

Pro-Energy: a novel energy prediction model for solar and wind energy-harvesting Wireless Sensor Networks

Alessandro Cammarano
Computer Science Department
Sapienza University of Rome
Rome, Italy
Email: cammarano@di.uniroma1.it

Chiara Petrioli
Computer Science Department
Sapienza University of Rome
Rome, Italy
Email: petrioli@di.uniroma1.it

Dora Spenza
Computer Science Department
Sapienza University of Rome
Rome, Italy
Email: spenza@di.uniroma1.it

Abstract—Energy harvesting is one of the most promising technologies towards the goal of perpetual operation of wireless sensor networks (WSNs). Environmentally-powered systems, however, have to deal with the variable behavior of ambient energy sources, which results in different amounts and rates of energy available over time. To alleviate the problem of the harvested power being neither constant nor continuous, energy prediction methods can be employed. Such models forecast the source availability and estimate the expected energy intake, allowing the system to take critical decisions about the utilization of the available energy. In this work, we present a novel energy prediction model, named Pro-Energy (PROfile energy prediction model), for multi-source energy harvesting WSNs, which is able to leverage past energy observations to provide accurate estimations of future energy availability. To assess the performance of our proposed solution, we use real-life solar and wind traces that we collected by interfacing TelosB nodes with solar cells and wind micro-turbines, as well as public available traces of solar and wind obtained from weather monitoring stations in the US. A comparative performance evaluation between Pro-Energy and energy predictors previously proposed in the literature, such as EWMA and WCMA, has shown that our solution significantly outperforms existing algorithms for both short and medium term prediction horizons, improving the prediction accuracy up to 60%.

Keywords—Energy predictions; Energy harvesting; Environmentally-powered WSNs; Solar-powered WSNs; Wind-powered WSNs; Green WSNs; Multi-source energy harvesting; Rechargeable sensors.

I. INTRODUCTION

Many works have demonstrated both the potential and the feasibility of applying energy-harvesting techniques to wireless sensor networks [1], [2], [3]. However, networks powered by ambient energy sources, such as solar light and wind, have to face the problem of dealing with a variable energy intake. For instance, solar-powered nodes experience significant changes in the power harvested over time, due to the diurnal cycle in solar energy, varying weather conditions, monthly trends and seasonal patterns. Moreover, the position of the nodes and the orientation of their solar cell strongly impact the energy intake: even if two nodes are physically co-located, often their harvesting rates significantly differ [4], [5].

The uncertainty in the energy availability provided by ambient sources raises new challenges in developing reliable and energy-efficient power-management solutions. Many works assume that accurate predictions about the future energy intake are available to the system, either by simply looking at the past history [6] or by employing more sophisticated energy predictors [7]. Knowledge about the behavior of energy sources over short and medium time frames is often needed to optimize the system and some solutions even rely on it to work well [8]. In general, having no such knowledge available will result in an under-performing system, in which nodes have no possibility to plan in advance how to spend the energy they will harvest in the near future. Predictions about the future energy availability also helps minimizing waste of energy. Nodes store the harvested energy in devices such as supercapacitors or rechargeable batteries, which are limited in both size and time, due to their limited capacity and to leakage and self-discharge, respectively. Developing energy predictors for WSNs that provide accurate future predictions over short (few minutes to half an hour) and medium (a few hours) time frames is very important, as they allow to exploit at best the available energy, minimizing the likelihood that important tasks are not executed due to lack of energy, and minimizing energy waste, i.e., cases in which the generated energy is neither used nor stored as the buffer is full.

So far, existing works, such as WCMA [9], have mainly focused on short-term predictions for solar energy harvesting. In WCMA, each day is discretized into N time slots and, at the end of each timeslot t , the energy prediction for timeslot $t + 1$ is derived.

Our goal has instead been to develop a general framework for multi-source (i.e., solar and wind) energy-harvesting systems such as [1], [10], which is able to accurately predict energy for forecasting horizons that are dynamically chosen based on the application needs.

In this paper, we make the following contributions:

- We present a novel energy prediction model, named Pro-Energy (PROfile Energy prediction model), which is able to leverage past energy observations to provide

accurate estimations of future energy availability, for both solar and wind harvesting WSNs.

- We discuss a variation of our solution which combines multiple energy profiles in order to improve predictions accuracy for short and medium-term forecasting horizons.
- We provide a performance evaluation framework for energy prediction models. To validate our solution, we use real-life traces of energy availability obtained by interfacing Telos B nodes with photovoltaic panels and wind micro-turbines. We also exploit two additional solar and wind datasets from the US National Renewable Energy Laboratory [11]. Our results show that Pro-Energy performs up to 60% better than other prediction models, being able to provide higher accuracy than EWMA and WCMA.
- We perform a thorough analysis of how varying key parameters of our scheme impacts on the predictions accuracy achieved for short and medium term prediction horizons.

The remainder of this paper is organized as follows. Related work is presented in Section II. In Section III we describe our prediction model, Pro-Energy. In Section IV we discuss a variation of our scheme that may be employed to achieve better predictions accuracy for short and medium term predictions. We perform a comparative performance evaluation of Pro-Energy and other energy prediction models proposed in the literature in Section V. In the same section, we also evaluate the impact of critical parameters on the performance of our prediction scheme. Finally, we present our conclusions in Section VI.

II. RELATED WORK

The Exponentially Weighted Moving-Average (EWMA) algorithm is a widely used solar energy prediction scheme proposed by Kansal et al. in [12], which is based on an exponentially weighted moving-average filter [13]. EWMA relies on the assumption that the energy available at a given time of the day is similar to the energy generation observed at the same time on the previous days. The amount of energy available during the past days is maintained as a weighted average, in which the contribution of older data is exponentially decaying. This approach is able to both exploit the diurnal cycle in solar energy and to adapt to seasonal variations, but leads to significant prediction errors in case of frequently changing weather conditions, i.e, when sunny and cloudy days are mixed.

In order to address this problem, a new estimation method, the Weather-Conditioned Moving Average (WCMA), has been proposed by Piorno et al. in [9]. The high prediction errors shown by EWMA when sunny and cloudy days alternate is due to the high impact that the weather conditions of the previous day have when estimating the energy generation for the current day. The WCMA prediction algorithm avoids this

effect by taking into account, when computing the prediction for a given timeslot, the average energy availability experienced in that slot in the previous days. Such average value is then scaled according to a weighting factor indicating how the weather conditions of the current day changed with respect to the previous days. In case of frequently changing weather conditions, WCMA is shown to obtain average prediction errors almost 20% smaller than EWMA.

Another estimation method based on a weighted sum of historical data is presented by Moser et al. in [14], to provide information to a predictive controller able to adapt parameters of an application. Their prediction algorithm assumes the solar power to be periodic on a daily basis. To estimate the energy which will be harvested in the next time interval it combines the value of the energy harvested during the current time interval with the energy harvested in the past (whose age is a multiple of days). As for EWMA, the contribution of older data is exponentially decaying. In [15], this predictor is shown to often result in underestimation of the forecast values, thus in high prediction errors.

The solution proposed by Noh et al. in [16] is similar to previous approaches. They use the EWMA model to keep track of the solar energy profile observed in the past. In order to account for short-term varying weather conditions, they also introduce a scaling factor to adjust future energy expectations. At the end of each slot, scaling is performed by computing the ratio between the actual energy harvested during the current timeslot and the energy predicted for the same timeslot, appropriately scaled for future timeslots that are far away in time.

Lu et al. addressed the problem of energy-harvesting prediction for real-time embedded systems (RTES) in [17]. They argue that accurate prediction of the energy intake in the near future is crucial for RTES, as the performance of optimization techniques depends on harvesting predictions. Thus, they investigate three common techniques in real time series prediction (regression analysis, moving average and exponential smoothing), showing that regression analysis has the best accuracy for energy predictions within a time horizon of 1 second. Their approach, however, works well for real-time energy predictions, but it is not designed for medium-term prediction horizons.

A completely different approach is proposed by Sharma et al. in [18]. The authors explore a system for solar and wind powered sensor node that is able to derive energy harvesting predictions based on weather forecast. More in detail, they observe that, when predicting energy availability at timescales between 3 hour to 3 days, using forecasting data provides higher accuracy than calculating energy predictions based on past observation. The reason they give for the scarce performance of traditional predictors is the fact that weather patterns are not consistent in many regions of the United States. They thus formulate a model for solar panel and wind turbine that is able to convert

Table I
TABLE OF NOTATIONS

Symbol	Explanation
N	Number of timeslots in a day
D	Number of energy profiles stored in the memory
E	Matrix of energy profiles ($D \times N$)
E_t^d	Harvested energy observed during timeslot t of day d
C	Power harvested during the current day
C_t	Power harvested during timeslot t of the current day
K	Number of past observations used to compute profiles similarity
\hat{E}_{t+1}	Predicted energy at timeslot $t + 1$ on the current day
α	Weighting factor for short-term prediction
F	Prediction horizon (number of time slots)
G	Correlation parameter
P	Number of energy profiles combined for energy predictions
WP	Weighted profile (combination of P profiles)
γ	Weighting factor for short and medium-term predictions
MAE	Mean Absolute Error
$MAPE$	Mean Absolute Percentage Error

weather forecast data into energy harvesting predictions. However, they compare the performance of their solution with simple energy predictors based on past observations; a comparison with state-of-the-art solutions is not presented. Beyond being applicable only in some application scenarios, periodic forecast transmissions introduce an overhead that impacts on network lifetime.

To the best of our knowledge, no solution has been provided so far which allows to dynamically choose the time horizon of forecasting based on the application needs. Pro-Energy is the only prediction algorithm that forecasts the future energy intake for both short and medium term prediction horizons using only the information collected by the nodes themselves.

III. PRO-ENERGY

In this section, we discuss a new energy prediction algorithm for wireless sensor networks, which uses past-days observations to derive predictions on the future energy intake for both short (few minutes to half an hour) and medium (a few hours) predictions horizons.

To reduce the time and memory overhead of the prediction model, each day is discretized into a given number, N , of equal-length timeslots and predictions are performed once per slot. The main idea of our energy prediction algorithm is to make use of harvested profiles representing the energy intake available during different types of “typical” days. For instance, days may be classified into sunny, cloudy or rainy and a characteristic profile may be associated to each of these categories.

The energy received during the current day is stored in a vector, C , of length N , containing the energy obtained during each of the past timeslots. A pool of energy profiles observed in the past is also maintained in a matrix E , of size $D \times N$. These profiles represent the energy obtained

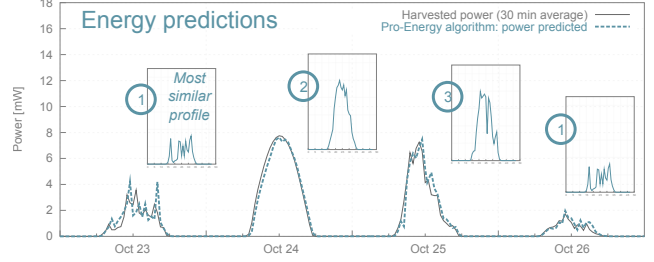


Figure 1. Pro-Energy: harvesting profiles representing energy observations from past days are used to predict the future energy availability.

during a given number, D , of typical days. Such energy profiles are used by Pro-Energy to forecast future energy intake over short and medium term time frames: Once per timeslot, Pro-Energy delivers energy predictions by looking at the stored profile that is the most similar to the current day. For each of these profiles, the similarity with the energy profile observed during the current day is computed as their mean absolute error (MAE), taking into account the last K energy observations. More formally, the stored profile, E^d , that, among the D typical profiles stored, is the most similar at timeslot t to the current day, C , is defined as follows:

$$E^d = \min_{E^d \in E} \sum_{i=t-K}^t \frac{1}{K} |C_i - E_i^d| \quad (1)$$

Figure 1 shows an example of application of the Pro-Energy algorithm over 4 days of solar predictions. During the initial timeslots of Oct 23, the first stored profile is selected among the typical ones, as it is the most similar to the portion of the current day observed so far. As the day goes on, the shape of the profile is scaled and adapted according to the new observations. Two other different profiles are used for predictions during days 2 and 3. Then, on the fourth day, the first profile is selected again as the most similar to the current observations.

If the weather conditions change during the current day, e.g., a sunny morning followed by a cloudy afternoon, we expect the most similar profile to be one of those stored which corresponds to a sunny day in the morning, and a cloudy one in the afternoon. In such cases, considering only the last K observations in Equation (1) has the effect of lowering the probability of choosing the wrong profile, while reducing at the same time the computational overhead of Pro-Energy.

Pro-Energy is made up of three components:

- 1) The **prediction module** delivers future predictions for short and medium term prediction horizons;
- 2) The **profile analyzer** selects, among the stored profiles, the one that is the most similar to the current day;
- 3) The **profile pool refresh** updates the pool of energy profiles, taking into account the age of profiles and their similarity.

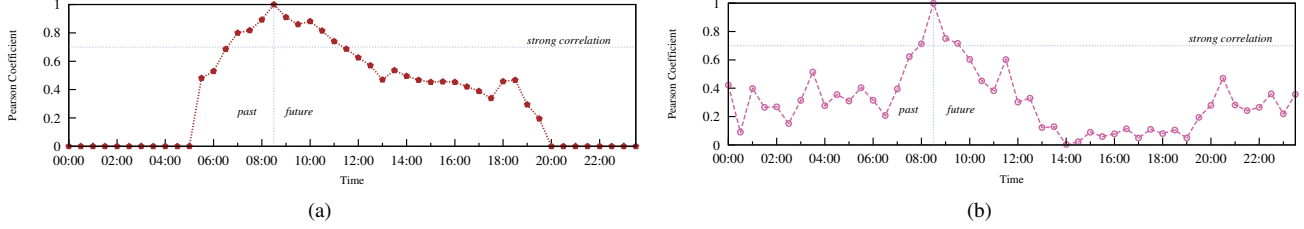


Figure 2. Pearson autocorrelation coefficient for (a) solar ORNL Dataset and (b) wind Bologna Dataset.

A. Short-term energy predictions

When delivering energy predictions, Pro-Energy tries to match the observations of the current day with one of the typical profiles stored in its pool.

Specifically, Pro-Energy computes the predicted value for the next timeslot based on a combination of the value for the next timeslot reported in the stored profile and the energy observed in the last timeslot, C_t . More formally, being E^d the stored profile that is the most similar (up to timeslot t) to the current day, the predicted energy intake for the next slot, $t + 1$, of the current day is computed as:

$$\hat{E}_{t+1} = \alpha \cdot C_t + (1 - \alpha) \cdot E_{t+1}^d \quad (2)$$

where:

- \hat{E}_{t+1} is the predicted energy intake in timeslot $t + 1$ of the current day;
- E_{t+1}^d is the energy harvested during timeslot $t + 1$ on the stored day d ;
- C_t is the energy harvested during timeslot t on the current day C ;
- α is a weighting factor, $0 \leq \alpha \leq 1$.

The weighting parameter, α , allows to combine the value reported in the stored profile with the current energy observation, i.e., the energy observed in the last slot, C_t .

B. Medium-term energy predictions

When computing short-term predictions, considering the correlation between two consecutive timeslots usually helps increasing the prediction accuracy. This approach, however, is not as effective when delivering medium term predictions. In fact, the correlation between the energy observed at time t and the one observed at time $t + \delta t$ generally decreases for increasing δt . To analyze such correlation, we discretized each day in 48 slots of 30 minutes duration and considered a given timeslot, t , which corresponds, in our example, to the 8:30 AM (results are similar for different timeslots). Figure 2 shows the Pearson correlation of the power observed during timeslot t and timeslot $t + \delta t$, $\forall \delta = \{0, \dots, 48\}$, for both solar and wind harvested data. The Pearson correlation coefficient ranges from -1 to 1. An absolute value of 1 implies a linear relationship between the energy observed at two different timeslots, while a value of 0 means that there is no linear correlation between them. It is evident from the

figure that, in the solar case, there is a *strong* correlation (i.e., correlation coefficient > 0.7) between the harvested power observed at 8:30 AM and the energy intake during the successive 4 – 6 slots (2 – 3 hours) (Fig. 2(a)). On the contrary, wind energy observed at 8:30 in the morning shows *strong* correlations only with the successive 1 – 2 slots (30 minutes to one hour) (Fig. 2(b)).

Figure 3 shows an example of the power harvested by a solar cell and a wind micro turbine over 8 days in August, which highlights that the wind energy intake is generally much more variable than the solar one.

Knowledge of correlation can be exploited for better prediction accuracy. To this purpose, we introduce a new parameter, γ , which determines the influence of the last energy observation while deriving predictions for the next F future slots.

Assuming that two slots at a distance equal to or greater than G show only a weak correlation, the γ parameter to be used when predicting the energy intake for the future slot i is defined as:

$$\gamma_i = \begin{cases} \alpha \cdot \left(1 - \frac{i-1}{G}\right), & \text{if } i \leq G \\ 0 & \text{if } i > G \end{cases} \quad \forall i, 1 \leq i \leq F$$

where:

- α is the weighting factor defined in Equation (2);
- i is the i^{th} timeslot in the future, with respect to the current slot, t ;
- G is the number of timeslots in the future which show a correlation above a given threshold with timeslot t ;
- F is the number of future timeslots for which Pro-Energy is delivering energy predictions.

The γ_i parameter plays a similar role in Equation (3) (defined in the following) as the weighting parameter α in Equation (2): It allows to combine the energy value of the stored profile with the current energy observation. However, the weight associated to the value observed during the current slot progressively decreases when computing predictions for timeslots that are further away in time. For timeslots that are more than G slots in the future from timeslot t , such weight, γ_{t+G} , is set to zero, as there is little or no correlation between the energy value observed at timeslot t and that observed at timeslot $t + G$.

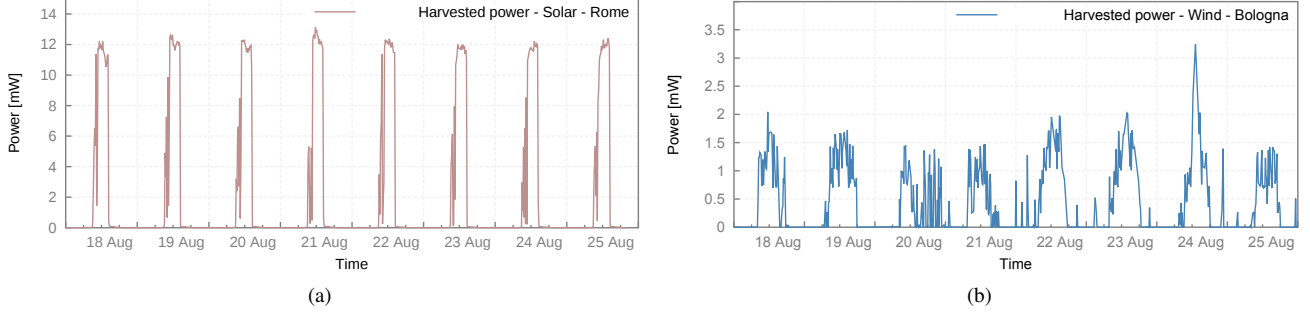


Figure 3. Power generated by our (a) solar cell and (b) wind micro-turbine over 8 days in August 2011.

Having defined such γ_i parameter, the medium term predictions are then computed by using a generalization of the short-term version (Equation 2):

$$\hat{E}_{t+i} = \gamma_i \cdot C_t + (1 - \gamma_i) \cdot E_{t+i}^d \quad (3)$$

Indeed, short-term predictions are computed according to Equation 3, where the prediction horizon F is equal to 1, $i = 1$ and $\gamma_i = \alpha$.

C. Stored profiles updates

Pro-Energy maintains a pool of D typical profiles, each ideally representative of a different weather condition. In order to adapt predictions to changing seasonal patterns, this pool has to be periodically updated. At the end of each day, Pro-Energy decides whether to update the pool of stored profiles with the profile observed during the current day, C , or not. Specifically, it jointly implements two update strategies:

- 1) If there is a profile in the pool that was stored longer than A days ago, substitute it with the profile observed during the current day, C .
- 2) If there are two profiles E^{d_1} and E^{d_2} in the pool that are very similar, i.e., their MAE is below a given threshold, substitute with C the one among the two that is the most similar to the current day. In case of multiple pairs of similar profiles we select among these pairs the most similar to C .

The first strategy allows to discard profiles that have become obsolete, while the second strategy allows to maintain a pool of profiles that are ideally representative of different weather conditions, by discarding profiles that are very similar.

IV. IMPROVING PREDICTION ACCURACY BY COMBINING MULTIPLE PROFILES

Pro-Energy predictions can be further improved through a technique which allows to combine multiple profiles together. Such method selects a set of P profiles, instead of a single one, among the D profiles stored in the E matrix and combines them to form a "weighted" profile WP .

The main idea behind the use of multiple profiles is to consider different possible evolutions of the current day. For instance, a sunny morning may be followed by a cloudy or rainy afternoon. While delivering medium and long term predictions, considering a single profile may lead to poor accuracy if significant variations occur in the weather conditions. On the contrary, considering multiple profiles allows to account for these potential changes, reducing the prediction error at the price of a small additional overhead.

Let $E^{d_1}, E^{d_2}, \dots, E^{d_P}$ be the ordered list of profiles that are most similar to the current day C , i.e., profiles with the smaller Mean Absolute Errors. The weighted profile WP , for the future slot $t + i$, $i \in \{1, 2, \dots, F\}$, is computed as:

$$WP_{t+i} = \frac{1}{P-1} \sum_{j=0}^P w_j \cdot E_{t+i}^{d_j} \quad (4)$$

where

$$w_j = 1 - \frac{MAE_k(E^{d_j}, C)}{\sum_{j=1}^P MAE_k(E^{d_j}, C)} \quad (5)$$

As for the previous case (Section III-B), energy predictions for the future slot $t+i$ are computed based on the value for such slot stored in the WP profile and on the energy harvested during the last time slot.

The predicted energy intake for the future slot, $t + i$, is then computed as:

$$\hat{E}_{t+i} = \gamma_i \cdot C_t + (1 - \gamma_i) \cdot WP_{t+i} \quad (6)$$

where:

- \hat{E}_{t+i} is the predicted energy in timeslot $t + i$ for the current day;
- C_t is the harvested energy during last timeslot;
- WP is the vector of the combination of timeslot $t + i$ of the P profiles;
- γ_i is a correlation factor for prediction of slot $t + i$.

Equation (6) is a generalization of Equation (3). Specifically Equation (3) is obtained by Equation (6) by using only one profile for future energy predictions.

V. PERFORMANCE EVALUATION

We evaluated the accuracy of the Pro-Energy predictor in many different settings, by using four different types of harvested energy datasets: 1) real-life solar data obtained from a testbed in Rome, Italy; 2) real-life wind data obtained from a testbed in Bologna, Italy; 3) traces of solar availability obtained from the National Renewable Energy Laboratory at Oak Ridge, Tennessee [11]; and 4) traces of wind availability obtained from the same source.

We obtained real-life solar data by interfacing Telos B motes [19] with photovoltaic cells (Figure 5(a)). A dedicated TinyOS application was developed to track the amount of energy generated by the harvesters every 30 seconds. The monitoring motes were deployed close to the window of our university building in Rome for 46 days reporting data under variable weather conditions and in different locations. Ten nodes were deployed in different locations, inside and outside the windows, with windows selected in offices with different orientations (west/east/south). In general, due to varying weather conditions, seasonal patterns and different node positions and solar cell orientations, the amount of energy harvested varied significantly over time (3–220 J per day). Real-life wind harvesting data were obtained from an outdoor testbed, located in Bologna, Italy, of Telos B motes equipped with micro wind turbines (Figure 5(b)), collecting data for 75 days. The other two datasets were obtained from the National Renewable Energy Laboratory at Oak Ridge, Tennessee, and consist of 90 days of solar and wind data collected with a granularity of one per minute.

A. Prediction algorithms evaluation

We evaluated the performance of Pro-Energy by comparing, for each timeslot, the amount of energy predicted with the energy actually harvested. The prediction error was calculated by using the Mean Absolute Percentage Error (MAPE) function [20]:

$$MAPE = \frac{1}{T} \sum \left| \frac{\bar{e}_t - \hat{e}_t}{\bar{e}_t} \right|, \quad (7)$$

where:

- \hat{e}_t is the energy predicted for timeslot t ;
- \bar{e}_t is the actual energy harvested during timeslot t ;
- T is the total number of timeslots over which the MAPE error is computed.

In general, T is lower than the total number of timeslots in a given dataset. When computing the MAPE error, we only consider the timeslots in which the energy intake is meaningful to evaluate the prediction accuracy [9]. For this reason, we discard from the MAPE calculation the timeslots in which the harvested power is low, i.e., it is less than the 10% of the maximum peak power of the day.

We compare the performance of Pro-Energy with that of two energy predictors previously proposed in the literature, EWMA and WCMA. In our experiments, we set

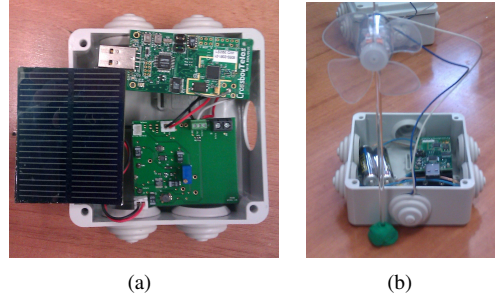


Figure 5. Telos B motes interfaced with (a) photovoltaic cell (b) and wind micro turbines.

$N = 48$. Thus, a whole day is represented by a vector of 48 timeslots, each corresponding to a 30 minutes interval. In order to perform a fair comparison, we set the coefficient of each prediction model to their optimal value, i.e., the ones minimizing the overall MAPE error, using the same energy traces. Since WCMA is designed to only deliver predictions for the next timeslot, we needed to extend it to make it able to perform predictions for different time horizons. Specifically, we modified WCMA so that, when it is asked to predict the energy intake at timeslot $t + \delta$, $\delta > 1$, it returns the average energy observed at timeslot $t + \delta$, computed over the last D_{wcma} days.

B. Accuracy of short and medium term energy predictions

Figure 4 shows the prediction error of Pro-Energy, WCMA and EWMA for four different solar and wind datasets and for different prediction horizons. Specifically, we report here results about the accuracy of both short and medium term energy predictions. The prediction error shown in the figure is the MAPE (Equation (7)) between the amount of energy predicted and the one actually observed, computed over the whole dataset.

The results highlight that the characteristics of the energy source impact the accuracy of prediction. As can be seen, wind energy (Fig. 4(c)- 4(d)) is more difficult to predict than solar energy (Fig. 4(a)- 4(b)), since it is less stable over time.

As can be expected, the prediction error of Pro-Energy increases for longer prediction horizons. In fact, correctly estimating the energy trend of the current day becomes much more difficult as the forecasting horizon increases. Pro-Energy however consistently outperforms EWMA and WCMA for both short and medium term solar energy predictions. In case of solar energy predictions with a prediction horizon of 1 hour, Pro-Energy performs $\approx 25\%$ better than EWMA and WCMA (Fig. 4(a)). The parameter setting in such scenario is summarized in Table III. More generally, for both short and medium-term energy predictions, Pro-Energy achieves a 5% – 26% reduction of the overall MAPE with respect to WCMA. The reduction of the MAPE error with respect to EWMA is between 9% – 43%.

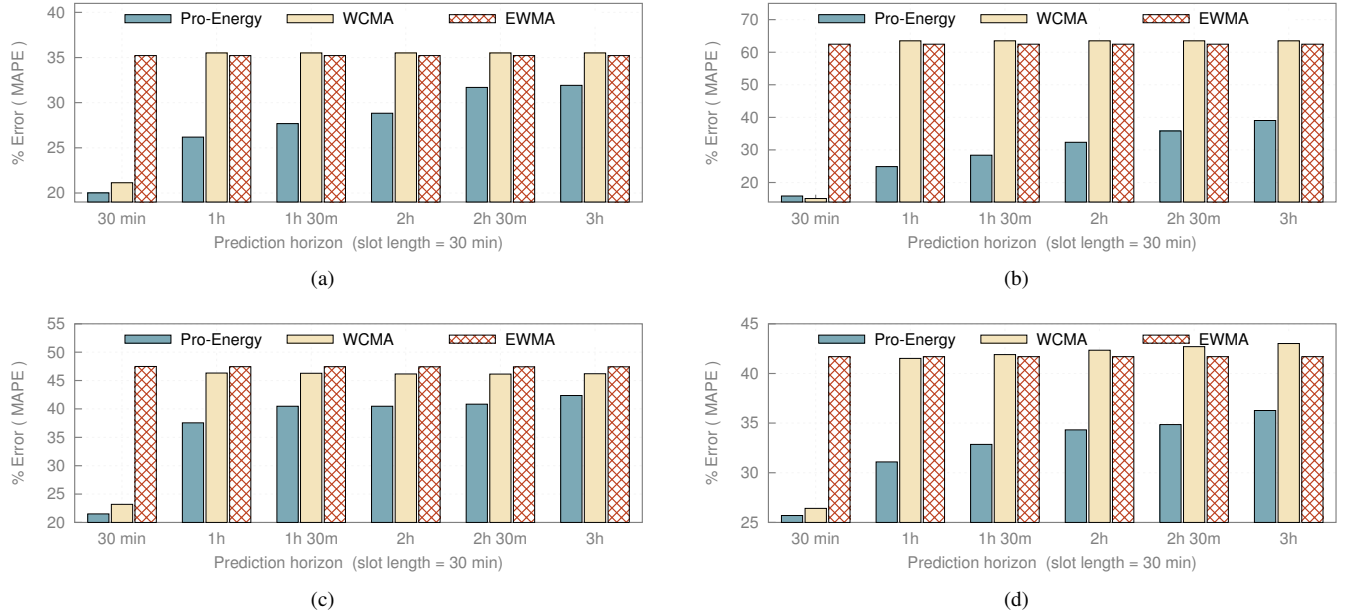


Figure 4. Short and medium-term energy predictions accuracy: comparison between Pro-Energy, WCMA and EWMA for (a) solar ROME dataset; (b) solar ORNL dataset; (c) wind BOLOGNA dataset and (d) wind ORNL dataset.

Such an improvement is even more significant in the solar-ORNL dataset (Fig. 4(b)). For energy predictions within a prediction horizon of 1 hour, Pro-Energy performs 60% better, by exhibiting a MAPE error that is almost one third of that of EWMA and WCMA. For medium-term energy predictions, Pro-Energy achieves an average reduction of the overall MAPE error of almost 50% with respect to WCMA and EWMA. In case of short-term predictions, the performance of Pro-Energy is comparable with that of WCMA while it is 75% better than EWMA's.

Differently from the other predictors, Pro-Energy achieves a good accuracy also when performing wind harvesting forecasting. For the Wind-Bologna dataset (Fig. 4(c)), Pro-Energy leads to a 7%–19% lower MAPE error than WCMA. The improvement is between 9% and 54% in case of EWMA. Figure 4(b) shows the improvement of Pro-Energy over WCMA and EWMA in the wind-ORNL dataset. The overall MAPE error achieved by Pro-Energy is 3% – 25% lower than WCMA and 13% – 38% lower than EWMA.

Figure 4 shows that EWMA is the worst performing predictor even if its performance are constant and do not depend on the prediction horizon. The reason is that EWMA does not use information about the current energy intake in slot t to adjust prediction in future slots $t + \delta$. At the end of each slot t , the only prediction that is updated is the one relative to the same slot t , which will be used for energy predictions on the next day. WCMA performs very well in delivering accurate short-term term energy predictions. Its prediction accuracy, however, degrades when handling medium term predictions. This is due to the fact that WCMA, being designed to only deliver predictions for the

next timeslot, does not exploit the current energy observation and the correlation between current observation and future ones to adjust future energy predictions over medium-term forecasting horizons. The remarkable improvement achieved by Pro-Energy over WCMA highlights the importance of considering the correlation between the current conditions and the future time slots for accurate medium term prediction.

C. Accuracy of energy predictions for varying parameter settings

The last part of our analysis is focused on the impact of varying parameter settings on the accuracy of the energy predictions. We focus our evaluation on the solar-ROME dataset, which well represents a practical application scenario. Due to space restrictions, we restrict our analysis considering three different predictions horizons: 30 minutes, 1 hour and 2 hour. We performed three sets of experiments, varying the value of the parameters α , K and P . While fixing a parameter, the value of the other parameters have been set so as to minimize the overall MAPE error (Table III).

Figure 6(a) displays the impact of varying the α parameter on the prediction accuracy of Pro-Energy, for 30 minutes, 1 hour and 2 hours energy prediction horizons. Figure 6(a) shows that the α parameter has a high impact on the total MAPE error in case of short-term predictions. In such case, the total error, computed over the whole solar dataset, is minimized for $\alpha = 0.5$. This corresponds to a balanced contribution between the energy value reported in the stored profile E_{t+1}^d and the last energy observation made during the current day, C_t . Higher values of the parameter, i.e.,

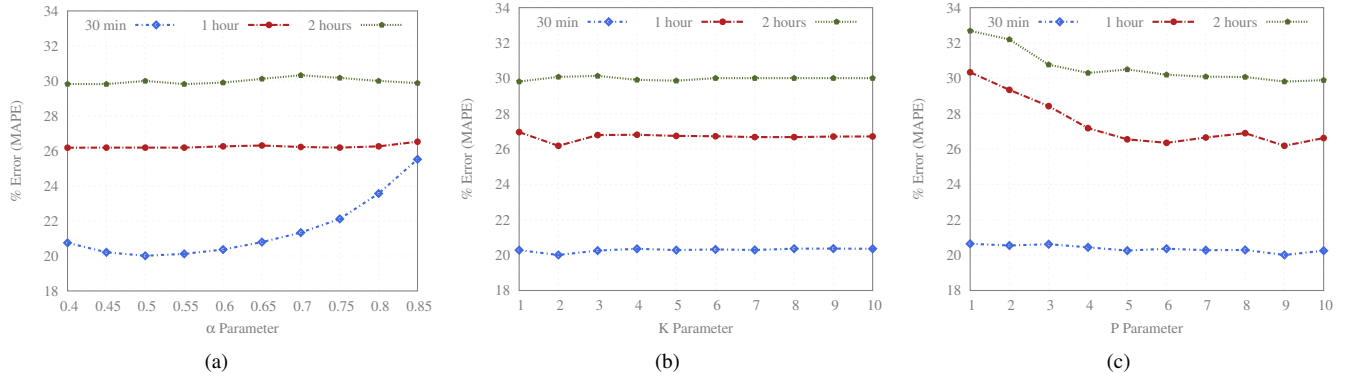


Figure 6. Impact of varying Pro-Energy parameters on the overall MAPE error for 30 minutes, 1 hour and 2 hours prediction horizons: (a) α parameter; (b) K parameter and (c) P parameter.

values of $\alpha > 0.75$, lead to performance degradations: The MAPE error increases of up to 5 percentage points. This is because in such settings Pro-Energy strongly relies on the energy trend expressed by the typical profiles, without effectively adapting them to the current weather conditions. When considering prediction horizons of one and two hours, instead, the performance of Pro-Energy are quite stable with respect to variations of the α parameter, as varying it changes the MAPE error only up to 0.5 percentage points. This is due to the fact that, for such prediction horizon, the value of the G parameter has a stronger impact than in the previous cases on the overall MAPE error.

Figure 6(b) shows that the value of K does not deeply influence the accuracy of energy predictions of short term energy prediction. In fact, for K parameter ranging from 1 to 10, the MAPE error increases only up to ≈ 0.5 percentage points. Such results suggest that, to characterize the future energy intake during the current day, it is enough to consider the energy harvested in the last few timeslots. In fact, the information collected during timeslots that are further away in time does not provide significant improvement in the accuracy of energy predictions. This characteristic allows Pro-Energy to use small values of K while performing MAE computations, thus reducing the overhead of computing similarity between different profiles.

Figure 6(c) shows how varying the number of profiles used by Pro-Energy impacts on the overall MAPE error. The overall trend is that increasing the number of profiles that Pro-Energy takes into account tends to reduce the prediction error. However, in case of short-term predictions, combining multiple profiles together does not have a very strong impact, as it improves the prediction accuracy only up to 0.5 percentage point. This means that the overhead of the Pro-Energy algorithm may be reduced in such case by choosing a small values of P , with limited impact on performance. On the contrary, the effect of the P parameter is much more evident for medium term predictions, as using a combination of multiple profiles can reduce the overall

Table II
PREDICTION ALGORITHMS OVERHEAD OF PRO-ENERGY, WCMA AND EWMA USING THE SOLAR-ROME DATASET

Prediction Algorithm	Number of multiplications
Pro-Energy (1 profile)	4369
Pro-Energy (2 profiles)	12917
Pro-Energy (3 profiles)	17047
Pro-Energy (4 profiles)	21081
Pro-Energy (5 profiles)	25019
WCMA	27931
EWMA	4416

MAPE error up to ≈ 4 percentage points with respect to the setting in which a single profile is used.

D. Pro-Energy Overhead

Table II compares the overhead of Pro-Energy with that introduced by other solutions, in terms of number of multiplications performed by each scheme, computed over the solar-Rome dataset. The values reported in the table refer to a setting of parameters in which $D = 10$ (number of energy profiles stored), $K = 7$ (number of slots used for comparing profiles), $F = 1$ (prediction horizon) and P (number of combined profiles) varies in $\{1, 2, 3, 4, 5\}$.

VI. CONCLUSIONS

In this paper, we have presented Pro-Energy a novel energy prediction model for multi-source energy harvesting WSNs, which is able to provide accurate predictions for short and medium term forecasting horizons. We have performed extensive validation of Pro-Energy using real-life traces of the harvested energy we have obtained by interfacing Telos B nodes with photovoltaic panels and wind micro-turbines. We have also exploited two additional solar and wind datasets from the US National Renewable Energy Laboratory. Our results show that Pro-Energy performs better than previous solutions such as EWMA and WCMA, with improvements in prediction accuracy which can be as high as 60%.

Table III
SOLAR ROME DATASET: PARAMETER SETTINGS OF PRO-ENERGY, WCMA AND EWMA FOR SHORT AND MEDIUM TERM PREDICTIONS.

	Pro-Energy						WCMA				EWMA	
	MAPE Error	α	D	K	G	P	MAPE Error	α	D	K	MAPE Error	α
30m	20.00	0.50	14	2	-	9	21.12	0.6	11	5	35.21	0.2
1h	26.18	0.40	14	2	5	9	35.51	-	10	7	35.21	0.2
1h30m	27.68	0.45	9	6	5	7	35.51	-	10	7	35.21	0.2
2h	29.81	0.45	14	1	6	9	35.51	-	10	7	35.21	0.2
2h30m	31.68	0.60	14	3	6	9	35.51	-	10	7	35.21	0.2
3h	31.92	0.55	18	3	7	9	35.51	-	10	7	35.21	0.2

ACKNOWLEDGMENTS

Dora Spenza is a recipient of the Google Europe Fellowship in Wireless Networking, and this research is supported in part by this Google Fellowship. This research was also sponsored in part by the FP7 project GENESI (Green sEnSOr NETworks for Structural monIToring), Contract n. 257916, and by the CHIRON ARTEMIS project, Contract n. 100228. We thank the MICREL lab, University of Bologna, for sharing their wind data sets.

REFERENCES

- [1] C. Park and P. Chou, "Ambimax: Autonomous energy harvesting platform for multi-supply wireless sensor nodes," in *Proc. of IEEE SECON 2006*, vol. 1, Reston, Virginia, USA, September 25-28, 2006, pp. 168–177.
- [2] V. Raghunathan, A. Kansal, J. Hsu, J. Friedman, and M. Srivastava, "Design considerations for solar energy harvesting wireless embedded systems," in *Proc. of ACM/IEEE IPSN 2005*, UCLA, Los Angeles, CA, USA, April 25-27, 2005, pp. 457–462.
- [3] F. Simjee and P. Chou, "Everlast: Long-life, supercapacitor-operated wireless sensor node," in *Proc. of ACM ISLPED 2006*, Tegernsee, Germany, October 4-6, 2006, pp. 197–202.
- [4] Z. A. Eu, H.-P. Tan, and W. K. Seah, "Opportunistic routing in wireless sensor networks powered by ambient energy harvesting," *Computer Networks: The International Journal of Computer and Telecommunications Networking archive (Elsevier)*, vol. 54, no. 17, pp. 2943–2966, December 2010.
- [5] T. Zhu, Z. Zhong, Y. Gu, T. He, and Z.-L. Zhang, "Leakage-aware energy synchronization for wireless sensor networks," in *Proc. of ACM MobiSys 2009*, Krakow, Poland, June 22-25, 2009, pp. 319–332.
- [6] C. Moser, D. Brunelli, L. Thiele, and L. Benini, "Lazy Scheduling for Energy Harvesting Sensor Nodes," in *From Model-Driven Design to Resource Management for Distributed Embedded Systems (Springer 2006)*, 2006, pp. 125–134.
- [7] C. Moser, J.-J. Chen, and L. Thiele, "Power management in energy harvesting embedded systems with discrete service levels," in *Proc. of ACM/IEEE ISLPED 2009*, San Francisco, CA, USA, August 19-21, 2009, pp. 413–418.
- [8] C. Moser, D. Brunelli, L. Thiele, and L. Benini, "Real-time scheduling for energy harvesting sensor nodes," *Real-Time Syst.*, vol. 37, no. 3, pp. 233–260, Dec. 2007.
- [9] J. Piorno, C. Bergonzini, D. Atienza, and T. Rosing, "Prediction and management in energy harvested wireless sensor nodes," in *Proc. of Wireless VITAE 2009*, Aalborg, Denmark, May 17-20, 2009, pp. 6–10.
- [10] M. Magno, S. Marinkovic, D. Brunelli, E. Popovici, B. O'Flynn, and L. Benini, "Smart power unit with ultra low power radio trigger capabilities for wireless sensor networks," in *Proc. of IEEE DATE 2012*, Dresden, Germany, March 12-16, 2012, pp. 75–80.
- [11] "NREL: Measurement and Instrumentation Data Center," 2011, <http://www.nrel.gov/midc/>.
- [12] A. Kansal, J. Hsu, S. Zahedi, and M. B. Srivastava, "Power management in energy harvesting sensor networks," *ACM Trans. Embed. Comput. Syst.*, vol. 6, no. 4, September 2007.
- [13] D. R. Cox, "Prediction by exponentially weighted moving averages and related methods," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 23, no. 2, pp. 414–422, 1961.
- [14] C. Moser, L. Thiele, D. Brunelli, and L. Benini, "Adaptive power management in energy harvesting systems," in *Proc. of IEEE DATE 2007*, Nice, France, April 16-20, 2007, pp. 773–778.
- [15] C. Bergonzini, D. Brunelli, and L. Benini, "Algorithms for harvested energy prediction in batteryless wireless sensor networks," in *Proc. of IEEE IWASI 2009*, Trani, Italy, June 25-26, 2009, pp. 144–149.
- [16] D. K. Noh and K. Kang, "Balanced energy allocation scheme for a solar-powered sensor system and its effects on network-wide performance," *Journal of Computer and System Sciences*, vol. 77, no. 5, pp. 917–932, September 2011.
- [17] J. Lu, S. Liu, Q. Wu, and Q. Qiu, "Accurate modeling and prediction of energy availability in energy harvesting real-time embedded systems," in *Proc. of IEEE IGCC 2010*, Chicago, IL, USA, August 15-18, 2010, pp. 469–476.
- [18] N. Sharma, J. Gummeson, D. Irwin, and P. Shenoy, "Cloudy computing: Leveraging weather forecasts in energy harvesting sensor systems," in *Proc. of IEEE SECON 2010*, Boston, Massachusetts, USA, June 21-25, 2010, pp. 1–9.
- [19] Crossbow Technology, "TelosB mote platform datasheet," <http://www2.ece.ohio-state.edu/~biky/ee582/telosMote.pdf>.
- [20] M. Ali, B. Al-Hashimi, J. Recas, and D. Atienza, "Evaluation and design exploration of solar harvested-energy prediction algorithm," in *Proc. of IEEE DATE 2010*, Dresden, Germany, March 8-12, 2010, pp. 142–147.