Creating Situational Awareness in Spacecraft Operations with the Machine Learning Approach

Zhenping Li[1]

ASRC Technical Services, 7515 Mission Drive, Lanham, MD. 20706

Abstract

This paper presents a machine learning approach for the situational awareness capability in spacecraft operations. There are two types of time dependent data patterns for spacecraft datasets: the absolute time pattern (ATP) and the relative time pattern (RTP). The machine learning captures the data patterns of the satellite datasets through the data training during the normal operations, which is represented by its time dependent trend. The data monitoring compares the values of the incoming data with the predictions of machine learning algorithm, which can detect any meaningful changes to a dataset above the noise level. If the difference between the value of incoming telemetry and the machine learning prediction are larger than the threshold defined by the standard deviation of datasets, it could indicate the potential anomaly that may need special attention. The application of the machine-learning approach to the Advanced Himawari Imager (AHI) on Japanese Himawari spacecraft series is presented, which has the same configuration as the Advanced Baseline Imager (ABI) on Geostationary Environment Operational Satellite (GOES) R series. The time dependent trends generated by the data-training algorithm are in excellent agreement with the datasets. The standard deviation in the time dependent trend provides a metric for measuring the data quality, which is particularly useful in evaluating the detector quality for both AHI and ABI with multiple detectors in each channel. The machine-learning approach creates the situational awareness capability, and enables engineers to handle the huge data volume that would have been impossible with the existing approach, and it leads to significant advances to more dynamic, proactive, and autonomous spacecraft operations.

1. Introduction

The situational awareness (SA) for a system or a mission refers the ability to perceive, comprehend, and predict its own states. SA enables real time detection and characterization of potential anomalies or unexpected behavior, and provides the actionable information of the system to the intelligent decision making tool or the management, which enables real time response essential to improve the system resiliency. For a satellite mission, the SA enables the predictions on the optimal operational states, such as its health and safety and the instrument performance, and monitors the incoming telemetry from spacecraft so that the potential anomalies can be detected and characterized in real time. The focus of this paper is to present the machine learning as a natural and systematic approach in enabling SA capabilities for spacecraft operations. A machine learning system represents a data model or an algorithm that can be trained and make predictions on data. The spacecraft datasets are generally time dependent, which can be characterized by the time dependent patterns. The machine learning captures the time dependent patterns of spacecraft datasets through data training, which are used to predict what to expect next. The incoming data are compared with the predicted value by a machine learning system, and if the difference is larger than a threshold level characterized by the noise in its dataset, it could indicate a potential anomaly.

The machine learning approach brings fundamental changes to spacecraft operations. The current spacecraft operations for maintaining the satellite health, safety, and the instrument performance and accuracy involve data trending, data monitoring, and engineering analysis processes. The data trending is a process to determine a true measure of a dataset that is statistically distinguished from random behavior. The statistical approach has been the standard to the data trending, which the trend of a dataset is characterized by its statistical properties, such as the mean and standard deviation for a given period. The data monitoring defines static threshold minimum and maximum limits for spacecraft telemetry values. Satellite operators typically determine spacecraft state-of-health by identifying telemetry excursions outside the static limits. To avoid high rates of "false alarms," telemetry limits are set wide enough to bound variations normal for spacecraft operations in space—that is, regular variations caused by environmental factors related to solar, ecliptic, charged particle, and magnetic forces, as well as predictable variations due to spacecraft operations driving power, momentum, and thermal influences. The static limits are much less sensitive in detecting operationally meaningful changes in telemetry points. For example, small changes in a temperature telemetry measured indicative of a progressing failure may be operationally undetected if those

[1] Zhenping.Li@asrcfederal.com

changes remain within generous limits established to bound the full range of temperature fluctuations driven by changing attitude orientation, solar illumination, and solar ecliptic angles over the life of the spacecraft. The engineering analysis involves the targeted review of the specific dataset to identify, characterize, comprehend, and workaround an anomaly or a failure, which is generally a tedious manual process. Thus, the existing approach for spacecraft operation is generally static, reactive, and manual processes. The machine learning approach provides an automated and integrated approach to the data trending, monitoring, and engineering analysis. Instead of the traditional statistical data trending, the machine learning approach performs the time dependent trending through the data training. The time dependent trend creates a dynamic limit to be used in the data monitoring, which has much tighter data bound than the traditional static limit. The dynamic limit defined by time dependent trends is highly sensitive to the data changes above the noise level, thus enables the potential anomalies to be identified and characterized in real or near real time. This leads to a more dynamic, proactive and autonomous spacecraft operations.

This paper is organized as the following. The section 2 provides a general framework of the machine learning approach to the spacecraft data. The challenges for developing a machine learning approach are to define a general representation for spacecraft datasets with arbitrary complexity and scales and the data training strategy that is systematic, accurate, efficient, and adaptive. Section 3 provides the operational concept of how machine-learning approach is used in spacecraft operations, which brings the fundamental changes to the spacecraft operations. Section 4 presents the examples of the machine learning approach in the instrument calibration processing for AHI[1]. Both AHI and ABI[2] on GOES-R series represents the next generation of Imager with the order of magnitude larger data volume than that for the Imager on the current GOES satellite, and the existing approach in data trending and monitoring can no-longer meet the challenges presented by AHI and ABI instruments. Our results shows that the machine learning approach provides not only the time dependent trending and dynamic data monitoring, but also a metric for evaluating the data quality. Finally, the summary and discussions of future works are given in Section 5.

2. The Machine Learning Approach

The data patterns for spacecraft telemetry data are characterized by their time dependence, and represent the optimal states in spacecraft operations. Generally, there are two types of the time dependent patterns for spacecraft datasets:

1. The absolute time pattern (ATP). The ATP datasets generally depend only on time, such as the temperature or power profiles for a spacecraft, whose time dependence is mainly influenced by the relative position or angle between the spacecraft and sun. For Earth orbit spacecraft, ATP datasets generally have diurnal behaviors. The ATP datasets can be characterized by time dependent functions $f(t)$.
2. The relative time pattern (RTP). The RTP datasets $\{d^r(t_i - t_c)\}$ are generally triggered by an event at a specific time t_c. An example of the event trigger is a command being sent to a spacecraft that leads to the change of its states, such as the orbit maneuver or data collections of its payload. The RTP datasets are characterized by the time dependent function $g(t - t_c)$, where t_c is the reference time triggered by an event on spacecraft and $t \geq t_c$.

The data training in a machine learning system is to find a time dependent function $f(t)$ or $g(t - t_c)$ so that the error function

$$e = \begin{cases} \frac{1}{2}\sum_i \left(d^a(t_i) - f(t_i)\right)^2 & \text{for APT datasets} \\ \frac{1}{2}\sum_i \left(d^r(t_i - t_c) - g(t_i - t_c)\right)^2 & \text{for RTP datasets} \end{cases} \quad [1]$$

is at minimum. The data training strategy generally involves two phases: the pattern discovery phase and the data-retraining phase. When a machine learning system is deployed into an operational environment, the system has no prior knowledge of the data patterns. The data training in this phase is a search process to capture the patterns of datasets with sufficient accuracy. The data retraining happens during the operational phase to capture the seasonal or long-term changes to the patterns in datasets. The machine learning system in this phase has the prior knowledge of dataset patterns, which could be used as the input for data retraining. The datasets with time dependent patterns are characterized by its time dependent trend $\{f(t), \sigma\}$ or $\{g(t - t_c), \sigma\}$, where σ represents the standard deviations of a dataset with respect to the time dependent function $f(t)$ or $g(t - t_c)$ for a given data training period:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(d^a(t_i) - f(t_i)\right)^2}. \quad [2]$$

The σ value also represents an important quality metric for its dataset. The value of a data point at the time t_i should be within the range

$$|f(t_i) - d^a(t_i)| < N\sigma_e \quad [3]$$

where the factor N is an integer defined by users. Eq. 3 could also be used to evaluate the incoming data during the real time to determine if a data point is within the bound defined by its noise level.

The challenges for a machine learning approach to the SA capability are to establish a representation that captures the complexity of data patterns with arbitrary scale, and to develop a training strategy and algorithm that is systematic, efficient, accurate, and adaptive. Two machine-learning algorithms have been developed for ATP datasets; Adaptive Trending and Limit Monitoring Algorithms (ATLMA)[3] and the neural network implementation[4]. ATLMA expresses a time dependent function $f(t)$ as the Fourier series, and the linear least square-fitting algorithm is implemented as the data-training algorithm for both phases. ATLMA is simple and efficient, and the data training with the linear square fitting always finds the minimum for the error functions defined in Eq. 1. However, ATLMA is limited to the continuous datasets dominated by the low frequency components in Fourier expansions. The neural network approach implements multi-layer feed-forward and back-propagation networks for time dependent functions, which is proven to be far more versatile that can be used for both continuous and non-continuous datasets. Fig.1 shows that the general network structure as the representation for time dependent functions, which is a two hidden layer network with single input and output node. The number of nodes at each hidden layer depends on the complexity of data patterns, which is directly related to the number of minimum or maximum of a datasets in a given trending period. Both input and output nodes represent the input time t and the output function $F(t)$. The output for the j-th node in the l-th layer is related to the output for i-node in the l-1-th layer by

$$O_j^l = F\left(\sum_i W_{i,j}^l O_i^{l-1}\right), \quad [4]$$

where $\{W_{i,j}^l\}$ is the weight parameter for l-th layer, j-th node, and the *tanh* function is used as the activation function for neurons for the data trending and monitoring;

$$F(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \quad [5]$$

The output function $F(t)$ in Fig. 1 has the range between -1 to +1. Thus, an additional preprocessing routine is required to convert the ATP function $f(t)$ with an arbitrary scale into the output function $F(t)$. The output function $F(t)$ also needs to be converted into the ATP function $f(t)$ during the post-processing phase so that the standard deviation σ in Eq. 2 can be calculated.

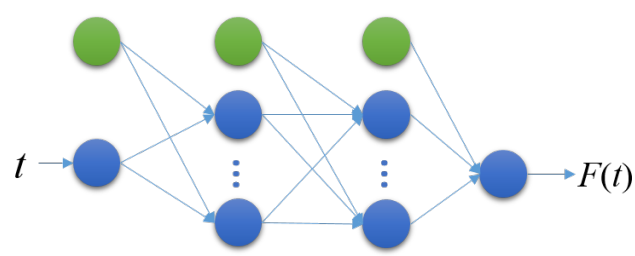

Figure 1 Neural network for time dependent patterns of the spacecraft datasets. The blue node represents a network node, and the green node corresponds to the bias node.

The neural networks are non-linear functions, which is why the networks capture complex data patterns that are either continuous or discontinuous. In the same time, the non-linear property of the neural networks also increases the complexity for data training during the initial pattern discovery phase. The standard data-training algorithm for neural networks is the gradient decent for the networks in Fig. 1. Our investigation shows that the gradient decent alone does not generate the sufficient accuracy in its training outcome. More sophisticated data training algorithm is the non-linear least square fitting algorithm, such as the Levenburg-Marquardt (LM) back-propagation algorithm[5]. The LM algorithm or similar non-linear least square-fitting algorithms depends on the initial inputs, and the wrong inputs may lead to the divergence of the least square fitting. Thus, the data-training algorithm for the networks shown in Fig. 1 in the initial pattern discovery phase involves the combination of the gradient decent and LM back-propagation. The gradient decent generates the input weight parameters for LM algorithm that is close to the minimum value, and the inputs for the initial gradient decent are the random number. LM least square fitting algorithm provides the refined solutions with sufficient accuracy. The data

retraining during the normal operations performs the minor adjustment to the network so that it is adaptive to the gradual pattern changes in the datasets. The efficiency of the data training algorithms becomes very important during the operation phase. The gradient decent[4] algorithm for the forward feeding and back-propagation network has been implemented during the retraining phase.

3. Operational Concept

Fig. 2 shows how a machine-learning system works in a ground system for datasets with ATP. The data model, data training and monitoring represent a machine learning system for spacecraft operations. The data training and monitoring in the machine learning approach are autonomous. The data model in the machine learning approach is represented by the time dependent trend of its dataset. The data model is trained periodically with the latest datasets in the data archive. During the normal operations, the overall data pattern changes are not expected to be large from one day to the next. Thus, the data retraining process uses the existing data model as the input, and the output of the data training updates the data model so that it contains the latest data patterns to ensure the accuracy of its predictions. The data monitor receives the real time data, and compare the value of the incoming data points with the predicted value from the data model. In case of the potential anomalies, the monitoring module generates warning messages to alert engineers for further troubleshooting or to an intelligent decision making system for an automated response. The real time data points are also ingested into the storage in a ground system, which are used as the input for the data training.

For the datasets with RTP, the data monitoring for a specific dataset is triggered by an event, such as a command or directive to spacecraft. The data model in this case is associated with a specific event. The monitoring module also monitors the triggering event to retrieve the time t_c and triggers the data monitoring for the RTP dataset. The triggering events could come from the spacecraft telemetry, command schedules generated by the scheduling system, or the event messages from the command and control system for the commands or directives being send to the spacecraft. Thus, the real time datasets shown in Fig. 2 may includes telemetry from the spacecraft, the intermediate products in the payload data processing system, the command schedule from the scheduling system, and the event messages from the command and control system. A service-oriented ground system architecture is crucial for enabling SA capabilities in a mission so that a standard interface and service protocols can be established for a machine learning system to receive the data from different components or subsystems in a ground system.

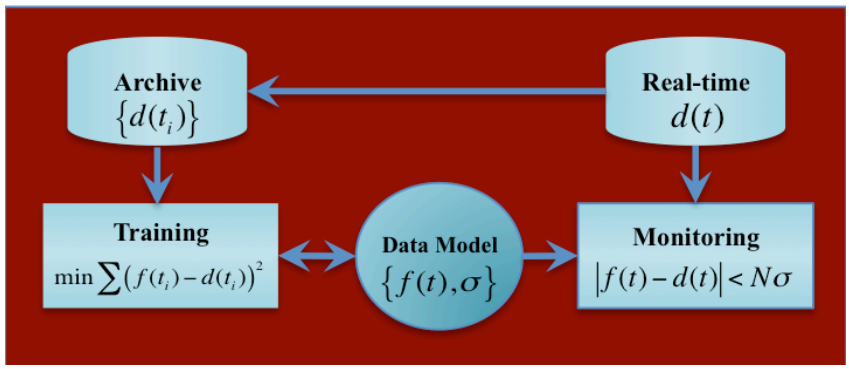

Figure 2 The operation concept of machine learning approach in spacecraft operations

The machine learning approach brings the fundamental changes to spacecraft operations, which are highlighted in Table 1. The statistical trending in the current approach becomes the data training for time dependent trend in the machine learning approach, which generates the dynamic data bound for data monitoring. The data monitoring becomes more dynamic comparing to the static limits in the traditional data monitoring, which is much more sensitive to the changes in a dataset above the noise level. The potential anomalies are automatically identified and characterized in the real or near real time. This leads to a more dynamic, proactive, and autonomous spacecraft operations.

Table 1 The current and machine learning approaches in spacecraft operations

Operations	Current Approach	Machine Learning
Data Trending	Statistical Trending to derive the statistical properties	Data Training to obtain the time dependent trend
Data Monitoring	Static monitoring (red/ yellow limit)	Dynamic, compare data with predicted value
Engineering Analysis	Manual	Automated.

4. Application to AHI Instrument Data Calibration Process

The initial studies on the machine learning algorithms have been concentrated on the calibration data on Geostationary Environment Operational Satellite (GOES) Imager[4] and Advanced Himawari Imager (AHI), which are used to demonstrate how the machine learning approach works with satellite datasets. The same approach can be extended easily to other processes such as the spacecraft health and safety. AHI like ABI is a latest generation of multi-channel Imager on Japanese Geostationary Himawari Satellite, which is designed to sense the radiant and solar reflected energy from sampled areas of Earth. AHI has the same configuration as the Advanced Baseline Imager (ABI) on the next generation GOES-R series spacecraft to be launched later this year, and the machine learning approach will be deployed in GOES-R ground system to monitor the ABI calibration process. The input data for the machine learning system are the intermediate radiometric variables in the instrument calibration processes, which is crucial to ensure the data quality and the radiometric accuracy. The data presented here are generated based on AHI L1A calibration statistics with the time range from 2015/003 to 2015/005, which is in the AHI in-orbit test period. The challenges for both AHI and ABI are that the amount of data to be monitored is order of magnitude larger; the number of detectors for each channel ranges from several hundreds to more than one thousand for both AHI and ABI comparing to just 2 detectors for infrared channels and 8 detectors for the visible channel for the current GOES Imager. The existing approach with manual trending and monitoring is no longer possible. The machine-learning approach provides an integrated and automated trending and monitoring operations, and potential anomalies can be detected and characterized in real time or near real time.

Fig. 3 shows an example of the data training output by ATLMA for the offset parameter of one of the detectors in the 6.9 μm channel. The offset parameter is one of the parameters used in converting the raw detector counts into the physical radiance used by scientific users, and the quality of this parameter can directly affect the radiometric accuracy. The training datasets generally cover two days in order to avoid the potential fluctuations due to the data noises from one day to the next. Fig. 3 shows excellent agreement between the time dependent trend and the actual dataset. The two orange lines in Fig. 3 represent the upper and lower limits determined by the standard deviation in Eq. 2, which form a very tight data bound. The tight data bound in Fig. 3 leads to the dynamic limit highly sensitive to the data changes above the noise level. After the time dependent trend is obtained through the data training, each data point in the training datasets is evaluated using Eq. 3. The data points that do not satisfy Eq. 3 could be potential anomalies, which are defined by two orange lines in Fig. 3. In fact, several data points in Fig. 3 are indeed outside data bound defined by the orange line, which could indicate a potential anomaly. The warnings of these potential anomalies are generated automatically to alert engineers for special attention. Machine-learning approach for data trending and monitoring enables engineers to look only for the datasets that may have potential problems or special interests instead of every datasets.

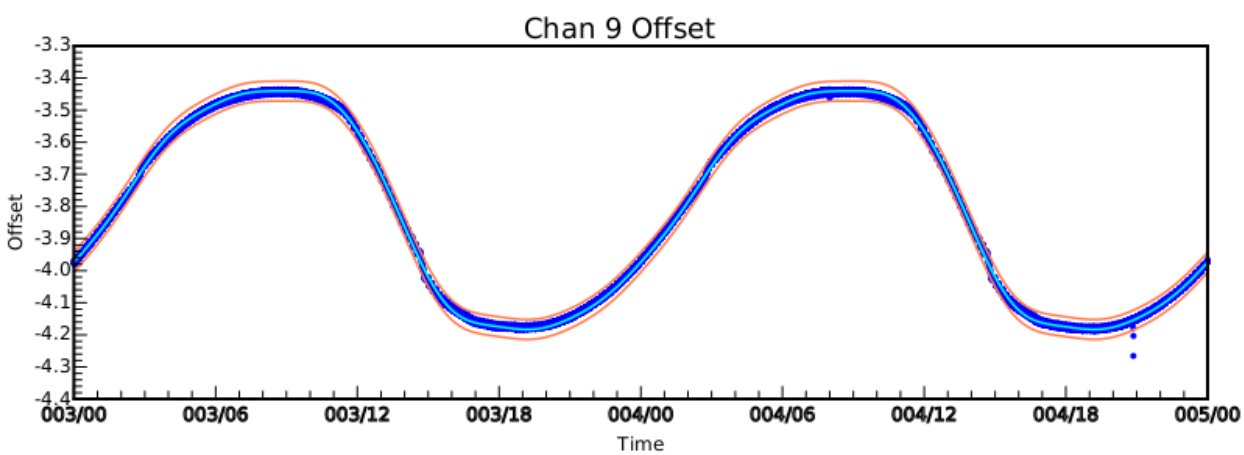

Figure 3 The Offset Parameter for 6.9 μm channel. The blue dots are the actual data points, the light blue line corresponds to the function $f(t)$ obtained from the data training with ATLMA algorithm, and the two orange lines represents the upper and lower limits defined by the standard deviations from the data training. The dataset has the time range from 2015/003/00:00:00 to 2015/005/00:00:00.

ATLMA is only providing good solutions for continuous datasets dominated by the low frequency components in Fourier expansions shown Fig 3. This is not always the case for spacecraft datasets. The neural network

implementation offers considerable advantages over ATLMA since it provides very good solutions for datasets that are both continuous and discontinuous. The study has shown that the networks with two hidden layers can be used to describe an arbitrary functions[7]. Fig. 4 shows the time dependent trend generated by a two-hidden layer neural network for the gain parameters in the 10.4 µm channel. The gain parameter is another one used in converting the raw detector counts into the physical radiance. The time dependent function for the gain parameter is less continuous comparing to that for the offset parameter shown in Fig 3, and there are sudden increases of the magnitude for gain parameters around 03 hour each day. The neural network solution consists of 6 nodes on the first hidden layer and 3 nodes on the second hidden layer, which generates excellent data training outcome with very tight data bound. The same results[4] have also been shown for the temperature parameters in GOES N-P Imager and Sounder dataset, which the discontinuities were shown around the satellite midnight.

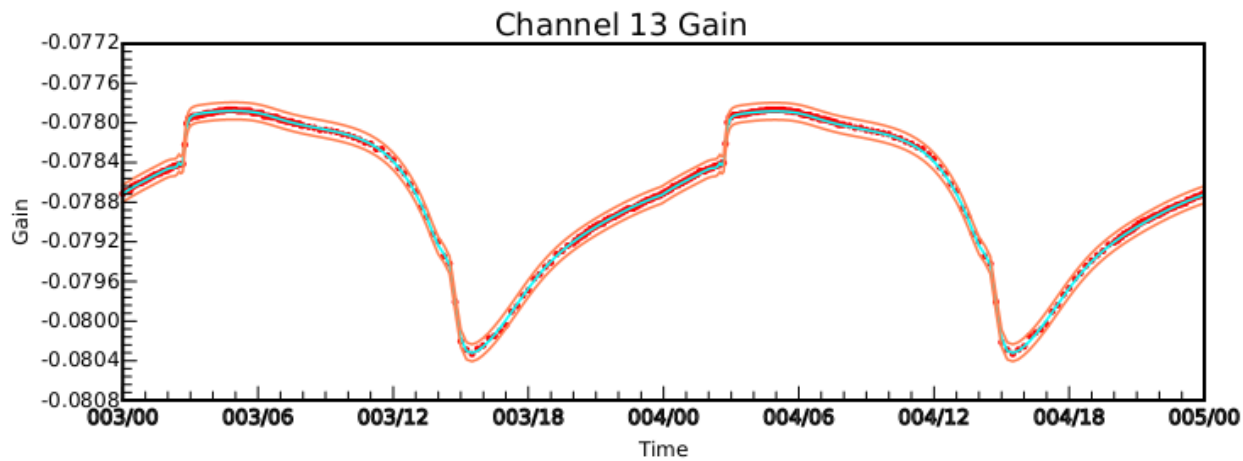

Figure 4 The gain parameter for the 10.4 µm channel. The red dots are the actual data points, and the green line is the time dependent function $f(t)$ using the neural network algorithm. The two orange lines represent the upper and lower data bounds defined in Eq. 3. The dataset has the time range from 2015/003/00:00:00 to 2015/005/00:00:00.

The time dependent trend from the data training in machine learning approach not only enables the dynamic data monitoring, but also provides users insights into the data quality generated by detectors in each channel. This would not be possible without accurate time dependent trends. The standard deviation σ in a time dependent trend characterizes the true noise level of a dataset, and it can be used as the metric for measuring data quality. The higher noise level for a dataset generated by a detector, the lower data quality of the detector is. After the time dependent trend is obtained for a dataset, the plot of the σ value as the function of the detector ID can be generated. Fig. 5 shows an example of the σ plot for the gain parameter in 10.4 µm channel, in which there are 412 detectors. The σ values for most of detectors are around 1.2E-5. There are two detectors with very significant higher σ values; in particular, the σ value for detector 271 is 1.3E-4, which is more than 10 times higher than the average value. The high σ value for a detector might be caused by the poor data training output, however, it may also be due the high noise level generated by its detector. Fig. 6 shows the gain parameter datasets for the detectors 270 and 271. The data for the detector 270 provides an example of low σ

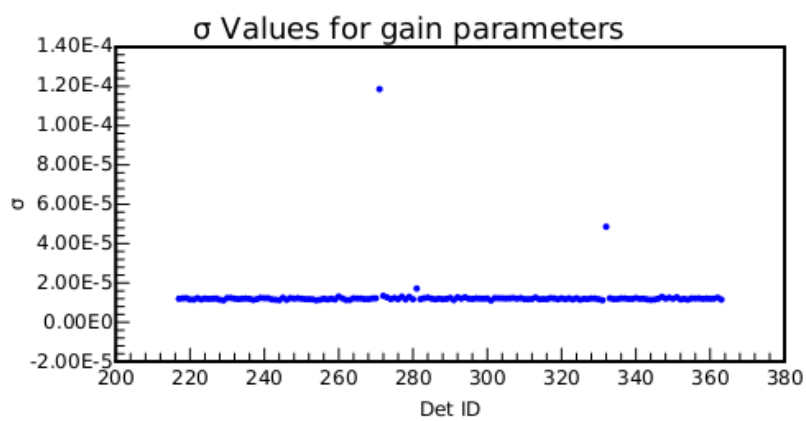

Figure 5 The σ values for the time dependent trend as the function of detector ID. There are 412 detectors in the 10.4 µm channel for AHI. The detector ID in this plot ranges from 216 to 364.

value in its time dependent trend, which is in excellent agreement with its time dependent function. The dataset for the detector 271 is much more noisy characterized by very high σ value. This shows that the σ value in the time dependent trend provides a good metrics for measuring the detector data quality.

For both AHI of Himawari series and ABI for GOES-R series, the number of the detectors for each channel ranges from more than 300 to more than 1000. Both AHI and ABI provide a multiple detectors for a detector row in each channel that enables the selection of the detector with better quality. The standard deviation σ in a time dependent trend provides a practical and systematic way to evaluate the detector quality. After the time dependent function $f(t)$ is obtained from the data training for each detector, its standard deviation is evaluated using Eq. 2. A statistical analysis can be performed for the σ values in each channel. A detector with very high σ value could indicate a low data quality.

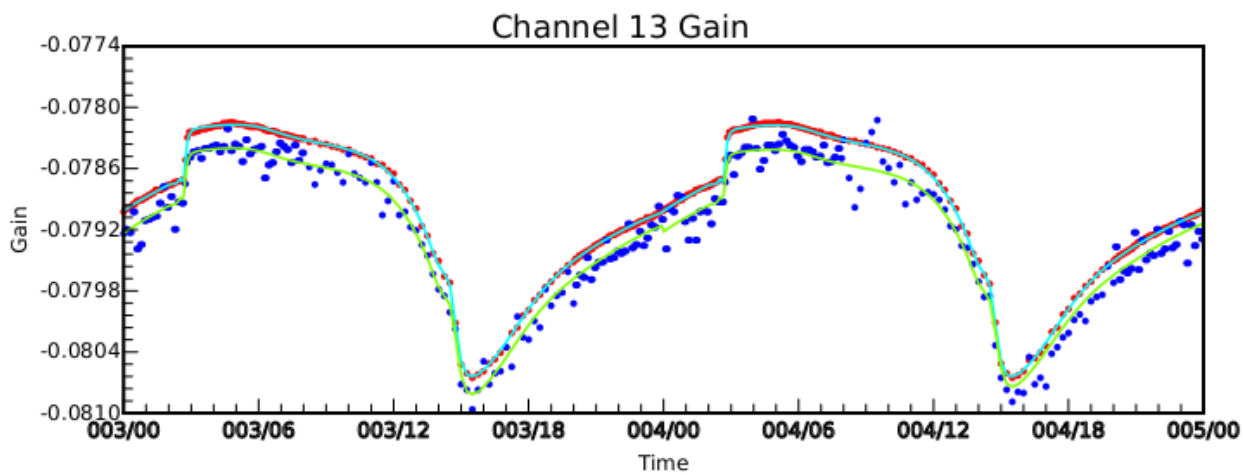

Figure 6 The gain parameters for the detector 270 and 271 in the 10.4 µm channel. The read and blue dots represent the data points for the detector 270 and 271 respectively. The two green lines correspond to their time dependent functions obtained from the data training in the machine-learning algorithm. The dataset has the time range from 2015/003/00:00:00 to 2015/005/00:00:00.

5. Summary and Future Work

Our initial studies present a framework for creating the SA capabilities using the machine-learning algorithms in spacecraft operations by establishing the representation for spacecraft datasets and developing accurate, efficient, and adaptive data training strategies. The results for both GOES N-P Imager and AHI calibration data show that the time dependent trends obtained through the data training provides excellent description of the spacecraft dataset with arbitrary complexity and scales. In particular, the neural network implementation with two hidden layers has shown to provide excellent solutions for both continuous and discontinuous datasets. The application to the AHI data shows that the machine-learning approach enables engineers to monitor high volume datasets and evaluate the instrument data quality and radiometric accuracy, which would have not been possible with the current approach in data trending and monitoring.

More investigations are needed to improve the data training efficiency especially for AHI and ABI data with very high data volume. There have been extensive investigations in the neural network literature to improve the data training efficiency, such as adaptive learning algorithm[8]. The challenge is how to improve the data training efficiency while maintaining the accuracy of the time dependent trend at the same time. The machine learning approach to the RTP datasets is another important area that needs to be investigated. The ability to predict and monitor the state changes due spacecraft command or directive is another crucial aspect of SA capability for spacecraft operations. Since the RTP datasets could be both continuous and discrete, the data representation and the data training strategies could be very different from the one presented here. Furthermore, the extension of the current approach to the datasets for Polar-orbit satellites remains to be done, which the different data patterns may emerge that may impact the data training strategies.

The results of time dependent trend for both GOES N-P and AHI show that the training outcomes from our data-training algorithm are in excellent agreement with the data trend. The tight data bounds shown in AHI results are

highly sensitive to data changes above the noise level, which lead to more dynamic data monitoring. This offers a very promising advance toward SA capabilities in spacecraft operations.

The machine learning approach presented in this paper has been implemented in an Intelligent Trending and Monitoring Toolkit, and it will be deployed in GOES-R ground system to monitor the instrument data calibration and navigation and registration processes.

The author is grateful to Bob Iacovazzi and Fred Wu at NOAA/STAR for providing the AHI instrument calibration data used in this paper, to Japan Meteorological Agency (JMA) for sharing with NOAA the AHI instrument calibration data under the collaboration between NOAA and JMA on ABI/AHI calibration/validation program, and to Dr. Fangfang Yu at NOAA/STAR for comments and suggestions.

Reference

[1] See http://www.data.jma.go.jp/mscweb/en/himawari89/space_segment/spsg_ahi.html.

[2] See http://www.goes-r.gov/spacesegment/abi.html.

[3] Zhenping Li, David Pogorzala, Ken Mitchell, J.P. Douglas, "Adaptive trending and limit monitoring algorithm for GOES-R ABI radiometric parameters" GSICS Quarterly Newsletter, Summer 2015 Issue, http://dx.doi.org/10.7289/V5XK8CHN#page9.

[4] Zhenping Li, J.P. Douglas, and Ken Mitchell, "Creating Situational Awareness in Satellite Data Trending and Monitoring", presentation at 32nd Space Symposium, April 11-12, 2016, Colorado Springs, Colorado. http://www.spacesymposium.org/tracks/technical-track/papers.

[5] H. Yu and B. M. Wilamowski, "Levenberg-Marquardt Training," in The Industrial Electronics Handbook, Vol. 5−Intelligent Systems, 2nd ed. (CRC Press, Boca Raton, 2011).

[6] Tom M. Mitchell, "Machine Learning", McGraw-Hill, (1997).

[7] G Cybenko, "Continuous Valued Neural Networks with Two Hidden Layers Are Sufficient", (Technical Report). Department of Computer Science, Tuffs University, Medford, MA. (1988).

[8] M. Riedmiller, "Advanced Supervised Learning in Multi-Layer Perceptrons –From Backpropagation to Adaptive Learning Algorithms", Computer Standards & Interfaces, 16 (1992), 265, http://dx.doi.org/10.1016/0920-5489(94)90017-5.