Representing Power Variability of an Idle IoT Edge Node in the Power State Model

Salma Tofaily*, Issam Raïs*, Otto Anshus*,
Department of Computer Science, UiT The Arctic University of Norway, Tromsø, Norway*
Corresponding authors: salma.tofaily, issam.rais@uit.no

***Abstract*—Simulations can be used to efficiently predict and explore energy consumption of nodes in cyber-physical and IoT systems. The Power State Model (PSM), widely used in simulators, uses only a single value for the energy consumption, for each power state of a node. However, for a given state (including the idle state) the actual consumed energy can vary. Consequently, PSM having a single value only per state does not accurately reflect the actual consumed energy.**

Previous research give us measured values for how the energy consumption actually varies for the idle power state, for three Raspberry Pi nodes. From the measurements of a single node, several traces of different sizes for energy consumption over time are extracted. These traces are also extracted in three scenarios considering cold start effects.

This paper proposes to update the PSM, by using the measured values for each extracted trace: (i) as an empirical distribution; (ii) as a percentile distribution; and (iii) as an average with a standard deviation. Simulations are done for each trace, to get predictions for each proposed update of the PSM.

The results show that the impact of changing the size of measurements used to build the model is 4.7 to 8.9 times higher than the impact of the proposed PSM updates. We conclude that increasing the calibration trace size increases PSM accuracy. Trade-off experiments between the size of calibration traces and the model accuracy helps to chose an informed trace size.

***Index Terms*—IoT, power, energy consumption, measurements, variability, power state model, calibration, simulation, Idle**

I. Introduction

Cyber Physical Systems (CPS) and IoT edge nodes can be energy constrained or powered by batteries in several contexts. It can be the case for environmental monitoring CPS [1]–[3], drones on a mission [4], edge systems monitoring the behavior of wild animals [5], and the Internet of Industrial things (IIoT) [6], for example. In several cases, replacing batteries could also be not possible during a long-term of operation [1], [5]. Furthermore, energy harvesting is not always an option [7]. In these contexts, energy efficiency and accurately estimating energy consumption is crucial in order to maintain long battery lifetime. It is important for the quality of service, availability and scalability of IoT and distributed edge systems.

Quantifying power usage can be based on measurements from physical nodes, data sheets, or on estimations of energy consumption using power models. Related works that measure power consumption on physical nodes show that variability of energy consumption on a single node [8], [9] and among multiple homogeneous nodes [8]–[11] exists for identical repeated workloads (e.g AI benchmarks [9], idle state, CPU and RAM intensive scenarios [8]).

This paper focuses on the variability that exists on a single node, for a unique state. This variability is not represented by the power state model (PSM), a widely used model in simulators to predict and estimate energy consumption [12]–[14]. PSM leverages the state-based operations representation that exists in simulators. In PSM, multiple states are defined for a node, where each state has a constant power value [14], [15]. In literature, PSM is calibrated from data-sheets [15] or from average power measurements [16] of experiments on physical nodes. When PSM is built from a physical node, the accuracy of energy consumption predictions increases [17]. Multiple works report a lower accuracy for the idle state [11].

This paper studies how to represent single node variability [8] in PSM, for the idle state. It also investigates how to calibrate the idle state in PSM, in order to have accurate energy consumption predictions. The focus is on the idle state as it is a crucial part of energy consumption.

This paper presents the following contributions:

- Comparing the accuracy of three updates of PSM to reflect variability on a single edge node, for the idle state, by representing real measurements using: (i) empirical distribution, (ii) percentiles distribution, and (iii) average and standard deviation.
- Evaluating the trade-off between the amount of power measurements used to calibrate PSM and the accuracy of the model, for the idle state.

Measurements from physical nodes from the FIT IoT-LAB [18] are extracted from [8].

This paper is organized as follows. Section II presents related work. Section III presents updates of the PSM. Section IV presents the trace re-player. Section V presents the experimental setup. Section VI presents evaluation metrics. Section VII presents results and observations. Finally, section VIII presents conclusion and future work.

II. Related work

Power models [17], [19]–[21] are used in literature to estimate energy consumption of CPS, Wireless Sensor Networks (WSN), or IoT edge systems and nodes. For the purpose of early studying, developing or evaluating new algorithms, architectures, and protocols, taking energy consumption into consideration is of utmost importance.

In literature, power consumption of a node is abstracted by PSM [13]–[15], [17], [20], [22]–[25]. This model is widely used for example in simulators or energy estimation methods [12]–[14], [22]. Instantaneous power of a node is represented as the sum of the power of each node component, in their current states [13], [24]. However, literature shows that power variability can exist for a specific state on a node [8], [9], including the idle state.

A. Simulations with the PSM and the static idle power

PSM is highly adopted in simulators. The power value of each state is set as part of its "calibration phase".

In [13], [25], SimGrid's PSM is used. It is a widely-adopted simulator of distributed systems. Each state is calibrated with one value. In [26], SimGrid is used to model WiFi power consumption. Energy consumption is modeled in four states: idle, transmit, receive, and sleep. In [22], an energy consumption framework for the network simulator NS-3 is presented, where all operations of node components is state based.

In [13], ESDS, a simulation framework for CPS, IoT and edge platforms is proposed and validated. It provides a plugin that models nodes energy consumption, based on the PSM. Each state can be calibrated with a constant power value. An arbitrary number of power states can be defined. Network interfaces energy consumption are modeled by several power states values for an interface. The accuracy of energy predictions in simulated systems will be related to the accuracy of power calibrations.

In [16], a simulator is built to study energy saving techniques for energy constrained CPS. It assumes that power usage for the idle state is constant. Average power is used to calibrate the idle state. Constant power value is multiplied by (i) the duration of the idle phase and (ii) the during of communication. Therefore, the accuracy of chosen power measurements in calibration affects the accuracy of energy consumption estimation of the simulator, for both idle state and communication phases.

In [17], a simulator for WSN, PowerTOSSIM, with a PSM is used. The error of energy consumption estimations is measured when the model is calibrated from the literature. Real experiments are conducted to calibrate the power model. One run for 60 seconds is done for each load. Measured accuracy for the proposed model is higher than from the literature. Low accuracy from the literature is linked to simulation limitations and mismatching experimental conditions between calibration and simulation experiments.

Previously presented works do not study the impact of the methodology to calibrate PSM.

B. Variability of power for a state

In [9], the variability of energy consumption on a single edge node and multiple homogeneous edge nodes is quantified for multiple AI benchmarks. It is shown that the variability on a single node is not avoidable for these benchmarks.

In [8], the variability of power and energy consumption is quantified, by thorough experiments, on a single edge node and multiple homogeneous edge nodes, for three scenarios and states: Idle, RAM intensive stress, and CPU intensive stress. The variability of energy observed on a single edge node is roughly estimated for a period of a month. It is considered equivalent to an idle up time of 2 h, 5.5 h, and 7 h for Idle, RAM and CPU scenarios, respectively.

Several works aim at increasing the accuracy of energy consumption measurements and predictions of PSM for IoT [9], [10], [14], [27]. However, the art is not exhaustive on representing the variability of power for a specific workload. In [10], the effectiveness of having energy measurements from every node is investigated. In [19], measurements from multiple homogeneous nodes in power models is shown to increase prediction accuracy. These works do not study the representation of power variability from a single node in PSM.

C. Considering power variability, in measurement methodologies

In [28], experiments are conducted running workloads on physical nodes, while measuring power. Each experiment is run for 15 seconds, only. In [29], the energy consumption of several states is measured for Raspberry Pi2 B, laptop, smartphone, and tablet. Energy measurements for each state are a maximum of 2 minutes. In [21], a mathematical power model is developed for Raspberry Pi B. Power usage is modeled per component. Node idle state is characterized by one run of 15 minutes. Average idle power is reported. In [30], calibration of power models for IoT edge nodes are done. Measurements of 15 minutes are done for each component. The duration is said to be chosen as a compromise between measurement duration and accuracy. Authors argue that the effects of short variations in system load on power consumption are difficult to include in any model. In these works, idle power is characterized based on the average power measurements, without repeating identical experiments.

In [11], experiments are run stressing several components, separately. Mathematical power models are built for node components. All distinct experiments are conducted for 15 minutes and repeated 3 times. Average energy consumption is reported for idle. The accuracy of idle predictions is the lowest reported accuracy in the paper. The chosen experimental protocol is not explained. In [31], a software power model is created and evaluated for Linux single board computers. For measuring the idle power, 30 runs are repeated in an experiment. The duration is not given. The reason for having multiple repetitions is to mitigate the effect of system background processes. Measurements show little noise during the idle state. The work considers power measurements to be reliable. These works do not investigate different sizes of power measurements, to calibrate PSM.

In [9], a dynamic size of measurements is taken to measure the energy consumption and variability of AI benchmarks. Five warm-up iterations of each benchmark are excluded to eliminate application-level cold-start effects. In [32], authors discard the results of the first interval in order to avoid cold start effects. These works do not analyze the size of initial measurements to be removed.

D. Summary

In the literature, for computing average idle power, a fixed number of power measurement is usually conducted. Without being experimentally studied, the duration of measurements is short and the number of runs is low. Initial measurements are commonly not removed. The amount of removed measurements are not studied.

Up to our knowledge, no related work investigates how to represent power variability in PSM. No work studies the effect of power measurement size on the accuracy of predictions.

This paper investigates multiple approaches to represent variability of power, by updating PSM, and calibrating it using different measurement methodologies. It studies how to represent single node idle power variability in simulations, in order to enhance prediction accuracy.

III. Updated versions of PSM

PSM calibrated with the average or median power is studied. As PSM associates only one static power value for a state (e.g, in [13]), three updates for the model are suggested and studied, in order to make it variability-aware. The aim is to find out if the prediction accuracy can increase, with the suggested models. Comparison is made with the static PSM.

The studied models are noted as follows.

(I) static_mean: PSM where the power of a state is described by a constant, the *average* power measurements from a physical node, a common practice [17].

(II) static_median: PSM where the power of a state uses the *median* of power measurements, from a physical node. The median is a representative of central tendency. It is rarely used in the literature to calibrate the PSM.

(III) var-empirical-dist: Updated PSM where the power of a state is described with an *empirical distribution* of power measurements. A frequency distribution of power measurements is recorded. Power predictions from the model are variable and follow the recorded empirical distribution.

(IV) var-unif-q1-q3: Updated PSM where predictions for a state are done from a *uniform distribution* between two values the 25th and the 75th percentiles of power measurements, from a physical node.

(V) var-unif-avg-stdev: Updated PSM where predictions for a state are done from a *uniform distribution* between the average and its standard deviation. Power measurements might not follow normal distributions. However, as average and standard deviation are used in literature to describe variability [8], [11], this model is used for comparison.

IV. Trace re-player

Existing models implemented in simulators do not take node variability into consideration in PSMs. An implementation of the previously exposed updates of the PSM, along with a trace re-player, is mandatory.

In a *replay experiment* (i) a calibration power trace, (ii) an original power trace, and (iii) a power model are needed. The *calibration trace* is a set of power measurements collected from experiments. It is used to calibrate the proposed PSM. The *original trace* is a set of timestamped power measurements collected from experiments. Experiments are done on a physical node. The delta between the prediction done by the proposed PSMs and the original trace is used to study models accuracy. The model outputs a power prediction at each timestamp. Consequently, the trace re-player outputs a new *predicted trace*, using the associated model. The duration of the predicted trace is the same as the duration of the original trace. A comparison between the total energy by the model predictions (i.e the predicted trace) and the original trace is done. The trace re-player facilitates studies of the accuracy of power predictions of the models. Automation scripts use the trace re-player to facilitate repeatability of experiments. It is open source on GitHub [1].

V. Experimental setup

A. Idle power measurements from physical nodes

To calibrate the power models, timestamped power measurements are used from [8], observing single node variability on physical IoT edge nodes, for the idle state. These measurements are from *3 identical Raspberry Pi 3B* nodes, noted rpi3-1, rpi3-3, rpi3-4. In [8], the followed experimental protocol is the following one: (i) instant power is measured every *0.2 second*, using an external and ultra precise power monitoring device, ina226 [2]; (ii) from each node, measurements are used from *10 identical experiments*, each of *100 iterations*; (iii) each iteration includes *60 seconds* of idle power measurements, preceded by *60 seconds* of cool-down.

Power measurements are cleaned as follows: *First*, the beginning and end of each iteration are automatically removed, from any interference due to the experimental protocol logging timestamps. The aim is to keep only idle power measurements. Consequently, the average duration for an idle *iteration* becomes *59.8 seconds*. *Next*, for each node, power measurements are concatenated chronologically to have one consequent idle trace. Each of the 3 generated traces constitutes *1000 iterations* and has idle power measurements for a duration of approximately *16.6 h*. These traces are noted as the *original traces*.

B. Simulation and replay

Calibration power traces: Multiple methodologies for extracting power measurements to calibrate the PSM from the original traces are applied.

The scenarios are the following: (A) all 1000 iterations, (B) from iteration 21 to 1000, and (C) from iteration 101 to 1000. A, B, and C represent scenarios without considering cold-start

[1] https://github.com/SalmaTofaily/TraceReplayer_TracesData_CPSCom23
[2] https://www.ti.com/product/INA226

effect, with considering cold-start effect, and with exhaustive consideration of cold-start effect, respectively.

In addition, for each scenario, a trade-off is explored between the size of the measurements used for calibration (i.e the size of the trace used to calibrate a given PSM) and the accuracy of the model predictions. Thus, the first extracted trace contains measurements from the first iteration only. The second trace contains measurements from the first two iterations, and so on. The last extracted trace is identical to the original trace.

Therefore, for each of the 3 nodes, multiple traces are used to calibrate PSMs. The total is 8640 different calibration traces (extracted from 3 physical nodes, with 3 distinct methodologies). In several works, one or few iterations of power measurements are considered. We aim at exploring the impact of power measurements size and different cold-start on models accuracy which is, up to our knowledge, not studied in related works [9], [28], [32].

Original power traces: In a replay experiment, the original trace and calibration trace are from the same node.

The original trace is used as a baseline to compute the delta in accuracy of the power models predictions.

Replay experiments: For each calibration trace, one replay experiment, simulating 16.6 hour, is conducted with each of the previously presented power models.

Given the power models and the calibration traces, 43200 replay experiments are conducted (8640 different calibration traces, each trace is used to build 5 models). These experiments represent around 27 years of real time experiments, for each node.

VI. Evaluation Metrics

In the formulas, the metrics time, power, and energy are in seconds (s), Watts (W), and Joules (J), respectively.

Delta energy: In a replay experiment, delta energy, noted $\Delta E(rExp)$, is the difference between energy consumption in the predicted trace, noted $E(predicTrace)$, and the original trace, noted $E(orgTrace)$. It is defined as follows:

$$\Delta E(rExp) = E(predicTrace) - E(orgTrace) \quad (1)$$

Negative $\Delta E(rExp)$ indicates that the power model in the replay experiment underestimates energy consumption.

Time estimation for delta energy: $equiv\Delta T(rExp)_{month}$ is a rough estimation of $\Delta E(rExp)$ in a month, noted $equiv\Delta E(rExp)_{month}$, translated as idle up-time. We make the assumption that corresponding original trace duration is increased to a month (30 days, each of 24 hours, each of 60 minutes, each of 60 seconds). They are defined as:

$$equiv\Delta E(rExp)_{month} = \frac{\Delta E(rExp) * 30 * 24 * 60 * 60}{duration(orgTrace)} \quad (2)$$

$$equiv\Delta T(rExp)_{month} = \frac{equiv\Delta E_{month}(rExp)}{AvgIdleP(orgTrace)} \quad (3)$$

$duration(orgTrace)$ is the duration of the original trace in the replay experiment (16.6 h), as $\Delta E(rExp)$ is observed for this duration. $AvgIdleP(orgTrace)$ is the average power of the original trace. It is computed as 1.3613 W, 1.3668 W, and 1.3599 W for the nodes rpi3-1, rpi3-3, rpi3-4, respectively.

Translating delta energy to time makes it easier to understand models' accuracy. It illustrates how far expectations of remaining lifetime for edge nodes can be, when power variability is not represented.

VII. Results and Observations

Idle power traces are replayed with PSM and the three suggested model updates. Delta energy between predicted and measured consumption is studied.

A. *Evaluating the impact of using an updated PSM*

For all studied nodes and scenarios, the plots of power models "static_mean", "var-empirical-dist" and "var-unif-avg-stdev" are overlapping. In the presented experiments, no gain in accuracy from using "var-empirical-dist" and "var-unif-avg-stdev" is achieved, compared to the widely used "static_mean". In scenario A (fig. 1(a), fig. 2(a), fig. 3(a)), prediction accuracy using "static_median" and "var-unif-q1-q3" is often slightly higher than prediction accuracy of other models.

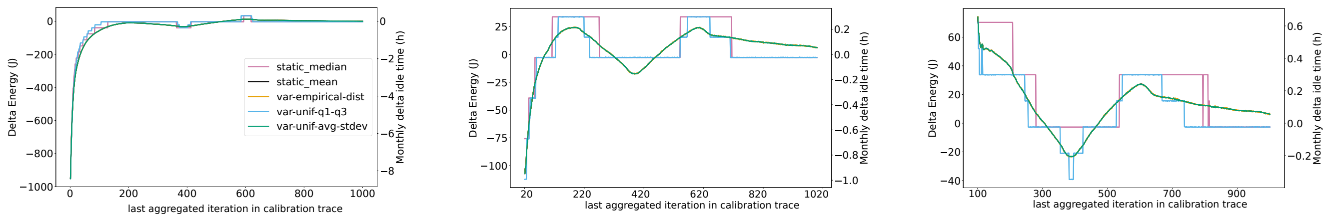
The maximum difference in $\Delta E(rExp)$ among replays with the same calibration trace size, and with different PSMs is 219 J, 94 J, and 77 J in scenarios A(fig. 2(a)), B(fig. 2(b)) and C(fig. 2(c)), respectively. Corresponding maximum difference in $equiv\Delta T(rExp)_{month}$ is 1.93 h, 0.83 h, and 0.68 h, respectively. These values represent the maximum impact of changing PSMs on accuracy. On the other hand, the maximum range of $\Delta E(rExp)$ when increasing the calibration trace size from 1 iteration to the maximum is 1168 J, 835 J, and 364 J, in scenarios A(fig. 3(a)), B(fig. 2(b)), and C(fig. 2(c)), respectively. Corresponding maximum difference in $equiv\Delta T(rExp)_{month}$ is 10.35 h, 7.36 h, and 3.21 h, respectively. These values represent the maximum impact of increasing the calibration trace size on accuracy.

In scenarios A, B, and C, the maximum change in delta energy caused by increasing the calibration trace size (from 1 iteration to the maximum) is 5.3, 8.9 and 4.7 times higher than the one caused by using different PSMs, for a specific calibration trace size. Changing the size of the calibration trace is more impact-full than using an updated version of PSM, for the studied idle state.

The impact of trace extraction methodology is higher than the usage of updated PSM. A relatively low benefit in accuracy can be achieved by having an informed decision when choosing between using the median or the mean, for the idle state. Improving predictions accuracy can be done by increasing the size of power measurements used to calibrate PSM.

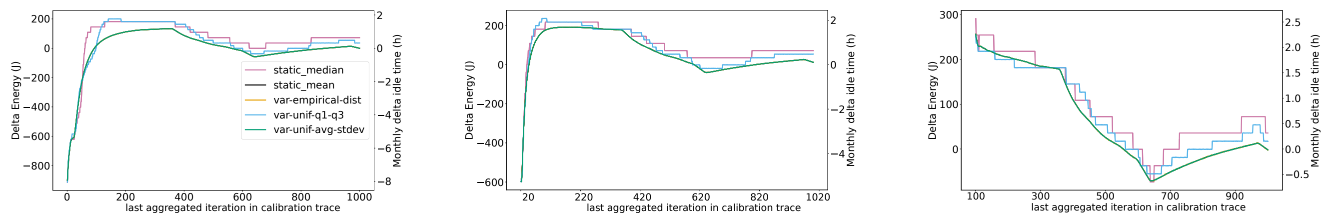
B. *Energy consumption when first power measurements are not discarded*

In scenario A (figs. 1(a), 2(a) and 3(a)), calibration trace size ranges between 1 and 1000 iterations, for each node.



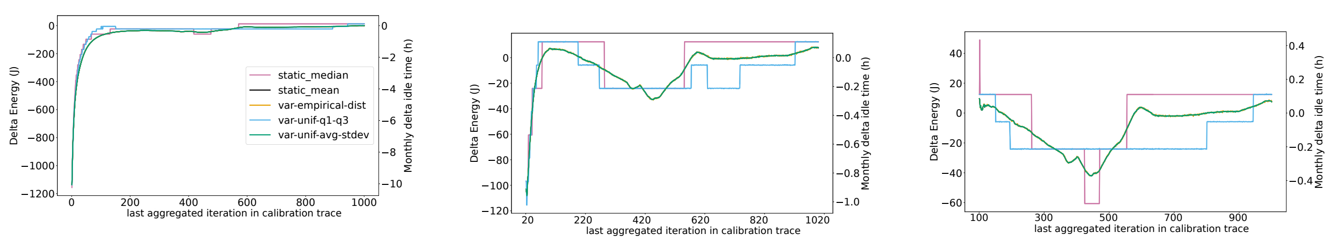
(a) Scenario A: Calibration traces from all iterations. (b) Scenario B: Calibration traces from iteration 21 to 1000. (c) Scenario C: Calibration traces from iteration 101 to 1000.

Fig. 1. Replay experiments. Power models are built from calibration traces recorded from rpi3-1, for the idle state. Power traces are extracted with 3 methodologies, using different sizes. Left y axis is delta energy, in Joules, between predicted and original traces. Right y axis is its equivalent idle up time, in a month. The "static_mean", "var-empirical-dist", and "var-unif-avg-stdev" plots overlap.



(a) Scenario A: Calibration traces from all iterations. (b) Scenario B: Calibration traces from iteration 21 to 1000. (c) Scenario C: Calibration traces from iteration 101 to 1000.

Fig. 2. Replay experiments. Power models are built from calibration traces recorded from rpi3-3, for the idle state. Power traces are extracted with 3 methodologies, using different sizes. Left y axis is delta energy, in Joules, between predicted and original traces. Right y axis is its equivalent idle up time, in a month. The "static_mean", "var-empirical-dist", and "var-unif-avg-stdev" plots overlap.



(a) Scenario A: Calibration traces from all iterations. (b) Scenario B: Calibration traces from iteration 21 to 1000. (c) Scenario C: Calibration traces from iteration 101 to 1000.

Fig. 3. Replay experiments. Power models are built from calibration traces recorded from rpi3-4, for the idle state. Power traces are extracted with 3 methodologies, using different sizes. Left y axis is delta energy, in Joules, between predicted and original traces. Right y axis is its equivalent idle up time, in a month. The "static_mean", "var-empirical-dist", and "var-unif-avg-stdev" plots overlap.

Across the three presented homogeneous nodes, when the model is built from the first iteration only (i.e first point on each curve), $\Delta E(rExp)$ ranges between -900 J and -1155 J, and $equiv\Delta T(rExp)_{month}$ ranges between -7.93 h and -10.23 h. As the number of iterations in the calibration trace increases toward 100, $\Delta E(rExp)$ gets closer to 0 J. At 100 iterations, $\Delta E(rExp)$ is between -70 J and 145 J, and $equiv\Delta T(rExp)_{month}$ is between -0.62 h and 1.28 h. Positive delta indicates that predicted energy is higher than the actual energy consumption.

In this scenario, when calibration traces are extracted from the first 100 iterations of real measurements, predicted energy is lower than the actual energy consumption, except for fig. 2(a). Accuracy of energy consumption prediction improves as calibration trace size increases. Predicted remaining idle uptime can be underestimated by 7.93 h up to 10.23 h, for a month, when the calibration is from the first iteration. When estimating energy consumption for power-constrained nodes, avoiding underestimations is important.

In the context of power-constrained CPS, IoT and edge systems, being as accurate as possible when it comes to energy prediction is mandatory. It is also the case in simulators, where the discovered delta is not taken into account. As it is at the node level, this error in accuracy can spread, especially in large simulations. To increase predictions accuracy, the size of power measurements used to calibrate PSM needs to be studied.

C. Trade-off between models accuracy enhancement and calibration trace size

In scenario A (figs. 1(a), 2(a) and 3(a)), when the calibration trace size is equal to 1000 iterations, $\Delta E(rExp)$ is near 0 J. However, in this case, the original and calibration traces are identical. As iterations count in the calibration trace increases, from 100 to 200, $\Delta E(rExp)$ improves in low percentages. However, adding more iterations do not always improve accuracy (eg, rpi3-3 when calibration trace size is increased to 600 iterations, in fig. 2(a)).

The percentage of maximum change in $\Delta E(rExp)$ represents the maximum change in delta energy for calibrations with traces up to a specified number of iterations, compared to the maximum change considering all iterations. In this scenario, 35.98 % to 72.87 % of the maximum change in $\Delta E(rExp)$ is achieved when calibrating from the first 20 iterations. When calibrating from the first 50, 60, and 100 iterations, 75.95 % to 91.6 %, 92.06 % to 99.92 % and 97.82 % to 100 % of the maximum change in $\Delta E(rExp)$ are achieved, respectively.

Results show that using more than 60 iterations, when no first measurements are removed, does not highly enhance the accuracy. These numbers are expected to be different for other homogeneous nodes. The difference is related to the observed variability.

In related literature, experimental protocols for collecting power measurements from edge nodes vary in number of iterations (1 [28], 3 [11], 30 [31]). The trade-off study done in this work helps to have a justified size of power measurements for calibrating PSM.

D. Insights about removing initial power measurements from calibration traces

In scenario B (figs. 1(b), 2(b) and 3(b)), idle power measurements are from iteration 21 to 1000 (i.e initial 20 iterations are removed from the calibration traces). When the calibration trace size is equal to 1 iteration, $\Delta E(rExp)$ is between -598 J and -75 J (on rpi3-3 and rpi3-1), and its corresponding $equiv\Delta T(rExp)_{month}$ is between -5.27 h and -0.66 h. Excluding the power measurements of the first 20 iterations in the calibration traces still result in having negative $\Delta E(rExp)$. Furthermore, as the number of iterations in the calibration trace increases, $\Delta E(rExp)$ enhances gradually to be positive for the first time when calibration trace size is 69, 22, and 42 iterations, for rpi3-1, rpi3-3, and rpi3-4, respectively. Important number of iterations are needed in order to increase the accuracy of the power models and minimize underestimations. Studying the size of initial power measurements to be removed is beneficial to increase PSM accuracy.

In scenario C (figs. 1(c), 2(c) and 3(c)), idle power measurements are from iteration 101 to 1000 (i.e the first 100 iterations are removed from the calibration trace). $equiv\Delta T(rExp)_{month}$ is between -0.64 h and 2.56 h. Removing initial 100 iterations from calibration traces minimizes underestimations in predictions, to be almost negligible. Maximum underestimated $equiv\Delta T(rExp)_{month}$ in scenario B (5.27 h) is higher than scenario C (0.64 h).

Part of the related work removes initial power measurements from monitoring data that is used for calibration in order to eliminate cold start effects [9]. The size of initial power measurements to be removed is not chosen based on experiments. Results reveal that accuracy is increased when the calibration skips a certain amount of iterations.

VIII. Conclusion

Energy is a limited resource for battery-powered CPS and IoT edge nodes. Previous research estimate energy consumption to evaluate energy efficient techniques, predicting and optimizing remaining lifetime. Accurately characterizing energy consumption is crucial in contexts where energy is limited or constrained, a multiplication of nodes are used, and it is hard to scale and maintain. These include CPS where IoT edge nodes are in rural areas, attached to animals in a forest, under hard weather conditions like snow, or attached to drones. PSM is widely used to estimate energy consumption for nodes, in simulators and energy estimation methods. PSM uses a static power value for a state. Typically, the methodology of taking energy measurements on a node is not backed up by preliminary experiments in several research works. For example, a few repetitions of small duration of idle measurements are assumed to be enough, without explanation. However, previous art shows that there is a power variability for a specific state of a single node [8].

This paper presents a study of how to modify and calibrate PSM, in order to more accurately represent variability on a single node, for the idle state. The study is based on real measurements on physical nodes. Three updates are studied for the model, to represent power variability. In addition, the impact of multiple sizes of power monitoring data on models accuracy are studied. Furthermore, the impact of three methodologies for considering cold start effects on models accuracy are explored. The results are presented for three homogeneous Raspberry Pi nodes.

For the studied idle state, results show that when first power measurements are not discarded, predictions underestimate actual energy consumption. In addition, predictions for a month, based on only few iterations create underestimations of around 10 hours, for a single node. The impact on the accuracy of the models, from changing calibration trace size, is 4.7 to 8.9 times higher than from the impact of using an updated PSM. In the scenario where no initial power measurements are discarded, adding more than 60 iterations does not enhance the accuracy.

We conclude that, for the studied idle state, using a large amount of power measurements for calibration increases PSM accuracy. It has more impact than using the proposed updates of the PSM to represent power variability, on a single node. A trade-off between the size of traces used for calibration and the model accuracy is underlined.

In simulators, the variability in power estimations might be propagated. Calibrating the idle state in PSM with caution can result in improved accuracy, as idle is an ongoing power consumption on an active node.

Future works include investigating the impact of representing variability in PSM for other states and for other nodes, especially when power variability is higher. In addition, we plan to build an automated framework to design experimental protocols to calibrate PSM.

References

[1] I. Raïs, J. M. Bjørndalen, P. H. Ha, K.-A. Jensen, L. S. Michalik, H. Mjøen, Ø. Tveito, and O. Anshus, "Uavs as a leverage to provide energy and network for cyber-physical observation units on the arctic tundra," in *2019 15th International Conference on Distributed Computing in Sensor Systems (DCOSS)*. IEEE, 2019, pp. 625–632.

[2] P. Giri, K. Ng, and W. Phillips, "Wireless sensor network system for landslide monitoring and warning," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 4, pp. 1210–1220, 2018.

[3] E. Kökten, B. C. Çalışkan, S. Karamzadeh, and E. G. Soyak, "Low-powered agriculture iot systems with lora," in *2020 IEEE Microwave Theory and Techniques in Wireless Communications (MTTW)*, vol. 1, 2020, pp. 178–183.

[4] J. Yao and N. Ansari, "Power control in internet of drones by deep reinforcement learning," in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, 2020, pp. 1–6.

[5] K. Nakagawa, D. Shimotoku, J. Kawase, and H. H. Kobayashi, "Sustainable wildlife dtn: Wearable animal resource optimization through intergenerational multi-hop network simulation," in *2021 17th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, 2021, pp. 193–199.

[6] J. Arshad, M. A. Azad, M. M. Abdeltaif, and K. Salah, "An intrusion detection framework for energy constrained iot devices," *Mechanical Systems and Signal Processing*, vol. 136, p. 106436, 2020.

[7] M. J. Murphy, M. Tveito, E. F. Kleiven, I. Raïs, E. M. Soininen, J. M. Bjørndalen, and O. Anshus, "Experiences building and deploying wireless sensor nodes for the arctic tundra," in *2021 IEEE/ACM 21st International Symposium on Cluster, Cloud and Internet Computing (CCGrid)*, 2021, pp. 376–385.

[8] S. Tofaily, I. Raïs, and O. Anshus, "Quantifying the variability of power and energy consumption for iot edge nodes," in *2023 19th International Conference on Distributed Computing in Smart Systems and the Internet of Things (DCOSS-IoT)*, 2023, pp. 577–584.

[9] H. A. Abdelhafez, H. Halawa, A. Almoallim, A. Ahmadi, K. Pattabiraman, and M. Ripeanu, "Characterizing variability in heterogeneous edge systems: A methodology & case study," in *2022 IEEE/ACM 7th Symposium on Edge Computing (SEC)*. IEEE, 2022, pp. 107–121.

[10] Q. Li, M. Martins, O. Gnawali, and R. Fonseca, "On the effectiveness of energy metering on every node," in *2013 IEEE International Conference on Distributed Computing in Sensor Systems*. IEEE, 2013, pp. 231–240.

[11] M. S. Aslanpour, A. N. Toosi, R. Gaire, and M. A. Cheema, "Wattedge: a holistic approach for empirical energy measurements in edge computing," in *Service-Oriented Computing: 19th International Conference, ICSOC 2021, Virtual Event, November 22–25, 2021, Proceedings 19*. Springer, 2021, pp. 531–547.

[12] K. Lahmar, R. Chéour, and M. Abid, "Wireless sensor networks: Trends, power consumption and simulators," in *2012 Sixth Asia Modelling Symposium*. IEEE, 2012, pp. 200–204.

[13] L. Guegan, I. Raïs, and O. Anshus, "Validation of esds using epidemic-based data dissemination algorithms," in *2023 19th International Conference on Distributed Computing in Smart Systems and the Internet of Things (DCOSS-IoT)*, 2023, pp. 277–284.

[14] A. Sabovic, C. Delgado, J. Bauwens, E. De Poorter, and J. Famaey, "Accurate online energy consumption estimation of iot devices using energest," in *Advances on Broad-Band Wireless Computing, Communication and Applications: Proceedings of the 14th International Conference on Broad-Band Wireless Computing, Communication and Applications (BWCCA-2019) 14*. Springer, 2020, pp. 363–373.

[15] I. Raïs, L. Guegan, and O. Anshus, "Impact of loosely coupled data dissemination policies for resource challenged environments," in *2022 22nd IEEE International Symposium on Cluster, Cloud and Internet Computing (CCGrid)*, 2022, pp. 524–533.

[16] I. Raïs, O. Anshus, J. M. Bjørndalen, D. Balouek-Thomert, and M. Parashar, "Trading data size and cnn confidence score for energy efficient cps node communications," in *2020 20th IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing (CCGRID)*, 2020, pp. 469–478.

[17] T. Trathnigg, M. Jürgen, and R. Weiss, "A low-cost energy measurement setup and improving the accuracy of energy simulators for wireless sensor networks," in *Proceedings of the workshop on Real-world wireless sensor networks*, 2008, pp. 31–35.

[18] C. Adjih, E. Baccelli, E. Fleury, G. Harter, N. Mitton, T. Noel, R. Pissard-Gibollet, F. Saint-Marcel, G. Schreiner, J. Vandaele *et al.*, "Fit iot-lab: A large scale open experimental iot testbed," in *2015 IEEE 2nd World Forum on Internet of Things (WF-IoT)*. IEEE, 2015, pp. 459–464.

[19] H. A. Abdelhafez, H. Halawa, M. O. Ahmed, K. Pattabiraman, and M. Ripeanu, "Mirage: Machine learning-based modeling of identical replicas of the jetson agx embedded platform," in *2021 IEEE/ACM Symposium on Edge Computing (SEC)*, 2021, pp. 26–40.

[20] A. D. Jun, S. Hong, W. Lee, K. Lee, I. Joe, K. Lee, and T.-J. Park, "Modeling and simulation of lora in opnet," in *Advanced Multimedia and Ubiquitous Engineering*, J. J. J. H. Park, S.-C. Chen, and K.-K. Raymond Choo, Eds. Singapore: Springer Singapore, 2017, pp. 551–559.

[21] F. Kaup, P. Gottschling, and D. Hausheer, "Powerpi: Measuring and modeling the power consumption of the raspberry pi," in *39th Annual IEEE Conference on Local Computer Networks*. IEEE, 2014, pp. 236–243.

[22] H. K. Wu, S. Nabar, and R. Poovendran, "An energy framework for the network simulator 3 (ns-3)," in *International ICST Conference on Simulation Tools and Techniques*, 2011.

[23] R. A. Bergamaschi and Y. W. Jiang, "State-based power analysis for systems-on-chip," in *Proceedings of the 40th annual Design Automation Conference*, 2003, pp. 638–641.

[24] A. Dunkels, J. Eriksson, N. Finne, and N. Tsiftes, "Powertrace: Network-level power profiling for low-power wireless networks," 2011.

[25] M. Quinson, C. Rosa, and C. Thiéry, "Parallel simulation of peer-to-peer systems," in *2012 12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (ccgrid 2012)*. IEEE, 2012, pp. 668–675.

[26] C. Courageux-Sudan, A.-C. Orgerie, and M. Quinson, "A wi-fi energy model for scalable simulation," in *24th IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, 2023.

[27] C. Wan, M. Santriaji, E. Rogers, H. Hoffmann, M. Maire, and S. Lu, "Alert: Accurate learning for energy and timeliness," in *Proceedings of the 2020 USENIX Conference on Usenix Annual Technical Conference*, 2020, pp. 353–369.

[28] C. N. Cabaccan, F. Reidj, and G. Cruz, "Power characterization of raspberry pi agricultural sensor nodes using arduino based voltmeter," in *2018 3rd International Conference on Computer and Communication Systems (ICCCS)*, 2018, pp. 349–352.

[29] G. Bekaroo and A. Santokhee, "Power consumption of the raspberry pi: A comparative analysis," in *2016 IEEE International Conference on Emerging Technologies and Innovative Business Practices for the Transformation of Societies (EmergiTech)*, Aug 2016, pp. 361–366.

[30] F. Kaup, S. Hacker, E. Mentzendorff, C. Meurisch, and D. Hausheer, "The progress of the energy-efficiency of single-board computers," *Tech. Rep. NetSys-TR-2018-01*, 2018.

[31] L. Ardito and M. Torchiano, "Creating and evaluating a software power model for linux single board computers," in *2018 IEEE/ACM 6th International Workshop on Green And Sustainable Software (GREENS)*, 2018, pp. 1–8.

[32] R. Kumar, K. Farkas, N. Jouppi, P. Ranganathan, and D. Tullsen, "Single-isa heterogeneous multi-core architectures: the potential for processor power reduction," in *Proceedings. 36th Annual IEEE/ACM International Symposium on Microarchitecture, 2003. MICRO-36.*, 2003, pp. 81–92.