# Preliminary Study of Disruption Prediction with Machine Learning: from Solar Plasma to Tokamaks

Carlos E. Falandes[1*], Reinaldo R. Rosa[1], Salatiel A. A. Jordão[1], Rubens A. Sautter[2], Luan O. Barauna[1], Jiro Kawamura[3], Pablo A. Medina Sanchez[4], Juan A. Valdivia[5], Daniel A. S. Mendes[1]

[1]National Institute for Space Research (INPE), S.J. dos Campos, São Paulo, Brazil
[2]Federal University of Technology (UTFPR), Brazil
[3]International Thermonuclear Experimental Reactor (ITER), France
[4]University of Los Andes, Colombia
[5]Universidad de Chile, Santiago, Chile

*E-mail: carlos.falandes@inpe.br

Keywords:

30 / 06 / 2025

## Abstract

Plasma disruptions represent critical risks both in magnetically confined fusion devices and in astrophysical environments such as the solar corona. Despite their different scales and configurations, these systems share a nonlinear magnetohydrodynamic behavior that can lead to instabilities and abrupt loss of confinement or energy release, respectively. Motivated by the challenge of effectively treating disruptions and anomalies in plasma dynamics, we present criteria and a methodology to simulate, detect and predict disruptions in plasma dynamics in different contexts. The methodology, which we call here AEDeep (Anomaly Explorer with Deep Learning) is based on the simulation and prediction of extreme fluctuation patterns generated by multiplicative cascades, compatible with the dynamics of plasma instability and prediction models based on recurrent neural networks (RNN). In this preliminary version, to treat disruptions in time series, we use the multifractal p-model and an LSTM (Long-Short Term Memory) model. We validate the methodology, in its simplest version, using data from solar EUV images observed by SDO/AIA and also from tokamak disruptions obtained by the JOREK simulator. For both cases, we show results evidencing the detection of precursors with good accuracy, early and precise. We also outline future directions for the implementation of the second strategy based on LSTM autoencoder, which will enable cutting-edge AI embedded in CubeSats for space weather forecasting and, in the case of tokamaks, the integration with advanced quantum sensing materials implanted in their magnets. These advances point to broader applications of AEDeep in the detection of spatiotemporal anomalies in physics and engineering.

## 1. Introduction

In recent years, the analysis of extreme fluctuations in plasmas has attracted increasing attention due to its relevance both for understanding astrophysical phenomena, such as X-class solar flares, and for advancing magnetic confinement technologies in tokamaks, aimed at generating nuclear fusion energy. These fluctuations, which represent atypical activities within a system, indicate behaviors that exceed expected intervals and are often interpreted as signals of unpredictability, as plasmas are dynamic systems highly susceptible to instabilities, whose formation and evolution are closely related to local physical conditions such as density, temperature, and magnetic field configuration.

Among these instabilities, disruptions stand out due to their significant impact on plasma stability, as a minimal disturbance in the magnetic field can cause nonlinear amplification, compromising the global stability of the system. In the astrophysical context, these instabilities trigger extreme events, such as X-class flares, where magnetic reconnection generates a catastrophic release of energy [1]. In confinement devices, such as tokamaks, instabilities pose significant challenges to plasma control, potentially resulting in substantial energy losses and irreversible damage to the reactor by releasing large quantities of energetic particles against its walls [2].

Since the formalization of this phenomenological concept, various approaches have been proposed to better understand extreme events (XEs). Despite the diversity of perspectives across different fields, there is still no unified theory of extreme fluctuations [3]. The two main classes of extreme events considered here are endogenous and exogenous extremes. The distinction between these two classes was introduced by Sornette, Deschtres, Gilbert, and Ageon (SDGA) [4], whose framework differentiates these fluctuations based on the scaling properties of their temporal evolution. In this context, dissipative processes caused by internal instabilities (endogeneity) and/or external disturbances (exogeneity) produce distinct fluctuation patterns in a time series. A crucial aspect of this approach is that statistical and spectral properties can be used to differentiate between endogenous and exogenous processes.

Among the disruptions most relevant to fusion devices are Edge Localized Modes (ELMs), nearly periodic phenomena that release energy and particles through filamentary structures characterized by high anisotropy. These events, particularly critical in H-mode (high-confinement mode) operation regimes, exhibit a well-defined filamentary morphology, with elongated toroidal structures that propagate radially into the scrape-off layer (SOL). According to A. Kirk et al. [5], these filaments, which are generated on a 100 $\mu s$ time scale and erupt from the outboard side before reconnecting to the plasma, providing strong evidence that ELMs are associated with filament like structures. This observation aligns with a theoretical model based on the “ballooning”' instability, which has applications in both solar and tokamak physics. Additionally, multiple pieces of evidence have highlighted the important role of turbulence in physical processes in solar eruptions, from particle acceleration to the

suppression of conductive cooling. Xie et al. [6] have shown that radio observations of density variation in solar wind have established a Kolmogorov-like spectrum for solar wind density disturbance. Using the Atmospheric Imaging Assembly (SDO/AIA) onboard the Solar Dynamics Observatory (SDO), they computed structure functions from EUV emission measurements, providing a new way to study turbulence in the solar atmosphere. Their findings suggest that the structure functions of intensity variations can be a proxy for turbulent fields in the plasma sheet, providing valuable insights into turbulence behavior. Such periodic eruptions, with their characteristic structure, can also be predicted using machine learning models, as demonstrated in the work of J.X. Zhu et al. [7], which uses deep neural networks to predict disruptions in tokamaks, aiming to anticipate their occurrence with enough time for mitigation measures to be applied.

The detection of anomalies in time series, such as those observed in complex dynamic systems like plasma, can be enhanced using models based on Long Short-Term Memory (LSTM) networks and autoencoders. These models have great potential to capture complex temporal relationships and detect subtle or anticipatory patterns of abrupt events, such as disruptions [8]. In particular, Akbarian et al. [9] proposed the use of the Autoencoder with LSTM (AE-LSTM) algorithm, combining the dimensionality reduction benefits of autoencoders with the LSTM's ability to model temporal dependencies, showing a significant improvement in anomaly detection compared to the traditional DNN-LSTM model. This unsupervised approach offers a promising strategy for early disruption prediction, with direct applications in the analysis of complex time-series data, such as plasma data, where early detection is crucial to prevent adverse events.

However, one of the biggest challenges for training machine learning models is the scarcity of high-quality experimental data. To overcome this limitation, the use of synthetic data generated by numerical simulations based on physical models is proposed. Physically, an XE can be represented as an energy transfer process from low to high frequencies, governed by a probability $p$, where energy is distributed over time intervals with probabilities $p$ and $(1 - p)$ [10]. In a balanced energy cascade, where $p = 1/2$, the distribution is uniform, resulting in non-extreme signals. In contrast, when the distribution is unbalanced, as in the extreme case $p = 0$, energy concentrates in one region, generating an XE. This process is described by the P-Model, which is closely related to the multifractal versus monofractal paradigm of analysis, determining the distribution of the singularity spectrum. Recent studies indicate that the cascade process is not uniformly distributed [11], suggesting the absence of a universal law governing certain cascade phenomena.

Thus, this work proposes the use of LSTM networks to estimate the probability of anomalous events occurring in plasma over upcoming time steps in a time series. For model training, synthetic data generated by the P-model were used, chosen for their multifractal characteristics compatible with the dynamics of extreme plasma fluctuations. After training, the model was applied to data from the JOREK

simulator and images from the SDO, aiming to assess its generalization capability across different contexts of plasma disruptions.

## 2. Methods

### 2.1. Methodology for Generating Extreme Fluctuations and Predicting Disruptive Events

A The strategy adopted in this work was based on the use of the multiplicative cascade p-model to generate time series with extreme fluctuations, representing, to some extent, the main patterns of extreme events (XEs) observed in real physical systems. Although the p-model does not capture all possible dynamics, it is widely recognized as a canonical model for generating multifractal time series and is therefore suitable for the proposed methodology. Synthetic time series were generated using vectors with 1024 points, aiming to ensure statistically representative samples and to enable the observation of distinct patterns, primarily of exogenous origin.

In the next step, Extreme Value Theory (EVT) was applied in conjunction with the Weibull distribution to identify anomalies in the time series, following the methodology proposed by Pereira E. S. et al. [12]. Based on the points classified as extreme, an LSTM-based neural network model was trained to estimate the probability that upcoming values in the series would correspond to extreme events.

Finally, shorter time series were used to evaluate the model's ability to predict extremes in both the synthetic data generated by the p-model and in real and simulated data: solar images from the Solar Dynamics Observatory (SDO) and disruption simulations in tokamaks produced by the JOREK code. This step aimed to assess the generalization capacity of the trained model across different domains of plasma dynamics.

### 2.2. Multiplicative Cascade Patterns

The *p-model* [10, 13] enables the generation of non-homogeneous cascades that realistically replicate the fluctuations found in stochastic time series. Originally introduced by Meneveau and Sreenivasan [13] as a framework for modeling turbulence-like cascades, it offers a novel interpretation of kinetic energy dissipation during the vortex fragmentation process within the inertial range of fully developed turbulence. The model is based on a specific scheme of asymmetric energy redistribution.

The *p-model* approach for a non-homogeneous turbulent-like cascade was proposed by Meneveau and Sreenivasan [13]. It gives a new insight into the kinetic energy dissipation in the cascading process of eddies in the inertial range of a fully developed turbulence, and it is based on the special case of weighted transfer.

From a theoretical point of view, the *p-model* is a generalized form of a two-scale cantor set with a balanced distribution of length which shows multifractal

properties of one-dimensional sections of the dissipation field. The generalized form starts from the classical view of an eddy cascade before the inertial range of a fully developed turbulence, where a flux of energy ($E_K$) of size $L_K$ cascades down along the inertial range and eventually dissipates when it reaches the Kolmogorov length scale $\beta$. Hence, in a cascade step, each eddy of size $L_K$ is divided into two equal parts, $L_K/2$ expressed as $L_{K_1}$ and $L_{K_2}$ each. However, in each cascade step, the flux of energy is distributed, as a probability, unequally in fractions of $p_1$ and $p_2$, where $p_2 = 1 - p_1$ and $p_1 + p_2 = 1$. This process is iterated over fixed $p_1$ until the eddies reach the Kolmogorov scale β.

In terms of a multifractal approach, the spatial energy distribution of the *p-model* multiplicative cascade can be described by [13]:

$$\alpha = \frac{log_2 p_1 + (\omega - 1) log_2 p_2}{log_2 l_1 + (\omega - 1) log_2 l_2} \quad (1)$$

and

$$f(\alpha) = \frac{(\omega - 1) log_2 (\omega - 1) - \omega log_2 p_2}{log_2 l_1 + (\omega - 1) log_2 l_2} \quad (2)$$

where ω is a term related to the Stirling approach for the generalized Cantor set [13], and $f(\alpha)$ is the multifractal spectrum. For the case of equal-sized eddies ( $l_1 = l_2 = 1/2$), ω can be expressed in terms of α, $p_1$, and $p_2$:

$$\omega = \frac{log_2 (p_2) - log_2 p_1}{\alpha + log_2 (p_2)} \quad (3)$$

We can further simplify $f(\alpha)$ by expressing it explicitly in terms of ω, which yields:

$$f(\alpha) = \frac{\omega log_2 (\omega) - (\omega - 1) log_2 (\omega - 1)}{\omega} \quad (4)$$

Substituting equation 3 into the expression 4 yield an equation for α which is only dependent to$p_1$, $p_2$, and α. We can use this proxy to construct a time series that displays an equivalent multifractal spectrum, but now in the time domain.

Deviations from the homogeneous cascade (p1=p2=0.5) manifest as abrupt changes in the frequency and magnitude of the time series, referred to as extreme events $XE$. When the level of amplitudes increases significantly due to internal factors, then the extreme event are called endogenous ($XE_{endo}$). In contrast, it is

called an extreme exogenous event ($XE_{exo}$) when external energy transfer or abrupt dissipation is the main cause of $XE$. To generate time series $XE_{endo}$ and $XE_{exo}$, we will adopt the SDGA formalism [4]. Typical exogenous and endogenous processes can be obtained with the *p-model* for $0.18 < p_1 < 0.28$ and $0.32 < p_1 < 0.42$, respectively. Figures 1 and 2 illustrate examples of these two classes of time series generated with the p-model.

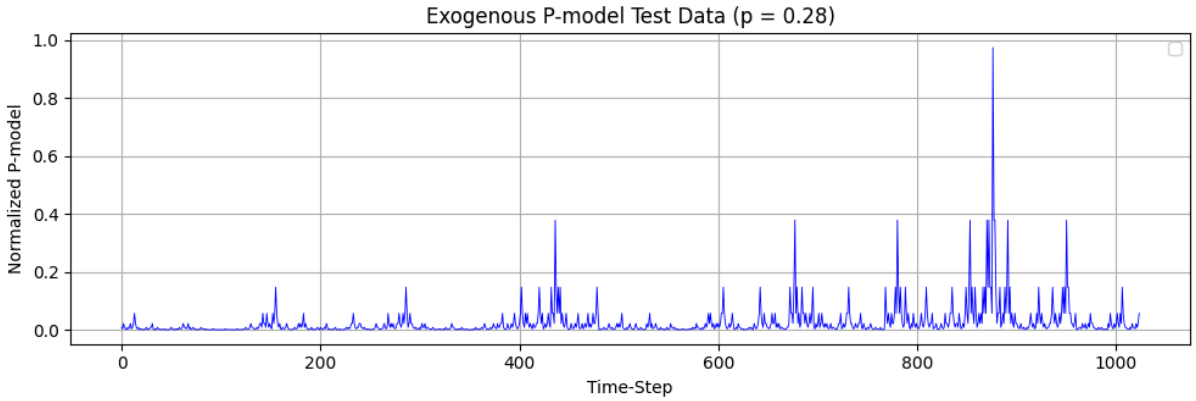


**Figure 1.** Time Series Simulated with the p-Model Showing Exogenous Extreme Events (XEexo), $p = 0.28$.

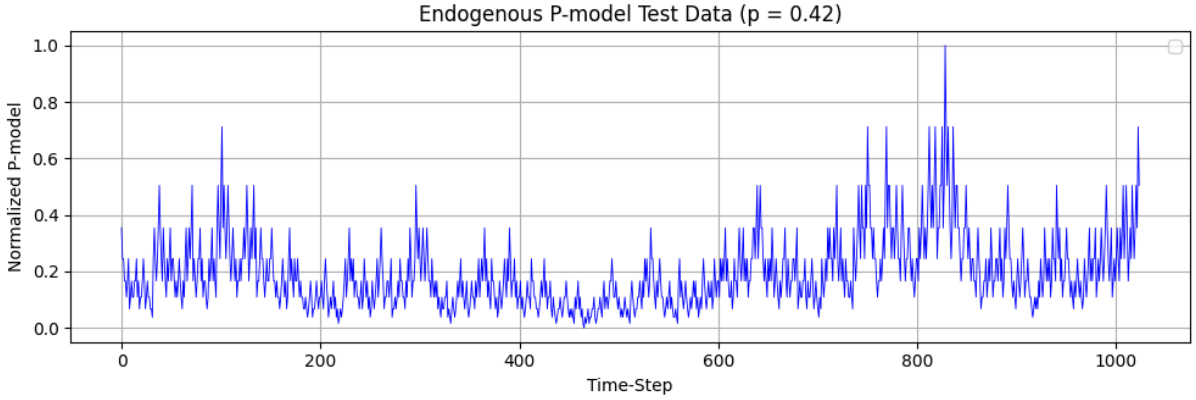


**Figure 2.** Time Series Simulated with the p-Model Showing Endogenous Extreme Events (XEendo), $p = 0.42$.

Although the *p-model* was originally inspired by turbulent eddy dissipation, its connection to classical turbulence metrics remains underexplored. Many applications rely on the Power Spectral Density $S(f)$, defined as:

$$S(f_j) = \frac{1}{f_j^{\beta}} \tag{5}$$

where, $f_j$ is a given frequency of the signal, and in a fully developed turbulent regime, the scaling exponent approaches Kolmogorov's $\beta = 5/3 \approx 1.667$. However, not all signals following a power-law spectrum exhibit extreme event statistics. This highlights the need for a model that explicitly incorporates the

power-law exponent as a tunable parameter, ensuring a more comprehensive description of dynamical behaviors.

### 2.3. Extreme Value Theory (EVT) and the Weibull Distribution

Extreme Value Theory (EVT) is a fundamental statistical approach for modeling the behavior of the extreme values of a random variable, being especially useful for anomaly detection in time series. This theory provides a coherent mathematical foundation for extrapolating the behavior of rare events beyond the observed range [14]. Moreover, EVT is widely used in applications where it is necessary to assess the risk of extreme events, offering a robust framework for modeling data in long-tail situations [15]. Its application is particularly relevant in contexts where the occurrence of outliers may represent critical phenomena, such as disruptions in plasmas or failures in dynamical systems.

In this context, the work of Pereira et al. [12] demonstrated the effectiveness of combining the Weibull distribution with EVT as a statistical tool to identify equipment failures and deviations in energy consumption patterns. Inspired by this approach, the present study employs EVT to detect anomalies in discrete time series using a two-parameter version of the Weibull distribution, specifically adjusted to model the distribution of extreme values. This approach was used to label anomalous events in the data generated by the P-model, providing an annotation base that was later used to train the LSTM network.

The cumulative distribution function (CDF) of the Weibull distribution, used to describe the probability that the random variable $x$ takes on a value less than or equal to $x$ , is given by Equation 6:

$$F(x;\lambda,k) = 1 - exp\left[-\left(\frac{x}{\lambda}\right)^{k}\right] \tag{6}$$

This model becomes particularly useful for anomaly detection in time series, where extreme values, characterized by significant deviations from the mean, are identified as potential anomalous events. In the detection algorithm, extreme values are identified using a sliding window of data and a moving average and standard deviation to define a threshold, as given by Equation 7:

$$threshold = \mu_{window} + threshold_{factor} \cdot \sigma_{window} \tag{7}$$

Extreme values are defined as those that exceed this threshold. From this selection of extreme values, the Weibull distribution is fitted using the maximum likelihood estimation (MLE) method, which estimates the scale $\lambda$ and shape $k$ parameters of the distribution to more accurately model the observed rare events.

The Weibull CDF is then used to compute the cumulative probability for new data points $x$. If this probability (i.e., the CDF) exceeds an anomaly threshold

(*anomaly_threshold*), the value is classified as an anomaly. The formula for the Weibull CDF used in anomaly detection is the same as previously presented in Equation 8:

$$CDF(x) = 1 - exp\left[-\left(\frac{x}{\lambda}\right)^{k}\right] \quad (8)$$

In the detection procedure, as illustrated in Figure 2, the defined threshold is relatively low, which results in the identification of a larger number of anomalies. While this increases the model's sensitivity, it poses the risk of including less intense peaks that may still be relevant. In later results, we discuss the importance of considering these less pronounced peaks during training, as they may represent important anomalous patterns which, if overlooked, could compromise the model's effectiveness in more complex scenarios.

### 2.4. Long Short-Term Memory (LSTM)

The Long Short-Term Memory (LSTM) model, introduced by Hochreiter and Schmidhuber [16], was designed to address the limitations of traditional recurrent neural networks (RNNs), which struggle to capture long-term dependencies due to the vanishing and exploding gradient problems. LSTM overcomes this challenge by incorporating a memory architecture that allows for the storage and selective control of information over time.

A key feature of LSTM is the introduction of \textit{memory cells}, which are governed by three essential gates: the *forget* gate, the *input* gate, and the *output* gate. The memory cell is designed to retain information for extended periods, enabling LSTM to capture long-term dependencies in temporal sequences.

The dynamics of LSTM networks can be described recursively at each time step $t$. The hidden state $h_t$ is computed based on the current input $x_t$ and the previous hidden state $h_{t-1}$, with modulation from the three main gates: *input*, *forget*, and *output*. Long-term memory is stored in the cell state $C_t$, which is updated at each time step through a weighted combination of incoming and retained information. This mechanism enables LSTM to maintain and update information over time, effectively overcoming the challenges faced by traditional RNNs in capturing long-term dependencies [16, 17].

Given the temporal nature of disruptions in plasmas—characterized by significant changes in plasma parameters shortly before the event—memory-capable models such as LSTM recurrent neural networks are particularly well suited for this task. The ability of LSTMs to capture long-duration temporal dependencies is crucial for predicting imminent disruptions.

Several recent studies have investigated this approach with promising results in magnetic confinement devices such as tokamaks. For instance, Vega et al. [2] used LSTM networks to predict disruptions in advance from time series of diagnostic

data; Zheng et al. [18] proposed a hybrid LSTM-based architecture with attention mechanisms to improve prediction robustness; and Agarwal et al. [19] explored various deep learning models, demonstrating that LSTMs outperform traditional approaches in the task of disruption prediction in plasmas.

### 2.5. Methodology Used to Prepare and Train the Data with the Goal of Predicting Future Disruption

The methodology used in this study can be divided into four main stages: time series generation, anomaly detection, LSTM model training, and model performance evaluation. Each of these stages is detailed below.

#### 2.5.1 Time Series Generation and Preprocessing:

To simulate the dynamic behavior of disruptive events, the previously discussed p-model was utilized. This model is widely applied to represent multifractal processes and generate time series with extreme fluctuations, which are relevant characteristics for the phenomena under study. Time series of fixed length 1024 points were generated using the exogenous parameter $p = 0.28$. The choice of this p value was based on prior studies indicating its adequacy in reproducing statistical patterns similar to those observed in real data of solar events and disruptions in tokamaks, including high variability and abrupt intensity peaks.

For the training dataset, 50 independent series were produced and concatenated to form the input time series for the model, totaling a series length of 51,200 points. This concatenation allows for greater pattern diversity and enhances the model's ability to capture different disruption dynamics. Each series exhibits stochastic variations inherent to the p-model, ensuring realism in the simulation of events.

Subsequently, the data were normalized by dividing by the maximum value of the series plus one, thereby constraining the values between 0 and 1, which facilitates the training process and improves model convergence.

For the test dataset using *p-model* data, an individual time series was generated, also with 1024 points and $p = 0.28$. This series was normalized following the same procedure applied to the training dataset.

#### 2.5.2 Anomaly Detection:

After generating the time series, an anomaly detection process based on Extreme Value Theory (EVT), as previously described, was applied. The Weibull Distribution was used to model the data and identify atypical behaviors within fixed windows of 100 points. For each window, the upper threshold was calculated as the mean plus 0.2 times the standard deviation of the data within that window ( $threshold\ factor = 0.2$). Values exceeding this threshold were classified as

extremes and used to fit the parameters of the Weibull distribution. Subsequently, for anomaly detection in new data, a cumulative probability threshold (anomaly threshold) of 0.85 was adopted, which allowed capturing a significant number of peaks, including those of lower magnitude. Including these less intense peaks, which precede major disruptive events, is crucial to improve the training performance of the LSTM network, especially considering data imbalance. This approach enables the model to learn and recognize temporal features and preliminary patterns that precede disruptive events, allowing for more sensitive and robust detection, even in the presence of subtle signals.

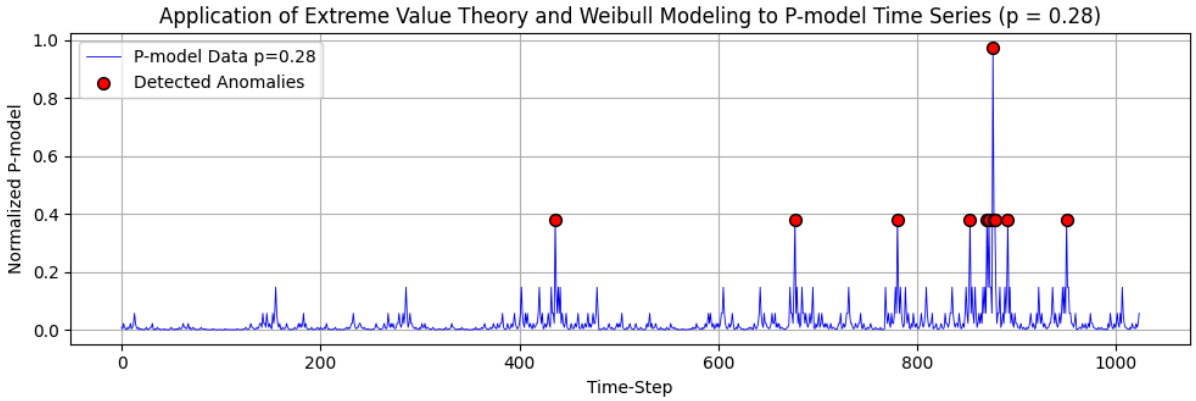


**Figure 3.** Time series generated by the *p-model* with 1024 points and detection of extreme values using the Weibull distribution.

The figure 3 illustrates the time series generated by the p-model, containing 1024 points, on which the model was tested. It is possible to observe the detection of extreme values in the series, highlighting one more pronounced peak alongside others of lesser intensity.

### 2.5.3 LSTM Model Training:

A Long Short-Term Memory (LSTM) network was employed to capture temporal relations and produce time series forecasting. The model architecture begins with a fully connected projection layer of 128 neurons to project the input data shape into the model's embedding dimension. Next, the model has a sequential structure in the form of six consecutive LSTM layers, each with an embedding dimension of 512. Lastly, another fully-connected projection layer made of 128 neurons transforms the learned features derived from the LSTM sequence to the final prediction space.

For every training iteration, a one-dimensional p-model sequence of 4096 data points was created to be used as the input sequence.

Before feeding the data into the network, the data was normalized via a min-max normalization. The scaling transformed the input values into the range [0, 1] by taking the 99th percentile of the data as the maximum. This was done to avoid the effect of possible outliers, with the minimum being zero by default according to the

p-model. The network was trained using a sliding window approach with a context window of 512 time steps which were applied iteratively to the 4096-point series. Optimization of model parameters was done using the Adam algorithm with a learning rate of 1 x $10^{-4}$ and batch size of 64.

### 2.5.4 Model Testing and Evaluation:

After training, the model was tested on three distinct domains in order to assess its robustness and generalization capability across different physical contexts of disruption. The first domain (i) corresponds to synthetic data generated by the p-model itself. A new time series of 320 points (for easier visualization) was created using the parameter $p = 0.28$, exhibiting multifractal characteristics and extreme fluctuations similar to those observed in real events. This series was kept separate from the training data and used exclusively for model evaluation.

The second domain (ii) is based on simulated heat flux data from tokamaks. The time series were extracted from simulated thermal emissivity images for the SPARC tokamak and adapted to the ITER project. These simulations are part of studies on disruption mitigation modeling, combining techniques such as massive gas injection (MGI) and the use of runaway electron mitigation coils (REMC). The dynamics of the thermal quench (TQ) process induced by MGI were simulated using the three-dimensional MHD code NIMROD, considering different valve configurations and REMC current drive scenarios. The resulting series reflect the temporal evolution of thermal emission during the critical stages of disruption, as described in Izzo et al. [20].

The third domain (iii) refers to real observational data acquired by the Solar Dynamics Observatory (SDO). In this case, the series were constructed from the average of the 10 maximum intensity values extracted over 317 windows of 10 seconds during the evolution of a solar eruption. These data reflect the dynamic activity associated with the formation and rupture of current sheets, which are central structures in magnetic reconnection processes in solar flares. The event is detailed in Chitta et al. [21].

The diversity of these three domains: (i) synthetic multifractal, (ii) magnetohydrodynamic simulation of confined plasma, and (iii) real heliospheric observations, provides a basis for investigating the model's applicability in different disruption regimes. Representative images of these time series and solar snapshots extracted from the SDO are presented in the results section, visually illustrating the variability and complexity of the data used. Next section presents the datasets used to validate the AEDeep methodology across three distinct domains: real solar observational data, numerical simulations from a tokamak system, and synthetic time series.

## 3. Data in the Context of AEDeep

In this preliminary study, we selected key physical variables from both fusion and solar plasma datasets, namely, heat flux in tokamaks (Figure 4) and intensity fluctuations in extreme ultraviolet (EUV) wavelengths observed by SDO (Figure 5). It is worth noting their intrinsic similarity to the energy cascade dynamics characterized by the p-model. These time series exhibit intermittent bursts, scale-dependent fluctuations, and non-Gaussian statistical signatures consistent with the energy distribution processes modeled by the p-model in both its endogenous (self-organized) and exogenous (externally driven) variants. The heat flux in magnetically confined plasmas reflects localized instabilities and transport events, while the EUV intensity captures sudden releases of energy in the solar corona, both serving as proxies for complex, multiscale energy dissipation. This similarity supports the methodological foundation of our approach, where synthetic time series generated via the p-model serve as training references for deep learning models tasked with detecting and predicting real disruption events.

To bridge the gap between synthetic modeling and real observational data, we developed AEDeep (Anomaly Explorer with Deep Learning), a framework that integrates time series generated by the p-model with Long Short-Term Memory (LSTM) neural networks for disruption detection and prediction. By tuning the p-model parameters, particularly those controlling the series length and energy fragmentation scale, we produced surrogate time series that closely mimic the statistical behavior of the case study data. This allowed us to train LSTM models under conditions that reflect the temporal structure and burstiness inherent in real plasma systems. Validation of the synthetic-to-real transfer was performed using criteria based on higher-order statistical moments represented in the Kullen-Frey parameter space and Power Spectral Density (PSD) analyses. These metrics ensured that the multifractal features and spectral characteristics of the synthetic and observed series were compatible. Full details of this validation process, including parameter sensitivity and comparative plots, are provided in the supplementary material.

### 3.1. Solar Data

Figure 4 shows the time series extracted from a solar eruption observed by the AIA instrument aboard the Solar Dynamics Observatory (SDO) in the 304 Å band. This was the most powerful flare observed during the solar cycle 25, on 2020 November 29 around UT 13:00 [21]. Although the active region 12786 is noted with relevance, the target of this analysis was the eruptive phenomenon that occurred at the limb. This is a rare phenomenon where a flux tube abruptly emerges from a flux rope in a disruptive manner. The fluctuation pattern of the maximum emission values in the region presents a similarity to the multiplicative cascade pattern for which the LSTM was trained with the p-model. The full video is added as supplementary material available on github [22].

The series was constructed by averaging the 10 highest intensity values within each of 317 time windows, each lasting 10 seconds, throughout the evolution of the solar event. The target region used for the analysis is highlighted in the first image with a yellow box.

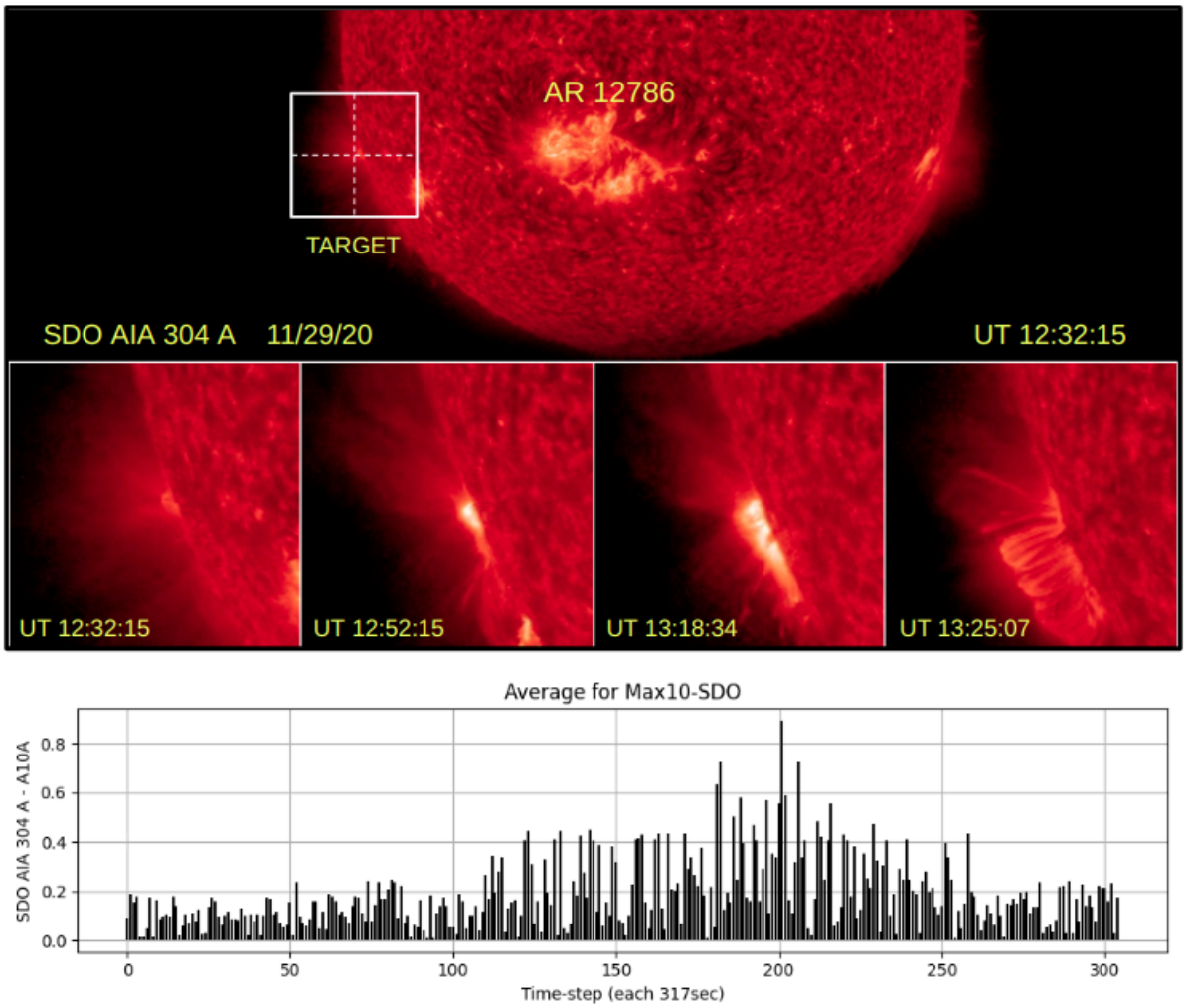


**Figure 4.** Selected frames from SDO/AIA images (304 Å) showing the evolution of the solar eruption. The yellow-marked region indicates the target area from which the time series was extracted.

### 3.2. Tokamak Data

The same methodology used for solar data was applied to obtain plasma data in a SPARC-to-ITER scenario. Figure 5 presents three snapshots from a high-resolution, non-resonant 6-valve simulation of plasma emissivity, taken at the 180º toroidal angle, up to the point of peak radiated power. This dataset is derived from a time- and power-scaled version of a NIMROD simulation [20] adapted for ITER conditions via AEDeep. The original simulation [20] was conducted in the context of SPARC and included two disruption mitigation strategies: Massive Gas Injection (MGI) and a Runaway Electron Mitigation Coil. The third frame (at 3.05 ms) illustrates the characteristic disruptive pattern, capturing the onset of plasma

instability. From this simulation, a representative time series of heat flux (thermal emission) was extracted, effectively capturing the dynamic evolution leading to the disruptive event.

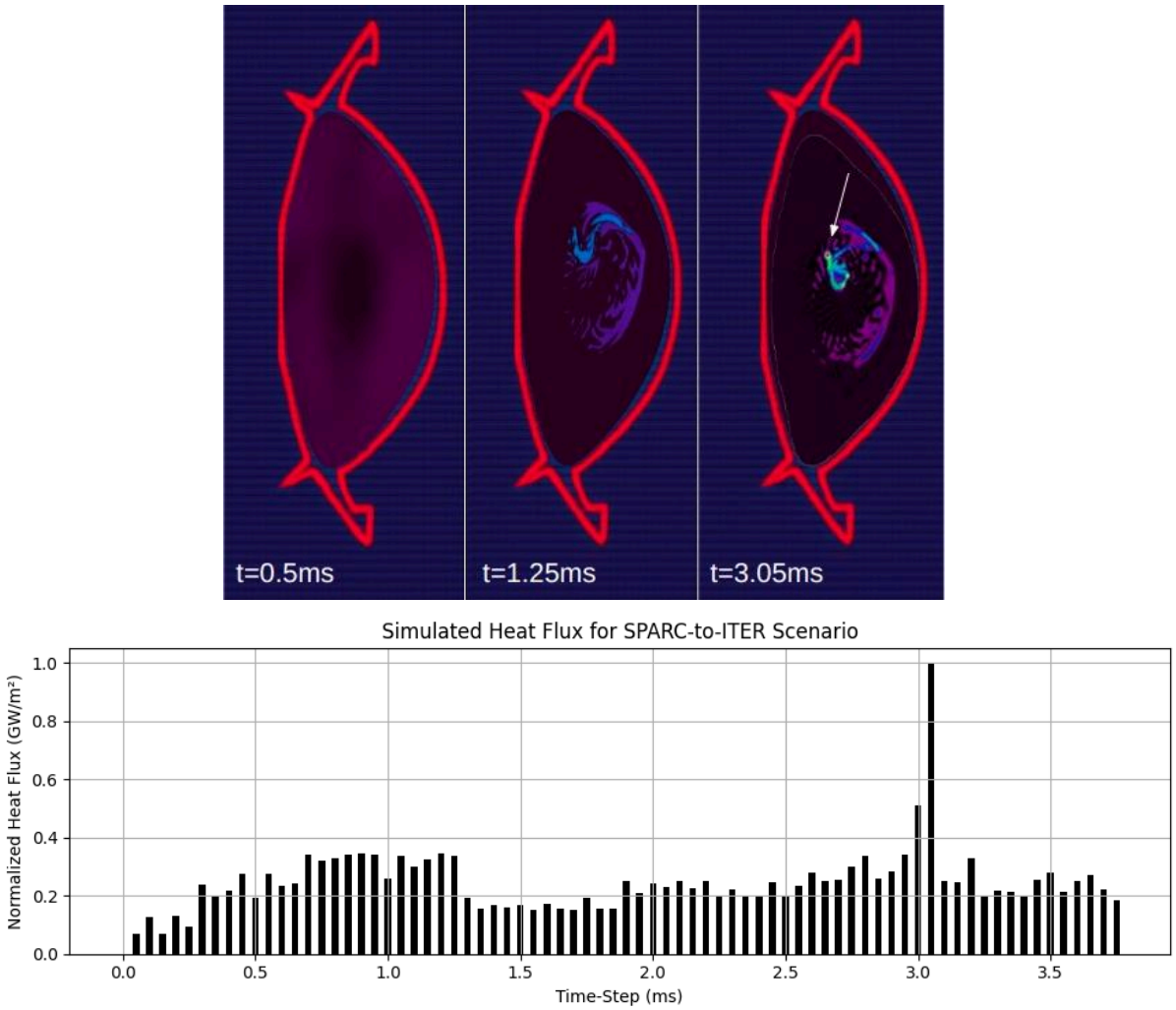


**Figure 5.** Snapshots from the SPARC-to-ITER simulation highlighting the plasma dynamics and disruption at 3.05 ms. The corresponding heat flux time series was extracted for analysis.

## 4. Results and Interpretation

In this section, we present the results obtained from the application of the AEDeep methodology across three distinct domains: (i) synthetic data generated by the p-model, (ii) thermal flux simulations in tokamaks (SPARC-to-ITER), and (iii) real observations of solar eruptions (SDO/AIA). The analysis focused on anomaly detection and the predictive capability of the LSTM model. Extreme Value Theory (EVT) and the Weibull distribution were employed for anomaly detection in the analyzed data.

The model's performance on synthetic data (p-model) was evaluated using an exogenous time series with parameter p=0.28 and 1024 time steps. Since the LSTM model was trained on time series generated by the p-model (also with p=0.28), which simulate multifractal fluctuations typical of extreme events, the best results were

observed in this domain due to the similarity between the training and test data. Figure 3 illustrates the series used, along with the anomaly detection performed using Extreme Value Theory (EVT) and the corresponding Weibull distribution.

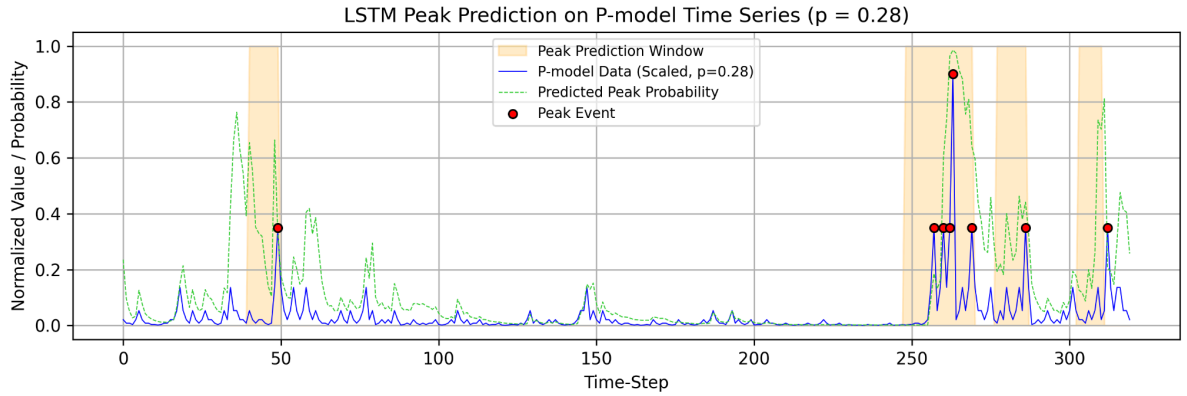


**Figure 6.** LSTM Anomaly Prediction (10-Step Horizon) for P-model Time Series (p=0.28).

Figure 6 presents the prediction results obtained by the LSTM model. The green dashed line represents the probability of an anomaly occurring within the next 10 steps of the time series. This means the model is anticipating the likelihood of an extreme event, in this case, a peak, happening within the following 10 points of the series.

The LSTM model demonstrated the ability to capture subtle temporal patterns that precede abrupt fluctuations. This ability is evidenced in the regions highlighted as peak prediction windows, which indicate the intervals where the model is expected to forecast the occurrence of future peaks. Since it was configured to estimate the probability of an anomaly occurring within the next 10 steps, this is the most appropriate window for prediction. Observing the graph, it can be seen that the model was able to correctly predict the peaks of extreme fluctuations generated by the p-model, with few instances of high probabilities in regions without actual peaks, indicating a low number of false positives. This suggests that the LSTM learned to recognize precursor dynamic patterns present in the multifractal structure of the synthetic series. It is also noticeable that the predicted probabilities increase shortly before the actual peaks occur, validating the model's ability to anticipate disruptions in advance.

This result highlights the model's sensitivity not only to major peaks but also to lower-intensity fluctuations that serve as early warning signals. This predictive performance was enhanced by the training strategy adopted, which included a significant number of moderate anomalies in the training set.

This procedure was also applied to other datasets, including solar eruptions observed with SDO/AIA and time series generated from plasma simulations in the SPARC-to-ITER tokamak. These datasets are presented in Section 3, where the data acquisition process is detailed. Initially, analyses were performed on SDO/AIA images in the 304 Å band. Some frames from the segment used to generate the time

series show the occurrence of a plasma disruption, as illustrated in the third image (UT 13:18:34) of Figure 4. The detected anomalies coincided with phases of magnetic energy accumulation, corroborating previous studies on magnetic reconnection. As shown in Figure 7, several anomalies were identified that, in fact, correspond to a single plasma disruption. The peaks detected before and after the event represent the intense variations that occur immediately prior to the disruption and during the reconnection of the magnetic field.

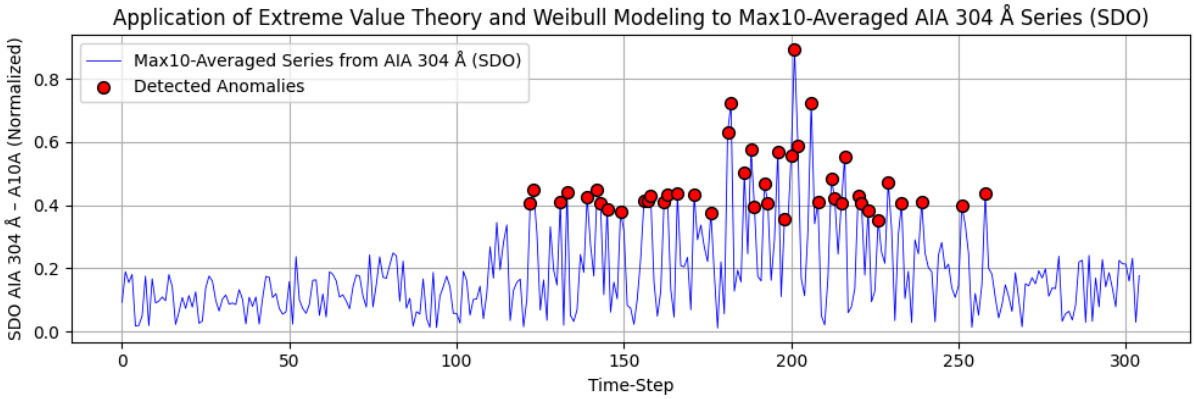


**Figure 7.** Time series of the solar eruption with detected anomalies, highlighting phases of magnetic energy accumulation and plasma disruption.

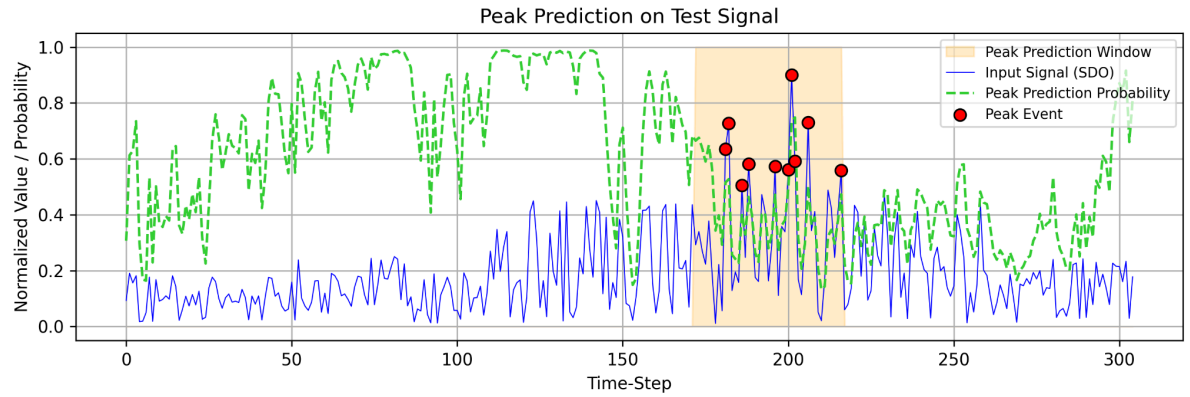


**Figure 8.** LSTM Anomaly Prediction (10-Step Horizon) for Max10-Average Series (SDO).

Although real data present greater complexity (e.g., instrumental noise), the model was initially able to predict peaks with a probability of around 60%. However, the results were not entirely satisfactory, as within the interval from 50 to approximately 150 time steps, a high probability of peak occurrence was indicated, yet no actual peaks were observed in the subsequent data. This may suggest several factors, including the possibility that the correlation between previous data and a peak event may, in fact, extend over a window larger than 10 time steps. Overall, the peak prediction window reached a maximum probability of about 80% for a peak to occur in the following data. These results suggest that models trained with the p-model may be generic enough to generalize to real-world time series; however, some adaptations are necessary due to the number of false positives observed.

These details are illustrated in Figure 8, indicating the model's potential for solar flare prediction in space weather systems.

For the simulated plasma data in the SPARC-to-ITER tokamak scenario, a similar approach was adopted. Figure 9 shows the heat flux time series extracted from the simulation, with anomalies detected by the AEDeep methodology highlighted. These anomalies correspond to dynamic events preceding the plasma disruption observed at 3.05 ms in the simulation images (Figure 5). The detection indicates abrupt variations in heat flux associated with instability and the onset of disruptive plasma behavior.

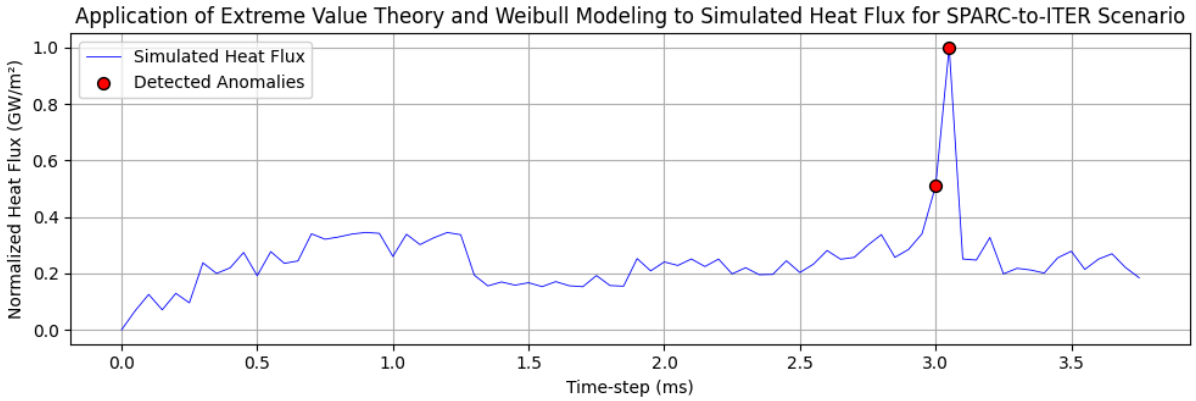


**Figure 9.** Time series generated from the SPARC-to-ITER tokamak simulation with detected anomalies, highlighting events related to plasma disruption.

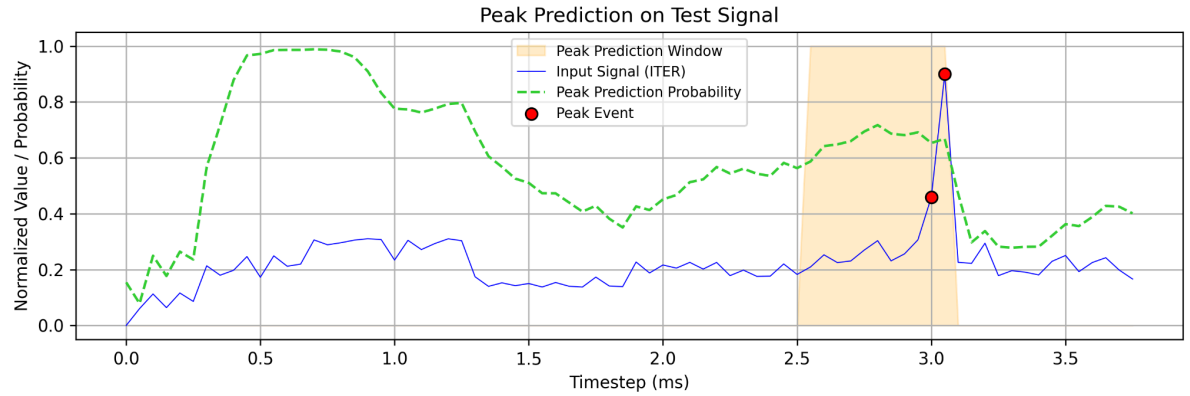


**Figure 10.** LSTM Anomaly Prediction (10-Step Horizon) for Simulated Heat Flux in SPARC-to-ITER.

Figure 10 shows the results obtained with the LSTM model trained on data from the tokamak simulation, confirming the effectiveness of AEDeep in predicting simulated disruptions, even under the complexity imposed by the physical model. This performance is highlighted within the *peak prediction window* (in yellow on the graph), where the predicted probability of a peak reaches 75%. It is worth noting that a similar scenario to that observed in the SDO/AIA data occurs between approximately 0.5 ms and 1.0 ms. Nevertheless, the model successfully generalizes to this situation, demonstrating its adaptability. These findings indicate that the training of the network can be further improved to make it even more robust in

predicting disruptions across different plasma types using mathematical models such as the *p-model*. This suggests that the proposed approach can be scaled and integrated into operational systems for disruption forecasting and control in tokamaks.

## 5. Concluding Remarks

The preliminary results obtained in this study were fully satisfactory for both case studies considered, demonstrating that the AEDeep methodology can be useful in characterizing anomalies and disruptions from time series that are compatible with the multiplicative cascade patterns (endogenous or exogenous) simulated by the p-model.

A common challenge in disruption analysis is the limited temporal length of time series associated with plasma instabilities. In both fusion and astrophysical contexts, disruptions often unfold over brief intervals, resulting in short datasets that hinder the application of classical statistical forecasting techniques. Frequentist approaches, in particular, become unreliable due to the non-Gaussian nature of the underlying distributions, dominated by extreme fluctuations and intermittent bursts, key signatures of disruptive events. In this context, the AEDeep framework offers a significant methodological advantage. By leveraging deep learning models, such as LSTM networks, AEDeep circumvents the need for long individual time series, instead requiring a sufficient number of representative samples for effective training. These samples can be synthetically generated using the p-model, whose capacity to replicate multifractal cascade dynamics enables the creation of realistic, statistically consistent datasets. This strategy allows AEDeep to generalize from synthetic to real data, making it a powerful tool for disruption prediction even in scenarios with limited observational data length.

In addition to testing AEDeep on a large sample set for both case studies covered in this preliminary work, the main goal of which is to introduce the methodology, we are advancing the method by implementing a 2D version of the p-model, the Gradient Pattern Analysis [23] and CyMorph [24] technique, and an LSTM-Autoencoder [15] architecture. Together, these three additional tools will enhance AEDeep by enabling exploration of anomalies and disruptions directly in the spatiotemporal domain, thus enabling the explicit identification of the location and evolution of the structural pattern of a disruption as it evolves. This advancement will open up new possibilities for cutting-edge applications, from space weather forecasting using AI-embedded CubeSats to real-time prediction and mitigation of disruptive anomalies in tokamaks. With the emergence of quantum sensing technologies, advanced smart materials can be integrated into CubeSats or incorporated into tokamak magnets, thus reinforcing, in the latter case, the technological basis necessary to ensure the full operational success of 21st century fusion reactors, vital to driving the global energy transition.

**Acknowledgments**

C.F., S.J., D.M., L.B., and R.S. acknowledge scholarship support from the Brazilian funding agencies CNPq and CAPES. R.R.R. and J.A.V. are grateful to FAPESP for Grant No. 21/15114-8 and C.E.ANID/Fondecyt Grant No. 1240697. C.E.F. thanks to the LSST-DA Data Science Fellowship Program, which is funded by LSST-DA, the Brinson Foundation, the WoodNext Foundation, and the Research Corporation for Science Advancement Foundation, as well as the Líderes Estudar Program from Fundação Estudar. His participation in the program has benefited this work. The authors thank NASA for providing the SDO image data used in this study.

## REFERENCES


[1] C. Meneveau and K. R. Sreenivasan, "The multifractal nature of turbulent energy dissipation," **Journal of Fluid Mechanics**, vol. 224, pp. 429–484, 1991.

[2] C. Meneveau and K. R. Sreenivasan, "Simple multifractal cascade model for fully developed turbulence," **Physical Review Letters**, vol. 59, no. 13, p. 1424, 1987.

[3] S. Hochreiter and J. Schmidhuber, "Long short-term memory," **Neural Computation**, vol. 9, no. 8, pp. 1735–1780, 1997. doi: 10.1162/neco.1997.9.8.1735

[4] F. Gers, J. Schmidhuber, and F. Cummins, "Learning to Forget: Continual Prediction with LSTM," **Neural Computation**, vol. 12, no. 10, pp. 2451–2471, Oct. 2000. doi: 10.1162/089976600300015015

[5] S. Coles, *An Introduction to Statistical Modeling of Extreme Values*, London: Springer-Verlag, 2001.

[6] A. Kirk et al., "Spatial and Temporal Structure of Edge-Localized Modes," **Phys. Rev. Lett.**, vol. 92, no. 24, p. 245002, Jun. 2004. doi: 10.1103/PhysRevLett.92.245002

[7] D. Sornette, F. Deschâtres, T. Gilbert, and Y. Ageon, "Endogenous versus exogenous shocks in complex networks: An empirical test using book sale rankings," **Physical Review Letters**, vol. 93, no. 22, p. 228701, 2004.

[8] L. de Haan and A. Ferreira, *Extreme Value Theory: An Introduction*, **Springer Series in Operations Research and Financial Engineering**, Springer, 2006. doi: 10.1007/0-387-34471-3

[9] K. Shibata and T. Magara, "Solar Flares: Magnetohydrodynamic Processes," **Living Reviews in Solar Physics**, vol. 8, no. 6, 2011. doi: 10.12942/lrsp-2011-6

[10] L. E. McPhillips et al., "Defining extreme events: A cross-disciplinary review," **Earth's Future**, vol. 6, no. 3, pp. 441–455, 2018. doi: 10.1002/2017ef000686

[11] J. Kates-Harbeck, A. Svyatkovskiy, and W. Tang, "Predicting disruptive instabilities in controlled fusion plasmas through deep learning," **Nature**, vol. 568, pp. 526–531, 2019. doi: 10.1038/s41586-019-1116-4

[12] J. X. Zhu et al., "Hybrid deep-learning architecture for general disruption prediction across multiple tokamaks," **Nuclear Fusion**, vol. 61, no. 2, p. 026007, 2021. doi: 10.1088/1741-4326/abc664

[13] A. Agarwal et al., "Deep sequence to sequence learning-based prediction of major disruptions in ADITYA tokamak," **Plasma Physics and Controlled Fusion**, vol. 63, no. 11, p. 115004, 2021. doi: 10.1088/1361-6587/ac234c

[14] J. Vega, A. Murari, S. Dormido-Canto, et al., "Disruption prediction with artificial intelligence techniques in tokamak plasmas," **Nature Physics**, vol. 18, pp. 741–750, 2022. doi: 10.1038/s41567-022-01602-2

[15] H. Akbarian, I. Mahgoub, and A. Williams, "Autoencoder-LSTM Algorithm for Anomaly Detection," in **2023 IEEE 20th Int. Conf. on Smart Communities (HONET)**, pp. 1–6, 2023. doi: 10.1109/HONET59747.2023.10374710

[16] W. Zheng, F. Xue, Z. Chen, et al., "Disruption prediction for future tokamaks using parameter-based transfer learning," **Communications Physics**, vol. 6, no. 1, p. 181, 2023. doi: 10.1038/s42005-023-01296-9

[17] E. S. Pereira, L. S. Marcondes, and J. M. Silva, "On-Device Tiny Machine Learning for Anomaly Detection Based on the Extreme Values Theory," **IEEE Micro**, vol. 43, no. 6, pp. 58–65, 2023. doi: 10.1109/MM.2023.3316918

[18] S. Mukherjee, S. D. Murugan, R. Mukherjee, and S. S. Ray, "Turbulent flows are not uniformly multifractal," **Physical Review Letters**, vol. 132, no. 18, 2024. doi: 10.1103/physrevlett.132.184002

[19] X. Xie, G. Li, K. K. Reeves, and T. Gou, "Probing turbulence in solar flares from SDO/AIA emission lines," **Frontiers in Astronomy and Space Sciences**, vol. 11, 2024. doi: 10.3389/fspas.2024.1383746

[20] V. A. Izzo, B. Stein-Lubrano, A. Battey, R. Sweeney, C. Hansen, and R. A. Tinguely, "Disruption mitigation modeling for the SPARC tokamak," **Physics of Plasmas**, vol. 32, no. 4, p. 042507, Apr. 2025. doi: 10.1063/5.0254080

[21] L. P. Chitta, E. R. Priest, and X. Cheng, "From Formation to Disruption: Observing the Multiphase Evolution of a Solar Flare Current Sheet," **The Astrophysical Journal**, vol. 911, no. 2, p. 133, Apr. 2021. doi: 10.3847/1538-4357/abec4d

[22] C. E. Falandes et al., "AEDeep-Disruption-Prediction: Machine Learning for Plasma Disruption Prediction in Solar and Tokamak Systems," **GitHub repository**, 2025. Available: https://github.com/Desduh/AEDeep-Disruption-Prediction

[23] R. R. Rosa, R. R. de Carvalho, R. A. Sautter, P. H. Barchi, D. H. Stalder, T. C. Moura, S. B. Rembold, D. R. F. Morell, and N. C. Ferreira, "Gradient pattern analysis applied to galaxy morphology," **Monthly Notices of the Royal Astronomical Society: Letters**, vol. 477, no. 1, pp. L101–L105, Jun. 2018. doi: 10.1093/mnrasl/sly054

[24] P. H. Barchi, R. R. de Carvalho, R. R. Rosa, R. A. Sautter, M. Soares-Santos, B. A. D. Marques, E. Clua, T. S. Gonçalves, C. de Sá-Freitas, and T. C. Moura, "Machine and deep learning applied to galaxy morphology: A comparative study," **Astronomy and Computing**, vol. 28, p. 100334, 2019. doi: 10.1016/j.ascom.2019.100334