A Low-cost System for High-frequency Solar Imagery and Power Data Acquisition

G F Bassous¹, R F Calili¹ and C R H Barbosa¹

¹ Postgraduate Program in Metrology, PUC-Rio, R. Marquês de São Vicente, 225, Gávea, Brazil

Email: calili@puc-rio.br

Abstract. With advances in solar energy research and increasingly accurate forecast techniques, intermittency no longer stands as a barrier to the adoption of solar energy. Coupling reliable data and knowledge on the inherent variability of the solar resource with advanced learning and forecast models, renewable energy can take an even bigger role in today’s energy paradigm. The objective of this work is to develop and test a low-cost data acquisition system able to provide relevant data for solar energy forecast models. The results yielded from the performed tests indicate high correlation between image derived attributes and power measurements 20 s ahead.

1. Introduction

For a long time, it’s been believed that the intermittency of Renewable Energy Sources (RES) is the main barrier to the massive adoption of greener energy sources, especially solar Photovoltaic (PV). In reality, evidence points towards politics and resistance to change from the traditional energy sector as being the main hindrance [1]. However, depending on the level of penetration, the inherent intermittence of PV energy may lead to several technical issues and inefficient harnessing of this resource [2,3], especially in systems operation, planning and scheduling.

In order to fully make use of solar PV energy, such problems caused by variability must be addressed. The first step in being able to address variability is to accurately forecast the behaviour of the PV system. The stochastic component of the variability can be split in terms of temporal and spatial resolution and, to each, its own optimal forecasting methods and models can be developed and applied [4].

The main issue preventing predictable PV operation is the presence of clouds casting shadows on the panel surface. Both theory and practice suggest that sky-image based approaches provide significant information for intra-minute and sub-kilometer forecasting [4,5], hence the decision to develop a low-cost system based on sky imagery. Passing clouds can trigger violent albeit short disturbances in global insolation at 1 minute intervals [6], but to accurately understand the effect of these disturbances on grid studies, higher frequency data must be used [7].

The goal behind having a low-cost system capable of providing data at the required frequencies is increasing the geographical dispersion of data acquisition sites, which can provide more abundant and relevant information on the subject. Thus, this manuscript presents the development and tests of a prototype for a low-cost data acquisition system capable of providing all-sky images with an acquisition frequency up to 1 Hz. Due to the lack of a suitable testbed capable of providing 1 Hz PV conversion data at the time of this study, this prototype also encompasses a 6 W solar panel.
This study is structured as follows: section 2 introduces sky imaging systems, design process and choices for hardware and software of the acquisition system. Section 3 contains information on testing circumstances and results obtained from measurement and data analysis. Finally, section 4 presents the manuscript conclusions and recommendations for furthering this investigation in future works.

2. Sky-imaging systems

Sky imagers or total sky imagers are devices capable of imaging the whole sky dome with the use of a 180° field of view (FOV) lens. Some have cameras directly pointed towards the sky, while others are pointed down to dome-shaped mirrors reflecting the entire sky. Sky imagers were originally developed for meteorological use, mainly cloud cover analysis. For that reason, many devices have solar-occlusion mechanisms to prevent the high intensity direct solar beam from saturating a portion of the image due to forward scattering [8]. Only in the past decade researchers have begun coupling sky imagers with solar forecasting, driving increased research focus on the area for intra-hour application [9].

Unfortunately, being of scientific grade, such equipment represents impeditive costs for the wide geographical dispersion needed for very short-term forecasting in an electrical system scale [10]. Even more so if one considers the solar resource availability in developing countries, whose currency inequivalence to the dollar or euro would further the inability to use these imagers in a national operation scale. Hence the need for affordable devices developed specifically for solar forecasting, with all the necessary functionalities cost-effectively wrapped in a small and simple equipment.

Some research centers have developed their own equipment tailored for solar forecasting, and their results show that the increased specificity of the devices yield more relevant information [8,10,11]. Another important feature for PV-specific sky imagers is the necessity of higher frequency acquisition compared to meteorological applications. Image acquisition frequency should be high enough to describe distinguishable events within the desired temporal resolution.

2.1. Hardware

In order to keep the costs as low as possible, the imaging system was built around a single board computer (SBC) [10]. The selected SBC was a Raspberry Pi 3B+ due to its easy access, low cost, ample documentation as well as easy access to Linux operating system. Connected to it via USB is an ELP-USBFHD01M-L180 camera consisting of a printed circuit board (PCB) module with a 1920 pixels x 1080 pixels CMOS OV2710 sensor and a 180° FOV lens.

In order to provide an electrical quantity related to solar generation on the same frequency as the image acquisition, a 20 cm x 15 cm solar panel with nominal capacity of 6 V and 1 A at 25°C was used. Current and voltage measurements were done using an Adafruit INA219 DC sensor and, since solar cell temperature has such a strong impact on conversion efficiency, a Maxim Integrated DS18B20 temperature sensor was attached to the bottom of the panel.

Figure 1 shows a basic 3D model of the equipment. The camera and solar panel sit atop the main support structure, enclosed by an acrylic dome and the 6 W PV panel used to acquire solar related quantities. On the bottom two housing units can be seen. One has the power supply components, 110 V AC to 5 V/3 A DC power supply for the SBC and minor cooling fans, plus a 5 V to 12 V DC-DC converter for the main cooling fan. The second unit houses the Raspberry Pi and INA219 DC sensor.
2.1.1. Measurement circuit
Both sensor boards have embedded analog-to-digital converters and are Raspberry Pi compatible. The temperature sensor was placed on an aluminium heat-exchange pad with the fins bent into contact with the entire sensor area in order to maximize heat transfer from the panel. The pad was then covered with packing foam to reduce heat loss to the atmosphere. As for the electrical quantities' measurement, the INA219 has an internal circuit capable of measuring DC up to 3.2 A by using a 0.1 Ω shunt resistor. A dichroic bulb was used as load on the solar panel, completing the circuit. Finally, the camera was powered and controlled via USB cable. Figure 2 shows a diagram of the measurement circuit.

Figure 1. Basic 3D model of data acquisition system.

Figure 2 Diagram of measurement circuit.
2.2. Software

The Raspberry Pi was running Raspbian Stretch, a Linux OS made for the SBC, with OpenCV 4.0 Python distribution and Python 3.7. OpenCV is an open source image processing and computer vision library available for several platforms. It was exclusively used for controlling the camera and saving the images. All post-processing was performed in Matlab 2017a. There are several Python libraries available for the INA219 and DS18B20 for use with the Raspberry Pi’s GPIO pins, the most notorious, and with largest online documentation being “pi_ina219” [12] and “w1thermsensor” [13]. The final program integrated these libraries to record images, power and temperature data on the required frequency for studying cloud induced intermittency on PV power.

As for the acquisition frequency, 1 Hz has proven to provide very useful information about the problem addressed [7]. However, one image per second for several days yields too much data, with part of it not being useable at all. Hence the decision to go with the principle of acquisition by exception. As the main goal is studying high-amplitude fast variations of power, the system keeps getting images and power measurements every second, but it does not record all the data. Only when triggered by a variation larger than a certain predefined threshold, does the system record the measured data.

In this case, the system keeps the previous ten measurements \( t_{10s} \ldots t_{1s} \) and calculates a moving average of the past 3 values, so if the absolute difference between the present measurement and the calculated average is larger than a certain threshold the data recording process begins. Aside from the past 10 s, it also records the present measurement \( t_0 \) and keeps recording for 4 more seconds \( t_{+1s} \ldots t_{+4s} \). Those 15 values with their corresponding images and time stamps, plus a temperature measurement, are then saved and the program goes back to listening for intermittence triggers. This data structure is called an “event” throughout this work. The reason for only one temperature measurement per event is that the sensor’s ADC often takes over 1 s to convert it, and this makes it impossible to achieve 1 Hz, but fortunately the temperature variations are minimal in only a couple seconds span due to the thermal inertia from the panel. Figure 3 presents the decision process and data flow from the acquisition software.

![Figure 3. Decision process and data flow from the acquisition software.](image)

3. Testing and results

The Data Acquisition System was placed on a balcony located at Lat 22.9354° S / Long 43.1756° W, at an altitude of 46 meters. Coordinates and altitude were obtained using Google Earth. Unfortunately, due to building obstruction, only the morning solar path was available, up to roughly 12 h 30 min.

Initially, the camera was protected by an acrylic dome; however, due to exposure to rain and consequently infiltration of water into the dome, too much condensation rendered possible sky images
unusable. A decision was made to remove the dome and to cover or expose the equipment depending on the forecast.

3.1. Testing
The first successful capture occurred on February 25th, 2019 at 09 h 21 min 48 s, with the acquisition period spanning until March 23rd, 2019 at 8 h 2 min 45 s. Records were taken for 12 non-straight days due to very clear or cloudy/rainy days in this period. Midway through the period, after analysing the first images, an attempt was made at darkening the images due to oversaturation of the solar region and a large area around it.

A neutral density filter was placed over the lens on March 14th, 2019 and, after adjusting the camera exposure parameter, the images appeared to have a smaller saturated area around the solar region. It also served as a shield protecting the camera from rain, so after its placement the equipment was no longer removed when it rained.

3.2. Results and data analysis
A total of 500 events were recorded over the twelve days. The first processing step was image subtraction to investigate the visible difference between each point in an event. For this, firstly the images were converted to grayscale, then the intensity matrices subtracted. The results depicted mostly noise, but on the scarce ones where more significant border information was available, the solar region was blacked out due to sensor saturation. This pointed towards an inadequacy of the event structure in yielding valuable information for a forecast model. Figure 4 shows image subtraction results for different sky conditions with a highlighted area tracing the solar region. Since the pixel differences are very small and therefore barely visible, the image contrast was increased for differences larger than 5.

Figure 4. Two examples of raw images from different cloudy days and three cropped image subtraction results for diverse sky conditions with highlighted solar regions.
Next step was to analyse the data as a continuous time series, only separating it by day due to the major differences between each one. The time series plots very clearly showed the fast high-amplitude variations as can be seen in figure 5.

Figure 5. Time series plot of power measurements for March 14th, 2019.

The next goal was to find some relationship between the images and power data. The chosen attribute to be analyzed first was the energy of the pixels in a region of interest (ROI) around the sun. To achieve that, a derived dataset was put together with the power measurements, temperature measurements and ROI energy. Several different ROI radii were used along with differential values as a function of a time interval \( \Delta t \). The differential values used were for power, temperature and 6 different energy values: absolute difference for each image channel and image subtraction for each channel. The image channels represent red, green and blue components of the colored images and the distinction between absolute difference and image subtraction is, whilst the former allows negative numbers, the latter does not.

The final group of variables were: power at \( t_0 \) and \( t_0 + \Delta t \); power difference between those two measurements; temperature at \( t_0 \) and \( t_0 + \Delta t \) (for some instances temperature values had to be filled in using a linear method); temperature difference between those two measurements, absolute difference between the ROI energies at \( t_0 \) and \( t_0 + \Delta t \); energy of the ROI from the image subtraction at \( t_0 \) and \( t_0 + \Delta t \).

In total, 70 combinations of \( \Delta t \) and ROI radius pairs were computed for each of the 12 days, with \( \Delta t = \{1; 2; 5; 8; 10; 15; 20; 30; 45; 60\} \) seconds and ROI radius = \{25; 50; 75; 100; 150; 200; 250\} pixels. Then, the correlation coefficients were calculated for each combination for each day, resulting in 840 matrices of 12x12 correlation coefficients between the aforementioned variables. In order to more easily evaluate the results, the Identity matrix was subtracted from the correlation matrices, then the first 6 columns were binarized between larger and smaller than 0.8, then summated columnwise. This step produced 840 vectors of size 6 with the amount of correlation coefficients larger than 0.8 in each of the first 6 columns.

Through these compiled metrics, it was possible to determine the combination with the highest count of higher (> 0.8) correlation coefficients between the first 6 variables and all twelve. The best results happened with \( \Delta t = 20 \) s and ROI radius = 150 pixels. The compilation of the 840 vectors is presented in table 1.
Table 1. Compilation of vector coefficients for each combination of $\Delta t$ and ROI radius.

| ROI RADIUS | $\Delta t$ | 2 | 5 | 8 | 10 | 15 | 20 | 30 | 45 | 60 |
|------------|------------|---|---|---|----|----|----|----|----|----|
| 25         | 0 | 2 | 9 | 12 | 17 | 19 | 8  | 4  | 19 |
| 50         | 0 | 2 | 9 | 11 | 17 | 20 | 8  | 8  | 25 |
| 75         | 0 | 2 | 9 | 12 | 18 | 21 | 8  | 7  | 25 |
| 100        | 0 | 2 | 9 | 12 | 18 | 20 | 9  | 7  | 24 |
| 150        | 0 | 2 | 9 | 12 | 18 | 25 | 9  | 4  | 22 |
| 200        | 0 | 3 | 9 | 12 | 16 | 23 | 7  | 4  | 17 |
| 250        | 0 | 3 | 11| 12 | 15 | 19 | 7  | 4  | 16 |

Figure 6 presents power measurements 20 seconds ahead from the reference point and ROI energy differences with $\Delta t = 20$ s for ROIs with 150-pixel radii. Values were normalized to help visualization. In blue, energy data from image subtraction and in orange, energy from absolute image difference.

Figure 6 makes the correlation between the attributes derived from image processing, – in blue and orange – and power measurements taken 20 s after the reference point – in green – visually apparent. Both curves vary proportionately on the same points save for some apparent outliers near 10:00. This was numerically determined in the previous step in data processing. Temperature measurements also boasted a good correlation coefficient but due to a low range of values they were visually inapparent and therefore kept off the plot to ensure clear visualization of the most critical relationship.
4. Conclusion and Future works
This work presented the development and testing process of a low-cost data acquisition system aimed at providing data for solar energy forecast models. The initial assumptions using an event-based approach did not yield relevant information, but upon further data analysis, high correlation was found between images and power differences at a horizon of 20 s. Future work should focus on three aspects:

- Increasing ruggedness, reliability and control of the data acquisition system;
- Validating the acquired data by using a nonlinear regression model such as a neural network; and
- Increasing the geometrical complexity of the problem by using power and temperature data from an array of PV panels.

Acknowledgements
The authors thank for the financial support provided by the Brazilian funding agencies CNPq, CAPES, FINEP, and FAPERJ. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brasil (CAPES), Finance Code 001.

References
[1] Sovacool B K 2009 The intermittency of wind, solar, and renewable electricity generators: Technical barrier or rhetorical excuse? Util. Policy 17 288–296
[2] Karimi M, Mokhli H, Naidu K, Uddin S, and Bakar A H A 2016 Photovoltaic penetration issues and impacts in distribution network - A review Renew. Sustain. Energy Rev. 53 594–605
[3] Denholm P and Hand M 2011 Grid flexibility and storage required to achieve very high penetration of variable renewable electricity Energy Policy 39 1817–1830
[4] Diagne M, David M, Lauret P, Boland J and Schmutz N 2013 Review of solar irradiance forecasting methods and a proposition for small-scale insular grids Renew. Sustain. Energy Rev 27 65–76
[5] Sobri S, Koohi-Kamali S and Rahim N A 2018 Solar photovoltaic generation forecasting methods: A review Energy Convers. Manag. 156 459–497
[6] Mills A and Wiser R 2010 Implications of Wide-Area Geographic Diversity for Short-Term Variability of Solar Power 1–10
[7] Lave M, Reno M J and Broderick R J 2015 Characterizing local high-frequency solar variability and its impact to distribution studies Solar Energy 118 327–337
[8] Urquhart B, Kurtz B, Dahlin E, Ghonima M, Shields J E and Kleissl J 2015 Development of a sky imaging system for short-term solar power forecasting Atmos. Meas. Tech. 8 875–890
[9] Yang D, Kleissl J, Gueymard C A, Pedro H T C and Coimbra C F M 2018 History and trends in solar irradiance and PV power forecasting: A preliminary assessment and review using text mining Solar Energy 168 60–101
[10] Richardson W, Krishnaswami H, Vega R and Cervantes M 2017 A low cost, edge computing, all-sky imager for cloud tracking and intra-hour irradiance forecasting Sustain. 9 1–17
[11] Gohari S M I et al 2013 Comparison of solar power output forecasting performance of the Total Sky Imager and the University of California, San Diego Sky Imager Energy Procedia 49 2340–2350
[12] Borrill C, Timmons T, van Staveren T, Kluiver T and Bauer S 2019 pi_ina219 [Online] Available: https://github.com/chrisb2/pi_ina219 [Accessed: 24-Jul-2019]
[13] Furrer T 2019 w1thermsensor [Online] Available: https://github.com/timofurrer/w1thermsensor [Accessed: 24-Jul-2019]