

A Challenging Dataset of Jet Engine Fault Scenarios

Christoforos Romesis^{1,*}, Stasinos Konstantopoulos^{1,†}

¹Institute of Informatics and Telecommunications, NCSR Demokritos, Patr.Gregoriou E & 27 Neapoleos Str, 15341 Agia Paraskevi, Greece

Abstract

This dataset description paper introduces a challenging sequence processing task. Specifically, the task is to recognize faults from aircraft gas turbine engine conditions where previous faults determine how the presently observed conditions should be interpreted. Several state-of-the-art deep-learning sequence processors have been tested on the task, and these preliminary results demonstrate that they cannot correctly model the phenomenon.

Keywords

time-series dataset, sequence processing, long-range dependencies, gas turbine faults

1. Motivation and Purpose

Time series forecasting and time series classification are the major *sequence processing* tasks, and both depend on a method's ability to identify time-based patterns in the data. When dealing with *long sequence* processing in particular, successful methods must identify non-local patterns such as trends, seasonality, and long-range dependencies. Observing large-scale and high-impact repositories such as the *Monash Time Series Forecasting Archive* [1] reveals an abundance of datasets that exhibit long-term trends and long-period seasonal patterns, but a lack of datasets that exhibit non-periodic long-range dependencies. Unsurprisingly, successful methods stem from the non-parametric statistics field or from the deep learning field.

The sequential integration of sub-symbolic and symbolic modules allows connectionist recognition of single events to interact with long-range dependencies between events. Such *neuro-symbolic* approaches are powerful, but to demonstrate their effectiveness they must be applied to datasets and tasks where the following conditions hold simultaneously: (a) the recognition of individual events in the raw data is a non-trivial pattern recognition task for which connectionist methods are more appropriate; and (b) this recognition is affected by unknown symbolic patterns that need to be discovered in conjunction with the sub-symbolic patterns themselves.

In this dataset description paper we present such a task where purely connectionist sequence processing performs substantially below the par established through the usual trend and seasonality tasks and where neuro-symbolic methods can demonstrate their effectiveness in challenging real-world applications.

2. Dataset Description

Jet engines play a vital role in aviation, and ensuring their safety and reliability is paramount. The harsh environments

they operate in can cause malfunctions, impacting aerothermodynamic measurements. Detecting faults in-flight from time-series data remains challenging due to limited instrumentation, similar fault effects, concurrent faults, and prior events. This paper introduces a dataset of simulated time-series measurements from a turbofan engine, including various common faults, to aid in developing fault detection strategies. This dataset is valuable for the Machine Learning community, presenting a complex engineering problem with unique features. It includes 2,410 time series, each with 3,600 time steps, where faults are introduced, and measurements are taken at every step. Some faults have nearly indistinguishable effects, posing challenges for differentiation, while others exhibit long-term dependencies. These attributes make the dataset suitable for research in sequential and time-stepwise classification and investigating long-term dependencies. The dataset comprises time series measurements from a low bypass ratio turbofan engine, typical of modern civil aviation engines, generated using an Engine Performance Model. The dataset has been archived and is publicly accessible from <https://doi.org/10.5281/zenodo.15856441>

2.1. Engine Performance Model

An EPM interrelates parameters that represent engine component health and operating conditions with measurements performed on an engine [2], and can be expressed through Equation (1):

$$\mathbf{Y} = \mathbf{g}(\mathbf{u}, \mathbf{f}) \quad (1)$$

where \mathbf{g} is a vector function representing the EPM, \mathbf{u} is a vector of measured quantities defining the engine's operating point, \mathbf{Y} is a set of measurements for condition monitoring, and \mathbf{f} is a vector of engine component *health parameters*. Typically, two health parameters are used for each component: the flow factor (SW), indicating the component's swallowing capacity, and the efficiency factor (SE), representing thermal efficiency. The application of such parameters for assessing engine component health has been discussed in [3]. A deviation of a health parameter from its nominal value signals a fault in that component. As extensively discussed in [4] the ratio of deviations between two health parameters can serve as a characteristic metric for faults such as fouling, increased tip clearance, erosion, and foreign object damage. The severity of the fault correlates with the level of this deviation. The deviation (or *delta*) of a health parameter $\Delta \mathbf{f}$ is defined as:

ANSyA 2025: 1st International Workshop on Advanced Neuro-Symbolic Applications, co-located with ECAI 2025.

*Corresponding author.

†These authors contributed equally.

✉ chris.romesis@iit.demokritos.gr (C. Romesis);

konstant@iit.demokritos.gr (S. Konstantopoulos)

🌐 <https://github.com/data-eng> (C. Romesis);

<https://github.com/data-eng> (S. Konstantopoulos)

🆔 0009-0001-6485-5548 (C. Romesis); 0000-0002-2586-1726

(S. Konstantopoulos)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

$$\Delta \mathbf{f} (\%) = \frac{\mathbf{f} - \mathbf{f}_o}{\mathbf{f}_o} \cdot 100\% \quad (2)$$

where \mathbf{f} denotes the value of the health parameter, and \mathbf{f}_o its nominal value.

Using the EPM, the nominal values of the available measurement set, denoted as \mathbf{Y}_o , can be computed for a specified operating point \mathbf{u} . These values correspond to the nominal health parameters, \mathbf{f}_o , and are derived via the following relation:

$$\mathbf{Y}_o = \mathbf{g}(\mathbf{u}, \mathbf{f}_o) \quad (3)$$

Once a measurement vector \mathbf{Y} is obtained from the engine, the relative deviation $\Delta \mathbf{Y}(\%)$ from its nominal value \mathbf{Y}_o is defined as:

$$\Delta \mathbf{Y}(\%) = \frac{\mathbf{Y} - \mathbf{Y}_o}{\mathbf{Y}_o} \cdot 100\% \quad (4)$$

To emulate realistic measurement conditions, random noise is superimposed on the measurement deviations. The noise levels considered are consistent with those typical of this instrumentation, as documented in [5].

2.2. Dataset Generation

The EPM used for dataset generation was developed in MATLAB. At each time step, the data includes a labeled vector of measurement deltas produced by the EPM at a specified operating point and a specific engine health condition. A total of 10 operating points were considered, covering a broad range of the engine's flight envelope. Throughout a time series, all time steps correspond to the same operating point. Additionally, five distinct health conditions were considered, as summarized in Table 1. The first health condition represents the nominal, healthy operation of the engine. The second health condition (coded as TIP) models an increased clearance of the compressor blades. As reported in [4], this fault results in a ratio of the related health parameter deviations, $\Delta SW/\Delta SE$, of 0.2; the actual deviations of ΔSW and ΔSE are around -1% and -5% , respectively. The remaining three conditions involve a mistuning of the inlet guide vanes of the compressor, damage caused by foreign objects within the compressor, and the concurrent occurrence of VGV and FOD faults. These conditions result in a ratio of -3 , 1 , and 0.2 , respectively. In all cases, deviations within the range of 0.7 to 1.3 times the aforementioned values are considered, simulating faults of varying severity. Notably, the effects of TIP faults and the combined VGV+FOD faults on the compressor's health parameters are the same, and therefore, measurements under these conditions follow the same distribution. Each dataset record is labeled according to the specific health condition present at that time step.

Each time series comprises 3,600 time steps. During this period, various health conditions may develop. A time series without any fault contains measurements indicative of healthy engine operation at all time steps. In other time series, the engine operates fault-free initially, but at a certain point, a TIP, VGV, or FOD fault occurs and persists until the end of the series. In some other time series, the engine also operates fault-free initially, but at a certain point, a brief period occurs during which a VGV fault is present. The

Table 1
Considered health conditions.

label	health condition	ΔSW	ΔSE	$\Delta SW/\Delta SE$
1	Healthy	0%	0%	N/A
2	TIP fault	-1%	-5%	0.2
3	VGV fault	+3%	-1%	-3
4	FOD fault	-4%	-4%	1
5	VGV+FOD faults	-1%	-5%	0.2

engine then continues to operate without faults until a later time step, when a combined VGV+FOD fault manifests and persists for the remainder of the time series. This scenario exemplifies a specific fault condition characterized by a severe VGV fault leading to FOD. The damage is attributed to the detachment of a particle from the VGV mechanism, such as a loosened bolt, which is subsequently ingested by the engine. The failure of the VGV mechanism may be preceded by a transient VGV fault event, serving as an early indicator of the impending failure. This scenario was selected because both the TIP and VGV+FOD faults exert similar impacts on engine performance. The key differentiator between the TIP fault and the VGV+FOD fault is the long-term dependency of the latter on the initial, instantaneous VGV fault.

3. Concluding Remarks and Next Steps

The dataset is used to evaluate the classification capabilities of sequential machine learning methods on time-series data with long-term dependencies. Preliminary results suggest the dataset challenges current techniques. For example, Figure 1 shows four classification methods trained on the dataset and tested on a time-series segment containing a VGV+FOD fault. The methods, include an RNN, a GRU, and an LSTM, together with an MLP, trained for point-to-point classification for comparison. All methods failed to detect the VGV+FOD fault, though they detected a precursor VGV fault at time step 740. Sequential methods have shown limited ability to use information from prior time steps, performing similarly to the point-wise MLP.

Acknowledgments

This research was co-funded by the European Union under GA no. 101135782 (MANOLO project). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or CNECT. Neither the European Union nor CNECT can be held responsible for them.

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

- [1] R. W. Godahewa, C. Bergmeir, G. I. Webb, R. Hyndman, P. Monero-Manso, Monash time series forecasting archive, in: Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks

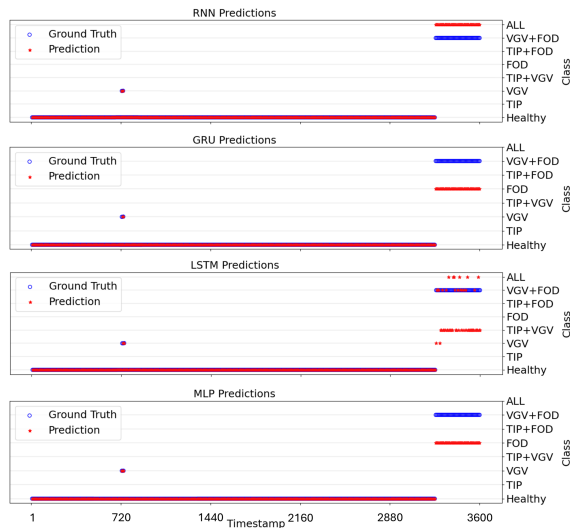


Figure 1: Health conditions detection during a time-series of the dataset.

- Track (Round 2), 2021. URL: <https://openreview.net/forum?id=wEc1mgAjU->.
- [2] A. Alexiou, K. Mathioudakis, Development of Gas Turbine Performance Models Using a Generic Simulation Tool, in: Volume 1: Turbo Expo 2005, ASMEDC, Reno, Nevada, USA, 2005. doi:10.1115/gt2005-68678.
 - [3] K. Mathioudakis, A. Stamatis, A. Tsalavoutas, N. Aretakis, Performance analysis of industrial gas turbines for engine condition monitoring, Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy 215 (2001) 173–184. doi:10.1243/0957650011538442, publisher: SAGE Publications.
 - [4] K. Mathioudakis, A. Alexiou, N. Aretakis, C. Romesis, Signatures of Compressor and Turbine Faults in Gas Turbine Performance Diagnostics: A Review, Energies 17 (2024) 3409. doi:10.3390/en17143409, publisher: MDPI AG.
 - [5] C. Romesis, N. Aretakis, K. Mathioudakis, Model-Assisted Probabilistic Neural Networks for Effective Turbofan Fault Diagnosis, Aerospace 11 (2024) 913. doi:10.3390/aerospace11110913, publisher: MDPI AG.