Anomaly and Failure Prediction in Gas-turbine Using Statistical Analysis

Shinjan Ghosh, Vipul Goyal, Mengyu Xu, Jayanta Kapat, Ladislav Vesely
Center for Advanced Turbomachinery and Energy Research, 12781 Ara Dr., Orlando, FL 32816, USA

OBJECTIVES
In the field of gas-turbines, the methods currently in operation rely on physics-based knowledge alone to predict the failure and to estimate the maintenance schedule. This reliance on physics-based methods alone tends to make the maintenance schedule overly conservative and thus costly. By incorporating the data-based algorithms in the operational parameter prediction (e.g., power, pressure, temperature) and failure detection of gas turbine, higher operational efficiency can be achieved. Moreover, these algorithms can be applied to system-wide data and thus uncover hidden pattern which may not be discovered using physics based methods alone.

To implement this, time-series data has been used from the combined cycle utility gas turbines consisting of three gas-turbine units and one steam turbine unit. In this study the attempt has been made to apply and integrate various statistical algorithms to predict the anomalies in the gas turbines using the time-series history of operational gas turbine units.

INTRODUCTION
In this analysis, 8 different data sets have been used with each having more than 200 variables and 200,000 time series values. These data sets are from different but identical gas turbines and have not been taken simultaneously. These data-sets are not known to have identical failure among them which makes it a very important consideration for classification based application because the classification algorithms in the initial phase of learning require significant amount of similar cases to be able to learn effectively. Moreover, the number of failures are known to be sparse, i.e. every data set has no more than 2 failures and there is no specified definition/criterion of failure/Anomaly for analysis, i.e. the definition of failure is broad in nature in the sense that the system deviated significantly from ideal behavior.

DATA ANALYSIS METHODS

In this study we establish statistical models on historical gas path data. Through these statistical models we find hidden patterns in the model with methods like Granger Causality Networks. Using these models we also achieve one-step-ahead prediction which is then used to analyze the prediction errors. The behavior is deemed as anomalous if the prediction error has a magnitude that exceeds 3 standard deviations.

In the figure above we show the nonlinear relationship between the temperature spread and each of the four predictors. In the graph on the top of the next column, the turbine cross sectional temperature spread is shown in black and the predicted temperature spread may be an exact but complex function of some unknown variables but if we can capture a significant variation in the spread using thermodynamic variables alone then it can be used to predict anomalous. The graph below reveals the nonlinearity between the temperature spread and each of the four predictors.

RESULTS

With the help of Threshold VAR model we can get the Granger Causality networks. These networks help us in determining the failure propagation in the case of anomaly detection. The graph below shows one such network. The number correspond to different variables and top and bottom correspond to values at consequent time. These networks only capture the linear relation among variables.

With the help of GAM models we can find the relation among variable for which no physics based relation is readily available. For example we can model the turbine cross sectional temperature spread as a function of turbine gas path variables. In reality the spread may be an exact but complex function of some unknown variables but if we can capture a significant variation in the spread using thermodynamic variables alone then it can be used to predict anomalies. The graph below reveals the nonlinearity between the temperature spread and each of the four predictors.

CONCLUSIONS

No one method is expected to cover all types of anomalies but a combination of statistical methods can be used to cover a wide swath of expected anomalies which can be easily detected. Still, to know whether the broad behavior is anomalous or not, we need continuous help of physics-based methods.

The approaches of data analytics are presented here, and the potential limitations are discussed, it should be pointed that a multi-faceted approach will be necessary to have a robust digital twin where failures can be anticipated in advance so that corrective actions can be taken. In other words, pure data-driven approach, may not be enough. A low-order, physics based model must be operated in tandem based on the latest system parameters in order to enhance and interpret the findings from the data-driven process.

REFERENCES