

Machine Learning-Based Anomaly Detection for ITER's Tokamak Systems Monitor: A Gyrotron Case Study

Joris Paret¹, Daniel Iglesias, Daniel Sabio Ruiz, Dilin Meloni², Andrea Antonione, Ruggero Bertazzoni³, Giuseppe Carannante⁴, Martino Ferrari, María Ortiz de Zúñiga⁵, Mario Cavinato, Francisco Sánchez Arcos, and Alfredo Portone⁶

Abstract—The Tokamak Systems Monitor (TSM) software will provide ITER operators with timely assessments of machine health, component lifetime, and early warnings of potential faults. Among its key functions, the TSM will employ data-driven anomaly detection methods to identify unexpected or abnormal behavior across a broad range of systems and diagnostics. This work presents an initial proof-of-concept for anomaly detection applied to gyrotron pulses, leveraging data from the European gyrotron prototype. The method combines dimensionality reduction and clustering techniques to identify deviations from expected operational patterns. This approach enables the detection of subtle anomalies that might otherwise go unnoticed. While this first demonstration focuses on intershot anomaly detection for gyrotrons, the methodology is designed to be adaptable to other systems and signals within the TSM, contributing to the improved maintenance strategies and reliability of the ITER tokamak.

Index Terms—Anomaly detection, diagnostics, gyrotron, ITER, machine learning (ML), monitoring.

I. INTRODUCTION

IN COMPLEX physical systems such as tokamaks, ensuring reliable and safe operation requires the ability to detect subtle signs that something may be going wrong, often before any damage occurs or performance is degraded. This challenge is compounded by the fact that unusual behaviors in sensor

data can take many different forms, and may not always trigger alarms or fall outside predefined operational thresholds.

To address this, anomaly detection techniques are used to identify patterns or events in data that deviate from expected or nominal behavior. Such anomalies may include increased noise, gradual drifts, sudden spikes, persistent offsets, or more complex issues such as the loss of correlation between multiple signals in a multivariate time series. Because of this diversity, machine learning (ML)-based approaches have emerged as powerful tools for anomaly detection [1]. They can automatically learn flexible representations of normal behavior and flag deviations without requiring rigid, manually defined rules. This is especially useful in high-dimensional settings, where complex and often nonlinear interactions between variables may render traditional threshold-based monitoring ineffective.

During tokamak operation, the timely detection of anomalies—whether in real time or intershot—is crucial for predictive maintenance and performance optimization. A system or subsystem may continue operating within its operational limits while already exhibiting early signs of failure, degradation, or suboptimal performance, which may not be immediately visible from the raw sensor data. As the number of monitored sensors increases into the thousands, as is expected for ITER, it becomes infeasible for operators to manually track every individual signal. Consequently, automated anomaly detection methods are essential for identifying abnormal behaviors in a scalable and reliable manner. While there is a growing number of studies applying such techniques to plasma behavior, particularly for disruption prediction [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], only a few have focused on monitoring the machine itself or its components. Among these, ML-based anomaly detection methods have been successfully tested or implemented in other fusion devices such as JET [12], DIII-D [13], and WEST [14], [15], [16], demonstrating their practical utility in this emerging area.

The Tokamak Systems Monitor (TSM) software suite [17] will collect and analyze sensor data across various subsystems of ITER. It includes a module for applying ML algorithms to detect anomalies in the instrumentation data. Two classes of anomaly detection algorithms are being developed within the TSM: 1) intershot algorithms that analyze entire pulses postmortem to determine whether they deviate from nominal

Received 31 July 2025; revised 21 November 2025; accepted 23 February 2026. Date of publication 8 May 2026; date of current version 19 June 2026. This work was supported by the ITER Organization. The review of this article was arranged by Senior Editor M. Kovari. (Corresponding author: Joris Paret.)

Joris Paret, Daniel Iglesias, Giuseppe Carannante, and Martino Ferrari are with the ITER Organization, 13067 St. Paul Lez Durance, France (e-mail: Joris.Paret@iter.org).

Daniel Sabio Ruiz is with Automaticity Solutions, 43880 Tarragona, Spain.

Dilin Meloni is with Massachusetts Institute of Technology, Cambridge, MA 02139 USA.

Andrea Antonione is with the Università degli Studi di Napoli Federico II, 80125 Naples, Italy.

Ruggero Bertazzoni is with Karlsruhe Institute of Technology (KIT), 76131 Karlsruhe, Germany.

María Ortiz de Zúñiga is with the Fusion for Energy, 08019 Barcelona, Spain, and also with the Universidad Nacional de Educación a Distancia (UNED), 28040 Madrid, Spain.

Mario Cavinato, Francisco Sánchez Arcos, and Alfredo Portone are with the Fusion for Energy, 08019 Barcelona, Spain.

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TPS.2026.3670510>.

Digital Object Identifier 10.1109/TPS.2026.3670510

conditions; and 2) time-resolved algorithms that monitor data in real time to flag anomalies as they occur, or allowing a more refined postmortem analysis.

This work focuses on the intershot anomaly detection approach, with a specific application to the thermal behavior of the gyrotron collector. We analyze temperature signals recorded from thermocouples mounted on the collector of the 1-MW European gyrotron industrial prototype. The characteristic ramp-up, plateau, and cool-down profile of these pulses is similar to that observed in other ITER subsystems, such as plasma-facing components, making the methodology applicable across other use cases covered by the TSM. Since the dataset is unlabeled, we employ unsupervised ML methods such as dimensionality reduction and clustering to identify pulses that deviate from the expected plateau-like pattern.

This article is organized as follows. Section II provides an overview of the TSM, its objectives, and its role in ITER operations. Section III describes the experimental setup and presents the dataset used in this study. Section IV details the anomaly detection pipeline, including the data preprocessing steps, feature extraction, and ML algorithm employed. Section V discusses the preliminary results obtained using unsupervised methods and highlights key findings from the analysis. Finally, Section VI offers concluding remarks and outlines directions for future developments within the TSM framework.

II. TOKAMAK SYSTEMS MONITOR

The TSM is a dedicated software suite developed as part of ITER's diagnostics to provide a comprehensive engineering assessment of the tokamak during its operational life. Its role is to monitor the mechanical, thermal, hydraulic, and electromagnetic responses of key systems inside the cryostat and auxiliary systems, both under normal operation and during transient off-normal events such as plasma disruptions, vertical displacement events, fast magnet discharges, and seismic events.

To meet the ITER Project Requirements, operational instrumentation is deployed across the machine to measure key signals related to system integrity, availability, and failure prediction. However, due to integration constraints and the limitations of sensor technologies, certain critical engineering quantities—such as forces, heat fluxes, or internal stress states—cannot be directly measured or may only be indirectly inferred from limited data. The TSM addresses this challenge by using discrete sensor data to reconstruct quantities that are not directly observable, leveraging numerical models and data-driven techniques to extract meaningful insights from available instrumentation.

At the highest level, the TSM fulfills five core functions.

- 1) Reconstruction of global engineering parameters using multiphysics numerical models across thermal, hydraulic, electromagnetic, dynamic, and structural domains.
- 2) Evaluation of safety margins and lifetime consumption based on reconstructed cyclic loading, to assess fatigue and residual life.

- 3) Detection and registration of anomalies that may not trigger interlocks but indicate degraded behavior.
- 4) Generation of an engineering database summarizing tokamak behavior, to support future fusion devices.
- 5) Calibration and validation of physics-based models using real-world data, ensuring consistency between numerical simulations and machine behavior.

This article focuses on the anomaly detection function of the TSM, which aims to identify unexpected behaviors across the tokamak's systems, either by evaluating full pulses after completion (i.e., intershot) or by analyzing data as it is generated during operation (i.e., time-resolved). Both approaches are essential: while time-resolved algorithms can raise immediate warnings during operations or a finer postmortem analysis of a pulse, intershot methods enable a broader classification and a direct comparison of multiple pulses. The present work explores the intershot anomaly detection strategy using thermal data acquired from the collector of the European 1 MW prototype gyrotron tested at the Falcon facility (Swiss Plasma Center).

Upon detecting an anomaly, the system flags it, saves it to a database, and reports it to the operators through the human-machine interface (HMI), enabling them to swiftly determine whether the anomaly is genuine and requires further investigation or if it can be dismissed. This mechanism supports expert feedback and iterative improvement or retraining of the algorithms.

III. EXPERIMENTAL SETUP AND DATASET

Gyrotrons are high-power microwave sources used in tokamaks for Electron Cyclotron Resonance Heating and Current Drive (ECRH&CD), where a relativistic electron beam interacts with a resonant cavity under a strong magnetic field to generate high-frequency RF power. This energy is extracted through an output window and guided into the plasma to resonantly heat electrons. After RF generation, the spent electron beam is directed into the collector, a critical component located at the top of the gyrotron, where its residual kinetic energy is dissipated as heat through active water cooling. See Fig. 1 for a simplified representation of the gyrotron and its main components.

The gyrotron considered in this work is an industrial prototype developed by Thales for the European ECRH&CD system in ITER. It operates at 170 GHz with a nominal output power of 1 MW in continuous wave conditions. It was tested and qualified under ITER-relevant conditions at the Falcon facility of the Swiss Plasma Center [18], [19]. The collector was instrumented with 50 thermocouples distributed across its surface to monitor temperature evolution during operation. A preliminary inspection of the dataset revealed that only $N_t = 33$ of these sensors were consistently acquired across all pulses. The remaining 17 thermocouples exhibited intermittent acquisition or permanently flat signals, for which no detailed information on their location or underlying cause was available. Because such missing or constant signals represent trivial anomalies that do not require data-driven analysis, this study focuses on the reliably recorded subset in order to investigate structural deviations in the thermal evolution of the collector.

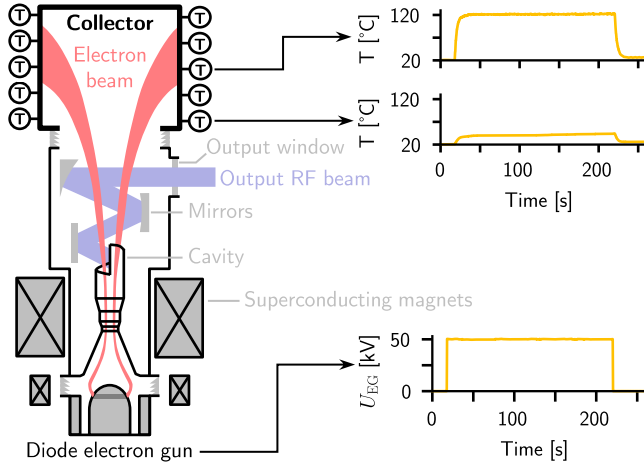


Fig. 1. Schematic view of a high-power gyrotron used for ECRH&CD in ITER. The gyrotron converts the kinetic energy of a relativistic electron beam (red beam) into RF power (blue beam), which is extracted through an output window and injected into the plasma. The remaining electron energy is deposited in the collector, where it is dissipated as heat via active cooling. An array of thermocouples (indicated by the circled T symbols) monitors the collector's surface temperature. The two plots on the top right show typical pulse data: the temperature evolution recorded by two thermocouples, one located near the center of the collector (where the heat flux is highest) and the other near the edge (where thermal loading is lower). The bottom right plot shows the applied acceleration voltage of the electron gun, U_{EG} , which follows a step-like profile consistent with the observed thermal response.

The collector exhibits a well-defined heating cycle characterized by three phases: a rapid ramp-up, a quasi-stationary thermal plateau, and a cool-down. This cycle is driven by the input voltage applied to the diode electron gun, which approximates a Heaviside step function. Anomalies may manifest as deviations from this expected shape, either globally or locally in one or more signals.

Each gyrotron pulse is represented as a set of multivariate time series, i.e., one for each of the N_t thermocouples. The signals are originally sampled at 20 Hz, then resampled at 10 Hz on a common time grid for consistency. Since thermocouples are slow-varying sensors, this reduction in sampling rate does not impact the quality of the data, and 10 Hz remains sufficient to capture all relevant thermal dynamics. The dataset comprises $N_p = 401$ such pulses acquired during the qualification campaign. Fig. 2 summarizes key descriptive statistics of the dataset in the form of boxplots. In particular, pulse durations are highly variable, ranging from 2 to 1000 s, with a strong right-skew: 75% of the pulses are shorter than 180 s, and the median duration is 60 s. The maximum collector temperature per pulse is more consistent, typically lying between 128 °C and 146 °C, with a median of 136 °C. The mean temperature across all thermocouples exhibits a similar spread, with most values between 59 °C and 70 °C, and a median of 65 °C. Finally, the temperature range (i.e., the difference between the hottest and coolest thermocouples during a pulse) varies from 65 °C to 158 °C, with a median of 121 °C.

While this work focuses on the gyrotron collector, its thermal dynamics are representative of other ITER systems with similar operational cycles, such as the diagnostics first

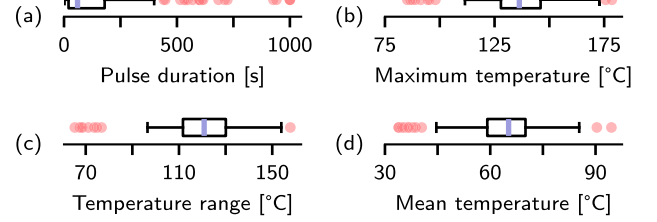


Fig. 2. Summary statistics of the gyrotron pulse dataset. Boxplots show the distribution across the 401 pulses of (a) pulse duration, (b) maximum temperature, (c) temperature range (maximum minus minimum across thermocouples), and (d) mean temperature (averaged over all thermocouples). The box extends from the first quartile (Q1) to the third quartile (Q3) of the data, with a blue vertical line at the median. The whiskers extend to the data points within 1.5 times the interquartile range (IQR = Q3 – Q1) from the box. Flier points beyond this range are shown individually with semi-transparent red markers.

wall or the blanket modules. Therefore, the anomaly detection approach developed here is applicable within the TSM framework, beyond the scope of this particular subsystem.

IV. METHODOLOGY AND WORKFLOW

In order to compare the N_p gyrotron pulses despite their varying durations, each multivariate time series must be transformed into a fixed-length feature vector $\{\mathbf{x}_n\}_{n=1}^{N_p}$. For each pulse $n \in \{1, \dots, N_p\}$ and each thermocouple signal $T_n^k(t)$ (with $k \in \{1, \dots, N_t\}$), we extract $n_\theta = 31$ time-independent statistical features $\{\theta_i^k\}_{i=1}^{n_\theta}$ using the TSFEL Python package [20]. These features include mean, standard deviation, skewness, entropy, slope, and many more.

The N_t sets of features of a pulse n are concatenated to form a single feature vector $\mathbf{x}_n$ of size $M = N_t \times n_\theta$

$$\mathbf{x}_n = (\theta_1^1, \dots, \theta_{n_\theta}^1, \dots, \theta_1^{N_t}, \dots, \theta_{n_\theta}^{N_t}). \quad (1)$$

These vectors are stacked to form the dataset matrix $\mathcal{D} \in \mathbb{R}^{N_p \times M}$

$$\mathcal{D} = \begin{pmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_{N_p} \end{pmatrix}. \quad (2)$$

All features in $\mathcal{D}$ are normalized to zero mean and unit variance. Because the feature space is high-dimensional ($M = 1023$), we apply two preprocessing steps to reduce redundancy among features and improve the tractability of the anomaly detection pipeline. First, we remove features that are strongly correlated with pulse duration, to avoid biasing the anomaly detection toward pulselength rather than waveform shape. Specifically, any feature whose Spearman correlation coefficient with pulse duration exceeds 0.5 across the dataset is discarded. This reduces the dimension of the dataset from $M = 1023$ to $M = 184$.

Second, we apply principal component analysis (PCA) to further reduce dimensionality while preserving the variance structure of the data. PCA identifies orthogonal directions—Principal Components (PCs)—along which the variance is maximized, projecting the data into a lower-dimensional subspace of $m \ll M$ dimensions. The number of retained

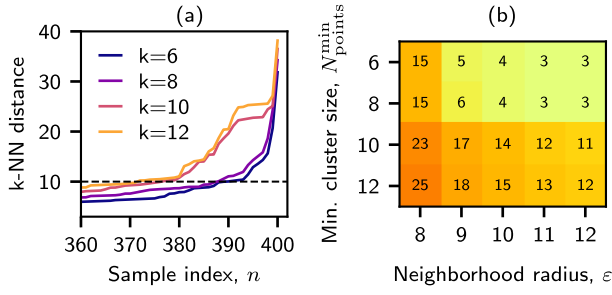


Fig. 3. DBSCAN parameter tuning. (a) k -distance graphs for multiple values of k , where the k -NN distance is the distance from each sample to its k th nearest neighbor in the 4-D PCA space (PC1–PC4). Curves show sorted k -distances over all samples, and the “elbow” region, used to guide the choice of the neighborhood radius ϵ , consistently appears near $\epsilon \approx 10$ for all k considered (horizontal dashed line). For clarity, the plot focuses on the index range where this onset is most visible (i.e., $n \in [360, 400]$). (b) Number of outliers detected by DBSCAN as a function of the neighborhood radius ϵ and the minimum cluster size, $N_{\text{points}}^{\min}$. The transition in outlier count between $N_{\text{points}}^{\min} = 8$ and 10 reflects the merger of two clusters into one under the more conservative setting.

components is typically chosen to preserve 80%–90% of the total variance. Throughout, we report each principal component’s explained variance ratio (EVR) to indicate its contribution to total variance. In our case, the feature space is effectively reduced to a 4-D space ($m = 4$), capturing the dominant structure of the dataset while keeping 82% of the original variance. This strong compressibility is due to high correlations among thermocouple signals located close to each other, which leads to redundant feature sets across sensors (i.e., $\theta_i^k \approx \theta_{i'}^{k'}$).

Anomaly detection is then performed on this reduced representation using the DBSCAN clustering algorithm [21]. DBSCAN identifies clusters as regions of high density in the feature space, distinguishing between core, border, and outlier samples. Samples (i.e., pulses) that do not belong to any cluster are labeled as outliers by the algorithm, and are interpreted as candidate anomalies pending further validation. The algorithm requires two parameters: a neighborhood radius ϵ and a minimum number of points $N_{\text{points}}^{\min}$ to define a dense region (i.e., a cluster). The value of $N_{\text{points}}^{\min}$ is typically set based on the expected minimum cluster size and the dimensionality of the data. A common heuristic is to set it as approximately twice the number of features [22]. Here, DBSCAN is applied on four features and $N_{\text{points}}^{\min}$ was set to ten after trying values in the range [6, 12]. Given this choice, ϵ is selected using a k -distance graph, where for each sample the distance to its k th nearest neighbor in the 4-D PCA space is computed, and all such distances are plotted in ascending order. The optimal ϵ corresponds to the “elbow” of the curve, indicating a sharp transition between dense clusters and sparse noise, as shown in Fig. 3(a). Based on this graph, a value of $\epsilon = 10$ was selected for this work. To assess the sensitivity of the outlier detection to parameter choices, the number of outliers was computed over a grid of ϵ and $N_{\text{points}}^{\min}$ values, as shown in Fig. 3(b). In particular, a noticeable change in outlier count occurs between $N_{\text{points}}^{\min} = 8$ and 10, which is due to a transition from two clusters to one. This behavior will be discussed further in the next section.

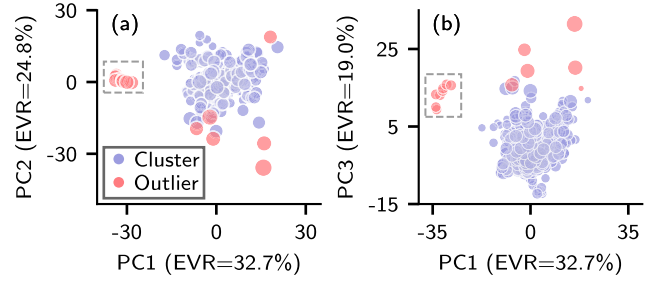


Fig. 4. DBSCAN clustering results projected onto pairs of PCs. (a) PC1 versus PC2 and (b) PC1 versus PC3. Clustering is performed in the full 4D PCA space (PC1–PC4) with $(\epsilon, N_{\text{points}}^{\min})$ chosen from Fig. 3 (here $\epsilon = 10$ and $N_{\text{points}}^{\min} = 10$). The legend distinguishes clustered points (blue) from outliers (red). Axis labels report the EVR for each PC. To mimic a 3-D perspective in these 2-D projections, point diameters are proportional to the coordinate along the third orthogonal PC [PC3 in panel (a) and PC2 in panel (b)] after min–max scaling. Gray dashed rectangles mark the compact group of short pulses that forms a separate cluster when $N_{\text{points}}^{\min} < 10$ but is classified as outliers when $N_{\text{points}}^{\min} = 10$ (see Section IV). Note that projections may visually distort actual distances, since clustering is computed in four dimensions.

DBSCAN is well-suited to this task because, unlike simpler methods such as K-means, it does not assume any predefined number of clusters and can naturally handle clusters of varying shapes and sizes. In our analysis, it operates in the 4-D PCA space, where the bulk of nominal pulses forms a compact cluster, while outliers (i.e., pulses with atypical thermal behavior) are clearly separated, as will be discussed in Section V. For comparison, we also tested the Isolation Forest algorithm [23], which yielded results largely consistent with those of DBSCAN, showing over 90% overlap in the identified outliers. However, Isolation Forest requires specifying a contamination factor (i.e., the expected fraction of anomalies), tested here between 1% and 5%, which is not ideal in unsupervised settings where the true proportion of outliers is unknown.

The entire pipeline—including feature extraction, feature filtering, PCA projection, and DBSCAN clustering—runs in approximately 15 s on a standard workstation using a single core. In practice, since the statistical features from previously processed pulses can be cached, only the newly acquired pulse needs to be processed. By parallelizing the computations and reusing cached data, the whole intershot anomaly detection can be completed in under one second after the pulse ends, providing near-instantaneous feedback to the operators.

V. RESULTS

The low-dimensional representation obtained from PCA reveals that the majority of pulses form a single dense cluster in the reduced feature space. When DBSCAN is applied to this 4-D space, it identifies 14 outlier pulses (out of 401) that deviate from the main cluster. Two 2-D projections of the results along the first three principal components (PC1–PC3), which together account for approximately 77% of the original variance, are shown in Fig. 4. For reference, PCs 1 through 4 cumulatively retain about 82% of the total variance in the dataset.

To interpret why certain pulses are flagged as outliers, we examine their average temperature profiles, denoted $\langle \hat{T}_n(\hat{t}) \rangle$.

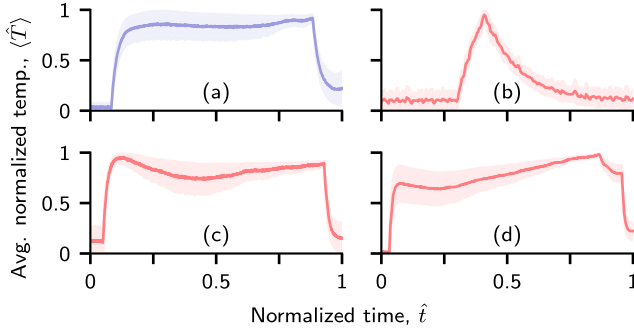


Fig. 5. Normalized average temperature profiles $\langle \hat{T}_n(\hat{t}) \rangle$ for four representative pulses. (a) Typical pulse identified as a core sample, displaying the expected ramp-up, plateau, and cool-down behavior. (b) Outlier pulse with a missing plateau due to its very short duration. (c) Outlier pulse showing an unexpected dip in temperature during the plateau phase. (d) Outlier pulse characterized by a progressive temperature rise instead of a flat plateau, followed by a two-step decay during cool-down. The shaded area around each curve indicates one standard deviation across the thermocouple signals, highlighting variability between sensors. For visualization purposes, the time axis has been extended: $t_{\min}$ and $t_{\max}$ are defined as 10 s before and 15 s after the application and shutdown of the acceleration voltage U_{EG} , respectively. This provides better visibility of the ramp-up and cool-down phases.

This profile represents the mean of all thermocouple signals from a given pulse after signal-wise normalization and temporal rescaling. Specifically, for each pulse n and each thermocouple k , we compute

$$\langle \hat{T}_n(\hat{t}) \rangle = \frac{1}{N_t} \sum_{k=1}^{N_t} \frac{T_n^k(\hat{t}) - \min(T_n^k)}{\max(T_n^k) - \min(T_n^k)} \quad (3)$$

where $\hat{t} \in [0, 1]$ is the normalized time, defined as $\hat{t} = (t - t_{\min}) / (t_{\max} - t_{\min})$. This transformation allows all pulses to be visualized on a common time axis, regardless of their original durations. Fig. 5 shows the average normalized profiles of one typical (core) pulse and three outlier pulses. The core pulse exhibits the expected ramp-up, plateau, and cool-down behavior, whereas the outliers clearly deviate from this shape. In particular, we observe a missing plateau, progressive drift, and premature decay, consistent with thermal irregularities.

As mentioned in the previous section, a noticeable change in DBSCAN behavior occurs between $N_{\text{points}}^{\min} = 8$ and 10, where the algorithm transitions from detecting two clusters to a single one. In Fig. 4, gray dashed rectangles highlight the compact group of short pulses on the left side of both projections, which forms a distinct cluster when $N_{\text{points}}^{\min} = 8$ but is classified as outliers under the more conservative value of $N_{\text{points}}^{\min} = 10$. Closer inspection of the pulses in this group reveals that they are all very short and lack a plateau phase, exhibiting only a ramp-up followed by an immediate cool-down. While this waveform difference is inherently linked to pulse duration, being flagged as outliers is not due to duration itself: the anomaly detection pipeline was specifically designed to remove features highly correlated with duration to avoid such bias. Instead, these pulses are flagged because their thermal evolution deviates structurally from the dominant pattern shared by most longer pulses, regardless of exact duration, as shown in Fig. 5(b). Importantly, these short pulses are expected operationally (e.g., during commissioning or early campaign phases) and are not pathological: they simply

represent a different regime within the dataset. This illustrates a broader point: in operational settings, different regimes of behavior should manifest as separate clusters in feature space once sufficient data is accumulated. Initially, rare regimes (like these very short pulses) may appear as outliers, but as more data is collected under those conditions, they may eventually form stable clusters and be recognized as distinct operational modes. The current choice of $N_{\text{points}}^{\min} = 10$ is thus a conservative decision that avoids prematurely assigning cluster status to a small group of samples. In the future, this framework may evolve from a binary classification (“normal” versus “outlier”) toward a richer taxonomy of operational states (e.g., “short,” “long,” “outlier”), with new clusters emerging as data accumulates.

These results confirm that the extracted statistical features are sensitive to global shifts in thermal behavior, and DBSCAN can identify full pulses that differ structurally from the nominal trend. The combined use of time-independent features, correlation-aware filtering, dimensionality reduction, and density-based clustering, therefore, enables the detection of a wide range of anomalies in the thermal profiles of gyrotron pulses. However, local or subtle anomalies affecting a single thermocouple may go undetected, as their influence can be diluted in the original high-dimensional feature space.

It is also important to emphasize that a pulse being flagged as an outlier does not necessarily imply a physical fault in the system. Unsupervised methods such as DBSCAN merely flag samples that are statistically distinct from the bulk of the dataset. For instance, the two-step cool-down observed in Fig. 5(d) may correspond to a physical intervention, such as the activation of an auxiliary cooling loop triggered by overheating, though this remains to be verified.

Since no labeled data is currently available, the anomaly detection pipeline presented in this work relies on unsupervised ML methods and is intentionally conservative. Its purpose is not to provide definitive fault classification but to highlight unusual pulses that merit further investigation by raising a warning in the TSM HMI for operator review. Over time, expert feedback on these flagged pulses will enable the refinement of detection algorithms and the eventual deployment of more specialized or supervised models.

VI. CONCLUSION AND FUTURE DEVELOPMENTS

This study presents an ML-based framework for the detection of anomalous thermal behaviors in gyrotron pulses, as part of the TSM software being developed for ITER. Despite the absence of labeled data, the proposed unsupervised method—based on time-independent features, dimensionality reduction via PCA, and clustering with DBSCAN—demonstrates the ability to distinguish nominal pulses from those exhibiting deviations from the expected thermal profile. By transforming each pulse into a fixed-length feature vector, the method enables consistent comparison across cycles of varying durations and effectively captures global anomalies such as atypical ramp-up, missing plateaus, or unusual cool-down sequences.

While encouraging, the current approach remains relatively coarse-grained. It evaluates each pulse as a whole and

provides no direct insight into the specific time intervals or thermocouple signals responsible for the detected anomaly. Moreover, localized or subtle irregularities, such as transient disturbances in a single sensor, may be diluted during the feature aggregation presented in (1). Improving the sensitivity of the detection pipeline will therefore be a key focus of future work.

A crucial next step will be the development of time-resolved algorithms capable of detecting anomalies within seconds of their occurrence if run during operations, or allowing a more detailed postmortem analysis of a pulse. Such methods would complement the intershot approach by enabling fine-grained temporal localization of anomalous behavior. Exploring deep learning models such as LSTM autoencoders, trained on sequential thermocouple data, may offer a path toward within-pulse anomaly detection.

Another important avenue for development involves incorporating operator feedback directly into the anomaly detection workflow via the HMI of TSM. This feedback loop will help distinguish true anomalies from benign deviations, guide the refinement of existing algorithms, and eventually enable supervised training. Improving the interpretability of the models [24], e.g., by identifying which features or sensors contributed most to a flagged anomaly, will also be essential. To this end, alternative algorithms such as Isolation Forest may be explored further in future iterations of the pipeline, as they provide continuous anomaly scores and are readily interpretable using tools like SHAP [25], although they inherently require a predefined contamination factor forcing the identification of a fixed proportion of outliers, as discussed in Section IV. Hierarchical extensions such as HDBSCAN [26], [27] also merit consideration, as they identify clusters across multiple density scales without the need for fixed parameter tuning and, importantly, provide a continuous probability-like measure of cluster membership, thereby combining the strengths of DBSCAN with the graded anomaly scoring offered by methods like Isolation Forest.

Although this work focused on the gyrotron collector, the monitoring approach and data processing pipeline are applicable to other ITER subsystems that exhibit similar thermal cycles. The anomaly detection module of the TSM software is currently under active development, and the first algorithms based on the methods outlined here are scheduled for integration in the near future.

Overall, this initial demonstration lays the groundwork for scalable, data-driven anomaly detection in complex fusion systems. It provides a strong starting point for further algorithmic refinement and broader deployment across ITER subsystems, and ultimately supports the goal of robust, autonomous health monitoring for the ITER machine.

AUTHOR CONTRIBUTIONS

Joris Paret—Conceptualization, Methodology, Software, Formal analysis, Investigation, Data Curation, Writing—Original Draft, Visualization, Project administration. Daniel Iglesias—Writing—Review and Editing, and Project administration, Supervision. Daniel Sabio Ruiz—Formal analysis. Dilin Meloni—Formal analysis. Andrea Antonione—Resources,

Data Curation. Ruggero Bertazzoni—Resources, Project administration. Giuseppe Carannante—Supervision. Martino Ferrari—Supervision. Marfa Ortiz de Zúñiga—Supervision. Mario Cavinato—Supervision. Francisco Arcos Sánchez—Supervision. Alfredo Portone—Supervision.

DISCLAIMER

The views and opinions expressed herein do not necessarily reflect those of the ITER Organization and Fusion for Energy.

DATA AND CODE AVAILABILITY

The dataset analyzed in this study is proprietary and cannot be shared publicly. However, the codes used to perform the data analysis and anomaly detection can be made available by the authors upon reasonable request.

ACKNOWLEDGMENT

The authors would like to thank the ITER Organization, Fusion for Energy, the Falcon team, and all collaborators who contributed to this research.

During the preparation of this work, they acknowledge the use of Microsoft 365 Copilot in order to improve writing efficiency and clarity. The AI-generated text was reviewed and edited by them, who take full responsibility for the content of the published article.

REFERENCES

- [1] A. B. Nassif, M. A. Talib, Q. Nasir, and F. M. Dakalbab, "Machine learning for anomaly detection: A systematic review," *IEEE Access*, vol. 9, pp. 78658–78700, 2021.
- [2] C. Rea, K. J. Montes, K. G. Erickson, R. S. Granetz, and R. A. Tinguely, "A real-time machine learning-based disruption predictor in DIII-D," *Nucl. Fusion*, vol. 59, no. 9, Sep. 2019, Art. no. 096016.
- [3] J. Kates-Harbeck, A. Svyatkovskiy, and W. Tang, "Predicting disruptive instabilities in controlled fusion plasmas through deep learning," *Nature*, vol. 568, no. 7753, pp. 526–531, Apr. 2019.
- [4] D. R. Ferreira, P. J. Carvalho, C. Sozzi, P. J. Lomas, and J. Contributors, "Deep learning for the analysis of disruption precursors based on plasma tomography," *Fusion Sci. Technol.*, vol. 76, no. 8, pp. 901–911, Nov. 2020.
- [5] G. Farias, E. Fabregas, S. Dormido-Canto, J. Vega, and S. Vergara, "Automatic recognition of anomalous patterns in discharges by applying deep learning," *Fusion Sci. Technol.*, vol. 76, no. 8, pp. 925–932, Nov. 2020.
- [6] B. H. Guo et al., "Disruption prediction on EAST tokamak using a deep learning algorithm," *Plasma Phys. Controlled Fusion*, vol. 63, no. 11, Nov. 2021, Art. no. 115007.
- [7] J. Vega, S. Dormido-Canto, R. Castro, J. D. Fernández, A. Murari, and J. Contributors, "Real-time disruption prediction in multi-dimensional spaces leveraging diagnostic information not available at execution time," *Nucl. Fusion*, vol. 64, no. 4, Apr. 2024, Art. no. 046010.
- [8] R. Rossi, M. Gelfusa, T. Craciunescu, I. Wyss, and J. Vega, "A hybrid physics/data-driven logic to detect, classify, and predict anomalies and disruptions in tokamak plasmas," *Nucl. Fusion*, vol. 64, no. 4, Apr. 2024, Art. no. 046017.
- [9] X. K. Ai et al., "Tokamak plasma disruption precursor onset time study based on semi-supervised anomaly detection," *Nucl. Eng. Technol.*, vol. 56, no. 4, pp. 1501–1512, Apr. 2024.
- [10] J. Kim, J. Lee, J. Seo, Y.-C. Ghim, Y. Lee, and Y.-S. Na, "Enhancing disruption prediction through Bayesian neural network in KSTAR," *Plasma Phys. Controlled Fusion*, vol. 66, no. 7, Jul. 2024, Art. no. 075001.
- [11] A. R. Saperstein, C. Rea, R. Sweeney, M. D. Boyer, G. L. Trevisan, and Y. Wei, "Design and development of an off-normal warning system for SPARC," *Nucl. Fusion*, vol. 65, no. 11, Nov. 2025, Art. no. 116007.

- [12] R. Skilton and Y. Gao, "Combining object detection with generative adversarial networks for in-component anomaly detection," *Fusion Eng. Design*, vol. 159, Oct. 2020, Art. no. 111736. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0920379620302842>
- [13] H. Anand, B. S. Sammulu, K. E. J. Olofsson, and D. A. Humphreys, "Real-time magnetic sensor anomaly detection using autoencoder neural networks on the DIII-D tokamak," *IEEE Trans. Plasma Sci.*, vol. 50, no. 11, pp. 4126–4130, Nov. 2022.
- [14] V. Gorse, R. Mitteau, and J. Marot, "Anomaly classification by inserting prior knowledge into a max-tree based method for divertor hot spot characterization on WEST tokamak," *Rev. Scientific Instrum.*, vol. 94, no. 8, Aug. 2023, Art. no. 083510.
- [15] E. Grelier, R. Mitteau, and V. Moncada, "Deep learning-based process for the automatic detection, tracking, and classification of thermal events on the in-vessel components of fusion reactors," *Fusion Eng. Design*, vol. 192, Jul. 2023, Art. no. 113636.
- [16] V. Gorse, E. Grelier, V. Moncada, and R. Mitteau, "Real-time monitoring system for detection and characterization of thermal events on WEST tokamak: Implementation and first results," *Fusion Eng. Design*, vol. 215, Jun. 2025, Art. no. 114960.
- [17] L. Maqueda et al., "Feasibility evaluation and pre-conceptual design of the iter tokamak systems monitor," *Fusion Eng. Design*, vol. 188, Mar. 2023, Art. no. 113435.
- [18] T. P. Goodman et al., "Tests and qualification of the European 1 MW, 170 GHz CW gyrotron in an ITER relevant configuration at SPC," in *Proc. 47th Int. Conf. Infr., Millim. Terahertz Waves (IRMMW-THz)*, Aug. 2022, pp. 1–2.
- [19] F. Braunmueller et al., "Results on the 1 MW CW 170 GHz gyrotron TH1509UA for ITER and DTT at the FALCON test stand," *EPJ Web Conf.*, vol. 313, p. 04005, Jun. 2024.
- [20] M. Barandas et al., "TSFEL: Time series feature extraction library," *SoftwareX*, vol. 11, Jan. 2020, Art. no. 100456.
- [21] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proc. 2nd Int. Conf. Knowl. Discovery Data Mining*, 1996, pp. 226–231.
- [22] J. Sander, M. Ester, H.-P. Kriegel, and X. Xu, "Density-based clustering in spatial databases: The algorithm GDBSCAN and its applications," *Data Min. Knowl. Discov.*, vol. 2, no. 2, pp. 169–194, 1998.
- [23] F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation forest," in *Proc. 8th IEEE Int. Conf. Data Min.*, Dec. 2008, pp. 413–422, doi: [10.1109/ICDM.2008.17](https://doi.org/10.1109/ICDM.2008.17).
- [24] Z. Li, Y. Zhu, and M. Van Leeuwen, "A survey on explainable anomaly detection," *ACM Trans. Knowl. Discov. Data*, vol. 18, no. 1, pp. 1–54, Sep. 2023, doi: [10.1145/3609333](https://doi.org/10.1145/3609333).
- [25] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Proc. 31st Int. Conf. Neural Inf. Process. Syst.* New York, NY, USA: Curran Associates Inc., 2017, pp. 4768–4777.
- [26] R. J. G. B. Campello, D. Moulavi, and J. Sander, "Density-based clustering based on hierarchical density estimates," in *Advances in Knowledge Discovery and Data Mining*. Berlin, Germany: Springer, 2013, pp. 160–172.
- [27] R. J. G. B. Campello, D. Moulavi, A. Zimek, and J. Sander, "Hierarchical density estimates for data clustering, visualization, and outlier detection," *ACM Trans. Knowl. Discovery Data*, vol. 10, no. 1, pp. 1–51, Jul. 2015, doi: [10.1145/2733381](https://doi.org/10.1145/2733381).