Influence of the trigger time window on the detection of gas turbine trip

Influence of the trigger time window on the detection of gas turbine trip

E Losi [1], **M Venturini** [1], **L Manservigi** [1] **and G Bechini** [2]

[1] Università degli Studi di Ferrara, Via G. Saragat 1, 44122, Ferrara, Italy
[2] Siemens Energy, Otto-Hahn-Ring 6, 81739, Munich, Germany

Corresponding author's e-mail address: enzo.losi@unife.it

Abstract. Gas turbine (GT) trip is one of the most disrupting events that affect GT operation, since its occurrence causes a reduction of equipment remaining useful life as well as revenue loss because of business interruption. Thus, early detection of incipient symptoms of GT trip is crucial to ensure efficient operation and lower operation and maintenance costs. This paper applies a data-driven methodology that employs a Long Short-Term Memory (LSTM) neural network and a clustering technique to identify the time point at which trip symptoms are triggered. The same methodology also partitions trips into homogeneous clusters according to their most likely trigger position. The methodology is applied to two real-world case studies composed of a collection of trips, of which the causes are different, taken from various fleets of industrial GTs. Data collected from twenty sensors during three days of operation before trip occurrence are analyzed. For each trigger scenario, this paper investigates different lengths of the training and testing time window (namely "trigger time window"), by considering up to 24, 18, 12 or 6 hours before and after the considered trigger position. The results demonstrate that longer time windows allow an improvement of the predictive capability.

1. Introduction

In the current competitive market, monitoring and diagnostics of gas turbines (GTs) require the implementation of a robust, efficient, and flexible predictive maintenance plan to ensure high reliability and productivity [1 - 2]. To this purpose, machine learning approaches can exploit the data of GT history to detect deviations of normal behavior and changes in GT health state before they impact customer operation and profitability [2]. However, the capability of the diagnostic process depends on data reliability, which can be affected by both feature and label noise. To tackle these issues, data should be pre-processed. Several tools are available in the literature, as for instance [3, 4] for feature noise and [5, 6] for label noise.

In the field of GT monitoring and diagnostics, GT trip is of great concern for both manufactures and users. In fact, trip is an unscheduled operational event during which a GT abnormally shuts down [7], thus leading to a direct impact on GT lifespan and revenue. In addition, since GT trip may be triggered by different causes, e.g., turbine vibrations and overspeed, increase in the gradients of exhaust gases or problems in fuel spray nozzles, its occurrence may not be infrequent. Therefore, predicting GT trip during its incipient phase would allow saving costs and improving GT reliability.

In the literature, very few studies dealt with GT trip prediction. The study [8] presented a tool for the analysis of the causes of GT trips. The study [9] aimed at predicting GT trips produced by a specific cause (hydraulic valve failure) by means of a physics-based model.

The authors of the current paper also tackled the challenge of predicting GT trip by exploiting operating data gathered from multiple GT assets and by applying machine learning (ML) techniques.

The study [10] presented a procedure for handling big data and a structured methodology aimed at partitioning GT transients. Subsequently, based on the transients classified in [10], the study [11] proposed a procedure for target data selection by identifying two datasets of transients, gathered from fleets of GTs installed in different sites, and developed a feature engineering methodology to generate and select the most suitable features to fit an ML model aimed at predicting GT trip. The Random Forest (RF) model developed in [12] allowed predicting GT trip with an accuracy in the range 75% - 85%. The paper [12] also reported a discussion about feature importance, by demonstrating that features related to compressor efficiency and specific fuel consumption likely embed most of the pieces of information useful to discriminate GT trip from normal operation. The study [13] investigated the fusion of five data-driven base models (k-Nearest Neighbors, Support Vector Machine, Naïve Bayes, Decision Trees, and Long Short-Term Memory neural networks) by means of voting and stacking, in order to increase base model accuracy and improve prediction robustness. The results demonstrated that the stacked model provided higher accuracy than base models and also outperformed voting. The study [14] applied a systematic statistical analysis to identify the most important variables, followed by the application of a novel machine learning technique known as temporal decision tree. The learned models were used to extract statistical rules for predicting trip. Finally, in the study [15], the authors developed a data-driven methodology aimed at disclosing the onset of trip symptoms. A Long Short-Term Memory (LSTM) neural network was employed as the classification model. The methodology provided the most likely trigger position for four clusters of trips within the two days before trip occurrence with a confidence in the range 66% - 97%. The outcome of the study [15] is the starting point of the analyses conducted in the current paper.

One of the most consolidated outcomes of the previous studies conducted by the authors of the current paper is that, given its inherent complexity and randomness of occurrence, the goal of GT trip prediction can be pursued by employing ML data-driven techniques. For this reason, a brief survey is reported below about the potential of applying neural network models to GT health monitoring. In particular, Long Short-Term Memory (LSTM) neural networks represent a powerful machine learning tool potentially suitable to GT diagnostics and prognostics, since they can exploit their inherent ability to encode temporal information.

Bai *et al*. [16] used an ensemble of LSTM regressors for detecting faults in marine GTs and compared LSTM network to other methods including support vector regression, single-layer feedforward neural network, extreme learning machine and Elman recurrent neural network by demonstrating the superiorities of LSTM network in fault detection. Thanks to the capability of learning long-term patterns in time-dependent data, LSTMs proved high reliability in predicting degradation of GT assets and estimating their remaining useful life. The study [17] demonstrated that LSTMs could predict the declining trend of performance of heavy-duty GT assets with maximum prediction error lower than 0.2%. Xiang *et al*. [18] employed LSTMs to construct health indicators for remaining useful life prediction of aircraft engines. The studies [19 - 22] documented other specific applications of LSTM to remaining useful life prediction of aircraft turbofan engines.

As anticipated, the current paper stems from the study [15] in order to further investigate the time evolution of trip symptoms and identify the most likely time window within which the symptoms of an incoming trip arise, thus potentially allowing early actions for trip prevention. In this paper, symptoms are intended as perturbances in the data that cannot be detected by means of classical approaches (e.g., physics-based thermodynamic models or frequency-domain vibration analysis) and thus require a data-driven approach to grasp signs embedded in the measured values.

More in detail, the current paper applies the data-driven methodology presented in [15], which is composed of two steps. The first step makes use of a data labeling approach to model a variation in GT health state due to the trigger of trip symptoms; multiple scenarios characterized by different trigger positions are investigated. An LSTM is employed as the predictive model to estimate the trigger time point. The second step of the methodology uses the responses of the trained models to isolate the most likely trigger position and cluster trips into homogeneous groups that reflect different trip causes. To validate the analyses carried out in this paper, two large field datasets comprising trip observations composed of three days before trip occurrence are considered. The main novel contribution of this paper is represented by the systematic analysis of different lengths of the training

and testing time window (namely "trigger time window"). First, the data-driven models developed in [15] by employing the longest available time window for each trigger scenario (up to 24 hours) are tested on unseen trips by reducing the trigger time window up to 1 hour in order to search for the most likely time window within which symptoms appear. Second, for each trigger scenario, the training of the data-driven models is also conducted by considering different trigger time windows, i.e., 24, 18, 12, 6 or 1 hour before and after the considered trigger position, while the test is always conducted by considering the longest available time window (up to 24 hours), in order to investigate the influence of the amount of training information on the prediction capability of the data-driven models.

2. Methodology

This section summarizes the procedure used for identifying the time point of the onset of GT trip symptoms developed in [15]. First, the rationale of the methodology is presented. Then, data labeling subsection describes the approach employed to model the onset of trip symptoms. Finally, observation partition subsection summarizes the identification of the most likely trigger time point and the partition of trips into homogeneous clusters according to the time point at which the symptoms arise.

2.1. Overview

The proposed data-driven methodology shown in Figure 1 aims at modeling a deviation in GT performance by means of a step variation in the label of available data and forces the predictive data-driven model (in this paper, LSTM) to learn hidden patterns and rules embedded in the data.

The time point of the onset of GT trip symptoms is usually unknown. Moreover, it can be reasonably expected that the considered trips are not homogeneous since trip causes can be disparate, e.g., failed sensors, compressor failure or combustion issues, and the time point of the onset of GT trip symptoms varies. For this reason, different trigger scenarios are modelled to partition GT trip observations into homogeneous clusters according to the respective most likely time point of occurrence of trip onset.

2.2. Data labeling

To identify the onset of trip symptoms, the proposed approach makes use of a parameter T (trigger) that quantifies the time before trip occurrence. For each trigger position, the approach makes use of two time windows of the same length, where symptoms are supposed to be absent before the trigger and present after the trigger. Thus, the data sequence before the time point T is labeled as "No trip" and is supposed to be representative of normal operation, while the data sequence after the time point T is labeled as "Trip" to model a deviation in GT performance and always starts right after the No trip sequence, as shown on the left hand-side of Figure 1. The length of the two time windows (expressed in hours) is denoted by L in this paper.

As discussed in the Case Study section, 72 hours of operation (i.e., three days) are taken into account for each trip observation, and the trigger T varies from 48 hours through 1 hour before trip occurrence, thus modeling $N_T = 48$ trigger scenarios in total in this paper. The first $(N_T - L_{max} + 1)$ scenarios have $L = L_{max}$. Then, the trigger moves forward timewise one hour at a time by decreasing the value of L in such a manner that the trigger is closer to trip occurrence, and L is equal to the number of remaining hours between the trigger and the time point at which trip occurs. Unlike the analyses made in [15], this paper investigates different values of L_{max} in order to provide a further insight about the detection of GT trip symptoms.

2.3. Training and testing

In the training phase, one classification model is trained for each scenario, i.e., 48 models are developed in total, by assuming that the same trigger position fits all trip observations. In this paper, the training phase is conducted by considering different values of L_{max}.

In the test phase, each trained model is tested on unseen trips in all the considered training scenarios. Therefore, for a given model, the trigger point moves from $T = 48$ through $T = 1$ one hour at a time, regardless of the trigger position used to train model, to search for the most likely trigger time point. For each trip observation, 48 predictions in total (equal to the total number of trained models)

are performed for each trigger position. In such a manner, the responses of different models are exploited to identify the most likely trigger position for any given observation.

2.4. Observation partition

To identify the most likely trigger time point of a given trip and cluster multiple trips into homogeneous groups, the observation partition methodology, sketched on the right hand-side of Figure 1, goes through the following steps:

- The quantity *Ipp* is calculated to quantify the confidence with which a trained model predicts the No trip label before the trigger and the Trip label after the trigger for a given trigger scenario and trip observation. Therefore, the higher the confidence of the predictions before and after the trigger, the higher *Ipp* (the maximum value is 1); otherwise, if at least one of the predictions before or after the trigger is not correct, the value of *Ipp* is set to 0.
- For each trip observation separately, the values of *Ipp* feed the *k*-medoids technique that groups trained models into three homogeneous subsets. Each subset contains those models that perform similarly by varying the trigger position. The least informative subset (i.e., the one that contains those models that do not provide valuable information on the detection of the most likely trigger position for the trip under analysis) is discarded.
- For each of the two remaining subsets, a change-point detection technique [23] isolates the range of consecutive trigger positions that has the highest mean and the barycenter of this range is the most likely trigger position T^* according to the models included in the considered subset.
- Based on the identified trigger position T^*, the observation under analysis is assigned to the *c*th cluster (four clusters in total), according to the criterion reported in Eq. (1):

$$\begin{aligned} c = 1 &\Leftrightarrow T^* \in [48, 36[\\ c = 2 &\Leftrightarrow T^* \in [36, 24[\\ c = 3 &\Leftrightarrow T^* \in [24, 12[\\ c = 4 &\Leftrightarrow T^* \in [12, \ \ 1] \end{aligned} \tag{1}$$

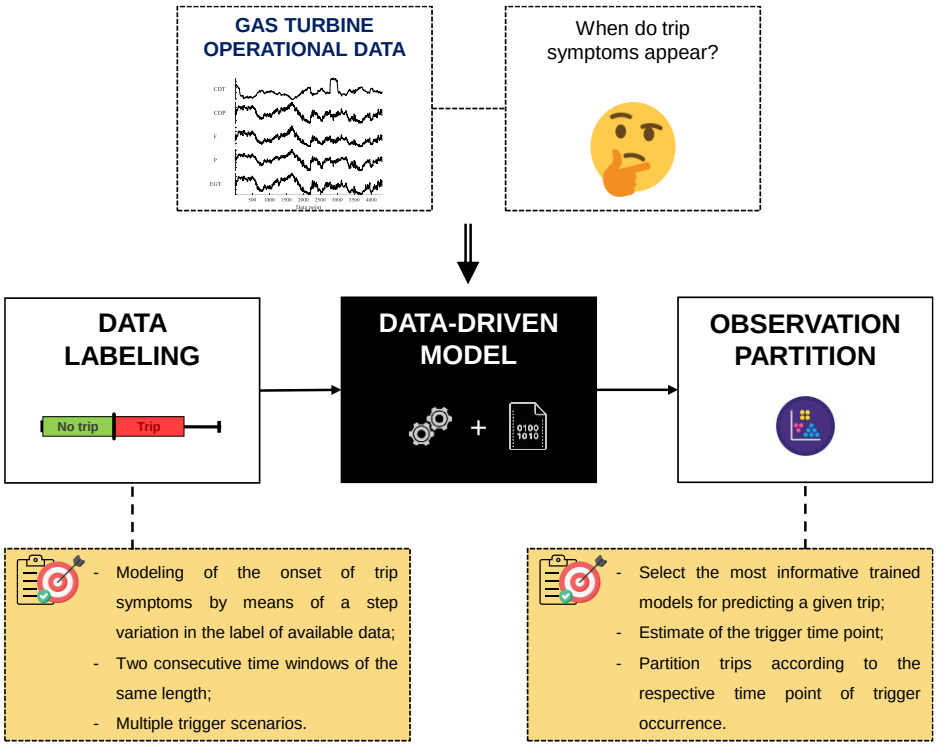

Figure 1. Methodology.

- Since two different cluster assignments may be possible, the best choice is the one that provides the highest value of the silhouette metric and the most informative subset of models used to predict the trip under analysis is obtained.
- To characterize the behavior of a given trip over multiple trigger positions, the mean of the *Ipp* values provided by the models included in the most informative subset of that trip is calculated.

2.5. Long short-term memory network

The data-driven model employed in the proposed methodology is an LSTM neural network [24], i.e., a recurrent neural network (RNN) capable of encoding temporal information and learning long-term dependencies embedded in multivariate timeseries data.

Based on the analyses carried out in [15], the neural network employed in this paper is composed of i) a sequence input layer, ii) three LSTM layers, iii) a fully connect layer, iv) a softmax layer and v) a classification output layer. To avoid overfitting, a dropout layer is inserted after each LSTM layer [25]. The number of hidden units is set equal to 200, 150, and 125 for the three LSTM layers, respectively. The dropout probability is set equal to 0.5 according to [25]. The LSTM is trained by using the ADAM stochastic optimization method and all the training parameters are in agreement with [26]. The training process stops if the maximum number of epochs is reached or the loss function falls below a given threshold, which in this paper is set equal to 10^{-3}.

3. Case study

3.1. Data description

Two real-world case studies, namely case study A and case study B, are considered in this paper. Dataset A is composed of 44 trip observations collected from four GTs and dataset B includes 39 trip observations in total from six GTs. Therefore, the two datasets have a comparable size. It is worth highlighting that both datasets contain some trip observations that occurred at the same time in different units of the same power plant.

For each observation, 20 measured variables, namely compressor outlet temperature, compressor outlet pressure, fuel flow rate, power output, and sixteen turbine outlet temperatures, are monitored with frequency equal to one minute during three days before trip occurrence and are employed in this paper to detect the trigger of trip symptoms.

3.2. Clusters

Based on the methodology summarized in Section 2, the paper [15] identified four different clusters of trips for both case study A and case study B. Table 1 reports the number of trips contained in each cluster for both case studies. It can be noted that trips of case study A are split almost uniformly among the four clusters; instead, cluster 3 of case study B contains more than half trip observations. In both case studies, cluster 3 and 4 contain the highest number of observations. This means that trip symptoms tend to more frequently arise closer to trip occurrence, thus challenging their detection and prediction.

Table 1. Cluster sizes [15].

Cluster	**Cluster size**	
	Case study A	*Case study B*
1	8	4
2	9	4
3	13	21
4	14	10
Total	44	39

3.3. Analyses

With respect to [15], in this paper two brand-new analyses, namely "Analysis #1" and "Analysis #2", which are summarized in Table 2, are carried out by varying the dimension of the trigger time window: the former considers different time windows during the test phase of the models developed in [15], while the latter consists of training new models by reducing the maximum length L_{max} of the trigger time window by keeping the same test configuration as in [15]. Therefore, Analysis #1 exploits the maximum amount of information during training, while it is progressively decreased in the test phase. Instead, for each trigger position, Analysis #2 trains models with different amounts of information while each model is subsequently tested by always exploiting the maximum amount of information.

More in detail, the first analysis considers the models developed in [15] and trained by considering a time window of length $L = L_{max} = 24$ hours before and after the trigger for scenarios from 1 through 25 and $L = T$, i.e., the number of remaining hours between the trigger and the time point at which trip occurs, from scenarios 26 through 48. For each trigger scenario, the models developed in [15] exploit the longest time window before and after the considered trigger time point. These models are then tested on new trips by progressively reducing the length of the time window before and after a given trigger position, from L_{max} to 1 for trigger positions $T = 48$ through $T = 24$ and from T to 1 when $T \leq 23$.

The second analysis is challenged to detect the onset of trip symptoms by reducing the length of the trigger time window when training the data-driven models. During the training phase, different values of L_{max} are considered, i.e., 24 (which is the case reported in [15] and considered in Analysis #1), 18, 12, 6 and 1. In fact, the more L_{max} decreases, the farther the sequence representative of Trip condition to trip occurrence and the lower the number of models trained by also including the hour right before trip occurrence. The test is conducted as in [15] by considering for each training scenario a maximum time window of 24 hours when the first 25 scenarios are simulated and equal to T for the remaining scenarios, i.e., by exploiting the maximum amount of information during test.

These new analyses are carried out to evaluate the influence of trigger window length on the reliability of the detection of trip symptoms. As can be grasped from Table 2, the training and test phase switch in the two Analyses. In fact, the test phase of Analysis #2 is performed in the same way as the training phase of Analysis #1, while some of the scenarios considered during the test phase of Analysis #1 are employed to train the models of Analysis #2. Therefore, the two analyses are complementary and allow to check whether most of information on GT trip symptoms is included in the day of trip occurrence.

4. Results and discussion

This Section presents and discusses the results for the two analyses described in Section 3. For both case studies, the results are presented in terms of *Ipp* by partitioning the available trip into the four clusters identified in [15]. The prototype of each cluster is calculated as the average of *Ipp* values among the trips included in a given cluster. In [15], a most likely trigger time point for each cluster was identified as the trigger position that provided the maximum of the prototype of that cluster (see Table 3). The values in Table 3 refer to the case $L_{max} = 24$.

Table 2. Analyses conducted in this paper.

		Analysis #1	**Analysis #2**
	Trigger position, T	Time window length, L	Time window length, L
Training	$T = \{48, 47, \ldots, L_{max}\}$	$L = L_{max} = 24$	$L = L_{max} = \{24, 18, 12, 6, 1\}$
	$T = \{L_{max}-1, \ldots, 1\}$	$L = T$	$L = T$
Test	$T = \{48, 47, \ldots, L_{max}\}$	$L = \{L_{max}, L_{max}-1, \ldots, 1\}$	$L = L_{max} = 24$
	$T = \{L_{max}-1, \ldots, 1\}$	$L = \{T, T-1, \ldots, 1\}$	$L = T$

Table 3. Most likely trigger time point of each cluster [15].

Cluster	**Most likely trigger position**	
	Case study A	*Case study B*
1	42	42
2	31	32
3	16	14
4	7	6

4.1. First analysis

Figures 2 and 3 report the *Ipp* values of the prototype of four clusters of trips for 48 different trigger positions as a function of the length L of the trigger time window for case study A and B, respectively.

In both case studies, regardless of the considered cluster, for a given trigger position T, the values of *Ipp* decrease by decreasing the length L of the trigger time window. The variation of *Ipp* is more evident if the considered trigger position T is equal or close to its most likely value (see Table 3); otherwise, the length L of the trigger time window is not influential on *Ipp*. This reasonably occurs since *Ipp* is sensitive to a change in the performance of the GT, while its value tends to 0 before the occurrence of this change or once this change is occurred. In fact, *Ipp* is generally lower than 0.3 for cluster 1 if T is lower than 24 or for clusters 3 and 4 if T is higher than 24.

It can be noted that, if the most likely trigger time point is considered, the decrease is more rapid passing from cluster 1 to cluster 4. In fact, for cluster 1, the value of *Ipp* remains as high as approximately 0.8 even if the length L is reduced from 24 to 19 or 13 for case study A or B, respectively; then the value of *Ipp* gradually decreases up to approximately 0.1 if $L = 1$. This means that symptoms of trips included in cluster 1 are likely to appear even 23 and 30 hours before trip occurrence for case study A and B, respectively. Instead, if clusters 3 and 4 are considered, the decrease of *Ipp* (at T equal to 16 or 7 for case study A and T equal to 14 and 6 for case study B) is higher than 0.3 and 0.2 for case study A and B, respectively, just by reducing the length of the trigger time window of one hour with respect to the maximum available value, which is equal to T.

In addition, Figures 2 and 3 point out that symptoms of trips of clusters 1 and 2 appear earlier and more gradually than those of clusters 3 and 4. In fact, the highest values of *Ipp* cover a time span that is larger than that of clusters 3 and 4 (multiple pairs of T and L provide an high value of *Ipp*), where *Ipp* achieves values close to or higher than 0.8 only if the length of the trigger time window is equal to its maximum value ($L = T$), i.e., it includes the hour of operation right before trip occurrence. For trips of clusters 3 and 4, symptoms arise abruptly when approaching the time point of trip occurrence.

4.2. Second analysis

Figures 4 and 5 show the prototypes of the four clusters by varying the amount of information used to train the predictive data-driven models. The black curves refer to the longest time windows ($L = 24$ for $T > 24$ scenarios and $L = T$ for $T \leq 24$) during training and were already reported in [15]. The main general outcome of this analysis is that, regardless of the cluster and the case study, the confidence with which the most likely trigger time point is provided for a given cluster reduces by reducing L_{max}. The overall behavior of the curves obtained with L_{max} lower than 24 reflects that of the reference curves ($L_{max} = 24$). However, in both case studies, the decrease of *Ipp* is higher than 0.2 just by reducing the maximum length of the trigger time window to 18. Moreover, if L_{max} is lower than 6 the values of *Ipp* are quite independent of the considered trigger position, and lower than 0.4 regardless of the considered trigger position.

By decreasing the value of L_{max} from 24 to 1 implies that i) the time window after the trigger of the first 24 scenarios ends farther from trip occurrence and ii) a smaller number of trained models exploits the data until the time point of trip occurrence. Since the decrease with respect to the reference case is significant, the results convey that most of valuable information to capture the signs of trip occurrence is included in the data points collected right before trip occurrence.

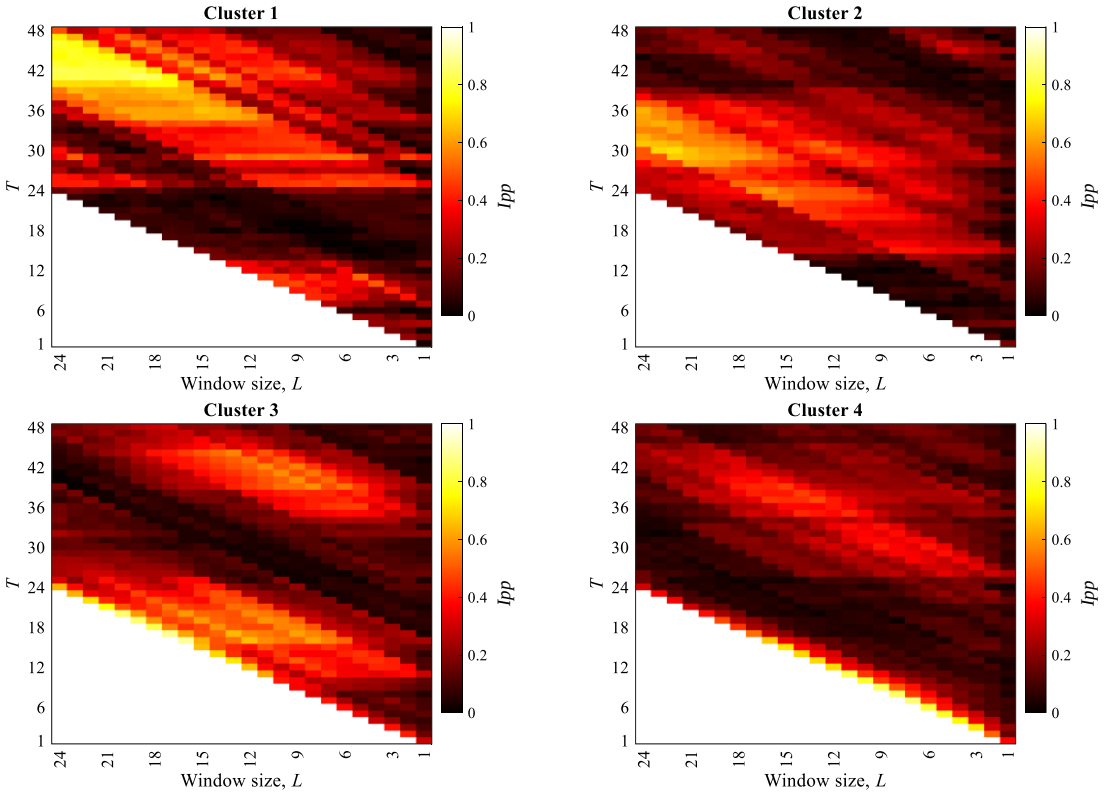

Figure 2. Analysis #1 for case study A. Colors reflect the value of *Ipp*.

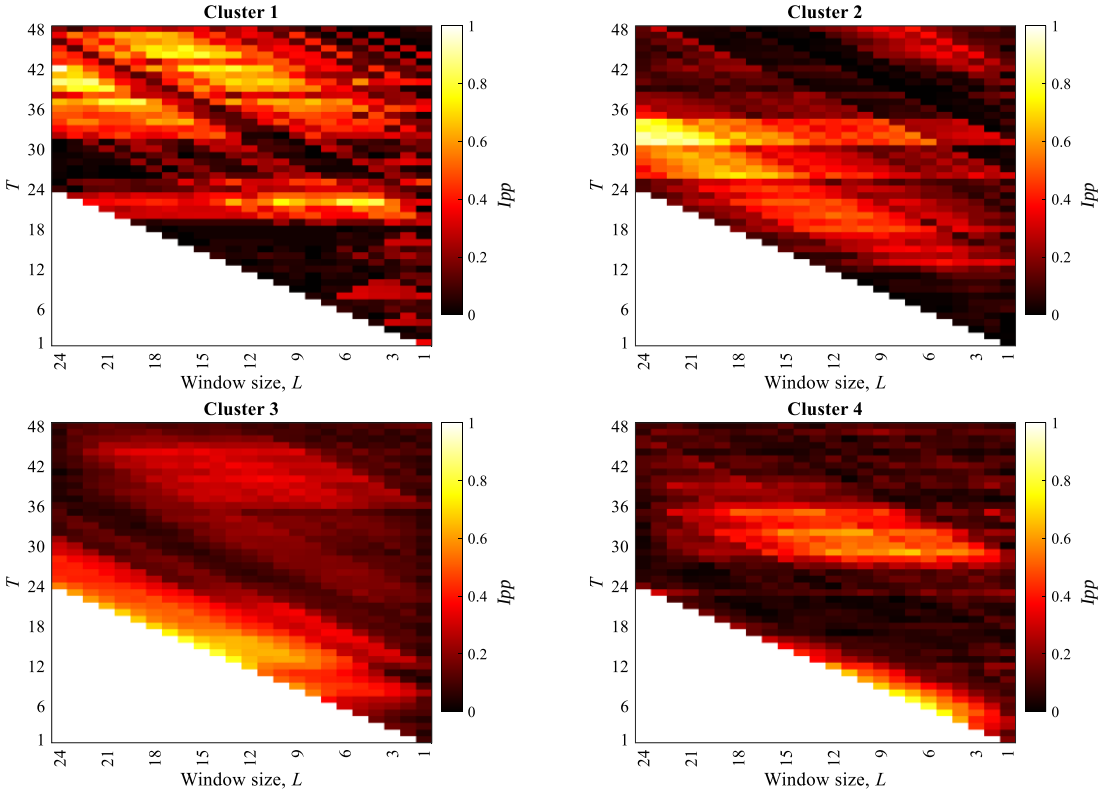

Figure 3. Analysis #1 for case study B. Colors reflect the value of *Ipp*.

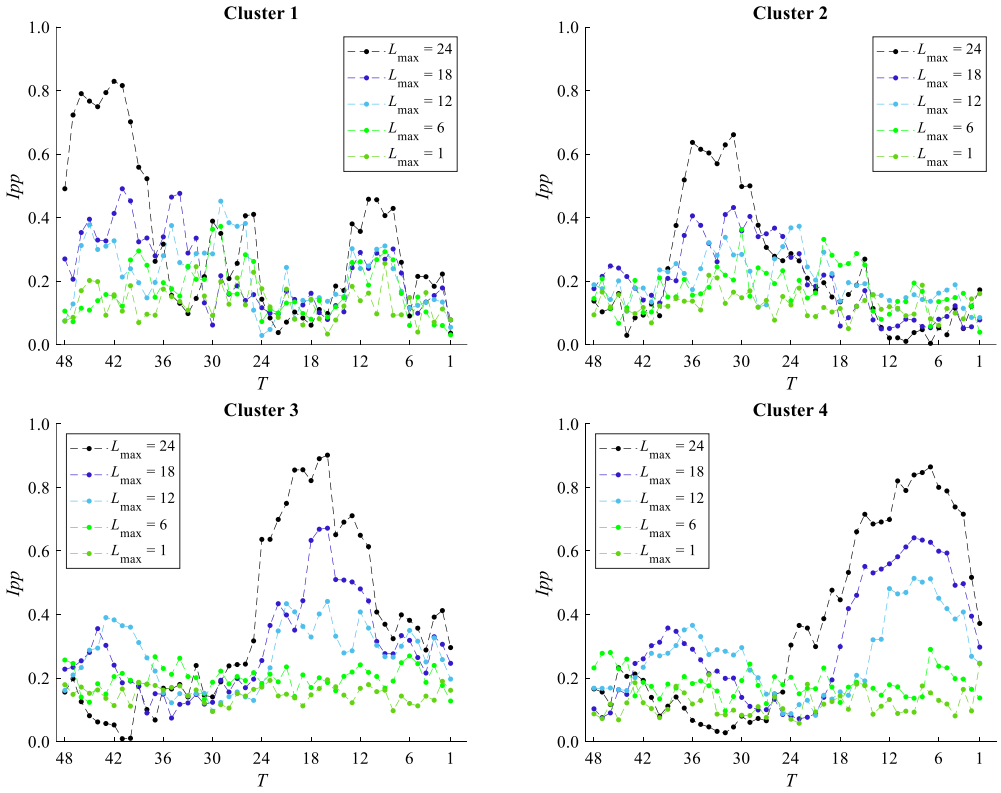

Figure 4. Analysis #2 for case study A: *Ipp* trends for different values of L_{max}.

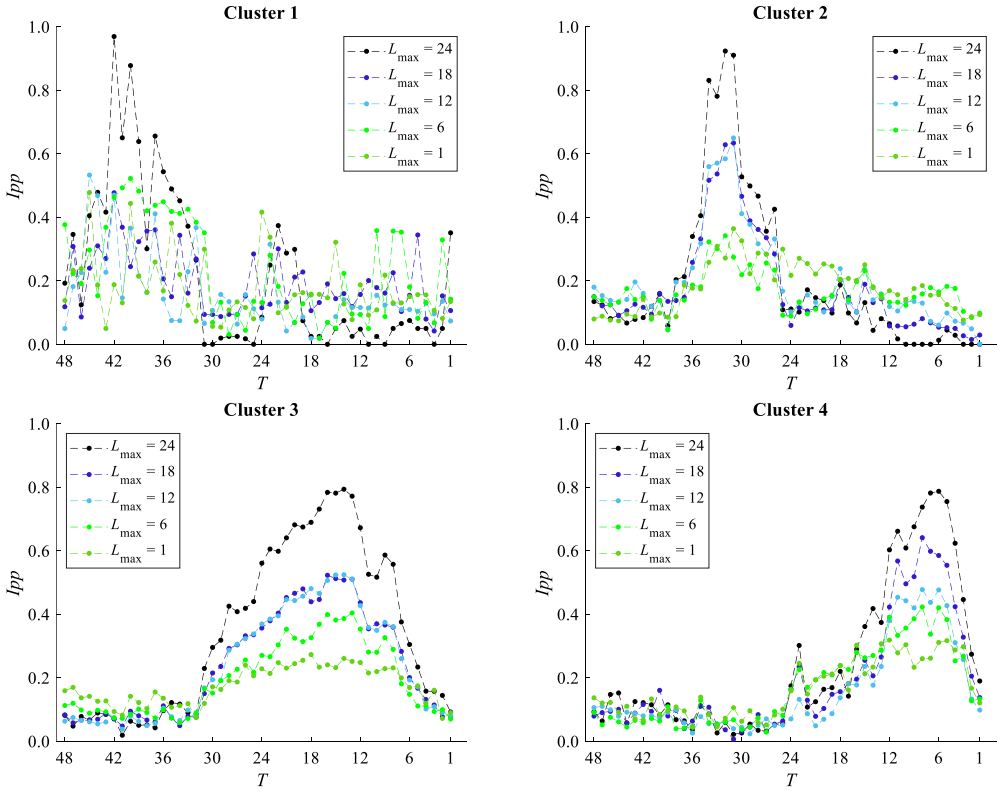

Figure 5. Analysis #2 for case study B: *Ipp* trends for different values of L_{max}.

5. Conclusions

This paper applied a data-driven methodology developed by the authors in a previous study by performing a systematic analysis of the influence of the trigger time window on the identification of the onset of gas turbine trip symptoms.

The data-driven methodology considered different trigger scenarios, from 48 hours to just 1 hour before trip occurrence, and two consecutive time windows of the same length, where the symptoms are supposed to be absent (before the trigger) and present (after the trigger). For each trigger scenario, a Long Short-Term Memory neural network was trained and tested on new trips to provide their most likely trigger time point.

The paper presented two different analyses, namely Analysis #1 and Analysis #2: the former considered the maximum amount of data during training and different trigger time windows were investigated during testing by progressively reducing the length of the time windows within which searching for the symptoms of gas turbine trip; the latter investigated different time windows during training for each trigger scenarios and then tested the models by always considering the longest available time window.

The results demonstrated that GT trip symptoms mainly arise on the day of trip occurrence, thus implying that the detection of trip symptoms is significantly influenced by the length of the two time windows before and after the trigger time point. Analysis #1 demonstrated that the confidence on the most likely trigger time point decreases from at least 0.8 to approximately 0.1 if the time window is reduced up to 1 hour. Analysis #2 demonstrated that the longest time windows have to be exploited during training in order to take into account the data that are closest to trip occurrence and contain most of trip symptoms. In fact, a reduction of the training window from 24 hours to 18 hours decreased the confidence on the most likely trigger time point by up to 0.3; this effect is magnified if the length of the training time window is further reduced. Therefore, the results of Analysis #1 and Analysis #2 validate each other. However, it has to be highlighted that performing Analysis #2 is more computationally expensive.

Acknowledgments

The authors gratefully acknowledge Siemens Energy for the permission to publish the results. The authors also wish to express sincere gratitude to Dr. Giuseppe Fabio Ceschini for inspiring and supporting the research activity about gas turbine trip prediction.

Nomenclature

c	cluster label
Ipp	prediction confidence
L	time window length
N_T	number of trigger scenarios
T	trigger
T^*	most likely trigger position

References

[1] Tahan M, Tsoutsanis E, Muhammad M, and Abdul Karim, Z A, 2017 *Applied Energy* **198** pp. 122 - 144.
[2] Wen Y, Rahman Md F, Xu H, Tseng T.-L B 2022 *Measurement* **187** 110276.
[3] Manservigi L, Venturini M, Ceschini G F, Bechini G, Losi E, *2020 ASME J. Eng. Gas Turbines Power* **142** 021009 (15 pages).
[4] Losi E, Venturini M, Manservigi L, Ceschini G F, Bechini G, 2019 *ASME J. Eng. Gas Turbines Power* **141**(11), 111019 (9 pages).
[5] Manservigi L, Venturini M, Losi E, Bechini G, Artal de la Iglesia J, 2022, *Machines* **10**, 228.
[6] Manservigi L, Murray D, Artal de la Iglesia, J, Ceschini G F, Bechini G, Losi E, Venturini M, 2022, *ISA Transactions*, **123** pp. 323-338.
[7] Bhargava R K, 2017, Technical Dictionary on the Gas Turbine Technology, Innovative Turbomachinery Technologies Corp, Katy, TX.

[8] Graichen, C. M., and Cheetham, W. E., "Case-Based Reasoning Approaches for Gas Turbine Trip Diagnosis.", Proc. ASME Turbo Expo 2007, ASME Paper No. GT2007-27856.
[9] Ravi, Y. B., Pandey, A., Jammu, V., "Prediction of Gas Turbine Trip due to Electro Hydraulic Control Valve System Failures.", Proc. ASME Turbo Expo 2010, ASME Paper No. GT2010-23228.
[10] Losi E, Venturini M, Manservigi L, Ceschini G F, Bechini G, Cota G, and Riguzzi F, 2021, *ASME J. Eng. Gas Turbines Power* **143**(3): 031014.
[11] Losi E, Venturini M, Manservigi L, Ceschini G F, Bechini G, Cota G, and Riguzzi F, "Data Selection and Feature Engineering for the Application of Machine Learning to the Prediction of Gas Turbine Trip", Proc. ASME Turbo Expo 2021, ASME Paper No. GT2021-58914.
[12] Losi E, Venturini M, Manservigi L, Ceschini G F, Bechini G, Cota G, and Riguzzi F, 2022 *ASME J. Eng. Gas Turbines Power* **144**(3), 031025 (13 pages).
[13] Losi, E., Venturini, M., Manservigi, L., Bechini, G., 2022, "Ensemble Learning Approach to the Prediction of Gas Turbine Trip", Proc. ASME Turbo Expo 2022, June 13 - 17, Rotterdam, The Netherlands. ASME Paper GT2022-80372.
[14] Bechini, G., Losi, E., Manservigi, L., Pagliarini, G., Sciavicco, G., Stan, E. I., Venturini, M., "Temporal Random Forest Applied to Gas Turbine Trip Prediction.", Proc. ASME Turbo Expo 2022, June 13 - 17, Rotterdam, The Netherlands. ASME Paper No. GT2022-82915.
[15] Losi, E., Venturini, M., Manservigi, L., Bechini, G., 2022, "Detection of the Onset of Trip Symptoms Embedded in Gas Turbine Operating Data", Proc. ASME Turbo Expo 2022, June 13 - 17, Rotterdam, The Netherlands. ASME Paper GT2022-80666.
[16] Bai M, Liu J, Ma Y, Zhao X, Long Z, Yu D, 2021 *Energies* **14** 13.
[17] Zhou H, Ying Y, Li J, Jin Y, 2021 *Advances in Mechanical Engineering*
[18] Sheng X, Yi Q, Jun L, Huayan P, Baoping T, 2016 *Reliability Engineering & System Safety* **216** 107927.
[19] Zhang J, Wang P, Yan R, and Gao R X, 2018, *Procedia CIRP*, **72**, pp. 1033 - 1038.
[20] Wu Y, Yuan M, Dong S, Lin L, and Liu Y, 2018 *Neurocomputing* **275** pp. 167 - 179.
[21] Zhao S, Zhang Y, Wang S, Zhou B, Cheng C, 2019 *Measurement* **146**, pp. 279 - 288.
[22] Zhang A, Wang H, Li S, Cui Y, Liu Z, Yang G, Hu J., 2018 *Appl. Sci.* **8** 2416.
[23] Killick R, Fearnhead P, Eckley I A, 2012 *Journal of the American Statistical Association* **107** pp. 1590 - 1598.
[24] Hochreiter S, and Schmidhuber J, 1997 *Neural computation* **9**(8) pp.1735 - 1780.
[25] Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R, 2014 *Journal of Machine Learning Research* **15** pp. 1929-1958.
[26] Diederik K, and Ba J, "Adam: A method for stochastic optimization.", 3rd International Conference for Learning Representations, San Diego, 2015. arXiv:1412.6980.