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 systemwide 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.

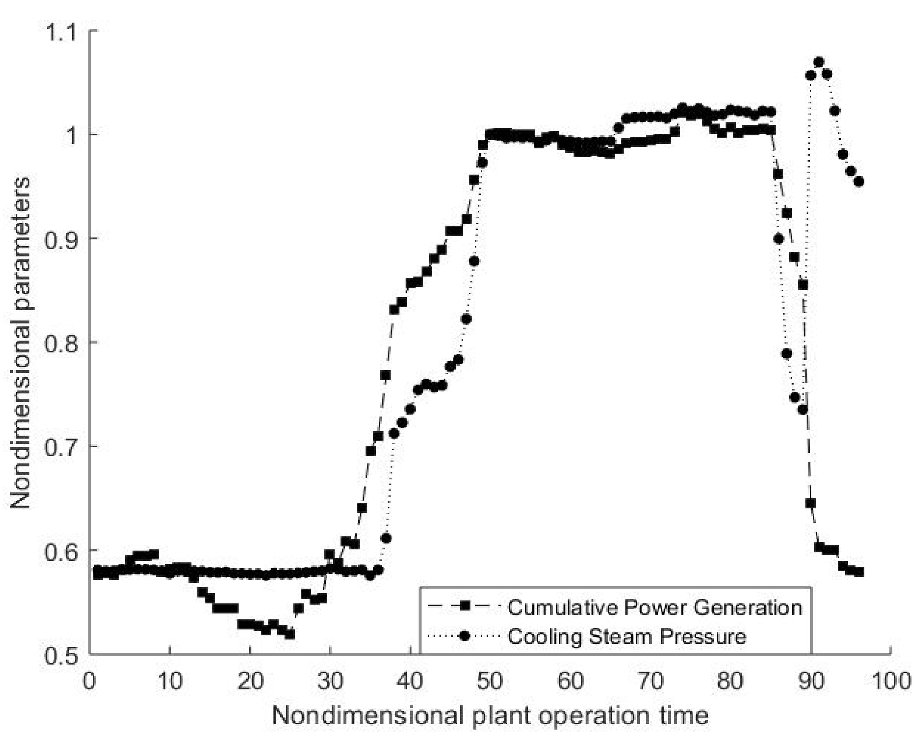

Figure 1. An example of unexpected change in cooling steam pressure.

The idea of method can be broadly explained with the help of figure above. In this figure cumulative combined-cycle power generation is shown along with cooling steam pressure in one particular gas turbine unit. After extensive data-analysis, if it is ascertained a priori that cooling steam pressure and power generation are highly correlated statistically then the unexpected increase in the cooling pressure at the time of powering down can be flagged as an anomaly much before pressure reaches to extremely high levels as seen at top-right corner of the graph.

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.

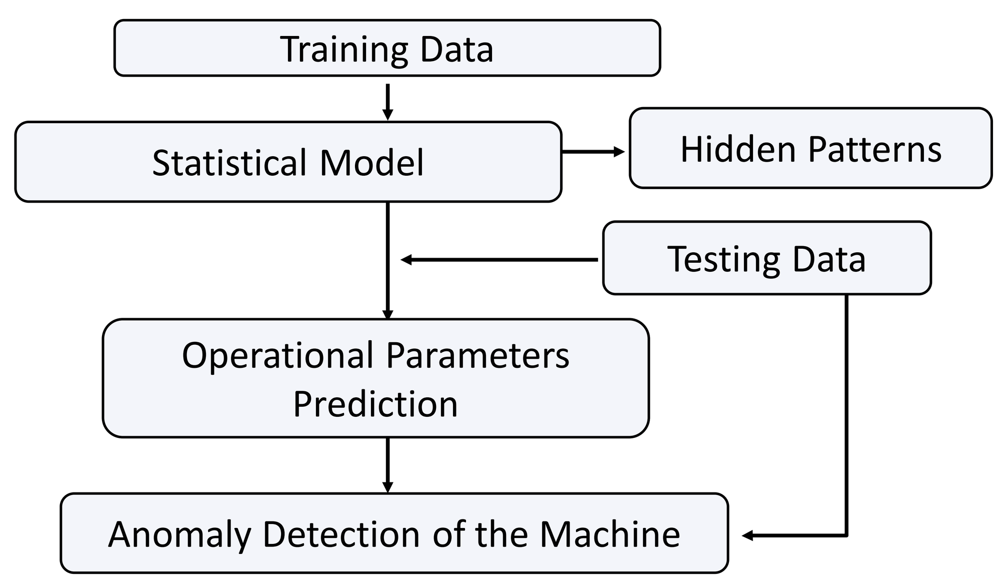

Figure 2. Predictive anomaly detection Procedure.

Two separate data-analysis methods have been used for anomaly prediction in the current study. The first method is a multi-stage threshold Vector Autoregressive Method. In this method we first segment the historical data into different stages according to exogenous variables such as rotating speed. Then the current values of the variables are modeled as the linear function of past values of the same variables and exogenous input variables. LASSO estimation is used to select the most important variable to reduce the dimensionality.

$$Y_t = c + \sum A_h Y_{t\text{-}h} + e_t + BX_t$$

Y_t – Output vector at time t,

$Y_{t\text{-}h}$ – Output vector at time t-h

e_t – Residual/error/noise, c – Intercept

A_h, B– Coefficient matrices,

X_t – Input vector at time t

The second method involves construction of a generalized additive model to predict the gas turbine performance. In our study turbine power, turbine temperature and turbine temperature spread are predicted under normal operation. The difference between predicted value and observed value are computed to detect the deviation from normal operating mode. The figure-5 on the right show the non-linear relationship between temperature spread and fuel flow, IGV angle, bypass valve position and combustion chamber pressure, respectively. The solid lines show the fitted function and the dashed lines show fitted function plus/minus two standard errors.

$$E(Y_t \mid X_{t,1}, X_{t,2}, \ldots\ldots X_{t,p}) = f(X_{t,1}, X_{t,2}, \ldots\ldots X_{t,p})$$

E(A|B) – Conditional expectation,

Y_t – Output variable at time t,

$X_{t,h}$ – Input variable h at time t

For simplification, assume no interaction among inputs

$$Y_t = c + f_1(X_{t,1}) + f_2(X_{t,2}) + \ldots\ldots f_p(X_{t,p}) + \varepsilon_t$$

f(X) – Smooth function of input X

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.

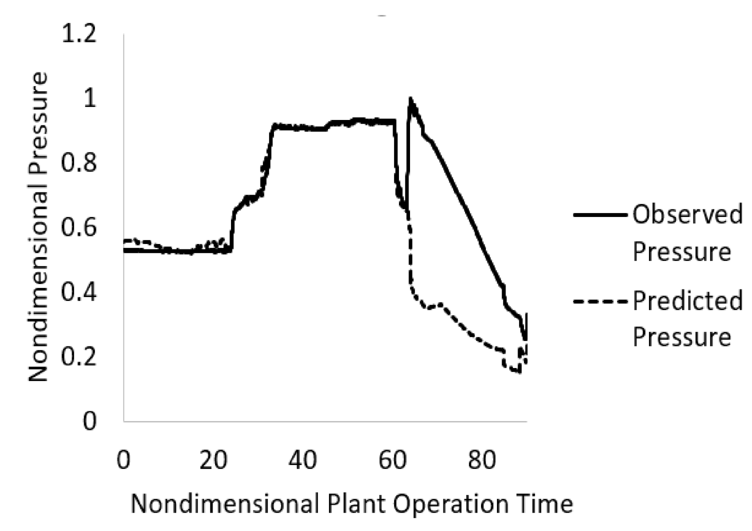

Figure 3. Non-dimensional one-step-ahead prediction in cooling steam pressure over a non-dimensional testing period.

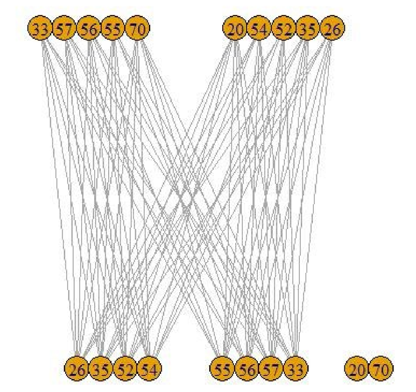

Figure 4. Granger Causality network between variables at the stage "on".

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 nonlinear relationship between the temperature spread and each of the four predictors.

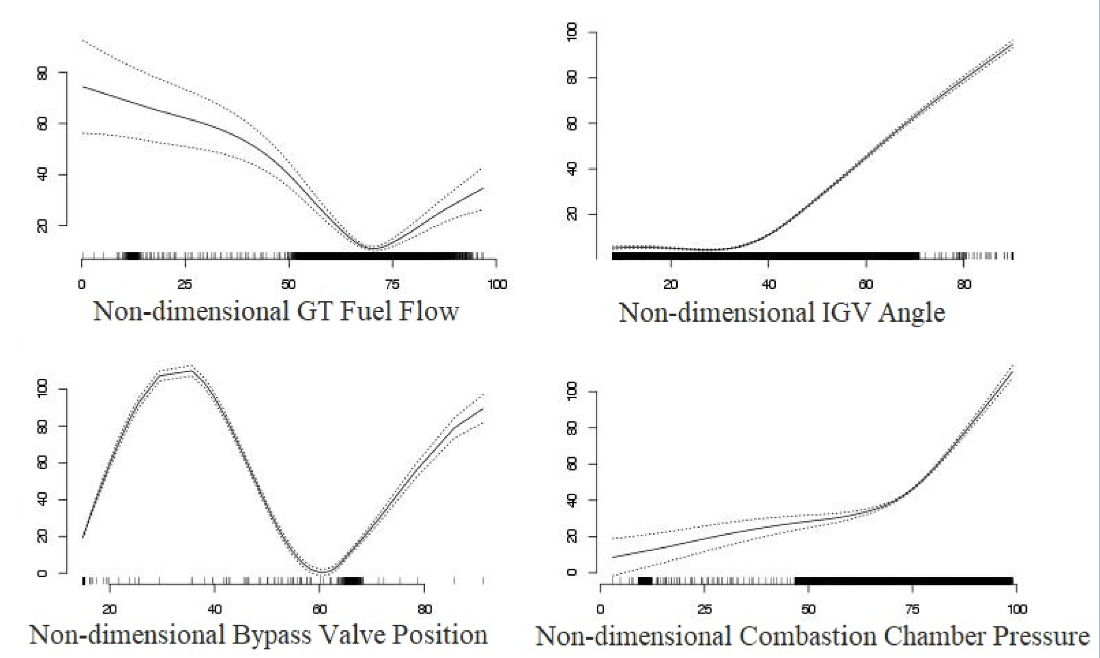

Figure 5. Relationship between the temperature spread and each of the four predictors, fuel flow, IGV angle, bypass valve position and combustion chamber pressure, respectively.

In the graph on the top of the next column, the turbine cross sectional temperature spread is shown in black and the predicted value from the GAM model are superimposed in the blue.

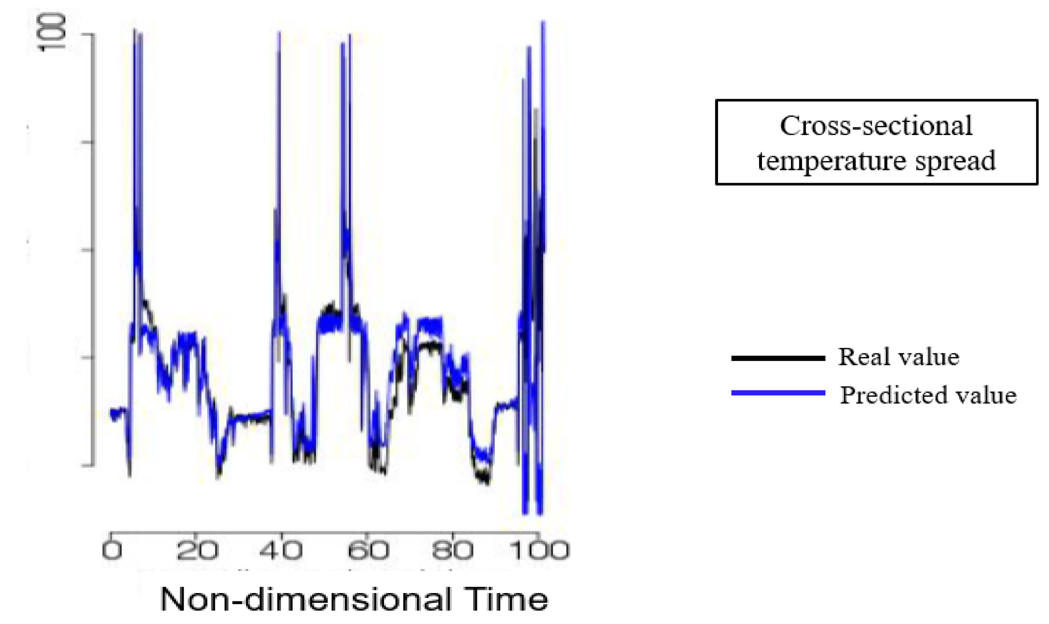

Figure 6. Non-dimensional prediction of the turbine cross sectional temperature spread from the GAM model.

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.

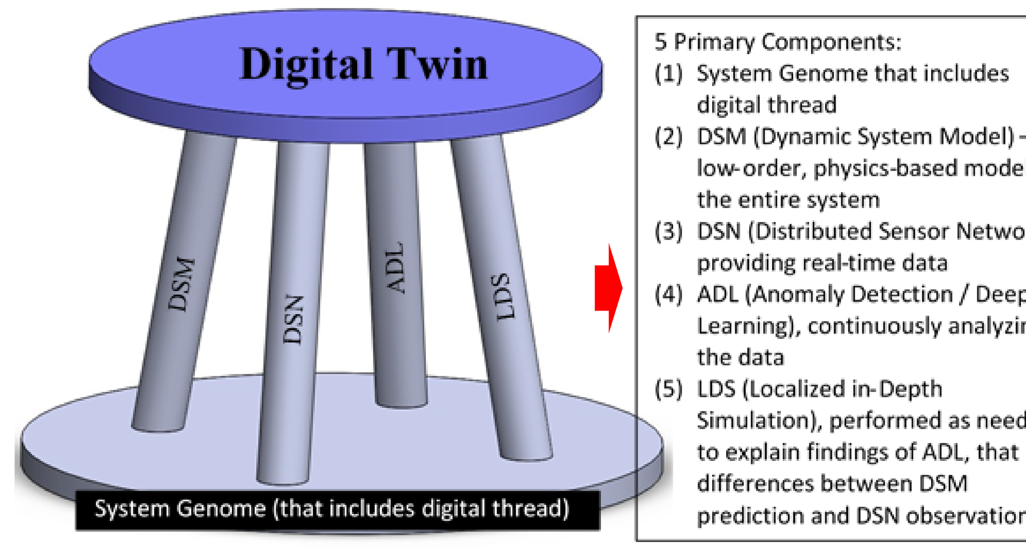

Figure 7. Components of digital twin with system genome.

REFERENCES

Goyal, V., Xu, M., and Kapat, J., "Use of Vector Autoregressive Model for Anomaly Detection in Utility Gas Turbines," *ASME Turbo Expo 2019: Turbomachinery Technical Conference and Exposition*, American Society of Mechanical Engineers, 2019

Goyal, V., Xu, M., and Kapat, J., "Statistical Modeling in Failure Detection in Gas Turbines," AIAA Propulsion and Energy 2019 Forum, 4088

CONTACT

vipulgoyal1211@knights.ucf.edu