Towards Energy-Efficient Buildings

Timothy Mulumba, Rita Sousa

Department of Civil & Urban Engineering
New York University

NeurIPS 2025: Tackling Climate Change with Machine Learning Workshop

Motivation & Summary

HVAC systems can account for **30–40%** of building energy use, so timely Fault Detection & Diagnosis (FDD) in chillers matters for cost and climate. We present a deployment-oriented **ARX→SVM** pipeline with:

- **Forgetting-factor tuning** to align estimator memory with macro-F1,
- **Clean ablations** vs. a static regression baseline at matched feature budgets,
- **Latency/memory audit**: sub-ms per-timestep RLS (recursive least squares) updates on commodity hardware,
- **Uncertainty & significance**: time-series-aware CIs and paired tests.

Results on ASHRAE RP-1043 [1] show competitive macro-F1 while preserving interpretability and runtime efficiency. (See paper for details.)

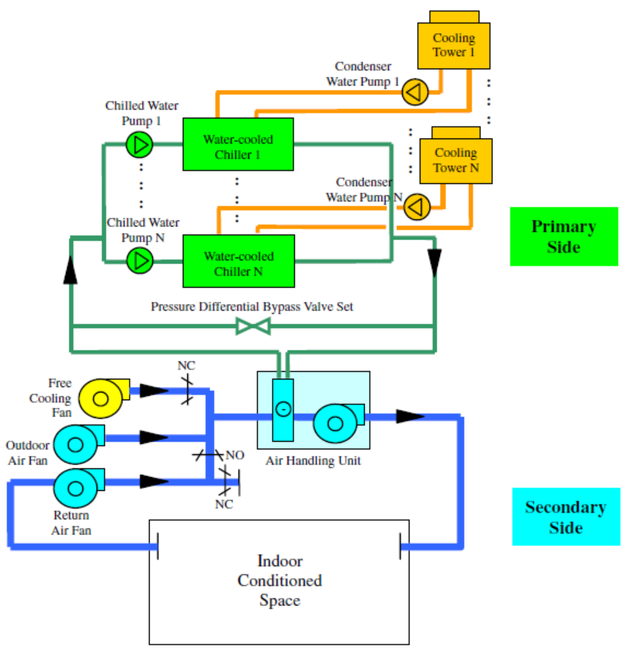

Figure 1: *Typical HVAC system.*

Data & Setup (ASHRAE RP-1043)

Public chiller dataset with normal/faulty operation at 2-min sampling; we respect temporal order (train first block, test later), tune hyperparameters on training-only validation, and report mean ±95% CI over repeated temporal splits. (Details in paper.)

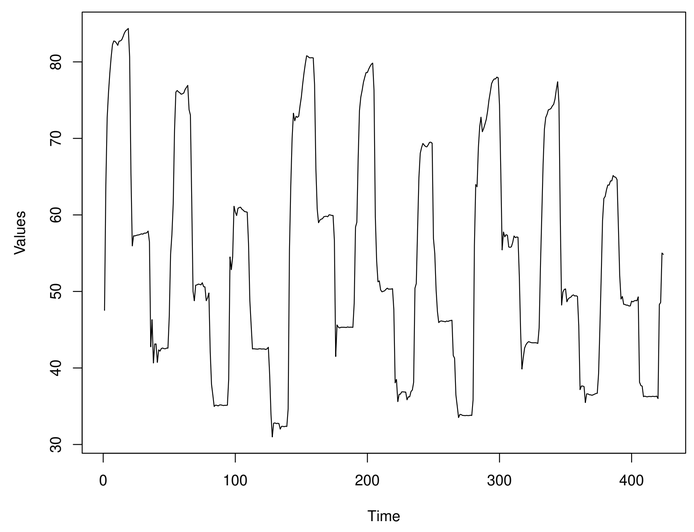

Figure 2: *Chiller power (kW) over time (illustrative segment). Temporal structure motivates ARX features.*

Method: Hybrid ARX→SVM

ARX model. With target $y(t)$ (chiller power) and exogenous inputs $x_k(t)$, we estimate

$$y(t) = \sum_{j=1}^{p} \alpha_j\, y(t-j) + \sum_{k=1}^{m} \sum_{\ell=0}^{q_k} \beta_{k,\ell}\, x_k(t-\ell) + \varepsilon(t).$$

We fit $\theta(t) = [\alpha, \beta]$ **online** via Recursive Least Squares (RLS) with exponential forgetting:

$$\lambda = \exp(-\Delta/\tau_f)\,,$$

linking the forgetting factor λ to a time constant τ_f (with sampling interval Δ). We select lags p, q_k using domain knowledge and **ACF/PACF** screening.

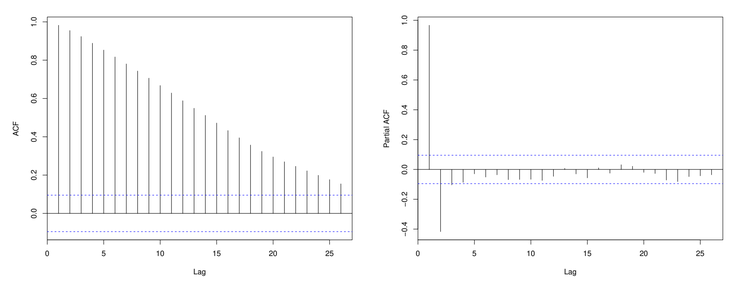

Figure 3: *ACF (left) and PACF (right) of chiller kW guide p, q_k.*

Classifier. At each time step we feed the ARX parameter vector $\theta(t)$ into an SVM, comparing against NB, LR, RF, XGBoost, and a regression-only baseline (no lags).

Offline: tune λ on validation F1 and train SVM

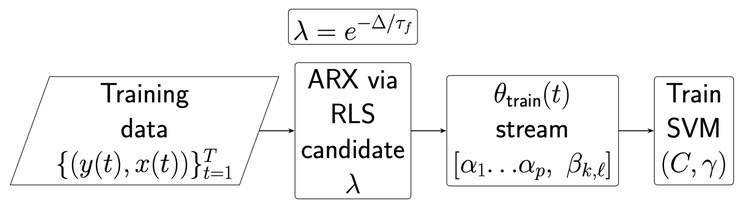

Figure 4: *Offline ARX→SVM inference path. $\theta(t)$ are ARX parameters. Choose $\lambda^\star$ to maximize validation macro-F1.*

Online: estimate $\theta(t)$ and classify.

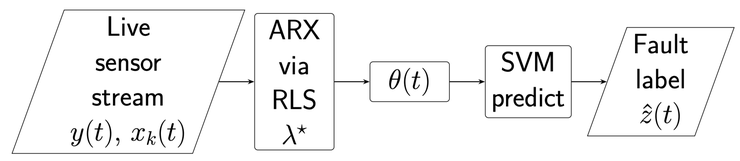

Figure 5: *Online ARX→SVM inference path.*

Tuning Memory for FDD

We **tune** λ to maximize validation macro-F1 (golden-section search). In our study, $\lambda^\star = 0.9935$ corresponds to ~ 180 samples; at a 2-min interval this is ~ 6 hours of memory, a good balance between reactivity and stability for evolving faults.

Results on RP-1043

Ablation: ARX features substantially outperform regression-only across classifiers.

Table 1: *Mean macro-F1 (%) ± 95% CI. RF vs. SVM CIs overlap; paired McNemar tests show no sig. difference at $\alpha = 0.05$.*

Classifier	ARX-based F1 (%)	Regression-only F1 (%)
Naïve Bayes	73.1 (±3.2)	58.4 (±4.1)
Logistic Regression	73.9 (±3.5)	59.8 (±3.1)
XGBoost	78.1 (±2.2)	68.4 (±3.3)
Random Forest	81.4 (±3.7)	66.9 (±3.5)
SVM (proposed)	**84.0** (±3.1)	69.2 (±3.6)

Runtime: Online RLS updates + SVM prediction are **sub-ms** per time step on a laptop; training the final SVM is modest. CIs use a *stationary/block bootstrap*; paired tests via *McNemar*.

Deployment & Maintenance

RLS forgetting adapts short-term dynamics; under concept drift, schedule periodic refresh or incremental SVM updates (simple triggers: rolling macro-F1 or calibration drift).

Climate Impact

Accurate, fast FDD for chillers reduces wasted energy and GHG emissions, and the same hybrid recipe extends to AHUs/FCUs for broader building impact.

References

[1] M. C. Comstock and J. E. Braun. Experimental data from fault detection and diagnostic studies on a centrifugal chiller. *Purdue University*, 1999.