Power Prediction via Module Temperature for Solar Modules Under Soiling Conditions

Salsabeel Shapsough(✉), Rached Dhaouadi, Imran Zualkernan, and Mohannad Takrouri

American University of Sharjah, Sharjah, UAE
sshapsough@aus.edu

Abstract. The ability to predict the output power of remote solar modules is key to successful wide-scale adoption of solar power. However, solar power is a direct product of its environment and can vary vastly from one location to another. Predicting generated power for a specific facility requires monitoring the output of the solar modules in the context of ambient variables such as temperature, humidity, solar irradiance, air dust, and wind. This is especially challenging in areas where soiling is a significant environmental variable. Soiling particles such as sand and dust can shade segments of the solar module, thus effectively reducing the amount of solar irradiance absorbed and, consequently, the power produced. Measuring soiling particles requires expensive equipment that can increase the cost of running the facility and therefore lower the total output. However, dust can also serve as a cooling layer that can reduce the temperature of the solar module and to a certain extent, reduce overheating. This observation can be used to correlate the amount of dust accumulated on the surface of the panel with its temperature. In this work, the module temperature and power output of a clean module and a dusty module are observed using an Internet of Things monitoring system. The data is used to train various machine learning and deep learning algorithms to eventually predict the output of a soiled module over time using only its temperature and a reference clean module.

Keywords: Photovoltaic · Module temperature · Prediction · Machine learning

1 Introduction

Wide-scale deployment of solar power faces significant challenges due to the source's dependence on its environment. Environmental elements such as temperature, shading, wind, and soiling can cause the power yield to deviate significantly from expected values [1]; thus, making it more challenging to adopt the solar power as a reliable energy source. Artificial intelligence is key to overcoming such obstacles by providing tools to build models that make it possible to understand the behavior of solar modules out in the field, as well as generate short-term and long-term predictions to efficiently plan production by matching load to demand [2]. The first step to modeling solar power is understanding factors that have the most influence on modules in a given setting. As the significance of

© ICST Institute for Computer Sciences, Social Informatics and Telecommunications Engineering 2020
Published by Springer Nature Switzerland AG 2020. All Rights Reserved
D.-J. Deng et al. (Eds.): SGIoT 2019, LNICST 324, pp. 85–95, 2020.
https://doi.org/10.1007/978-3-030-49610-4_7

different elements may vary from one location to another, it is critical when modeling a solar facility to take into account its operating conditions as accurately as possible. This is where remote an in-situ monitoring comes in [3]. Building accurate models requires collection of a solar facility's output as well as context over a long period of time. However, since the cost of monitoring instrumentation can increase the cost of large-scale facilities and reduce the overall efficiency of production, the type of data collected should be considered carefully. Monitoring should focus on observing elements that have the most significant impact in that specific setting.

In desert areas, soiling represents the biggest obstacle to solar energy [4]. The presence of soil particles on the surface of the panel can influence its behavior. It is well-known that the accumulation of sand and dust particles on the surface of the panel can reduce the amount of solar irradiance absorbed by the panel, effectively reducing the output power [5]. One way to quantify the level of soil on the surface of a panel is by simulating panels through a glass plate and measuring the optical losses due to the dust accumulated on the plate [6, 7]. However, such studies become less feasible as the facility in question increases in scale as the distribution of the aerosol index (AI) varies spatially and temporally. Alternatively, several studies investigated the approximation of the soiling index using output power relative to meteorological conditions such as aerosol particle size [8], mechanics [9], and composition [10, 11]. While accuracy is relatively high in such works, the cost of the instrumentation required to perform such studies may exceed the benefit.

Most recent and existing work focuses instead on measuring the influence of soil particles. Commonly, the performance of a clean panel with a soiled one and use the loss percentage to calculate a "soiling index" [12] that is indicative of the soil accumulation on a soiled module's surface. The soiling index is then used to predict the power output of soiled modules given the module's temporal and environmental context. While soil's influence as a shading agent has been studied in numerous works, it has not yet been characterized as the complex phenomenon it is. For one, module's surface temperature is a function of solar irradiance absorbed by the module as well as ambient temperature [13]. By shading the module and reducing the amount of solar irradiance absorbed, the accumulation of soil on the surface of a module can interfere with heat transfer between the module and its environment and have a cooling effect [14]. The output of a solar module depends on the temperature of its cells as much as it depends on solar irradiance incidence. Standard Test Conditions (STC) denote temperature and irradiance values at which a module outputs maximum power. In desert areas, module overheating can significantly reduce the output of the module. Cooling due to soil accumulation, however, may have an opposite effect. It is therefore beneficial to investigate the correlation between soil's influence on module temperature and output power.

This work aims to investigate the relationship between soiling and module temperature, and their consequent influence on output power. First, a low-cost monitoring system using Internet of Thing technologies to remotely monitor the performance and context of solar module. The data collected over the period of ten weeks is then used to build various machine learning models to predict the output of a soiled module using only its temperature and a reference clean module.

2 Hardware Setup

The test setup consisted of two identical 100 W monocrystalline solar panels, shown in Fig. 1 (a). One was cleaned every week, while the other was left to gather dust and other soiling elements for ten weeks. The panels' performance and context were measured using an IoT-based monitoring system discussed in [15]. Performance was measured by performing IV tracing once every hour to extract the maximum power point, short circuit current, and open circuit voltage of each panel. The context, on the other hand, included the ambient temperature and solar irradiance. The module temperature was measured using twenty low-cost one-wire temperature sensors distributed across each module. While the sensors do not measure every single cell temperature, the twenty cells that were measure were selected as to cover each region of the module. The layout of the temperature sensors on the back of each panel is shown in Fig. 1 (b).

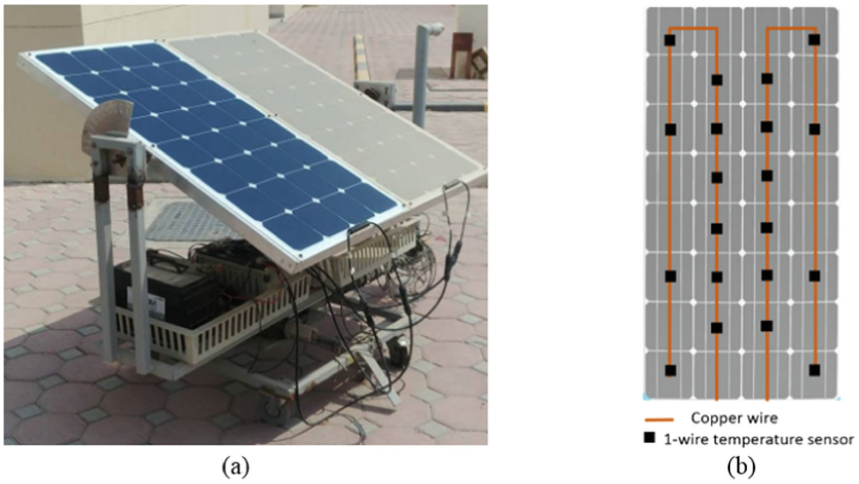

Fig. 1. (a) Test setup (b) one-wire temperature sensors layout

3 Data Analysis

At the end of the testing period, the dataset consisted of over 800 entries denoting the cell temperatures and power output of each panel, solar irradiance and ambient temperature readings, timestamp, and number of days since the beginning of the experiment. In the first phase of analysis, the average module temperature was calculated by averaging the individual cell temperatures. The behavior of individual cell temperatures will be explored in future works.

3.1 Soiling and Module Cooling

First, Fig. 2 shows the general trend in ambient temperature as well as the two modules' temperatures using all the readings observed during the test period, as well as with filtering out solar noon values. Solar noon refers to the time of the day where irradiance is at its highest value. It usually occurs between 12:00 pm and 01:00 pm and is a common method of studying module performance at maximum possible yield. Both figures show a general rise in temperatures as time goes by and the season progresses more towards summer (July). As the dusty panel was not clean at the beginning of the testing period, the temperature difference between the two panels appears early on as the dusty panel starts off with a slightly lower temperature.

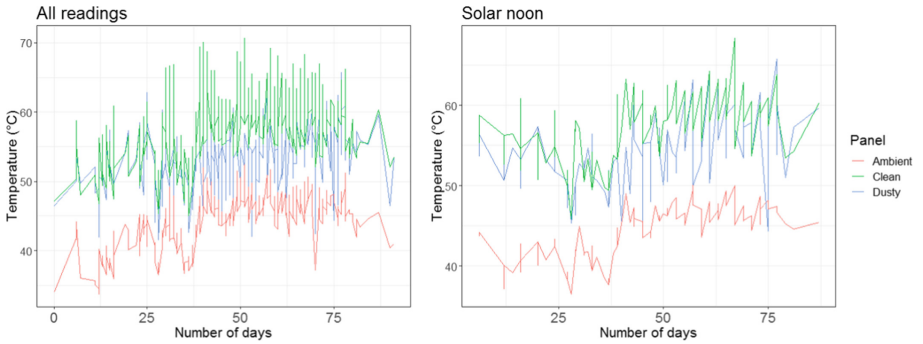

Fig. 2. Ambient, clean module, and dusty module temperatures vs. number of days since the beginning of the test period

Next, a heatmap was created for each panel which shows the distribution of average module reading and maximum power throughout the test period (Fig. 3). Given the level of dust accumulation is a function of time, the number of days since the beginning of the test period is shown as well. As discussed earlier, the ambient temperature, and consequently the modules' temperatures, increased with time. This is observable as most of the reading at the beginning are clustered around the 50 °C point but begin to shift towards 60 and 70 °C. Consequently, as the module overheats it also experiences a decrease in output power. While the shift is clear in the clean panel readings, the dusty panel exhibits a different trend. The maximum power decreases with time, while increase in temperature is generally less noticeable. While the decrease of power can be attributed to the accumulation of dusty which effectively lowers the amount of solar irradiance absorbed by the panel, it is likely that the dust also serves as a cooling mechanism which prevents the module overheat exhibited by the clean panel.

The relationship between ambient and each module's temperature can be furthered explored by running a correlation analysis. The dataset was first divided based on time of the day in order to eliminate time as a variable. As shown in Fig. 4, there is a strong correlation between ambient temperature and the clean module's temperature throughout the day.

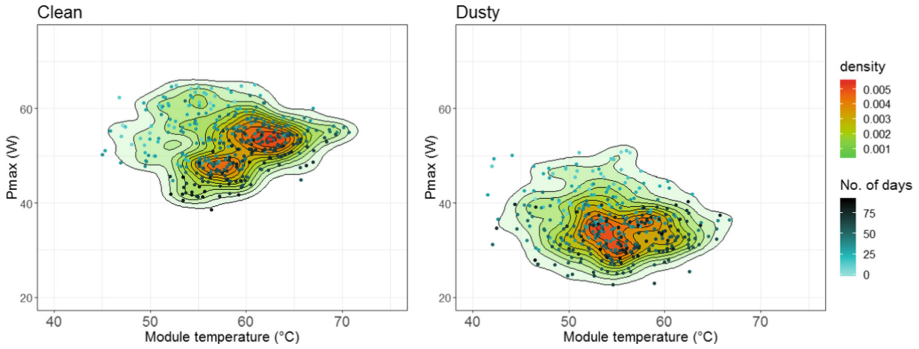

Fig. 3. Statistical distribution of average module temperature and maximum power for the clean and dusty panels

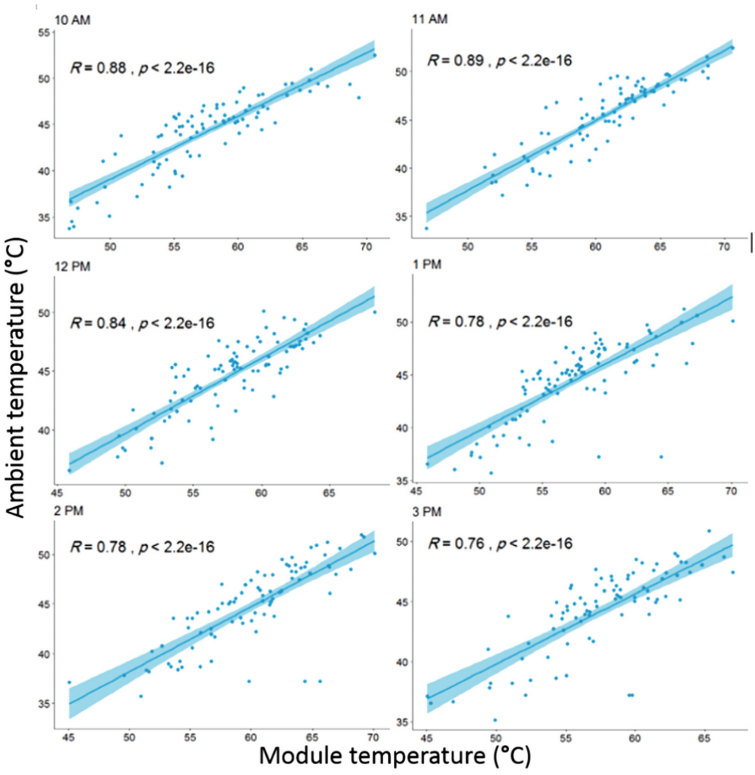

Fig. 4. Correlation between ambient temperature and the clean module's temperature at each hour of the day

Looking at the dusty panel readings in Fig. 5, however, shows that the correlation between the ambient temperature and the dusty module's temperature is not as strong as

the clean panel. It is safe to assume that the addition of dust particles alters the amount of heat transferred from the ambience to the dusty module.

Figure 6 compares the statistical distribution of the two panels' average temperature by time of day. As shown in the graph, the average temperature of the dusty module is consistently lower by up to 6 °C. Readings from both panels were subjected to Shapiro normality test. The results showed that with the exception for the 3:00 pm and 12:00 pm readings, the module temperature was generally normally distributed. Two-sided T-tests and Kruskal-Wallis test showed that there is a statistically significant difference between the distribution of the two modules' readings. The results are shown in Table 1.

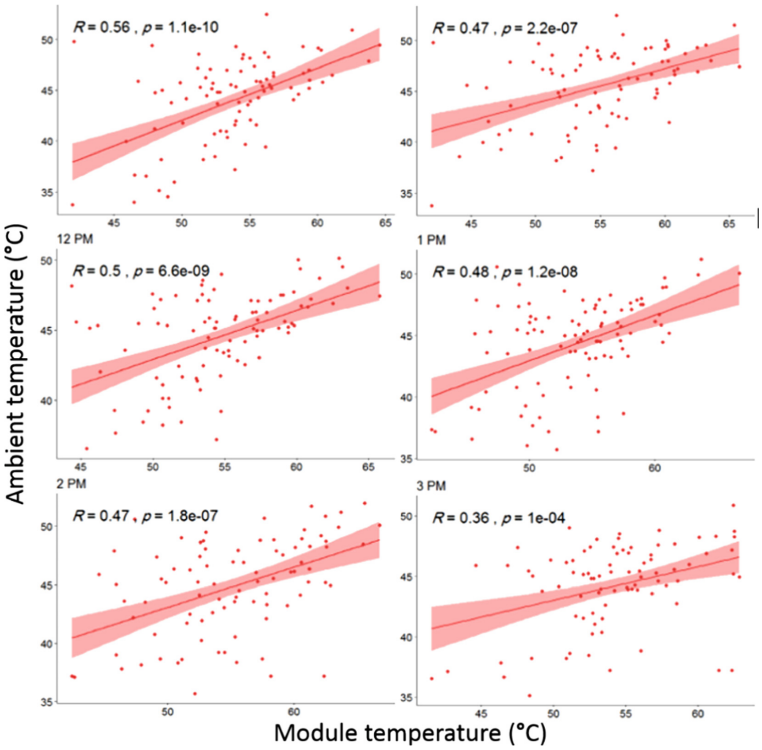

Fig. 5. Correlation between ambient temperature and the dusty module's temperature at each hour of the day

The statistical difference between the two modules' temperatures suggest that it might be possible to use the module temperature as an indication of dust levels. If true, it could be possible to use the difference in temperature between a dusty module and a clean one to predict the dusty module's output, if the clean module's temperature is known. Clean PV modules exhibit a fairly linear dependence on solar irradiance. This was confirmed in [16], where a linear regression model was used to predict the output of

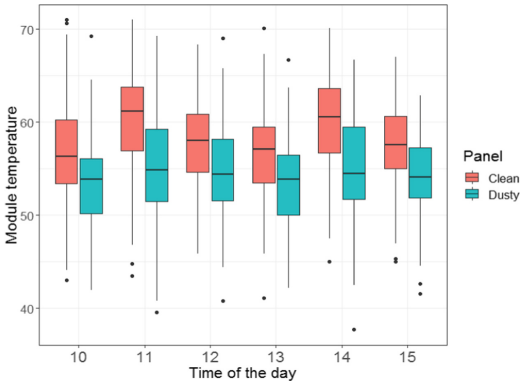

Fig. 6. Comparing statistical distribution of the two modules' temperatures by time of day

a clean module using only solar irradiance. Given a reference clean panel whose output and temperature are known, the goal is to approximate the output of a dusty module out in the field with low cost hardware and little and no interference with the dusty module's circuitry.

Table 1. Results of statistical analysis of the module temperature distributions

			10 am	11 am	12 pm	1 pm	2 pm	3 pm
Shapiro	Clean	W	0.982	0.980	0.977	0.992	0.988	0.979
		P_value	0.134	0.108	0.037	0.670	0.428	0.078
	Dusty	W	0.989	0.984	0.989	0.988	0.984	0.974
		P_value	0.472	0.225	0.461	0.325	0.196	0.029
T-test	Module temperature	T-value	5.143	8.636	NA	6.238	7.121	NA
		Df	220.03	211.1	NA	255.6	22.59	NA
		p-value	$6.0e-7$	$1.4e-15$	NA	$1.8e-9$	$1.5e-11$	NA
Kruskal-Wallis	Module temperature	Chi-squared	227	215	239	257	227	221
		df	171	165	183	191	173	169
		P_value	0.0027	0.0054	0.0034	0.0010	0.0037	0.0044

3.2 Power Prediction via Module Temperature

Several models were built in an attempt to predict the output of a dusty panel using a reference clean module and low-cost hardware at the dusty module's site. The input variables considered were the maximum power of a reference clean panel, the difference in temperature between the dusty and clean panels, the number of days, and time of the day. The structure of the neural network is shown in Fig. 7.

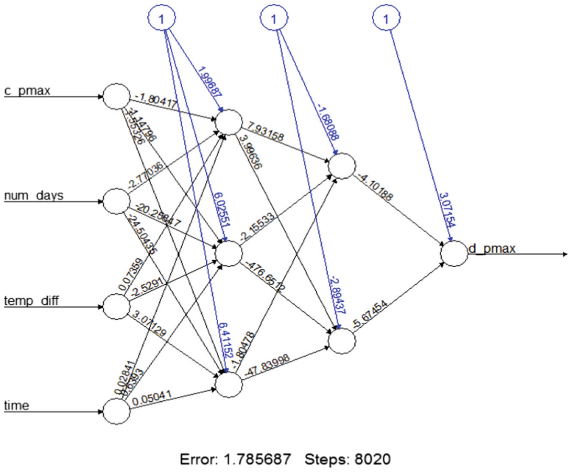

Fig. 7. Neural network used to predict the output of a dusty panel

The training and testing were done via 10-fold cross validation. the dataset was divided into 10 random subsets, and each subset was used once to test while rest of the dataset was used to train for a total of 10 full training and testing runs. The process was also repeated 10 times to decrease bias. The average Root Mean Square Error (RMSE) and correlation coefficient (R) were calculated by averaging the statistics of all runs. Table 2 shows an evaluation of the models built to predict the maximum power of a dusty module using only its average temperature and information about a reference clean module.

Table 2. Results of running various machine learning algorithms to predict the maximum power point of a dusty module

	Back propagation neural network	Random forest	Random tree	Linear regression	Lazy IBK	Regression by discretization	Multilayer perceptron
R	0.8684	0.9846	0.9846	0.8407	0.9089	0.9461	0.8622
MAE	0.0572	0.0230	0.0125	0.0913	0.0477	0.0462	0.0837
RMSE	0.0735	0.0359	0.0355	0.1095	0.0859	0.0657	0.1039

As the table shows, the algorithms with the lowest root mean square error are the random trees and random forest, with an RMSE of 0.036 and 0.035, respectively. To compare with existing literature on power prediction, Table 3 presents a comparison of various models presented in literature. In addition to varying algorithms, a different assortment of context variable was used in each work. As shown in the table, the results

achieved by the proposed model are comparable with other proposed models that utilize more elaborate context variables.

Table 3. Performance comparison of example models in literature for power prediction

Source	Inputs	Context	Algorithm	RMSE	MAE	R
[17]	Cumulative dust, temperature, humidity, module temperature, irradiance, dust, wind speed	N/S	Linear regression	0.074	0.050	0.926
	Temperature, module temperature, irradiance, dust	N/S	Linear regression	0.095	0.070	0.877
	Irradiance, wind speed, module temperature, temperature	N/S	Linear regression	0.094	0.069	0.878
	Cumulative dust, temperature, humidity, module temperature, irradiance, dust, wind speed	N/S	M5P tree	0.070	0.048	0.934
	Temperature, module temperature, irradiance, dust	N/S	M5P tree	0.090	0.068	0.891
	Irradiance, wind speed, module temperature, temperature	N/S	M5P tree	0.086	0.062	0.899
[18]	Irradiance, temperature, total cloud amount, low cloud amount	Cloudy	GA-BP	0.066	0.056	–
	Irradiance, temperature, total cloud amount, low cloud amount	Cloudy & rainy	GA-BP	0.1286	0.1332	–
[19]	Measured power, irradiance predictions, humidity prediction, temperature prediction	N/S	SVM	0.150	–	–
	Measured power, irradiance predictions, humidity prediction, temperature prediction	N/S	SVM-DF	0.135	–	–
	Measured power, irradiance predictions, humidity prediction, temperature prediction	N/S	ANN	0.180	–	–
[20]	Operating voltage, short circuit current, reverse current, series resistance	Sunny	DE	–	0.0494	–
	Operating voltage, short circuit current, reverse current, series resistance	Sunny	PSO	–	0.0460	–
	Operating voltage, short circuit current, reverse current, series resistance	Sunny	ECSS	–	0.0345	–
	Operating voltage, short circuit current, reverse current, series resistance	Cloudy	DE	–	0.0118	–
	Operating voltage, short circuit current, reverse current, series resistance	Cloudy	PSO	–	0.0112	–
	Operating voltage, short circuit current, reverse current, series resistance	Cloudy	ECSS	–	0.0065	–
[21]	Irradiance, temperature	N/S	GAPSOBP	–	0.0100	–
	Irradiance, temperature	N/S	D-GAPSOBP	–	0.0070	–
	Irradiance, temperature	N/S	GABP	–	0.0100	–
	Irradiance, temperature	N/S	D-GABP	–	0.0073	–

4 Conclusion

In this work, a low-cost Internet of Things system was used to observe the output and context of two modules, one dusty and one clean. The measured data was used to predict the output of a dusty module using only the temperature of the dusty module, the temperature and power output of a clean module, and temporal data. Seven different algorithms were trained and tested. At the end of the evaluation, it was found that the random tree and random forest algorithms were able to predict the output with an RMSE of around 3%, while the output of a backpropagation neural network had an RMSE of around 7%. While the performance is not as good as some of the artificial intelligence discussed in some literature, the simplicity of the design in terms of low financial cost as well as low interference with the modules' circuitry motivates to further pursue the method for large scale farms. It is expected that with a larger dataset and more accurate temperature readings the error rate my drop and the method will therefore prove to be highly efficient in the long run.

Acknowledgment. The research reported here was supported in part by grant #SCRI-18-02, and Petrofac-Chair grant from the American University of Sharjah, UAE.

References

1. Fouad, M.M., Shihata, L.A., Morgan, E.I.: An integrated review of factors influencing the performance of photovoltaic panels. Renew. Sustain. Energy Rev. **80**, 1499–1511 (2017). https://doi.org/10.1016/j.rser.2017.05.141
2. de Freitas Viscondi, G., Alves-Souza, S.N.: A systematic literature review on big data for solar photovoltaic electricity generation forecasting. Sustain. Energy Technol. Assess. **31**, 54–63 (2019). https://doi.org/10.1016/j.seta.2018.11.008
3. Rahman, M.M., Selvaraj, J., Rahim, N.A., Hasanuzzaman, M.: Global modern monitoring systems for PV based power generation: a review. Renew. Sustain. Energy Rev. **82**, 4142–4158 (2018). https://doi.org/10.1016/j.rser.2017.10.111
4. Yang, L., Gao, X., Lv, F., Hui, X., Ma, L., Hou, X.: Study on the local climatic effects of large photovoltaic solar farms in desert areas. Sol. Energy **144**, 244–253 (2017). https://doi.org/10.1016/j.solener.2017.01.015
5. Maghami, M.R., Hizam, H., Gomes, C., Radzi, M.A., Rezadad, M.I., Hajighorbani, S.: Power loss due to soiling on solar panel: a review. Renew. Sustain. Energy Rev. **59**, 1307–1316 (2016). https://doi.org/10.1016/j.rser.2016.01.044
6. Burton, P.D., Boyle, L., Griego, J.J.M., King, B.H.: Quantification of a minimum detectable soiling level to affect photovoltaic devices by natural and simulated soils. IEEE J. Photovoltaics **5**, 1143–1149 (2015). https://doi.org/10.1109/JPHOTOV.2015.2432459
7. Burton, P.D., King, B.H.: Determination of a minimum soiling level to affect photovoltaic devices. In: 2014 IEEE 40th Photovoltaic Specialist Conference (PVSC), pp. 0193–0197 (2014). https://doi.org/10.1109/PVSC.2014.6925529
8. Pulipaka, S., Kumar, R.: Analysis of soil distortion factor for photovoltaic modules using particle size composition. Sol. Energy **161**, 90–99 (2018). https://doi.org/10.1016/j.solener.2017.11.041
9. Figgis, B., Ennaoui, A., Ahzi, S., Rémond, Y.: Review of PV soiling particle mechanics in desert environments. Renew. Sustain. Energy Rev. **76**, 872–881 (2017). https://doi.org/10.1016/j.rser.2017.03.100

10. Figgis, B., et al.: Investigation of factors affecting condensation on soiled PV modules. Sol. Energy **159**, 488–500 (2018). https://doi.org/10.1016/j.solener.2017.10.089
11. Ilse, K.K., et al.: Comprehensive analysis of soiling and cementation processes on PV modules in Qatar. Solar Energy Mater. Solar Cells **186**, 309–323 (2018). https://doi.org/10.1016/j.solmat.2018.06.051
12. Javed, W., Guo, B., Figgis, B.: Modeling of photovoltaic soiling loss as a function of environmental variables. Sol. Energy **157**, 397–407 (2017). https://doi.org/10.1016/j.solener.2017.08.046
13. Coskun, C., Toygar, U., Sarpdag, O., Oktay, Z.: Sensitivity analysis of implicit correlations for photovoltaic module temperature: a review. J. Clean. Prod. **164**, 1474–1485 (2017). https://doi.org/10.1016/j.jclepro.2017.07.080
14. Salari, A., Hakkaki-Fard, A.: A numerical study of dust deposition effects on photovoltaic modules and photovoltaic-thermal systems. Renew. Energy **135**, 437–449 (2019). https://doi.org/10.1016/j.renene.2018.12.018
15. Shapsough, S., Takrouri, M., Dhaouadi, R., et al.: Using IoT and smart monitoring devices to optimize the efficiency of large-scale distributed solar farms. Wirel. Netw. (2018). https://doi.org/10.1007/s11276-018-01918-z
16. Shapsough, S., Dhaouadi, R., Zualkernan, I.: Using linear regression and back propagation neural networks to predict performance of soiled PV modules. Presented at the 9th International Conference on Sustainable Energy Information Technology (SEIT), Halifax, Canada, August 2019
17. Benhmed, K., 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 (2018). https://doi.org/10.1109/IRSEC.2018.8702880
18. Meng, X., Xu, A., Zhao, W., Wang, H., Li, C., Wang, H.: A new PV generation power prediction model based on GA-BP neural network with artificial classification of history day. In: 2018 International Conference on Power System Technology (POWERCON), pp. 1012–1017 (2018). https://doi.org/10.1109/POWERCON.2018.8601567
19. Buwei, W., Jianfeng, C., Bo, W., Shuanglei, F.: A solar power prediction using support vector machines based on multi-source data fusion. In: 2018 International Conference on Power System Technology (POWERCON), pp. 4573–4577 (2018). https://doi.org/10.1109/POWERCON.2018.8601672
20. Huang, C., Huang, Y., Yang, S., Huang, K., Chen, S.: Parameter estimation and power prediction for PV power generation using a multi-agent algorithm. In: 2019 IEEE International Conference on Industrial Technology (ICIT), pp. 679–684 (2019). https://doi.org/10.1109/ICIT.2019.8755090
21. Li, J., Wang, R., Zhang, T., Zhang, X., Liao, T.: Predicating photovoltaic power generation using an improved hybrid heuristic method. In: 2016 Sixth International Conference on Information Science and Technology (ICIST), pp. 383–387 (2016). https://doi.org/10.1109/ICIST.2016.7483443