A Hybrid Model for on-line Detection of Gas Turbine Lean Blowout Events

Matteo Iannitelli[1], Carmine Allegorico[2], Francesco Garau[3] and Marco Capanni[4]

[1,2,3,4]*Baker Hughes, a GE Company, Firenze, 50127, Italy*

matteo.iannitelli@bhge.com
allegorico.carmine@bhge.com
francesco1.garau@bhge.com
marco.capanni@bhge.com

ABSTRACT

Modern dry low NOx combustors can target very low emissions levels by operating at a lean air/gas ratio. However, ultra-lean combustion is extremely susceptible to thermoacoustic combustion instabilities and Lean Blowout (LBO), which can lead to large pressure oscillations in the combustor and decreased durability of components.

Conventional on-board diagnostics embedded in the Unit Control Panel (UCP) of a Gas Turbine (GT), continuously check the health status of the combustion section at a high scan rate and raise alarms when abnormal conditions occur. While ensuring protection and control, UCP control logics may not provide precise indications on the nature of the issue and further troubleshooting, also using specific tools, is typically required.

In a changing environment where Industrial Internet of Things (IIoT) is offering a chance to drive productivity and growth, online Monitoring and Diagnostic (M&D) software and services on connected units are becoming strategic to increase asset availability and reliability, as well as reducing maintenance costs.

In this paper, we present a hybrid analytic, which combines physics-based and data-driven models, for the detection of Lean Blowout conditions on Gas Turbines equipped with Dry Low NOx multi-can combustion system. Regarding the data-driven model, we face a problem of classification and exploit dimensionality reduction to reduce the number of variables under consideration. During the development, different techniques are tested and benchmarked.

The analytic is trained on real LBO events and finally is deployed in a production environment to process incoming on-line data acquired from monitored units. Obtained results are presented.

1. INTRODUCTION

Rotating equipment and in particular gas turbines are often among the most critical items in Oil&Gas plants serving various applications, such as Liquefied Natural Gas (LNG), pipeline, refinery and petrochemical.

Today, one of the most important aspects for the operators is to ensure the highest level of availability of the engine during its entire life-cycle. Unscheduled shutdown of the gas turbines can have impact on the whole plant downtime with associated significant loss of production.

In this era of connected devices, plants operators often rely on connecting the most critical assets to remote diagnostic centers to support fast troubleshooting and optimization of operations & maintenance. A typical M&D workflow can be described as follow: data from on-board sensors is continuously stored in remote and/or local databases/cloud; the quality of the acquired sensors is checked; various analytics automatically process the incoming data to detect deviations from the expected behavior and to calculate asset performance KPIs; in the event that anomalies are identified, a dedicated analysis is started with subject matter experts and operators, and corrective actions are provided to site (Ozgur et al., 2000).

Baker Hughes, a GE company, implements a similar approach for its monitored fleet, consisting of more than 1000 rotating equipment. This large number of asset requires that all the analytics running in production have a high detection rate and low false positives. Furthermore, a continuous improvement process is required to sustain the fleet growth and to increase the detection capabilities.

Given the critical nature of this context, the ability to promptly detect potential anomalies and provide robust diagnosis with corrective actions in few hours is one of the major challenges facing M&D today and is mission-critical for all operators. The key factors that determine the effectiveness of the M&D service are: deep knowledge of the assets, quality of data and analytics performance.

2. THE LEAN BLOWOUT EVENT

Starting from 1970s, when the emission controls were originally introduced (Pavri & Moore, 2001), Countries regulations have pushed the GT manufacturers to design and develop combustors capable to meet emission requirements. Regulations on pollutants, have become increasingly severe over time and technology moved from systems based on water/steam injection to Dry Low NOx (DLN) technology (Davis & Black, 2000). Nowadays modern combustors can target single digit NOx and CO emissions, with ultra-lean premixed combustion in very narrow air/fuel equivalence ratio. However, ultra-lean combustion is extremely susceptible to thermoacoustic instabilities and lean blowout phenomenon.

In heavy duty gas turbines equipped with DLN multi-can combustors, LBO manifests itself with partial or complete flame-out of one or more combustion chambers and can lead to different consequences, based on the severity of the causes that activated it. Incipient LBO can randomly occur and disappear without significant impact on gas turbine operation, e.g. during load or combustion mode transients (Rebosio et al., 2011), or can take place with complete flame-out and consequent unscheduled engine stop. The latter results in an unavailability of the gas turbine, with accelerated degradation of parts and eventually loss of production.

Potential sources of LBO have been identified in variation of the fuel gas composition, wrong fuel split, improper operation/tuning of control components, instrumentation failure or shift in calibration and issues of the combustion hardware.

Tuning of the DLN system, i.e. optimizing the fuel streams distribution in the combustion chambers over the whole operating range, is therefore required to find the best trade-off between emissions level, margin from LBO and acceptable level of dynamics. Even if the DLN system is properly tuned, extreme changes in ambient condition (in general any factors impacting the fuel/air ratio) may potentially impact the emissions or reduce the LBO margin (Muruganandam et al., 2005).

When a severe LBO event occurs, the control system automatically shuts down the engine for “High Exhaust Spread” or “Loss of Flame”. However, clear evidences of LBO and the related corrective actions cannot be provided without the implementation of specific detection algorithms. The alarms generated by the control system in these cases are not very effective for the troubleshooting activity, with consequent impact on the time required to figure-out and correct the issue. Under these circumstances, the availability of a remote monitoring service can be extremely useful.

3. METHODS

Modern science and engineering is historically based on the usage of physics-based models, which are usually developed during the design phase of complex systems. The biggest advantage of using physics-based models is that they are based on physical equations and provide a sufficient explanation of the problem under analysis. Unfortunately, these models require substantial engineering time to be developed and in some cases, accurate models cannot be obtained due to complex or unknown physics of the system.

On the other hand, the recent exponential growth of data is supporting the development and the diffusion of new approaches purely based on data. These data-driven models are built more easily by collecting measurements recorded over the operating range of the system, and relationship between the sensor measurements are learned or embedded in the model architecture through mathematical techniques. Peculiarity of these data-driven models is that they are only accurate in the learned space, so if the system operation changes significantly, the model is forced to extrapolate and the result cannot be trusted.

The combination of data-driven and physics-based models is called hybrid modeling. This hybrid framework, if well designed, has the advantage of exploiting the strengths of both methodologies resulting in much better overall performance (Hines et al., 2008).

In this paper we present a hybrid approach, where the challenging task of detecting LBO signatures in operational data is demanded to a data-driven model to simplify the modeling approach; then the results are further validated by a physics-based block, designed on the observations of real LBO events.

3.1. Feature Engineering

The feature engineering process is a fundamental step of the development of any intelligent analytic. It consists in transforming raw data into usable set of predictors (features) that better represent the underlying problem and make the machine learning algorithms work. Since each problem is domain specific, the selection of the best features is often the deciding factor of the resulting performance.

For the problem of detecting lean blowout events of gas turbine combustors, the measurements of the thermocouples at the exhaust section, operating parameters and other calculated variables must be considered. In this work, the feature engineering process consisted in selecting the minimum number of predictor variables for the model development. A further simplification of the problem was then obtained by applying techniques of dimensionality reduction.

In statistics, the dimensionality reduction is a strategy that allows to convert data from high dimensional space to low dimensional space. The most popular methods are Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). Both methods were used to achieve better model's performance, reducing the risk of overfitting and improving computational efficiency.

3.2. Dimensionality Reduction Methods

The Principal Component Analysis is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. The PCA transformation aims to find the directions of maximum variance within a high-dimensional dataset, and projects these directions into a new orthogonal subspace Z with dimensions k equal or lower than the original space (typically $k<<p$). The principal components of the new subspace Z can be interpreted as the directions of maximum variance along which the observations vary the most. All subsequent major components have the highest possible variance, since they are not correlated to the other main components as they are orthogonal. The PCA technique is widely used in various fields, mainly in exploratory data analysis to reduce the dimensionality of a high-dimensional problem, and for making predictive models by increasing computational efficiency and reducing the degree of overfitting.

Similarly to PCA, the Linear Discriminant Analysis (LDA) is a statistical method that looks for linear combinations of variables which best explain the original dataset. The main difference from the PCA is that, under certain assumption, LDA uses the class information, when already present in the training dataset, to maximize the separability between classes. Therefore, LDA analysis, if compared to PCA, is a superior technique for extracting features in a classification task (Duda, Hart and Stork, 2001).

3.3. Classification

In machine learning and statistics, classification is the problem of classifying examples into given set of categories based on past observations. Today, many classification techniques, or classifiers, are available in literature. Some of these are: Logistic Regression, Artificial Neural Networks, K-Nearest Neighbors, Decision Tree and Support Vector Machines (James et al., 2013). In this work we investigated Logistic Regression and Decision Trees.

Logistic regression algorithm is one of the most widely known algorithms for classification of linear and binary problems and, with appropriate techniques, it can also be extended to problems of multiclass classification. In a binary classification problem, where the response y falls into one of two categories, 0 or 1, logistic regression uses the logistic or sigmoid function $g(z) = \frac{1}{1+e^{-z}}$ to predict the probability that y belongs to a particular category.

Unlike other classification algorithms, decision tree models divide the original dataset by learning decisions based on the answers to a series of questions. Decision trees try to solve the problem by using a tree representation, where each internal node of the tree corresponds to an attribute (feature), and each leaf node corresponds to a class label. The process starts from the root of the tree and recursively divides the data according to the feature that produces the best split. The recursion is completed when all the samples in a node belong to the same class, or when further splitting no longer adds value to the predictions. In practice, this can produce a very deep tree with many nodes, which can easily generate overfitting if not properly controlled with appropriate techniques. Nevertheless, the biggest advantage of the decision tree algorithms, is that they are simple to understand and interpret (Raschka, 2015).

4. DESIGN AND DEVELOPMENT

The analytic presented in this work is designed for heavy-duty gas turbines with DLN multi-can combustion system. As previously explained, LBO events may occur abruptly and may last even a few seconds. This requires the analytic to run on operational data with sampling frequency of at least one second.

If a LBO event occurs, as the combustor loses the flame, the gas/air mixture continues entering the combustion chamber with reduced or absent burning taking place, thus producing two immediate consequences: load reduction and arise of a cold spot in the Exhaust Gas Temperature (EGT) profile. The magnitude of the power loss will depend on the load level that the gas turbine was delivering just before the event, and on the severity of the event itself, e.g. partial or complete flame-out of one or more combustion chambers. The distortion of the EGT profile, typically involves more than one thermocouple (TC) reading, as also shown by Allegorico and Mantini (2014). In fact, the number of thermocouples of these gas turbines is around twice the number of combustors, thus it follows that multiple adjacent readings showing deviations can be indicative of a real combustion issue. An example of distorted EGT profile can be seen in Figure 1, where the blue dashed line shows the normal operation, while the red line shows the cold spot area typical of a LBO event.

From above observations, it can be deduced that the detection of LBO events requires at least the identification of the following patterns:

- substantial power drop, in terms of absolute value and first derivative. Depending on engine application (generator or mechanical drive), measurements necessary are the electrical power or the shaft speed.

- asymmetric EGT profile showing a cold spot area, with at least two TCs readings colder than the average temperature.
- high spread (i.e. maximum – minimum) among the EGT, as direct consequence of the cold spot.

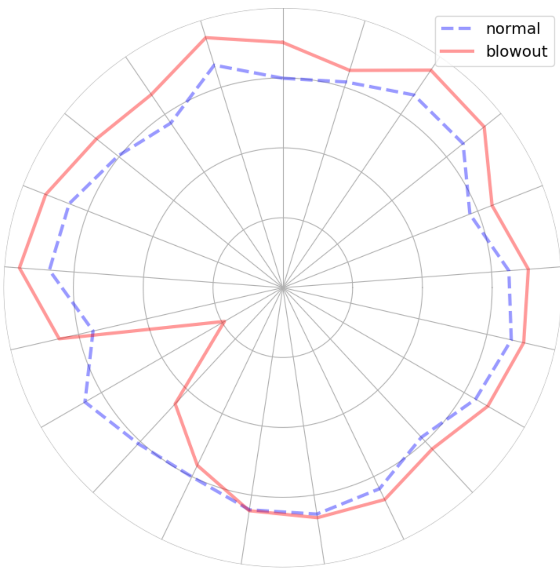

Figure 1 - Blowout effect on EGT profile

As already explained, the hybrid model presented here is composed of two parts: one physics-based and the other data-driven. The physics-based model detects the sudden drop of power by verifying that the accounted power loss and its backward finite difference are below certain (negative) thresholds, which have been derived from the analysis of real LBO events. The second part instead, is the data-driven classifier, which detects the presence of cold spots and high spread in the EGT profile. This data-driven model has been trained over real data taken from known blowout events. The entire code has been developed with Python language and the predictive models have been built with the scikit-learn library for machine learning (Pedregosa et al., 2011).

4.1. Data-driven classifier

The availability of large amount of data is an important factor when building data-driven models; however, in real-world applications, positive events may be a limited number (Allegorico & Mantini, 2014). To overcome this problem, sometimes additional data can be generated using artificial data synthesis methods (Surendra & Mohan, 2017).

For the training and testing of our data-driven classifier, we exploited M&D data opportunely selected from about 20 real events of variable length, ranging from 15 seconds to 6 minutes. The sampling rate of these datasets was 1 sample per second.

The categories to be classified are two: fault-free and blowout. For most classifiers, a better performance is achieved if the training dataset is balanced among the classes. Even if the availability of data is much higher for the first category, we manually selected a number of fault-free samples comparable to those including blowout signature.

As discussed in section 3.1, when setting a classification problem, one of the first steps is the features selection. Based on the failure signature that has been described in Section 4, we selected the following set of four measurements for the detection of cold spot and high spread conditions:

- the lowest TC reading and its two adjacent TCs;
- median of the spreads calculated subtracting the maximum temperature and the four lowest ones.

This set of four variables is the outcome of an iterative process based on the analysis of the results of different classification tests and which proved to be the most representative and robust.

Since instrumental failures may compromise model results and generate fault alerts (e.g. broken thermocouples get negative full-scale values), we used a set of standard algorithms of data pre-processing, currently implemented in our M&D systems, to detect any potential instrumental issue. Example of checks performed are the signal range, excessive noise or signal freeze: all unreliable samples are excluded from calculations.

Before model training, data is standardized subtracting the mean from each feature and then dividing the value of each feature by its standard deviation. Standardization is useful to not get conditioned by the feature magnitude while computing model parameters. After this step, a split between training and test dataset is performed. Considering the volume of data, we decided to keep 80% of dataset for training and remaining 20% for testing purposes.

At this point, dimensionality reduction techniques are applied and benchmarked, specifically the Principal Component Analysis and the Linear Discriminant Analysis.

By applying PCA to the training dataset, it came out that the first two principal components retain about 96% of the entire problem variance. Hence, it was convenient to reduce the original 4-dimensional space $\mathbf{R}^4$ to a 2-dimensional space $\mathbf{R}^2$, without significant information loss.

The score plot of Figure 2 shows the projection of the data onto the first two PCs, where the green round markers represent normal conditions, while the red stars are the blowout conditions. For classification purpose, a linear decision boundary could be enough accurate to separate the two categories. To this purpose, we used the logistic regression and the decision tree algorithms, described in Section 3.3.

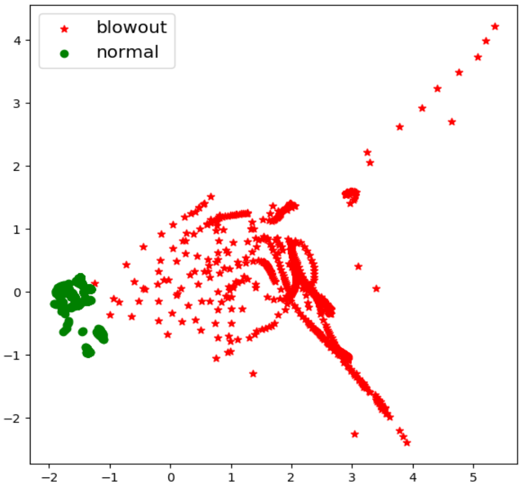

Figure 2 - Dataset transformed in bi-dimensional space using PCA.

For the implementation of the first method, the scikit-learn class called *LogisticRegression* was used with its standard parameters: first order polynomial decision boundary, default value for the regularization parameter λ, and the "liblinear" optimization solver, which implements a conjugate gradient method (Fan et al., 2008).

For the decision tree model, we selected the information entropy as the criterion to measure the quality of a split, and tuned the minimum samples per split and the maximum depth of the tree to avoid overfitting.

As explained before, we also implemented the LDA method to reduce the dimensionality of the problem. To make sure that the hypothesis under which LDA is applicable are valid, data was double-checked for outliers and a good balance between classes samples was ensured. The dimensionality of the LDA output is expected to be the number of problem classes - *1*: since we are handling a binary classification problem, the output is mono-dimensional.

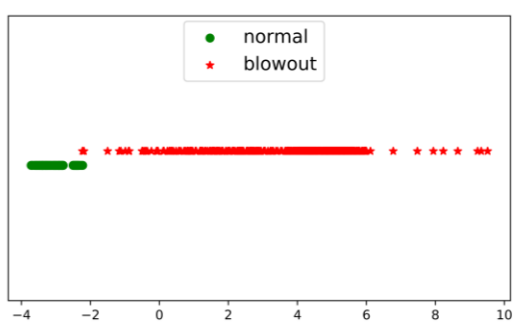

Figure 3 - Dataset transformed in mono-dimensional space using LDA.

Projection of the samples in the output space is shown in Figure 3, where green round markers represent the normal class, while the red stars are the blowout samples. The two sets are vertically offset for a better visualization.

The separation between the two categories is quite evident. By setting the threshold $x = -2$ as a separation boundary, we got only one sample wrongly classified.

5. RESULTS ON REAL DATA AND IMPLEMENTATION

The three data driven models, PCA + Logistic Regression (LR), PCA + Decision Tree (DT) and LDA + threshold (th) were validated on real M&D operational data.

Purpose of the validation was to check if each sample is correctly placed in the right class; performance results are summarized in a confusion matrix (James at al., 2013). For this purpose, we used several batches of data, sampled 1 per second, containing real blowout events. Results obtained for each of the three models are shown in the confusion matrix of Table 1.

		Target	
	Prediction	Blowout	Normal
PCA + LR	Blowout	17.2%	1.5%
	Normal	0.6%	80.7%
PCA + DT	Blowout	17.6%	0.5%
	Normal	0.2%	81.7%
LDA + th	Blowout	17.2%	0.1%
	Normal	0.6%	82.1%

Table 1 - Confusion Matrices, percentages of the total amount of data

A deeper analysis of these results revealed that most False Negatives (FN), namely the blowout observations classified as normal (lower left of each submatrix), were due to some samples wrongly labelled during the manual selection of the training data. In particular, some portions of the test dataset are transient's data and they could neither be categorized as "normal", nor as "blowout". Instead, the False Positives (FP), namely the normal samples predicted as blowout (upper right of each submatrix), are due to the wrong labelling of the samples near the blowout event and the improper classification of other combustion issues different from LBO.

Some additional metrics, precision and recall, are also calculated. Precision measures the probability that an event selected by the classifier is relevant (true), while recall represents the classifier capability to retrieve a relevant (true) event. They are defined as follows:

$$precision = \frac{true\ positive}{true\ positive + false\ positive}$$

$$recall = \frac{true\ positive}{true\ positive + false\ negative}$$

A combination of them is the metric F1-score, which is the harmonic mean of the two, formulated as:

$$F1 = 2 \cdot \frac{precision \cdot recall}{precision + recall}$$

In this study we avoided using the *accuracy* metric, which is defined as the ratio between well-predicted and overall cases, since a predictive model may have high accuracy, but be useless (Valverde-Albacete & Peláez-Moreno, 2014).

The overall metrics are reported in Table 2 and show that LDA + th model is the better candidate to build the hybrid-model, since it gives slightly better performance.

	Precision (%)	*recall* (%)	*F1-score* (%)
PCA + LR	92.2	96.6	94.3
PCA + DT	97.2	98.7	97.9
LDA + th	99.7	96.7	98.2

Table 2 - Precision, recall and F1-score for the three methods.

The final hybrid-model, which results from the aggregation of the physics-based and the data-driven part, is then validated by verifying that the algorithm is able to correctly detect some known LBO events. Purpose of this diagnostic analytic is to raise an alarm when a LBO event is detected; to avoid repeated alarms, the alert is triggered only if all the conditions are verified for at least 5 seconds. An example of a successful catch is in Figure 4, which shows respectively the EGT readings, calculated spreads, power and shaft speed, classifier output and hybrid model output. Note that a failed thermocouple (TC) is identified and the LBO event lasts less than 10 seconds.

The final hybrid-model, tested on all real available LBO events, proved capable of capturing all cases without false positives, and it achieved following performance: 100% precision *TP/(TP+FP)* and 0% False Discovery Rate *FP/(FP+TP)*.

The analytic was finally deployed on a proprietary platform to run on the on-line operational data acquired on all gas turbines under monitoring. Since this analytic has to process a huge amount of data, to maximize computational resources we implemented a cascade architecture, where in the upper layer a simple code uses 1-minute data to check if the turbine is running, and only under such condition the blowout analytic is executed on faster data (1 per second).

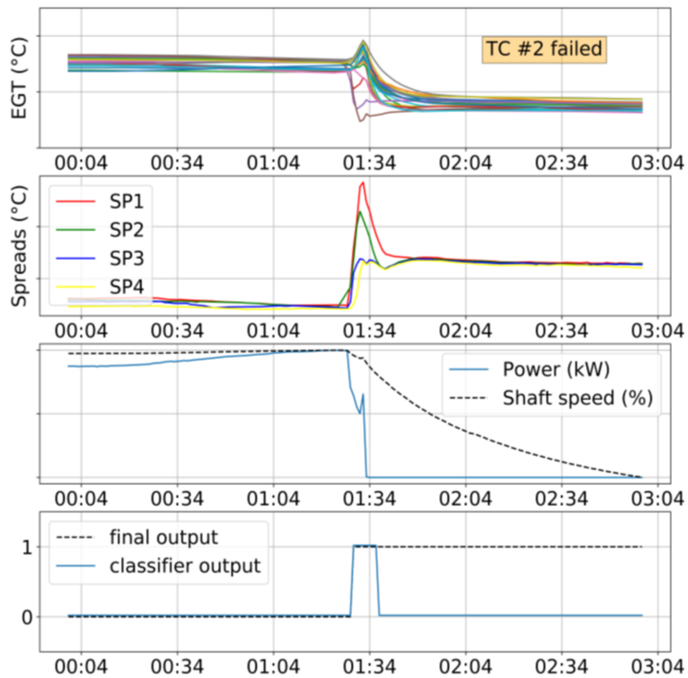

Figure 4 - Blowout event and its identification

6. CONCLUSIONS

Nowadays, monitoring and diagnostic of gas turbines and other critical equipment is getting growing importance in the industry. An effective M&D service can only be possible by merging together the knowledge of subject matter experts and intelligent analytics in a single workflow. Within this context, the analytic presented in this work has been designed to enable the quick identification of the lean blowout phenomenon which occurs on modern gas turbines equipped with Dry Low NOx multi-can combustion system. The advantages of such analytic are:

- speed up and simplify troubleshooting activities;
- catch the events that could remain undiscovered.

The latter typically allows corrective and preventive actions to be taken more effectively; in fact, about 15% of the observed LBO didn't cause a machine unscheduled stop, since the combustor was able to relight and the unit continued running: recognizing the phenomenon in this condition would have been extremely difficult.

To build this capability, a hybrid model has been developed and it consists of two main components: a physic-based model and a data-driven model. The first model checks the occurrence of a power drop, while the second one uses a machine learning classifier to detect blowout signatures in the EGT profile. Training and testing data are properly selected from a proprietary database of known cases of combustion blowout. The classification problem is solved by firstly reducing the dimensionality of the problem, and then implementing a classifier algorithm. Three different methods have been tested: PCA + Logistic Regression, PCA + Decision Tree and LDA + threshold. It was demonstrated that the last approach is the one giving slightly better performance, so it was selected to be implemented, together with the physics-based model, in the M&D platform. Appropriate precautions have also been taken to manage any problems of data quality and to avoid overload of computational resources due to the amount of data being processed.

Future research focuses on the identification of blowout causes, which could lead to further acceleration of the current troubleshooting process, and on the use of this technology for the control and optimization of gas turbine combustion. Possible enhancements of this analytic will probably require the use of additional sensors, such as combustion chamber pulsations, CO and NOx emissions and other. Another improvement could be the use of the swirl angle to automatically identify the combustor chamber affected by blowout.

ACKNOWLEDGEMENT

We would like to thank Francesca Barbara Rebosio for the review of the paper.

REFERENCES

Allegorico C., Mantini V., (2014). A data-driven approach for on-line gas turbine combustion monitoring using classification models. *European Conference of the Prognostics and Health Management Society*, July 8-10, Nantes, France.

Ozgur D., Lakshminarasimha A., Rucigay R., Morjaria M., Sanborn S., (2000). Remote Monitoring and Diagnostics System for GE Heavy Duty Gas Turbines. *Proc. ASME. 78569; Volume 3: Heat Transfer; Electric Power; Industrial and Cogeneration*, May 08, Munich, Germany.

Duda R.O., Hart P.E. and D.G. Stork (2001). Pattern Classification, *Journal of Classification*, Vol. 24, Issue 2, pp. 305-307.

Cohen J., & Anderson T. (1996). Experimental investigation of near-blowout instabilities in a lean, premixed step combustor, *34th Aerospace Sciences Meeting and Exhibit*, Jan 15-18, Reno, USA.

Davis L. B., & Black S. H. (2000). Dry Low NOx combustion systems for GE Heavy-Duty Gas Turbines. GE Power System report GER-3568G.

Fan R.E., Chang K.W., Hsieh C.J., Wang X.R., & Lin C.J. (2008). LIBLINEAR: A library for large linear classification. *Journal of Machine Learning Research*, vol.9, pp. 1871-1874.

Muruganandam T. M., Nair S., Scarborough D., Neumeier Y., Jagoda J., Lieuwen T., Seitzman J., Zinn B., (2005). Active Control of Lean Blowout for Turbine Engine Combustors, *Journal of Propulsion and Power*, Vol. 21, No. 5, pp. 807-814.

Pavri R., & Moore G. D. (2001). Gas turbine emissions and control. GE Power System report GER-4211.

Pedregosa et al. (2011). Scikit-learn: Machine Learning in Python, *Journal of Machine Learning Research*, vol.12, pp. 2825-2830.

Rebosio F., Di Domenico M., Noll B., Aigner M. (2011). Numerical analysis of combustion instability mechanisms in a lean premixed can combustor, *49th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition*, July 4-7, Orlando, USA.

Surendra H., Mohan H. S. (2017). A review of synthetic data generation methods for privacy preserving data publishing, *International Journal of Scientific & Technology Research*, vol.6, issue 3.

Valverde-Albacete F. J., & Peláez-Moreno C. (2014). 100% classification accuracy considered harmful: the normalized information transfer factor explains the accuracy paradox. *PLoS ONE 9(1): e84217*. doi: 10.1371/journal.pone.0084217.

Hines J.W., Garvey D., Seibert R., & Usynin A. (2008). Technical Review of On-Line Monitoring Techniques for Performance Assessment: Theoretical Issues. *NUREG/CR-6895, ORNL/TM-2007/188, Volume 2*.

Raschka S., (2015). Python Machine Learning. Packt Publishing, Birmingham, UK, isbn 1783555130.

James G., Witten D., Hastie T., & Tibshirani R., (2013). An Introduction to Statistical Learning. NY, USA: Springer Science + Business Media.

BIOGRAPHIES

Matteo Iannitelli is a Data Scientist at Baker Hughes, a GE Company, Firenze, Italy. He received his Master Degree in Automation Engineering from University of Pisa, 2011. He currently works in Monitoring & Diagnostic department developing and implementing algorithms for predictive maintenance and failure identification. He is also leading the functional development of an internal big data analytics platform, and is involved in pilot projects on risk analysis and maintenance.

Carmine Allegorico is a Principal engineer and experienced data scientist at Baker Hughes, a GE Company, Firenze, Italy. He received his Master Degree in Mechanical Engineering from University of Napoli Federico II. In his current role, Carmine is a technical point of reference for the analytics discipline providing engineering guidance to other teams, helping to train new engineers and keeping abreast of industry trends and issues. He provides consulting during the development and implementation of advanced solutions for the on-line diagnostic and predictive maintenance, coordinates the creation of internal processes and support the adoption of new platforms and technologies.

Francesco Garau is a Service Engineer at Baker Hughes, a GE Company, Firenze, Italy. He received his Master Degree in Mechanical Engineering from Politecnico di Torino. He started collaborating with GE Oil & Gas since 2008 as Diagnostic Engineering in the RM&D iCenter in Firenze. In that position he also led the activity for signals configuration management on the new units connected remotely. Since 2014 he is part of the Product Service Team as Heavy-Duty Gas Turbine specialist. In this role, he continues collaborating actively with the RM&D supporting analytics development and fleet management as well as with Product Design Team for the continuous improvement of gas turbine product.

Marco Capanni is a Service Engineer at Baker Hughes, a GE Company, Firenze, Italy. He received his Degree in Mechanical Engineering from University of Firenze. He started collaborating with GE Oil & Gas since 2011 as Diagnostic Engineering in the RM&D iCenter in Firenze.