Poster Abstract: Measuring Predictability of Ambient Energy Harvesting with Calibrated Probabilistic Forecasting

Ding Huo
TU Darmstadt University
ding.huo@tu-darmstadt.de

Marco Zimmerling
TU Darmstadt University
marco.zimmerling@tu-darmstadt.de

***Abstract*—Battery-free devices harvest ambient energy, but environmental dynamics make harvested power volatile. When future energy is predictable, devices can schedule sensing and communication proactively; otherwise, they must act conservatively. We define environment predictability as the history-conditioned uncertainty of future harvested energy, namely the irreducible component of forecasting error. To estimate this uncertainty as a lower bound on achievable prediction error, we compare three estimators: a k-nearest-neighbors residual estimator, an information-theoretic estimator, and a calibrated probabilistic neural forecaster. The neural forecaster outputs a predictive distribution, allowing us to assess calibration directly; when calibration holds, its predictive variance provides an empirical estimate of the noise floor and thus of the fundamental limits of predictability. Here, calibrated probabilistic forecasting serves not as a deployable runtime solution, but as a tool for offline predictability analysis.**

***Index Terms*—Energy harvesting, intermittent computing, probabilistic forecasting, aleatoric uncertainty, calibration, mixture density networks, k-nearest neighbors, information-theoretic estimation**

I. Motivation and Overview of Approach

Energy-harvesting devices benefit when future energy availability can be anticipated: they can then schedule sensing, computation, and communication proactively rather than acting conservatively. This motivates a principled measure of predictability based only on a device's observed energy history. Our goal is not to build a deployable runtime predictor, but to use offline analysis to estimate the history-conditioned irreducible uncertainty over a given horizon. Unlike prior work that targets hour–day horizons [2], we focus on ms–s horizons required by battery-free devices with tiny energy buffers.

Let X denote a history window of length h from the harvested-energy time series. Let Y denote the next p future steps, where $Y^{(j)}$ is the harvested energy j steps ahead. We quantify predictability through the variance decomposition:

$$\mathrm{Var}(Y^{(j)}) = \mathrm{Var}(\mathbb{E}[Y^{(j)} \mid X]) + \mathbb{E}[\mathrm{Var}(Y^{(j)} \mid X)]. \quad (1)$$

The second term is the irreducible aleatoric uncertainty given the history X. It represents a mean squared error (MSE) noise floor that no predictor using only X can outperform in expectation.

This work was funded by the LOEWE initiative (Hesse, Germany) within the emergenCITY center [LOEWE/1/12/519/03/05.001(0016)/72].

We consider three estimators of the history-conditioned noise floor. A **k-nearest-neighbors residual estimator (kNN residual estimator)** directly approximates $\mathbb{E}[\mathrm{Var}(Y^{(j)} \mid X)]$ by gathering samples with similar history windows X and measuring the variance of their corresponding futures $Y^{(j)}$. An **information-theoretic estimator (IT estimator)** instead estimates the mutual information $I(X; Y^{(j)})$ in the embedded space and converts it into an MSE lower bound via the Gaussian rate-distortion relation:

$$\mathrm{MSE}_{\text{avg}} \geq \frac{1}{p} \sum_{j=1}^{p} \mathrm{Var}(Y^{(j)})\, 2^{-2I(X;Y^{(j)})}. \quad (2)$$

Both estimators therefore rely on neighborhood structure in a reconstructed history space. With finite, noisy traces, this makes them sensitive to regime mixing, sparse neighborhoods for longer histories, and heuristic choices such as the history representation and neighbor count k.

In contrast, a **calibrated probabilistic neural forecaster** learns task-relevant representations directly from the history window and outputs a predictive distribution. From this distribution, we estimate the irreducible term through the averaged

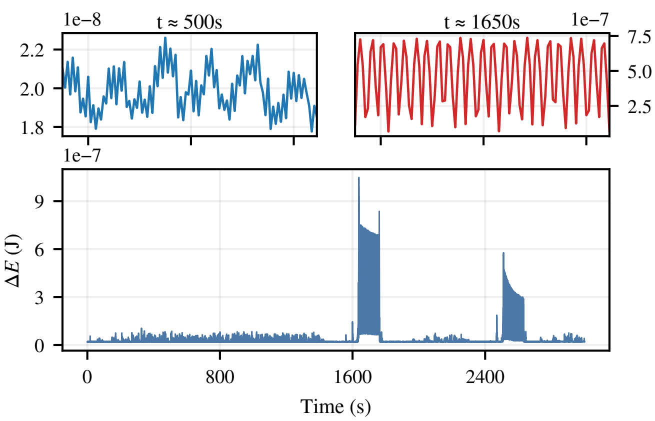

Fig. 1. Top: two zoom-in views centered at 500 s and 1650 s, each with a 1 s window. Bottom: the full integrated harvested-energy increment sequence ΔE (10 ms windows).

local predictive variance:

$$\widehat{B} \triangleq \frac{1}{p}\sum_{j=1}^{p} \widehat{\mathbb{E}}_X\left[\mathrm{Var}_\theta\Big(Y^{(j)} \mid X\Big)\right]. \tag{3}$$

where $\widehat{B}$ denotes the estimated history-conditioned noise floor. We interpret $\widehat{B}$ as credible only when the forecaster is well calibrated [3], i.e., when predicted probabilities match empirical frequencies. We assess calibration using the Probability Integral Transform (PIT), which should be close to uniform when observations fall into predicted quantiles at the correct rate. When calibration holds, CRPS provides a complementary measure of sharpness, and the predictive variance can be interpreted as an empirical estimate of the irreducible uncertainty.

II. Preliminary Results

In this section, we evaluate these three estimators on synthetic and real-world datasets to assess how well they estimate the irreducible uncertainty.

A. Data and settings

We evaluate two time-series settings. First, a synthetic trace serves as a controlled sanity check: we generate a smooth nonstationary signal with added i.i.d. Gaussian noise ($\sigma = 10^{-9}$) to establish a known ground truth. Second, we analyze a real-world trace from a piezoelectric harvester mounted on an industrial washing machine [1] (Fig. 1). We derived the energy time series by integrating raw 100 kHz power samples into 10 ms increments.

For both datasets, the task is to predict the next $p = 5$ steps (50 ms). We set the history window $h = 50$ for the synthetic case and evaluate $h \in \{9, 15, 30\}$ for the real trace. We split the data into Train/Val/Test sets and normalize inputs using training statistics.

B. Models and estimators

We compare the estimated noise floors against the test MSEs of two deterministic point predictors, a Multi-Layer Perceptron (MLP) and a Long Short-Term Memory (LSTM) network. If a lower bound is accurate, the MSEs of these predictors should approach it. Our probabilistic neural forecaster uses an LSTM backbone to parameterize a Gaussian Mixture distribution. We train via negative log-likelihood loss. Provided the model is well-calibrated, we compute the irreducible uncertainty $\widehat{B}$ (Eq. (3)) analytically from the learned distribution.

C. Results

1) Synthetic Trace: On the synthetic trace, we inject Gaussian noise with variance $\sigma^2 = 10^{-18}$. Our target is the history-conditioned uncertainty $\mathbb{E}_X[\mathrm{Var}(Y^{(j)} \mid X)]$, which can exceed σ^2 because a finite history window X may not fully determine the noiseless future. Using Eq. (3), the probabilistic forecaster yields an MSE noise floor of $1.177 \times 10^{-18}\,\mathrm{J}^2$, compared with $1.498 \times 10^{-18}\,\mathrm{J}^2$ from the kNN residual estimator and $2.484 \times 10^{-18}\,\mathrm{J}^2$ from the information-theoretic estimator. Among the three, the neural estimate is the closest to the known synthetic noise floor.

TABLE I
Results on the real-world washer piezoelectric trace with $p = 5$ (50 ms). MSE values are reported in $\times 10^{-18}$.

h	MLP	LSTM	kNN Residual est.	IT est.	Calibrated Neural est.
9	5.443	6.985	22.099	1156.000	5.318
15	3.818	2.921	5.983	1014.423	2.786
30	4.260	2.247	3.021	1170.324	3.375

2) Washer Trace: Table I reports the real-world washer-trace results for $p = 5$, including the test MSEs of the deterministic predictors and the estimated noise floors. On the washer trace, the calibrated neural estimate decreases from $5.318 \times 10^{-18}\,\mathrm{J}^2$ at $h = 9$ to $2.786 \times 10^{-18}\,\mathrm{J}^2$ at $h = 15$, suggesting that additional history reduces the apparent noise floor over this range. At $h = 15$, the forecaster is well calibrated, with PIT mean/var $= 0.5047/0.0838$ close to the ideal $0.5/0.0833$, and it also achieves a low CRPS of 6.21×10^{-10}. We therefore interpret its predictive variance as a credible estimate of the history-conditioned noise floor. This estimate is close to the MLP and LSTM test MSEs, suggesting only a small reducible gap at this (p, h) scale, whereas the kNN residual and IT estimates remain much looser. For the largest history window, however, the PIT departs from uniformity, so we treat the corresponding variance-based estimate as unvalidated, possibly due to data sparsity at longer histories or model misspecification.

III. Conclusion and Future Work

These preliminary results suggest that calibrated probabilistic forecasting is a promising tool for offline predictability analysis in energy-harvesting traces. In our experiments, the neural estimate is closest to the known noise floor on the synthetic trace, and on the real washer trace it provides a plausible estimate of the history-conditioned noise floor when calibration holds. This, in turn, helps separate irreducible uncertainty from the remaining reducible error of deterministic predictors. At the same time, our study is still limited to one synthetic and one real-world trace, and our current validation framework relies mainly on global diagnostics such as PIT and CRPS, which may miss local miscalibration when biases cancel out across regions. Broader evaluation and more rigorous validation methods therefore remain important future work.

References

[1] K. Geissdoerfer and M. Zimmerling, "Learning to Communicate Effectively Between Battery-free Devices," in *Proc. 19th USENIX Symposium on Networked Systems Design and Implementation (NSDI 22)*, Renton, WA, USA, Apr. 2022, pp. 419–435. [Online]. Available: https://www.usenix.org/conference/nsdi22/presentation/geissdoerfer

[2] B. Buchli, F. Sutton, J. Beutel, and L. Thiele, "Dynamic Power Management for Long-Term Energy Neutral Operation of Solar Energy Harvesting Systems," in *Proc. 12th ACM Conf. Embedded Network Sensor Systems (SenSys '14)*, Memphis, TN, USA, Nov. 2014, pp. 31–45, doi:10.1145/2668332.2668333.

[3] T. Gneiting, F. Balabdaoui, and A. E. Raftery, "Probabilistic forecasts, calibration and sharpness," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 69, no. 2, pp. 243–268, 2007, doi:10.1111/j.1467-9868.2007.00587.x.