

Technical Report: Energy Source Prediction for Wireless Mesh Sensor Networks

Anthongy Harris, Robert Cox, and Asis Nasipuri
Department of Electrical and Computer Engineering
University of North Carolina, Charlotte

1 Summary

This document describes the design and implementation of a solar energy prediction algorithm for wireless mesh sensor networks (WMSNs). Section 2 provides an overview of general concepts that will be built upon throughout. These general concepts range from assumptions made from a global standpoint to issues impacting harvestable energy by an individual mote. In Section 3, the nuts and bolts behind the prediction approach are explored and explained in detail. Finally, Section 4 shows the final algorithm used in the prediction model for each node within a wireless mesh sensor network as well as quantifiable results showing the benefits of such a predictor.

2 Problem Definition

Solar energy predictions are used to determine attainable solar energy at a given location and are applicable to a wide range of applications. One such applications is a solar harvesting wireless mesh sensor network, in which the network attempts to operate in what is known as a neutral energy state, i.e. the energy into and out of the system are equal. For a network-wide power scheme, like cooperative power control (CPC), to work effectively, day ahead predictions are required about available harvestable energy at each node. Accordingly, this document begins with an overview of general concepts that are required prior to a predictor of these sorts can be developed. These general concepts will vary from how to determine a regional value for harvestable energy to issues that can arise when attempting to make predictions from node-to-node.

2.1 Review of Appropriate Concepts

Photovoltaic cells convert solar energy into electrical energy. In silicon-based devices, photons enter the semiconductor structure. The absorbed energy creates electron-hole pairs. Single cells produce voltages and current densities on the order of 0.6V and 10mA/cm², respectively. The delivery of appreciable power thus requires an appropriate series and parallel combinations of individual cells. Photovoltaic device design is a complex subject, and a complete treatment is far beyond the scope of this work. It is important in this context, however, to provide at least a brief discussion on the fundamental energy conversion process.

Most importantly, individual photovoltaic cells are designed to respond to a particular subset of wavelengths in the solar radiation spectrum. Figure 1 shows the spectral response of a typical cell, as well as the solar spectrum at both the top of the atmosphere and at sea level. Note that the spectral response is defined as the ratio of the output current to the incident solar power at a given wavelength. The rate of energy conversion is thus directly related to the total radiant flux incident on the panel. Given that panel size is highly variable, we often refer to radiant flux density. This quantity is termed irradiance and is specified in units of W/m². For reference, Table 1 contains common radiometric terms and their units.

Although rudimentary, the preceding analysis clearly shows that the target for any solar-harvesting algorithm should be the total solar irradiance incident upon a particular panel. Given the inherent variability in panel placement, it is useful to focus on a standardized measurement known as global horizontal irradiance (GHI). This quantity is the total radiant flux on a horizontal surface and it represents the combined effect of

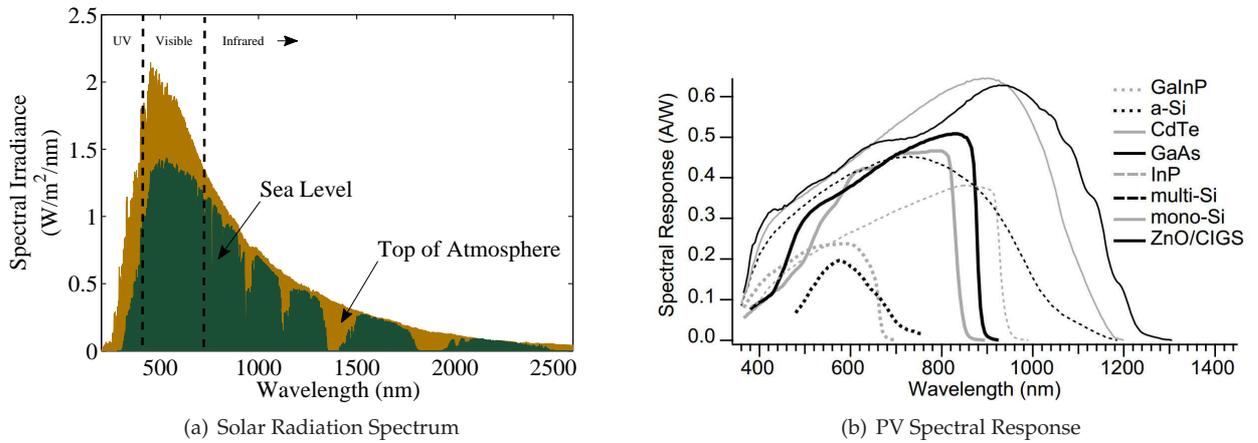


Figure 1: (a) Solar radiation spectrum at the top of the atmosphere [15] and at sea level [1]. (b) Spectral response of selected PV cells [5].

Table 1: Radiometric quantities and units.

Quantity	Symbol	Unit
Radiant Flux	Φ	W
Irradiance	I	W/m^2
Radiant Energy	Q	J
Energy Density	E	J/m^2

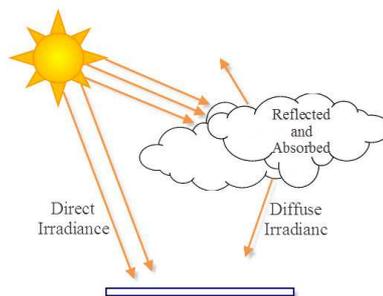


Figure 2: Pictorial representation of GHI.

the two different sources of radiation shown in Fig. 2. The first of these is known as direct normal irradiance (DNI), and it is a measure of the total radiant flux passing through a plane normal to the sun’s direct rays. An additional component results from the sunlight that is scattered as it passes through the atmosphere or reflected from the ground. The measure of this flux on a horizontal surface is known as diffuse horizontal irradiance (DHI). GHI is the sum of this quantity and the projection of the direct normal irradiance onto the horizontal surface, i.e.

$$\text{GHI} = \text{DHI} + \text{DNI} \times \cos(\Theta_Z). \tag{1}$$

Figure 3 shows that Θ_Z , which is termed the solar zenith angle, is defined between the geometric center of the sun’s disk and a line horizontal to the surface. Figure 4 shows the anticipated evolution of GHI over the course of a day. Noting that as the zenith angle approaches 90° GHI approaches 0.

2.2 Practical Factors Affecting the Prediction Problem

The previous section showed that GHI is a useful measure of the radiant flux at a given horizontal surface, and thus its prediction based on the movements of the Earth and its atmospheric conditions would seem to be a reasonable starting point for the development of a prediction algorithm. In reality, however, GHI is also affected by other practical issues. Consider, for instance, a network of rechargeable wireless sensor nodes deployed to monitor some outdoor infrastructure. Since node placement is driven by sensing needs, it is likely that individual nodes will see variations resulting from the effects of shading, reflections, and non-horizontal orientation.

Figure 5 clearly demonstrates the practical issues affecting GHI using the example of two nodes in EPIC-Roofnet, which is a small network deployed atop the Energy Production and Infrastructure Center (EPIC)

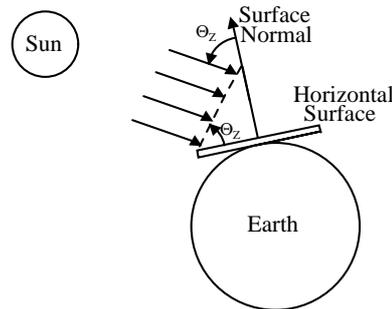


Figure 3: Geometrical representation of Θ_Z for a horizontal surface.

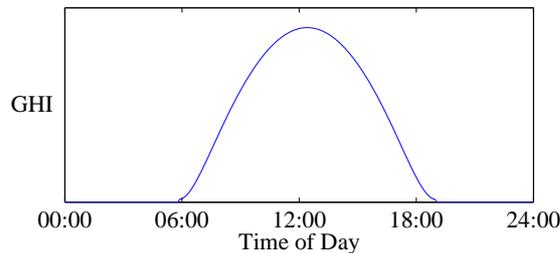


Figure 4: Anticipated trend of GHI over the course of a day.

at UNC Charlotte. As shown, node 152 has an un-obstructed view of the sky, whereas node 157 is placed between two large exhaust vents that provide significant shading throughout much of the day. Furthermore, note that the latter node is oriented vertically, meaning that it will naturally see a reduced contribution from the direct normal component of the radiant flux. Both nodes are equipped with pyranometers measuring the GHI at their respective locations. Figure 5(c) and 5(d) show two representative measurement series recorded on a clear day in April 2013. Note that the GHI observed by node 152 follows closely to the expected general shape. It is interesting to note that the peak value is approximately 1100 W/m^2 . It can be shown, however, that the combined effects of diffuse and direct irradiance should not yield a total GHI above approximately 1000 W/m^2 at the surface of the Earth [13]. Reflections from the surrounding environment have thus caused the total radiant flux to be larger than that typically expected from theoretical predictions. Similarly, note that the vertical orientation and shading cause a significantly different pattern to be observed at node 157. Note that the shading is minimized slightly during a short period between about 10:30 and 11:30 AM. Such patterns were observed frequently throughout our previous field deployments [14].

In addition to the spatial effects noted above, individual nodes are also subject to various temporal variations. This phenomena can be classified into high and low frequency categories. Weather patterns, for instance, can cause high frequency fluctuations as passing clouds can momentarily reduce direct irradiance. Figure 6 demonstrates this effect by presenting the observed GHI at node 152 on three different days. Such high frequency phenomena, which can be observed even on relatively clear days, are known to be extremely difficult to predict on time scales more than several minutes in advance [12]. Whereas weather variations are high frequency, seasonal variations are considered to be low frequency. Seasonal variations are caused by fluctuations due to slow changes within the Earth's rotation about the sun. Furthermore, with respect to GHI, it can be observed that the amount of irradiance to be present within a day is directly related to the time at which the sun rises and sets. Thus Fig. 7 shows that the maximum irradiance value will vary about seasonal variations at a particular location. It should be noted that days with less daylight hours have a lower maximum irradiance value whereas the conjugate could be said about days with more daylight hours.

3 Developing a Prediction Algorithm

The science described in Section 2 suggests that one way to predict the amount of solar energy at a given point is to begin by first predicting the GHI at that point. Since GHI is a measure of the rate at which solar energy is transferred, it is thus an instantaneous quantity and should be viewed as a function of time, i.e. $\text{GHI}(t)$. The total energy per unit area delivered in a given time interval T is thus defined as

$$E = \int_T \text{GHI}(t) dt. \quad (2)$$

To develop an energy prediction algorithm, we thus begin by predicting the time evolution of GHI at a particular point. We then integrate the resulting time series over a predefined time window. In this case, we seek to predict the energy per unit area absorbed over a given day. By definition, we assume our days to begin at midnight local time and we compute the energy density by integrating GHI over the ensuing 24 hours. Each day at midnight, the calculation resets. The result is a discrete time series consisting of one scalar value per day. The process is illustrated graphically in Fig. 8. On any given day n , the predicted energy density is thus

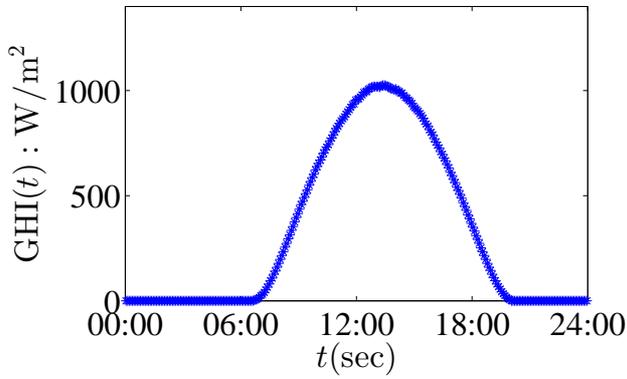
$$E[n] = \int_T \text{GHI}(t) dt, \quad (3)$$



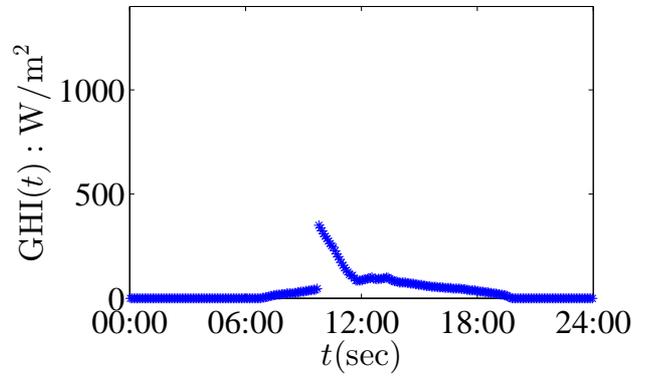
(a) Unobstructed View (Mote 152)



(b) Obstructed View (Mote 157)



(c) Power Profile of Mote 152



(d) Power Profile of Mote 157

Figure 5: (a) and (b): Pictorial representations of where Motes 152 and 157 reside in EPIC-Roofnet respectively. (c) and (d): The typical clear-sky power profile for each mote with respect to their particular locations.

where the interval T is defined to be the 24-hour window beginning at midnight. This definition will later be shown to be convenient in the context of the data available from the National Weather Service (NWS).

Direct calculation of this time series however is challenging because many factors effect it, including

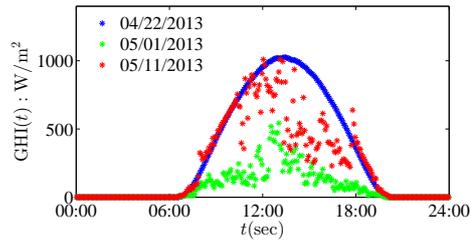


Figure 6: Mote 152's observed harvestable solar energy for three separate days.

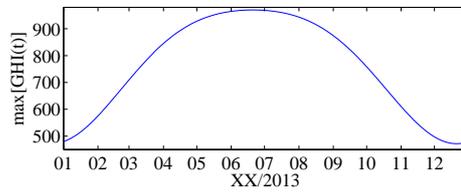


Figure 7: Maximum GHI within a given day over a year.

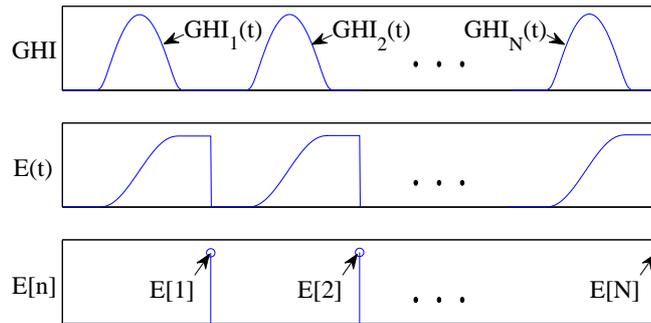


Figure 8: Progression from GHI in the continuous time series (top graph) to the discrete time series energy density (bottom graph) for the n^{th} day.

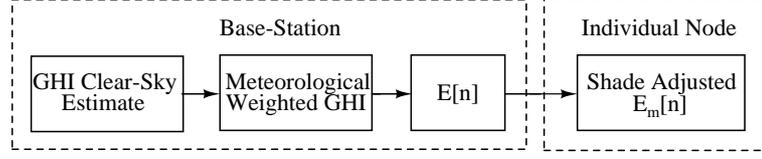


Figure 9: A typical prediction about the n^{th} days expected energy density at an individual node.

absorption by the ozone layer, Rayleigh scattering by airborne molecules, and absorption by water vapor. Since all of these effects are difficult to model, simplifications are needed. One common approach is to develop so-called clear-sky models that predict GHI at a given surface under clear conditions. Even with this simplification, researchers have posited several models with differing levels of complexity and hence accuracy. The interested reader is referred to Refs. [19, 6, 18, 11, 7, 20, 17] for a more through discussion.

For purposes of this research it has been proposed that the prediction about an individual nodes daily energy will follow such a process as described in Fig. 9. Where a Clear-Sky model will be chosen to estimate a theoretical GHI curve for the n^{th} day. At which point a meteorological weighting system will be implemented to down weight the clear-sky model. Where the total estimated energy about the n^{th} day can be calculated utilizing Eq. 3. Up to this point the estimated energy is still defined about a region, so each individual node will derive its own weighting about the estimated energy as to predict its individual energy density.

3.1 GHI Clear-Sky Estimate

Daneshyar-Paltridge-Proctor (DPP) developed a clear-sky model in 1978, that only depends on the solar zenith angle of the sun. Utilizing spherical trigonometry the solar zenith angle for any latitude and longitude is found to be solvable such that

$$\cos(\Theta_Z) = \cos(\phi) \cos(\delta) \cos(\omega) + \sin(\phi) \sin(\delta), \quad (4)$$

where ϕ is the latitude of the site during solar noon on either the spring or fall equinox, δ is the declination angle or the angle between a line pointing from the center of the earth to the sun and the solar equator, and ω is an expression for the local time as an angle. The American Society of Civil Engineers (ASCE) describe the declination angle as [2]

$$\delta = 23.45 \sin(x), \quad (5)$$

where

$$x = \frac{360^\circ}{365} [\text{DOY} + 284]. \quad (6)$$

DOY is a numerical representation for ‘Day of Year’ where DOY = 1 is denoted as January, 1st and December, 31st is either DOY = 365 or 366 for non-leap and leap years respectively. Note that Eq. 5 is a representation for the northern hemisphere only, to represent the declination with respect to the southern hemisphere, Eq. 6 has to be modified such that 284 is subtracted from DOY. Lastly the solar zenith angle calculation takes into effect the daily hour angle which can be represented as

$$\omega(\text{deg}) = (\text{solarTime}[\text{Hour}] - 12) \times 15, \quad (7)$$

where

$$\text{solarTime}[\text{Hour}] = \text{LocalTime} + (\text{MeridianTime} - \text{LocalTime}) \times 4 + \text{EoT}[\text{min}], \quad (8)$$

$$\text{EoT}[\text{min}] = 9.87 \sin(2x) - 7.53 \cos(x) - 1.5 \sin(x), \quad (9)$$

and x here is the same as stated in Eq. 6 [2]. It should be noted that the hour angle is a representation of the local time in which 15 degrees is equivalent to one hour. Utilizing the solar zenith angle, the DPP clear-sky model can be written in terms of the direct irradiance and diffuse irradiance [16, 4], defined respectively as

$$\text{DNI}_{\text{clear}}(t) = 950.2(1 - \exp -0.075(90^\circ - \Theta_Z)) \quad (10)$$

and

$$\text{DHI}_{\text{clear}}(t) = 14.29 + 21.04 \left(\frac{\pi}{2} - \Theta_Z \frac{\pi}{180} \right). \quad (11)$$

The full DPP model is described as the combination of Eqs. 10 and 11, such that it can be represented in a form similar to that of Eq. 1, as

$$\text{GHI}_{\text{clear}}(t) = \text{DHI}_{\text{clear}}(t) + \text{DNI}_{\text{clear}}(t) \times \cos(\Theta_Z). \quad (12)$$

It is suffice to say that even though the DPP model only relies on the solar zenith angle, DPP have made a reasonable assumption by stating that the GHI on a clear-sky basis is highly dependent upon the solar zenith angle. This is considered to be true, due to that the smaller the angle is the less atmosphere the solar radiation has to traverse causing less attenuation in attainable solar energy.

3.2 Meteorological Weighted GHI

Clear-sky models are useful in defining the theoretical obtainable GHI curve for any location at any point in time. It is however, unreasonable to assume that this model will work for all weather conditions, i.e. cloudy days. For this reason researchers have been working on clear-sky to weather-condition solar energy models that take into affect future and current weather conditions. Most of which are developed using various fitting techniques correlating meteorological data and that of the measured GHI. Most notably for this is Richard Perez, where he has defined a correlation between the clear-sky to weather-condition solar energy model utilizing only cloud coverage as exemplified by [18]

$$\text{GHI}(t) = \text{GHI}_{\text{clear}}(t) \left(1 - 0.87 \left(\frac{C_p(t)}{100} \right)^{1.9} \right), \quad (13)$$

where C_p is defined as the predicted cloud coverage in percentage for the desired location. This fit was found to be most suitable for North America as expressed in [18], in which several geographical locations were explored to determine the "best-fit" correlation between $\text{GHI}_{\text{clear}}$ and GHI. To utilize Perez's correlation efficiently, predictions of sky coverage for a particular geographical location are needed, in which the NWS provide through the National Digital Forecast Database (NDFD).

The predictions that the NWS provide is within respect to the percentage of sky coverage by opaque clouds. Where opaque clouds are defined as clouds that are thick enough to block direct irradiance but still translucent enough to allow diffuse irradiance to traverse through. The NWS divides North America into varying regions such as the Mid-Atlantic where the predictions from the NWS are provided in a grid like format against latitude and longitude, seen in Fig. 10. These predictions are provided in three hour interval predictions for the first 72 hours in advance and then six hour intervals for the following 96 hours in advance. Each 'pixel' within the grid-format prediction has a spatial resolution of 5 kilometers. For this reason the 'pixel' of importance for a region is defined by the nearest neighbor pixel for the given locale.

Prior to an individual node, within the network, being able to calculate its particular energy density, the base-station must first carry out an integration about the estimated weather conditioned GHI. For this it is possible to implement Eq. 3 to obtain $E[n]$ on the n^{th} day. It should be noted that the NDFD database gives

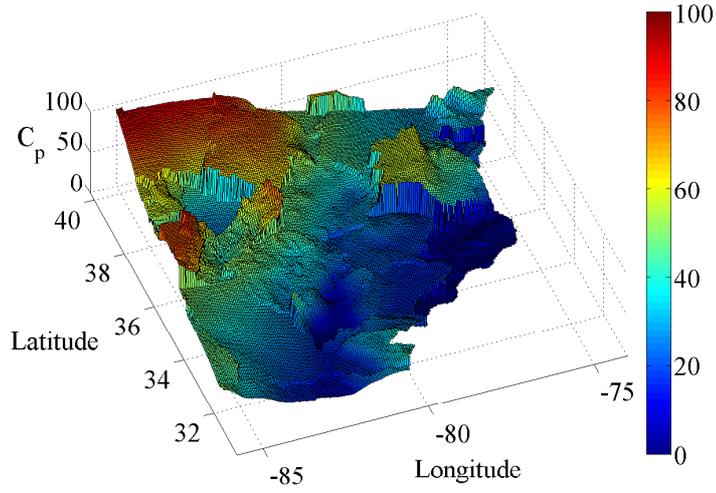


Figure 10: Three dimensional representation of predicted cloud coverage for the Mid-Atlantic.

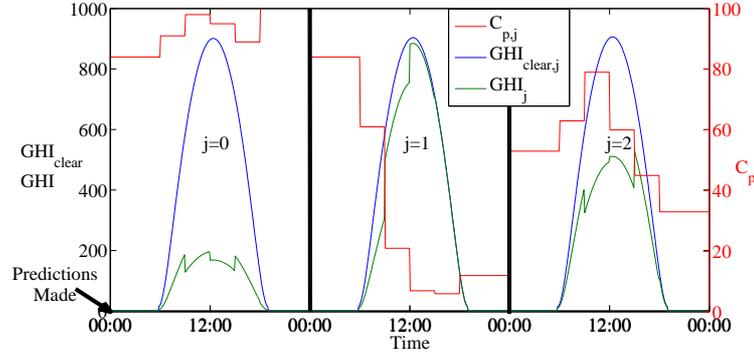


Figure 11: Three day out prediction of $GHI_{clear}(t)$, $C_{p,j}(t)$, and $GHI_j(t)$ for the n^{th} day.

3 hour increments for the future 72 hours worth of time about cloud coverage. From this it is suggested that if all 72 hours were to be utilized it is possible to estimate the regional energy density for up to 72 hours in advance. For purposes of reducing the required amount of computational power at the base-station, the cloud-coverage is considered to be piecewise constant about the prediction interval for the day, which can be seen in Fig. 11. At which point the base-station can now calculate the expected energy density for the region for the j^{th} day ahead of the n^{th} day represented as $E_j[n]$. Where j is defined as the number of days ahead the estimation is made, i.e. day of estimation is denoted as $j = 0$, one day ahead estimation is denoted as $j = 1$, and so forth. Accordingly, the regional energy densities may be represented in vector-matrix form,

$$\underline{E}[n] = [E_0[n], E_1[n], E_3[n]]^T, \quad (14)$$

which can thus be transmitted from the base station to each individual node within the network.

3.3 Adjust for Local Shading

It is postulated that there must exist a system that is capable of transforming the regional energy density into an individual node specific energy density for the same n^{th} day. Such a system is thought to be a summation about a scaled version of the previous ρ regional energy densities taking on the form of

$$\hat{E}_m[n] = \sum_{k=0}^{\rho-1} \omega_k E[n-k]. \quad (15)$$

This form can be expanded to allow for the estimation about the j^{th} day ahead of the n^{th} day. Thus expanding Eq. 15, an estimate can be calculated at each individual node about the available solar energy density for the j^{th} day ahead given by,

$$\hat{E}_{m,j}[n] = \sum_{k=0}^{\rho-1} \omega_{k,j} E_j[n-k]. \quad (16)$$

Utilizing Eq. 16 as the estimation of solar energy for the j^{th} day ahead, it is possible to define the error for the n^{th} day about the j^{th} prediction day ahead with respect to the individual nodes measured energy density, $E_{m,j}[n]$, represented as

$$e_j[n] = E_{m,j}[n] - \hat{E}_{m,j}[n]. \quad (17)$$

In which a cost function represented as $J_j[n]$ can be developed for the system in terms of this error, $e_j[n]$, and a weighting function λ limited to a range of $(0, 1]$. As λ approaches 1, higher weight is given to previous data and inversely can be said as λ approaches 0. Therefore the cost function, $J_j[n]$, can be represented as the sum of the product between the error-squared and the weighting factor λ as such

$$J_j[n] = \sum_{i=0}^n \lambda_j^{n-i} |e_j[i]|^2, \quad (18)$$

leading to an exponentially weighted least squares (EWLS) estimation to be made about the unknown desired coefficients. Substituting Eq. 16 into Eq. 17 it is possible to rewrite the cost-function as

$$J_j[n] = \sum_{i=0}^n \lambda_j^{n-i} \left[E_{m,j}[i] - \sum_{k=0}^{\rho-1} \omega_{k,j} E_j[i-k] \right]^2. \quad (19)$$

At this point it is possible to directly solve for the desired coefficients, knowing that the coefficients must be defined as the value that minimizes the overall cost of the system, represented as

$$\underline{\omega}_j[n] = \underset{\omega}{\operatorname{argmin}}[J_j[n]], \quad (20)$$

where $\underline{\omega}_j[n]$ is $[\omega_{0,j}[n], \omega_{1,j}[n], \dots, \omega_{\rho-1,j}[n]]$. Solving Eq. 20, it can be shown that the optimal coefficients are equivalent to

$$\underline{\omega}_j[n] = \left(\sum_{i=0}^n \lambda^{n-i} \underline{E}_j \underline{E}_j^T \right)^{-1} \left(\sum_{i=0}^n \lambda^{n-i} \underline{E}_j E_{m,j} \right), \quad (21)$$

such as to allow for the system to have the minimalist cost.

From Eq. 21, it is realizable that the solution for the coefficients can be represented as the product of the inverse auto-correlation matrix about the predicted energy densities and the cross-correlation vector

between the predicted energy density and the measured energy density for a particular node, represented respectively as

$$\underline{\underline{R}}_j[n] = \sum_{i=0}^n \lambda^{n-i} \underline{E}_j \underline{E}_j^T \quad (22)$$

and

$$\underline{r}_j[n] = \sum_{i=0}^n \lambda^{n-i} \underline{E}_j E_{m,j}, \quad (23)$$

where \underline{E}_j is equal to $[E_j[n], E_j[n-1], \dots, E_j[n-\rho+1]]^T$. Plugging back into Eq. 21 the solution for the unknown coefficients can be represented as

$$\underline{\omega}_j[n] = \underline{\underline{R}}_j[n]^{-1} \underline{r}_j[n]. \quad (24)$$

Even though Eq. 24 allows for the coefficients to be solved for a particular number ρ coefficients about N data points, it should be noted that it typically requires a massive amount of data points as to obtain accurate estimations on the coefficients leading to a large demand on overhead and memory. To reduce the overhead required it is proposed that the EWLS estimator should be solved for recursively, creating an exponentially weighted recursive least squares (EWRLS) estimation, allowing for lower overhead and memory usage. To take the coefficients solution and make it recursive, Eq. 24 must be represented in terms of the previous auto and cross-correlation matrices.

Applying a few simple manipulations to the auto and cross-correlation matrices, shown in Eqs. 22 and 23 respectively, the correlations can be re-written in terms of their previous values as

$$\begin{aligned} \underline{\underline{R}}_j[n] &= \lambda \sum_{i=0}^{n-1} \lambda^{n-1-i} \underline{E}_j \underline{E}_j^T + \underline{E}_j \underline{E}_j^T \\ &= \lambda \underline{\underline{R}}_j[n-1] + \underline{E}_j \underline{E}_j^T \end{aligned} \quad (25)$$

and

$$\begin{aligned} \underline{r}_j[n] &= \lambda \sum_{i=0}^{n-1} \lambda^{n-1-i} \underline{E}_j E_{m,j} + \underline{E}_j E_{m,j} \\ &= \lambda \underline{r}_j[n-1] + \underline{E}_j E_{m,j} \end{aligned} \quad (26)$$

respectively. However, it is necessary to obtain the inverse of the auto-correlation in terms of this previous data. To obtain this inversion the Woodbury inversion formula stated to be

$$(\underline{\underline{A}} + \underline{u}\underline{v}^H)^{-1} = \underline{\underline{A}}^{-1} - \frac{\underline{\underline{A}}^{-1} \underline{u}\underline{v}^H \underline{\underline{A}}^{-1}}{1 + \underline{v}^H \underline{\underline{A}}^{-1} \underline{u}}, \quad (27)$$

allows for the inversion of the sum of two $m \times m$ matrices [8]. If we allow A to be $\underline{\underline{R}}_j[n-1]$, u as \underline{E}_j , and v^H as \underline{E}_j^T the inverse auto-correlation can be described as

$$\underline{\underline{R}}_j[n]^{-1} = \lambda^{-1} \underline{\underline{R}}_j[n-1]^{-1} - \frac{\lambda^{-2} \underline{\underline{R}}_j[n-1]^{-1} \underline{E}_j \underline{E}_j^T \underline{\underline{R}}_j[n-1]^{-1}}{1 + \lambda^{-1} \underline{E}_j^T \underline{\underline{R}}_j[n-1]^{-1} \underline{E}_j}. \quad (28)$$

For simplifications purposes we will allow $\underline{\underline{P}}_j[n]$ to represent by $\underline{\underline{R}}_j^{-1}[n]$ at time n and what is commonly known as the gain vector to be [9]

$$\underline{g}_j[n] = \frac{\lambda^{-1} \underline{\underline{P}}_j[n-1] \underline{E}_j}{1 + \lambda^{-1} \underline{E}_j^T \underline{\underline{P}}_j[n-1] \underline{E}_j}. \quad (29)$$

Plugging these changes of variables into Eq. 28, the auto-correlation matrix can further be simplified to

$$\underline{P}_j[n] = \lambda^{-1} \left[\underline{P}_j[n-1] - \underline{g}_j[n] \underline{E}_j^T \underline{P}_j[n-1] \right]. \quad (30)$$

Interestingly the gain vector, $\underline{g}_j[n]$, can be inversely solved by simplifying Eq. 29 and substituting Eq. 30 into this simplification, resulting in [9]

$$\underline{g}_j[n] = \underline{P}_j[n] \underline{E}_j. \quad (31)$$

Utilizing Eqs. 26 and 30, it is possible to solve for the current unknown coefficients in terms of its past value, shown to be

$$\underline{\omega}_j[n] = \underline{P}_j[n-1] r_j[n-1] - \underline{g}_j[n] \underline{E}_j^T \underline{P}_j[n-1] r_j[n-1] + E_{m,j} \underline{g}_j[n], \quad (32)$$

where $\underline{R}_j[n-1] r_j[n-1]$ are representations of the previous filter coefficients, $\omega_j[n-1]$. Thus allowing for the coefficients to be represented as a combination of the scaled a-priori error and the previous coefficients described as

$$\underline{\omega}_j[n] = \underline{\omega}_j[n-1] + \underline{g}_j[n] [E_{m,j} - \underline{E}_j^T \underline{\omega}_j[n-1]]. \quad (33)$$

For this predictor to work properly it should be noted that each individual node would need to implement Eqs. 29, 30, and 33 along with a prior estimate of the total energy density for the day by the mote described as

$$\hat{E}_{m,j}^- = \underline{E}_j^T \underline{\omega}_j[n-1]. \quad (34)$$

For purposes of initialization, $\underline{P}_j[0]$ is to be set to an identity matrix with dimensions $\rho \times \rho$ and $\underline{\omega}_j[0]$ should be initialized to all ones in a $\rho \times 1$ vector.

3.4 EWRLS Parameter Optimization

Utilizing the estimator described in Section 3, the optimal values for λ and ρ can be solved for by using a subset of training data about the best and worst case scenarios, as described in Section 2.2. Using one months' worth of estimated energy densities for the desired location along with the measured energy densities per-node, the optimal values are determined by finding the desired λ and ρ values that minimize the combined Root Mean Squared Error (RMSE),

$$\widetilde{RMSE} = \sqrt{RMSE_{152}^2 + RMSE_{157}^2}, \quad (35)$$

about the two nodes. The RMSE for motes 152 and 157, $RMSE_{152}$ and $RMSE_{157}$ respectively, are found by

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n |E_{m,j}[i] - \hat{E}_{m,j}[i]|^2}, \quad (36)$$

where $\hat{E}_{m,j}$ is determined from the predictor as stated in section 3.3 with ρ varying between 1 and 6 in steps of 1, as it must be an integer, and λ varying between 0.01 and 1 in steps of 0.01. This in turn allows for a three-dimensional representation of the combined RMSE, \widetilde{RMSE} , to be created, at which the optimal values for λ and ρ are defined as the point that creates the global minima upon the surface, shown in Fig. 12. Reviewing this Fig. 12, the global minima was found to be where $\lambda = 0.84$ and $\rho = 1$.

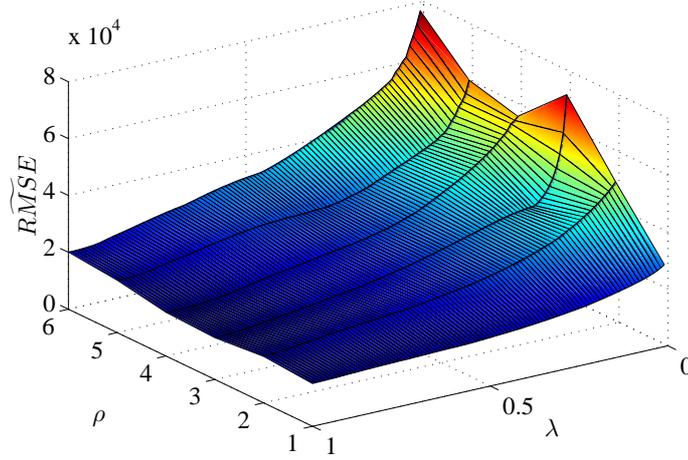


Figure 12: Combined Root Mean Squared Error for Motes 152 and 157 with respect to λ and ρ .

The general form predictor derived in section 3.3, can now be simplified by plugging in the optimal values of λ and ρ , resulting in

$$\hat{E}_{m,j}[n] = E_j(n)\omega_j[n-1], \quad (37)$$

$$g_j[n] = \frac{P_j[n-1]E_j[n]}{0.84 + E_j^2[n]P_j[n-1]}, \quad (38)$$

$$P_j[n] = 1.19 \times P_j[n-1] (1 - g_j[n]E_j[n]), \quad (39)$$

and

$$\omega_j[n] = \omega_j[n-1] + g_j[n] \left(E_{m,j}[n] - \hat{E}_{m,j}[n] \right). \quad (40)$$

At which point, every matrix and vector calculation has been reduced to singular discrete values, in turn reducing the required memory needed. Meeting one of the two stipulations of the desired predictor. To meet the second requirement it is necessary to reduce the computational burden, per-node. This is realizable by reordering the predictor to minimize the number of multiplications and/or additions per instruction cycle. Reducing the required number of manipulations per instruction cycle, will allow the Tiny operating system to become uninterrupted in its tasking as each instruction cycle within the predictor will be able to run between interrupt cycles. Utilizing simple substitution the predictor may be formulated as such that

$$\hat{E}_{m,j}[n] = E_j[n]\omega_j[n-1], \quad (41)$$

$$\alpha_j[n] = P_j[n-1]E_j[n], \quad (42)$$

$$\beta_j[n] = 0.84 + E_j[n]\alpha_j[n], \quad (43)$$

$$g_j[n] = \frac{\alpha_j[n]}{\beta_j[n]}, \quad (44)$$

$$P_j[n] = P_j[n-1] - \alpha_j[n]g_j[n], \quad (45)$$

$$P_j[n] = P_j[n] \times 1.9, \quad (46)$$

$$\zeta_j[n] = g_j[n] \left(E_{m,j}[n] - \hat{E}_{m,j}[n] \right), \quad (47)$$

and

$$\omega_j[n] = \omega_j[n - 1] + \zeta_j[n]. \quad (48)$$

Thus creating a maximum of one multiplication and addition per instruction cycle, allowing each instruction to be completed within 4-8 clock cycles, dependent upon if variables need be stored or loaded from RAM [3].

4 Experimental Results

For assessment of the EWRLS predictor, a total of 64 days worth of energy densities are to be used across the two scenarios presented in Section 2.2 obtained through the EPIC-Roofnet network. Two parameters about performance and accuracy are to be utilized, one of which is the RMSE as defined in Eq. 36, and the other to be of the percent difference between the predicted energy density and that which was measured by the desired node. The percent difference is seen to be a key metric of consideration when viewing the benefit of a predictor. This is said to be true because a predictor that is seen to under predict versus over predict, causes the power scheme for the network to work more conservatively rather than wasting energy that it does not have. For this reason, a percent difference error calculation can be conducted in such a way that lower weight is given to under predictions, represented as

$$\text{err}[n] = \frac{\text{abs}\left(E_m[n] - \hat{E}_m[n]\right)}{E_m[n]} \times 100\%. \quad (49)$$

From this it can be seen that $\text{err}[n]$ will only have a percent greater than 100% if and only if the estimated energy density is greater than twice that of the obtained energy density for the individual node. Utilizing these two metrics for assessment about performance for day of energy predictions, it can be observed through testing that the EWRLS predictor was able to obtain a mean percent difference of roughly 27% between the two particular nodes and a RMSE of 15908 and 2844 for node 152 and 157 respectively. The predictions for the two nodes can be seen in Fig. 13. Note that the predictions in which a significant error can be seen for a particular node, can be seen across both nodes, shown in Figs. 13(c) and 13(d). Thus the initial estimation about cloud coverage provided by the NDFD is said to of been incorrect, leading to a miss calculation of the regional energy density. It should be noted that, the predictions about cloud coverage are said to have an accuracy within 25 to 35 percent of actual, thus compounding the overall error within the EWRLS predictor [10].

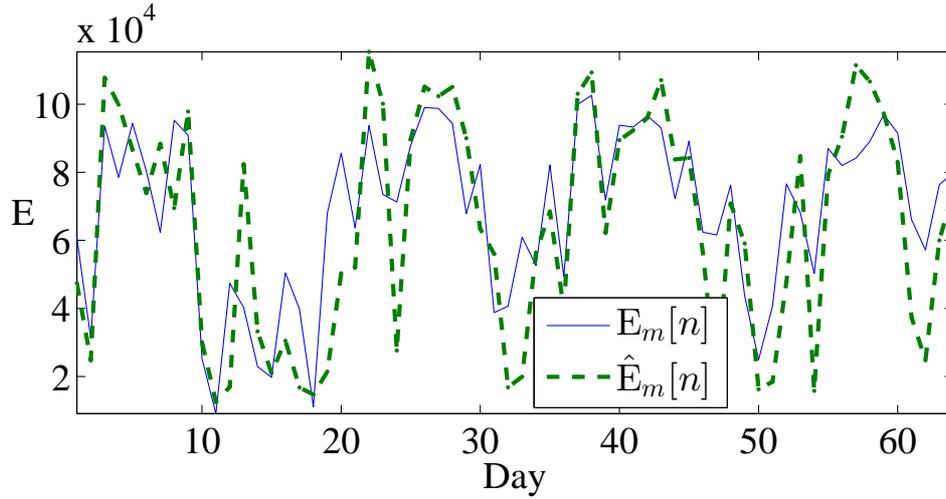
It is possible to compare the EWRLS predictor against previously defined predictors for up to three days in advance, as to really see the benefit of the EWRLS predictor. For a full description of previously defined predictors the interested reader is pointed to App. ???. For purposes of comparison, a persistence predictor along with an exponentially weighted moving average (EWMA) predictor will be utilized. A persistence predictor, is only reliant of the previous days energy density. in which the predicted energy density about the n^{th} day is equivalent to the previous days energy density, i.e.

$$\hat{E}_{m,j}[n + 1] = E_{m,j}[n]. \quad (50)$$

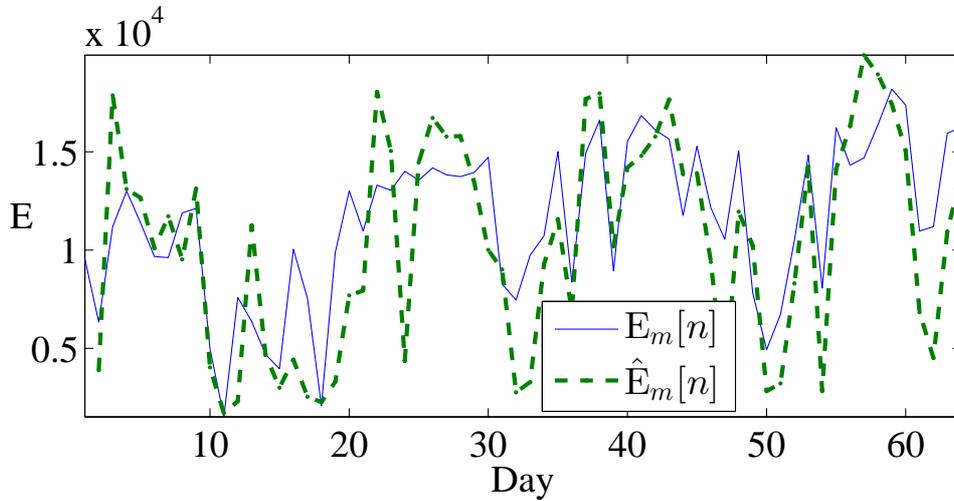
But where as the persistence predictor is only based on the previous days energy density, the EWMA predictor assumes a weighted moving average about the previous days energy densities. Thus eliminating the high frequency component about the fluctuations in energy density day-to-day. Such a predictor can be represented as

$$\hat{E}_{m,j}[n + 1] = \gamma E_{m,j}[n] + (1 - \gamma)\hat{E}_{m,j}[n], \quad (51)$$

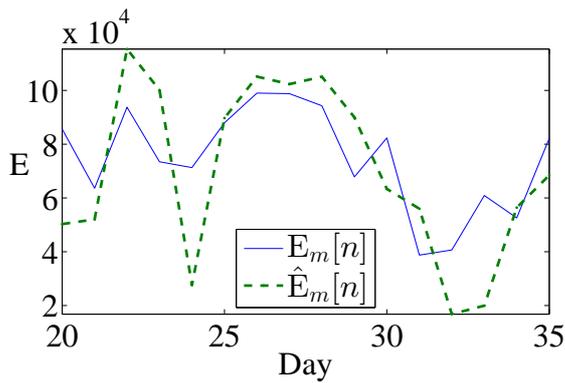
where γ varies between (0, 1], for purposes of comparison γ is set to be 0.25.



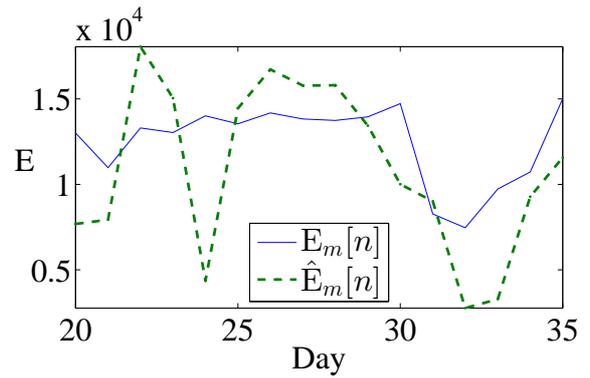
(a) Day Of Prediction Node 152



(b) Day of Prediction Node 157



(c) Close up of Node 152 from Day 20 to 35



(d) Close up of Node 157 from Day 20 to 35

Figure 13: (a) and (b): EWRLS predictions for day of energy for Nodes 152 and 157, respectively. (c) and (d): Close up view of day 20 to 35 for nodes 152 and 157, respectively.

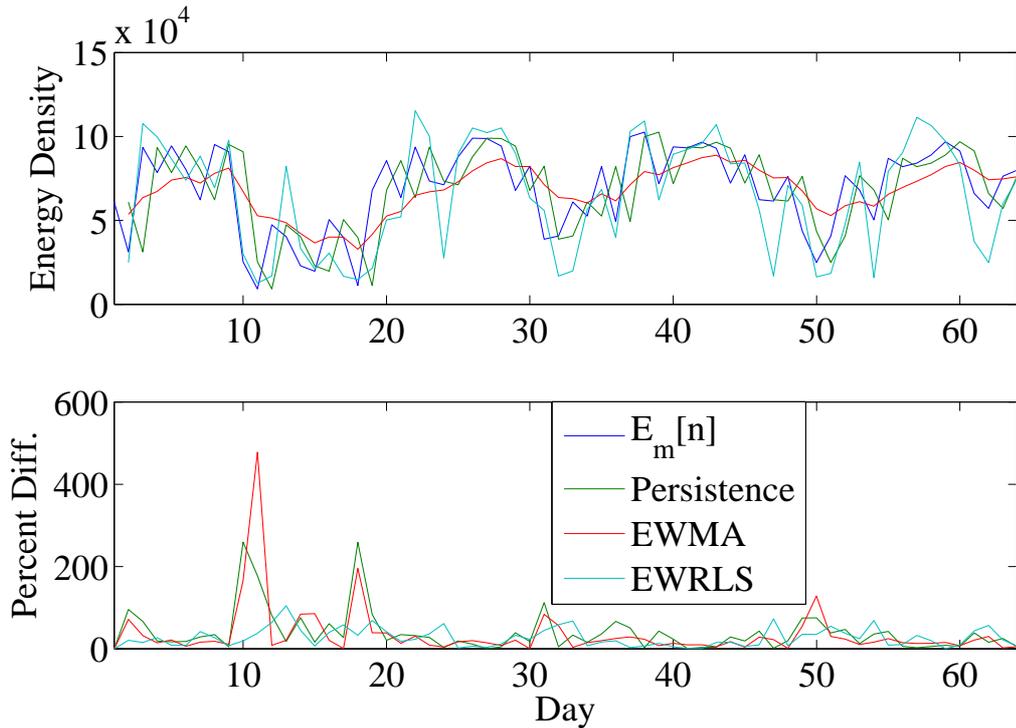


Figure 14: Observed percent difference between Persistence, EWMA, and EWRLS predictors.

Comparing the three predictors against one another in day of predictions, shown in Fig. 14, it can be seen that the EWMA and the persistence predictor have significant error at times. Whereas the EWRLS predictor can be seen to have a minimal amount of error. The persistence and EWMA predictor have a maximum error of 259% and 637% respectively, whereas the EWRLS predictor has a maximum error of 104%. From this it can be concluded that the EWRLS predictor has a significant decrease in the amount of error between the measured and predicted energy densities at a particular node. A full comparison of the three predictors about nodes 152 and 157, for up to 72 hours out, can be explored in Tables 2 and 3, respectively.

References

- [1] Standard tables for reference solar spectral irradiances: Direct normal and hemispherical on 37 tilted surface, 2003.
- [2] Rick G Allen. *The ASCE standardized reference evapotranspiration equation*. American Society of Civil Engineers, 2005.
- [3] Atmel. *ATmega48A/PA/88A/PA/168A/PA/328/P Complete*. Atmel Corporation, g edition, 02 2013.
- [4] M Daneshyar. Solar radiation statistics for iran. *Solar Energy*, 21(4):345–349, 1978.
- [5] Halden Field. Solar cell spectral response measurement errors related to spectral band width and chopped light waveform. In *Photovoltaic Specialists Conference, 1997., Conference Record of the Twenty-Sixth IEEE*, pages 471–474. NREL, IEEE, 1997.

Table 2: Comparison of Persistence, EWMA, and EWRLS predictors for day of to 2 day out predictions, for Mote 152.

Type	Days Out	Memory	RMSE	Mean Diff.	Peak Diff.
Persistence	0	2 Bytes	19055	39.60%	259.11%
	1	2 Bytes	23159	58.93%	896.61%
	2	2 Bytes	22794	54.10%	944.32%
EWMA	0	2 Bytes	19434	47.50%	637.48%
	1	2 Bytes	21794	53.37%	790.80%
	2	2 Bytes	23275	55.42%	755.52%
EWRLS	0	16 Bytes	15908	27.34%	104.52%
	1	16 Bytes	18122	32.63%	125.94%
	2	16 Bytes	20540	39.34%	335.28%

Table 3: Comparison of Persistence, EWMA, and EWRLS predictors for day of to 2 day out predictions, for Mote 157.

Type	Days Out	Memory	RMSE	Mean Diff.	Peak Diff.
Persistence	0	2 Bytes	2945	36.09%	262.29%
	1	2 Bytes	3514	50.28%	701.00%
	2	2 Bytes	3295	43.93%	685.96%
EWMA	0	2 Bytes	3021	40.26%	526.02%
	1	2 Bytes	3281	44.03%	625.02%
	2	2 Bytes	3480	45.08%	599.69%
EWRLS	0	16 Bytes	2844	27.14%	76.82%
	1	16 Bytes	2683	26.15%	99.62%
	2	16 Bytes	3629	37.36%	276.92%

- [6] Pinde Fu and Paul M Rich. Design and implementation of the solar analyst: an arcview extension for modeling solar radiation at landscape scales. In *Proceedings of the 19th annual ESRI user conference, San Diego, USA*, 1999.
- [7] Christian A. Gueymard. Clear-sky irradiance predictions for solar resource mapping and large-scale applications: Improved validation methodology and detailed performance analysis of 18 broadband radiative models. *Solar Energy*, 86(8):2145–2169, 2012.
- [8] William W Hager. Updating the inverse of a matrix. *SIAM review*, 31(2):221–239, 1989.
- [9] Monson H. Hayes. *Statistical Digital Signal Processing and Modeling*. John Wiley and Sons, INC., 2009.
- [10] ISCCP. International satellite cloud climatology project, 03 2013.
- [11] Ricardo Marquez and Carlos F.M. Coimbra. Forecasting of global and direct solar irradiance using stochastic learning methods, ground experiments and the nws database. *Solar Energy*, 85(5):746–756, 2011.
- [12] Andrew Mills, Mark Ahlstrom, Michael Brower, Abraham Ellis, Ray George, Tom Hoff, Benhamin Kroposki, Carl Lenox, Nicholas Miller, Joshua Stein, and Yih huei Wan. Understanding variability and uncertainty of photovoltaics for integration with the electric power system. *The Electricity Journal*, 2010.
- [13] Daryl R Myers. Solar radiation modeling and measurements for renewable energy applications: data and model quality. *Energy*, 30(9):1517–1531, 2005.
- [14] Asis Nasipuri, Robert Cox, James Conrad, Luke Van der Zel, Bienvenido Rodriguez, and Ralph McKosky. Design considerations for a large-scale wireless sensor network for substation monitoring. In *Local Computer Networks (LCN), 2010 IEEE 35th Conference on*, pages 866–873. IEEE, 2010.
- [15] Heinz Neckel et al. The solar radiation between 3300 and 12500 Å. *Solar Physics*, 91(2):205–258, 1984.
- [16] GW Paltridge and D Proctor. Monthly mean solar radiation statistics for australia. *Solar Energy*, 18(3):235–243, 1976.
- [17] Richard Perez, Pierre Ineichen, Kathy Moore, Marek Kmiecik, Cyril Chain, Ray George, and Frank Vignola. A new operational model for satellite-derived irradiances: description and validation. *Solar Energy*, 73(5):307–317, 2002.
- [18] Richard Perez, Kathleen Moore, Steve Wilcox, David Renne, and Antoine Zelenka. Forecasting solar radiation preliminary evaluation of an approach based upon the national forecast database. *Solar Energy*, 81(6):809–812, 2007.
- [19] Matthew J. Reno, Clifford W. Hansen, and Joshua S. Stein. Global horizontal irradiance clear sky models: Implementation and analysis. Technical report, Sandia National Laboratories, 2012.
- [20] Frank Vignola, Peter Harlan, Richard Perez, and Marek Kmiecik. Analysis of satellite derived beam and global solar radiation data. *Solar Energy*, 81(6):768–772, 2007.