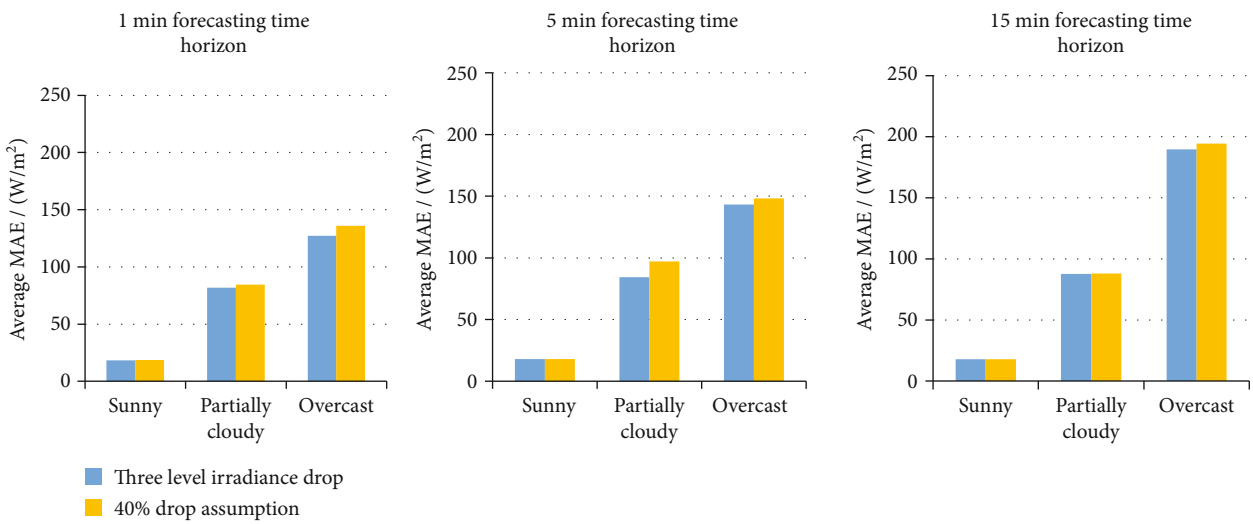


Figure 18: Comparison of MAEs obtained for the three-level cloud categorization and with a single drop level assumption (40% drop).

Figure 12(b) shows the tracked cloud points over six iterations, and Figure 12(c) shows the average motion vectors of the selected points.

Figures 13(a) and 13(b) show the histograms of pixel velocity magnitudes and angles of the selected cloud. The magnitude of the velocity that related to the highest point count (bin count) in the velocity magnitude histogram and the angle of the velocity that related to the highest point count in the angle histogram was assigned as the cloud moving velocity vector.

Figure 14(a) shows the *RGB* raw image corresponding to the overcast sky condition, and Figure 14(b) shows the identified location of the sun (marked in red marker). Separately identified blobs, selected cloud points inside the cloud blob, and tracked cloud points are shown in Figures 14(c)–14(e). Motion vectors of the selected cloud points and the histograms for velocity angles and velocity magnitudes are shown in Figures 14(f)–14(h), respectively.

3.2.6. Cloud Base Height (CBH) Calculation. Then, the cloud base height calculation process was implemented using the images captured from camera 1 and camera 2. According to the FOV (=140/2) of the cameras and the distance (=1.92 km), the minimum CBH ("h_m") was calculated as 350 m.

To get CBHs, initially, the cloud boundary points on images obtained from camera 1 were found using Matlab® "*bwboundaries*" as shown in Figure 15(a). Then, cloud boundary points closer to the center of the image were selected in such a way that the distance of the cloud boundary to the center of the image is less than 250 pixels (the radius of the full image is 366 pixels). Out of those points, 10 points were randomly selected as shown in Figures 15(b) and 15(c) for the calculation of CBH. Figure 15(d) shows a single point that was selected for the illustration. By changing the CBH from 400 m to 10 km with 200 m steps, the coordinates of the selected boundary point $[x, y]$ were mapped on to sky image 2 $[(x'_{400}, y'_{400}), (x'_{600}, y'_{600}), \cdots, (x'_{10000}, y'_{10000})]$ as in Figure 15(e). As seen in Figure 15(f), for the selected point, there are similar maximum cross-correlation outputs for different CBH values. Then, the distance between the center points of the search windows and the corresponding maximum cross-correlation points were mapped in the same plot of Figure 15(f) (brown color) and the minimum distance was assigned as the CBH.

3.2.7. Area of the Sky Image Obtained at Location 1 That Creates a Shadow on the PV Plant, Time for Occlusion, and Irradiance Forecast. The selected points inside the cloud are shown in Figure 16(a), and the image location ("*P*") that creates a shadow on the PV plant at location 1 is shown in Figure 16(b). Then, the cloud area required for the irradiance forecasting was selected according to the direction of the cloud motion vector and the image location ("*P*") as shown in Figure 16(c). Since the images were obtained via a wide-angle lens, to get the corresponding area in the image, the backward mapping function of the camera was applied for the points and obtained as shown in Figure 16(d).

The time for occlusion was obtained according to the cloud moving velocity. Then, the irradiance profile was forecasted by multiplying the clear day irradiance value with the irradiance dropping factor of the cloud category as shown in Figure 17. Figures 17(a)–17(c) show the irradiance profile obtained from the onsite (for location 1) forecasting model for the three forecasting time horizons (1 min, 5 min, and 15 min), and Figures 17(d)–17(f) show the irradiance profile obtained from the multiple-site forecasting model (for PV system at location 3) for the three forecasting time horizons. The irradiance is forecasted from 8.15 am to 3.40 pm with 1 min granularity.

3.3. Performances of the Forecasting Model. Sunny days, overcast days, and partially cloudy days that are randomly selected from 09/04/2019 to 12/09/2019 were used to illustrate the performance of 1 minute, 5 minutes, and 15 minutes ahead irradiance forecasts of two forecasting models: onsite model and multiple-site model.

Table 4: RMSE, MAE, and TS for irradiance forecasts and irradiance forecasts at location 3 (with 40% irradiance drop assumption) (Sn: sunny; Oc: overcast; Pc: partially cloudy) (W: west; E: East; S: south; N: north; NW: northwest; SE: southeast; SW: southwest; NE: northeast).

Date Sky condition Wind direction	Time horizon	The forecasting model introduced in this paper							Forecasting model in [10]			
		Forecast with three-level cloud categorization			Forecast with 40% drop assumption [10, 14]			TS/(%)				
		RMSE/(Wm^{-2})	MAE/(Wm^{-2})	SF	RMSE/(Wm^{-2})	MAE/(Wm^{-2})	SF		RMS/(Wm^{-2})	MAE/(Wm^{-2})	SF	TS/ (%)
07/09/2019	1 min	30	27	-5.67	31	27	-5.68	91	19	15	-3.03	93
(Sn)	5 min	93	40	-2.90	83	38	-2.45	91	18	14	0.27	93
W/W	15 min	31	28	0.56	32	28	0.56	93	17	13	0.76	93
08/09/2019	1 min	18	13	-4.11	18	13	-4.11	92	32	28	-8.31	92
(Sn)	5 min	18	13	0.23	18	13	0.23	91	31	28	-0.41	92
W/W	15 min	20	14	0.74	20	14	0.74	92	31	27	0.52	91
11/09/2019 [45]	1 min	19	14	-3.69	19	15	-3.69	91	46	42	-10.4	95
(Sn)	5 min	19	15	0.17	19	15	0.17	91	43	40	-0.96	90
W, SW/S	15 min	20	15	0.74	20	15	0.74	91	43	40	0.32	90
28/07/2019	1 min	145	84	-1.39	150	88	-1.47	78	147	80	-1.43	73
(Pc)	5 min	110	60	0.25	112	62	0.24	81	152	85	-0.03	71
W, NW/S,SE	15 min	152	84	0.04	152	85	0.04	71	158	83	0.01	73
29/07/2019	1 min	114	63	-1.36	108	65	-1.23	76	182	128	-2.74	57
(Pc)	5 min	139	77	-0.12	131	78	-0.06	73	213	162	-0.71	45
W, NW/W, SE	15 min	157	97	0.08	154	100	0.10	61	226	177	-0.29	38
30/07/2019	1 min	110	62	-3.8	111	54	-3.88	75	88	61	-2.87	72
(Pc)	5 min	73	43	-0.73	67	34	-0.63	66	125	93	-1.98	55
SW/S	15 min	66	42	0.23	56	31	0.34	58	114	82	-0.34	58
10/09/2019 [45]	1 min	167	100	-0.99	168	101	-0.98	72	194	130	-1.3	66
(Pc)	5 min	187	113	-0.43	186	114	-0.43	70	178	124	-0.35	63
W/SW	15 min	118	66	0.25	117	66	0.26	68	176	116	-0.10	68
06/09/2019	1 min	199	146	-3.12	194	145	-3.01	73	204	156	-3.20	62
(Oc)	5 min	194	142	-0.67	190	140	-0.63	68	192	146	-0.64	50
N, NE/E, NE	15 min	173	116	-0.19	167	111	-0.14	59	191	148	-0.28	42
12/09/2019 [45]	1 min	177	127	-3.6	177	135	-3.61	72	254	208	-5.59	59
(Oc)	5 min	202	146	-1.26	196	151	-1.20	70	247	207	-1.76	61
W/W	15 min	171	113	-0.24	166	119	-0.21	33	248	211	-0.77	31

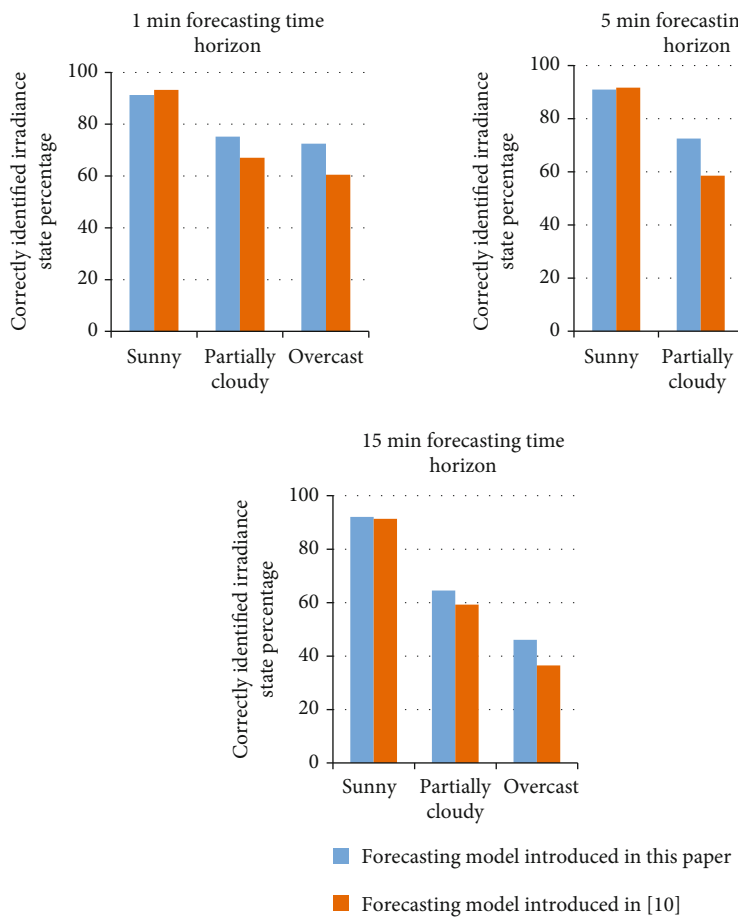


Figure 19: Percentage of correctly identified true irradiance state by the forecasting model introduced in this paper and by the method in [10].

The images captured between 8.30 am and 4.30 pm (local time) (with timestamps) were used for irradiance forecasts. For a single day, the experimental dataset contains 2880×2 images from two sky cameras (which were time-synchronized) and PV power measurements obtained from the two locations. The irradiance is forecasted from 8.30 am to 4.30 pm with 1 min granularity. Therefore, for each day, 8 hours $\times 60$ min forecasts will be generated (a total of 480 data points were forecasted).

Only the data obtained from camera 1 was used for the onsite forecasting model, and forecasts were generated for the PV system at location 1. For the multiple-site forecasting model, data obtained from both cameras were used, and camera 2 was used to calculate the CBH only. From this model, the irradiance forecasts were obtained for location 3, which is 2 km away from camera 1.

The measured PV power values obtained from the rooftop solar PV plants were transformed into the irradiance using the capacities of both rooftop PV plants to compare the forecasted irradiance with the measured irradiance.

3.3.1. Onsite Forecasting. RMSE, MAE, skill factor, and the percentage of true irradiance state identification of onsite irradiance forecasting results are given in Table 3.

Further, to compare the effectiveness of applying three different irradiance dropping levels (found according to the cloud color properties) with a constant irradiance dropping factor, irradiance forecasts for 40% drop assumption (as used in [10, 14]), RMSE, MAE, and SF were calculated. As shown in Table 3, errors in the single irradiance drop assumption are higher than that of the forecasts obtained from three-level cloud categorization for partially cloudy days and overcast days. Figure 18 compares the average of the MAE obtained from three-level cloud categorization with the single drop level assumption method for a sunny day, partially cloudy day, and overcast day and shows that it is lower in the three-level cloud categorization.

3.3.2. Multiple-Site Irradiance Forecasting. Table 4 shows RMSE, MAE, SF, and TS of the forecasting results obtained for a location (location 3) 2 km away from camera 1. It compares the individual cloud movement-based tracking model introduced in this paper with the tracking method used in [10], where the entire cloud area is assumed to be moving at a uniform velocity throughout the image (without considering the individual cloud movement).

Figure 19 compares the average percentage of correctly identified true irradiance state (TS) for sunny, partially cloudy, and overcast days from the forecasting model introduced in this paper and from the method in [10] for the three forecasting time horizons. As in Table 4 and Figure 19, when considering the percentage of correctly identification of

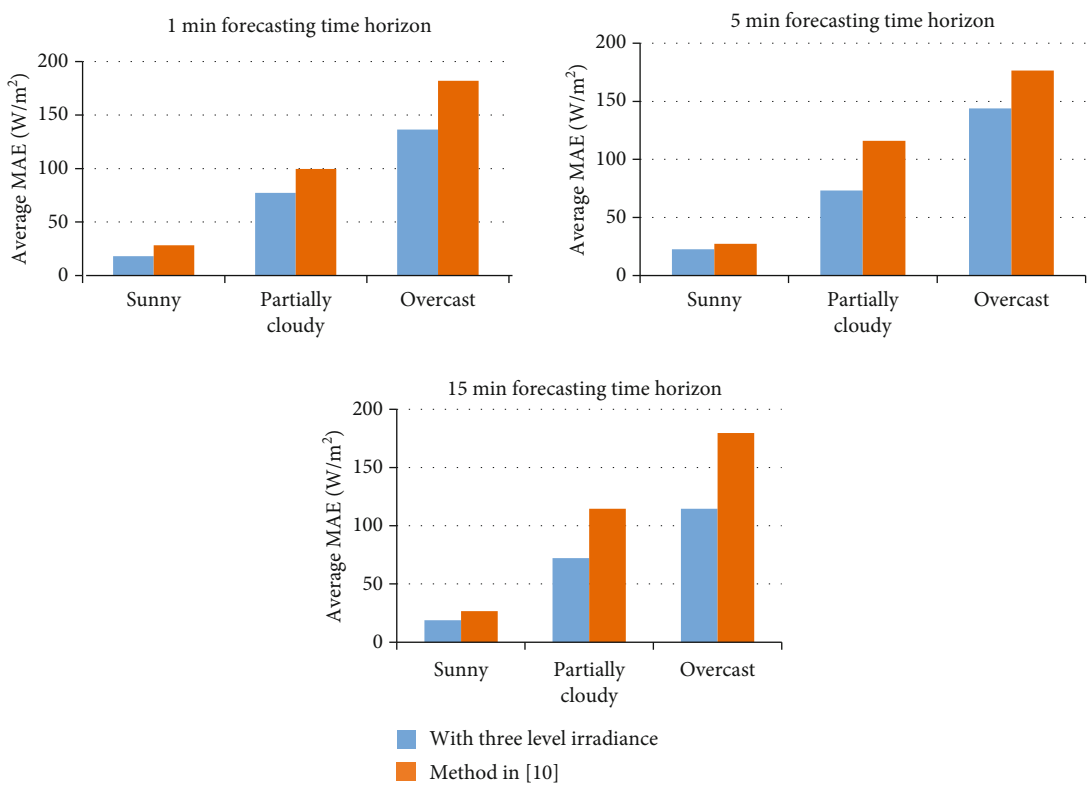


Figure 20: Comparison of average MAEs obtained for the three-level cloud categorization and single drop level assumption (40% drop) and for the forecasting model introduced in [10].

Table 5: Comparison of average RMSE.

Method	Average RMSE (Wm^{-2})
Irradiance mapping methods trained using BPNN [33]	117
Irradiance mapping methods trained using SVM [33]	116
Our cloud motion-based forecasting model	108

irradiance state (TS), the individual cloud movement tracking-based forecasting model introduced in this paper performed better than the method in [10] (which assumed that all clouds move in the same velocity).

Figure 20 shows the average MAEs obtained from the three-level cloud categorization introduced in this paper, single drop level assumption (40% drop), and MAEs for the results obtained from the forecasting model introduced in [10] for three forecasting time horizons. Highest MAE was received from the forecasting model introduced in [10].

When considering the skill factor, the SF remains positive for 15 min forecasting time horizon for location 3 for sunny and partially cloudy days, and this implies the model introduced in this paper outperforms the persistent model.

Reference [33] introduces a solar irradiance forecasting model based on surface irradiance mapping. In that model, initially, the mapping relationship between the information of the cloud pixels and irradiance was established, and then a sky image-irradiance mapping model is developed. When establishing the mapping model, *RGB* values of the circular sky region, which will cover the location of the sun on the image in the next 10 min, were extracted as model input, while the corresponding irradiance was selected as the output. With the sky image—irradiance mapping methods trained using BPNN and SVM, average RMSE values of 117 Wm^{-2} and 116 Wm^{-2} were achieved. In comparison, the method introduced in this paper achieved an average RMSE of 108 Wm^{-2}, as shown in Table 5. Therefore, the forecasting model introduced in this paper performs better than [33].

4. Conclusion

According to the current trend, the percentage share of solar PV in electricity generation will be drastically increased in the coming years. Hence, impacts on power systems due to variations of solar PV power generation need to be addressed. As a solution, in this paper, a novel short-term irradiance forecasting algorithm is presented. The forecasting model presented here makes use of an inexpensive ground-based sky imaging system. The inexpensive camera with a wide-angle lens captures a broad area of the sky; thus, it increases the forecasting time duration. As the shape of the clouds changes with time, a cloud segment tracking method using a cross-correlation algorithm was introduced instead of the optical flow algorithm based on cloud feature point tracking.

Furthermore, the cloud in the different layers in the sky has different moving velocities. Therefore, the proposed individual cloud tracking method allows finding the moving

velocity of each cloud in different layers. Further, the results obtained from the proposed method were compared with a method that assumes all clouds move at an average velocity and proved that the method introduced in this paper performs better.

Irradiance forecasts were obtained for 1 minute, 5 minutes, and 15 minutes using the proposed method. The method has the capability of forecasting the onsite irradiance changes 1 minute in advance with 80% of accuracy and 5 minutes in advance with 78% of accuracy. For the results obtained for 15 minutes, the forecasting time horizon has a positive skill factor indicating that the method introduced in this paper is better than the persistent model.

Since the PV power generation can be directly obtained from the irradiance forecasts, the short-term irradiance forecasts are helpful to overcome the problems caused due to the intermittency of solar PV generation. For example, the forecasting of pending energy shortfalls is useful for the management of PV inverters and energy storage systems. Further, PV forecasts can be used to predict the network status to control and manage the smart inverters and smart transformers. Furthermore, accurate solar irradiance forecasts derived from the proposed method helps to schedule solar PV generation in large interconnected networks. The proposed image-based method shown to perform best for the highly volatile condition in the partially cloudy situation is promising. Further, a hybrid method based on the type of day predicted beforehand may be even more useful. Therefore, the forecasting method proposed in this paper should be further developed considering the above applications while integrating with network models, market models, and intelligent control models.

Abbreviations

NEM:	Australian National Electricity Market
CCM:	Cross-correlation method
CBH:	Cloud base heights
R:	Red
B:	Blue
G:	Green
RBR:	Red to blue ratio
SP:	Sunshine parameter
GHI:	Global horizontal irradiance
RBD:	Difference of each pixel red and blue values
FOV:	Camera field of view
TSIs:	Total sky imagers
USI:	The University of California, San Diego designed and developed a sky imager system
ANN:	Artificial neural network
VOF:	Variational optical flow
UCSD sky imager:	Camera developed by the University of California, San Diego
SBRC:	Sustainable Buildings Research Centre
RPI:	Raspberry Pi
Y:	Image pixel's luma component
Cb:	Blue-difference chroma component
Cr:	Red-difference chroma component
Δt:	The average time between two image frames
WPR:	White pixel ratio
h:	Minimum cloud base height
RMSE:	Root mean square error
MAE:	Mean absolute error
SF:	Skill factor.

Data Availability

The sky images and irradiance data used to support the findings of this study have been deposited in the Mendeley Data repository (doi:10.17632/cb8t8np9z3.2).

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgments

The authors would like to acknowledge the collaboration between the University of Wollongong and the University of Peradeniya as a part of the ongoing Joint Ph.D. program.

References

[1] A. Sayigh, "Renewable energy – the way forward," *Applied Energy*, vol. 64, no. 1-4, pp. 15–30, 1999.

[2] REN21, "Renewables 2019 global status report," August 2019, https://www.ren21.net/gsr-2019.

[3] "Australian Photovoltaic Institute • Market Analyses," December 2020, https://pv-map.apvi.org.au.

[4] "Australian Energy Update 2020|energy.gov.au," December 2020, https://www.energy.gov.au/publications/australian-energy-update-2020.

[5] W. D. Lubitz, "Effect of manual tilt adjustments on incident irradiance on fixed and tracking solar panels," *Applied Energy*, vol. 88, no. 5, pp. 1710–1719, 2011.

[6] H. M. Diagne, P. Lauret, and M. David, "Solar irradiation forecasting: state-of-the-art and proposition for future developments for small-scale insular grids," 2012, June 2016, https://hal.archives-ouvertes.fr/hal-00918150/.

[7] Y. Su, L.-C. Chan, L. Shu, and K.-L. Tsui, "Real-time prediction models for output power and efficiency of grid-connected solar photovoltaic systems," *Applied Energy*, vol. 93, pp. 319–326, 2012.

[8] J. Guo, X. Xu, W. Lian, and H. Zhu, "A new approach for interval forecasting of photovoltaic power based on generalized weather classification," *International Transactions on Electrical Energy*, vol. 29, no. 4, article e2802, 2019.

[9] U. K. Das, K. S. Tey, M. Seyedmahmoudian et al., "Forecasting of photovoltaic power generation and model optimization: a review," *Renewable and Sustainable Energy Reviews*, vol. 81, pp. 912–928, 2018.

[10] C. W. Chow, B. Urquhart, M. Lave et al., "Intra-hour forecasting with a total sky imager at the UC San Diego solar energy testbed," *Solar Energy*, vol. 85, no. 11, pp. 2881–2893, 2011.

[11] W. VanDeventer, E. Jamei, G. S. Thirunavukkarasu et al., "Short-term PV power forecasting using hybrid GASVM technique," *Renewable Energy*, vol. 140, pp. 367–379, 2019.

[12] L. Olatomiwa, S. Mekhilef, S. Shamshirband, K. Mohammadi, D. Petković, and C. Sudheer, "A support vector machine-firefly algorithm-based model for global solar radiation prediction," *Solar Energy*, vol. 115, pp. 632–644, 2015.

[13] M. Ding, L. Wang, and R. Bi, "An ANN-based approach for forecasting the power output of photovoltaic system," *Procedia Environmental Sciences*, vol. 11, pp. 1308–1315, 2011.

[14] K. Stefferud, J. Kleissl, and J. Schoene, "Solar forecasting and variability analyses using sky camera cloud detection & motion vectors," in *2012 IEEE Power and Energy Society General Meeting*, pp. 1–6, San Diego, CA, USA, 2012.

[15] R. B. J. Driesen, "Distributed generation: challenges and possible solutions," Que, 2006.

[16] P. F. Lyons, P. Trichakis, R. Hair, and P. C. Taylor, "Predicting the technical impacts of high levels of small-scale embedded generators on low-voltage networks," *IET Renewable Power Generation*, vol. 2, no. 4, pp. 249–262, 2008.

[17] C. Arrigoni, M. Bigoloni, I. Rochira et al., "Smart distribution management system: evolution of MV grids supervision & control systems," in *2016 AEIT International Annual Conference (AEIT)*, pp. 1–6, Capri, Italy, 2016.

[18] Y. Chaibi, M. Malvoni, A. Chouder, M. Boussetta, and M. Salhi, "Simple and efficient approach to detect and diagnose electrical faults and partial shading in photovoltaic systems," *Energy Conversion and Management*, vol. 196, pp. 330–343, 2019.

[19] M. Karakose, M. Baygin, K. Murat, N. Baygin, and E. Akin, "Fuzzy based reconfiguration method using intelligent partial shadow detection in PV arrays," *International Journal of Computational Intelligence Systems*, vol. 9, no. 2, pp. 202–212, 2016.

[20] F. Spertino, A. Ciocia, P. D. Leo, G. Malgaroli, and A. Russo, "A smart battery management system for photovoltaic plants in households based on raw production forecast," in *Green Energy Advances*, D. Enescu, Ed., IntechOpen, 2018.

[21] "Australian Energy Market Operator," *Australian Energy Market Operator*, 2013, April 2018, http://www.aemo.com.au/.

[22] W. Rich, H. Krishnaswami, R. Vega, and M. Cervantes, "A low cost, edge computing, all-sky imager for cloud tracking and intra-hour irradiance forecasting," *Sustainability*, vol. 9, no. 4, p. 482, 2017.

[23] D. Kumar, H. D. Mathur, S. Bhanot, and R. C. Bansal, "Frequency regulation in islanded microgrid considering stochastic model of wind and PV," *International Transactions on Electrical Energy Systems*, vol. 29, no. 9, article e12049, 2019.

[24] S. R. West, D. Rowe, S. Sayeef, and A. Berry, "Short-term irradiance forecasting using skycams: motivation and development," *Solar Energy*, vol. 110, pp. 188–207, 2014.

[25] M. Cervantes, H. Krishnaswami, W. Richardson, and R. Vega, "Utilization of low cost, sky-imaging technology for irradiance forecasting of distributed solar generation," in *2016 IEEE Green Technologies Conference (GreenTech)*, pp. 142–146, Kansas City, MO, USA, 2016.

[26] Y. Chu, H. T. C. Pedro, and C. F. M. Coimbra, "Hybrid intra-hour DNI forecasts with sky image processing enhanced by stochastic learning," *Solar Energy*, vol. 98, pp. 592–603, 2013.

[27] H. Yang, B. Kurtz, D. Nguyen et al., "Solar irradiance forecasting using a ground-based sky imager developed at UC San Diego," *Solar Energy*, vol. 103, pp. 502–524, 2014.

[28] H. Huang, J. Xu, Z. Peng et al., "Cloud motion estimation for short term solar irradiation prediction," in *2013 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pp. 696–701, Vancouver, BC, Canada, 2013.

[29] D. Bernecker, C. Riess, E. Angelopoulou, and J. Hornegger, "Continuous short-term irradiance forecasts using sky images," *Solar Energy*, vol. 110, pp. 303–315, 2014.

[30] P. Wood-Bradley, J. Zapata, and J. Pye, "Cloud tracking with optical flow for short-term solar forecasting," in *Proceedings of the 50th Conference of the Australian Solar Energy Society*, Melbourne, 2012.

[31] C. W. Chow, S. Belongie, and J. Kleissl, "Cloud motion and stability estimation for intra-hour solar forecasting," *Solar Energy*, vol. 115, pp. 645–655, 2015.

[32] Z. Peng, D. Yu, D. Huang, J. Heiser, S. Yoo, and P. Kalb, "3D cloud detection and tracking system for solar forecast using multiple sky imagers," *Solar Energy*, vol. 118, pp. 496–519, 2015.

[33] F. Wang, Z. Xuan, Z. Zhen et al., "A minutely solar irradiance forecasting method based on real-time sky image- irradiance mapping model," *Energy Conversion and Management*, vol. 220, article 113075, 2020.

[34] J. P. Lewis, "Fast normalized cross-correlation," *Vision Interface*, vol. 10, pp. 120–123, 1995.

[35] M. J. Ahmad and G. N. Tiwari, "Solar radiation models—a review," *International Journal of Energy Research*, vol. 35, no. 4, pp. 271–290, 2011.

[36] P. Ganesan, V. Rajini, B. S. Sathish, V. Kalist, and S. K. Khamar Basha, "Satellite image segmentation based on YCbCr color space," *Indian Journal of Science and Technology*, vol. 8, no. 1, p. 35, 2015.

[37] W. W. Boles, M. Kanefsky, and M. Simaan, "A reduced edge distortion median filtering algorithm for binary images," *Signal Processing*, vol. 21, no. 1, pp. 37–47, 1990.

[38] M. C. Allmen and P. Kegelmeyer Jr., "The computation of cloud base height from paired whole-sky imaging cameras," *Machine Vision and Applications*, vol. 9, no. 4, pp. 160–165, 1997.

[39] E. Kassianov, C. N. Long, and J. Christy, "Cloud-base-height estimation from paired ground-based hemispherical observations," *Journal of Applied Meteorology*, vol. 44, no. 8, pp. 1221–1233, 2005.

[40] D. A. Nguyen and J. Kleissl, "Stereographic methods for cloud base height determination using two sky imagers," *Solar Energy*, vol. 107, pp. 495–509, 2014.

[41] D. M. L. H. Dissawa, M. P. B. Ekanayake, G. M. R. I. Godaliyadda, J. B. Ekanayake, and A. P. Agalgaonkar, "Cloud motion tracking for short-term on-site cloud coverage prediction," in *2017 Seventeenth International Conference on Advances in ICT for Emerging Regions (ICTer)*, pp. 1–6, Colombo, Sri Lanka, 2017.

[42] J. Kleissl, *Solar Energy Forecasting and Resource Assessment*, Academic Press, 2013.

[43] L. H. Dissawa, G. M. R. I. Godaliyadda, M. P. B. Ekanayake, J. B. Ekanayake, and A. P. Agalgaonkar, *Sky Image Based Short-term Solar Power Forecasting Model*, Vidulka National Energy Symposium, Sri Lanka, 2019.

[44] "Climate Wollongong," *meteoblue*July 2020, https://www.meteoblue.com/en/weather/historyclimate/climatemodelled/wollongong_australia_2171507.

[45] L. Dissawa, D. Robinson, A. Agalgaonkar et al., "Sky images and PV power measurements for irradiance forecasting," *Mendeley Data*, vol. 2, 2021.