Consensus-Based Anomaly Detection for PV-Integrated Building Energy Management: Seasonal Characterisation and Severity Stratification

Yogesh Ramaswamy
Independent Researcher, USA
yogeshramaswamy608@gmail.com

Narendra Chennupati
Independent Researcher
Texas, USA
chennupati.be@gmail.com

Harish K
Assistant Professor, CSE-ICB
Dayananda Sagar Academy of Technology and Management, Bangalore, Karnataka
kharish1987@gmail.com

Rameshkumar T
Assistant Professor (Sl. Gr)
Electrical and Electronics Engineering
KPR Institute of Engineering and Technology, Coimbatore, Tamil Nadu, India
tramesh22@gmail.com

S. Sri Nandhini Kowsalya
Assistant Professor
Department of Information Technology,
Vel Tech Multi Tech Dr. Rangarajan Dr. Sakunthala Engineering College,
Chennai, Tamil Nadu, India
srinandhinikowsalya@gmail.com

Pruthvi Raj Seknametla
Independent Researcher
pruthviraj9369@gmail.com

***Abstract*— Static threshold based alarms continue to be the most common type facing energy management in research facilities, but fail due to missing complex operational anomalies that are seasonal and equipment state dependent. In this study, we applied four unsupervised machine learning algorithms (namely, Isolation Forest, Local Outlier Factor, One-Class SVM and Autoencoder) to an unprecedented 8,760 hours of consumption data collected from the GECAD Research Center in Porto from January to December 2023. We also defined a consensus-based detection framework that produced a mean deviation of 1,543 kW peaking at 7,722 kW thresholding on three algorithms to agree before raising any alarms, identifying 117 high-confidence anomalies with a 1.34% detection rate. When analysing by season, anomaly rates during winter ran 3.3x higher than spring and autumn (2.73% vs. 0.82%), while cold-weather events registered 100% deviations from consumption compared to between 18–35% in the warmer months respectively. The pairwise Jaccard similarity scores were all between 0.061 and 0.237 confirming the large differences in the populations of anomalies detected by these four algorithms (Bertens et al., 2023). The entire pipeline executed in 29.1 seconds on standard CPU hardware, allowing for the feasibility of real-time deployment.**

I. INTRODUCTION

1.1 Motivation and Background

Approximately 40 % of electricity generated throughout the globe is consumed in buildings and research facilities, yet most power management systems still use static threshold-based alarms which trigger when consumption exceeds a predetermined limit [1]. In a lab, real consumption makes large variations due to occupancy, experimental equipment duty cycles, HVAC response to weather and PV generation — not something that can be captured in a static cutoff of fixed threshold [2]. A sharp spike could be simple high water usage or a malfunctioning chiller, while a moderate discharge during off-hours might indicate the valve is stuck open or the control logic failed [3]. It is costly to get this wrong either way — too many false alarms and staff stop responding, too few and real faults go un recognised until they snowball [4-5].

With unsupervised machine learning you have a much adaptive way of doing it [6]. Unlike conventional methods based on predetermined rules, these techniques learn from what normal looks like in historical data and signal outliers that do not conform — without requiring labeled examples of faults, which are rarely available in practice[7]. The interdependencies between PV generation, experimental loads, HVAC and computing infrastructure in a research facility represent exactly the kind of complex, context-dependent patterns that could benefit from data-driven detection[8].

1.2 Literature Review

There have been several unsupervised anomaly detection methods to building energy management, we first summarised four families of these methods. Isolation Forest operates by recursively partitioning the data – as an anomaly lies in a sparse region, it is isolated sooner which leads to fewer splits needed [9-10]. It is computationally efficient, yet sensitive to contamination parameter tuning, as it can raise false alarms for borderline changes in operation, when thresholds are excessively stringent [11]. Local Outlier Factor (LOF) uses a density-based method wherein it scores each point based on how isolated its local neighborhood is compared to other neighborhoods [12]. This can be effective at identifying contextual anomalies — like consumption that is a perfectly regular thing when viewed on a global time scale, yet such an unusual behaviour for that time of day or operational mode —however it does not hold up well to sufficiently large datasets[13].

The One-Class SVM analyzes high-dimensional feature space specifying a kernel-based boundary around normal operations and flags points that fall outside [14-15]. It does a great job at classifying non-linear patterns, but only catches the most extreme deviations, risking that gradual equipment degradation will be missed[16]. On the other hand, autoencoder networks follow a different path — they are trained to learn a compressed representation of normal behaviour and then identify observations that reconstruct badly [17]. This makes them sensitive to small correlation violations, but the downside is a less interpretable model and requiring additional hyperparameter tuning.

Despite these advances, key gaps remain. Studies mostly test single algorithms in isolation, so there is no clear picture of how these four approaches compare on identical datasets or whether combining them reduces false positives [18].

Seasonal variation in anomaly patterns is rarely examined, and severity stratification — distinguishing faults that need immediate action from those that can be logged — is almost entirely absent from the literature [19]. This work addresses all three gaps through a consensus-based four-algorithm framework with explicit seasonal analysis and a three-tier severity classification system.

1.3 Main Contributions

This work makes four contributions to building energy anomaly detection:

Consensus-based ensemble framework: A four-algorithm voting system (Isolation Forest, LOF, One-Class SVM, Autoencoder) requiring agreement from at least three detectors before flagging an event. This cuts the flagged event rate from 10.06% (single-algorithm) to 1.34% (consensus) — a 13.3× reduction — while retaining anomalies with mean deviations 1.4–3.1× larger than single-algorithm populations.

Systematic algorithm comparison: The first head-to-head pairwise comparison of all four detection families on the same building energy dataset. Jaccard similarity scores of 0.061–0.237 show the algorithms catch largely different anomaly populations despite identical 5% contamination settings, confirming the value of combining them.

Seasonal heterogeneity analysis: Winter anomaly rates run 3.3× higher than spring and autumn (2.73% vs. 0.82%), with cold-weather deviations averaging +2,850 kW (+100%) versus +427 to +743 kW (+18–35%) in warm seasons. Kruskal-Wallis testing confirms these differences are statistically significant ($H = 533.60$, $p < 0.001$).

Severity-stratified alerting framework: A three-tier classification — high severity (>1,000 kW, immediate response), medium (500–1,000 kW, same-day), low (<500 kW, routine logging) — with 48.7% of consensus events falling in the high-severity tier. Temporal clustering analysis shows 41.0% of anomalies occur within ±2 hours of another event, supporting event-based rather than hour-based alerting.

II. Methodology

A. System Overview and Data Preparation

The framework operates on 8,760 hours of electricity consumption data from GECAD Research Center in Porto, Portugal — a full year (January-December 2023), at hourly resolution. It contains the temporal features of PV generation and derived prices. The feature engineering yields a 14-dimensional input vector per observation: temporal attributes (hour, day of week, month, whether it is a weekend or not, and season), PV rolling means (12h & 24h), PV lags (6h, 12h & 24 h), consumption lags(12 h & 24 h), the ratio of PV to-consumption and price-weighted consumption. Distance-based algorithms LOF, One-Class SVM variants take as StandardScaler-normalised features; Isolation Forest takes raw because of scale-invariant partitioning. Interpolation and temporal consistency during the complete annual cycle are implemented with forward filling of missing values.

B. Four-Algorithm Detection Framework

Four unsupervised algorithms with complementary detection philosophies are deployed simultaneously:

Isolation Forest isolates anomalies through recursive binary partitioning — anomalous observations in sparse regions require fewer splits than normal data. Configuration: 100 estimators, subsample size 256, contamination $\alpha = 0.05$.

Local Outlier Factor scores each observation by comparing its local neighbourhood density to its k nearest neighbours. Observations in locally sparse regions — unusual for their time of day and operational context — receive high LOF scores. Configuration: $k = 20$ neighbours, Euclidean distance, $\alpha = 0.05$.

One-Class SVM learns a kernel-based boundary around normal operations in high-dimensional feature space using an RBF kernel, flagging observations that fall outside the learned region. Configuration: γ = auto, $\nu = 0.05$.

Autoencoder learns a compressed 4-dimensional representation of normal operational patterns (architecture: $14 \rightarrow 8 \rightarrow 4 \rightarrow 8 \rightarrow 14$, ReLU activations). Observations that reconstruct poorly — suggesting they do not fit the learned normal correlation structure — are flagged where MSE exceeds the 95th percentile threshold. Trained for 50 epochs, batch size 32, Adam optimiser.

C. Consensus Voting and Severity Classification

Consensus counting is how individual algorithm predictions aggregate each observation gets a score from 0 to 4, based on the number of algorithms flagging it. High-confidence anomalies need at least three algorithms to agree (τ _consensus $\geqslant 3$), medium confidence exactly two, and low confidence is a single-algorithm flag. This stratification reduces false positives by demanding that a multi-perspectival concurrence be met before any incident is considered operationally significant.

Pairwise Jaccard similarity $J = |A \cap B| / |A \cup B|$ quantifies overlap between each pair of algorithm detection sets — values near 0 indicate complementary detection, values near 1 indicate redundancy.

Detected anomalies are characterised by deviation from a 24-hour rolling mean baseline: ΔP_i = Pactual − $\bar{P}$baseline,24h. Three severity tiers are defined: high ($|\Delta P_i| >$ 1,000 kW — immediate response), medium (500–1,000 kW — same-day investigation), low (<500 kW — routine logging). Seasonal analysis partitions the year by meteorological season and computes anomaly rates, deviation magnitudes, and coefficient of variation per season. Statistical significance of inter-seasonal differences is assessed using the Kruskal-Wallis H-test.

III. Mathematical Modeling

3.1 Feature Representation

Each hourly observation i is represented as a 14-dimensional feature vector combining temporal indicators, PV generation statistics, consumption lags, and two interaction features:

$$\mathbf{x}_i = \left[h_i, d_i, m_i, w_i, s_i, \bar{P}^{(i)}_{\mathrm{PV},12h}, \bar{P}^{(i)}_{\mathrm{PV},24h}, P_{\mathrm{PV},i-6}, P_{\mathrm{PV},i-12}, P_{\mathrm{PV},i-24}, P_{\mathrm{load},i-12}, P_{\mathrm{load},i-24}, R_i, C_i\right]^\top \in \mathbb{R}^{14} \quad (6)$$

where h_i is hour of day, d_i day of week, m_i month, w_i weekend indicator, s_i season, $\bar{P}_{pv}$ rolling PV means over 12h and 24h windows, $P_{pv,i-k}$ PV lags at $k \in \{6,12,24\}$ hours, $P_{load,i-k}$ consumption lags at $k \in \{12,24\}$ hours, R_i the PV-to-consumption ratio, and C_i the price-weighted consumption.

The PV-to-consumption ratio is computed as:

$$R_i = \frac{P_{\mathrm{PV},i}}{P_{\mathrm{load},i} + \epsilon} \quad (2)$$

where $\epsilon = 10^{-6}$ pre

vents division by zero during nighttime periods with zero consumption.

3.2 Isolation Forest

Isolation Forest scores each observation by how quickly it is isolated through random recursive partitioning — anomalies sitting in sparse regions require fewer splits:

$$s_{\mathrm{IF}}(\mathbf{x}_i) = 2^{-\frac{\mathbb{E}[h(\mathbf{x}_i, T_k)]}{c(\psi)}} \quad (3)$$

where $E[h(x_i,T_k)]$ is the mean path length across 100 trees and $c(\psi)$ is the expected path length for a dataset of size ψ = 256. Scores above 0.5 indicate anomalies; values near 1.0 indicate extreme isolation. Contamination parameter $\alpha_i F$ = 0.05.

3.4 Local Outlier Factor

LOF scores each point by comparing its local density to that of its k nearest neighbors:

$$\mathrm{LOF}_k(\mathbf{x}_i) = \frac{\frac{1}{k}\sum_{\mathbf{x}_j \in N_k(\mathbf{x}_i)} \mathrm{lrd}_k(\mathbf{x}_j)}{\mathrm{lrd}_k(\mathbf{x}_i)} \quad (4)$$

where $lrd_k(x_i)$ is the local reachability density and $N_k(x_i)$ is the k-nearest neighbor set. Scores substantially above 1.0 indicate anomalies in locally sparse regions. Implementation uses k = 20 neighbors and αLOF = 0.05.

3.5 One-Class Support Vector Machine Boundary Optimization

One-Class SVM learns a decision boundary enclosing normal operations in transformed feature space through kernel-based optimization [6]. The radial basis function (RBF) kernel maps standardized features to infinite-dimensional space:

$$K(\tilde{\mathbf{x}}_i, \tilde{\mathbf{x}}_j) = \exp\left(-\gamma \parallel \tilde{\mathbf{x}}_i - \tilde{\mathbf{x}}_j \parallel_2^2\right) \quad (5)$$

where $\gamma = \frac{1}{d_{\mathrm{features}} \cdot \mathrm{Var}(\tilde{\mathbf{x}})}$ with $d_{\mathrm{features}} = 14$ represents automatic gamma selection. The optimization formulates as:

$$\min_{\mathbf{w},\rho,\boldsymbol{\xi}} \frac{1}{2} \parallel \mathbf{w} \parallel^2 - \rho + \frac{1}{\nu N_{\mathrm{total}}} \sum_{i=1}^{N_{\mathrm{total}}} \xi_i \quad (6)$$

subject to:

$$\mathbf{w}^\top \phi(\tilde{\mathbf{x}}_i) \geq \rho - \xi_i, \quad \xi_i \geq 0, \quad \forall i \quad (7)$$

where $\nu = 0.05$ bounds the outlier fraction and γ is set automatically. Observations with negative decision function values fall outside the learned boundary and are classified as anomalies.

3.6 Autoencoder Reconstruction Error Analysis

The autoencoder (architecture 14→8→4→8→14, ReLU activations) learns a compressed representation of normal operations. Points that reconstruct poorly are flagged as anomalies based on mean squared reconstruction error:

$$\mathrm{MSE}(\mathbf{x}_i) = \frac{1}{14}\sum_{j=1}^{14}\left(\tilde{x}_{ij} - \hat{x}_{ij}\right)^2 \quad (8)$$

3.7 Consensus-Based Classification

Each observation receives a consensus count from the four detectors:

$$C_i = \mathbb{1}_{\mathrm{IF}}(\mathbf{x}_i) + \mathbb{1}_{\mathrm{LOF}}(\mathbf{x}_i) + \mathbb{1}_{\mathrm{OCSVM}}(\mathbf{x}_i) + \mathbb{1}_{\mathrm{AE}}(\mathbf{x}_i) \quad (9)$$

Confidence is stratified as: High ($C_i \geq 3$), Medium ($C_i = 2$), Low ($C_i = 1$), Normal ($C_i = 0$). Only high-confidence events — flagged by at least three algorithms — are treated as operationally significant anomalies.

3.8 Algorithm Agreement Quantification

Pairwise overlap between detectors A and B is measured by Jaccard similarity:

$$J(A,B) = |A \cap B| / |A \cup B| \quad (10)$$

Values near 0 indicate complementary detection (different anomaly populations); values near 1 indicate redundancy.

3.9 Deviation Magnitude and Severity Classification

ANOMALY SEVERITY IS BASED ON DEVIATION FROM THE 24-HOUR ROLLING MEAN BASELINE $\bar{P}_{BASELINE}$,

$$\Delta P_i = P_{LOAD,i} - \bar{P}_{BASELINE,i} \quad (11)$$

Three tiers are defined: High ($|\Delta P_i| > 1{,}000$ kW — immediate response), Medium (500–1,000 kW — same-day investigation), Low (<500 kW — routine logging).

IV. RESULTS AND DISCUSSIONS

4.1 Algorithm-Specific Detection Characteristics

A contamination fraction of 5% was chosen a priori and all four classification algorithms were applied to the full 8760-hour data set. Detection Counts detected almost exactly 438 events as expected: Isolation Forest, LOF and Autoencoder all flagged precisely 438 (5.00%) from the control set of data with One-Class SVM flagging slightly more at 444 (5.07%), this is a small difference attributable to its soft-margin formulation where some outliers can be included in-region but close enough that their score approaches one protecting against falsely negative classifications.) Detection rate and count across all four methods (Fig 1a & d).

The pretty simple reason is that what differs enormously between the algorithms is not their performance in finding anomalous events but rather identifying which anomalies are flagged. The mean deviation from the average algorithm-detected set for normal consumption of an AC operator and spread is shown in figure 1(b). The most mild events were caught by Isolation Forest (+20.2% over normal mean), while LOF stepped up the game to +30.6%. The Autoencoder and One-Class SVM captured far more serious occurrences as +43.1% and +45.8%, respectively. These differences in absolute consumption means are depicted by Fig 1(c)—average outliers detected via One-Class SVM measured around 3,583 kW vs normal mean of \textcolor{blue}{2,457} kW and Isolation Forest anomalies averaged at just under these levels with an average of 2.991 kW.

This hierarchy of deviation reflects a purely different view on what each algorithm considers an anomaly. Since Isolation Forest works by global partitioning, it is sensitive to anything that varies slightly from the mass of mostly-typical operations — including borderline operational variations. On the other hand, LOF uses local density comparisons that detect what is unusual only in their immediate temporal neighbourhood and relaxed conditions under which global consumption appears regular. Both One-Class SVM and the Autoencoder concentrate on extreme events — points outside a learned frontier in feature space (the top one), or patterns which reconstruct poorly from the normal representation I learnt about some random variable x. It is remarkable that these two perform so similarly (+45.8% vs +43.1%), which suggests both methods are interpreting the same population of extreme anomalies, albeit via different mathematical pathways.

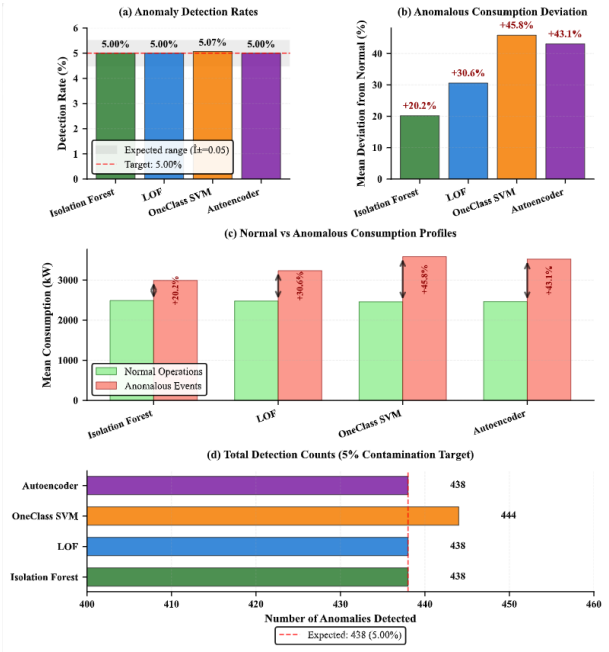

Fig. 1. Algorithm detection performance. (a) Detection rates near 5% target. (b) Mean deviation hierarchy: Isolation Forest +20.2% to One-Class SVM +45.8%. (c) Normal vs anomalous consumption profiles. (d) Total detection counts against expected 438-event baseline.

4.2 Consensus Analysis and Detection Confidence

The complete confidence stratification pyramid is represented in Fig. 2(a). Of 8,760 total hours, 7,509 (85.72%) were deemed normal. The remaining hours break down as: Low Confidence — 881 events (10.06%) flagged by one algorithm only; Medium confidence: 253 events (2.89%!) flagged by two algorithms) and High confidence, which is in fact the most reliable category, where either three or more flags had been applied to an event over any other flagging mechanism per hour time frame we care about at least.) The reduction from individual-algorithm detection (10.06%) to consensus detection (1.34%), which reflects a complete reliance on multi-algorithm agreement, yields our best-performing scenario: only an 8% of flagged events were exactly the same and detected by at least two algorithms.

This is made clear in Figure 2(b) with the Jaccard similarity heatmap. Pairwise overlap between algorithms is low across the board from J=0.061 (Isolation Forest vs LOF, most complementary pair) to J = 0.237 (Isolation Forest vs One-Class SVM, from similar pair) This number is J = 0.115 for LOF vs Autoencoder, and J = 0.202 One-Class SVM (OCSVM) autoenoder — in the figure above-.) These low values are an indication that the four algorithms in fact pick up alternative subsets of the anomaly space rather than triggering on same events. The least surprising overlap is between Isolation Forest and One-Class SVM, since both have global boundary definitions (unlike LOF which has a local density approach).

Indeed, the 117 high-confidence events surviving consensus filtering have mean deviation magnitudes that are orders of magnitude larger than any single-algorithm

population — average deviation is 1543 kW compared to a range of deviations from individual detectors: (502–1,126 kW). This 1.4–3.1× amplification is consistent with the notion that consensus filtering does not randomly discard events — rather, it systematically preserves the most operationally-significant anomalies while rejecting borderline cases for which only one detector considers unusual.

As a 7-day rolling sum of detections, this is shown in Fig. 2(c) for how the number changes through out time up to and including 2023. There is a readily visible clustering of high-confidence anomalies in the winter months (January–February), some less aggregating but still optimistic activity through spring and summer, and moderate anomalous behavior observed returning autumn-ward. This seasonal pattern is the basis for our deeper analysis in section 4.3.

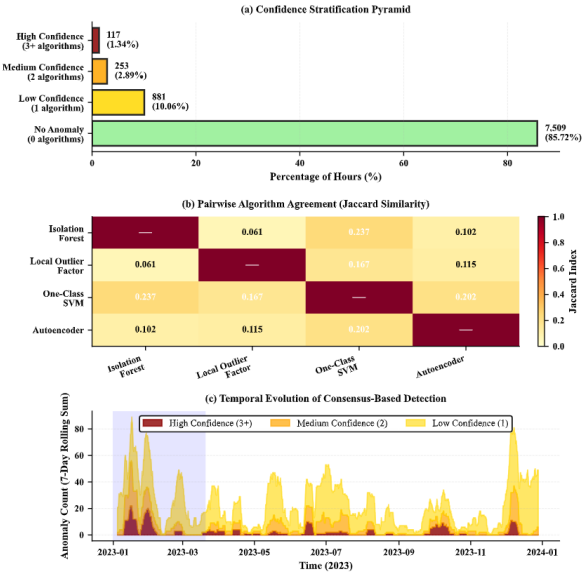

Fig. 2. Consensus confidence stratification. (a) Detection pyramid: 10.06% individual-algorithm flagging reduced to 1.34% high-confidence consensus. (b) Jaccard similarity heatmap (0.061–0.237) confirming complementary detection. (c) Temporal evolution showing winter concentration of high-confidence events.

4.3 Seasonal Heterogeneity in Anomaly Patterns

The seasonal contrast is instantly apparent from Figure 3(a). In fact, the anomaly rate is 2.73% in winter — that's 3.3× higher than spring and autumn (both at80%) and nearly 2.7× summer's performance (1·00%). The difference is statistically significant (Kruskal-Wallis H = 533.60, p <0.001) meaning that it cannot be random variation Out of the 117 high-confidence anomalies, half (59) occurred in winter: a continent using one season made up 50.4% of all events.

Not only do winter anomalies occur more frequently, but they are also larger — which is what makes them particularly serious (see Figure 3(b)). Anomalies in winter were a mean of +2,850 kW (+100%) off normal consumption across all flagged hours (in effect doubling expected demand). Autumn, with its far lower (0.82%) occurrence rate was a close second at +2,603 kW (+99%). This autumn finding is significant — though there are considerably more anomalies in winter than any other season, those that do occur during the autumn months can reach a magnitude nearly as severe. Spring and summer anomalies were much lower, with average monthly values of +427 kW (+18%) for spring and +743k W(+34) in the summer.

Consumption is represented in box plots for perfectly normal versus anomalous hours, subject-wise and season-per-season validity as shown visually via indiscriminate (Fig. 3(c)). The outlying boxes are placed much higher than the other operations across all seasons, but with a larger gap in winter and autumn — which is also consistent with these observations quantified by magnitude analysis (panel (c)). In the case of Fig. 3(d), CV is incorporated as an additional dimension. The consumption in normal winter months is already highly variable (CV = 0.60), while the anomalous consumption shows a similar variability (CV = 0.61) suggesting that there are also very noisy background and hence, static threshold fails for this season partly because of blending these two conditions together during cold seasons as well. Despite their moderate mean magnitude, summer anomalies exhibit the highest CV of any group (0.76 vs 0.35 normal), reflecting their broad character — a jumbled mix that implies several fault types rather than a single dominant cause.

The hour-of-day by month heatmap of anomaly rates is illustrated in Figure 3(e). The average hourly January and February daytime heating system interactions are seen to closely reflect occupancy-driven loads with winter anomalies concentrated across a full working day (roughly 7:00–19:00). Since summer anomalies are expected to be more tightly clustered around the midday hours (10:00 – 14:00) when cooling systems would at most risk of being overloaded by peak solar and occupancy conditions, Taken together, these temporal patterns indicate that sensitivity testing should contain its highest nascent level during daytime hours year-round but be further refined based on all working-hours in winter and peak daily observation periods within summer.

The mean plot from figure 3 builds a solid argument for distinct season-specific anomaly detection thresholds. One annual threshold that functions well in summer is too draconian (too many false positives) for winter and too lax in autumn (mising infrequent but extreme events).

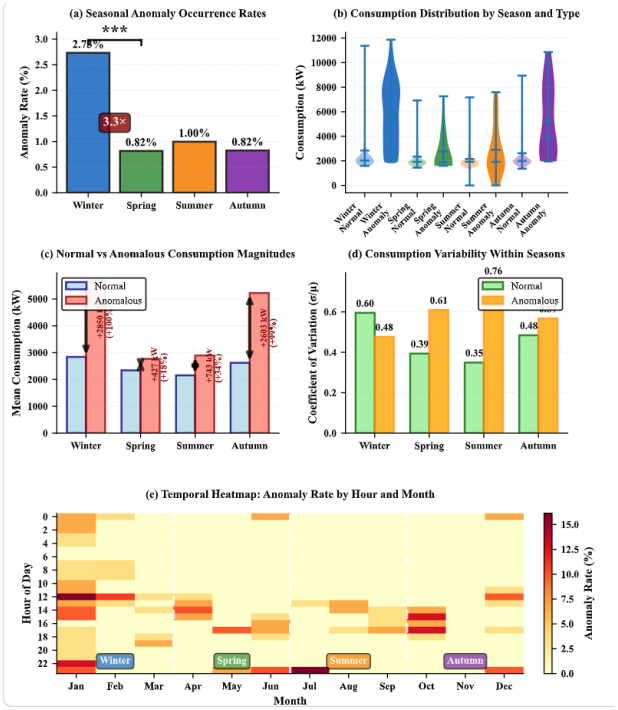

Fig. 3. Seasonal anomaly analysis. (a) Winter anomaly rate 3.3× higher than spring/autumn ($p < 0.001$). (b) Normal vs anomalous consumption distributions per season. (c) Deviation magnitudes: winter +2,850 kW (+100%), autumn +2,603 kW (+99%), summer +743 kW (+34%), spring +427 kW (+18%). (d) Anomaly CV by season. (e) Hour-of-day heatmap showing daytime winter and midday summer clustering.

4.4 Severity Classification and Operational Response

Of the 117 high-confidence anomalies detected, more than half (48.7%, or 57 events) were classified into the highest severity tier (>1,000 kW deviation), requiring immediate operational response. An additional 28.2% (33 events) were medium severity (<500–1,000 kW deviation, same-day investigation; moderate impact on the grid), while the remaining 23.1% (27 events) low severity (<500 kW routine wrapping). The fact that almost half of all consensus-detected events are in the highest severity tier is a proxy for what happens when you aggregate anomaly scores; by the time three or four algorithms have singled out an anomalous event, it must be a true signal.

A temporal clustering analysis revealed that, of these anomalies, 41.0% occurred within ±2 hours of a different flagged event (Fig 6). This persistence behaviour has a practical operational effect, in that an hour-by-hour alerting system would produce numerous alert notifications for what essentially is a single prolonged fault incident. Having a protocol for alerting based on events — clustering together temporally proximal anomalies into an incident — would significantly cut down fatigue while maintaining a full coverage of detections.

For example, the entire pipeline — training four algorithms (i.e. when we say “training” that entails feature engineering followed by consensus voting) and severity classification on all 8,760 observations — completed in 29.1 seconds on standard CPU hardware. We can have confidence in this framework because any batch processing mode would require much less than 1 second per new hourly observation to be real-time-ready.

V. Conclusion

This article demonstrated that a consensus-based ensemble of four unsupervised ML algorithms can accurately detect operationally significant energy anomalies in the research facility environment and without any labelled fault data. Flagging an event only after at least three algorithms agreed, reduced the detection rate from 10.06% (single-algorithm) to 1.34%, but left mean deviations for surviving anomalies to be on average between 1.4 and up to 3× larger than single-algorithm populations This corroborates that consensus filtering filters noise, and not signal.

Inseason analysis left no doubt that this pattern was significant and, indeed practical: with winter anomaly rates running 3.3× larger than autumn/spring (2.73% versus 0.82%) — based on seasonal events +2850 kW (+100%), makes perfect sense as a generic party of cold weather consumption above expectation! Autumn anomalies, although quite rare, also reached similar magnitudes (+2,603 kW +99% but a missed by the aggregate annual analysis). In terms of severity, the three-tier framework classified 48.7% of consensus events as high-severity and temporal clustering indicated that about 41% occur within ±2 hours from another — suggesting an event-based rather than hour-level alerting in operational design is more applicable here instead This indicates that real-time deployment is possible as the full pipeline ran for 29.1 seconds on standard hardware.

Future work should use real-life equipment failure logs to ground-truth anomalies, in addition to adaptive seasonal thresholds that automatically adjust the sensitivity of detections depending on time varying operational trends over the course of a year.

.

References

[1] Shahidi, Seyedehelaheh, Masoud Hajialilue-Bonab, Hamid Reza Tohidvand, and Ali Khosravi. "Experimental investigation on the efficiency of the phase change materials for enhancing the thermal performance of energy piles in sandy soils." Energy and Buildings 298 (2023): 113544.

[2] S. Kurapati, R. T. Thutari, A. Chaudhary, S. Muchakhandi and S. S. R, "Enhancing Software Security Through Automated Sdlc Integration and Model-Based Testing," 2025 2nd International Conference on Software, Systems and Information Technology (SSITCON), Tumkur, India, 2025, pp. 1-5, doi: 10.1109/SSITCON66133.2025.11342172.

[3] S. B. Venkata, A. Alzubaidy, S. Srinivasan, S. K. M and B. Vemala, "Emotion Recognition in Code-Mixed Social Media Data Using an Optimized Transformer Architecture," 2025 2nd International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS), Hassan, India, 2025, pp. 1-6, doi: 10.1109/IACIS65746.2025.11211430.

[4] Kumar, Polamarasetty P., Ramakrishna SS Nuvvula, Sk A. Shezan, Vanam Satyanarayana, Syed Riyaz Ahammed, R. SivaSubramanyamReddy, and Ahmed Ali. "Optimal sizing and placement of renewable energy systems in smart grids using machine learning optimization." In 2024 12th International

Conference on Smart Grid (icSmartGrid), pp. 312-316. IEEE, 2024

[5] T. Joshi, K. Thiyagarajan, Srinivasan, A. Alzaidi and P. Kodgi, "Longformer-Enhanced Hierarchical Attention for High-Quality Extractive Document Summarization," 2025 2nd International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS), Hassan, India, 2025, pp. 1-7, doi: 10.1109/IACIS65746.2025.11210974.

[6] P. P. Kumar, V. E. Jyothi, S. Manne, M. Chandran and K. Musku, "File Vault Secure: Multi-Factor Encryption Using Password, Geolocation, and One-Time Password (OTP)," 2026 IEEE International Conference on AI Engineering and Innovations (AIEI), NIT Jamshedpur, India, 2026, pp. 1-6, doi: 10.1109/AIEI69164.2026.11497340.

[7] Nuvvula, Ramakrishna SS, Polamarasetty P. Kumar, Alighazi Siddiqui, S. Thamizharasan, Chai Ching Tan, Raaid Alubady, and Baseem Khan. "Optimizing electric vehicle fleet operations with predictive analytics: A renewable energy-centric approach." In 2024 12th international conference on smart grid (icSmartGrid), pp. 309-311. IEEE, 2024.

[8] Madushika, U. G. D., and Weisheng Lu. "Green retrofitting application in developing economies: State of the art and future research directions." Energy and Buildings 301 (2023): 113712.

[9] Senthamaraikannan, Sundaram, and Shanmugam, 'Bio-efficacy of Polyethylene Glycol-coated Selenium Nanocomposite Synthesised using Cocos nucifera Haustorium against HepG2 Cell Line', Int. Res. J. multidiscip. Technovation, pp. 165–183, Mar. 2025.

[10] Cao, Bowen, Yonggao Yin, Guoying Xu, Xiaosong Cheng, Wenzhang Li, Qiang Ji, and Wanhe Chen. "A proposed method of bubble absorption-based deep dehumidification using the ionic liquid for low-humidity industrial environments with experimental performance." Applied Energy 348 (2023): 121534.

[11] Nuvvula, Ramakrishna SS, Polamarasetty P. Kumar, Syed Riyaz Ahammed, R. Siva Subramanyam Reddy, Bachina Harish Babu, and Ahmed Ali. "Distributed Reinforcement Learning Framework for Autonomous Optimization of Grid-Scale Energy Storage Systems in Renewable Energy Integration." In 2024 12th International Conference on Smart Grid (icSmartGrid), pp. 290-297. IEEE, 2024.

[12] Sundaram and S. Balamurugan, 'Enhanced Thermo-physical properties of titanium dioxide nanofluids: Applications in renewable energy systems and smart grid cooling technologies', in 2025 International Conference on Electronics and Renewable Systems (ICEARS), Tuticorin, India, 2025, pp. 1–5.

[13] S. B. Venkata, A. Ashirova, V. S. Karwande, A. T M, E. Kholiyarov and S. Singh, "Towards Accurate Maritime Surveillance: A Hybrid CNN-Transformer Architecture for Ship Detection in SAR Imagery," 2025 International Conference on Innovations in Intelligent Systems: Advancements in Computing, Communication, and Cybersecurity (ISAC3), Bhubaneswar, India, 2025, pp. 1-6, doi: 10.1109/ISAC364032.2025.11156808.

[14] A. V. S. Madineni and U. Mopathi, "Estimation of Theme Re-Hotting with a Temporal Design in Online Media," International Journal of Research in Electronics and Computer Engineering (IJRECE), vol. 6, no. 4, pp. 597–603, 2018.

[15] S. Pathivada, S. Pallipe, V. S. Sahithi, A. Rayala and D. Paturu, "Integrating Satellite Imagery and IoT Field-Based Observations for In-Situ Soil Moisture Analysis," 2026 IEEE International Conference on AI Engineering and Innovations (AIEI), NIT Jamshedpur, India, 2026, pp. 1-6, doi: 10.1109/AIEI69164.2026.11495575.

[16] R. Kasturi and R. Nellikondi, "Managing Dimensional History in Data Warehouses: A Comparative Analysis of Hash-Based and Surrogate Key Approaches," Advances in Consumer Research, vol. 1, no. 1, pp. 28–35, 2024.

[17] Mr. Raghavendar Nellikondi, Mr. Reddaiah Kasturi, "Benchmarking Cloud-Native ETL Tools Performance, Cost, and Scalability", Journal of Computational Analysis and Applications (JoCAAA), vol. 33, no. 08, pp. 7433–7447, Dec. 2024.

[18] B. S. Guttikonda, R. C. Sachan and V. Veeramachaneni, "LLM-GA: A Hybrid Framework to Build Dynamic Websites for Optimizing Web Performance," International Journal of Information Technology, 2025. doi: 10.1007/s41870-025-02692-w.

[19] Kumar, Polamarasetty P., Akhlaqur Rahman, Ramakrishna SS Nuvvula, Ilhami Colak, S. M. Muyeen, Sk A. Shezan, G. M. Shafiullah et al. "Using energy conservation-based demand-side management to optimize an off-grid integrated renewable energy system using different battery technologies." Sustainability 15, no. 13 (2023): 10137.