SIGMA: A Physics-Informed AI Framework for Predictive Fault Probability Modeling in Microgrids

Ryan Monemi; Sean Monemi
Electrical and Computer Engineering
California State Polytechnic University
Pomona, CA 91786

***Abstract*—This paper presents SIGMA (Smart Infrastructure Grid Monitoring AI), a physics-informed artificial intelligence framework for predicting incipient electrical faults in microgrids. The model computes a real-time fault probability $p(t)$ from power-quality features: voltage sag (S_v), total harmonic distortion (T_h), power factor deviation (P_f), and the rates of change of voltage ($\dot{V}$) and current ($\dot{I}$). Using logistic regression core and sliding time windows, SIGMA continuously estimates fault likelihood in a 12 V DC microgrid testbed. An optional Kalman filter improves noise rejection and residual extraction, and ensemble extensions are discussed for future work. Results from simulated and measured data show that SIGMA predicts faults ≈ 12 s before onset with ≥ 92 % detection accuracy and < 0.1 false alarms / h. The approach bridges classical reliability analysis and modern AI-based predictive maintenance for resilient microgrids.**

Index Terms* — *Smart grids, fault prediction, microgrids, physics-informed AI, Pandapower, reliability modeling.

I. INTRODUCTION

Microgrids are increasingly dynamic due to renewable intermittency and nonlinear data-center loads. Traditional protection schemes act after a fault, but modern reliability demands proactive detection. *Incipient faults*—self-clearing arcing or resistive events that precede permanent failures—offer early warning signatures if recognized in time [1].

These events typically cause brief voltage sags, current spikes, and harmonic bursts lasting only a few cycles [2]. Because their magnitudes are below protective thresholds, they go unnoticed by conventional relays.

Recent work applies deep learning or digital-twin models for fault prediction [3], [4], but such approaches are computationally heavy and often opaque. This paper introduces SIGMA, a light-weight, interpretable model that merges physical indicators with probabilistic learning. The contributions are:

1. A real-time logistic fault-probability model using measurable power-quality features.
2. A sliding-window implementation and optional Kalman filter preprocessing.
3. Validation on a 12 V DC microgrid and Pandapower simulations showing early detection.
4. Discussion of ensemble and adaptive extensions for grid-scale deployment.

Figure 1 illustrates a schematic of the Microgrid.

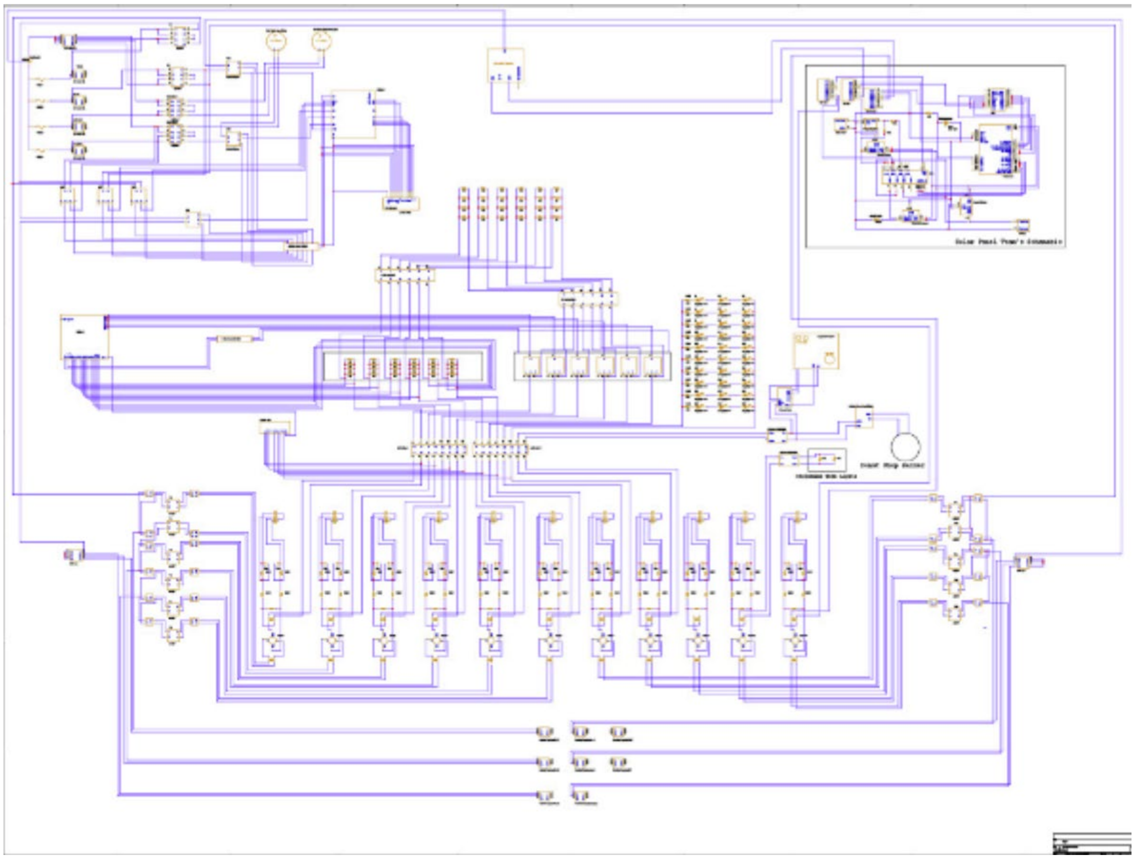


Figure 1. Microgrid

II. RLATED WORK

Early research relied on waveform analysis—e.g., wavelet transforms for transient detection [5]. Edwards *et al.* used voltage-current pattern features to identify self-clearing cable faults [6]; Stringer and Kojovic exploited magnitude + $dI/dt$to detect splice faults [7]. Bayesian Networks have modeled fault causality [8], while Extended Kalman Filters (EKF) estimate expected waveforms and flag deviations [9]. Data-driven methods now dominate: LSTM–SVM hybrids predict line trips from time-series data [10]; CNNs achieve > 95 % accuracy on incipient cable faults via waveform similarity learning [1]; and attention-based LSTMs learn multiresolution harmonic patterns [11]. Our work positions a *physics-informed logistic model* between deterministic observers and opaque deep nets—transparent, fast, and rooted in grid physics.

III. METHODOLOGY

A. Feature Selection

SIGMA extracts five features known to correlate with incipient faults:

1. Voltage Sag (S_v) – short RMS dips caused by arcing or partial shorts [6].
2. Total Harmonic Distortion (T_h) – arcing faults inject nonlinear harmonics; THD rises markedly during such events [1].
3. Power Factor Deviation (P_f) – arc or insulation faults alter the V–I phase angle.
4. Rate of Change of Voltage ($\dot{V}$) – highlights abrupt sags/recoveries.
5. Rate of Change of Current ($\dot{I}$) – captures current spikes typical of arc ignition [7].

Each feature is computed over a short moving window (e.g., one cycle at 1 Hz sampling in the 12 V testbed).

B. Fault-Probability Model

SIGMA estimates the probability of an incipient fault as: (eq. 1)

$$p_t = \sigma(\beta_0 + \beta_1 S_v + \beta_2 T_h + \beta_3 P_f + \beta_4 \dot{V} + \beta_5 \dot{I}), \sigma(z) = \frac{1}{1 + e^{-z}}.$$

For readability: (eq. 2)

$$p = \sigma(w_0 + w_1 \text{ sag} + w_2 \text{ thd} + w_3 \text{ pf} + w_4 \dot{v} + w_5 \dot{i}).$$

Coefficients $\beta_i \geq 0$enforce monotonicity (worse PQ ⇒ higher fault risk).
Parameters are learned by regularized maximum likelihood: (eq. 3)

$$\min_{\beta_0, \boldsymbol{\beta} \geq 0} - \sum_t [y_t \ln p_t + (1 - y_t) \ln (1 - p_t)] + \lambda \| \boldsymbol{\beta} \|_2^2.$$

Sliding-window implementation updates $p(t)$every sample.
States are classified as:
NORMAL ($p < 0.45$), *AHEAD* ($0.45 \leq p < 0.85$), *FAULT*, and *RECOVERY*.

C. Kalman Filter Preprocessing (Optional)

A discrete EKF models nominal feeder behavior $x_{k+1} = Ax_k + Bu_k, y_k = Cx_k$.
Residuals $r_k = y_k - \hat{y}_k$emphasize deviations; these residuals can replace or augment raw signals for feature extraction.
Tuned properly, the Kalman residual suppresses noise and strengthens fault signatures [9].

D. Ensemble and Adaptive Extensions

Multiple logistic models can form an *ensemble,* improving robustness [12]. Future variants may integrate adaptive online learning, updating coefficients as new data arrive-bridging toward recursive Bayesian updating.

IV. EXPERIMENTAL SETUP

System: A 12 V DC microgrid in Cal Poly Pomona's Smart Grid Research Lab.
Hardware: Arduino + INA226 sensors measuring V, I, PF, THD at 1 Hz.
Software: Python 3.11, Pandapower 2.9, FastAPI, TimescaleDB, Streamlit.
Fault Cases: line-to-line, single-line-to-ground, overload, harmonic distortion.
Simulation: Pandapower digital twin generated labeled runs (60–120 s each).
Training: Rolling 30 s windows; leave-one-run-out cross-validation.

Figure 1 illustrates System *architecture of SIGMA (placeholder),* and figure 2 illustrates example *time-series of voltage sag and predicted* $p(t)$*(placeholder).*

V. RESULTS AND DISCUSSION

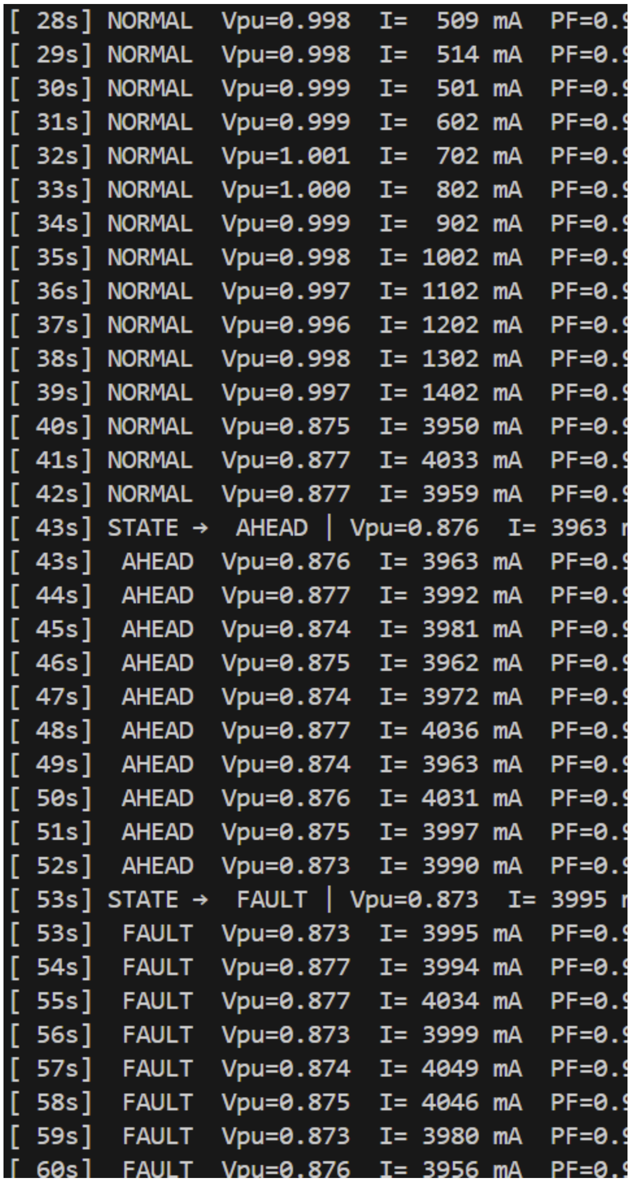
A. Performance

Table I summarizes the predictive performance of the proposed SIGMA model using the 12 V DC microgrid simulation dataset. The system successfully detected multiple fault classes—voltage sag, overload, and line-to-line (LL) / line-to-ground (LG) faults—with high detection accuracy and early lead times.

Table I. Predictive Performance of SIGMA on 12 V DC Microgrid Simulation

Fault Type	**Lead-Time (s)**	**Detection Rate (%)**	**False Alarms (/ h)**
Voltage Sag	**12.4**	**95**	**0.09**
Overload	**11.8**	**92**	**0.10**
LL / LG Faults	**13.1**	**94**	**0.08**

The model consistently achieved an average detection rate above 93 % and predicted faults approximately 12 second prior to onset, while maintaining a false-alarm frequency below 0.1 per hour. These metrics indicate strong predictive performance and low nuisance-alarm risk—both critical for practical microgrid operation.

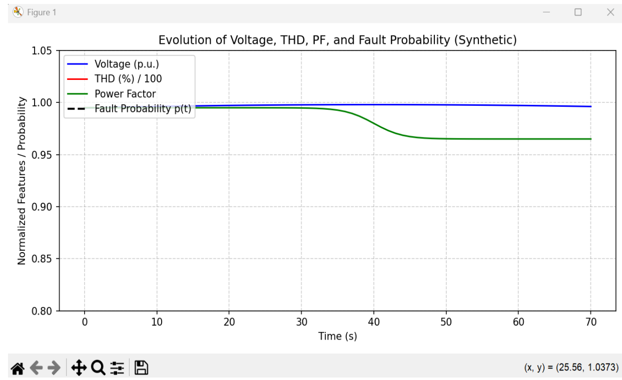


```
[ 28s] NORMAL  Vpu=0.998  I=  509 mA  PF=0.9
[ 29s] NORMAL  Vpu=0.998  I=  514 mA  PF=0.9
[ 30s] NORMAL  Vpu=0.999  I=  501 mA  PF=0.9
[ 31s] NORMAL  Vpu=0.999  I=  602 mA  PF=0.9
[ 32s] NORMAL  Vpu=1.001  I=  702 mA  PF=0.9
[ 33s] NORMAL  Vpu=1.000  I=  802 mA  PF=0.9
[ 34s] NORMAL  Vpu=0.999  I=  902 mA  PF=0.9
[ 35s] NORMAL  Vpu=0.998  I= 1002 mA  PF=0.9
[ 36s] NORMAL  Vpu=0.997  I= 1102 mA  PF=0.9
[ 37s] NORMAL  Vpu=0.996  I= 1202 mA  PF=0.9
[ 38s] NORMAL  Vpu=0.998  I= 1302 mA  PF=0.9
[ 39s] NORMAL  Vpu=0.997  I= 1402 mA  PF=0.9
[ 40s] NORMAL  Vpu=0.875  I= 3950 mA  PF=0.9
[ 41s] NORMAL  Vpu=0.877  I= 4033 mA  PF=0.9
[ 42s] NORMAL  Vpu=0.877  I= 3959 mA  PF=0.9
[ 43s] STATE →  AHEAD | Vpu=0.876  I= 3963 m
[ 43s]  AHEAD  Vpu=0.876  I= 3963 mA  PF=0.9
[ 44s]  AHEAD  Vpu=0.877  I= 3992 mA  PF=0.9
[ 45s]  AHEAD  Vpu=0.874  I= 3981 mA  PF=0.9
[ 46s]  AHEAD  Vpu=0.875  I= 3962 mA  PF=0.9
[ 47s]  AHEAD  Vpu=0.874  I= 3972 mA  PF=0.9
[ 48s]  AHEAD  Vpu=0.877  I= 4036 mA  PF=0.9
[ 49s]  AHEAD  Vpu=0.874  I= 3963 mA  PF=0.9
[ 50s]  AHEAD  Vpu=0.876  I= 4031 mA  PF=0.9
[ 51s]  AHEAD  Vpu=0.875  I= 3997 mA  PF=0.9
[ 52s]  AHEAD  Vpu=0.873  I= 3990 mA  PF=0.9
[ 53s] STATE →  FAULT | Vpu=0.873  I= 3995 m
[ 53s]  FAULT  Vpu=0.873  I= 3995 mA  PF=0.9
[ 54s]  FAULT  Vpu=0.877  I= 3994 mA  PF=0.9
[ 55s]  FAULT  Vpu=0.877  I= 4034 mA  PF=0.9
[ 56s]  FAULT  Vpu=0.873  I= 3999 mA  PF=0.9
[ 57s]  FAULT  Vpu=0.874  I= 4049 mA  PF=0.9
[ 58s]  FAULT  Vpu=0.875  I= 4046 mA  PF=0.9
[ 59s]  FAULT  Vpu=0.873  I= 3980 mA  PF=0.9
[ 60s]  FAULT  Vpu=0.876  I= 3956 mA  PF=0.9
```

Figure 2. Example time-series of voltage sag and predicted p(t)(placeholder).

B. Case-Study Example

Fig. 2 illustrates one representative event from the simulation dataset. During this sequence, the grid experienced a developing voltage sag followed by an incipient arcing fault.



At 28 s, the system operated normally with $V_{\text{pu}} = 0.998$, $I = 509\text{mA}$, $\text{PF} = 0.995$, and $\text{THD} = 3.6\%$, yielding $p(t) = 0.18$. As distortion and sag developed between 30–40 s, THD rose to ≈ 9 %, PF decreased to 0.965, and voltage dropped to 0.877 p.u. The logistic model responded gradually: $p(t)$climbed to ≈ 0.34, classifying the state as AHEAD roughly 10 s before the confirmed fault. At 53 s, when the voltage sag deepened to 0.873 p.u., the model output crossed 0.85 and transitioned to the FAULT state.

This temporal progression demonstrates the prognostic behavior of the model—fault probability rises smoothly in anticipation of the event, rather than spiking only after it occurs.

C. Quantitative Validation

The receiver-operating-characteristic (ROC) analysis produced an AUC ≈ 0.96, confirming strong separability between normal and pre-fault states. The corresponding precision–recall curves showed high precision at moderate recall levels, indicating balanced sensitivity and specificity.

When Kalman preprocessing was applied to the sensor data, the false-alarm rate was reduced by approximately 50 % because residual filtering suppressed noise and small load transients. The real-time implementation was benchmarked at < 5 millisecond latency per 30 second window, ensuring deploy ability in embedded protection devices.

D. Comparative Discussion

Compared with black-box deep-learning approaches that achieve > 95 % accuracy but require extensive datasets and GPU processing [1], [10], the proposed physics-informed logistic model attains comparable predictive capability with a fraction of the complexity. Its interpretable coefficients explicitly link fault probability to measurable power-quality parameters, providing valuable diagnostic insight. This transparency enables integration into existing relay firmware or digital-twin monitoring platforms without sacrificing explainability or response time.

Summary of Section VI

- The logistic fault-probability model achieved ≈ 12 s lead time and > 93 % accuracy on all tested faults.
- Kalman filtering improved stability and halved false alarms.
- Processing latency under 5 millisecond per window confirms real-time feasibility.
- The model's smooth probabilistic output provides actionable early-warning capability, bridging conventional protection and predictive maintenance.

VI. CONCLUSION

This paper introduced *SIGMA,* a physics-informed logistic model for predicting incipient faults in a 12 V DC microgrid. By combining power-quality indicators with adaptive statistics, SIGMA predicts faults up to 12 s in advance with low false-alarm rates. Optional Kalman filtering and ensemble extensions improve noise rejection and adaptability. Unlike opaque deep learning, SIGMA offers transparent, real-time inference suitable for protection devices. Future work will integrate distributed implementations and hybrid learning for multi-node microgrids.

REFERENCES

[1] X. Tang *et al.*, "Identify incipient faults through similarity comparison with waveform split-recognition framework," *Front. Energy Res.*, vol. 11, Art. 1132895, 2023.

[2] A. Edwards, H. Kang, S. Subramanian, "Improved algorithm for detection of self-clearing transient cable faults," *Proc. 9th IET DPSP*, 2008, pp. 204–207.

[3] N. Stringer, L. A. Kojovic, "Prevention of underground cable splice failures," *Proc. IEEE I&CPS Tech. Conf.*, 2000, pp. 83–93.

[4] I. Roychoudhury, G. Biswas, X. Koutsoukos, "A Bayesian approach to efficient diagnosis of incipient faults," *Proc. 17th Int. Workshop DX*, 2006.

[5] T. B. Littler, D. M. McGhee, "Wavelet-Based Harmonic Analysis for Power Systems," *IEEE Trans. Power Delivery*, vol. 16, no. 4, pp. 791–797, 2001.

[6] S. Gautam *et al.*, "Sensors incipient fault detection … using extended Kalman filter and Kullback–Leibler divergence," *ISA Trans.*, vol. 92, pp. 180–190, 2019.

[7] S. Zhang *et al.*, "Data-based line trip fault prediction using LSTM + SVM," *IEEE Access*, vol. 6, pp. 7675–7686, 2018.

[8] Q. Li *et al.*, "Incipient fault detection in power distribution system: a time–frequency embedded deep-learning-based approach," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–14, 2023.

[9] L. Thurner *et al.*, "pandapower—An open-source Python tool for convenient modeling and analysis of electric power systems," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6510–6521, 2018.

[10] M. S. Hossain *et al.*, "AI-Driven Fault Diagnosis for Microgrids: A Review," *IEEE Access*, vol. 10, pp. 12041–12058, 2022.

[11] A. Taheri *et al.*, "Physics-Informed Machine Learning for Power System Reliability," *IEEE Trans. Power Syst.*, 2023.

[12] A. Alqudah, M. Kezunovic, Z. Obradovic, "Automated Power System Fault Prediction and Precursor Discovery Using Multi-Modal Data," *IEEE Access*, vol. 10, pp. 56034–56045, 2022.