OPEN ACCESS
Saudi Journal of Engineering and Technology
Abbreviated Key Title: Saudi J Eng Technol
ISSN 2415-6272 (Print) |ISSN 2415-6264 (Online)
Scholars Middle East Publishers, Dubai, United Arab Emirates
Journal homepage: https://saudijournals.com

Original Research Article

Statistical Analysis of Solar Power Generation Patterns and Capacity Utilization: A Time-Series Study Using SPSS

Jawaher Abdulla Alshamsi[1]*

[1]Department of Electrical Engineering

DOI: https://doi.org/10.36348/sjet.2026.v11i05.002 | **Received:** 04.03.2026 | **Accepted:** 30.04.2026 | **Published:** 05.05.2026

***Corresponding author:** Jawaher Abdulla Alshamsi
Department of Electrical Engineering

Abstract

The current paper is a full statistical study of solar power generation and capacity utilisation using a huge sample (161,864 half-hourly observations) of the photovoltaic grid supply infrastructure in the United Kingdom. The primary goal is to analyse the dynamics in the production of solar energy and compare the effectiveness of using capacity under seasonal and diurnal variations. The information, acquired with the help of the Kaggle open-data platform, represents the data about the actual generation in megawatts (MW), the actual capacity, lower and upper confidence limits of the generation forecast, and the inferred data, such as the percentage of the capacity utilisation, and the range of the prediction interval. The statistical tests on descriptive statistics, Pearson correlation analysis, one-way analysis of variance (ANOVA) with a post hoc Tukey HSD, and multiple linear regression were done using IBM SPSS Statistics. The results show that the mean solar output varies significantly across the seasons, with the highest mean (M = 1959.49 MW, SD = 2362.23) and lowest (M = 492.23 MW, SD = 1057.79) in summer and winter, respectively. Diurnal analysis indicated the afternoon hours showed the highest generation (M = 2866.20 MW), and low generation in the night (M = 11.17 MW). The overall capacity utilisation was just 10.04, and this implies that the installed photovoltaic infrastructure was not well used. The outcomes present some practical information that may be utilised by grid operators, renewable energy planners and energy policymakers to utilise solar energy optimally in the national power systems.

Keywords: Solar Power Generation, Capacity Utilization, SPSS, ANOVA, Pearson Correlation, Renewable Energy, Time-Series Analysis, Photovoltaic Systems, Seasonal Variation, Diurnal Patterns.

1. INTRODUCTION

The trend towards renewable sources of energy in the world has become an imperative in response to the increase in climate change, the exhaustion of fossil fuels and the energy demand. Solar photovoltaic (PV) systems are one of the fastest-growing and most widely used technologies in the world today for generating energy. Solar energy is also free and clean, and thanks to its cost-effectiveness, which is increasing, solar energy is abundant and has become the focus of government, researchers, and industry stakeholders (Ahmadi *et al.*, 2018). Solar PV systems have been growing exponentially in capacity within the past two decades, with global capacity reaching the hundreds of gigawatts. However, the intermittency and stochasticity of solar radiation are the only causes of certain grid integration and energy planning problems (Notton *et al.*, 2018).

It is the knowledge of the temporal patterns that the solar energy resources are controlled by that makes the efficient exploitation of solar energy resources possible. These patterns are influenced by a combination of astronomical, meteorological, and geographical factors, including the level of sun irradiance, the presence of clouds, the atmospheric conditions, and the direction and tilt of PV panels (Kalogirou, 2023). The fact that the length of the day varies, and the sun elevation varies, leads to great differences in the output of the generation in different months and seasons, but the diurnal cycles introduce predictable but variable differences in each day. These patterns are vital to proper forecasting, sound grid management, and optimal capacity planning due to the knowledge of these patterns (Sharadga *et al.*, 2020).

The previous research has conducted a comprehensive investigation into the solar energy generation pattern from varying perspectives. The literature review of power system planning models by Emmanuel *et al*. (2020) was aimed at assessing the flexibility of a high level of penetration of solar energy, and it is worth mentioning that powerful statistical models are necessary to characterise the variability of the solar output. The photovoltaic solar energy generating unit inventory developed by Kruitwagen *et al*., (2021) has a global scale that provides the most extensive spatial coverage of solar installations ever. The article by Su *et al*., (2017) compared the trends in the utilisation of solar energy in the context of the different typologies of districts, which was carried out through multi-objective optimisation, demonstrating the importance of considering the context. Moreover, Mahjabeen and Mohammad (2023) studied the application of artificial intelligence in the photovoltaic system, in which intelligent analysis tools would be handy in improving the management of solar energy.

Despite these advancements, there is still a need to have a simple statistical analysis that can be simply replicated and understood by energy practitioners and policymakers without necessarily possessing intricate computational abilities. Machine learning and deep learning approaches have shown potential in predicting solar energy (Akhter *et al*., 2019; Shih *et al*., 2019; Lai *et al*., 2018), yet classical statistical approaches, such as ANOVA, correlation analysis, and regression modelling, can also deliver a lot, with increased transparency and interpretability. Furthermore, the capacity utilisation aspect of the solar PV systems or the percentage of the actual generation to the installed capacity, is a little-researched parameter that can offer crucial data regarding the effectiveness of the systems and their adequacy in planning (Victoria *et al*., 2021).

The UK is quite an interesting case study where solar energy can be studied. The UK has also registered astronomical solar PV installation in spite of its relatively greater latitude and cloud cover, which has also led to its emergence as one of the largest solar energy markets in Europe. The UK grid supply point (GSP) monitoring system provides extremely fine temporal data of solar generation, which enables a highly detailed statistical investigation of the generation pattern. This work takes advantage of this rich data to conduct a rigorous statistical analysis with the assistance of IBM SPSS Statistics, which is an easily accessible statistical package which enables the reproduction of the studies in any research setting.

The study goals of the research are tri-fold: 1) to present the descriptive statistical properties of the variables of solar power generation and capacity utilization in the UK grid; 2) to identify and estimate the value of seasonal and diurnal variations in the output of solar generation by analysing the variables of the ANOVA; 3) to build a regression model, which takes into account the predictive relationships between the variables and the output By fulfilling these objectives, this research will contribute to the constantly changing body of knowledge about renewable energy optimization, and will aid in making evidence-based decisions in solar energy planning (Osman *et al*., 2023).

2. EXPERIMENTAL SECTION

2.1 Data Source and Description

The study uses a series of solar power generation data of the United Kingdom photovoltaic grid supply infrastructure on the Kaggle open-data site. The initial data consisted of 161,864 observations per half hour that documented comprehensive generation readings of the whole solar fleet of the UK. The following were the main variables that were covered in every observation: actual solar power generation in megawatts (Generation_MW), lower confidence limit of the generation forecast (LCL_MW), upper confidence limit of the generation forecast (UCL_MW), installed photovoltaic capacity in megawatt-peak (Installed_Capacity), and effective capacity (Capacity_MWP). The data has sufficient data to examine both seasonal and diurnal patterns because the temporal resolution of the data provides a 30-minute granularity (Sengupta *et al*., 2024).

2.2 Data Preparation and Variable Construction

Data was prepped using Python 3.x in order to be input into SPSS. The database field was the datetime, which had time variables e.g. Hour (023), Month (12), Day. Two categorical variables were derived for analysis: Season, coded as 1 = Winter (December–February), 2 = Spring (March–May), 3 = Summer (June–August), and 4 = Autumn (September–November); and TimeOfDay, coded as 1 = Night (00:00–05:59), 2 = Morning (06:00–11:59), 3 = Afternoon (12:00–17:59), and 4 = Evening (18:00–23:59). Two derived continuous variables were calculated, Capacity Utilization (%) which was (Generation_MW / Installed_Capacity) x 100, which is the percentage of installed capacity that is used at the point at which the value is observed and Prediction Interval Width (MW) which is UCL_MW - LCL_MW, the range of the forecast uncertainty. The processed data was then stored as a comma-separated values (CSV) file, formatted as a decimal value to be properly numeric-parsed in SPSS.

2.3 Statistical Analysis Methods

All statistical analyses were done using IBM SPSS Statistics (Version 29). This analytical methodology comprises four approaches. Firstly, all continuous variables (measures of central tendency (mean), the dispersion (standard deviation, minimum, maximum) and the distribution form (skewness, kurtosis) were also processed by descriptive statistics. Second, bivariate correlations between the continuous variables were done using Pearson product-moment correlation coefficients with two-tailed significance levels at 0.01. Third, significant differences between solar generation output were compared by the use of one-way analysis of variance (ANOVA) between seasonal categories and time-of-day categories. A Levene assumption of homogeneity of variances check was done, and Tukey Honestly Significant Difference (HSD) was used post hoc when the omnibus F-test was determined significant. Fourth, the dependent variable was Generation MW and Hour, Month, and Installed Capacity were considered as predictor variables, and the forced entry method was chosen to perform multiple linear regression analysis. Regression diagnostics were the evaluation of collinearity by the Variance Inflation Factor (VIF) and tolerance values, the Durbin-Watson statistic to check autocorrelation, and the evaluation of the residuals using histograms and normal probability plots (Chen *et al.*, 2018).

3. RESULTS AND DISCUSSION

3.1 Descriptive Statistics

The descriptive statistics of the key continuous variables in the data are reflected in Table 1. The mean solar power output was 1270.48 MW (SD = 2001.90) across the 161,864 observations, with a minimum of 0.00 MW on a night and low-irradiance day, and a maximum of 11,012.20 MW on a sunny summer day. The large standard deviation relative to the mean indicates that there is a high degree of variability of solar generation output that characterises solar energy systems and is susceptible to both diurnal and seasonal variations (Notton *et al.*, 2018). The skewness (1.696) of the Generation_MW variable was positive, suggesting that the distribution is skewed towards the right, with the higher values concentrated towards the left, with the tail of higher values on the right. The kurtosis (2.171) value reveals that the distribution is flatter than normal; thus, it is a platykurtic distribution, which is bi-modal because of close-zero values of solar output at night and varying values of output at other times of the day.

Table 1: Descriptive Statistics of Principal Variables (N = 161,864)

Variable	Mean	SD	Min	Max	Skewness	Kurtosis
Generation_MW	1270.48	2001.90	0.00	11012.20	1.696	2.171
Installed_Capacity	12737.47	2206.95	5453.95	16445.20	-0.949	1.287
Capacity_Utilization (%)	10.04	15.61	0.00	75.00	1.610	1.713
Prediction_Interval	26.95	38.80	0.00	706.99	2.226	11.753
Hour	11.50	6.92	0	23	0.000	-1.204
Month	6.41	3.48	1	12	0.028	-1.233

The capacity utilisation averaged only 10.04% (SD = 15.61), ranging from 0% to 75%. This value represents that the solar PV fleet in the UK is currently operating at a tenth of its installed nameplate capacity on average. This finding follows the intermittency of the generation of solar energy and agrees with a reported capacity factor of solar installations in northern latitudes in Europe (Gernaat *et al.*, 2021). The mean of the installed capacity was 12,737.47 MWp(SD=2206.95) with a negative skewness (-0.949), indicating that the distribution of installed capacity is skewed towards the left, and the negative skewness shows the upward trend of the installed capacity as the period of observation occurred. The Prediction Interval variable (with the highest kurtosis of 11.753) was identified as the variable that signifies the uncertainty of the forecast and the average of 26.95 MW (SD = 38.80) with heavy tails and spikes of extreme forecast uncertainty (Willis, 2018).

3.2 Correlation Analysis

Table 2 shows the Pearson correlation table of the main continuous variables. Generation MW and Capacity utilization have the greatest correlation (r = 0.981, p <.001) as they are supposed to, being derivatives of each other. There was also a significant relationship between Generation MW and Prediction Interval (r = 0.838, p <.001): the greater the generation, the greater the uncertainty in the forecast. The implications of this relationship to grid management are that, at the time when solar is high, the uncertainty in the forecast is also high, and more balancing capacity is needed during the high solar hours (Newbery *et al.*, 2018).

Table 2: Pearson Correlation Matrix

Variable	Hour	Month	Gen_MW	Inst_Cap	Cap_Util	Pred_Int
Hour	1.000	.000	.057**	.000	.059**	.063**
Month	.000	1.000	-.043**	.089**	-.050**	-.040**
Generation_MW	.057**	-.043**	1.000	.081**	.981**	.838**
Inst_Capacity	.000	.089**	.081**	1.000	-.023**	.077**
Cap_Utilization	.059**	-.050**	.981**	-.023**	1.000	.823**
Pred_Interval	.063**	-.040**	.838**	.077**	.823**	1.000

*** Correlation is significant at the 0.01 level (2-tailed).*

Hour had weak and statistically significant positive correlations with Generation mW (r = 0.057, p < .001), Capacity Utilisation (r = 0.059, p <.001) and Prediction Interval (r = 0.063, p <.001). The correlation between Hour and generation is positive but weak and reflects that there is some slight tendency towards an increase in generation during later hours, and this trend can be explained by the existence of a maximum in the generation of solar irradiance in the afternoon. The month had weak negative relationships with Generation mW (r = -0.043, p <.001) and Capacity Utilisation (r = -0.050, p <.001) and a weak positive relationship with Installed Capacity (r = 0.089, p <.001). The correlation between the month and generation is negative and may not be intuitive, yet it points to the coding structure, where the later months (autumn and winter) tend to have lower generation. The monthly capacity-positive association is the growth of installed solar capacity during the time of observation (He *et al*. 2020). The correlation coefficients of the temporal variables and generation output are not strong, which demonstrates the nonlinearity of the solar energy production process that cannot be well represented with the simple linear correlation but with the categorical tests, including ANOVA (Sepulveda *et al*., 2018).

3.3 Analysis of Variance

3.3.1 Seasonal Variation in Solar Generation

The differences between solar power generation in the four seasons were tested by one-way ANOVA. A test to examine the homogeneity of the variances reported by Levene found that the assumption was not met, F(3, 161860) = 11987.35, p < .001, which it would otherwise have as in the case of solar generation data, the variance is larger during the seasons of high irradiance. However, at large and approximately equal sample sizes, ANOVA is not sensitive to the violation of this assumption, as it was in the current study (Dowling *et al*., 2020).

ANOVA revealed that season was significant in the production of solar power, F(3, 161860) = 5091.99, p <.001. Table 3 gives the descriptive statistics of the two sets of seasons. Summer exhibited the highest mean generation (M = 1959.49, SD = 2362.23), followed by Spring (M = 1704.00, SD = 2317.50), Autumn (M = 950.48, SD = 1627.59), and Winter (M =492.23, SD = 1057.79). This is consistent with the solar irradiance patterns in the UK, where there are longer days and greater solar elevation angles during summer months, resulting in much more energy gathered by photovoltaic systems (Chel & Kaushik, 2018).

Table 3: Descriptive Statistics for Generation_MW by Season

Season	N	Mean (MW)	SD	Min	Max
Winter (1)	41,854	492.23	1057.79	0.00	8408.20
Spring (2)	40,975	1704.00	2317.50	0.00	11012.20
Summer (3)	39,742	1959.49	2362.23	0.00	10760.70
Autumn (4)	39,293	950.48	1627.59	0.00	9853.38
Total	161,864	1270.48	2001.90	0.00	11012.20

Tukey HSD post hoc tests revealed that all the seasonal differences were found to have significant differences at p <.001. The greatest difference was noted between summer and winter (1467.27 MW, p < .001), then Spring and Winter (1211.77 MW, p < .001). The statistical significance (not as large in magnitude) between summer and spring (p <.001) was also statistically significant, which implies that even adjacent seasons with similar irradiance patterns produce statistically dissimilar values of generation. This was confirmed by the analysis of homogeneous subsets that showed each season is a distinct set and there is no overlap of any seasonal sets (Haegel *et al*., 2019).

3.3.2 Diurnal Variation in Solar Generation

A second one-way ANOVA was employed to determine whether there was a difference in the solar generation between four categories of time of day. Results indicated that the diurnal period effect on generation output was even greater, F(3, 161860) = 34125.90, p = 0.001, and the F-test is much larger than the seasonal ANOVA, indicating that the diurnal variation is more likely to cause the change in the generation outputs than the seasonal one. Table 4 presents the descriptive statistics of the time of day.

Table 4: Descriptive Statistics for Generation_MW by Time of Day

Time of Day	N	Mean (MW)	SD	Min	Max
Night (1)	40,474	11.17	65.53	0.00	947.86
Morning (2)	40,462	2107.92	2144.58	0.00	10740.50
Afternoon (3)	40,465	2866.20	2262.26	0.00	11012.20
Evening (4)	40,463	96.90	314.51	0.00	2671.83
Total	161,864	1270.48	2001.90	0.00	11012.20

The afternoon period (12:00–17:59) recorded the highest mean generation at 2866.20 MW (SD = 2262.26), followed by the morning period (06:00–11:59) at 2107.92 MW (SD = 2144.58). Evening generation (18:00–23:59) dropped substantially to 96.90 MW (SD = 314.51), while nighttime generation (00:00–05:59) was nearly negligible at 11.17 MW (SD = 65.53). The non-zero values of small nighttime generation are supposedly residual values of grid monitoring equipment or the minimum twilight generation during the long summer daylight hours. The tests were conducted to ensure that there was no statistically significant difference between time-of-day categories at p =.001, Tukey HSD posthoc tests were conducted. Afternoon vs. Night exhibited the highest difference in mean (2855.03 MW, p <. 001), with the maximum and minimum solar generation dramatically different. These results are consistent with the fundamental physics of solar energy conversion, according to which the photovoltaic performance directly depends on the incident solar irradiance, which is the greatest in the afternoon when the sun is at the highest point (Carrillo *et al*., 2019; Chen *et al*., 2019).

3.4 Capacity Utilization Analysis

The data on the capacity utilisation by season and time of day are represented in Table 5. The seasonal analysis reveals that Summer has the highest mean capacity utilisation of 15.41, followed by Winter with 13.60, Autumn with 7.37 and Winter with 3.94. These values indicate that there is a huge variation in the efficiency of solar energy harvesting in the UK across seasons, with the utilisation of the summer in the UK nearly four times greater than that of winter. The direct consequence of this observation on the energy storage requirements and grid balancing strategies is that, because of the seasonal generation potential and the energy demand requirement all through the year, there exists the need to have complementary energy sources and storage facilities during the low-utilisation period (Sepulveda *et al*., 2018).

Table 5: Capacity Utilization (%) by Season and Time of Day

Category	Mean (%)	SD	Min	Max
Winter	3.94	8.37	0.00	59.73
Spring	13.60	18.22	0.00	75.00
Summer	15.41	18.24	0.00	71.78
Autumn	7.37	12.52	0.00	64.98
Night	0.09	0.50	0.00	6.58
Morning	16.58	16.54	0.00	72.64
Afternoon	22.70	17.49	0.00	75.00
Evening	0.78	2.50	0.00	21.64

The daylight pattern of solar energy conversion efficiency reaffirms the concentration of effective solar energy conversion. The maximum utilisation was that of the afternoon (22.70), and utilisation at night was almost negligible (0.09). The use was also high at 16.58 even during the morning hours, meaning that there was high solar energy conversion during the daytime hours. The peak instantaneous use of 75.00 per cent was during spring and afternoon hours, which are nearly ideal conditions of high irradiance and maximum availability of the system. The overall utilisation of 10.04% at all times, though, underscores the natural problem of temporal concentration of solar energy and the necessity to create complementary technologies of generation and storage to capture all the potential of installed photovoltaic capacity (He *et al*., 2020).

3.5 Multiple Linear Regression Analysis

A multiple linear regression with three predictors was used to predict solar power generation: Hour, Month, and Installed_Capacity. The statistical significance of the model was high (F(3, 161860) = 675.12, p <.001); the coefficient of determination was low (R2 =.012, Adjusted R2 =.012), indicating that the linear combination of these predictors had a small percentage variance in solar generation output (1.2 per cent). The value of Durbin-Watson was 0.027, which formed a strong positive autocorrelation in the residuals, which is not unexpected considering the time-series nature of the observed data, i.e. half-hourly.

Table 6: Multiple Linear Regression Coefficients

Predictor	B	SE	β	t	Sig.	VIF
(Constant)	281.524	30.818	—	9.135	<.001	—
Hour	16.573	0.714	.057	23.200	<.001	1.000
Month	-29.184	1.426	-.051	-20.469	<.001	1.008
Installed_Capacity	0.077	0.002	.085	34.386	<.001	1.008

Dependent Variable: Generation_MW; $R^2 = .012$; $F(3, 161860) = 675.12$, $p < .001$

Table 6 has the regression coefficients. The three predictors, all of which are statistically significant, are given as p <.001. The most significant predictor was Installed_Capacity (=.085), and each MWp of installed capacity resulted in a 0.077 MW increase in output. Hour was a good predictor (=.057), and an increase of one hour was associated with a 16.573 MW greater generation, the highest of the afternoon generation. Month was a negative predictor (= -.051), with the higher the months elapsed, the lower the MW was, which was likely caused by the decrease in the peak summer generation to lower values in the autumn and winter of the temporal scope of the dataset. The collinearity diagnostics revealed that all predictors had VIF values close to 1.000, and it indicated that the problems of multicollinearity did not exist (Zhang and Chen, 2022).

The low R 2 value is intriguing and not a weakness. It demonstrates that fundamental time-related indicators and capacity measures are insufficient to describe the variability of solar power, and meteorological factors, such as solar irradiance, cloud cover, temperature, and aerosol concentration in the air, are extremely important to model photovoltaic power. These are some of the unexplained variables that are not part of the existing data and are the biggest cause of generation variability. Such a finding is in line with the previous research that weather-dependent variables explain the majority of the variability in solar output, and the temporal variables are just proxy variables and fail to predict well (Akhter *et al*., 2019; Sharadga *et al*., 2020). It is likely that new studies based on real-time meteorological data would achieve much higher explanatory power.

4. CONCLUSION

The paper has provided a detailed statistical study of the design pattern of solar power production and the capacity utilisation of the photovoltaic infrastructure of the United Kingdom photovoltaic grid using 161,864 half-hourly measurements. Those findings of the analyses, made possible with the assistance of IBM SPSS Statistics, gave several valuable insights that have practical implications regarding energy planning and policy.

The descriptive analysis revealed that there was a high variability in solar generation output (M = 1270.48 MW, SD = 2001.90), and overall capacity utilisation was 10.04%. One-way ANOVA tests showed very significant and significant seasonal and diurnal effects on solar production. The seasonal analysis showed that Summer produced nearly 4 times as much (1959.49 vs 492.23 MW) as Winter, and the diurnal analysis showed that Afternoon generation (2866.20 MW) was more than Nighttime generation (11.17 MW) by approximately 257 times. All pairs were significant at $p < .001$, and each category of time was a homogeneous subset. In the Pearson correlation analysis, it was found that the correlations between generation output and capacity utilisation ($r = .981$) and prediction interval width ($r = .838$) and temporal variables were found to be weak, which implies that the meteorological factors dominate in determining the generation output. The regression model, though statistically significant ($p < .001$), only explained 1.2 per cent of the variance, which confirms that temporal variables can not be used as a predictor of solar generation and that weather-related variables should be included in prediction models.

These findings have important implications for renewable energy planning. The average capacity utilisation (10.04) is low, which means that there is much more to tap into the solar energy harvesting by increasing the panel orientation, tracking, and location. Both the seasonal and diurnal variations are high, which explains the need to have complementary energy storage facilities and diversified renewable energy portfolios to ensure that the grid is reliable throughout the year. In the subsequent research, the methodology should be combined with meteorological variables, more advanced models of time series modelling should be employed, and the examination of a variety of geographic regions should be carried out to develop more comprehensive predictive models to streamline solar energy production.

Acknowledgements

I acknowledge that publicly available data on the Kaggle open-data platform has been applied in this study. The analysis was done using IBM SPSS Statistics software.

REFERENCES

- Afroz. (2024). *Solar power generation data* [Data set]. Kaggle. https://www.kaggle.com/datasets/pythonafroz/solar-power-generation-data
- Ahmadi, M. H., Ghazvini, M., Sadeghzadeh, M., Alhuyi Nazari, M., Kumar, R., Naeimi, A., & Ming, T. (2018). Solar power technology for electricity generation: A critical review. Energy Science & Engineering, 6(5), 340–361.
- Akhter, M. N., Mekhilef, S., Mokhlis, H., & Mohamed Shah, N. (2019). Review on forecasting of photovoltaic power generation based on machine learning and metaheuristic techniques. IET Renewable Power Generation, 13(7), 1009–1023.
- Carrillo, A. J., González-Aguilar, J., Romero, M., & Coronado, J. M. (2019). Solar energy on demand: a review on high temperature thermochemical heat storage systems and materials. Chemical Reviews, 119(7), 4777–4816.
- Chel, A., & Kaushik, G. (2018). Renewable energy technologies for sustainable development of energy efficient building. Alexandria Engineering Journal, 57(2), 655–669.
- Chen, C., Kuang, Y., & Hu, L. (2019). Challenges and opportunities for solar evaporation. Joule, 3(3), 683–718.
- Chen, Y., Wang, Y., Kirschen, D., & Zhang, B. (2018). Model-free renewable scenario generation using generative adversarial networks. IEEE Transactions on Power Systems, 33(3), 3265–3275.
- Dowling, J. A., Rinaldi, K. Z., Ruggles, T. H., Davis, S. J., Yuan, M., Tong, F., ... & Caldeira,
- K. (2020). Role of long-duration energy storage in variable renewable electricity systems. Joule, 4(9), 1907–1928.
- Emmanuel, M., Doubleday, K., Cakir, B., Marković, M., & Hodge, B. M. (2020). A review of power system planning and operational models for

flexibility assessment in high solar energy penetration scenarios. Solar Energy, 210, 169–180.
- Gernaat, D. E., de Boer, H. S., Daioglou, V., Yalew, S. G., Müller, C., & van Vuuren, D. P. (2021). Climate change impacts on renewable energy supply. Nature Climate Change, 11(2), 119–125.
- Haegel, N. M., Atwater Jr, H., Barnes, T., Breyer, C., Burrell, A., Chiang, Y. M., ... & Bett, A. W. (2019). Terawatt-scale photovoltaics: Transform global energy. Science, 364(6443), 836–838.
- He, G., Lin, J., Sifuentes, F., Liu, X., Abhyankar, N., & Phadke, A. (2020). Rapid cost decrease of renewables and storage accelerates the decarbonization of China's power system. Nature Communications, 11(1), 2486.
- Kalogirou, S. A. (2023). Solar energy engineering: processes and systems. Elsevier.
- Kruitwagen, L., Story, K. T., Friedrich, J., Byers, L., Skillman, S., & Hepburn, C. (2021). A global inventory of photovoltaic solar energy generating units. Nature, 598(7882), 604–610.
- Lai, G., Chang, W. C., Yang, Y., & Liu, H. (2018, June). Modeling long-and short-term temporal patterns with deep neural networks. In The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval (pp. 95–104).
- Mahjabeen, F., & Mohammad, A. (2023). Revolutionizing solar energy: The impact of artificial intelligence on photovoltaic systems. International Journal of Multidisciplinary Sciences and Arts, 2(3).
- Newbery, D., Pollitt, M. G., Ritz, R. A., & Strielkowski, W. (2018). Market design for a high-renewables European electricity system. Renewable and Sustainable Energy Reviews, 91, 695–707.
- Notton, G., Nivet, M. L., Voyant, C., Paoli, C., Darras, C., Motte, F., & Fouilloy, A. (2018). Intermittent and stochastic character of renewable energy sources: Consequences, cost of intermittence and benefit of forecasting. Renewable and Sustainable Energy Reviews, 87, 96–105.
- Osman, A. I., Chen, L., Yang, M., Msigwa, G., Farghali, M., Fawzy, S., ... & Yap, P. S. (2023). Cost, environmental impact, and resilience of renewable energy under a changing climate: a review. Environmental Chemistry Letters, 21(2), 741–764.
- Sengupta, M., Habte, A., Wilbert, S., Gueymard, C. A., Remund, J., Lorenz, E., ... & Jensen, A. (2024). Best practices handbook for the collection and use of solar resource data for solar energy applications.
- Sepulveda, N. A., Jenkins, J. D., De Sisternes, F. J., & Lester, R. K. (2018). The role of firm low-carbon electricity resources in deep decarbonization of power generation. Joule, 2(11), 2403–2420.
- Sharadga, H., Hajimirza, S., & Balog, R. S. (2020). Time series forecasting of solar power generation for large-scale photovoltaic plants. Renewable Energy, 150, 797–807.
- Shih, S. Y., Sun, F. K., & Lee, H. Y. (2019). Temporal pattern attention for multivariate time series forecasting. Machine Learning, 108(8), 1421–1441.
- Su, S., Lu, H., Zhang, L., Alanne, K., & Yu, Z. (2017). Solar energy utilization patterns for different district typologies using multi-objective optimization: A comparative study in China. Solar Energy, 155, 246–258.
- Victoria, M., Haegel, N., Peters, I. M., Sinton, R., Jäger-Waldau, A., Del Cañizo, C., ... & Smets, A. (2021). Solar photovoltaics is ready to power a sustainable future. Joule, 5(5), 1041–1056.
- Willis, H. L. (2018). Distributed power generation: planning and evaluation. CRC Press.
- Zhang, S., & Chen, W. (2022). Assessing the energy transition in China towards carbon neutrality with a probabilistic framework. Nature Communications, 13(1), 87.