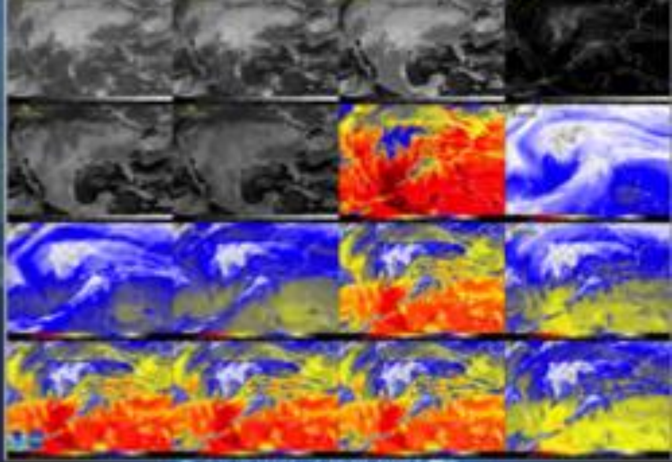


Monitoring GOES-R ABI Radiometric Performance with a Machine Learning System

Zhenping Li, Ken Mitchell, J.P Douglas, Biruh Tesfaye, Dave Pogorzala
Arctic Slope Technical Services (ASTS) and IAI.

ASRC FEDERAL

Agenda

- Why Machine Learning solutions
- Machine Learning (ML) Approach to the systems with large number of detectors or sensors.
- Application in the GOES-R ABI radiometric performance monitoring
- Summary

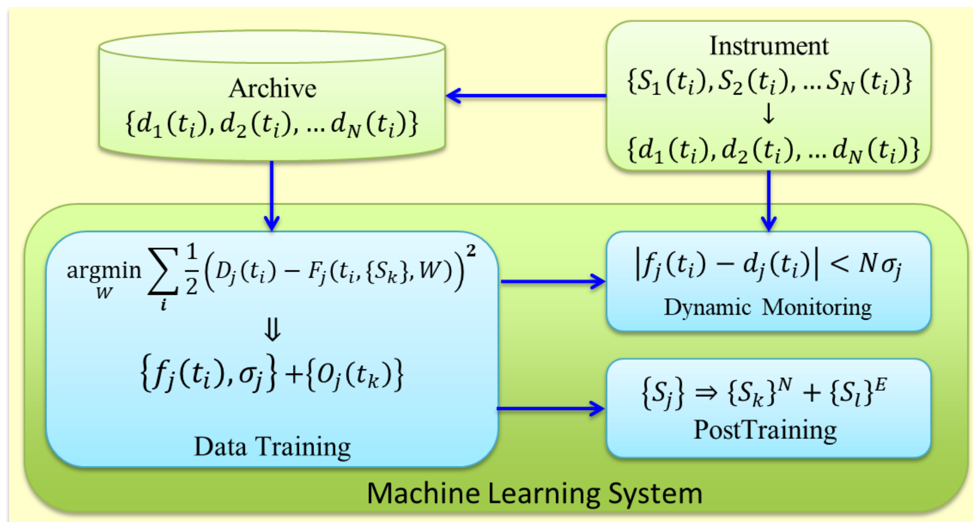
Why Machine Learning (ML) in Radiometric Operations

- Instrument Radiometric Operation monitors the instrument calibration process,
  - Monitoring the radiometric performance in ground system
  - Assess the detector/sensor data quality.
  - Anomaly detection, and provide support in troubleshooting.
- Very Large number of datasets to be monitored
  - GOES-R ABI has over 7000 active detectors comparing to total 16 detectors in GOES N-P Series.
    - The number of datasets to be monitored is over 20k.
  - The OPS Concept for GOES N-P Radiometric Operation no-longer works
    - Mostly manual process.
- Machine Learning approach automated the radiometric operations
  - Provides actionable information to engineers for paying close attention to specific datasets.
  - Enables very quick turnaround in troubleshooting in case of anomalies.
  - Provides the quality assessment of sensor/detector.

- An instrument with large set of sensors/detectors is a dynamic system:
  - Time dependent: the state variables $\{s_j(t_i)\}$ is a function of time
  - Non-deterministic: datasets $\{d_j(t_i)\}$ are measurements of their state variable $\{s_j(t_i)\}$ with the Gaussian probability distributions.
- Data Training generates time dependent trends for datasets $\{d_j(t_i)\}$:

$$\{f_j(t_i, \{s_k\}), \sigma_j\}$$

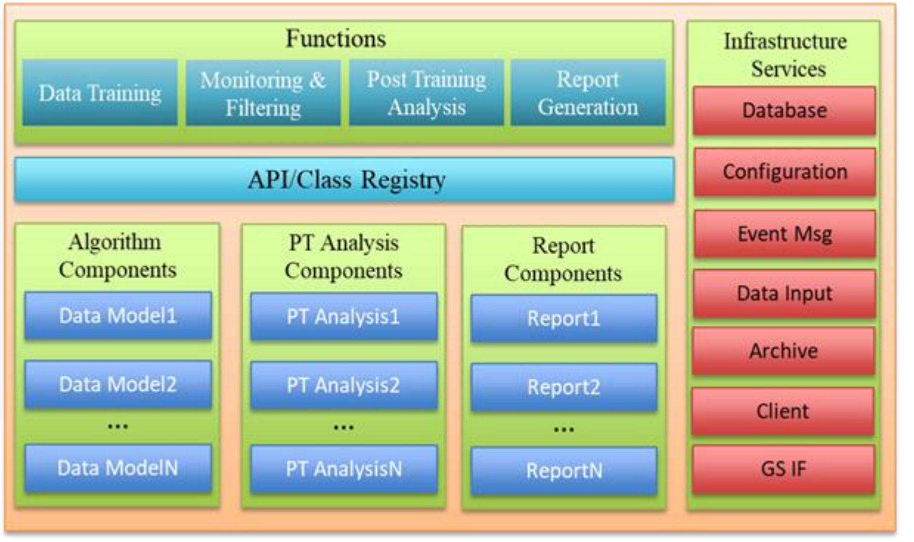
- Dynamic monitoring compares the values of an incoming dataset with its predictions.
- Post training analysis (clustering) separates the normal datasets from anomalous or low quality datasets based on the data training outputs.
  - Developing data quality metrics
  - Generally implements clustering techniques



Advanced Intelligent Monitoring System (AIMS)

- The ML solution for ABI is developed on a scalable and extensible ML platform
  - Can be Easily extended to radiometric process in other satellite remote sensing instruments
- Scalable and extensible Enterprise Architecture
- Separate the common services and infrastructure from the mission specific components
- ML algorithms for data training and post training are treated as plugin and play components with standard API.
- Provides flexibility to select algorithms for datasets with specific patterns

- Other ML Solutions on AIMS:
  - Satellite health and safety telemetry monitoring for LEO Satellite (NPP/JPSS) and Geosynchronous Satellite (GOES)



Data Training Operation Concepts

- Time-dependent trending in operational environments is performed periodically in sessions to ensure that it captures both short-term data patterns and long-term changes.
- The trending period $T_0 = t_f - t_i$ must be long enough to capture the data patterns.
- Two neighboring trending sessions overlap to ensure the continuity and the stability of the trending outcomes.
- The output of the trending session *M* is used as the input of the trending session *M* +1 to improve trending efficiency.

Data Training for Instrument Radiometric Performance Monitoring

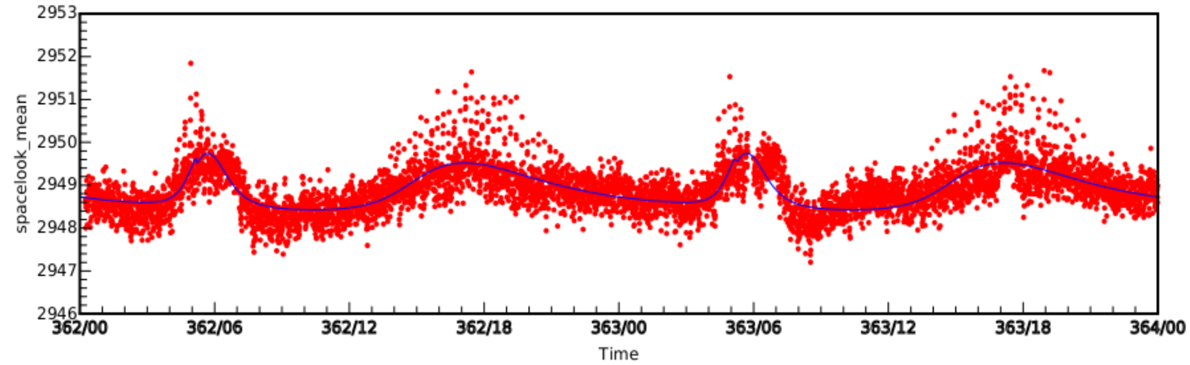
- Treat the datasets in radiometric calibration process as time series data
  - The spacelook value for each detector,
  - ICT(black body) values for detectors in IR channel.
  - The gain parameters for each channel
- A large number of datasets to be trained
  - Manage both GOES-16 and GOES-17 ABI instruments
  - There are more 20k datasets for each ABI instrument to be trained
    - More than 40k datasets for two ABI instruments
- Multi-Data Model Implementation
  - Neural Networks, Fourier Expansion Models
  - Ensure accuracy in training outcomes, efficiency in training algorithms and robustness in handling defective data.
  - GOES-16 and GOES-17 have different algorithms due to different complexities in data patterns.

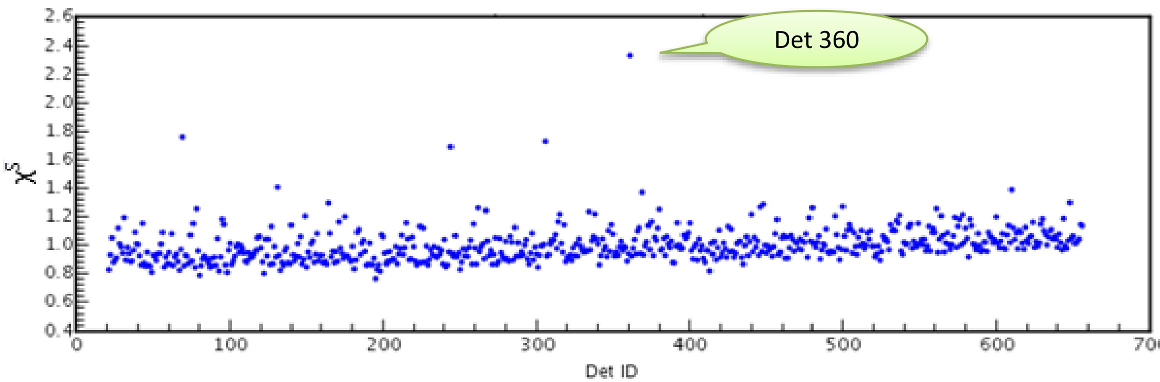
Post Training Analysis

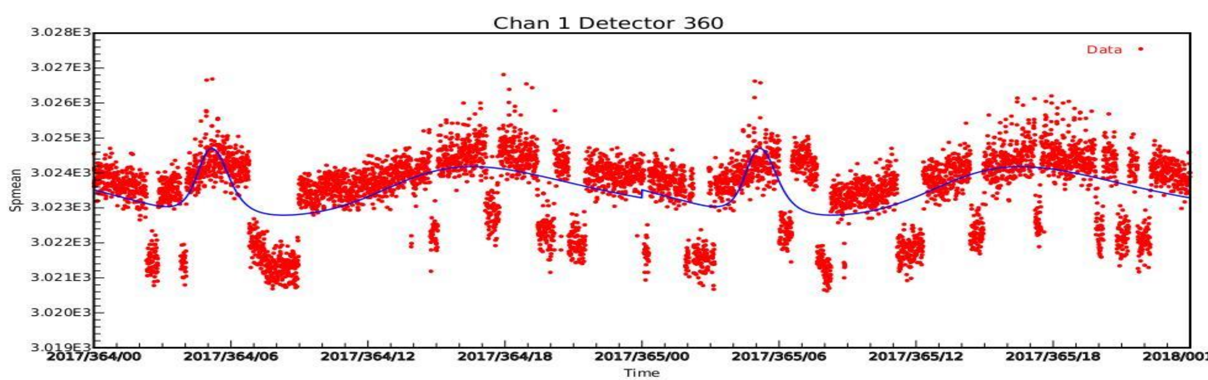
- Obtains actionable information from large number of data training outputs
  - The data training output for each dataset consists of time dependent trends and list of outliers.
  - Actionable information includes detector quality and anomaly characterization (if an anomaly exist)
- The clustering technique is implemented: $\{S_j\} = \{S_j\}^N + \{S_j\}^E$
  - Majority of detectors in each channel are normal, and have common characteristics → form their own clusters.
  - Few anomalous ones are treated as noises→don't form clusters.
- Develop data quality metrics for clustering purpose
  - Outlier clusters: $\chi_j^O = \sum_i \left(\frac{\delta_i^W}{T}\right) + \frac{N^E}{N^W} \sum_i \left(\frac{\delta_i^E}{T}\right)$
  - The temporal change measure the change in $\{\sigma_j\}$ in the consecutive training sessions: $\chi_j^T = \frac{\sigma_j^M}{\sigma_J^{M-1}}$
  - The spatial change measures the relative data quality for a dataset group: $\chi_j^S = \frac{\sigma_j}{\frac{1}{N}\sum_{k\in\{j\}} \sigma_K}$
- An anomaly for a dataset in the machine learning is defined as the "**unexpected data pattern changes**"

Variables in the Visible Channels

- Spacelook datasets use the neural network (4 nodes in the 1st hidden layer, 2 nodes in the 2nd).
- The spatial change plot generated in post training analysis process has $\chi_j^S \approx 2.4$ for the detector 360.
- The spacelook data points for detector 360 during the trending period show much larger fluctuations.
  - Larger noise level in its time dependence
- There are few detectors in both GOES-16 and 17 with high level of noises in both visible and IR channels

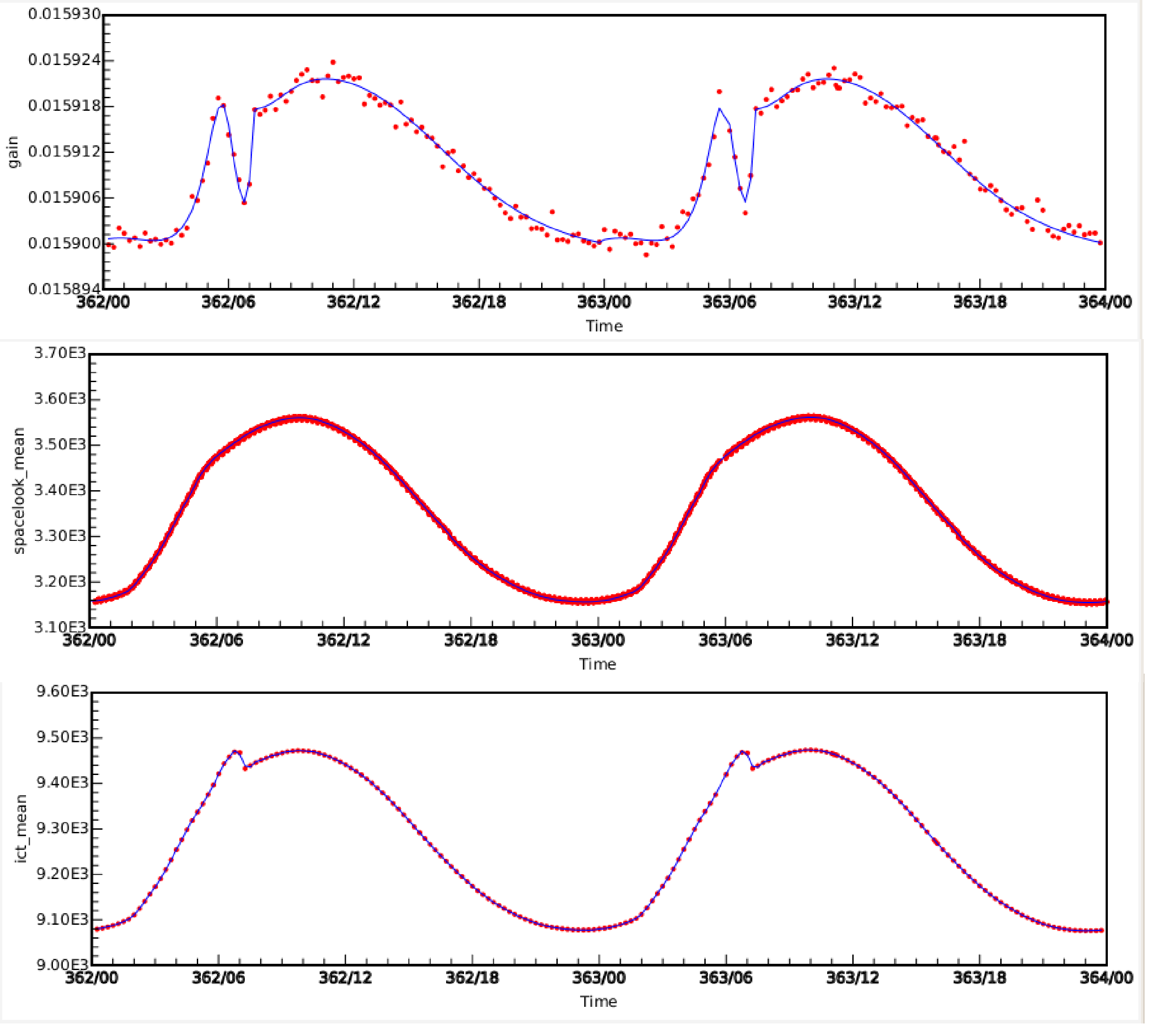






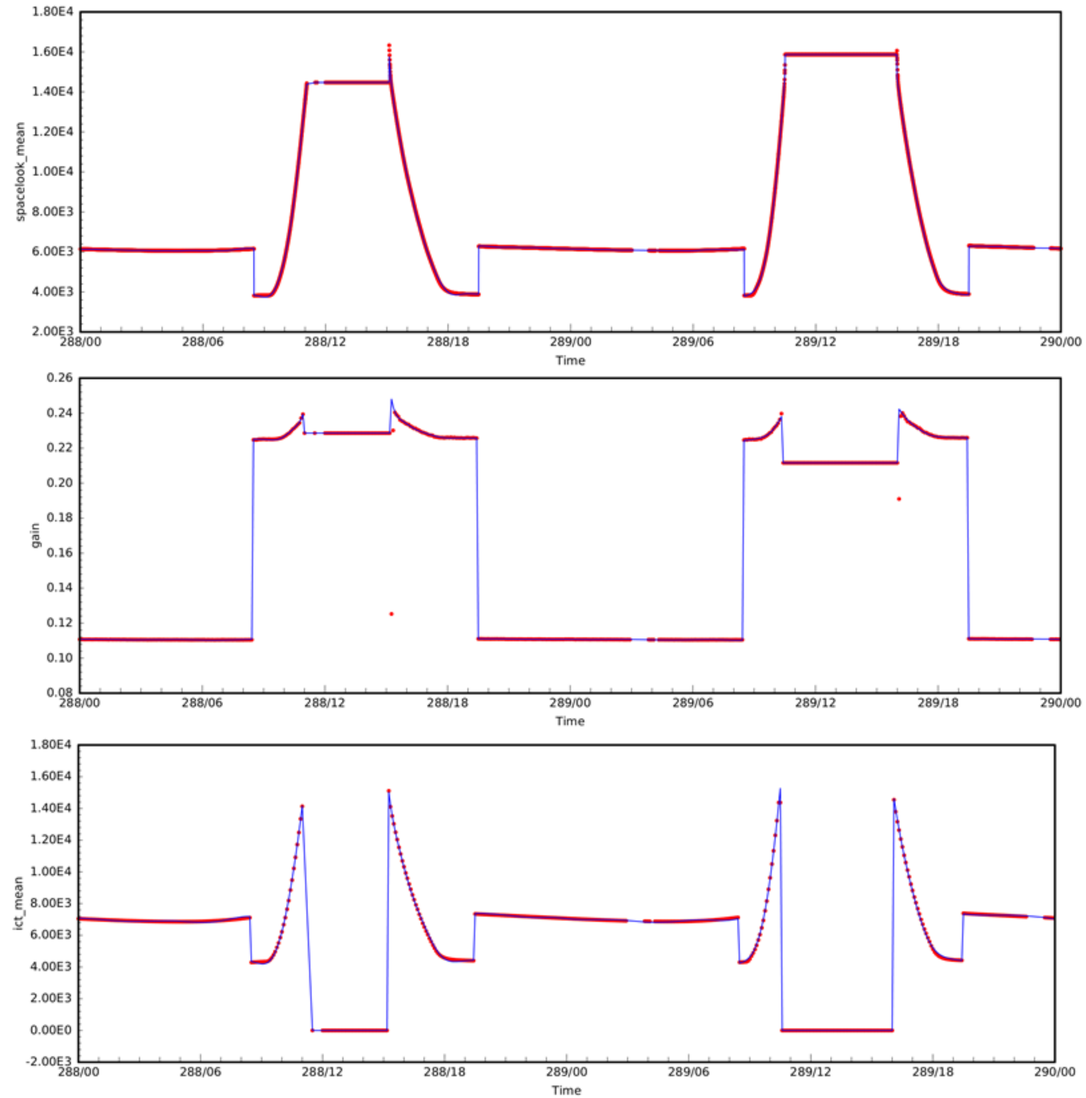
Training Outputs in the Infrared Channels for GOES-16

- Gain data uses the neural network (6 nodes in the 1st hidden layer, 3 in the 2nd). Using the Fourier Expansion Model may lead to overfitting by including noise as part of the regular data pattern.
- Both spacelook_mean and ict_mean variables use the Extended Fourier Expansion Model to account for daily scale changes. This model has explicit linear dependence on time in its expansion amplitudes.



Training Outputs for GOES-17 Infrared Channels

- GOES-17 data in infrared channels have double settings and latching.
  - Makes data training more challenging
- Stitching algorithm has been developed to 'repair' the data before the actual data training
  - Make data continuous
- The latched data points are removed from data training sets.
  - Present in channel 8 to 12.

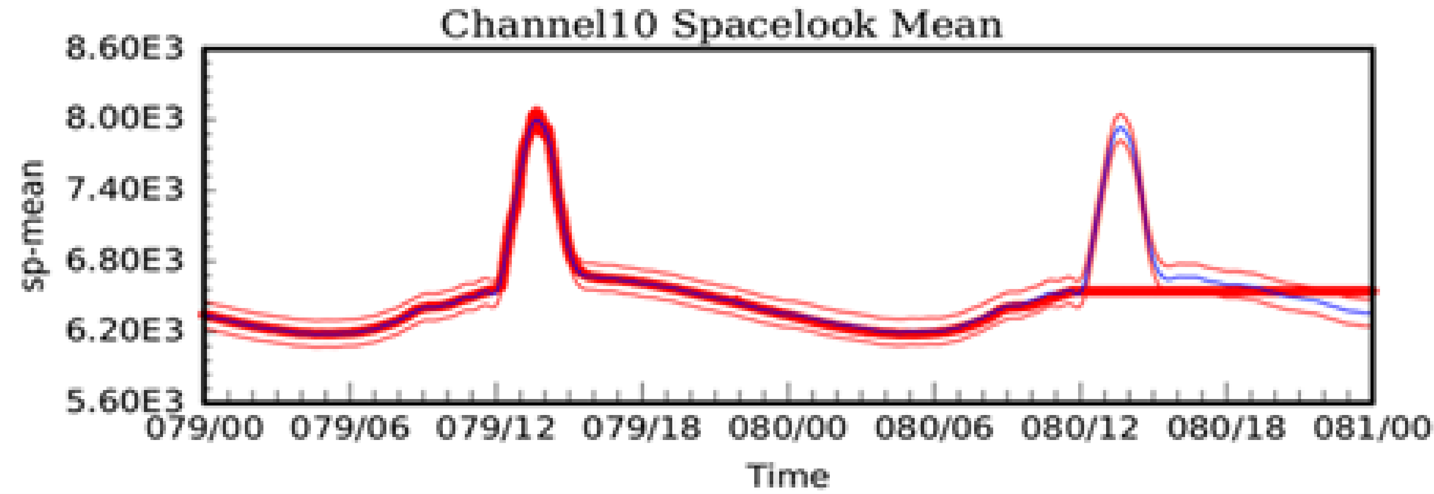


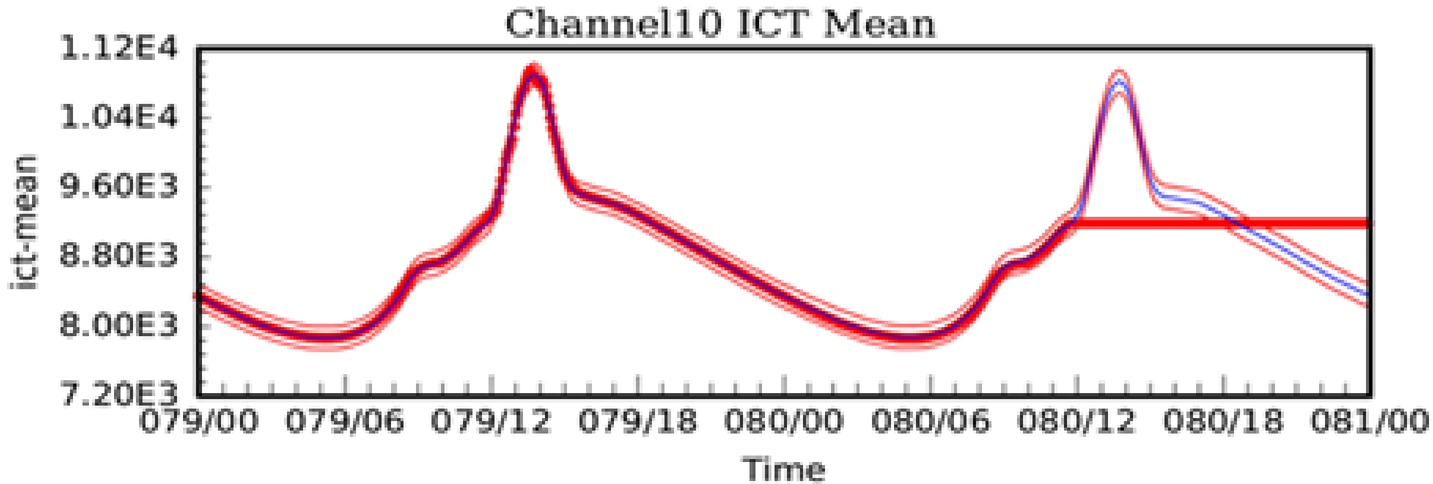
GOES-17 Channel 12 detector 26 training outputs for spacelook (top), gain(middle) and ICT mean(bottom)

An Unexpected Pattern Change Example

Both spacelook and ICT values became constant around 2019/080/10Z for detector 231 in GOES-17 channel 10, which leads to striping in channel 10 images in the same time period.

The constant data lead to large deviations from the predictions by machine learning data models.

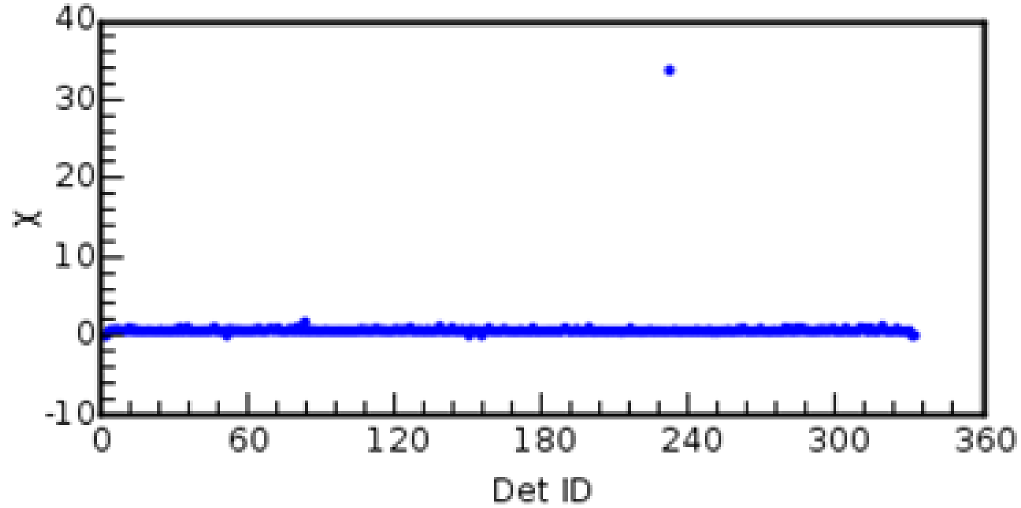


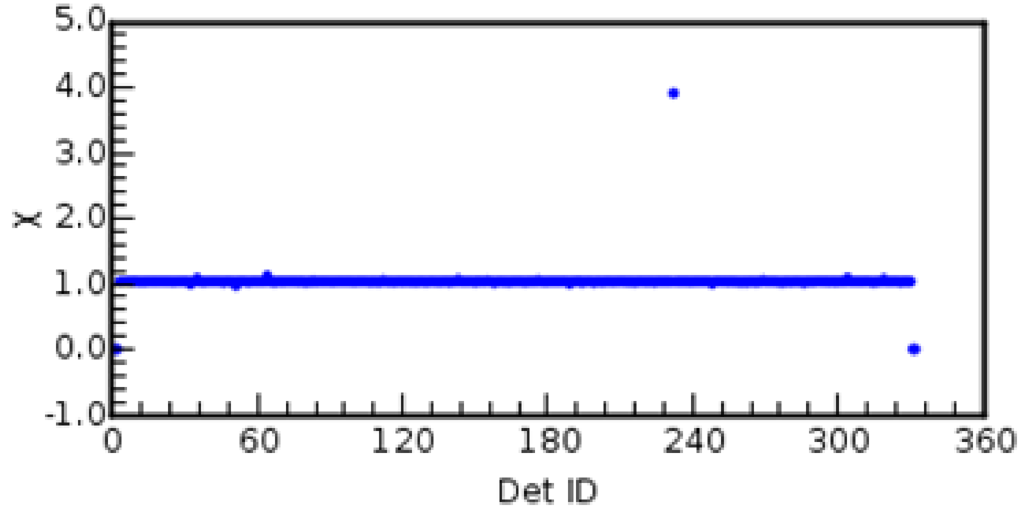


The pattern change cannot be detected with the existing approach using the static red/yellow limits.

The Output from the Post Training Analysis

- The pattern changes for a dataset leads to two distinct patterns in a training period.
- The temporal change metric become very large due to the larger standard deviations in training output.
- The spatial change metric also become very large.





The temporal changes for gain and ICT data show elevated values for detector 231.

Operation Status Report

Status Report for Session 2018003

DataSet Name	Status
g16-channel1	NORMAL
g16-channel2-1	TRAINING ERROR
g16-channel2-2	TRAINING ERROR
g16-channel2-3	TRAINING ERROR
g16-channel2-4	TRAINING ERROR
g16-channel2-5	TRAINING ERROR
g16-channel3	NORMAL
g16-channel4	NORMAL
g16-channel5	WARNING
g16-channel6	WARNING
g16-channel7	NORMAL
g16-channel8	NORMAL
g16-channel9	NORMAL
g16-channel10	NORMAL
g16-channel11	NORMAL
g16-channel12	ERROR
g16-channel13	NORMAL
g16-channel14	NORMAL
g16-channel15	NORMAL
g16-channel16	WARNING
g16-temp	WARNING

AIMS Data Training Output

The DataSet g16-channel12 Status for Session 2018002

Summary

The following elements have elevated χ^S, χ^T, or χ^O values, which require further investigation.

ID	χ^S	χ^T	χ^O
290	10.404085	0.6413882	0.0
291	22.653366	0.6001036	0.0

g16-channel11: AIMS Day to Day Summary Report

Date	σ Mean Value				Above Threashold Detectors	
	spacelook_mean	spacelook_stddev	gain	ict_mean	Warning	Error
2019/151	1.05	0.07580	0.000001815	0.5791		
2019/150	0.9641	0.07575	0.000002081	0.2751		
2019/149	0.9355	0.07537	0.000001561	0.1549	99	
2019/148	0.9472	0.07526	0.000001621	0.3040		
2019/147	0.9264	0.07516	0.000001712	0.1420	99	
2019/146	0.9355	0.07519	0.000002148	0.1614	99	

- Can be customized to meet the specific requirements from Engineers
- The left column are hyper links (left graph) that enable users to navigate into detailed status of a specific channels (right).
  - The detector ID column provides the hyperlinks for the plots of a specific detectors.
- Engineers only need to look at the status report for a given sessions.
- Provides data plots of data and ML predictions

Summary

- The Machine Learning approach automates engineering analysis and provides actionable information for engineers.
  - The root cause of an anomaly could be determined in minutes.
    – Significantly improves the system resilience.
  - The same approach can be extended to other remote sensing instruments.
- The short term trending (data training in two day period) provides information on detector quality and anomaly detections.
  - Long term trending (data training in one year or longer period) provides information on seasonal behavior and sensor degradation,
    – still needs to be investigated.
- AIMS enterprise architecture enables easy integration of customized algorithms for GOES-17 data.