CRANFIELD UNIVERSITY

Simone Togni

An integrated combined methodology for the online gas turbines performance-based diagnostics and signal failure isolation

School of Aerospace, Transport and Manufacturing
Department of Power and Propulsion

PhD
Part-Time
Academic Year: 2019 - 2020

Main supervisor: Dr Nikolaidis Theoklis
Co-supervisor: Dr Suresh Sampath

CRANFIELD UNIVERSITY

School of Aerospace, Transport and Manufacturing
Department of Power and Propulsion

PhD
Part-Time

Academic Year 2019 - 2020

Simone Togni

An integrated combined methodology for the online gas turbines performance-based diagnostics and signal failure isolation

Main supervisor: Dr Nikolaidis Theoklis
Co-supervisor Dr Suresh Sampath
November 2019

This thesis is submitted in partial fulfilment of the requirements for the degree of MPhil/PhD

ABSTRACT

The target of this research is the performance-based diagnostics of a gas turbine for the online automated early detection of components malfunctions with the presence of measurements malfunctions. The research proposes a new combination of multiple methodologies for the performance-based diagnostics of single and multiple failures on a two-spool engine. The aim of this technique is to combine the strength of each methodology and provide a high rate of success for single and multiple failures with the presence of measurement malfunctions – measurement noise. A combination of Kalman Filter, Artificial Neural Network, Neuro-Fuzzy Logic and Fuzzy Logic is used in this research in order to improve the success rate, to increase the flexibility and the number of failures detected and to combine the strength of multiple methods to have a more robust solution. The Kalman Filter has in his strength the measurement failure treatment, the artificial neural network the simulation and prediction of reference and deteriorated performance profile, the neuro-fuzzy logic the estimation precision, used for the quantification and the fuzzy logic the categorization flexibility, which are used to classify the components failure. All contributors are also a valid technique for online diagnostics, which is a key objective of the methodology. In the area of gas turbine diagnostics, the multiple failures in combination with measurement issues and the utilization of multiple methods for a 2-spool industrial gas turbine engine has not been investigated extensively.

This research investigates the key contribution of each component of the methodology and reaches a success rate for the component health estimation above 92.0% and a success rate for the failure type classification above 95.1%. The results are obtained with the first configuration, running with the reference random simulation of 203 points with different level of deterioration magnitude and different combinations of failures type. If a measurement noise 5 times higher than the nominal is considered, the component health estimation drop to a minimum of 70.1% (reference scheme 1) while the classification success rate remains above 88.9% (reference scheme 1).

Moreover, the speed of the data processing – minimum 0.23 s / maximum 1.7 s per every single sample – proves the suitability of this methodology for online diagnostics.

The methodology is extensively tested against components failure and measurement issues. The tests are repeated with constant simulations, random simulation and a deterioration schedule that is reproducing several months of engine operations.

ACKNOWLEDGEMENTS

This thesis is dedicated to my wife Yaxi. Her encouragement, support, patience and sacrifice, have been outstanding during the entire phase of my studies. From her, I always had the right words in the right moments and when I asked for advice she always gave the best to help me going forward. Together, we shared the good moments, but also the sacrifices this extra work implied.

A special mention goes to my little daughter, Gaia. She came to the world during my studies and she has been, by far, the best achievement of my life. She brought to our family such a joy and extra power, who helped me completing these studies. Her shiny smile is such effective to blow away every problem or fatigue. I hope my dedication and effort would be of a good example for her as her joy and love for life have been a vital example for me.

I would like to thank my main supervisor, Prof. Nikolaidis Theoklis for his support both technical and personal. He has constantly addressed me in the right direction and helped me achieving objectives beyond my expectations. Moreover, he offered me the opportunity to participate in conferences and publish papers that taught me so much.

I also would like to thank my co-supervisor Prof. Suresh Sampath for his support and for his counsels. His advice, in particular, helped me critically thinking my work and step forward.

A special mention goes also to Prof. Sethi Bobby who first offered me the opportunity to undertake this PhD.

Finally, I also would like to thank my family for being always such proud of me and making me feel I did something important. Their sacrifices and support thorough out the years, allowed me being in the position of writing this thesis.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

ANFIS	Adaptive Neuro-Fuzzy Inference
ANN	Artificial Neural Network
BBN	Bayesian Belief Network
EGT	Exhaust Gas Temperature
EHM	Engine Health Monitoring
EKF	Extended Kalman Filter
ELM	Extreme Learning Machine
FF	Fuel Flow
FL	Fuzzy Logic
GA	Genetic Algorithm
GPA	Gas Path Analysis
GT	Gas Turbine
HP	High Pressure
HPC	High-Pressure Compressor
HPT	High-Pressure Turbine
HRSG	Heat Recovery Steam Generator
IATA	International Air Transport Association
KF	Kalman Filter
LP	Low Pressure

LPC	Low-Pressure Compressor
LPT	Low-Pressure Turbine
MISO	Multiple Input Single Output
MLKF	Multiple Layers Kalman Filter
MM	Multiple Measurements
MRO	Maintenance Repair Overhaul
NFL	Neuro-Fuzzy Logic
O&M	Operation and Maintenance
QPSO	Quantum-Behaved Particle Swarm Optimization
SA	Simulated Annealing
SLKF	Single Layer Kalman Filter
SVN	Support Vector Machine
TBO	Time Between Overhauls
TOPSIS	Technique for Order of Preferences

1.0 Introduction

1.1 Health monitoring from early stages to today

The evolution of the engine health monitoring went through a few fundamental stages in the last few decades. Before any electrical device could be introduced and knowledge of the gas turbine was still not wide enough for the users, the maintenance was unplanned. The monitoring was limited to visual inspections and on the run to failure technique [1]. This policy is very risky and could not represent a solution especially for aero engines, implying a high degree of safety.

The following step was the preventive maintenance that consists of scheduled maintenance intervals for a fixed number of operating hours. As reported by Marinai [2], in the early 50s the intervals of maintenance of reciprocating engines could be as low as 1000 operating hours, compared to the 30000 operating hours reported for the Solar Turbine Mars 90. Despite the increase of TBO, preventive maintenance has still margin for cost savings. Some replacements, in fact, might occur before the part is effectively at the end of its safe lifetime, and some replacements might occur too late, while the engine has already encountered an unplanned event.

With the advent of the electric devices in the 70s, the health monitoring system was introduced first in the aero-engines. Volponi [3] reported that the system was introduced in the aircraft and the parameters could be recorded. In the beginning, only flight parameters could be recorded, but after some additional parameters of inlet and outlet condition of the compressor and the temperature became available. It is from the 80s, when the electronic made a step forward, that the monitoring enhanced its capabilities. In fact, the data

were recorded on the inflight monitoring system and set to the ground for further analysis and for the predictive maintenance. The predictive maintenance consists of the prediction of the health status of the engine based on its operating data.

The advantages provided by the predictive maintenance are economical as reported in section 1.3, but also the safety can be significantly improved by interpreting early enough key information from the gas turbine.

1.2 Maintenance costs present and future scenario

The importance of health monitoring in the gas turbine industrial system and in the aero engines has grown in the last two decades. One of the motivations behind this growth is the economic advantage. As reported by Verbist et al. [4] in the last years, the technical enhancements of the gas turbines technology decreased fuel consumption. However, due to the fuel price increase over the same period of time, the fuel cost still counts for one-third of the operating expenses. The authors clarify that to reduce the operating costs, the engines are demanded to operate longer and reduce maintenance costs.

The contribution that health monitoring can attack is the maintenance cost of the engine, that represents an important portion of the operating costs. Looking at the civil aero engines, Marinai et al. [5] reported that the engines are responsible for the 26% of the operating costs and out of that the maintenance and overhaul count for 31%. The rest of the operating costs are coming from depreciation, 35% and from the fuel 34% (Figure 1).

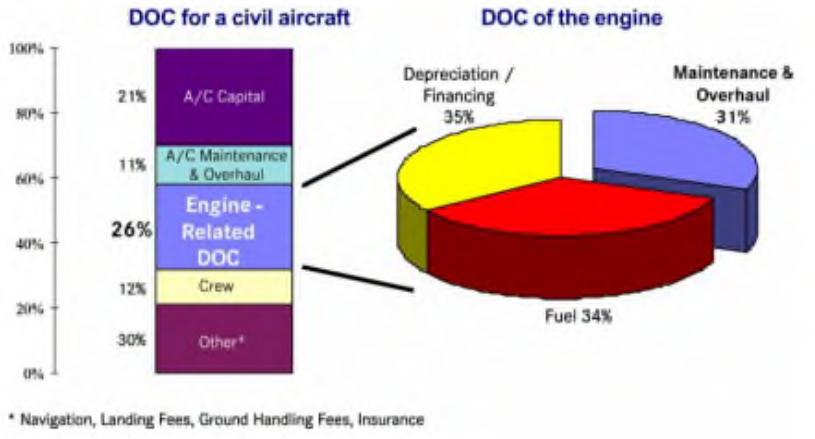

Figure 1 Overall costs and maintenance cost on civil engines – Courtesy of [5]

The same trend is seen in the industrial gas turbines where according to Grace [6] the planned maintenance costs on industrial gas turbines and combined cycle plants have a portion of 56% of the total O&M costs that represents the 7% of the overall project cash flow (Figure 2). The unplanned maintenance costs have been also quantified and they correspond to 8.3% of the total O&M costs (Figure 3).

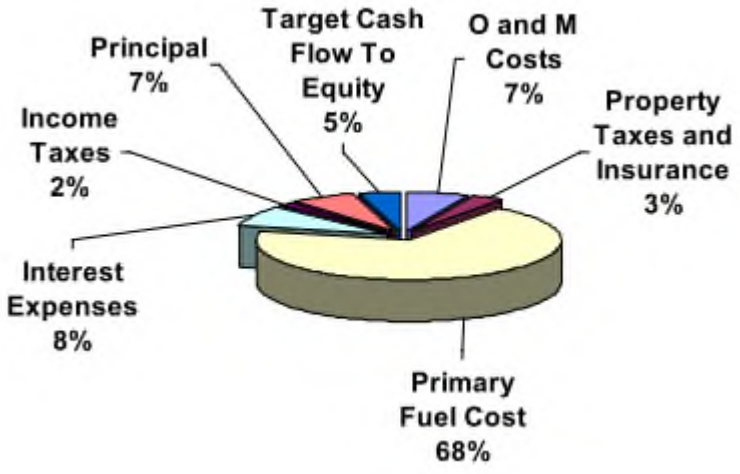

Figure 2 Distribution of project cash flow costs – Courtesy of [6]

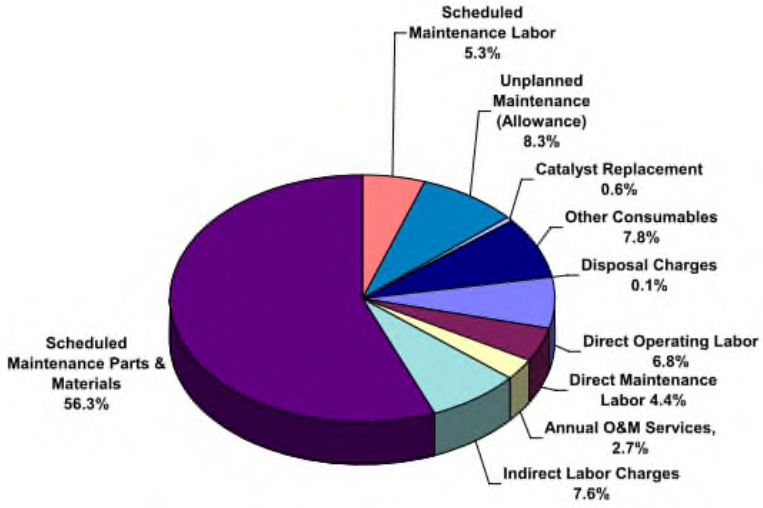

Figure 3 Distribution of maintenance costs – Courtesy of [6]

The main power plant components contributing to the cost considered by Grace in his publication [6] are the gas turbine, the HRSG and the steam turbine. The distribution of the outage time that they report is 79% for the gas turbine, which is by far the major contributor, 13% for the HRSG and 8% for the steam turbine.

Still, on the unplanned events, Grace et al. [7] conducted statistic studies for 3000 E and F class engines over a 15 years period of time. The authors concluded that the unplanned maintenance cost can reach 8% of the O&M costs, or 2% of net revenue income and the loss of revenue can reach the 15% of the O&M cost or 5% of net revenue income.

Looking at the future trends, IATA [8] provides in his report the figures of the global maintenance repair and overhaul spent in 2016, the date of the report, and proposes a forecast for 2026. The figure shows that the MRO costs will increase by circa 50% in ten years and the engine will be responsible for 41% of the total costs (Figure 4). It is clear from these numbers that the MRO is an increasing topic for the years to come and IATA also add that the big data analytics to support maintenance prediction, will be a growing topic.

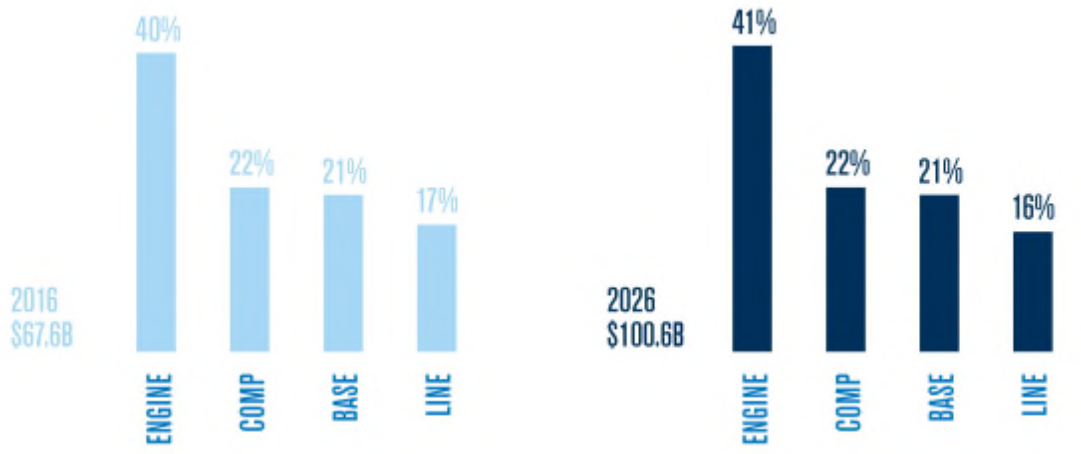

Figure 4 Distribution of maintenance costs – Courtesy of [8]

1.3 Benefits of engine health monitoring

As confirmed by Volponi [3] the maintenance strategy running two decades ago, was based on scheduled maintenance. The rapid improvement of the monitoring and diagnostics capabilities provided the possibility to the industrial parties to move from the scheduled maintenance strategy to the maintenance performed conditions. The author is providing evidence for the airlines, but the same advantages can be considered for the industrial-scale gas turbines.

The new capability supplied by the health monitoring systems is the provision of early information about the health status of the engine. The benefits of this information can be evaluated from different perspectives.

- As clarified in section 1.2 the economic advantage is clearly an additional value. In this case, the earlier knowledge of the status of the gas turbine can reduce maintenance costs;
- The efficiency costs related can be also optimized by better planning the engine components cleaning;
- The safety can be increased by knowing the status of the engine before any fatal failure;

- The lifetime extension is a possible benefit if there is sufficient information from the monitoring system to neglect the original overhaul schedule.

The benefits provided by each health monitoring techniques must be leveraged with the costs related to them. In fact, some of the techniques require additional signals that can come only from the installation, at a cost, of additional probes. The additional information provided must be interpreted in a manual or automated way. This step comes also with the cost of the monitoring system and of the specialist committed to do it.

2.0 Gas Turbine Diagnostics

2.1 Diagnostics techniques

2.1.1 Visual inspection

The visual inspection consists of the examination of the gas turbine part with the aid of optical assistance boroscopes or with the unaided eyes. The first can be used to access small areas or to access the internal part of the machine without opening it. The second is used once the machine is open or to visualize external parts and might be indicated to detect visible failures like corrosion, part detachment and superficial fouling/erosion.

2.1.2 Radiography

The radiography makes use of X-ray to detect hidden failures on internal components. A common example of this application is the inspection of the aero gas turbine blades for the detection of defects in the cooling passages. This technique can be integrated with the visual inspection to combine the analysis of the surfacing and the detection of internal defects.

2.1.3 Thermography analysis

The thermography is the analysis of the temperature scale based on a colour scale. By making use of cameras, the image of the object under investigation is combined with temperature (Figure 5). This combination of information allows detecting internal and external parameters by analysing the external surface. A hot spot detected by the thermal camera could detect wear in some parts of the gas turbine or inadequate lubrication in a bearing.

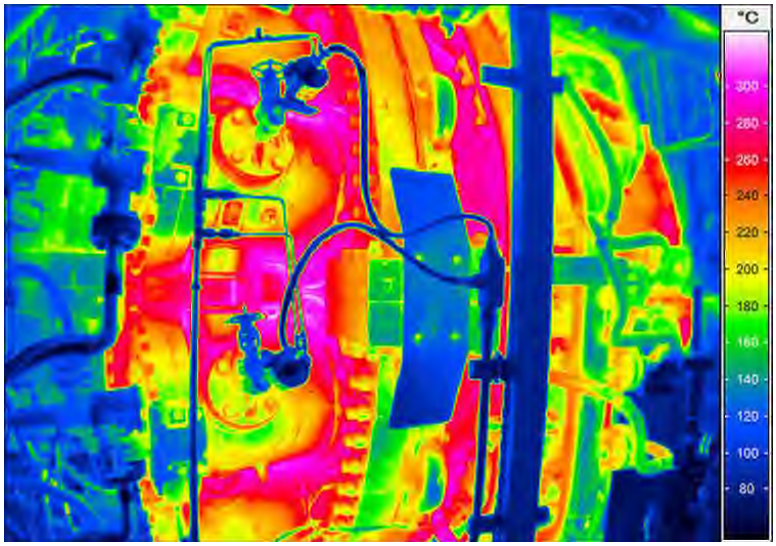

Figure 5 Thermographic images – Courtesy of InfraTec available at https://www.infratec-infrared.com/thermography/industries-applications/thermal-optimisation/

2.1.4 Vibration analysis

The vibration analysis is one of the early techniques implemented on the rotating machinery for the diagnostics. As detailed in the book “Machinery vibration and rotordynamics” [9], the first dynamic problem that has been encountered was the critical speed, where the speed and the natural frequency of the rotor match, generating high vibration levels. This first challenge shown the forces a rotor is subjected to that have been later detailed by Rankine. The governing equation (1) that allowed Rankine to calculate the eigenfrequency of a rotor is:

$$\omega = \sqrt{\frac{k}{m}} \qquad (1)$$

Where ω is the eigenfrequency of the rotor, k is its stiffness and m is its mass. This equation is the foundation of the vibration analysis that checks which force is the cause of vibration. From the base equation, the only force involved the excitation is the mass of the system that while rotating generate a force at a certain frequency, the function of the rotational speed. Other forces are generated by unbalance, misalignment, torsional coupling, shaft bending,

shaft cracks, gear wear, rub, hydrodynamic forces, aerodynamic forces, component damage or looseness. These are causes of vibration and can be detected by the vibration analysis.

The vibration manifests itself with amplitude, that defines the magnitude of the vibration, the phase, that defines when the vibration occurs relative to another or to a reference and the frequency, that define the repetition rate of the vibration. The amplitude can be measured in terms of displacement, velocity or acceleration and for each of these measurements, a different type of sensor will be used. Once the information is available throughout a monitoring system, the root cause can be determined by analysing the magnitude of the vibration, its frequency and the phase.

2.1.5 Lube oil analysis

The lube oil analysis found its principle on the observation of the wear particles raised by the use of the gas turbine or by some additional malfunctions. The foundation is that the wear particles would deposit in the machine and later collected by the lubrication system. Once the sample is collected, the characteristics of the debris and the lube oil system can be determined.

A typical example described in the “Handbook of condition monitoring” [10] reports a case of gas turbine lube oil monitoring. The turbine is reported to have large plain white metal journal bearings. The spectrographic analysis used to evaluate the wear debris detect an increasing level of lead metal and copper. This is reported to be a clear indication of the white metal wear and its detection allow to prevent additional damages of the rotor shaft.

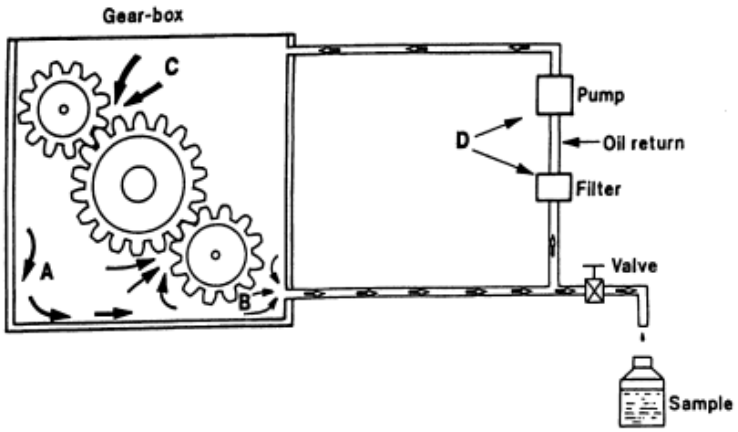

Figure 6 Debris collection in the lubricating system – Courtesy of [10]

2.1.6 Performance monitoring

The methods detailed before are mostly focused on the detection of the gas turbine components failure that is remarked by signals or information provided by the machine or by the monitoring system. However, a gas turbine is already able to provide signals and information that can be used to draw conclusions on the status of the machine. For instance, the parameters already measured on a gas turbine are the pressure, the temperature, the fuel mass flow, the thrust in case of aero engines and the power for industrial gas turbines. Combining these parameters in a thermodynamic model, it is possible to derive additional information such as efficiency. Comparing the calculated results with the reference model, it is possible to calculate a delta from the reference values and determine the health status of each component.

While acting on internal parameters, this technique is able to provide information about the component status well before the failure. By comparing the physical model with the reference one, in fact, it is possible to draw the deterioration profile, from the new installation up to the latest recorded sample. For this reason, the performance evaluation is not only used for the diagnostic evaluation, but also for the economic evaluation of the maintenance feasibility.

The available existing techniques, together with their advantages and drawbacks are well listed and explained by Marinai et al. [5]. Each of the options explained has his own advantage and drawback but no one of them

alone is able to solve all the problems the gas path analysis should deal with. To analyse them in detailed a dedicated section will follow (section 3.3).

3.0 Contribution to knowledge

3.1 Introduction

Before clarifying the additional knowledge, this thesis aims to give, it is important to remark the motivation behind the growing interest around the performance-based diagnostics and behind the data analytics in general.

The first reason to be mentioned is economical. As exposed in section 1.2 in-fact, the growing cost of the fuel, together with the cost competitiveness, the engines are expected to run longer, to reduce the costs of exercise and to reduce the maintenance costs. To accomplish to this requirement the diagnostics moved throughout the years from preventive diagnostics techniques, where the maintenance is scheduled based on a fixed number of operating hours, to the predictive diagnostics, where the status of the engine can be predicted based on the data recorded. This transition is justified by the trade-off between cost and number of failures. In fact, preventive maintenance can anticipate many failures but is causing additional maintenance costs. On the other hand, reactive maintenance has high costs due to the increasing amount of failures. Therefore, the optimum falls into the predictive maintenance technique (Figure 7).

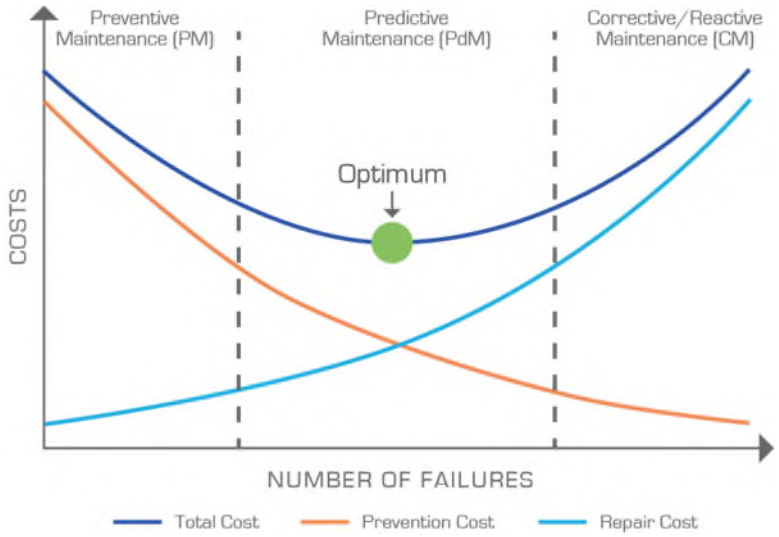

Figure 7 Cost vs failure trade-off generic graph – Courtesy of CBM services, available at https://ivctechnologies.com/2017/08/29/reactive-preventive-predictive-maintenance/reactive-preventative-predictive-maintenance/

The same conclusion can be drawn by looking at the cost of repair vs time, where knowing in advance the failure decreases the final price (Figure 8).

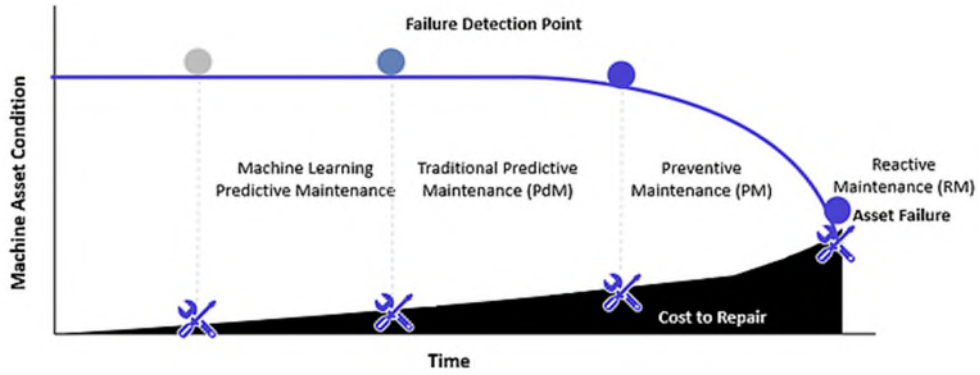

Figure 8 Cost vs failure trade-off generic graph – Courtesy of Presenso, available at https://www.presenso.com/single-post/2017/05/24/the-economics-of-the-smart-factory-how-does-machine-learning-lower-the-cost-of-asset-maintenance-part-1/

The key facilitators who allowed this technology to gain interest are the computational capability, that has been exponentially growing over the last few decades and the increasing amount of data from the engine acquired and recorded by the monitoring systems. With the available techniques exposed in section 2.1 it is then possible to conceive a robust combination, suitable for:

- online monitoring and diagnostics, for predictive maintenance of components and sensors failures.

The availability of new data and the availability of additional computation power increased the potential of the gas turbine diagnostics. If the target from the initial stages was to have an early warning of the status of the machine, in the upcoming years the target has been the diagnostics of multiple failures. Moreover, the target has grown also in terms of accuracy meaning the correct the detection of the type of failure and the correct quantification of it. The last few years have seen also a growing interest in the investigation of the gas turbine failures together with the measurements failures. These topics have been often studied separately, also for the specificity of the arguments, but few papers started to propose solutions for that. Starting from the existing background, a new methodology should then target at least:

a. Diagnostics of single and multiple failures of gas turbine components;
b. Diagnostics of the gas turbine also with the presence of measurement issues;
c. Speed to be suitable for online monitoring.

3.2 The way to gas path analysis

The performance-based gas path analysis is a topic that has been studied in the last 40 years since Urban [11] defined the possibility of making diagnostics on the gas turbines components, based on the performance parameters. In his concept, the health status of a gas turbine could be reflected by its key performance parameters like for example the efficiency of the compressor or of the turbine. By this means, he was trying to anticipate the detection of a certain event, before any catastrophic failure could cause a forced shut down of the machine. Before Urban and its newly introduced gas path analysis the other consolidated techniques that had taken place were:

a. Visual inspection;
b. Radiography;

c. Thermography analysis;
d. Vibration analysis;
e. Lube oil analysis.

The visual inspection, the radiography and the thermography are certainly not suitable for online monitoring and they might not be so effective for the predictive analysis of gas turbine components. Due to their nature, they would be more suitable for preventive maintenance or for analysis after failure events.

The vibration analysis can be used for the predictive maintenance, but as described in section 2.1.4 the indication of malfunction may come once the component is already deteriorated. Moreover, by keeping the information from the bearings, it might not be possible to isolate the status of a single component.

A similar conclusion can be drawn by the lube oil analysis who is able to provide indications once the engine is starting to experience wear and collect debris in the lube oil circuit and might also not be able to isolate the deterioration of the components (Figure 9).

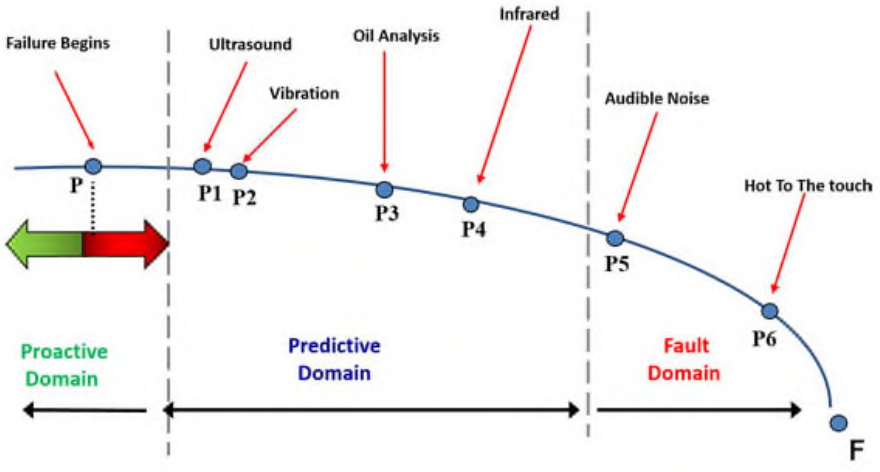

Figure 9 P-F (potential failure, functional failure) diagram for diagnostics techniques – Courtesy of bearing news available at https://www.bearing-news.com/mechanical-remote-monitoring-with-ultrasound

The performance-based diagnostics proposes to detect loss of performances and malfunctions from the early stages of operation with several clear advantages:

a. identify which part of the engine is subjected to shortages and plan the outage accordingly;
b. identify the delta from the baseline and evaluate the magnitude of the component shortage;
c. prevent unplanned outages;
d. extend the lifetime of certain components.

The advantages listed might not be obvious from outside, but they are from the user perspective. For instance, once an engine starts to show vibrations toward the limits, or its lube oil is showing a growing amount of debris, it would be an indication that some additional inspections and or replacement are needed. In such a case, knowing which part of the engine has failed, can be crucial information to get the right parts and to save unnecessary transportation of other parts.

Looking at the aero engines, it is a common thing that the deterioration of the engines is different among the planes. It is then important, in order to know how to manage the fleet schedule, which engine is performing better and what is the cause of the shortage. If this is known the user or the contractor, have the possibility to optimize the flight schedule.

Even if the unplanned shut down are now reduced compared to a decade ago, it is always an undesired and expensive event. Therefore, every measure might result as convenient to be used to prevent it as much as possible.

Talking about the engine operation it can be that a certain engine utilize can affect more one or another component. With the vibration diagnostics and with the lube oil analysis it is difficult to make a connection between the component shortage and the engine operation. However, this is possible with the performance-based diagnostics.

3.3 Open points for single techniques

Due to the lack of computational resources and to the novelty of the topic, the gas path analysis started investigating the single techniques. The first one has been the linear gas path analysis, considered by Urban [11] in one of his early publications and all the others followed. Applying a single technique offers the advantage of keeping the methodology lean and simple in terms of performance/execution time and in terms of set up. However, each methodology has limitations, that causes restrictions in the gas path analysis capabilities. Bechini [1] provided a comprehensive rank of the potential of each methodology, while challenged with the detection of single failure and multiple failures, that has been updated considering the last years of research on this topic Table 1. The capabilities evaluated during the assessment are the accuracy, the speed, the reliability with noise or with few measurements available, the possibility of data fusion and the flexibility. The rank is based on an overall research overview and can be taken as a high-level indication for pros and cons. A fundamental element is the speed that is a key enabler for the online monitoring system. The accuracy is also crucial as the analysis must show alerts if a component is experiencing a malfunction, and they must avoid false alarms that could trigger unnecessary actions. The reliability has been evaluated against noise in the measurements and with the presence of few measurements, meaning a few information about the engine status. The characteristic of a technique, instead of another, should be able to overcome the lack of missing data or the decreased quality of it. This is a clear advantage in a real situation where noise and measurement errors can be common. Looking at the growth of data availability, fusing data can represent a strong feature to make the information more robust. Finally, the flexibility is always a desired characteristic especially thinking about the possible expansion of the methodology capabilities – e.g. diagnostics of a new type of failures – or to the combination with other methods.

	Accuracy	Speed	Reliability		Data Fusion	Flexibility
			Noise	Few Meas		
LGPA	1.5	5	1.5	1.5	1.5	5
NLGPA	2	5	1.5	1.5	1.5	5
KF	3	5	2.5	3	4	3
EKF	3.5	4	3.5	3	4	2
ANN	2	4	3.5	4	3.5	1.5
BB	1.5	2.5	4	4	4.5	2.5
GA	3	1.5	4	4	3.5	3
FL	3	3	3.5	4	4	5

Table 1 Rank of suitable methodologies for gas path analysis for accuracy, speed, reliability, data fusion, and flexibility. 0 minimum or non-performing, 5 maximum or perfectly performing – Main sources from [1]

Combining the information from literature, it is possible to draw the advantages and disadvantages of each methodology in order to get a clear picture of what is the potential and what are the limitations of every single method.

3.3.1 Linear gas path analysis and non-linear gas path analysis

The linear gas path analysis consists of a linear relationship between the gas turbine parameters and the gas turbine health parameters. The governing equation (2) of this relation is:

$$\Delta z = H \cdot \Delta x \tag{2}$$

Where x is the matrix reflecting the physical changes, z is the key diagnostics indicator and H is the exchange rate. Being the gas turbine behaviour mostly nonlinear, the linear gas path analysis is an approximation that can result in a loss of information or to non-accurate results.

The linear gas path analysis can be applied for the detection of single and multiple failures and can still be used in combination with other techniques to identify a part of a more complex problem.

The linear gas path analysis is the technique used by Urban [11] while proposing this methodology for predictive maintenance. In that frame, he has

shown the potential of the methodology with the tools of that time. By now, since more powerful techniques are available, and the computational power is increased, the linear gas path analysis is not precise enough on the approximation of nonlinear problems.

The non-linear approach instead reduces the residual error by introducing non-linear equations to simulate the cause-effect relationships among the measurements and the gas turbine health parameters. It has been presented by Escher [12] in the frame of his PhD thesis at Cranfield University and described by Ogaji [13]. The graphical illustration is reported in Figure 10 and shows a higher accuracy if compared to the linear approach.

Figure 10 Simplified comparison of linear and non-linear approaches – Courtesy of [13]

However, as reported by Escher [12] the deviations cannot be distinguished from a possible measurement discrepancy. Within measurement discrepancy, it can be included the noise, especially if it has value comparable to the modelling error, and the bias that would cause false conclusions.

Strengths

- Easy to set up;
- Suitable for isolation of single and multiple failures;
- Can quantify the performance loss of each component.

Limitations

- Not accurate as it is approximating the problem to a linear one;
- The low accuracy can prevent the detection of certain types of failures;
- Not capable to deal with measurements noise and bias.

3.3.2 Kalman filter

As well described in the book “Kalman Filter, theory and practice using Matlab” [14] the KF has been called the linear least mean squares estimator because it minimizes the mean-squared estimation error for a linear stochastic system using noisy linear sensors. It has also been called the linear quadratic estimator because it minimizes a quadratic function of estimation error for a linear dynamic system with white measurement and disturbance noise. It can be used to better estimate a dynamic system, therefore, is suitable to many physical applications. The governing equations of the KF are:

$$\dot{\hat{x}_k^-} = f(\hat{x}_{k-1}, u_{k-1}) \tag{3}$$

Where u is the control vector and x_{k-} is the a-priori state estimation.

$$\dot{\hat{x}_k^-} = A\hat{x}_{k-1} + \omega_{k-1} \tag{4}$$

Where A is the state Jacobian gain, ω is a white noise process covariance and Q being the model noise covariance.

$$\dot{P_k^-} = AP_{k-1}A^T + Q \tag{5}$$

$$\dot{K_k} = P_k^- H^T / (HP_k^- H^T + R) \tag{6}$$

$$\dot{\hat{x}_k} = \hat{x}_k^- + K_k \left(z_k - \hat{x}_k^-\right) \tag{7}$$

$$P_k^{\cdot} = (1 - K_k\ H)P_k^{-} \quad (8)$$

Where P^- is the predicted error covariance, K is the Kalman gain, x_k is the updated state estimation calculated using the Kalman gain multiplied by the difference between the a-priori state estimation x_k^- and the measurement z_k. H is the observation model and R the covariance of the observation noise.

Among the most common applications, one is the signal estimation with Gaussian distributed noise around the signal. Applied in the field of GPA, the KF is used to estimate the difference between the predicted value and the measured one or to estimate any measurement error.

Kerr et al. [15] used the KF for online estimation of deterioration techniques on the gas turbine components. The authors built a methodology able to reproduce the gas turbine key parameters. These parameters are then compared with the baseline for deterioration comparison. The model can predict the parameters with a deviation between 2% and 15% recorded during a rapid deviation. The methodology is also able to deal with measurement errors with 2% order of magnitude but is not able to handle the higher error and consistent bias that the authors say are not likely to be encountered due to the redundant measurements. Compared to the non-linear model, this paper is an enhancement in terms of accuracy but is not able to provide accurate diagnostics on the components failure, the failure is quantified in terms of delta, but the data have to be post processed by a user, the success rate is not proved and the methodology is not able to deal with bias in the measurements.

Kobayashi et al. [16] proposed a bank of KF to detect and correct sensor and actuator measurements errors. The scheme is named bank because one KF is dedicated to each individual measurement. Applying this methodology, the authors are able to correct bias in the measurements and detect the effect of foreign object damage. The detection of the foreign object damage with measurement bias result being correct as of the simulation after the filter

collapse on the unbiased line. This result indirectly shows the potential of the KF for the measurement noise and bias estimation and correction. The filter in-fact is fast and reliable and is suitable for online applications. However, the authors report, the smearing effect has been encountered during the simulation as far as measurement errors have been transmitted also to unbiased measurements.

Looking at the KF applied to data estimation and filtering Simon [17] compared different types of approaches for data filtering applied to the aircraft engine health estimation. The author processed three types of KF: LKF, UKF and EKF and compared the estimation accuracy while measurement noise and bias were also present. The results showed that the health estimation, performed after with a dedicated tool, consistently improved with the application of the KF and in particular with the EKF. This type of filter end also being the best compromise in terms of computational effort vs performance. The limitation outlined by the author is the incorrect estimation of the parameters with large biases.

Still, on the measurement noise, Lu et al. [18] studied the information fusion of different Kalman Filters in order to increase the robustness of the diagnostics in case of measurements failure. The authors state that the KF fusion is capable of isolating the measurement noise and bias. The best technique the authors found is the UKF with information fusion that is outperforming the EKF and the LKF. The authors report that the time of processing is in the order of seconds which makes the technique suitable for online diagnostics. This time remains at low values also thanks to the fusing technique implemented, preferred to the slower bank of KF. This paper sets the possibility and the convenience of making use of different measurements for the isolation of the measurements noise and bias. On the other hand, it remarks again the limit of a single technique, even in this detailed architecture, to fulfil all the objectives pretended from a performance-based diagnostics technique. In fact, the multiple failure components diagnostics is not investigated, and the quantification is yet not feasible.

An information fusion performed with the KF has been studied by Sun et al. [19] who presented a methodology to combine the information of multiple sensors and reduce the level of disturbance related to each probe.

Strengths

- Can deal with measurement noise;
- Its speed is acceptable for online monitoring;
- Prior knowledge can be included to improve the power of the recursiveness.

Limitations

- Requires a fine-tuning and prior information to do that properly;
- The smearing effect of the error over other components;
- The solution is accurate only if the system of equations is correctly set up.

3.3.3 Artificial neural network

The artificial neural network principle is inspired by brain connections in the human head. The human brain, in fact, works with a large number of connected elements called neurons, passing information to each other. The communication within the neuron works with four sub-elements: the dendrites, the axon, the cell body and the synapse (Figure 11). Their arrangement within the neuron and the strength of the synapse determines the response within the brain. Some of the connections are established at born age and are part of the basic functions. Instead many of other connections are modified with time within the learning process with the experience each individual face in life. Looking at the human brain, it is noticeable that the learning process is much faster in the earlier stages of life, possibly because all the synapses are very flexible and there is a large margin of additional education. In the later stages, instead, the learning process is slower as many connections are established. In this case, the additional learning requires a bigger effort as it must add new information on strong established connections.

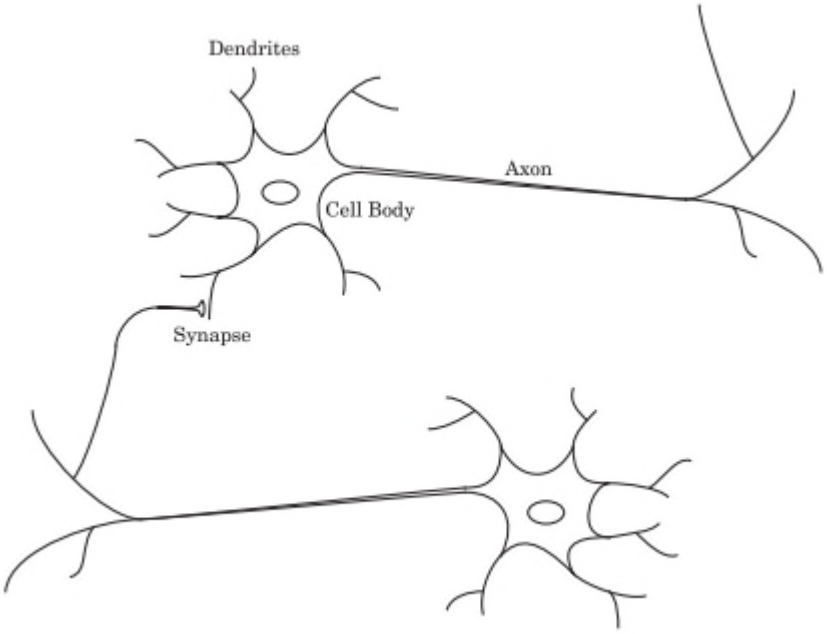

Figure 11 Drawing of human neuron – Courtesy of [20]

As mentioned at the beginning of the section, the aim of the artificial neural network is not exactly replicate the human brain behaviour but is to take inspiration out of its mechanism. In particular, what is meant to be used is its efficiency on operating parallel neurons at the same time.

Based on this premise, the mathematical equation base of the artificial neural network is set as:

$$a = f(w^T p - b) \tag{9}$$

Considering a scalar input p, the neuron output can be calculated by tuning the weight w and the bias b. The weight can be associated with the strength of the synapse, the cell body is reflected by the sum of the transfer function and the result n can be related to the signal on the axon. The transfer function can be linear or not linear according to the problem the artificial neural network is exposed to. An artificial neural network built in this way is called “single input”. This type of structure does not make use of the parallel problem solving, that has been described as a clear advantage of the human brain. To perform a parallel calculation the weight is converted to a matrix W as:

$$n = \sum_{j=1}^{R} w_{1j} \cdot p_j - b = W^T P - b \quad (10)$$

$W^T P$ is the result of the sum of the weights times the input scalar. This type of structure is a multi-input neuron.

The artificial neural network cannot only expand vertically, meaning in parallel but also horizontally, meaning in series. In this case, the $W^T P$ + p is reproduced in series according to:

$$n = W_2^T P(W_1^T P - b_1) + b_2 \quad (11)$$

Once supplied with data the artificial neural network can be trained. In this phase, the weights and the bias are tuned to match the results provided. The minimization technique can also vary in order to optimize the convergence and the match. Once the artificial neural network is trained it is able to provide the response based on the scalar input p.

The artificial neural network is a very practical technique used to make estimation instead of a physical model. In the frame of gas path analysis is often used to predict the calculated GT parameters like the efficiency of the components. An example of usage for GT parameters prediction for gas path analysis comes from Joly et al. [21]. Their final scope was to evaluate the effectiveness of the ANN on the prediction of the efficiency and mass flow capacity for single and double failure. Their conclusion was that the ANN is capable of correctly predicting the single failure but is subjected to higher deviations while two components are failing. Additionally, the authors mention that a higher number of components failure would further increase the uncertainty. There are several open points or opportunities left on this paper. One already mentioned is the detection of multiple failures with more than two components. The engine health estimation is another open point. In fact, if the efficiency decay can be predicted, the status of the component is not defined. Moreover, there's no mention of the measurement issues, that could be part of the engine normal run.

In a more recent paper, Asgari et al. [22] compared various ANN variants in order to establish the best fit for a one spool engine. The ANN prediction was satisfactory on several variants confirming the robustness of the technique.

Vatani et al. [23] proposed the ANN for the prognostic degradation trend of non-linear dynamics of a single spool aircraft engine. This study investigated the possibility of estimating the lifetime of the engine, provided the actual status and the historical trend. Despite the limitation to the one spool engine, this paper confirms the role of the ANN in predictive maintenance.

Strengths

- It is able to build a model starting from data only;
- The speed is suitable for online monitoring in the prediction mode;
- It is suitable for problems with measurement noise included.

Limitations

- It requires data for the training and time to analyse all of them;
- The model might not be reliable if the prediction is done outside the range of training;
- If the prediction is unexpected the reason can be looked on the result only as the model could be hardly investigated.

3.3.4 Bayesian belief network

The Bayesian belief network is based on probabilistic knowledge for a failure or an event to come. It bases its prediction on previous knowledge from where it is able to create a network. This knowledge is stored into probability tables built during the set-up phase.

The nodes reflect the variables invoked by the problem. On one side there are the causes of an event and on the other side the effect or evidence of this cause. Cause and effect are connected by the arcs that represent the casual

relationship. Supposed the cause is a measurement change in the GT, and the effect is the deterioration of one or more components, a change in one measurement can be reflected in one or more GT components (Figure 12). The peculiarity of the BBN is the capability of evaluating the probability of measurable changes to be caused by a certain type of event.

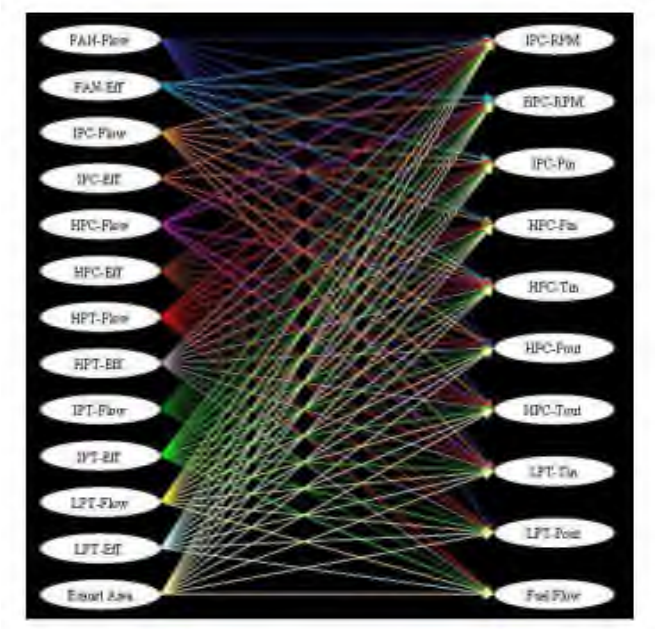

Figure 12 Typical BBN layout for gas path analysis – courtesy of [24]

The evaluation of the probability is based on the Bayes theorem for conditional probability:

$$p\left(\frac{x}{z}\right) = \frac{p(x)p(\frac{z}{x})}{p(z)} \qquad (12)$$

Where x represents the independent variable and z the dependent vector.

The applications of the BBN are also several due to the capability of relating event changes with possible causes, situation ideal to detect GT failures based on measurements variations.

Looking at a single methodology with Bayesian belief network one interesting work from Pu et al. [25] was reported. In this paper, a single spool industrial gas turbine was investigated for single failure together with

measurement bias. The challenge that the authors overcame is the diagnostics of failures together with measurement issues, however, the paper did not go beyond the single failure due to the possible interaction between measurement failure and the effect of the deterioration on the measurements. However, the authors could reach a success rate of 97.1% for the specified case.

In a similar way, Kestner et al. [26] established one way, based on Bayesian belief network helped by a thermodynamic reference model, to detect the deterioration of the gas turbine components – the focus is on the compressor – together with a measurement bias. The methodology is able to detect both the deterioration and the measurement bias while offline. The open points identified by the authors are the increase of the accuracy also for other types of failures and to set up a base to further testing the methodology out of real cases. Other points, independent from the paper as such but related to the methodology are the possibility of detection of this type of failures for online points and the detection of multiple failures.

Strengths

- Engine model hardware changes can be easily introduced;
- It can deal with multiple failures;
- It accepts different types of data – discrete, continuous, qualitative.

Limitations

- Cannot deal with sensor bias or there's no evidence yet;
- Long set up required;
- It requires high computational capabilities to deal with multiple failures and might result in inappropriate for online monitoring.

3.3.5 Genetic algorithms

The genetic algorithm is one of the most common techniques used in computational science since decades to find a solution within complex

problems. The process of the GA mimics the species evolution process selecting the most “fit” one at each step. With the first step, a sufficiently large population is created. The selection is random and covers a certain range. The solutions are evaluated and ranked according to their fitness and filtered for the next step. A mutation is also considered to allow other solutions to raise and to avoid local minima. The governing objective function can be reported as:

$$J(x) = \sum_{i=1}^{m} \frac{[z_j - h_j(x)]^2}{(z_{obj}\sigma_j)^2} \quad (13)$$

Where z is the signals or measurements, h is the corresponding simulated value, z_{obj} is the reference value and σ is the standard deviation.

The GA is an excellent technique for probabilistic estimation and is suitable for a wide variety of problems. However, its main drawback is that it requires a long computational process. Therefore, it is not suitable for online data processing rather for offline data processing. The most recent scientific publications confirm this conclusion while reporting analysis of gas turbine deterioration offline (post processing of data). It is the case of the paper proposed by Mo et al. [27] whose focus is the correct estimation of the maintenance time based on performance indexes. To do so, the authors evaluated a GA combined with a SA and compared the results with the condition-based maintenance and the proactive maintenance schedule. The evaluation has been done for the whole machine but considers the contribution of the component deterioration and the measurement deterioration.

Another relevant paper about GA has been presented by Qingcai et al. [28] who analysed the effect of deterioration on a three-shaft industrial gas turbine at full and part load. The paper shows that the GA was able to correctly capture the loss of efficiency but also the variation of the physical parameters. The information provided by the GA constitute a base for the component health estimation, however, they also tell that the GA cannot do it on his own

since the deviations should be further worked out to provide an indication of the status of each component.

Strengths

- It uses probabilistic rules rather than deterministic allowing running on almost all the problems;
- The method uses a global search procedure, therefore, does not stop on local minima;
- It can deal with single and multiple failures with a limited number of information.

Limitations

- The estimation time makes it unsuitable for online diagnostics purposes;
- The set-up is demanding and requires an experienced person;
- The speed can limit the capability of fusing big amount of data.

3.3.6 Fuzzy logic

The fuzzy logic is a rule-based approach, founded on the formulation of novel algebra, enable decision making processes to be performed. The fuzzy logic relies on the formulation of the fuzzy algebra that is based on a new definition of fuzzy sets and logical operators. To define all the rules that will compose the fuzzy logic elements must be defined:

- input and output fuzzy sets, that defines the sets of rules that will be defined;
- fuzzy rules, that defines the relation between the input and the output;
- membership function, that defines how the value is mapped in the definition space.

At this step for a certain set of inputs, rules are defined and it has to be defined how the output should be processed (Figure 13). The steps making it are:

- fuzzification, that process the inputs in the fuzzies;
- application, that apply the fuzzy operators;
- implication, that decides how the output from the application are merged;
- aggregation, that decides how all the rules are merged together
- defuzzification, that provides that final output of the entire set of rules.

Due to its possibility on combining rules together, but also on the flexibility of allowing complex structures, it is used in the gas path analysis to isolate the types of failure in the frame of multiple failure detection.

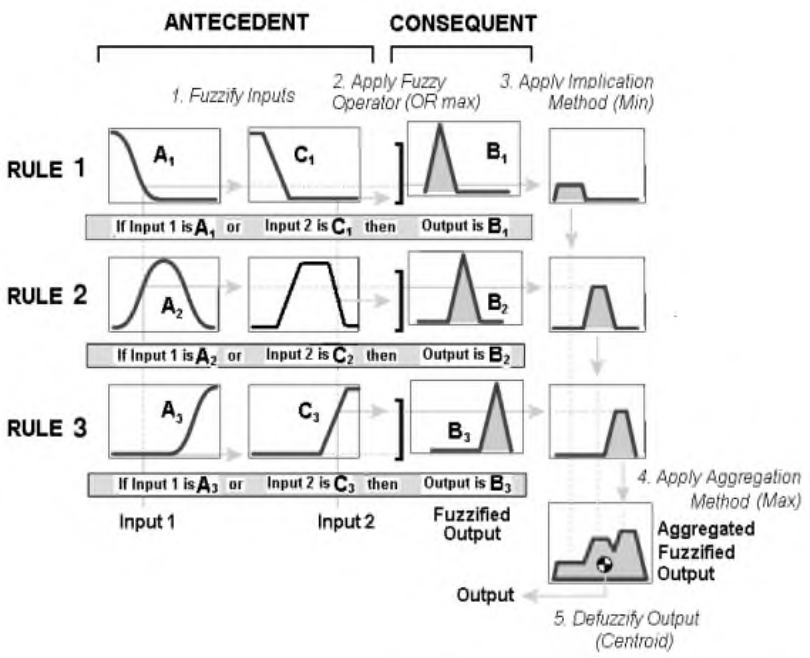

Figure 13 Fuzzy logic scheme – Courtesy of [1]

The fuzzy logic can be used in the performance-based diagnostics for the fault isolation and for the components health estimation. Ogaji et al. [29] used the fuzzy logic on a three spool aero gas turbine to isolate the components failure while a non-better specified measurement noise was included. Their work was able to show the feasibility of the detection of the efficiency deviation and the flow capacity deviation. However, the authors did not make use of the

FL for the component health estimation and focused their attention on single component failure.

In a more recent paper, Barbosa et al. [30] set up a methodology based on FL for the failure isolation. The engine evaluated by the authors is a one spool industrial gas turbine. The authors focused on two types of failure that are the compressor surge and the clogged filter. By looking at the deviations of the internal parameters the authors have been able to provide a severity reflecting the status of the type of failure and the deviation from the physical parameter. This paper also outlined that there's the need for a dedicated FL for each type of failure which means that additional types of failure can be included by increasing and incorporating other FL.

In another publication, Eustace [31] used the FL to isolate the faults of an aero engine. The types of fault considered in the paper are 17 and they are all related to the hardware, excluding any measurement implication. The author assigns a probability of a specific failure to happen and sort them by magnitude. The first failure is the most likely to happen, while the following are less likely. The failures are categorized from identified to eliminated, passing through other steps like possible and unlikely. Despite there's a severity reported, the authors do not explicitly assign a health status to the component or to the specific failure. On the overall, this paper shows the flexibility and the capability of the FL of detecting and isolating several types of failure.

Strengths

- It is suitable when there's no precise model describing the system;
- It is very flexible, and they can be used for approximation of different types of models;
- The setup and prediction phase is fast and suitable for online monitoring.

Limitations

- The fact that can be model-free can affect the robustness of the method;

- The set up is demanding and requires an experienced person;
- The prediction lack in precision outside the range of training.

3.4 Open points on combined techniques

As reported in section 3.3, the use of a single methodology is not enough to fulfil the detection of multiple failures, together with measurements failure, with sufficient speed for an online diagnostic and with the capability of correctly interpret the difference against the baseline. The major difficulty is to have a single methodology flexible enough to isolate the measurement failures and gas turbine failures. Another type of limitation is the computational time, that cannot increase at the point of compromising the online detection capability. In addition to that, several authors comply that the detection of a deviation parameter is not a univocal sign of failure, but it has to be addressed and quantified to reflect the type and the magnitude of the failure detected. Since for each methodology, open points remain, the research moved to the combination of techniques, with the objective of merging the strength of each contributor and expand their capabilities.

The paper published by Sampath et al. [32] described how to detect multiple failures including measurement noise and bias. The methodology was built with an auto-associative neural network used to isolate the bias and a combined genetic algorithm, artificial neural network, employed to detect multiple failures. The authors reported that GA is the most accurate for the detection of the failure but is requiring hours to reach the conclusions. In the attempt of reducing the working time, a pre-layer of auto-associative neural network has been established. The layer has been able to isolate gas turbine failures (single and multiple) and measurements issues. The GA has been used in sequence to refine the detection. The authors support that the result is efficient in terms of detection accuracy, that is above 90%, but the amount of time required is still not suitable for online applications. In addition to that, the quantification does not provide a clear interpretation of the failure but is limited to the detection of the efficiency delta with respect to the base case.

Viharos et al. [33] proposed a comparison of different neuro-fuzzy solutions and their applications. This paper, not related to the GT diagnostics, is cited here because it shows how a discrepancy from a baseline, can be isolated and quantified to reflect a failure magnitude. For instance, the neural network is used to establish the transfer function between the inputs, that can be the deviation from the baseline and the output that can be the quantification of the failure or its category. The authors found that the combination of the neural network learning, together with the fuzzy logic, reduces the setup time and improves the quality of the detection. In addition to that, fuzzy logic can be manually implemented to include some user-based rules.

In another paper, Wang et al. [34] introduced a series of fuzzy logic, which is coupled with the TOPSIS methodology. The fuzzy logic is used in two layers to create the decision-making rules and to set and calculate the weights associated to each contributor. The TOPSIS is used to make the overall performance assessment and take the final decision. One of the peculiar things about the methodology is that is making use not only of the performance evaluation but is also considering other information like lube oil analysis, vibration analysis and boroscope inspection. It is obvious that making use of information that can provide information later in time Figure 9, delay also the potential of the performance-based diagnostics. On the other hand, it might be an added value for certain types of failure. The result achieved by the authors is a quantification of the engine health in a relative scale 0 to 1 and an indication of the actions to take Figure 14.

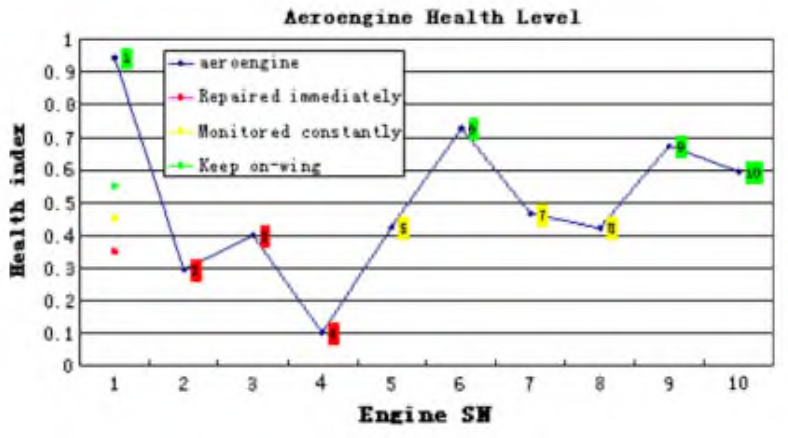

Figure 14 Health index assessment – Courtesy of [34]

Despite being relevant, the work does not consider the health status of single components of the gas turbine and omit that analysis of the gas turbine in the presence of measurement issues or, in the specific case, wrong data from boroscope, vibration analysis or lube oil analysis.

The paper published by Dewallef et al. [35] proposes a combination of Bayesian Belief with KF in order to benefit from their mutual advantages. The Bayesian Belief is used to improve the information the KF is using for the prediction of the failure in the presence of measurement uncertainties within 3σ. The results are proposed for the single methodologies separated and for the combined methodology and they show that the combination is reducing false alarms and improving the health parameter prediction. The detection focuses on the single component failure and on the health parameters value. Moreover, the open point highlighted by the authors is the missing information fusion that could further improve the prediction. It must also be highlighted, that some deviations from the nominal value are still present, even with the combination applied.

Within the paper [36] the fuzzy logic has been coupled with the backward elimination not only for the diagnostics but also for the remaining lifetime estimation. This paper does not investigate in detail the multiple failure scenarios, but it provides a cue for the need of exact modelling for the lifetime

estimation. On top of that, it offers a proposal on how to quantify the convenience of maintenance intervention.

A recent paper, taking advantage of the increasing knowledge and power of the neural network Yang et al. [37] applied it to the diagnostics of a gas turbine. In particular the Quantum-Behaved Particle Swarm Optimization – QPSO – is used to optimize the weights and the bias of the neural network. The result is then used for the failure classification. The authors compared different solutions and found that the combination of QPSO and the extreme learning machine – ELM – is outperforming the others. This confirms the great potential of ELM and overcomes, the authors add, the ill condition problems because of wrong selection of weights and bias. The performance reached on the classification is 92% on the single failure and below 88% on the multiple failures (Figure 15). This result sets a target for the multiple failure classification and the potential of the ELM. It has to be remarked that measurements errors are not included in the study and the quantification of the failure is not explicitly mentioned.

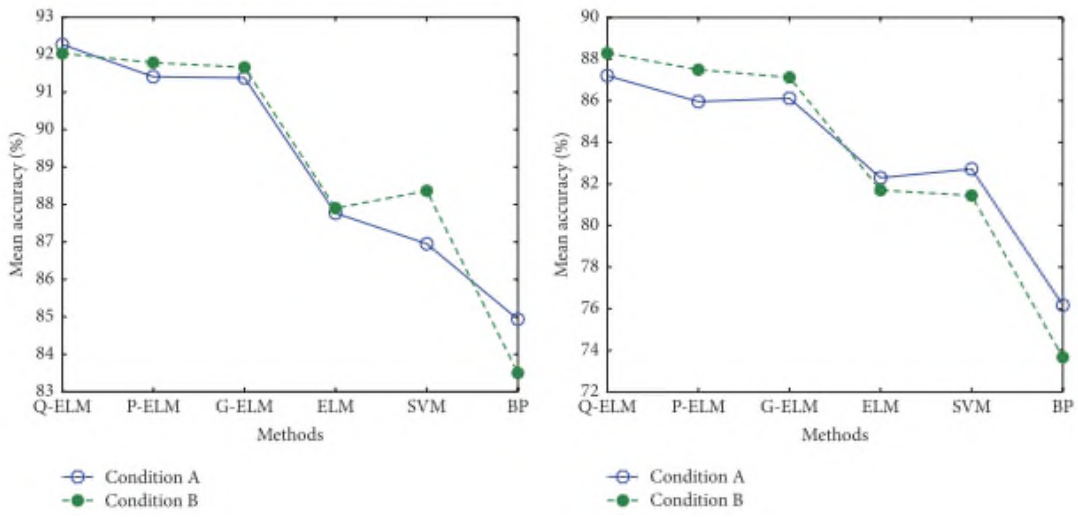

Figure 15 Mean classification accuracy for SF (left) and MF (right) – Extract courtesy of [37]

Finally, Li et al. [38] coupled neural network with support vector machine for the quantification and classification of the gas turbine failures with standard

white noise. The authors set a back-propagation artificial neural network for the performance deviation from the baseline estimation. This information has been used by the authors as a base for the quantification of the component's failure. In sequence to that, the authors introduced an SVM, a variant of machine learning for the classification of the gas turbine failure types. The SVM has the peculiarity of being characterized by a multi-kernel that the authors found to be more robust. To test this methodology, the authors collected the EGT, the FF and N2 speed available data from the CFM56 engine. Based on that they have been able to classify multiple failures. The success rate ranges from 57% to 90% depending on the type of failure. The authors claim that a higher number of signals considered could improve the detection rate. On top of that, it has to be added that the research does not consider any measurement failure in addition to the normal noise that the source data can have. Additionally, despite the failure is classified, there is no health status set up or detected.

A key topic to consider in the review of the scientific publications is the methodology proof against real data and/or simulated data. As reported by the NASA [39] the gas turbine engines deteriorate while being in service. This deterioration varies from engine to engine and in a few cases, it might experience a failure. The type of failure can also vary from engine to engine according to many factors – manufacturing, ambient conditions, how the engine has been operated. Provided this environment, it appears unfeasible to validate an EHM (engine health monitoring) methodology only based on real engine data. For this reason, Simon et al. [39] created a platform, called C-MAPSS simulating an aero gas turbine engine, offered to benchmark the EHM technique and to challenge the diagnostics methodology to correctly detect the available failure. The types of failure that have been included are single component failures and measurements issues. This platform has been used by Simon et al. [40] to test four different methodologies provided by a pool of research institutes and universities working on the EHM. The methodologies have been tested against the data provided for single failure and measurement issues and reached a maximum correct classification rate

of 46.7%. The user feedback is written in the paper conclusion remark that the data should include more realistic measurement noise covariance values and different outlier. On top of that, and based on the other literature review, the missing point outlined in the paper is the multiple failure simulation.

3.5 Summary of research gaps from the literature

The literature review detailed in section 3.3 and section 3.4 described what has been done in literature from the early stages of the research in the gas path analysis. The objectives explored in the literature that coincide with the potential of the gas path analysis are:

- Detection of single failure: detect the failure of a single gas turbine component and do not raise false alarms on the others;
- Detection single and multiple failures: detect the failure of single component and of multiple components failing at the same time;
- Classification of different types of failure: distinguish between different failures type ranging from fouling, erosion, corrosion, tip clearance, foreign object damage, cooling leakage, burner failures;
- Quantification of the failures and remaining lifetime estimation: associate the failure magnitude to remaining lifetime for maintenance schedule (health estimation);
- Isolation of measurements failure: isolate the noise and the failure of one or more signals that could lead to false failure detection;
- Online diagnostics: each sample processing time should be adequate for online diagnostics;
- Methodology proving against real data: test the methodology against real data or against proved and realistic simulated data;
- Data fusion: fuse information from multiple measurements or from multiple sources to reinforce the overall detection.

The achievements reached in literature are quite exhaustive, however, as outlined in section 3.4 there are still few open points in each paper giving

opportunities for additional research. The main missing points from the literature are:

- Detection of multiple failures together with measurements issues: the methodologies are focusing either on the detection of measurements failure or in the detection of components failure. The combination of objectives has been found only in one paper [32] which was not suitable for online diagnostics;
- Combine the detection of multiple failures with the health state estimation: the health state estimation has been investigated by Wang et al. [34] but focusing only on overall failure and omitting the presence of measurement noise and bias.

The study of the scientific papers allowed to draw up a list of most suitable techniques for each module of the methodology. The conclusions of that are:

- Kalman filter: several works remarked the power of the KF for data analytics including data fusion. The results are accurate, and the execution time is suitable for online applications;
- Artificial neural network: it is very accurate on learning the gas turbine performance behaviour and on predicting it based on the inputs;
- Fuzzy logic: it has been used to determine the health status of the gas turbine components or of the overall gas turbine. Its set up can be coupled with the artificial neural network leading to a neuro-fuzzy system. On the other side, its flexibility can be used to add rules for additional failures.

It is implicit that a technique should couple the strength of multiple methodologies. Therefore three candidates are selected here for the execution of three different scopes. While specifying that, another missing point from literature can be described:

- The combination of KF, artificial neural network and fuzzy logic has not been investigated yet in literature in the gas path analysis environment.

Its interest, as said, resides in the combination of multiple methodologies to mutually compensate the failing of each contributor and expand the capabilities of the methodology as a whole.

3.6 Objectives

A methodology is intended to offer an easy way to detect the components deterioration or failure [3]. The final set up for health monitoring should include the features remarked by the experts in this field [84]:

- interface with the increasing amount of data available from the engines; integrate new sensor suites and capabilities;
- precise modelling of the baseload and part load conditions; leverage all available information including user-specific inputs;
- have a practical design. What is meant to be detected is the deterioration of every single component – single failure – and the combination of components deterioration – multiple failures;
- measurements failure shall also be part of the simulation as they are relevant in any working engine.

Combining the intent of the methodology with the open points not investigated yet from literature, the established objectives are:

a. Detection of single and multiple failures also in the presence of measurement failures. The measurement issue established is the noise that is one of the most common failure encountered in the GT measurement package and generally in the probes. To declare the objective fully reached, a minimum success rate of 90% has been established. This rate is in line with the levels proposed by the literature review in particular with the papers dealing with multiple failures scenario [37];
b. Combine the component failure isolation – single and multiple failures – with the component health estimation. The methodology

should not only isolate the deterioration of the single component but account for the mutual impact of one component on another and establish the health status of every single component;

c. Establish a methodology able to deal with multiple failures and measurement issues while working online. The target to reach in terms of time is in the order of seconds per sample processed;
d. Fuse the multiple measurements available for each location. This opportunity is most suitable for industrial gas turbine, where redundant measurements are available;
e. Test the methodology against different situations to prove its robustness. The testbed should include different types of failure, different levels of deterioration, single multiple and no failures and measurements errors.

4.0 Methodology

As defined in section 3.6 the technique established is focused on enhancing the scientific understanding of multiple components failure of gas turbines together with the measurements failure. While doing that it has to treat the incoming data for noise, predict the efficiency and the main gas turbine parameters and offer a health status of each component of the engine.

The steps used to determine the selected technique are divided into three phases:

a. Selection: the study of the methodologies and the selection of the contributors;
b. Implementation: the process and methodology constructions within a coding environment;
c. Testing: the creation of a testing procedure for the established methodology.

Selection

- Define the most promising techniques to be used among all the possibilities. This step is obtained analysing the potential of each technique based on the dedicated literature review and if necessary, testing it. This step has to consider also the fact the methodology should be capable to operate online;
- Select the combinations able to compensate for all the weak points. The selection is done analysing the strength and weaknesses of each technique and merging them while compensating the weaknesses.

Implementation

- Implement a process flow able to achieve the expected results while combining techniques. For example, two techniques could be made acting in series to process the data one after the other or in parallel to process the data at the same time.

Testing

- Define the inputs data able to simulate the real engine failures (compressor fouling, turbine fouling, turbine erosion) and test the methodology against it until the results are consistent with the expectations;
- Test the methodology against random inputs to prove the robustness of the results;
- Test the methodology with measurement noise to verify the viability of the noise isolation block;
- Test the methodology against the simulation of a real schedule to verify how the methodology could react in a real environment;
- Iterate the process until the results are consolidated.

Among all the possible solutions, the combination that has been selected is the Kalman filter in combination with the artificial neural network, the neuro-fuzzy logic and the fuzzy logic.

The KF module has been introduced upfront to filter the data and make the diagnostics precise and accurate. The KF has been selected to compensate for measurements noise. As reported by Simon [17] in fact, it is successful in GT applications providing good results both for noise reduction and moderate bias compensation. Moreover, it is offered and studied in many variants and for many applications. Among others, the possibility of fusing multiple signals, which is one of the objectives outlined in section 3.6.

The ANN has been considered because of its high strength and potential with multiple failures, high dimensional cases that make it applicable to complex problems and flexible to deal with different hardware. As outline by Fentaye et al. [82] in fact, it is widely used in the latest researches, to overcome the drawbacks of the model-based methodologies. Compared to those, in fact, the ANN can work also with a limited number of sensors, it is extremely fast and can easily adapt to any type of engine. The disadvantage mentioned by Bechini [1] of possible unreliability out of the training range, can be bypassed with an extensive training available from simulated data provided by the physical model.

The NFL, instead, has been considered for its precision, as also confirmed by Viharos et al. [33], that allow an accurate relation of input variation with output quantification compensating, in the same time, the disadvantage of poor precision of the ANN. Moreover, the ANFIS variant makes the set up very fast and easily adaptable to every engine architecture.

Finally, the FL will be used to classify the types of failure and it has been selected because of its flexibility. As reported by Meher et al. [76] in fact, there are several typologies of failures that an engine can encounter, and the KF allow integrating them once the key parameters are known and set. Moreover, the flexibility of the KF leave also space to some user experience rules, that may engine or case-specific.

The final scheme selected is then composed by KF, ANN, NFL and FL.

The structure behind the methodology drawn so far is divided into two main sections, preparation and prediction, and driven by Matlab (Figure 16).

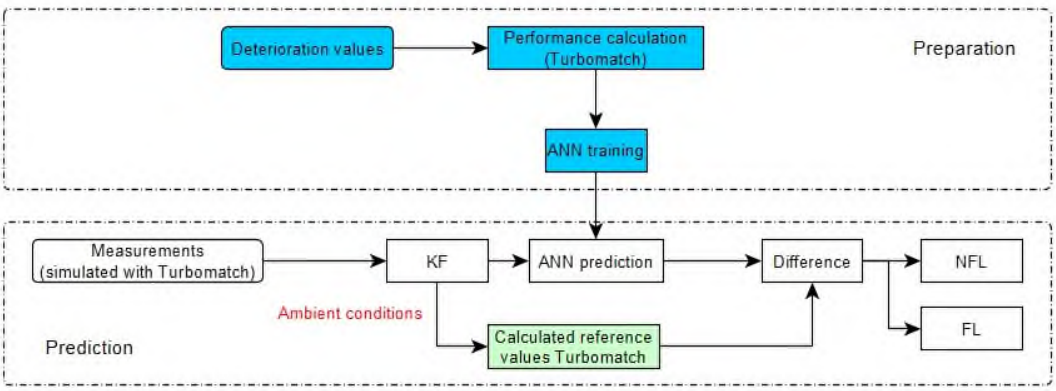

Figure 16 Structure of the methodology

The preparation section includes the performance calculation and the ANN training phase (Figure 17).

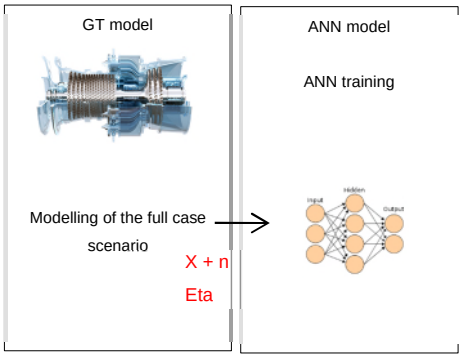

Figure 17 Preparation of the database for the ANN training and base for the simulation and storage of the ANN for future prediction

The first part of the preparation section is the performance calculation done with Turbomatch. Turbomatch is a gas turbine cycle modeller solver and is based on a code able to simulate engine blocks (i.e. compressor, combustor and turbine) and to simulate the deterioration of the components singularly or separately. Turbomatch code is transferred in Matlab and executed from its workspace. The output data are then available in the Matlab environment for the additional phases. The data coming out from Turbomatch are based on assumptions on reduced mass flow, and efficiency variation is taken from the literature ([42] - [50]) that are expected to represent the real behaviour of the engine while subjected to the compressors and turbines failure (deterioration

values in Figure 16). For example, the effect of compressor fouling is the reduction of reference mass flow and efficiency. The effect of turbine erosion is the increase in reduced mass flow and the reduction of efficiency. Once the code is filled with input data, Matlab calls Turbomatch that runs the input files generating two sets of outputs:

- Defective input data X: results of deteriorated or malfunctioning points at a certain imposed ambient condition;
- Reference input data X_{ref}: reference of non-deteriorated points for the same ambient condition imposed.

The data produced to represent the behaviour of the gas turbine under several circumstances from new conditions, to single failure and to multiple failures. In the second part these data are used in two areas:

- As training base to initially train the ANN-based on the full set up of failures;
- As a non-deteriorated reference to calculate the difference between the values predicted by the ANN and the engine at the new status.

The second part of the preparation section is the ANN training. The input of this section is the data from the simulation on the deteriorated engine, that are used to train the network. The range of input points is optimized with the design of experiments in a way to provide reliable results with an acceptable computational effort. The network is a cascade forward neural network that has the same structure of the better-known feed-forward back propagation neural network, with an additional channel from the n^{th} input to the n+2 channel. The architecture of the network is MISO for the efficiency estimation (multiple inputs/single output) that shown to be the most precise in the prediction phase and MIMO (multiple inputs/multiple outputs) for the reference deterioration estimation. The inputs considered for the ANN structure are of two types and lead to two possible variants:

a. Considering the power as possible measurement: in this case, the power measurement is actively included in the ANN;
b. Excluding power. This second option will lead to scheme 3.

The two variants are considered to provide the methodology with the necessary flexibility to work on configurations where the power is not available, or cannot be directly linked to the GT output.

The prediction section is the core of the methodology and includes the data analytics, where the noise n is removed or at least reduced (rn is the reduced noise) from the measurements values, the ANN prediction, where the efficiency values are predicted and the quantification and classification of the failure (Figure 18).

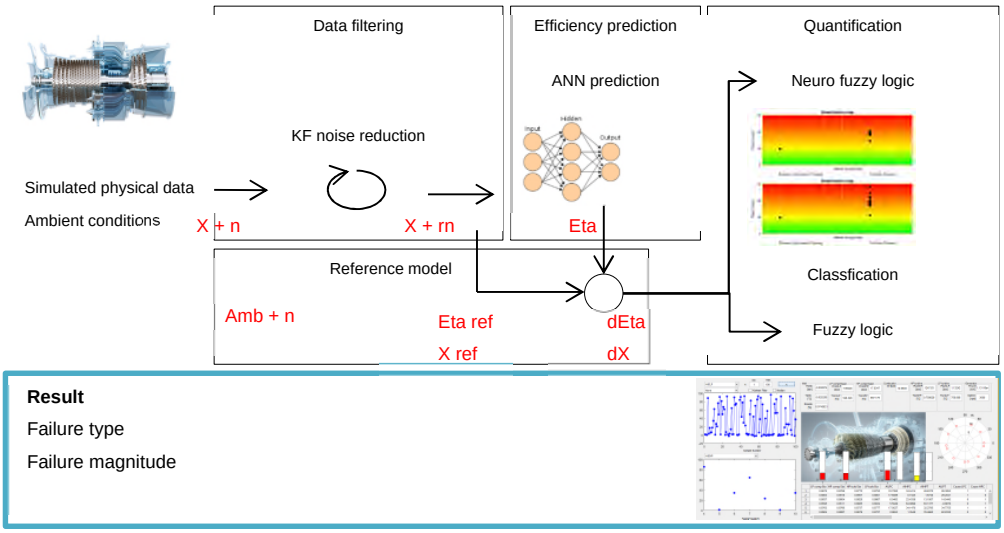

Figure 18 Flow chart representing the concept behind the diagnostics tool – Turbomatch used as reference (Scheme 1 - Scheme 3) – Where X are the measurements, n is the noise, rn is the remaining noise, dEta and dX are the delta efficiency and measurement, Amb are the ambient conditions and ref is the reference of the non deteriorated engine

The first part of the prediction section is the data filtering that is done with the Kalman filter. This block is based on a linear KF and is used to reduce the measurement noise.

The second part is the artificial neural network prediction. Based on the measured inputs, typically pressure and temperature at the main sections and power (if used), the trained network is able to calculate the efficiencies of the components. Practically the network is replacing Turbomatch but in addition, it can predict the output of a deteriorated engine, without needing the deterioration grade to be specified. The reference data instead can be calculated in two ways:

- With Turbomatch taking the ambient conditions as input and the deterioration factors as 1 (with 1 corresponding to the new engine conditions);
- With the ANN itself that, starting from the internal measurements, can be used to predict the deteriorated values and the associated undeteriorated reference.

The first option has the advantage of being more precise as the source model is used to calculate. However, the risk is to include erratic ambient measurements that can mislead the reference. The second option has the advantage of not using the ambient conditions but as it's an approximation of the physical model, it has less precision. These two options lead to two possible schemes that can be set up.

The third part is the component of health estimation. This part takes as input the differences between the calculated reference of a non-deteriorated engine and the results coming from the artificial neural network. The differences are calculated for the physical measurable parameters – pressure and temperature dX – and for the calculated parameters – efficiencies dEta – and this is done for each component of the gas turbine. The difference determines the level of malfunction of a component. For example, if the pressure and the temperature at the compressor exit vary from the nominal level, the compressor is most likely affected by deterioration. On the same way the turbine pressure and temperature, and the combustor fuel mass flow would be affected by the compressor behaviour even if perfectly working.

However, the training behind the neuro-fuzzy logic would account for this and would calculate the efficiency of the components as non-deteriorated. The differences of each component are used as input for the neuro-fuzzy logic that finally quantifies the malfunction level. The quantification is based on a severity scale that goes from 0 (no deterioration or perfect component health) to 100 (maximum deterioration or bad component health). The relationship between the alarm level and the deterioration is established via a cause-effect association within the fuzzy logic technique. The value of the maximum deterioration is related to the higher range that has been found in literature, but the ultimate goal is to relate the maximum value with what is predicted to be the lifetime limit for that particular component. Details about the quantification are reported in the paper submitted to the ISABE 2017 [41]. After the quantification is done, the classification is performed via fuzzy logic. In this section, the alarm points coming from the quantification are taken and categorized. The categorization is done through two boxes: one for the compressor fouling, turbine fouling; another for the turbine erosion. These two boxes are proposed for each component in the gas turbine.

Based on the possible options listed above there are three possible schemes, for this methodology, that can be used:

a. Scheme 1: the first scheme is composed by the Kalman filter placed at the beginning to filter for noise; after it comes the ANN used to predict the performance parameters; the values from Turbomatch are used as reference to calculate the difference that will be used by the fuzzy logic to make the quantification first and the classification after (Figure 18). This scheme is considered as the baseline for this methodology whereas the others are considered variants;
b. Scheme 2: the second scheme uses the Kalman filter at the beginning as per scheme 1; after it also comes the ANN that is used to predict the performance values of the deteriorated engine and to predict the performance reference of the new engine as well; the difference is then calculated among the two ANN values and passed to the fuzzy logic for

the quantification and classification phase (Figure 19). This scheme has the advantage of being faster, since Turbomatch must not be called from the routine, and is not using the ambient conditions;

c. Scheme 3: the third scheme starts also from the Kalman filter that is used upfront; the second part, the ANN is built without the power measurement. The scheme as such does not change, but the block is different. The prediction of the deteriorated engine performance is done with the ANN while the reference is from Turbomatch as per scheme 1. The difference is then used for quantification and classification via the fuzzy logic (Figure 18). This scheme is necessary to give the methodology the flexibility to operate with engines with power available as measurement and engines with power not available.

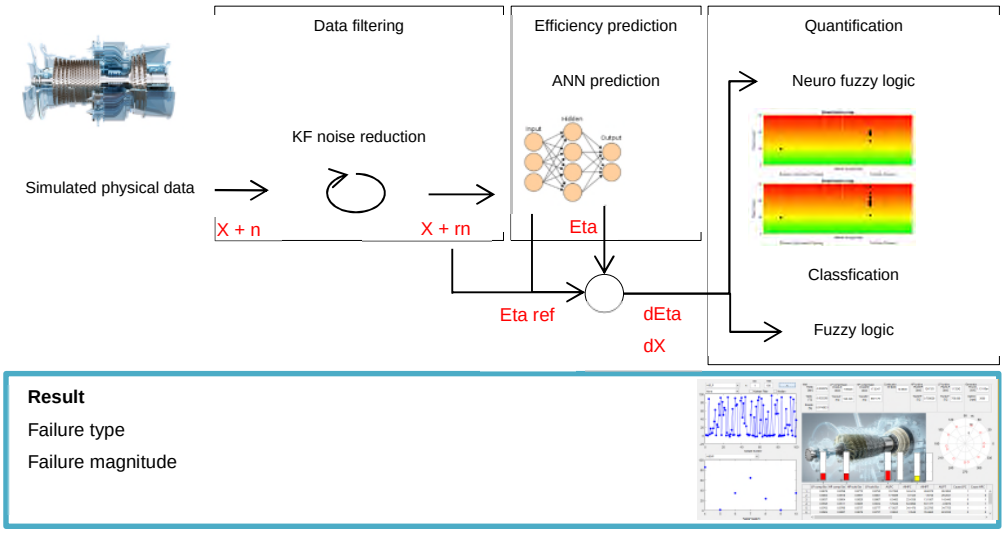

Figure 19 Flow chart representing the concept behind the diagnostics tool – ANN used as reference (Scheme 2) – Where X are the measurements, n is the noise, rn is the remaining noise, dEta and dX are the delta efficiency and measurement, Amb are the ambient conditions and ref is the reference of the non deteriorated engine

4.1 Gas turbine performance modelling

The gas turbine considered is a 2 spool industrial gas turbine of small size providing 11.9 MWe power output with a pressure ratio of 17. The gas turbine has two compressors, one LP and one HP, two turbines, one HP and one LP one burner and one extraction for the cooling system. The efficiency values of the compressor and of the turbine are taken from [75] that proposes values of an engine with a pressure ratio of 17. The overall values are taken from the freely published values of an engine of that size. In particular, the pressure ratio is 17, the power is 11.9 MWe, the exhaust temperature is 485°C and the inlet mass flow is 41.6 kg/s. The cooling is modelled with one extraction after the HP compressor. The amount of cooling air at each pressure level has been tuned to match the exhaust gas temperature. The performance values are modelled in Turbomatch the thermodynamic cycle modeller built and maintained in Cranfield [77].

The gas turbine is modelled at reference load and ISA ambient conditions (15°C, 1.013 bar, 60% RH). In fact, the control concept of the gas turbine is not modelled with this simulation, as the information is not available in the open literature and they would be anyhow engine specific. The modelling type used by Turbomatch is based on components maps. Imposed the reference efficiency, the curves are taken to fit the gas turbine pressure ratio and are selected from the generic available. A better representation would be with engine specific curves, but this will no longer be necessary when the gas turbine data will feed the diagnostics system.

4.2 Deterioration profile simulation

As clarified in the objectives, a key point to prove the validity of the EHM is testing against real data or simulated data. However, as reported by Simon, et. al [39], the reliability reached by the engines as per today makes unlikely

for a gas turbine to experience multiple failures. Therefore, it is challenging to have a sufficient palette of real data including distinguished failures for testing. It is because of it, that a simulation of gas turbine malfunctions has been taken into account. The types of deterioration considered are the compressor fouling, the turbine fouling and the turbine erosion and this is applicable for a total of four components. The combinations of the types of failures and the number of components lead to a certain number of combinations and among those 24 have been selected for this simulation (Table 2). The combinations include no failure, meaning that the engine performs as per design, failure of single components and failure of multiple components.

	LP comp fouling	HP comp fouling	HP turbine fouling	HP turbine erosion	LP turbine fouling	LP turbine erosion
Case-0						
Case-1	X					
Case-2		X				
Case-3			X			
Case-4				X		
Case-5					X	
Case-6						X
Case-7	X	X				
Case-8	X		X			
Case-9	X			X		
Case-10	X				X	
Case-11	X					X
Case-12		X	X			
Case-13		X		X		
Case-14		X			X	
Case-15		X				X
Case-16			X			X
Case-17				X	X	
Case-18	X	X		X		
Case-19	X	X			X	
Case-20	X	X				X
Case 21	X	X	X		X	
Case 22	X	X	X			X
Case 23	X	X	X		X	
Case 24	X	X	X			X

Table 2 Deterioration combination: the deterioration is simulated in all the components taking into account the ratio reported in the literature to make the simulation realistic [50]. Based on the overall literature review [42] to [74], the ratio of efficiency decay and mass flow degradation is 1 to 2 (Table 3).

To have a simulation close to reality, the open-source data has been investigated. The goal is finding the ratio between the flow capacity and the

efficiency that has to be included in the model while simulating the deterioration. Looking at the compressor side and at the fouling type of failure, there is not a unique answer on the ratio Fentaye et al. [82] summarize several papers and report a ratio A/B that ranges from 1:2 to 1:3, where A is the efficiency and B is the flow capacity. In line with that and looking at industrial gas turbines Qingcai et al. [28] report a ratio of 1:3 for the compressor fouling and of 1:2 for the compressor erosion. On the same page, Hepperle et al. [66] report data whose ratio is 1:2. Finally, looking at aero-engine data, Verbist et al. [62] report a ratio of 1:2. Looking a the turbine fouling, the same ratio of 1:2 is supported with engine tests by Stromberg [78] and is also confirmed by Verbist et al. [62] who are reporting the data of the compressor together with the turbine. For the turbine erosion, the main difference resides in the flow capacity change which is increasing instead of decreasing. However, as confirmed by Qingcai et al. [28], the ratio remains 1:2. Based on this portion of the literature review, it is assumed in the simulation that the ratio between the efficiency and flow capacity is 1:2 (Table 3). Qingcai et al. [28] also reported that the deterioration magnitude of the erosion is less aggressive than the deterioration caused by the fouling. It is unlikely to define what the ratio could be, since this is very much dependent upon the type of engine and the conditions it is experiencing, therefore, for this engine the ratio is set as 1:2, meaning that the fouling is causing twice as much deterioration as the erosion.

	Δη (A)	ΔΓc (B)	Relation A/B
Compressor fouling	↓	↓	1:2
Turbine fouling	↓	↓	1:2
Turbine erosion	↓	↑	1:2

Table 3 Deterioration ratio

The range of deterioration that is taken into account is between 0.0% and 7.7% to give room to the ANN to cover all possible conditions. The deterioration levels that are reported from the literature in fact rarely go beyond 5.0%. The worst-case scenario is necessary to make sure that the

EHM will be able to perform in an environment it has been tested to. To reflect this deterioration conditions, Turbomatch deterioration parameters reported in Table 79 in the appendix has been set.

To test the robustness of the methodology dedicated tests have been conducted. The base for the test is the simulated engine described in section 4.1. The main variants that are considered to validate the methodology are:

a. Failure of one or multiple components;
b. Variation of the degradation magnitude;
c. Variation of the level of noise applied to the measurements.

To consider all these aspects, three sets of tests have been set:

a. Constant deterioration, on a single case over x number of points;
b. Random deterioration including different magnitude and different combinations over x numbers of points;
c. Deterioration scenario over a period of time, simulating a close to the real behaviour of a gas turbine over an x period of time.

The first option is a constant deterioration and is performed in a case with 4 failures. The type of failure on the compressor is the fouling and on the turbine is the erosion. This type of simulation is used as a preliminary check for the Kalman Filter to observe its efficacy in the presence of measurement noise. Being the deterioration constant, it is known what the detection should provide and only a measurement error could cause a discrepancy to that. For this case is preferable having higher values of deterioration to be able to still distinguish between the deterioration and the measurement influence. Therefore, the magnitude of 7.4% is selected because is close to the edge of the maximum available values.

The second option is the random deterioration that is used to prove the robustness of the methodology. In this case, at every point, the magnitude varies, so the type of failure. The reference magnitude is allowed to vary between 0.15% to 7.4%, while the failure type can be one of the 24 pre-

established in the modelling (Table 2). This scenario is not likely to happen in reality but is exclusively set up to test the robustness of the methodology.

The third option is a deterioration scenario over a period. The intent is to replicate, starting from the information collected in literature, the gradual deterioration of the gas turbine. To set up this type of deterioration, the open literature has been revised looking for long term deterioration profiles of industrial gas turbines. The deterioration proposed by the researchers varies consistently according to the type of engine considered and the condition encountered by the gas turbine. For instance, Bakken et al. [59] report a deterioration of 2% efficiency on the compressor over 800 OH. Hepperle et al. [66] instead report a compressor deterioration of 1.5% over nine months of time and Kurz et al. [67] write about a deterioration of 2% over 3000 OH. These papers represent the extreme on one side, and the normal deterioration – gradual – on the other. To account for all of these possible conditions three deterioration magnitude have been considered – low L1, medium L2 and high L3. The different deterioration levels are applied to all the components in a sequence (Table 4) where the deterioration level is gradually increased while proceeding with the samples. The ratio among the efficiency and capacity is kept as 1:2 as in all the other previous simulations. Moreover, even if the deterioration data are explicitly referring to the compressor, it is assumed that the levels of deterioration are the same both in the compressor and in the turbine.

	LP comp	HP comp	HP turb	LP turb	Samples
1	L1	L1	L1	L1	500
2	L1	L1	L1	L2	500
3	L1	L1	L2	L2	500
4	L1	L2	L2	L2	500
5	L2	L2	L2	L2	500
6	L2	L2	L2	L3	500
7	L2	L2	L3	L3	500
8	L2	L3	L3	L3	1000
9	L3	L3	L3	L3	500

Table 4 Deterioration sequence for the deterioration scenario over a period of time

In respect to the pre-established sequence, the resulting deterioration profile is shown in Figure 20. The resulting most deteriorated component is the LP turbine as it's the one increasing its deterioration first and the level of reference deterioration reached is 6.5% absolute. The resulting deterioration is multiple failures of all the components and the type of failure that is imposed is the compressor and turbine fouling.

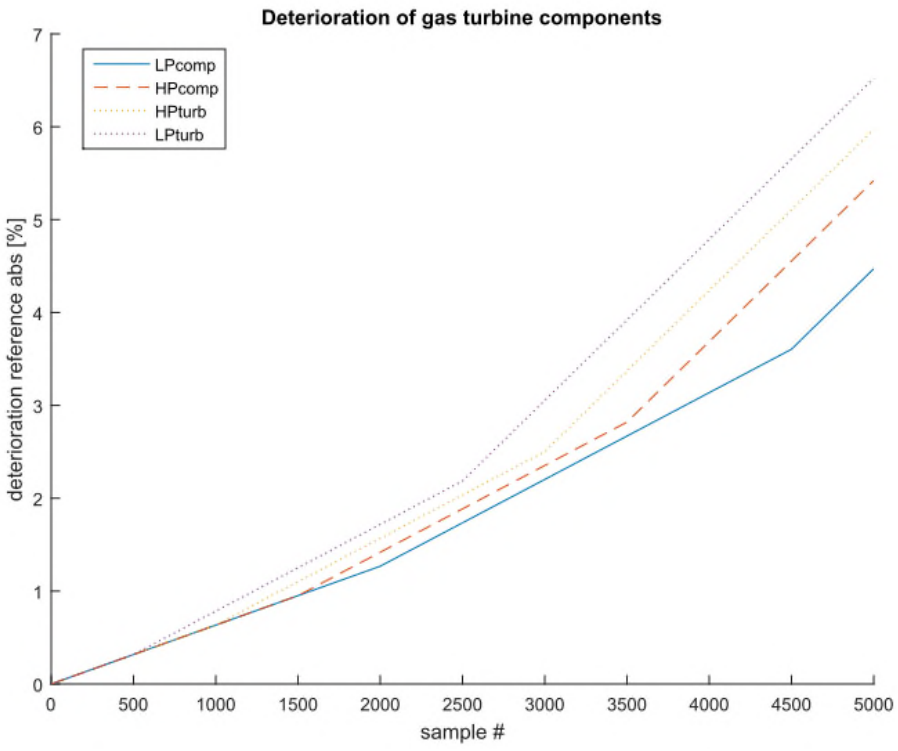

Figure 20 Deterioration reference for the gas turbine components

The scope of this profile is to approach the real profile of the gas turbine and to observe if the methodology is able to correctly quantify and classify this behaviour, according to the imposed exchange rate on the fuzzy logic. There are several points to take into account for this test. The first is the data selection used for the training that can be chosen between the random simulation as per and the random selection plus the dedicated deterioration as per Figure 20. The second is to effectively see how the methodology is behaving under a standard deterioration profile.

4.3 Measurement selection

A key point for the performance-based diagnostics is the measurement selection. The measurements need to reflect the status of the engine, to be sufficiently accurate and to have a frequency response high enough to describe the malfunction the engine is subjected to. The selection of the measurements varies from engine to engine according to the technical limits – e.g. the maximum temperature a probe can be subjected to – the costs and the application of the engine – typically industrial engine for power production or aero engine. The maximum temperature has an influence on the location of the probes. The cost has an influence on the number of measurements installed. The type of engine finally, has a big influence on the selection due to largely different technical limitations, to the different type of application and to a different environment.

The typical measurement equipment for an industrial gas turbine has been described by Jiang et al. [79] and it includes:

a. ambient conditions
    - pressure, temperature and relative humidity
b. intake
    - pressure loss, temperature
c. compressor

- inlet pressure and temperature and outlet pressure and temperature
- inlet guide vane position

d. combustion
    - fuel mass flow, fuel temperature and fuel dynamics

e. turbine
    - wheel space temperature and speed

f. exhaust
    - temperature and pressure loss

g. generator
    - power output, frequency and power factor

The measurements types and locations are summarized in Table 5.

Measurements	
LP compressor inlet pressure	p_1
LP compressor inlet temperature	T_1
LP compressor inlet relative humidity	RH_1
LP compressor exhaust pressure	p_2
LP compressor exhaust temperature	T_2
HP compressor exhaust pressure	p_3
HP compressor exhaust temperature	T_3
Mass flow rate	mf
HP turbine exhaust pressure	p_5
HP turbine exhaust temperature	T_5
LP turbine exhaust pressure	p_6
LP turbine exhaust temperature	T_6
Power	P
Rotational speed spool 1	N_1
Rotational speed spool 2	N_2

Table 5 Set of typical measurements of a two-spool industrial engine [79]

It is also foreseen that for some configurations of the industrial scale gas turbines, the power measurement is not available. In this case, the set of measurements remain the same, but the power is excluded.

The number of measurements can vary depending on the engine type. According to the literature references [80] the amount can vary between 1 and

5 for each location. In addition to that, the literature reports that at the exhaust location the number of probes can raise up to 18 [81].

Based on what previously reported, the set of measurements can be summarized in Figure 21.

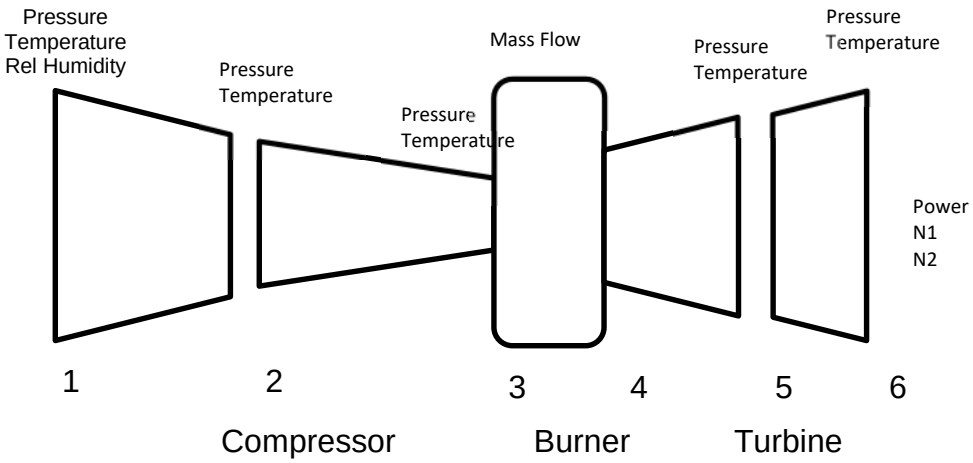

Figure 21 Set of measurements of a two-spool industrial gas turbine

4.4 Measurement uncertainty

In order to build a set of data close to reality, the measurement uncertainty has to be selected to be representative of the real values of the field engines. An exhaustive summary has been presented by Joly et al. [21] (Table 6). This summary refers to an aviation gas turbine.

Engine parameter	Description	Noise (± % range)
N_1	LP relative spool-speed	0.03
N_2	IP relative spool-speed	0.03
N_3	HP relative spool-speed	0.02
P_3	Fan's exit total-pressure (bar)	0.10
P_5	LPC's exit total-pressure (bar)	0.10
T_3	Fan's exit total-temperature (K)	0.40
T_5	LPC's exit total-temperature(K)	0.40
T_{10}	LPT's inlet-temperature(K)	0.40
T_{12}	IPT's inlet-temperature (K)	0.40
T_{13}	HPT's inlet-temperature (K	0.40
Wf	Fuel flow (kg/s)	0.40

Table 6 Noise associated with the sensors – reference to Joly et al. [21]

The level of noise has to be adequate as a too high value might unnecessarily penalize the EHM system, while a too low value might result in not being representative of the real world. Considered that, the level of uncertainty considered in the simulation is 0.4% for the temperature and 0.1% for the pressure. However, for simplicity, the level of noise will be referred to 0.4% while it'll be implicit that the relationship between pressure noise and temperature noise is 1:4.

As the intention of this work is also to detect the measurements failure, a set of measurements failure has been established. Verbist et al. [4] considered, for testing purposes, a set of measurement noise of 0.5%, 1.0%, 2.0% and 4.0%. Based on this reference, the measurement noise considered varies between 0.4% and 2.0% (an increase of 5 times of the nominal value) with increments of 0.4%.

4.5 Kalman filter set up for data analytics

The KF is set up to filter for the measurement noise on the probes. This module makes use of multiple values and makes linear KF self-iteration based on multiple measurements. The number of measurements considered for each location of the gas turbine (location 1 to 6 in Figure 21) is 5 ([80]) for

each measurement except for Power, T_6 and the mass flow. The measurements are 1 for the power, not shown in Figure 21, 18 for the exhaust temperature T_6 as per [81] and 1 for the fuel mass flow. This structure allows the methodology to be flexible for industrial gas turbines, where multiple measurements are normally available. It must be remarked that, since the KF is set up for multiple measurements, it is not applied to single measurements like the fuel flow and power.

The type of KF used for the data filtering is the LKF and it can be detailed as follows:

$$\hat{x}_k^{-\cdot} = A\hat{x}_{k-1} + \omega_{k-1} \tag{14}$$

Where ω is a white noise process covariance with spectral density Q(t) with Q being the model noise covariance.

$$P_k^{-\cdot} = AP_{k-1}A^T + Q \tag{15}$$

$$K_k^{\cdot} = P_k^- H^T / (HP_k^- H^T + R) \tag{16}$$

$$\hat{x}_k^{\cdot} = \hat{x}_k^- + K_k \left(z_k - \hat{x}_k^-\right) \tag{17}$$

$$P_k^{\cdot} = (1 - K_k H) P_k^- \tag{18}$$

The LKF has been selected because the iteration is among multiple measurements of the same quantities, therefore the linear estimation is the most appropriate. Within the architecture, the LKF comes as the first element and has the objective of filtering the measurement noise and combine the information of multiple signals.

The parameters tuned to obtain the best filtration of the KF are the model noise covariance, Q and the observation model covariance R.

The data processing is proposed in two ways – multiple-layer Kalman filter and single layer Kalman filter - according to the scheme shown in Figure 22 and Figure 23. The initial set up consists of five measurements that have to be fused in a single output. With the multiple layer KF the measurements are

divided in three sets: [1,2,3]; [3,4,5] and [5,3,1]. The iteration is done first on each set of measurements, considering the first information of the set to be the initial estimate and after for all the samples taken. Once the first iteration is completed, three new inputs are available for the second layer. Again, the iteration starts, and a unique output is computed. Being the first information the starting point and the last the endpoint, the first and last information are different among the three cases. This is necessary to avoid giving the priority to one or another measurement signal. The scheme is slightly different at the TET where 18 measurements are planned [81]. The change is on the sets of the first layer that are divided in three as: [1,…,6], [7,…,12], [13,…,18].

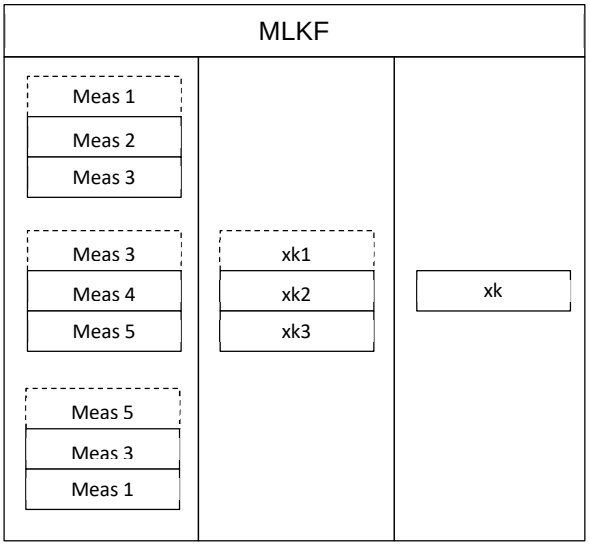

Figure 22 Multiple layer linear KF set up. The dashed lines show the measurement that is used for previous state estimation – Meas 1, Meas 3 and Meas 5 in the first layer, x_{k1} in the second layer

With the SLKF one level is removed and the iteration is done fusing all the five measurements at once. Again, with TET the scheme is different than the one presented in Figure 23 as 18 measurements are used instead of 5.

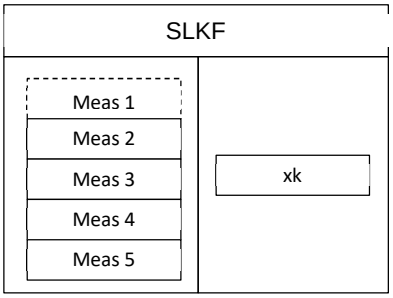

Figure 23 Single layer linear KF set up. The dashed line shows the measurement used for previous state estimation – Meas 1

The KF has been tested for a constant reduction of 7.4% of the efficiency on each component with over 203 samples, varying the level of reference noise from 0% to 2%. The maximum level of noise is 5 times higher than the reference noise that is set to 0.4% according to [21]. The type of failure injected in the compressor is the fouling and the erosion in the turbine (Table 7).

	Fouling	Erosion	Noise [%]
LPC	7.4%	-	0.0;0.4;0.8;1.2;1.6;2.0
HPC	7.4%	-	0.0;0.4;0.8;1.2;1.6;2.0
HPT	-	3.7%	0.0;0.4;0.8;1.2;1.6;2.0
LPT	-	3.7%	0.0;0.4;0.8;1.2;1.6;2.0

Table 7 Failure characterization

According to the charts shown in Figure 24 referring to the temperature, the MLKF leads to a maximum reduction of 83% of the standard deviation. The reduction moves to 76% if the SLKF is used meaning that the second layer of the KF is worth 7%. The maximum reduction is 32% for MLKF and 36% for SLKF if the pressure is considered instead (Figure 25). This means that for the pressure the SLKF results in a higher reduction of the noise. However, the MLKF is better on overall for all the measurements and at different noise levels.

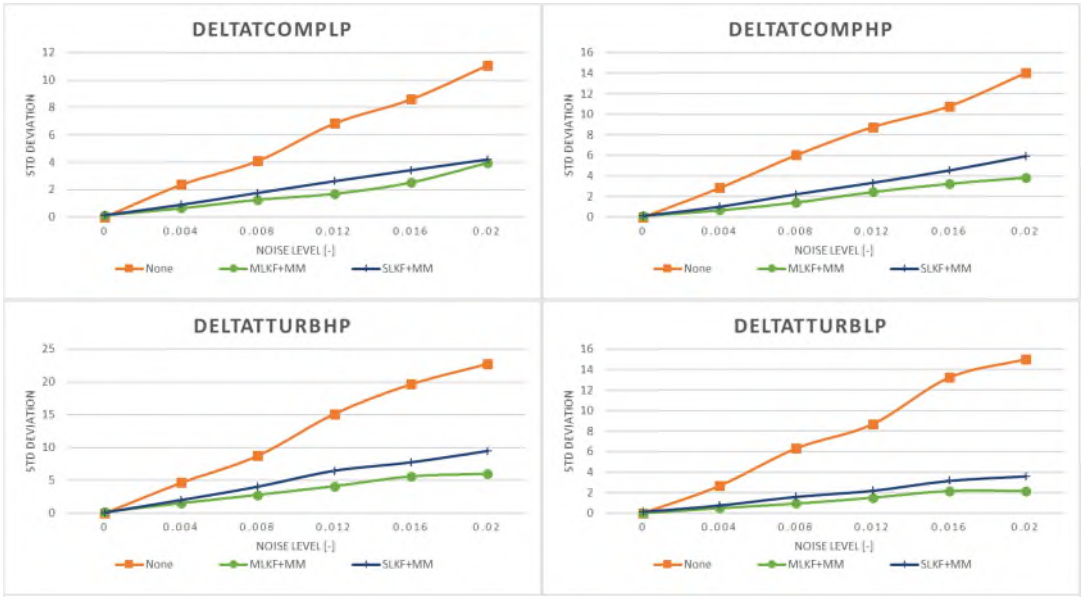

Figure 24 Temperature standard deviation for different noise levels [0-2%] – MLKF vs SLKF vs None

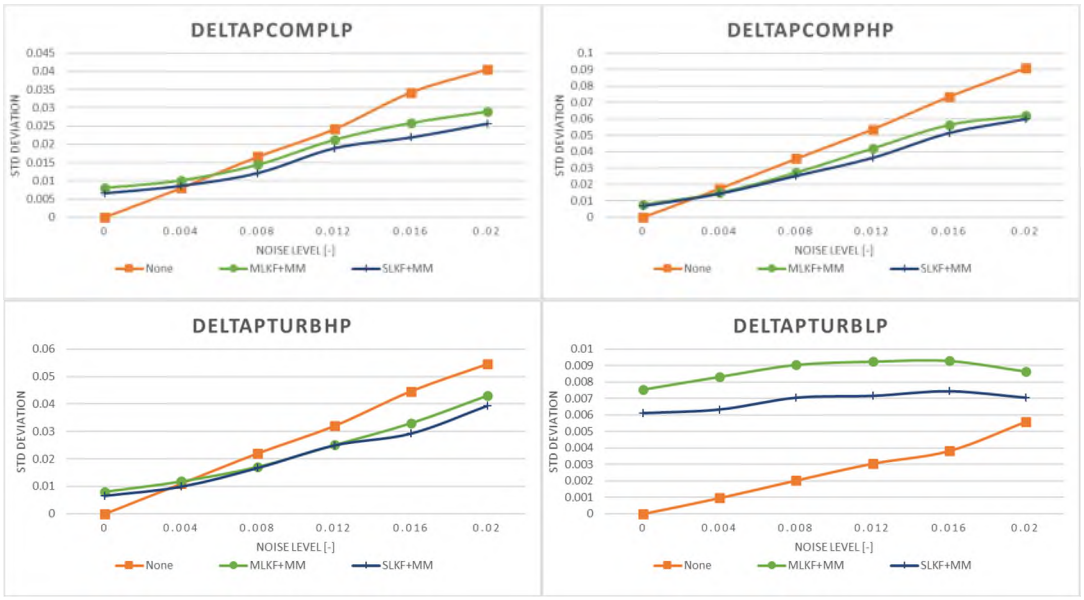

Figure 25 Pressure standard deviations for different noise levels [0-2%] – MLKF vs SLKF vs None

It is implicit that both the MLKF and SLKF require multiple measurements, therefore they are not applied to the power and to the mf measurements, where the probes per location are set to 1.

4.6 Artificial neural network for performance prediction

The ANN is used to predict the key performance parameters based on the inputs received. The base equations and a practical application are well described by Tengeleng in his publication about cascade forward backpropagation neural networks for estimating rain parameters with raindrop size distribution [83].

The equations governing the artificial neural network are:

$$\boldsymbol{n} = \sum_{\boldsymbol{j=1}}^{\boldsymbol{R}} \boldsymbol{w_{1j}} \cdot \boldsymbol{p_j} - \boldsymbol{b} = \boldsymbol{W^T P} - \boldsymbol{b} \tag{19}$$

Where W is the weight matrix containing the input of each neuron and P is the vector. b is the bias assigned to each neuron. Once the output of the integrator n is calculated, the output of the neuron is given by:

$$\boldsymbol{a} = \boldsymbol{f}(\boldsymbol{n}) = \boldsymbol{f}(\boldsymbol{W^T P} - \boldsymbol{b}) \tag{20}$$

The learning process of each neuron of the ANN is calculated as:

$$\Delta \boldsymbol{w_{ij}}(\boldsymbol{t}) = \boldsymbol{w_{ij}}(\boldsymbol{t}+\boldsymbol{1}) - \boldsymbol{w_{ij}}(\boldsymbol{t}) \tag{21}$$

Where $w_{ij}(t)$ is the weight connecting the neuron i at the entry j at the time t. The $w_{ij}(t+1)$ stands for the next entry and the difference between the two is the weight update during the learning process.

The feed-forward back propagation neural network is built with M number of layers (Figure 26). Starting from the output of one neuron, the equation can be modified as:

$$\boldsymbol{a} = \boldsymbol{f^k}\left(\boldsymbol{W^k a^{k-1}} - \boldsymbol{b^k}\right) \boldsymbol{k} = \boldsymbol{1}\text{:}\boldsymbol{M} \tag{22}$$

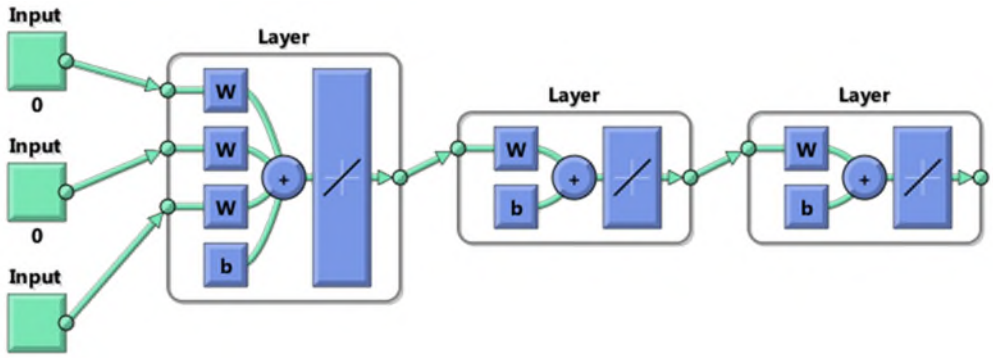

Figure 26 Three layers feed-forward neural network built-in Matlab® – Source Matlab®

The backpropagation takes place if a combination of entries and output are present. In this problem, the entries are the measurements selected in Table 8 and the output are the efficiencies and the deterioration reference parameters calculated with Turbomatch and it can be expressed as:

$$\{(P_d, d_q)\}, q = 1{:}\,Q \tag{23}$$

Where P_d are the entry and d_q are the output. At each instant t, a vector P(t) can be backpropagated to obtain an output vector a(t). Within the learning process the error between the output a(t) and the desired output d(t) is generally minimized through the root mean square function:

$$F(X) = E[e^T(t) \cdot e(t)] \tag{24}$$

Where e is the error, E is the mean of values grouping the sets of weights and bias and X is the vector grouping the sets of weights and bias. The parameter X can be optimized with the steepest falling gradient equation for the weight:

$$\Delta w_{ij}(t) = -\eta \frac{\partial \hat{F}}{\partial w_{ij}^k} \tag{25}$$

And for the bias:

$$\Delta b_i(t) = -\eta \frac{\partial \hat{F}}{\partial b_i^k} \tag{26}$$

Back to the governing equation of the ANN the expression is:

$$n_i^k = \sum_{j=1}^{S^{k-1}} w_{ij}^k \; a_j^{k-1} - b_i^k \tag{27}$$

Defining the sensitivity δ_i^k of $\hat{F}$ of the neuron i of the layer k as:

$$\delta_i^k = -\eta \frac{\partial \hat{F}}{\partial n_i^k} \tag{28}$$

The final expression of the variation of the weight is:

$$\Delta W^k(t) = -\eta \delta^k(t)(a^{k-1})^T(t) \tag{29}$$

For the bias the equation is:

$$\Delta b^k(t) = \eta \delta^k(t) \tag{30}$$

The type of ANN selected for this study is the cascade forward neural network (Figure 27) that is working in a similar as the feed-forward back propagation neural network but is adding a connection between the input and the n+1 layer. Both layouts have been tested, but the second one provided on overall better results.

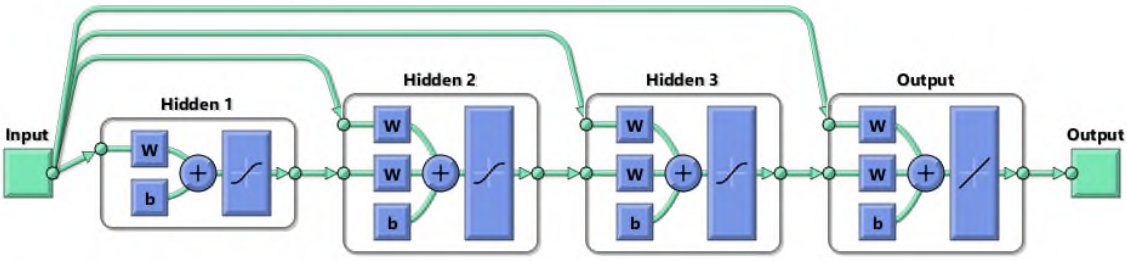

Figure 27 Cascade forward neural network with 3 hidden layers – Source Matlab®

The neural network is set up in three hidden layers that proved to be effective while keeping the computational time reasonable. The most time demanding phase is the learning where the bias and weights are adjusted according to the inputs and the output provided. It must be reminded that this phase is done offline and once the ANN is trained it can be used for the online

prediction. The measurements available and used for the ANN are listed in Table 5 and they include all the locations of the gas turbine and the power. Among the excluded measurements, the exhaust pressure p_6 is not used because of its uncertainty, also generated after the Kalman Filtering Figure 25. The fuel mass flow is a key parameter for the heat balance of the gas turbine but it has been excluded for its uncertainty and the high dependability on other parameters like the composition, the temperature and the pressure, that might affect the ANN prediction. Finally, the ambient conditions have been excluded because of the difficulty in correcting them against measurement issues and the consequent possibility of prediction errors in the case of measurement failures. The pressure instead, is considered in the scheme 1 and scheme 2 but excluded in scheme 3. The resulting set of measurements used is reported in Table 8.

		Scheme1	Scheme2	Scheme3
LP compressor inlet pressure	p_1	No	No	No
LP compressor inlet temperature	T_1	No	No	No
LP compressor inlet relative humidity	RH_1	No	No	No
LP compressor exhaust pressure	p_2	Yes	Yes	Yes
LP compressor exhaust temperature	T_2	Yes	Yes	Yes
HP compressor exhaust pressure	p_3	Yes	Yes	Yes
HP compressor exhaust temperature	T_3	Yes	Yes	Yes
Mass flow rate	mf	No	No	No
HP turbine exhaust pressure	p_5	Yes	Yes	Yes
HP turbine exhaust temperature	T_5	Yes	Yes	Yes
LP turbine exhaust pressure	p_6	No	No	No
LP turbine exhaust temperature	T_6	Yes	Yes	Yes
Power	P	Yes	Yes	No

Table 8 Set of measurements included in the ANN

As specified in section 4.3, the power measurement is included in the ANN if available, while is not included in the other cases. The output of the first four ANN is the efficiencies of each component. Whereas scheme 1 and 3 includes the deteriorated efficiency only, scheme 2 includes also the reference efficiency, that will replace the reference mode. Instead, the output of the last

two ANN is the deterioration factors related to compressor and turbine components. For instance, the pressure loss, the efficiency decay and the mass flow reduction (Table 9). The optimal set up has been calculated with a specific ANN for each efficiency, therefore for each component. Differently, it has been observed that the deterioration parameters can be grouped. Therefore, there are two ANN, one for the compressors and another for the turbines.

ANN id	Scheme 1/Scheme 3	Scheme 2
1	LP compressor efficiency	LP compressor efficiency LP compressor efficiency reference
2	HP compressor efficiency	HP compressor efficiency HP compressor efficiency reference
3	LP turbine efficiency	LP turbine efficiency LP turbine efficiency reference
4	HP turbine efficiency	HP turbine efficiency HP turbine efficiency reference
5	LP Comp Efficiency deterioration	LP Comp Efficiency deterioration
	LP Comp Massflow deterioration	LP Comp Massflow deterioration
	LP Comp Pressure Ratio deterioration	LP Comp Pressure Ratio deterioration
	HP Comp Efficiency deterioration	HP Comp Efficiency deterioration
	HP Comp Massflow deterioration	HP Comp Massflow deterioration
	HP Comp Pressure Ratio deterioration	HP Comp Pressure Ratio deterioration
6	HP Turb Efficiency deterioration	HP Turb Efficiency deterioration
	HP Turb Massflow deterioration	HP Turb Massflow deterioration
	HP Turb dh/T deterioration	HP Turb dh/T deterioration
	LP Turb Efficiency deterioration	LP Turb Efficiency deterioration
	LP Turb Massflow deterioration	LP Turb Massflow deterioration
	LP Turb dh/T deterioration	LP Turb dh/T deterioration

Table 9 Set of ANN for the parameter prediction

The inputs provided to the ANN, in the absence of real data, need to come from the model described in section 4.1. To have a good prediction, the data needs to necessarily cover the operating range of the engine and all its possible deterioration. Among the limitations of the ANN, in fact, there's the inability of making reliable prediction outside the range of training. To avoid this problem the cases considered for the deterioration include all the possible failures and a maximum level of the deterioration higher than the maximum reachable.

The type of deterioration selected is gradual, meaning that the simulation of the degradation goes from zero to the minimum value with a certain number of points. The number of points has an influence on the quality of the prediction, but they have also an influence on the speed of ANN training. The combination of the right amount of points, the gradual deterioration and the reproduction of all the cases should have the intention to provide the proper set of data to avoid overfitting and underfitting.

To check the prediction of the ANN with a different number of points per case a test has been performed. The number of points considered is 20;50;100;150;200 for a total of 5 tests. The results show that the standard deviation reaches its minimum at 150 samples (Figure 28). It has also to be remarked that the deviation increases at 200 samples meaning that an increase in the number of samples is no more beneficial. This trend is particularly emphasised in the LP compressor efficiency, but it is also valid for the other predictions.

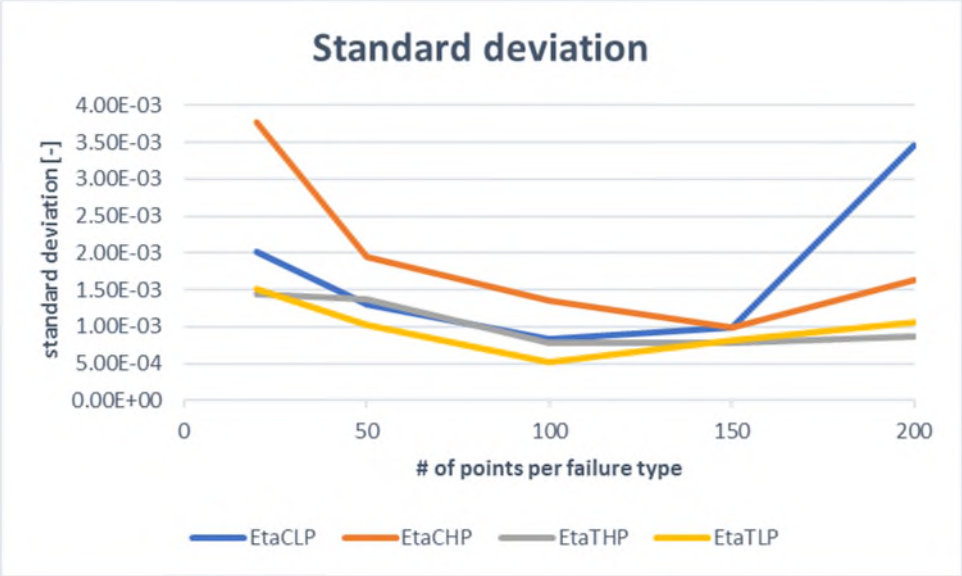

Figure 28 Standard deviation of the efficiency prediction of the gas turbine components

The second factor to consider while varying the number of points is the training time. It is true that the training is mainly done offline, but it should not be unnecessarily high. To evaluate it the highest time is set as reference

(100%) and the other results are scaled on that value. From the results obtained, it is clear that after 100 points, the ANN training time is consistently increasing moving from 14% relative time at 100 points to 100% at 150 points (Figure 29). Interestingly the training time at 200 points is 76%, so lower than the time at 150 points. This means that the ANN converged faster but looking at the standard deviation it converged not exactly at the optimal solution.

Considered these results, the number of samples recommended for this case is 100. The growing computational time, in fact, does not justify the minor improvements in the standard deviation.

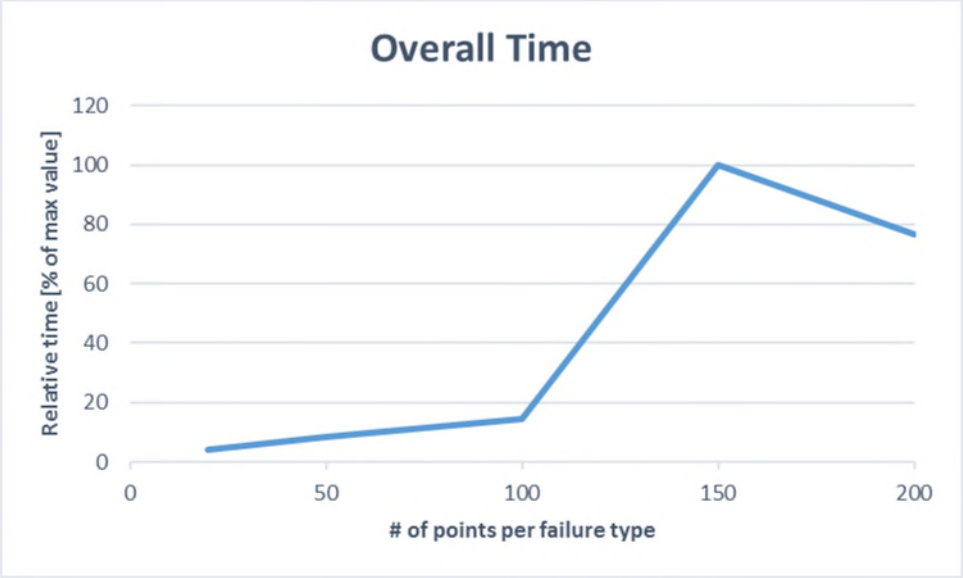

Figure 29 Relative ANN training time vs number of points per failure

The same test has to be repeated with the ANN excluding the power. The number of samples used is the same as the preceding case, but the power is excluded from ANN as available measurement. The resulting standard deviation produces results very close to those obtained with the power included in the measurements (Figure 30). Similarly, to the case with power, one deviation diverges at 200 samples to a value of 1.2E-2. This time the diverging parameter is the turbine efficiency.

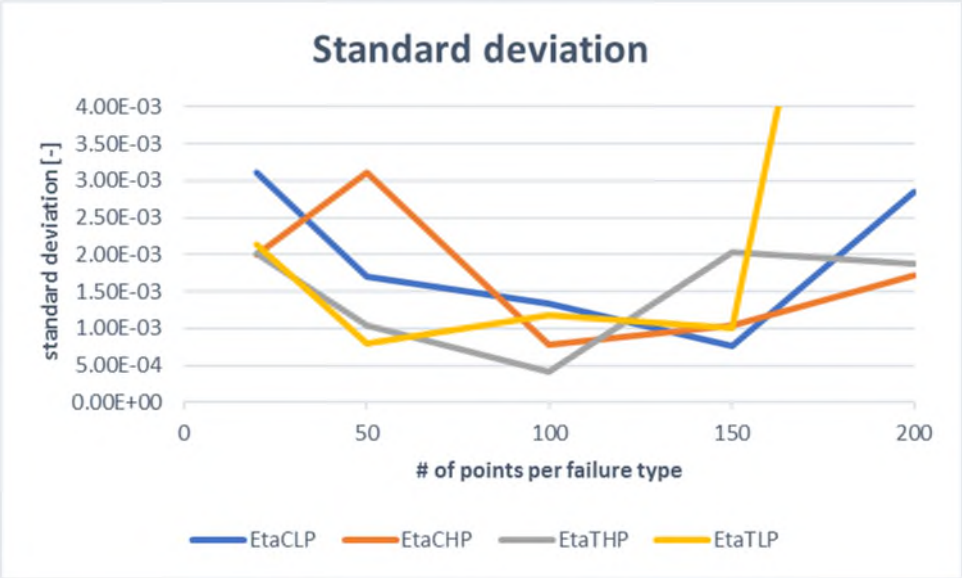

Figure 30 Standard deviation of the efficiency prediction of the gas turbine components

The computational time has also a similar behaviour moving from 20% at 100 points to 100% at 150 points. In this case, the difference between the case with 150 points and the case with 200 points is less pronounced (Figure 31).

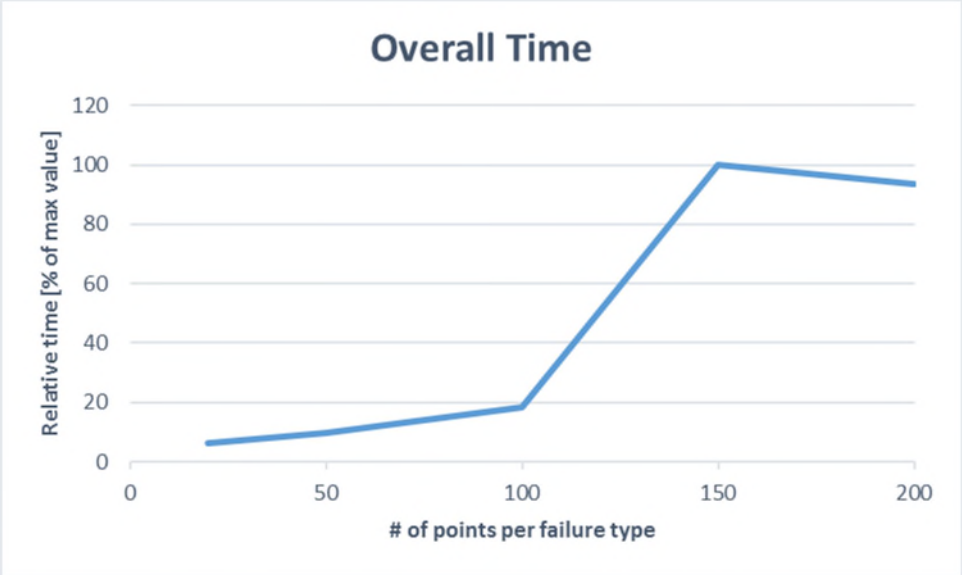

Figure 31 Relative ANN training time vs number of points per failure

With the input set up with 100 points per failure type and the failure characterization of Table 10, a test on the ANN prediction has been

performed. The predicted points instead are 200 to provide a sufficiently large population of the statistical evaluations.

	Fouling	Erosion	Noise [%]	Points
LPC	7.4%	-	0.0;0.4;0.8;1.2;1.6;2.0	200
HPC	7.4%	-	0.0;0.4;0.8;1.2;1.6;2.0	200
HPT	-	3.7%	0.0;0.4;0.8;1.2;1.6;2.0	200
LPT	-	3.7%	0.0;0.4;0.8;1.2;1.6;2.0	200

Table 10 Failure characterization for the ANN deviation testing

The test has been repeated with two instrumentation configurations: with the power measurement included in the set of measurements, with the power measurement excluded from the set of measurements. The results of the simulation including the power measurement (Figure 32) report a maximum standard deviation of 1.0E-2 absolute on the LP turbine. This value is achieved at 2.0% reference measurement error while at the reference noise level the standard deviation decreases by one order.

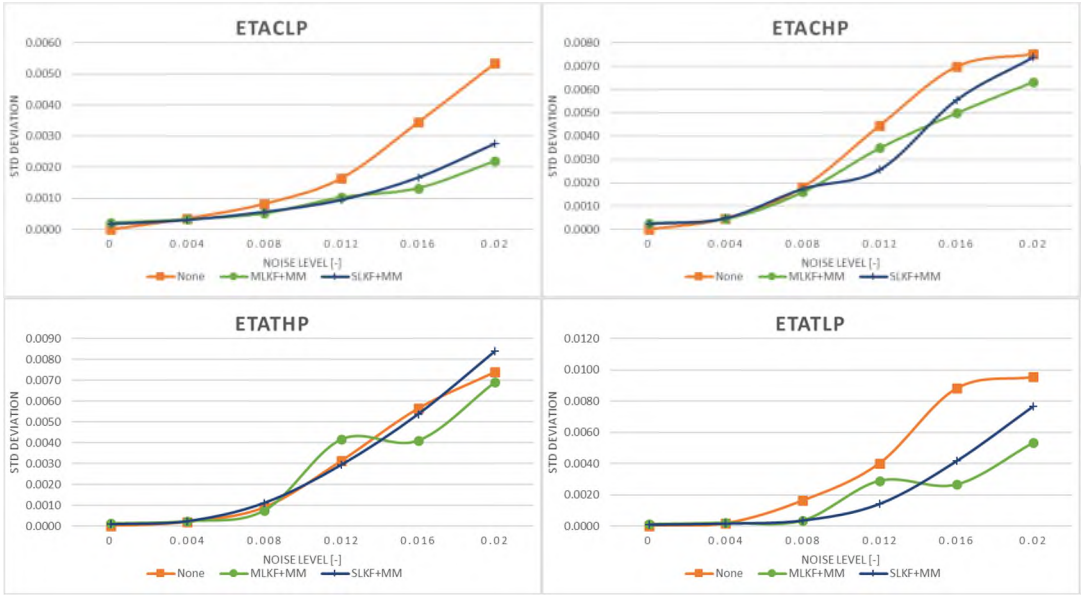

Figure 32 Standard deviation of the efficiency prediction of the gas turbine components – Power included in the measurements

If the power is excluded from the measurements, the efficiency standard deviation at 2.0% reference measurement noise reaches a maximum of 3.5E-

2 on the HP compressor (Figure 33). This deviation has the potential to affect the results of the diagnostics that will be investigated in the Results chapter.

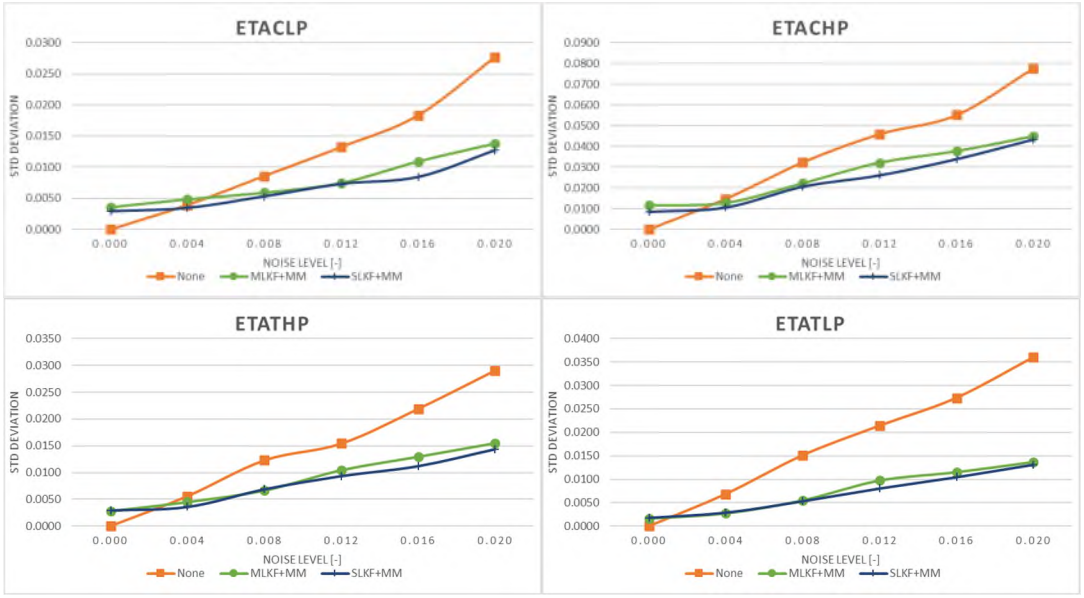

Figure 33 Standard deviation of the efficiency prediction of the gas turbine components – Power excluded from the measurements

4.7 Neuro-Fuzzy logic failure quantification

The first scope of a robust performance-based diagnostics is the correct quantification of the component failure. Correct quantification means precisely relate the severity of an event, as fouling for example, to the engine health status. This is a prerogative of the intelligent engines as reported by Volponi [3]. With this methodology, the scale varies from 0 to 100 where 0 means no deterioration detected on the component and 100 means maximum deterioration detected on the component. To judge the methodology successful, no false alarm should be raised. False alarm, in fact, could mislead the final user leading to excessive maintenance or to a decrease of thrust on the methodology.

The range of efficiency variation will be between 0.0% to 7.7% component deterioration as per literature review [42] - [74]. This range varies from no alarm (0.0%) to an extreme deterioration (7.7%). The minimum level of the

scale – 0 - should alert but not alarm the user while the maximum level on the scale – 100 - should alarm the user and drive him to immediate actions. This scale would ideally be related with the EH, however being the data taken from open literature and being this knowledge engine specific, this remains an open point to be investigated with real data or mechanical simulations. Moreover, the scale may vary within the tests to prove the flexibility of the methodology on a different understanding of the EH or on different engine scenario.

The quantification match is done via ANFIS (adaptive-network-based fuzzy inference system) that couples the Tagaki-Sugeno-Kang fuzzy logic structure and the artificial neural network training and prediction capabilities. This methodology is widely used and as confirmed by Viharos [33] is among the most precise solutions. As shown in Figure 34 the structure is MISO providing one single output for multiple inputs.

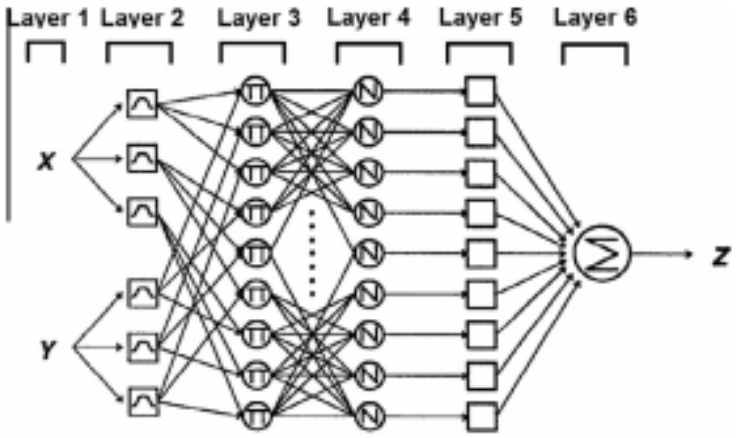

Figure 34 ANFIS structure [33]

The data used for the fuzzy logic set up via the ANFIS tuner is the combination of a single failure, multiple failures and no failures. The use of the ANFIS allows the tuning across all the possible scenarios and a better overall result.

4.8 Fuzzy logic failure classification

The failure classification is the sorting of the component failures into sub-categories. For instance, once the component failure has been quantified, the category of the deterioration is assessed. This part of the research has been already studied by Eustace [31] and Ogaji et al. [29] and defined to be the future trend of the gas path analysis by Volponi [3]. The approach used by the mentioned researches and considered here is the bank of fuzzy logic, that is nothing but a pile of fuzzy logic, driving the diagnosed points into the correct subcategories. The starting point for the bank of fuzzy logic construction is the analysis of the effect of the deterioration on the single physical parameter [42] - [50]. The effect on the measurements and on the calculated parameters is briefed in Table 12, considering 8 deterioration types listed in Table 11. From the effect shown it can be seen that the deterioration on the compressor is affecting the turbine, so the deterioration on the turbine is affecting the compressor and the single effect has to be isolated by the fuzzy logic.

	LP comp fouling	HP comp fouling	HP turb fouling	HP turb erosion	LP turb fouling	LP turb erosion
1	x					
2		x				
3			x			
4				x		
5					x	
6						x
7	x	x	x		x	
8	x	x		x		x

Table 11 Types of deterioration considered for the thermodynamic value change

The structure selected for the isolation of the failures starts from the quantified values, taken from the neuro-fuzzy logic. After the bank of fuzzy logic is addressing the failure to the proper location and so classify it (Figure 35).

	Cases							
	1	2	3	4	5	6	7	8
EtaLPC	↓	↓	↘	↓	↓	→	↓	↓
p2	↓	↑	↑	↑	↑	↓	↑	↓
T2	↘	↑	↑	↑	↑	↓	↑	→
EtaHPC	→	↓	↓	↘	→	↓	↓	↓
p3	↓	→	↑	↓	↘	→	↓	↓
T3	→	↑	↑	→	↑	↘	↑	→
FuelFlow	↓	↓	↓	→	↓	→	↓	↓
EtaHPT	→	↘	↓	↓	→	↘	↓	↓
p5	↓	→	↘	↘	↑	↓	→	↓
T5	→	→	↓	↘	↑	↘	↗	→
EtaLPT	→	→	→	→	↓	↓	↓	↓
p6	↓	→	↑	↓	↓	↑	↓	↘
T6	↑	→	↘	↘	↑	↑	↑	↑
Power	↓	↓	↓	↓	↓	↓	↓	↓

Table 12 Effect of deterioration on physical components of single parameters (ref 7.4% deterioration) - - ↑↓ Variation above-below 2% relative; ↗↘ variation above-below 1% relative; → variation within ± 1% relative

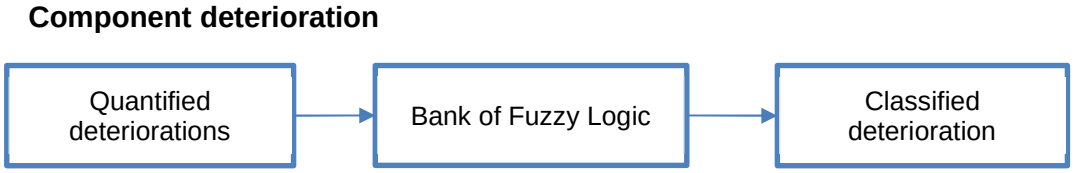

Figure 35 Classification scheme

To graphically pack the engine health estimation, a chart including colour code – for the severity estimation – and categories – for the failure classification – has been created (Figure 36).

The severity estimation is based on a chromatic scale that immediately reflects the magnitude of the failure. The scale is built to range from 0 to 100 as per section 4.7 where 0 is no deterioration and 100 is 7.7% of the component deterioration. However, to consider a hypothetical deterioration higher than 7.7% or quantification higher than 100, the represented scale is extended to 150. The colour coding is built with the traffic light concept moving from green at zero magnitudes, to red at 150. This colour coding is not

necessarily related to a status of the GT component, since the mechanical severity of the failure is not available, however, referring to the deterioration section 4.2 7.7% is on the extreme limit seen in literature, therefore, should raise a critical alarm coded as red.

The classification is divided into three main areas. On the left-hand side compressor fouling, erosion and corrosion together with turbine fouling and corrosion are located. Referring to Table 12 and to the literature review, in fact, they are leading to similar physical effects. The simulation though will include only the compressor and turbine fouling, excluding the compressor and turbine corrosion which remains an open point. On the right-hand side, turbine erosion is placed. The central area is free for other types of failure like for example the measurement errors classification. However, this area is just conceptually free since nothing is programmed from the code to be there yet. This chart is still available in the simulation process but could be replaced with the GUI, that is built with a similar concept.

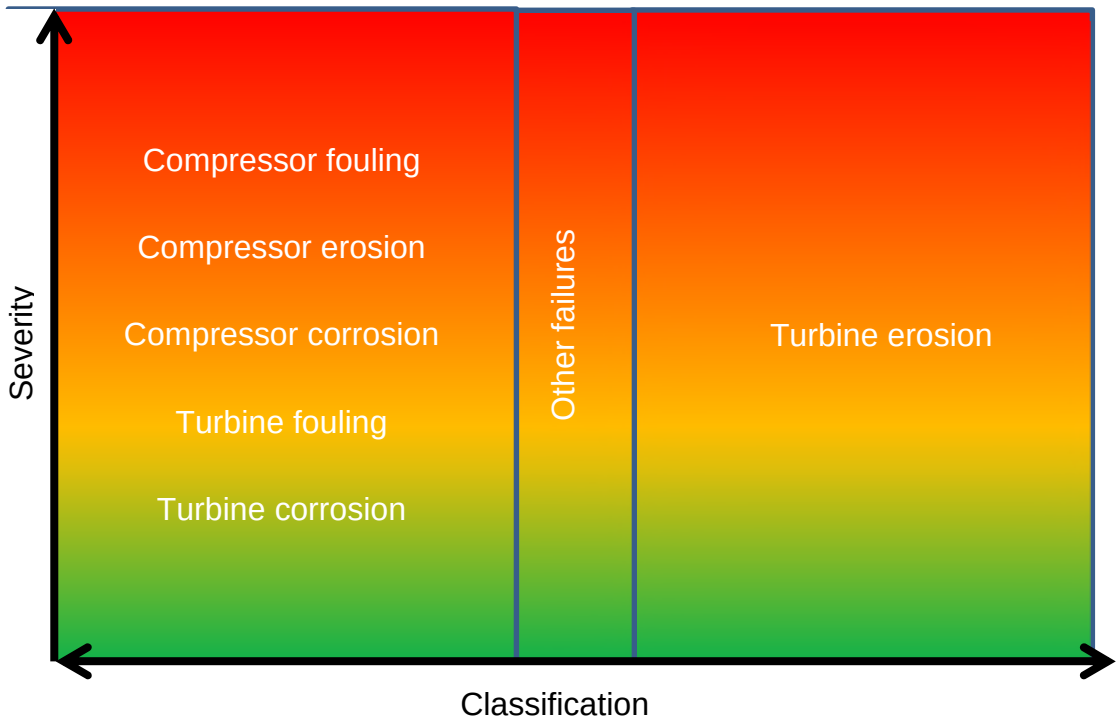

Figure 36 Classification chart

5.0 Results with the reference engine

The reference engine used as a test base for the methodology described in chapter 4.0 is a two-spool industrial engine shown in Figure 37, providing a power of 11.9 MWe. The simulated components are the LP and HP compressor, the HP and LP turbine, the burner and the cooling system, while the diagnostics focus on the LP and HP compressor and on the HP and LP turbine.

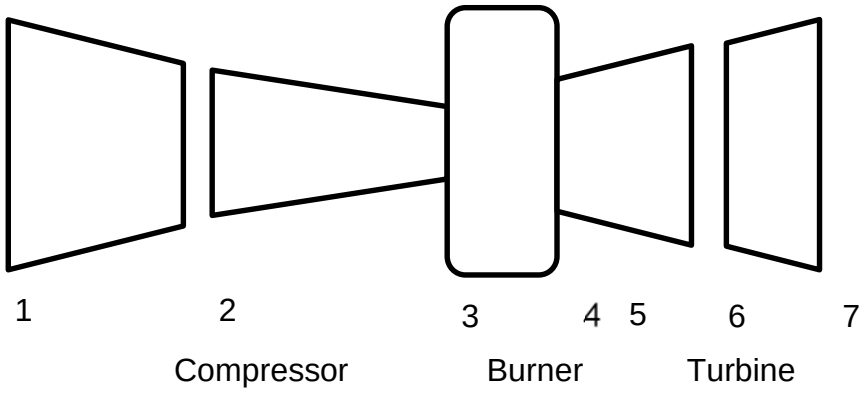

Figure 37 Gas turbine layout and reference numbering

As reported in section 4.2, the methodology will be tested with three different scenarios.

The first scenario is the constant deterioration with 7.4% degradation on the compressor (fouling) and 3.7% degradation on the turbine (erosion); multiple failures (LP compressor fouling, HP compressor fouling, HP turbine erosion, LP turbine erosion); variation of the noise level (Table 10). In this case, the test is combined with measurement noise. The measurements noise cases considered varies from 0.0% to 2.0% with discrete steps of 0.4%.

The second scenario is the random deterioration with degradation within 0.15% - 7.4%; single, multiple and no failures. This test will provide evidence of the robustness of the methodology since it is, as mentioned in section 4.2 the stress test. In this case, the measurement noise is also included. Since the main focus of this test is not the KF and the data filtering section only two extreme noise level have been included: the nominal - 0.4% - and the maximum - 2.0%.

The third and final scenario is the deterioration profile of several months. In this case, multiple failures (LP compressor fouling, HP compressor fouling, HP turbine fouling, LP turbine fouling). With this case, the noise set at the nominal value of 0.4%. This test will be performed with another set of NFL, trained to include the deterioration schedule, in coherence with the ANN combined training. This choice is done to simulate a different understanding of the component health estimation and to prove the flexibility of the NFL on different values set up.

The output of the tests is the success rate. For the quantification, the simulated point is counted if it lies within 3σ standard deviation and for the classification is counted if classified in the right category (also if the quantification is outside the 3σ standard deviation). The standard deviation 1σ is calculated from a dry run with nominal noise (0.4%) and constant deterioration of multiple components. The calculated value is ±2.06 for 1σ and therefore, ±6.18 for 3σ. Since this acceptance range is set, the quantification plots are also adapted with a lower threshold of 5. The value 5 is nothing but the rounding of the 3σ calculated value.

There are several assumptions to be considered before the results are presented that can be summarized here as follows in Table 13

#	Assumption	Test case/scheme of application
1	The noise is normally distributed around the mean value	All cases All schemes
2	The first 7 points of the simulation are not considered	All cases All schemes
3	During the classification success rate evaluation, the points are considered if the quantification is above the threshold value of 2, corresponding to the min deterioration rate of 0.15%	All cases All schemes

Table 13 Assumptions for the test cases

On top of the assumptions, there are also remarks about the set of scheme components (Table 14).

#	Remarks	Test case/scheme of application
1	The configuration without KF is set with one measurement per location. The configurations with the SLKF and MLKF combine the multiple measurements at each location when available	All cases All schemes
2	The LKF is tuned according to the type of test that is performed	All cases All schemes
3	The FL for the classification of the failure is unchanged for all the cases	All cases All schemes
4	The neuro FL for the quantification of the failures is common for the test cases 1 and 2 but is dedicated for the test case 3	All cases All schemes

Table 14 Remarks for the test cases

5.1 Scheme 1

Scheme 1 (Figure 18) uses Turbomatch model as reference and includes the power in the measurements package.

5.1.1 Constant deterioration

5.1.1.1 Measurements noise

The constant deterioration leads to a decrease of the efficiency on all the components and on all the 203 simulated points. The reference decreases, as said in chapter 5.0 is 7.4% on the compressor and 3.7% on the turbine, but the impact has a different magnitude after the LP compressor (Figure 38). The efficiency is obtained with the ANN module, using the simulated measurement values as input.

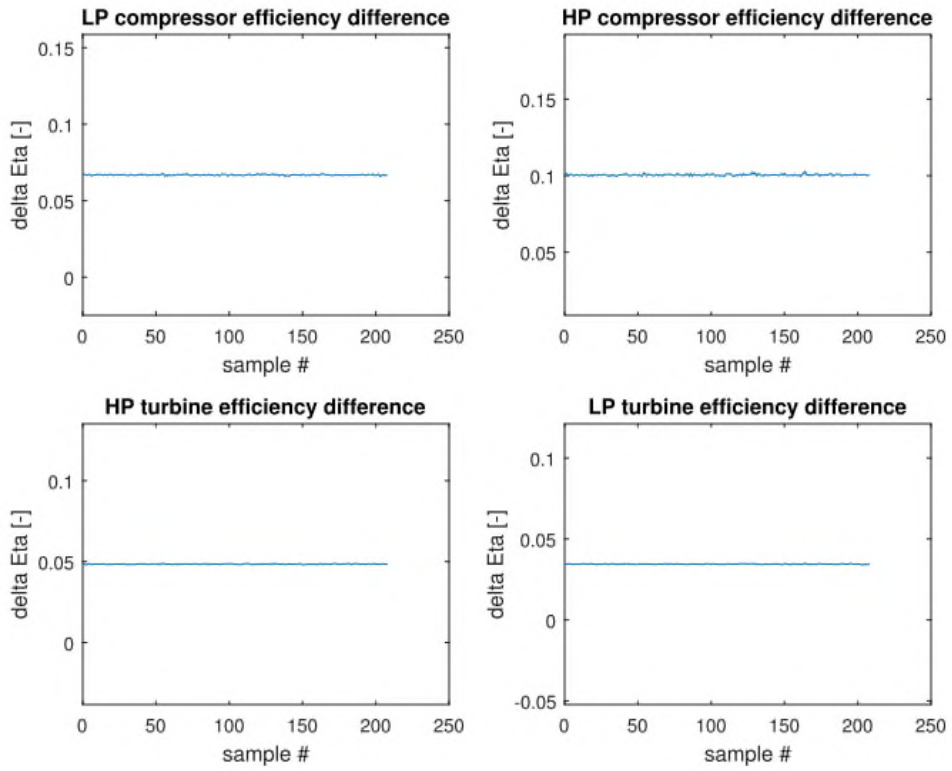

Figure 38 Deterioration imposed on the gas turbine components – 0.4% noise

In the first test, as briefed in all the components are subjected to a constant deterioration (fouling on the compressors and erosion on the turbines). Meanwhile, increasing measurement noise is implanted to verify if the methodology can address it and correctly quantify and classify the failures in that situation. As mentioned before, the points analysed will be considered successful only if within 3σ standard deviation and if correctly allocated in the right failure type. Therefore, being the deterioration set to 7.4% for the compressor fouling and 3.7% for the turbine erosion, all the points will have to have a corresponding quantification of 96. On the other side, the compressor points will have to end in the fouling category and the turbine points in the erosion field. The results obtained with the first combination considered – without KF applied – show a constant value around 96 for all the points that are in line with the expectations (Figure 39).

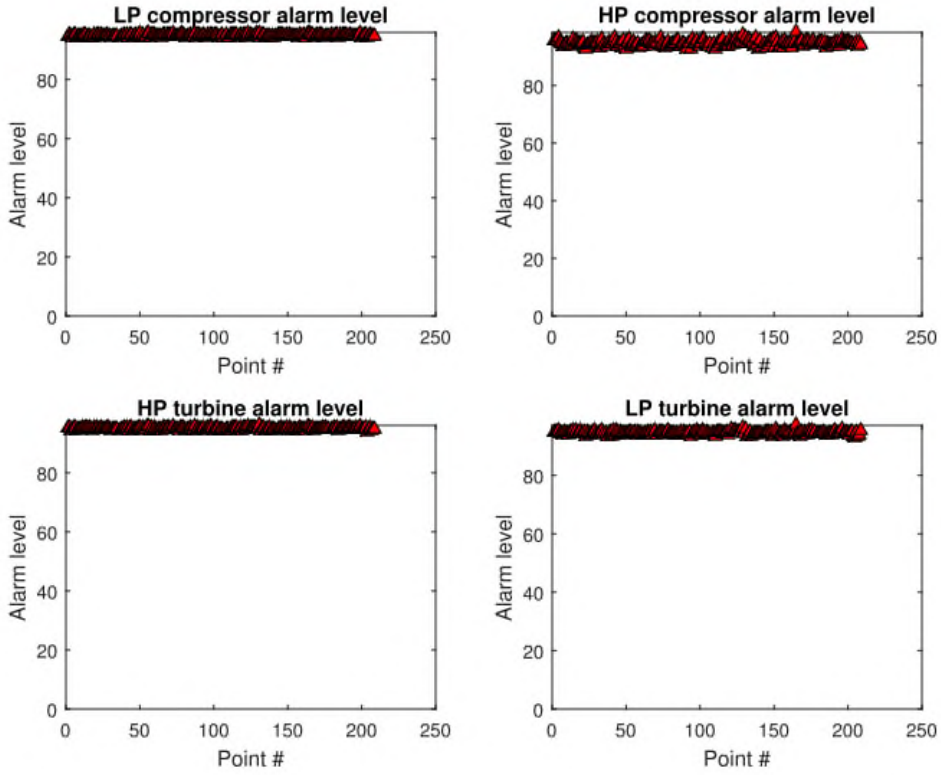

Figure 39 Quantification of the gas turbine failure – 0.4% noise no pre-filtering

The results are then grouped in a single graph that quantifies, giving a component health status, and classify, defining the type of failure. Considered

that the type of failure is kept constant for this first simulation, the values should end in the same section of the graph. The results obtained confirm that the values are correctly quantified, and all the samples are classified in the appropriate category (Figure 40).

As the magnitude of the deterioration is constant, and the parameter that varies is the measurement noise, the component that makes the difference in this phase is the KF that, as shown in section 4.5, is reducing the measurement noise. The effect takes place starting from the nominal noise of 0.4%. With this level of noise, the gas turbine malfunction is still correctly quantified and classified (Figure 40 and Figure 41).

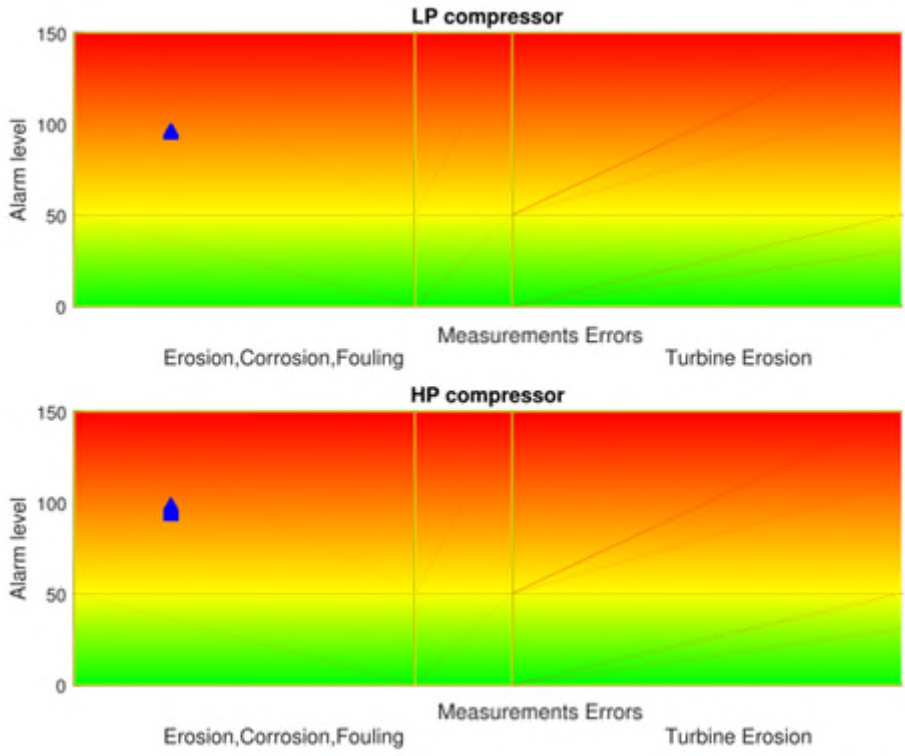

Figure 40 Quantification and classification of the compressor – 0.4% noise no pre-filtering

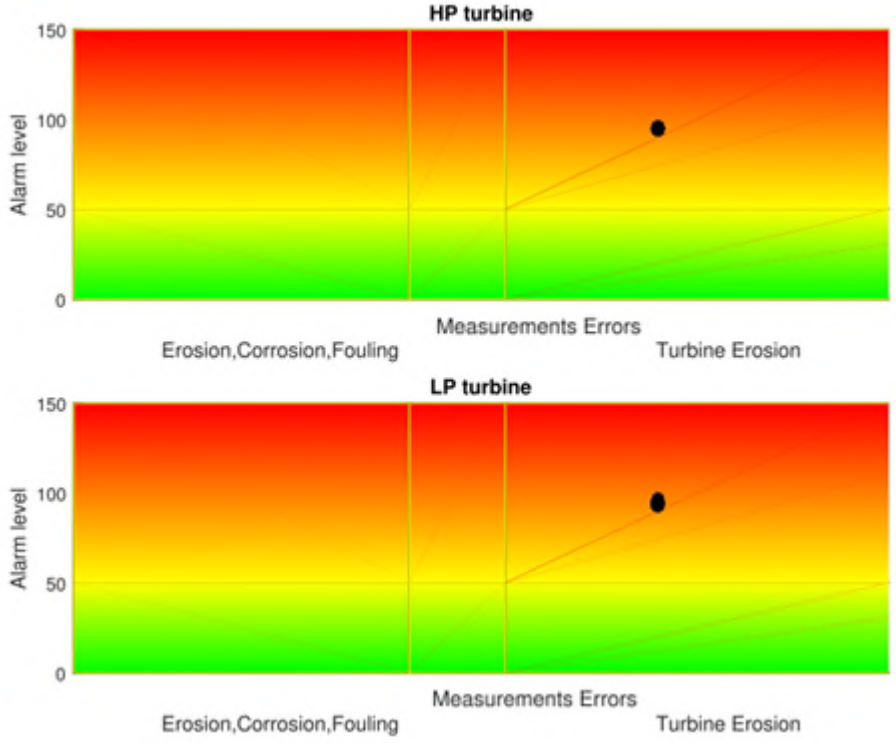

Figure 41 Quantification and classification of the turbine – 0.4% noise no pre-filtering

However, if the level of noise increases, there is a clear disturbance on the quantification and of the classification. Taking into account the 2.0% measurement noise, for example, the quantification instead of being around 96, ranges between 50 and 140 where 50 corresponds to a medium averaged engine health status and 140 corresponds to an extremely highly deteriorated engine (Figure 42 and Figure 43). The reason behind this increase is obviously driven by the noise, that is causing oscillation of the physical value and of the predicted value around its average. Translated this means that the health status is underestimated in one case, leading to slower reaction on the maintenance actions, while is overestimated in another case leading to faster reactions. The classification instead, leads to several false categorizations as some samples fall into the wrong category. This is valid especially for the turbine fouling/erosion since the KF has to distinguish between the two. This result is also related to the measurement noise that causes the delta mass flow reference deterioration used for the failure classification (Table 81), to

diverge from its exact value, misleading the KF. In this case, the maintenance would be addressed on a wrong action.

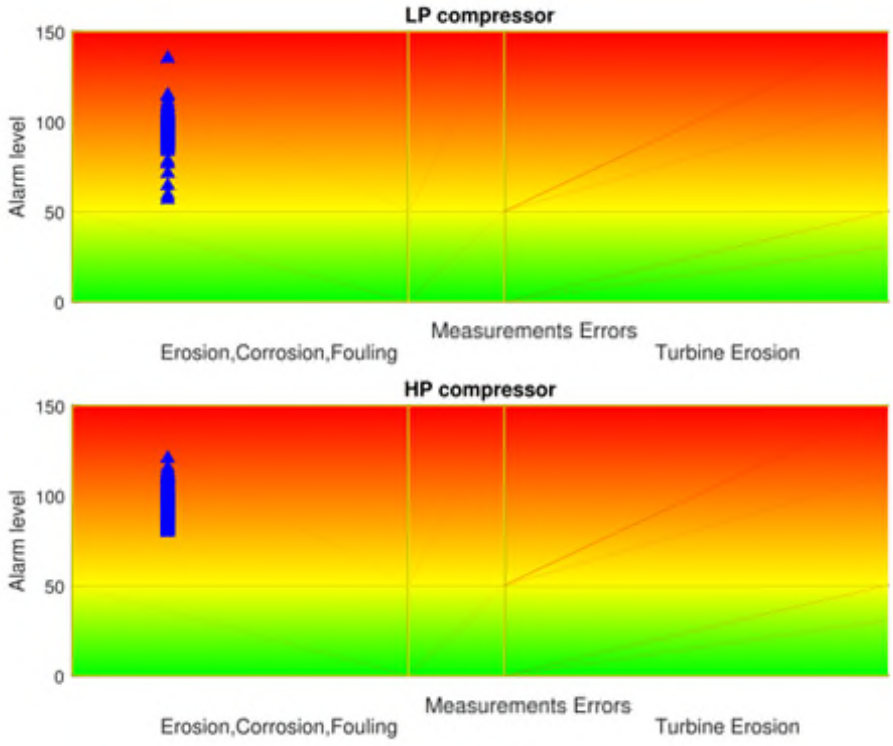

Figure 42 Quantification and classification of the compressor – 2.0% noise no pre-filtering

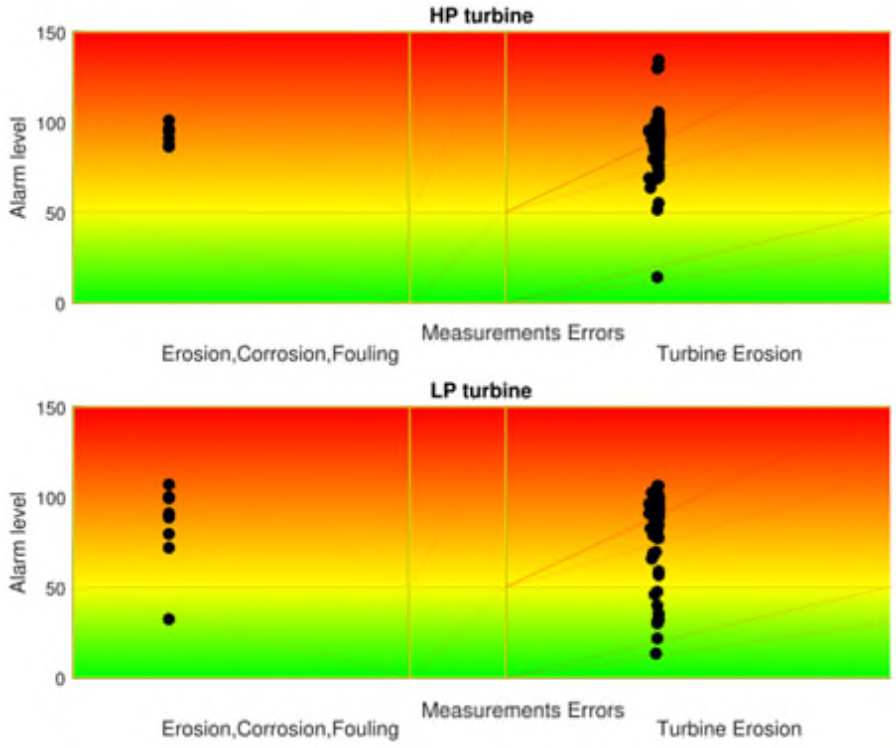

Figure 43 Quantification and classification of the turbine – 2.0% noise no pre-filtering

The module dedicated to the data analytics/data filtering is the KF described in section 4.5. The variants studied in this thesis are the SLKF and the MLKF. Looking at the results including the SLKF with 2.0% measurement noise, the spread of the magnitude is decreased, leading to more robust diagnostics. On the compressor side the spread head between 80 and 110 with only one point at 50 (Figure 44). The worst situation is on the LP turbine, where the spread is still within 40 and 110 (Figure 45).

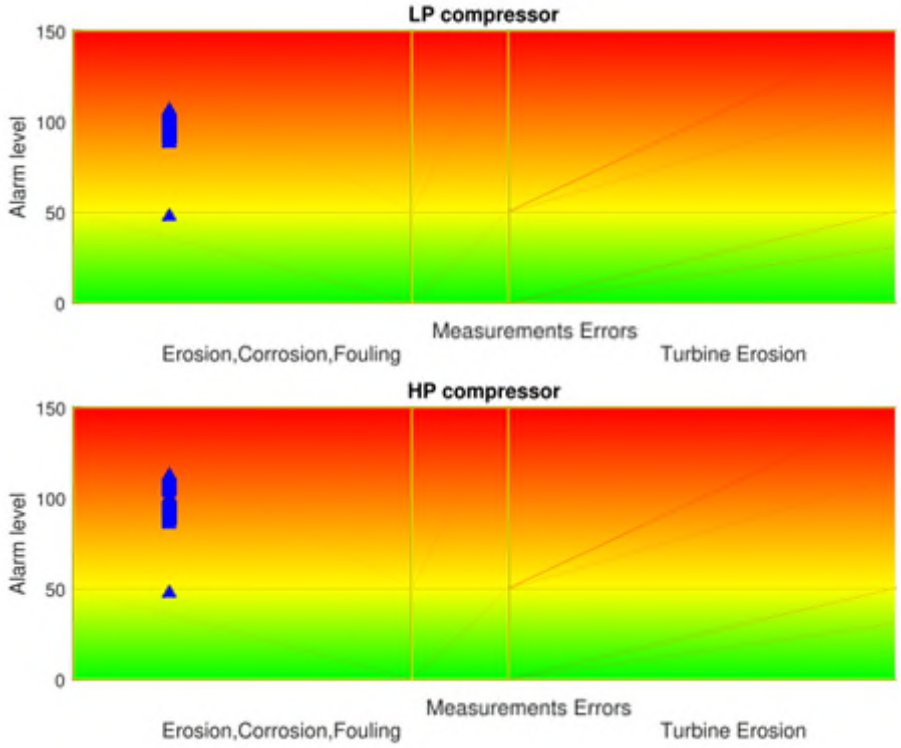

Figure 44 Quantification and classification of the compressor – 2.0% noise SLKF

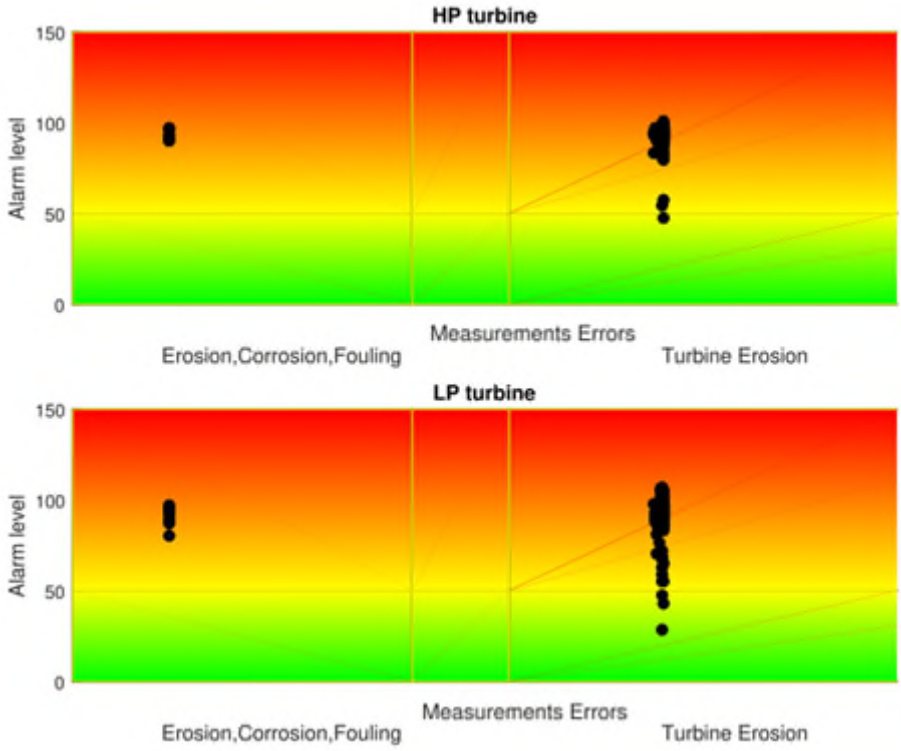

Figure 45 Quantification and classification of the turbine – 2.0% noise SLKF

Looking at the success rate (calculated points within 3σ compared to the reference points), the overall results of the first test show a rate with all the combinations above 99.5%, both for 0.0% noise and for the nominal noise of 0.4%. This result shows that this technique is capable of correctly detecting the multiple failures on almost all the samples.

While the level of noise is increasing, the data analysis and correction starts to become more important. At 0.8% reference noise in-fact, the success rate on the quantification decrease to 95.0% if no KF is applied. With the maximum noise level of 2.0%, the rate of success is decreased to 58.8% without the KF block. The rate increases to 72.4% if also the MLKF is employed and to 73.4% if the SLKF is used with the SLKF having a higher minimum (Table 15). However, it must be noticed that the MLKF is performing better on overall reaching a maximum quantification success rate – with 2.0% reference measurement noise - of 91.0%.

These results confirm that the KF block can reduce the noise, as already seen in section 4.5 and improve the success rate up to 23.6% absolute delta (ref MLKF HP comp). It must be remarked that this delta does not match with the relative delta improvement on the measurement noise since the other blocks such as the ANN and the FL are involved.

	Noise	0.0%	0.4%	0.8%	1.2%	1.6%	2.0%
None	LP comp	100%	100%	100%	95.5%	79.9%	76.9%
	HP comp	100%	100%	95.0%	84.9%	69.3%	58.8%
	HP turb	100%	100%	100%	95.5%	77.4%	66.3%
	LP turb	100%	100%	98.0%	86.4%	65.3%	63.3%
SLKF	LP comp	99.5%	99.5%	98.5%	97.0%	89.4%	82.4%
	HP comp	99.5%	99.5%	98.0%	96.0%	88.4%	75.4%
	HP turb	99.5%	99.5%	99.5%	98.0%	94.0%	83.9%
	LP turb	99.5%	99.5%	99.0%	94.0%	85.4%	73.4%
MLKF	LP comp	99.5%	99.5%	99.0%	94.5%	92.5%	86.4%
	HP comp	99.5%	99.5%	98.5%	94.0%	90.5%	82.4%
	HP turb	99.5%	99.5%	99.5%	97.0%	96.0%	91.0%
	LP turb	99.5%	99.5%	99.0%	92.5%	91.5%	72.4%

Table 15 Success rate for the failure quantification

Looking at the graphs instead, it can be observed that the success rate is above 90% for the MLKF combination at noise up to 1.6%. After, the

component with a lower success rate is the LP turbine (Figure 48). With the SLKF the 90% level is kept up to 1.2% measurement noise (Figure 47). Without filters instead, the 90% rate can be achieved only up to 0.8% reference noise (Figure 46). This means that the methodology can deal with noise up to 1.6% while providing the component health estimation (quantification) with a success rate above 90% (ref. MLKF).

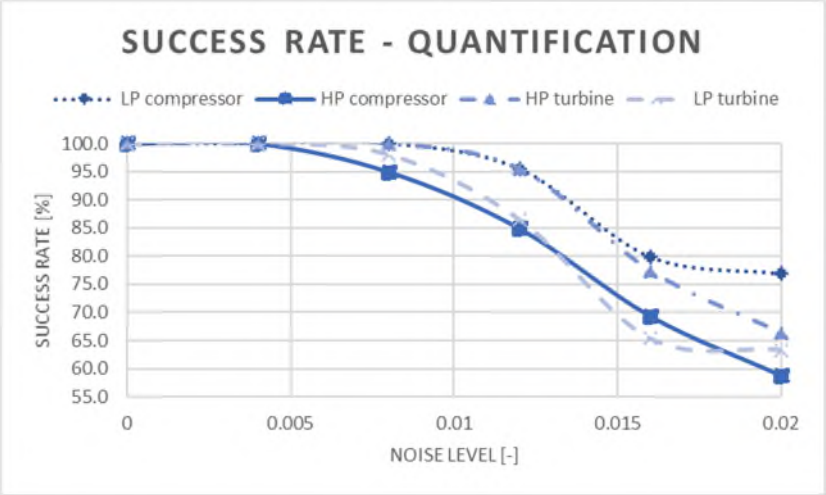

Figure 46 Success rate for the failure quantification – Combination No KF

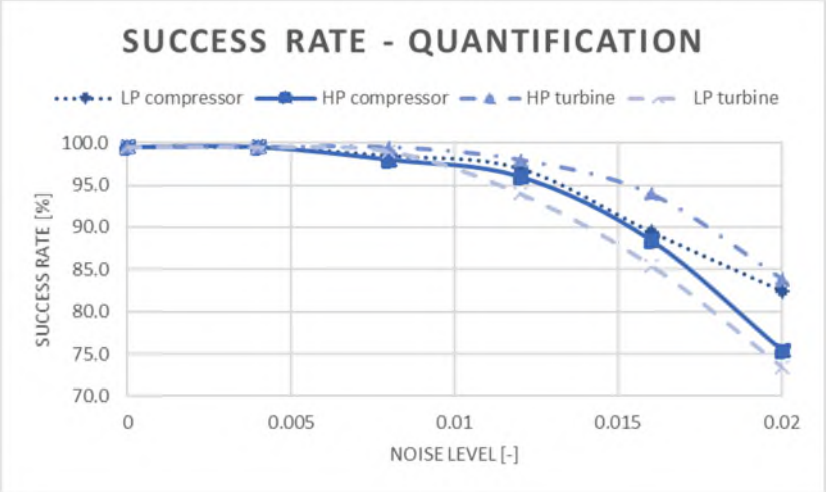

Figure 47 Success rate for the failure quantification – Combination SLKF+ANN+NFL+FL

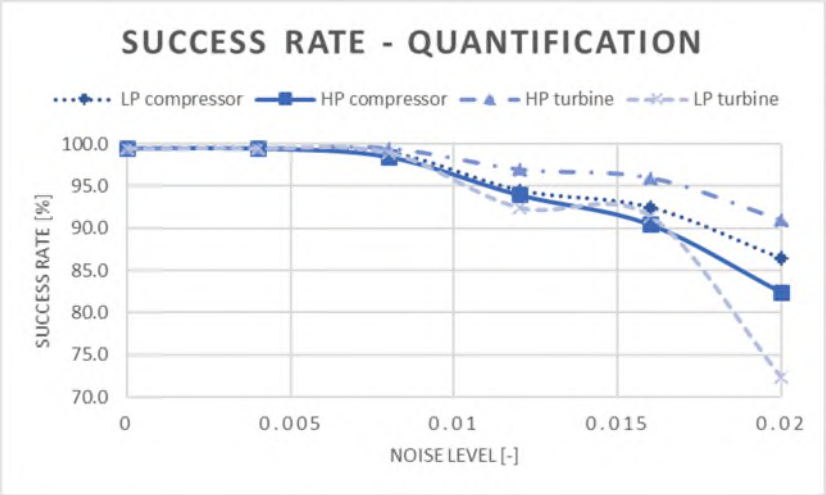

Figure 48 Success rate for the failure quantification – Combination MLKF+ANN+NFL+FL

The results of the classification with nominal noise show a success rate above 93.0% without filtering and at 100% with MLKF. With the maximum noise, the success rate decreases to 86.0% if no filtering is used and increases again to 92.4% with MLKF. Remarkable results are also achieved with SLKF, where the minimum success rate is 90.9% with 2.0% reference measurement noise (Table 16).

These results confirm that the methodology can deal with noise up to 2.0% while providing a classification with rates above 90% (ref. SLKF and MLKF). Moreover, the KF has shown to be a key contributor since is improving the success rate up to 7.4% (ref. MLKF LP turbine).

On the overall, the methodology is capable of correctly quantify and classify a GT multi failure also with the presence of measurement noise. However, the methodology can deal with noise up to 1.6% while keeping the success rate above 90% both for quantification and classification.

	Noise	0.0%	0.4%	0.8%	1.2%	1.6%	2.0%
None	LP comp	100%	100%	100%	100%	100%	99.0%
	HP comp	100%	93.0%	93.0%	94.5%	98.0%	97.0%
	HP turb	100%	100%	100%	99.5%	95.5%	94.0%
	LP turb	100%	100%	100%	99.0%	92.5%	86.0%
SLKF	LP comp	100%	100%	100%	100%	100%	100%
	HP comp	100%	100%	96.0%	90.4%	92.9%	94.9%
	HP turb	100%	100%	100%	100%	99.5%	92.9%
	LP turb	100%	100%	100%	100%	99.0%	90.9%
MLKF	LP comp	100%	100%	100%	100%	100%	100%
	HP comp	100%	100%	98.5%	96.0%	97.5%	92.4%
	HP turb	100%	100%	100%	99.5%	98.5%	92.9%
	LP turb	100%	100%	100%	99.5%	98.5%	93.4%

Table 16 Success rate for the failure classification

It must be noticed that the classification is less affected by the noise compared to the quantification. The minimum rate in-fact is 86.0% at 2.0% measurement noise, compared to the 58.8% for the quantification. These results are justified by the higher threshold of the classification FL that is less compromised by the noise (Figure 49). While the quantification must relate each point to a quantity, in fact, the classification has to place the point in the right category. However, the classification rate is also benefitting from the data filtering either SLKF or MLKF. Additionally, the HP compressor rate with SLKF shows a decay at 1.2% measurement noise (Figure 50). The effect is lower if the MLKF is used (Figure 51). The reason for this decay resides in a higher uncertainty of the HP compressor efficiency that, even if the noise is lower, peaked up. This effect is more pronounced at 1.2% noise and is then slowly resolved.

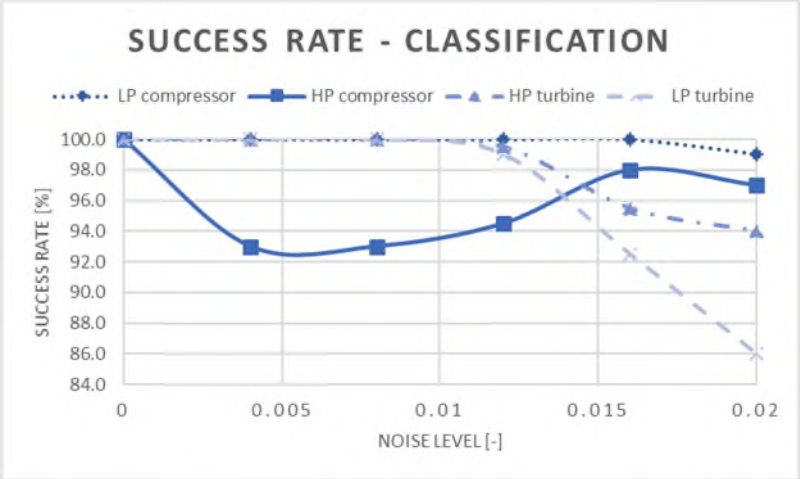

Figure 49 Success rate for the failure classification – Combination No KF

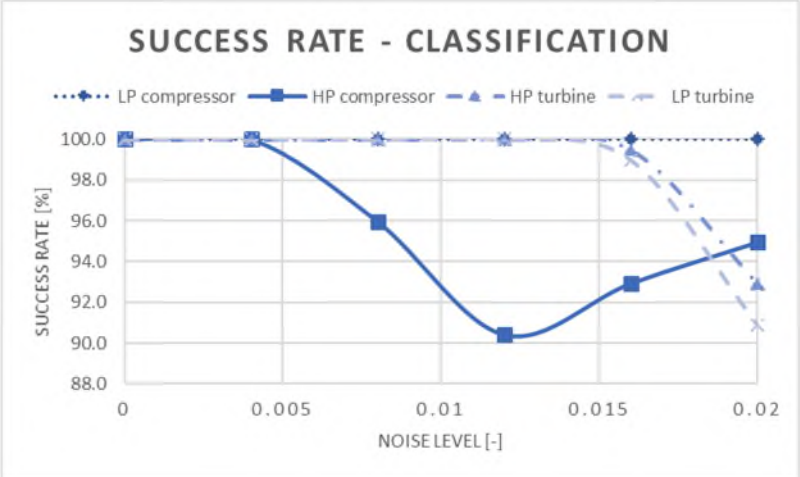

Figure 50 Success rate for the failure classification – Combination SLKF+ANN+NFL+FL

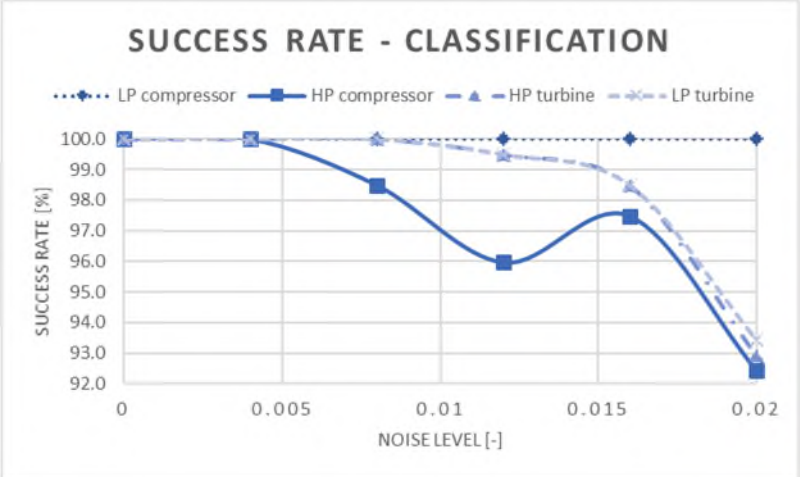

Figure 51 Success rate for the failure classification – Combination MLKF+ANN+NFL+FL

5.1.2 Random deterioration

The random deterioration consists of 203 points with deterioration level between 0.15% and 7.4% (Figure 52). As described in chapter 5.0 this is the stress test of the methodology and its scope is the validation of the methodology under severe conditions. The failure type, in fact, will change from one sample to the other moving from single failure to multiple failures, changing the failures types and varying the failure magnitude between 0.15% and 7.4%. This case is much worse than what could happen in reality and that's why a success rate above the target with this test represents a strong proof and achievement for the methodology.

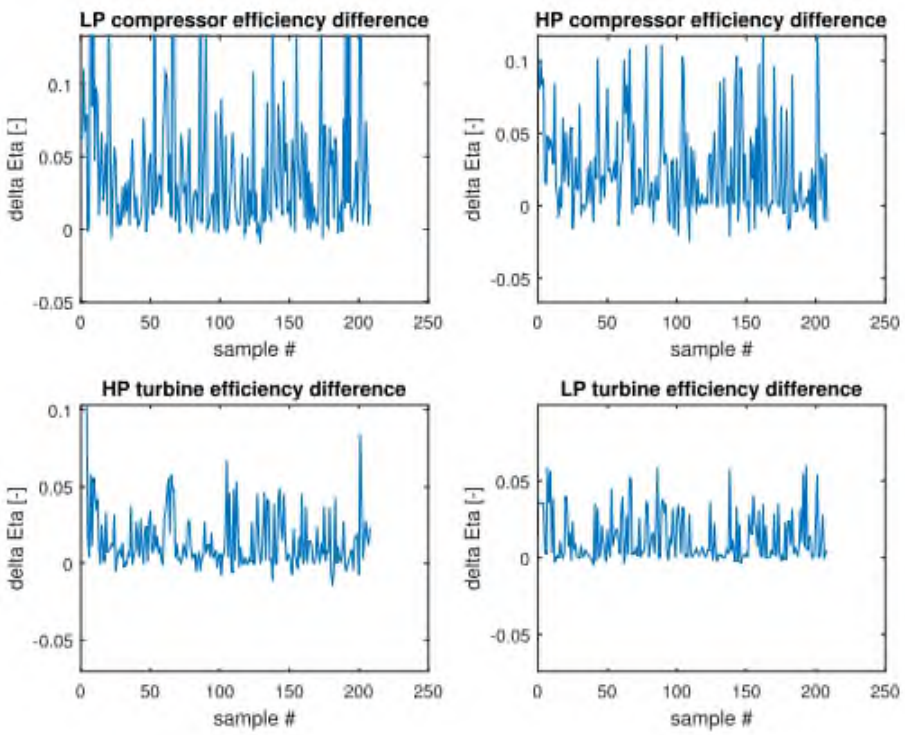

Figure 52 Deterioration imposed on the gas turbine components – 0.4% noise

5.1.2.1 No pre-filtering

The first case to be tested is the one without filtering. This case includes the structure with the ANN+FL but excludes the KF for the measurement noise isolation and for the multiple measurements processing. The level of noise for this test is the nominal of 0.4% as the aim is to isolate the robustness of the

methodology with random deterioration. The aspects to be tested are the components health estimation (quantification) while the magnitude of the failure varies randomly and the classification of single and multiple components failure together with different types of failure. The classification, in this section, becomes the most relevant to be tested as the types of failure are augmented to 6 (including LP compressor fouling, HP compressor fouling, HP turbine fouling, HP turbine erosion, LP turbine fouling and LP turbine erosion) and the number of combinations increases from 1 to 24.

The results of the quantification (Figure 53) show the range of variation of the quantification between 5 and 96. The lowest value, 5, is selected as a threshold to avoid a superposition of plot between the noisy results and meaningful results. The value 5 corresponds to the lower rounding of the 3σ value calculated during the dry run.

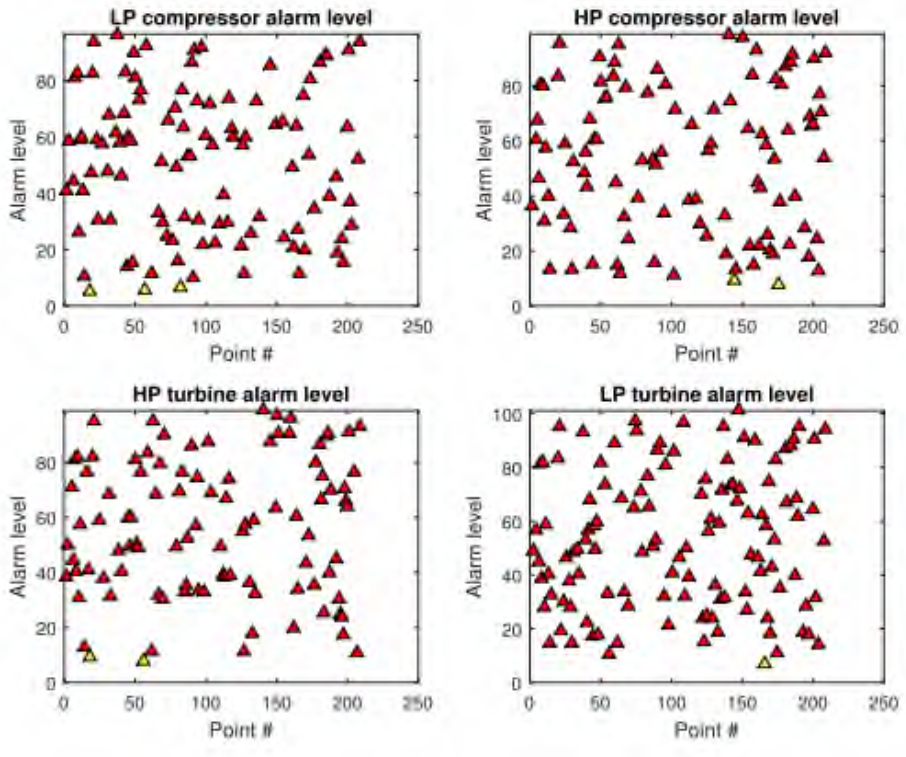

Figure 53 Quantification chart in the case of random deterioration – 0.4% noise no pre-filtering

Looking at the numbers obtained, the quantification success rate ranges from a minimum of 96.5% on the HP compressor to a maximum of 100% on the HP turbine (Table 17).

Component	Success rate
LP comp	99.0%
HP comp	96.5%
HP turb	100%
LP turb	98.0%

Table 17 Success rate for the failure quantification for random simulation – 0.4% noise no pre-filtering

The classification rate instead, ranges from 95.5% for the LP turbine to the 100% on the HP turbine fouling and erosion and on the LP turbine fouling (Table 18). It must be specified that the success rate includes single and multiple failures.

Failure type	Success rate
LP comp fouling	97.1%
HP comp fouling	96.8%
HP turb fouling	100%
HP turb erosion	100%
LP turb fouling	100%
LP turb erosion	95.5%

Table 18 Success rate for the failure classification for random simulation – 0.4% noise no pre-filtering

The same test has been repeated for a different level of noise that in this case has been selected to be 2.0%. This level of noise is chosen as it is the maximum value among those tested in section 5.1.1 and, therefore, represents the worst case for the random simulation. In this case, the spread of the malfunction quantification is increased from the expected range 5-96 to a range of 5-140 (Figure 54) and this gives a sign that the quantification is deviating from the expected rate.

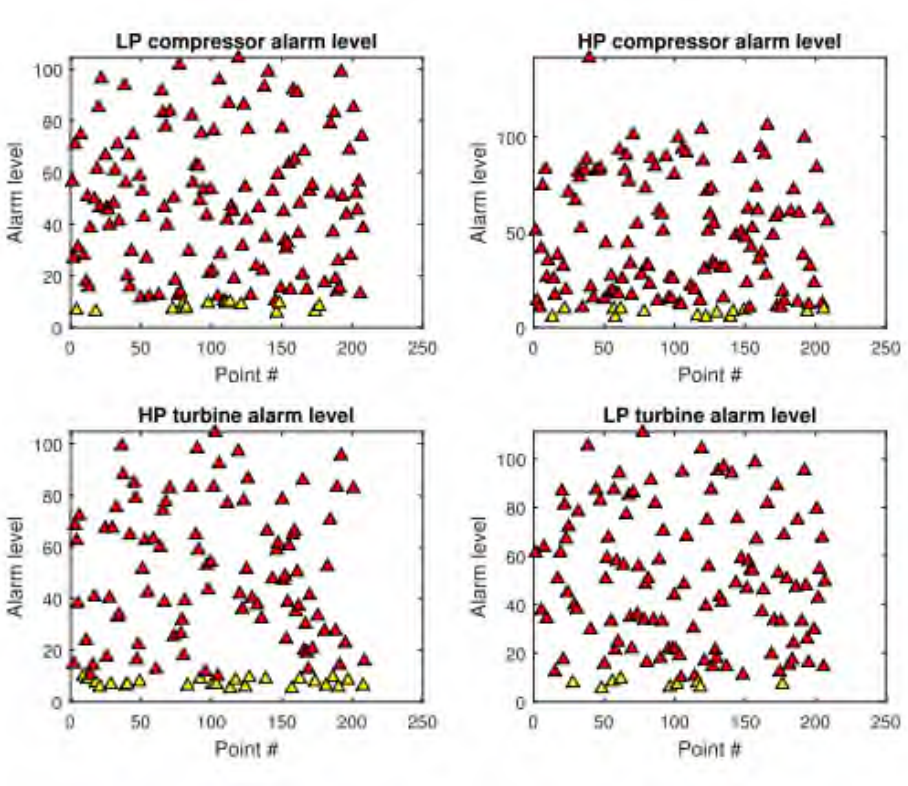

Figure 54 Quantification chart in the case of random deterioration – 2.0% noise no pre-filtering

The consequence of this deviation from the exact quantification is a decrease of the success rate that ranges, in this case, from 56.2% on the HP compressor to 74.1% on the LP turbine (Table 19).

Component	Success rate
LP comp	56.7%
HP comp	56.2%
HP turb	59.2%
LP turb	74.1%

Table 19 Success rate for the failure quantification for random simulation – 2.0% noise no pre-filtering

This reduction on the success rate is also reflected in an increasing number of deviation spots (Figure 55). The deviation from the reference value is in both directions since the noise is acting randomly around the mean parameter.

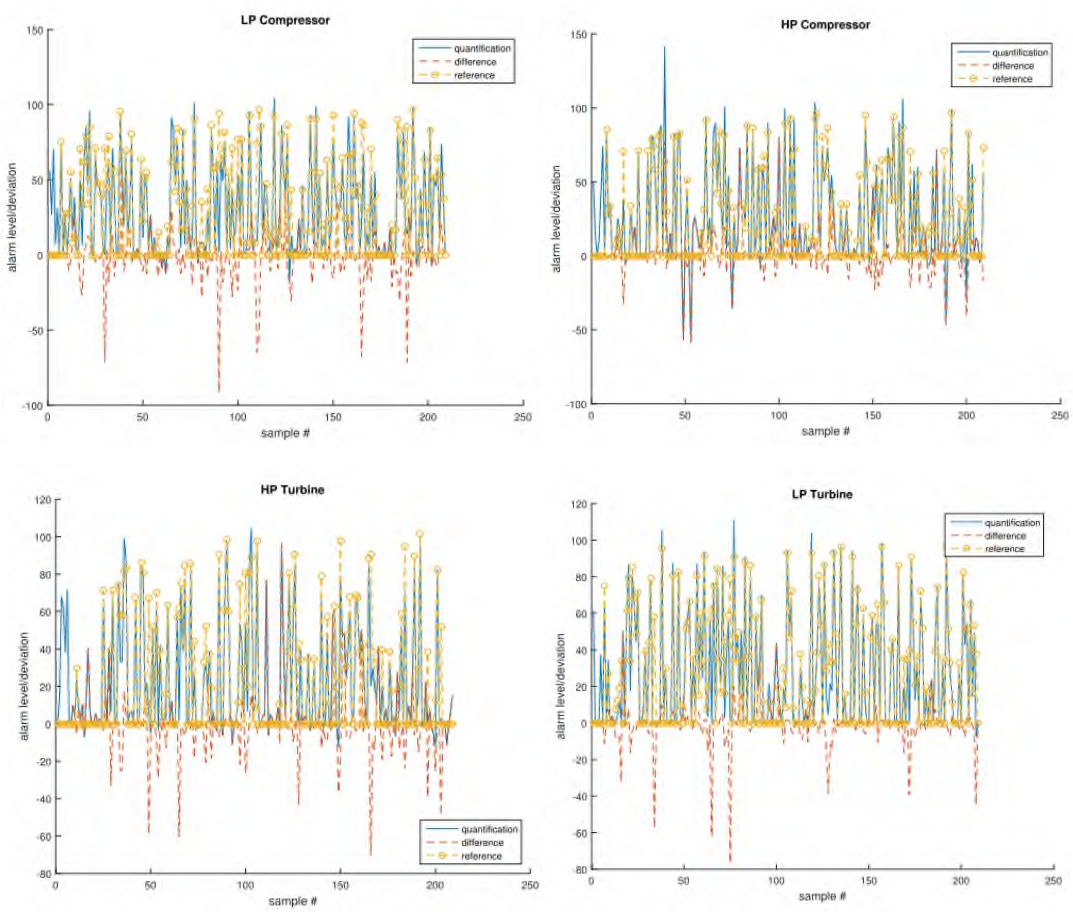

Figure 55 Absolute quantification and deviation from the reference for compressor and turbine components – 2.0% noise no pre-filtering

The classification rate instead, ranges from 86.5% for the HP turbine erosion to the 100% on the HP compressor and HP turbine fouling (Table 20). Compared to the quantification the impact on the success rate is lower, but the impact can reach a maximum of 13.5%. The lesser effect is justified by the higher threshold the classification KF has, as already mentioned in section 5.1.1.1.

Failure type	Success rate
LP comp fouling	99.1%
HP comp fouling	100%
HP turb fouling	100%
HP turb erosion	86.5%
LP turb fouling	100.0%
LP turb erosion	88.3%

Table 20 Success rate for the failure classification for random simulation – 2.0% noise no pre-filtering

The results obtained confirm the capability of the methodology of quantifying the failure magnitude and classifying the type of failure with a random simulation of single and multiple failures with varying magnitude and nominal noise of 0.4%. The results above the 90% target set are not achieved instead, if the level of noise increases to 2.0%, confirming the necessity of a section for the measurement filtering (KF).

5.1.2.2 Single Linear Kalman Filter and measurement fusion

During this test, the KF is included before the diagnostics phase. Provided the improvement of the data over the constant case reported in section 4.5 an increase of the quantification and classification rate is expected. The set-up built around the KF allows the processing of multiple measurements placed at the same location in two ways: with a single layer and with multiple layers. The preliminary check described in section 4.5 shown better performances for the MLKF rather than the SLKF and these results must be checked also for the random deterioration case. The portion analysed in this section is the SLKF that consist of one layer devoted to combine all the information coming from every single measurement and to filter out the measurement noise.

The results of the quantification repeat the same trend observed without pre-filtering ranging within 5 and 96 (Figure 56). It must be mentioned that as the points are selected randomly, the samples are not one to one comparable among the tests.

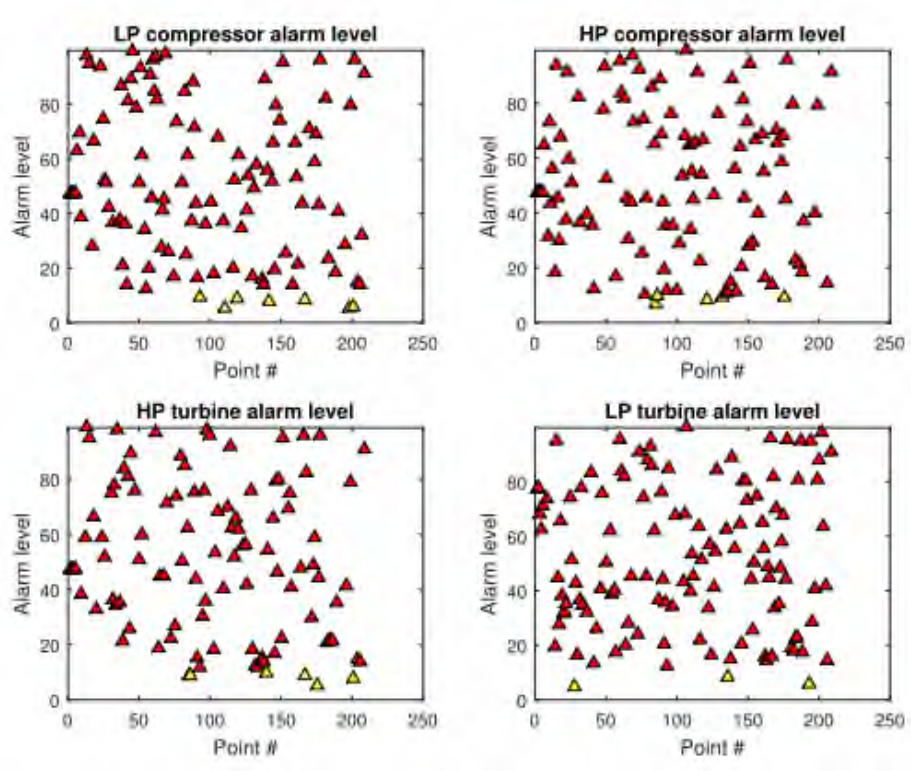

Figure 56 Quantification chart in the case of random deterioration – 0.4% noise SLKF

The quantification success rate ranges from a minimum of 97.5% on the HP compressor to a maximum of 99.0% on the LP compressor (Table 21). Compared to the test without filtering there's almost no difference. This is what it is expected since, with nominal noise, the reduction of measurement noise has been shown to be negligible section 4.5.

Component	Success rate
LP comp	99.0%
HP comp	97.5%
HP turb	98.5%
LP turb	98.5%

Table 21 Success rate for the failure quantification for random simulation – 0.4% noise SLKF

The classification rate instead, ranges from 94.3% for the LP compressor fouling to the 100% on HP turbine fouling and LP turbine fouling. The results are in line with the case without the filter (Table 22) with a small worsening on the LP compressor fouling and on the HP turbine erosion.

Failure type	Success rate
LP comp fouling	94.3%
HP comp fouling	96.7%
HP turb fouling	100%
HP turb erosion	96.6%
LP turb fouling	100%
LP turb erosion	95.3%

Table 22 success rate for the failure classification for random simulation – 0.4% noise SLKF

Moving to the case with 2.0% reference noise, the quantification ranges between 5 and 130 (Figure 57). Considering that the expected quantification is from 5 to 96, a consequence of the quantification success rate is expected.

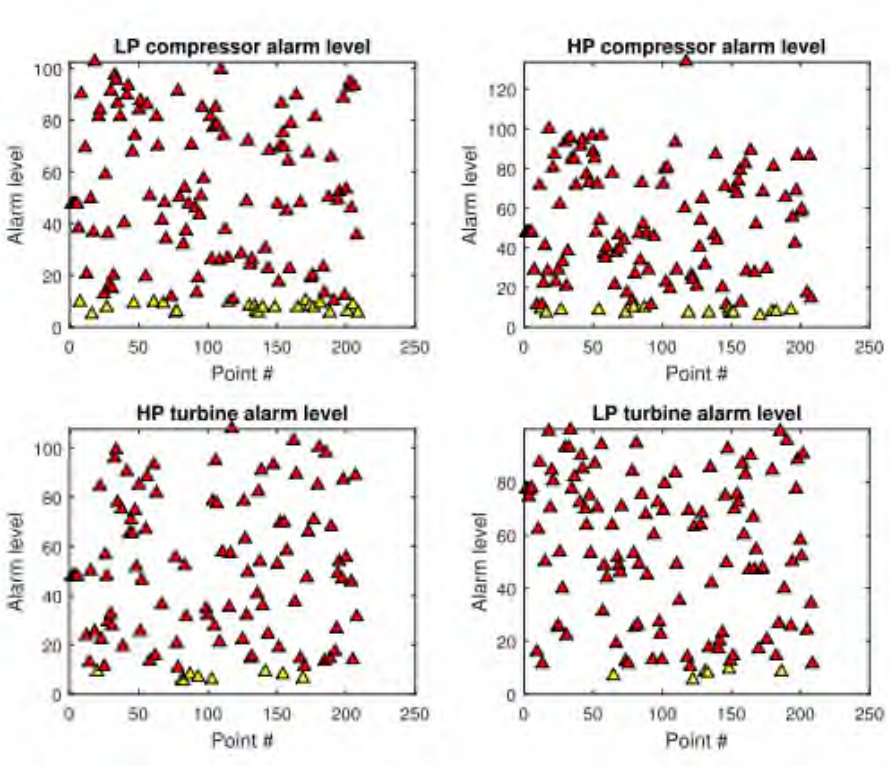

Figure 57 Quantification chart in the case of random deterioration – 2.0% noise SLKF

The resulting quantification success rate varies, from 69.0% on the HP compressor to 87.6% on the LP turbine with an improvement on all the components compared to the case without filtering (Table 23). These results confirm the improvement seen with the constant deterioration 5.1.1 while introducing the SLKF. However, as reported for the case with constant deterioration, the target is not reached.

Component	Success rate
LP comp	73.1%
HP comp	69.2%
HP turb	70.6%
LP turb	87.6%

Table 23 Success rate for the failure quantification for random simulation – 2.0% noise SLKF

This reduction on the success rate can be also observed in the graph comparing the reference with the predicted value, that is showing an increasing number of deviation spots (Figure 58).

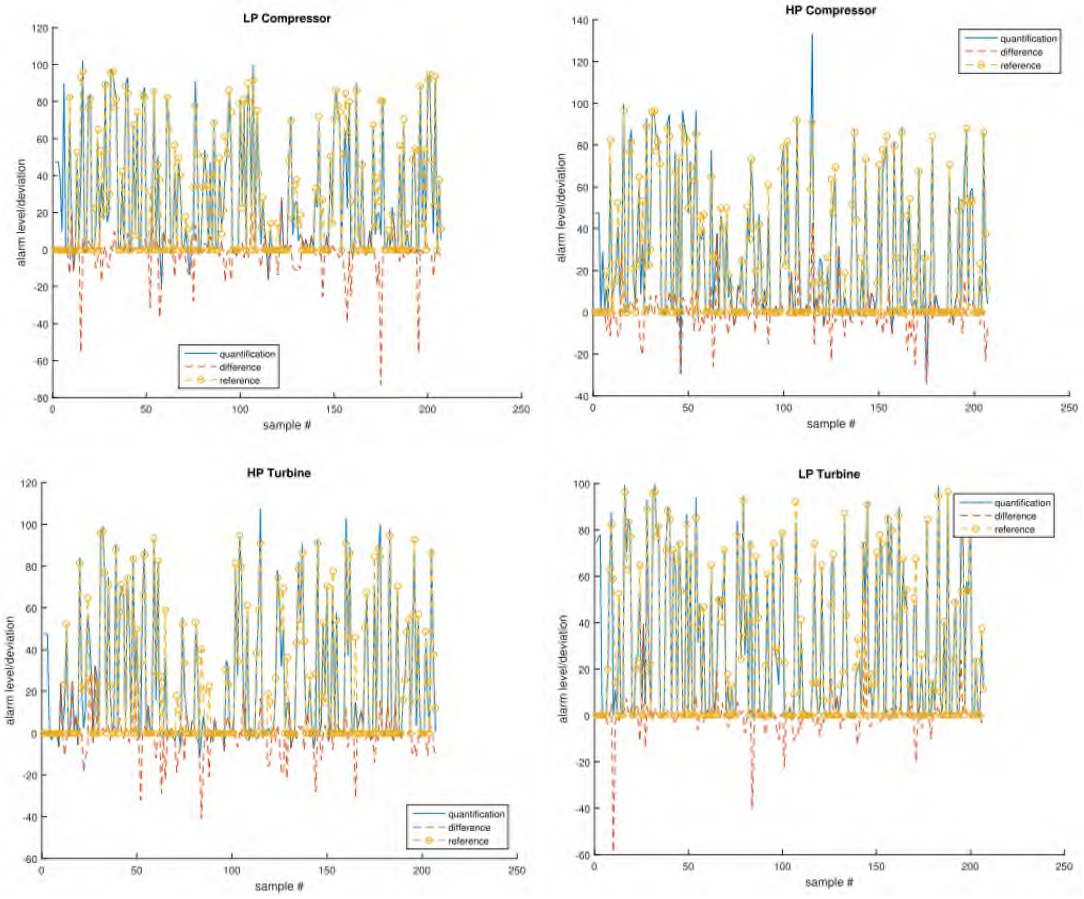

Figure 58 Absolute quantification and deviation from the reference for compressor and turbine components – 2.0% noise SLKF

The classification rate instead, ranges from 89.5% for the LP turbine erosion to the 100% on the HP compressor fouling, HP turbine fouling and LP turbine fouling (Table 24). The weakest results are on the turbine erosion that

is relying on the mass flow reference deterioration only, that while challenged with this high noise lead to errors on some points. The other results are above the target confirming the higher robustness of the classification module.

The results shown in this section, say that the methodology can perform multiple component health estimation and detect among them the specific type of failure. The success rate though is above the targeted 90% only for the case with nominal noise while with a 2.0% noise the target is reached on the classification side only (provided 0.5% gap on the LP turbine erosion). Some work has to be done on the component health estimation that does not reach the target at 2.0% measurement noise.

Failure type	Success rate
LP comp fouling	97.1%
HP comp fouling	100%
HP turb fouling	100%
HP turb erosion	92.5%
LP turb fouling	100%
LP turb erosion	89.5%

Table 24 Success rate for the failure classification for random simulation – 2.0% noise SLKF

5.1.2.3 Multiple Layer Kalman Filter and measurement fusion

The other scheme analysed in this section is the MLKF that consist of two layers devoted to combine all the information coming from every single measurement and to filter out the measurement noise. The aim of the second layer is to further make use of the filtered information to improve the final result.

The first results of the quantification performed with the nominal noise of 0.4% repeat the same trend observed without pre-filtering ranging within 5 and 96 (Figure 59). As for the case with the SLKF, no improvements are expected since the filtering effect on the measurement noise is negligible at this stage.

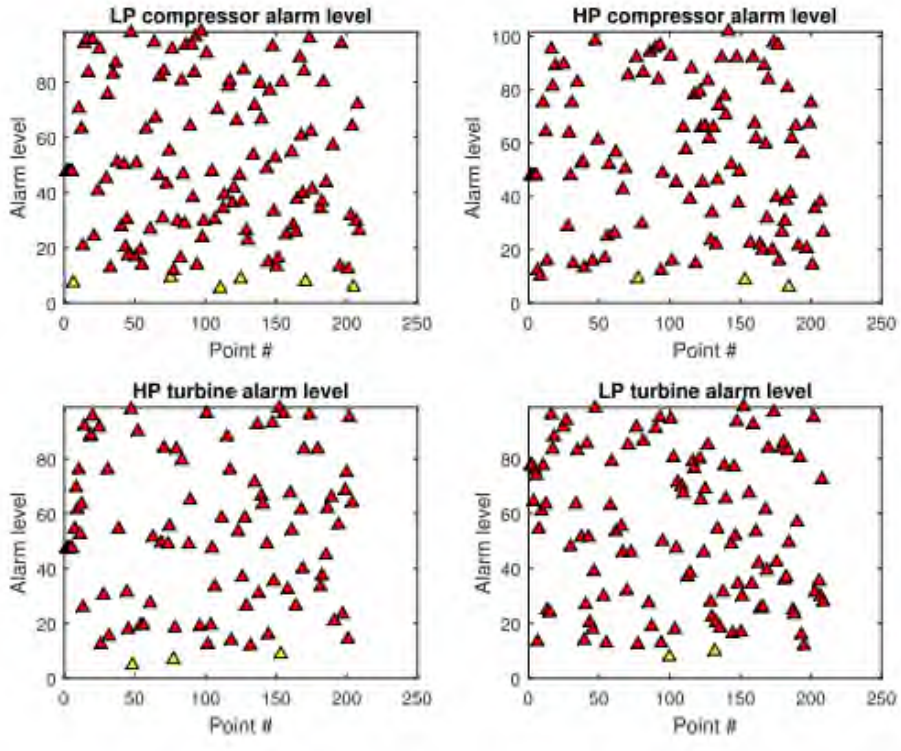

Figure 59 Quantification chart in the case of random deterioration – 0.4% noise MLKF

The quantification success rate instead, ranges from a minimum of 92.0% on the HP compressor to a maximum of 99.5% on the LP turbine (Table 25). Again, compared to the test without filtering there's almost no difference except for a small decay on the HP compressor rate.

Component	Success rate
LP comp	97.0%
HP comp	92.0%
HP turb	98.5%
LP turb	99.0%

Table 25 Success rate for the failure quantification for random simulation – 0.4% noise MLKF

The classification rate instead, ranges from 95.1% for the HP turbine erosion to the 100% on the HP turbine fouling and LP turbine fouling (Table 26). The results are in line with the cases without KF and with the SLKF.

Failure type	Success rate
LP comp fouling	98.3%
HP comp fouling	97.9%
HP turb fouling	100%
HP turb erosion	95.1%
LP turb fouling	100%
LP turb erosion	96.9%

Table 26 Success rate for the failure classification for random simulation – 0.4% noise MLKF

Looking at the case with 2.0% noise level instead, the quantification ranges between 5 and 105 (Figure 60), which is already a sign of improvement, also seen with the SLKF.

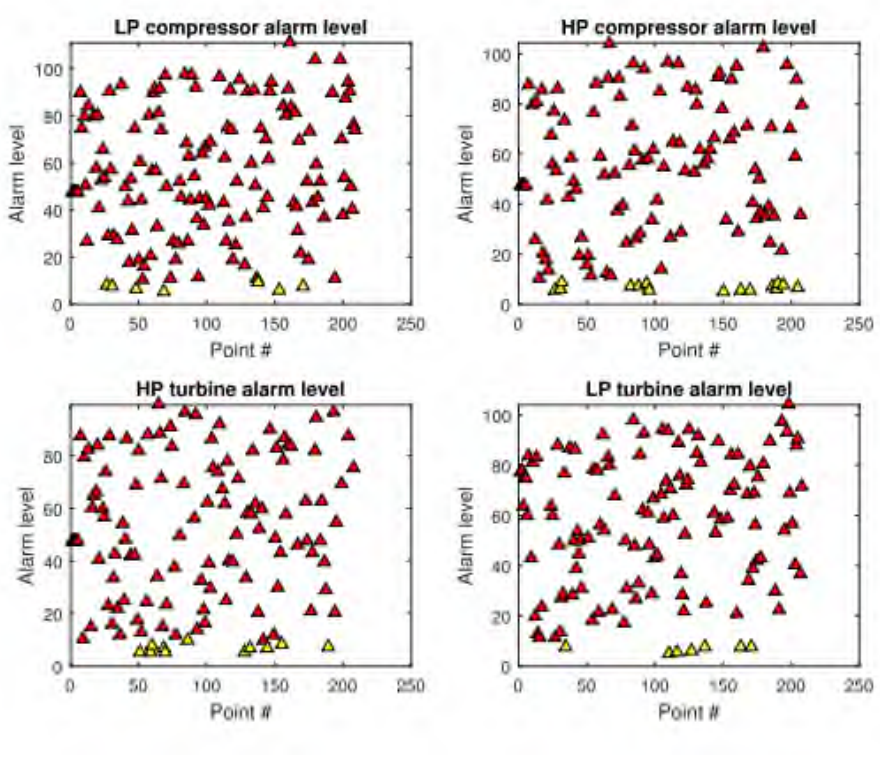

Figure 60 Quantification chart in the case of random deterioration – 2.0% noise MLKF

The resulting success rate varies, from 70.1% on the HP turbine to 83.3% on the LP turbine with an improvement on all the components compared to the case without filtering (Table 27).

Component	Success rate
LP comp	76.1%
HP comp	73.6%
HP turb	70.1%
LP turb	83.1%

Table 27 Success rate for the failure quantification for random simulation – 2.0% noise MLKF

The improvement in the success rate for the failure quantification is also reflected in the decrease of deviations spots (Figure 61).

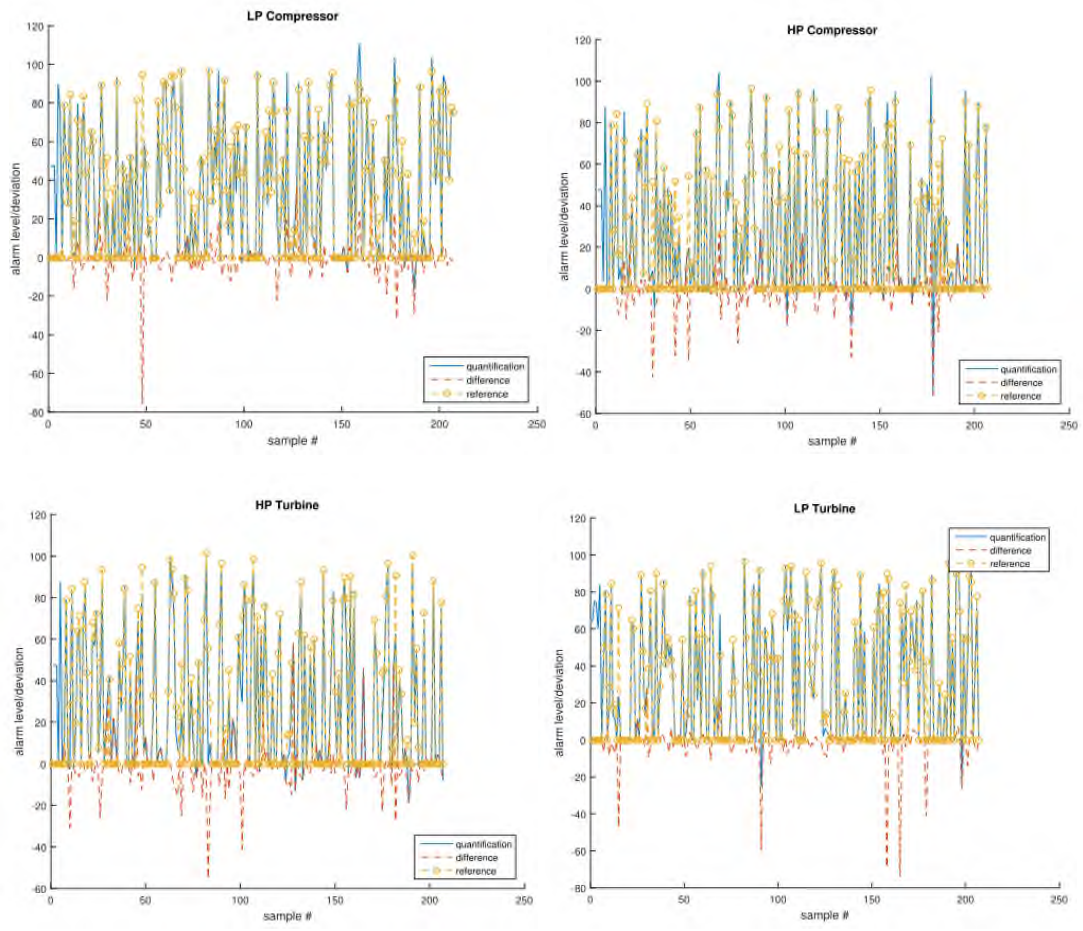

Figure 61 Absolute quantification and deviation from the reference for compressor and turbine components – 2.0% noise MLKF

The classification rate ranges from 88.9% for the LP turbine erosion to 100% on the HP compressor fouling and LP turbine fouling (Table 28). The improvement seen here is in line with the case with SLKF applied even if the classification success rate is already better placed also without filtering.

Failure type	Success rate
LP comp fouling	97.7%
HP comp fouling	100%
HP turb fouling	96.0%
HP turb erosion	90.7%
LP turb fouling	100%
LP turb erosion	88.9%

Table 28 Success rate for the failure classification for random simulation – 2.0% noise MLKF

The results shown in this section, confirm that the methodology can perform multiple component health estimation (quantification) and can detect among them the specific type of failure (classification). This achievement is in line with ultimate goal foreseen by Volponi [3] in his past, present and future analysis which is the automated detection of the gas turbine components health. It must be remarked that the prefixed target can be achieved only for the case with 0.4% noise, while is missed for the case with 2.0% noise. In this case, even if the improvement is clearly visible, some additional work has to be done. Nevertheless, considered the challenge of this test, the results are still remarkable.

5.1.3 Deterioration schedule

After having checked the methodology against a constant deterioration with measurement noise and against a random deterioration with different levels of noise, the methodology is tested against a deterioration schedule. This schedule has been already described in section 4.2 and out of the initial profile with 5000 points, 203 samples reflecting the deterioration between 0% and 75% of maximum magnitude have been extracted. For instance, the final reference deterioration on the LP compressor is 3.3% while is 5.1% on the LP turbine. Since the deterioration is meant to represent a possible real degradation over months of the gas turbine, the aim here is to see how the methodology is reacting in such conditions. The entire simulation is carried over with the reference noise of 0.4%

The numerical data provided by the deterioration schedule mimic the gas turbine behaviour, therefore, the extracted points represent the

measurements coming from it (Figure 21). This means that only pressure, temperature and power values are available.

5.1.3.1 Combined training

The combined training is made up by the random simulation and by the deterioration schedule. The random simulation consists of the 24 combinations described in section 4.2 and reported in Table 2. The deterioration schedule consists of the 5000 points reflecting a period of time of the gas turbine behaviour.

The necessity of the data combination comes from the unsatisfactory results obtained with the use of the 24 combinations only (Figure 62). This combination in-fact is made up of 100 points for each deterioration type and cannot be sufficient to predict points among the 5000 foresee in the deterioration schedule. On the other hand, the increase of the number of points among the 24 combinations is leading to memory issues, so the final solution is the merge of the 24 combinations, together with the deterioration schedule in the ANN training.

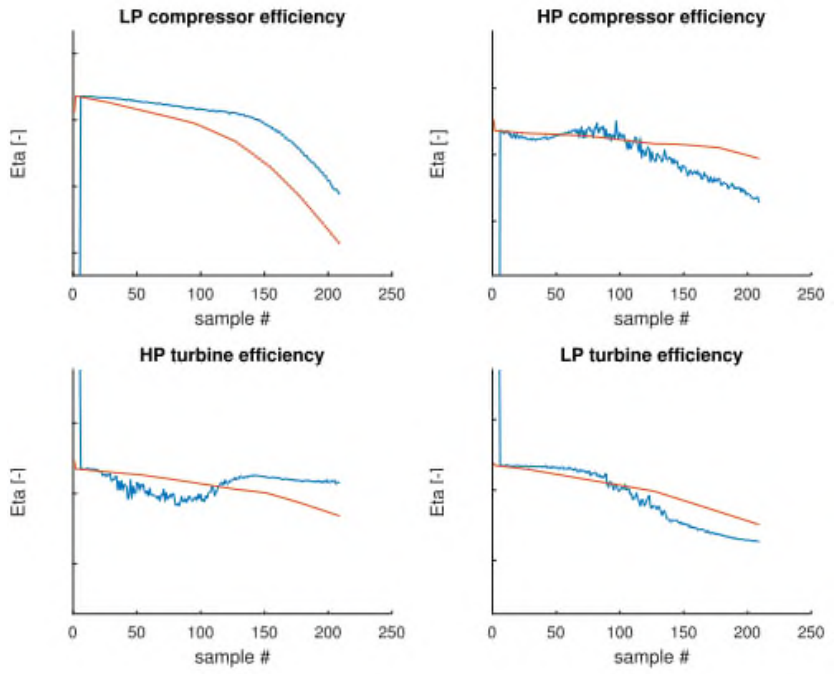

Figure 62 Reference (red line) vs prediction (blue line) of components efficiency

The result of this merge is an increased number of points available to the ANN. The consequence is an increased number of data available that could improve the ANN quality prediction but could also cause overfitting on some points that could reduce the quality of the prediction.

As a single type of multiple failures has been implemented (fouling on all the components), it is possible to observe its effect on the efficiency of the components.

The impact on the efficiency is very clear and it is a decrease in all the components (Figure 63). It is worth to mention that the most affected component is the LP compressor which is suffering from the deterioration of the compressor itself and from the deterioration of the upcoming components. On the other hand, it is also worth to notice that the HP compressor efficiency is remaining almost constant and decreases a bit more only in the last part when the deterioration is more pronounced. It has also to be remarked that the deterioration profile imposed (Figure 20) is not exactly reflected onto the efficiency profiles due to the mutual interaction of the components. From that, it is clear that the delta efficiency from the reference values cannot be used to represent the health status of each component. For instance, judging the health status from the HP compressor efficiency delta would lead to a much less impacted component. Instead, judging it from the LP compressor would lead to a much higher impact. In reality, the deterioration profile is identical at the beginning and diverges only at the end by 0.5% reference deterioration.

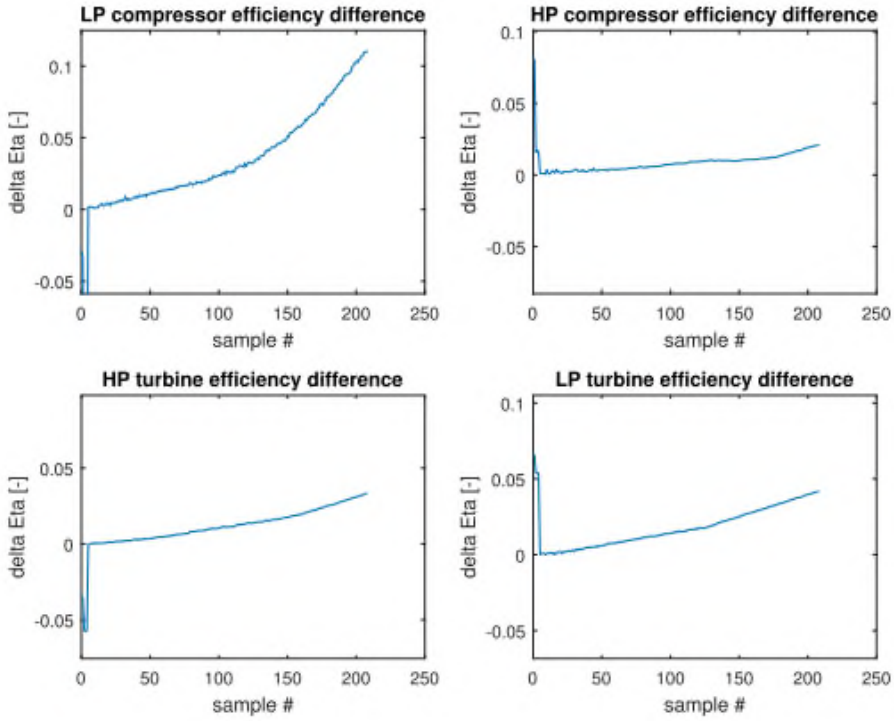

Figure 63 Deterioration imposed on the gas turbine components starting from the deterioration schedule

Once the efficiencies and physical values deviations are processed by the fuzzy logic, they deal with the quantification described in section 4.7. Based on the deterioration profile imposed in Figure 20, the expectation is a growing quantification magnitude trend starting from the LP compressor moving to the LP turbine. Based on the deterioration selected for this problem - 3.3% on the LP compressor and 5.1% on the LP turbine - the prediction is to have an alarm level of 43 on the LP compressor and 66 on the LP turbine (Figure 64).

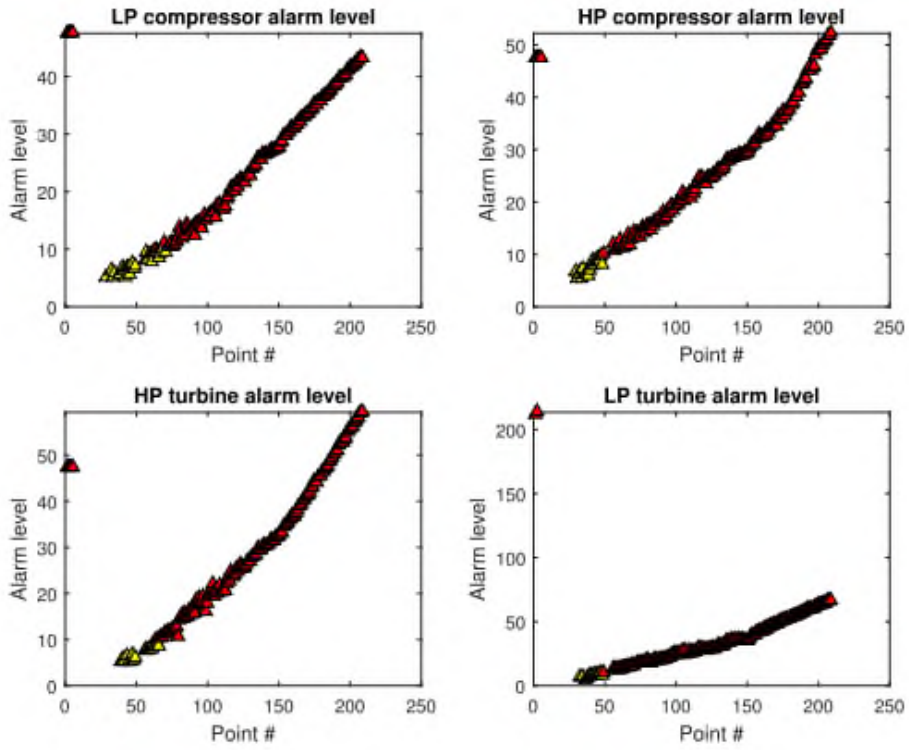

Figure 64 Quantification chart in the case of scheduled deterioration

The quantification is exactly matched as reflected in the resulting quantification success rate, whose result is 100% for all the components (Table 29).

Component	Success rate
LP comp	100%
HP comp	100%
HP turb	100%
LP turb	100%

Table 29 Success rate for the failure quantification for scheduled simulation – Noise is set to 0.4%

These perfect results are also reflected in the small deviation between the component health estimation (quantification) and its reference (Figure 65). It must be remarked that a residual deviation is present in the HP compressor especially between sample 50 and sample 120. There, an additional turning of the NFL could further improve the results.

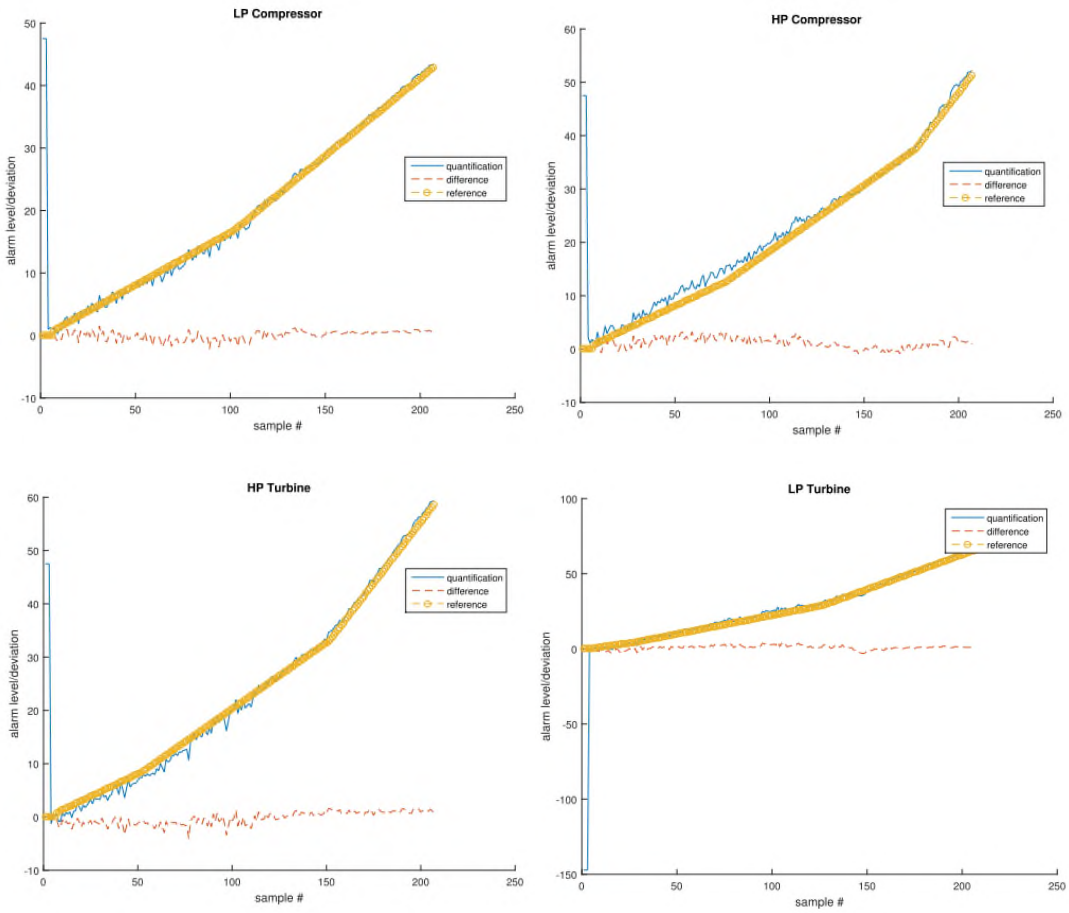

Figure 65 Absolute quantification and deviation from the reference for compressor and turbine components – 0.4% noise MLKF

The next step to be tested is the classification which is easier in this case as the type of failure remains the same for the entire simulation – compressor and turbine fouling on all the components (Figure 66 and Figure 67). The resulting success rate ranges from 89.1% on the HP turbine to 97.0% on the HP compressor (Table 30). Despite the easiness of the problem, the target is not fully met therefore some additional tuning could be done in the FL to better match these points.

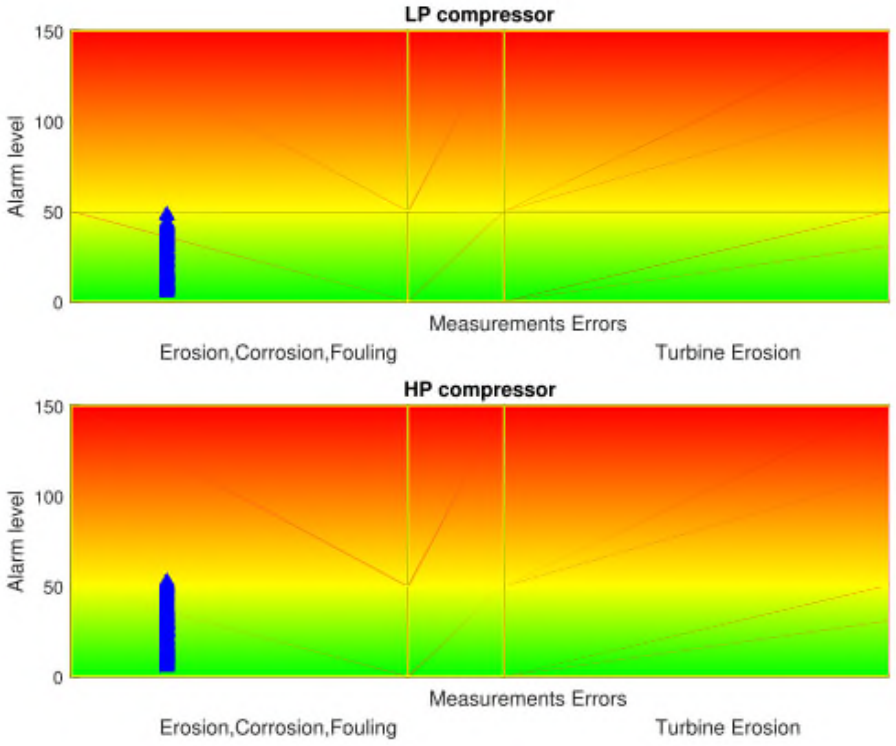

Figure 66 Classification chart for the compressor

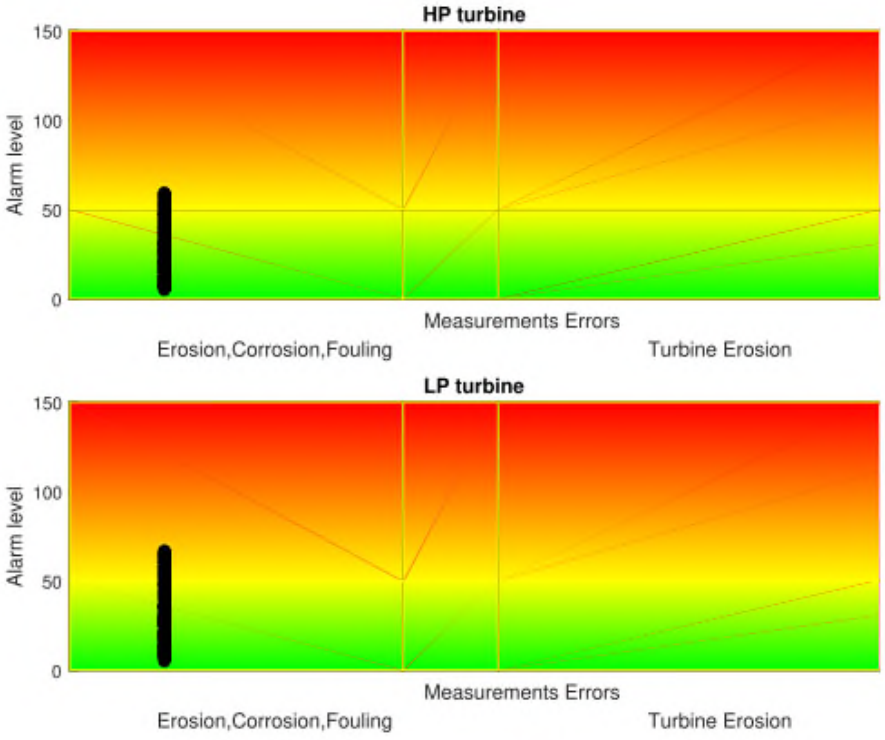

Figure 67 Classification chart for the turbine

Failure type	Success rate
LP comp fouling	93.6%
HP comp fouling	97.0%
HP turb fouling	89.1%
LP turb fouling	93.6%

Table 30 Success rate for the failure classification for scheduled simulation – 0.4% noise MLKF

These results confirm that the data combination is a necessary step in order to let the ANN make the right prediction of the efficiency and to let the rest of the methodology make an accurate prediction. Moreover, they confirm one of the weaknesses of the ANN outlined in section 4.6, which is the high uncertainty of the ANN while predicting outside the trained path. Additionally, the results confirm the feasibility of the failure characterization also with a normal deterioration run of several months, even if some additional turning is required on the HP compressor fouling.

As reported at the beginning of the chapter, this simulation makes use of the full deterioration profile as input during the ANN training phase and of a sample of it during the prediction phase. Therefore, it can be said that the prediction is done among data used in the ANN training phase. However, it is not always realistic to know beforehand the deterioration profile of the gas turbine. In this case, the real data of the deterioration schedule should be manually processed and integrated offline into the ANN training. The fact that the ANN can predict only within points provided during the training is a limitation already stated by Bechini [1] in its analysis.

5.2 Scheme 2

The difference introduced by the second scheme is on the reference value of the non-deteriorated engine that is not provided by the thermodynamic model made with Turbomatch but rather by the ANN. This way the ambient temperature reference is no more necessary, so the Turbomatch calculation. The advantage is an even faster calculation, and the use of internal measurements only. The disadvantage is a possible reduction of accuracy in the reference model, since it is coming from the ANN and not from the official thermodynamic model. The block subjected to this change is the ANN that is trained in the first part including the deteriorated engine and the reference engine. The new scheme is affecting the delta calculated as the difference between the reference value and the deteriorated value:

$$\Delta x = x_{ref} - x_{deteriorated}$$

Where x is the physical parameter under investigation. The parameter changed with this scheme is x_{ref} which is not coming from the calculated Turbomatch reference, but is predicted, together with the $x_{deteriorated}$, by the ANN.

5.2.1 Constant deterioration

5.2.1.1 Measurements noise

The constant deterioration has identical characteristics as for scheme 1 – section 5.1. The scope of this test is to see how this scheme is reacting to the noise and if the reference provided by the ANN, instead of Turbomatch, is causing any uncertainty.

The first parameter to check with this scheme is the quantification, that should end at the level of 96 for each gas turbine component. The result

obtained is very close to that seen with scheme 1 therefore in line with the prediction (Figure 68).

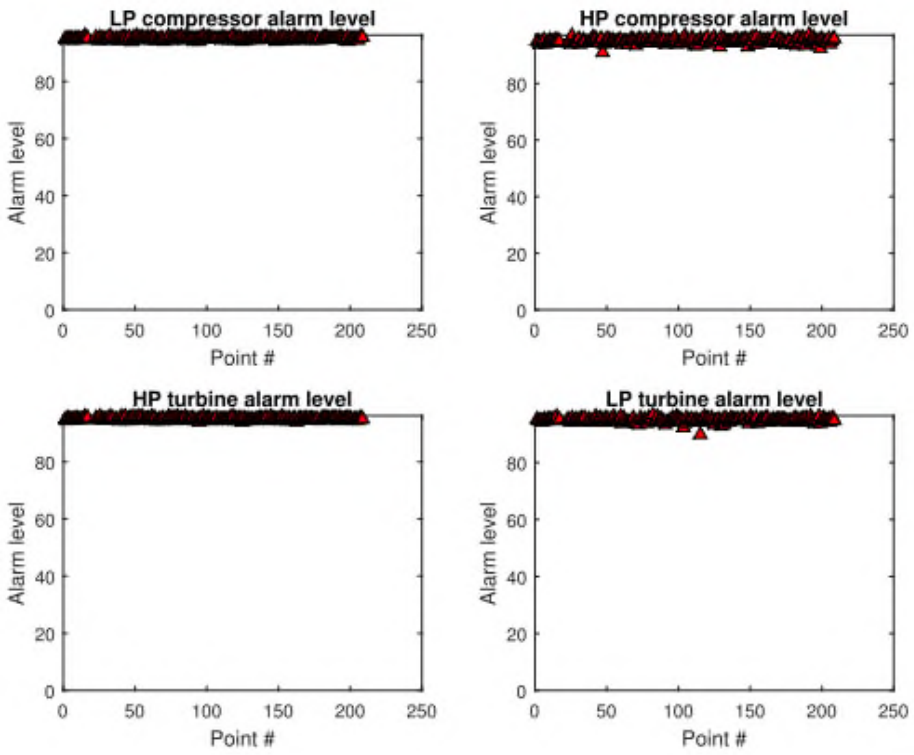

Figure 68 Quantification of the gas turbine failure – 0.4% noise no pre-filtering

The cumulative graph, proposing the quantification and the classification, shows that the types of failure are correctly classified as compressor fouling and turbine erosion (Figure 69 and Figure 70).

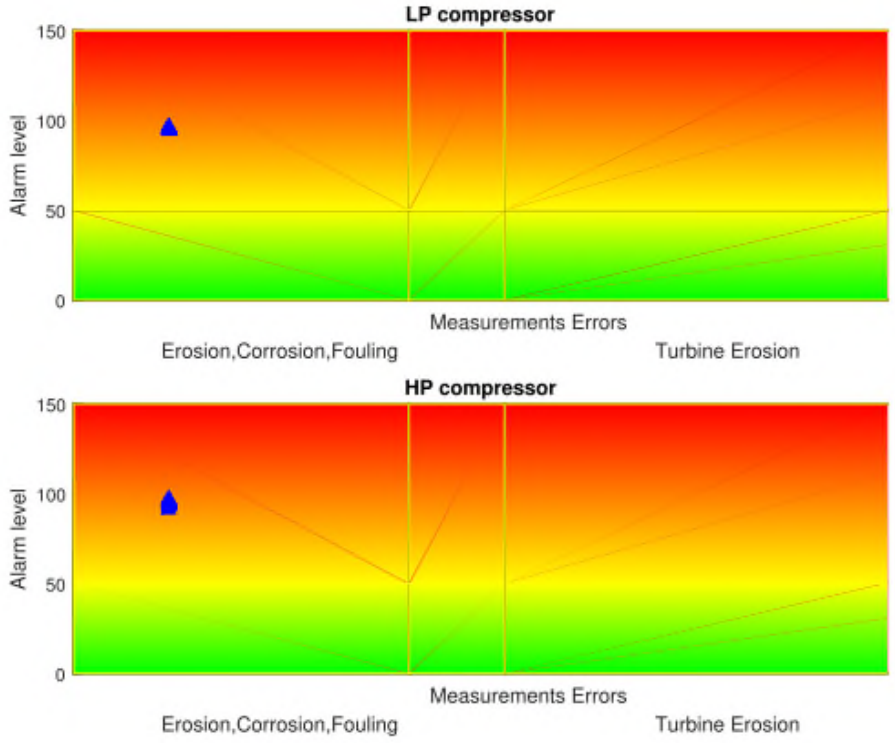

Figure 69 Quantification and classification of the compressor – 0.4% noise no pre-filtering

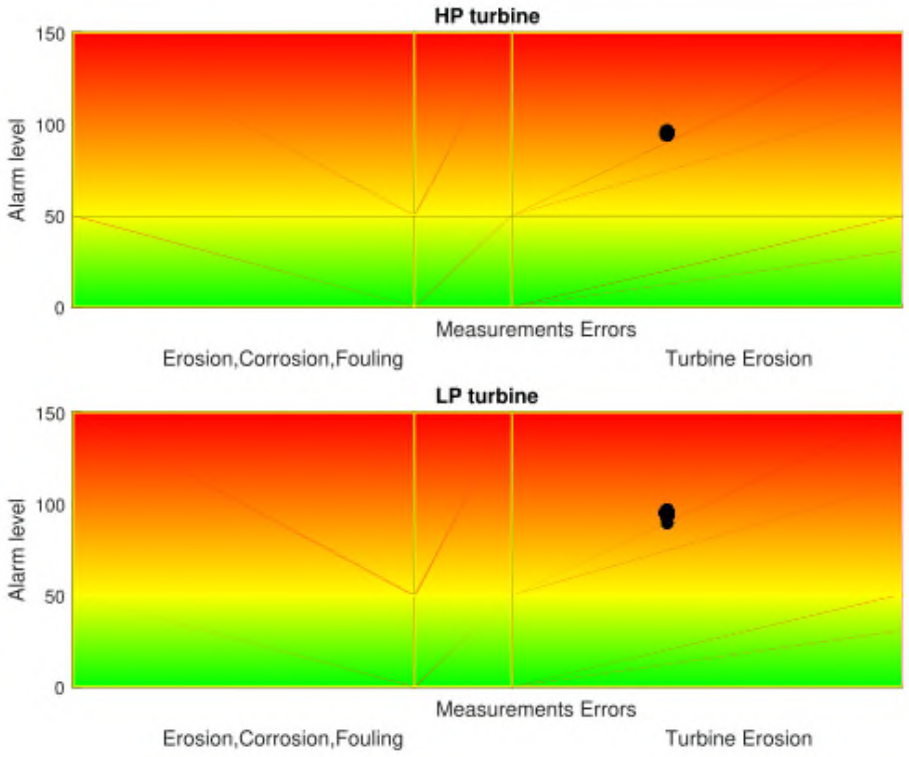

Figure 70 Quantification and classification of the turbine – 0.4% noise no pre-filtering

With the level of measurement noise set to 2.0%, the quantification is affected by spreading between 20 and 140 on the compressor and between 10 and 130 on the turbine (Figure 71 and Figure 72). The classification instead looks correct on the compressor but shows some false alarms on the turbine. This behaviour is similar to what seen with scheme 1 and the turbine classification looks having less false alarms.

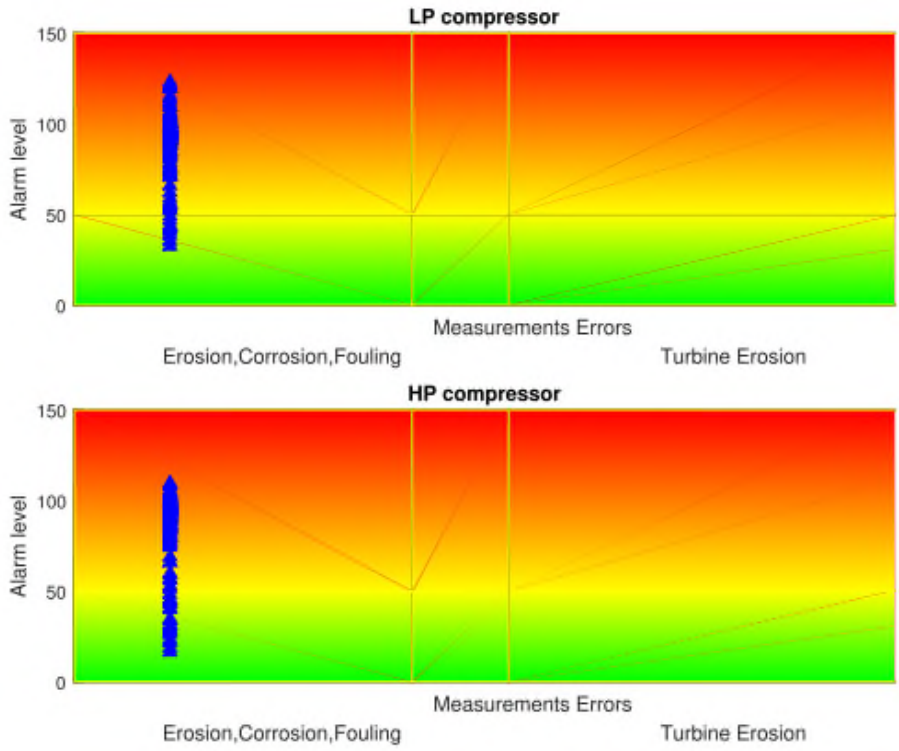

Figure 71 Quantification and classification of the compressor – 2.0% noise no pre-filtering

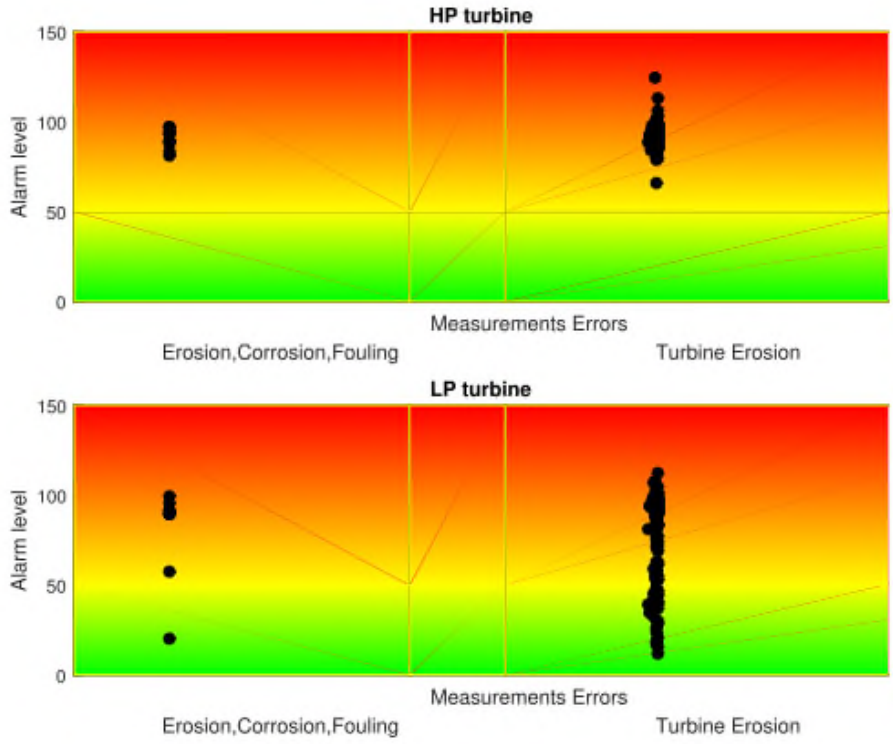

Figure 72 Quantification and classification of the turbine – 2.0% noise no pre-filtering

The next item to be compared is the scheme 2 with the data filtering module that, as per section 5.1.1.1, are the SLKF and the MLKF. The most interesting condition to be compared is with 2.0% measurement noise. In that case, the quantification ranges between 20 and 110 on the compressor and between 30 and 110 on the turbine. The reason behind it, as described for scheme 1, is the increase of the noise, that is directly reflected in the component health estimation. This means that the quantification is not as much overestimated as in the case without the KF, but there are still some samples not correctly quantified (Figure 73 and Figure 74). On the classification side instead, some false alarms remain.

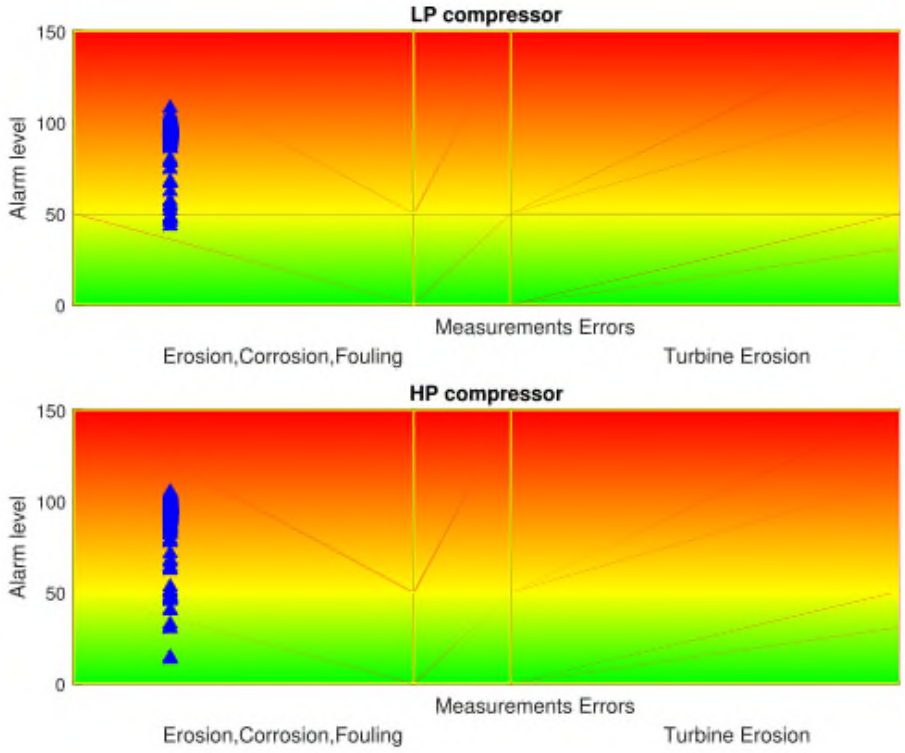

Figure 73 Quantification and classification of the compressor – 2.0% noise MLKF

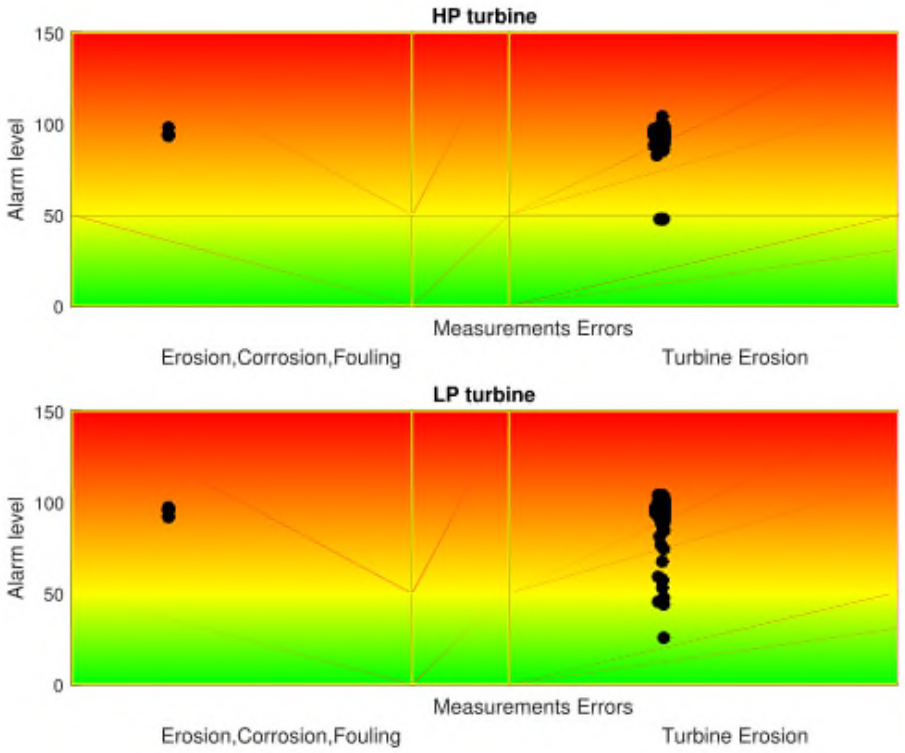

Figure 74 Quantification and classification of the turbine – 2.0% noise MLKF

The overall results of the first test show a quantification success rate of 100% for 0.0% and 0.4% noise in all the component and with all combinations. This is in line with what obtained for the scheme 1 – 99.5% - and confirms the capability of this methodology on classifying multiple failures while providing the correct health estimation also without the use of the Turbomatch reference.

While the level of noise is increasing, the trend tends to be similar to what seen for scheme 1. At 0.8% reference noise in-fact, the quantification success level decreases to 96.0% if no KF is applied. With the maximum noise level of 2.0%, the success rate is decreased to 51.7% without the KF block. The success rate though increases to 72.6% if also the SLKF is employed and to 70.6% if the MLKF is used (Table 31).

	Noise	0.0%	0.4%	0.8%	1.2%	1.6%	2.0%
None	LP comp	100%	100%	100%	93.5%	76.6	59.7
	HP comp	100%	100%	99.0%	89.6%	67.7%	51.7%
	HP turb	100%	100%	100%	96.5%	91.5%	79.6%
	LP turb	100%	100%	96.0%	84.6%	64.2%	54.2%
SLKF	LP comp	100%	100%	100%	96.5%	84.6%	82.6%
	HP comp	100%	100%	100%	94.0%	78.1%	72.6%
	HP turb	100%	100%	100%	100%	99.0%	97.0%
	LP turb	100%	100%	100%	97.0%	92.0%	78.1%
MLKF	LP comp	100%	100%	100%	98.5%	89.6%	78.1%
	HP comp	100%	100%	99.5%	95.0%	84.6%	70.6%
	HP turb	100%	100%	100%	100%	99.5%	97.5%
	LP turb	100%	100%	100%	98.0%	89.6%	80.1%

Table 31 Success rate for the failure quantification

Looking at the same values reported in the graphs instead it can be seen, once again, that the quantification success rate decreases starting from 0.8% (Figure 75). At 1.2% reference noise level, the success rate without KF is already below 90% and decreases to 51.7% with 2.0% reference noise level. The SLKF and MLKF improve this trend and the success rate remains above 90% until 1.2% reference noise level. The success rate starts then to decrease but remains above 70.6% at 2.0% reference noise level (Figure 76 and Figure 77). Compared to scheme 1 the result is slightly worse since the 90% success

rate was granted up to 1.6% noise. This result remarks that while flexibility is possible, the results might not be as accurate.

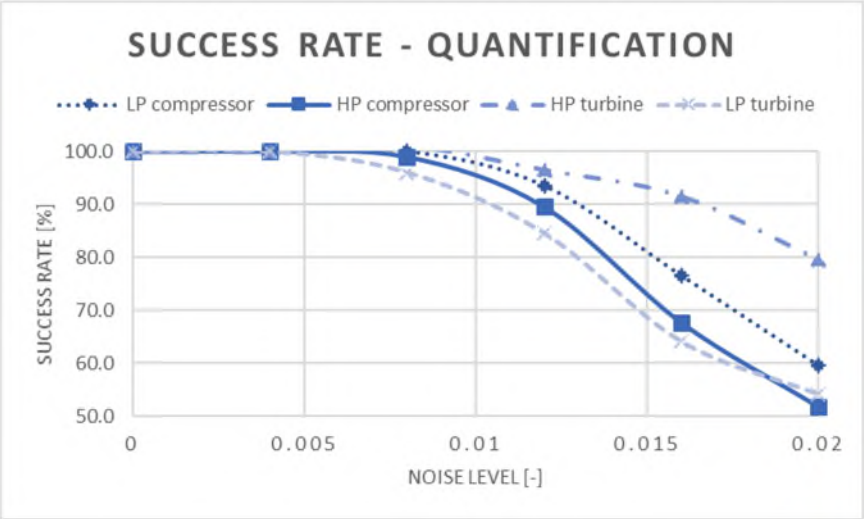

Figure 75 Success rate for the failure quantification – Combination No KF

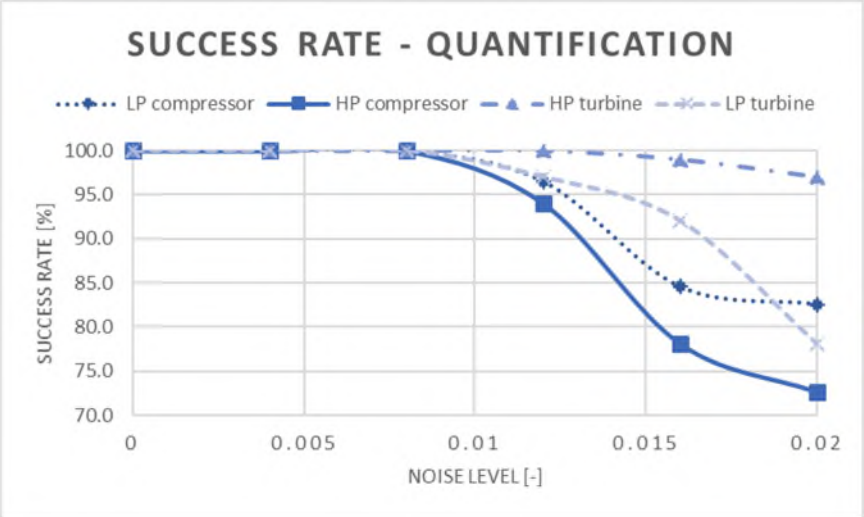

Figure 76 Success rate for the failure quantification – Combination SLKF+ANN+NFL+FL

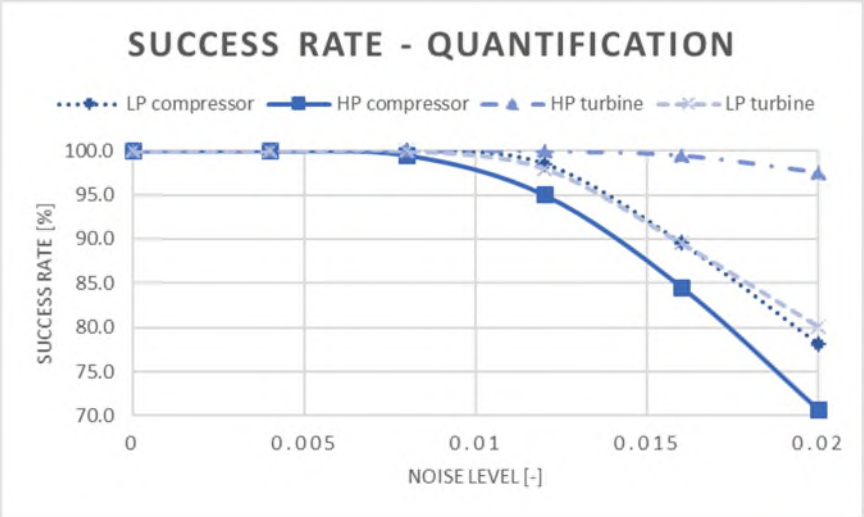

Figure 77 Success rate for the failure quantification – Combination MLKF+ANN+NFL+FL

The results of the classification with 0.4% noise show a success rate above 93.6% without filtering and at 100% with MLKF. With the maximum noise of 2.0%, the success rate decreases to 91.1% if no filtering is used and increases again to 94.0% with MLKF. The results of the SLKF are in line with the MLKF, as the minimum success rate is 94.0% for both (Table 32).

	Noise	0.0%	0.4%	0.8%	1.2%	1.6%	2.0%
None	LP comp	100%	100%	100%	100%	100%	100%
	HP comp	100%	93.6	93.1%	94.6%	97.5%	97.0%
	HP turb	100%	100%	100%	99.5%	97.5%	91.1%
	LP turb	100%	100%	100%	99.5%	96.6%	91.1%
SLKF	LP comp	100%	100%	100%	100%	100%	100%
	HP comp	100%	100%	96.5%	92.0%	95.5%	94.0%
	HP turb	100%	100%	100%	100%	99.5%	97.5%
	LP turb	100%	100%	100%	100%	98.5%	97.5%
MLKF	LP comp	100%	100%	100%	100%	100%	100%
	HP comp	100%	100%	96.0%	93.5%	88.6%	94.0%
	HP turb	100%	100%	100%	100%	99.0%	98.0%
	LP turb	100%	100%	100%	100%	98.5%	97.0%

Table 32 Success rate for the failure classification

Like the scheme 1, the classification is less affected by the noise compared to the quantification (Figure 78). The minimum rate in-fact is 91.1% at 2.0% measurement noise, compared to the 51.7% for the quantification. The decay on the HP compressor success rate for the SLKF instead, is also seen here, where the minimum classification rate is at 1.6% measurement noise with a

value of 92.0% (Figure 79). A similar effect is also seen on the MLKF scheme, but the minimum success rate is 88.6% at 1.6% measurement noise (Figure 80). The root cause of this decay, as explained for scheme 1, is an increase of the uncertainty on the HP compressor efficiency.

The overall results obtained confirm that is possible to correctly quantify and classify a GT multi failure also with the presence of measurement noise, by using the ANN both for the reference and for the deteriorated values prediction. However, the methodology can deal with noise up to 1.2% while keeping the success rate above 90% both for quantification and classification.

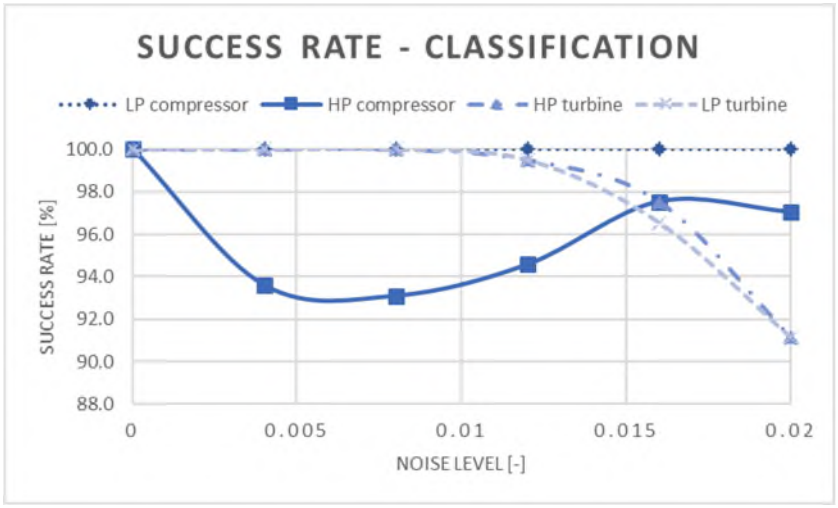

Figure 78 Success rate for the failure classification – Combination No KF

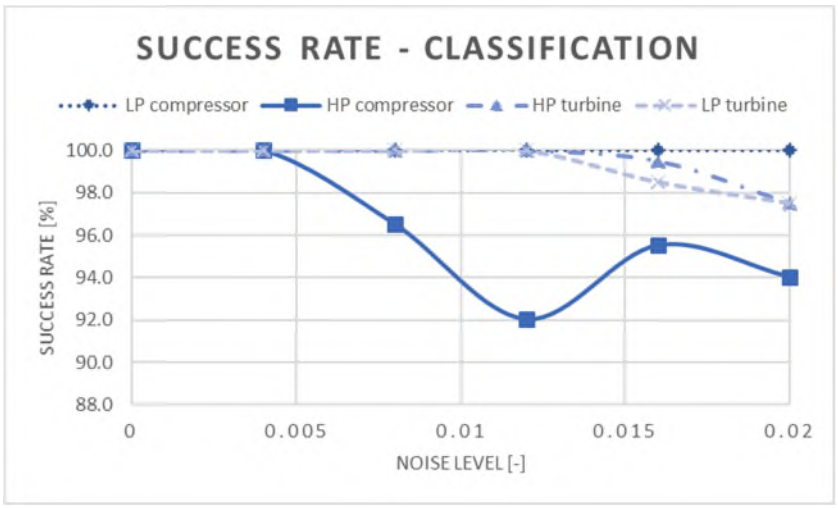

Figure 79 Success rate for the failure classification – Combination SLKF+ANN+NFL+FL

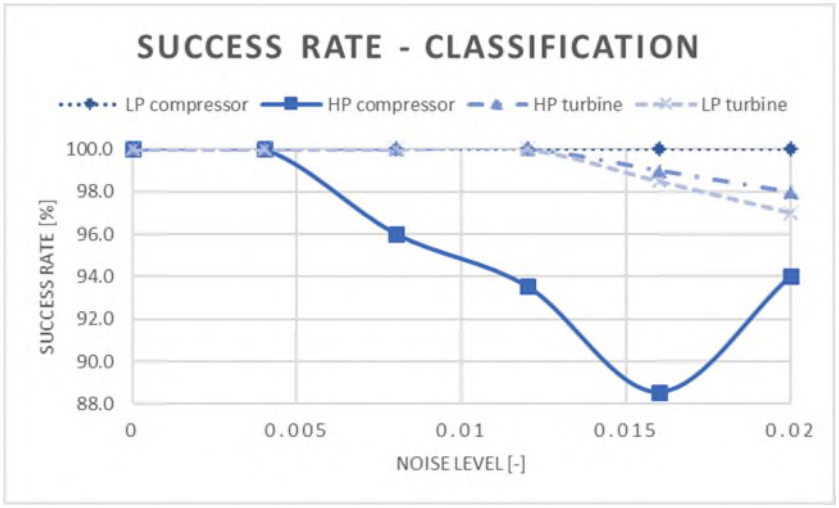

Figure 80 Success rate for the failure classification – Combination MLKF+ANN+NFL+FL

5.2.2 Random deterioration

The random deterioration consists of 203 points with a deterioration level between 0.15% and 7.4% and is identical to what presented with the scheme 1 section 5.1.2 (Figure 81). However, the problem might be different as the samples are selected randomly within the defined range.

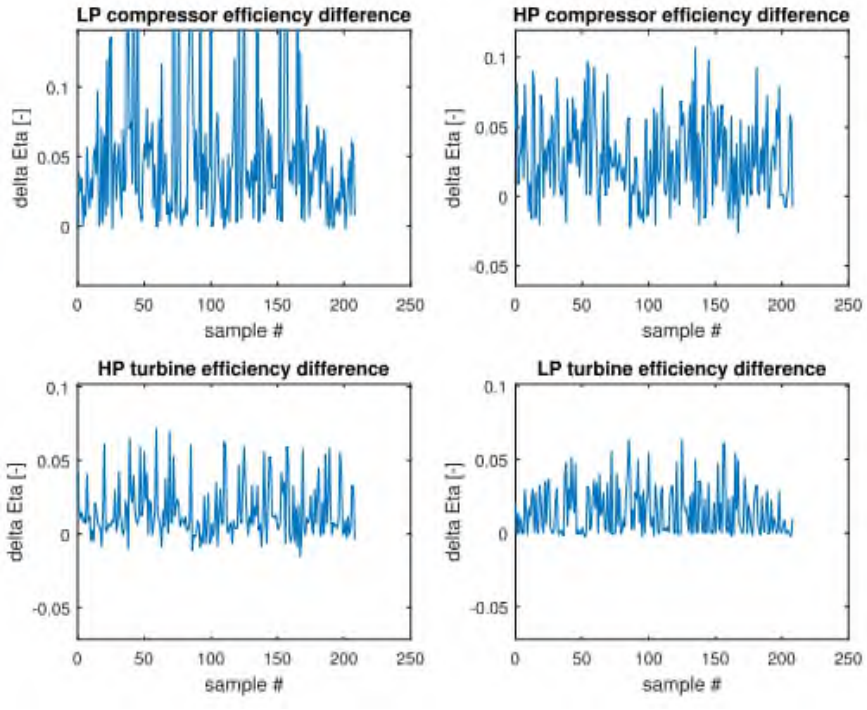

Figure 81 Deterioration imposed on the gas turbine components – 0.4% noise

5.2.2.1 No pre-filtering

The case without the pre-filtering includes the ANN+FL and excludes the KF processing. In this case, in fact, only one measurement per position is considered.

The results of the quantification (Figure 82) are visually very close to the one seen for the scheme 1 section 5.1.2.1 as the range of variation of the quantification is within 5 and 96 (Figure 82).

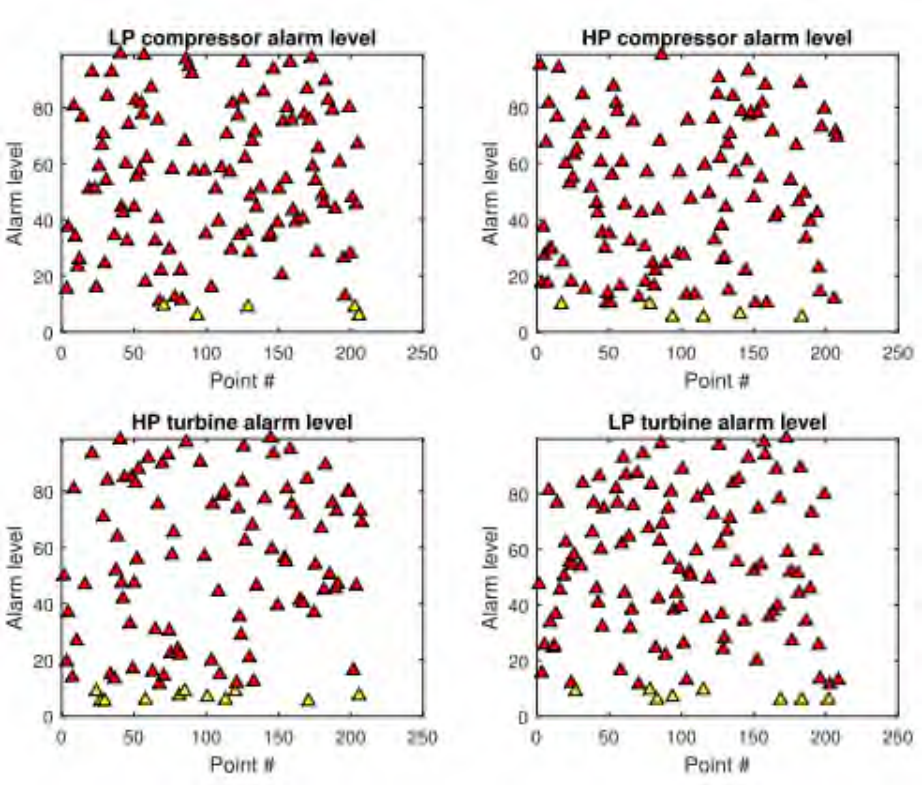

Figure 82 Quantification chart in the case of random deterioration – 0.4% noise no pre-filtering

The quantification success rate ranges from a minimum of 93.5% on the HP turbine to a maximum of 99.0% on the LP compressor (Table 33).

Component	Success rate
LP comp	99.0%
HP comp	96.0%
HP turb	93.5%
LP turb	98.5%

Table 33 Success rate for the failure quantification for random simulation – 0.4% noise no pre-filtering

The classification rate instead, ranges between 93.8% on the HP turbine erosion to the 100% on the HP and LP turbine fouling (Table 34).

Failure type	Success rate
LP comp fouling	93.9%
HP comp fouling	97.1%
HP turb fouling	100%
HP turb erosion	93.8%
LP turb fouling	100%
LP turb erosion	96.7%

Table 34 Success rate for the failure classification for random simulation – 0.4% noise no pre-filtering

The same test repeated for the noise level at 2.0% shows a quantification range between 5 and 200 instead of the 5 to 96 foreseen (Figure 83).

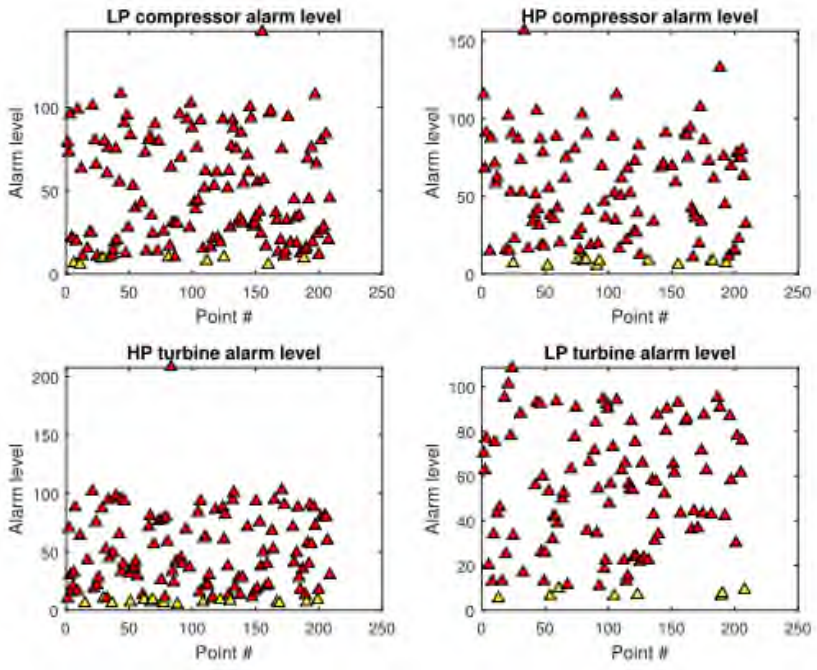

Figure 83 Quantification chart in the case of random deterioration – 2.0% noise no pre-filtering

The consequence of this deviation is reflected in the quantification rate that decreases down to 56.7% on the HP turbine. The best result instead is achieved on the LP turbine where the success rate is 72.6% (Table 35).

Component	Success rate
LP comp	58.7%
HP comp	62.7%
HP turb	56.7%
LP turb	72.6%

Table 35 Success rate for the failure quantification for random simulation – 2.0% noise no pre-filtering

This reduction on the success rate can be also observed in the graphs that are showing an increasing number of deviation spots (Figure 84). The deviation from the reference value is in both directions since the noise is acting randomly around the mean parameter.

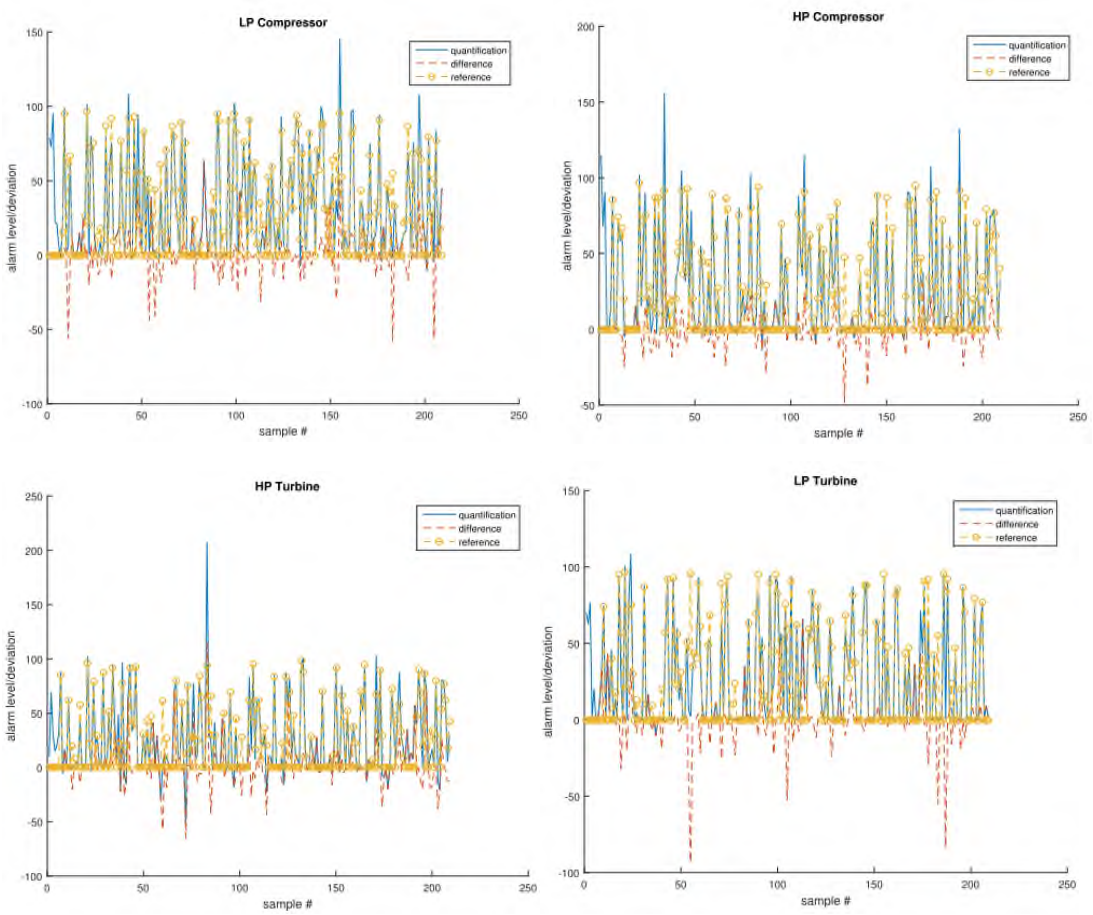

Figure 84 Absolute quantification and deviation from the reference for compressor and turbine components – 2.0% noise no pre-filtering

The classification rate with 2.0% reference noise instead, ranges from 94.4% for the LP compressor fouling to the 100% on the LP turbine fouling (Table 36). The results are better than scheme 1, where the minimum success rate sets at 86.5% on the HP turbine erosion.

Failure type	Success rate
LP comp fouling	94.4%
HP comp fouling	97.8%
HP turb fouling	100%
HP turb erosion	94.7%
LP turb fouling	100%
LP turb erosion	96.0%

Table 36 Success rate for the failure classification for random simulation – 2.0% noise no pre-filtering

5.2.2.2 Single Linear Kalman Filter and measurement fusion

As already seen in section 5.1.2.2, the SLKF is applied before the ANN and the FL to filter for measurement noise.

The results of the quantification repeat the same trend observed in the case without pre-filtering ranging within 0 and 96 (Figure 85). It must be mentioned that as the points are selected randomly, the samples are not directly comparable among the tests.

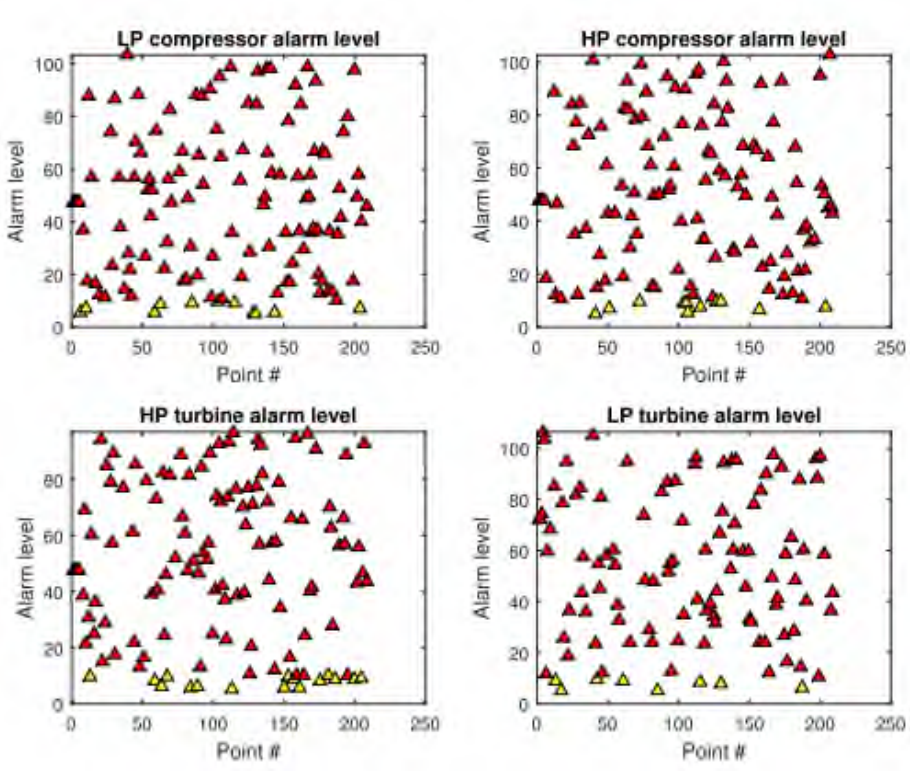

Figure 85 Quantification chart in the case of random deterioration – 0.4% noise SLKF

The quantification success rate ranges from a minimum of 98.0% on the HP compressor and LP turbine to a maximum of 99.0% on the LP compressor (Table 37). This result is in line with what obtained with scheme 1 and is an improvement compared to the case without filter, where the minimum success rate was 93.5% on the HP turbine.

Component	Success rate
LP comp	99.0%
HP comp	98.0%
HP turb	98.5%
LP turb	98.0%

Table 37 Success rate for the failure quantification for random simulation – 0.4% noise SLKF

The classification rate instead, ranges from 89.1% for the LP turbine erosion to the 100% on the HP and LP turbine fouling (Table 38). This result is slightly below the target of 90% set for this work and is up to 6% worse than its homologous in scheme 1.

Failure type	Success rate
LP comp fouling	98.6%
HP comp fouling	97.1%
HP turb fouling	100%
HP turb erosion	89.8%
LP turb fouling	100%
LP turb erosion	89.1%

Table 38 Success rate for the failure classification for random simulation – 0.4% noise SLKF

Looking at the case with 2.0% measurement noise, the quantification ranges between 5 and 105 (Figure 86).

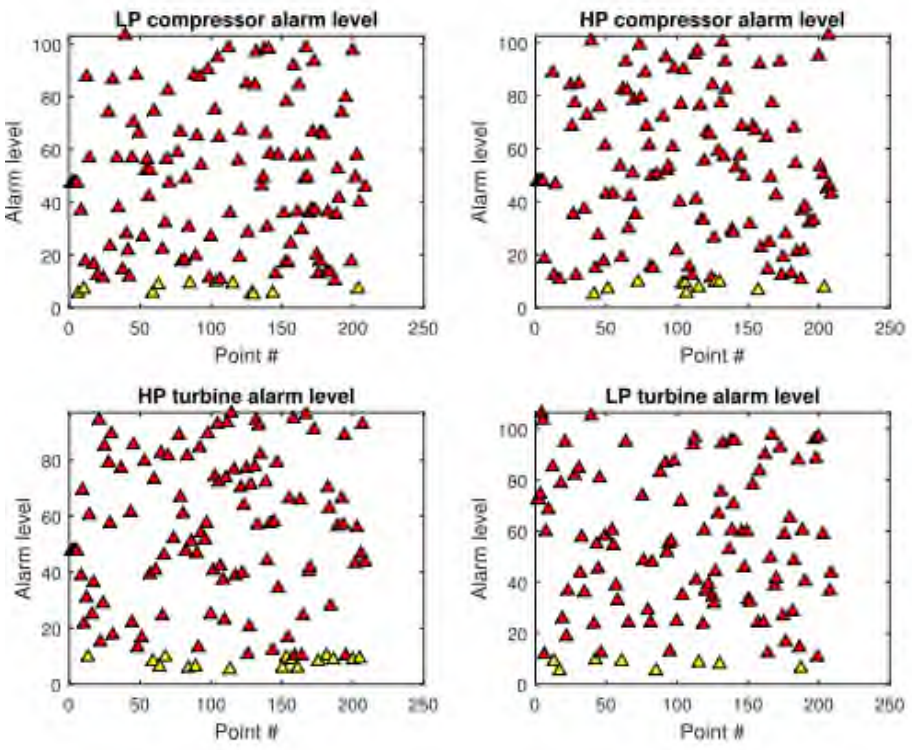

Figure 86 Quantification chart in the case of random deterioration – 2.0% noise SLKF

The resulting success rate varies, between 70.1% on the HP compressor to 78.6% on the LP turbine with an improvement on all the components compared to the case without filtering (Table 39).

Component	Success rate
LP comp	73.6%
HP comp	70.1%
HP turb	73.6%
LP turb	78.6%

Table 39 Success rate for the failure quantification for random simulation – 2.0% noise SLKF

This reduction on the success rate can be also observed in the graph that is showing an increasing number of deviation spots (Figure 87).

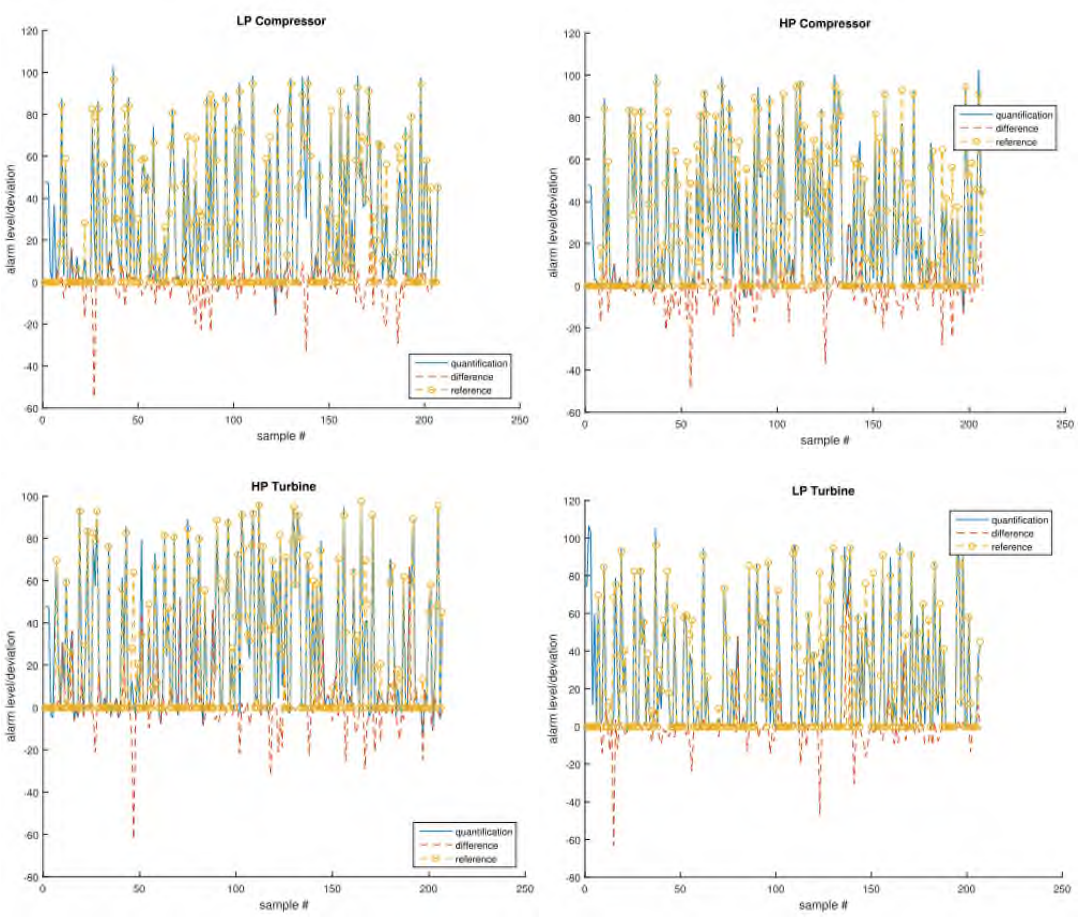

Figure 87 Absolute quantification and deviation from the reference for compressor and turbine components – 2.0% noise SLKF

The classification rate ranges between 86.0% for the LP turbine erosion to 100% on the LP and HP turbine fouling (Table 40).

Failure type	Success rate
LP comp fouling	95.8%
HP comp fouling	97.1%
HP turb fouling	100%
HP turb erosion	93.8%
LP turb fouling	100%
LP turb erosion	86.0%

Table 40 Success rate for the failure classification for random simulation – 2.0% noise SLKF

5.2.2.3 Multiple Layer Kalman Filter and measurement fusion

The other scheme analysed in this section is the MLKF that consist of two layers devoted to merge all the information coming from every single measurement and to filter out the measurement noise. The aim of the second layer is to further make use of the filtered information to improve the final result.

The results of the quantification with 0.4% reference noise, repeat the same trend observed in the case without pre-filtering ranging within 5 and 96 (Figure 88).

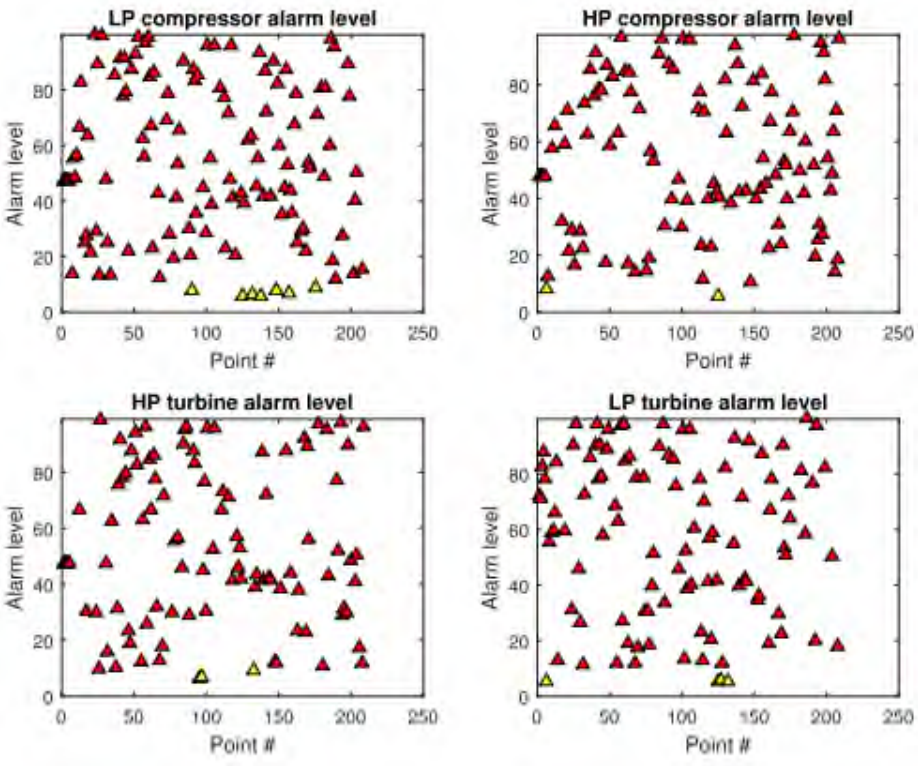

Figure 88 Quantification chart in the case of random deterioration – 0.4% noise MLKF

The quantification success rate ranges from a minimum of 97.5% on the HP and LP turbine to a maximum of 100% on the LP compressor (Table 41). Compared to the case without a filter, there is already an improvement but compared to the case with SLKF there is almost no change.

Component	Success rate
LP comp	100%
HP comp	98.5%
HP turb	97.5%
LP turb	97.5%

Table 41 Success rate for the failure quantification for random simulation – 0.4% noise MLKF

The classification rate instead, ranges from 93.2% for the LP turbine erosion to 100% on the HP turbine fouling (Table 42). This result is in line with scheme 1 and with the configuration with SLKF and scheme 2.

Failure type	Success rate
LP comp fouling	93.7%
HP comp fouling	98.2%
HP turb fouling	100%
HP turb erosion	98.2%
LP turb fouling	100%
LP turb erosion	93.2%

Table 42 Success rate for the failure classification for random simulation – 0.4% noise MLKF

Looking at the case with 2.0% reference noise, the quantification ranges between 5 and 105 (Figure 89), result in line with the case with SLKF.

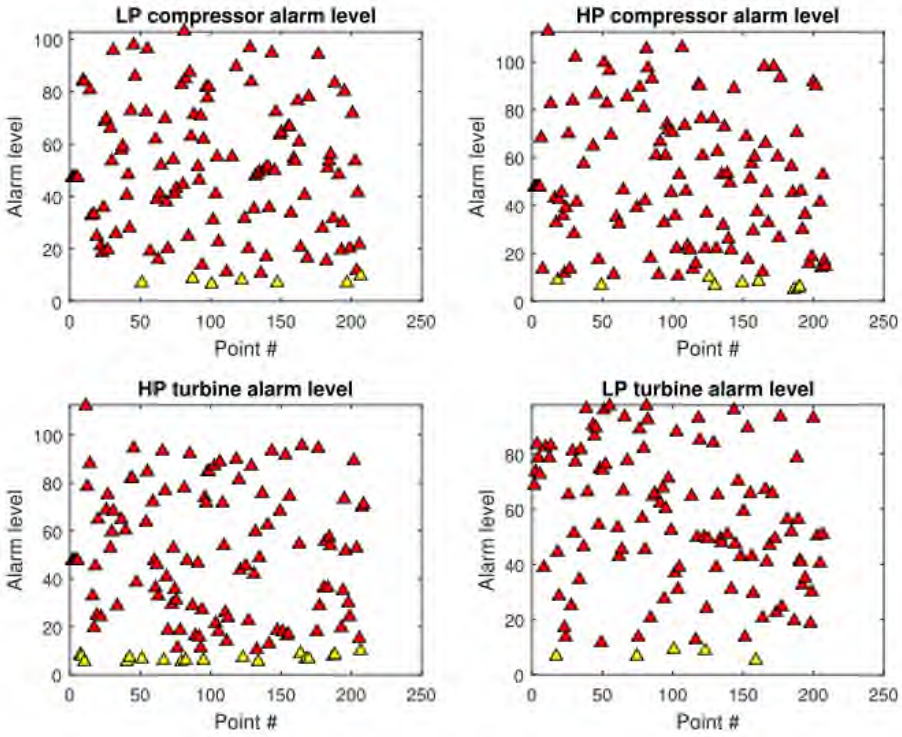

Figure 89 Quantification chart in the case of random deterioration – 2.0% noise MLKF

The resulting success rate varies, from 65.7% on the HP turbine to 78.6% on the LP turbine (Table 43). Compared to scheme 1, the success rate has a worsening in all the components. Compared with the SLKF seen before, there are variations on the LP compressor, where the success rate increased by 3% and on the HP turbine where the success rate surprisingly decreases by almost 8%.

Component	Success rate
LP comp	76.1%
HP comp	70.1%
HP turb	65.7%
LP turb	78.6%

Table 43 Success rate for the failure quantification for random simulation – 2.0% noise MLKF

The lower rate seen in the HP turbine is also reflected in the higher deviations from the reference quantification value (Figure 90).

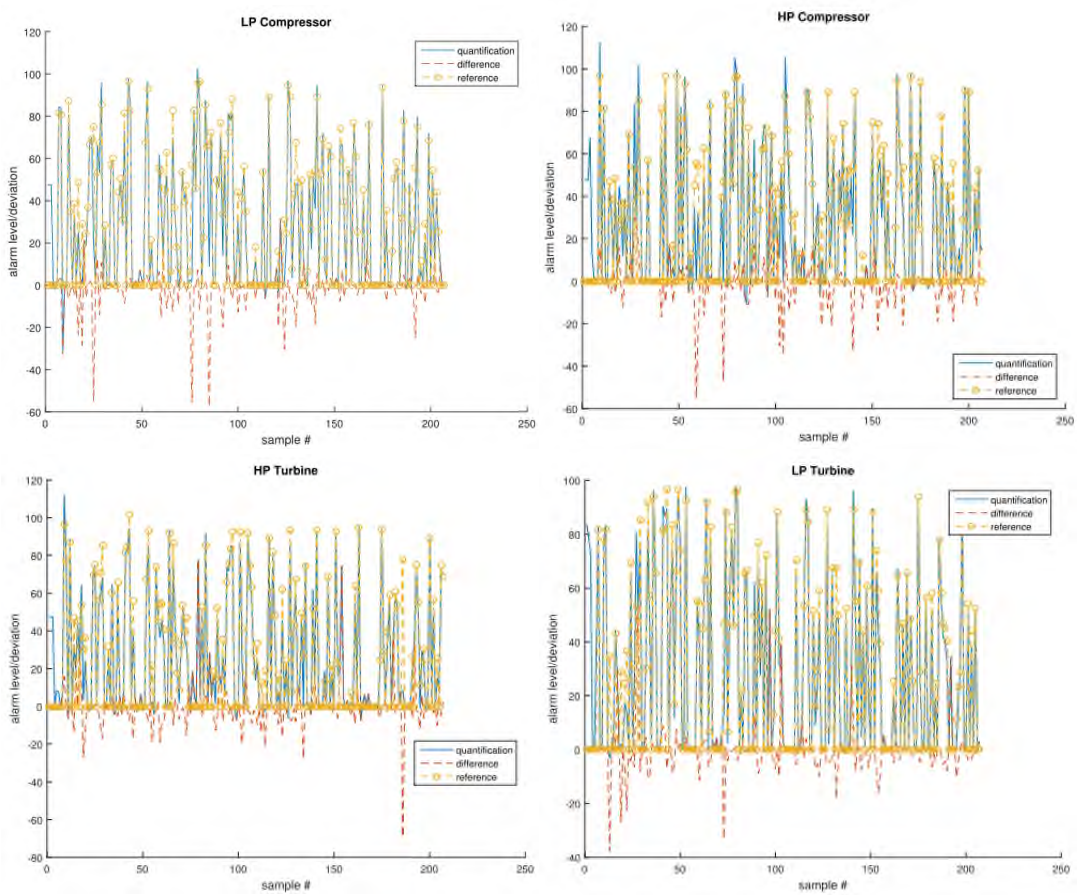

Figure 90 Absolute quantification and deviation from the reference for compressor and turbine components – 2.0% noise MLKF

The classification rate instead, ranges between 91.7% for the HP turbine erosion to the 100% on the HP and LP turbine fouling (Table 44).

Failure type	Success rate
LP comp fouling	98.6%
HP comp fouling	98.5%
HP turb fouling	100%
HP turb erosion	91.7%
LP turb fouling	100%
LP turb erosion	93.9%

Table 44 Success rate for the failure classification for random simulation – 2.0% noise MLKF

The analysis of the random section confirms that the methodology can perform the multiple components of health estimation while classifying different type of failures, also without the use of Turbomatch. The

thermodynamic model instead is fully taken over by the ANN that calculates the reference value and the deteriorated value.

5.2.3 Deterioration schedule

The deterioration schedule consists of 203 samples with deterioration ranging from 0.0% to 5.1%. The schedule of the deterioration is identical to the one presented in section 5.1.3.1. The simulation, as in scheme 1, is carried over with 0.4% measurement noise.

5.2.3.1 Combined training

The combined training is made up by the random simulation and by the deterioration schedule. The result of this merge is an increased number of points available to the ANN. As already remarked in section 5.1.3.1 the consequence can be an increased number of data available that could improve the ANN quality prediction but could also be overfitting on some points that could worsen the prediction. With this scheme, it is important to check the reaction of the ANN while predicting the deteriorated values and their references.

Once the efficiencies and physical values deviations are processed by the fuzzy logic, they deal with the quantification described in section 4.7. Based on the deterioration selected for this problem, 3.3% on the LP compressor and 5.1% on the LP turbine, the prediction is to have 43 as alarm level on the LP compressor and 66 on the LP turbine (Figure 91). Comparing the graphs with the predicted numbers, they look closely with the prediction and in line with scheme 1.

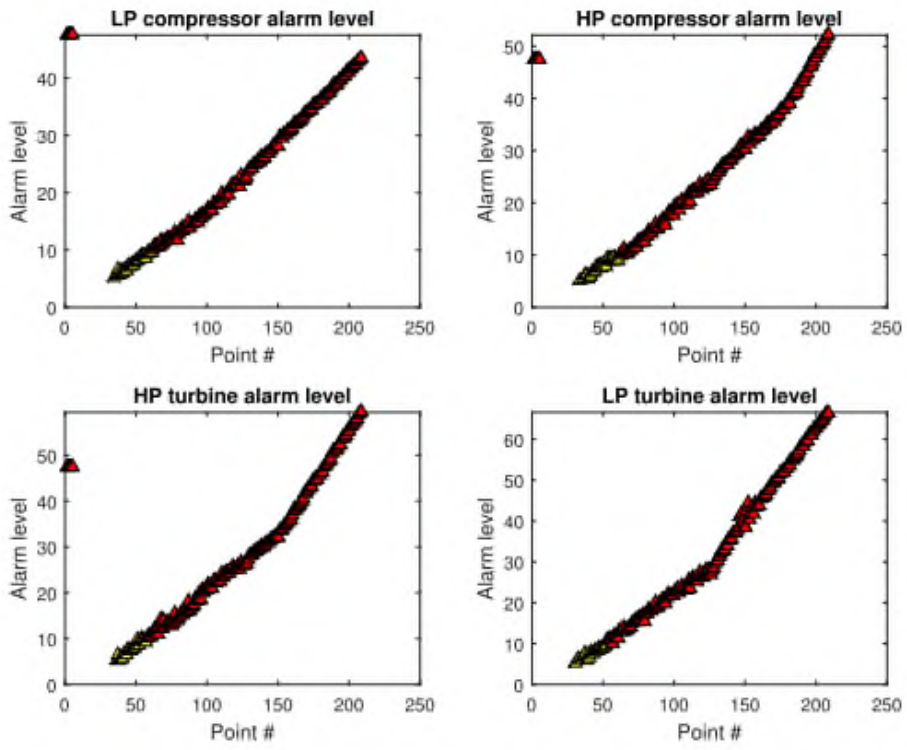

Figure 91 Quantification chart in the case of scheduled deterioration

The resulting quantification success rate is 100% for all the components (Table 45).

Component	Success rate
LP comp	100%
HP comp	100%
HP turb	100%
LP turb	100%

Table 45 Success rate for the failure quantification for scheduled simulation – 0.4% noise MLKF

The very high success rate is also confirmed by the tight alignment between the predicted component health status and the reference (Figure 92). Moreover, the lines look even tighter than scheme 1, due to the better reference provided by the ANN.

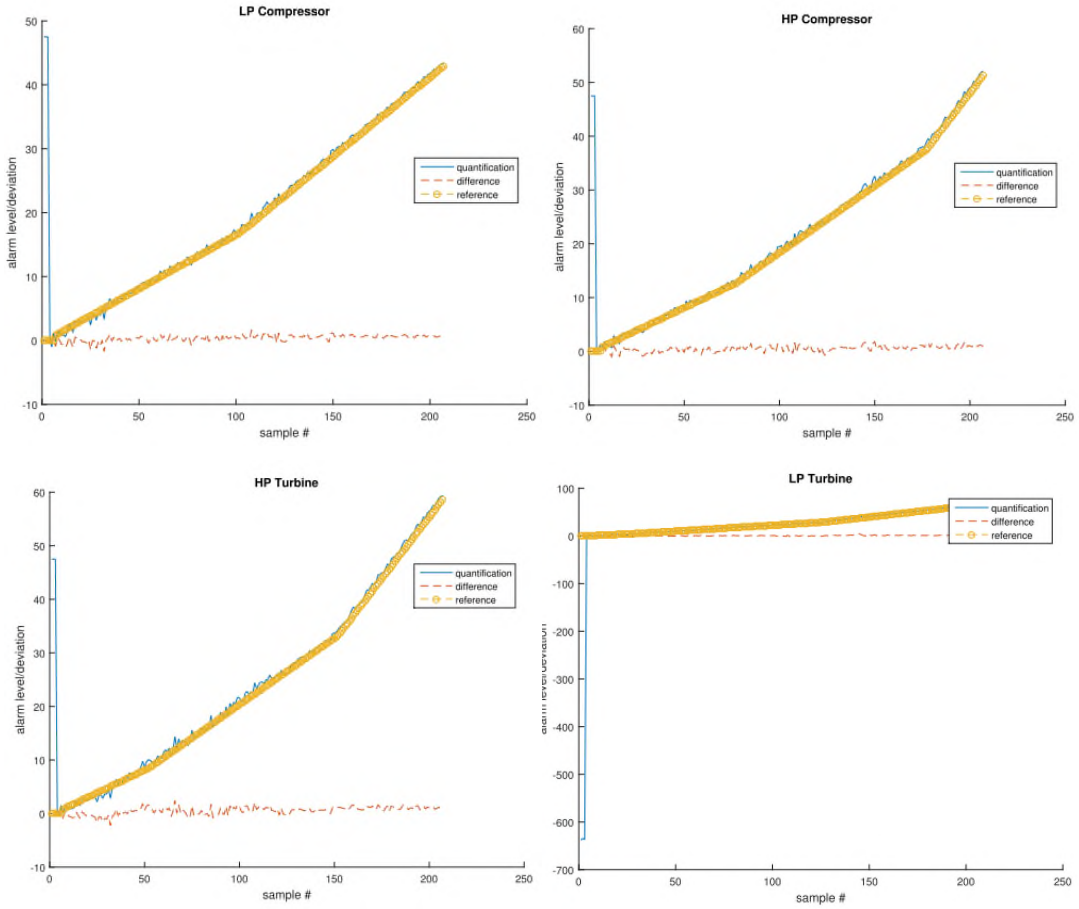

Figure 92 Absolute quantification and deviation from the reference for compressor and turbine components – 0.4% noise MLKF

The next step to be tested is classification. The resulting success rate ranges from 94.1% on the HP turbine to 95.5% on the HP compressor (Table 46). Despite these results above the target, one dot is visible in the erosion section, a sign of false alarm (Figure 93 and Figure 94).

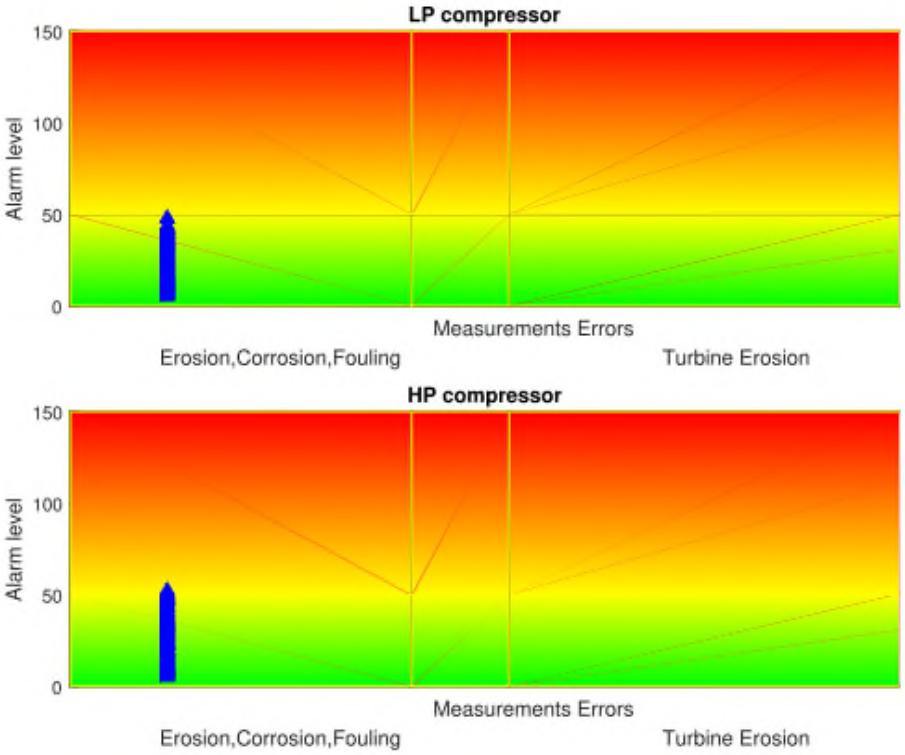

Figure 93 Classification chart for the compressor

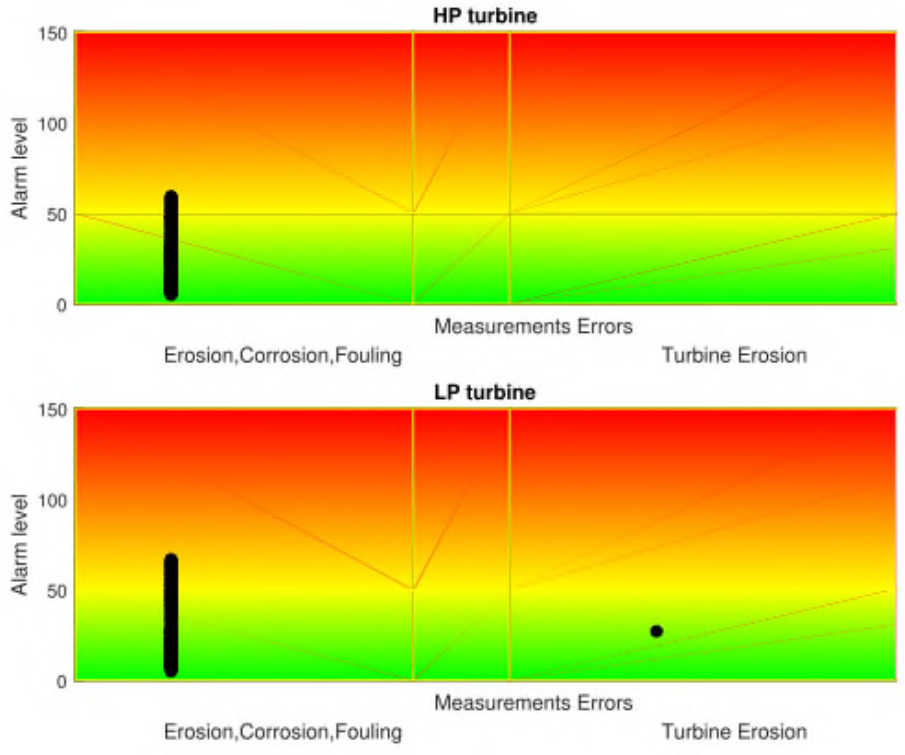

Figure 94 Classification chart for the turbine

Failure type	Success rate
LP comp fouling	94.6%
HP comp fouling	95.5%
HP turb fouling	94.1%
LP turb fouling	94.6%

Table 46 Success rate for the failure classification for scheduled simulation – 0.4% noise MLKF

These results confirm the feasibility of the failure characterization with a deterioration schedule of several months and without the Turbomatch reference model. Moreover, scheme 2 outline that some uncertainties can be even improved by moving the prediction to the ANN instead of relying on the ambient based reference.

5.3 Scheme 3

Scheme 3 uses Turbomatch model as a reference but excludes the power. The power, in fact, might not be available as a direct measurement of the GT output and a replacement would be needed. The noise is included as in all other measurements.

5.3.1 Constant deterioration

5.3.1.1 Measurements noise

The constant deterioration has identical characteristics as for scheme 1 – section 5.1.

The first parameter to check with this scheme is the quantification, that should end at the level of 96 for each gas turbine component. The quantification obtained with the first combination considered – without KF applied – is in line with the expectations and with the results of scheme 1 and scheme 2 (Figure 95). It must be remarked that the LP turbine quantification has a higher variation with on-point reaching the level of 80 in some samples.

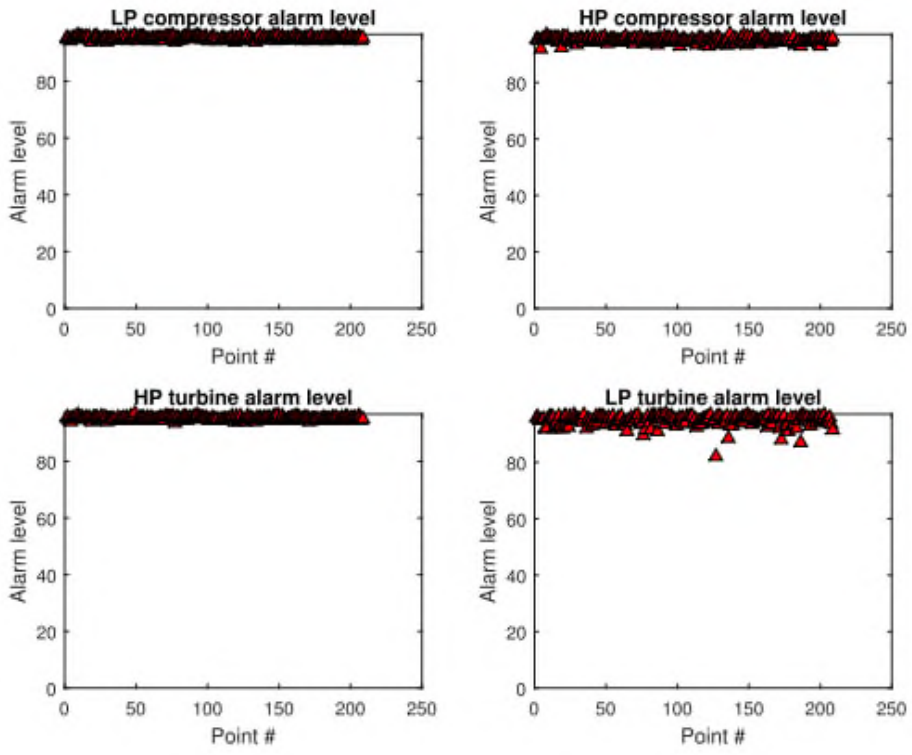

Figure 95 Quantification of the gas turbine failure – 0.4% noise no pre-filtering

Considered that the given type of failure is fouling on the LP and HP compressor and erosion on the HP and LP turbine, the classification chart shows the correct classification being all samples of the compressor on the fouling side and all the samples of the turbine on the erosion side (Figure 96 and Figure 97).

The increased quantification variation outlined before is reflected also in the combined chart (Figure 97).

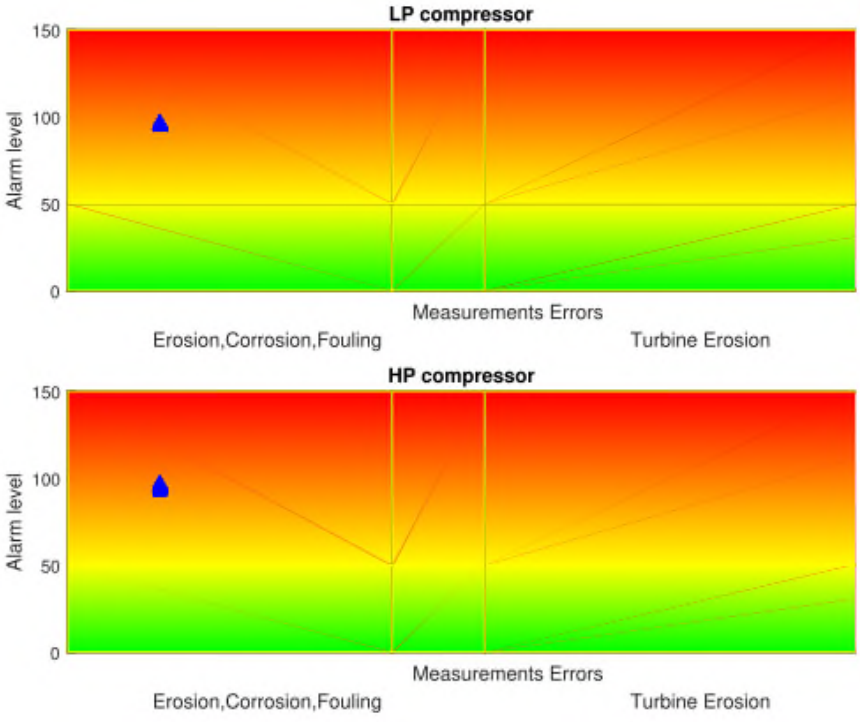

Figure 96 Quantification and classification of the compressor – 0.4% noise no pre-filtering

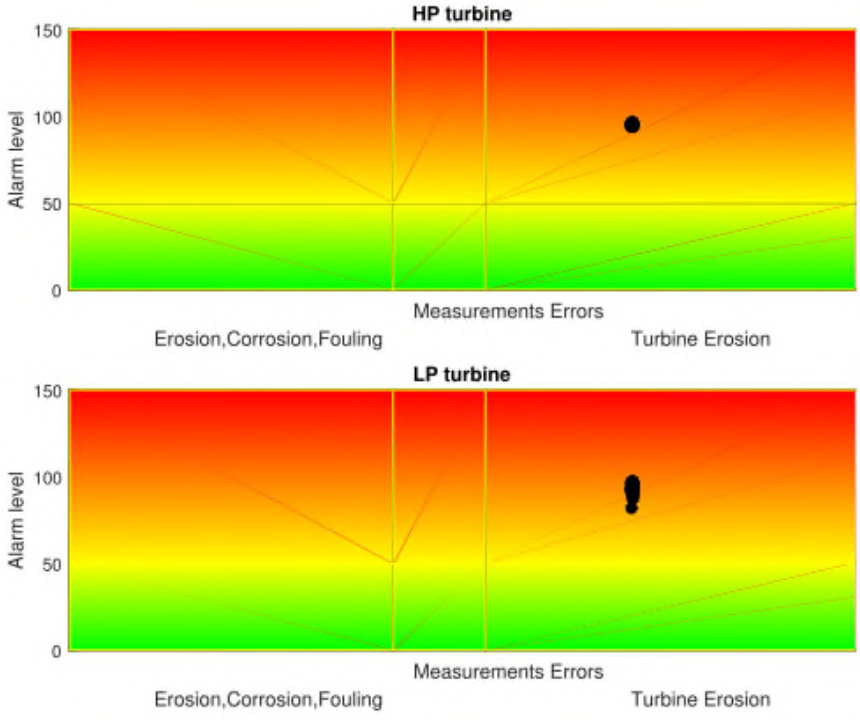

Figure 97 Quantification and classification of the turbine – 0.4% noise no pre-filtering

However, if the level of noise increases, there is a clear disturbance on the quantification and of the classification success rate. Taking into account the 2.0% measurement noise example, the quantification instead of being around 96, ranges from 40 to 150 where 40 corresponds to a medium deteriorated component and 150 corresponds to a highly deteriorated component (Figure 98 and Figure 99). This reaction is directly related to the noise that while increasing, causes the measurement values to increase their standard deviation around the mean value. Translated this means that the health status is underestimated in one case, leading to slower reaction on the maintenance actions, while is overestimated in another case leading to faster reactions. The classification instead, leads to false categorizations as some samples fall into the wrong category. The reason is also related to the increased noise as clarified in scheme 1 section 5.1. In this case, the maintenance would be addressed on a wrong action.

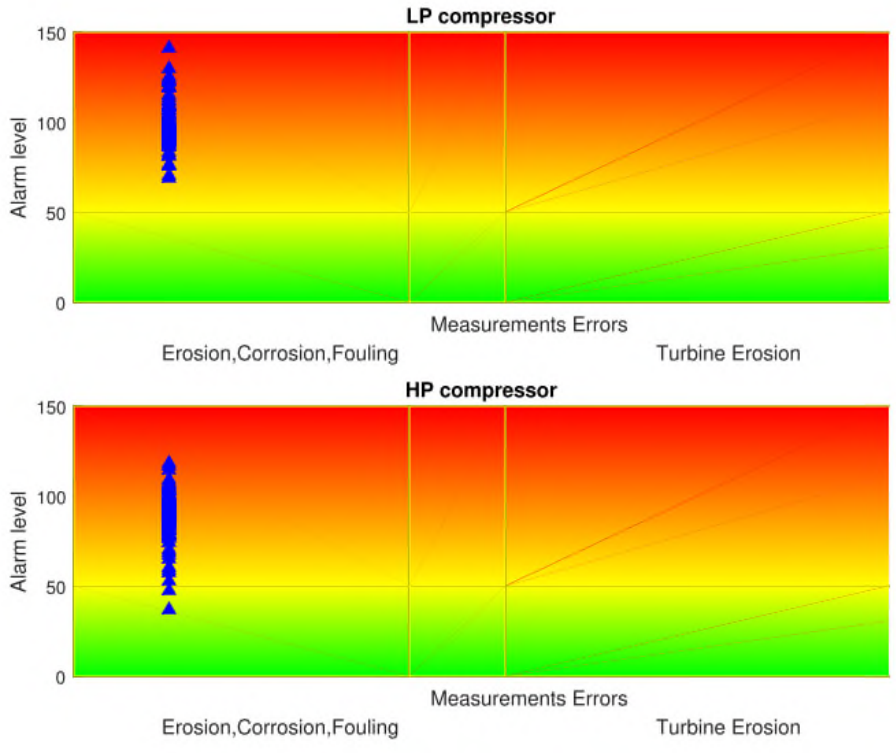

Figure 98 Quantification and classification of the compressor – 2.0% noise no pre-filtering

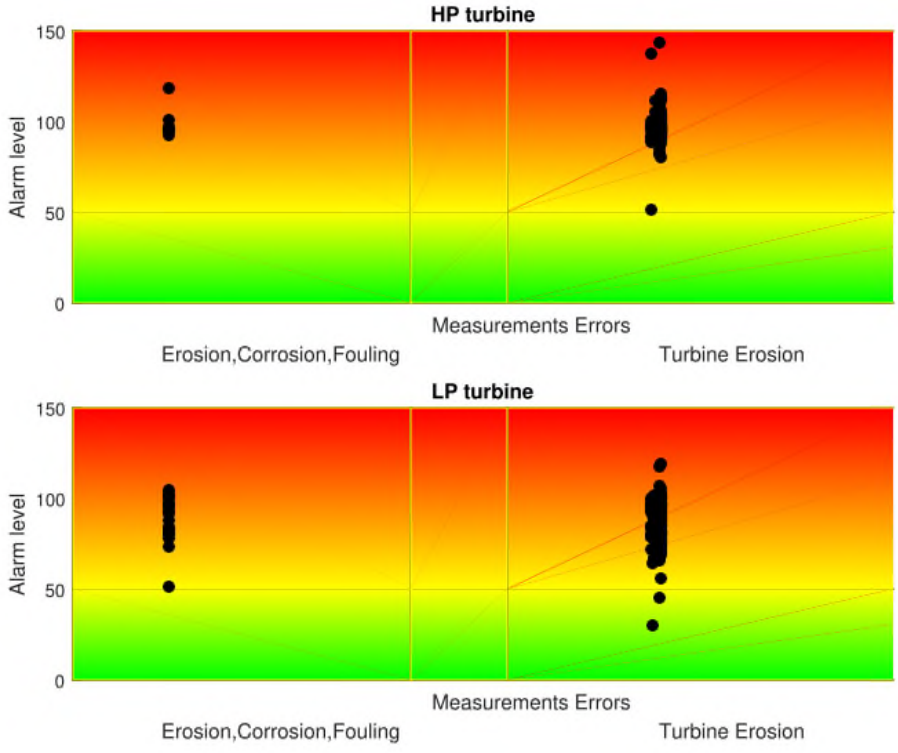

Figure 99 Quantification and classification of the turbine – 2.0% noise no pre-filtering

Looking at the results including the MLKF with 2.0% measurement noise, the spread of the magnitude is decreased to the range 50 – 110 (Figure 100 and Figure 101). However, some false alarms persist on the turbine side (Figure 101).

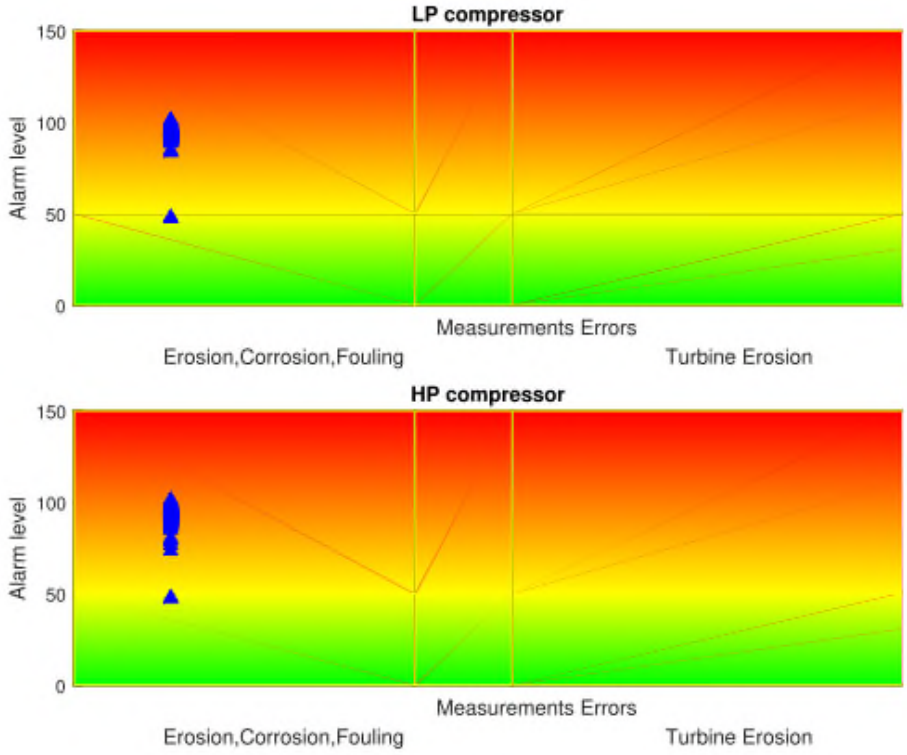

Figure 100 Quantification and classification of the compressor – 2.0% noise MLKF

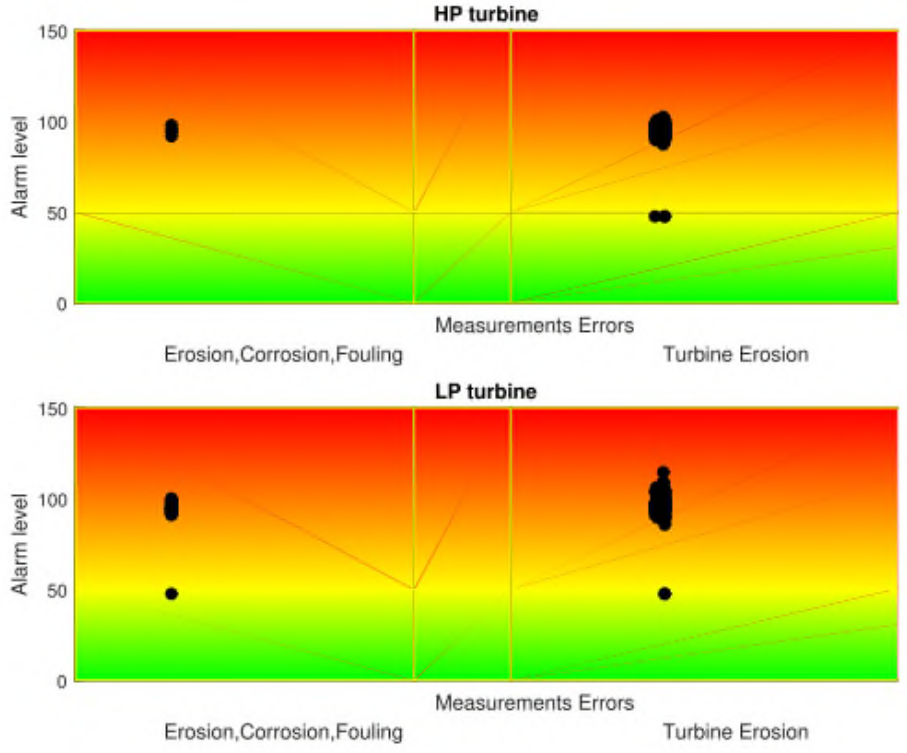

Figure 101 Quantification and classification of the turbine – 2.0% noise MLKF

The overall results of the first test show a quantification success rate of 100% for 0.0% noise and for the nominal noise of 0.4% with all the combinations. This result proves that this methodology can work also without power measurement. This achievement confirms that the methodology is applicable also where the power is not a variable for example in industrial gas turbines configurations that cannot measure the power produced by the gas turbine standalone.

While the level of noise is increasing, the data analysis and correction starts to become more important. At 0.8% reference noise in-fact, the success level on the quantification decrease to 90.5% if no KF is applied. With the maximum noise level of 2.0%, the rate of success is decreased to 39.8% without the KF block. The rate increases to 78.6% if also the SLKF is employed and to 84.6% if the MLKF is used (Table 47). These rates are even better than obtained with scheme 1 and scheme 2.

	Noise	0.0%	0.4%	0.8%	1.2%	1.6%	2.0%
None	LP comp	100%	100%	97.0%	93.5%	83.6%	72.6%
	HP comp	100%	100%	99.5%	92.0%	79.1%	67.7%
	HP turb	100%	100%	97.0%	95.5%	87.1%	78.6%
	LP turb	100%	100%	90.5%	70.1%	58.2%	39.8%
SLKF	LP comp	100%	100%	100%	100%	97.5%	94.5%
	HP comp	100%	100%	100%	100%	98.0%	90.5%
	HP turb	100%	100%	100%	99.5%	97.5%	94.5%
	LP turb	100%	100%	100%	94.5%	90.5%	78.6%
MLKF	LP comp	100%	100%	100%	100%	99.5%	94.0%
	HP comp	100%	100%	100%	99.5%	96.5%	92.0%
	HP turb	100%	100%	100%	100%	97.5%	96.0%
	LP turb	100%	100%	99.0%	96.0%	89.1%	84.6%

Table 47 Success rate for the failure quantification

Looking at the graphs instead, it can be seen that the success rate is above 90% for the SLKF combination at noise up to 1.6%. After, the component with a lower success rate is the LP turbine (Figure 103). The combination with MLKF falls below 90% at 1.6% reference noise (89.1%) but stays above 84.6% at 2.0% reference noise (Figure 104). This result is not far from the 90% target and is outperforming both scheme 1 and scheme 2. This means that the inclusion of power measurement can lead to higher uncertainties

especially because, being the power a single measurement, it cannot be treated into the KF section. Without filters, the 90% rate can be achieved only up to 0.8% reference noise (Figure 102).

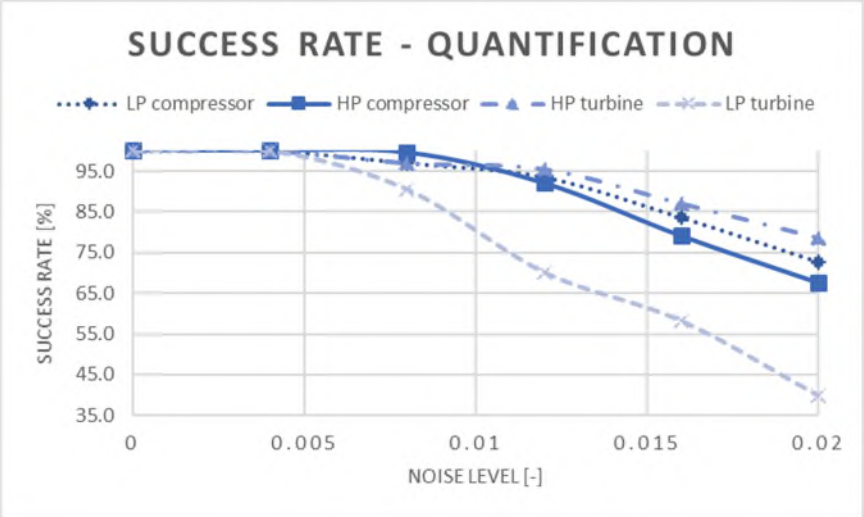

Figure 102 Success rate for the failure quantification – Combination No KF

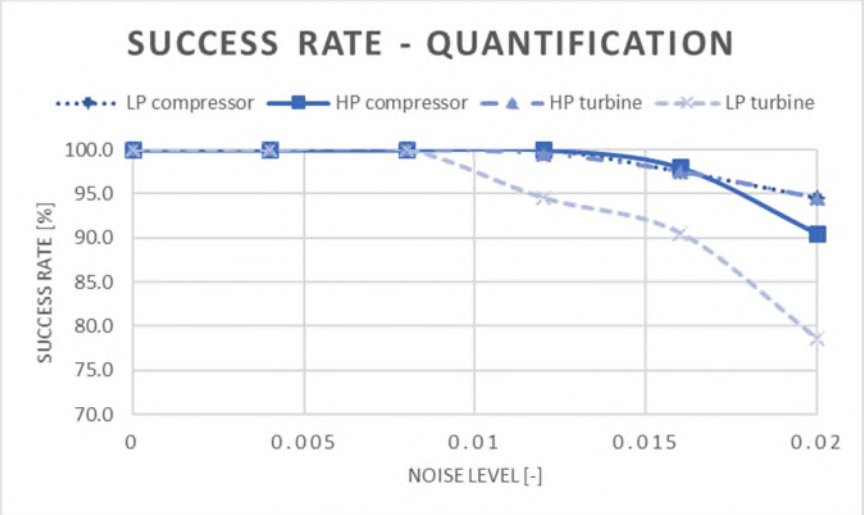

Figure 103 Success rate for the failure quantification – Combination SLKF+ANN+NFL+FL

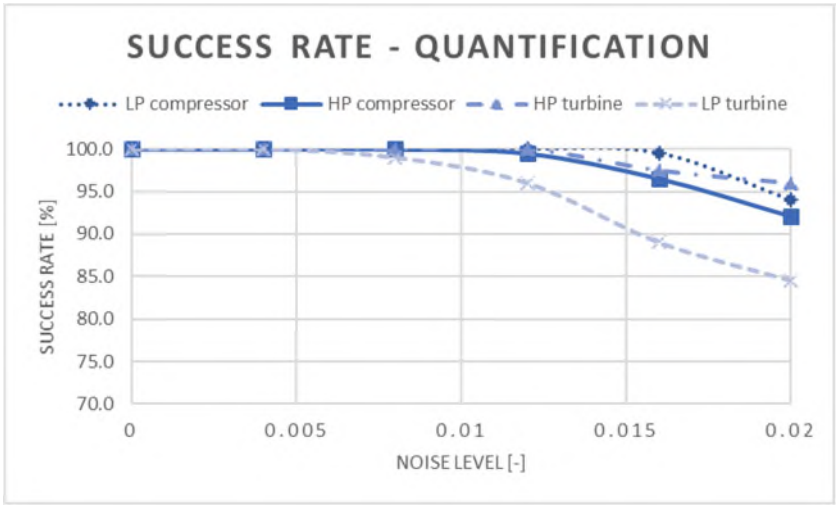

Figure 104 Success rate for the failure quantification – Combination MLKF+ANN+NFL+FL

The results of the classification with nominal noise show a success rate above 95.6% without filtering and at 100% with MLKF. With the maximum noise, the success rate decreases to 89.7% if no filtering is used and increases again to 89.6% with MLKF. Additionally, for this scheme, the MLKF classification success rate is better than the case without filter and with SLKF (Table 48).

	Noise	0.0%	0.4%	0.8%	1.2%	1.6%	2.0%
None	LP comp	100%	100%	100%	100%	100%	100%
	HP comp	100%	95.6%	93.6%	95.6%	95.6%	95.6%
	HP turb	100%	100%	100%	100%	99.0%	96.1%
	LP turb	100%	100%	100%	97.5%	94.1%	89.7%
SLKF	LP comp	100%	100%	100%	100%	100%	100%
	HP comp	100%	100%	100%	97.5%	95.5%	94.5%
	HP turb	100%	100%	100%	99.0%	98.5%	98.5%
	LP turb	100%	100%	99.0%	96.5%	93.5%	89.6%
MLKF	LP comp	100%	100%	100%	100%	100%	100%
	HP comp	100%	100%	100%	100%	100%	100%
	HP turb	100%	100%	100%	99.0%	99.5%	98.0%
	LP turb	100%	100%	100%	96.0%	94.5%	89.6%

Table 48 Success rate for the failure classification

It must be remarked that the classification is less affected by the noise compared to the quantification (Figure 105). The minimum rate in-fact is 89.6% at 2.0% measurement noise, compared to the 39.8% for the quantification. Compared to scheme 1 and scheme 2, the local minimum at

1.2% measurement noise is not visible and the success rate decreases start from 1.2% and decrease gradually (Figure 106).

The results of this scheme confirm the feasibility of the multiple components of health estimation with the presence of measurement noise, also without the power measurement included in the ANN training. Not only, but the results with the MLFK also provide an engine health estimation not fat from the target, with a success rate of 84.6% (Figure 107). On the other hand, the classification success rate is slightly worsened but is also very close to the target.

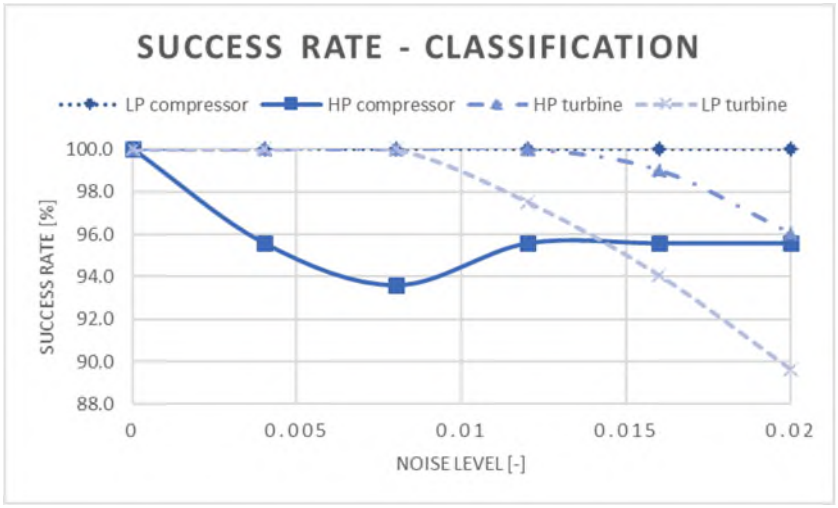

Figure 105 Success rate for the failure classification – Combination No KF

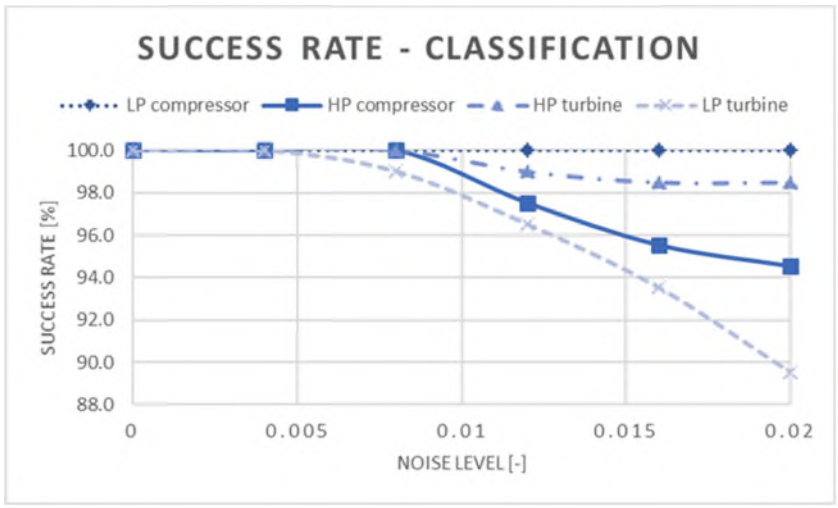

Figure 106 Success rate for the failure classification – Combination SLKF+ANN+NFL+FL

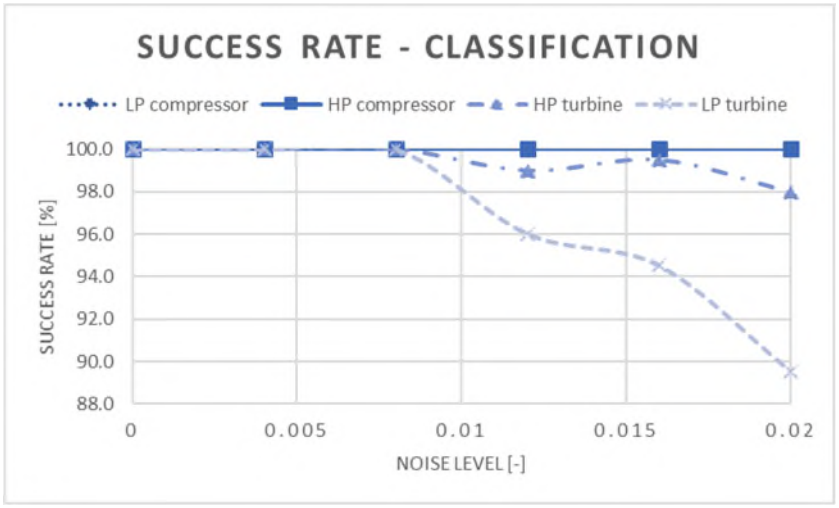

Figure 107 Success rate for the failure classification – Combination MLKF+ANN+NFL+FL

5.3.2 Random deterioration

The random deterioration consists of 203 points with deterioration level between 0.15% and 7.4% (Figure 108). The set-up of the deterioration is identical as the one set for scheme 1 and scheme 2 but the types and the sequences may differ due to the random nature of the selection.

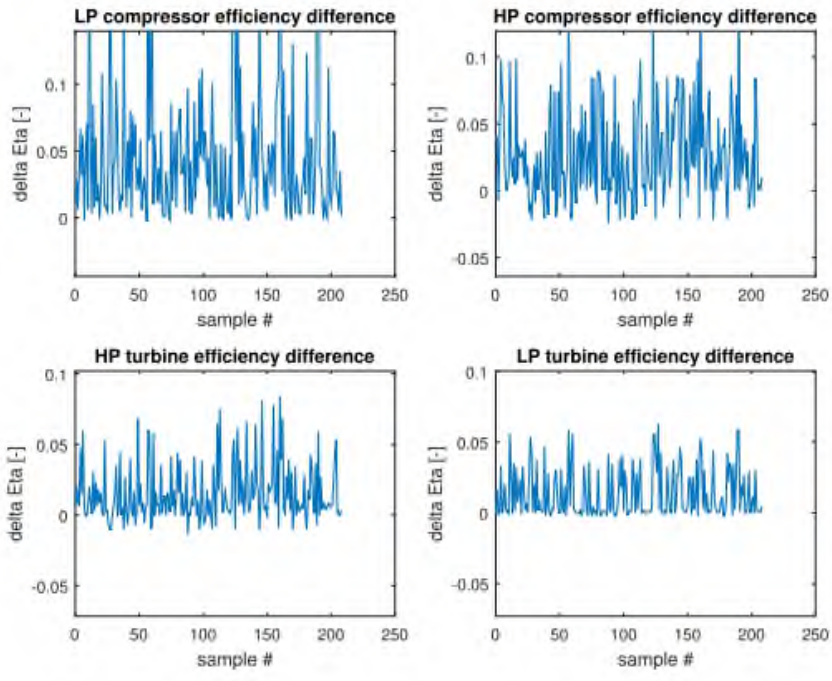

Figure 108 Deterioration imposed on the gas turbine components – 0.4% noise

5.3.2.1 No pre-filtering

The first case to be tested is the one without filtering. This case includes the structure with the ANN+FL but excludes the KF for the measurement noise isolation with multiple measurements combination. The level of noise for this test is the nominal – 0.4% - and the maximum among those selected for the constant deterioration - 2.0%.

The results of the quantification (Figure 109) show the range of variation of the quantification from 5 to 96 in line with the prediction and in line with scheme 1 and scheme 2.

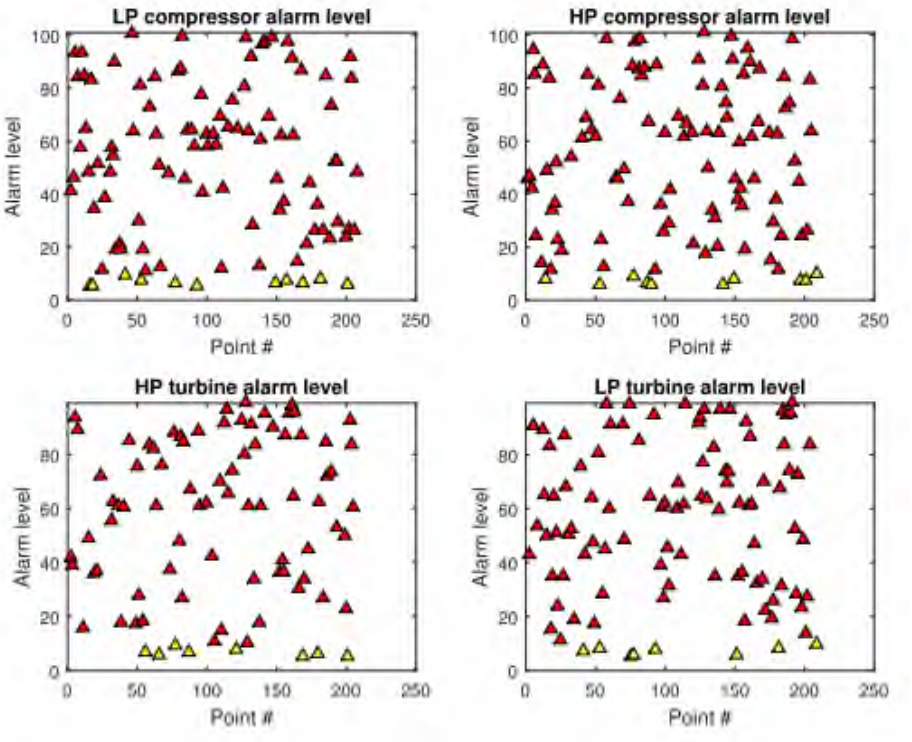

Figure 109 Quantification chart in the case of random deterioration - 0.4% noise no pre-filtering

The quantification success rate ranges from a minimum of 96.5% on the HP compressor to a maximum of 98.0% on the LP compressor (Table 49).

Component	Success rate
LP comp	98.0%
HP comp	96.5%
HP turb	97.0%
LP turb	97.5%

Table 49 Success rate for the failure quantification for random simulation – 0.4% noise no pre-filtering

The classification rate instead, ranges from 90.2% for the HP turbine erosion to the 100% on the HP and LP turbine fouling (Table 50).

Failure type	Success rate
LP comp fouling	92.9%
HP comp fouling	98.0%
HP turb fouling	100%
HP turb erosion	90.2%
LP turb fouling	100%
LP turb erosion	94.1%

Table 50 Success rate for the failure classification for random simulation – 0.4% noise no pre-filtering

The same test, repeated with 2.0% measurement noise, show a deviation from the expected range 5-96 as the samples range between 5 and 150 (Figure 110).

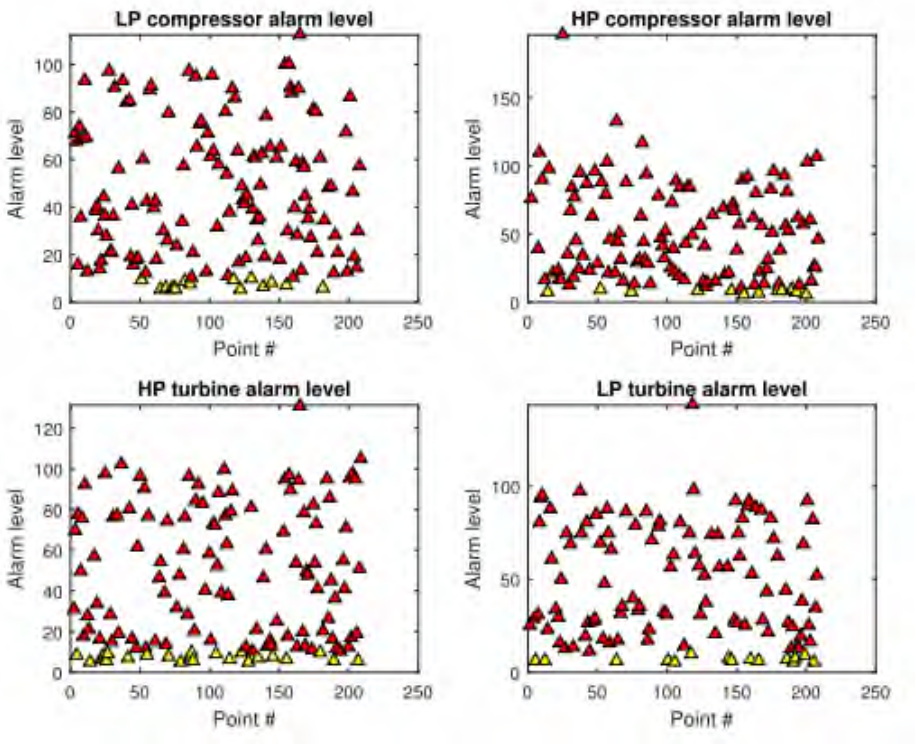

Figure 110 Quantification chart in the case of random deterioration – 2.0% noise no pre-filtering

The consequence of this deviation from the exact quantification is a decrease of the success rate that ranges, in this case, from 51.7% on the HP turbine to 64.7% on the LP turbine (Table 51). These rates are lower compared to scheme 1 and scheme 2 confirming the increasing sensitivity of the methodology once the number of signals available decrease. This conclusion is also in line with the results of the constant deterioration.

Component	Success rate
LP comp	60.2%
HP comp	53.2%
HP turb	51.7%
LP turb	64.7%

Table 51 Success rate for the failure quantification for random simulation – 2.0% noise no pre-filtering

This reduction on the success rate can be also observed in the graph that is showing an increasing number of deviation spots (Figure 111).

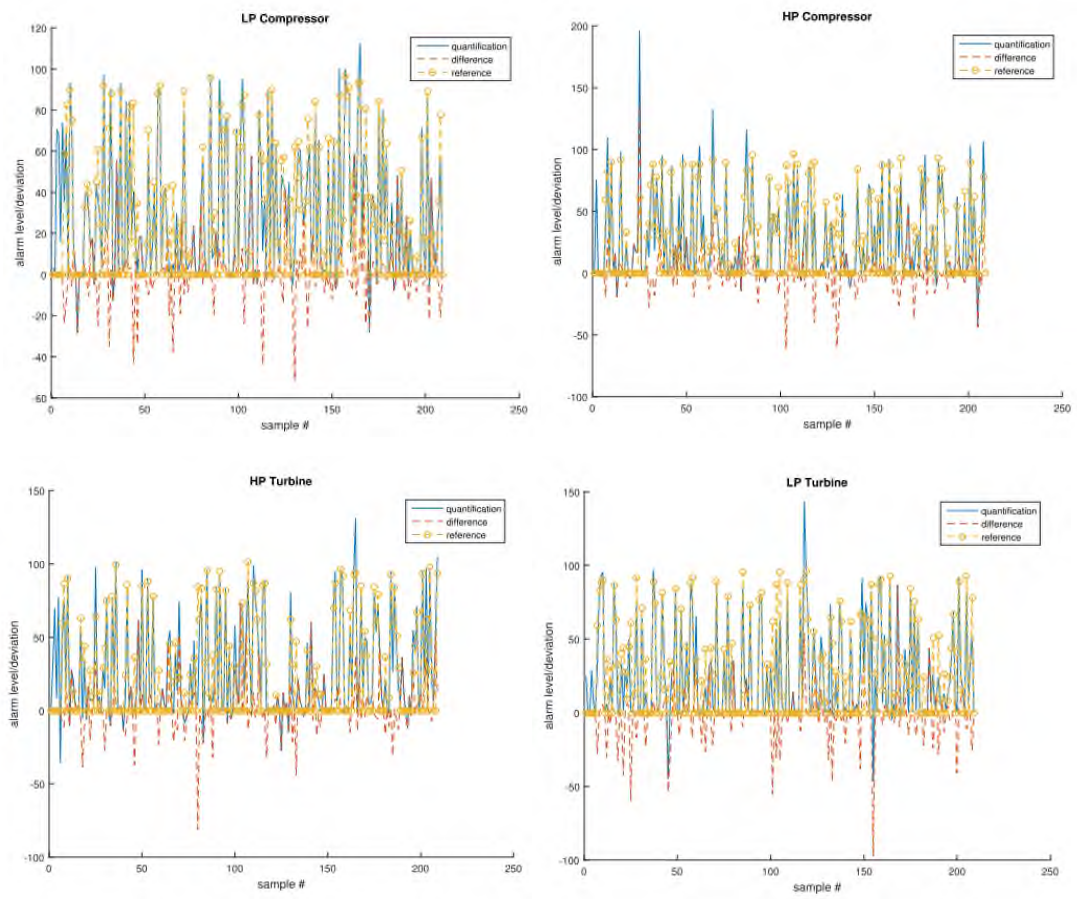

Figure 111 Absolute quantification and deviation from the reference for compressor and turbine components – 2.0% noise no pre-filtering

The classification rate instead, ranges from 76.5% for the LP turbine erosion to 100% on the HP turbine fouling (Table 52). Compared to the quantification the impact on the success rate is lower, but the impact is still around 17%.

Failure type	Success rate
LP comp fouling	99.0%
HP comp fouling	97.8%
HP turb fouling	100%
HP turb erosion	87.0%
LP turb fouling	97.7%
LP turb erosion	76.5%

Table 52 Success rate for the failure classification for random simulation – 2.0% noise no pre-filtering

The results obtained confirm that is possible to quantify and classify single and multiple failures with variable failure magnitude, also with a reduced number of probes, for instance without the power measurement. However, the uncertainty is a bit higher and some results have a lower success rate. Moreover, with 2.0% measurement noise, the target cannot be reached both for the quantification and for the classification, confirming the importance of the KF module.

5.3.2.2 Single Linear Kalman Filter and measurement fusion

The portion analysed in this section is the SLKF that consist of one layer devoted to combine all the information coming from every single measurement and to filter out the measurement noise.

The results of the quantification repeat the same trend observed in the case without pre-filtering ranging within 5 and 96 (Figure 112).

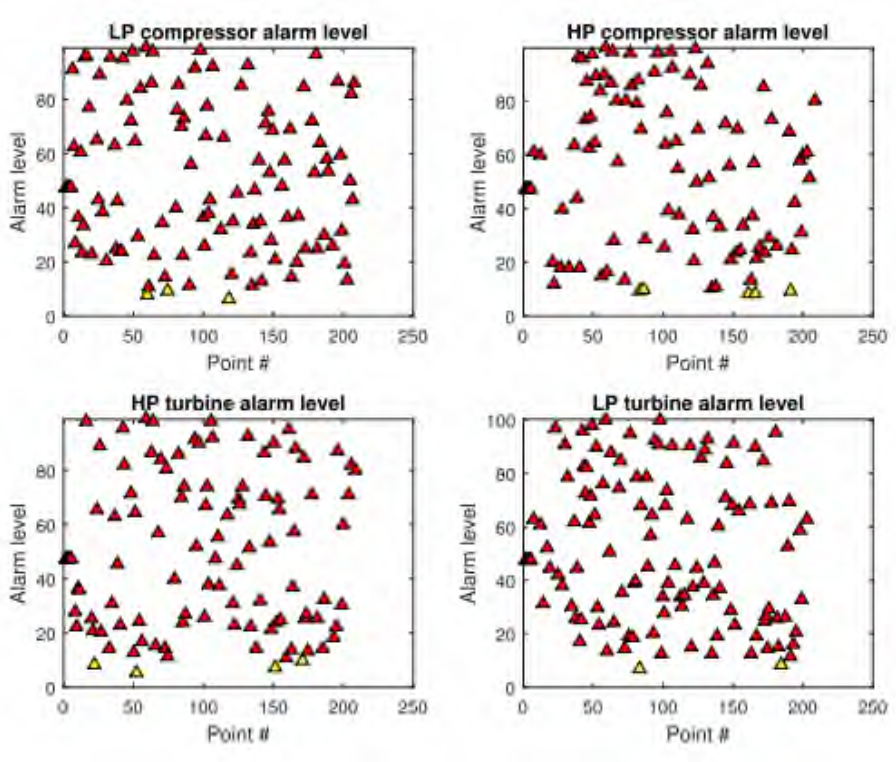

Figure 112 Quantification chart in the case of random deterioration – 0.4% noise SLKF

The quantification success rate ranges from a minimum of 95.0% on the HP turbine to a maximum of 98.5% on the LP turbine (Table 53). This result is in line with the previous simulation without the SLKF and is also in line with the other schemes.

Component	Success rate
LP comp	97.5%
HP comp	96.0%
HP turb	95.0%
LP turb	98.5%

Table 53 Success rate for the failure quantification for random simulation – 0.4% noise SLKF

The classification rate instead, ranges from 96.5% for the HP turbine erosion to 100% on all other failures, except for the LP turbine erosion (Table 54).

Failure type	Success rate
LP comp fouling	100%
HP comp fouling	100%
HP turb fouling	100%
HP turb erosion	96.5%
LP turb fouling	100%
LP turb erosion	98.3%

Table 54 Success rate for the failure classification for random simulation – 0.4% noise SLKF

Looking at the case with 2.0% noise, the quantification range varies between 5 and 110 (Figure 113), which looks already deviating from expectations.

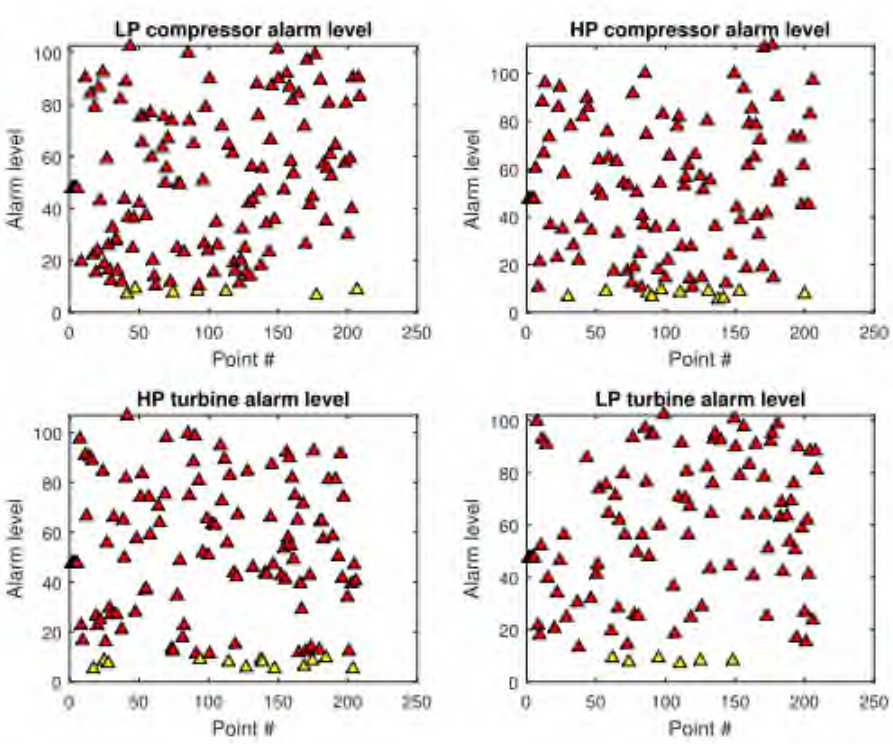

Figure 113 Quantification chart in the case of random deterioration – 2.0% noise SLKF

The resulting success rate varies, from 64.7% on the HP compressor to 73.1% on the LP turbine with an improvement on all the components compared to the case without filtering (Table 55).

Component	Success rate
LP comp	65.7%
HP comp	64.7%
HP turb	67.7%
LP turb	73.1%

Table 55 Success rate for the failure quantification for random simulation – 2.0% noise SLKF

The improvement in the success rate is also reflected in the reduction of the number of deviation spots in Figure 114.

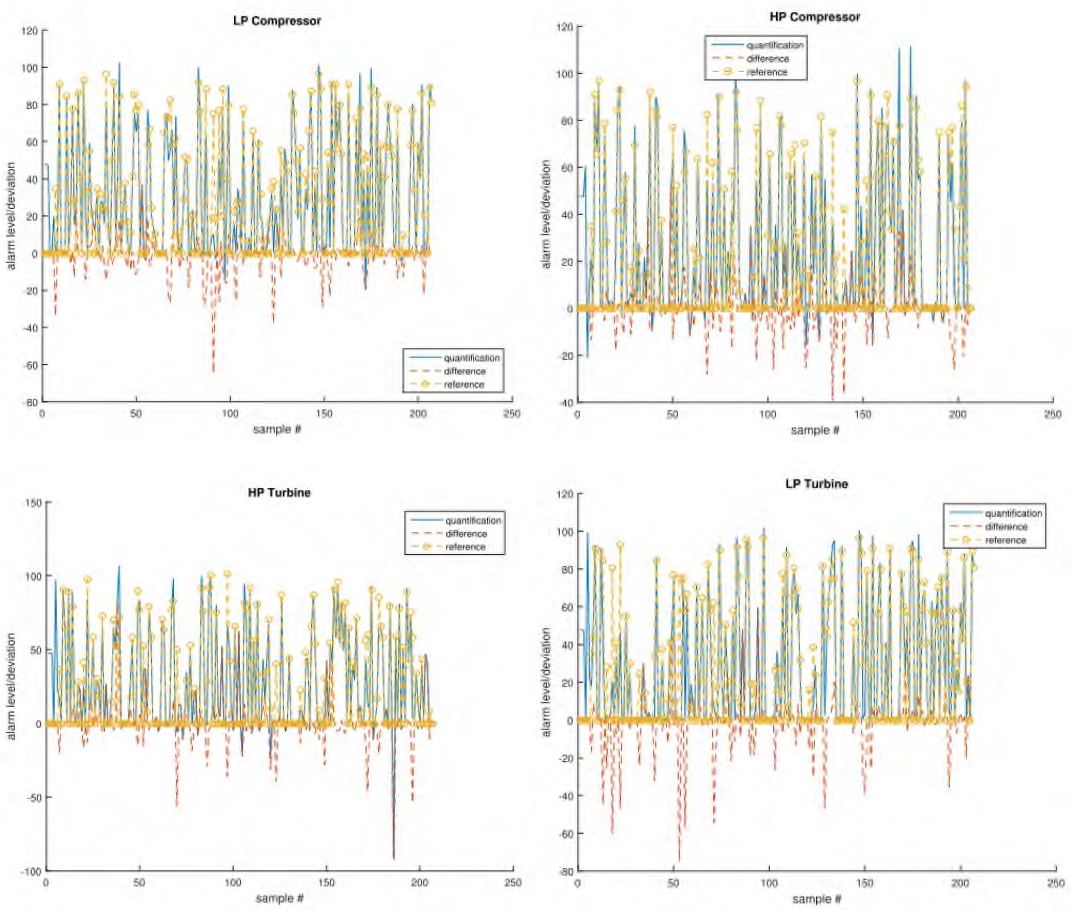

Figure 114 Absolute quantification and deviation from the reference for compressor and turbine components – 2.0% noise SLKF

The classification rate ranges from 83.0% for the LP turbine erosion to 100% on the HP and LP turbine fouling (Table 56). This result is already an improvement compared to the case without KF, even if the target is missed for the LP turbine erosion.

Failure type	Success rate
LP comp fouling	98.7%
HP comp fouling	94.0%
HP turb fouling	100%
HP turb erosion	92.3%
LP turb fouling	100%
LP turb erosion	83.0%

Table 56 Success rate for the failure classification for random simulation – 2.0% noise SLKF

5.3.2.3 Multiple Layer Kalman Filter and measurement fusion

The other scheme analysed in this section is the MLKF that consists of two layers devoted to combine all the information coming from every single measurement and to filter out the measurement noise.

The results of the quantification repeat the same trend observed in the case without pre-filtering ranging within 5 and 96 (Figure 115). As for the case with the SLKF, with the standard level of noise, no improvements are expected since the improvement of the measurement noise is negligible.

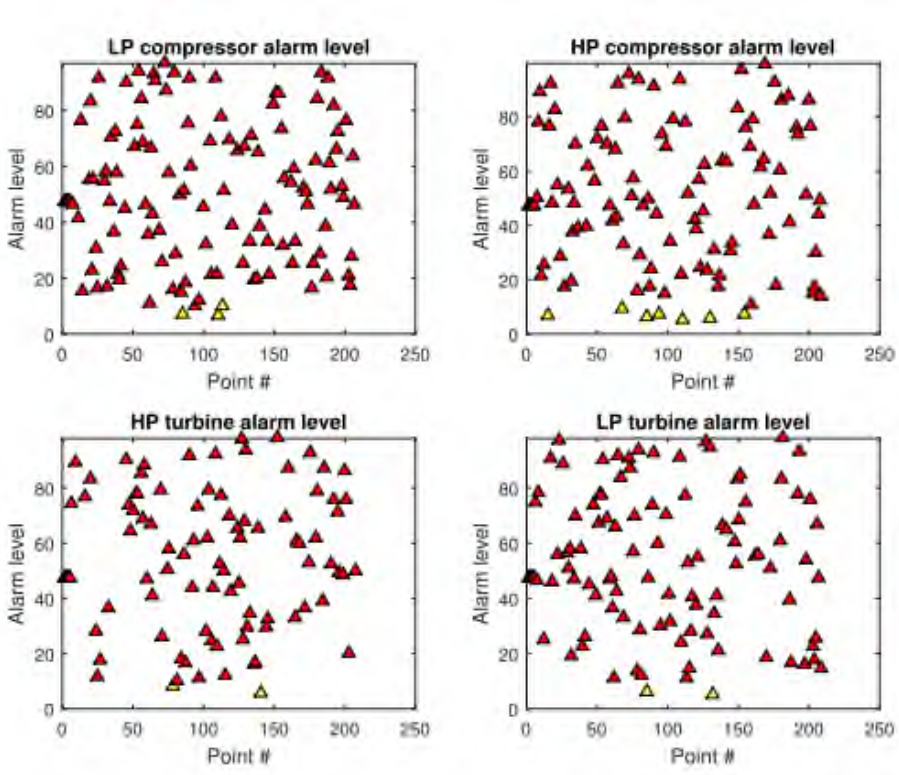

Figure 115 Quantification chart in the case of random deterioration –0.4% noise MLKF

The quantification success rate ranges from a minimum of 95.5% on the LP turbine to a maximum of 99.0% on the LP compressor and HP turbine (Table 57). These results are in line with scheme 1 and scheme 2.

Component	Success rate
LP comp	99.0%
HP comp	97.0%
HP turb	99.0%
LP turb	95.5%

Table 57 Success rate for the failure quantification for random simulation – 0.4% noise MLKF

The classification rate instead, ranges from 95.0% for the HP turbine erosion to 100% on the HP turbine and LP turbine fouling (Table 58). This result is slightly better than the case without KF but is worse than the case without SLKF. However, considered the small variation among the two cases, the deviation can be imputable to the random nature of the problem.

Failure type	Success rate
LP comp fouling	96.8%
HP comp fouling	96.2%
HP turb fouling	100%
HP turb erosion	95.0%
LP turb fouling	100%
LP turb erosion	96.2%

Table 58 Success rate for the failure classification for random simulation – 0.4% noise MLKF

The quantification range with 2.0% noise level ranges between 5 and 105 (Figure 116). This chart is not far from the target and looks in line with the other schemes.

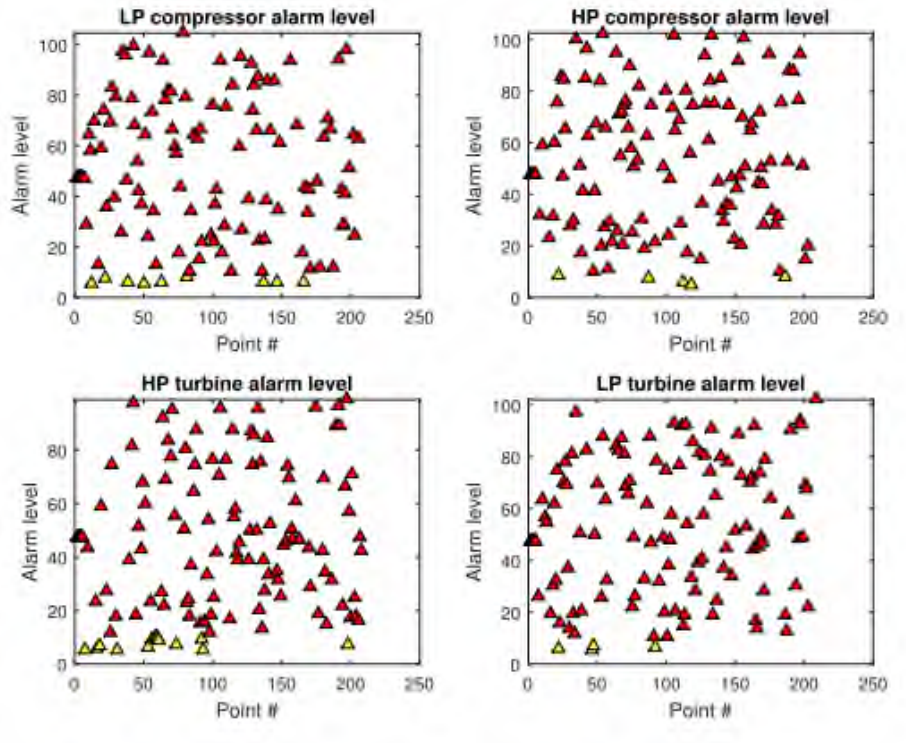

Figure 116 Quantification chart in the case of random deterioration – 2.0% noise MLKF

The resulting success rate varies, from 72.1% on the HP turbine to 80.1% on the LP turbine. This result is in line with scheme 1 and is better than the scheme 2 (Table 59). The message of the quantification, compared to the other schemes, is that the removal of the power measurement from the equation does not change the results. Moreover, the uncertainty introduced by the reference, that relies on the ambient conditions, looks very well compensated by the KF. In fact, scheme 2 has similar results than schemes 1 and 3.

Component	Success rate
LP comp	82.1%
HP comp	74.1%
HP turb	72.1%
LP turb	80.1%

Table 59 success rate for the failure quantification for random simulation – 2.0% noise MLKF

The related deviation is increased compared to the case with nominal noise (Figure 117).

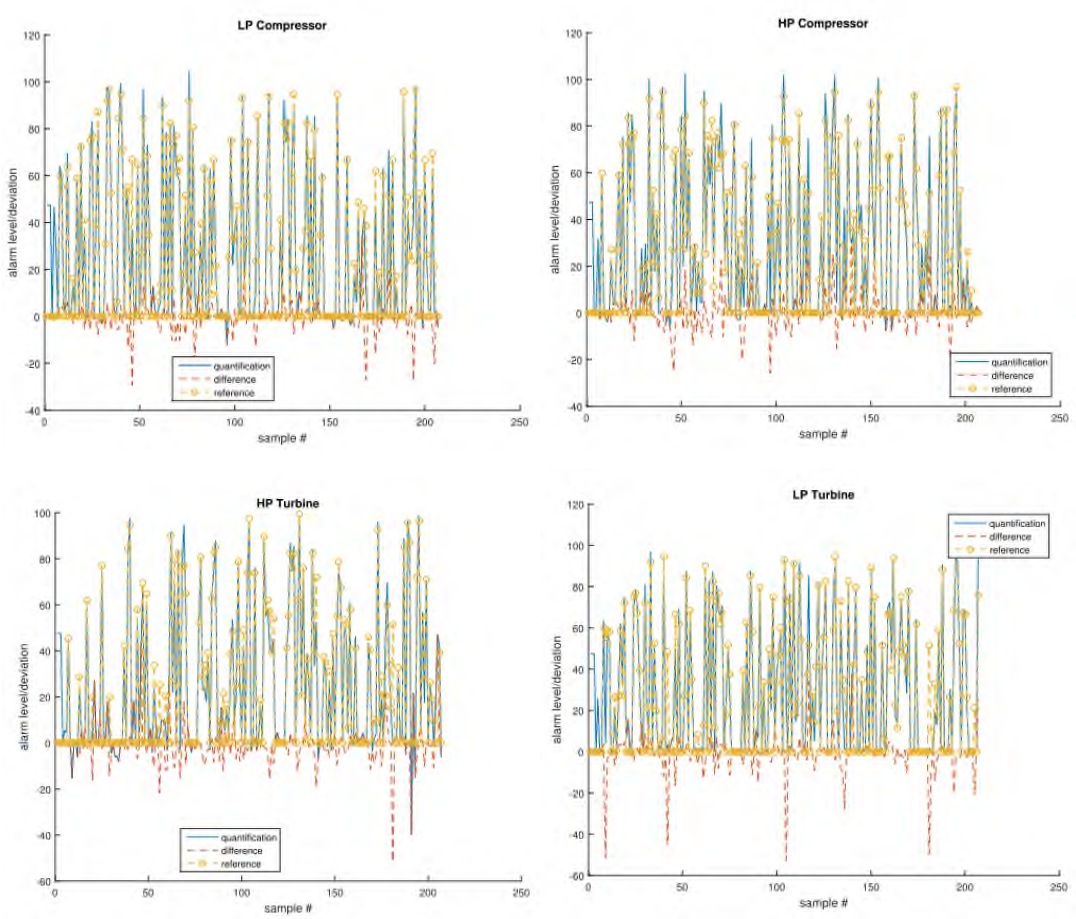

Figure 117 Absolute quantification and deviation from the reference for compressor and turbine components – 2.0% noise MLKF

The classification rate ranges from 84.7% for the LP turbine erosion to 100% on the HP and LP turbine fouling (Table 60). The success rate is in line with the case with SLKF, but worse than the other schemes especially the scheme 2 that has a success rate always above the target.

Failure type	Success rate
LP comp fouling	95.2%
HP comp fouling	95.9%
HP turb fouling	100%
HP turb erosion	94.4%
LP turb fouling	100%
LP turb erosion	84.7%

Table 60 Success rate for the failure classification for random simulation – 2.0% noise MLKF

The results obtained in this section confirm that it is possible to perform the multiple component health estimation together with multiple failure estimation. In fact, even the removal of the power from the measurements may cause an increase of the uncertainty, the overall results are very close to the other schemes and in some cases even better.

5.3.3 Deterioration schedule

The deterioration schedule is also tested for scheme 3. Considered that the ANN is the most important point to be tested, this test is crucial to see if the accuracy is reliable also without the power measurement. As per scheme 1, out of 5000 samples 200 have been extracted until 75% of the entire deterioration. For instance, the reference deterioration on the LP compressor is 3.3% while is 5.1% on the LP turbine.

5.3.3.1 Combined training

The inputs of the combined training are identical to those used in scheme 1 and they include the full combination of deterioration plus the schedule of 5000 reflecting a possible normal operation of a gas turbine.

Once the efficiencies and physical values deviations are processed by the fuzzy logic, they deal with the quantification described in section 4.7. Based on the deterioration selected for this problem, 3.3% on the LP compressor and 5.1% on the LP turbine, the prediction is to have 43 as alarm level on the LP compressor and 66 on the LP turbine (Figure 118). The results obtained are in line with scheme 1 and scheme 2 and they confirm that a correct component health estimation without the power measurement is possible.

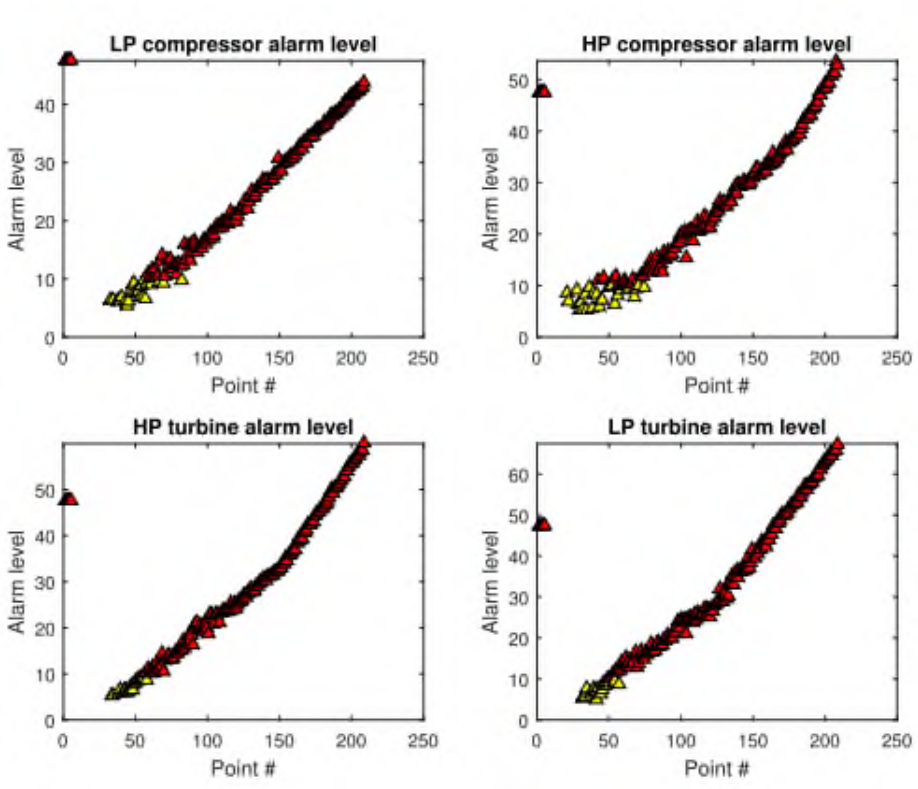

Figure 118 Quantification chart in the case of scheduled deterioration

The resulting quantification success rate is 100% on all the components (Table 61). This last result is well above the target of 90% and is in line with the other schemes.

Component	Success rate
LP comp	100%
HP comp	100%
HP turb	100%
LP turb	100%

Table 61 Success rate for the failure quantification for scheduled simulation – 0.4% noise MLKF

The high success rate is confirmed by the small deviations between the predicted component health status and the reference value (Figure 119). Some more deviations are present on the first samples of the HP compressor where the failure magnitude is low and the measurement uncertainties, together with the prediction uncertainties play a bigger role.

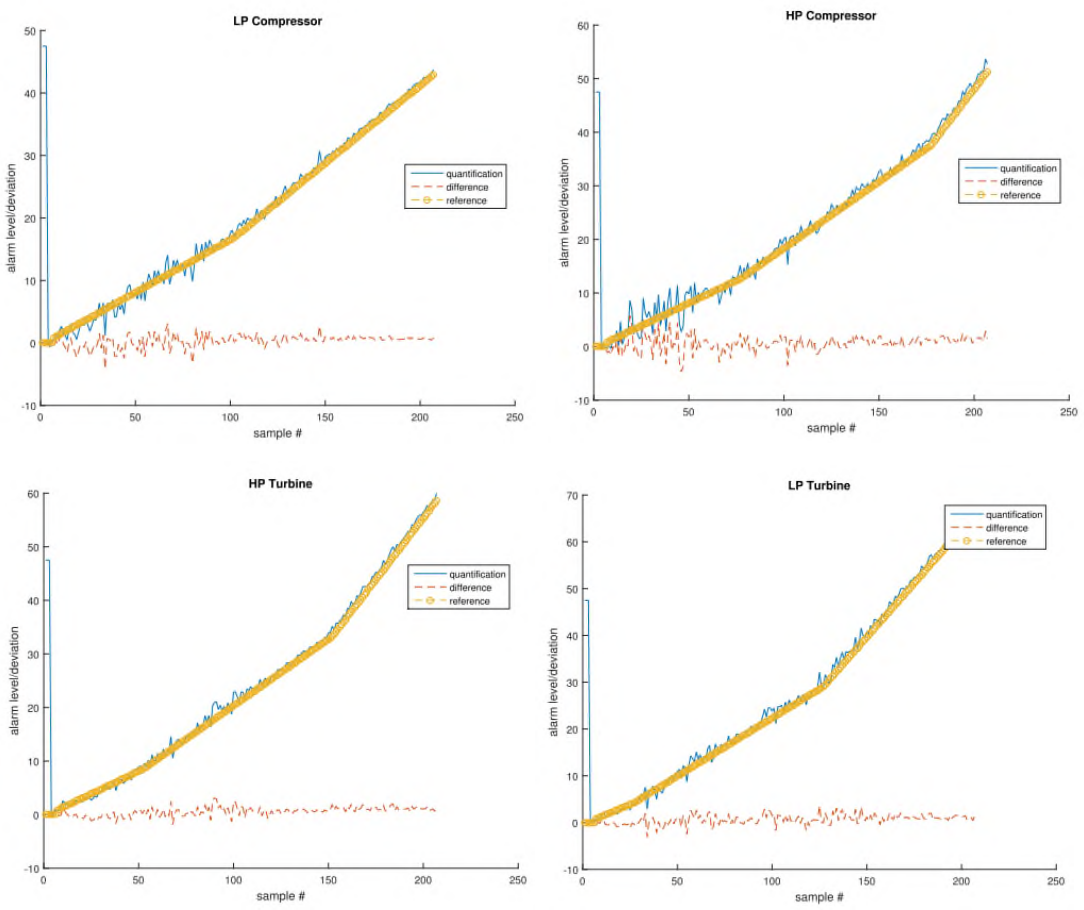

Figure 119 Absolute quantification and deviation from the reference for compressor and turbine components – 0.4% noise MLKF

The next step to be tested is classification. The resulting success rate ranges from 93.1% on the LP compressor to 95.5% on the HP turbine fouling (Table 62). Despite the very good results obtained, two points are incorrectly classified and visible in the HP turbine (Figure 120 and Figure 121). The root cause here is not easy to follow since there are no visible signs of deviations nor of uncertainties. However, since the points have low magnitude, the deviation can be caused by the measurement noise.

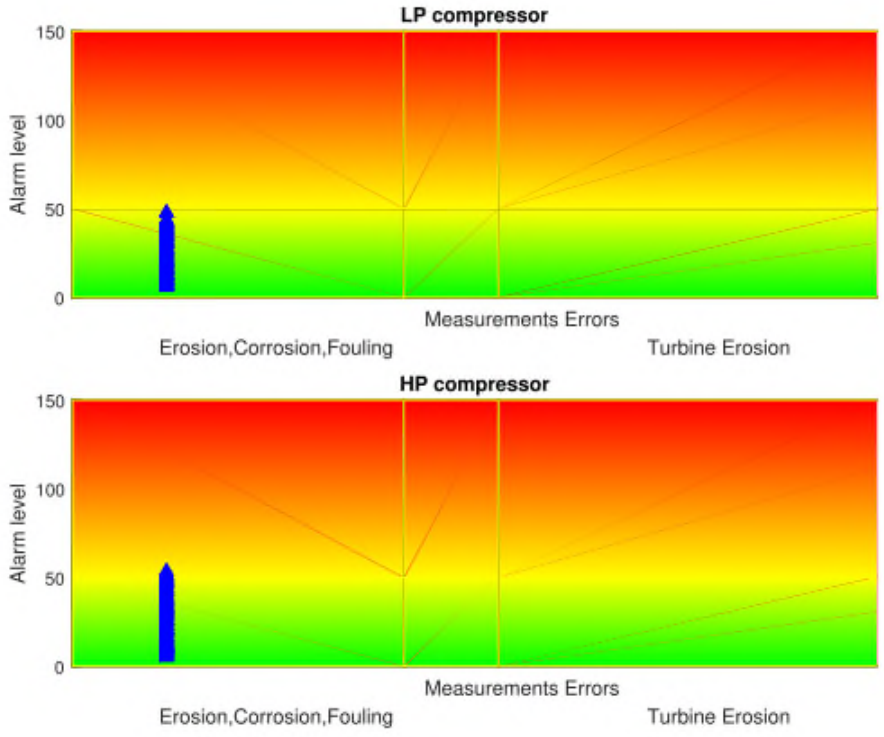

Figure 120 Classification chart for the compressor

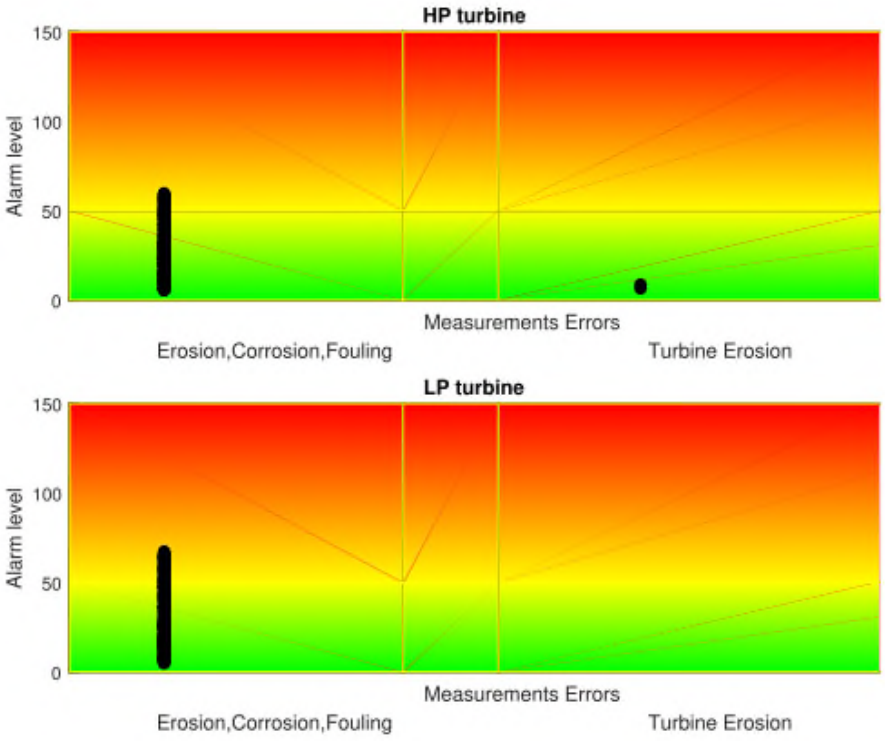

Figure 121 Classification chart for the turbine

Failure type	Success rate
LP comp fouling	93.1%
HP comp fouling	94.1%
HP turb fouling	95.5%
LP turb fouling	95.0%

Table 62 Success rate for the failure classification for scheduled simulation – 0.4% noise MLKF

These results confirm the feasibility of the failure characterization also with scheme 3 without the power used as a measurement. Additionally, it is interesting to notice that the scheme 3 is performing slightly better than scheme 1, showing that the power measurement is not essential.

6.0 Results with the test engine

The results reported so far try to stress as much as possible the methodology in order to validate it. However, the engine considered for all these tests is unique and the entire methodology is suited around it. Therefore, the methodology has been tested with one additional engine, whose data have been provided by the Cranfield power and propulsion department based on open-source information. The intent is to see if the methodology is robust enough to provide the same results quality with other engines, or if it's only working for a single-engine.

The engine code provided has the same architecture as the base engine used for the simulation, but it has a higher power output of circa 43 MWe and a different pressure ratio distribution. For this engine, three different tests are available: constant deterioration, random deterioration and deterioration schedule. However, the preferred deterioration that is selected for this test is random. In fact, the constant deterioration is mainly focused on the measurements error handling that is a peculiarity of the measurements rather than the engine as such. Opposite, the deterioration schedule is a good test to predict the long-term behaviour of an engine but is limited to a single type of failure and it does not provide a full picture.

The schemes available for this additional engine are still three: scheme 1 including the power among the possible measurements and using Turbomatch as reference model; scheme 2 using the ANN to predict both the reference values and the deteriorated values; scheme 3 excluding the power from the measurements and using Turbomatch as the reference model. It is decided to use all the schemes for the comparison to analyze the effect of these on a different engine.

The steps followed to reach the full detection of the multiple failures, starting from the engine code are similar to those used to set up the reference engine at the birth of the methodology and include tuning of all the methodology main components:

a. Set up of the reference data used as training for the ANN. The number of cases included in the training is identical as these used for the reference engine. The main difference here is on the engine model;
b. The KF section for the data filtering remained unchanged;
c. The ANN has been re-trained with the new data. This is a must for every new engine since the prediction of the deteriorated data relies on the previously supplied data;
d. The neuro-fuzzy logic for the data quantification has been re-trained since there are different mechanism between the deltas on efficiency and measurements and the severity of the failure. Moreover, to prove the flexibility of the NFL module and to test the engine on data closer to the possible reality, the ratio between the alarm level and the deterioration rate has been changed. In the case, 100 corresponds to 5% fouling and 2.5% erosion;
e. The fuzzy logic for the classification instead, remained unchanged since the characteristics of the type of the failure detected are identical among the two engines.

The data filtering options selected for the validation of this engine is the MLKF which has demonstrated being the most promising with the reference engine. On the other hand, since all the schemes demonstrated having drawbacks and advantages compared to the others, they have been all selected and compared.

The test used for the validation of the methodology against this other engine is the random deterioration which has demonstrated, with the reference engine, being the most demanding. As before, the number of points that will be detected is 203, taken randomly from the simulated values. The results

confirm what has been reached by the reference engine showing a success rate above the target for scheme 1 and scheme 2 and close to the target for scheme 3, where the minimum success rate is 88.1%. The best results are achieved with scheme 2, where the success rate is even above 93.0% (Table 63).

Quantification - Constant deterioration 0.4% measurement noise				
		Scheme 1	Scheme 2	Scheme 3
MLKF	LP comp	94.0%	93.0%	90.5%
	HP comp	98.5%	100%	94.5%
	HP turb	91.5%	95.5%	88.1%
	LP turb	97.0%	96.0%	91.5%

Table 63 Quantification with constant deterioration, 0.4% measurement noise, MLKF

The classification over the 203 random points over the 24 possible combinations, partly confirm the quality of the results reached in the quantification phase with results above 79.7%. The results, in fact, are above the target for the HP and LP compressor fouling and for the HP and LP turbine fouling. However, the turbine erosion falls below the target both on HP and LP, with the best result on scheme 3, where the success rate is above 85.0% (Table 64). Since the classification FL has not been changed, some additional tuning would further improve the status of the turbine erosion classification rate.

Even if the target is not fully reached for the turbine fouling, this test confirms the capabilities of the methodology on detecting single and multiple failures and on estimating the health status of each component.

Classification - Random deterioration 0.4% measurement noise				
		Scheme 1	Scheme 2	Scheme 3
MLKF	LP comp fouling	96.8%	91.7%	96.8%
	HP comp fouling	100%	100%	98.2%
	HP turb fouling	100%	100%	100%
	HP turb erosion	86.4%	86.4%	88.9%
	LP turb fouling	100%	100%	100%
	LP turb erosion	79.7%	81.4%	85.0%

Table 64 Classification with constant deterioration, 0.4% measurement noise, MLKF

7.0 Analysis of the results

The previous chapters presented the results that the methodology was able to achieve under different conditions. The objective is to correctly quantify, establishing the component health, and classify assessing the type of failure, the malfunction of the gas turbine. To reach this goal a methodology composed by KF, ANN, NFL and FL has been established. This combination, together with the ambition and the flexibility pretended for this methodology, lead to three types of schemes:

- Scheme 1: is the base methodology including the KF for the data filtering, the ANN for the component efficiency prediction, the Neuro-Fuzzy for the failure quantification and the engine health estimation and the FL for the failure classification. With this scheme, the reference for the delta is determined by the thermodynamic module calculated with Turbomatch;
- Scheme 2: in scheme 2 the components and the sequence of the methodology are identical to the scheme 1. However, since Turbomatch wanted to be excluded from the diagnostics. The ANN then is providing the deteriorated value together with its reference;
- Scheme 3: in scheme 3, the components of the methodology are the same as for scheme 1. In this case, the power measurement is excluded to make the methodology suitable for configurations where the power measurement is not available.

The conditions established to test the robustness and the capabilities of the methodology are 3: constant deterioration, random deterioration and deterioration schedule. The objectives of each test are detailed in Table 65.

The objectives of each test		
Constant deterioration Variation of measurement noise	Random deterioration with measurement noise	Deterioration schedule
The capability of the methodology to quantify and classify the GT malfunctions	Stress test to determine the robustness of the methodology on the quantification and classification	Test the methodology with conditions close to reality
See the capability to isolate the measurements noise	See the capability to isolate the measurements noise	Check the effect of the ANN training on the prediction

Table 65 Objectives of each test

The tests conducted with the constant deterioration and nominal noise, shown that it is possible to reach a quantification success rate of 100% with all the schemes. The best result is obtained with the configuration without the KF or with the SLKF while the configuration with MLKF has a success rate of 99.5% with scheme 1 (Table 66).

Quantification - Constant deterioration 0.4% measurement noise				
		Scheme 1	Scheme 2	Scheme 3
None	LP comp	100%	100%	100%
	HP comp	100%	100%	100%
	HP turb	100%	100%	100%
	LP turb	100%	100%	100%
SLKF	LP comp	99.5%	100%	100%
	HP comp	99.5%	100%	100%
	HP turb	99.5%	100%	100%
	LP turb	99.5%	100%	100%
MLKF	LP comp	99.5%	100%	100%
	HP comp	99.5%	100%	100%
	HP turb	99.5%	100%	100%
	LP turb	99.5%	100%	100%

Table 66 Quantification with constant deterioration and 0.4% measurement noise

In the same way, the results of the classification shown that is possible to categorize the failures with a success rate of 100% with all the schemes. The best results are achieved with the SLKF and with the MLKF. Without the KF, the success rate is decreased on the HP compressor leading to a success rate of 93.0% with scheme 1, 93.6 with the scheme 2 and 95.6% with scheme 3 (Table 67).

Classification - Constant deterioration 0.4% measurement noise				
		Scheme 1	Scheme 2	Scheme 3
None	LP comp fouling	100%	100%	100%
	HP comp fouling	93.0%	93.6%	95.6%
	HP turb erosion	100%	100%	100%
	LP turb erosion	100%	100%	100%
SLKF	LP comp fouling	100%	100%	100%
	HP comp fouling	100%	100%	100%
	HP turb erosion	100%	100%	100%
	LP turb erosion	100%	100%	100%
MLKF	LP comp fouling	100%	100%	100%
	HP comp fouling	100%	100%	100%
	HP turb erosion	100%	100%	100%
	LP turb erosion	100%	100%	100%

Table 67 Classification with constant deterioration and 0.4% measurement noise

At this point, the methodology proved to be able to correctly estimate the engine health status and to correctly classify the type of failure imposed, that in this case was static. By increasing the level of noise, the results showed how the methodology has been capable of reacting. Provided the 2.0% noise level, the SLKF and the MLKF imply an increase of the success rate up to 39 points. The scheme that is showing the best results is the scheme 3 with MLKF who is leading to a success rate between 84.6% and 94.0% (Table 68). However, the same scheme is reaching the lower success rate if no filter is applied, proving the influence of the power measurement for this methodology with and without the KF. If the power, in fact, is supporting on achieving better results without the KF, on the other hand, it reveals itself as a driver even after the filtering. Being the power a single measurement, in fact, it cannot be processed by the KF so established.

Quantification - Constant deterioration 2.0% measurement noise				
		Scheme 1	Scheme 2	Scheme 3
None	LP comp	76.9%	59.7%	72.6%
	HP comp	58.8%	51.7%	67.7%
	HP turb	66.3%	79.6%	78.6%
	LP turb	63.3%	54.2%	39.8%
SLKF	LP comp	82.4%	82.6%	94.5%
	HP comp	75.4%	72.6%	90.5%
	HP turb	83.9%	97.0%	94.5%
	LP turb	73.4%	78.1%	78.6%
MLKF	LP comp	86.4%	78.1%	94.0%
	HP comp	82.4%	70.6%	92.0%
	HP turb	91.0%	97.5%	96.0%
	LP turb	72.4%	80.1%	84.6%

Table 68 Quantification with constant deterioration and 2.0% measurement noise

As already reported in the result section, the classification of this single failure does not show the full capability of the methodology but provides a good indication of the effect of the measurement noise on the success rate. With this environment, the configuration that is showing the best results is the scheme 2 with SLKF showing success rates from 94.0% to 100% above 90.0% declared as the target (Table 69). This second section confirms the effectiveness of the methodology also with the presence of measurement issue – in this case, noise. The key contributor is the KF which is consistently improving the success rate. The results are still below the target of 90.0% in the quantification part, but above 90.0% on three components. Moreover, the results are above 90% until 1.6% reference noise.

Classification - Constant deterioration 2.0% measurement noise				
		Scheme 1	Scheme 2	Scheme 3
None	LP comp fouling	99.0%	100%	100%
	HP comp fouling	97.0%	97.0%	95.6%
	HP turb erosion	94.0%	91.1%	96.1%
	LP turb erosion	86.0%	91.1%	89.7%
SLKF	LP comp fouling	100%	100%	100%
	HP comp fouling	94.9%	94.0%	94.5%
	HP turb erosion	92.9%	97.5%	98.5%
	LP turb erosion	90.9%	97.5%	89.6%
MLKF	LP comp fouling	100%	100%	100%
	HP comp fouling	92.4%	94.0%	100%
	HP turb erosion	92.9%	98.0%	98.0%
	LP turb erosion	93.4%	97.0%	89.6%

Table 69 Classification with constant deterioration and 2.0% measurement noise

Comparing the three schemes with the random deterioration and with nominal noise the quantification success rate is above the target with all the schemes. The best results are achieved by the cases with the KF, either SLKF or MLKF, but there is not a clear best combination among those since the success rate is overall very high (Table 70).

Quantification - Random deterioration 0.4% measurement noise				
		Scheme 1	Scheme 2	Scheme 3
None	LP comp	99.0%	99.0%	98.0%
	HP comp	96.5%	96.0%	96.5%
	HP turb	100%	93.5%	97.0%
	LP turb	98.0%	98.5%	97.5%
SLKF	LP comp	99.0%	99.0%	97.5%
	HP comp	97.5%	98.0%	96.0%
	HP turb	98.5%	98.5%	95.0%
	LP turb	98.5%	98.0%	98.5%
MLKF	LP comp	97.0%	100%	99.0%
	HP comp	92.0%	98.5%	97.0%
	HP turb	98.5%	97.5%	99.0%
	LP turb	99.0%	97.5%	95.5%

Table 70 Quantification with random deterioration and 0.4% measurement noise

Looking at the classification the results provided different answers among the schemes. One reason can be found in the nature of the test case, that does not allow to compare the three schemes directly, another one in the different characteristics of the schemes. The classification rate can fall below 90% only with scheme 2 and, surprisingly, with the SLKF applied. This depends on the additional uncertainty created by the reference that moved

from Turbomatch to the ANN. However, with the scheme 1 the classification rate is above 95.1% for the MLKF and above 94.3% for the SLKF. Scheme 2 instead, have a success rate above 93.2% when the MLKF is applied and also without any filter with a minimum success rate of 93.8% (Table 71).

Classification - Random deterioration 0.4% measurement noise				
		Scheme 1	Scheme 2	Scheme 3
None	LP comp fouling	97.1%	93.9%	92.9%
	HP comp fouling	96.8%	97.1%	98.0%
	HP turb fouling	100%	100%	100%
	HP turb erosion	100%	93.8%	90.2%
	LP turb fouling	100%	100%	100%
	LP turb erosion	95.5%	96.7%	94.1%
SLKF	LP comp fouling	94.3%	98.6%	100%
	HP comp fouling	96.7%	97.1%	100%
	HP turb fouling	100%	100%	100%
	HP turb erosion	96.6%	89.8%	96.5%
	LP turb fouling	100%	100%	100%
	LP turb erosion	95.3%	89.1%	98.3%
MLKF	LP comp fouling	98.3%	93.7%	96.8%
	HP comp fouling	97.9%	98.2%	96.2%
	HP turb fouling	100%	100%	100%
	HP turb erosion	95.1%	98.2%	95.0%
	LP turb fouling	100%	100%	100%
	LP turb erosion	96.9%	93.2%	96.2%

Table 71 Classification with constant deterioration and 0.4% measurement noise

The results with 2.0% reference noise and constant deterioration show that going from scheme 1 to scheme 3, the success rate gradually decreases. This means that the noise plays a bigger role, once the available information is less – power measurement removed in scheme 3 – or the uncertainty is higher – reference predicted by the ANN instead of calculated by the thermodynamic model in scheme 2. However, with the random simulation, there is no clear outstanding since the results are very close (Table 72). This depends on the nature of the problem, that being random, does not facilitate the direct comparison. Moreover, in absolute terms, the quantification success rate obtained with the random test is lower than the rate obtained with the constant deterioration.

In the same way, the classification obtained with the SLKF and MLKF is worse compared to the constant case. The first reason is that the problem to

be solved is more severe and the KF ends up being less effective. The second reason is that, due to the random nature of the problem, both the SLKF and MLKF parameters have been modified to increase the response of the KF while moving from one sample to the other, at the price of a reduction of the measurement noise reduction capabilities. It is also interesting to notice that with the constant deterioration, scheme 3 is the top performer, while with the random simulation there's no clear outstanding scheme. This message remarks the importance of an extended testing set, to make robust conclusions for a complete methodology or for variants of it.

Quantification - Random deterioration 2.0% measurement noise				
		Scheme 1	Scheme 2	Scheme 3
None	LP comp	56.7%	58.7%	60.2%
	HP comp	56.2%	62.7%	53.2%
	HP turb	59.2%	56.7%	51.7%
	LP turb	74.1%	72.6%	64.7%
SLKF	LP comp	73.1%	73.6%	65.7%
	HP comp	69.2%	70.1%	64.7%
	HP turb	70.6%	73.6%	67.7%
	LP turb	87.6%	78.6%	73.1%
MLKF	LP comp	76.1%	76.1%	82.1%
	HP comp	73.6%	70.1%	74.1%
	HP turb	70.1%	65.7%	72.1%
	LP turb	83.1%	78.6%	80.1%

Table 72 Quantification with random deterioration and 2.0% measurement noise

The classification success rate, as also observed for the constant deterioration case, is less affected by the measurement noise, compared to the quantification results. However, the success rate reaches 84.7% with scheme 3 if the MLKF is applied. This result is improved to 94.4% with scheme 2 without the filter. Remarkable results, still with scheme 2, are also obtained with the MLKF where the success rate is above 91.7% (Table 73). Translated, this means that the classification section can reach the target also with the random simulation and 2.0% measurement noise.

This section remarks that the methodology can quantify and classify the 24 random deterioration, with 6 different types of failure and with variable deterioration magnitude with a success rate above 90% for the 0.4% reference measurement noise. The result is decreased to 70.1% (reference scheme 2)

in terms of quantification but can stay above 90.0% in terms of classification if 2.0% measurement noise is applied. This means that the goal is not fully achieved, but there is an improvement of more than 10% introduced by the KF module.

Classification - Random deterioration 2.0% measurement noise				
		Scheme 1	Scheme 2	Scheme 3
None	LP comp fouling	99.1%	94.4%	99.0%
	HP comp fouling	100.0%	97.8%	97.8%
	HP turb fouling	100.0%	100%	100%
	HP turb erosion	86.5%	94.7%	87.0%
	LP turb fouling	100.0%	100%	97.7%
	LP turb erosion	88.3%	96.0%	76.5%
SLKF	LP comp fouling	97.1%	95.8%	98.7%
	HP comp fouling	100.0%	97.1%	94.0%
	HP turb fouling	100.0%	100%	100%
	HP turb erosion	92.5%	93.8%	92.3%
	LP turb fouling	100.0%	100%	100%
	LP turb erosion	89.5%	86.0%	83.0%
MLKF	LP comp fouling	97.7%	98.6%	95.2%
	HP comp fouling	100.0%	98.5%	95.9%
	HP turb fouling	96.0%	100%	100%
	HP turb erosion	90.7%	91.7%	94.4%
	LP turb fouling	100.0%	100%	100%
	LP turb erosion	88.9%	93.9%	84.7%

Table 73 Classification with random deterioration and 2.0% measurement noise

The random test, being the stress test, has also been taken as the base for the speed comparison among the schemes and in absolute terms. The values that are compared are per sample and divided into the contributors of the methodology: the KF section used to pre-process the data, the ANN used to predict the performance values of the GT, the calculation section done through Turbomatch used to determine the reference values and the FL block that includes the NFL for the component health estimation and the FL for the failure classification. The total time varies depending on the type of scheme used and on the configuration of the KF. Scheme 2, for instance, does not need the Turbomatch calculation as a reference and can process one sample in less than half a second while including the MLKF. The other schemes instead can process one point in a maximum of 1.7 s (Table 74). It can be noticed that the MLKF is taking a bit more time than the SLKF for instance around 0.44 s per sample vs 0.16 s per sample. Overall, the methodology shows that can

process one sample, depending on the option used and with the KF active, between 0.23 s and 1.7 s. This makes the methodology suitable for the online diagnostics as targeted.

The execution time of the schemes – Time per sample				
		Scheme 1	Scheme 2	Scheme 3
		[s]	[s]	[s]
None	KF	-	-	-
	ANN	1.8E-03	1.8E-03	1.0E-03
	Turbomatch	9.1E-01	-	9.0E-01
	FL	5.4E-02	5.8E-02	5.0E-02
	Total	**9.7E-01**	**6.1E-02**	**9.5E-01**
SLKF	KF	1.6E-01	1.7E-01	2.1E-01
	ANN	1.7E-03	2.0E-03	1.2E-03
	Turbomatch	9.0E-01	-	9.0E-01
	FL	5.8E-02	5.8E-02	6.7E-02
	Total	**1.1E+00**	**2.3E-01**	**1.2E+00**
MLKF	KF	4.4E-01	3.4E-01	3.5E-01
	ANN	1.7E-03	2.1E-03	1.0E-03
	Turbomatch	1.2E+00	-	1.0E+00
	FL	5.7E-02	5.9E-02	5.3E-02
	Total	**1.7E+00**	**4.0E-01**	**1.4E+00**

Table 74 Execution time of the schemes

The last comparison from the result section is with the deterioration schedule that is done with the ANN training including the random data plus the deterioration profile. Comparing the three schemes instead it can be seen that the component health estimation is almost perfect with all the schemes (Table 75).

In this section though, the block under investigation is the ANN and in particular the preparation phase. It is clear that having extra data would lead to a more detailed prediction, but there's also a risk of overfitting if too many data on a single condition are given. On the other hand, the training with data integrated during the process and not properly treated may lead to a decrease of the prediction accuracy. In this case, it can be clearly seen that the architecture built is robust enough to provide data with good quality able to train the ANN to make correct predictions.

From this comparison, it is very clear how the ANN prediction is dependent upon the quality of the given data. In particular, dedicated training is always key for good diagnostics. This is also a clear demonstration that the methodology can be used in real cases, while data from a gas turbine are supplied, provided that all the necessary figures are previously integrated into the ANN training.

Quantification – Deterioration schedule 0.4% measurement noise				
		Scheme 1	Scheme 2	Scheme 3
Combined training	LP comp	100%	100%	100%
	HP comp	100%	100%	100%
	HP turb	100%	100%	100%
	LP turb	100%	100%	100%

Table 75 Quantification with scheduled deterioration and 0.4% measurement noise

The second parameter to be tested is the classification success rate. In this case, the success rate is above the target and is very similar for all the configurations (Table 76). The only result below the target is recorded on the HP turbine fouling where the success rate falls at 89.1%. This result though is very close to the target.

Classification – Deterioration schedule 0.4% measurement noise				
		Scheme 1	Scheme 2	Scheme 3
Combined training	LP comp fouling	93.6%	94.6%	93.1%
	HP comp fouling	97.0%	95.5%	94.1%
	HP turb fouling	89.1%	94.1%	95.5%
	LP turb fouling	93.6%	94.6%	95.0%

Table 76 Classification with scheduled deterioration and 0.4% measurement noise

This last point confirms that the methodology can reach the target of 90% success rate both for the quantification and for the classification also in a pseudo-real environment.

Finally, the tests on another engine, support the goodness of the methodology that is capable of achieving the quantification success rate above 90%. The only small gap is present on scheme 3, where the success rate falls down to 88.1% on the HP turbine (Table 77).

Quantification - Random deterioration 0.4% measurement noise				
		Scheme 1	Scheme 2	Scheme 3
MLKF	LP comp	94.0%	93.0%	90.5%
	HP comp	98.5%	100%	94.5%
	HP turb	91.5%	95.5%	88.1%
	LP turb	97.0%	96.0%	91.5%

Table 77 Quantification with constant deterioration and 0.4% measurement noise

The same quality of the results has been achieved with the classification even if the success rate of the HP turbine erosion and of the LP turbine erosion falls below 90% (Table 78). There, a dedicated KF with additional tuning would certainly improve the results. This confirms, once again, that the methodology is suitable for the diagnostics of single and multiple failures.

Classification - Random deterioration 0.4% measurement noise				
		Scheme 1	Scheme 2	Scheme 3
MLKF	LP comp fouling	96.8%	91.7%	96.8%
	HP comp fouling	100%	100%	98.2%
	HP turb fouling	100%	100%	100%
	HP turb erosion	86.4%	86.4%	88.9%
	LP turb fouling	100%	100%	100%
	LP turb erosion	79.7%	81.4%	85.0%

Table 78 Classification with constant deterioration and 0.4% measurement noise

8.0 Summary and Conclusion

8.1 The accomplishment of the objectives

As already reported in section 3.6 the objectives set for the research are:

a. Detection of single and multiple failures also in the presence of measurement issues with a success rate above 90%;
b. Combine the component failure isolation – single and multiple failures – with the component health estimation;
c. Establish a methodology able to deal with multiple failures and measurement issues while working online;
d. Combine the multiple measurements available for each location;
e. Test the methodology under different conditions to prove its robustness.

It has to be reminded that the first objective is one of the main open points in literature since it is challenging to be able to detect multiple failures while the signals are disturbed by measurement issues. To achieve the first objective a methodology composed by KF, ANN, NFL and FL have been set up. Out of this scheme, three variances have been built and the results have been compared. The results with the nominal noise – 0.4% - shown that the methodology is capable of detecting among 24 different types of failure with a variable magnitude that can randomly vary between 0.15% and 7.4% with a success rate above the target both for the component health estimation, quantification, and the classification of the type of failure. The methodology proved also to be able to deal with measurement errors as it is capable of keeping the quantification and classification success rate above 90% with measurement noise up to 1.6%, four times more than the nominal one. Some

additional work and some improvements are necessary when the level of noise reaches 2.0%. In this case, in fact, the success rate falls below 90% settling to a minimum of 70.1% in the case of random simulation for the quantification (reference scheme 1). The situation is better on the classification side where a result of 88.9% is reached (reference scheme 1).

It is implicit in the quantification success rate, that a component health estimation has been set up in the methodology. The dedicated section responsible for that is the neuro-fuzzy that relates the efficiency reduction to the component health status. The health status is set up in a range between 0 and 100 where 0 correspond to a new component and 100 to the maximum established deterioration. The alarm level has been set up based on literature and does not necessarily correspond to the real health status of the engine, since every engine may react differently to the failure and since there is no engine specific data that justifies that. As already clarified in section 4.5 the neuro-fuzzy accounts for the mutual interaction of one component to another. For instance, in the single failure, it considers a 1 to 1 relationship among efficiency decay and component health estimation, while in the multiple failures it accounts for the effect of one component on another. To make this possible a dedicated neuro-fuzzy is set up at each component and tuned considering each of the 24 possible combinations. By quantifying with a success rate above 90% (reference scheme 1 with 0.4% reference noise tested against the random deterioration), the objective is fulfilled. Once the health level is established the methodology proved to correctly classify the 24 combinations of failure and the 6 types of failure with a success rate above 90% (reference scheme 1 with 0.4% reference noise tested against the random deterioration). Both results have been folded into a graph with traffic light colour code able to provide direct feedback to the user on the status of the engine component. The graph gives a clear indication and removes errors related to a possible wrong data interpretation.

The next point targeted was is online processing of the data. The requirement is to keep the computational size in the order of a normal

computer or laptop and the processing time per point in the order of seconds. Provided that the current study is targeting baseload points only, the order of seconds is sufficient to capture all the points and mid fast changes due to components failures. The result obtained by the methodology presented is 1.7 s with scheme 1 including MLKF for the data pre-filtering, ANN and FL. This result has been obtained on a normal laptop and is therefore reproducible on any common computer. This result is well within the target and confirms that the objective is fully achieved. As well explained within the chapters describing the contribution of each section of the methodology, each contributor has been specifically selected to deal with the specific topic while reacting in a time short enough for the online suitability.

The combination of the multiple measurements available at each location has been studied in the dedicated section 4.5 and in the result chapter 5.0 while comparing the results with and without KF. It has been calculated that the combination of the measurements, applied through the KF methodology, was able to reduce the measurement noise up to 83%. This result was reflected in a better quantification of the health of the component, moving the success rate above 90% from 0.8% (ref. no KF, scheme 1) to 1.6% (ref. MLKF, scheme 1).

The last objective of the study is its testing. The ideal testing would be in a real environment, providing deteriorations, different types of failure on different conditions, with noise and errors on the measurements. However, as already remarked in section 4.2 the level or reliability of the gas turbines is such that is not possible to have data reproducing all the possible failures and engine could experience. Additionally, even if available, it is difficult to determine beforehand what type of failure a certain block of data can represent if the data are affected by measurement noise and or components failure. Therefore, it is always necessary to integrate the real data with simulated scenarios to prove the methodology against. In the present study, the decision went toward the full simulation of the data reflecting real deterioration status.

To make sure the methodology is proven under several conditions and the single faults contributors are isolated, the tests set up are 3:

a. Constant deterioration;
b. Random deterioration;
c. Deterioration schedule.

The constant deterioration has been established to check the effect of the noise within a type of failure that is known. In the case of the noise, the effect of the KF has been tested and proved that the MLKF is able to provide a significant improvement of the success rate and keep it above 90% until 1.6% reference noise.

The random deterioration is the stress test for the methodology because it is including the deterioration magnitude varying between 0.15% and 7.4%, all the 24 possible combinations and all the 6 failures type. By showing a success rate both in terms of quantification and classification above 90%, it implies the methodology is reliable and can handle an extreme scenario. It must, in fact, be remarked that the selection is not gradual but can brutally move from a single failure with 0.15% deterioration to multiple failures with an extreme deterioration of 7.4%. The random deterioration has been coupled with the measurement noise in order to further prove how the methodology can react under two combined effects.

The scheduled deterioration reproduces several months of a gas turbine regular deterioration. There are few peculiarities in this simulation: the schedule is based on open information from literature and includes, as per the random and constant deterioration, a ratio between efficiency and mass flow decay; the gradient of deterioration is changing with the time to verify if the methodology follows this trend; the gradient changes first on the LP turbine and last on the LP compressor ending with a higher degradation on the LP turbine; the nominal noise is included meaning that the signals reflect a real situation. These peculiarities allow establishing if the methodology is properly assessing the health of the single components over the established period of

time. The methodology resulted in a quantification success rate of 100% on all the schemes. The classification instead is above the target for all the components and schemes with an exception on the HP turbine fouling and scheme1 where the classification falls at 89.1%. In this case, the target is fully achieved for two schemes, and very close for scheme 1.

8.2 Novelties of the methodology

Chapter 3.0 clearly reported what are still open points from the literature review. The objectives described in section 3.6 mirrored them and expanded including additional points of interest of this research. However, focusing on the pure novelties brought by this work, the main points to mention are:

a. Online detection of single and multiple failures with the presence of measurement issues;

There is a clear research gap in the detection of the multiple failures, including measurements failures, and in the health estimation of the single components. The speed of detection (order of seconds per sample), combined with the precision of the component health estimation, fill this gap and open a research opportunity in that direction.

b. Heath estimation of multiple components in a gas turbine while subjected to single failures, multiple failures and measurement issues;

With this newly built technique, the failure is not only detected, but the health status of every single component is determined. This means that for every single component it would be possible to estimate its precise status and take predictive actions on it. Moreover, the technique can deal with measurements issues, while providing quality results.

c. The combination of KF for the data analysis, the ANN for the performance prediction and the FL/neuro Fuzzy logic for the quantification and the classification of the failures

Looking at the methodology as such, this combination represents also a novelty not seen yet in literature. While most of the newly proposed techniques started to combine two contributors (e.g. ANN and GA), the combination of these methodologies – KF, ANN, NFL and FL – is not yet seen in the literature for GPA investigation. Moreover, this research motivates and outlines the advantages of this combination showing the contribution of each section.

8.3 Further work recommended

The current work provides a strong base for the GPA and components of health estimation, also with the presence of measurement issues, suitable for online diagnostics.

One of the topics that could not be explored fully is testing. As described earlier, the methodology has been tested in an environment that is close to reality, but it does not represent it fully. Therefore, an important topic of additional work recommended, is the further testing of the methodology with real data, reflecting the deterioration of one or more components. Those data would ideally include measurements issues and should represent a long-term engine run in order to capture different scenarios. However, since it is unlikely to receive data clearly representing all the types of failure (as detailed in section 4.2 and reported by Simon et al. [39]), it would be interesting to blend simulated data and real data trying to find the correct mix between ANN and Neuro-Fuzzy training and the success rate.

Still on the testing, the deterioration profile simulating several months of GT run provided successful results for the combined training, while the random training option, the efficiency prediction was unsatisfactory – section 5.1.3. However, due to memory limitations, it is currently not possible to consistently increase the number of samples. Therefore, it would be interesting to

investigate the effect of the increase of the number of samples, necessary to overcome the underfit seen without the combined training. The same should be applied both to the ANN and to the NFL.

Looking at the methodology side in section 4.8 it has been reported that the Fuzzy Logic for failure classification has been selected for its flexibility, required for the implementation of different rules, for the detection of other failures. In the current work, the FL provided excellent results for the detection of compressor fouling and for the detection of turbine erosion. However, as reported by Meher et al. [76] there are several other failures that could be investigated like: tip clearance increase, the sealing degradation and the air leakage, the airfoil roughness, the inlet and outlet delta pressure increase. These failures should be modelled in the preparation phase, included in the KF and reported in the cumulative alarm level/failure type chart. Moreover, as implicitly set up in the cumulative alarm level/failure type chart, the non-recoverable measurements failures could be also isolated and reported to make sure the diagnostics are not driven by an erratic signal.

The current results reflect a full load status of the engine meaning that it is working at its full regime. On top of that, the variation of the ambient conditions does not go far from the ISA values. Additional work should include different ambient conditions and operating regimes in the training phase and try to make predictions and components health estimation also with these additional statuses.

Looking at additional features that could even further complete this work it is the physical connection of the alarm level to the component health estimation. By this, the alarm level could be directly related to a component status. An additional step forward can be the lifetime estimation, that in the same way as per the component health estimation, is related to the connection to the component real status.

This methodology has been conceived to be flexible enough to adapt to several types of engine. In this research, the focus went on industrial gas

turbines with two-spool in particular. However, the methodology can be expanded to analyse configurations with more components (e.g. three spools) or for different applications (e.g. aero engines). In these cases, the architecture will remain identical, but the single contributors will adapt to the new engine design. For instance, the KF will have to include the additional/different measurements, the ANN will be trained and adapted according to the measurements available. Additionally, the prediction will have to include the new components. Finally, the NFL and the FL will have to be structured to evaluate properly all the components and their mutual interaction.

REFERENCES

[1]. G. Bechini (2007). Performance diagnostic and measurement selection for on-line monitoring of gas turbine engines. Ph.D. Thesis Cranfield University

[2]. L. Marinai (2004). Gas-path diagnostics and prognostics for aero-engines using fuzzy logic and time series analysis. PhD Thesis Cranfield University

[3]. Volponi, A. J. (2014). Gas turbine engine health management: Past, present, and future trends. Journal of Engineering for Gas Turbines and Power. https://doi.org/10.1115/1.4026126

[4]. Verbist, M. L., Visser, W. P. J., Van Buijtenen, J. P., & Duivis, R. (2011). Gas path analysis on KLM in-flight engine data. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/GT2011-45625

[5]. Marinai, L., Probert, D., & Singh, R. (2004). Prospects for aero gas-turbine diagnostics: A review. Applied Energy. https://doi.org/10.1016/j.apenergy.2003.10.005

[6]. Grace, Dale S. "Combined-Cycle Power Plant Maintenance Costs." Proceedings of the ASME Turbo Expo 2008: Power for Land, Sea, and Air. Volume 2: Controls, Diagnostics and Instrumentation; Cycle Innovations; Electric Power. Berlin, Germany. June 9–13, 2008. pp. 1007-1016. ASME.

[7]. Grace, D., & Christiansen, T. (2012). Risk-based assessment of unplanned outage events and costs for combined-cycle-plants. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/GT2012-68435

[8]. Airline maintenance cost executive commentary, An Exclusive Benchmark Analysis (FY2016 data) - IATA's Maintenance Cost Task Forceb. https://www.iata.org/whatwedo/workgroups/Documents/MCTF/MCTF-FY2016-Report-Public.pdf

[9]. Vance, J., Zeidan, F., & Murphy, B. (2010). Machinery Vibration and Rotordynamics. In Machinery Vibration and Rotordynamics. https://doi.org/10.1002/9780470903704

[10]. Handbook of Condition Monitoring. (1998). In Handbook of Condition Monitoring. https://doi.org/10.1007/978-94-011-4924-2

[11]. Urban, L. A. (1973). Gas path analysis applied to turbine engine condition monitoring. Journal of Aircraft. https://doi.org/10.2514/3.60240

[12]. Escher, P.C. (1995). Pythia: An Object-Oriented Gas-path Analysis Computer Program for General Applications. Ph.D. Thesis, Cranfield University

[13]. S. Ogajiye (2003). Advanced Gas-path Fault Diagnostics for Stationary Gas Turbines. PhD Thesis Cranfield University

[14]. Grewal, M. S., & Andrews, A. P. (2001). Kalman Filtering : Theory and Practice Using MATLAB California State University at Fullerton. In Theory and Practice. https://doi.org/10.1002/0471266388

[15]. Kerr, L. J., Nemec, T. S., & Gallops, G. W. (1992). Real-time estimation of gas turbine engine damage using a control-based kalman filter algorithm. Journal of Engineering for Gas Turbines and Power. https://doi.org/10.1115/1.2906571

[16]. Kobayashi, T., & Simon, D. L. (2003). Application of a bank of Kalman filters for aircraft engine fault diagnostics. American Society of Mechanical Engineers, International Gas Turbine Institute, Turbo Expo (Publication) IGTI. https://doi.org/10.1115/GT2003-38550

[17]. Simon, D. (2008). A comparison of filtering approaches for aircraft engine health estimation. Aerospace Science and Technology. https://doi.org/10.1016/j.ast.2007.06.002

[18]. Lu, F., Wang, Y., Huang, J., Huang, Y., & Qiu, X. (2018). Fusing unscented Kalman filter for performance monitoring and fault accommodation in gas turbine. Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering. https://doi.org/10.1177/0954410016682269

[19]. Sun, S. L., & Deng, Z. L. (2004). Multi-sensor optimal information fusion Kalman filter. Automatica. https://doi.org/10.1016/j.automatica.2004.01.014

[20]. Chernyatin, A. N., & Ostroukhov, M. Y. (1968). Nural Network Design 2nd edition. Metallurgist. https://doi.org/10.1007/BF00738424

[21]. Joly, R. B., Ogaji, S. O. T., Singh, R., & Probert, S. D. (2004). Gas-turbine diagnostics using artificial neural-networks for a high bypass ratio military turbofan engine. Applied Energy. https://doi.org/10.1016/j.apenergy.2003.10.002

[22]. Asgari, H., Chen, X., Menhaj, M. B., & Sainudiin, R. (2013). Artificial neural network-based system identification for a single-shaft gas turbine. Journal of Engineering for Gas Turbines and Power. https://doi.org/10.1115/1.4024735

[23]. Vatani, A., Khorasani, K., & Meskin, N. (2015). Health monitoring and degradation prognostics in gas turbine engines using dynamic neural networks. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/GT2015-44101

[24]. Kadamb A. (2003). Bayesian belief network for aero gas turbine module and system fault isolation – MSc thesis – Cranfield University

[25]. Pu, X., Liu, S., Jiang, H., & Yu, D. (2013). Sparse bayesian learning for gas path diagnostics. Journal of Engineering for Gas Turbines and Power. https://doi.org/10.1115/1.4023608

[26]. Kestner, B. K., Lee, Y. K., Voleti, G., Mavris, D. N., Kumar, V., & Lin, T. P. (2011). Diagnostics of highly degraded industrial gas turbines using Bayesian networks. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/GT2011-45249

[27]. Mo, H., Sansavini, G., & Xie, M. (2018). Performance-based maintenance of gas turbines for reliable control of degraded power systems. Mechanical Systems and Signal Processing. https://doi.org/10.1016/j.ymssp.2017.10.021

[28]. Qingcai, Y., Li, S., Cao, Y., & Zhao, N. (2016). Full and part-load performance deterioration analysis of industrial three-shaft gas turbine based on genetic algorithm. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/GT2016-57120

[29]. Ogaji, S. O. T., Marinai, L., Sampath, S., Singh, R., & Prober, S. D. (2005). Gas-turbine fault diagnostics: A fuzzy-logic approach. Applied Energy. https://doi.org/10.1016/j.apenergy.2004.07.004

[30]. Barbosa, R., & Ferreira, S. (2012). Industrial gas turbine diagnostics using fuzzy logic. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/GT2012-68767

[31]. Eustace, R. W. (2007). A real-world application of fuzzy logic and inlfluence coefficients for gas turbine performance diagnostics. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/GT2007-27442

[32]. Sampath, S., & Singh, R. (2006). An integrated fault diagnostics model using genetic algorithm and neural networks. Journal of Engineering for Gas Turbines and Power. https://doi.org/10.1115/1.1995771

[33]. Viharos, Z. J., & Kis, K. B. (2015). Survey on Neuro-Fuzzy systems and their applications in technical diagnostics and measurement. Measurement: Journal of the International Measurement Confederation. https://doi.org/10.1016/j.measurement.2015.02.001

[34]. Wang, J., Fan, K., & Wang, W. (2010). Integration of fuzzy AHP and FPP with TOPSIS methodology for aeroengine health assessment. Expert Systems with Applications. https://doi.org/10.1016/j.eswa.2010.05.024

[35]. Dewallef, P., Romessis, C., Léonard, O., & Mathioudakis, K. (2006). Combining classification techniques with Kalman filters for aircraft engine diagnostics. Journal of Engineering for Gas Turbines and Power. https://doi.org/10.1115/1.2056507

[36]. Kumar, A., Shankar, R., & Thakur, L. S. (2018). A big data driven sustainable manufacturing framework for condition-based maintenance prediction. Journal of Computational Science. https://doi.org/10.1016/j.jocs.2017.06.006

[37]. Yang, X., Pang, S., Shen, W., Lin, X., Jiang, K., & Wang, Y. (2016). Aero Engine Fault Diagnosis Using an Optimized Extreme Learning Machine. International Journal of Aerospace Engineering. https://doi.org/10.1155/2016/7892875

[38]. Li, Z., Zhong, S. S., & Lin, L. (2017). Novel gas turbine fault diagnosis method based on performance deviation model. Journal of Propulsion and Power. https://doi.org/10.2514/1.B36267

[39]. Simon, D. L., Bird, J., Davison, C., Volponi, A., & Iverson, R. E. (2008). Benchmarking gas path diagnostic methods: A public approach. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/GT2008-51360

[40]. Simon, D. L., Borguet, S., Léonard, O., & Zhang, X. (2014). Aircraft engine gas path diagnostic methods: Public benchmarking results. Journal of Engineering for Gas Turbines and Power. https://doi.org/10.1115/1.4025482

[41]. S. Togni, N. Theoklis, S. Sampath (2017). The application of multiple methodologies for the diagnostics of components multiple failure. ISABE 2017

[42]. Kurz, R., Brun, K., & Wollie, M. (2008). Degradation effects on industrial gas turbines. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/GT2008-50020

[43]. Valero, A., Correas, L., Zaleta, A., Lazzaretto, A., Verda, V., Reini, M., & Rangel, V. (2004). On the thermoeconomic approach to the diagnosis of energy system malfunctions Part 2. Malfunction definitions and assessment. Energy. https://doi.org/10.1016/j.energy.2004.03.008

[44]. Morini, M., Pinelli, M., Spina, P. R., & Venturini, M. (2008). Influence of blade deterioration on compressor and turbine performance. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/GT2008-50043

[45]. Morini, M., Pinelli, M., Spina, P. R., & Venturini, M. (2009). CFD simulation of fouling on axial compressor stages. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/GT2009-59025

[46]. Lakshminarasimha, A. N., Boyce, B. P., & Meher-Homj, C. B. (1994). Modeling and analysis of gas turbine performance deterioration. Journal of Engineering for Gas Turbines and Power. https://doi.org/10.1115/1.2906808

[47]. Morini, M., Pinelli, M., Spina, P. R., & Venturini, M. (2010). Numerical analysis of the effects of non-uniform surface roughness on compressor stage performance. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/GT2010-23291

[48]. Industrial Gas Turbine Performance: Compressor Fouling and On-Line Washing - U. Igie, P. Pilidis, D. Fouflias, K. Ramsden, P. Laskaridis - Journal of Turbomachinery – October 2014

[49]. Zhou, D., Zhang, H., & Weng, S. (2015). A new gas path fault diagnostic method of gas turbine based on support vector machine. Journal of Engineering for Gas Turbines and Power. https://doi.org/10.1115/1.4030277

[50]. Mohammadi, E., & Montazeri-Gh, M. (2014). Simulation of full and part-load performance deterioration of industrial two-shaft gas Turbine. Journal of Engineering for Gas Turbines and Power. https://doi.org/10.1115/1.4027187

[51]. Kurz R., B. K. (2001). Degradation in gas turbine systems. Journal of Engineering for Gas Turbines and Power. https://doi.org/10.1115/1.1340629

[52]. Igie, U., Diez-Gonzalez, P., Giraud, A., & Minervino, O. (2016). Evaluating Gas Turbine Performance Using Machine-Generated Data: Quantifying Degradation and Impacts of Compressor Washing. Journal of Engineering for Gas Turbines and Power. https://doi.org/10.1115/1.4033748

[53]. Venturini, M., & Therkorn, D. (2013). Application of a statistical methodology for gas turbine degradation prognostics to alstom field data. Journal of Engineering for Gas Turbines and Power. https://doi.org/10.1115/1.4024952

[54]. Stalder, J. P. (2001). Gas turbine compressor washing state of the art: Field experiences. Journal of Engineering for Gas Turbines and Power. https://doi.org/10.1115/1.1361108

[55]. Morini, M., Pinelli, M., Spina, P. R., & Venturini, M. (2008). Influence of blade deterioration on compressor and turbine performance. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/GT2008-50043

[56]. Mathioudakis, K., Kamboukos, P., & Stamatis, A. (2002). Turbofan performance deterioration tracking using non-linear models and optimization techniques. American Society of Mechanical Engineers, International Gas Turbine Institute, Turbo Expo (Publication) IGTI. https://doi.org/10.1115/GT2002-30026

[57]. Bringhenti, C., & Barbosa, J. R. (2008). Effects of turbine tip clearance on gas turbine performance. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/GT2008-50196

[58]. Ghenaiet, A., Tan, S. C., & Elder, R. L. (2004). Experimental investigation of axial fan erosion and performance degradation. Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy. https://doi.org/10.1243/0957650041761900

[59]. Bakken, L. E., & Skorping, R. (1996). Optimum operation and maintenance of gas turbines offshore. ASME 1996 International Gas Turbine and Aeroengine Congress and Exhibition, GT 1996. https://doi.org/10.1115/96-GT-273

[60]. Meher-Homji, C. B., Chaker, M., & Bromley, A. F. (2009). The fouling of axial flow compressors - Causes, effects, susceptibility and sensitivity. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/GT2009-59239

[61]. Brekke, O., & Bakken, L. E. (2010). Accelerated deterioration by saltwater ingestion in gas turbine intake air filters. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/GT2010-22455

[62]. Verbist, M. L., Visser, W. P. J., & Van Buijtenen, J. P. (2013). Experience with gas path analysis for on-wing turbofan condition monitoring. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/GT2013-95739

[63]. Tarabrin, A. P., Schurovsky, V. A., Bodrov, A. I., & Stalder, J. P. (1998). Influence of axial compressor fouling on gas turbine unit perfomance based on different schemes and with different initial parameters. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/98-GT-416

[64]. Schneider, E., Bussjaeger, S. D., Franco, S., & Therkorn, D. (2010). Analysis of compressor on-line washing to optimize gas turbine power plant performance. Journal of Engineering for Gas Turbines and Power. https://doi.org/10.1115/1.4000133

[65]. Igie, U., Pilidis, P., Fouflias, D., Ramsden, K., & Lambart, P. (2011). On-line compressor cascade washing for gas turbine performance investigation. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/GT2011-46210

[66]. Hepperle, N., Therkorn, D., Schneider, E., & Staudacher, S. (2011). Assessment of gas turbine and combined cycle power plant performance degradation. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/GT2011-45375

[67]. Kurz, R., & Brun, K. (2011). Fouling mechanisms in axial compressors. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/GT2011-45012

[68]. Maiwada, B., Mu 'az, N. I., Ibrahim, S., & Musa, S. M. (2016). Impacts of Compressor Fouling On the Performance of Gas Turbine. International Journal of Engineering Science and Computing. https://doi.org/10.4010/2016.514

[69]. Yang, H., & Xu, H. (2014). The new performance calculation method of fouled axial flow compressor. Scientific World Journal. https://doi.org/10.1155/2014/906151

[70]. Schroth, T., Rothmann, A., & Schmitt, D. (2007). Nutzwert eines Dreistufigen Luftfiltersystems mit Innovativer Technologie für Stationäre Gasturbinen. VGB PowerTech.

[71]. Cavarzere, A., & Venturini, M. (2011). Application of forecasting methodologies to predict gas turbine behavior over time. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/GT2011-45710

[72]. Wilcox, M., & Brun, K. (2011). Gas turbine inlet filtration system life cycle cost analysis. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/GT2011-46708

[73]. Diagnostic Methods for an Aircraft Engine Performance - E. L. Ntantis, P. N. Botsaris - Journal of Engineering Science and Technology – November 2015

[74]. Simulation of performance deterioration in eroded compressors – D. Singh, A. lamed, W Tabakoff - The American Society Of Mechanical Engineers – June 1996

[75]. Pathirathna, K. A. B. (2013). Gas Turbine Thermodynamic and Performance Analysis Methods Using Available Catalog Data (Dissertation). Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:hig:diva-17474

[76]. Meher-Homji, Cyrus B.; Chaker, Mustapha A.; Motiwala, Hatim M. (2001). Gas Turbine Performance Deterioration.. Texas A&M University. Turbomachinery Laboratories. Available electronically from http : / /hdl .handle .net /1969 .1 /163330

[77]. J. Palmer and T. Nikolaidis (2015), "The TURBOMATCH Scheme for Gas Turbine Performance Calculations (ver. 2.0) - User Manual,"

[78]. W. J. Stromberg (1981). Performance deterioration based on simulated aerodynamic loads test, JT9D jet engine diagnostics program. NASA CR-165297

[79]. Jiang, X., & Foster, C. (2014). Plant performance monitoring and diagnostics - Remote, real-time and automation. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/GT2014-27314

[80]. Saravanamuttoo, H I (1990). Recommended Practices for Measurement of Gas Path Pressures and Temperatures for Performance Assessment of Aircraft Turbine Engines and Components. AGARD-AR-.245

[81]. Gülen, S. C., Griffin, P. R., & Paolucci, S. (2002). Real-time on-line performance diagnostics of heavy-duty industrial gas turbines. Journal of Engineering for Gas Turbines and Power. https://doi.org/10.1115/1.1413465

[82]. Fentaye, Amare & Baheta, Aklilu & Gilani, Syed & Kyprianidis, Konstantinos. (2019). A Review on Gas Turbine Gas-Path Diagnostics: State-of-the-Art Methods, Challenges and Opportunities. Aerospace. 6. 83. 10.3390/aerospace6070083.

[83]. Tengeleng, S., & Armand, N. (2014). Performance of using cascade forward back propagation neural networks for estimating rain parameters with rain drop size distribution. Atmosphere. https://doi.org/10.3390/atmos5020454

[84]. Garg, Sanjay, Simon, Donald L. (2012). NASA Technical Reports Server (NTRS) 20150009565: Challenges in Aircraft Engine Control and Gas Path Health Management

[85]. Abdul Ghafir, M. F., Li, Y. G., & Wang, L. (2012). Creep life prediction for aero gas turbine hot section component using Artificial Neural Networks. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/GT2012-68856

[86]. Basso, M., Giarré, L., Groppi, S., & Zappa, G. (2005). NARX models of an industrial power plant gas turbine. IEEE Transactions on Control Systems Technology. https://doi.org/10.1109/TCST.2004.843129

[87]. Loboda, Igor & Yepifanov, Sergiy & Felshteyn, Yakov. (2009). An Integrated Approach to Gas Turbine Monitoring and Diagnostics. International Journal of Turbo Jet Engines. 26. 111-126. 10.1515/TJJ.2009.26.2.111.

[88]. Simon, D. L., & Armstrong, J. B. (2012). An integrated approach for aircraft engine performance estimation and fault diagnostics. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/GT2012-69905

[89]. Dewallef, P., & Borguet, S. (2013). A methodology to improve the robustness of gas turbine engine performance monitoring against sensor faults. Journal of Engineering for Gas Turbines and Power. https://doi.org/10.1115/1.4007976

[90]. Kumano, S., Mikami, N., & Aoyama, K. (2011). Advanced gas turbine diagnostics using pattern recognition. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/GT2011-45670

[91]. Kanelopoulos, K., Stamatis, A., & Mathioudakis, K. (1997). Incorporating neural networks into gas turbine performance diagnostics. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/97-GT-035

[92]. Loboda, I., Yepifanov, S., & Feldshteyn, Y. (2007). A generalized fault classification for gas turbine diagnostics at steady states and transients. Journal of Engineering for Gas Turbines and Power. https://doi.org/10.1115/1.2719261

[93]. Volponi, A. J., DePold, H., Ganguli, R., & Daguang, C. (2003). The use of kalman filter and neural network methodologies in gas turbine performance diagnostics: A comparative study. Journal of Engineering for Gas Turbines and Power. https://doi.org/10.1115/1.1419016

[94]. Fast, M., Assadi, M., & De, S. (2008). Condition based maintenance of gas turbines using simulation data and artificial neural network: A demonstration of feasibility. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/GT2008-50768

[95]. Lee, S. M., Choi, W. J., Roh, T. S., & Choi, D. W. (2007). Defect diagnostics of gas turbine engine using hybrid SVM-artificial neural network method. Collection of Technical Papers - 43rd AIAA/ASME/SAE/ASEE Joint Propulsion Conference. https://doi.org/10.2514/6.2007-5109

[96]. Song, T. L., & Speyer, J. L. (1985). A Stochastic Analysis of a Modified Gain Extended Kalman Filter with Applications to Estimation with Bearings Only Measurements. IEEE Transactions on Automatic Control. https://doi.org/10.1109/TAC.1985.1103821

[97]. Drécourt, J. P., Madsen, H., & Rosbjerg, D. (2006). Bias aware Kalman filters: Comparison and improvements. Advances in Water Resources. https://doi.org/10.1016/j.advwatres.2005.07.006

[98]. Novoselov, R. Y., Herman, S. M., Gadaleta, S. M., & Poore, A. B. (2005). Mitigating the effects of residual biases with Schmidt-Kalman filtering. 2005 7th International Conference on Information Fusion, FUSION. https://doi.org/10.1109/ICIF.2005.1591877

[99]. Ogaji, S. O. T., & Singh, R. (2002). Advanced engine diagnostics using artificial neural networks. Proceedings - 2002 IEEE International Conference on Artificial Intelligence Systems, ICAIS 2002. https://doi.org/10.1109/ICAIS.2002.1048094

[100]. Yilmaz, Y., Kurz, R., Özmen, A., & Weber, G. W. (2015). A new algorithm for scheduling condition-based maintenance of gas turbines. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/GT2015-43545

[101]. Li, Y. G. (2008). A genetic algorithm approach to estimate performance status of gas turbines. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/GT2008-50175

[102]. Zedda, M. (1999). Gas turbine engine and sensor fault diagnosis. PhD Thesis - School of mechanical engineering, Cranfield University

[103]. Zedda, M. and Singh R. (1999). Gas turbine engine and sensor fault diagnosis. ISABE 99-7238, 13th ISABE, Florence, Italy

[104]. Elfghi, F. M. (2016). A hybrid statistical approach for modeling and optimization of RON: A comparative study and combined application of response surface methodology (RSM) and artificial neural network (ANN) based on design of experiment (DOE). Chemical Engineering Research and Design. https://doi.org/10.1016/j.cherd.2016.05.023

[105]. Bezerra, M. A., Santelli, R. E., Oliveira, E. P., Villar, L. S., & Escaleira, L. A. (2008). Response surface methodology (RSM) as a tool for optimization in analytical chemistry. Talanta. https://doi.org/10.1016/j.talanta.2008.05.019

[106]. Ferreira, S. L. C., Bruns, R. E., Ferreira, H. S., Matos, G. D., David, J. M., Brandão, G. C., … dos Santos, W. N. L. (2007). Box-Behnken design: An alternative for the optimization of analytical methods. Analytica Chimica Acta. https://doi.org/10.1016/j.aca.2007.07.011

[107]. Franceschini, G., & Macchietto, S. (2008). Model-based design of experiments for parameter precision: State of the art. Chemical Engineering Science. https://doi.org/10.1016/j.ces.2007.11.034

[108]. Buzzi-Ferraris, G., & Forzatti, P. (1983). A new sequential experimental design procedure for discriminating among rival models. Chemical Engineering Science. https://doi.org/10.1016/0009-2509(83)85004-0

[109]. Li, Y. G. (2002). Performance-analysis-based gas turbine diagnostics: A review. Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy. https://doi.org/10.1243/095765002320877856

[110]. O'Brien, W., Kurzke, J., Rudnitski, D., Horobin, M., Csavina, F., Masson, P., … Evans, R. (2007). Performance Prediction and Simulation of Gas Turbine Engine Operation for Aircraft, Marine, Vehicular and Power Generation. In NATO Research and Technology Organisation.

[111]. Kanelopoulos, K., Stamatis, A., & Mathioudakis, K. (1997). Incorporating neural networks into gas turbine performance diagnostics. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/97-GT-035

[112]. A. Apostolidis (2015). Turbine cooling and heat transfer modelling for gas turbine performance simulation. PhD Thesis Cranfield University

[113]. Sun, J., Zuo, H., Liang, K., & Chen, Z. (2016). Bayesian Network-Based Multiple Sources Information Fusion Mechanism for Gas Path Analysis. Journal of Propulsion and Power. https://doi.org/10.2514/1.B35658

[114]. Ntantis, E. L., & Botsaris, P. N. (2015). Diagnostic methods for an aircraft engine performance. Journal of Engineering Science and Technology Review. https://doi.org/10.25103/jestr.084.10

[115]. Lu, F., Huang, J., & Lv, Y. (2013). Gas path health monitoring for a turbofan engine based on a nonlinear filtering approach. Energies. https://doi.org/10.3390/en6010492

[116]. Dewallef, P., Romessis, C., Léonard, O., & Mathioudakis, K. (2006). Combining classification techniques with Kalman filters for aircraft engine diagnostics. Journal of Engineering for Gas Turbines and Power. https://doi.org/10.1115/1.2056507

[117]. Kyriazis, A., Tsalavoutas, A., Mathioudakis, K., Bauer, M., & Johanssen, O. (2009). Gas turbine fault identification by fusing vibration trending and gas path analysis. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/GT2009-59942

[118]. Zhou, D., Zhang, H., & Weng, S. (2015). A new gas path fault diagnostic method of gas turbine based on support vector machine. Journal of Engineering for Gas Turbines and Power. https://doi.org/10.1115/1.4030277

[119]. D. Simon (2012). Challenges in aircraft engine gas path health management. Presented at turbo expo 2012

[120]. Tahan, M., Tsoutsanis, E., Muhammad, M., & Abdul Karim, Z. A. (2017). Performance-based health monitoring, diagnostics and prognostics for condition-based maintenance of gas turbines: A review. Applied Energy. https://doi.org/10.1016/j.apenergy.2017.04.048

[121]. Allen, C. W., Holcomb, C. M., & De Oliveira, M. (2017). Gas turbine machinery diagnostics: A brief review and a sample application. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/GT2017-64755

[122]. Pérez-Ruiz, J. L., Loboda, I., Miró-Zárate, L. A., Toledo-Velázquez, M., & Polupan, G. (2017). Evaluation of gas turbine diagnostic techniques under variable fault conditions. Advances in Mechanical Engineering. https://doi.org/10.1177/1687814017727471

[123]. Palade, V., Patton, R. J., Uppal, F. J., Quevedo, J., & Daley, S. (2002). Fault diagnosis of an industrial gas turbine using neuro-fuzzy methods. IFAC Proceedings Volumes (IFAC-PapersOnline). https://doi.org/10.3182/20020721-6-es-1901.01632

[124]. Li, K., Zhang, Q., Wang, K., Chen, P., & Wang, H. (2016). Intelligent condition diagnosis method based on adaptive statistic test filter and diagnostic bayesian network. Sensors (Switzerland). https://doi.org/10.3390/s16010076

[125]. Pourbabaee, B., Meskin, N., & Khorasani, K. (2016). Robust sensor fault detection and isolation of gas turbine engines subjected to time-varying parameter uncertainties. Mechanical Systems and Signal Processing. https://doi.org/10.1016/j.ymssp.2016.02.023

[126]. Jia, F., Lei, Y., Lin, J., Zhou, X., & Lu, N. (2016). Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data. Mechanical Systems and Signal Processing. https://doi.org/10.1016/j.ymssp.2015.10.025

[127]. Sina Tayarani-Bathaie, S., & Khorasani, K. (2015). Fault detection and isolation of gas turbine engines using a bank of neural networks. Journal of Process Control. https://doi.org/10.1016/j.jprocont.2015.08.007

[128]. Verma, R., Roy, N., & Ganguli, R. (2006). Gas turbine diagnostics using a soft computing approach. Applied Mathematics and Computation. https://doi.org/10.1016/j.amc.2005.02.057

[129]. Zio, E., & Gola, G. (2009). A neuro-fuzzy technique for fault diagnosis and its application to rotating machinery. Reliability Engineering and System Safety. https://doi.org/10.1016/j.ress.2007.03.040

[130]. Zhou, D., Wei, T., Zhang, H., Ma, S., & Wei, F. (2018). An information fusion model based on dempster-shafer evidence theory for equipment diagnosis. ASCE-

ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering. https://doi.org/10.1115/1.4037328

[131]. Yang, Q., Cao, Y., Yu, F., Du, J., & Li, S. (2017). Health estimation of gas turbine: A symbolic linearization model approach. Proceedings of the ASME Turbo Expo. https://doi.org/10.1115/GT201764071

[132]. Lu, F., Ju, H., & Huang, J. (2016). An improved extended Kalman filter with inequality constraints for gas turbine engine health monitoring. Aerospace Science and Technology. https://doi.org/10.1016/j.ast.2016.08.008

[133]. Coraddu, A., Oneto, L., Ghio, A., Savio, S., Anguita, D., & Figari, M. (2016). Machine learning approaches for improving condition-based maintenance of naval propulsion plants. Proceedings of the Institution of Mechanical Engineers Part M: Journal of Engineering for the Maritime Environment. https://doi.org/10.1177/1475090214540874

[134]. Alaswad, S., & Xiang, Y. (2017). A review on condition-based maintenance optimization models for stochastically deteriorating system. Reliability Engineering and System Safety. https://doi.org/10.1016/j.ress.2016.08.009

[135]. Panov, V. (2015). Gas turbine performance diagnostics and fault isolation based on multidimensional complex health vector space. 11th European Conference on Turbomachinery Fluid Dynamics and Thermodynamics, ETC 2015.

[136]. Lu, F., Qian, J., Huang, J., & Qiu, X. (2017). In-flight adaptive modeling using polynomial LPV approach for turbofan engine dynamic behavior. Aerospace Science and Technology. https://doi.org/10.1016/j.ast.2017.02.003

[137]. Feng Lu and Tiebin Zhu and Jinquan Huang (2014). Multi-sensor data fusion using least square support vector regression for missing data online recovery. 50th AIAA/ASME/SAE/ASEE Joint Propulsion Conference

[138]. Li, Y. G., & Nilkitsaranont, P. (2009). Gas turbine performance prognostic for condition-based maintenance. Applied Energy. https://doi.org/10.1016/j.apenergy.2009.02.011

[139]. Dewallef, P., & Borguet, S. (2013). A methodology to improve the robustness of gas turbine engine performance monitoring against sensor faults. Journal of Engineering for Gas Turbines and Power. https://doi.org/10.1115/1.4007976

[140]. Simon, D. L., & Rinehart, A. W. (2016). Sensor Selection for Aircraft Engine Performance Estimation and Gas Path Fault Diagnostics. Journal of Engineering for Gas Turbines and Power. https://doi.org/10.1115/1.4032339

[141]. Brun K., K. R. (2001). Measurement uncertainties encountered during gas turbine driven compressor field testing. Journal of Engineering for Gas Turbines and Power. https://doi.org/10.1115/1.1340628

APPENDIX

Driving parameters for the Turbomatch deterioration modelling

The parameters used to model the 24 deteriorated combinations are listed in Table 79 and reflect the ratio reported in the literature and summarized in Table 3.

1	LP Comp Efficiency deterioration
2	LP Comp Massflow deterioration
3	LP Comp Pressure Ratio deterioration
4	HP Comp Efficiency deterioration
5	HP Comp Massflow deterioration
6	HP Comp Pressure Ratio deterioration
7	HP Turb Efficiency deterioration
8	HP Turb Massflow deterioration
9	HP Turb dh/T deterioration
10	LP Turb Efficiency deterioration
11	LP Turb Massflow deterioration
12	LP Turb dh/T deterioration

Table 79 Turbomatch deterioration parameters

Condition based monitoring graphical user interface

The graphical user interface is built to help the user assessing the deterioration of each component of an engine. The mechanism behind the GUI is the one shown in Figure 18 including the measurements filtering section, the artificial neural network training and prediction and the fuzzy logic qualification and quantification. The characteristics planned for this GUI are: be easy to use, provide a simple and immediate answer to the user, provide an additional level of detail for the user to investigate the failures, show the regular measurements, summarize all the failures and the related parameters in a table, provide space and flexibility to user-defined rules.

The GUI is divided into the sections that reflect the requirements planned above (Figure 122). In section 2 six bars are inserted: the first two on the compressor side reflect the LP and HP compressor status. The bar is colour

coded (with traffic light code) with green lemon colour if the alarm is between 5 and 20, yellow colour if the alarm level is between 20 and 50, orange if the alarm is between 50 and 90, and red colour if the alarm level is above 90. On the turbine side, four bars are drawn. Two for the HP turbine and two for the LP turbine as the fouling/corrosion and the erosion effect are distinguished. The left bar indicates fouling/corrosion while the right bar indicates the erosion.

In section 1 two charts are drawn. On the upper side, the chart is used to plot two possible parameters among the available within a specified range of time. The parameters available are those at each location - reference Figure 21 - and in addition to that the efficiency and the alarm level of each component. The plots allow the user to perform a graphical diagnostic and to cross-compare parameters in order to reach the right conclusion. The lower graph is used to plot real-time value. In the alarm level of the LP, the compressor is shown for points 9 and 10. The same level is also reported in the bar graph in section 2.

In section 3 the real-time measurements for each location are reported. This information can be useful for reference but will rarely provide a diagnostics indication. From there though it might be seen if a measurement is drifting or has other types of errors.

Section 4 shows a polar plot reporting the exhaust temperature difference. The difference is calculated as:

$$T_{diff} = T_i - (T_{est}) \quad (31)$$

Where T_i is the temperature of the single probe and T_{est} is the estimated temperature after the KF.

Section 5 contains a table that summarizes all the diagnostics information. The first four columns report the efficiency of the components. The next four columns report the alarm level of each component and the last four to indicate the cause of the failure. The code is 1 for the fouling, 0 for no failure and 2 for the erosion. The last column, that is not appearing below, indicates the exhaust temperature spread.

Section 6 has one tick box for KF. If ticked it triggers the multiple measurements and the measurement filtering shown in Figure 21, if not it works on a single measurement.

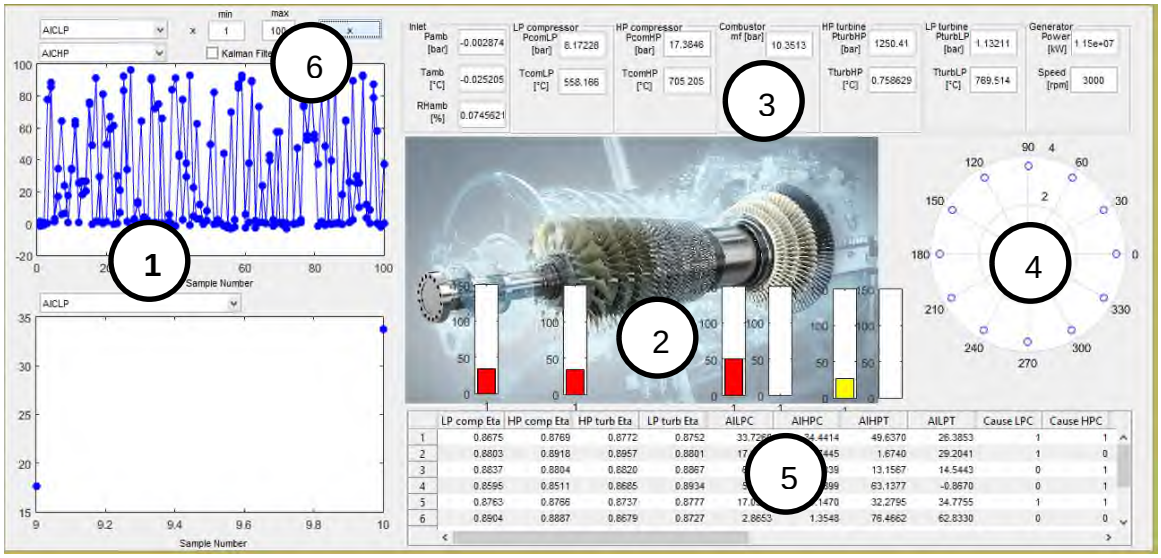

Figure 122 Graphical user interface

Set up the FL for failure quantification and for failure classification

The set up for the failure quantification relies on the deltas from the reference conditions. The selection is based on what have been experienced to be the most impactful parameters for that specific simulation. There is a set up that is valid for the failure quantification (Table 80) and another that is used for the failure classification (Table 81).

Failure quantification					
		LP comp	HP comp	HP turb	LP turb
LP compressor exhaust pressure	Δp2	X	X		
LP compressor exhaust temperature	ΔT2	X	X		
LP compressor efficiency	ΔEtaCLP	X		X	X
HP compressor exhaust pressure	Δp3		X		
HP compressor exhaust temperature	ΔT3	X	X	X	
HP compressor efficiency	ΔEtaCHP	X	X	X	X
HP turbine exhaust pressure	Δp5				
HP turbine exhaust temperature	ΔT5				
HP turbine efficiency	ΔEtaTHP			X	X
HP turbine massflow deterioration	ΔmfTHPDeg				
LP turbine exhaust pressure	Δp6				
LP turbine exhaust temperature	ΔT6	X	X	X	
LP turbine efficiency	ΔEtaTLP			X	X
LP turbine mass flow deterioration	ΔmfTLPDeg				
Power	ΔP				

Table 80 Fuzzy logic failure quantification set up

Failure classification							
		LP comp foul	HP comp foul	HP turb foul	HP turb erosion	LP turb foul	LP turb erosion
LP compressor exhaust pressure	Δp2						
LP compressor exhaust temperature	ΔT2	X					
LP compressor efficiency	ΔEtaCLP	X					
HP compressor exhaust pressure	Δp3						
HP compressor exhaust temperature	ΔT3		X				
HP compressor efficiency	ΔEtaCHP		X				
HP turbine exhaust pressure	Δp5						
HP turbine exhaust temperature	ΔT5			X			
HP turbine efficiency	ΔEtaTHP			X			
HP turbine massflow deterioration	ΔmfTHPDeg				X		
LP turbine exhaust pressure	Δp6						
LP turbine exhaust temperature	ΔT6					X	
LP turbine efficiency	ΔEtaTLP					X	
LP turbine mass flow deterioration	ΔmfTLPDeg						X
Power	ΔP						

Table 81 Fuzzy logic failure classification set up