GOES-R Supervised Machine Learning

Ronald O.S. Adomako[1] and Michael Mandel[2]

[1]Department of Computer Science, Data Science Engineering, City College of New York, CUNY; NOAA-CESSRST
[2]Brooklyn College, CUNY

December 2020

1 Introduction

The GOES-R series is a product line of two satellites (GOES-16 "East" and GOES-17 "West"). GOES-17 is susceptible to a Loop-Heat-Pipe (LHP) phenomenon where during Fall and Spring U.S. Continental seasons, there are times of day where the infrared bands records inaccurate readings from the Advanced Baseline Imager (ABI). This occurs from joint astronomical behavior and position of the GOES-17 satellite. This calibration issue occurs when the LHP instrument fails to radiate the heat of the sun out of ABI. Predictive Calibration (pCal) is an algorithm developed by the National Oceanic Atmospheric Agency (NOAA) to correct the readings of GOES-17 [Lindstrom, 2019]. pCal is a regression that corrects for the average temperature in a region of interest where a threshold of points may be susceptible to LHP [Yu et al., 2019]. pCal is an equation. In this project we explore a multi-layer perceptron (MLP), neural network, to train a model using various sample size. We compare our correlation scores, mean square error (MSE), and mean absolute (MAE) between pCal and MLP models. The quality label is an associated value to a measurement taken that describes the reliability of the measurement. There are sixteen channels per satellite, specialized to take measurements of various properties.

GOES stand for Geostationary Observational Environmental Satellite and both GOES-West and GOES-East make up the GOES-R series [NASA, 2019].

1.1 Problem

1.1.1 What

One approach is a synthetic break-down of GOES-16 measurements to simulate the skew of GOES-17 measurement with the hope of reverse engineering algorithm to generate a correct GOES-17 measurement. Another approach is to compile a visualization library of temperature histogram to study behavior GOES-R radiance and anomalies. Finally, taking both of these approaches into consideration, we do the next best thing: machine learning. We explore an MLP with of 100 layers and 500 iterations under supervised learning. With the size of the data set we approach the domain of deep learning.

One pixel can represent the size of a city on a global scale. The objective is to make accurate predictions of pixel to pixel transformations. Currently, state-of-art predictions is limited to average temperature of a region of interest, which can represent the size of a country.

1.1.2 Why

Impact: The Cooperative Institute for Meteorological Satellite Studies (CIMSS) has requested histogram plots as an added feature to their public database. The values recorded from the operational data set is used by the National Weather Service (NWS) for weather forecasting. Novelty: previous work revealed that single-mode, bi-modal, and tri-modal features occur in the histogram analysis [Adomako et al., 2020]. We are identifying radiance anomalies using machine learning and will incorporate artificial intelligence into GOES. This project cross-validates work done by CIMSS and the GOES community at large using different methodology. These are the beginning steps to deploy cloud based architecture to run ten terabytes of data on household devices. Tractability: findings are available to the open-source community for tracking and reference. Presentations, proceedings, and publishing are in collaboration with NOAA Center for Earth Science Systems and Remote Sensing Technology (NOAA-CESSRST) and CIMSS lab.

2 Related Work

In Figure 1 we see the region of interest (ROI). The latitudinal range is from -109.59326ºE to -102.40674ºE and the longitudinal range is from 8.94659ºN to

Figure 1: Region of Interest for GOES

-8.94656ºN. These extents were transcribed from the operational full disk netcdf files. For a full disk image the time and band is the same for all pixels in that file. The band is the channel that satellite is tuned in for a radiance reading. The operational data is a cleaned version of the raw satellite data for organizations such as NWS. We can round to five significant digits for consistency.

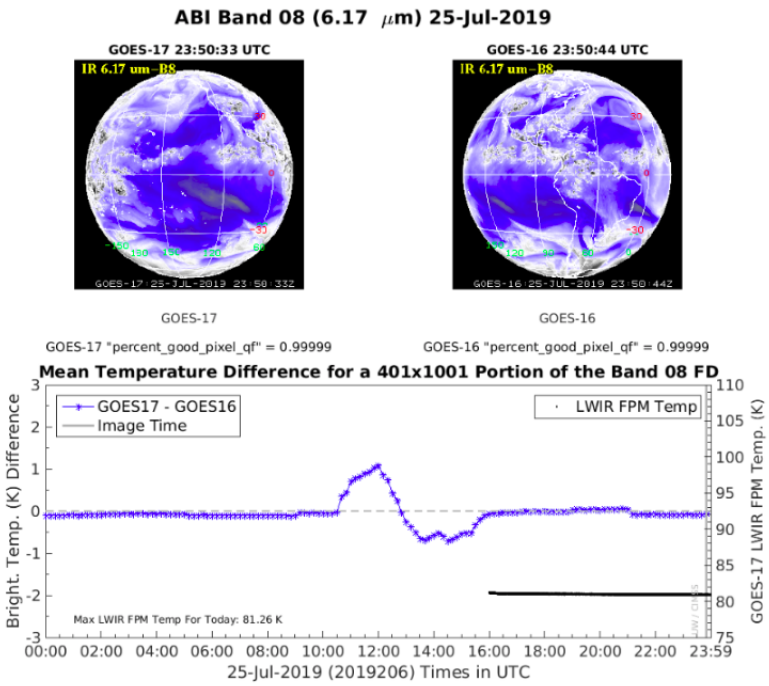

Figure 2: Temperature difference for GOES-R in ROI

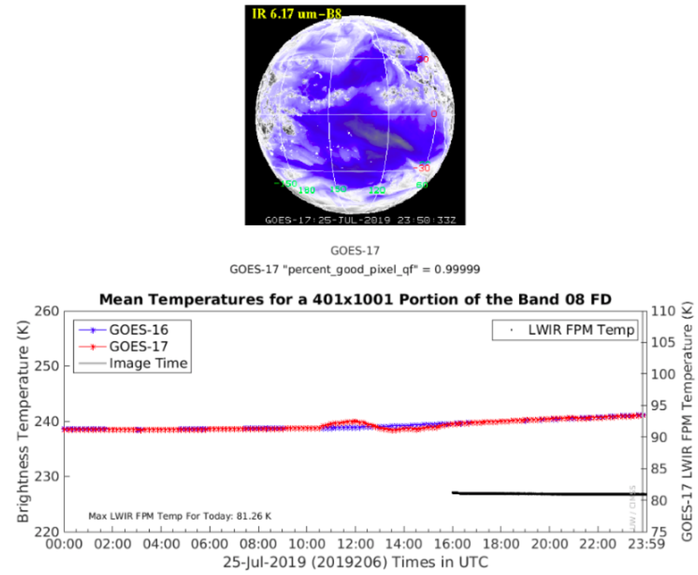

Figure 3: Average temperature comparison for GOES-R in ROI

Studies of the LHP problem have looked at a conversion of the radiance (L) recordings in brightness temperatures as shown in Figures 2 and 3 ([CIMSS, 2017, Smith, 2017]). The temperature (T) is interpreted in Kelvin using a discrete formation of Plank's law where $fk_{1,2}$ and $bc_{1,2}$ are functions of central frequency ([Wang, 2020]):

$$T = \frac{\frac{fk_2}{log^{-1}(\frac{fk_1}{L}+1)-bc_1}}{bc_2} \tag{1}$$

$$L = \frac{fk_1}{\frac{(e)(fk_2)}{(bc_1+(bc_2*T))-1}} \tag{2}$$

July 25, 2019 was chosen as the date to observe the LHP phenomenon because it was the last day until GOES switched to pCal corrections for operational data (Figure 4 and 5). This date is chosen for a similar comparison on July 26, 2019 where the pCal corrections are expected to have the same underlying LHP behavior for a day difference. The pre-operational data set is not publicly available and some metadata on of the satellite instrument is proprietary. We expect that the degree of the LHP phenomenon does not change significantly in one day so that we can make a one-to-one comparison between MLP and pCal.

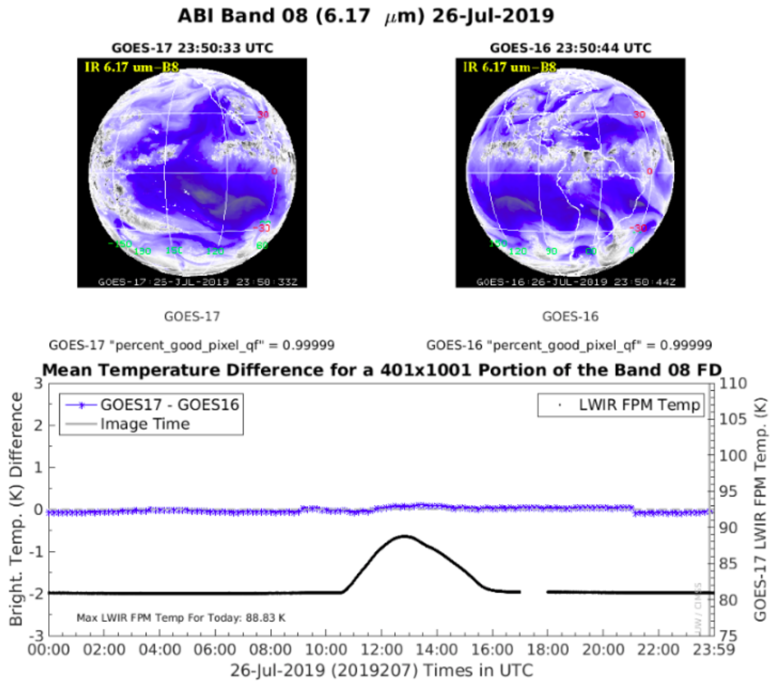

Figure 4: pCal temperature difference for GOES-R in ROI

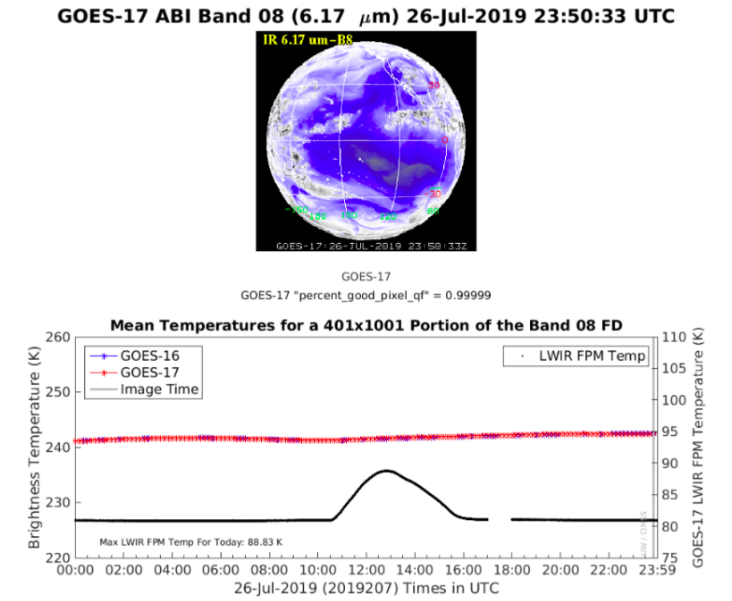

Figure 5: pCal average temperature comparison for GOES-R in ROI

3 Approach

In a previous study we were able to measure a true-positive (hit rate) up 82 percent when we formalized the LHP phenomenon as a classification problem of five categories. NOAA provides data quality flags (DQF), effectively a confidence interval, for how accurate each temperature measurements are. In this formulation we chose similar inputs where included class categories of one through five to represent the number of standard deviations that a measurement was from the mean. The lower the measurement the more accurate. We mapped this onto the five categories for DQF where zero is represented as the best quality. Similarly, the GOES-16 DQF measurement was chosen as ground truth.

In a discrete view of the data, we also examined a histogram mapping for extreme cases of LHP behavior where there are gaps of data. We were able to rule noise effect of clouds in the atmosphere. Once filtering the pixels using a cloud mask we compared remaining pixels and found results to be statistically significant. When comparing averages over the region of interest, on average, clouds statistically do not change your results (Figure 6).

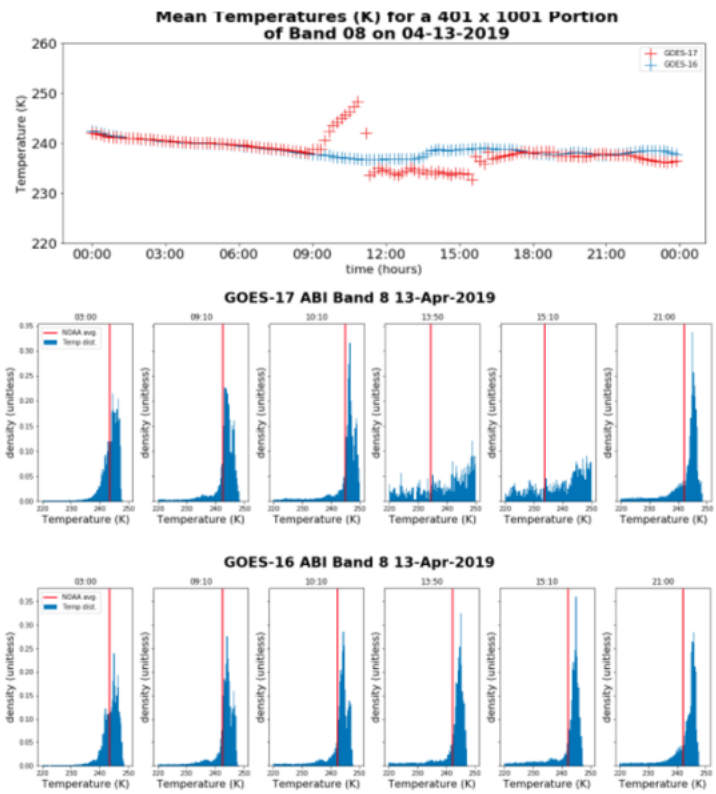

Figure 6: Average temperature plot with select histogram distribution

In a machine learning classification experiment on the temperature data the for various training models the test error was very high (Figure 7), because for numerical data, a regression analysis is more appropriate.

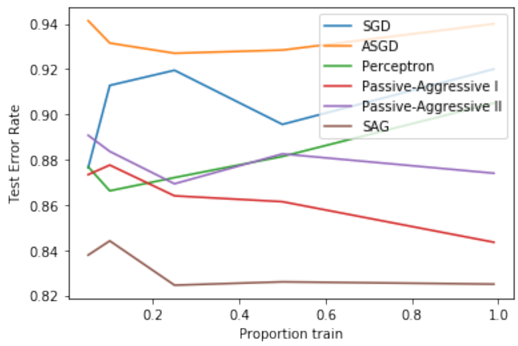

Figure 7: Average temperature plot with select histogram distribution

From our classification experiment on temperatures we learned that the Passive-Aggressive I (linear loss function) and Stochastic Average Gradient (SGD, logistic regression) were the best performing machine learning algorithms on GOES-R and needed to train a minimum around 20 percent of the data.

3.1 How

In our machine learning algorithm we consider the following inputs ([Adomako, 2020]):

time	longitude	latitude	band	G17 Temp	G17 std–dev	G17 mean

The output is the GOES-East measurement:

G16 Temp

3.2 Design

1. Discretize region of interest into pixels
   - Longitude and latitude
   - 401 (columns) by 1001 (rows) pixelation
2. Run a Regression
   - Multilayer Perceptron neural network
   - 100 layers, 500 iterations

3. Compare R-square, MSE, and MAE
   - MLP vs Pcal
   - Trend of training points as function of sample size

4 Results

For July 25 and July 26, 2019 we chose the 1200:00 UTC timestamp to compare our results. The 1200:00 UTC timestamp has is in a range for a maximum number erroneous. These full disk image time stamps are done in 10 minute or 15 minute intervals for each file. A full disk image file captures the entire extent of the GOES-R imagery and is necessary to obtain the ROI. We want to know the limits of the MLP and pCal models. For the vast amount of data in each full disk image a thorough analysis is done on one file to observe trends before generalizing an approach to many full disk image files. From Figure 8 we see that 100 testing points out of 401401 is too small a sample size. 1000 test points would be sufficient for learning but the errors are too high and the correlation is poor. We have a further degradation in our accuracy after 10000 points are trained. The cross-validation is taken at 10 percent: a 1 to 9 ratio for testing and training sets respectively for a given sample size. Notice even in the case of oversampling beyond 10000 points our correlation ratio remains constant near 1 showing that the models are correct and not over-fitting. For full disk images where there is the LHP behavior machine learning still suffers. The accuracy becomes poor because we pick too many poor samples. Between 1000 and 10000 pixels is the sweet spot for machine learning on GOES-R this was observed for full disk image files where the LHP behavior was not present. Figure 9 depicts tables for a closer inspection.

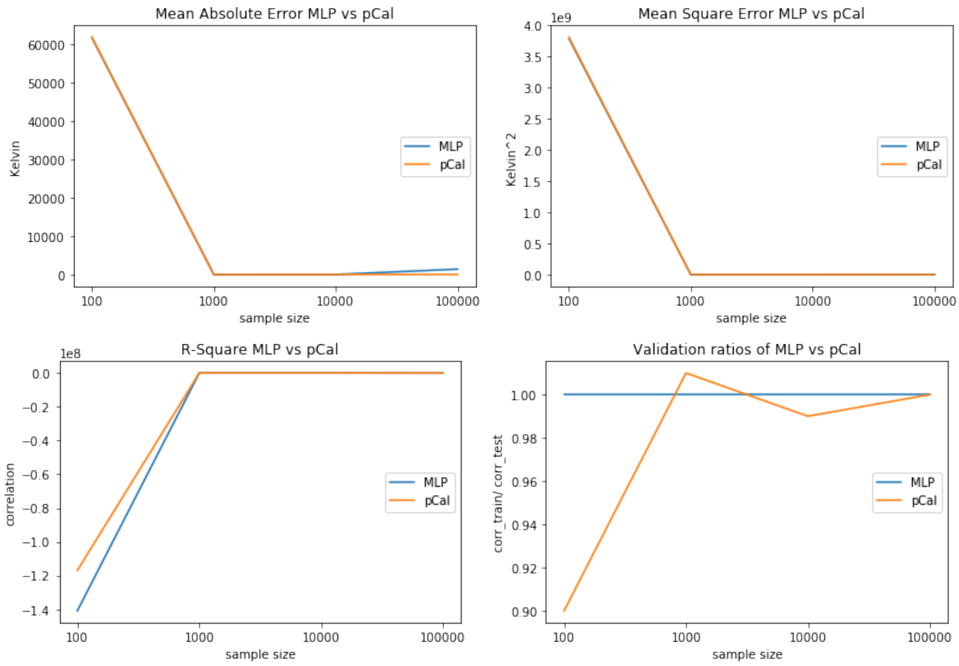

Figure 8: Correlation and error of Multilayer Perceptron vs Predictive Calibration

Metric 100pts	MLP	pCal
MAE	61576.82	61729.15
MSE	3.73	4.5
R-Square	13.39	4.67
Validation ratio	1401.98	12.01
Metric 1000pts	MLP	pCal
MAE	3791704747.71	3810506987.43
MSE	23.95	34.1
R-Square	187.69	35.85
Validation ratio	1965559.63	177.83
Metric 10000pts	MLP	pCal
MAE	-140866512.25	-116985450.28
MSE	-0.31	-0.32
R-Square	-7.27	-0.33
Validation ratio	-83760.29	-4.67
Metric 100000pts	MLP	pCal
MAE	1	0.9
MSE	1	1.01
R-Square	1	0.99
Validation ratio	1	1

Figure 9: Metrics of Multilayer Perceptron and Predictive Calibration

5 Conclusion

The MLP will accurately predict pixels from the image that it is trained on, which is sufficient when GOES-16 is available. For non-LHP behavior full disk image files you could train on the ROI and then predict the rest of the image. Future work will determine whether we can train the model once sampling various files and use it to determine both temperatures where GOES-16 and GOES-17 do not overlap and new images, especially ones with the LHP problem.

5.1 Discussion

The aim of the first part of this study was determine whether we can automate a transformation of GOES-17 temperatures to correct temperatures estimated by GOES-16 (Figure 10)

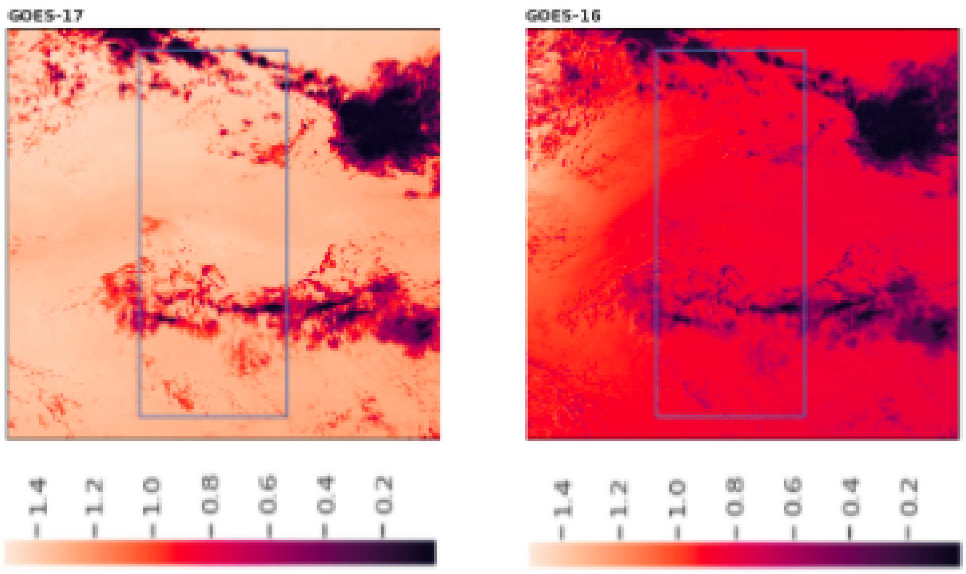

Figure 10: GOES-R Transformation of Brightness Temperature

We found that a regression is an appropriate model for numerical data and determined the range of optimal sample points to be between 1000 and 10000. Altogether, We found limitation on training using files that exhibit the LHP phenomenon. We experimented with a few classification models with limited success. This topic commands further research interest as better, automated predictions of the weather leads to better quality of life and is a topic of daily interest.

5.2 Future Work

Continuing work on GOES-R will include computer visioning and anomaly detection of images (Figure 11). This will be artificial intelligence for CIMSS lab to auto-detect which files are defective before generating and publishing images. The median full disk image file is about 20 megabytes, they are generated every 10 minutes to 15 minutes, for every day of the year. It is costly for an engineer to spot check each one of these and both computationally and professionally expensive to load defective files. Using a U-Net training model, we will determine a tensorflow or pytorch workflow that labels 30 sample images pixel-to-pixel of known image anomalies (artifacts). U-net is a convolutional neural network traditionally developed for biomedical imaging. This approach is adapted for GOES-R. The image artifacts are "caterpillar tracks" and "shark fins". The "shark fins" are most prevalent and will be the aim for detection as a proof of concept. The "caterpillar track" is out of scope.

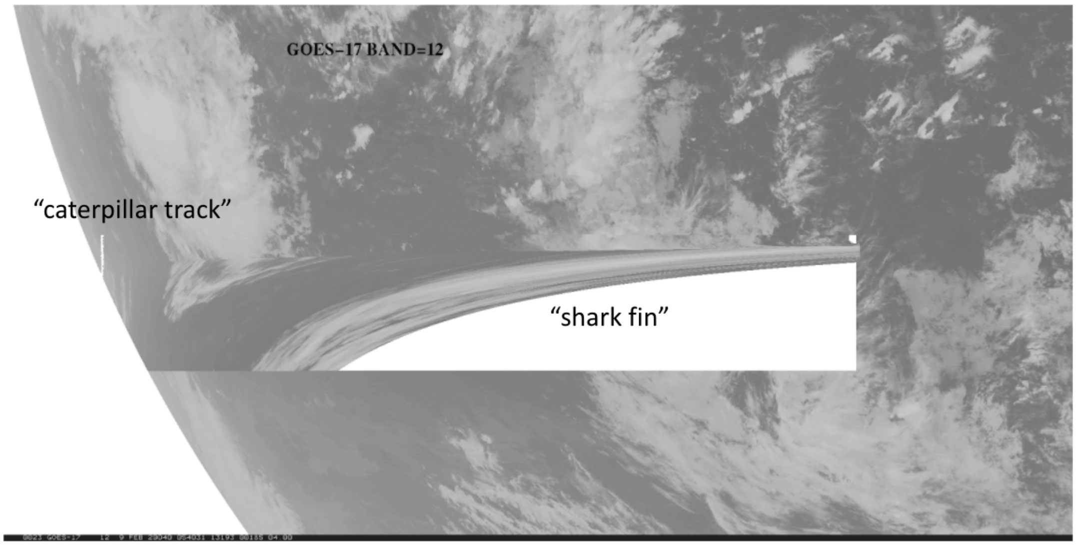

Figure 11: Artifact of both GOES-16 and GOES-17 (courtesy of T. Schmit

References

[Adomako, 2020] Adomako, R. O. (2020). capstone. In *Data Science Engineering thesis; https://github.com/adomakor412/capstone*.

[Adomako et al., 2020] Adomako, R. O., Grossberg, M. D., Schmit, T., and Aizenman, H. (2020). Distribution of goes-16 and goes-17 abi radiance anomalies. In *AMS 100th American Meteorological Society Annual Meeting; 16th Annual Symposium on New Generation Operational Environmental Satellite Systems*.

[CIMSS, 2017] CIMSS (2017). Goes abi realtime imagery. In *GOES Calibration, GOES-16/17 Comparison*.

[Lindstrom, 2019] Lindstrom, S. (2019). Predictive calibration is now operational for goes–17. In *CIMSS Satellite Blog*.

[NASA, 2019] NASA (2019). National aeronautical space association- goes satellite network. In *GOES-R*.

[Smith, 2017] Smith, T. J. (2017). A closer look at the abi on the goes-r series. In *Bulletin of the American Meteorological Society, Volume 98: Issue 4*.

[Wang, 2020] Wang, Z. (2020). On–orbit calibration and characterization of goes-17 abi ir bands under dynamic thermal condition. In *SPIE Digital library*.

[Yu et al., 2019] Yu, F., Wu, X., Yoo, H., Wang, Z., Qian, H., and Shao, X. (2019). Radiometric calibration performance of goes-17 advanced baseline imager (abi). In *Proc. SPIE 11127, Earth Observing Systems XXIV, 111271C*.