Final Report

Project Title:
Real-Time Health Monitoring for Gas Turbine Components Using Online Learning and High-Dimensional Data

Submitted/Revised: 12/29/2021

WORK PERFORMED UNDER AGREEMENT

DE-FE0031288
Project Period: 10/01/2017 – 09/30/2021

SUBMITTED BY

Georgia Tech Research Corporation
505 Tenth St. NW
Atlanta, GA – 30332-0420
DUNs #: 09-739-4084

PRINCIPAL INVESTIGATOR

Principal Investigator: Nagi Gebraeel
P: 404.894.00541
F: 404.385.6178
nagi@gatech.edu

SUBMITTED TO

U. S. Department of Energy
National Energy Technology Laboratory

Omer Bakshi
omer.Bakshi@netl.doe.gov

1 Executive Summary

This report provides a summary of the research activities conducted over the four years of this project. These research activities were performed by three teams:

- Combustion Team: A team of engineers located in Georgia Tech that conducted experiments on combustor test-rigs. Team members include: Timothy Lieuwen, Benjamin Emerson, Nicholas Rock, and Raghul Manosh Kumar
- Turbine Team: A team of engineers located in Pennsylvania State University that conducted experiments on the turbine test-rig. Team members include: Reid Berdanier, Karen Thole, and Eric Deshong
- Data Analytics Team: A team of engineers located in Georgia Tech that aided both teams in designing experiments and performed data analytics for condition monitoring tasks. Team members include: Nagi Gebraeel, Kamran Paynabar, and Benjamin Peters

The report is organized as follows. Section 2 discusses the overall objectives of the project and outlines the data analytics framework used to address challenges related to the combustor and the turbine. Section 3 discusses the development of data analytics methodologies related to two combustor faults: lean blowout and centerbody degradation. Section 4 discusses the development of two data analytics methodologies for monitoring cooling faults in the gas turbine. Furthermore, the development of a new dataset for future analysis is also discussed. Section 5 outlines the dissemination of results for this project. The report concludes with a summary of accomplishments and potential for future work.

2 Objectives and Approach

The objective of this project is to enable the development of a Big Data analytics framework for critical gas turbine components through a systematic experimental program that leverages unique industry-class turbine test rigs. As mentioned in the proposal, the analytics framework consists of four key components:

- Data curation procedures that tackle data storage, data quality assessments, and integrity checks.
- Feature engineering, which revolves around using different data transformation algorithms guided by physics-based models to develop high-fidelity fault features that can be leveraged by subsequent fault detection algorithms and prognostic models.
- Machine learning-based fault detection and diagnostics algorithms constitute the third component of our analytics methodology.
- Remaining life prognostics of critical gas turbine components will be the focus of the fourth component of our methodology, prognostics and predictive analytics.

Advanced gas turbine test facilities will be interrogated using state-of-the-art instrumentation techniques to build fault signatures and data trends for key combustor and turbine faults. Data generated from a combustor test rig (Georgia Tech) and a turbine test rig (Penn State) during both normal operation and with “seeded” faults serve as the basis for the Big Data sets. The test conditions in the two test facilities include common, critical events that occur in the operation of

power plants. Figure 2.1 provides an overview of the research approach and has been extracted from the proposal narrative.

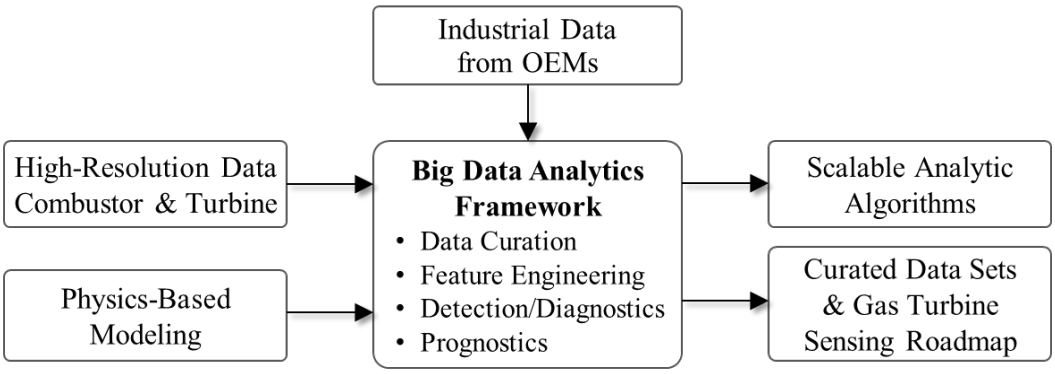

Figure 2.1. Research Objective, Scope, and Deliverables

3 Combustor

Modern emission standards have resulted in a transition from non-premixed combustion to premixed combustion. While enabling a reduction in nitrogen oxide (NOx) emissions, this transition has rendered modern combustors more susceptible to operational faults such as lean blowout and combustor instabilities. Therefore, it is important to detect precursors to these faults to maintain operability of the combustor. In this section, we highlight the contributions of both the combustion team and data analytics team as it pertains to combustor fault detection. Notably, we discuss experimentation and data analytics methodologies related to lean blowout and combustor centerbody erosion, the latter of which results from combustor instabilities.

3.1 Lean Blowout

Lean blowout, or the failure to stabilize a flame in a combustor, is a phenomenon that narrows combustor operability margins. The ground power generation heavy frame gas turbine industry is faced with this issue as combustor blowout will trip the plant and cause costly operational down time. This problem is also relevant to other industries, such as gas turbine engines for aircraft. The working solution is to maintain a combustor at an equivalence (fuel-to-air) ratio ϕ well above the lean blowout limit ϕ_{LBO}. However, modern low NO_x systems operate at low equivalence ratios by design for emissions reasons. Advanced lean blowout detection would enable increased gas turbine turndown and therefore wider operability range. Advanced detection methods require detailed combustor rig data and improved understanding of lean blowout physics. Lab-scale experiments are helpful for this purpose, as they provide optical access and flexibility in a way that is impossible in an actual engine.

In this study, the combustor test-rig located at the Ben T. Zinn Combustion Laboratory at Georgia Tech was used to acquire lean blowout data. This test-rig is displayed in Figure 3.1.

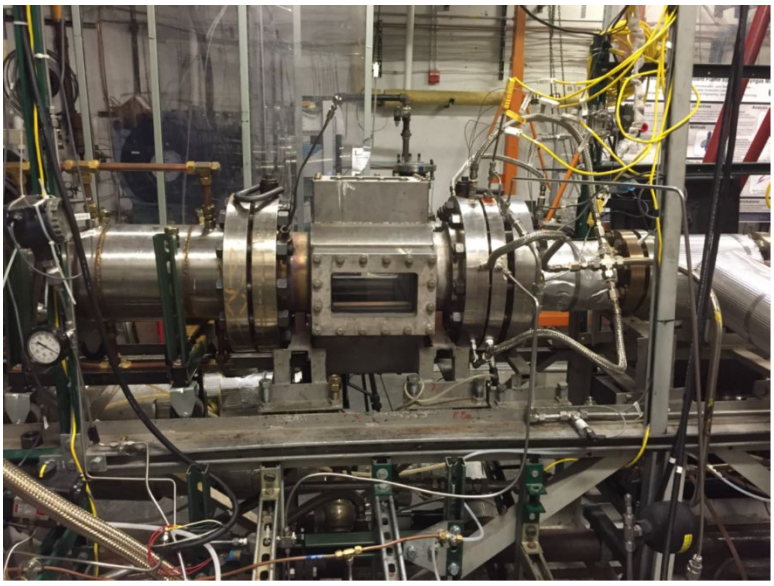

Figure 3.1: Combustion Test-rig 1

OH* chemiluminescence time series data was recorded using a photomultiplier tube, which aggregates the light intensity in its field of view into a singleton point. Data was sampled at a rate of 10 kHz over a 50 second interval. During this interval, the equivalence ratio was gradually reduced from $\phi - \phi_{LBO} \approx 0.1$ to blowout. A realization of this data is shown in Figure 3.2.

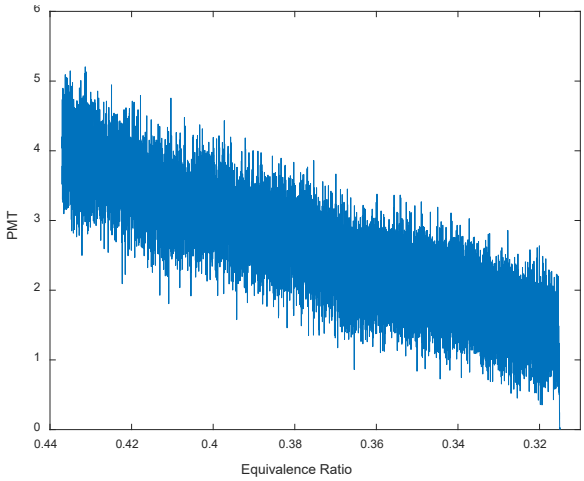

Figure 3.2: OH* Chemiluminescence Signal Recorded by PMT

Ten replications of this process were recorded for ten fuels and two air temperatures (300 K and 450 K). This dataset was utilized to develop a data analytics framework for detecting precursors to lean blowout. However, application of the framework was limited to the conventional fuel. Therefore, application of the framework to the alternative fuels is a topic for future work.

The data analytics framework consists of constructing a control chart to monitor the combustion flame as the equivalence ratio is reduced. In Statistical Process Control (SPC), data is collected while the system is in a healthy state and operating under normal operating conditions. The statistical distribution of this data serves as a baseline for the system. The control chart is constructed such that statistically significant deviations from the baseline distribution are detected.

These alarms can then be investigated as either chance anomalous occurrences or as results of system faults. An assumption commonly made in SPC is that data samples are statistically independent of each other while being identically distributed. However, the signal in Figure 3.2 is highly autocorrelated and consists of trends in both the mean and variance. Constructing a conventional control chart without accounting for the properties of the data inevitably leads to a high false alarm rate. Therefore, we model the autocorrelation and the nonstationary behavior in the signal using a time series model, ARIMA(2,1,1)-IGARCH(1,1). The formulation for this model is shown below:

$$\nabla X_i = \mu + \phi_1 \nabla X_{i-1} + \phi_2 \nabla X_{i-2} + \theta_1 a_{i-1} + a_i$$
$$a_i \sim N\left(0, \sigma_i^2\right) \quad (3.1.1)$$
$$\sigma_i^2 = Var(a_i | a_{i-1}) = \alpha_0 + \alpha_1 a_{i-1}^2 + \eta_1 \sigma_{i-1}^2, \alpha_0 > 0, \alpha_1, \eta_1 \geq 0, \alpha_1 + \eta_1 = 1$$

In Equation 3.1.1, $\nabla X_i = X_i - X_{i-1}$. The advantage of this modeling framework is that the sequence $Z_i, i = 1,2, \ldots$, where $Z_i = \frac{a_i}{\sigma_i} \sim N(0,1)$ for all i, consists of uncorrelated, identically distributed random variables. Thus, the observed values of this sequence are suitable for constructing the control chart.

To estimate the parameters of the ARIMA(2,1,1)-IGARCH(1,1) model, the ten realizations are equally split into training and testing. A time series model is fitted for each training realization using the first half of the OH* chemiluminescence signal. This corresponds to when the flame is stable. The parameters for the five training realizations are then averaged to obtain a global time series model. This global model is then applied to the five testing realizations.

To detect precursors to blowout, we construct an Exponentially Weighted Root Mean Squared error (EWRMS) control chart. To remove residual autocorrelation, we collect rational subgroups of size $b = 10$. The EWRMS control chart monitors the squared deviation of the process $\bar{Z}_k, k = 1,2, \ldots$ from the expected value of zero. The equation for the monitoring statistic is as follows:

$$S_k = \sqrt{(1-\gamma) S_{k-1}^2 + \gamma \bar{Z}_k^2}, k = 1,2, \ldots, [b^{-1} T_n] \quad (3.1.2)$$

where $S_0 = b^{-\frac{1}{2}}$ and $\gamma = 0.2$. The control limits are determined by first establishing an acceptable false alarm rate, $\alpha = 0.0027$ and then computing the $\frac{100\alpha}{2}$th percentile and the $\frac{100(1-\alpha)}{2}$th percentile of the sequence S_k when the flame is stable. An example of control charts for air temperatures of 300 K and 450 K are shown in Figure 3.3 and Figure 3.4 respectively.

The control chart elicits an alarm when the monitoring statistic breaches one of the control limits. We notice that as blowout becomes more imminent, the frequency of alarms increases. We leverage this to imbue each alarm with a probability that it is a true precursor to blowout. For each air temperature, a 2-parameter exponential distribution is fitted to the equivalence ratio values at the time of alarms for the training realizations. This is shown in Figure 3.5.

For the test realizations, once an alarm is observed, we condition on the event that the equivalence ratio is less than the current value which causes the probability of a true alarm to approach unity as the equivalence ratio reduces toward blowout. This is shown in Figure 3.6.

This true alarm probability can aid turbine operators in decision making. Furthermore, future work can focus on developing utilizing this true alarm probability to control equivalence ratio automatically to prevent blowout.

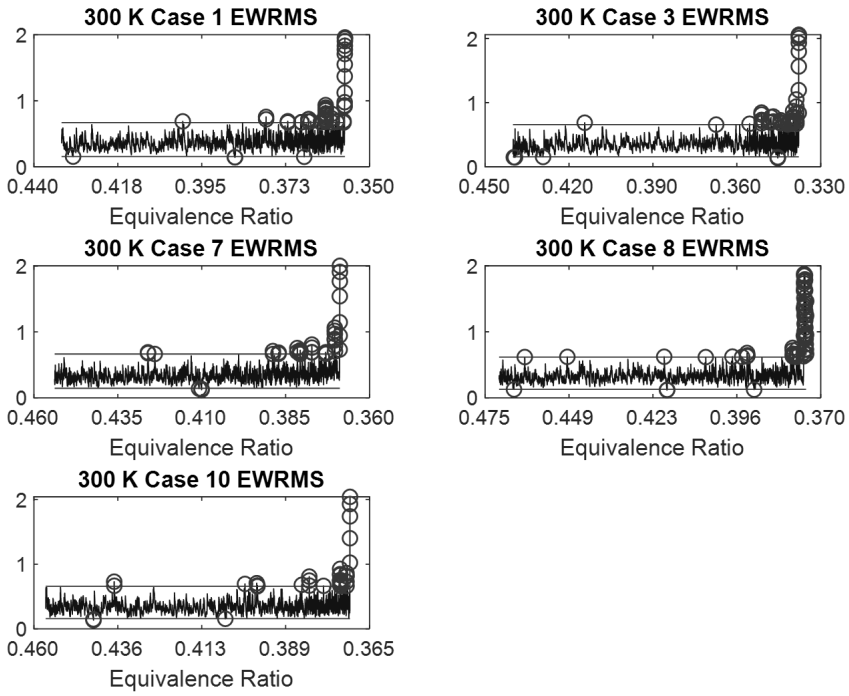

Figure 3.3: EWRMS Control Charts of Training Realizations for 300 K Air Temperature

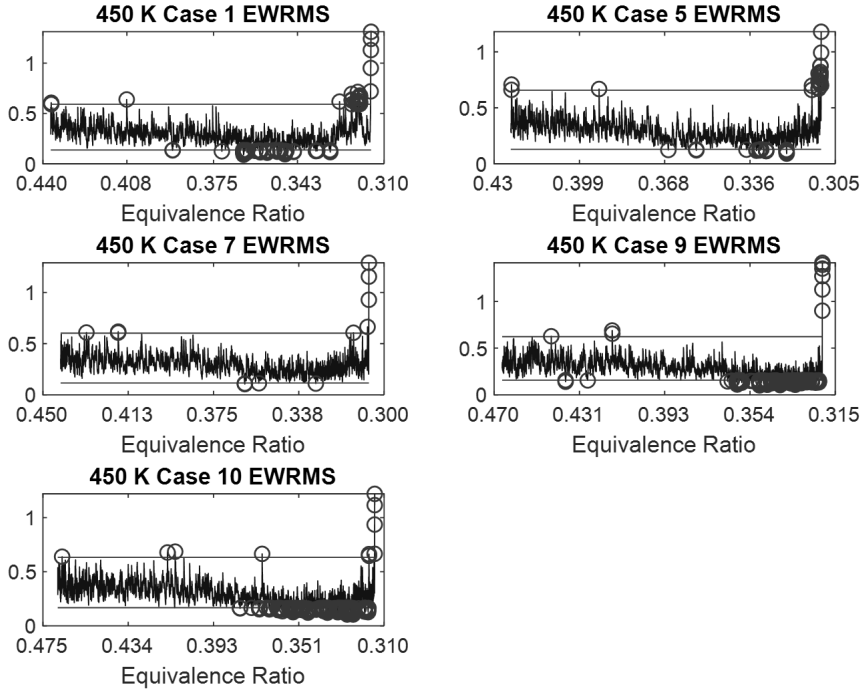

Figure 3.4: EWRMS Control Charts of Training Realizations for 450 K Air Temperature

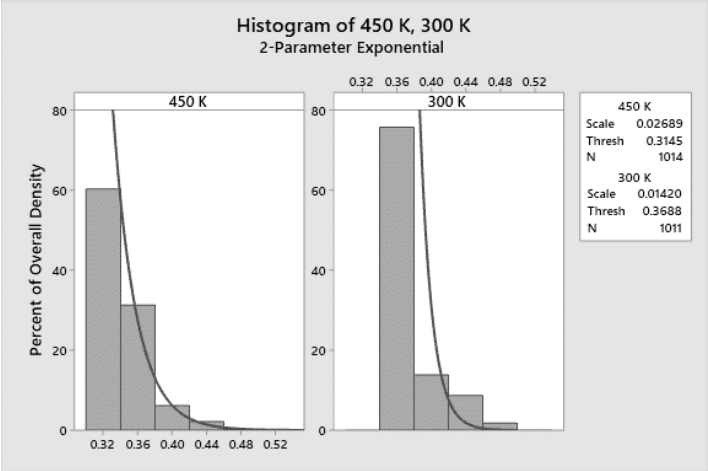

Figure 3.5: Fitted Distributions for Air Temperatures 300 K and 450 K

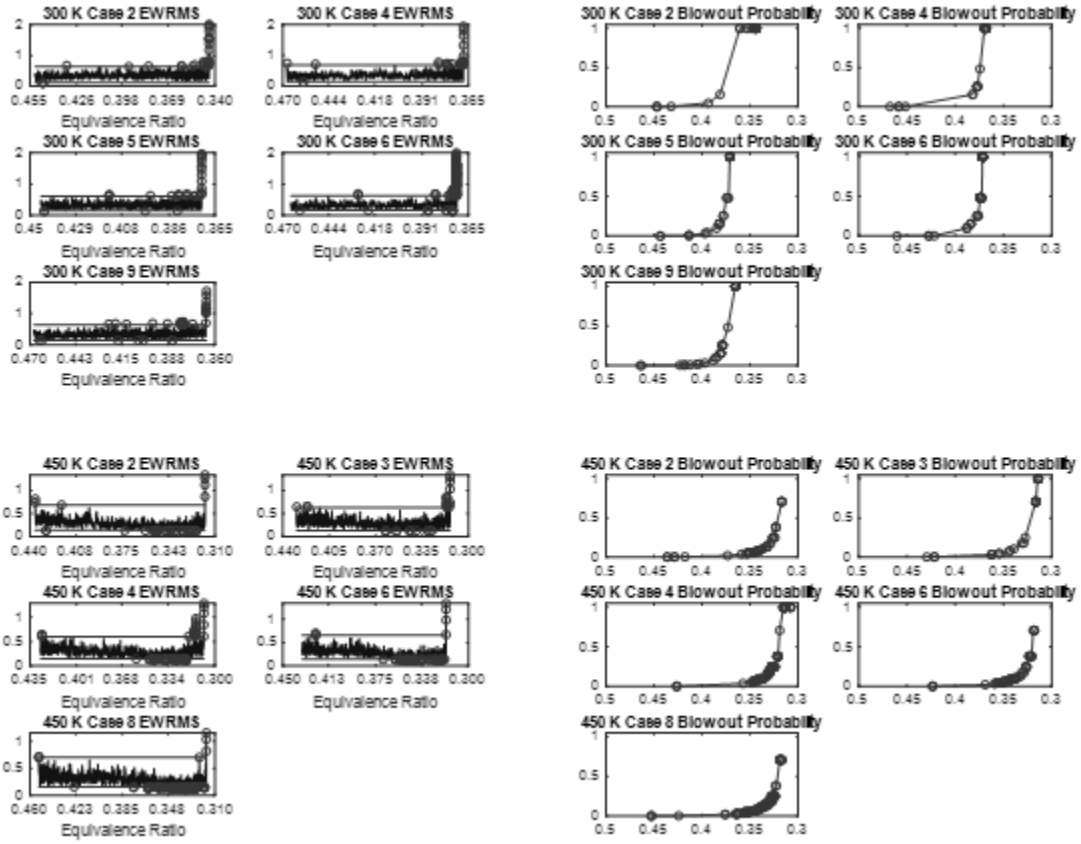

Figure 3.6: Control Charts for 300 K (Top) and 450 K (Bottom)

3.2 Centerbody

In addition to lean blowout, the transition to pre-mixed combustion leaves modern combustors susceptible to instabilities that can result in component degradation. For example, thermal stresses can cause degradation to the combustor centerbody, a component responsible for maintaining a stable combustor flame and for protecting combustor hardware from the flame. Therefore, a study

was conducted to develop a methodology for diagnosing centerbody degradation. This methodology consisted of placing 10 acoustic sensors around a combustor test rig as shown in Figure 3.7.

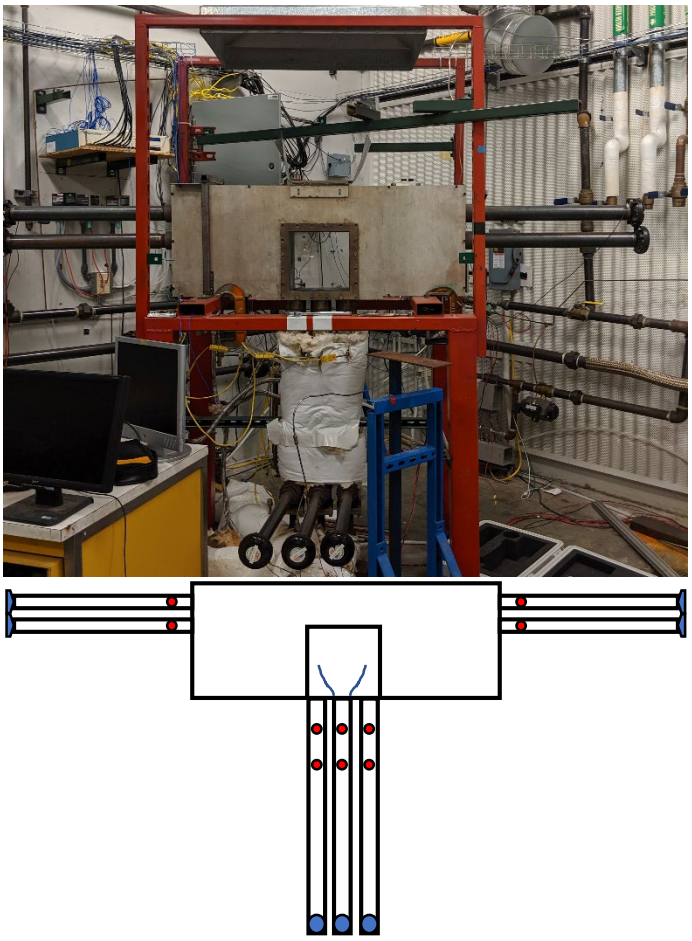

Figure 3.7: Combustor Test-rig with Acoustic Sensors

An experiment was conducted where 10 samples of data were recorded from the 10 sensors for four different levels of centerbody degradation. The degraded centerbodies were manufactured to lengths of 87.4 mm, 76.2 mm, 63.5 mm, and 50.8 mm. Each data sample consisted of monitoring the pressure around the flame using the acoustic sensors for 5.1 seconds at a sampling rate of 40 kHz. An example of a data sample is shown in Figure 3.8.

The data analytics approach consists of the following methodology:

- Feature Extraction: This step consists of transforming the raw sensor data to extract salient features for modeling. A Discrete Wavelet Transform (DWT) was utilized since the time series is nonstationary. This transform decomposes the signal into discrete frequency ranges, where high frequencies are characterized by a high number of coefficients and low frequencies are characterized by a low number of coefficients. To extract features, we compute the energy in each frequency range by calculating the sum of squared coefficients within their respective frequency ranges. An example of this feature extraction is shown in Figure 3.9.

- Sensor Selection: In the previous step, two of the sensors are removed because they malfunctioned during the experiment. To select the appropriate features for monitoring, we employ a hierarchical feature selection approach. The first step is to select an optimal subset of sensors. The reason is that some sensors may either be redundant or noninformative regarding degradation. Inclusion of these sensors introduces noise in the model and can result in reduced prediction accuracy. To perform sensor selection, we use multi-class logistic regression with adaptive group lasso penalty. This type of regression maps the features to the space of discrete degradation classes. The penalty causes the regression coefficients for features extracted from nonsignificant sensors to shrink to zero, thereby removing them from the analysis. The sensor selection can be shown in Figure 3.10 below.
- Individual Feature Selection: After selecting significant sensors, the optimal set of individual features are selected from the remaining sensors. This is achieved using multi-class logistic regression with lasso penalty. The result of this is shown in Figure 3.11.
- Diagnosis: The individual feature selection showed that a small number of features can be used for predicting centerbody degradation. To perform diagnosis, the model fitted in the previous step is used to predict the probability that the centerbody is in one of the four states. The state with the highest probability is selected for diagnosis.

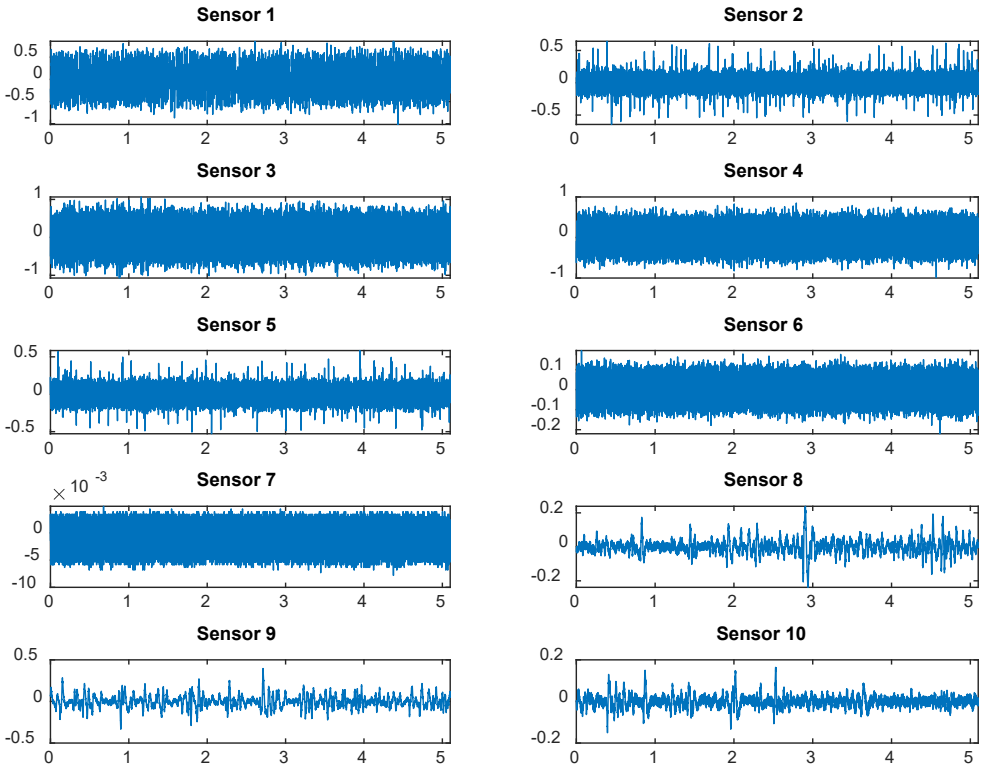

Figure 3.8: Acoustic Sensor Data for Combustor Flame

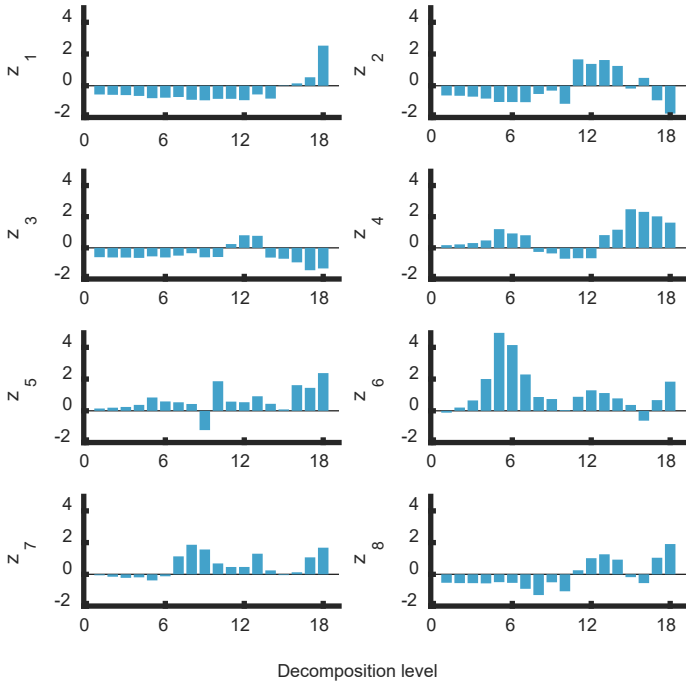

Figure 3.9: Component Feature Vectors

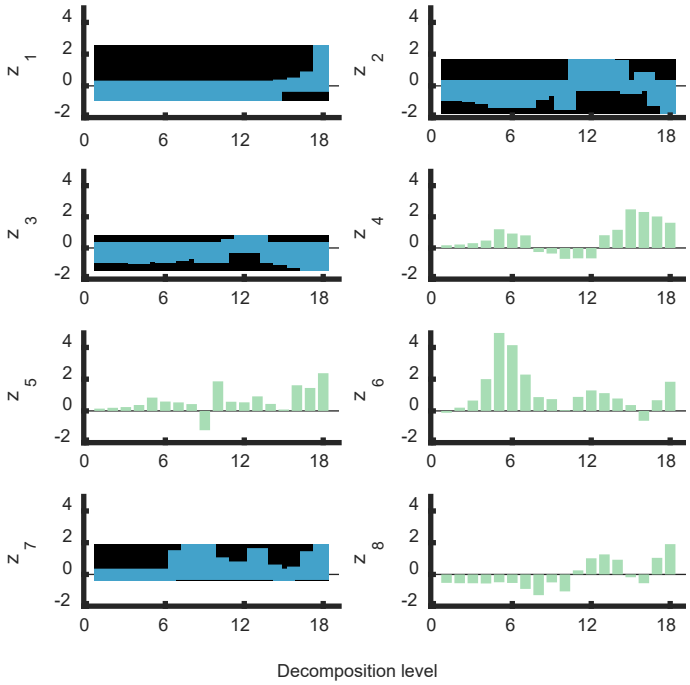

Figure 3.10: Component Feature Vectors Showing the Selected Sensors (Blue) and Rejected Sensors (Green)

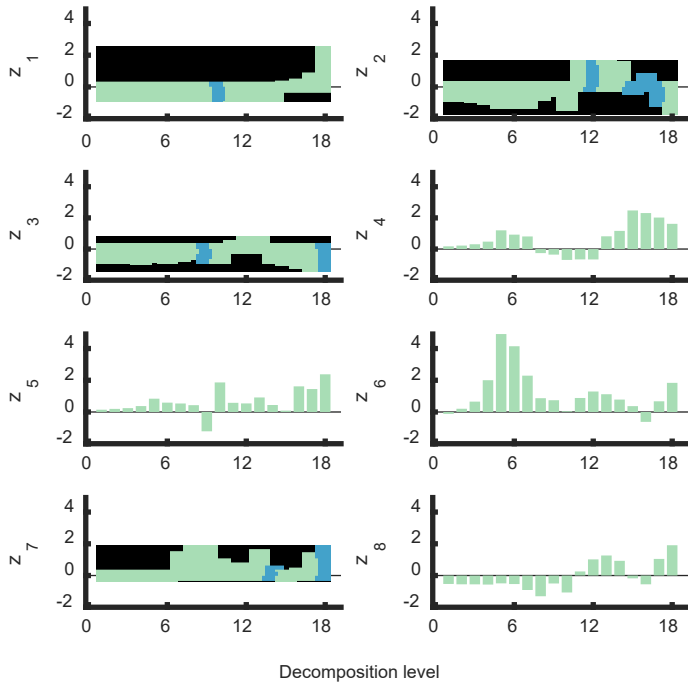

Figure 0.11: Component Feature Vectors, z_i, Showing the Selected Features (Blue) and the Rejected Features (Green) Corresponding to Each Sensor

To assess the effectiveness of the proposed methodology, three hypotheses were tested. First, we tested whether the proposed methodology was better than using a model with no feature selection. The baseline model consisted of a multi-class logistic regression model with ridge penalty. This model shrinks regression coefficients but does not reduce them to zero. Second, we tested whether our methodology was robust to changes in the training sample size. The results for both tests are shown in Table 3.1.

Table 3.1: Classification Accuracy for Models Using N Total Observations for Training

***N* (# Obs/Class)**	**Model**	
	Baseline model	**Hierarchical Feature Selection**
20 (5)	0.985	1.000
16 (4)	0.974	0.999
12 (3)	0.961	0.994
8 (2)	0.882	0.956

Table 3.1 demonstrates that the proposed model outperforms the baseline model. Furthermore, it is more robust to changes in the sample size as it maintains a 95% accuracy even when using only two observations per class for training. The final test was to determine which set of sensors was more accurate: the sensors located on the tubes (longitudinal) or those placed on the support beams (transverse). The results showed that when the training size was $N = 20$, the longitudinal sensors had an overall accuracy of 99.75% whereas the tangential sensors had an overall accuracy of 85.6%. Therefore, the sensors placed on the tubing are more effective than those on the support beams. This concluded the work conducted on the combustion section. In the next chapter, we discuss research efforts related to the turbine section.

4 Turbine Section

High main gas path temperatures are important for ensuring turbine efficiency. However, these temperatures can exceed component melting temperatures. To counter this, modern turbines are equipped with cooling systems to preserve the health of the turbine components despite the high temperatures. For example, turbine blades are now manufactured with coolant holes so cold air from the compressor can maintain a blade temperature below the material melting point. For components below the main gas path, cooling efforts are more novel. They involve using coolant flow from the compressor to create a seal to protect the under-platform region from main gas path ingress. The challenges associated with these cooling strategies is in monitoring their effectiveness. Several instruments for monitoring blade temperature and signs of material warping like blade tip clearance have been utilized. However, these techniques have not been utilized for monitoring the coolant flow from the upstream compressor. Therefore, research efforts focused on developing a framework for monitoring the coolant flow rate using infrared images of the blade. For the underplatform region, there is a dearth of monitoring strategies published in the literature. Therefore, we propose two approaches to monitor the effectiveness of the coolant seal using time-resolved pressure signals recorded either near the rim seal or above the casing. Finally, we conclude the section by briefly discussing an experiment to monitor tip clearance using the principles of experimental design. While no data analytics have been published from this effort, the dataset generated serves as an opportunity for future research teams to develop condition monitoring methodologies. All experimental work was conducted in the Steady Thermal Aero Research Turbine (START) facility at Pennsylvania State University. A diagram of this facility is shown in Figure 4.1.

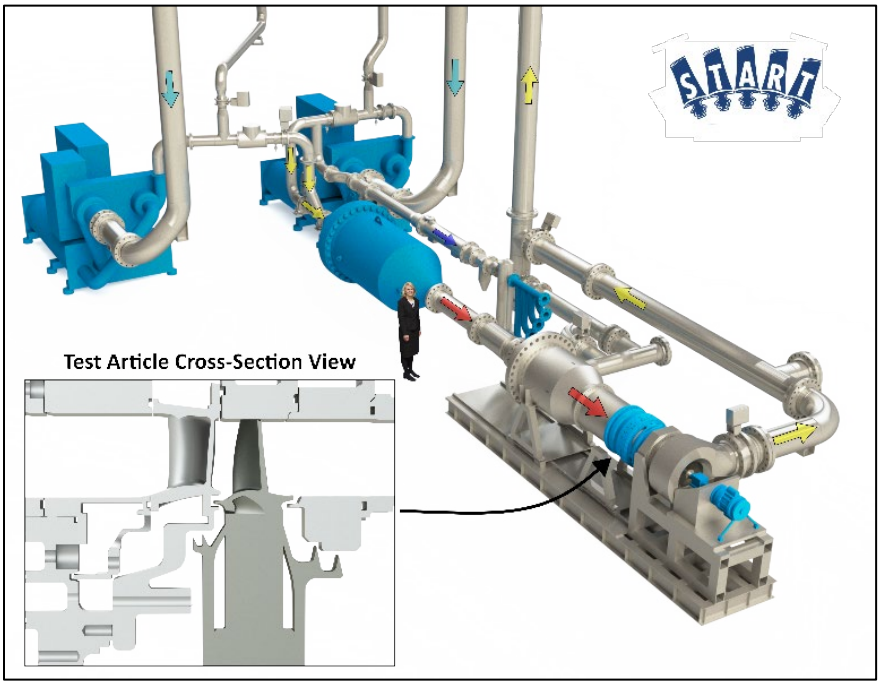

Figure 4.1: START Lab

4.1 Blade Coolant

Blade cooling flow is necessary to prevent high-temperature erosion of turbine blades. An example of this erosion is shown in Figure 4.2. To avoid this erosion, turbine blades are designed to allow cooling flow through manufactured holes as shown in Figure 4.3.

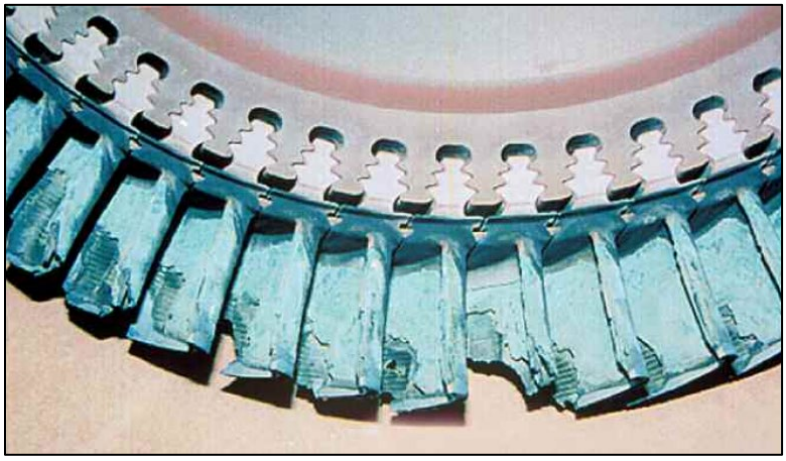

Figure 4.2: Turbine Blade Erosion

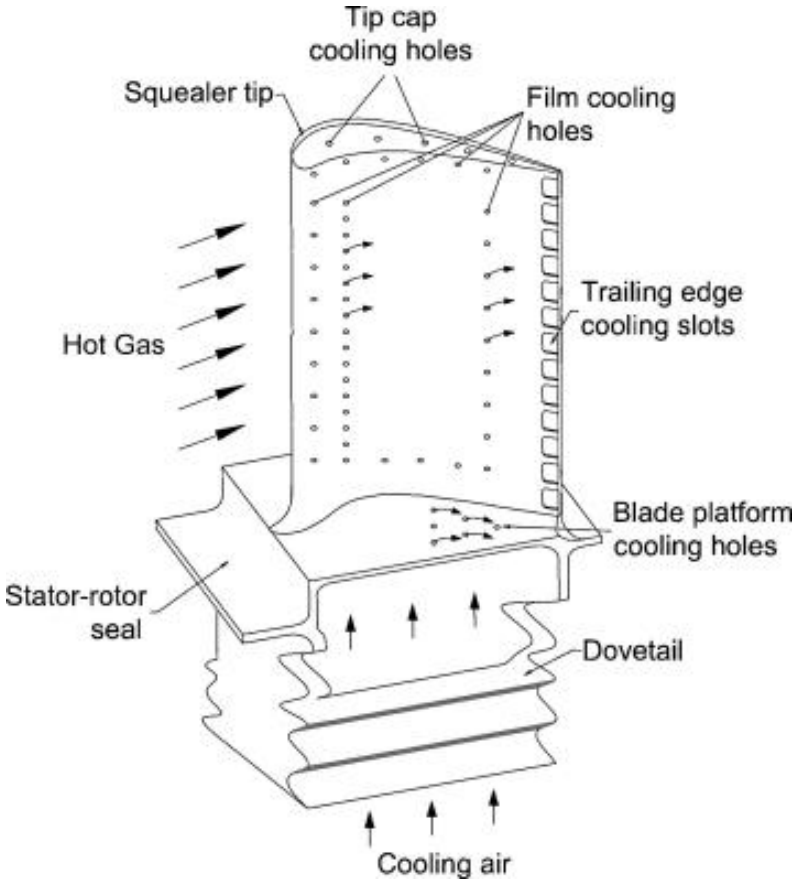

Figure 4.3: Turbine Blade Diagram

For this experiment, we utilize infrared imaging of the turbine blade to monitor the health of the coolant system. For example, if an upstream fault occurs in the compressor, we want to determine whether this type of fault can be detected via the temperature distribution of the blade. A diagram of the imaging is shown in Figure 4.4.

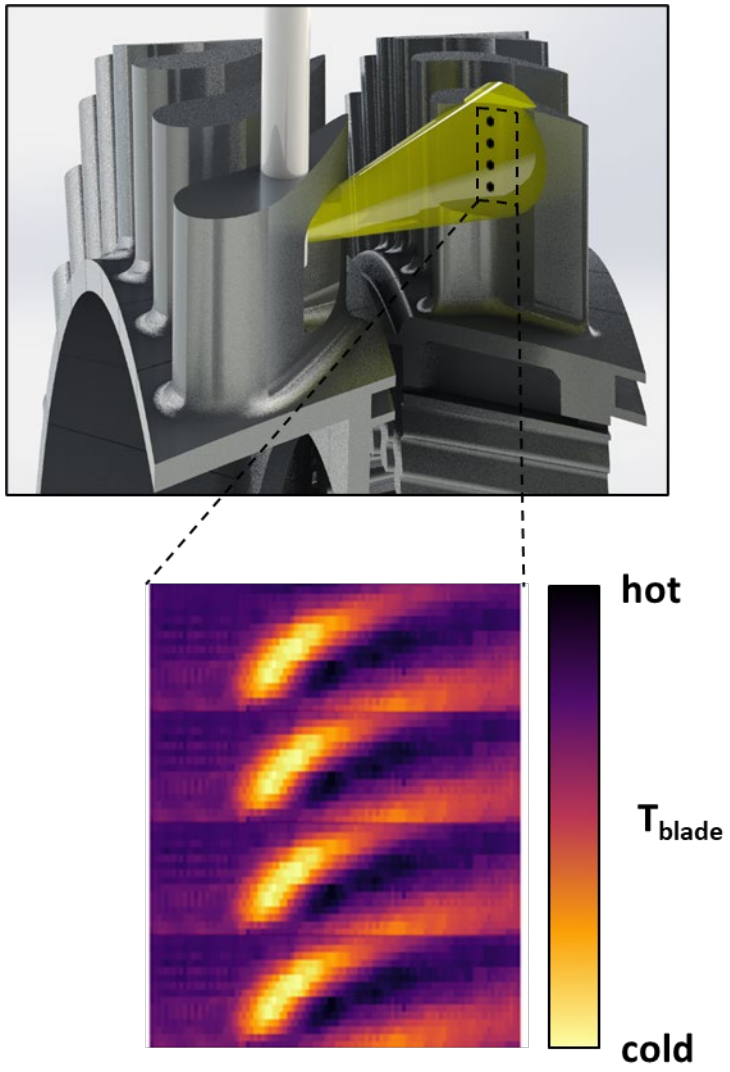

Figure 4.4: Infrared Imaging of Blade

Imaging of the blade was recorded for four main gas path temperatures and five coolant flow rates. A diagram showing the set points for these measurements is shown in Figure 4.5.

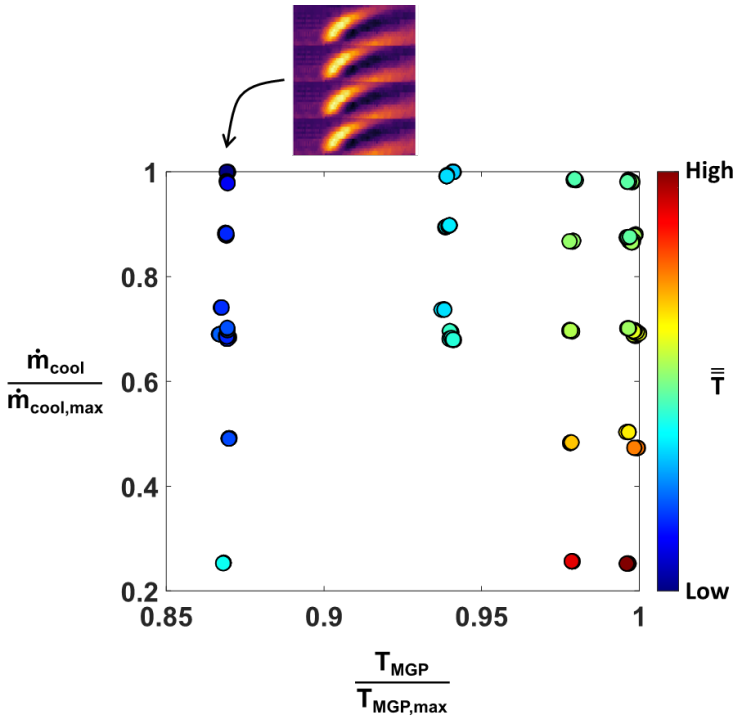

Figure 4.5: Experimental Set Points for Imaging

Note that in Figure 4.5, there are large deviations amongst the coolant flow rates despite the set points being the same. This is due to latent factors influencing the measurements. The existence of these measurements influenced the experimental design demonstrated in Section 4.3. Part of this project was to determine how important including these factors into the model was in predicting the coolant flow rate. The model used was a linear regression with lasso penalty, whose formulation is shown below:

$$\min_{(\boldsymbol{\beta_0},\boldsymbol{\beta})\in\mathbb{R}^{P+1}} \frac{1}{2N}\sum_{i=1}^{N}(y_i-\beta_0-\boldsymbol{x}_i^T\boldsymbol{\beta})^2+\lambda\|\boldsymbol{\beta}\|_1 \qquad (4.1.1)$$

In Equation 4.1.1, y_i is the ith observed coolant flow rate, β_0 is the bias term, and $\boldsymbol{x}_i \in \mathbb{R}^P$ is the vector of predictors. For this project, $\boldsymbol{x}_i$ is a vector comprised of either just features from the image or of features combined with ancillary measurements. When other measurements are included, so too are the interaction terms. For example, if main gas path temperature is included in the model, $\boldsymbol{x}_i = \left(\bar{\bar{T}}_i, T_{MGP,i}, \bar{\bar{T}}_i \times T_{MGP,i}\right)^T$, where $\bar{\bar{T}}_i$ is the average intensity of the infrared image. In general, three feature extraction techniques were implemented:

- 2D Avg: an overall average of the image pixels
- 1D Avg: the set of column wise averages
- Image: using all image pixels as features

As with the combustor experiments, the data was partitioned into a training and test set, where the models were fitted using the training data and the accuracy of the models was determined based on the test data.

Since ancillary factors had a significant effect on the data, we wanted to first see how our modeling methodology worked on an ideal dataset where variance of these ancillary factors was minimal. Then, we included the remaining data to test how inclusion of the ancillary measurements into the model can reduce the prediction error introduced by these non-homogeneous datapoints. The prediction root mean squared error (RMSE) for all models fitted is shown in Figure 4.6.

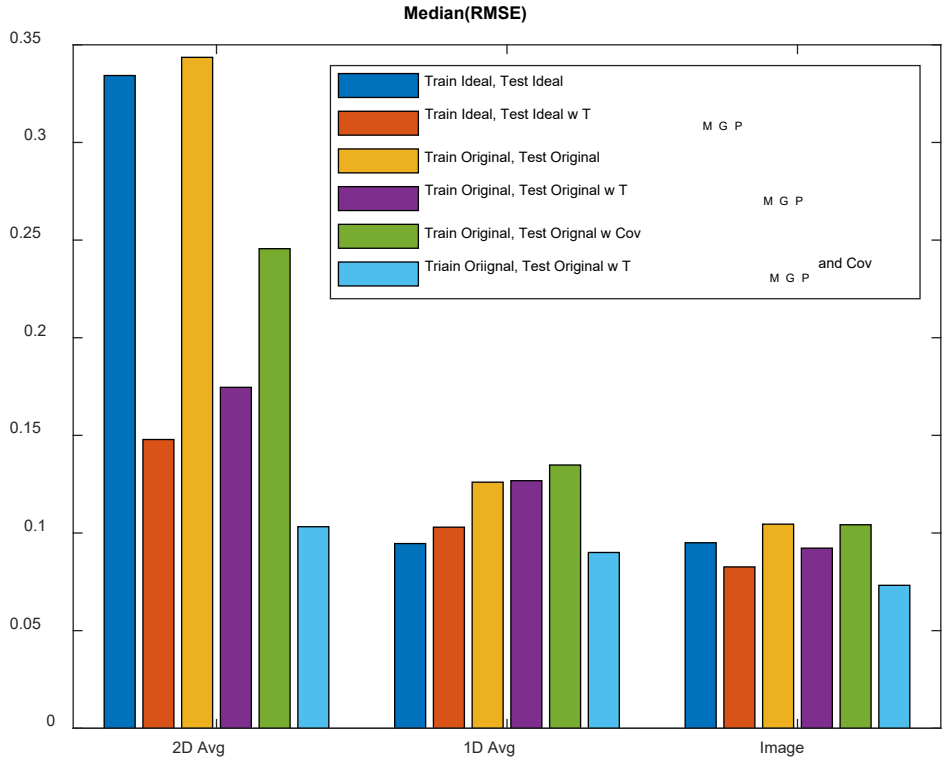

Figure 4.6: Comparison of Coolant Flow Rate Prediction Accuracy

By comparing the median root mean squared error, we can draw the following conclusions:

- The 1D Avg and Image features result in more accurate models than the 2D Avg
- Inclusion of the main gas path temperature into the model only improves accuracy for the 2D Avg feature
- Modeling using the ideal dataset results in less prediction error than modeling using the original dataset and not including any ancillary measurements. This is true, even if including the main gas path temperature as a predictor.
- Including the ancillary measurements with the main gas path temperature results in the most accurate model for all feature extraction techniques. However, the benefit for 1D Avg and 2D Avg is minimal. Therefore, the added benefit may be outweighed by the cost of extra sensing capability.

Another interesting discovery was made when looking at the regression coefficients for the model fitted using just the Image features on the original dataset. The coefficients that were nonzero were aggregated in two distinct clusters on the image as shown in Figure 4.7.

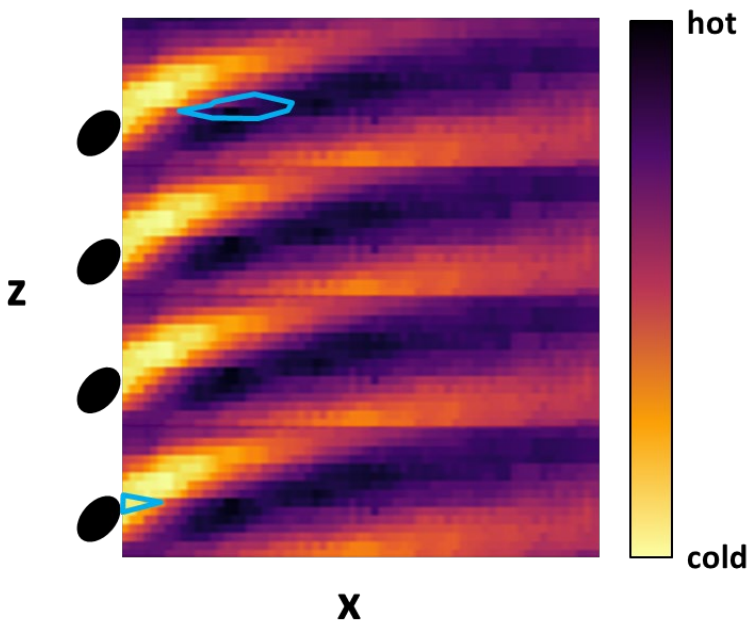

Figure 4.7: Regression Coefficient Clusters

The two regions correspond to a cold region from the bottom-most cooling hole and the hot region below the upper-most cooling hole. This indicates potential for targeting more cost-effective sensors to these regions of the blade. As shown in Figure 4.8, the use of four pixels spanning both regions can perform on par with the optimal model.

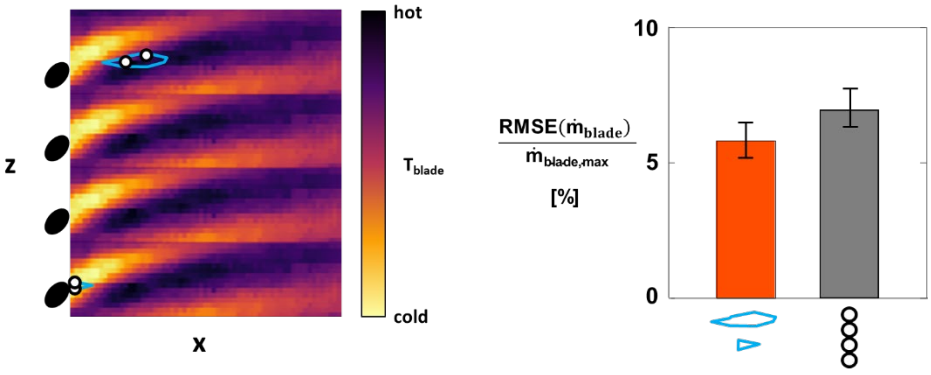

Figure 4.8: Comparison Between Optimal Model and 4-Feature Model

4.2 Sealing Effectiveness

In addition to monitoring blade coolant flow, the START Lab was used to conduct an experiment to aid in the development of a framework for monitoring the ability of purge flow to prevent ingress into the wheelspace region. A diagram of the cross-section of the gas turbine showing the location of the sensors is shown in Figure 4.9.

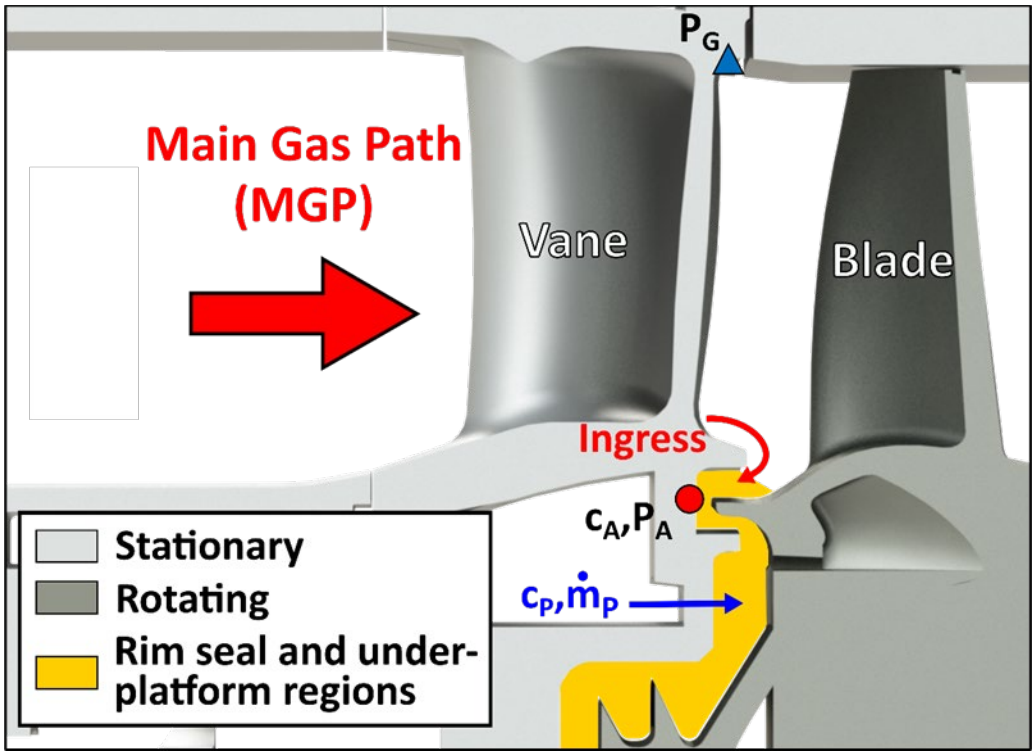

Figure 4.9: Cross-section of Gas Turbine

In this experiment, 9 samples were recorded at 9 different purge flow conditions. Data was sampled from pressure sensors located at the rim seal (P_A) and the outer casing (P_G). A gas tracer method was used to measure the sealing effectiveness for a given purge flow condition. The purge flow settings along with the sealing effectiveness values are shown in Figure 4.10.

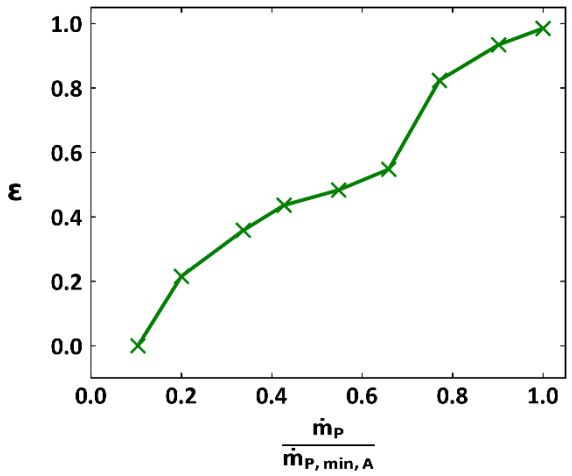

Figure 4.10: Sealing Effectiveness vs. Purge Flow Rate

To extract features from the pressure signals, a fast Fourier transform is employed. Then the energy within frequency bands centered at integer multiples of the disk rotating frequency is computed to generate features. An example of this transform for both sensor locations is shown in Figure 4.11.

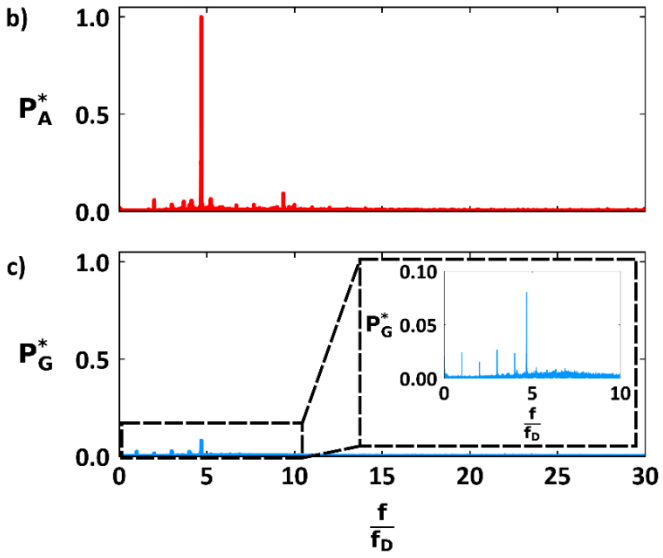

Figure 4.11: Fast Fourier Transform

In Figure 4.11, there is a noticeable attenuation in the amplitude from the rim seal to the outer casing. We expect a higher prediction accuracy using features from the rim seal vs. features from the casing. However, we are interested in whether the outer casing features are usable for monitoring.

For this study, two diagnosis methodologies are proposed. The first is an off-the-shelf technique that uses linear regression with lasso to map the features extracted from the fast Fourier transform to the sealing effectiveness. Upon observing the regression coefficients in Figure 4.12, we notice that the coefficient related to the dominant frequency in Figure 4.11 $\left(\frac{f}{f_D} = 5\right)$ is zero.

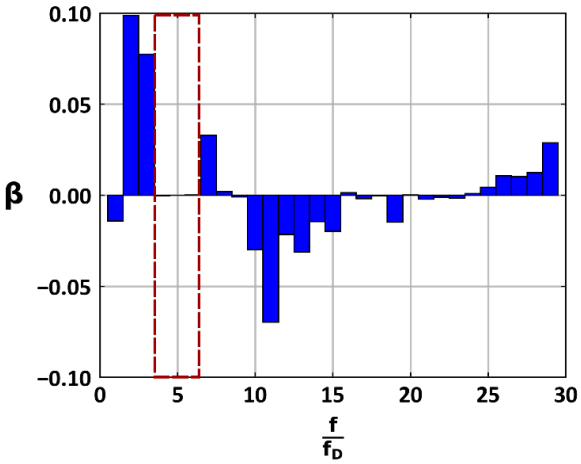

Figure 4.12: Lasso Regression Coefficients

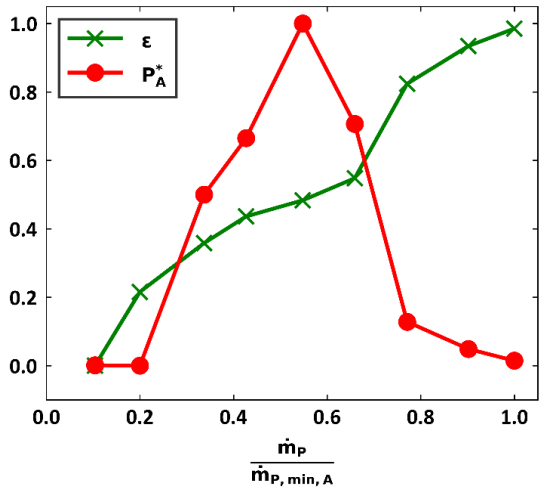

Figure 4.13: Purge Flow and Fluid-Dynamic Instability

However, this feature corresponds to a fluid-dynamic instability that whose correlation with the sealing effectiveness changes after the sealing effectiveness is greater than 0.5. This can be seen in Figure 4.13.

To obtain a more accurate model for sealing effectiveness, a two-step approach is employed. First a coarse model that predicts whether the sealing effectiveness is less than 0.5 or greater than 0.5 is fitted using all features. This model utilizes logistic regression with lasso, where logistic regression maps the features to a binary state. Then using the dominant frequency and its next four harmonics (and their log transforms), linear regression with lasso is fitted. However, this model includes a covariate corresponding to whether the sealing effectiveness is less than 0.5 or greater than 0.5. This splits the model into two models. Therefore, given the result from the first step, a granular prediction of sealing effectiveness is obtained using the linear regression model in the second step. The result of these two strategies is shown in Figure 4.14.

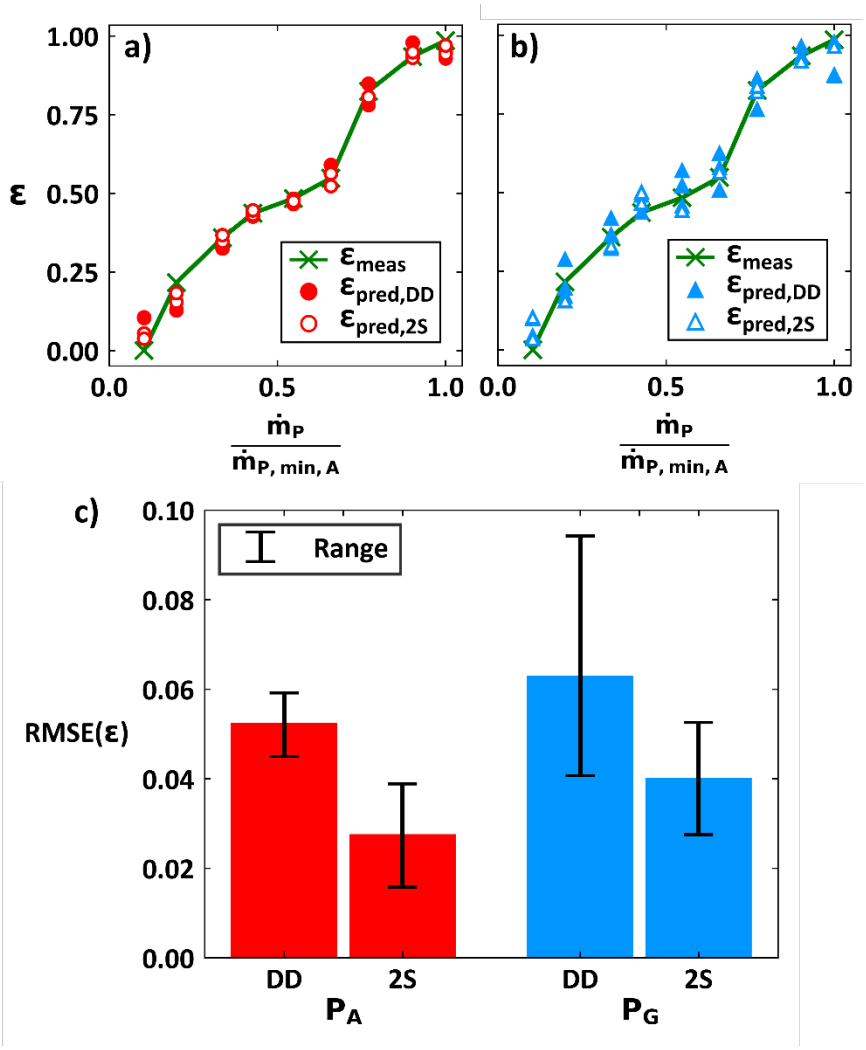

Figure 4.14: Prediction Results for a) Rim Seal, b) Outer Casing, and the c) RMSE

From Figure 4.14, the two-step approach (2S) significantly reduces the prediction error from the off-the-shelf method (DD). Furthermore, despite having a higher RMSE than the rim seal sensor, the outer casing sensor still enables accurate prediction of the sealing effectiveness.

4.3 Tip Clearance

Tip clearance has an inversely proportional relationship with gas turbine efficiency. If tip clearance can be monitored using temperature or pressure measurements, then that would increase functionality of these sensors. Therefore, we are interested in whether time-resolved over rotor pressure and temperature sensors can be used to predict tip clearance. Part of the capabilities of the START Lab is a magnetic bearing that can move the rotor in either cardinal direction. A diagram of the turbine blade and the sensor layout is shown in Figure 4.15.

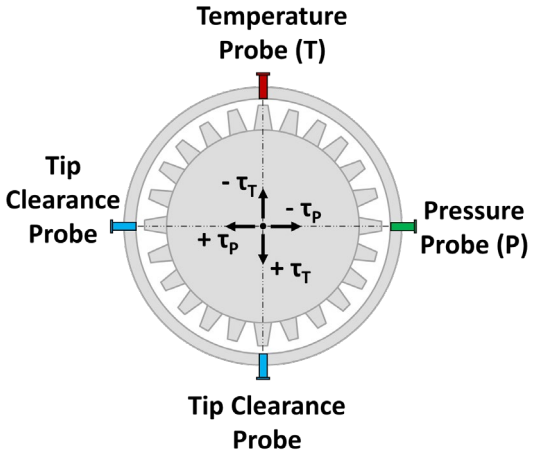

Figure 4.15: Tip Clearance Sensor Layout

Data analytics related to this study involved first normalizing the pressure and temperature by the inlet parameters and the coolant parameters. The prediction accuracy of the normalized sensor features can then be compared to the raw sensor features. A discrete wavelet transform is used to extract features from the sensor signals. Then linear regression with lasso is used to map the features to the tip clearance. A model is fit for features extracted from the nonnormalized signals and the normalized signals. Furthermore, two ensemble methods are employed to combine the results from the three models. The first method selects the sensor with the lowest prediction error on the validation set. The second method uses a weighted average of the tip clearance predictions, where the weights are inversely proportional to the prediction error for that sensor. The results of this study are summarized in Figure 4.16.

The RMSE plots show that there is not a discernable difference in prediction accuracy amongst the various sensor normalization techniques. The pressure sensor is far less susceptible to outliers compared to the temperature sensor. However, predictions near the boundaries of the sealing effectiveness show a high degree of error.

This project was expanded upon to account for measurements collected under various turbine operating conditions. Specifically, three types of measurements were highlighted as potential candidates for diagnostic modelling: blade pressure ratio (BPR), exit temperature (T_{exit}), and time-resolved blade tip pressure ratio (BTPR). All these measurements are sensitive to both changes in tip clearance, as well as changes in facility operating conditions. To facilitate the development of a model which can account for the interdependency of tip clearance and turbine operating

conditions, four boundary conditions were varied across 16 runs with independent tip clearance variations executed at each run, as shown in Figure 4.17.

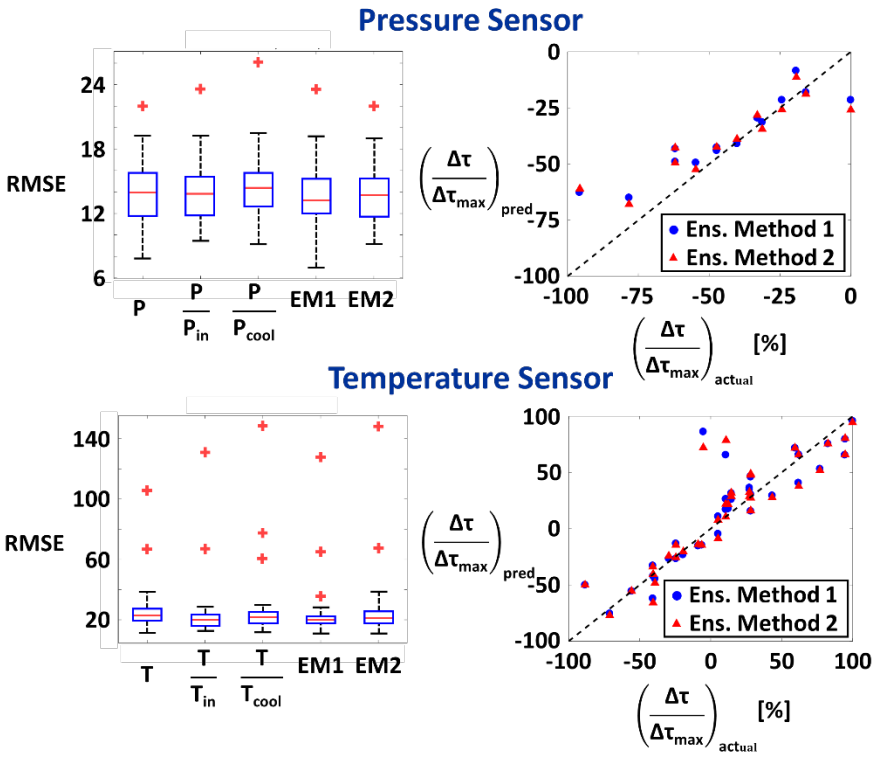

Figure 4.16: Prediction Accuracy for Tip Clearance Study

Test Matrix

Total-to-static pressure ratio

Run	T_{in} [-]	m_{cool} [-]	TSPR [-]	Speed [-]
1	A	A	A	A
2	A	A	A	B
3	A	A	B	A
4	A	A	B	B
5	A	B	A	A
6	A	B	A	B
7	A	B	B	A
8	A	B	B	B
9	B	A	A	A
10	B	A	A	B
11	B	A	B	A
12	B	A	B	B
13	B	B	A	A
14	B	B	A	B
15	B	B	B	A
16	B	B	B	B

A = Design condition 1
B = Design condition 2

Tip Clearances

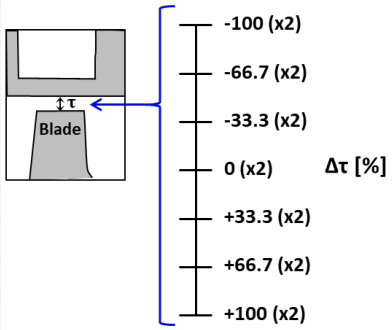

Figure 4.17. Tip Clearance Diagnostics Study Test Matrix

Beginning with BPR, Figure 4.18 shows the data alongside the facility cross-section with the sensors highlighted. The relationship between BPR and tip clearance (τ) is linear across all runs (operating conditions). Notably, the data are falling into two distinct clusters, which is driven by the two TSPR set points, as indicated in Figure 4.18. Within these clusters, the x and y shifts are caused by variations in the other 3 independent variables (inlet temperature, coolant flow rate, and speed). When considering BPR with respect to the other diagnostic measurements, BPR is relatively simple in that it does not capture any special or temporal variation.

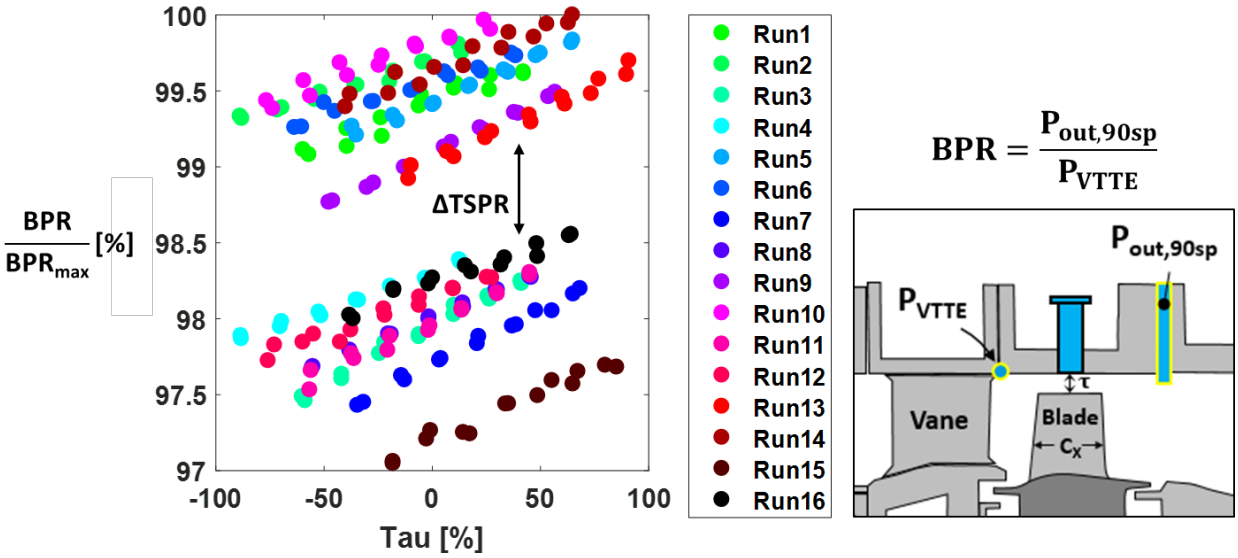

Figure 4.18: Left) Equation and Measurement Locations Used to Quantify BPR. Right) BPR vs. Tip Clearance (τ) at All Run Conditions

Next, a subset of the exit temperature (T_{exit}) measurements are shown in Figure 4.19. Importantly, these T_{exit} measurements were collected at multiple locations between 50% and 100% span to quantify the rotor exit temperature profile. These profiles are shown in Figure 4.19 for the concentric tip clearance case across all run conditions. The most pronounced shift in the data is caused by the variations between the two inlet temperature set points, as noted in Figure 4.19. The radial distribution of the T_{exit} measurements may be specifically useful for this application because the most pronounced tip clearance effects will be observed in the measurements near the tip, while the measurements near the midspan (50% span) will be largely insensitive to the tip clearance variations.

Lastly, an example of the BTPR data is shown in Figure 4.20. Importantly, the tip pressure sensor (P_{tip}) is a fast-response sensor that was sampled at frequencies greater than the blade passing frequency. This high frequency response enables an examination of the over-rotor pressure on a blade-by-blade basis. When examining the frequency spectrum (bottom left of Figure 4.20), the blade passing frequency and its harmonics are clearly observable. This diagnostic measurement is important because it may capture tip clearance effects that cause high frequency fluctuations in the flow field.

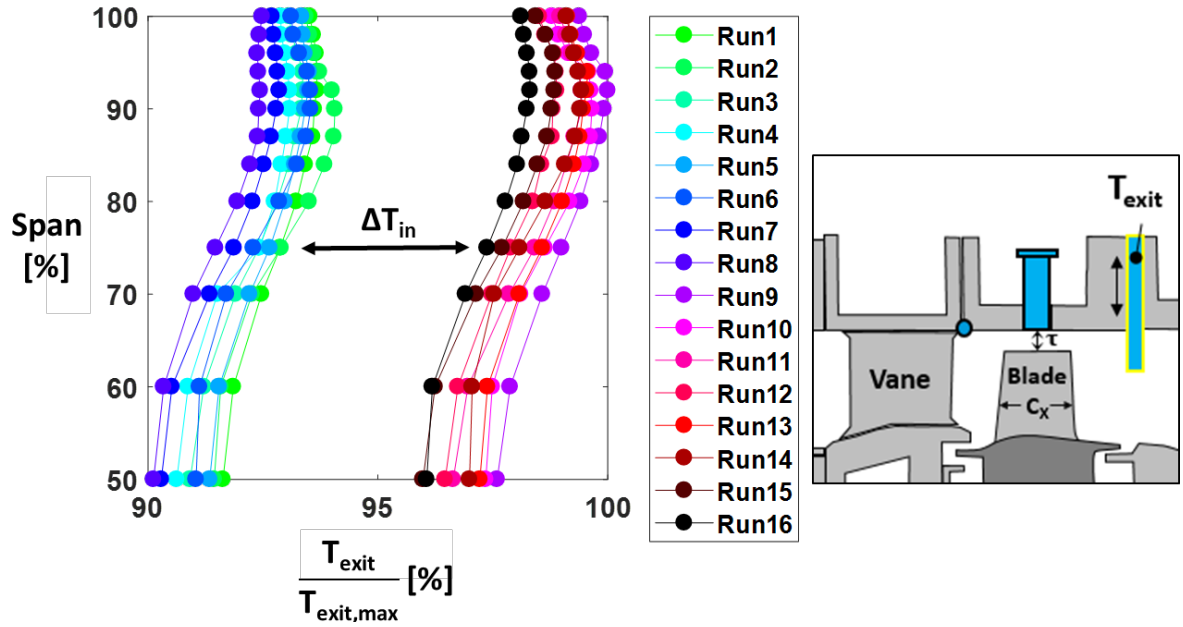

Figure 4.19: Left) Spanwise Exit Temperature Profiles for the Concentric Case Across All Run Conditions. Right) T_{exit} Measurement Location

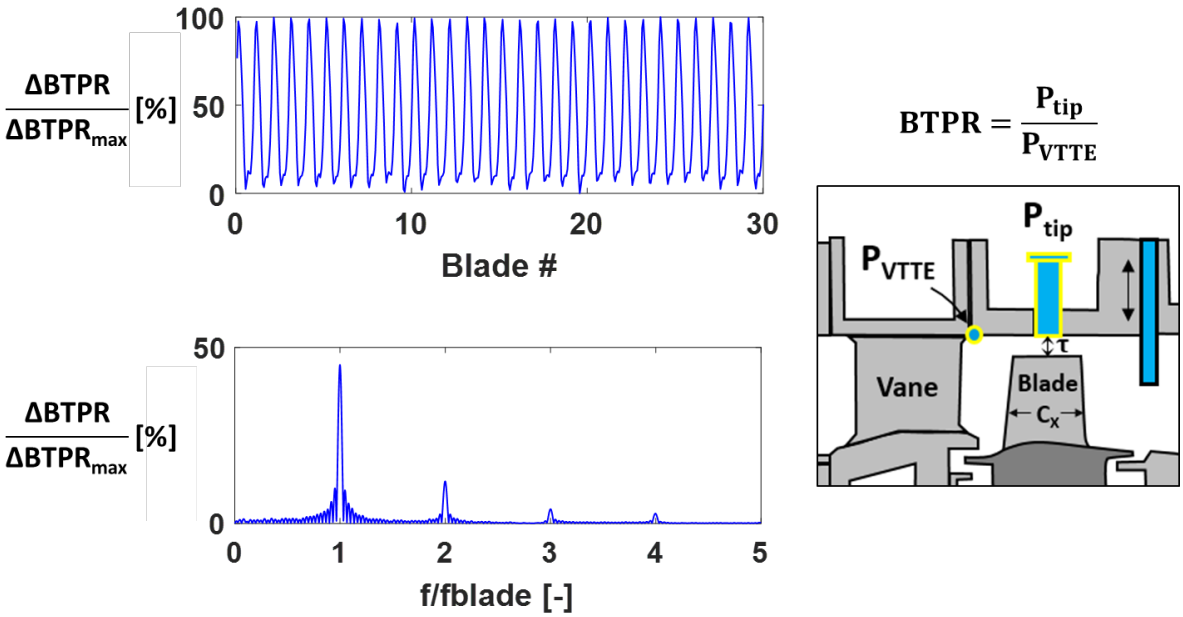

Figure 17.20 Top left) BTPR Waveform for the First 30 Turbine Blades. Bottom Left) BTPR Frequency Spectrum. Right) Equation and Sensors Used to Compute BTPR

Altogether, this suite of diagnostic measurements will shed light on the sensor and modelling requirements for prediction of blade tip clearance using engine representative sensors. The unique characteristics of each diagnostic measurement will be useful for separating operating condition

changes from tip clearance changes when using the sensor measurements as inputs to generate clearance predictions.

5 Dissemination of Project Results

A comprehensive list of conferences/journals in which our findings were presented/published is provided below:

5.1 Posters and Presentations

1. Poster: 2017 University Turbine Systems Research Project Review Meeting (Pittsburgh, PA), Title: Real time health monitoring for gas turbine components using online learning and high-dimensional data
2. Poster: 2018 Annual Project Review Meeting for Crosscutting Research Portfolios (Pittsburgh, PA), Title: Title: Real time health monitoring for gas turbine components using online learning and high-dimensional data
3. Presentation: 2018 University Turbine Systems Project Review Meeting (Daytona Beach, FL), Title: Real time health monitoring for gas turbine components using online learning and high-dimensional data
4. Presentation: 2019 Annual Project Review Meeting for Crosscutting, Rare Earth Elements, Gasification and Transformative Power Generation (Pittsburgh, PA), Title: Real time health monitoring for gas turbine components using online learning and high-dimensional data
5. Presentation: 2019 University Turbine Systems Research (UTSR) Project Review Meeting (Orlando, FL), Title: Real time health monitoring for gas turbine components using online learning and high-dimensional data

5.2 Conference Publications

1. Kumar, R. M., Peters, B., Emerson, B., Paynabar, K., Gebraeel, N., Lieuwen, T. (2020). Data driven fault detection of premixer centerbody degradation in a swirl combustor. *Proceedings of ASME Turbo Expo 2020: Turbomachinery Technical Conference and Exposition GT2020*

2. Correlating time-resolved pressure measurements with rim sealing effectiveness for real-time turbine health monitoring by Eric T. Deshong, Benjamin Peters, Reid A. Berdanier, Karen A. Thole, Kamran Paynabar, and Nagi Gebraeel, Submitted to: *Proceedings of ASME Turbo EXPO 2021: Turbomachinery Technical Conference and Exposition GT 2021*

5.3 Journal Publications

3. Peters, B., Rock, N., Emerson, B., Gebraeel, N., & Paynabar, K. (2021). Data Analytics Method for Detecting Extinction Precursors To Lean Blowout In Spray Flames. **Accepted** *Combustion Science and Technology*. 2020

4. DeShong E., Peters B., Berdanier R. A., Thole K. A., Paynabar K., and Gebraeel N., "Correlating Time-Resolved Pressure Measurements with Rim Sealing Effectiveness for Real-Time Turbine Health Monitoring", **Accepted** *ASME Journal of Turbomachinery*, 2022.

6 Conclusion

This project enabled the development of a symbiotic relationship between industry professionals and data scientists. The benefit for industry professionals is that they are introduced to the concepts of experimental design, statistical process control, and data analytics that either provide credibility to previously held beliefs or illuminate interesting phenomena that motivates further investigation. Conversely, the data scientists are provided feedback on the applicability of their theoretical modeling frameworks to real-world systems. The results of this project demonstrated that Big Data analytics is applicable for developing effective models for condition monitoring of industrial systems. This was shown by a control chart methodology for detecting precursors to lean blowout, a hierarchical feature selection methodology for diagnosing centerbody degradation, a two-step data-driven/domain knowledge-based approach for monitoring sealing effectiveness and using infrared imaging for detecting the state-of-health of cooling systems.

Despite the project concluding, there is still potential for future work. In the lean blowout study, data was collected for multiple alternative fuels. The OH* chemiluminescence data for these fuels is yet to be analyzed. Furthermore, the tip clearance data set collected using the designed experiment is also yet to be analyzed. Therefore, there are opportunities for future research groups to continue with what was accomplished in this project. Future work would also include an attempt to include prognostics. The limitation is the need for data failure data. Since this is not easily attainable, alternative approaches to developing a prognostic dataset can be discussed by future research teams.

		Milestone Status Report					
Milestone #	**Task ID**	**Milestone Title**	**Milestone Description**	**Anticipated Start Date**	**Anticipated Completion Date**	**% Completed**	**Description (progress toward achieving milestone, explanation of deviation from plan, etc.)**
A	2.0	Combustor test conditions outlined	This milestone will establish the test conditions (operational settings such as temperatures, pressures, etc) for two types of tests: operational faults and hardware faults	Oct-17	Mar-18	100%	Test conditions have been completed for LBO testing and hardware fault testing.
B	2.0	Combustor test articles manufactured	This milestone encompases the manufacture of the two combustor test rigs. This includes combustor comonents with seeded hardware faults.	Mar-18	Sep-18	100%	Completed manufacture of operational faults rig (blowout rig). Built the hardware faults rig. Manufacturing of components with seeded faults is nearly complete.
C	2.1, 2.2, 2.3	Measure baseline combustor data	This task involves measurement of typical combustor data (temperatures and pressures) during healthy operation and with faults of varying severity.	Oct-18	Mar-20	100%	All LBO data have been completed, and all baseline data have been captured with hardware faults rig.
D	2.1, 2.2, 2.3	Measure over-instrumented combustor data	This task will augment the datasets from milestone C with additional data that are not typically available from fielded systems.	Jan-19	Jun-20	100%	All LBO data have been completed and all of the hardware faults cases have also been completed.
E	3.0	Turbine test conditions outlined	The operating conditions (inlet conditions, speed, etc.) are defined for the turbine test cards.	Oct-17	Mar-18	100%	Turbine test conditions have been defined for the initial test campaigns.
F	3.1, 3.2	Identification of installation and operation methods for thermal measurements	Install techniques for HFGs, operation requirements for thermal imaging system.	Oct-17	Feb-20	100%	Thermal imaging system has been fully characterized on the bench, and rig integration is underway. Thermal imaging work was completed in Q2FY20. HFG measurements were successfully conducted as part of tip clearance testing toward Milestone H.
G	3.3	Validation of CO_2 sampling	Confirm measurement technique and multi-channel system utilizing CO2 as a gas tracer.	Oct-17	Dec-17	100%	Measurement technique fully validated. Several parallel test campaigns have been completed using this method with high rates of success.
H	3.4	Tip clearance performance assessment	Evaluation of tip clearance effects on turbine performance.	Apr-18	Feb-20	100%	Local clearance variations measured for one hardware case, including time-resolved data; FRAP has been fully integrated; test data have been collected to complete milestone; additional training data are targeted in Q3-Q4FY20 for improved model development with GT-Data team.
I	4.1	Initial AI model trained and passed success criteria	Develop virtual sensing framework (for turbine section)	Apr-18	Sep-20	100%	This milestone originally focused on enhancing the accurcy of estimating tip clearance in the turbine section using commercial inexpensive sensors. In Q9, we identfied another opportunity, which has expanded the scope of this milestone and made us reopen this milestone with an expected completion date of 9/1/2020. **Expanded Scope:** Data features extracted from high-resolution sensors placed inside the turbine section to measure coolant efficiency from a fully-instrumented test-rig are being leveraged to guide feature/frequency selection from spectral data acquired by commercial grade sensors placed on the outside of the turbine casing.
J	5.1	Cloud-based storage architecture created	Identify design requirements for uploading and processing data on AWS Compute nodes	Oct-17	Dec-17	100%	Prleiminary skeleton architecture created on AWS system with upload/download verifications using synthetic data.
K	5.2	Initial "sensor selection" algorithm developed	Develop an algorithm that selection the appropriate features after implementing feature extraction algorithm	Apr-18	Sep-18	100%	Satistical regularization techniques were used tor educe the number of variables generated by Wavelett anaylsis of PMTI data related to Lean Blowout. Wavelet analysis is one of the proposed feature extraction algorithms.
L	5.3	Online fault detection algorithm passed success criteria	Develop a preliminary Machine Learning algorithm that utilizes data to detect candidate combustor and/or turbine faults	Jan-19	Oct-20	100%	A data-driven Machine Learning algorithm has been trained using Photomultiplier tube (PMT) data to detect the initial precursors of lean blowout. This was completed as planned on Mar-19. However we are extending this milestone to encompass additional Combustor and Turbine faults which have now expanded the scope of this milestone. Furthermore, due to campus closures (COVID-19), the new expected completion date will be Decmber 2020.
M	5.4	Component life predictions ready and passed success criteria	Develop prognostic capability for predicting remaining lifetime of a turbine component.	Oct-19	Dec-20	100%	Several studies were conducted on both combustor and turbine sections for different degradation and fault severity levels. Due to practical and experimental limitations, it was prohibitively expensive to run components to failure, hence it was not practical to obtain a "degradation rate". However, different severity levels were tested and hence degradation rates can be obtained based on real-world repair data.