



Exploration & **SPACE**  
Communications

*Utilization of Unsupervised Anomalies Detector as a Tool  
For Managing the TDRS Constellation at GSFC*

Presenter: Kenneth Y. Ma

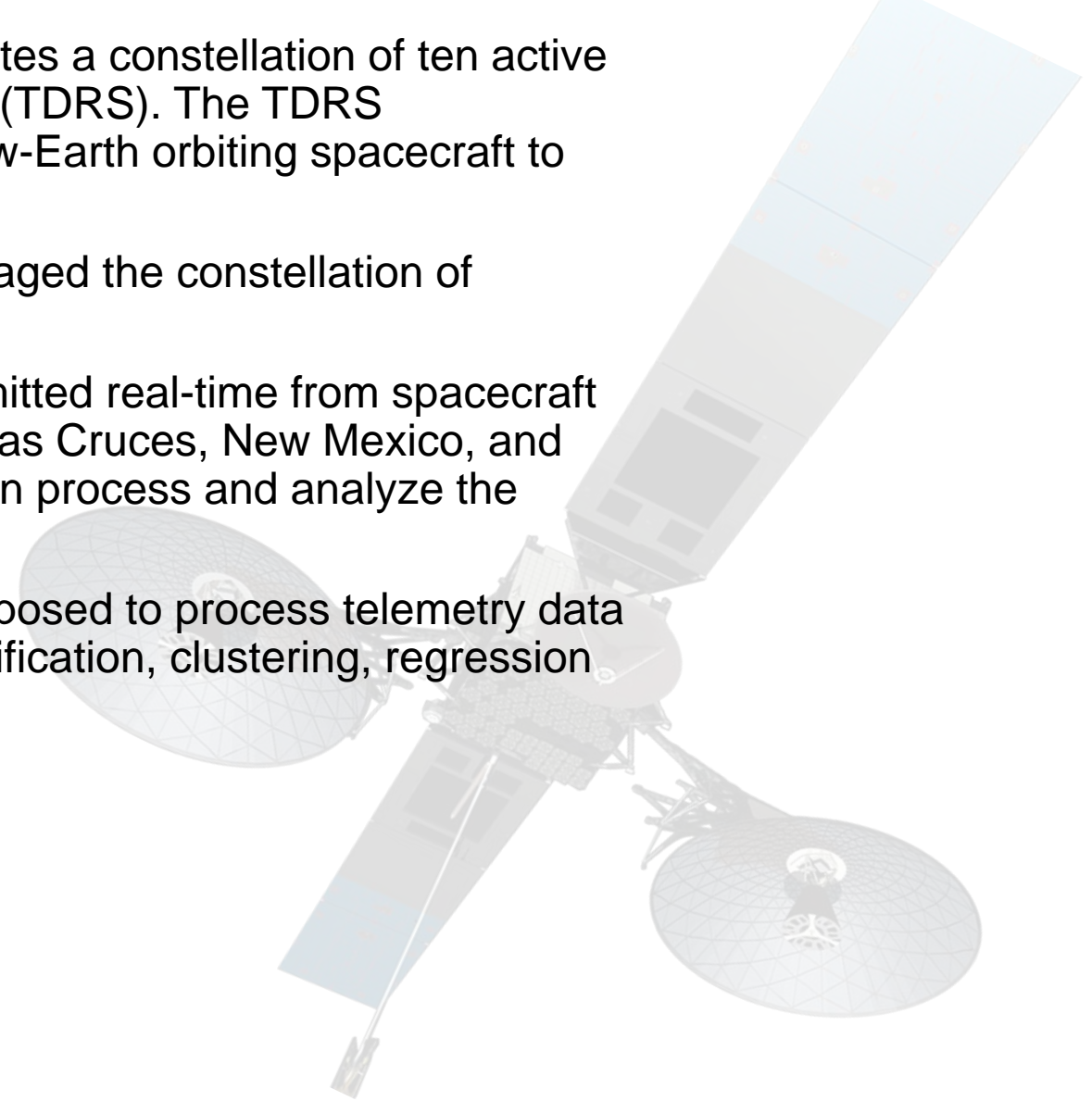
*More than you ever imagined...*



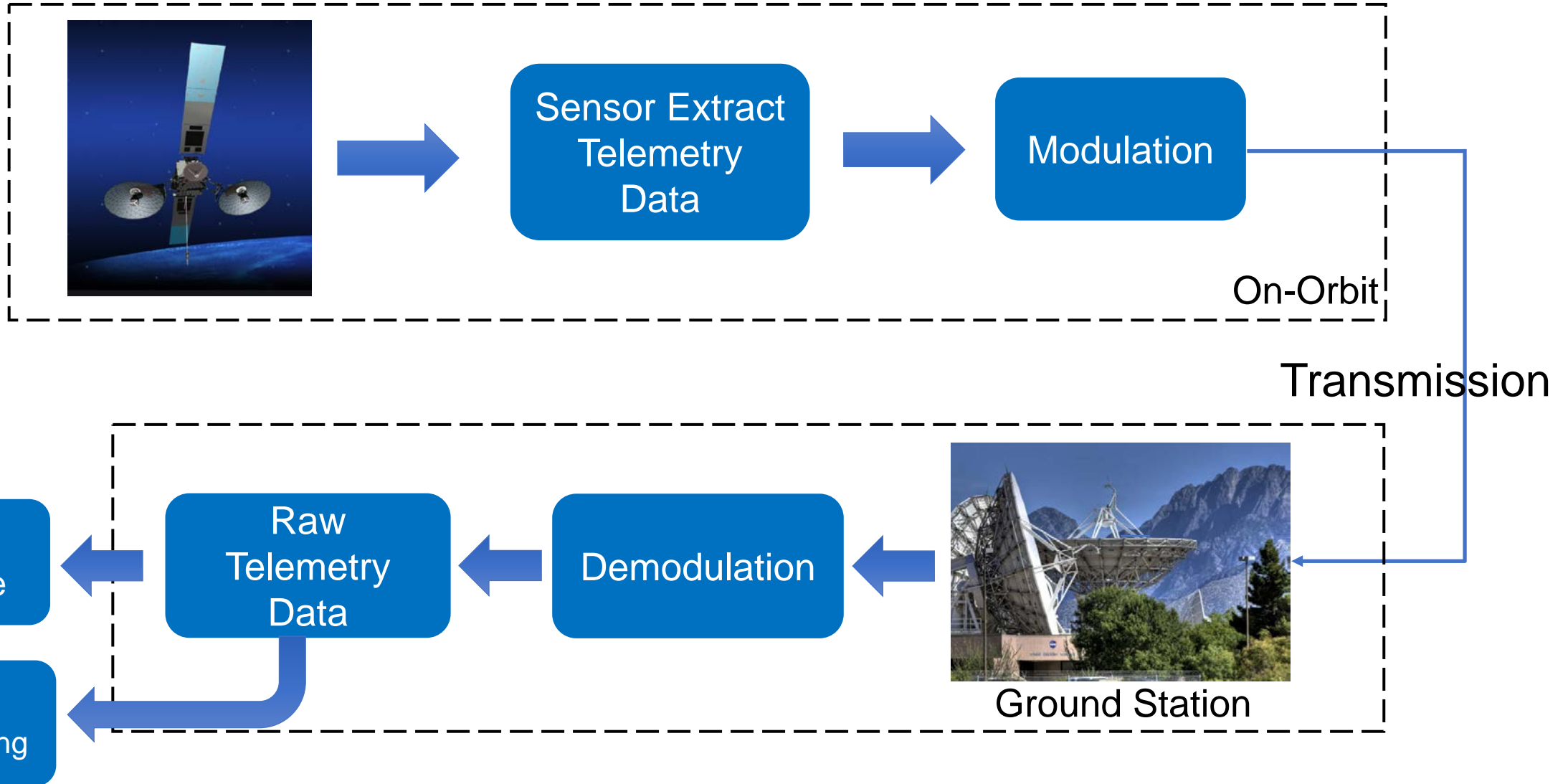
# Introduction



- NASA's Goddard Space Flight Center (GSFC) operates a constellation of ten active geosynchronous Tracking and Data Relay Satellites (TDRS). The TDRS constellation provides relay communications from low-Earth orbiting spacecraft to the primary ground stations.
- The Space Network project office at GSFC has managed the constellation of spacecraft over the past 30 years.
- Thousands of gigabytes of telemetry data are transmitted real-time from spacecraft to a ground station at the White Sands Complex in Las Cruces, New Mexico, and recorded as a historical dataset so that engineers can process and analyze the events occurring on-orbit.
- In recent years, telemetry mining (TM) has been proposed to process telemetry data by using data mining (DM) techniques such as classification, clustering, regression or anomaly detection.



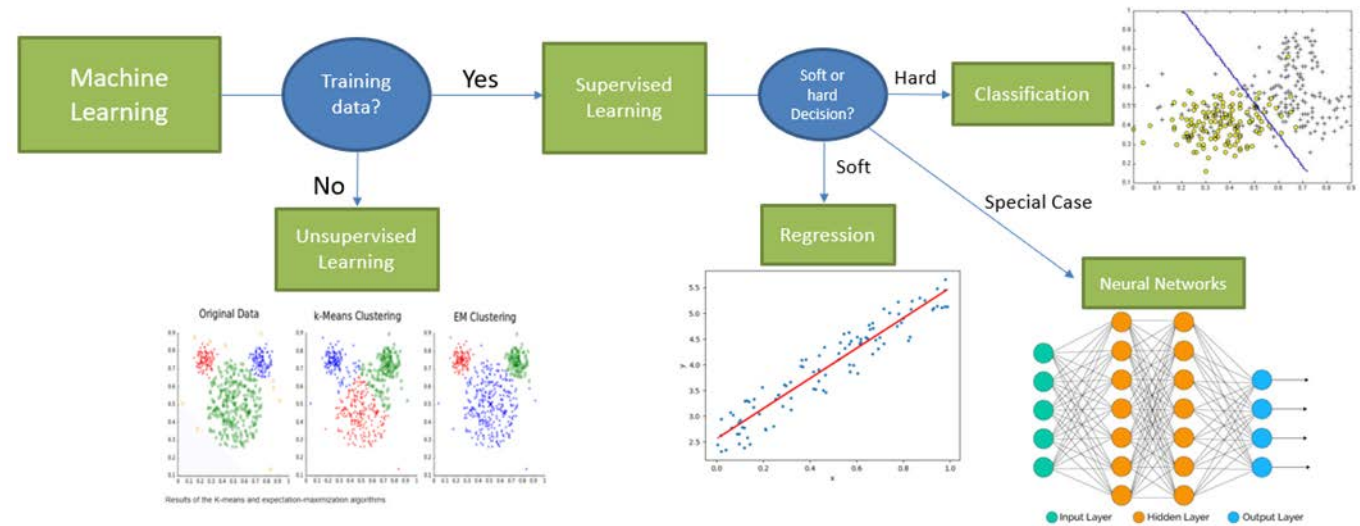
# Example Of Telemetry Data Transmission Model



# Telemetry Data Processing Review



- Traditional Limit Checking: The easiest and most widely used method for telemetry data processing is limit checking, which measures the standard deviation (SD) of each individual parameter.
- Expert system or model-based method.
- Artificial intelligence (AI) system.
- Machine Learning:
- Classification or Regression.
- Deep Learning:
- Different kinds of Neural Network.



# Telemetry Data Quantification Analysis (TDQA)

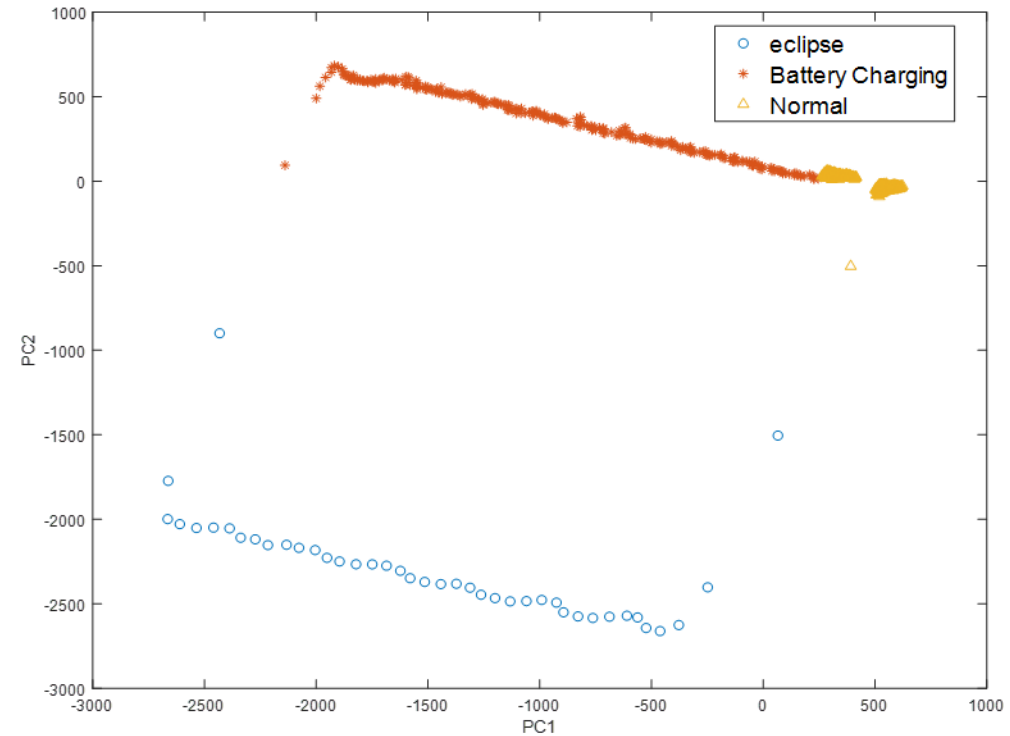


- TDQA is quite different from traditional analysis currently being used, such as expert model and limit checking.
- TDQA makes use of hundreds of different telemetry parameters to characterize spacecraft's health situation and function of behavior by using quantification algorithm.
- It provides many advantages and benefits that current analysis method cannot.

# Application in TDQA



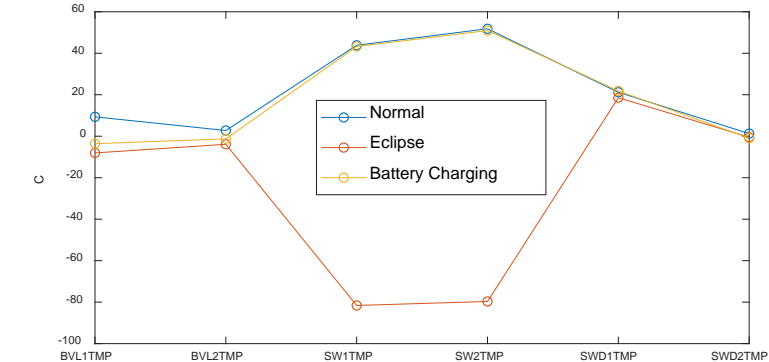
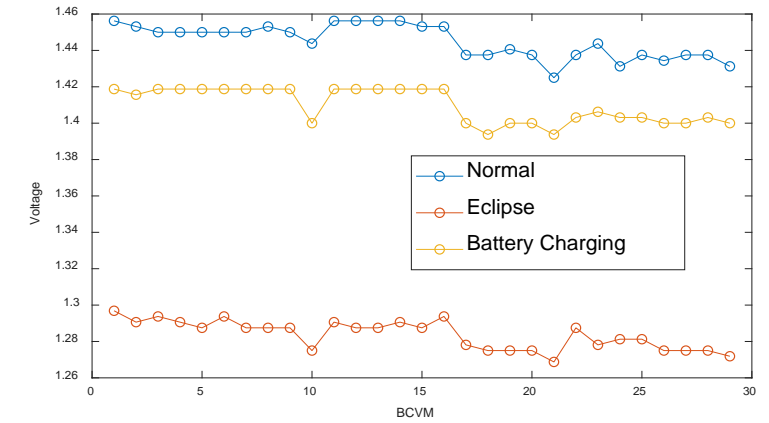
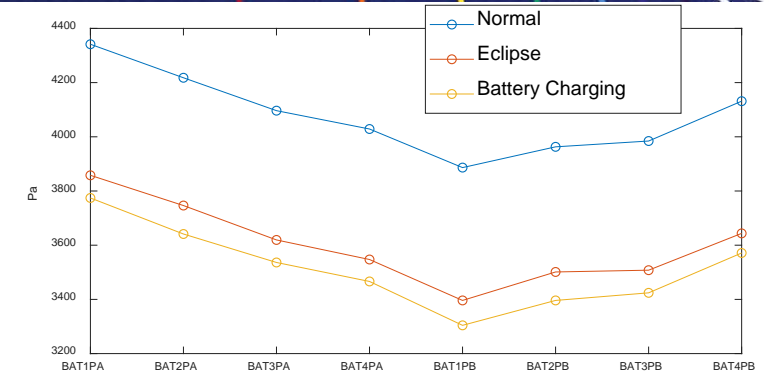
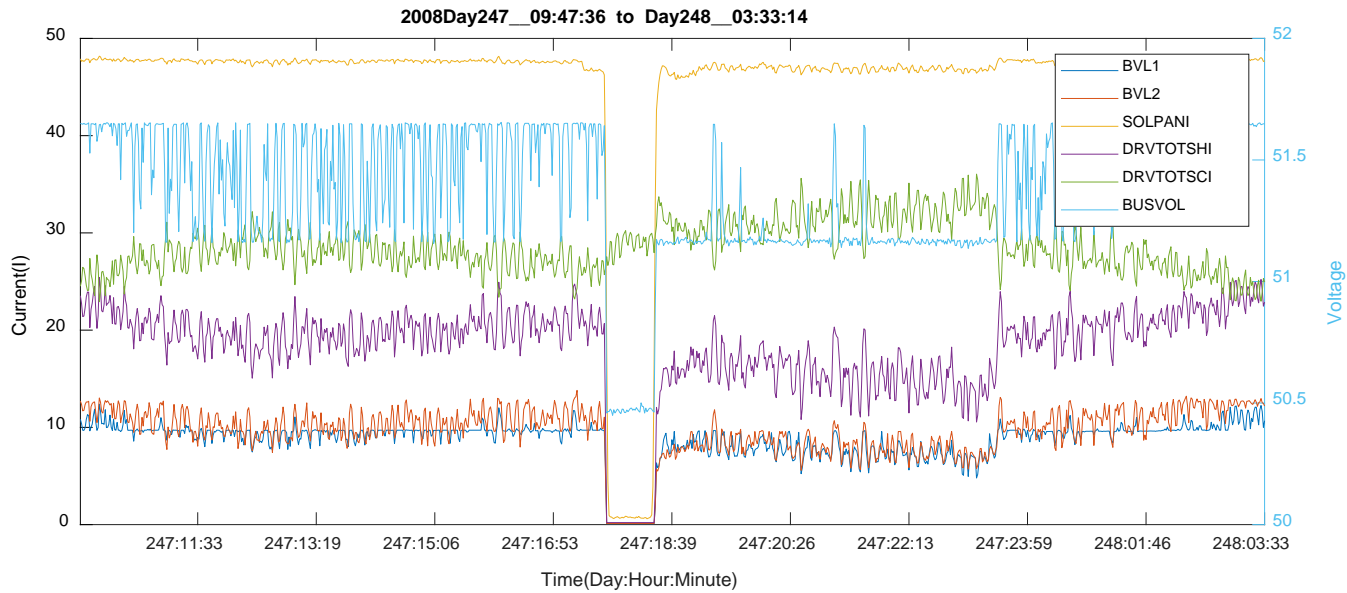
- Application in TDQA
  - Subsystem anomalies detection
  - Spacecraft behavior classification
  - Subsystem anomalies quantification
  - Battery quality quantification
- Advantage of TDQA
  - Fast analyzing
  - Soft decision rather than hard decision
  - Real time processing



# Telemetry Feature Vector



- Different parameters are connected as a vector to represent the characteristic of spacecraft in each time. The correlation between different parameters then can be included in data analysis.



# Reed-Xiaoli Detector (RXD)



- Anomaly detector (AD) is based on the assumption that the occurrence of anomalous data is very **rare** compared to normal data.
- K-RXD ( $\delta^{k-AD}$ ), which takes advantage of removing Second order statistic  $\mathbf{K}$  from original data set and calculates the anomaly score based on Mahalanobis Distance (MD):

$$\delta^{k-AD} = (\mathbf{r} - \boldsymbol{\mu})^T \mathbf{K}^{-1} (\mathbf{r} - \boldsymbol{\mu}), \quad \mathbf{K} = \frac{1}{N} \sum_{i=1}^N (\mathbf{r}_i - \boldsymbol{\mu})(\mathbf{r}_i - \boldsymbol{\mu})^T$$

- R-RXD ( $\delta^{r-AD}$ ), where the covariance matrix  $\mathbf{K}$  is replaced by correlation matrix  $\mathbf{R}$ :

$$\delta^{R-AD} = \mathbf{r}^T \mathbf{R}^{-1} \mathbf{r}, \quad \mathbf{R} = \frac{1}{N} \sum_{i=1}^N \mathbf{r}_i \mathbf{r}_i^T$$

- In real time application, the detected anomaly data should not be counted into correlation matrix:

$$\tilde{\mathbf{R}} = \mathbf{R} - \sum_{\mathbf{r}_i \in \Delta} \mathbf{r}_i \mathbf{r}_i^T$$



# Signature Discriminability Measurement (SDM)



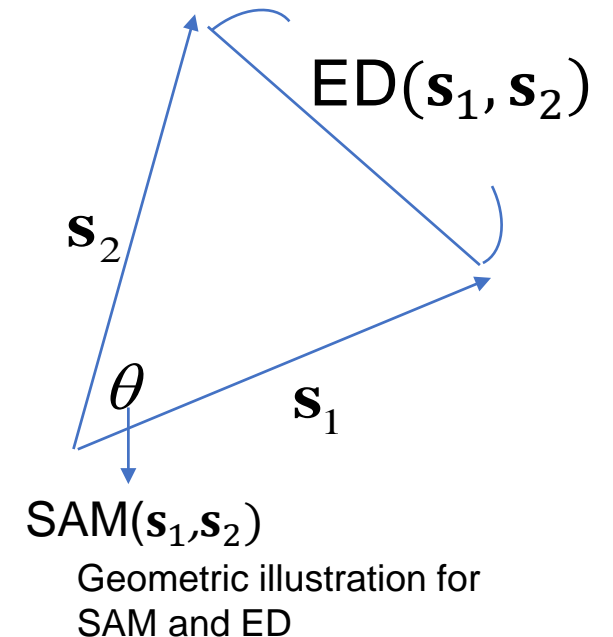
- SDM is primarily based on measuring the divergence between two input signals. This can be significantly useful to compute discriminability between health subsystem signal and failed substance signal.
- Euclidean Distance (ED):

$$\mathbf{s}_1 = (s_{11}, s_{12}, \dots, s_{1L})^T \quad \mathbf{s}_2 = (s_{21}, s_{22}, \dots, s_{2L})^T$$

$$ED(\mathbf{s}_1, \mathbf{s}_2) = \|\mathbf{s}_1 - \mathbf{s}_2\|_2 = \sqrt{\sum_{l=1}^L (s_{1l} - s_{2l})^2}$$

- Signature Angle Mapper (SAM):

$$SAM(\mathbf{s}_1, \mathbf{s}_2) = \cos^{-1} \left( \frac{\langle \mathbf{s}_1, \mathbf{s}_2 \rangle}{\|\mathbf{s}_1\| \|\mathbf{s}_2\|} \right)$$



# Signature Discriminability Measurement (SDM)



- Signature Information Divergence (SID) method using relative entropy in information theory to compute discriminability between two signals.

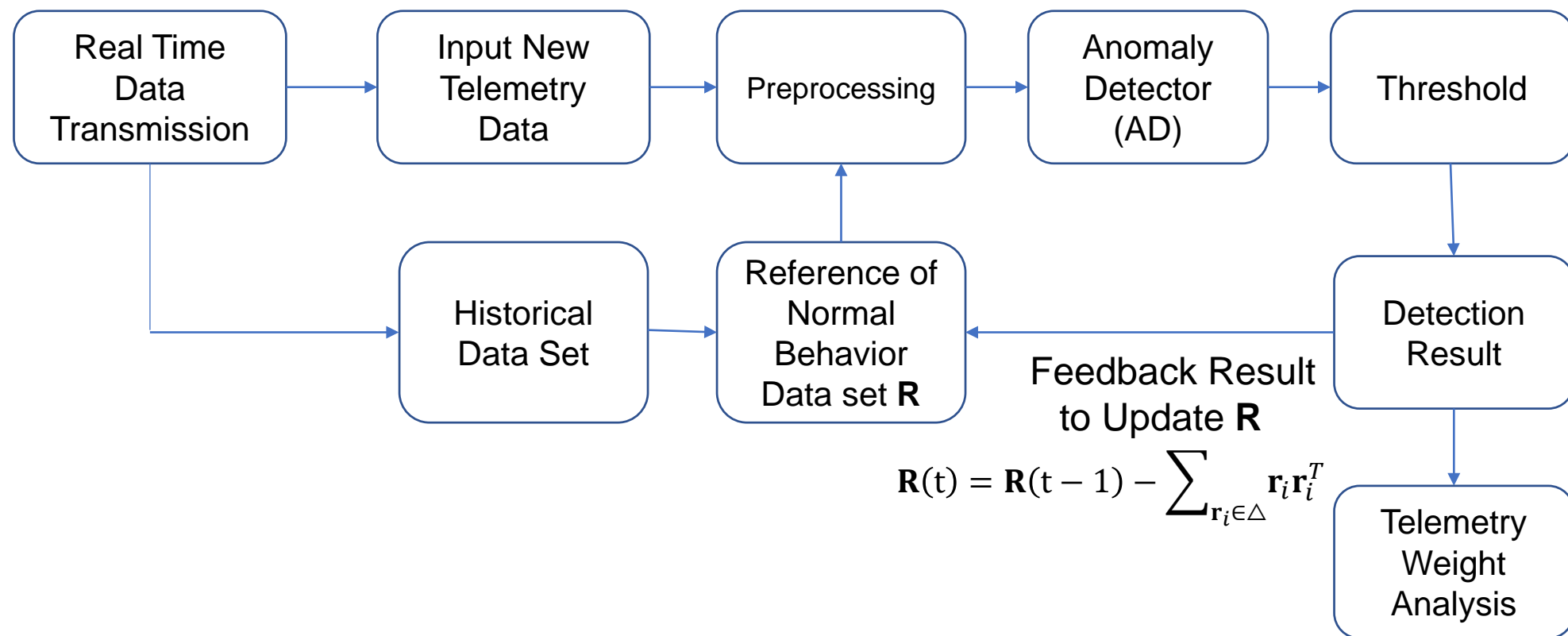
$$p_{1l} = s_{1l} / \sum_{l=1}^L s_{1l} \quad p_{2l} = s_{2l} / \sum_{l=1}^L s_{2l}$$

$$D(\mathbf{s}_1 \parallel \mathbf{s}_2) = \sum_{l=1}^L p_{1l} \log(p_{1l} / p_{2l})$$

$$D(\mathbf{s}_2 \parallel \mathbf{s}_1) = \sum_{l=1}^L p_{2l} \log(p_{2l} / p_{1l})$$

$$\text{SID}(\mathbf{s}_1, \mathbf{s}_2) = D(\mathbf{s}_1 \parallel \mathbf{s}_2) + D(\mathbf{s}_2 \parallel \mathbf{s}_1)$$

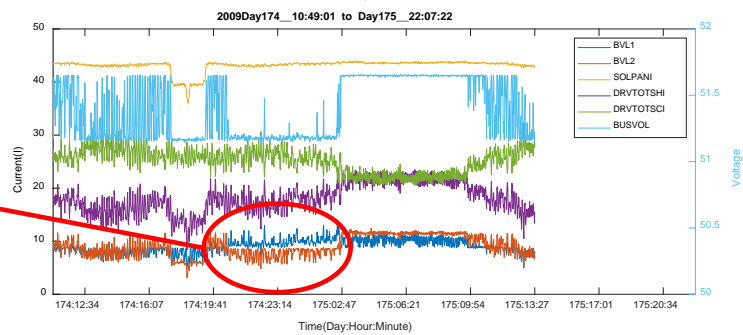
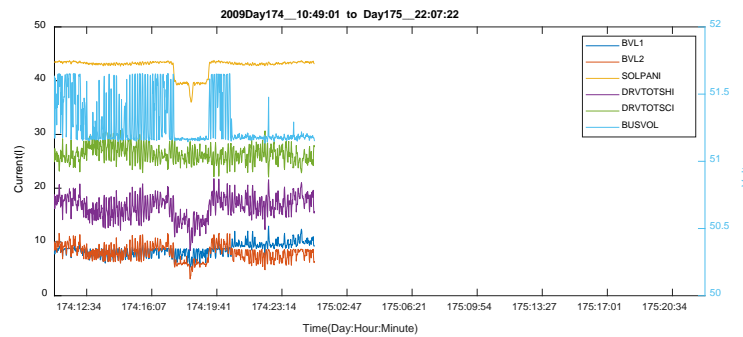
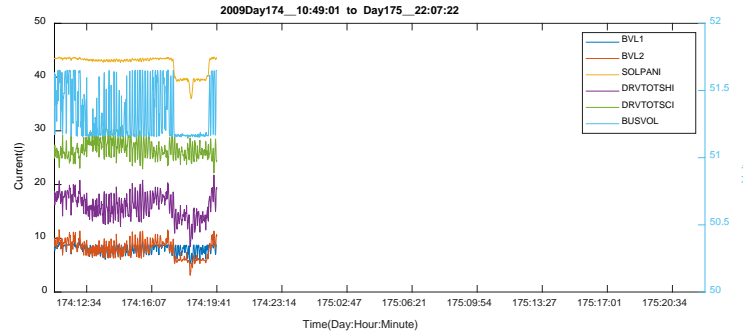
# Flow Chart of Real-Time Anomaly Detection



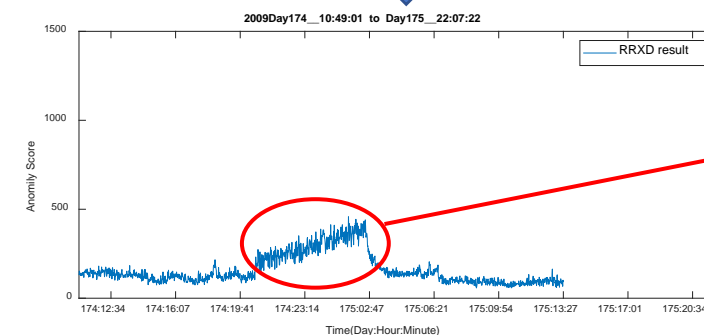
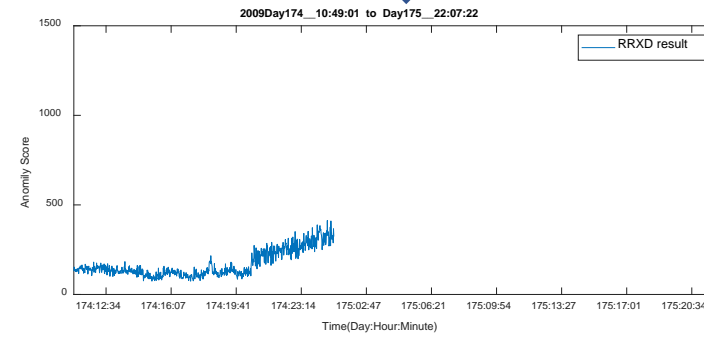
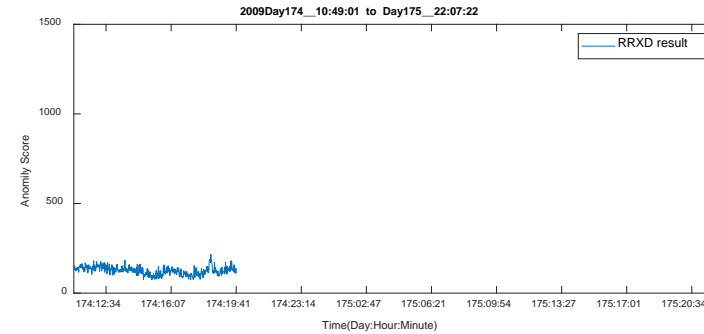
# Real-Time Anomaly Detection Processing



## Telemetry Data



## Detection Results



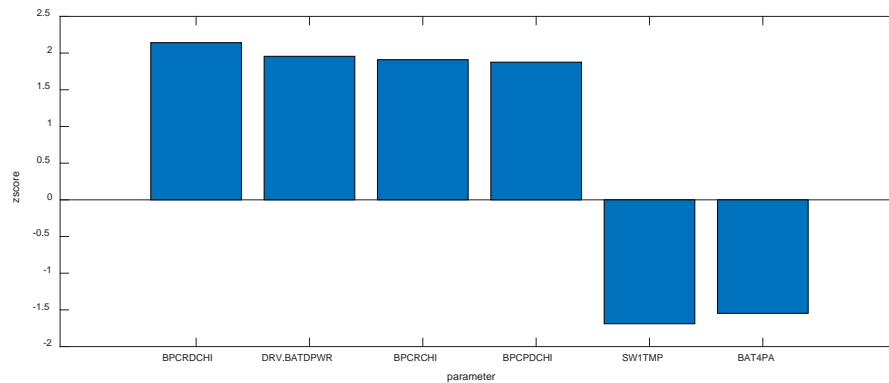
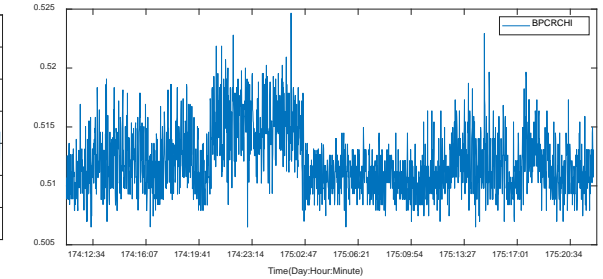
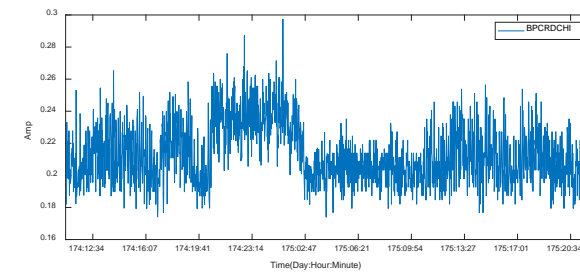
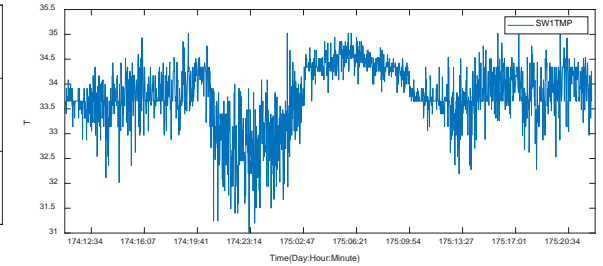
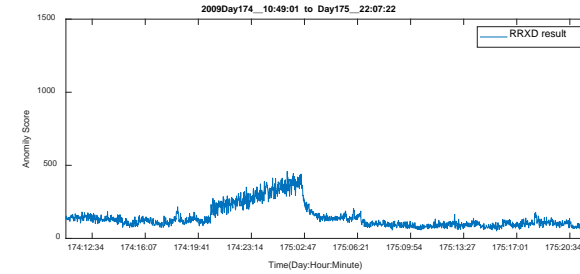
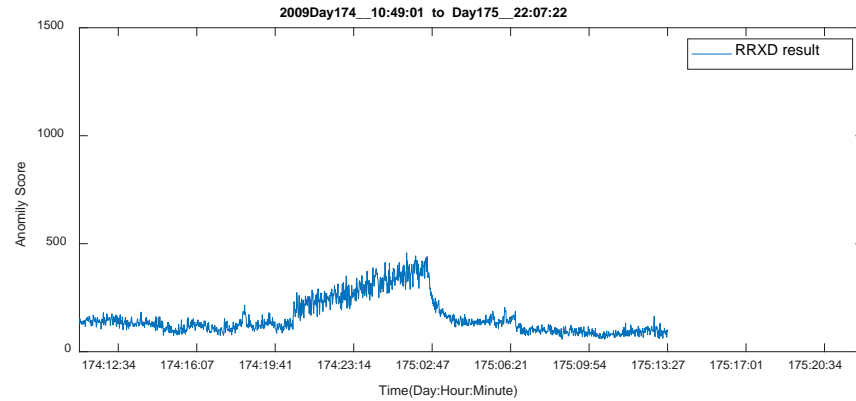
Anomaly

Anomaly

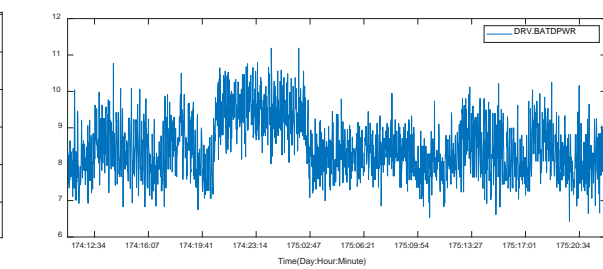
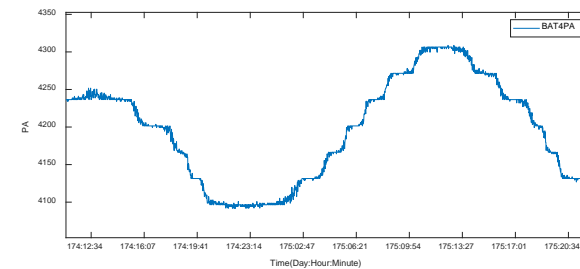
# Parameter Contribution Analysis



- Z-score is the calculation of how many standard deviations  $\sigma$  of each parameter is from mean point  $\mu$ . This approach can be implemented to quantify the contribution for each parameter to detect anomaly.



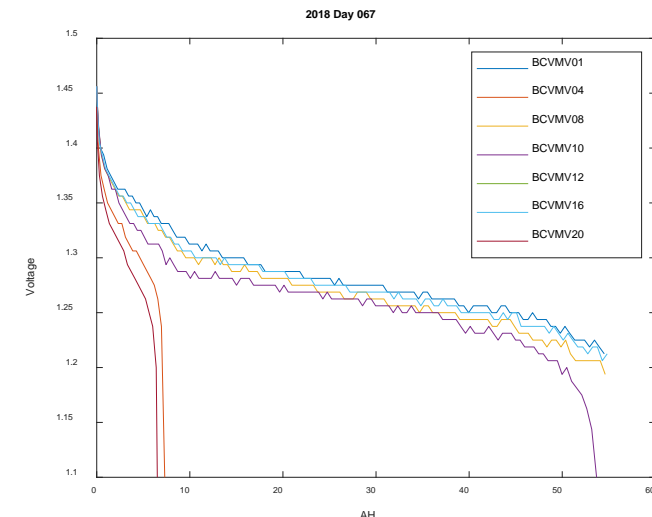
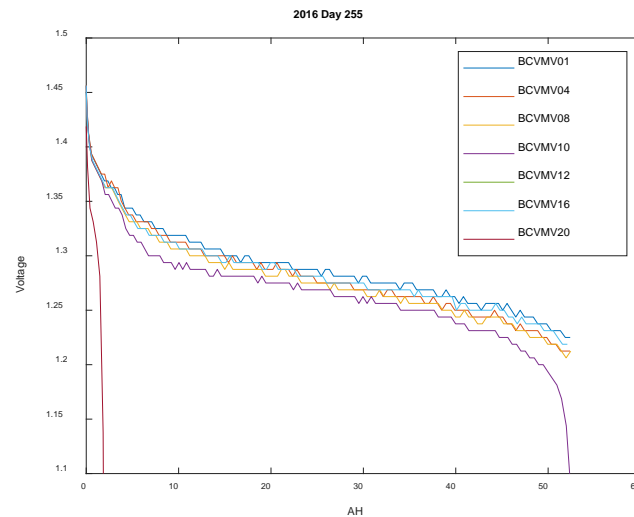
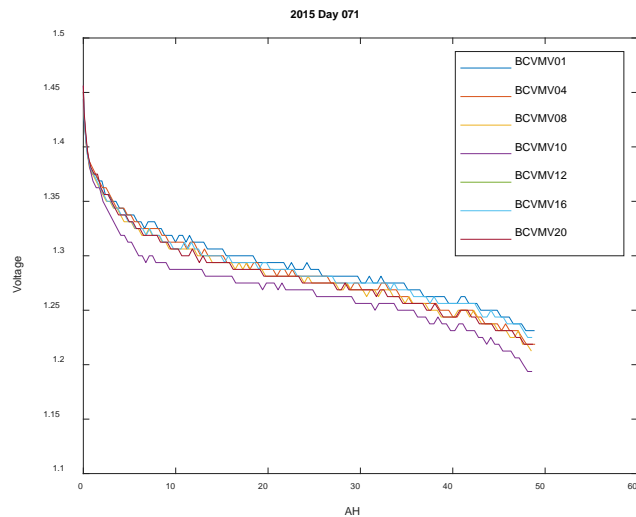
Extract data



# Battery Issues in TDRS Spacecraft



- One critical issue of TDRS batteries is the challenge of battery cell capacity estimation when the spacecraft is operating on orbit. Since the batteries never fully discharge in each eclipse condition, telemetry can only observe a partial V/Ah curve. Consequently, the current quality of the battery cells is hard to predict until the failure of a cell appears.
- Each Generation 2 TDRS spacecraft battery contains 4 packs with 29 Nickel-Hydrogen battery cells.

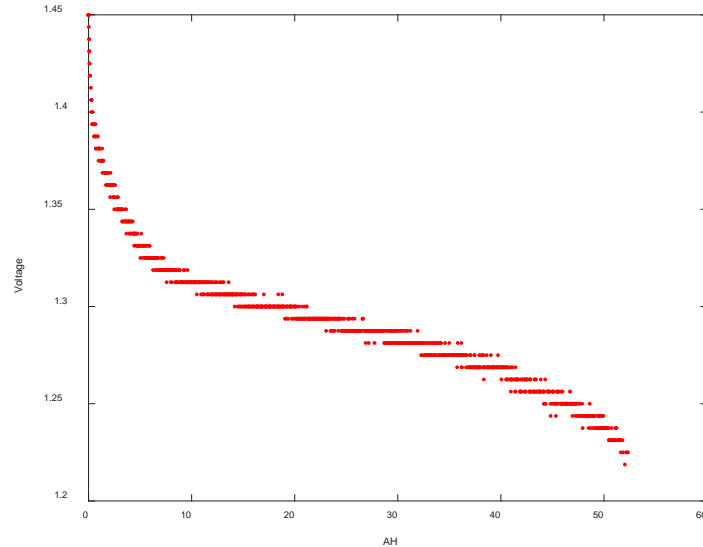


# Non-Linear Regression Approach in Estimation of True Cell Signal

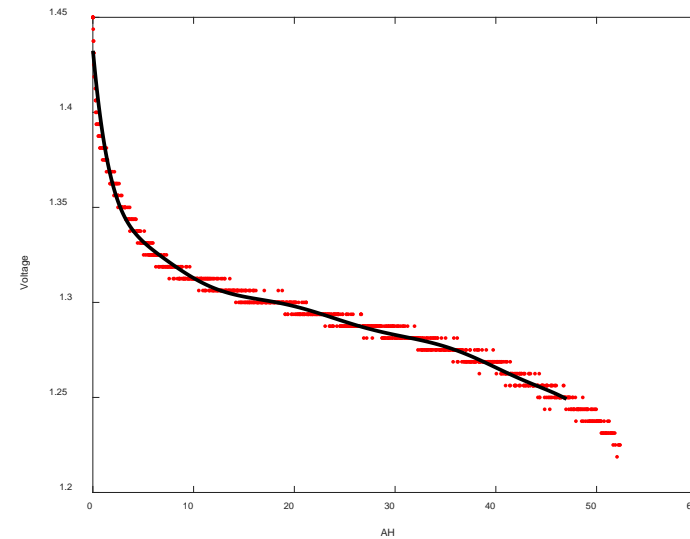


- Statistically, linear regression is an approach to modeling the relationship between a scalar response and one or more variables.
- Linear or nonlinear regression function:

$$y = \sum_{i=1}^N \alpha_i x_i^c + \alpha_o \text{ for } c \geq 1$$



Observed Cell Voltage Data

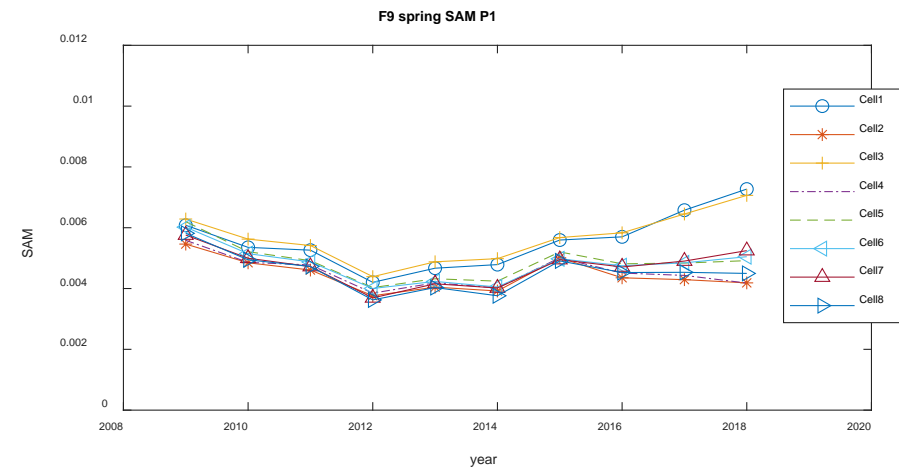
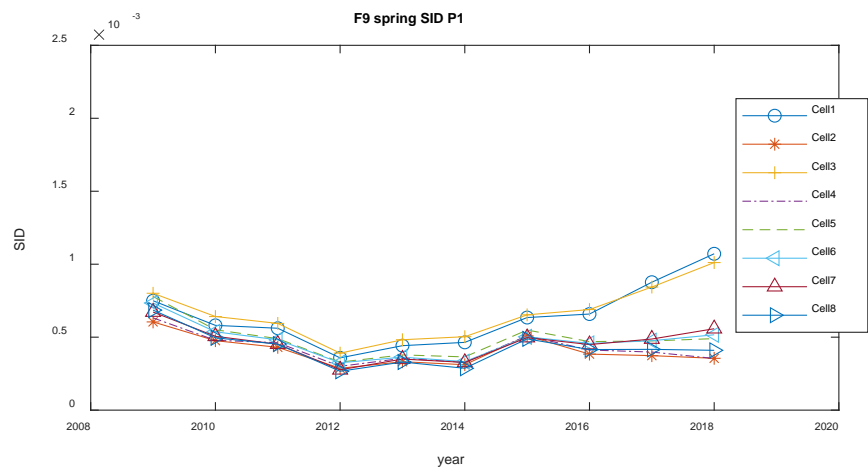
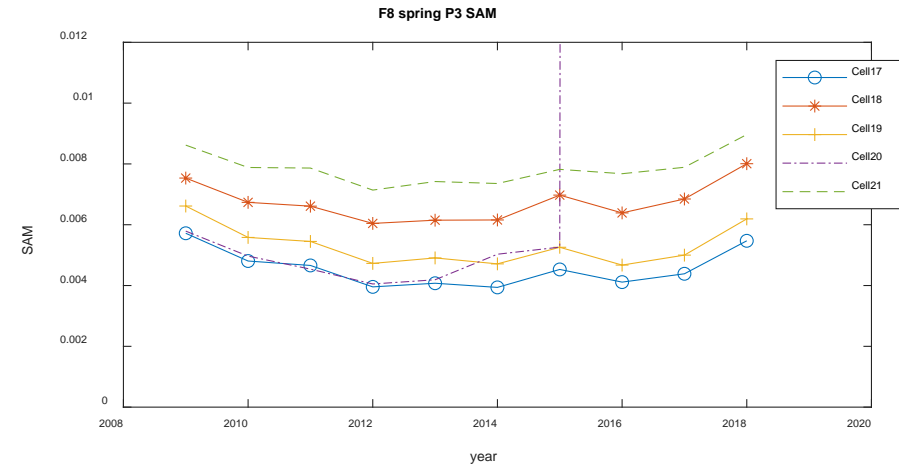
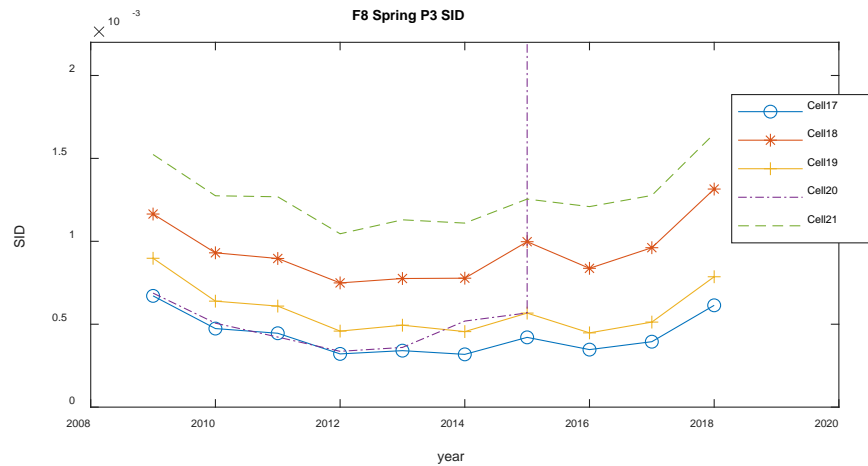


Estimated Cell Voltage Data

# SDM Battery Cell Performance Analysis



- SDM applied to quantify current performance of battery cell. The cell with abnormal discriminability when compared to the rest of the cells in same pack indicates the anomalous event. These anomalies could be the failure of a scanner, a chemical substance leak or other unknown events.





# Conclusion



- For the sensor that collects telemetry data in real time, the best tool to evaluate data is an unsupervised algorithm.
- In the event that telemetry data is not collected, sample statistics are generally available. In that case, more powerful machine learning algorithms can be applied to classified anomaly data in each subsystem.
- Noise is a critical issue in telemetry data processing. In the future, more powerful tools such as a Kalman filter could be introduced to eliminate noise for each parameter.
- Reduce the storage requirement of historical telemetry data.
- Continue to develop algorithms for analyzing each subsystem to improve the efficiency in monitoring spacecraft quality.