

Article

Predicting Overload Risk on Plasma-Facing Components at Wendelstein 7-X from IR Imaging Using Self-Organizing Maps

Giuliana Sias ^{1,*}, Emanuele Corongiu ^{1,*}, Enrico Aymerich ¹, Barbara Cannas ¹, Alessandra Fanni ¹, Yu Gao ², Bartłomiej Jabłoński ³, Marcin Jakubowski ², Aleix Puig Sitjes ², Fabio Pisano ^{1,*}
and W7-X Team [†]

¹ Department of Electrical and Electronic Engineering, University of Cagliari, 09123 Cagliari, Italy; enrico.aymerich@unica.it (E.A.); barbara.cannas@unica.it (B.C.); fanni@unica.it (A.F.)

² Max-Planck-Institut für Plasmaphysik, 17491 Greifswald, Germany; yu.gao@ipp.mpg.de (Y.G.); marcin.jakubowski@ipp.mpg.de (M.J.)

³ Department of Microelectronics and Computer Science, Lodz University of Technology, 93-005 Lodz, Poland; jablonskiba@dmcs.pl

* Correspondence: giuliana.sias@unica.it (G.S.); emanuele.corongiu@unica.it (E.C.); fabio.pisano@unica.it (F.P.)

[†] Collaborators/Membership of the W7-X Team is provided in the Acknowledgments.

Abstract: Overload detection is crucial in nuclear fusion experiments to prevent damage to plasma-facing components (PFCs) and ensure the safe operation of the reactor. At Wendelstein 7-X (W7-X), real-time monitoring and prediction of thermal events are essential for maintaining the integrity of PFCs. This paper proposes a machine learning approach for developing a real-time overload detector, trained and tested on OP1.2a experimental data. The analysis showed that Self-Organizing Maps (SOMs) are efficient in detecting the overload risk starting from a set of plasma parameters that describe the magnetic configuration, the energy behavior, and the power balance. This study aims to thoroughly evaluate the capabilities of the SOM in recognizing overload risk levels, defined by quantizing the maximum criticality across different IR cameras. The goal is to enable detailed monitoring for overload prevention while maintaining high-performance plasmas and sustaining long pulse operations. The SOM proves to be a highly effective overload risk detector. It correctly identifies the assigned overload risk level in 87.52% of the samples. The most frequent error in the test set, occurring in 10.46% of cases, involves assigning a risk level to each sample adjacent to the target one. The analysis of the results highlights the advantages and drawbacks of criticality discretization and opens new solutions to improve the SOM potential in this field.

Keywords: overload detection; Self-Organizing Maps; Wendelstein 7-X; plasma-facing components; real-time monitoring; infrared diagnostics



Academic Editor: Enrico Zio

Received: 30 April 2025

Revised: 30 May 2025

Accepted: 11 June 2025

Published: 18 June 2025

Citation: Sias, G.; Corongiu, E.; Aymerich, E.; Cannas, B.; Fanni, A.; Gao, Y.; Jabłoński, B.; Jakubowski, M.; Puig Sitjes, A.; Pisano, F.; et al. Predicting Overload Risk on Plasma-Facing Components at Wendelstein 7-X from IR Imaging Using Self-Organizing Maps. *Energies* **2025**, *18*, 3192. <https://doi.org/10.3390/en18123192>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

W7-X is the world's largest stellarator-type fusion device [1], designed to demonstrate the feasibility of stellarators as a future power plant concept. Due to its stellarator-type structure, W7-X features a three-dimensional helical architecture with five-fold modular symmetry [2–4]. The main objective of W7-X is to achieve steady-state plasma operation at fusion-relevant parameters, addressing key challenges, such as plasma confinement, stability, and heat exhaust. The protection of plasma-facing components during long-pulse plasma operation is also one of the main concerns for ITER operation and lifetime, as discussed in the ITER Physics Basis [5,6].

One of the critical aspects of ensuring the safe and efficient operation of W7-X is the development of reliable overload risk detectors [7]. The exhaust of thermal energy from the confined plasma is predominantly managed by an island divertor system, which channels the heat toward ten specialized divertor units—distributed equally between the upper and lower sections of the device. Each unit integrates vertical and horizontal targets, specifically engineered to tolerate intense thermal loads, with design limits reaching up to 10 MW/m^2 . In Figure 1, one of the ten divertor units is shown.

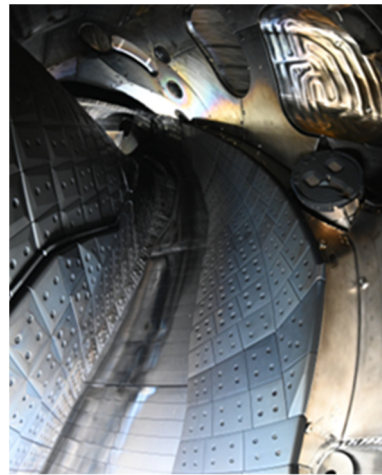


Figure 1. View of a divertor unit.

The need for developing effective overload risk detectors is underscored by the potential risks associated with thermal overloads, which can compromise the structural integrity of key elements, particularly those in direct contact with the hot plasma [8].

In this context, a key challenge in nuclear fusion research is finding a good trade-off between safety [9] and high-performance plasmas. In fact, achieving high performance often involves operating at conditions close to the stability limits of the plasma, which increases the risk of thermal overloads [10]. Therefore, it is crucial to develop systems that can predict and mitigate these risks in real-time [11], ensuring both the safety of the reactor and the achievement of high-performance plasma conditions. Additionally, thermal overload detection may entail monitoring the performance of the present cooling systems, which are active on the 10 water-cooled divertors of W7-X, to guarantee the efficient dissipation of heat from critical components and maintain a safe operating temperature range [12,13].

Automated systems using deep learning algorithms, such as Cascade R-CNN and SORT [14,15], are under development in nuclear fusion devices, e.g., at WEST, to detect, track, and classify thermal events. These systems are trained on datasets of thermal events and are capable of identifying the hottest zones, which are critical for preventing damage to PFCs [16].

An infrared (IR) image analysis system is currently under deployment at W7-X for real-time protection of the divertors. The system aims to prevent thermal overloads that could damage PFCs, leading to machine downtimes and repair costs. The thermal overload detection (TOD) system processes calibrated IR images of PFCs acquired every 10 ms to produce an alarm signal for the Fast Interlock System (FIS), thresholding an overload risk indicator, calculated among the different IR cameras. The calibration, made to map digital levels to apparent temperatures, includes correcting for non-uniformity, bad pixels, and temperature drift. Pre-processing filters out bad pixels and averages surface temperatures to reduce fluctuations. The detection stage adjusts the temperature thresholds to anticipate overloads and triggers the FIS to stop heating systems if an overload is detected. A real-time

constraint to produce an overload alarm is 0.11 s, which entails acquisition, calibration, detection, and interlock delays [11].

This study focuses on developing a machine learning (ML)-based tool for monitoring the overload risk at W7-X, specifically using the Self-Organizing Map (SOM) [17], a nonlinear dimensionality reduction technique that preserves input space similarities while mapping data onto a lower-dimensional (typically 2D) representation [18]. Two-dimensional mappings have been widely applied in nuclear fusion for classification and prediction tasks [19–22], including disruption prediction [23,24], and electromagnetic studies [25]. In the context of thermal overload detection, SOMs offer intuitive visual insights that are easily interpretable by experimentalists, aiding in the development of effective thermal safety and control policies.

The primary aim of this paper is to demonstrate that the multidimensional operational space of W7-X, as explored through the analyzed experimental campaigns, can be effectively mapped onto a 2D Self-Organizing Map (SOM) and proposed as a real-time overload risk detector. To support this, a criterion is introduced to correlate the map's composition with different overload risk levels. As a foundational step toward a probabilistic approach, a multi-level risk scale is defined to estimate the probability of overload for each experimental sample.

Furthermore, this paper highlights the added value of a SOM-based detector, particularly in terms of output interpretability. The 2D map enables tracking the operational point in relation to the evolution of input features, providing insight into the underlying causes of the detector's response and helping to interpret classification errors. This level of interpretability is rarely achievable with conventional machine learning algorithms.

This paper is organized as follows. In Section 2, both the characteristics of the input space and the ratio used to define the target are described. In Section 3, the fundamentals of the SOM algorithm are reported. In Section 4, the performance analysis is reported, and the results are discussed. In Section 5, the main causes of the classification errors are analyzed. Finally, the conclusions are drawn in Section 6.

2. Methodology

2.1. Experimental Dataset

A database of 17 0D plasma parameters, for a list of 69 experiments taken from the OP1.2a campaign, was developed. Among all the experiments available from the OP1.2a campaign, only items containing all the selected plasma parameters were retained.

The signals were selected based on their impact on the presence of a thermal overload. W7-X utilizes a variety of diagnostic and control input signals to monitor and influence plasma behavior. A key aspect is the magnetic configuration, which is governed by the currents in different sets of coils: the planar coils, non-planar coils, trim coils, and control coils. These coil systems collectively shape the magnetic field geometry, enabling W7-X to explore optimized stellarator configurations and control edge conditions, such as island divertors. In particular, the upper and lower divertors control coil currents, together with the toroidal current, measured using Rogowski coils, affect the distribution of heat loads on the divertors while impacting the overall magnetic confinement and stability of the plasma [26–28]. While W7-X is primarily a current-less stellarator, any induced or driven toroidal current is relevant for stability and confinement studies. Heating power is provided through two main systems: Electron Cyclotron Resonance Heating (ECRH) and Neutral Beam Injection (NBI). These systems inject energy into the plasma to raise its temperature and maintain the required conditions for fusion-relevant plasma behavior. To assess the plasma's energy content, diamagnetic energy measurements are used. These are derived from magnetic diagnostics and provide a direct estimate of the total thermal energy stored in the plasma. Finally, the total radiated power is measured using bolometers positioned to

provide both horizontal and vertical views of the plasma. These measurements are critical for energy balance calculations and for understanding radiative losses due to impurities and other processes within the plasma. Together, these input signals form the foundation of W7-X's operational diagnostics, enabling accurate control, monitoring, and analysis of plasma discharges. A time window for each experiment was defined based on the toroidal current value and by removing potential outliers and experiments with missing values. A features' evaluation was made in order to discard correlated signals reducing the space dimensionality. In particular, control coil currents were aggregated into one feature, by averaging all the upper divertors and lower divertors control coil currents. Also, the additional heating powers (ECRH and NBI) were aggregated into a single feature providing the total heating power. Regarding the bolometer, the total radiated power calculated from both the horizontal (H) bolometer camera and the vertical (V) bolometer camera were used. The list of features used to train the SOM-based detector is reported in Table 1.

Table 1. List of the input features.

Feature	Unit
Averaged control coil current	A
Toroidal current	A
Five trim coil (TC) currents	A
Five non-planar coil (NPC) currents	A
Averaged planar coil current	A
Total input power (P_{INP})	W
Total radiated power from H bolometer camera	W
Total radiated power from V bolometer camera	W
Diamagnetic energy (W_{DIA})	J

As can be seen, the SOM does not receive explicit input regarding IR properties. The temperature data from the IR cameras is used solely to define the target for the overload detector.

2.2. Criticality

In this paper, the overload detector's target is defined as the maximum criticality across all the different IR cameras. The criticality (c) is an overload risk indicator, calculated for each IR camera image, pixel, and time instance, as the ratio between the pixel temperature T and a dynamic threshold T_{th} :

$$c(x, y, t) = \frac{T(x, y, t)}{T_{th}(x, y, t)} \quad (1)$$

where (x, y) is the pixel position and t is the time instant [8].

Before calculating the criticality and the related threshold T_{th} , several filters and morphological operations are applied to the temperature taken from the IR camera in order to reduce the effect of noise peaks. The temperature derivative is then estimated as the temperature increment from the previous frame divided by the frame time $\Delta t = 10$ ms. After this, the temperature threshold T_{th} is computed to avoid a temperature overshoot of the temperature limit within the reaction time of the system $t_r = 0.16$ s. Given the high noise in the temperature derivative calculation, a protective mechanism is introduced, and

the threshold is set as the minimum between the computed threshold and an upper bound $T_{th}^{max}(x, y)$ dependent on the material properties:

$$T_{th}(x, y, t) = \min \left\{ T_{limit}(x, y) - t_r \frac{dT}{dt}(x, y, t), T_{th}^{max}(x, y) \right\} \quad (2)$$

where $T_{limit}(x, y)$ is the limit temperature of the component in the related pixel (see Figure 3 in [8]).

Finally, a moving average filter is applied to the criticality calculated in (1) in order to filter out the fast transients and spikes.

A criticality greater than or equal to 1 means that if there are no heat-flux changes in the short term, an overload will occur within or before the reaction time, i.e., without the possibility of avoiding it in due time. Thus, a binary classifier based on the criticality thresholding is employed: when $c \geq 1$, an alarm is triggered, and the experiment is interrupted.

2.3. Machine Learning Target

To prevent overload during high-performance operations, a tool is proposed to monitor approaching overload by detecting increasing risk levels. This allows the control system to plan different avoidance actions depending on the detected risk level.

The diagnostic signals from 69 shots in the database were downsampled to a common time base, matching the IR camera's sampling time (10 ms). This resulted in a dataset of 36,256 samples, with 70% used for training the SOM and the remaining 30% for performance testing.

In this paper, a scale of five risk levels for thermal overload is defined as the target. Compared to a simple binary classification of risk (i.e., safe operation vs. overload), this approach allows each risk level to incorporate preventive actions that are more aligned with those of neighboring levels. The discretization of criticality into a finite number of risk levels serves as an initial step toward a probabilistic framework, in which each sample will eventually be assigned an estimated overload probability.

The five risk levels were defined by equally distributing the dataset samples among five ranges of criticality, as summarized in Table 2. Since both higher and lower criticality samples are typically scarce in each experiment, they were aggregated into two larger intervals.

Table 2. Risk levels (LVs) and related ranges of criticality.

Risk Level	Range	Label
1	$c \leq 0.68$	LV1
2	$0.68 < c \leq 0.75$	LV2
3	$0.75 < c \leq 0.83$	LV3
4	$0.83 < c \leq 0.94$	LV4
5	$c > 0.94$	LV5

Since balancing a dataset makes training an ML model easier, because it helps prevent the model from becoming biased towards one risk level, the training samples were selected in order to cover equally the five ranges of criticality, as shown in Figure 2. Only 15 shots reached the highest risk level of thermal overload. Note that the higher level LV5 mostly covered the high overload range (i.e., $c \geq 1$) defined by the binary classifier used at W7-X [8].

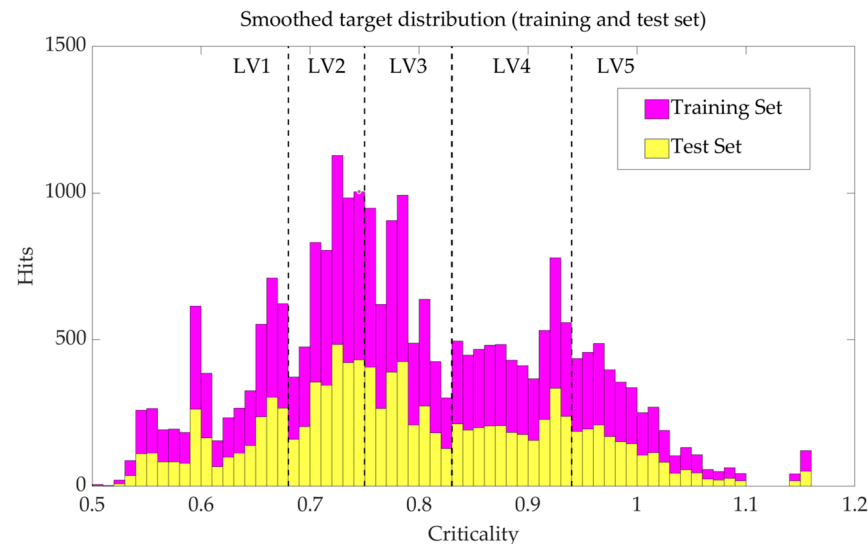


Figure 2. Distribution of training and test samples among the five ranges of criticality.

3. Self-Organizing Maps

The Self-Organizing Map (SOM) [17,29] is a nonlinear dimensionality reduction technique that generates a low-dimensional representation of data while preserving the topology of the input space. A SOM is a neural network with two layers, i.e., the input layer and the output layer, with the nodes of the input layer directly connected to those of the output layer. The output layer is arranged as a bounded grid of neurons, referred to as prototypes, each having coordinates in both the original high-dimensional space and a lower-dimensional grid (see Figure 3). The SOM is trained iteratively using competitive nonsupervised learning, where each data point is assigned to its Best Matching Unit (BMU), i.e., the closest prototype in the input space. The BMU and its neighbors are updated to move towards the data point, following an adaptive learning rule governed by a neighborhood function that decreases over time. This process allows prototypes to gradually organize themselves within the data distribution while preserving topological relationships rather than strict Euclidean distances. Once training is complete, new data points can be mapped to the closest prototype, effectively clustering the input space. A SOM requires careful selection of parameters, such as the grid size, learning rate, and neighborhood function, to ensure meaningful organization of the data.

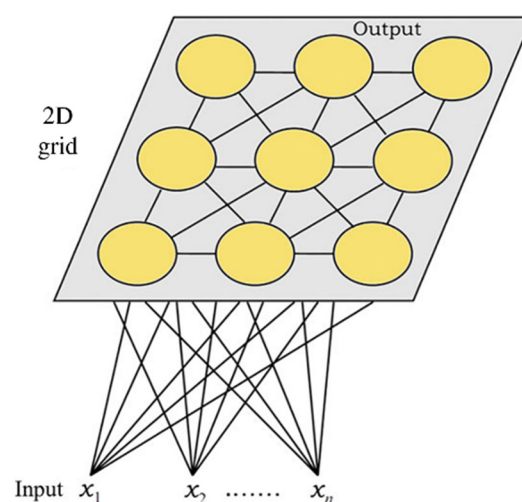


Figure 3. Example of the architecture of a 2D SOM. The output layer is arranged as a bounded grid of neurons, while the nodes of the input layer are directly connected to those of the output layer.

Beyond dimensionality reduction, SOMs are widely used for unsupervised learning tasks, such as clustering, visualization, and anomaly detection. Unlike other clustering methods, SOMs can provide an intuitive and interpretable 2D (or 3D) grid that highlights the relationships between clusters, such as the one shown in Figure 3. In addition, SOM interpretability can be enhanced by the so-called component planes, consisting of the representation of the distribution of each input feature on the grid, through color-coding [30]. Plotting the component planes for all input features allows for correlating each variable's behavior with the resulting clustering, assisting in obtaining a better understanding of the characteristics of specific map regions and in discovering the underlying dependencies [31]. Indeed, by comparing the component planes, the input correlation can be easily investigated by visual inspection. If the outlook of the component planes is similar, the inputs are strongly correlated.

This study used a 2D SOM, which is a commonly used tool for clustering and classification tasks [32]. The architecture was optimized by using the toolbox presented in [33].

4. Overload Detector

4.1. W7-X Operational Space Mapping

A 2D grid with a hexagonal topology was selected to map the W7-X operational space, represented by the 17D experimental dataset described in Section 2.1, into one that is easily visualized in two dimensions. The map size, as suggested in [33], is a function of the number N of samples defining the input space. However, to optimize the map size for the specific application, multiple maps were trained with the number of clusters varying within the range ($245 \div 4814$). For each map, the clustering quality was evaluated based on the clustering error, which quantifies the degree to which samples of different risk levels are grouped within the same cluster. Note that the SOMs were trained by an unsupervised algorithm; thus, no information was provided about the target risk levels when the clustering was performed. The risk level value was assigned to each sample after the training with the aim of interpreting the attained clustering. To this purpose, the clustering error was evaluated by assigning a label to each cluster depending on its composition in terms of risk levels. The clusters containing samples from a single risk level were labeled with their own level value, and the label corresponded to the risk level. The clusters containing samples from at least two risk levels were labeled as mixed clusters, and they did not have a unique risk level assigned. Finally, clusters with no samples were labeled as empty clusters, and they did not have any risk level assigned. Therefore, the SOM clustering error was expressed by the fraction of mixed clusters. For each map size, Figure 4a reports the fraction of clusters for each defined risk level (LV1 in light blue, LV2 in light orange, LV3 in orange, LV4 in red, and LV5 in dark red), the mixed cluster in black, and the empty clusters in grey. The diagram shows that, for a size larger than 2000 clusters, the empty cluster fraction has roughly flat behavior, while, as expected, the mixed one slightly decreases. The optimum clustering should reduce the clustering error, and thus, the mixed cluster rate, while limiting the empty one. Indeed, a large number of empty clusters reduces the gathering of clusters with the same level in macro-clusters. Thus, a map size larger than 2000 meets both aims. Since mixed clusters can contain up to five risk levels, different types of mixed clusters can be defined depending on the number of levels in the clusters. Figure 4b reports the fraction of the mixed clusters involving two (in light orange), three (in orange), four (in red), and five (in dark red) risk levels, with a growing magnitude of clustering fault. For each map size, most of the mixed clusters contain only two levels, while the percentage of the ones mixing five risk levels is always a few units. For a map size larger than about 3000 clusters, the fraction of mixed clusters with the lowest degree of fault is about 80%. For a map size larger than 4000, the fraction of empty clusters

slightly increases; therefore, the map size of 3822 was selected as the optimum one. The selected map is arranged in a rectangular lattice with 147 columns by 26 rows. Each cluster is represented by a hexagonal map unit.

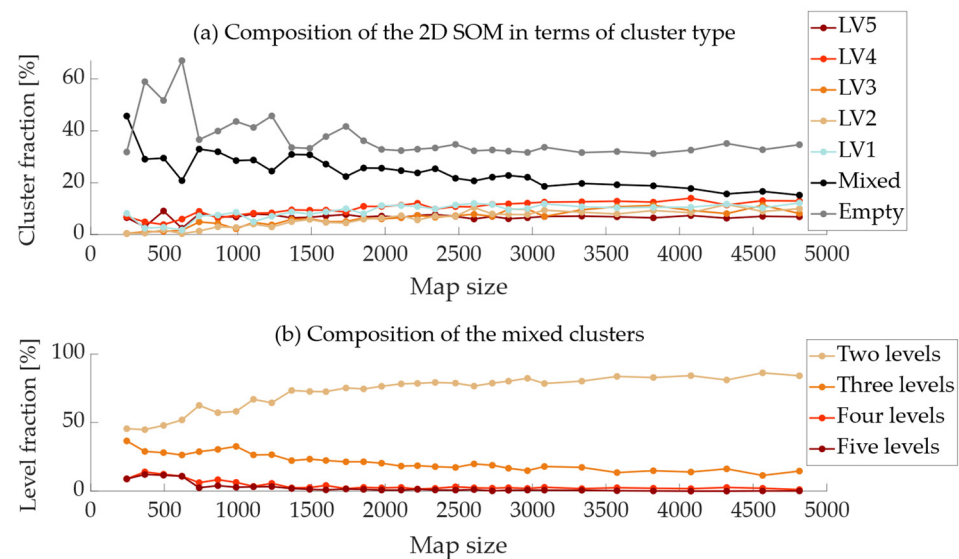


Figure 4. (a) Composition of the SOM in terms of cluster types, expressed as the fraction of the total map size. Clusters containing only samples belonging to one risk level: LV1, light blue line; LV2, light orange line; LV3, orange line; LV4, red line; and LV5, dark red line. Clusters containing more than one level (mixed): black line. Clusters containing no samples (empty): gray line. (b) Composition of the mixed cluster. Mixed cluster type as a fraction of the mixed cluster. Clusters containing 2 levels in light orange, 3 levels in orange, 4 levels in red, and 5 levels in dark red.

The composition of the selected map in terms of cluster types is detailed in Table 3. Figure 5 reports the selected 2D map, where the clusters containing samples from LV1 are colored in light blue, those containing samples from LV2 are colored in light orange, and those containing samples from LV3, LV4, and LV5 are orange, red, and dark red, respectively.

Table 3. Composition of the selected SOM in terms of cluster type, as a fraction of the total number of map clusters.

Map Clusters	Empty [%]	Mixed [%]	LV1 [%]	LV2 [%]	LV3 [%]	LV4 [%]	LV5 [%]
3822	31.21	18.81	10.52	9.26	11.22	12.51	6.46

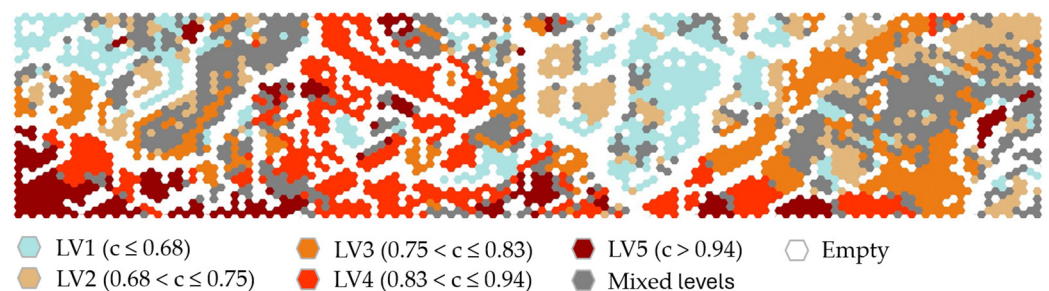


Figure 5. Selected SOM with 3822 clusters: 147 columns by 26 rows. The clusters are colored depending on their composition in terms of risk levels: LV1 in light blue, LV2 in light orange, LV3 in orange, LV4 in red, and LV5 in dark red. The clusters with mixed levels are grey, and the empty clusters are white.

The mixed clusters are colored grey, while the empty clusters are colored white. For the selected SOM, Table 4 reports the mixed cluster type as a fraction of the total number of mixed clusters (719), defined based on the number of different risk levels contained inside. About 90% of the mixed clusters mixed up only two risk levels. The quality of the operational space clustering can be assessed by the fraction of mixed clusters, which is approximately 18%. The clustering ability to identify risk levels was evaluated on the test set using standard classification metrics, as detailed in the following section.

Table 4. Mixed cluster composition, in terms of risk levels, for the selected SOM (3822 clusters). The composition was evaluated as a fraction of the total mixed ones.

Mixed Clusters	Two Levels [%]	Three Levels [%]	Four Levels [%]	Five Levels [%]
719	89.89	14.88	2.09	0.14

4.2. Classification Performance

To evaluate the performance of the selected SOM as an overload risk detector, both the training and test sets were projected on the map, and the expertise of the map as a risk level classifier was evaluated. To compute the classification performance, at each projected sample, the risk level of its BMU was assigned. Thus, the proposed criterion does not assign any risk level to the samples projected into empty clusters, while an unambiguous risk level label must be defined for the mixed clusters. In this paper, a criterion is proposed to assign a risk level to each mixed cluster, which consists of either assuming the most frequent among the risk levels or their mean value when the cluster shows a balanced risk level composition. By applying the proposed criterion, 90.82% of the mixed cluster is labeled with the already defined risk levels, while the remaining 9.18% is assigned to four new intermediate levels, i.e., LV1.5, LV2.5, LV3.5, and LV4.5, resulting from averaging the existing ones. Table 5 reports the SOM composition after assigning the risk level to the mixed clusters, also including the details on the cluster fraction belonging to each new intermediate level. On the SOM, most of the clusters are recolored with the five color shades (from light blue to dark red) used in Figure 4 to identify LV1, LV2, LV3, LV4, and LV5, while the ones labeled with the intermediate levels are left in grey, as shown in Figure 6.

Table 5. SOM composition after the assignment of a risk level to mixed clusters, for the SOM with 3822 clusters. An intermediate level is defined between each pair of contiguous risk levels defined in Table 2.

Map Clusters	Empty [%]	Mix [%]	LV1 [%]	LV1.5 [%]	LV2 [%]	LV2.5 [%]	LV3 [%]	LV3.5 [%]	LV4 [%]	LV4.5 [%]	LV5 [%]
3822	31.21	0.00	13.74	0.52	15.04	0.60	15.80	0.29	14.89	0.31	7.59

The 2D map, encoded as shown in Figure 6, was used to assess the classification performance by comparing the detected risk level with the target one. Table 6 reports the risk level classification performance in terms of correct detections, mistakes, and unknowns, as a fraction of the total number of samples for both the training and test sets. A classification is correct when the SOM-detected risk level matches the target one; otherwise, it is considered a mistake. It is unknown when the sample is assigned to an empty cluster.

The confusion matrices reported in Figure 7 allow us to visualize the classification performance, detailing, for each target risk level, the samples classified in each output level, as a fraction of the total number of samples with the considered risk level [34]. Each row index represents the target risk level (from 1 to 5 with step 1), while the column index

represents the SOM-detected risk level (from 1 to 5 with step 0.5); the last column reports the sample fraction projected in empty clusters for each risk level. The matrix elements with equal column and row indexes (in green) report the correct classifications, while the remaining cells report the wrong classifications.

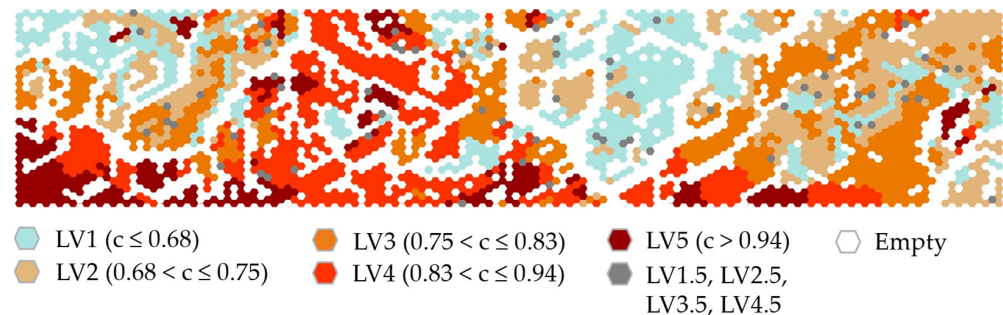


Figure 6. Selected SOM after mixed cluster interpretation. Most of the mixed clusters are recolored following the color code defined in Figure 4. The clusters with intermediate risk levels are grey, and the empty clusters are white.

Table 6. Sample classification attained by assigning to each sample the risk level of the BMU.

Classification	Training [%]	Test Set [%]
Correct	89.48	86.72
Wrong	10.52	12.17
Unknown	0	1.11

About 1% of the test samples at each risk level are projected into the empty clusters, resulting in 1.11% of the total number of test samples. For continuous overload risk monitoring, a risk level also needs to be assigned to the test samples projected into the empty clusters. To this aim, when a sample is projected into the empty cluster, the risk level of the second BMU is assigned in place of the first one. The proposed criterion allows for reducing the number of samples with an unknown risk level to 0.11% while improving the overall correct classification to 87.52%. Figure 8 reports the updated confusion matrix for the test set, where the performance enhancement attained for each risk level is shown.

Figure 9 reports the level error distributions, evaluated as a difference between the target risk level and the SOM output. The level error ranges between -4 and 4 , and its distribution allows for a global evaluation of the classification performance. Indeed, it shows that the SOM performs very well on the training set, with 89.48% correct classifications of the five risk levels: most of the misclassifications are because of adjacent classes mix up (8.87%); only for 1.75% of the samples, the absolute value of the level error is larger than one. On the test set, the SOM shows a good generalization capability, with 87.52% of correct classifications of the five risk levels. For only 1.91% of the samples, the absolute value of the level error is larger than one, while 10.46% of the samples are classified as belonging to adjacent risk levels. Thanks to the criterion adopted to assign a risk level to empty clusters, only 0.11% of the samples remain unclassified.

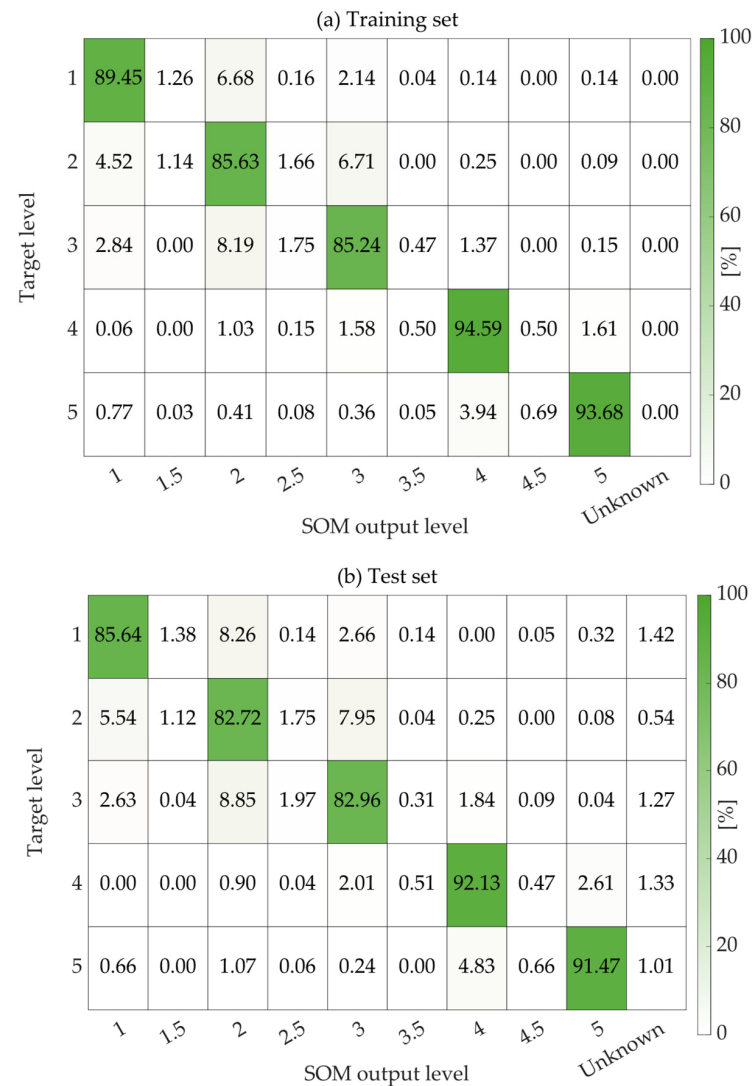


Figure 7. Confusion matrices for the training set (a) and test set (b). Each row shows the SOM-detected risk levels (column labels) for each target level (row label) as a percentage of samples belonging to the same target level.

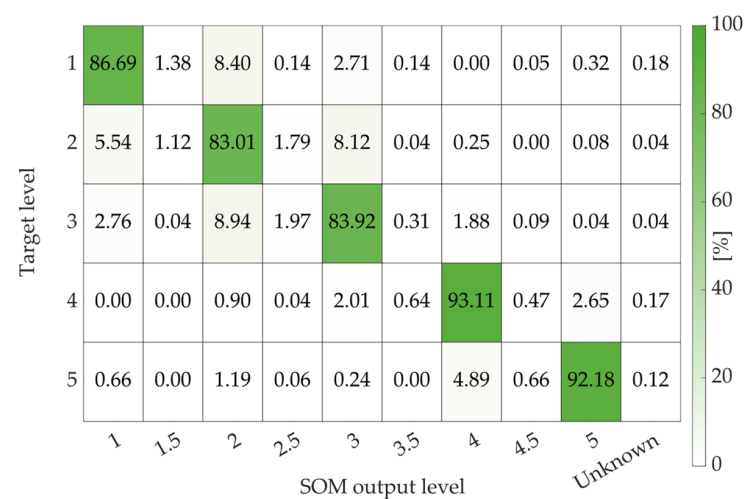


Figure 8. Confusion matrix for the test set after applying the criterion for interpreting the projection into empty clusters.

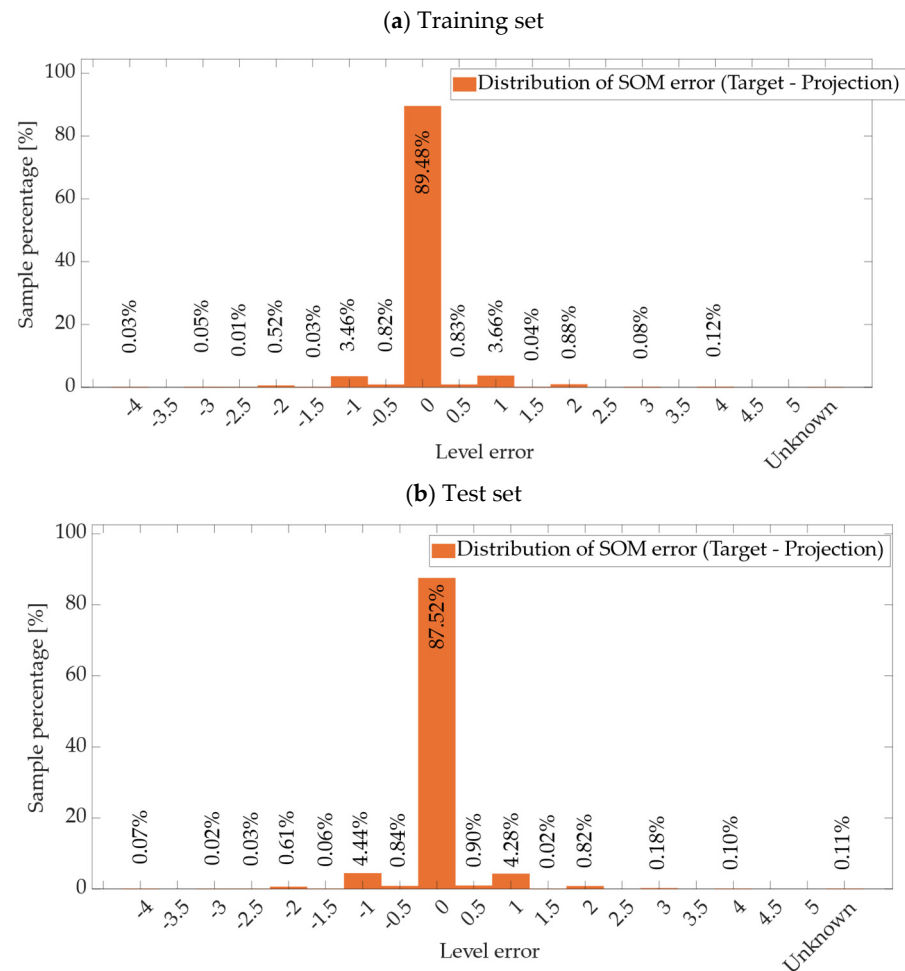


Figure 9. Level error distribution for (a) training and (b) test set. The level error is evaluated as the difference between the target and the SOM-detected risk level.

5. Pulse Tracking

The projection of a pulse on the SOM can be visualized on the 2D map, allowing us to display the temporal evolution of the operating point across the macro-clusters with different risk levels. A pulse trajectory can be obtained by tracking the projection of the test samples onto the map over time. Each projection involves identifying the Best Matching Unit (BMU), i.e., the cluster with the closest prototype vector. On the map, each test sample is represented as a dot positioned on the best-matching cluster. As an example, Figure 10 reports the trajectory of the eXperiment Program (XP) 20171205.026, which evolves in the right bottom corner of the selected SOM (see Figure 6). The yellow and the black dots represent the starting and the ending points of the trajectory, respectively. The shot starts in an LV2 region (light orange), passing through the LV3 (orange) and LV4 (red) regions and, finally, ending in an LV5 cluster (dark red).

Figure 11 compares the criticality behavior with respect to the SOM output for the experiment program 20171205.026. Subplot (a) reports the target criticality (in red) with respect to the target risk levels (in black). Subplot (b) reports the SOM output (in black), and subplot (c) reports the classification error (in red), evaluated as the difference between the target and the SOM output. The criticality monotonically increases from 0.770 to 1.026 in the first part of the pulse and then monotonically decreases to 0.985. The SOM clearly identified the three levels of risk defined in the criticality range; indeed, the error signal is zero. The SOM output is not monotonic at all; some spikes appear mainly in the transition between two risk levels due to jumps of the trajectory across the boundary delimiting two

macro-clusters with adjacent risk levels. As an example, Figure 12a shows a zoomed-in view of the pulse trajectory reported in Figure 10; only the trajectory of five samples around the transition time of 14.35 s is shown. It can be observed that the operating point oscillates between five adjacent clusters, with three clusters corresponding to LV4 and the remaining two to LV5.

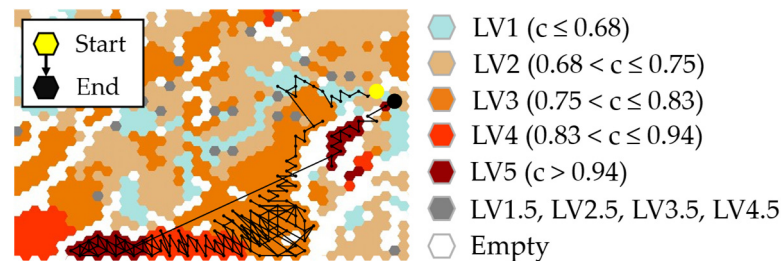


Figure 10. Projection of XP 20171205.26 on the selected SOM 2D map. The pulse trajectory is depicted by the black line, with the starting and ending points of the pulse highlighted by a yellow and a black point, respectively.

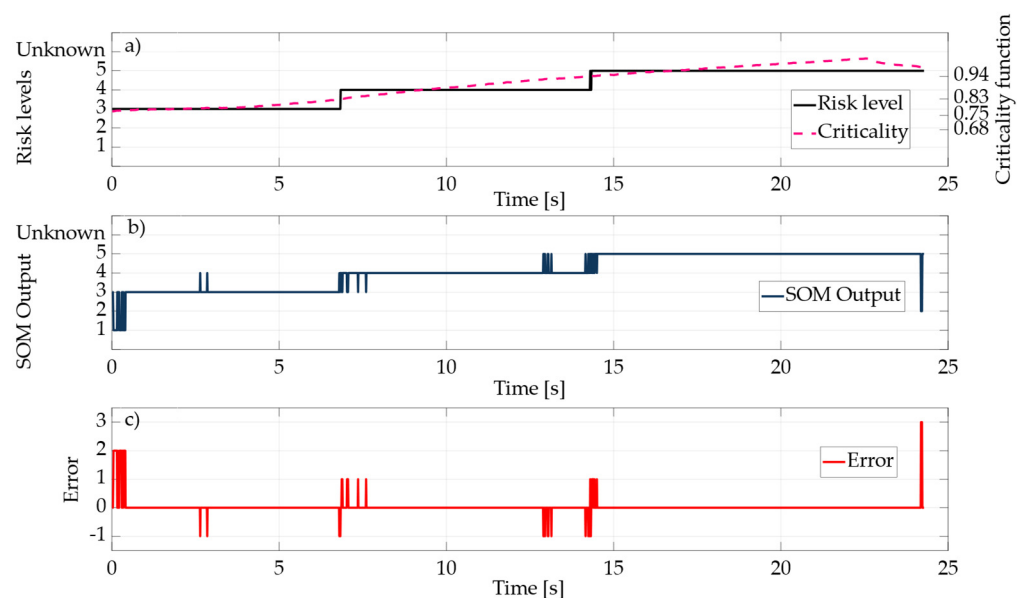


Figure 11. XP 20171205.026. (a) Criticality (red dashed line) compared to the target risk level behavior (black line). (b) SOM output (black line). (c) Classification error as the difference between the target risk level and the SOM output (in red).

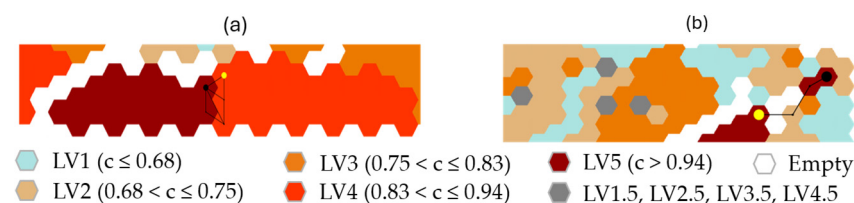


Figure 12. XP 20171205.026. (a) The pulse trajectory evolution around the risk level transition occurring at 14.5s. The fluctuation in the SOM output between LV4 and LV5 is due to the trajectory evolution among neighboring clusters with adjacent risk levels. (b) The pulse trajectory evolution around 24.2 s. The fluctuation in the SOM output between LV2 and LV4 is due to the trajectory evolution among neighboring clusters with non-adjacent risk levels.

Gaps between non-adjacent risk levels are observable at the beginning and the end of the experiment. At the end, a jump between LV5 and LV2 and then a jump back to LV5 occurs because the trajectory evolves over time across a region where a mix of clusters

with low risk levels are present, as shown in Figure 12b, where the last 360 ms of the trajectory shown in Figure 10 is reported. The root of the errors can be investigated by analyzing the component planes, which show the distribution of the input features on the map. This allows us to straight link the risk levels defined on the map to the behavior of the inputs. Thus, the risk level assigned to the test points by the trajectory evolution in the map can be interpreted. As an example, the region in which the last 360 ms of the XP 20171205.026 trajectory evolves is characterized by low values of both the total input power (P_{INP}) and diamagnetic energy (W_{DIA}), as clearly highlighted by the magenta square on the component planes shown in Figure 13a and Figure 13b, respectively. This behavior can be representative of both a low overload risk condition and the starting and/or ending phases of the experiments. This explains why both the map regions are characterized by LV2 and the last part of the experimental trajectory is projected right there. Indeed, Figure 12b visualizes the part of the experiment following the switch-off of the additional heating systems. This example shows the potential in terms of the result interpretability of the SOM output.

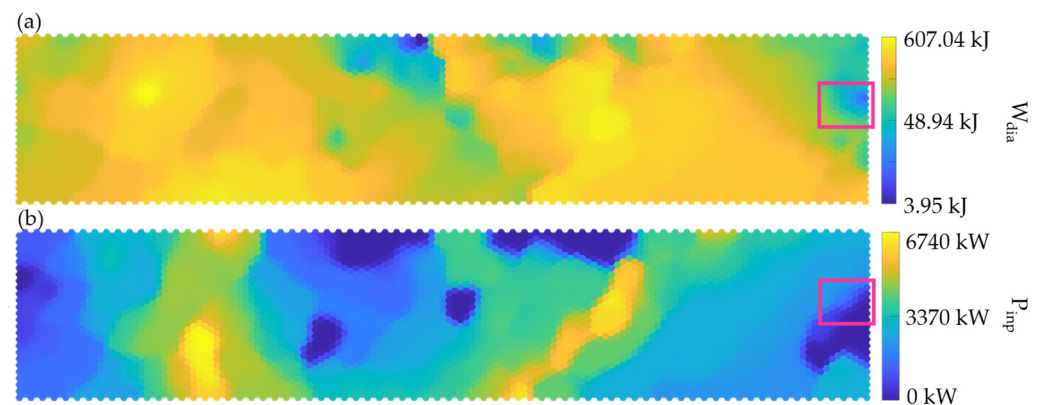


Figure 13. (a) W_{DIA} component plane of SOM. (b) P_{INP} component plane of SOM. The pulse trajectory evolution around 24.2 s, reported in Figure 12b, is highlighted by the magenta square. This region is characterized by the low value of both W_{DIA} and P_{INP} .

Figure 14 reports the trajectory of XP 20171205.014, which evolves in the middle-top region of the selected SOM. Figure 15 compares the criticality behavior with respect to the SOM output. Subplot (a) reports the target criticality (in red) compared to the target risk levels (in black). Subplot (b) reports the SOM output (in black), and subplot (c) reports the classification error (in red), evaluated as the difference between the target and the SOM output. For this experiment, the criticality always remains below 0.766, starting from 0.596 and then coming back to 0.696 at the end of the shot. In this case, the risk levels from 1 to 3 are experienced. The shot trajectory evolves in a small region, well delimited from the rest of the map by a white cluster border. As for XP 20171205.026, the classification errors occur just behind the transition between the two risk levels.

Figure 16 reports the trajectory evolution around the transition occurring at 1.45 s, between LV2 and LV3. The fluctuations in the SOM output are due to a repeated jump of the trajectory between a light-orange cluster and an orange cluster.

Generally speaking, the spikes in the SOM output and the oscillation at the transition times do not correspond to fast changes in the input variables but are caused by the quantization of the target, which results in a fragmentation of the 2D map into several adjacent macro-clusters.

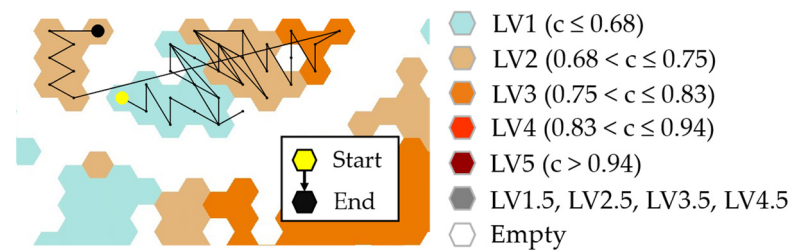


Figure 14. Projection of XP 20171205.14 on the selected SOM. The pulse trajectory is depicted by the black line. The pulse starting and ending points are highlighted by a yellow and a black point, respectively.

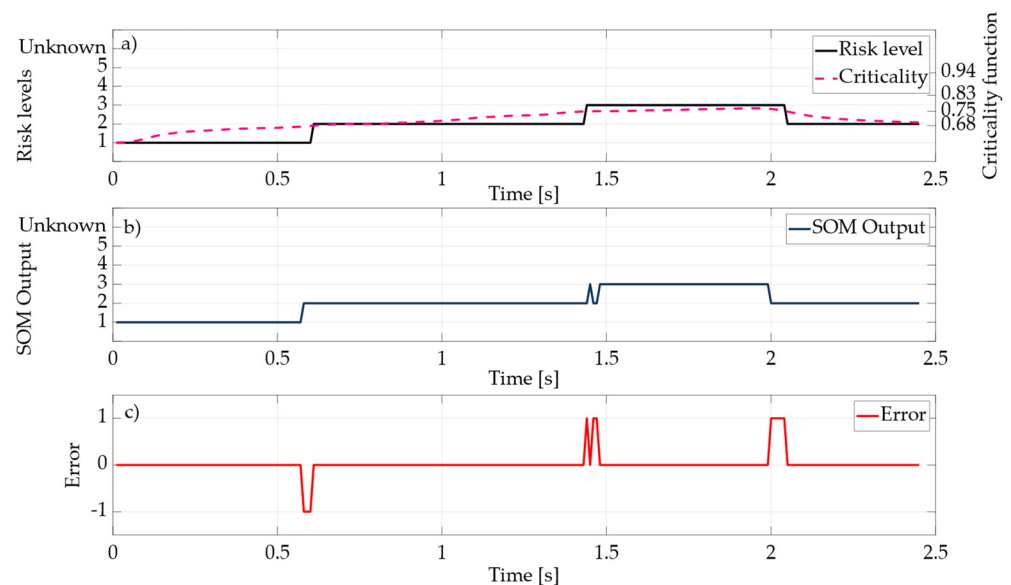


Figure 15. XP 20171205.014. (a) Criticality (red dashed line) compared to the target risk level behavior (black line). (b) SOM output (black line). (c) Classification error as the difference between the target risk level and the SOM output (in red).

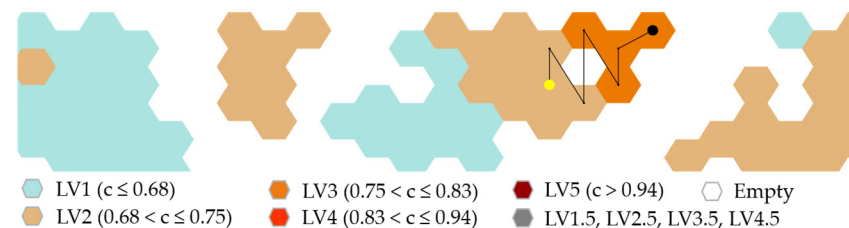


Figure 16. XP 20171205.014. The pulse trajectory evolution around the transition level that occurs at 1.45 s. The fluctuation in the SOM output between LV3 and LV4 is due to the trajectory evolution among neighboring clusters with adjacent risk levels.

6. Conclusions

In this paper, a method for mapping the W7-X operational space onto a 2D grid by means of Self-Organizing Maps (SOMs) was proposed. The peculiarities of SOMs in preserving the topological structure of the operational input space were considered, while representing the data onto a discrete and interpretable 2D grid. This enabled easy visual exploration of high-dimensional data by a 2D map, allowing us to track the experiments and monitor the overload risk during W7-X operation. At this preliminary stage, the so-called criticality was quantized into five risk levels in order to test the ability of the SOM in clustering experimental conditions with similar thermal load conditions.

The SOM proved to be an effective tool for detecting the overload risk, with good performance on both the training and test sets. It correctly identified the overload risk level for 87.52% of the samples, with less than 1% of the samples left unclassified. The most frequent error on the test set, occurring in 10.46% of the cases, was assigning a risk level to the sample that is adjacent to the target one. Only 1.91% of the samples were misclassified with a non-adjacent risk level. These results are promising for the application of SOMs in the real-time detection system at W7-X, where accurate and timely predictions are crucial to prevent damage to the PFCs.

The analysis of the results allowed us to highlight the prospects of future work. The identified common errors suggest areas for improving the potential of this tool. The next step will focus on limiting the effect of the level transitions occurring between two contiguous risk levels. For this purpose, macro-regions with an assigned overload will be defined to limit the output oscillation among contiguous clusters in the same map region. In addition, the risk level definition could be optimized based on the control actions to be undertaken to limit inaccurate responses of the control system when oscillation among adjacent classes occurs. Meanwhile, an assertion time can be optimized to reduce the oscillation output among non-adjacent risk levels, in accordance with the overload detection needs, reducing the harsh risk level misclassification. Implementing these actions could significantly enhance the reliability of the overload detection system at W7-X, contributing to the safe and efficient operation of the stellarator.

Author Contributions: Conceptualization, G.S.; Software, E.C.; Formal analysis, G.S. and E.C.; Investigation, E.A., B.C. and A.F.; Resources, M.J.; Data curation, E.A., W7-X team, Y.G., B.J. and A.P.S.; Writing—original draft, G.S. and F.P.; Writing—review & editing, B.C. and A.F.; Supervision, G.S. and F.P.; Project administration, M.J.; Funding acquisition, F.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by EUROfusion Consortium, funded by the European Union via the Euratom Research and Training Programme, Grant Agreement No. 101052200-242 EUROfusion.

Data Availability Statement: The data are contained within this article.

Acknowledgments: This work was carried out within the framework of the EUROfusion Consortium. The views and opinions expressed are, however, those of the authors only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them. See Grulke et al., 2024. (<https://doi.org/10.1088/1741-4326/ad2f4d>) for the W7-X Team.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Spitzer, L., Jr. The Stellarator Concept. *Phys. Fluids* **1958**, *1*, 253–264. [[CrossRef](#)]
2. Pedersen, T.S.; König, R.; Krychowiak, M.; Jakubowski, M.; Baldzuhn, J.; Bozhnikov, S.; Fuchert, G.; Langenberg, A.; Niemann, H.; Zhang, D.; et al. First Results from Divertor Operation in Wendelstein 7-X. *Plasma Phys. Control. Fusion* **2018**, *61*, 014035. [[CrossRef](#)]
3. Wolf, R.C.; Alonso, A.; Äkäslompolo, S.; Baldzuhn, J.; Beurskens, M.; Beidler, C.D.; Biedermann, C.; Bosch, H.-S.; Bozhnikov, S.; Brakel, R.; et al. Performance of Wendelstein 7-X Stellarator Plasmas during the First Divertor Operation Phase. *Phys. Plasmas* **2019**, *26*, 082504. [[CrossRef](#)]
4. Grulke, O.; Albert, C.; Belloso, J.A.A.; Aleynikova, P.; Aleynikova, K.; Alonso, A.; Anda, G.; Andreeva, T.; Arvanitou, M.; Ascasibar, E.; et al. Overview of the First Wendelstein 7-X Long Pulse Campaign with Fully Water-Cooled Plasma Facing Components. *Nucl. Fusion* **2024**, *64*, 112002. [[CrossRef](#)]
5. ITER Physics Expert Group on Divertor; ITER Physics Expert Group on Divertor Modelling and Database; ITER Physics Basis Editors. Chapter 4: Power and Particle Control. *Nucl. Fusion* **1999**, *39*, 2391. [[CrossRef](#)]

6. Loarte, A.; Lipschultz, B.; Kukushkin, A.S.; Matthews, G.F.; Stangeby, P.C.; Asakura, N.; Counsell, G.F.; Federici, G.; Kallenbach, A.; Krieger, K.; et al. Chapter 4: Power and Particle Control. *Nucl. Fusion* **2007**, *47*, S203. [\[CrossRef\]](#)
7. Bongiovì, G.; Böckenhoff, D.; Carls, A.; Endler, M.; Fellingner, J.; Geiger, J. Assessment of Potential Heat Flux Overload of Target and First Wall Components in Wendelstein 7-X Finite-Beta Magnetic Configurations and Choice of Locations for Temperature Monitoring. *Fusion Eng. Des.* **2020**, *161*, 111902. [\[CrossRef\]](#)
8. Puig Sitjes, A.; Jakubowski, M.; Naujoks, D.; Gao, Y.; Drewelow, P.; Niemann, H.; Fellingner, J.; Moncada, V.; Pisano, F.; Belafdil, C.; et al. Real-Time Detection of Overloads on the Plasma-Facing Components of Wendelstein 7-X. *Appl. Sci.* **2021**, *11*, 11969. [\[CrossRef\]](#)
9. Lukacs, M.; Williams, L.G. Nuclear Safety Issues for Fusion Power Plants. *Fusion Eng. Des.* **2020**, *150*, 111377. [\[CrossRef\]](#)
10. Guo, Z. Nuclear Fusion: Overview of Challenges and Recent Progress. *NHSJS Reports* **2024**. Available online: <https://nhsjs.com/2024/nuclear-fusion-overview-of-challenges-and-recent-progress/> (accessed on 10 June 2025).
11. Jabłoński, B.; Puig Sitjes, A.; Makowski, D.; Jakubowski, M.; Gao, Y.; Fischer, S.; Winter, A. Implementation and Performance Evaluation of the Real-Time Algorithms for Wendelstein 7-X Divertor Protection System for OP2.1. *Fusion Eng. Des.* **2023**, *190*, 113524. [\[CrossRef\]](#)
12. Jakubowski, M.; Drewelow, P.; Fellingner, J.; Puig Sitjes, A.; Wurden, G.; Ali, A.; Biedermann, C.; Cannas, B.; Chauvin, D.; Gamradt, M.; et al. Infrared Imaging Systems for Wall Protection in the W7-X Stellarator (Invited). *Rev. Sci. Instrum.* **2018**, *89*, 10E116. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Fellingner, J.; Lippmann, U.; Greve, H.; Alhashimi, M.; Schülke, M.; Äkaslompolo, S.; Drewelow, P.; Jakubowski, M.; König, R.; Lorenz, A. Design of Endoscopes for Monitoring Water-Cooled Divertor in W7-X. *Fusion Eng. Des.* **2020**, *158*, 111841. [\[CrossRef\]](#)
14. Elgendy, M. *Deep Learning for Vision Systems*; Manning Pubns Co.: Shelter Island, NY, USA, 2020; ISBN 978-1-61729-619-2.
15. Bengio, Y. Learning Deep Architectures for AI. *Found. Trends Mach. Learn.* **2009**, *2*, 1–127. [\[CrossRef\]](#)
16. Grelier, E.; Mitteau, R.; Moncada, V. Deep Learning and Image Processing for the Automated Analysis of Thermal Events on the First Wall and Divertor of Fusion Reactors. *Plasma Phys. Control. Fusion* **2022**, *64*, 104010. [\[CrossRef\]](#)
17. Kohonen, T. Self-Organized Formation of Topologically Correct Feature Maps. *Biol. Cybern.* **1982**, *43*, 59–69. [\[CrossRef\]](#)
18. Roweis, S.T.; Saul, L.K. Nonlinear Dimensionality Reduction by Locally Linear Embedding. *Science* **2000**, *290*, 2323–2326. [\[CrossRef\]](#)
19. Aledda, R.; Cannas, B.; Fanni, A.; Pau, A.; Sias, G. Improvements in Disruption Prediction at ASDEX Upgrade. *Fusion Eng. Des.* **2015**, *96–97*, 698–702. [\[CrossRef\]](#)
20. Aymerich, E.; Fanni, A.; Pisano, F.; Sias, G.; Cannas, B.; JET Contributors; WPTE Team. A Self-Organised Partition of the High Dimensional Plasma Parameter Space for Plasma Disruption Prediction. *Nucl. Fusion* **2024**, *64*, 106063. [\[CrossRef\]](#)
21. Pau, A.; Fanni, A.; Carcangiu, S.; Cannas, B.; Sias, G.; Murari, A.; Rimini, F.; JET Contributors. A Machine Learning Approach Based on Generative Topographic Mapping for Disruption Prevention and Avoidance at JET. *Nucl. Fusion* **2019**, *59*, 106017. [\[CrossRef\]](#)
22. Cannas, B.; Fanni, A.; Murari, A.; Pau, A.; Sias, G.; JET EFDA Contributors. Overview of Manifold Learning Techniques for the Investigation of Disruptions on JET. *Plasma Phys. Control. Fusion* **2014**, *56*, 114005. [\[CrossRef\]](#)
23. Cannas, B.; Fanni, A.; Sonato, P.; Zedda, M.K.; JET-EFDA Contributors. A Prediction Tool for Real-Time Application in the Disruption Protection System at JET. *Nucl. Fusion* **2007**, *47*, 1559–1569. [\[CrossRef\]](#)
24. Cannas, B.; Fanni, A.; Pautasso, G.; Sias, G.; Sonato, P. An Adaptive Real-Time Disruption Predictor for ASDEX Upgrade. *Nucl. Fusion* **2010**, *50*, 075004. [\[CrossRef\]](#)
25. Scarpari, M.; Minucci, S.; Sias, G.; Lombroni, R.; Buxton, P.F.; Romanelli, M.; Calabrò, G. ST40 Electromagnetic Predictive Studies Supported by Machine Learning Applied to Experimental Database. *Sci. Rep.* **2024**, *14*, 27074. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Gao, Y.; Geiger, J.; Jakubowski, M.W.; Drewelow, P.; Endler, M.; Rahbarnia, K.; Bozhnikov, S.; Otte, M.; Suzuki, Y.; Feng, Y.; et al. Effects of Toroidal Plasma Current on Divertor Power Depositions on Wendelstein 7-X. *Nucl. Fusion* **2019**, *59*, 106015. [\[CrossRef\]](#)
27. Pisano, F.; Cannas, B.; Fanni, A.; Sias, G.; Gao, Y.; Jakubowski, M.; Niemann, H.; Sitjes, A.P. Learning Control Coil Currents from Heat-Flux Images Using Convolutional Neural Networks at Wendelstein 7-X. *Plasma Phys. Control. Fusion* **2020**, *63*, 025009. [\[CrossRef\]](#)
28. Gao, Y.; Geiger, J.; Jakubowski, M.W.; Pandey, A.; Bozhnikov, S.; Feng, Y.; Endler, M.; Baldzuhn, J.; Perseo, V.; Kremeyer, T.; et al. Heat and Particle Exhaust in High-Performance Plasmas in Wendelstein 7-X. *Nucl. Fusion* **2024**, *64*, 076060. [\[CrossRef\]](#)
29. Kohonen, T. The Self-Organizing Map. *Proc. IEEE* **1990**, *78*, 1464–1480. [\[CrossRef\]](#)
30. Vesanto, J.; Himberg, J.; Alhoniemi, E.; Parhankangas, J. Self-Organizing Map in Matlab: The SOM Toolbox. In Proceedings of the Matlab DSP Conference 1999, Espoo, Finland, 16–17 November 1999; pp. 35–40.
31. Kohonen, T. *Self-Organizing Maps*; Springer Series in Information Sciences; Springer: Berlin/Heidelberg, Germany, 2001; Volume 30, ISBN 978-3-540-67921-9.
32. Miljković, D. Brief Review of Self-Organizing Maps. In Proceedings of the 2017 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, 22–26 May 2017; pp. 1061–1066.

33. SOM Toolbox Homepage. Available online: <http://www.cis.hut.fi/somtoolbox/> (accessed on 27 May 2025).
34. Tharwat, A. Classification Assessment Methods. *Appl. Comput. Inform.* **2021**, *17*, 168–192. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.