Deep Learning at the Edge for Operation and Maintenance of Large-Scale Solar Farms

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

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

Abstract. Real-time monitoring of large-scale solar farms is one important aspect of reliable and secure deployment of 100% renewable energy-based grids. The ability to observe sensors on solar panels using Internet of Things (IoT) technologies makes it possible to study the behavior of solar panels under various conditions and to detect anomalous behaviors in real-time. Such technologies make it possible for grid administrators to make informed decisions in reacting to anomalies such as panel damage, electrical errors, monitoring hardware decay, or malicious data injection attacks. Smart edge devices offer an opportunity to reduce the cost of continuously sending data for anomaly detection by performing analytics on local edge device within a given farm and sending only the result of the analysis back to datacenters. This paper presents the design and evaluation of a low-cost edge-based anomaly detection system for remote solar farms using Raspberry Pi and deep learning. The design was implemented and tested using real-life observations from a solar monitoring system under soiling conditions. The experiments showed that it is possible to run real-time anomaly detection algorithms on edge devices with little overhead in terms of power consumption and utilization of computational resources, making it an ideal system for large-scale implementation.

Keywords: Solar power · Neural networks · Edge analytics · Fog computing

1 Introduction

Renewable energy holds the potential to replace carbon-based fossil fuels as the main energy source for future cities. In the light of the recent rapid global urbanization, integration of clean renewable energy sources is more important than ever to reduce damage to environmental resources and ensure reliable and sustainable population growth through upcoming decades [1, 2]. As a result, the recent years have witnessed a massive increase in renewable energy installations all over the world, including solar, wind, and geothermal energy facilities. The International Renewable Energy Agency (IRENA) reported a total of over 2.35 TW of global installed capacity, with hydro, wind, and solar amounting to 50%, 24%, and 20%, accordingly [3]. The numbers showcase a growth of 7.9% from the previous year, with 55% of new capacity attributed to new solar installations alone, of which 70% is attributed to new solar installations in Asia.

Y.-B. Lin and D.-J. Deng (Eds.): SGIoT 2020, LNICST 354, pp. 27–44, 2021.
https://doi.org/10.1007/978-3-030-69514-9_4

Countries in the GCC with a desert climate such as the United Arab Emirates and Saudi Arabia are some of the biggest investors in solar energy in the region [4]. With an average daily sunshine hour of 10 h/day and a Global horizontal Irradiance index (GHI) that can go as high as 2.12 MWh/m^2/year, such countries hold the potential to generate massive amounts of solar power which can eventually replace their current dependence on carbon-based fossil fuels [5, 6]. Nonetheless, integration of solar energy sources into main grids remains relatively low. In a country with a desert climate such as the United Arab Emirates, for example, its 22.5 MW of installed solar capacity amounts to only 0.49% [4]. Despite the exponential growth in new solar installations exhibited in the GCC over the past few years, a 100% renewable energy grid remains a distant reality. This is mainly due to the challenges that solar installations face in desert climates such as overheating and soiling [7, 8]. As a result, great research efforts have been dedicated to increasing and optimizing the performance of solar modules in such environments.

One of the main challenges to solar energy integration is the intermittent stochastic nature of the source [9]. The output of a solar module is a direct product of its environment, most notably solar irradiance and temperature. The intermittent nature of weather conditions is reflected through fluctuations in power output which can propagate through main girds, causing inconveniences to power planning at best, and damage to critical assets at worst. Furthermore, solar modules are often located in remote and harsh environments where they are susceptible to damage due to environmental conditions such as overheating, surfaces scratching, and material decay. Such faults can cause serious issues such as module mismatch or open circuit, thus significantly reducing the power output of an installation. Furthermore, once a solar installation has been integrated into the grid, it is vital that the state of the installation and its output at any given moment is known to the system in order to ensure reliable power planning [10]. As a result, faults must be detected and rectified with minimal delay in order to prevent fluctuations from propagating through the grid and causing service interruptions or damage to critical assets [11]. Anomaly detection is a key component of operation and maintenance (O&M) in automated systems. It is especially critical for IoT systems where autonomous response to system failures such as hardware malfunction, software errors, or security breaches is key to ensuring reliable operation.

Anomaly detection can be performed using supervised or unsupervised methods, where the output can either be a label ("normal" or "anomaly") or a score depicting the likelihood a reading is an anomaly [12]. However, performing real-time anomaly detection requires large amounts of real-time data to be transmitted from solar farms over wireless networks back to central datacenters. For example, an edge device that measures the performance of an individual solar panel generates a message which is around 200 Bytes in size. Similarly, an edge device that monitors the environmental conditions in a facility generates a message size of 1000 Bytes [13]. In an installation where edge devices send data at a rate of 1 message/minute, two devices alone observing over a period of 12 h can generate a network load of 864 kB/12 h. Given the scale of recent solar installations such as the 2 GW Al Dhafra project in Abu Dhabi [14], the amount of data required for real-time monitoring is likely to pose a challenge in terms of data transportation. As a result, recent research in various IoT applications has shifted to edge computing and fog computing as a way to perform analytics on edge devices local to the installation [15], thus reducing the requirements of wireless communication, as well as data processing and storage at datacenters. However, this requires edge hardware that is both low in cost and capable of running sophisticated neural networks in real time.

This paper presents and evaluates an edge analytics environment that uses Raspberry Pi to detect anomalies in solar power in real time for large-scale distributed solar farms.

2 Anomaly Detection and Solar Power

2.1 Intermittency in Solar Energy

Solar energy is a product of its environment. The amount of energy generated by a solar module depends not only on the amount of irradiance absorbed by the surfaces, but also the module temperature. When solar modules are produced, the maximum power is generated at standard testing conditions where irradiance is 1000 W/m^2 and ambient temperature is 25 °C. Any decrease of irradiance or change in ambient temperature can reduce the amount of energy generated. In a real-life solar farm, numerous conditions can influence the power output. Meteorological conditions such as shading [16–18], haze, or fog [19], can significantly reduce the amount of irradiance that is absorbed by the module. In high temperature weather regions, hot temperatures can result in module overheating, which can shift the operating point of a module thus preventing it from operating at maximum power. A complex environmental phenomenon that combines several environmental elements including irradiance, temperature, humidity, and wind level is soiling. Soiling is defined as the accumulation of particles such as dust, dirt, snow, or bird droppings on the surface of a solar module, effectively reducing the amount of solar irradiance that can be absorbed [20]. Soiling represents a major hindrance to solar energy adoption in desert regions such as the GCC region. A study conducted in 2013 on solar energy generation in the smart city Masdar [21] reported a power loss of at least 40% due to regular dust storms [22].

However, while fluctuations in meteorological conditions are key to solar power generation and can significantly influence the performance of large-scale remote solar installations, they seldom require immediate intervention. Instead, navigating solar power generation through adversarial weather conditions can be done via proactive operation and maintenance. For example, overheating can be prevented using module cooling [23]. Soiling loss, on the other hand, can be avoided in a cost-effective way by optimizing cleaning methods and schedules, as shown in [24, 25]. As for factors that cannot be controlled such as shading due to clouds and fog, mathematical models as well as machine learning and deep learning methods are used to predict power fluctuation due to shading [17, 26] and plan energy reserves accordingly, which can be applied to clean modules as well as modules under soiling conditions [27].

2.2 Anomalies in Solar Energy Systems

While fluctuations in power can be severe depending on ambient conditions, in this work, such fluctuations are not considered "anomalies" due to the fact they are an inherent part of the energy source's nature. Anomalies, instead, are faults that may occur within modules or data that describes modules. Anomalies can be loosely divided into three main categories: module faults, monitoring system faults, and cyberattacks.

Array Anomalies. Array anomalies refers to faults that can occur to module bodies, module interconnection, or internal electrical connections within the farm. The first type of array anomalies is module damage and decay. Over time, PV modules can develop localized overheating or "hot spots", where a mismatch between cells within one module causes a cell or more to overheat [28]. The phenomenon can occur when parts of the module are shaded by a clouds, soiling elements, or shadows from surrounding objects such as buildings or other modules for prolonged periods of time [29]. The development of hot spots can degrade the performance of the module, damage its integrity, and cause irreversible malfunctions [30]. The same issue can occur at a higher level when a mismatch occurs between several PV modules in a string [31–33], which in the long run can accelerate module degradation. Furthermore, faults can occur within the internal connections in the farm where a configuration error, short circuit, or open circuit can bring down full strings of modules [29].

System Anomalies. The second type of anomalies is corrupted or missing readings. This type of anomaly is common in monitoring systems where values can be lost or corrupted during data collection or transmission [34]. Internet of Things systems are prone to such anomalies partially due to their inherent low-resource nature which makes it difficult and costly to deploy sophisticated hardware and software components in remote installations [35, 36]. Damage to sensors, hardware malfunctions, software exceptions, and network disconnections can generate anomalous data that, if allowed to propagate through the system, may cause disruptions or result in false decision making. Distributed solar farms represent a particularly challenging case in data reliability due to their remote nature. In addition to standard IoT challenges, solar installations often exist in harsh weather conditions [37] where sensors and processing hardware are prone to damage and decay. In such cases, early detection of system faults, or anomalies, is key to preventing false readings from propagating through grids and causing damage to assets.

Solar Power Data Security. The third type of anomalies is cyberattacks. Energy systems are some of the most critical assets in a given community. However, studies show that 54% of cyberattacks on infrastructure are directed at energy systems [38]. Isolated solar microgrids and grid-integrated solar installations are susceptible to such attacks. Cyberattacks can target solar monitoring devices, the network through which data is transmitted, or the data itself [39]. The more highly populated and further distributed a solar energy system is, the harder it is to enforce appropriate security measures end-to-end [40]. Furthermore, modern energy systems are highly heterogenous in terms of type of hardware and data allowed within the system. Protecting such systems requires sophisticated security suites to address all vulnerabilities at each point in the system. It is therefore key to implement attack detection and prevention throughout the system. While active attacks such as Denial-of-Service [40] can be immediately detected, making it easier to rectify within a short time limit, passive attacks on data may go unnoticed. "False data injection" refers to a type of attack where an intruder gains access to the systems network and proceeds to transmit false data disguised as legitimate readings [41]. In an energy system where meter readings and monitoring systems dictate the flow of power through the system [38], false information can result in damages that range from energy theft, to inconvenient minor disruptions in services, and all the way to severe damage to critical assets and drastic financial losses. Anomalies in this context refer to illegitimate readings injected by a cyber attacker in order to disrupt system operations. The ability to immediately detect and correctly classify cyberattack anomalies makes it possible to contain the attack and prevent false information from propagating throughout the system.

Anomaly Detection in Solar Power Systems. Existing anomaly detection methods for solar monitoring generally follow one of two approaches [29, 42]: model-based anomaly detection, and data-based anomaly detection. Model-based anomaly detection uses information from the healthy module to build a virtual model of the healthy module which operates in parallel with the real-life PV module. From that moment on, real-time measured values are compared to predicted healthy values to look for deviations from expected behavior. In a variation of this method, the model is built offline based on the module's characteristics and the environment. Using the virtual model, error threshold values are then set for the voltage and the current. From that point on, real-time readings are compared to the threshold, and any point beyond the minimum or maximum expected performance is marked as an anomaly. Examples of existing work are shown in Table 1.

In data-based approaches, on the other hand, machine learning (ML) and deep learning (DL) algorithms are trained and used to detect anomalies. In this approach, previous knowledge of the domain theory is not required as analysis focuses on trends and patterns in the dataset independent of domain-based assumptions. Furthermore, several works on anomaly detection extend to using ML and DL in anomaly classification. The latter is rather useful in determining the source of the anomaly, which is key for implementing cost-effective countermeasures.

While existing work on anomaly detection in solar monitoring systems covers a variety of anomalies and detection methodologies, they often assume data with a somewhat clear distinction between "valid" readings and anomalies. However, in operating conditions where environmental elements such as soiling exist and skew performance, detecting real anomalies from expected degradation becomes more complex. Furthermore, while solar panels in one location are expected to give a relatively-identical performance, contextual elements such its surface temperature [48], short-term shading [17], or aerosol particle concentration [49, 50] can vary based on its location and skew the module's "normal" behavior from identical behavior assumed at manufacturing and give it a unique profile. In such cases, treating models for individual modules with unique characteristics and behaviors makes it possible to detect slight changes in behavior that may otherwise be neglected in generalized performance modeling.

Table 1. Summary of some of key existing work on anomaly detection in solar power

Paper title	year	Input data	Method	Outputs
Hierarchical Anomaly Detection and Multimodal Classification in Large-Scale Photovoltaic Systems [42]	2019	SCADA data (current, voltage) recorded every minute	Detection: Local Context-Aware Anomaly Detection using AutoGMM Classification: SVM, Bagging, XGBoost	Anomaly detection & classification: Sensor bias/aging, building shading, hotspot/glass breakage, grass shading, surface soiling
Anomaly Detection of Solar Power Generation Systems Based on the Normalization of the Amount of Generated Electricity [43]	2015	Electric current	Offline comparison with normal distribution of historical data	Anomaly detection
Online fault detection in PV systems [44]	2015	Irradiance, module temperature,	Predict max power of healthy panel and compare	Anomaly detection
What's Wrong with my Solar Panels: a Data-Driven Approach [45]	2015		Detection: Comparison with expected output Classification: comparison with coefficient of variation, trees, SVM, KNN	Anomaly detection & classification: partial shading, full cover
Expected output calculation based on inverse distance weighting and its application in anomaly detection of distributed photovoltaic power stations [46]	2020	Station power, grid voltage, grid current	Reverse distance weighting method	Anomaly detection & classification: open circuit, short circuit
Anomaly detection and predictive maintenance for photovoltaic systems [47]	2018	Irradiance, temperature, AC power	Compares ANN predicted output with measured output	Anomaly detection for predictive maintenance
Shading prediction, fault detection, and consensus estimation for solar array control [26]	2018	Current, voltage	Clustering through Expectation Maximization	Anomaly detection & classification: arc faults, ground faults

3 Proposed Design

The design shown here is based on the study presented in [13] which evaluated the feasibility of using low-cost IoT edge devices and long-range low-power wireless networks to facilitate the real-time monitoring of large-scale and remote solar farms. As shown in Fig. 1, each solar module in a string can be equipped with a low-cost microcontroller such as ESP32 [51] which can interface with current, voltage, and surface temperature sensors in order to observe the state of the module. Module observations can be sent over a local WiFi network to a gateway which then forwards all readings to a datacenter over a long-range network such as LoRaWAN [52]. Within the farm, Message Queueing Telemetry Transport (MQTT) [53] is used by the devices to send information from module observers to the RPi gateway. A lightweight MQTT broker such as Mosquitto [54] can be hosted on the RPi, enabling it to receive observations from any module that is added to the system in real-time and with little configuration overhead.

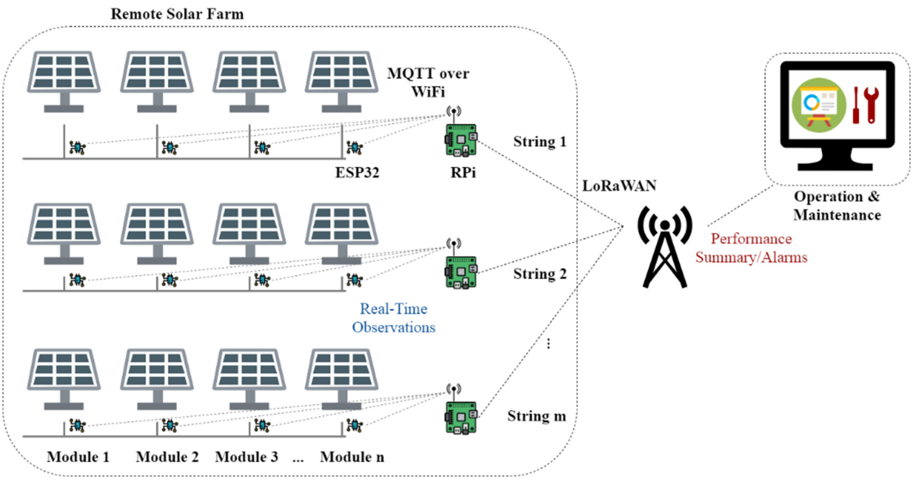

Fig. 1. Proposed edge architecture

At the datacenter, information can be processed and stored. However, as discussed earlier, sending observations from all modules in real-time generates large amounts of data which exponentially increases the network bandwidth requirements. Delegating all data processing to datacenters also significantly increases the cost of running and maintaining the datacenters and creates a single point of failure. Alternatively, utilizing an edge computer such as the Raspberry Pi (RPi) [55] as a gateway for each string makes it possible to pre-process and analyze observations from the modules locally. Not only is the RPi capable of reading and processing data from several module observers in a string, but the edge computer is able to run deep learning algorithms to predict performance based on modules' context and detect anomalies in performance. Furthermore, separate neural network models can be created for each module and trained to its own unique characteristics to be used to generate custom reports and detect anomalies that are specific to its profile.

4 Evaluation

4.1 Data Collection

Information describing the performance of two modules as well as their context was collected over the period of three months. While one of the modules was cleaned on weekly basis, the other module was left to experience soiling conditions. The two modules at the end of the testing period are shown in Fig. 2. The goal was to experiment with and compare anomaly detection in the case of clean modules as opposed to modules under soiling conditions. Hourly maximum power measurements were recorded using the IV tracing system described in [56]. Additionally, solar irradiance, and module temperature were also observed. The number of days since the beginning of the experiment was also used as a variable as it referred to the period of time for which the dusty panel has not been cleaned. This was used as a variable to roughly express the level of dust on the surface of the panel. The dataset was cleaned for invalid readings due to hardware faults. The final dataset consisted of 3308 observations containing performance and context information for each panel.

Fig. 2. The dusty and clean solar panels at the end of the testing period

4.2 Model Architecture

Observations collected over the period of the experiment were used to train a simple regression neural network that predicts the output of a remote solar panel under soiling conditions using the panel's context and a reference clean panel. This is because while real-time power can be measured in a remote solar farm, a single panel's performance is dependent on panels configuration in a string or an array. Alternatively, a better predictor of how much power a panel is able to produce in a given context is obtained through performing IV tracing [57], where voltage sweep is applied to panel starting from short-circuit to open-circuit in order to obtain short-circuit current, open-circuit voltage, as well as maximum power point and its corresponding voltage and current values. However, IV tracing requires disconnecting the panel from the rest of the array while tracing is performed. Such process is unfeasible in real-world operational conditions. Alternatively, the model built here aims to enable prediction of maximum power point using only maximum power from a reference clean model and the difference in context between the two panels. This relationship has been previously explored and a similar model was built in [58] (shown in Fig. 3). The same model can be used for anomaly detection by recognizing deviation from expected behavior using operational output power of a panel. The model can be trained to predict expected operational output of a remote solar panel based on its context and a reference panel. Since operational output of the remote panel can be measured in a real-world setting, comparing predicted output and observed output provides means to detect anomalous behaviors in real-time.

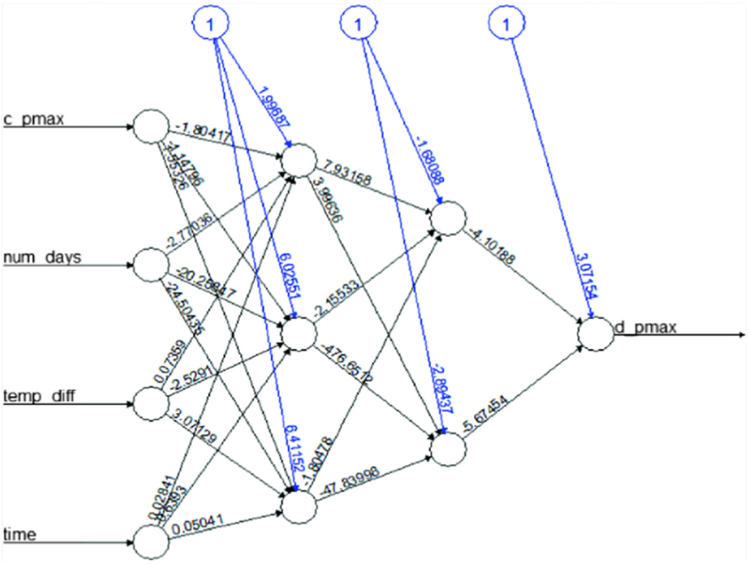

Fig. 3. Neural network architecture [58]

The model was first trained on the RPi using 80–20 training-validation split, with mean squared error (mse) and coefficient of determination (r^2) as training metrics. The formulas for mse and r^2 are shown in Eqs. 1 and 2.

$$mse = \frac{1}{n}\sum\nolimits_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2 \tag{1}$$

$$r^2 = \frac{\sum_i e_i^2}{\sum_i (y_i - \overline{y})^2} \tag{2}$$

The result of training the model for 1000 epochs is shown in Fig. 4 and Fig. 5. Once the model has been trained, an inference model was created to run on the RPi for real-time maximum power prediction. A regular Tensorflow model and a TFlite model were created to be implemented on the RPi.

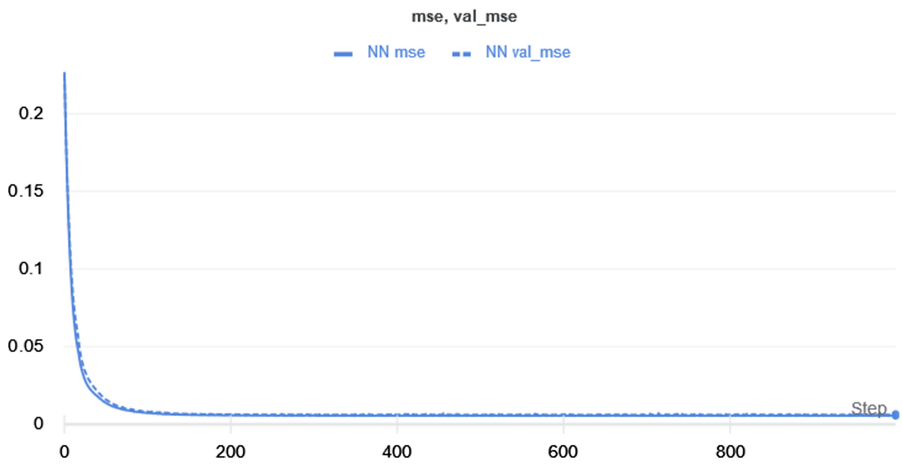

Fig. 4. Training and validation Mean Squared Error

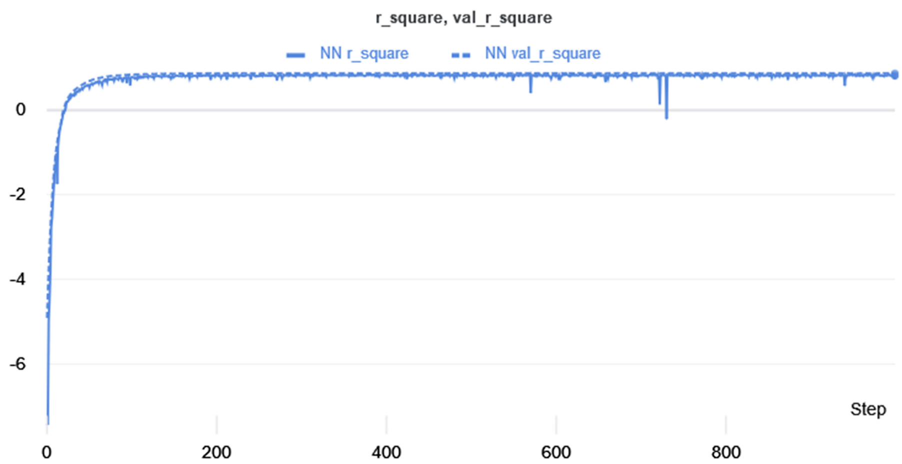

Fig. 5. Training and validation coefficient of determination

4.3 Inference Model Performance Evaluation

Experimental Setup. While the model performed well in terms of power prediction, the feasibility of running it on a RPi gateway in a solar farm depends on the amount of resources required by the RPi, as well as the number of panels that a single RPi can monitor and analyze. An experiment was designed to test the performance of the RPi and the required resources for running the algorithm as the number of panels increases. The experimental setup is shown in Fig. 6.

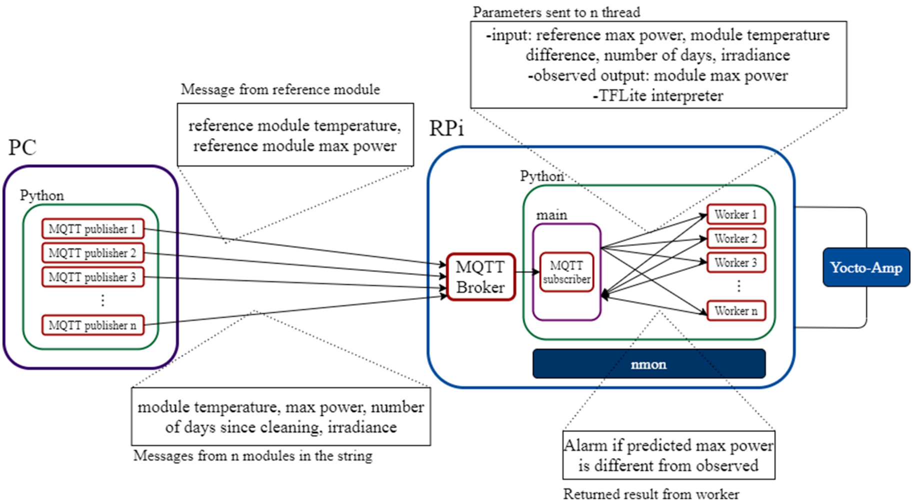

Fig. 6. Model inference performance evaluation experimental setup

As shown in Fig. 6, a python script running on a PC was used to simulate panels publishing real-time observation via MQTT over WiFi. A python script running on the RPi implements an MQTT subscriber which receives all observations from panels in its string. When an observation is received, the script uses information from the panel as well as a reference panel to combine the four inputs to the model which are the irradiance, number of days since the remote panel has been cleaned, difference in the two panels' surface temperatures, and the power output of the clean panel. The script launces a thread which uses the Tensorflow [59] model or TFLite [60] interpreter to predict the output of the remote panel and compare it with measure output. The thread then returns with an error code if deviation is detected. The error code can then be published to trigger appropriate action. A Yocto-Amp [61] is used to measure the current consumption of the RPi under a controlled voltage value of 5.0 V, while nmon [62] is used to monitor the RPi's internal resources such as CPU utilization. The number of clients, or modules in one string, was varied in power of 2 starting from 1 client, all the way to 512 clients, when possible. Each panel published at a frequency of 1 message/minute.

Experimental Results. Power consumption values and CPU% utilization while the RPi was running the inference model using TF and TFLite are shown in Fig. 7 and Fig. 8. As expected, there is a significant difference between running the inference by loading the TF model as opposed to using a TFLite version of the model. Furthermore, the TF model is only usable for up to 256 clients, at which point the system's resources including CPU and active memory get depleted, causing it to crash. A closer look into the RPI CPU utilization while serving one module as opposed to serving 16 modules is shown in Fig. 9 and Fig. 10. The cycles seen in the graph correspond to the publishing cycle of 1 batch of messages every minute. As shown in Fig. 10, almost 100% of the CPU is utilized most of the time. Furthermore, the busy state of the CPU extends for the full minute, causing it to overlap with the new batch of messages, which then causes a huge overload to the system's resources.

On the other hand, using a TFLite model for inference proves to be highly feasible as it consumes very little resources, where under 35 Wh are required to power the edge device for 12 h. The power consumption was calculated for a maximum of 12 h because that is the estimated max of sun hours in a day. As performing analysis during nighttime is not useful, the edge device can be turned off or switched to lower-power mode. The CPU utilization is also less than 40%, allowing the RPi enough free resources to run other reliability tasks. As shown in Fig. 11, the active cycles of the inference are cleanly separated and the CPU is able to complete the analysis task early enough so that it does not overlap with the following batch of messages, as opposed to the case with using a TF model.

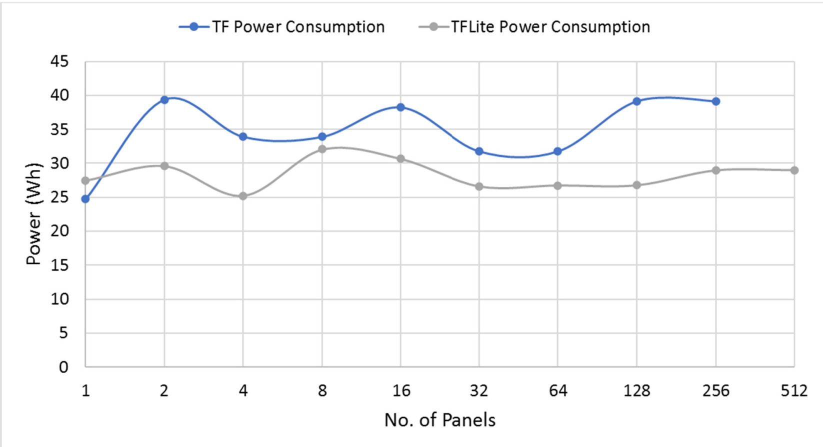

Fig. 7. Power consumption by the RPi while running inference for 12 h

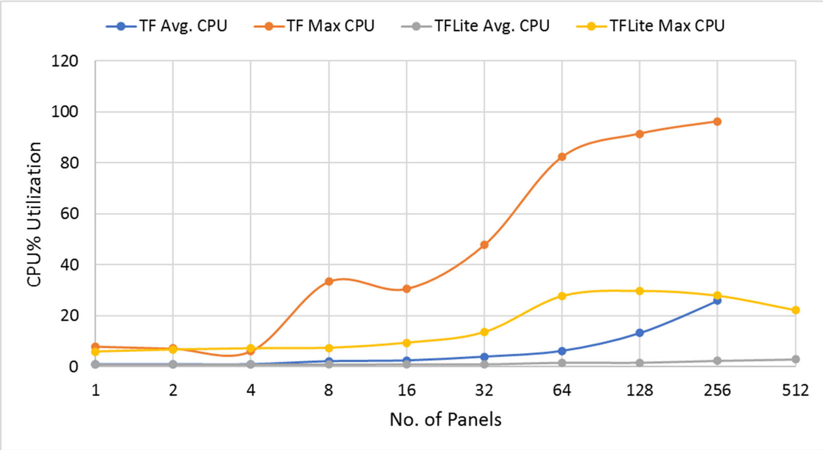

Fig. 8. CPU% utilization of the RPi while running inference

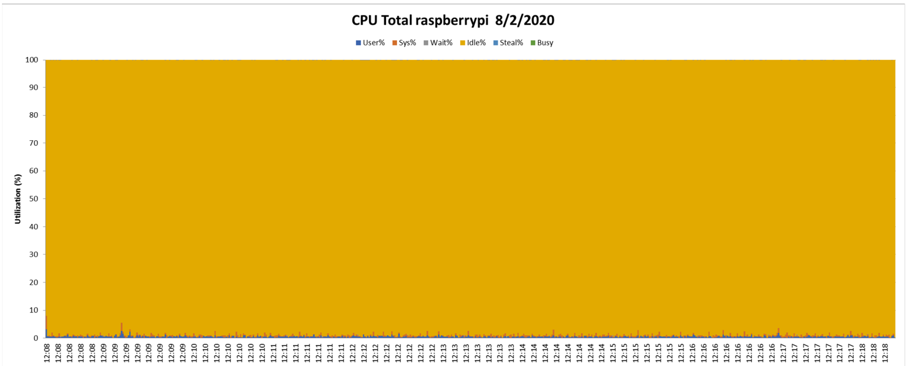

Fig. 9. RPi CPU utilization % while experimenting with 1 client for 5 min - TF

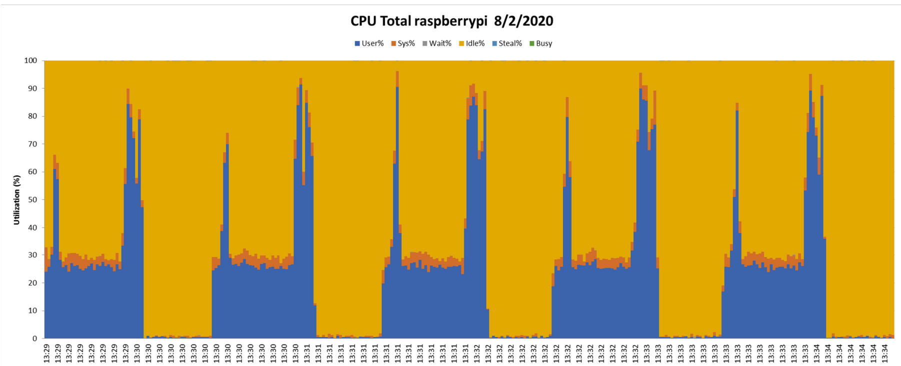

Fig. 10. RPi CPU utilization % while experimenting with 256 clients for 5 min – TF

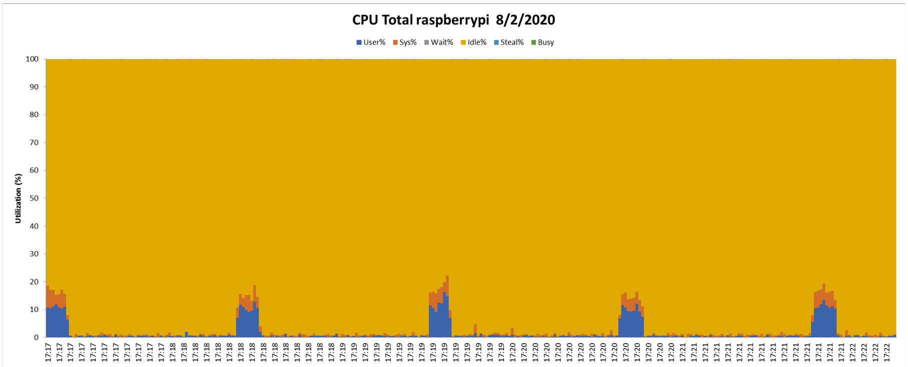

Fig. 11. RPi CPU utilization % while experimenting with 512 clients for 5 min – TFLite

The experimental results indicate that it is possible to use a low-cost edge computing device such as the RPi to perform real-time analytics within local solar farms and detect and react to anomalous behaviors. The design can be further evolved by running TFLite models on more constrained, cheaper devices such as ESP32 devices for monitoring the modules, making the design more modular and further reducing the data communication overhead.

5 Conclusion

Real-time solar monitoring and analysis holds the key to reliable and secure grid integration. It also pushes for the possibility for solar energy to completely replace fossil-based fuels it the future. However, one key issue with solar power is its intermittency and sensitivity to volatile environment elements such as shading and heat. Solar panels are also prone to performance degradation due to natural phenomena such as soiling as well as physical withering in harsh environments. The ability to detect such anomalies using IoT and deep learning makes it possible to swiftly and appropriately react to various issues in order to ensure reliable power generation. This paper has shown that it is possible to use data-driven real-time anomaly detection on the edge using deep learning in conjunction with data from an IoT solar monitoring network. The experiments show that it is possible to run a high number of power prediction and anomaly detection models on an edge device such as the Raspberry Pi with little overhead in required resources. The work can be further developed by evaluating the possibility of running even more complex algorithms on various types of low-resource edge devices used in IoT systems such as ESP32.

References

1. Islam, M.T., Huda, N., Abdullah, A.B., Saidur, R.: A comprehensive review of state-of-the-art concentrating solar power (CSP) technologies: current status and research trends. Renew. Sustain. Energy Rev. **91**, 987–1018 (2018). https://doi.org/10.1016/j.rser.2018.04.097

2. Nuortimo, K., Härkönen, J., Karvonen, E.: Exploring the global media image of solar power. Renew. Sustain. Energy Rev. **81**, 2806–2811 (2018). https://doi.org/10.1016/j.rser.2017.06.086
3. Renewable Capacity Statistics 2019. https://www.irena.org/publications/2019/Mar/Renewable-Capacity-Statistics-2019, Accessed 14 Oct 2019
4. Renewable Energy Market Analysis: GCC 2019. https://www.irena.org/publications/2019/Jan/Renewable-Energy-Market-Analysis-GCC-2019, Accessed 25 Jan 2019
5. Treyer, K., Bauer, C.: The environmental footprint of UAE's electricity sector: combining life cycle assessment and scenario modeling. Renew. Sustain. Energy Rev. **55**, 1234–1247 (2016). https://doi.org/10.1016/j.rser.2015.04.016
6. Jamil, M., Ahmad, F., Jeon, Y.J.: Renewable energy technologies adopted by the UAE: prospects and challenges – a comprehensive overview. Renew. Sustain. Energy Rev. **55**, 1181–1194 (2016). https://doi.org/10.1016/j.rser.2015.05.087
7. 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
8. Touati, F., Al-Hitmi, M., Bouchech, H.: Towards understanding the effects of climatic and environmental factors on solar PV performance in arid desert regions (Qatar) for various PV technologies. In: 2012 First International Conference on Renewable Energies and Vehicular Technology, pp. 78–83 (2012). https://doi.org/10.1109/REVET.2012.6195252.
9. Fouad, M.M., Shihata, L.A., Morgan, E.I.: An integrated review of factors influencing the perfomance of photovoltaic panels. Renew. Sustain. Energy Rev. **80**, 1499–1511 (2017). https://doi.org/10.1016/j.rser.2017.05.141
10. Pandey, A.K., Tyagi, V.V., Selvaraj, J.A., Rahim, N.A., Tyagi, S.K.: Recent advances in solar photovoltaic systems for emerging trends and advanced applications. Renew. Sustain. Energy Rev. **53**, 859–884 (2016). https://doi.org/10.1016/j.rser.2015.09.043
11. Rikos, E., Tselepis, S., Hoyer-Klick, C., Schroedter-Homscheidt, M.: Stability and power quality issues in microgrids under weather disturbances. IEEE J. Selected Top. Appl. Earth Obs. Remote Sens. **1**, 170–179 (2008). https://doi.org/10.1109/JSTARS.2008.2010557
12. Ahmed, M., Naser Mahmood, A., Hu, J.: A survey of network anomaly detection techniques. J. Netw. Comput. Appl. **60**, 19–31 (2016). https://doi.org/10.1016/j.jnca.2015.11.016
13. Shapsough, S., Zualkernan, I.: Designing an edge layer for smart management of large-scale and distributed solar farms. In: Kim, K.J., Kim, H.-Y. (eds.) Information Science and Applications. LNEE, vol. 621, pp. 651–661. Springer, Singapore (2020). https://doi.org/10.1007/978-981-15-1465-4_64
14. Campbell, M.: World's largest solar project to provide record-low energy tariffs. https://www.euronews.com/living/2020/05/05/world-s-largest-solar-project-to-provide-record-low-energy-tariffs, Accessed 29 June 2020
15. Khan, W.Z., Ahmed, E., Hakak, S., Yaqoob, I., Ahmed, A.: Edge computing: a survey. Fut. Gener. Comput. Syst. **97**, 219–235 (2019). https://doi.org/10.1016/j.future.2019.02.050
16. Belhachat, F., Larbes, C.: A review of global maximum power point tracking techniques of photovoltaic system under partial shading conditions. Renew. Sustain. Energy Rev. **92**, 513–553 (2018). https://doi.org/10.1016/j.rser.2018.04.094
17. Das, S.K., Verma, D., Nema, S., Nema, R.K.: Shading mitigation techniques: state-of-the-art in photovoltaic applications. Renew. Sustain. Energy Rev. **78**, 369–390 (2017). https://doi.org/10.1016/j.rser.2017.04.093
18. Ramli, M.A.M., Twaha, S., Ishaque, K., Al-Turki, Y.A.: A review on maximum power point tracking for photovoltaic systems with and without shading conditions. Renew. Sustain. Energy Rev. **67**, 144–159 (2017). https://doi.org/10.1016/j.rser.2016.09.013

19. Said, Z., Arora, S., Bellos, E.: A review on performance and environmental effects of conventional and nanofluid-based thermal photovoltaics. Renew. Sustain. Energy Rev. **94**, 302–316 (2018). https://doi.org/10.1016/j.rser.2018.06.010
20. Andenæs, E., Jelle, B.P., Ramlo, K., Kolås, T., Selj, J., Foss, S.E.: The influence of snow and ice coverage on the energy generation from photovoltaic solar cells. Sol. Energy **159**, 318–328 (2018). https://doi.org/10.1016/j.solener.2017.10.078
21. Masdar City Solar Photovoltaic Plant. https://masdar.ae/en/MasdarCleanEnergy/Projects/MasdarCitySolarPhotovoltaicPlant, Accessed 27 May 2019
22. Crot, L.: Planning for sustainability in non-democratic polities: the case of Masdar City. Urban Stud. **50**, 2809–2825 (2013). https://doi.org/10.1177/0042098012474697
23. Nath, M., Singh, D.: A Review on Performance Improvement of Solar Photovoltaic using Various Cooling Methods (2019)
24. Al-Housani, M., Bicer, Y., Koç, M.: Experimental investigations on PV cleaning of large-scale solar power plants in desert climates: comparison of cleaning techniques for drone retrofitting. Energy Convers. Manag. **185**, 800–815 (2019). https://doi.org/10.1016/j.enconman.2019.01.058
25. Luque, E.G., Antonanzas-Torres, F., Escobar, R.: Effect of soiling in bifacial PV modules and cleaning schedule optimization. Energy Convers. Manag. **174**, 615–625 (2018). https://doi.org/10.1016/j.enconman.2018.08.065
26. Katoch, S., Muniraju, G., Rao, S., Spanias, A., Turaga, P., Tepedelenlioglu, C., Banavar, M., Srinivasan, D.: Shading prediction, fault detection, and consensus estimation for solar array control. In: 2018 IEEE Industrial Cyber-Physical Systems (ICPS), pp. 217–222 (2018). https://doi.org/10.1109/ICPHYS.2018.8387662
27. Shapsough, S., Dhaouadi, R., Zualkernan, I.: Using Linear regression and back propagation neural networks to predict performance of soiled PV modules. Procedia Comput. Sci. **155**, 463–470 (2019). https://doi.org/10.1016/j.procs.2019.08.065
28. Kim, K.A., Krein, P.T.: Reexamination of photovoltaic hot spotting to show inadequacy of the bypass diode. IEEE J. Photovoltaics **5**, 1435–1441 (2015). https://doi.org/10.1109/JPHOTOV.2015.2444091
29. Pillai, D.S., Blaabjerg, F., Rajasekar, N.: A comparative evaluation of advanced fault detection approaches for PV systems. IEEE J. Photovoltaics **9**, 513–527 (2019). https://doi.org/10.1109/JPHOTOV.2019.2892189
30. Guerriero, P., Daliento, S.: Toward a hot spot free PV module. IEEE J. Photovoltaics **9**, 796–802 (2019). https://doi.org/10.1109/JPHOTOV.2019.2894912
31. Manganiello, P., Balato, M., Vitelli, M.: A survey on mismatching and aging of PV modules: the closed loop. IEEE Trans. Ind. Electron. **62**, 7276–7286 (2015). https://doi.org/10.1109/TIE.2015.2418731
32. Niazi, K.A.K., Yang, Y., Sera, D.: Review of mismatch mitigation techniques for PV modules. IET Renew. Power Gener. **13**, 2035–2050 (2019). https://doi.org/10.1049/iet-rpg.2019.0153
33. Liu, G., Yu, W., Zhu, L.: Experiment-based supervised learning approach toward condition monitoring of PV array mismatch. Transm. Distrib. IET Gener. **13**, 1014–1024 (2019). https://doi.org/10.1049/iet-gtd.2018.5164
34. Azimi, I., Pahikkala, T., Rahmani, A.M., Niela-Vilén, H., Axelin, A., Liljeberg, P.: Missing data resilient decision-making for healthcare IoT through personalization: a case study on maternal health. Fut. Gener. Comput. Syst. **96**, 297–308 (2019). https://doi.org/10.1016/j.future.2019.02.015
35. Mukhopadhyay, S.C., Suryadevara, N.K.: Internet of Things: challenges and opportunities. In: Mukhopadhyay, S.C. (ed.) Internet of Things. SSMI, vol. 9, pp. 1–17. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-04223-7_1

36. Chakraborty, T., Nambi, A.U., Chandra, R., Sharma, R., Swaminathan, M., Kapetanovic, Z.: Sensor identification and fault detection in IoT systems. In: Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems, pp. 375–376. Association for Computing Machinery, Shenzhen (2018). https://doi.org/10.1145/3274783.3275190
37. Bouaichi, A., et al.: In-situ evaluation of the early PV module degradation of various technologies under harsh climatic conditions: the case of Morocco. Renew. Energy **143**, 1500–1518 (2019). https://doi.org/10.1016/j.renene.2019.05.091
38. Kimani, K., Oduol, V., Langat, K.: Cyber security challenges for IoT-based smart grid networks. Int. J. Crit. Infrastruct. Prot. **25**, 36–49 (2019). https://doi.org/10.1016/j.ijcip.2019.01.001
39. Otuoze, A.O., Mustafa, M.W., Larik, R.M.: Smart grids security challenges: classification by sources of threats. J. Electric. Syst. Inf. Technol. **5**, 468–483 (2018). https://doi.org/10.1016/j.jesit.2018.01.001
40. Shapsough, S., Qatan, F., Aburukba, R., Aloul, F., Ali, A.R.A.: Smart grid cyber security: Challenges and solutions. In: 2015 International Conference on Smart Grid and Clean Energy Technologies (ICSGCE), pp. 170–175 (2015). https://doi.org/10.1109/ICSGCE.2015.7454291.
41. Mohammadpourfard, M., Sami, A., Weng, Y.: Identification of false data injection attacks with considering the impact of wind generation and topology reconfigurations. IEEE Trans. Sustain. Energy **9**, 1349–1364 (2018). https://doi.org/10.1109/TSTE.2017.2782090
42. Zhao, Y., Liu, Q., Li, D., Kang, D., Lv, Q., Shang, L.: Hierarchical anomaly detection and multimodal classification in large-scale photovoltaic systems. IEEE Trans. Sustain. Energy **10**, 1351–1361 (2019). https://doi.org/10.1109/TSTE.2018.2867009
43. Akiyama, Y., Kasai, Y., Iwata, M., Takahashi, E., Sato, F., Murakawa, M.: anomaly detection of solar power generation systems based on the normalization of the amount of generated electricity. In: 2015 IEEE 29th International Conference on Advanced Information Networking and Applications, pp. 294–301 (2015). https://doi.org/10.1109/AINA.2015.198.
44. Platon, R., Martel, J., Woodruff, N., Chau, T.Y.: Online fault detection in PV systems. IEEE Trans. Sustain. Energy **6**, 1200–1207 (2015). https://doi.org/10.1109/TSTE.2015.2421447
45. Gao, P.X., Golab, L., Keshav, S.: What's wrong with my solar panels: a data-driven approach. In: EDBT/ICDT Workshops (2015)
46. Shi, Y., et al.: Expected output calculation based on inverse distance weighting and its application in anomaly detection of distributed photovoltaic power stations. J. Clean. Prod. **253**, 119965 (2020). https://doi.org/10.1016/j.jclepro.2020.119965
47. De Benedetti, M., Leonardi, F., Messina, F., Santoro, C., Vasilakos, A.: Anomaly detection and predictive maintenance for photovoltaic systems. Neurocomputing **310**, 59–68 (2018). https://doi.org/10.1016/j.neucom.2018.05.017
48. Magalhães, P.M.L.P., Martins, J.F.A., Joyce, A.L.M.: Comparative analysis of overheating prevention and stagnation handling measures for photovoltaic-Thermal (PV-T) systems. Energy Procedia **91**, 346–355 (2016). https://doi.org/10.1016/j.egypro.2016.06.282
49. Caballero, J.A., Fernández, E.F., Theristis, M., Almonacid, F., Nofuentes, G.: Spectral corrections based on air mass, aerosol optical depth, and precipitable water for PV performance modeling. IEEE J. Photovoltaics. **8**, 552–558 (2018). https://doi.org/10.1109/JPHOTOV.2017.2787019
50. Balarabe, M.A., Tan, F., Abdullah, K., Nawawi, M.N.M.: Temporal-spatial variability of seasonal aerosol index and visibility—aA case study of Nigeria. In: 2015 International Conference on Space Science and Communication (IconSpace), pp. 459–464 (2015) https://doi.org/10.1109/IconSpace.2015.7283769.
51. ESP32 Overview | Espressif Systems. https://www.espressif.com/en/products/hardware/esp32/overview, Accessed 20 Oct 2018

52. What is the LoRaWAN® Specification? https://lora-alliance.org/about-lorawan
53. Banks, A., Gupta, R.: MQTT Version 3.1. 1. OASIS Standard (2014)
54. Mosquitto MQTT Server
55. Raspberry Pi - Teach, Learn, and Make with Raspberry Pi. https://www.raspberrypi.org/, Accessed 19 May 2017
56. Shapsough, S., Takrouri, M., Dhaouadi, R., Zualkernan, I.: An IoT-based remote IV tracing system for analysis of city-wide solar power facilities. Sustain. Cities Soc. **57**, 102041 (2020). https://doi.org/10.1016/j.scs.2020.102041
57. Muñoz, J., Lorenzo, E.: Capacitive load based on IGBTs for on-site characterization of PV arrays. Sol. Energy **80**, 1489–1497 (2006). https://doi.org/10.1016/j.solener.2005.09.013
58. Shapsough, S., Dhaouadi, R., Zualkernan, I., Takrouri, M.: Power prediction via module temperature for solar modules under soiling conditions. In: Deng, D.-J., Pang, A.-C., Lin, C.-C. (eds.) SGIoT 2019. LNICSSITE, vol. 324, pp. 85–95. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-49610-4_7
59. TensorFlow. https://www.tensorflow.org/, Accessed 20 July 2020
60. TensorFlow Lite | ML for Mobile and Edge Devices. https://www.tensorflow.org/lite, Accessed 20 July 2020
61. Yocto-Amp - Tiny isolated USB ammeter (AC/DC). https://www.yoctopuce.com/EN/products/usb-electrical-sensors/yocto-amp, Accessed 27 Mar 2017
62. nmon for Linux. https://nmon.sourceforge.net/pmwiki.php, Accessed 28 Mar 2017