



# **Anomaly Detection for GONG Doppler Imagery Using a Binary Classification Neural Network**

Mitchell Creelman, Kiran Jain, Niles Oien and Thomas M. Wentzel

*NISP, National Solar Observatory, Boulder, CO 80303, USA*

Technical Report No. NSO/NISP-2024-001

---

---

## **Abstract**

One of the products of the National Solar Observatory's Integrated Synoptic Program (NISP) is the farside seismic map which shows the magnetic activity on the unobserved side of the Sun. The production of these rudimentary maps began in 2006, and they have since proven to be a valuable tool in tracking solar activity which cannot be directly observed from the earth's surface. The continuous tracking of solar active regions allows space weather forecasters to monitor critical solar events which may have larger economic and societal impacts here on Earth. In an effort to improve these maps, several steps are underway through the Windows on the Universe project (WoU) funded by the NSF. One of these steps is to improve the quality assurance measures for the images collected at individual sites throughout the GONG network and is used to develop the farside maps. To this end, we have designed a binary classification neural network to determine which of these site images should and should not be included in the farside pipeline that produces the end product maps. This convolutional neural network is a highly effective and computationally efficient method of significantly improving the quality of the farside maps currently produced by the NISP program.

# Contents

<b>Abstract</b> . . . . .	<b>1</b>
<b>1 Background</b>	<b>3</b>
1.1 How binned-down images are generated . . . . .	4
1.2 What is the concern? . . . . .	4
<b>2 Image Quality Assurance and Control</b>	<b>4</b>
2.1 Current QAQC Filter . . . . .	5
2.2 Proposed QAQC Filter . . . . .	5
<b>3 Training the Network</b>	<b>5</b>
3.1 Sampling . . . . .	5
3.2 Network Design . . . . .	7
3.3 Hyperparameters . . . . .	9
<b>4 Performance Validation</b>	<b>9</b>
4.1 Test 1: Random Sample . . . . .	9
4.2 Test 2: RMS Rejections . . . . .	9
<b>5 Effects of Implementation</b>	<b>10</b>
5.1 Impact on Network Duty Cycle . . . . .	11
5.2 Impact on Site DC . . . . .	15
5.2.1 Learmonth (1e) . . . . .	16
5.2.2 Udaipur (ud) . . . . .	17
5.2.3 El Teide (td) . . . . .	18
5.2.4 Cerro Tololo (ct) . . . . .	19
5.2.5 Big Bear (bb) . . . . .	21
5.2.6 Mauna Loa (ml) . . . . .	23
<b>6 Case Studies for Farside Map Effects</b>	<b>24</b>
6.1 Noise Reduction in Merged fqi Images . . . . .	24
6.2 Noise Reduction in Farside Maps . . . . .	26
<b>7 Network Updates and Improvements</b>	<b>28</b>
7.1 Re-training . . . . .	28
7.2 Network Design Updates . . . . .	29
<b>8 Summary</b>	<b>29</b>
<b>Acknowledgements</b>	<b>29</b>
<b>References</b>	<b>29</b>

# 1 Background

The Global Oscillation Network Group (GONG) is a network of ground-based solar observatories run by the National Solar Observatory’s Integrated Synoptic Program (NISIP) which is aimed at collecting uninterrupted solar observations (Harvey et al., 1996; Hill, 2018). The network, established in 1995, is comprised of six observatories whose locations were chosen in order to create a redundant system capable of collecting solar observations with a mean duty cycle (DC) of over 90%. After a careful survey, observatories were established at Learmonth, Australia (1e), Udaipur, India (ud), El Teide, in the Canary Islands (td), Cerro Tololo, Chile (ct), Big Bear, California (bb), and Mauna Loa, Hawaii (ml) (Hill et al., 1994). The location of these observatories, shown in Figure 1, were carefully chosen to provide sufficient latitudinal and longitudinal coverage for consistent observational overlap between adjacent sites. A recent study by Jain et al. (2021b) shows that the network is capable of collecting observations from as many as four sites simultaneously depending on the time of year.

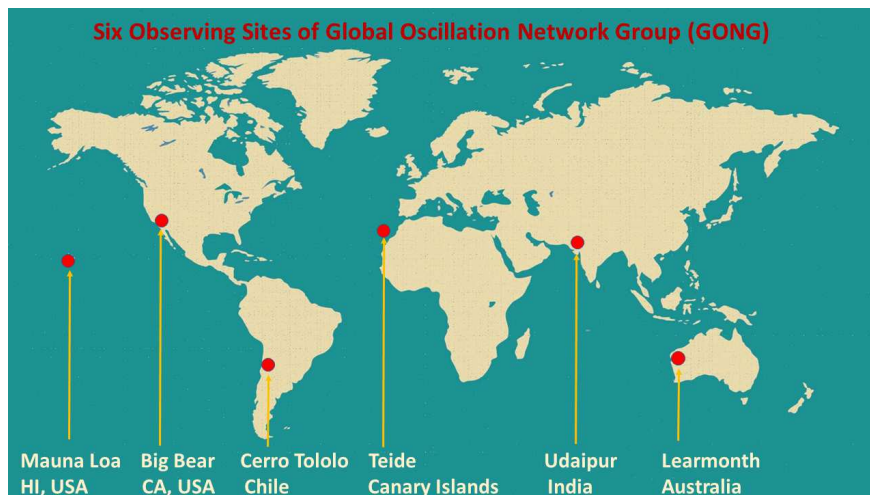


Figure 1: *Global distribution of all six GONG network observatories.*

While multiple data products are derived from the observations of current GONG network<sup>1</sup>, this report is primarily focused on the line-of-site low-resolution Dopplergrams which are produced and used to track the magnetic activity on the farside of the Sun. These Dopplergrams, or *fqi* images, are  $215 \times 215$  pixel images which are produced from the high-resolution images collected at individual sites. The individual site minute-resolution Dopplergrams are then merged together to form a nearly continuous series of images. If there are no images for a given minute, the produced Dopplergram is left empty. The fraction of 24 hr day for which the observations are available is known as the duty cycle (DC), with a range between 0 and 1. The Dopplergrams are then used to determine phase shift values that are utilized to create helioseismic maps of the farside produced by the NISP program (Lindsey and Braun, 2000; Braun and Lindsey, 2001; González Hernández et al., 2007, 2010).

The phase shift maps are employed for a wide variety of space weather forecasting applications (e.g., Jain et al., 2023). In particular, they serve to help monitor the presence of known solar active regions, as well as the development of new regions on the farside of the Sun. Through developing improved farside helioseismic maps, we can improve the prediction and warning times for solar activity events, such as flares, coronal mass ejections (CMEs), and increase our understanding of the precursors for these events.

<sup>1</sup><https://gong2.nso.edu/archive/patch.pl>

One such example can be found in the sudden increase in the irradiance values during the early phase of the current solar cycle 25. While the frontside hemisphere saw a reduction in magnetic activity, a large active region which had emerged on the farside was found responsible for this sudden rise in the irradiance (Jain et al., 2021a). These farside seismic maps have also been used to create global magnetic index (e.g., González Hernández et al., 2011, 2014), which is a useful proxy for the magnetic field models.

### 1.1 How binned-down images are generated

To generate single-site binned-down  $215 \times 215$  pixel *fqi* images, the near-real time  $860 \times 860$  pixel full-resolution *vqi* images are run through the Image Reduction and Analysis Facility (IRAF)/GONG data Reduction and Analysis Software Package (GRASP)’s *mkfsi* routine to fit a 2-d surface fit to the image and then subtract the obtained surface from the image to remove the solar rotation. A Gaussian filter is then applied before the images are down sampled to their final resolution. These *fqi* images from sites are transferred to *iQR* Keep at the NISP Data Center where they are processed by *qr-filter* to place them in the *oQR* Keep with altered header and added image statistics. The next step is to merge all *oQR* single-site images for a given minute to create merged *mr-fqi* image. During this process, the *oQR* single-site *fqi* images are rotated, flipped, and scaled to a standard radius prior to merging (see Toner et al., 2003; Hughes et al., 2016, for general description).

### 1.2 What is the concern?

The primary goal of this work is to develop a consistent and reliable methodology for excluding anomalous site images in the merged *fqi* images and thus the downstream farside data products. Figure 2 displays examples of the normal and erroneous images. The possible causes of anomalous images vary from malfunctioning calibration sequences, instrumental failures, or processing issues. The inclusion of anomalous Dopplergrams in the computation of farside seismic maps can have adverse consequences on the signal-to-noise ratio without affecting the duty cycle and rendering the final maps ineffective.

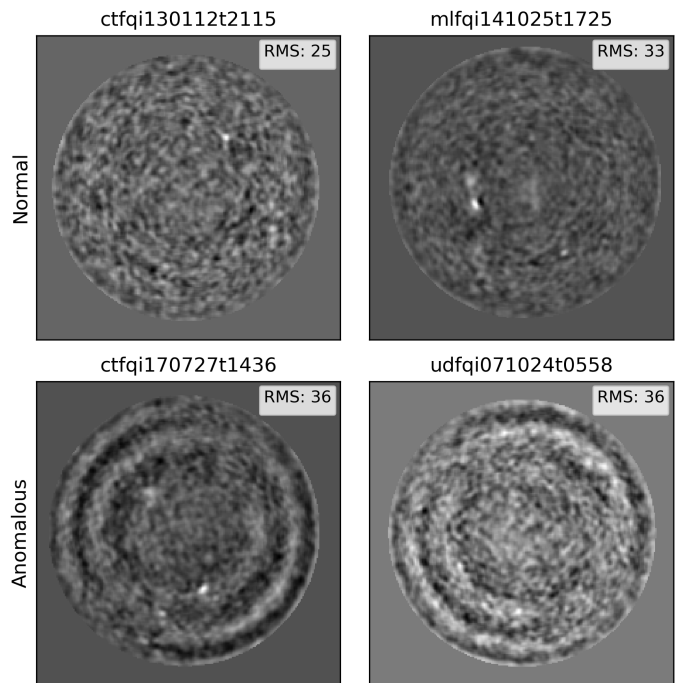


Figure 2: *Examples of normal (Top) and anomalous site images (Bottom) which would pass the current RMS site threshold.*

## 2 Image Quality Assurance and Quality Control

Implementing an effective quality assurance and quality control (QAQC) parameters for the site-specific *fqi* products is the overall focus of this report. Currently, the only QAQC method in place is a thresholding of Root Mean Square (RMS) values. As will be covered in more detail later in this report, this methodology is largely ineffective, allowing many anomalous images to be included while excluding valuable data. The alternative proposed here is the implementation of a binary classification neural network for detection of

anomalous `fqi` images. While not a perfect solution, such networks have been shown to be highly effective for anomaly identification tasks (Khan et al., 2021).

## 2.1 Current QAQC Filter

As mentioned, there is currently only one quality control parameter set on the `fqi` images generated at any individual site. This is a RMS filter with a threshold of 60. An inspection of the source code for the `quick_reduce` script shows that this filter was implemented on July 23, 2014. Prior to this date, all images irrespective of their quality were included in the image-merge process and sometimes produced poor quality data products. This filter diverts all images with the RMS exceeding the threshold and stores them in the `Oubliete Keep` at sites. These images are then discarded after 41 days without being transferred to the NISP Data Center. This type of filter fails to include vital images during the high solar activity periods as well as to identify many of the anomalous images during the low-activity period. Figure 2 displays examples of four site images with comparable RMS values where two of these images are erroneous. Such a poor quality control measure has led us to explore the use of a binary classification neural network as an anomaly detection measure to replace the current filters.

## 2.2 Proposed QAQC Filter

For this application, we developed a filter using a simple binary classification network to label an `fqi` image as normal or anomalous. The developed filter seeks to prevent anomalous site-`fqi` images from being included in the `mr_fqi` products by identifying anomalous site images during the transition between the `iQR` and `oQR` keeps at the NISP Data Center. By excluding anomalous site-`fqi` images from the farside pipeline, we seek to improve the signal-to-noise ratio in the farside maps, thus providing a more reliable data product to the users relying on the NISP data.

The network was designed using the PyTorch Python package (Ketkar and Moolayil, 2021). PyTorch is a machine learning programming package that allows for a high degree of customization in the design and troubleshooting of the neural networks.

# 3 Training the Network

## 3.1 Sampling

As mentioned above, the `fqi` products in the farside pipeline are velocity images of nearside of the Sun, and the phase shift values are calculated to determine the location of the active regions which are not directly observed from the Earth (Lindsey and Braun, 2000).

The efficacy of machine learning algorithms is highly dependent on the quality of the datasets that they are trained on. Though imbalanced datasets can be used to effectively train networks (Korkmaz, 2020), balanced datasets can significantly improve the performance of machine learning algorithms (Divovic et al., 2021). However, this was complicated by the fact that we wanted to include a mix of normal and anomalous data within our training dataset. All of the data used in the training process were manually classified using an in-house tool developed by Kevin Shuman (Shuman, 2023).

To find anomalous data, we started with all images in the `fqi` dataset which had a variance of over 1600. Additionally, a random sample was taken of the `oQR` keep, with 5 random images taken from 20 days for a random 200 months within the lifespan of the GONG network. This left us with 110,593 high variance

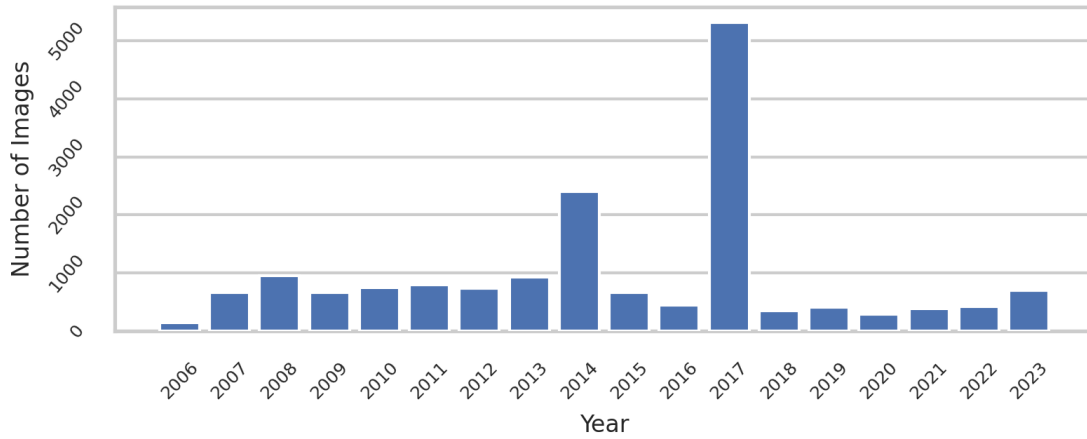


Figure 3: Number of training dataset images per year.

images, and 20,000 randomly sampled images. 21,268 of these images were then hand-sorted into normal or anomalous classes to be used to train the machine learning algorithm.

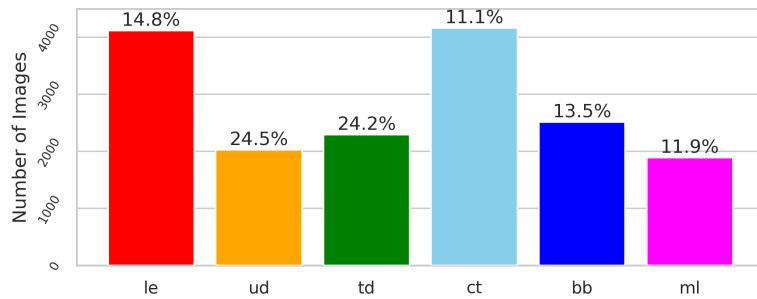


Figure 4: Distribution of individual site contributions to the 17014 images in the training dataset.

The year-wise distribution of the images used in training the network is shown in Figure 3. Using a training/testing split of 80% to 20%, 17,014 images were used to train the network while 4,254 images to test the performance of the classification network. The distribution of individual site images in the dataset is shown in Figure 4.

A source of imbalance in the training dataset was the rate of site occurrences between normal and anomalous images. This is illustrated in Figure 5 where ct is disproportionately represented in the training dataset at a rate nearly twice the other sites. While ct is the site with the largest sampling, it also appears as though ct and le produce a disproportionate number of anomalous images overall.

It appears that le made up nearly half of the anomalous images present in the training dataset. While improved sampling strategies may result in a more even distribution of sites among normal images, such discrepancies in the anomalous images would most likely remain. Such differences in occurrence rate by site are certainly something to be monitored in the flagged images caught by the fqj filter during deployment.

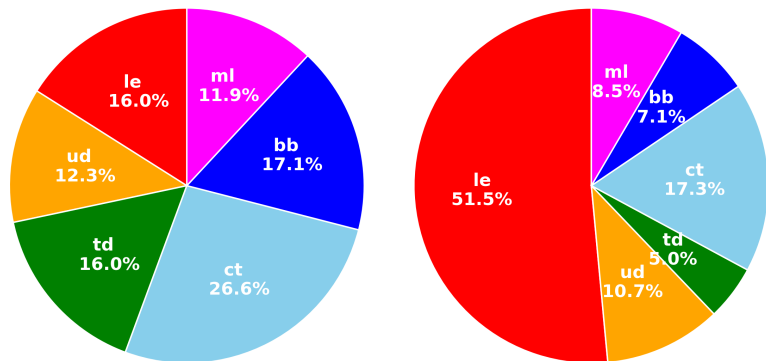


Figure 5: Distribution of 13098 normal (left) and 3916 anomalous (right) images by individual site for training dataset.

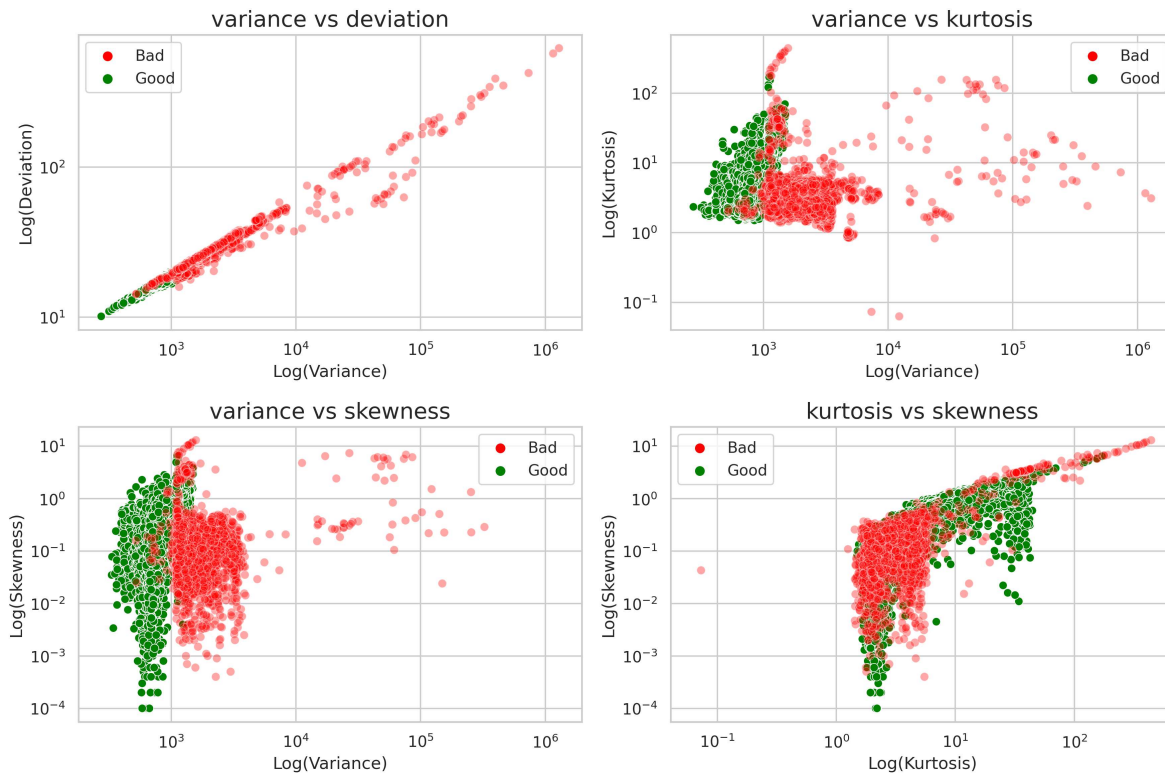


Figure 6: Statistics for training dataset highlighting the differences between normal (green) and anomalous (red) images.

Various statistical parameters for these images are plotted in Figure 6. Although the variance seems to be a good measure for identifying anomalous images, it becomes less effective during the high activity periods. In general, the erroneous images have higher variance as well as the RMS values relative to the normal images. The statistical overlap between the images during high-activity periods and the anomalous images often was seen in the database and it became one of the main motivations for the development of this filter.

### 3.2 Network Design

The design of the filter, outlined in Figure 7, is based on the neural network technique of image classification (Wang, 2022; Yal, 2023). However, while most classification algorithms look at a wide variety of classes (Yal, 2023), ours reduces the images into two bins: normal images and the anomalous images (Vailaya et al., 1998). The network design is simple, consisting of 3 convolutional layers, each followed by a leaky ReLU activation function and a max pooling. The leaky ReLU activation function was chosen for its better accuracy when compared to the commonly used ReLU activation function (Dubey and Jain, 2019). The network is ended with a sigmoid activation function due to its compatibility with more shallow networks (Wang, 2022).

One of the most important choices in designing the network was the ability to weight the output classes in order to minimize the number of images falsely classified as anomalous. The reason for this concern is the statistical similarity between anomalous images and normal images during periods of high activity. This

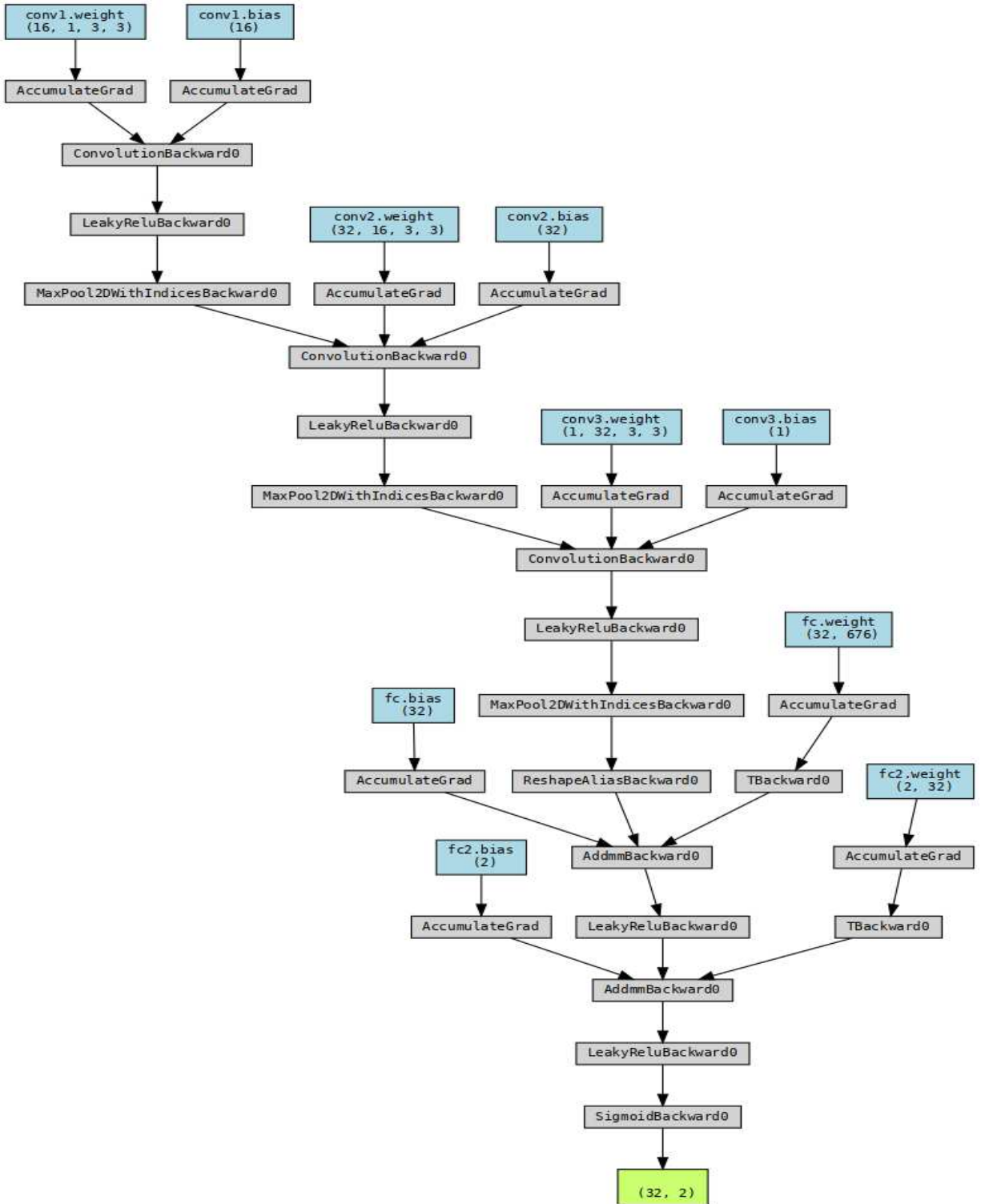


Figure 7: Outline of fqi Network design. This is based on a basic convolutional neural network which reduces the input image to a binary classification to determine whether the image is normal or anomalous.



means that if we allow the network to be more liberal in its tendency to classify images as anomalies, we run the risk of the network excluding large time-series of high-activity periods. To help prevent this, the network was designed to output a  $2 \times 1$  vector containing scores for each possible class as discussed by Khan (2019). These scores range between a value of 0 and 1 for each class.

### 3.3 Hyperparameters

While a variety of hyperparameters were explored for the training regiment, the final training regiment was set for 150 epochs with a batch size of 32, and a learning rate of 0.0008. The anomalous images were weighted at 1.6 while the normal images were weighted at 1.0 in order to minimize the occurrence of false anomaly classifications.

## 4 Performance Validation

There were two experiments performed in order to validate the efficacy of the neural network filter. The first of these was a random sampling of 3250 site `fqi` images from the keep which were then classified by the filter and also manually checked. The second experiment was an analysis of 617 images were rejected at sites due to the RMS values above the threshold as discussed in Section 2.1 and had been moved to `Oubliete`.

### 4.1 Test 1: Random Sample

The first experiment was the utilization of the `fqi` network to classify 3250 randomly sampled images from the storage keep. Of the 3250 randomly sampled images, 1 was found to be corrupt and unusable, leaving 3249 images to be used for the experiment. These images were then passed through the `fqi` network, which classified the individual site images as either normal or anomalous. We also manually inspected the 3249 images to ground-truth the dataset. Of the 3249 images, 98.8 % were correctly classified as normal images, and 0.8 % were correctly classified as anomalous. In addition, the network incorrectly classified remaining 0.4 % images. Of these 12 images, 10 were anomalous images which the network labeled as normal and 2 were normal images which the network misclassified as anomalous (Table 1). Overall, the network correctly identified normal images with an accuracy of 99.9 % and correctly identified anomalous images with an accuracy of 71.4 % This is in line with the intended results, as the network’s training regime was biased towards classifying images as normal in order to prevent discarding large amounts of good data due to the statistical similarities between anomalous and high-activity images as seen in the training dataset statistics (Figure 6).

Table 1: *Confusion matrix for random sampling of fqi images from keep to validate the fqi Network. Of the 3250 images selected, one image was found to be corrupt and thus dropped from the overall sample.*

	Actual Normal	Actual Anomaly
Predicted Normal	3212	10
Predicted Anomaly	2	25

### 4.2 Test 2: RMS Rejections

The second experiment was comparing the rejection performance of the `fqi` network to the current RMS threshold applied at sites before they are transferred to `iQR` and merged into `mr-fqi` products. In this thresh-

old, images are rejected if they have a RMS over 60. As mentioned above, the main drawback of this technique is that it can result in the rejection of images that are normal during high activity periods or allowing erroneous images to pass through during the low-activity periods.

For this experiment, we used 617 discarded site images that were stored in Oubliete. These images were manually checked to determine whether or not they had been correctly classified as anomalous by the thresholding methodology. Of these 617 images, 91.6 % of these images were incorrectly discarded and contained valuable data, while 8.4 % of the images were correctly discarded for containing artifacts (Table 2). It should be noted that it did not determine the performance of the RMS threshold on images which were labeled as normal and this would have required the manually scanning of all data produced by the GONG network, which is outside of the current scope of this project.

The next step was to pass all discarded images through the newly developed fqi network. The network labeled 92.1 % of the images as normal and labeled 7.9 % of the images as anomalous. The normal image predictions had an accuracy of 99.6 %, while the anomalous predictions had an accuracy of 90.4 %. This gives the network an overall accuracy of 98.7 % with a false positive rate of 0.4 %, and a false negative rate of 9.6 % (Table 2).

The comparison is stark. Next to the 8.4 % accuracy of the RMS threshold, the neural network showed an overall accuracy of 98.9 %. Thus, we can say that implementing it as a site-based fqi filter would provide a sizable improvement in the quality of `mr fqi` images and by extension the farside seismic maps produced by the GONG network.

Table 2: Test result comparison for RMS filtering (left) and the fqi network (right). The normal predictions for the thresholding method are excluded due to the volume of images, placing the category outside the scope of this study.

Threshold	Actual Normal	Actual Anom.	fqi Net	Actual Norm.	Actual Anom.
Pred. Norm	N/A	N/A	Pred. Norm	563	5
Pred. Anom	565	52	Pred. Anom	2	47

## 5 Effects of Implementation

There are two main areas of concern for implementing the fqi filter. The first is what is the effect it has on duty cycle. If the filter proves to take large sections of data, and reduce the duty cycle of the GONG network to below the generally accepted usability threshold of 0.8 or 80 %, then we would not consider it adequate as a QAQC measure. The second area of concern is the effect it has on the farside maps. If the application of the filter results in significantly lower quality of the farside maps, it would not be considered adequate as a QAQC measure.

In order to determine the effects of the developed fqi filter on duty cycle, we examined all images collected between 2010-01-01 and 2014-12-31 for, a total of 1826 days, or 2,629,440 minutes. It is important to mention that our data set include only those observations that were transferred to the NISP Data Center. All rejected site images were lost due to an existing criterion discussed in Section 2.1 and were unrecoverable. Details of the contributions from various sites are given in Table 3. It is evident from the table that all sites provide independent as well as overlapping observations. Though there are significant overlaps between different sites, the sites observed independently for about one-third of the total time. The maximum

Table 3: Rate of simultaneous observations from 2010-01-01 to 2014-12-31. This covered a period of 1826 days, or 2,629,400 minutes. The image counts for the individual sites show how often each site observed as a group of 0, 1, 2, 3, or 4 sites with the % of time as part of a group in parentheses.

# of Sites	minutes	le	ud	td	ct	bb	m1
0	153309 (5.83 %)	0	0	0	0	0	0
1	880548 (33.49 %)	288664 (33.81 %)	99175 (19.05 %)	192205 (24.05 %)	137100 (15.57 %)	81304 (10.14 %)	82100 (11.07 %)
2	1116679 (42.47 %)	452873 (53.05 %)	347749 (66.80 %)	374118 (46.81 %)	391754 (44.50 %)	326827 (40.78 %)	340037 (45.85 %)
3	432654 (16.45 %)	111027 (13.00 %)	73256 (14.07 %)	187757 (23.49 %)	305556 (34.71 %)	347117 (43.31 %)	273249 (36.84 %)
4	46250 (1.76 %)	1170 (0.14 %)	402 (0.08 %)	45080 (5.64 %)	45848 (5.21 %)	46250 (5.77 %)	46250 (6.24 %)

independent observations came from the le site. It is worth mentioning that erroneous images during the independent observations can have unfavorable effect on the duty cycle. Figure 8 illustrates the distribution of multiple-site observations during the day. One can easily notice that the overlaps between multiple sites are not random but these follow a specific pattern depending upon the season and the time during the day. The heatmap, shown in Figure 9, summarizes the number of overlapping observations between different pairs of sites. Since the observations are often overlapped, it is important to quantify the effect, if any, of fqi filter on the overall network duty cycle. We also discuss the performance of each site during the test period. These effects are the concentration of the next two subsections of this report.

In the next section, we look at the effects of filtering out anomalous fqi images on the farside maps. There, we show multiple examples of farside maps where the duty cycle either remained unchanged, or only slightly reduced with the application of fqi filter, and examine the effect excluding anomalous images has on the signal-to-noise ratio of these developed maps.

## 5.1 Impact on Network Duty Cycle

The impact of the fqi filter on the network duty cycle is an important concern for the implementation of this newly developed quality assurance tool as well as for the performance of the GONG network. If the network is too aggressive in its labeling of anomalous images, we risk losing large time series of good data. This is particularly difficult due to the overlap in statistical similarities between anomalous and high-activity images (Figure 6). However, if the network lets too many images pass through, then it will be ineffective as a QAQC measure. An analysis of the impacts of the fqi filter during our study period shows that it has been able to remove a significant number of anomalous images from the test dataset without seriously impacting the overall duty cycle of the GONG network. We summarize total number of rejected images at each site in Table 4 and Figure 10 displays the daily distribution of these rejected site images. We discuss two types of images in the following sections; the *unfiltered* and *filtered* images are the images prior and after passing through the proposed fqi filter, respectively.

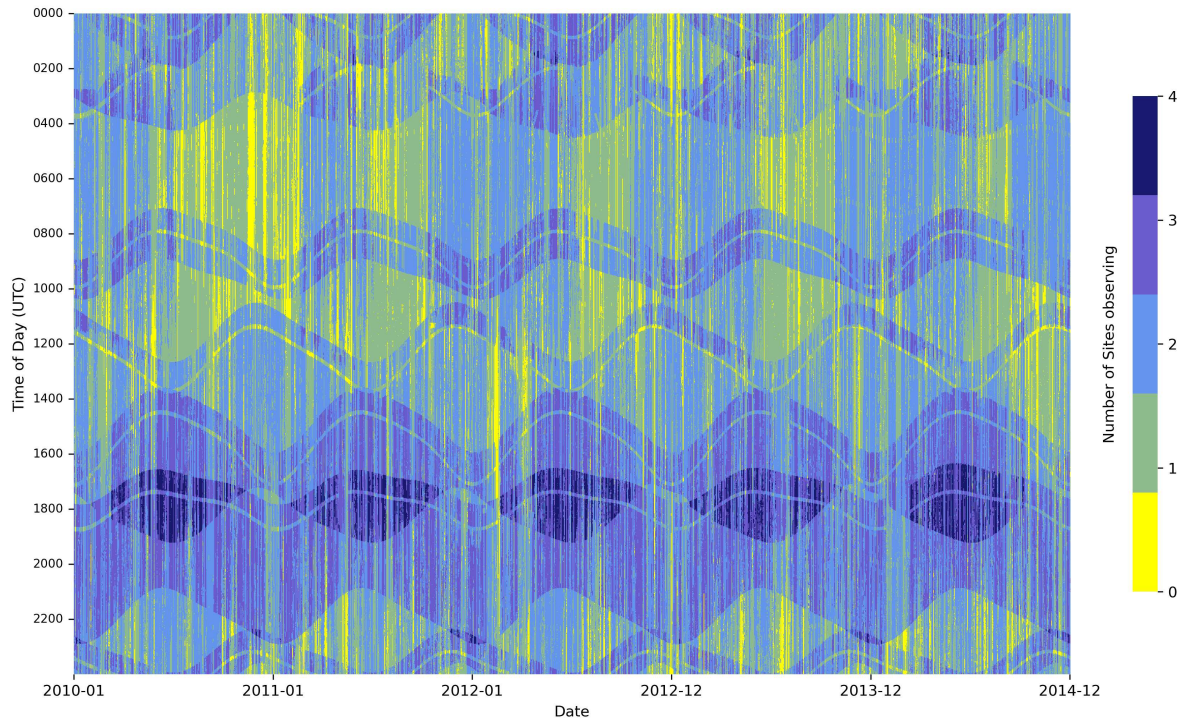


Figure 8: Minute-resolution figure showing the number of sites observed simultaneously during the day for the entire study period. Note that the temporal and seasonal patterns are clearly visible through density fluctuations.

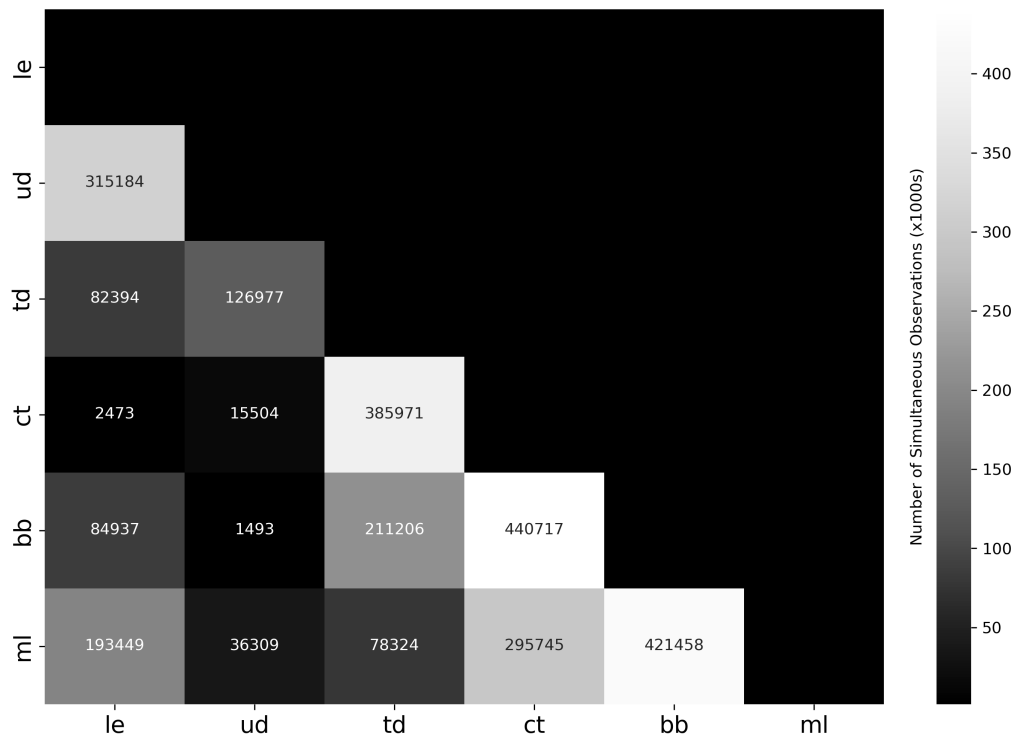


Figure 9: Heatmap showing the frequency of simultaneous observations for all site combinations for the testing period.

Table 4: Total number of images contributed by each site and the number of rejected images by the fqj filter. The % values indicate the percentage of all rejected images contributed from each site.

Site	Total Images (Unfiltered)		Rejected Images (Filtered)		
	Total #	%	Total #	% of site Images	% of All Rejections
le	853734	18.50	624	0.07	4.10
ud	520582	11.30	1454	0.28	9.50
td	799160	17.40	6510	0.81	42.50
ct	880258	19.10	328	0.04	2.10
bb	801498	17.40	5974	0.75	39.00
m1	741636	16.10	425	0.06	2.80
Total:	4596868	100.00	15315	0.33	100.00

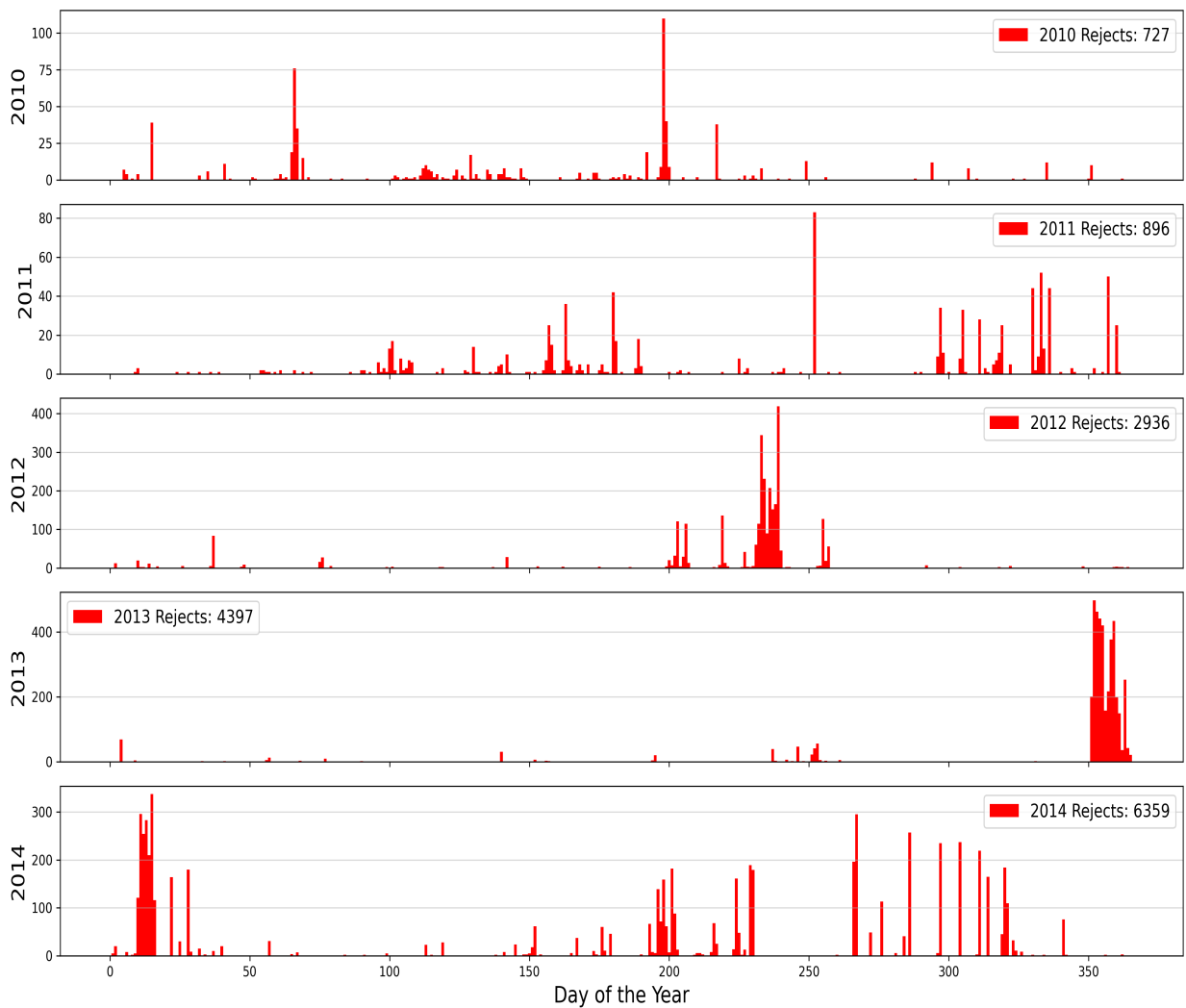


Figure 10: Number of images rejected each day of the sample period. Note the high volume events for each year which make up periods of consistent anomaly production.

In short, we find that the overall impacts on the GONG network’s duty cycle were minimal after passing all site images through the *fqi* filter. From Table 4, it can be seen that of the 4,596,868 images in the test dataset, the network excluded only 0.33 % of the total dataset. The daily distribution of rejected site images and the yearly total rejections for the testing period, exhibited in Figure 10, clearly suggests the

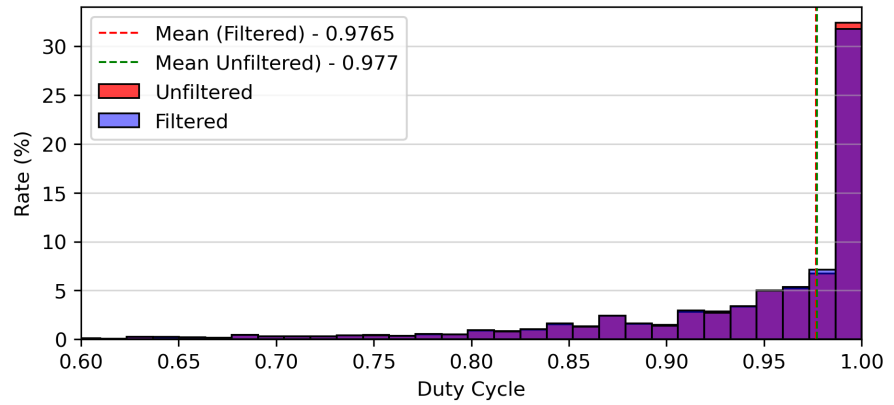


Figure 11: *Changes in network duty cycle due to *fqi* network implementation on the test dataset. The dashed lines show mean duty cycles for filtered and unfiltered images.*

sporadic nature of these rejections and the occurrence of these anomalous events is unpredictable. We notice that the maximum rejections happened during 2014. This resulted in the slight reduction of mean duty cycle from 94.17 % to 94.12 %, and the median duty cycle from 97.70 % to 97.65 % (Figure 11). Due to significant overlaps between different sites, there were only 145 days when the network duty cycle was affected by the exclusion of erroneous images. Of these days, as shown in Figure 12, the duty cycles were reduced by 0 – 1 %, 1 – 2 %, 2 – 4 % and 4 – 5 %, for 126, 7, 8 and 3 days, respectively. It is visible from Figure 13 that there was only one day when the duty cycle for the GONG network reduced by 10 %.

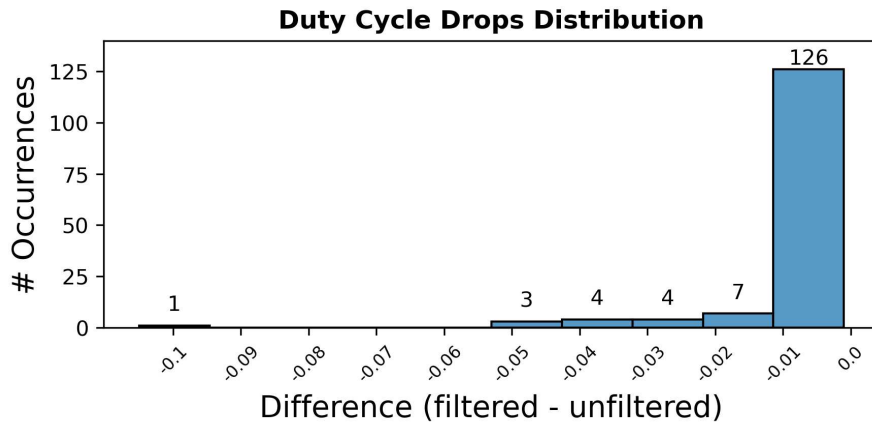


Figure 12: *Drops in duty cycle due to application of the *fqi* filter excluded 15,315 of the 4,596,868 images, affecting 145 days.*

The relatively minimal impact of the exclusion of bad site images on the network duty cycle is attributed to the distributed and redundant nature of the GONG network (Hill et al., 1994). While there are temporal coverage patterns of the overall network as shown in Figure 8 and also discussed in detail by Jain et al. (2021b), the network did not collect any observations for 5.7 % of the total minutes during our

study period. Overall, total independent observations from all individual sites in the network were 33.4 % of the total examined minutes, the simultaneous observations by two, three and four sites were 42.7 %, 16.5 % and 1.6 %, respectively. It is worth mentioning that the network’s duty cycle is most vulnerable to impact from the exclusion of site-images during the independent observing periods. These opportunities are especially prevalent during the period between July and October when the *ud* site is prone to be down due to the Indian monsoon or any other site being down due to instrument failures or other reasons (Jain et al., 2021b). The specifics of which sites often observe at the same time will be covered in the site-by-site analysis in the next section.

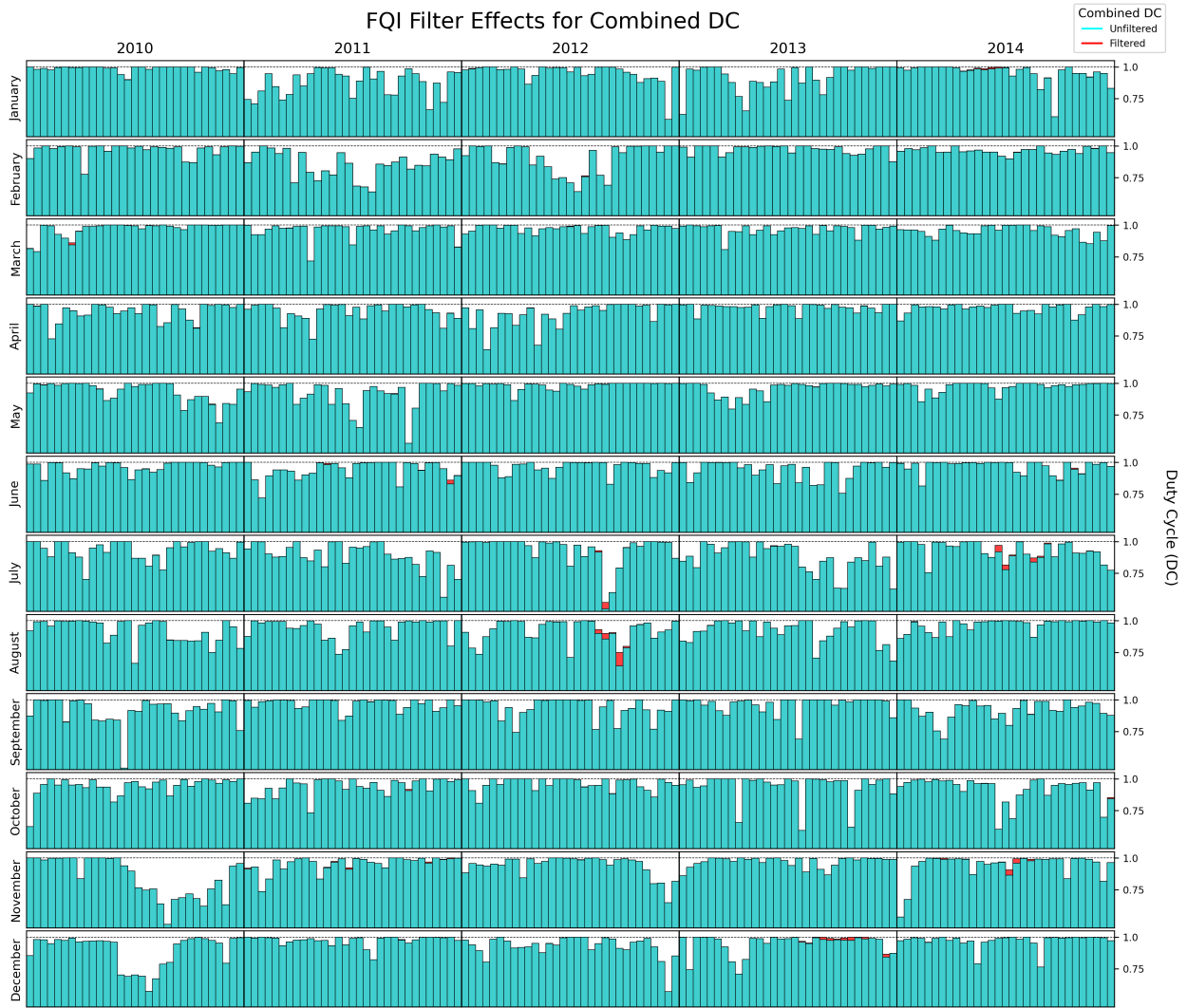


Figure 13: *Daily duty cycle for the network-merged images during the testing period from 2010 through 2014. The effects of the fqi filter are shown in red and the lower y limit shown is 0.45.*

## 5.2 Impact on Site DC

While the potential impact of the fqi filter on the duty cycle of entire network is of obvious importance, we also investigate its effects on the duty cycles for individual sites. We display distribution of duty cycle at each site in Figure 14. Aside from the obvious concern, that the overall resiliency of the network is dependent on its component sites, we also want to explore when and which sites produced anomalous images and what could be the probable reason. For reliably identifying anomalous image occurrences at each site in near-real time, this fqi filter may be proven useful for setting up a flag for an instrument malfunctioning if a large number of images are rejected. This information may be useful for the NISP operations team who may be able to identify and troubleshoot the regularly occurring problems in a timely manner. It should be noted that, due to the time discrepancy between data collection and when this test was conducted, any site images which have been excluded at the sites were not included for the purpose of DC analysis. However, as the prior performance comparison showed, we can expect the results to be comparably favorable.

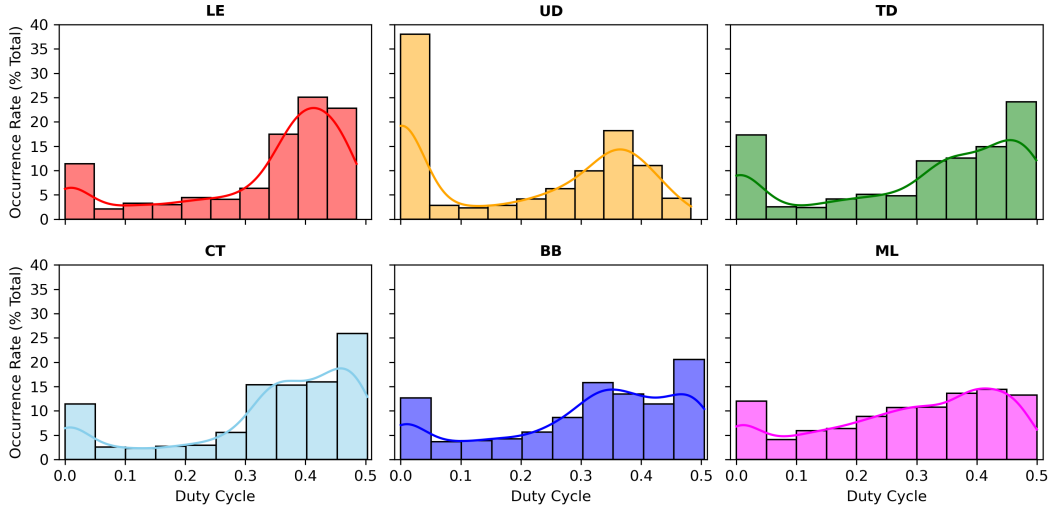


Figure 14: Histogram of the distributions of duty cycles for each site before the *fqi* filter is applied.

### 5.2.1 Learmonth (1e)

Learmonth (1e) had a mean daily duty cycle of 32.47 % for the testing period and as exhibited in Figure 14, for the majority of its days, the duty cycle was between 0.30 and 0.50. The details of total observations and the rejected images are given in Table 4. Overall it collected 853,734 minutes worth of observations, which makes 18.50 % of the total observations for the testing period. Of these total minutes observed by 1e, about one-third observation had no overlap with any other site and more than half were overlapped by one other site. Rest of the observations were gathered with two or three other sites (see Table 3). It is further evident from Figure 9 that the maximum overlap of 1e was with ud, followed by ml, td, bb, and ct, respectively.

From Table 4, it can be seen that of all the images collected by 1e, 624 (0.07 %) images were rejected by the *fqi* filter for being anomalous. This comprises 4.1 % of all the anomalous images found by the *fqi* network. Figure 15 provides number of the monthly rejections each year. We did not find any outstanding events contributing to anomalous images from the 1e site. This is also confirmed by a lack of visible drops in the daily duty cycle plotted in Figure 16. Overall, the application of the *fqi* filter reduced the mean duty cycle of the 1e site from 32.47 % to 32.45 % (Table 5) and did not discard all the observations for a given day.

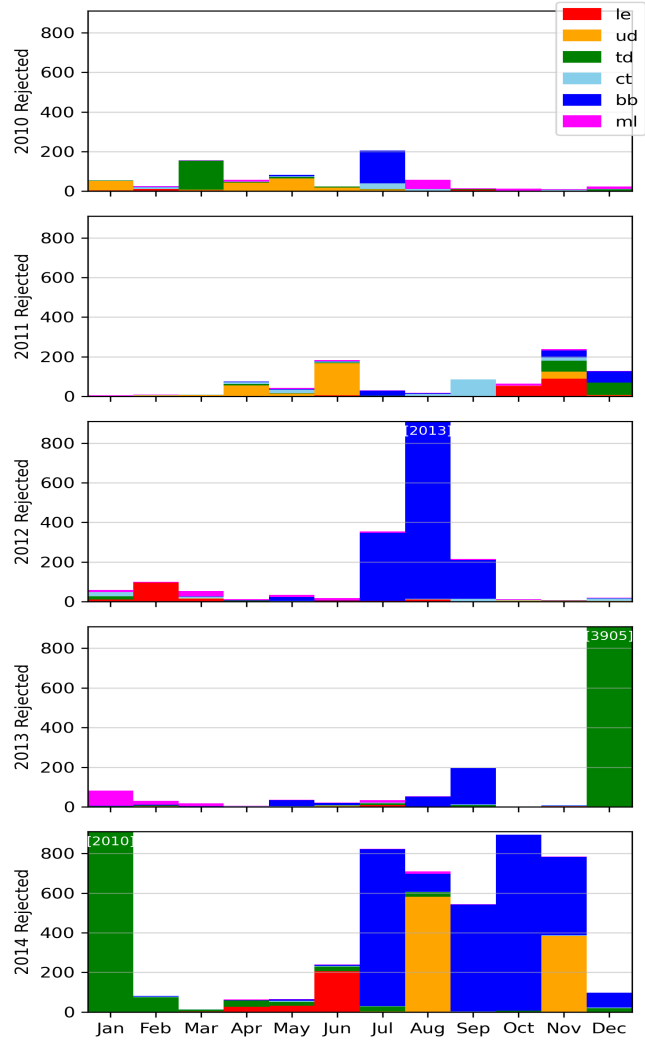


Figure 15: Images rejected by month for the sample period. Site contributions are shown in different colors.



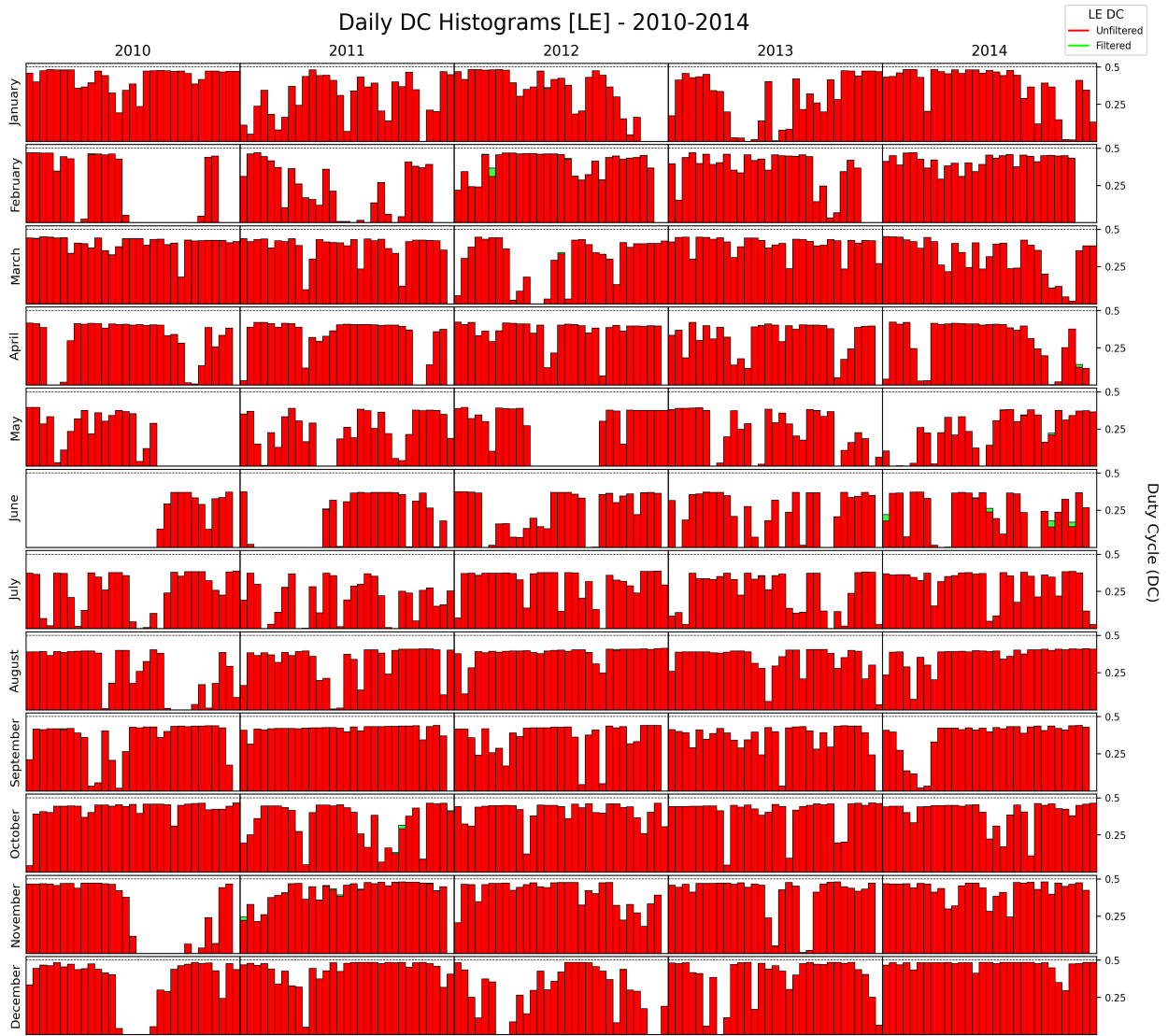


Figure 16: Daily duty cycle for the 1e site for the testing period from 2010 through 2014. The effects of the filter on duty cycle are shown in green. The major drops in duty cycles are seen on one day in February 2012 and several days in June 2014.

### 5.2.2 Udaipur (ud)

As illustrated in Figure 14, while a majority of the non-zero days for the ud site had a duty cycle between 0.3 and 0.5, there are a disproportionate number of days when it did not contribute any observation to the GONG network. This is due to the annual monsoon season, which forces ud site to shutdown for a couple of months every year. These periods are clearly visible in Figure 17. In addition, there was a disruption which lasted for several months from mid 2010 to the beginning of 2011 due to a fire near the ud site. These long outages adversely affected the mean and median duty cycles of ud for the testing period of five years and the mean and median daily duty cycles were 19.80% and 23.85 %, respectively, as summarized in Table 5. Of the 520,582 minutes that the ud site observed (Table 4), 19.05 % were independent without overlapping with any other site. About two-third of ud observations were collected simultaneously with only one other site. There were relatively smaller number of observations that had overlaps with two or more other sites.

Table 5: Duty cycle statistics for filtered vs. unfiltered images for each of the GONG sites during the testing period of five years. The percentage statistics are taken to 2 decimal places as many of the changes are too minute to be seen otherwise.

Site	Filter	Mean DC [%]	Median DC [%]	Standard Deviation [%]	Maximum DC [%]	Minimum DC [%]	# of Days Dropped to 0 % DC
le	Unfiltered	32.47	38.20	14.86	48.50	0	
	Filtered	32.45	38.20	14.87	48.50	0	0
ud	Unfiltered	19.80	23.85	17.27	48.30	0	
	Filtered	19.74	23.85	17.29	48.30	0	6
td	Unfiltered	30.39	35.50	17.12	49.90	0	
	Filtered	30.15	35.45	17.30	49.90	0	6
ct	Unfiltered	33.48	37.60	15.37	50.30	0	
	Filtered	33.46	37.60	15.37	50.30	0	1
bb	Unfiltered	30.48	34.00	15.81	50.50	0	
	Filtered	30.25	33.80	15.88	50.50	0	3
ml	Unfiltered	28.20	31.00	15.00	50.00	0	
	Filtered	28.19	31.00	15.00	50.00	0	0

Details of these overlapping observations are given in Table 3. As illustrated in Figure 9, the ud site most often observed in conjunction with two neighboring sites, i.e. le and td sites.

A total of 1,545 (0.30 %) out of 520,582 images were identified as anomalous by the fqi filter. These made up 9.50 % of the total anomalous images found for the testing period. Figure 15 shows that there were two major periods in August and November in 2014 when most rejections happened. In addition, the filter brought down the duty cycle to zero only for 5 days. While one of these days was during an expected operational period, many of these days were during a low-performance period during the month of August 2014. This is clearly visible in Figure 17 where we display the daily duty cycle for the entire test period. Also, the fqi filter reduced the mean duty cycle of the ud from 19.80 % to 19.74 % but did not affect the median DC. A comparison between the statistics of filtered and unfiltered images is given in Table 5.

### 5.2.3 El Teide (td)

The El Teide (td) site had a mean and median daily duty cycles of 30.39 % and 17.12 %, respectively, for the testing period. As exhibited in Figure 14, the non-zero duty cycle days for the td site were disproportionately distributed towards 0.50 with most of the observing days being reported as above a duty cycle of 0.30 and nearly 25 % of reported days being between 0.45 and 0.50. While td does not experience the seasonal drops in observations similar to ud, it provided good coverage during the months of April to September (Figure 18). As summarized in Table 3, the statistical analysis of the total observations for 799,160 minutes at td suggests that it collected independent observations for 24.05 % of those minutes, and the simultaneous observations with one, two and three other sites for 46.81 %, 23.49 % and 5.64 %, respectively. From Figure 9, we notice that the td site was most often observed with the ct and bb sites.

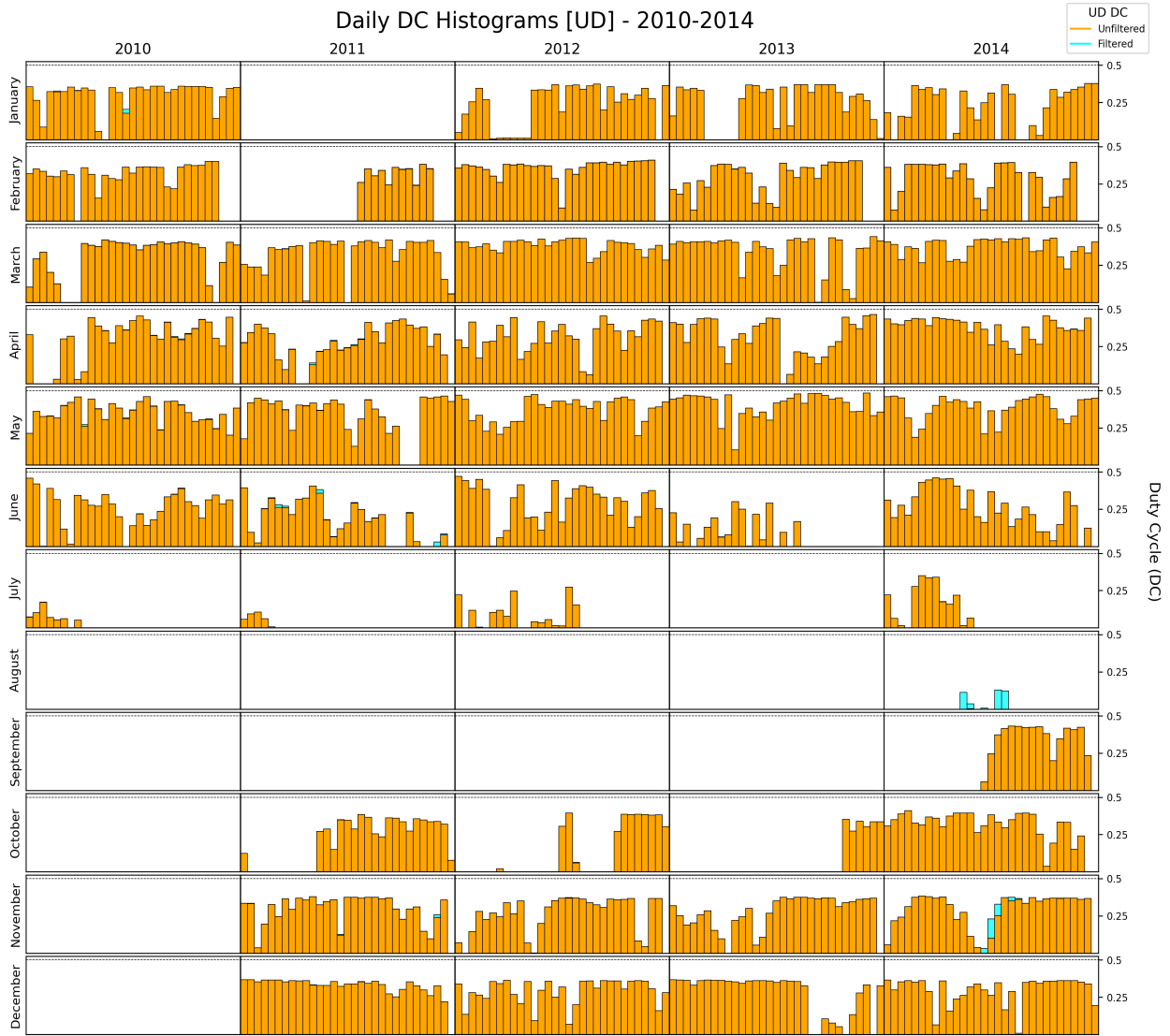


Figure 17: Daily duty cycle for the ud site for the testing period from 2010 through 2014. The effects of the fqi filter on duty cycle are shown in cyan. The major drops in duty cycles are apparent in August 2014 and November 2014.

From Table 4, we note that the implementation of fqi filter rejected 6510 observations. These rejected images accounted for 0.81 % of the total observations at td and 42.51 % of the all anomalous images identified by the fqi network, making the td site the largest contributor of anomalous images for the testing period. As a result of the filter's application, the duty cycles for 6 days completely reduced to 0 (Table 5). It is also observed from Figure 15 that the vast majority of the anomalous images were generated during December 2013 through February 2014. The application of the fqi filter reduced the mean duty cycle of the td site from 30.39 % to 30.15 %, however increased the median duty cycle from 17.12 % to 17.30 %.

#### 5.2.4 Cerro Tololo (ct)

The Cerro Tololo (ct) site had an average daily duty cycle of 33.76 % and a median duty cycle of 37.60 % for the testing period. It is demonstrated in Figure 14 that the non-zero duty cycle days for the ct site were

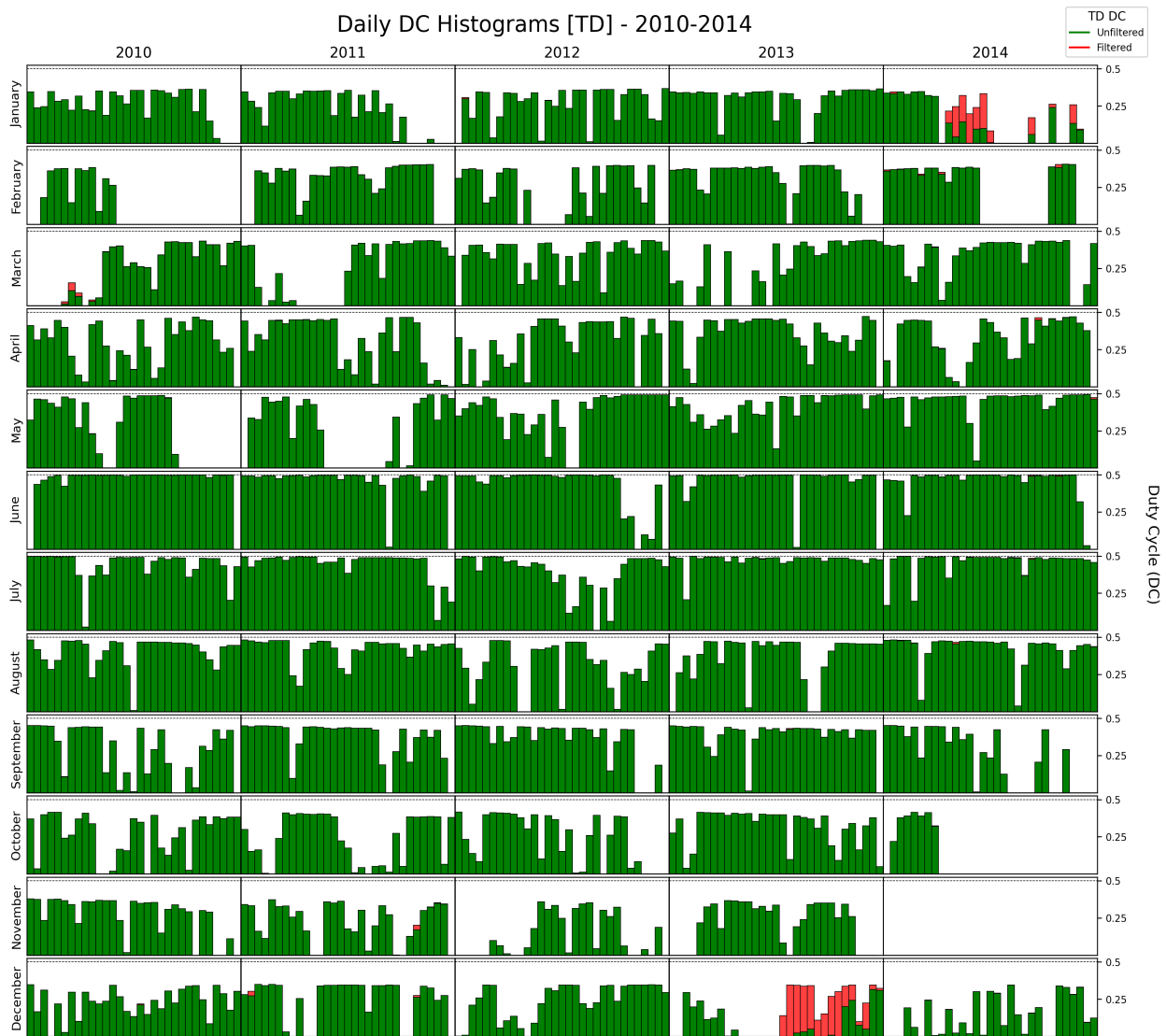


Figure 18: *Daily duty cycle for the td site for the testing period from 2010 through 2014. The effects of the fqi filter are shown in red. The significant drop in duty cycles are apparent in December 2013 and January 2014.*

largely above 0.30 with the majority of days reporting a duty cycle between 0.45 and 0.50. During the testing period, the ct site produced consistent observations but with some notable lapses in September of 2010 and 2011, and February and April of 2012. These gaps are clearly visible in Figure 19. The application of the fqi filter resulted in one day only having its duty cycle reduced entirely to zero. Of the 880,258 images collected at ct, 328 (0.04 %) of those images were identified as anomalous by the neural network (Table 4), making it the lowest contributing site to the anomalous images and accounted for 2.14 % of the total anomalous images found during the testing period.

Of the 880,258 minutes observed by ct during the testing period, it observed independently for 15.57 % of the total minutes. In addition, its observations were overlapped with one, two and three other sites for 44.50 %, 34.71 %, and 5.21% of the total minutes, respectively. It is also seen from Figure 9 that the ct site

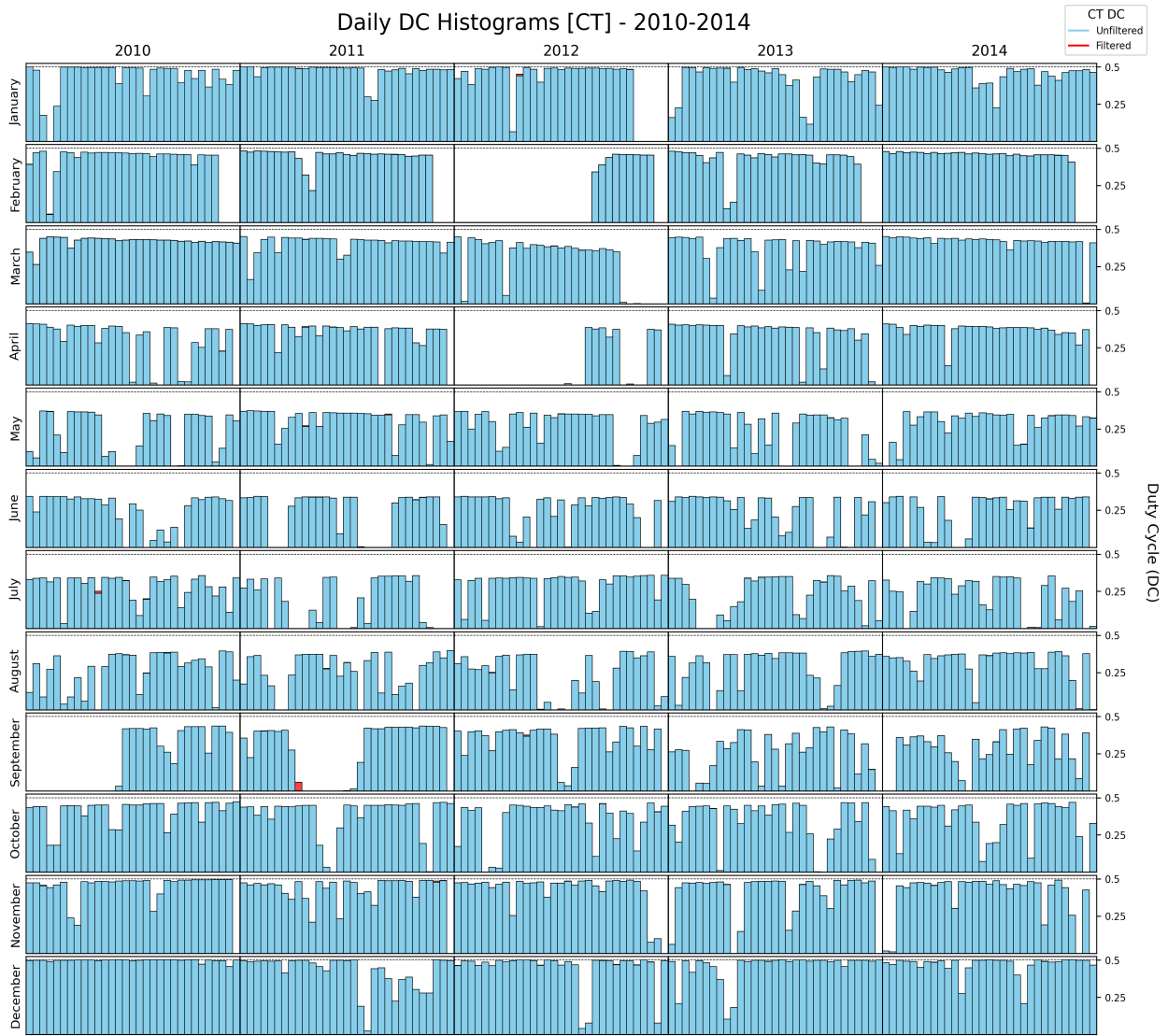


Figure 19: Daily duty cycle for the ct site for the testing period from 2010 through 2014. The effects of the fqi filter are shown in red. The drop in the duty cycle is apparent in on one day September 2011 where the duty cycle dropped to 0.

observed most frequently with the td and bb sites. The application of the fqi filter reduced the mean duty cycle for ct from 33.48 % to 33.46 % and did not affect the median duty cycle.

### 5.2.5 Big Bear (bb)

The average daily duty cycle of Big Bear (bb) site was 30.48 % with a median duty cycle of 34.00 %. As shown in Figure 14, a little over 10 % of the reported days for the testing period were reported as having a duty cycle of 0, while 20 % of the reported days had a duty cycle between 0.45 and 0.50 for the bb site. From Figure 20, we find that the bb site provided relatively good coverage in the months of May and June. Of the 801,498 images from bb, 5974 (0.75 %) of them were rejected for being anomalous (Table 4). This comprised 39.01 % of all anomalous images found for the testing period, making the bb site the second largest contributor of the erroneous images in the network. There were two major events, as shown in Fig-

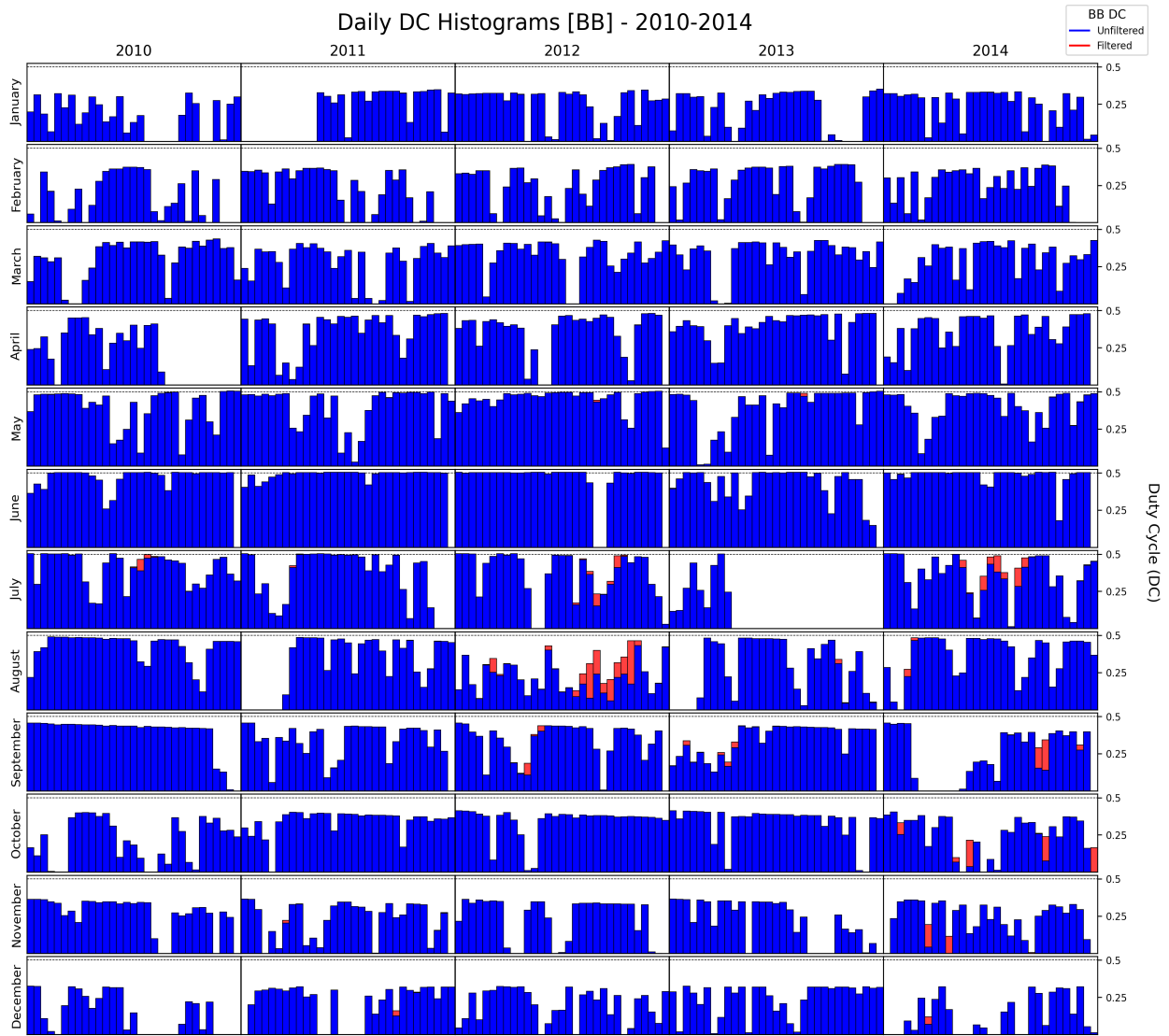


Figure 20: *Daily duty cycle for the bb site for the testing period from 2010 through 2014. The effects of the fqi filter are shown in red. The drops in duty cycles are apparent in July 2010, July – September 2011, August – September 2013 and July – December 2014.*

ure 15, which accounted for the majority of the rejected images from the bb site, the first occurring between July and September of 2012, and the second occurring between July and November of 2014. It is important to mention that the RMS filter became operational in July 2014 as discussed in Section 2.1, but it did not exclude all the anomalous images after its implementation and allowed these images to get ingested in downstream pipelines and related data products. The application of the fqi filter on the bb site images resulted in reducing the duty cycle to 0 for 3 days as listed in Table 5.

Of the 801,498 minutes observed at the bb site during the testing period, it provided independent observations for 10.14 % of those minutes, and was one of two, three and four sites observed at the same time for 40.78 %, 43.31 % and 5.77 % of the total minutes, respectively (Table 3). It is noticed from Figure 9 that the bb site most often observed with the ml and ct sites. Due to the implementation of the fqi filter, the mean

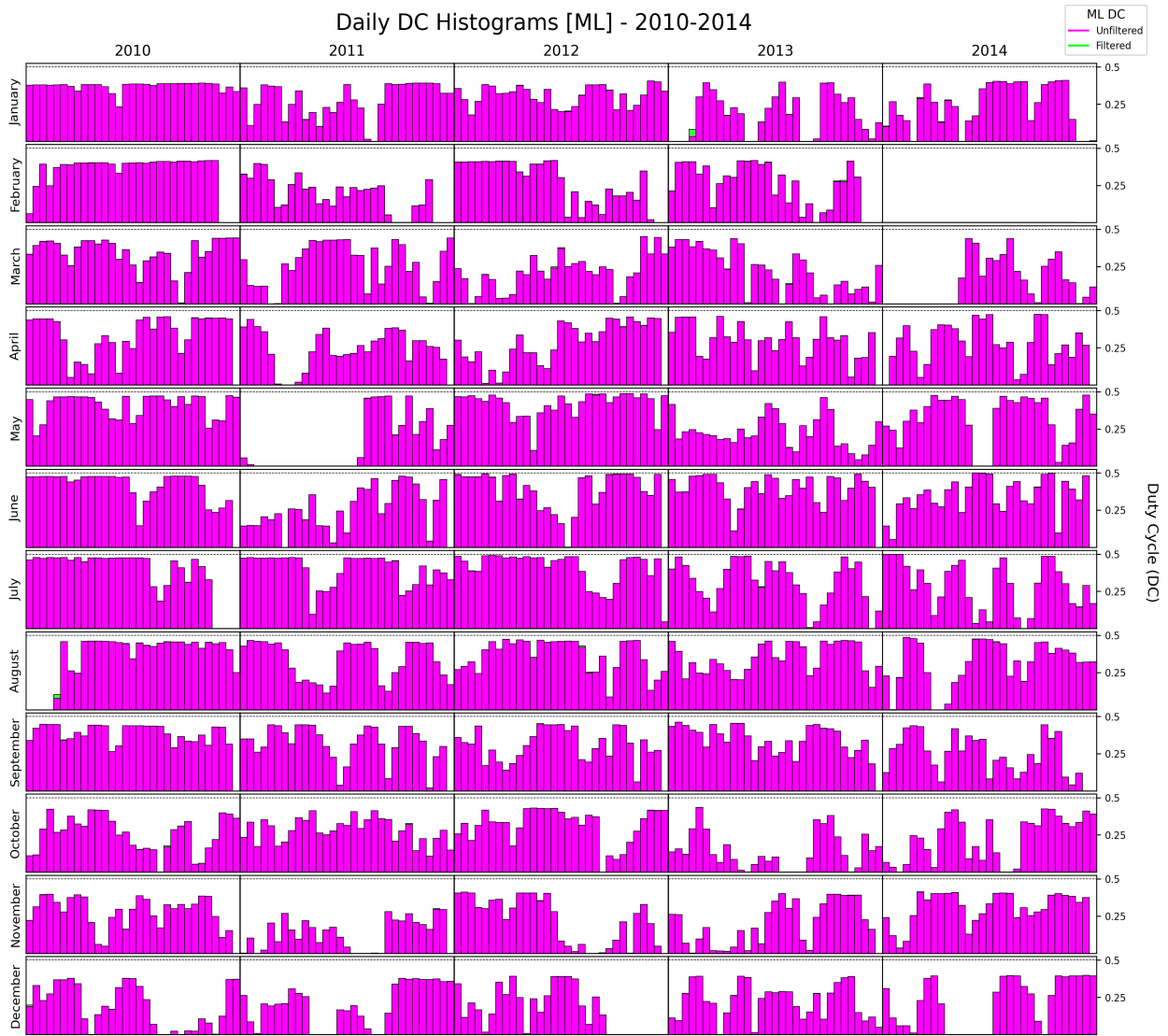


Figure 21: *Daily duty cycle for the m1 site for the testing period from 2010 through 2014. The effects of the filter are shown in green. There were no major drops in the duty cycle due to the applications of fqi filter, however one can notice changes in August 2010 and January 2013.*

daily and median duty cycles were reduced from 30.48 % to 30.25 %, and 34.00 % to 33.80 %, respectively.

### 5.2.6 Mauna Loa (m1)

The last site to be considered is Mauna Loa (m1). The daily duty cycles for the testing period are shown in Figure 21. It had an average daily duty cycle of 28.20 % with a median duty cycle of 31.00 %. As presented in Figure 14, the m1 site had a DC of 0 for a little over 10 % of the reported days with the majority of its days being reported above 0.30. For the testing period, m1 had relatively consistent performance except for two large outages occurring in May of 2011, and February/March of 2014. Additionally, observations collected during October through December were relatively less due to its geographic location. Nevertheless, this site has crucial importance because of its location in the middle of the Pacific ocean and provides vital support for overlapping observations to its neighboring sites. Of the 741,636 images collected by the m1 site, 425

(0.06 %) of these images were identified as anomalous by the network (Table 4). These comprised 2.78 % of all the anomalous images found during the testing period. These images were rejected during multiple small epochs as seen in Figure 15 and the implementation of the *fqi* filter did not result in the duty cycle of any day being brought down to zero for this site (Table 5).

The *m1* site collected observations for 741,636 minutes in total and 11.07 % of them were independent observations. As given in Table 3, 45.9 % of its observations were overlapped by one other site, 36.8 % and 6.2 % were overlapped with two and three other sites, respectively. An examination of Figure 9 reveals that the *m1* site had maximum overlap with *bb* and then with *ct* sites. Due to the implementation of the *fqi* filter on *m1* site images, the mean daily duty cycle reduced from 28.20 % to 28.19 %, and the median duty cycle remained unchanged at 31.00 %.

## 6 Case Studies for Farside Map Effects

### 6.1 Noise Reduction in Merged *fqi* Images

In order to illustrate the impact of single image exclusions on merged *fqi* images, we have selected three examples from the database where *fqi* filter identified erroneous site images. These were taken at 20100805t2005, 20141031t1742, and 20140625t0502, representing examples of some of the common types of anomalies seen in the site *fqi* images.

1. **20100805t2005** The first example includes a fully contaminated site image from *m1* shown in the top left panel of Figure 22. These types of artifacts often start at one corner of the site images and progress until they cover the entire observed hemisphere. There were two more sites, *ct* and *bb*, that provided observations at the same time. When anomalous *m1* image was combined with good *ct* and *bb* images, it contaminated the merged product (*mr fqi*) shown in the top right panel of the same figure, completely erasing the valuable Doppler signals. The application of *fqi* filter correctly identified *m1* image as anomalous and rejected its inclusion in the merged image. By eliminating the anomalous image, shown in the bottom row, we obtained a good quality merged *mr f6i* image that can be reliably used in the farside pipeline or any other scientific study.

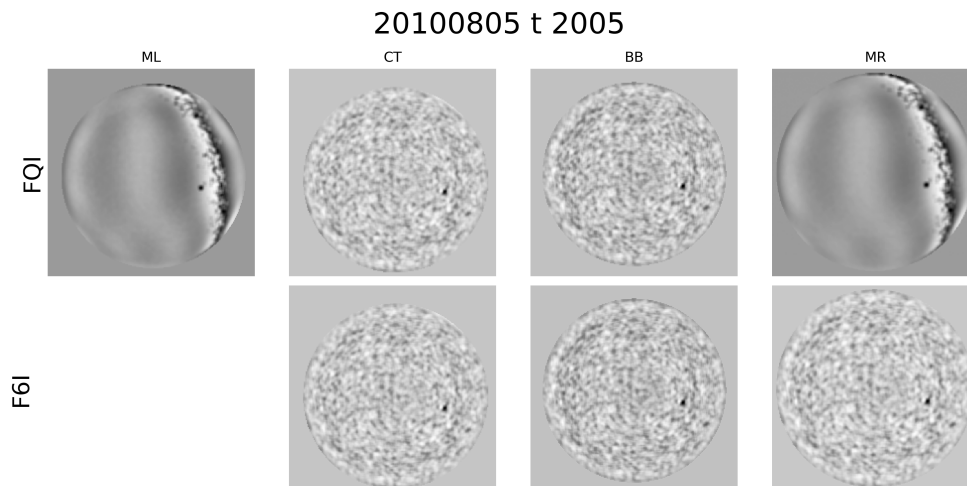


Figure 22: *Effects of including (top row) and excluding (bottom row) anomalous site data in merged fqi image for 20100805t2005. The m1 image was identified by the fqi filter network. The merged images without contaminated data are labeled F6I instead of FQI for the traditional pipeline.*



2. **20141031t1742** Another example shows a corrupted image with the so-called dinner plate artifacts. These artifacts are a ring pattern which often appear from time to time and disappear suddenly. This is a more common type of artifact seen in the site images that may affect either full or half disk. Its origin is not yet fully understood, however it adversely affects the quality of the merged images. As visible in Figure 23, the rings in bb image translated to the final merged image, introducing subdued noise in the ring pattern. The fqi filter eliminated the bb image from the list of images to be merged for this timestamp, thus allowed us to produce a merged image without any artifact.

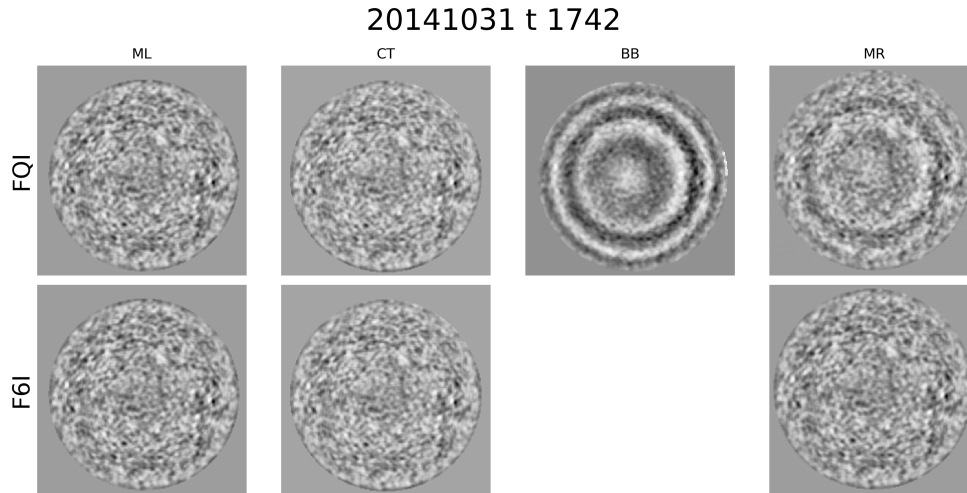


Figure 23: Same as Figure 22, but for the merged image at 20141031t1742 excluding the bb image.

3. **20140625t0502** The third example of the contaminated site fqi image includes a pressed plate artifact. These often appear suddenly and can completely erase the data from good images when included in merged (mrfqi) images. Due to their significant impact, it is important for the network to identify and exclude these images, as they can introduce significant noise from a single point. By excluding such artifacts, we are able to produce clean mrfqi images. Figure 24 shows an image from 1e with this type of artifact. Although ud image at the same time was absolutely normal, the inclusion of abnormal 1e produced corrupted merged image. The result of using the fqi filter for this timestamp as shown in the bottom row.

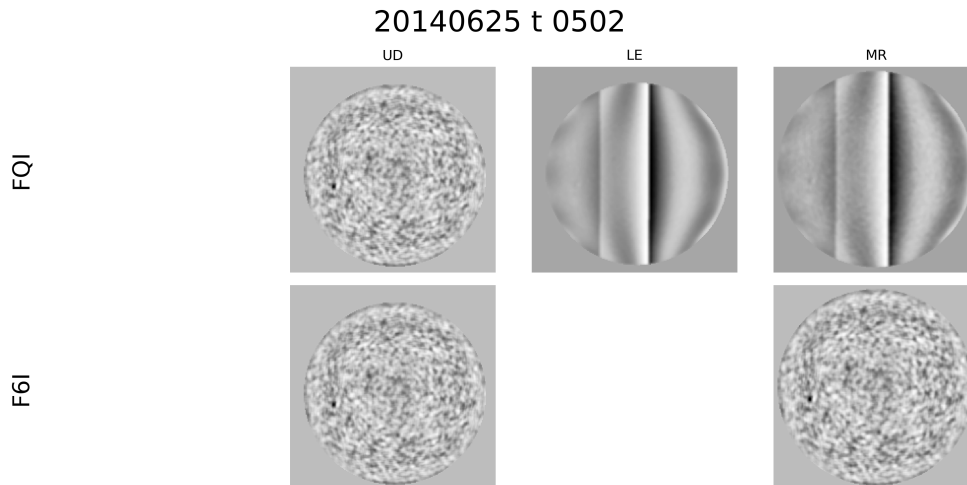


Figure 24: Same as Figure 22, but for the merged image at 20140625t0502 excluding the 1e image.

## 6.2 Noise Reduction in Farside Maps

While the effects of the *fqi* network are not visible in every farside map, it is more prominent during the periods when the *fqi* filter identified several anomalous images and the seismic maps were extremely noisy. The network filter significantly reduced the noise in exchange of a small drop in duty cycle. In order to demonstrate the efficacy of the small number of images excluded by the *fqi* filter, we discuss three examples of the farside seismic maps at 1200 UT on 20111126, 20111202, and 20120206. These maps are calculated using 1440 merged images. The statistical effects of the implementation of the filter on farside maps are summarised in Table 6.

Table 6: Summary statistics for production (*fqm*) and cleaned (*f6m*) sample phase shift maps.

Date	Map Type	Number of Input Images	Duty Cycle	Phase Shift (radians)			Standard Deviation	Signal-to-noise Ratio
				Min	Max	Mean		
20111126t1200	<i>fqm</i>	1395	0.97	-3.34	2.738	0.025	0.401	0.063
	<i>f6m</i>	1383	0.96	-0.206	0.133	-0.017	0.051	0.338
20111202t1200	<i>fqm</i>	1430	0.99	-0.691	0.973	0.027	0.173	0.155
	<i>f6m</i>	1430	0.99	-0.2	0.134	-0.015	0.050	0.307
20120206t1200	<i>fqm</i>	1239	0.86	-0.594	1.064	0.049	0.155	0.318
	<i>f6m</i>	1235	0.86	-0.241	0.185	-0.003	0.055	0.054

1. **20111126t1200** The first example is the farside seismic map with the timestamp of 20111126t1200. This map was initially computed with 44 contaminated site *fqi* images from the *td* site and the exclusion of these images resulted in the loss of 12 merged *fqi* images. The unfiltered images introduced significant noise in the farside as displayed in the left panel of Figure 25, however the barring these images not only resulted in a striking visual improvement in clarity (right panel in the same figure), but also a statistically noticeable improvement. From Table 6, it can be seen that the standard deviation dropped from 0.401 to 0.051 radians, indicating a far more narrow band of readings. Additionally, the minimum and maximum values changed from  $-3.34$  to  $-0.206$  radians and  $2.738$  to  $0.133$  radians, respectively. The values with unfiltered images were excessive and had no physical significance. The mean being reduced from  $0.025$  to  $-0.017$  radians is more in line with the predominantly negative phase shift values that we expect to see from the helioseismic maps. The signal-to-noise ratio also increased significantly from  $0.063$  to  $0.338$ .

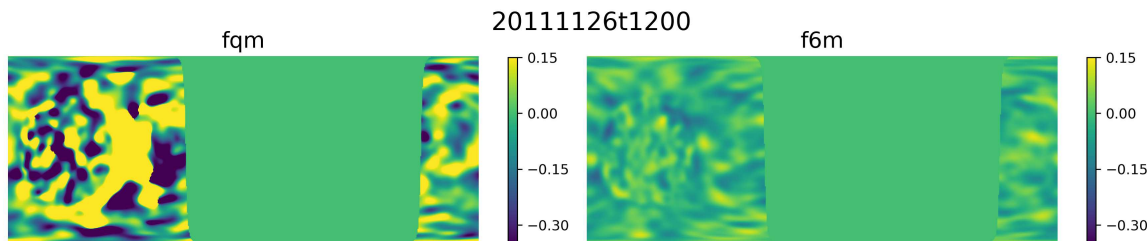


Figure 25: Farside phase-shift maps using contaminated (*fqm*) and cleaned (*f6m*) merged site images for 20111126t1200. This *fqm* map was contaminated by anomalous site images which were identified by the *fqi* network. The scale depicts the phase-shift values in radians with the frontside of the sun masked out in green.

2. **20111202t1200** The second example is the farside seismic map at 20111202t1200. This map, displayed in Figure 26, also has visible and statistical improvements when filtered images are used. The contaminated 83 images from the *le* and *ml* sites produced a noisy map as shown in the left panel and the exclusion of which did not reduce the number to good merged *fqi* images. Similar to previous case, the exclusion of erroneous images dropped the standard deviation from 0.173 to 0.050 radians, indicating a more accuracy in the farside map. The minimum and maximum values also modified from  $-0.691$  to  $-0.200$  radians and from 0.973 to 0.134 radians, respectively. The modified values are within the expected range of phase shifts and the mean changed from 0.027 to  $-0.015$  radians. Again, this shows the predominantly negative phase shift values that we expect from the helioseismic maps. The signal-to-noise ratio again doubled from 0.155 to 0.307 (Table 6).

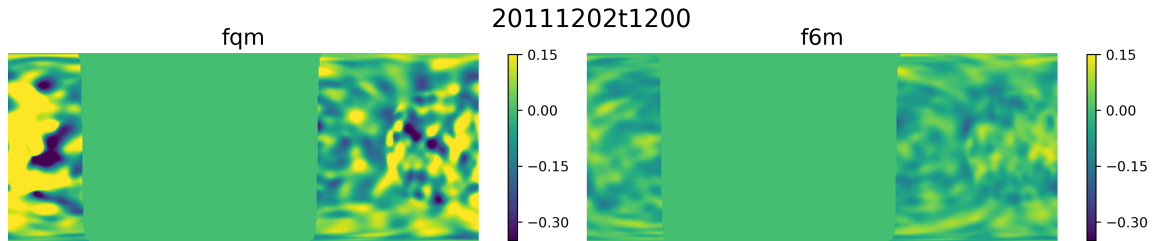


Figure 26: Same as Figure 25, but for 20111202t1200.

3. **20120206t1200** The last illustration is the map from 20120206t1200 shown in Figure 27. This is the most interesting map because it appears to be a seismic maps for the abnormally high-activity period at first glance. However, a close inspection of the images used to produce this map reveals that the *fqm* map was contaminated by 44 anomalous images from the *td* site and the exclusion of these images resulted in the loss of four merged *fqi* images. The result of the exclusion of these images produced a map with far more consistency that lost large phase-shift values which it initially displayed. The maximum phase-shift value was reduced from 1.064 to 0.185 radians, while it's minimum was also brought in from  $-0.594$  to  $-0.241$  radians. This also reduced the standard deviation from 0.155 to 0.055 radians. However, the signal-to-noise ratio reduced from 0.318 to 0.54 (see Table 6). This map is an excellent example of the necessity of the *fqi* neural network filter. While it clearly had anomalous images as its components, statistical analysis of the map itself would not have revealed it to be faulty.

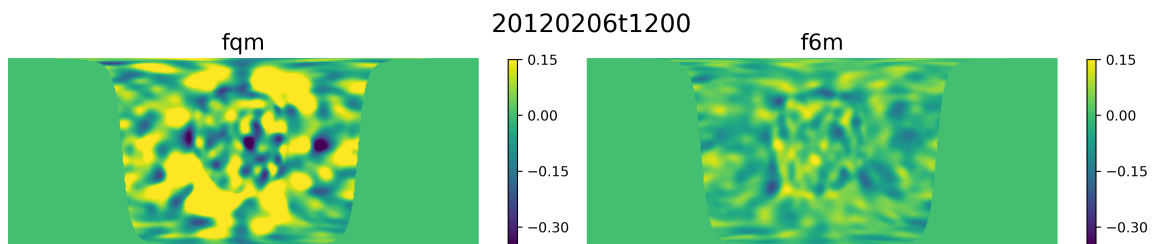


Figure 27: Same as Figure 25, but for 20120206t1200.

These examples show the adaptability of the discriminatory network and its ability to provide quality assurance and control measures in a variety of statistically novel situations. We further checked the credibility of *f6m* maps by comparing them with those produced by both Helioseismic and Magnetic Imager farside pipelines based on the helioseismic holography (a similar technique used by the NISP farside pipeline) and

the time-distance techniques, and found a close correspondence between them. It should be noted that HMI has several advantages being in space and the images are less prone to any artifacts.

## 7 Network Updates and Improvements

There are two methods through which the performance of the current fqi network can be improved. The first of these is re-training, during which the network weights would be updated through training on newly labeled data. The second of these methods is a network re-design, in which the current network architecture would be replaced by an updated architecture design.

### 7.1 Re-Training

In order to make network re-training and updates as easy as possible, a Jupyter notebook has been created and uploaded to the NISP Data Center's GitLab under the name `fqi_Network_Retraining`. While the README file in the repository contains specific step-by-step instructions, the seven steps of implementing new weights are also outlined in this document. This explanation will assume that the user has already locally cloned the retraining repository and has the appropriate environmental packages installed.

1. **Load the newly labeled data:** Load the newly labeled normal and anomalous data into `.lst` files so that it can be included in the training regime. The notebook assumes that a number of images which have not been previously used in training the network have been manually labeled into Normal and Anomalous classifications. Directories to the new anomalous and normal data should be collected and stored in two files named `anom_new.lst` and `norm_new.lst` respectively. These should be stored in the `data/new_data/` subdirectory of the retraining regime. The re-training regime is designed to take directories from the `oQR` directory. These files will be combined with old training data in order to update the fqi network weights. At the end of the re-training process, there is an opportunity to combine the new and old datasets so that the process can be repeated as many times as possible. The more data that is added to the re-training data, the better the updated results will be. It is recommended that re-training be conducted with at least a 1 % expansion on the current dataset, but more is always better.
2. **Ensure old network weights are available:** Ensure that the old network weights are available. These are stored as the `.pt` file from which the fqi Network loads its weights in deployment. This file needs to be stored in the `params/old_params` directory, and does not need to have a specific name as the re-training regime will pick up any file stored in this directory.
3. **Run the training regime:** Run the training regime. So long as the environment is properly set up, this should only entail running the Jupyter notebook. All of the required packages are stored in the first cell of the notebook.
4. **Evaluate the network performance:** Evaluate the network performance. New network performance can be seen at the bottom of the Jupyter notebook, where the network statistics, confusion matrix, and a visual sample of said matrix results can be seen. If there are patterns of images which the network has trouble identifying, then the suggestion is to add more of those images to the new dataset and re-run the program before re-deploying. It should be noted that the goal is not absolute perfection, as this would be indicative of over-training and may, counter-intuitively, lead to poor performance in deployment, but rather the minimization of false-negative results so as to avoid the taking of large sections of high-activity periods.

5. **Make adjustments to the network:** If network performance is not sufficiently improved, make adjustments. These include gathering more data, adjusting hyperparameters such as learning rate, batch size, or number of epochs, and double-checking the sample for incorrectly labeled images.
6. **Update the network weights:** Update the network weights. The new weights will be stored in `params/new_params/Spotty_Net_Update.pt`. This file can be renamed in accordance with the decided naming conventions for the fqi Network weights and replaced in the appropriate file.
7. **Merge the new and old data:** Once the network performance is satisfactory, merge the new and old data to prepare the re-training regime for future updates. Replace the files `anom_new.lst`, `anom_old.lst`, `norm_new.lst`, `norm_old.lst` with the files found under `new_data` of the same name. After this is done, you may delete the `new_data` directory.

## 7.2 Network Design Updates

The second option for improving the performance of the fqi network would be a fundamental re-design. While the specifics of a network re-design are outside the scope of this discussion, it would broadly entail re-training on the existent dataset with an updated network architecture. Such a campaign would be more intensive than a simple weight update and thus should be considered carefully. However, given the rapid pace of improvement in the field of machine learning, new network designs may provide significant performance improvements. New strategies such as GANs or other fundamental improvements in architectures should be considered for performance improvements (Di Mattia and et al., 2021; Xia and et al., 2022).

In the event the the network is re-designed, both the weights file and the network architecture files should be replaced. Once these are replaced and the associated calls in the code have been updated, that should be sufficient to run the network on the new architectures.

## 8 Summary

The binary classification network developed and analyzed in this report has proven to be a highly effective method for reducing the impact of anomalous site-images on the noise of NISP’s farside helioseismic maps. This neural network-based classification methodology is resilient enough to accurately identify anomalous images despite their statistical variability. It is also more selective than the current thresholding technique, preventing the exclusion of large amounts of valuable high-activity data. The spacial variability of the GONG network’s observational sites also provide adequate resiliency to compensate for the majority of anomaly production events at any single site. When we also consider the potential for re-training and re-design, the fqi network has proven to be invaluable and adaptable QAQC methodology for the NISP farside program. We hope that the performance and adaptability of this new methodology serves as a proof of concept for the implementation of similar measures for other NISP data products in the future.

**Acknowledgements:** This work utilises GONG data obtained by the NSO Integrated Synoptic Program, managed by the National Solar Observatory, which is operated by the Association of Universities for Research in Astronomy (AURA), Inc. under a cooperative agreement with the National Science Foundation (NSF) and with contribution from the National Oceanic and Atmospheric Administration (NOAA). The GONG network of instruments is hosted by the Big Bear Solar Observatory, High Altitude Observatory, Learmonth Solar Observatory, Udaipur Solar Observatory, Instituto de Astrofísica de Canarias, and Cerro Tololo Interamerican Observatory. This work is partially supported by the Windows On the Universe (WoU) grant funded by NSF.

## References

- Braun, D. C. and Lindsey, C. (2001). Seismic Imaging of the Far Hemisphere of the Sun. *Astrophys. J. Lett.*, 560(2):L189–L192. <https://ui.adsabs.harvard.edu/abs/2001ApJ...560L.189B>.
- Di Mattia, F. and et al. (2021). A survey on gans for anomaly detection. *arXiv*.
- Divovic, P., Obradovic, P., and Mistic, M. (2021). Balancing imbalanced datasets using generative adversarial neural networks. In *2021 29th Telecommunications Forum (TELFOR)*. <https://www.semanticscholar.org/paper/Balancing-Imbalanced-Datasets-Using-Generative-Divovi%2C4%27-Obradovi%20c/4ab76dfc9afacfe319c26221ccfa33615483defd>.
- Dubey, A. K. and Jain, V. (2019). Comparative study of convolution neural network’s relu and leaky-relu activation functions. *Lecture Notes in Electrical Engineering*, pages 873–880. [https://link.springer.com/chapter/10.1007/978-981-13-6772-4\\_76](https://link.springer.com/chapter/10.1007/978-981-13-6772-4_76).
- González Hernández, I., Díaz Alfaro, M., Jain, K., Tobiska, W. K., Braun, D. C., Hill, F., and Pérez Hernández, F. (2014). A Full-Sun Magnetic Index from Helioseismology Inferences. *Solar Phys.*, 289(2):503–514. <https://ui.adsabs.harvard.edu/abs/2014SoPh..289..503G>.
- González Hernández, I., Hill, F., and Lindsey, C. (2007). Calibration of Seismic Signatures of Active Regions on the Far Side of the Sun. *Astrophys. J.*, 669(2):1382–1389. <https://ui.adsabs.harvard.edu/abs/2007ApJ...669.1382G>.
- González Hernández, I., Hill, F., Scherrer, P. H., Lindsey, C., and Braun, D. C. (2010). On the success rate of the farside seismic imaging of active regions. *Space Weather*, 8(6):06002. <https://ui.adsabs.harvard.edu/abs/2010SpWea...8.6002G>.
- González Hernández, I., Jain, K., Tobiska, W. K., and Hill, F. (2011). The far-side solar magnetic index. In *GONG-SoHO 24: A New Era of Seismology of the Sun and Solar-Like Stars*, volume 271 of *Journal of Physics Conference Series*, page 012028. <https://ui.adsabs.harvard.edu/abs/2011JPhCS.271a2028G>.
- Harvey, J. W., Hill, F., Hubbard, R. P., Kennedy, J. R., Leibacher, J. W., Pintar, J. A., Gilman, P. A., Noyes, R. W., Title, A. M., Toomre, J., Ulrich, R. K., Bhatnagar, A., Kennewell, J. A., Marquette, W., Patron, J., Saa, O., and Yasukawa, E. (1996). The Global Oscillation Network Group (GONG) Project. *Science*, 272(5266):1284–1286. <https://ui.adsabs.harvard.edu/abs/1996Sci...272.1284H>.
- Hill, F. (2018). The Global Oscillation Network Group Facility—An Example of Research to Operations in Space Weather. *Space Weather*, 16(10):1488–1497. <https://ui.adsabs.harvard.edu/abs/2018SpWea..16.1488H>.
- Hill, F., Fischer, G., Forgach, S., Grier, J., Leibacher, J. W., Jones, H. P., Jones, P. B., Kupke, R., Stebbins, R. T., Clay, D. W., Ingram, R. E. L., Libbrecht, K. G., Zirin, H., Ulrich, R. K., Webster, L., Hieda, L. S., Labonte, B. J., Lu, W. M. T., Sousa, E. M., Garcia, C. J., Yasukawa, E. A., Kennewell, J. A., Cole, D. G., Zhen, H., Su-Min, X., Bhatnagar, A., Ambastha, A., Al-Khashlan, A. S., Abdul-Samad, M.-S., Benkhaldoun, Z., Kadiri, S., Sánchez, F., Pallé, P. L., Dhalde, O., Solis, H., Saá, O., and González, R. (1994). The Global Oscillation Network Group Site Survey - Part Two. *Solar Phys.*, 152(2):351–379. <https://ui.adsabs.harvard.edu/abs/1994SoPh..152..351H>.
- Hughes, A. L. H., Jain, K., Kholikov, S., and the NISP Solar Interior Group (2016). GONG ClassicMerge: Pipeline and Product. *arXiv e-prints*, page arXiv:1603.00836. <https://ui.adsabs.harvard.edu/abs/2016arXiv160300836H>.

- Jain, K., Lindsey, C., Adamson, E., Arge, C. N., Berger, T. E., Braun, D. C., Chen, R., Collado-Vega, Y. M., Dikpati, M., Felipe, T., Henney, C. J., Hoeksema, J. T., Komm, R. W., Leka, K. D., Marble, A. R., Martinez Pillet, V., Miesch, M., Nickisch, L. J., Pevtsov, A. A., Pizzo, V. J., Tobiska, W. K., Tripathy, S. C., and Zhao, J. (2023). Seismic Monitoring of the Sun’s Far Hemisphere: A Crucial Component in Future Space Weather Forecasting. In *Bulletin of the American Astronomical Society*, volume 55, page 189. <https://ui.adsabs.harvard.edu/abs/2023BAAS...55c.189J>.
- Jain, K., Lindsey, C., and Tripathy, S. C. (2021a). What is Exceptional about Solar Activity in the Early Phase of Cycle 25? *Research Notes of the American Astronomical Society*, 5(10):253. <https://ui.adsabs.harvard.edu/abs/2021RNAAS...5..253J>.
- Jain, K., Tripathy, S. C., Hill, F., and Pevtsov, A. A. (2021b). Continuous Solar Observations from the Ground-Assessing Duty Cycle from GONG Observations. *Publications of the Astronomical Society of Pacific*, 133(1028):105001. <https://ui.adsabs.harvard.edu/abs/2021PASP..133j5001J>.
- Ketkar, N. and Moolayil, J. (2021). Automatic differentiation in deep learning. In *Deep Learning with Python*, pages 133–145. [https://link.springer.com/chapter/10.1007/978-1-4842-5364-9\\_4](https://link.springer.com/chapter/10.1007/978-1-4842-5364-9_4).
- Khan, A. S., Ahmad, Z., Abdullah, J., and Ahmad, F. (2021). A spectrogram image-based network anomaly detection system using deep convolutional neural network. *IEEE Access*, 9:87079–87093.
- Khan, F. (2019). How to apply weighted loss to a binary segmentation problem? <https://discuss.pytorch.org/t/how-to-apply-weighted-loss-to-a-binary-segmentation-problem/35317>.
- Korkmaz, S. (2020). Deep learning-based imbalanced data classification for drug discovery. *Journal of Chemical Information and Modeling*, 60(9):4180–4190.
- Lindsey, C. and Braun, D. C. (2000). Seismic Images of the Far Side of the Sun. *Science*, 287(5459):1799–1801. <https://ui.adsabs.harvard.edu/abs/2000Sci...287.1799L>.
- Shuman, K. (2023). Good-bad-image-gui. <https://nispdc.nso.edu/gitlab/kshuman/good-bad-image-gui>.
- Toner, C. G., Haber, D., Corbard, T., Bogart, R., and Hindman, B. (2003). An image merge for GONG+. In Sawaya-Lacoste, H., editor, *GONG+ 2002. Local and Global Helioseismology: the Present and Future*, volume 517 of *ESA Special Publication*, pages 405–407. <https://ui.adsabs.harvard.edu/abs/2003ESASP.517..405T>.
- Vailaya, A., Jain, A., and Zhang, H. J. (1998). On image classification: City images vs. landscapes. *Pattern Recognition*, 31(12):1921–1935. <https://www.semanticscholar.org/paper/On-image-classification1f93539714f512e437c45a7a9d0a6d5a7675e>.
- Wang, Y. (2022). Deep convolutional neural networks for facial expression recognition. *International Journal of Neural Network*, 3(4). <https://scholar-press.com/papers/1049>.
- Xia, X. and et al. (2022). Gan-based anomaly detection: A review. *Neurocomputing*, 493:497–535.
- Yal, O. G. (2023). Image classification in 10 minutes with mnist dataset. <https://towardsdatascience.com/image-classification-in-10-minutes-with-mnist-dataset-54c35b77a38d>.