Research Article


Maximus Byamukama,1 Geoffrey Bakkabulindi,1 Roseline Akol,1 and Julianne Sansa-Otim2

1Department of Electrical and Computer Engineering, College of Engineering, Design, Art and Technology, Makerere University, Uganda
2Department of Networks, College of Computing and Information Sciences, Makerere University, Uganda

Correspondence should be addressed to Maximus Byamukama; maximus.byamukama@cedat.mak.ac.ug

Received 4 December 2018; Accepted 6 March 2019; Published 27 August 2019

Copyright © 2019 Maximus Byamukama et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The development of perpetually powered sensor networks for environment monitoring to avoid periodic battery replacement and to ensure the network never goes offline due to power is one of the primary goals in sensor network design. In many environment-monitoring applications, the sensor network is internet-connected, making the energy budget high because data must be transmitted regularly to a server through an uplink device. Determining the optimal solar panel size that will deliver sufficient energy to the sensor network in a given period is therefore of primary importance. The traditional technique of sizing solar photovoltaic (PV) panels is based on balancing the solar panel power rating and expected hours of radiation in a given area with the load wattage and hours of use. However, factors like the azimuth and tilt angles of alignment, operating temperature, dust accumulation, intermittent sunshine and seasonal effects influencing the duration of maximum radiation in a day all reduce the expected power output and cause this technique to greatly underestimate the required solar panel size. The majority of these factors are outside the scope of human control and must be therefore be budgeted for using an error factor. Determining of the magnitude of the error factor to use is crucial to prevent not only undersizing the panel, but also to prevent oversizing which will increase the cost of operationalizing the sensor network. But modeling error factors when there are many parameters to consider is not trivial. Equally importantly, the concept of microclimate may cause any two nodes of similar specifications to have very different power performance when located in the same climatological zone. There is then a need to change the solar panel sizing philosophy for these systems. This paper proposed the use of actual observed solar radiation and battery state of charge data in a realistic WSN-based automatic weather station in an outdoor uncontrolled environment. We then develop two mathematical models that can be used to determine the required minimum solar PV wattage that will ensure that the battery stays above a given threshold given the weather patterns of the area. The predicted and observed battery state of charge values have correlations of 0.844 and 0.935 and exhibit Root Mean Square Errors of 9.2% and 1.7% for the discrete calculus model and the transfer function estimation (TFE) model respectively. The results show that the models perform very well in state of charge prediction and subsequent determination of ideal solar panel rating for sensor networks used in environment monitoring applications.

1. Introduction

Environment monitoring systems are devices with electronic sensors and sensor networks that are deployed outdoors to quantify weather elements and are typically powered using solar energy. The traditional sizing technique for solar PV panels (solar panels) is a computation involving the hours of sunshine per day, the wattage of the loads, and the power output of the solar panel. The solar panel size is determined from

$$P_S = \frac{P_L t}{N}$$

(1)

where $P_L$ is the wattage for the load that draws power for $t$ hours and $N$ is the expected number of sunshine hours per day, which varies by season. This power generated by a solar
panel depends on the geographical location of the installation site, season of the year, time of the day, and the current local weather conditions. As such, the power output from the panel at any given time can be quite varied. Any calculation, hence, that is based on a blanket number representing total sunshine hours in order to calculate the energy produced from the panel, will be inherently flawed if these factors are not quantified to make the necessary adjustments.

The panel size estimated from this equation can be improved using analytical techniques. Analytical techniques are largely used to estimate the amount of energy harvested by the solar panel using statistical weather models, such as the approach used in [1]. These are then coupled with engineering models of the components inside. Analytical techniques are meant to obtain an error factor to account for the short term weather variations that affect the power output and other system-specific variations such as

(i) temperature effects: higher temperatures have been long known to reduce the solar panel efficiency [2, 3] and the charge and discharge rate for many battery technologies [4–6];

(ii) component efficiency: electronics like charge controllers, DC-DC converters, and regulators consume varying power from depending on the level of input voltages from the solar panel and operating currents.

Using analytical techniques is effective in modelling individual components in a system but modelling the system as a whole is often a challenging task. One challenge is that they often model the load and its power electronics in an unsophisticated way that excludes the detail in the system behavior which, in the case of environment monitoring sensor networks, may be caused by changes of data acquisition intervals, network connection, transmission times, etc. that deeply influence the energy demand of the system. As such, designers are changing the sizing philosophy to use actual recorded and highly localized solar radiation data to size solar systems for given locations [7, 8]. Apart from [9], all the related work we have seen is similar to the latter and deals with sizing solar systems for large scale generation for domestic usage and grid supply. We have not found any work dealing with sizing energy harvesting units of any type for very low power DC systems such as those in wireless sensor networks. The work in [9] though is a simulation and inherently does not involve all the complex interactions of the factors that have been stated.

For such loads, a more reliable technique would have to be evidence-based—where measurements of the actual solar insolation and the actual accumulated energy of the battery caused by the insolation under a given load in a real deployment are used to establish any input-output relationships. These measurements can act as input and output data that can be used to model the whole gateway device as a black box using a variety of techniques. The model can then be used with historic weather data for a given area to predict future battery state of charge (SoC) and the solar panel size that will be required to minimize the likelihood of the battery being drained below this SoC.

Section 2 discusses the preliminary experimentation that gave the motivation for this work. It also discusses the experimental setup that was used to collect the data to be analyzed. Section 3 discusses the system identification and modelling techniques being proposed. It also discusses the results and prediction accuracies for both models. We then use the more robust model to show how future battery SoC can be generated using a reference solar insolation profile. Section 4 concludes and gives recommendations for further work.

2. Materials and Methods

2.1. Preliminary Experimentation. The motivation for this research comes from the preliminary observations of the battery SoC and voltage profiles of two WSN gateways. The first deployment was a low power WSN gateway, designed and implemented following the guidelines in [10]. The gateway uses the RS mote [11] as a sink node and the Electron 3G [12] as the uplink device. The system consumes about 55mW (15 mA at 3.7 V). The second gateway is an embedded Linux gateway based on the Raspberry Pi [13] and using the same mote as a sink node. The system power consumption was about 1250 mW (250 mA at 5.0 V) which evaluates to approximately 30000 mWh per day. The two gateways are installed at 0.3292°N, 32.5710°E.

Using (1) and considering a daily solar cycle, the first system requires 1320 mWh per day. This experiment lasted 13 days in late February 2018. The average peak sunshine hours during this time were 5.65, calculated using the NASA Daily Agroclimatology solar insolation dataset [14] for these coordinates. With these hours of peak sunshine, the ideal panel size is calculated as 233 mW from (1). The second gateway was deployed for 6 days in November 2017, in which the number of peak sunshine hours was 5.20. The ideal panel size for this setup is 5.764 W from (1).

Figure 1 shows the battery SoC profile for a 2000 mAh Li-ion battery powering the first system in which a 1W solar panel was used as the energy harvesting unit. This panel size is over 4 times the calculated value of 220mW. Figure 2 shows the voltage of the Lead Acid battery that was used to power the second system. Here, we used a 30W solar panel as
the energy harvesting unit, over 5 times the required panel size. This experiment lasted 6 days in November 2017. The expected charge profile in both cases should indicate periods of discharge at night followed by full charge during the day to 100% and 14V, respectively. However, the trend of the profile indicates that the energy accumulated during the charging time was not sufficient to meet the load demand even with panels sized with factor of ×4 and ×5, respectively. The peaks in the charging profile reveal that the batteries did, in fact, accumulate charge, but it was not sufficient to sustain the operation of the WSN gateways. These experiments were carried out in the second half of February and November. The rainfall season in Uganda is biannual and the seasons in March to May (MAM) and September to November (SON) are characterized by limited intermittent sunshine, more rainfall, and cloudy days [15, 16]. We believe that the short term variations in solar insolation at a given location, especially in rainy seasons, account greatly for the observed results and that proper solar panel sizing must involve some modelling of the relationship between the actual observed solar radiation and the subsequent change in battery SoC it imparts.

2.2. Experiment Setup. The experiment setup consists of a modified implementation of the low power gateway already discussed in Section 2.1. The device consists of a sink node, an SD card for local storage, an uplink device, a 2000mAh LPI03450 Li-ion battery, a TP4056 battery protection module, and a 2W solar panel with dimensions 110 × 136 mm. Figure 3 shows the physical implementation. The uplink collects battery state of charge information approximately every 3 minutes. The solar radiation data is measured by the SP-Lite Silicon Pyranometer, part of an industry-grade automatic weather station from ADCON telemetry [17] managed by the Uganda National Meteorological Authority. We collected over 16500 evenly distributed samples for the battery SoC over a period of 45 days. The solar radiation data was recorded every 15 minutes. The 15-minute interval for recording solar radiation is a fixed setting that could not be changed in this experiment. Transmitting to the gateway are about 5 sensor nodes sending weather data reports such as temperature, humidity, wind speed and direction, atmospheric pressure, soil moisture, and soil temperature. The gateway uploads data after receiving 700 reports. Each report is approximately 150 bytes long.

3. System Identification and Modelling

3.1. Theory. System identification refers to the process of generating a mathematical model of a system’s nature by observing the input and output signals it generates in a known period of time. The behavior of standard systems, like a mass on a spring, can be modelled using mathematical techniques without the need to perform an actual experiment, if various variables like the mass, spring length, elasticity, etc. are known. In many other cases, however, the behavior of a system cannot be modelled mathematically either because its internal components are unknown or because their interactions are too complex to analyze effectively. In these cases, the system can be regarded as a black box, and only these observed inputs and outputs are used to estimate its behavior. The methodology of system identification follows the following steps:

(i) Identify the input into and output from the system.
(ii) Sample the inputs and outputs at known periods of time.
(iii) Split the observed data into two (or more) datasets. These datasets are used separately for building the model and for validating it.
(iv) Visualize the input and output datasets simultaneously. This step is important because it gives the ability to make an initial classification as to whether the system is linear or nonlinear.
(v) Generate a mathematical function that produces outputs from observed inputs, while minimizing the error between these outputs and actual observed outputs. This is an iterative step and is usually achieved by a computer program.
(vi) Apply this function to the validation dataset and observe differences between actual and observed outputs. These differences are the prediction errors.
(vii) Repeat (iv) and (v) until prediction errors are acceptable in the experiment context.

Systems may be purely linear, purely nonlinear, or linear with varying extents of nonlinear distortion. Classification is thus an important part of this process because it will enable the designer to decide whether established techniques of modelling may be sufficient or if a new technique is needed.

3.2. Classification of the WSN Gateway. The ideal size of solar panel needed to sustain the gateway operation will be one which, given a particular insolation profile, can guarantee a minimum battery SoC for the system in a desired time range. The solar insolation modulates the electrical power from the solar panel. This power is both consumed by the
Figure 3: Physical implementation of an environment monitoring WSN gateway showing the sink node, uplink, and power system components.

Figure 4: (a) Variation of solar radiation and power output from photovoltaic cell during the day. (b) Scatter plot showing linearity of this variation.

gateway device and accumulated into the battery in varying proportions which determine the state of charge at any given time. As such, the battery SoC is the response variable and can be looked at as the output.

While the actual cause of change in the battery SoC is the electrical power output from the photovoltaic panel, it has been shown already in [18, 19] that, in general, photovoltaic electrical power output varies linearly with the incident solar radiation. Figure 4, adapted from [18], illustrates this relationship for a particular photovoltaic cell. The correlation between the two is very strong and, as such, solar insolation can be regarded as the input signal. Figure 5 further shows how the battery SoC responds to the incident solar radiation profile over four selected days in June 2018.

The power consumption of the system is roughly constant; it is in a low power state consuming approximately 15 mA for about 60 minutes and small regular bursts of 280-400 mA for about 40 seconds during data transmission to the server. It therefore has constant power consumption approximately 99% of the time. From this information, we expect the relationship between solar insolation and battery state of charge to be approximately linear and time invariant. Figure 6 confirms our expectation. Here, we show the scatter plots of the variation in Figure 5 for a single day, split into the charging phase and discharging phase. The relationship in Figure 6 shows a very strong positive linear correlation between the solar insolation incident on the solar panel and the corresponding battery SoC.

We conclude thus by classifying our system as being Linear Time Invariant (LTI).

3.3. Modelling Techniques. In the techniques presented, we used half the data, about 23 days, to estimate model the system and the other half was used as validation data to test the model. A 3-minute interval dataset was generated from the solar radiation data by running a spline interpolation which produces data points corresponding to the points at which the SoC was measured while maintaining the exact shape of the original profile.

3.3.1. Discrete-Calculus Technique

(i) Theory. The term discrete calculus is used to show that the method concerns itself with the accumulation of the areas under the solar radiation curve and the incremental change this imparts on the battery SoC. Solar radiation data, in
$\frac{W}{m^2}$, is easily converted to $W$ by multiplying by the area of the PV panel. Figure 7 illustrates the theory behind this technique.

Considering a small unit of time, $\Delta t$, it is possible to evaluate the solar energy, in Wh, that arrives at the surface of the panel. This energy is equivalent to the area of the shaded trapezoid and is calculated to be

$$E = \frac{1}{2} \Delta t (W_1 + W_2)$$

(2)

where $W_1$ and $W_2$ are the observed incident solar radiation values at a time $t$ and $t + \Delta t$, respectively. This energy $E$ yields an equivalent electrical output energy $E_e = \eta E$ where $\eta$ is the solar panel efficiency. If $P$ is the power consumption of the load, the electrical energy is equal to the energy consumed by the load, $P\Delta t$, and the energy accumulated by the battery to cause the change in the battery state of charge equal to $\Delta C$. The change in battery state of charge is in $Wh$ and has the same dimensions as energy. Hence,

$$E_e = \eta E = P\Delta t + \Delta C$$

(3)

In a time-invariant system, one in which the power consumption does not change, the term $P\Delta t$ is a constant. The solar panel efficiency can be assumed to be also constant. In
practice, this efficiency varies slightly with temperature of the PV module. In [20, 21], it has been shown that the reduction in overall output efficiency is only about 1-2% in the 25° – 60° C temperature range.

Hence, it is observed that the change in battery SoC has a linear relationship with the incident solar insolation as per (4).

\[ \Delta C = \eta E - P \Delta t \]  

(4)

It can be seen from (4), that when \( E = 0 \), for example, during nighttime, the change in battery state of charge is a reduction equivalent to the energy consumed by the gateway device.

In an ideal time-invariant system, a plot of \( \Delta C \) vs. \( E \) should yield a straight line and thus a correlation of 1.

(ii) Results. Starting at a time \( t \), the algorithm in the model takes an arbitrary value of \( \Delta t \) and calculates the areas of subsequent trapezoidal strips in the dataset, which are equivalent to \( \eta E \), and then determines the incremental change in the battery SoC, \( \Delta C \), by subtracting successive SoC values at the various points \( t + n \Delta t \) where \( n = 1, 2, 3 \ldots \).

Figure 8 shows a scatter plot of the observed values of \( \Delta C \) on the vertical axis and \( \eta E \) on the horizontal axis. The measured linear correlation between the two is 0.605, which is classified as a moderate positive correlation. The limitations in Section 3.6 explain the factors that affect this correlation. The algorithm then generates a linear equation of the line of best fit between the two datasets. This linear equation is used to generate the expected change in \( \Delta C \) given a known value of solar insolation.

These errors can be mapped against \( \{ P_i \} \) to establish a line best fit that expresses the errors as a function of the predicted value,

\[ \epsilon_i = P_i - O_i \]  

(5)

These errors can be evaluated for all values in \( \{ P_i \} \) and the optimized predicted values \( \{ P_i^o \} \) are generated using the equation

\[ P_i^o = P_i - f(P_i) \]  

(7)

This reduced the observed RMSE by up to 2.8% from 11.86 to 9.20. The correlation coefficient of 0.844 shows that the model is also quite strong in predicting the charge-discharge pattern and places times the maxima and minima very well.
3.3.2. Transfer Function Estimation

(i) Theory. In LTI systems, transfer functions are used to generate outputs from known inputs. In many cases, transfer functions can be derived mathematically. However, when a system behavior is unknown but its input-output data is available, the system’s transfer function can be estimated and used for prediction of output given known inputs. Figure 10 shows how the gateway setup in Figure 3 can be interpreted as a black box and modelled as a single-input single-output (SISO) control system. The input $x(t)$ is the solar radiation incident on the solar panel in Watts. The output $y(t)$ is the observed battery state of charge. The model of the behavior of the black box, $h(t)$, is the system transfer function in the time domain. Usually, system identification is carried out in the frequency domain. For discrete time systems, the input, output, and transfer functions are $X(z)$, $Y(z)$, and $H(z)$ in the frequency (Z-) domain, respectively.

The transfer function was estimated using the tfest function in MATLAB (version 2018a), with an arbitrary selection of 3 poles and 2 zeros for stability. We then used the compare function to generate the predicted output.

(ii) Results. Figure 11 shows the variation of the predicted and the actual observed battery SoC values in the validation period. There was a strong positive correlation of 0.934 and an RMSE of 1.695.

3.4. Discussion. The correlation coefficient shown in Figures 8, 9, and 10 is the Pearson product-moment correlation coefficient and is a measure of the strength of the linear relationship between two variables. The correlation values of 0.844 and 0.935 indicate that the models are very accurate in predicting the pattern of the state of charge profile. They indicate that the models predicted the charging and discharging times of the battery and placed the peaks and

\[ r = 0.844 \]
\[ \text{rmse} = 11.86 \]
\[ r_m = 0.844 \]
\[ \text{rmse}_m = 9.20 \]
valleys of the profile very well. This can also be visually observed in Figures 9 and 11. A visual analysis reveals that the models overestimated a few SoC maxima and underestimated some minima but had accurate predictions on most of the SoC range. The most important prediction error statistic is the root-mean-square error (RMSE). The calculated RMSE values of 9.2 and 1.70 are equivalent to 9.2% and 1.70%, respectively, since the SoC varies from 0 to 100. This means that, on average, for the TFE technique, for example, we can expect only a 1.7% error margin in the predicted dataset. The discrete calculus model shows a weaker performance than the transfer function estimation (TFE) model with the difference in RMSE being 7.5 units. This means that, on average, the model will underestimate or overestimate the SoC by 7.5 units compared to the TFE technique. However, there are some primary advantages to this technique:

(i) It is very fast and will allow designers to run quick simulations at lower accuracy. The tfest function is an inbuilt MATLAB function with over 500 lines of code and runs multiple iterations until a good fit is obtained. It depends on several other inbuilt functions. Because of this, the tfest code took an average of 33.1 seconds to run to completion while the discrete calculus model lasted only 5.1 seconds on average on the same machine. This is advantageous when designers plan to simulate and deploy very many sensor nodes over a large expanse of land with varying solar radiation profiles.

(ii) The discrete calculus model can be coded easily in another language. The TFE technique relies on complex preliminary algorithms to estimate initial values of the transfer function coefficients such as those in [22].

3.5. Solar Panel Size Simulations. During the development of either model, a standard Li-ion battery of known capacity and solar panel of known wattage and dimensions are used. The model is specific to this system. Two adjustments need to be included in the simulation code for systems with different battery capacities and different solar panel sizes.

1. For a model developed with a battery of capacity $C_1$, to simulate with a new battery of capacity $C_2$, the observed successive changes in the output SoC, $\Delta C$, need to be multiplied by $C_1/C_2$. These new differences are then, together with the initial SoC, used to generate the new SoC profile.

2. To simulate and observe the output SoC for a different panel size, the input solar insolation is multiplied by the ratio of the new (required) panel size to the one used in the development of the model. This is an accurate approach because it has already been proven in Section 3.2 that electrical power output varies linearly with solar insolation. In our model, for example, the panel size used is $110 \times 136$ mm and its wattage is 2W. Its area is $149.6 \, \text{cm}^2$. This area is multiplied by the solar radiation in $\text{W/m}^2$ to obtain the solar insolation in $\text{W}$. To model the system with a solar panel of $3 \, \text{W}$, the panel area is multiplied by 1.5 to give $224.4 \, \text{cm}^2$. The solar insolation data to be used is generated using the same solar radiation data but multiplied with the new panel area.

Figure 12 shows one a backward prediction using the TFE technique for the period Jan–June 2018 starting from an initial SoC of 100% and with the solar panel size changed to 3W. We notice two things. First, the SoC is predicted to reach values of up to 130% correctly implying an abundance of energy from
the larger panel and secondly, there is a negative trend from March to May, in line with the low solar insolations of the season as stated in [15, 16]. The estimated transfer function can be used to back-predict the SoC profile using historic solar radiation data. A possible application of back-prediction is to analyze the performance of the system in seasons that are well known to have poor sunshine. By obtaining averages of historic data spanning back a few years, a representative insolation profile can be generated and used as the input into this model. The corresponding output is then used as a reliable estimate for future state of charge values for a given system. To ensure that the battery SoC never goes below a given threshold, the designer iteratively increases or decreases the panel size and runs the simulation again.

In conclusion, transfer function theory in control engineering generates reliable results when extended to solar panel sizing applications using measurements of solar radiation and accumulated energy in the battery.  

3.6. Limitations. Since this was an outdoor uncontrolled experiment, we are aware of some limitations that could have introduced some nonlinearities and time-variance and affected the accuracy of our results. These limitations are the reason why the observed correlation in Figure 8 is 0.605 instead of the expected 1.0.

Foremost, the official solar insolation data from the National Meteorological Authority is measured every 15 minutes and yet the SoC is measured approximately every 3 minutes. Solar radiation is quite dynamic, especially in cloudy days and may change many times during this interval. As such, it is likely that the calculation in (2) underestimates or overestimates some values in the dataset used to generate the model. This affects the TFE technique as well. The solution to this is to use high-resolution solar radiation values.

Secondly, the WSN is deployed with a cellular gateway. The power consumed during transmission will vary depending on the cellular signal strength since the connection time and time to upload data will increase. These factors were outside the scope of our control and the errors introduced cannot be modelled mathematically.

Thirdly, some reports from the sensor nodes may not reach the gateway because of poor signal strength. This causes the gateway to wait a bit longer in between some transmissions and shorter in others. Hence, the power consumed will vary slightly.

4. Conclusions

Autonomous gateways and sensor nodes in environment monitoring wireless sensor networks can be modelled as linear systems. Using the solar insolation profile incident on the solar photovoltaic panel as the input and the observed battery state of charge as the output, the input-output relationship can be effectively evaluated by using the suggested techniques of transfer function estimation and discrete calculus. The two techniques both give strong prediction accuracy and low error magnitudes between observed and predicted values. The discrete calculus technique can be used for fast rough estimation and the transfer function estimation technique can be used for simulations where accuracy is more important than speed.

Data Availability

The data and source code used in this research have been provided and may be made freely available.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was funded by the Norwegian Agency for Development Cooperation (NORAD) under the WIMEA-ICT project under grant number UGA-13/0018.
**Supplementary Materials**

obj3.mat contains tabular data of timestamps, voltage, solar irradiance, and battery state of charge values as measured in the experiment. obj3dc.m and obj3tfe.m contain the source code for the two techniques presented in this paper. (Supplementary Materials)

**References**


