Real-time Health Monitoring for Gas Turbine Components using Online Learning and High Dimensional Data

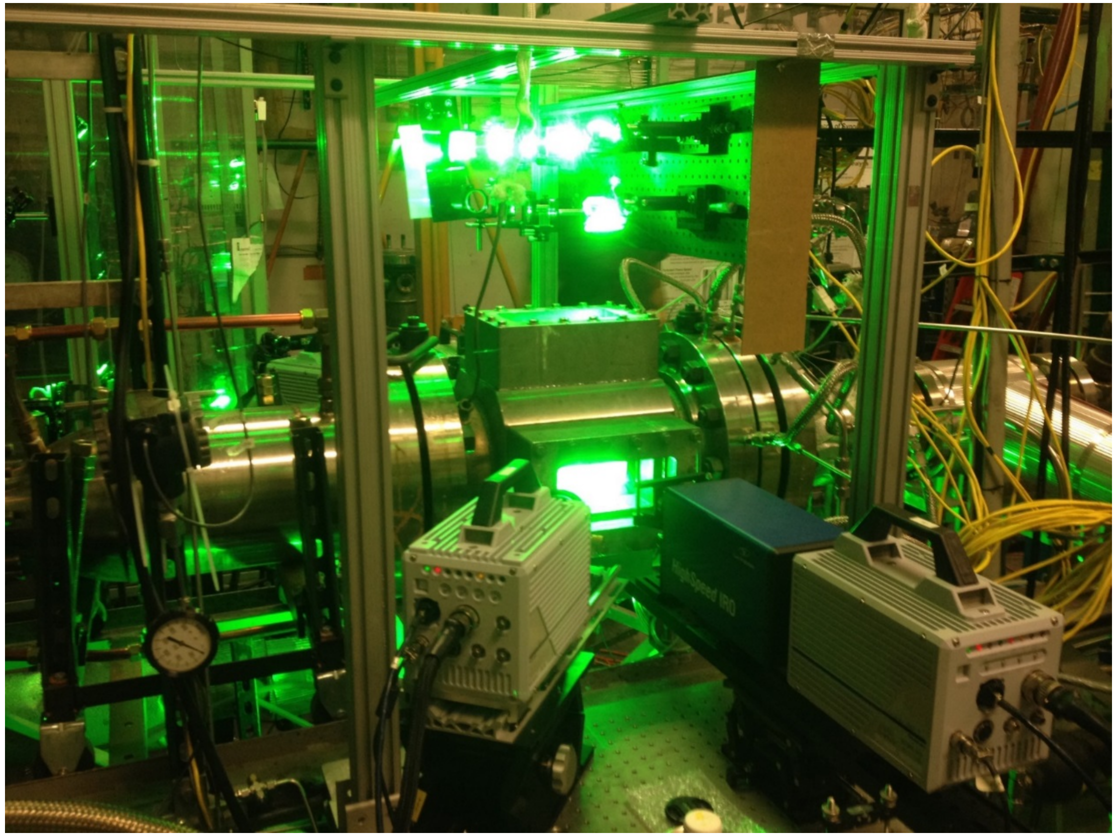
Benjamin Peters

Introduction

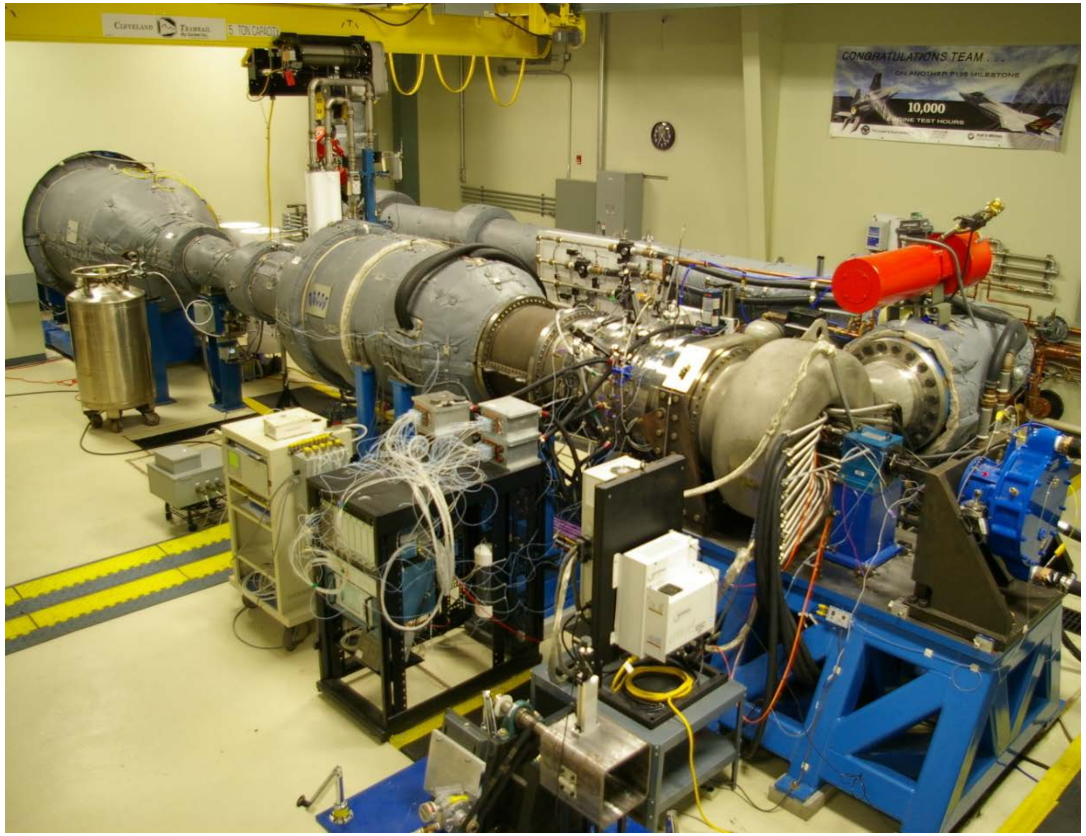
- Big Data analytics holds potential for enabling efficient and reliable operation of power generating gas turbines
  - Reduces unplanned outages
  - Allows for intelligent planning of preventative maintenance and repairs
  - Increases longevity of hardware
- Modern gas turbines are equipped with several hundred sensors
  - Condition Monitoring
  - Volume of data renders conventional analytic techniques ineffective
- **Objective is to develop a Big Data analytics framework for critical gas turbine components.**

Gas Turbine Test-rigs

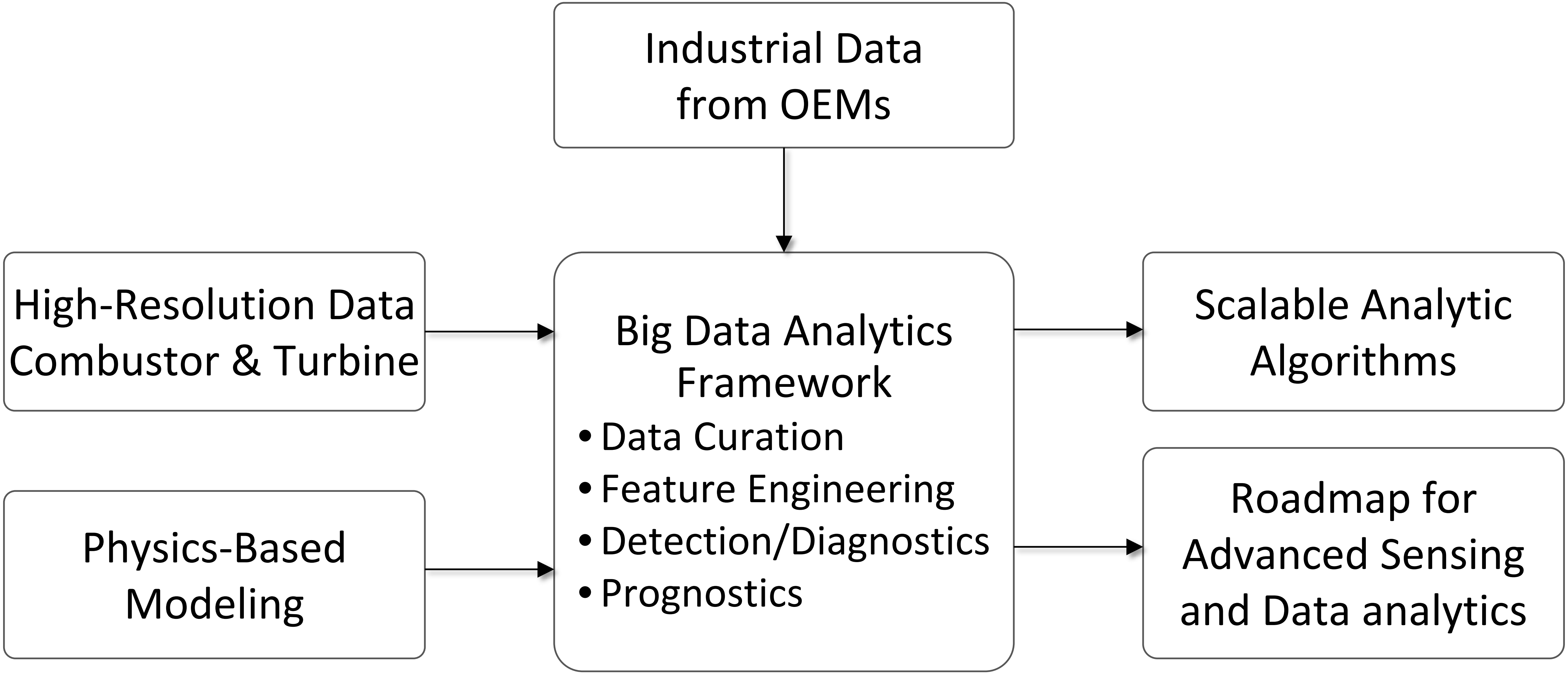
Georgia Tech Combustor Fault Rig



PSU START Turbine Facility



Research Objective, Scope, and Deliverables



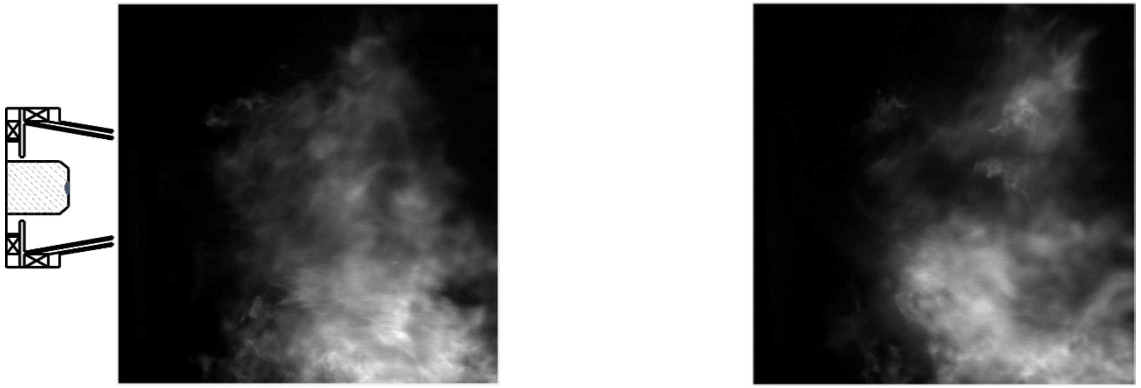
Faults

- Combustor
  - Combustor blowout aka Lean Blowout (LBO) or blowoff
    - Refers to loss of combustor flame
    - Occurs with little warning
    - Causes an immediate trip in the plant → long downtimes
- Gas Turbine
  - Coolant Loss
    - Leads to ingestion of hot air from main gas path to underplatform components
    - Catastrophically hot temperatures

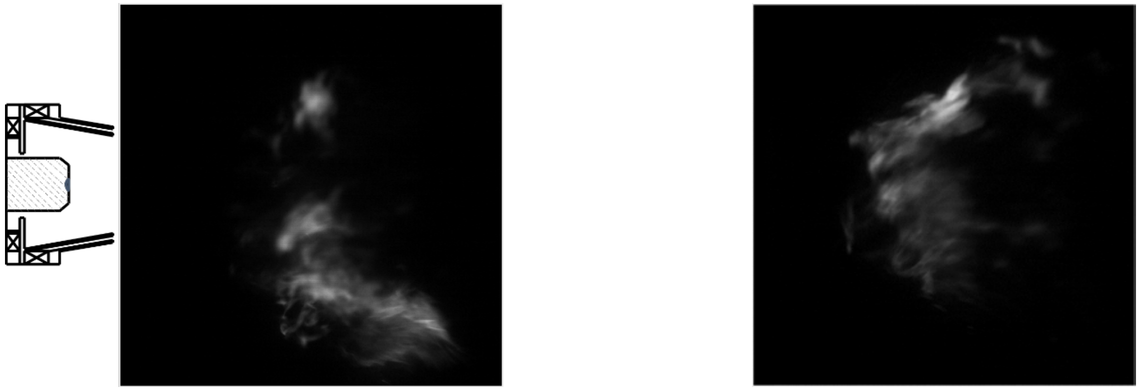
Detection of early onset of combustor blowout using control charts

Lean Blowout

- High speed chemiluminescence images demonstrate high level precursors to blowout
  - Dimming of flame intensity
  - High volatility of the flame
- However, there are no sufficient physics-based models for extracting precursors to blowout
- Therefore, we turn to a data-driven approach



Far from LBO, Φ=0.52

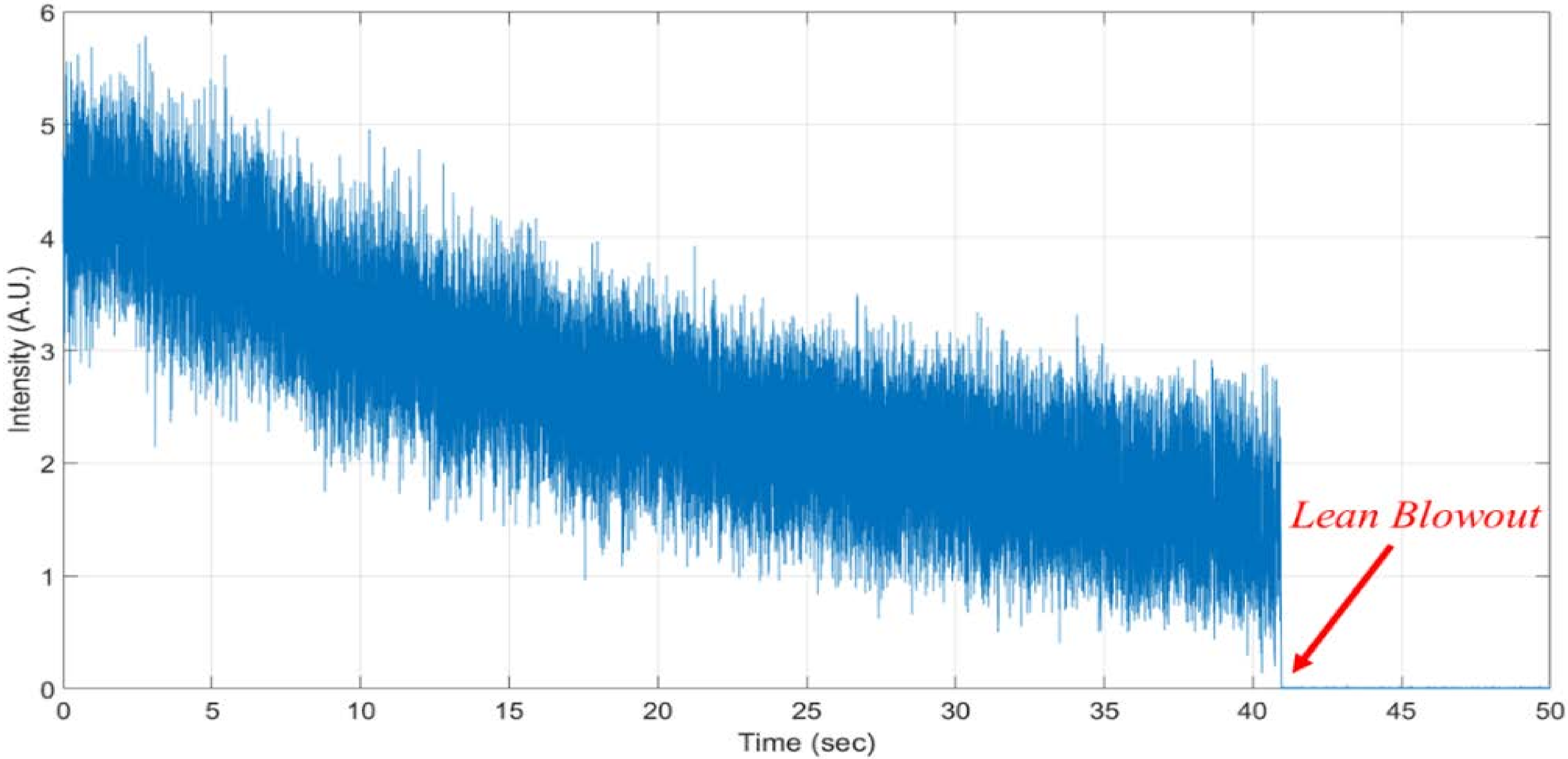


Near LBO, Φ=0.33

Experiment

- Procedure
  - Starting from a high fuel-to-air ratio [equivalence ratio (EQR)]
    - Flame is monitored as operator manually decreases EQR until flame blows out
    - Repeated 10 times for 10 fuels and 2 Air Temperatures
- Data
  - Photomultiplier Tube (10 kHz)
    - Aggregation of light intensity into a singleton point
    - Low resolution, but high frequency and can capture the entirety of the experiment
  - High speed images (4 kHz)
    - High resolution, but lacks ability for long, sustained monitoring due to lack of storage capacity

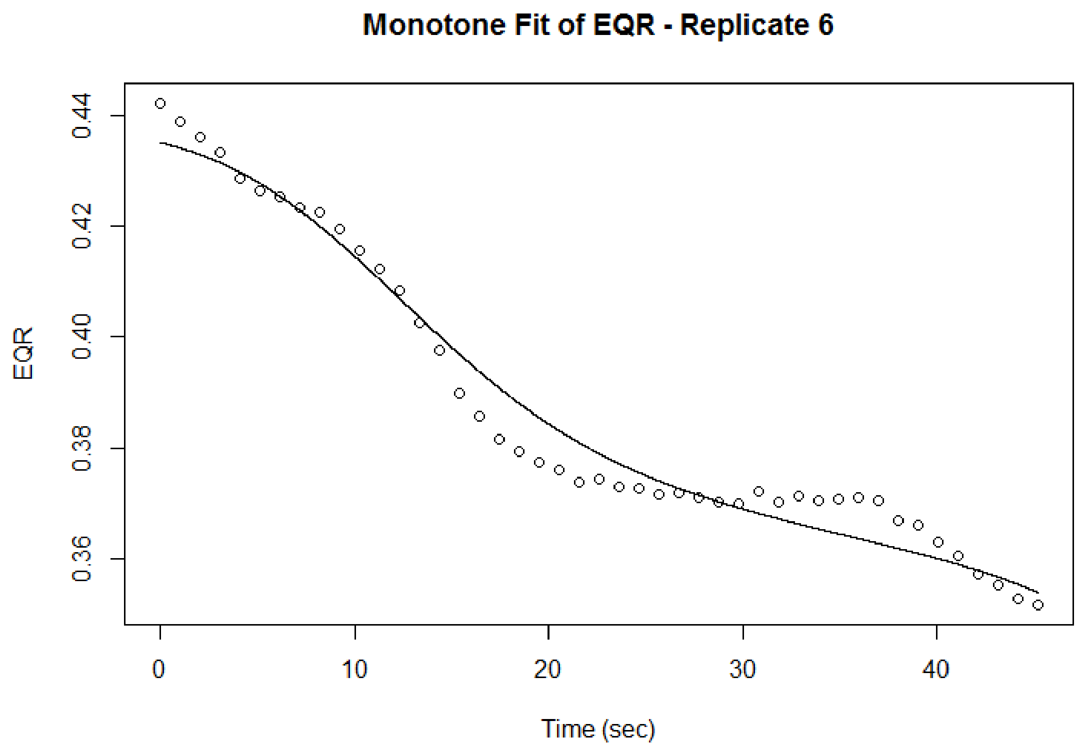
PMT Signal



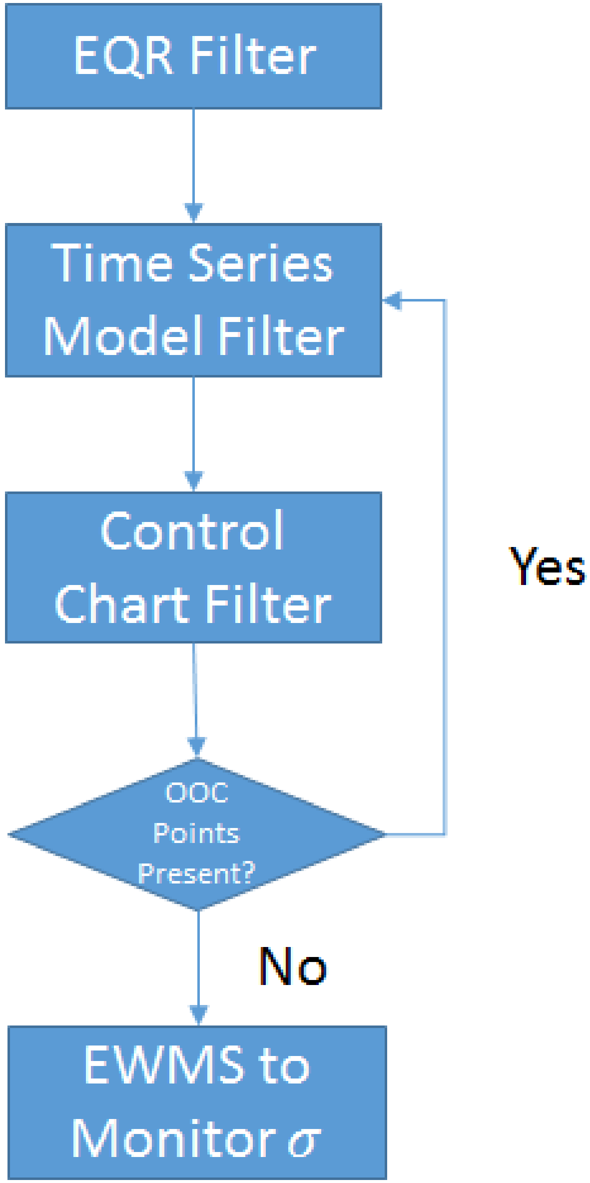
Intensity (A.U.)
0
1
2
3
4
5
6
0
5
10
15
20
25
30
35
40
45
50
Time (sec)
Lean Blowout

Data-driven methodology for LBO detection

- Objective: Using PMT data, develop a monitoring procedure for early detection and prediction of lean blowout
  - Considering the EQR level
- Preprocessing
  - Non-overlapping moving average of size 10 to reduce noise in PMT data
  - EQR is sampled at 1 Hz
    - Up-sample using monotone smoothing



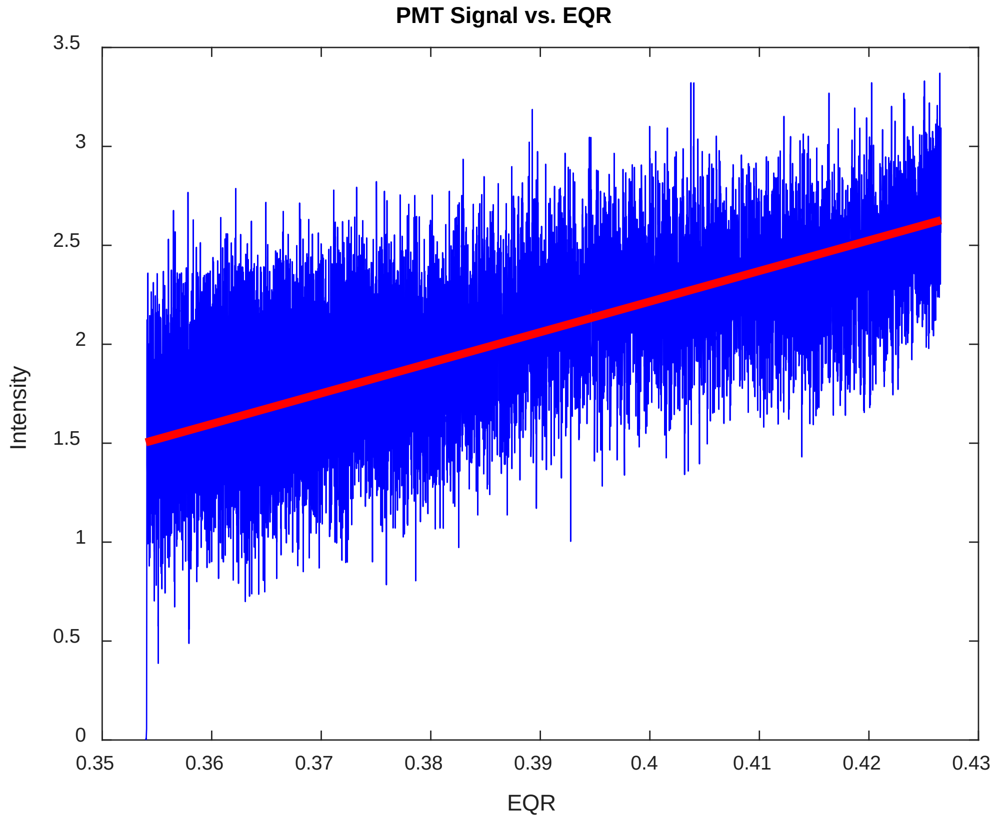
Methodology Overview



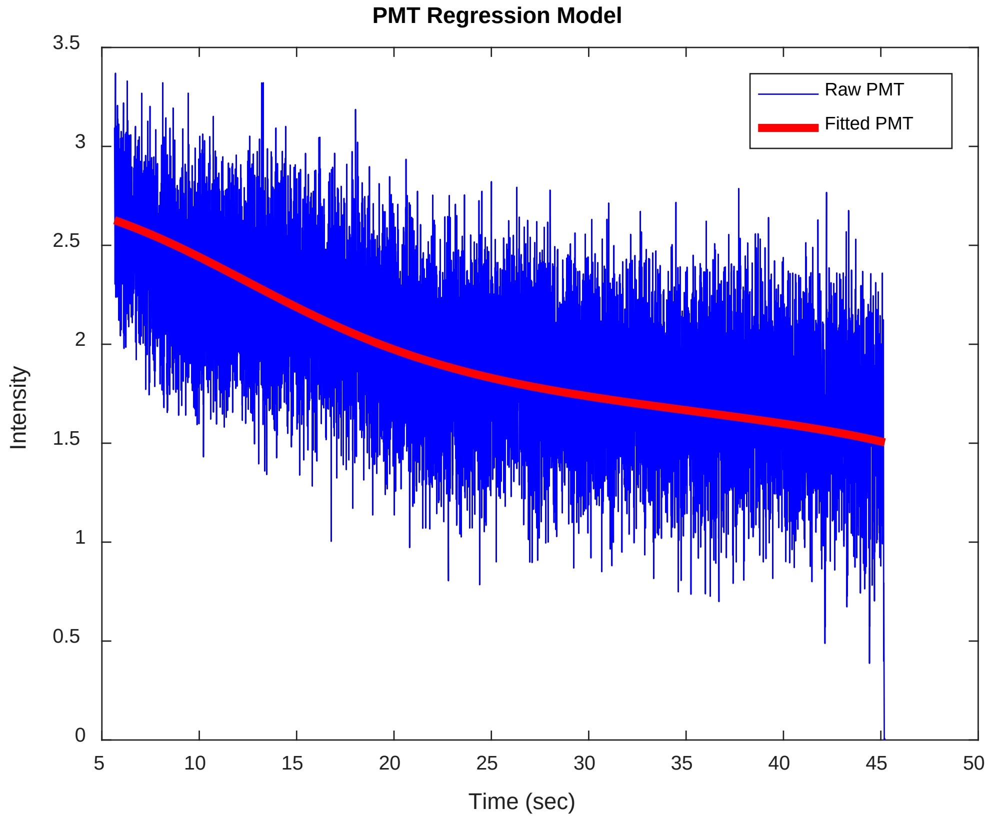
1. As the trend in the PMT signal is direct consequence of the EQR change, we filter out the effect of the EQR on the PMT.
2. The autocorrelation of regression residuals is removed using time-series models.
3. The outliers in training data are detected and removed using Shewhart control charts.
4. As PMT signals become more volatile close to lean blowout, an EWMS control chart is used to detect change in the variance.

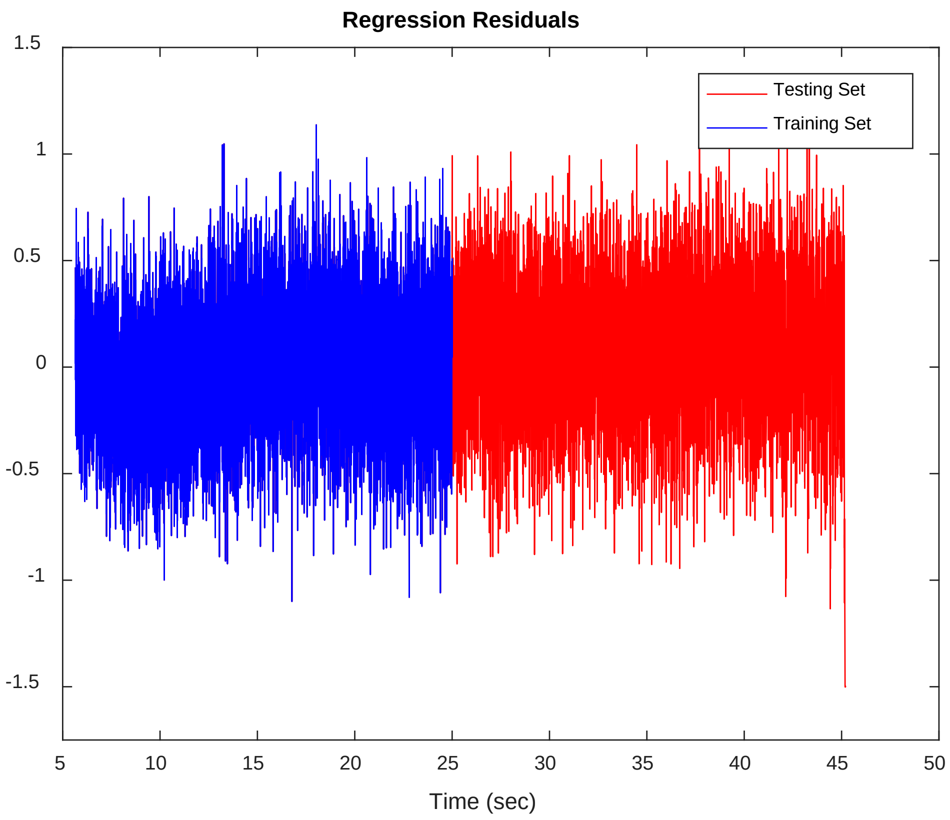
EQR Filter

- We assume the other replications represent a historical dataset and regress the PMT signal on the EQR signal for each of the replications
- For the fielded replication, the PMT signal is estimated by averaging the prediction of the models made at the current EQR value



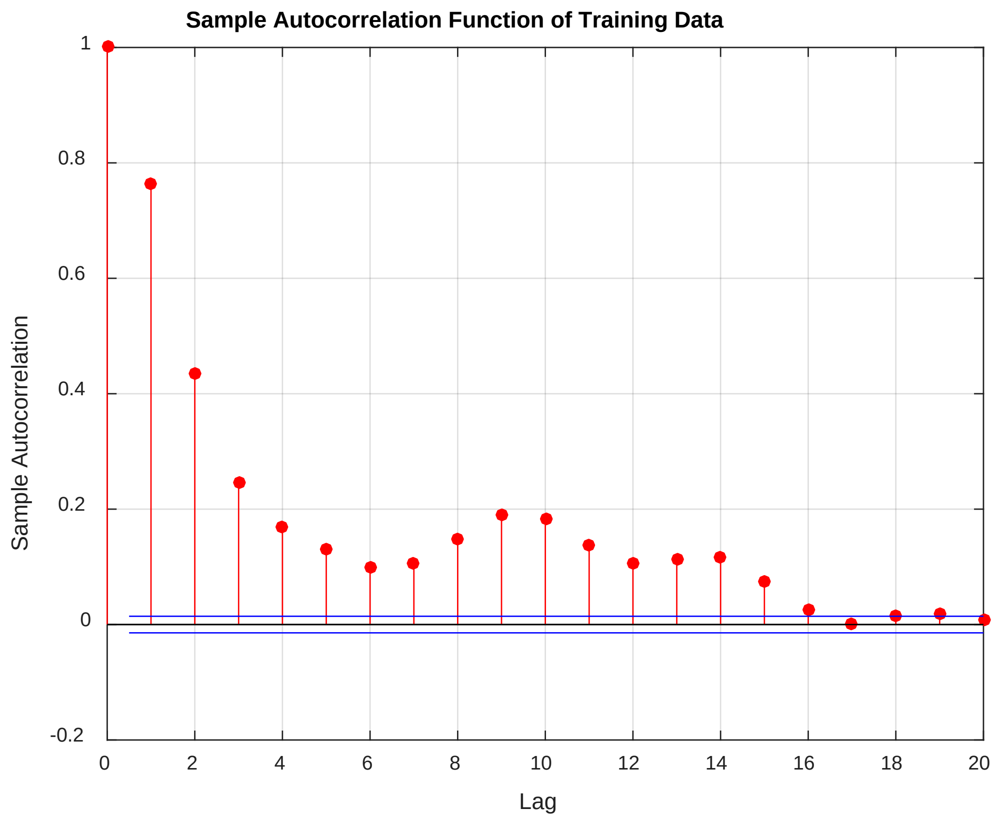
Split into Training and Testing





Time Series Model

- Autocorrelation is the correlation of a signal with a lagged copy of itself
  - Function of the lag
- Time Series Modeling is an approach to remove this autocorrelation
- We propose using ARMA-GARCH



ARMA-GARCH

- Let $Y_1, Y_2, \dots, Y_t, \dots, Y_N$ denote the training regression residuals
- ARMA(P,Q)-GARCH(U,V)

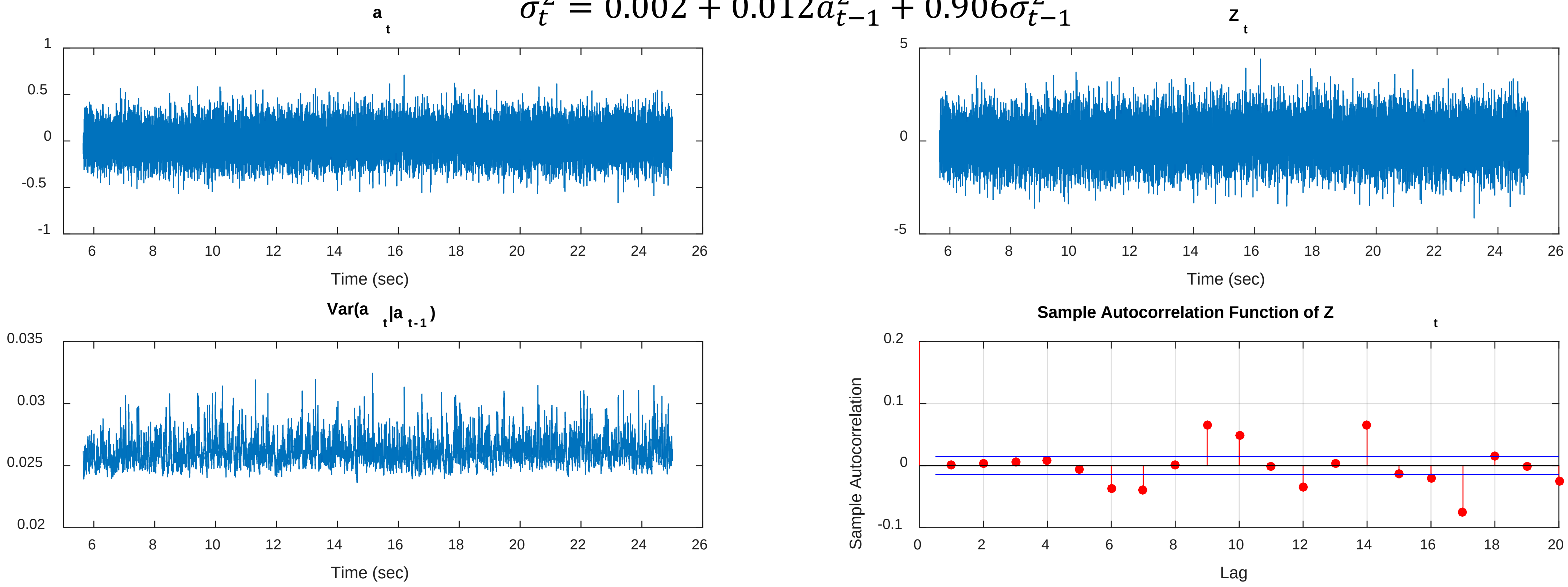
$$Y_t = \mu + \sum_{i=1}^{P} \phi_i Y_{t-i} + a_t - \sum_{j=1}^{Q} \theta_j a_{t-j}$$

$$a_t = \sqrt{\sigma_t^2 Z_t}, \sigma_t^2 = Var(a_t | a_{t-1}, a_{t-2}, \dots), Z_t \sim NID(0,1)$$

$$\sigma_t^2 = \alpha_0 + \sum_{u=1}^{U} \alpha_u a_{t-u}^2 + \sum_{v=1}^{V} \beta_v \sigma_{t-v}^2$$

Fit ARMA(1,4)-GARCH(1,1)

$$Y_t = 0.935X_{t-1} + a_t + 0.162a_{t-1} - 0.389a_{t-2} - 0.295a_{t-3} - 0.146a_{t-4}$$
$$a_t = \sigma_t Z_t, Z_t \sim N(0,1)$$
$$\sigma_t^2 = 0.002 + 0.012a_{t-1}^2 + 0.906\sigma_{t-1}^2$$



Shewhart Control Chart

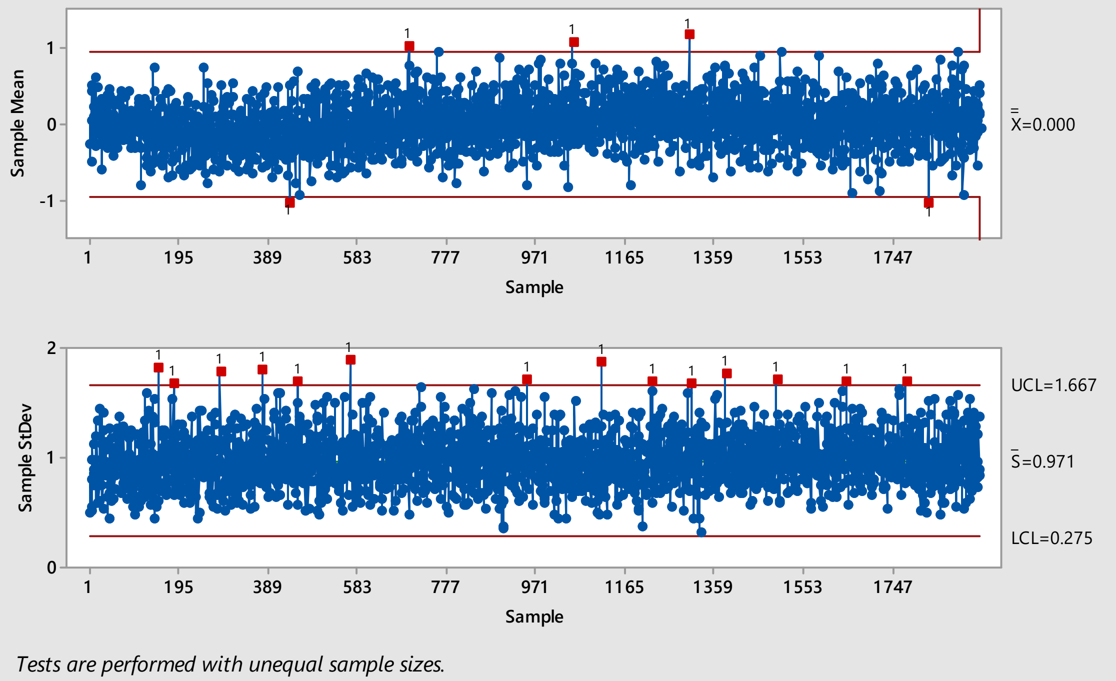
- $\bar{X}$ Control Chart
- $\bar{X}_i = \frac{1}{b}\sum_{t=(i-1)b+1}^{ib} X_t$
- $CL = \sum_{i=1}^{\frac{N}{b}} \bar{X}_i$
- $[LCL, UCL] = CL \mp 3\frac{\bar{s}}{c_4}$
- [1] c_4 is a function of the batch size and can be found in reference tables. c_4 when $b = 10$ is 0.9727

- S chart
- $s_i = \frac{1}{b-1}\sum_{t=(i-1)b+1}^{ib}(X_t - \bar{X}_i)^2$
- $CL = \bar{s} = \sum_{i=1}^{N} s_i$
- $[LCL, UCL] = CL \mp 3\frac{\bar{s}}{c_4}\sqrt{1 - c_4^2}$

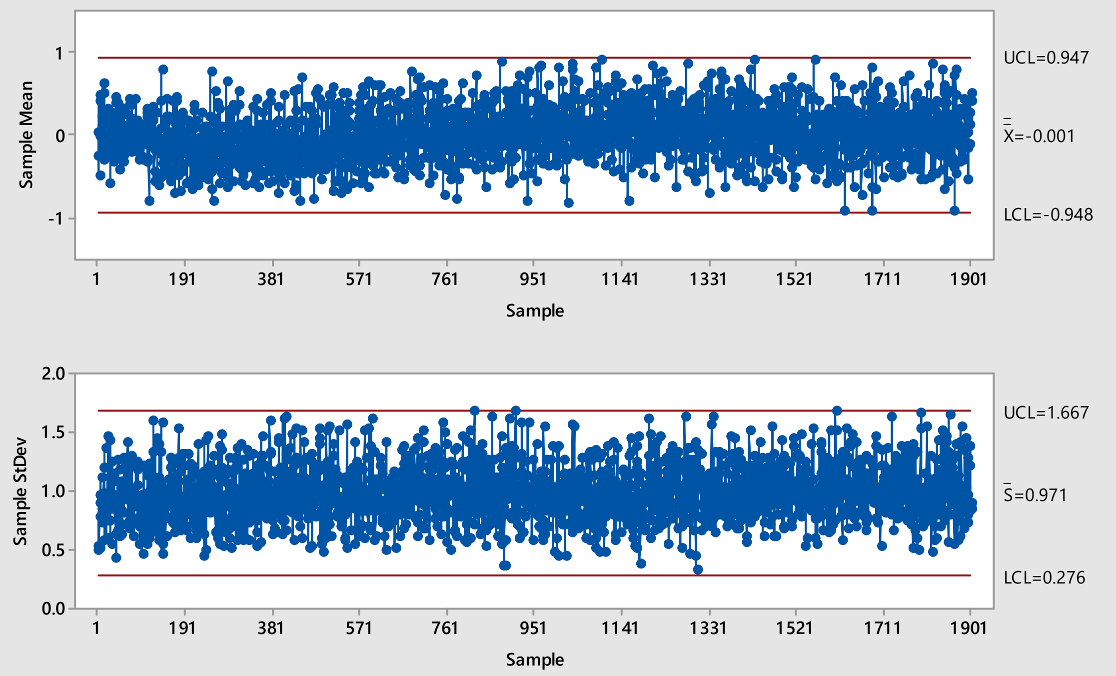
Control Chart on Z_t Process

$$Y_t = 0.935X_{t-1} + a_t + 0.156a_{t-1} - 0.391a_{t-2} - 0.295a_{t-3} - 0.143a_{t-4}$$
$$a_t = \sigma_t Z_t, Z_t \sim N(0,1)$$
$$\sigma_t^2 = 0.0254 + 0.017a_{t-1}^2$$

Xbar-S Chart of Training Data Before Filtering



Xbar-S Chart of Trianing Data After Filtering

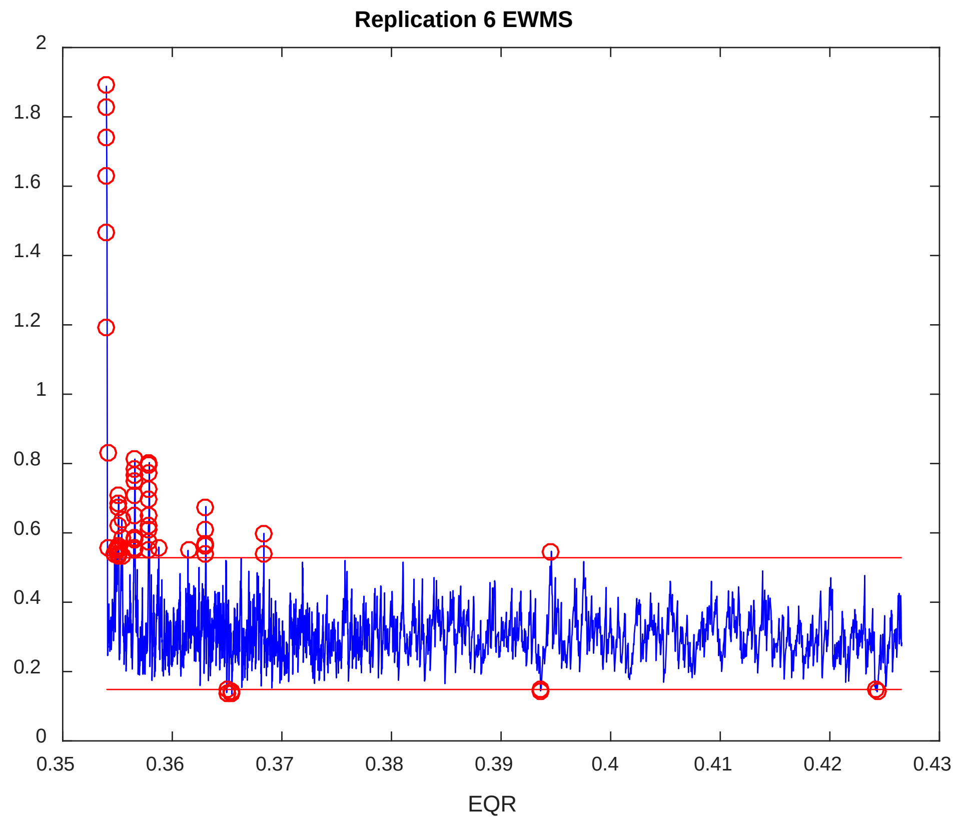


Exponentially Weighted Mean Squared deviation (EWMS)

- Test statistic

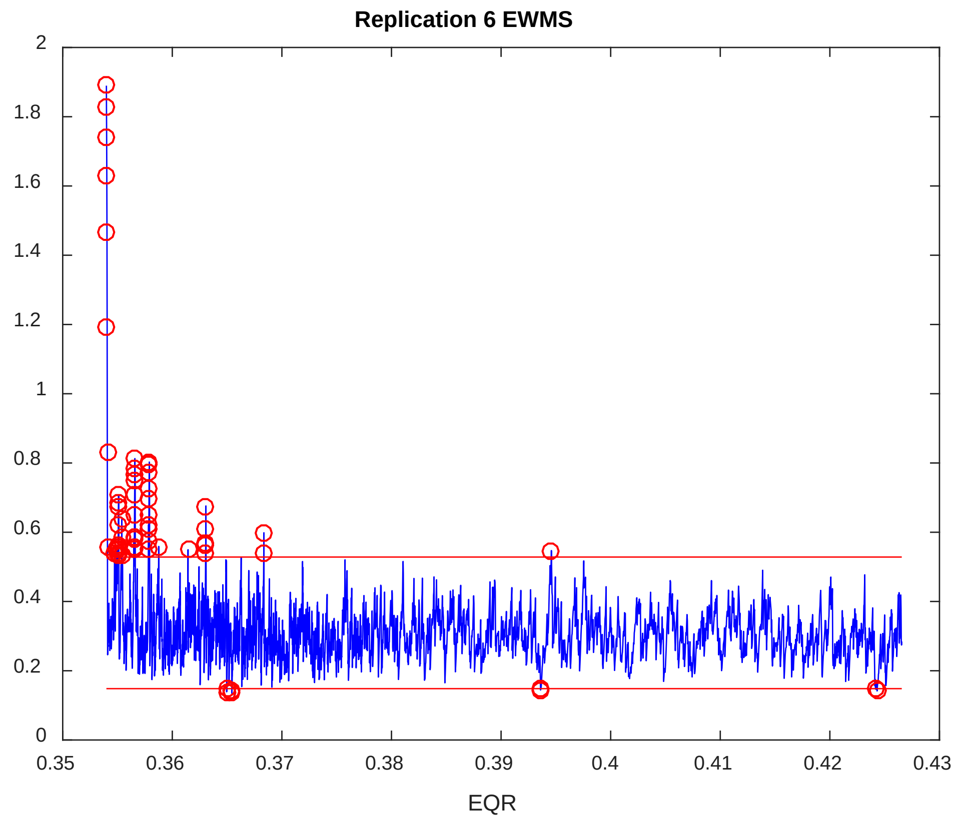
  $S_k^2 = (1-\gamma)S_{k-1}^2 + \gamma(\bar{Z}_k - \mu_0)^2$
- $\gamma \in (0,1]$
- S_0^2: Initial estimate of mean squared error

- Since $Z \sim N(0,1)$, $\bar{Z} \sim N\left(0, \frac{1}{b}\right)$
- For $b = 10$, $\mu_0 = 0$ and $S_0^2 = 0.1$

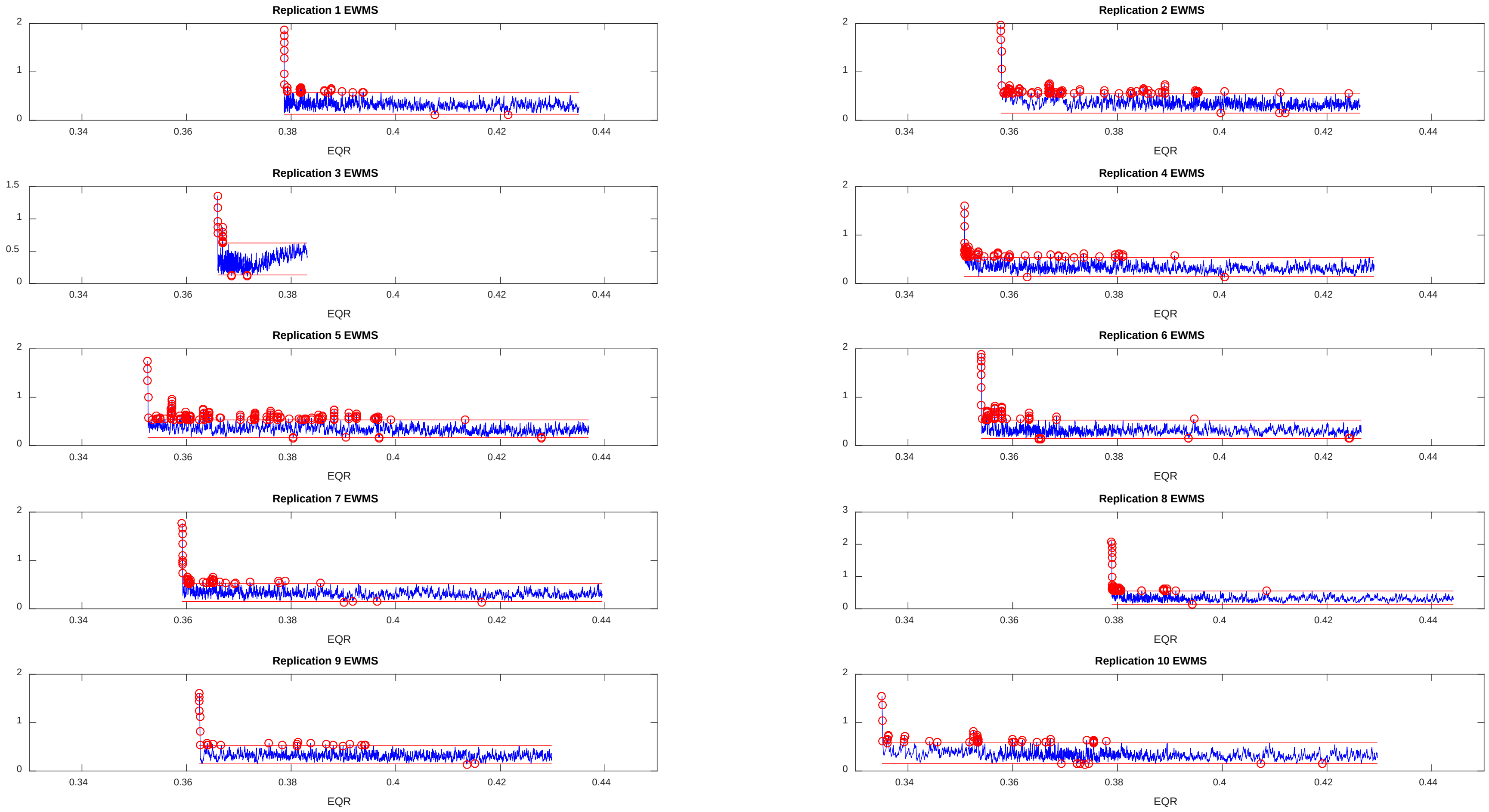


Exponentially Weighted Mean Squared deviation (EWMS)

- Control Limits
  - Compute test statistic for training data
  - For a selected confidence level α
  - $UCL = \alpha 100^{\text{th}}$ percentile
  - $LCL = (1-\alpha)100^{\text{th}}$ percentile



EWMS Control Charts

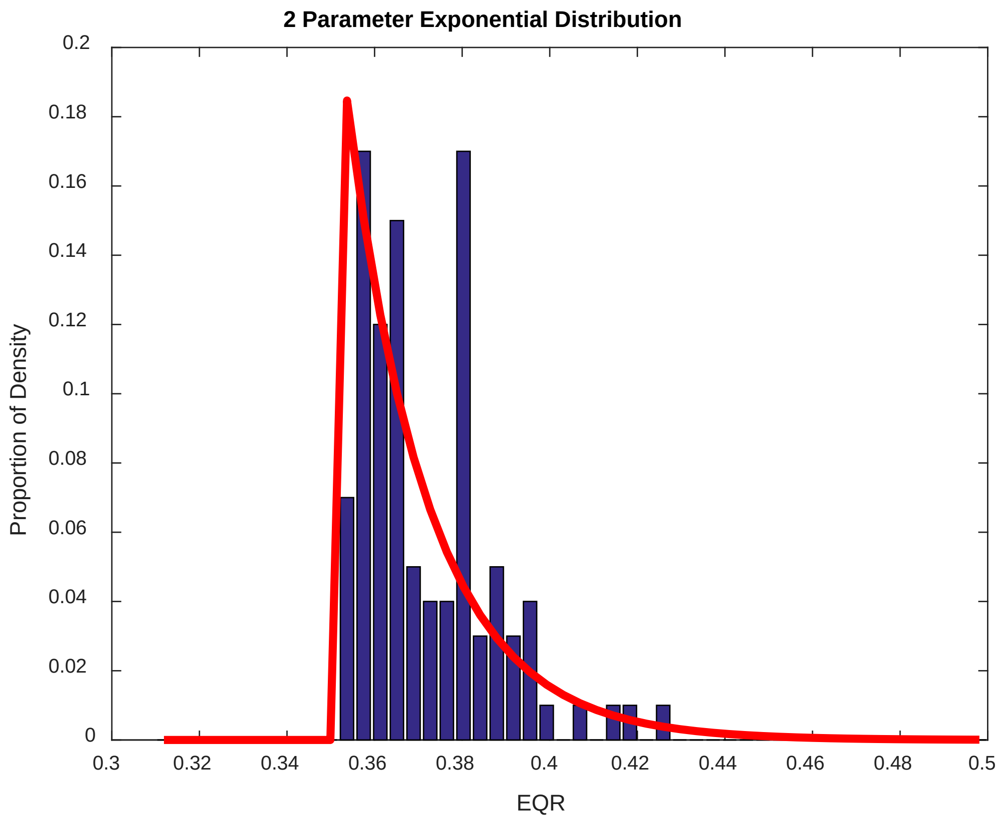


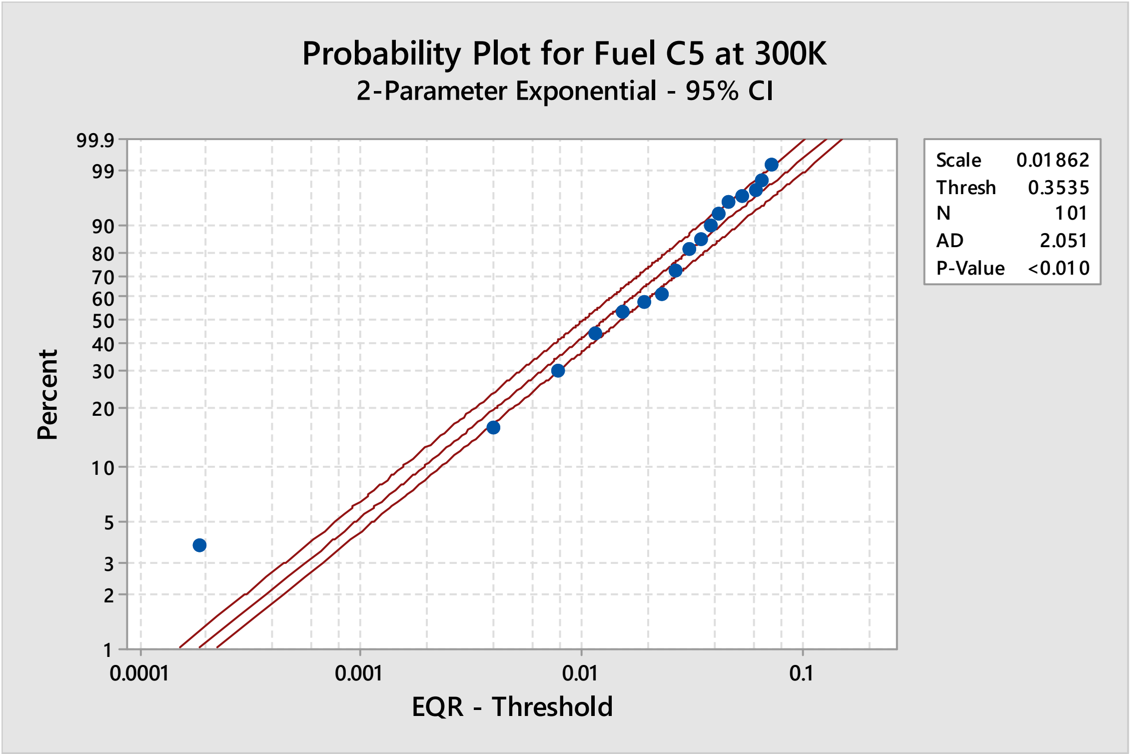
Estimating Probability of Blowout

- Using the alarms from the EWMS control chart, we estimate the probability of a blowout event
- Design a histogram (Specify bin size and number of bins)

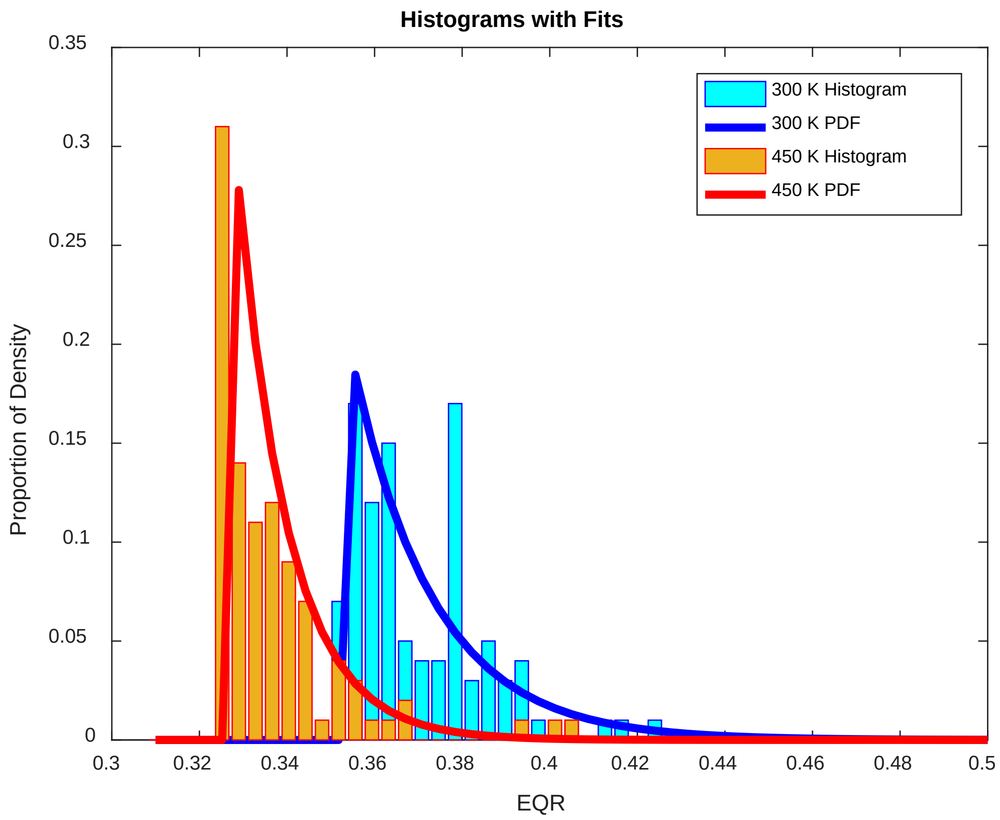
1. Threshold the alarm EQR
   - Median of blowout EQR for all replications
   - All EQR less than the median are set to this value
   - Prevents case where probability of a blowout decreases as EQR decreases
2. For each replication, compute proportion of alarms that occur within each bin
3. Average these frequencies across all replications

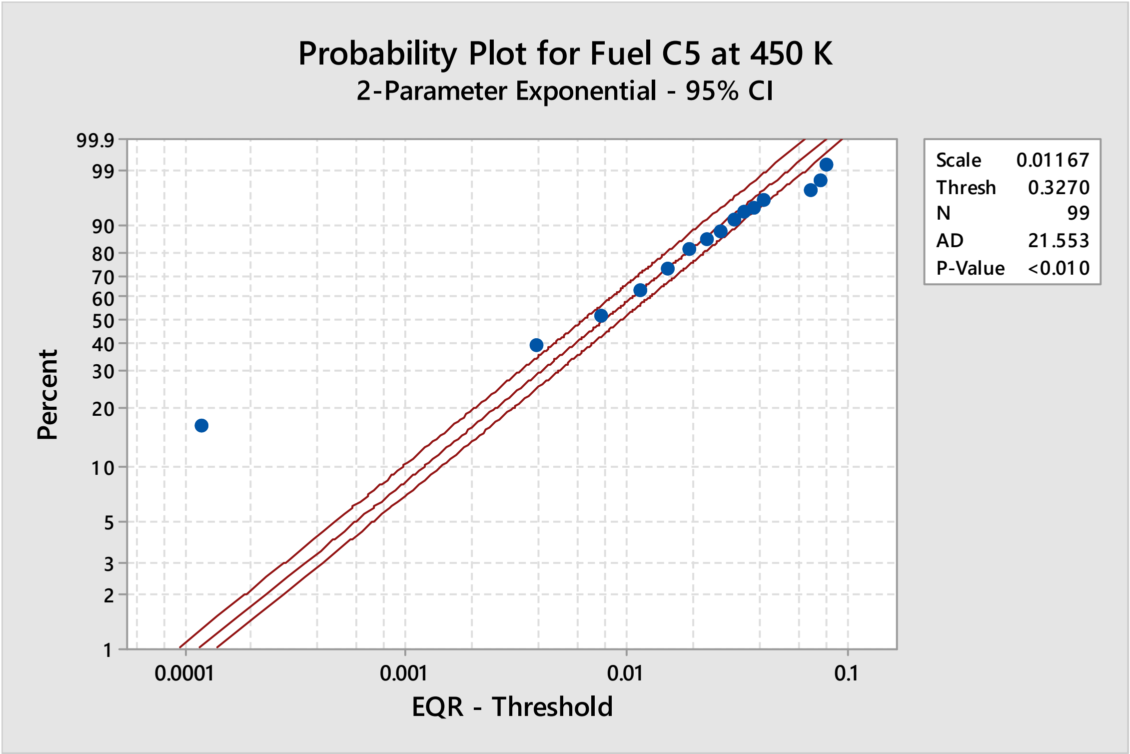
Histogram and Model Fitting





Comparison with Different Air Temperature





Parameter Estimation

- PDF: $f(t) = \lambda\, exp\big(-\lambda(t-\delta)\big), \lambda > 0, t \geq \delta$
- CDF: $P(T < t) = 1 - exp\big(-\lambda(t-\delta)\big)$
- MLE Estimates: $\hat{\delta} = \min(t_1, \dots, t_N), \hat{\lambda} = \frac{1}{\bar{t}-\hat{\delta}}$

Extension to Multiple Operating Conditions

- Let T denote a random variable
- Let $X = (X_1, \dots, X_P)$ denote a set of covariates on which the parameters of the distribution depend.
- PDF
  - $f(t|X = x) = \lambda(x)\exp\left(-\lambda(x)\big(t - \delta(x)\big)\right)$
  - $\lambda(x) > 0, \delta(x) \leq t$

Data Summary

- Data: $\{t_i, x_i\}_{i=1}^N$, $t_i \in \mathbb{R}^{D_i}, x_i \in \mathbb{R}^P$
- D_i - number of alarms in observation i
- Map the covariates to a set of basis functions e.g. $\phi(x_i) = \left(1, x_i^T\right)^T$
- Define $\phi_i \coloneqq \phi(x_i)$
- Let $\delta(x_i) = \phi_i^T \alpha$
- Let $\lambda(x_i) = \exp(\phi_i^T \beta)$ since $\lambda(x_i) > 0$
- Maximum Likelihood:

$$\arg\max_{\alpha,\beta} \prod_{i=1}^{N} \prod_{j=1}^{D_i} \exp(\phi_i^T \beta) \exp\left(-\exp(\phi_i^T \beta)\left(t_{i,j} - \phi_i^T \alpha\right)\right)$$

$$subject\ to\text{:}\ \phi_i^T \alpha \le t_{i,j} \forall j, i = 1, \dots, N$$

Likelihood Function

- Maximum Likelihood:

$$\arg\max_{\alpha,\beta} \prod_{i=1}^{N} \prod_{j=1}^{D_i} \exp(\phi_i^T \beta) \exp\left(-\exp(\phi_i^T \beta)\left(t_{i,j} - \phi_i^T \alpha\right)\right)$$

$$subject\ to:\ \phi_i^T \alpha \leq t_{i,j} \forall j, i = 1, \ldots, N$$

- Log-likelihood

$$\arg\max_{\alpha,\beta} \sum_{i=1}^{N} \sum_{j=1}^{D_i} \left(\phi_i^T \beta - \exp\left(\phi_i^T \beta\right)\left(t_{i,j} - \phi_i^T \alpha\right)\right)$$

$$subject\ to:\ \phi_i^T \alpha \leq t_{i,j} \forall j, i = 1, \ldots, N$$

Estimate α

- Log-likelihood

$$\arg\max_{\alpha,\beta} \sum_{i=1}^{N} \left(D_i \phi_i^T \beta - \exp\left(\phi_i^T \beta\right) \sum_{j=1}^{D_i} t_{i,j} + D_i \exp\left(\phi_i^T \beta\right) \phi_i^T \alpha \right)$$

$$subject\ to:\ \phi_i^T \alpha \leq t_{i,j} \forall j, i = 1, \dots, N$$

- Fixing β, this is an increasing function of $\phi_i^T \alpha$. Therefore, the log-likelihood is maximized when $\phi_i^T \alpha$ is as large as possible.
- We have the system of equations:
- $\phi_i^T \alpha = \min\left(t_{i,j}\right), i = 1, \dots, N$

Estimate β

- Fix α

$$\frac{\partial}{\partial\beta}\sum_{i=1}^{N}\left(D_i\phi_i^T\beta - D_i\exp\left(\phi_i^T\beta\right)\bar{t}_i + D_i\exp\left(\phi_i^T\beta\right)\phi_i^T\alpha\right) = 0$$

$$\sum_{i=1}^{N} D_i\left(1-\left(\bar{t}_i - \phi_i^T\alpha\right)\exp\left(\phi_i^T\beta\right)\right)\phi_i = 0$$

$$\left(\bar{t}_i - \phi_i^T\alpha\right)\exp\left(\phi_i^T\beta\right) = 1$$

$$\exp\left(\phi_i^T\beta\right) = \frac{1}{\left(\bar{t}_i - \phi_i^T\alpha\right)}$$

- System of Equations:

$$\phi_i^T\beta = -\ln\left(\bar{t}_i - \phi_i^T\alpha\right), i = 1, \dots, N$$