Interior Exploration Using Seismic Investigations, Geodesy, and Heat Transport (InSight) Mission

Insight Fluxgate Magnetometer (IFG)

InSight IFG Data Calibration Description

Prepared by

Steven Joy (in-flight calibration lead)

Katheryn Rowe (ground calibration lead)

Version 1.0

Contents

1 Introduction 5

1.1 Document Change Log 5

1.2 Purpose 5

2 Ground Calibration 6

3 In-flight Calibration 11

4 Data Artifacts 15

4.1 Single Point Spikes 15

4.2 Square Wave Steps 16

4.3 Square Wave Steps with Bounding Fluctuations 19

4.4 Irregular steps, ramps, etc. between 11 and 12.5 hours MLST 21

4.5 Other artifacts 22

4.6 Artifacts in the high time resolution data 23

5 Data versions and Documentation updates 26

Appendix 1: IFG Sensor Temperature 27

Appendix 2: IFG Electronics Temperature 28

Appendix 3: Fixed Solar Array Currents 29

Appendix 4: Total Solar Array Current 30

Table 1: Document change log 5
Table 2: Scale Factors 6
Table 3: IFG Performance (24 bit digitization) 6

Figure 1: Noise Levels below 0.3 nT (shown in yellow) 7
Figure 2: Sensor stability over 24 hours 7
Figure 3: Transfer Function (between DC-10 Hz) 8
Figure 4: Transfer function (between DC-120Hz) 8
Figure 5: Electronics thermal stability - Offset temperature dependence 9
Figure 6: Electronics thermal stability - Gain temperature dependence 9
Figure 7: Sensor head thermal stability - Offset temperature dependence 10
Figure 8: Sensor head thermal stability - Gain temperature dependence 10
Figure 9: Early landed IFG data with ground calibration applied 11
Figure 10: Diurnal variation in IFG temperatures and solar array currents 11
Figure 11: IFG Coherence with temperature and fixed solar array currents 12
Figure 12: IFG data after decorrelation applied 13
Figure 13: IFG data after decorrelation on a 75 nT scale 13
Figure 14: Single point spike examples 15
Figure 15: Examples of square wave steps 17
Figure 16: Blow up of the termination of a selected square wave step 18
Figure 17: Square wave steps with bounding fluctuations 20
Figure 18: Irregular steps, ramps, and other fluctuations between 11 and 12.5 hours MLST 21
Figure 19: Various other types of artifacts 22
Figure 20: High time resolution toggling near a baseline step 24
Figure 21: Polynomial fit to the IFG sensor temperature data 27
Figure 22: IFG Electronics temperature data and fits 28
Figure 23: Fixed solar array current shape model, Dec 14-31, 2018 29
Figure 24: Fixed solar array current and models, January 2019 30
Figure 25: Total Solar Array Current on two days in March, 2019 31
Figure 26:Model of the total solar array current 32
Figure 27: Total solar array current data and model 32

1 Introduction

1.1 Document Change Log

Table 1: Document change log

Version	Change	Date	Affected portion
1.0	Initial draft	April 29, 2019	All

1.2 Purpose

This document provides the details of the IFG data calibration process and examples of the various types of issues that have been identified in the raw data, and the data after the ground calibration has been applied. This document provides much more detail than the IFG Archive SIS, although some of the contents of the SIS are repeated here. The purpose of the SIS is to provide a description an overview of the dataset, the data processing, and the structure of the archive. This document focuses exclusively on the calibration process and processing. For details about the archive organization, contents, naming conventions, etc. the reader is referred to the IFG archive SIS.

IFG data are processed by first applying the ground calibration which results converts the raw data numbers to nanoTesla and orthogonalizes the sensor data giving data in the instrument frame. The resulting data show large diurnal variations that are well correlated with the instrument temperatures and the solar array currents. The in-flight calibration first removes the component of the diurnal variation that is correlated with these spacecraft sources. Once the low frequency corrections have been applied, the procedure is to identify high frequency variations and remove them when possible.

2 Ground Calibration

Prior to delivery to the project for spacecraft integration, the IFG instrument response was characterized over a range of temperatures similar to the expected Martian environment that could be achieved in the laboratory setting. The following tables and charts summarize the results of those efforts.

Table 2: Scale Factors

Data Type	Scaling	Units
X-Axis	Divide by 145.6	nT
Y-Axis	Divide by 141.4	nT
Z-Axis	Divide by 141.7	nT
HK0 (+8V)	Divide by -181.23, add 20.91	V
HK1 (AGND)	Divide by -64, add 38.31	V
HK2 (SH Temp)	$9E\text{-}05x^2\text{-}0.576x+803.43$	ºC
HK3 (EU Temp)	Divide by -7.3, add 333.5	ºC
HK4 (+13V)	Divide by -219.3, add 21.8	V

Scale factors are used to convert engineering data number to physical units.

Table 3: IFG Performance (24 bit digitization)

Dynamic Range of X-Axis	±20,600nT
Dynamic Range of Y-Axis	±21,200nT
Dynamic Range of Z-Axis	±21,200nT
Offset of X-Axis @ 24°C	-71.6nT
Offset of Y-Axis@ 24°C	-57.3nT
Offset of Z-Axis @ 24°C	-0.5nT
Data Rate of X, Y, Z	20 samples per second
Bandwidth of X-Axis	3dB @ 9Hz
Bandwidth of Y-Axis	3dB @ 9Hz
Bandwidth of Z-Axis	3dB @ 9Hz
Noise levels of X, Y, Z	Less than 30pT/√Hz @ 1Hz

Figure 1: Noise Levels below 0.3 nT (shown in yellow)

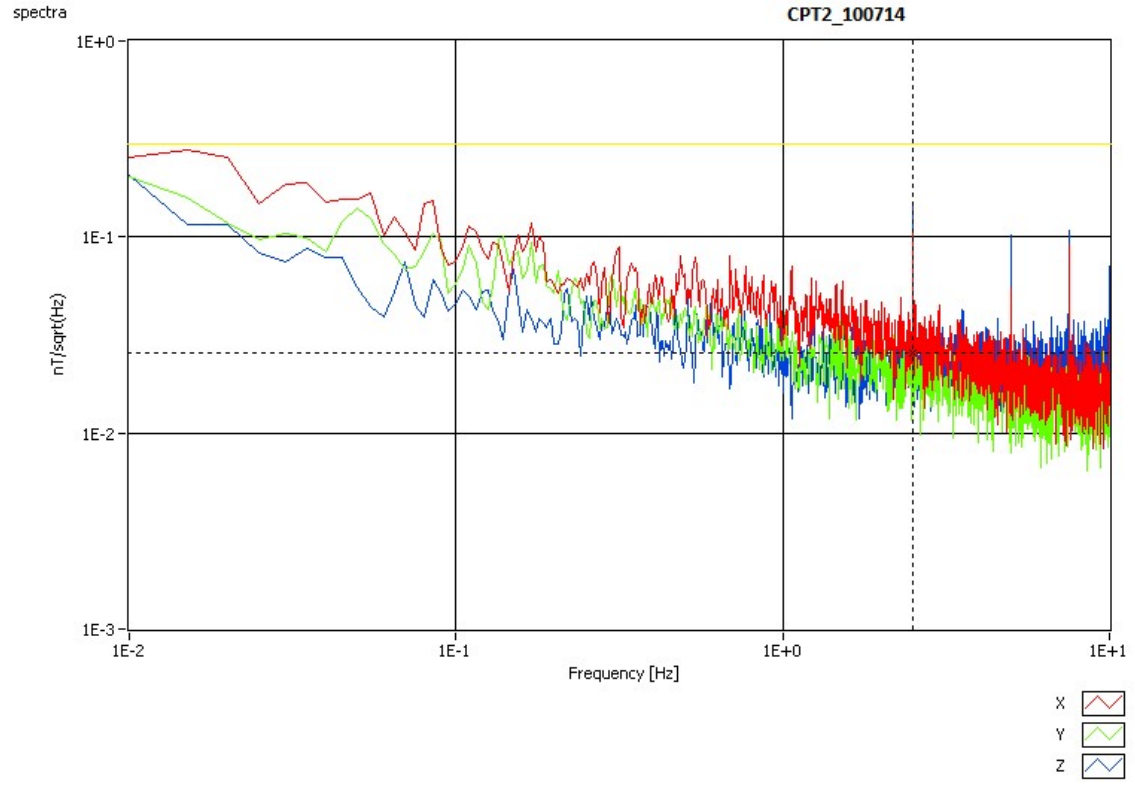
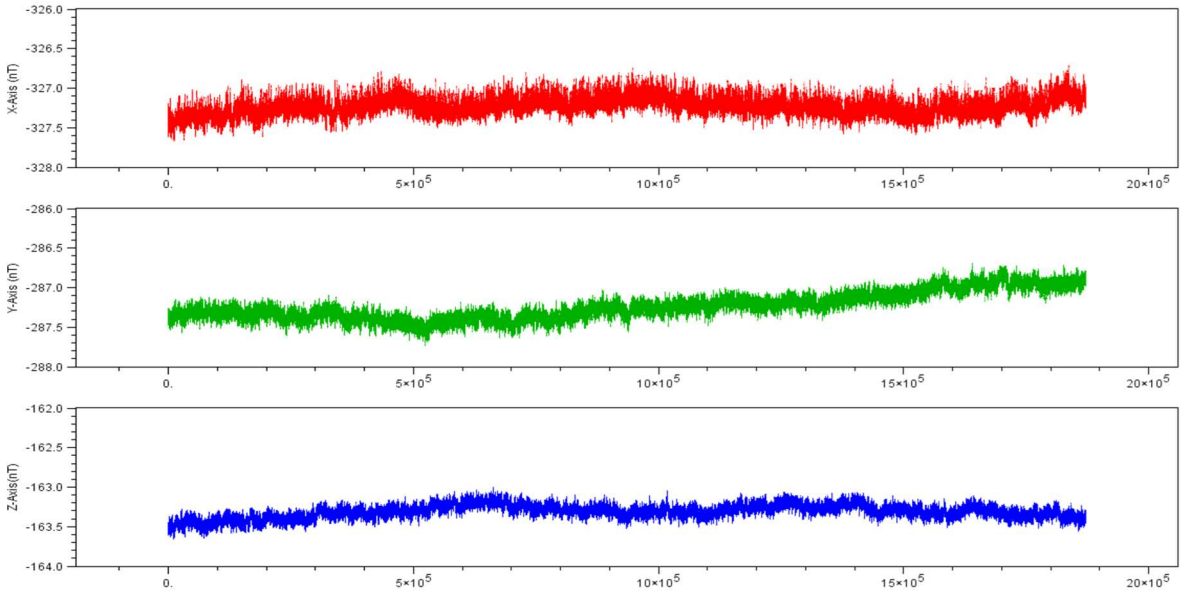


Figure 2: Sensor stability over 24 hours



Transfer function: The digital filter used in the IFG is a 400 point boxcar filter, with a gain of $400/2^9$.(Also can be represented as a coefficient of 1/512, with 400 points). The digital filter dominates the end-to-end transfer function.

Figure 3: Transfer Function (between DC-10 Hz)

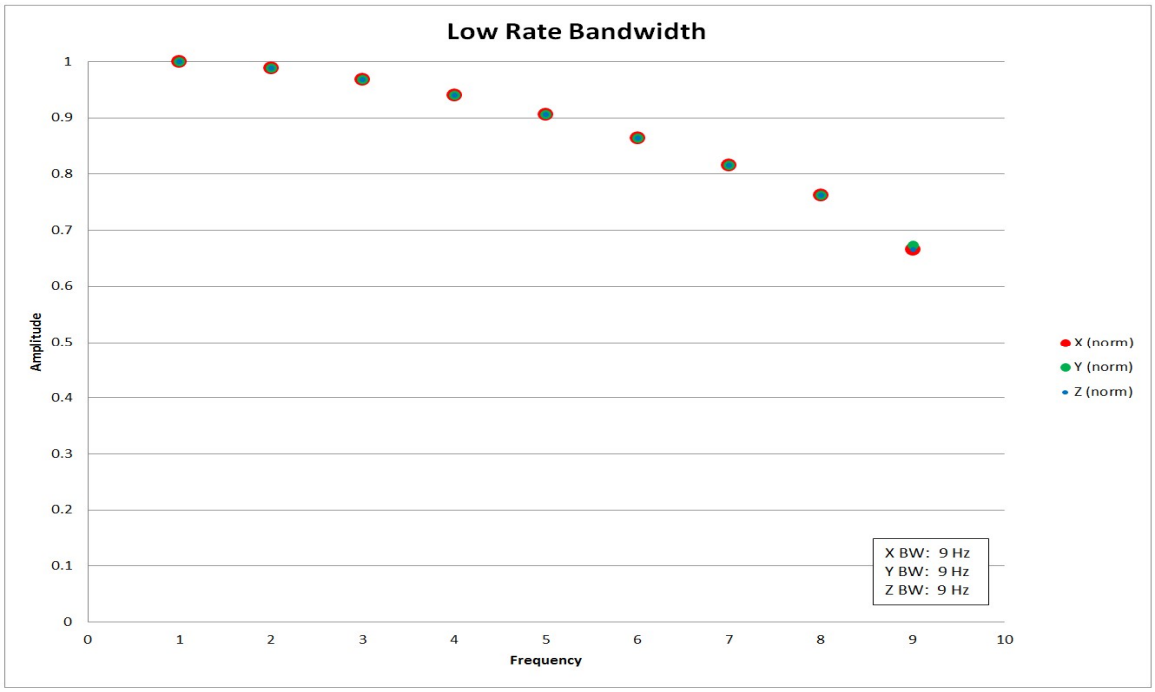


Figure 4: Transfer function (between DC-120Hz)

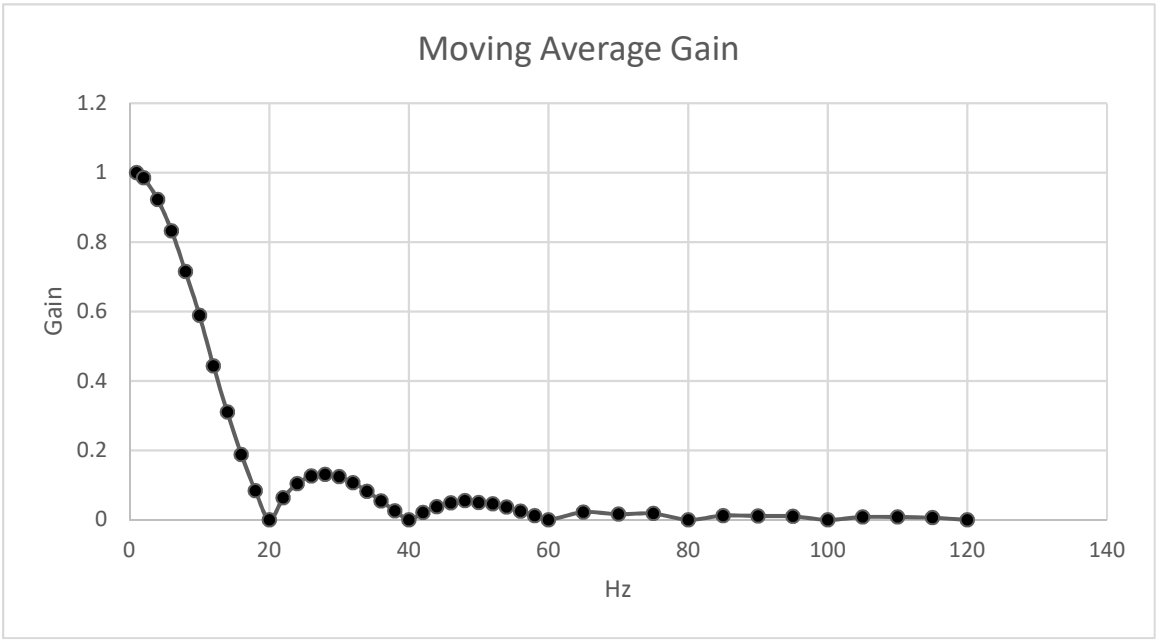


Figure 5: Electronics thermal stability - Offset temperature dependence

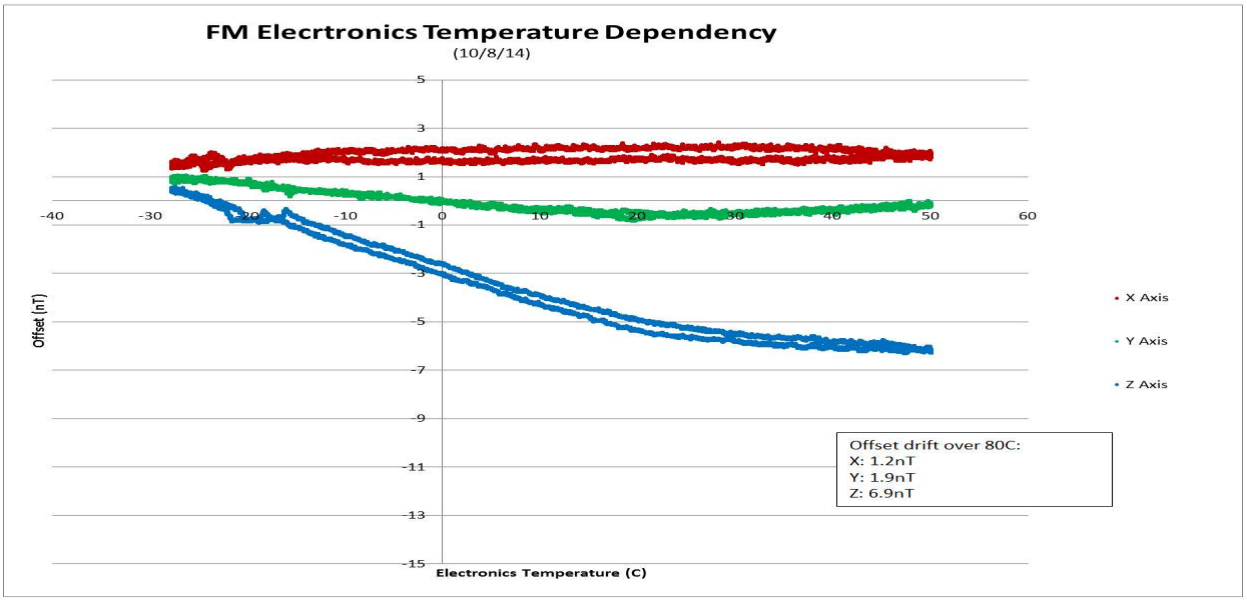
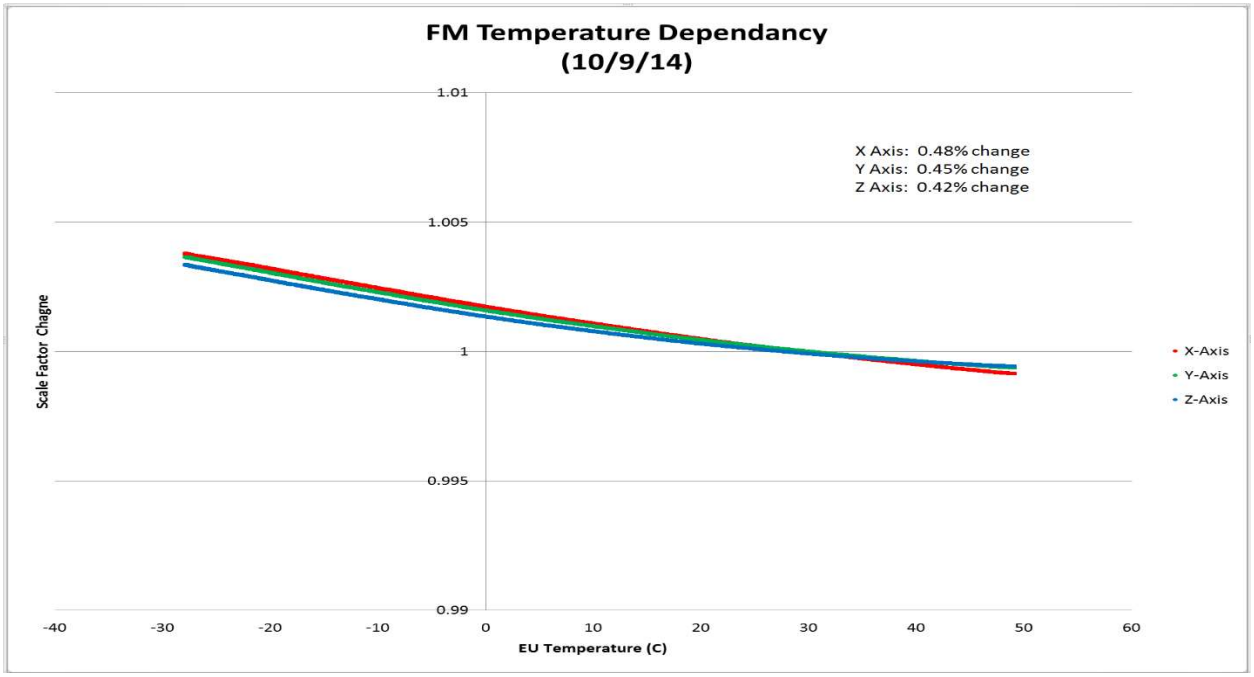


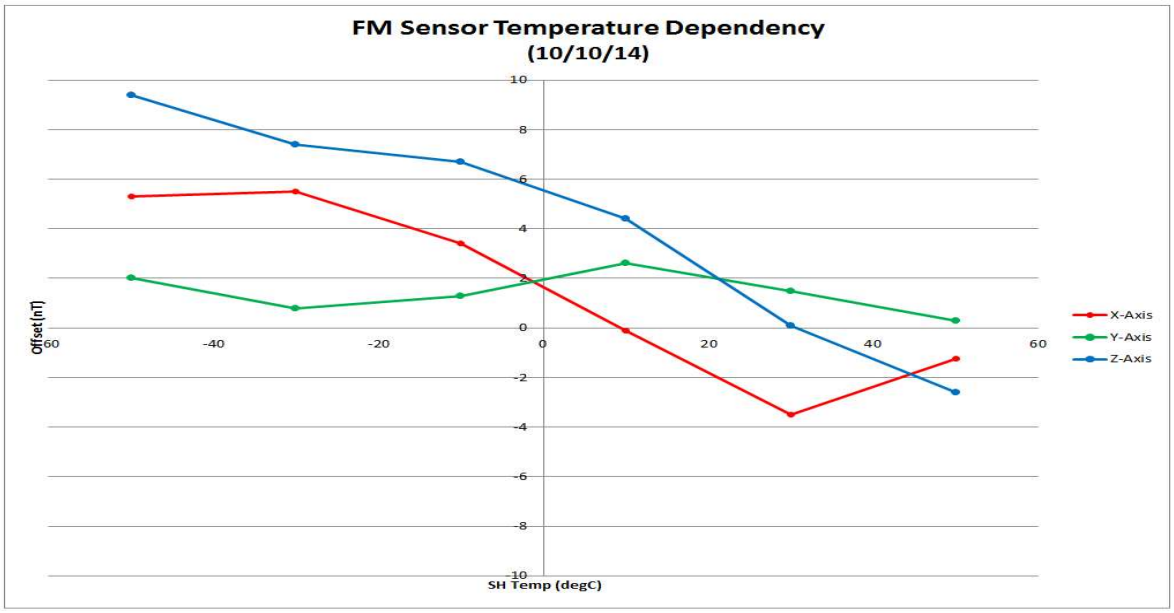
Figure 6: Electronics thermal stability - Gain temperature dependence



Offset X: $y = -0.0002x^2 + 0.0073x + 1.8809$ Offset Y: $y = 0.0005x^2 - 0.0275x - 0.1188$ Offset Z: $y = 0.0009x^2 - 0.1085x - 3.0463$	Gain X: $y = 3E\text{-}07x^2 - 7E\text{-}05x + 1.0017$ Gain Y: $y = 4E\text{-}07x^2 - 6E\text{-}05x + 1.0016$ Gain Z: $y = 4E\text{-}07x^2 - 6E\text{-}05x + 1.0013$

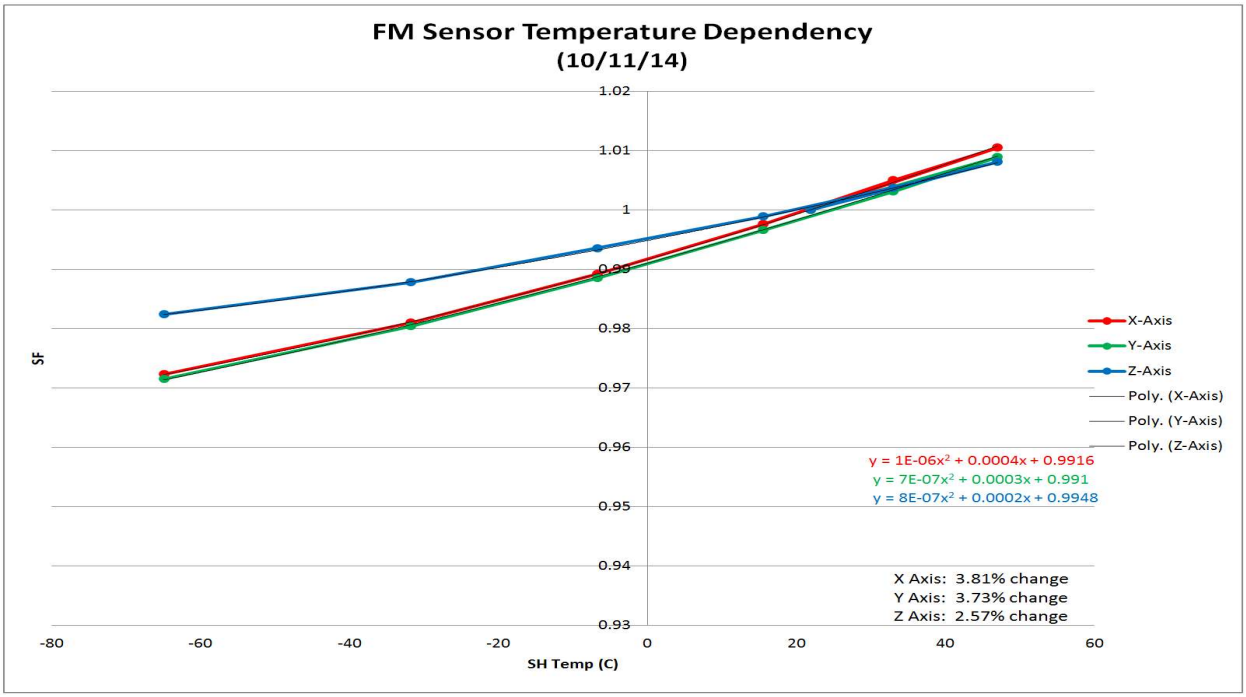
Note that a 0.5% change in a 1,000 nT field is 5 nT over the 80 ºC change in temperature.

Figure 7: Sensor head thermal stability - Offset temperature dependence



Offset X: $y = 0.0002x^2 - 0.0904x + 1.2953$ Offset Y: $y = -0.0003x^2 - 0.0073x + 1.75$ Offset Z: $y = -0.0008x^2 - 0.1203x + 5.1656$	Gain X: $y = 1E\text{-}06x^2 + 0.0004x + 0.9916$ Gain Y: $y = 7E\text{-}07x^2 + 0.0003x + 0.991$ Gain Z: $y = 8E\text{-}07x^2 + 0.0002x + 0.9948$

Figure 8: Sensor head thermal stability - Gain temperature dependence

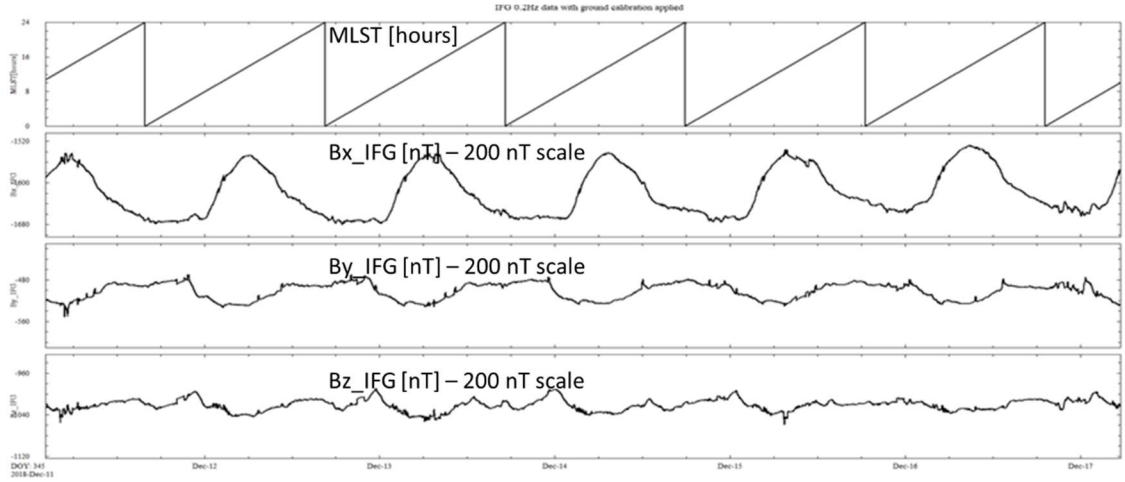


Note that a 3.8% change in a 1,000 nT field is 38 nT over the 120 ºC change in temperature, and that the ground calibration did not extend below -70 ºC.

3 In-flight Calibration

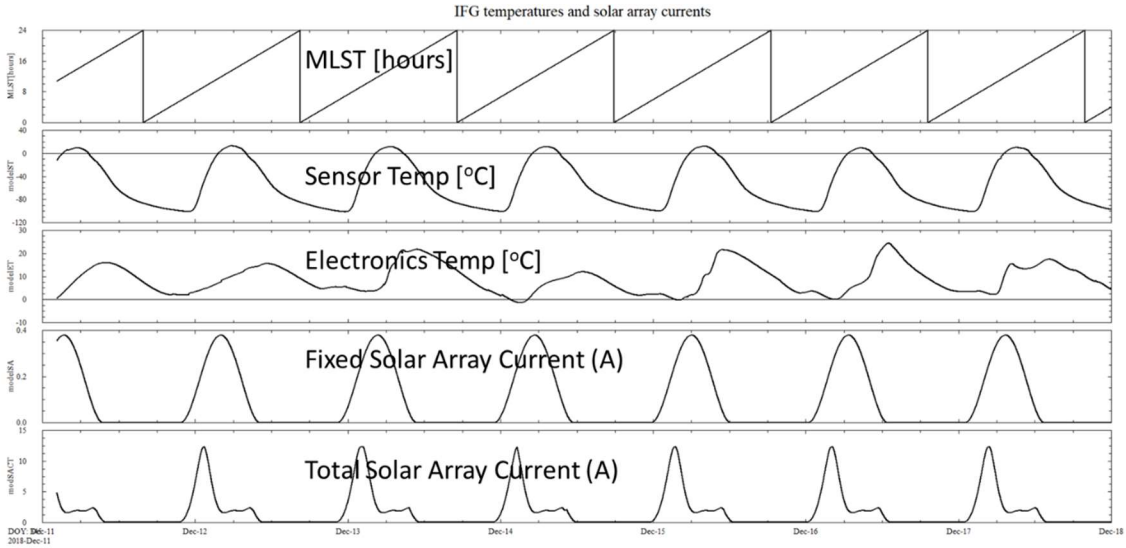
Initially, the InSight IFG were processed just using the ground calibration values. However, the resulting field values showed an unexpectedly large diurnal variation when only processed to this level. Figure 9 shows a few SOLs of data shortly after landing with just the ground calibration applied. The three IFG data panels (bottom 3) show the data on a 200 nT vertical scale.

Figure 9: Early landed IFG data with ground calibration applied



The IFG team initially showed that there was a strong correlation between the temperatures and the observations and the data were initially decorrelated with the temperature data. However, the resulting data still showed larger than expected diurnal variations. After some further investigation, the data were show to also be correlated with the solar array current data. Figure 10 shows the diurnal variation in the IFG temperatures and solar array currents and Figure 11

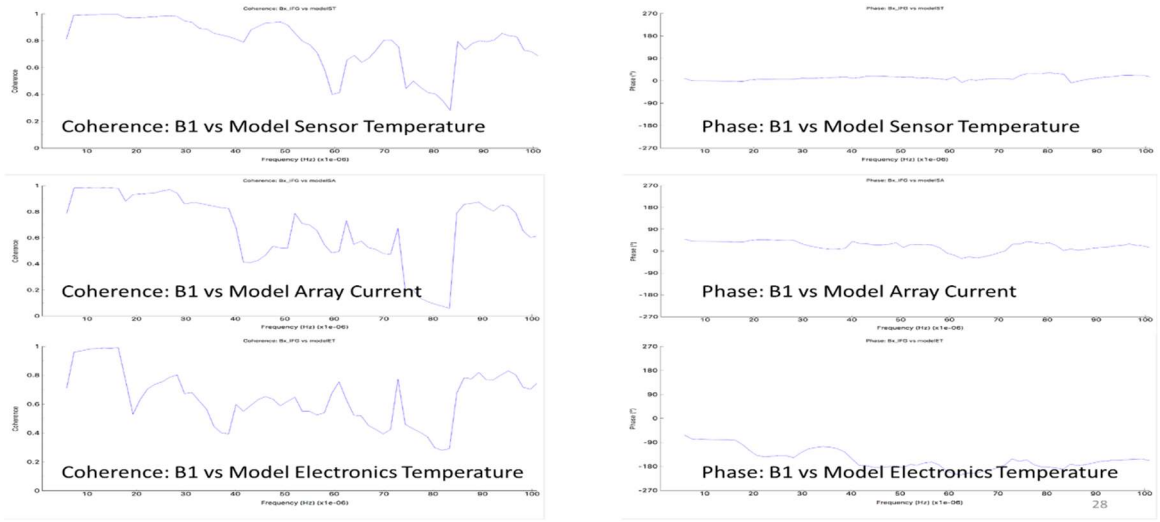
Figure 10: Diurnal variation in IFG temperatures and solar array currents



shows the coherence (left) and phase (right) between the sensor (top) and electronics (bottom) temperatures and the fixed solar array currents (middle). Note that while the sensor temperature and solar array currents are in phase with the IFG data, the electronics temperature is out of

phase. Additional analysis showed that there was an additional correlation with the total solar array current as well.

Figure 11: IFG Coherence with temperature and fixed solar array currents



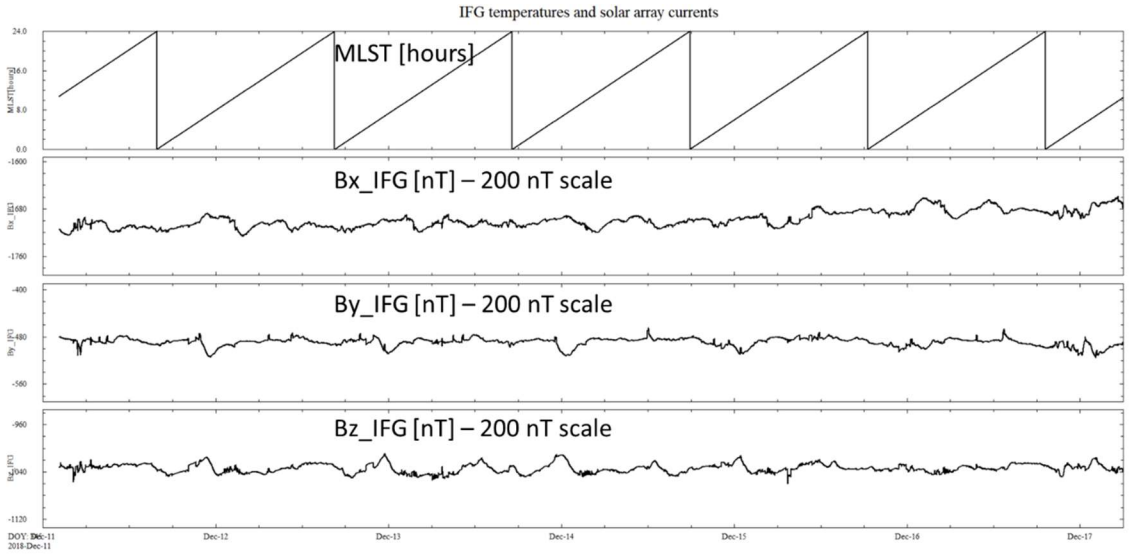
The IFG data are decorrelated with the temperature and current data by assuming that the diurnal variation associated with these parameters can represented as a linear function of the four variables, and the subtracting the resulting function from the data. The daily perturbation is calculated as the difference between the observed field and the mean value observed at 20:00 TLST. The local time was selected because the solar array currents have gone to zero and the temperatures (sensor temp ~ -60 ºC, elect temp ~20 ºC) are within the range that was calibrated in the laboratory (see Figures 5-8). The mean value of the field that is used for this calculation is (-1645, -500, -1045) nT in the IFG frame. These values are subtracted from the data and then the residuals (dB_i) are fit as a linear function of the temperatures and solar array currents of the form:

$$dB_i = C_{0,i} + C_{1,i} * ST + C_{2,i} * ET + C_{3,i} * FSAC + C_{4,i} * TSAC$$

where the C are the constants determined by the fit, ST and ET are the sensor and electronics temperatures (modelSA and modelET) in the ºC respectively, and FSAC and TSAC are the fixed and total solar array (modelSA and modSACT) currents in Amps respectively. During the spacecraft commissioning, the spacecraft environment was changing frequently. We found that a given set of coefficients could be used for only a few days to weeks. The fit coefficients and time range of applicability is provided with the archive in the file called PolynomialFits.txt (LID: urn:nasa:pds:insight-ifg-mars:document:polynomial-fits) in the document collection of the insight-ifg-mars bundle. We are hopeful that the frequency at which new fits and models of the current systems used in the fitting process will decrease once commissioning is complete. In theory, after that time there should only be seasonal changes and those attributable to dust on the solar arrays.

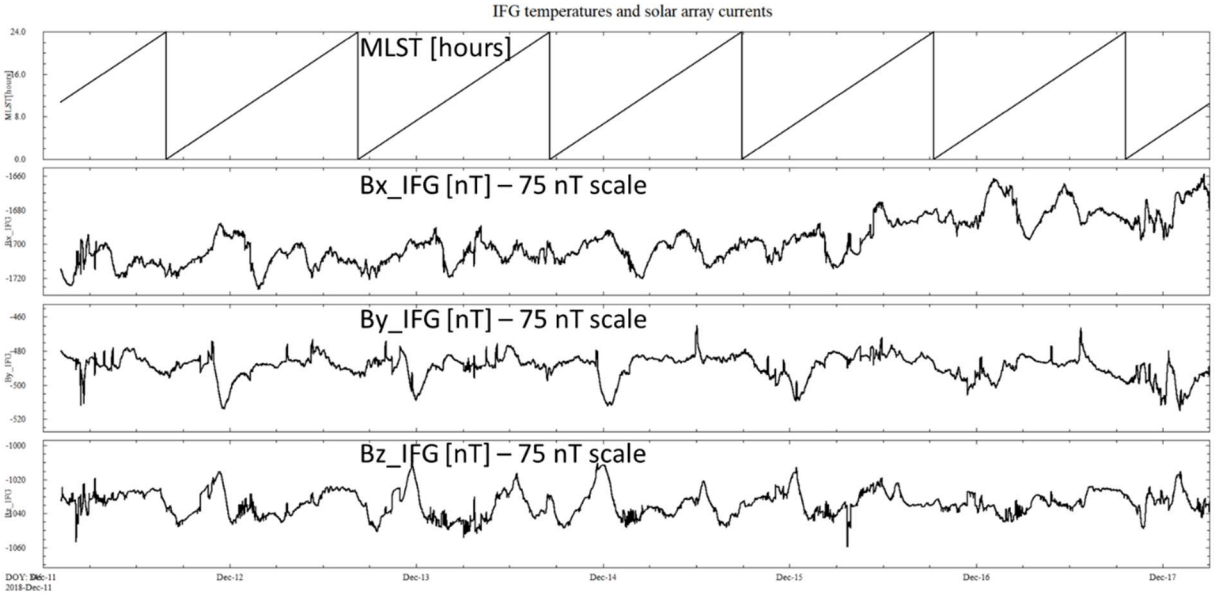
Figure 12 shows the same time interval as Figures 9 and 10, and the same vertical scales, after the calculated dBi's have been subtracted from the data. Figure 13 shows the same time interval on a 75 nT scale. After correction, a diurnal variation in the field that is on the order of 35 nT remains. Of this residual, some portion is likely to be associated with one or more current systems on the spacecraft that have not yet been identified. The IFG team has looked at the

Figure 12: IFG data after decorrelation applied



various current data that are included in the spacecraft engineering and ancillary data provided and have not identified any channels that are well correlated with the residual diurnal signal in the IFG data. This lack of correlation leads us to believe that most of the residual variation in the data are not the result of spacecraft sources.

Figure 13: IFG data after decorrelation on a 75 nT scale



The computation of the delta **B** associated with these thermal and current variations requires a continuous measure of those parameters which is not provided in the InSight spacecraft engineering and ancillary data. Models of the TLST variations in these parameters have been developed to provide the continuous input required. The appendices of this document describe

the computation of the parameters used (modelST, modelET, modelSA, modSACT) for the polynomial fits.

Data that have been corrected for the very low frequency diurnal variations correlated with spacecraft phenomenon are archived in IFG and spacecraft coordinates in the “partially-processed” data collection (urn:nasa:pds:insight-ifg-mars:data-ifg-partially-processed). The data in this collection are provided for the purpose of decorrelating the high frequency signals observed by the IFG with those observed by the SEIS experiment. These data contain significant residual high frequency fluctuations that result from a combination of the natural environment and those attributable to variations in the spacecraft current systems (due to heater cycling, communication systems, etc.) which we collectively refer to as data artifacts.

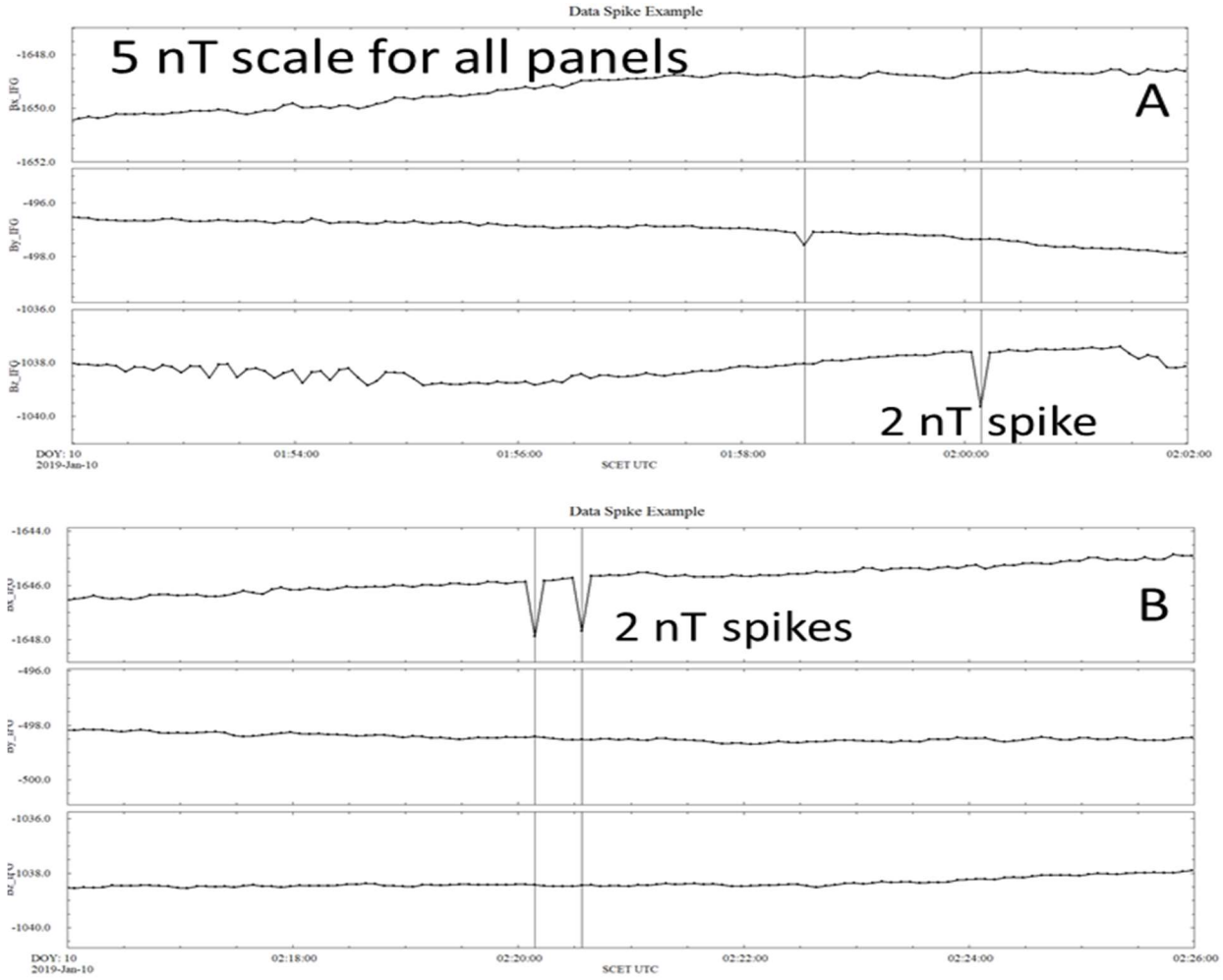
4 Data Artifacts

Prior to the first public data release, several types of artifacts have been observed in the partially processed IFG data. Unfortunately, there has not been sufficient time since the start of landed operations to develop the codes required to automatically detect and correct these issues. The following section describes the types of issues that are present in the data that we hope to be able to correct, in full or in part, in the near future. When a substantially improved data set is available, a complete new version of the calibrated data will be redelivered to the PDS and any previous versions will be superseded.

4.1 Single Point Spikes

The first type of artifact found in the data that we will describe are single point data spikes. The source of these spikes is unknown. While spikes are observed in all components, typically only a single component is impacted at a time. Spike amplitudes are typically only a few nanoTesla. This either implies that spikes are not the result of random bit flips from cosmic rays or that we

Figure 14: Single point spike examples



have not yet acquired a large enough data set to observe the full range of statistical variation of this type. Figure 14 shows two examples of single point data spikes. Both panels show 10 minutes of data in IFG coordinates with a 5 nT scale and the actual data points shown as dots along the traces. In the top panel (A), there is an approximately 2 nT downward spike in the Bz component at 02:00:10 (vertical line) and another smaller fluctuation in the By component at 01:58:35 (vertical line). While the drop in the Bz component is large enough to be considered an artifact (spike), the amplitude of the By component fluctuation is not. It is of the same magnitude as the Bz fluctuations that occur a few minutes earlier. The bottom panel (B) shows a pair of spikes of about 2 nT in the Bx component in close temporal proximity. Note that neither the spikes in the top or bottom panels occur in more than a single component.

Single point data spikes can be identified by comparing the value of a component sample to proceeding and following samples. If the sample differs more that some amount (1.5 nT TBC) from both preceding and following samples, and the preceding and following samples themselves do no vary by mode than some threshold (0.25 nT TBC), then the sample is identified as a spike and its value is replaced by the average of the preceding and following samples. Mathmatically, if

$(\text{abs}(B_{i,j} - B_{i,j-1}) > 1.5)$ and $(\text{abs}(B_{i,j} - B_{i,j+1}) > 1.5)$ and $(\text{abs}(B_{i,j-1} - B_{i,j+1}) < 0.25)$

then

$B_{i,j} = (B_{i,j-1} + B_{i,j+1})/2$

where the (i) subscript refers to the field component (1,2,3) and the (j) subscript refers to the sample number or time step.

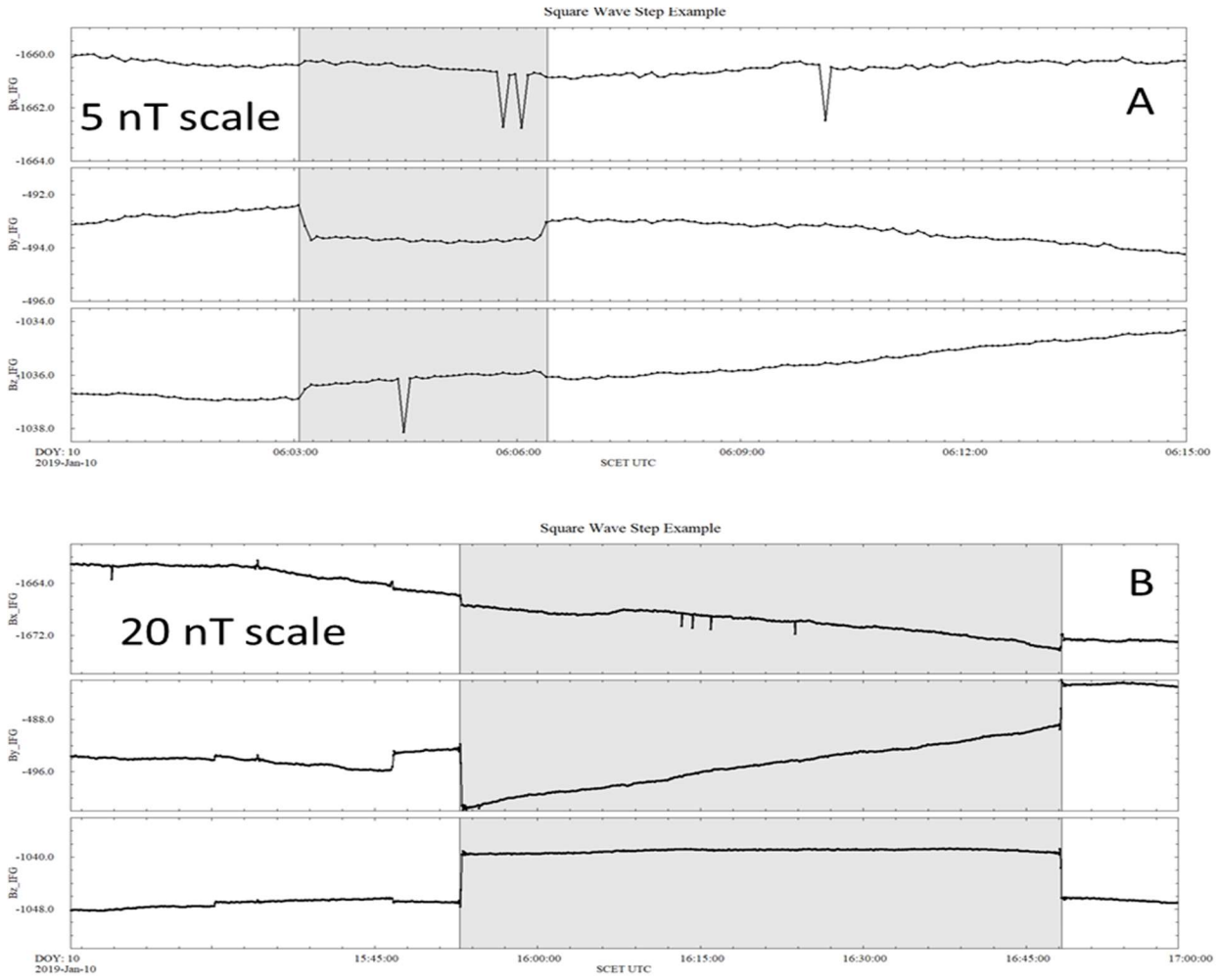
Single point data spikes should be corrected/removed before attempting to identify or removed more complex interference structures. After identification and correction, the data quality flag (dqf) for the sample (8^{th} element from the right) should be set to 1 (Issue corrected in any/all components).

4.2 Square Wave Steps

Square wave steps are defined as intervals where two or more field components abruptly shift (increase or decrease) at the same time and then shift by approximately the same amount in the opposite sense some time later. The sense of the shift (initial increase or decrease) is not same between components in general. Figure 15 shows two examples of square ware steps. In the top panel (A), the By and Bz components (IFG frame) shift in the opposite sense and then return to near their original levels after about 3 minutes. The shifts are small, roughly 0.8-1.5 nT in By and 0.5 nT in Bz. The negative enhancement of the By component at the onset is larger than the return at the termination. The same is true for the Bz component, although somewhat less apparent since the amplitude of the shift is smaller.

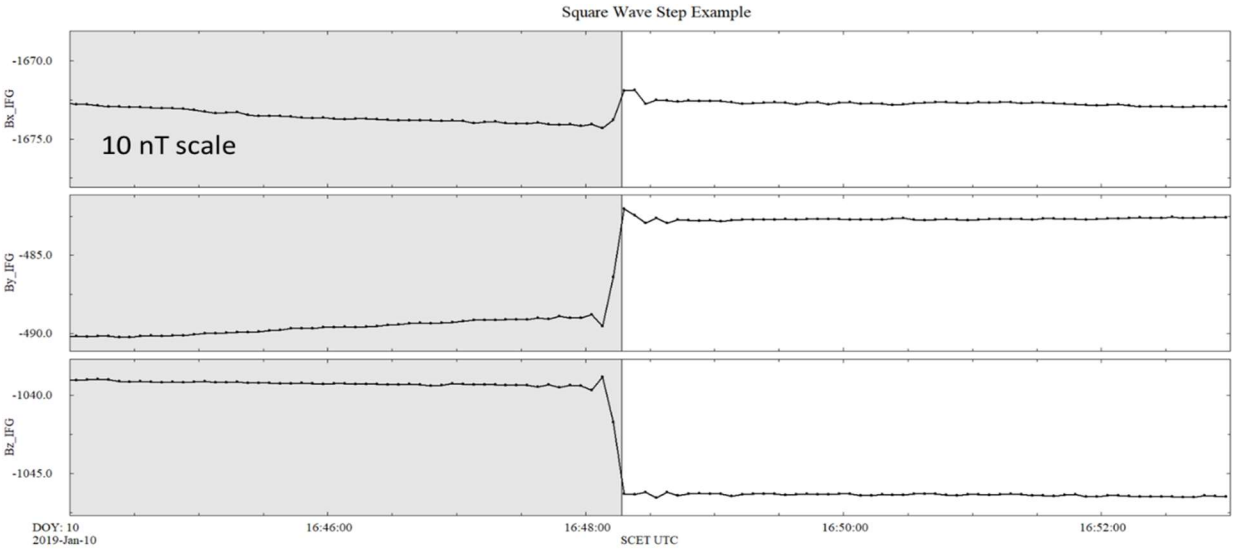
Panel B shows a much larger (7-10 nT in By and Bz, 1.6 nT in Bx) and longer (15:53 – 16:48) shift (shaded) that impacts all three components in the IFG frame that immediately following a shorter, smaller shift that is not shaded that primarily only impacts the By and Bz components

Figure 15: Examples of square wave steps



(~15:46). In both panels of Figure 15, the "step" duration is longer than the time between samples (5 sec) so that there is at least one sample part way between the top and bottom of the step. Square wave step identification and mitigation can be further complicated by short transient events near the onset or termination of the events. Figure 16 focuses in on the termination of the event show in panel B of Figure 15. In this figure you can see that there appears to be ringing of the FIR filter associated with the termination of this artifact. This ringing will complicate the determination of both the timing and amplitude of the artifact termination. Artifact removal may be incomplete or result in the introduction of additional artifacts. For the purpose of determining the amplitude of a potential step in a component, the differences between averages over a small number of points, maybe five (TBC) before ($B_{i,j-7}$ to $B_{i,j-2}$) and after ($B_{i,j+2}$ to $B_{i,j+7}$) the step should be computed. This window needs to be small enough to not eliminate trends in the data and long enough to any filter ringing that might be associated with the step.

Figure 16: Blow up of the termination of a selected square wave step



Square wave steps can be identified in an automated fashion by applying the following algorithm.

1) Onsets can be identified by comparing samples $B_{i,j-1}$ to $B_{i,j+1}$ where the subscript (i) indicates the component and (j) the sample time/number. The initial amplitude of the step at sample/time j ($A_{i,j}$) is $A_{i,j} = B_{i,j+1} - B_{i,j-1}$. If the absolute value of at least one initial $A_{i,j}$ is greater than 0.5nT – AND- the absolute value of the amplitude of another initial $A_{k\neq i,j}$ is at least 0.3 nT, then vector at time/sample j is a potential onset sample with step amplitudes of $A_{i,j}$. The onset sample (j) is recorded and the actual amplitudes are computed using the averaging method described above.

2) If a potential onset has been identified, then search for a termination sometime in the next 120 minutes (TBC) by comparing samples $B_{i,j+n-1}$ to $B_{i,j+n+1}$ where (j+n) indicates a sample/time greater than (j) and the potential step amplitudes ($A_{i,\,j+n}$) are computed for samples (j+n). Potential termination steps and their amplitudes are determined as described for onsets. To be considered a step termination, the amplitude of the return steps must have the opposite sign as the onset amplitudes and be within 20% (TBC) of the onset amplitude $-0.8 * A_{i,j} \leq A_{i,\,j+n} \leq -1.2 * A_{i,j}$ for all (i).

3) The identification and correction algorithms will need to account for the "stacking" of events, as appears to be happening in the event identified in Figure 15, panel B where a second (or more) step(s) occur before the field returns to its unperturbed state. In this time interval, both the By (+2.9 nT) and Bz (-0.4 nT) components see a shift at 15:46 that would be identified as an onset event. It is unlikely that the slight shift in the Bx component (-0.2 nT) at this time would be considered since, given the downward trend in the data, the difference does not meet criteria (1) above. The potential termination event at 15:52 in components By (-9.5 nT) and Bz (+7.6 nT) amplitudes are too large to be identified as a termination, using the definition in (2) above. At this time, there is also a potential onset event in the Bx component (-1.6 nT). Since the events at 15:46 and 15:53 are not an onset/termination pair, they should be considered as separate onset events, and their amplitudes summed dBx = 0 - 1.6 = -1.6 nT, dBy = 2.9 – 9.5 = -6.6 nT, dBz = -0.4 + 7.6 = 7.2 nT. The next potential termination event occurs at 16:48 where there are shifts in

all three components (1.8, 6.4, -6.9) nT. When the summed amplitudes of the onset events (-1.6, -6.6, 7.2) are compared to amplitudes of the event at 16:48 (1.8, 6.4, -6.9), the termination criteria (2) are met. We refer to square wave steps that have multiple onsets as "compound square wave" events while those that have only a single onset and termination are called "simple square wave" events.

If a simple square wave event is identified, it can be at least partially removed by applying the following steps.

1) Begin by setting the values of the data immediately before and after the step to the average values used for the amplitude determination:
$B_{i,j-7}$ to $B_{i,j}$ are set to the average value of samples $B_{i,j-7}$ to $B_{i,j-2}$
$B_{i,j+1}$ to $B_{i,j+7}$ are set to the average value of samples $B_{i,j+3}$ to $B_{i,j+7}$
$B_{i,j+n-7}$ to $B_{i,j+n}$ are set to the average value of samples $B_{i,j+n-7}$ to $B_{i,j+n-2}$
$B_{i,j+n+1}$ to $B_{i,j+n+7}$ are set to the average value of samples $B_{i,j+n+3}$ to $B_{i,j+n+7}$

2) For samples (j+1) to (j+n), the linearly varying amplitude between the onset and termination is subtracted from the field components. The linearly varying amplitude can be represented as
$linA_{i,k} = A_{i,j} + dA_{i,k}$
where k=j, j+n and $dA_{i,k} = (k-j)*(A_{i,j+n} – A_{i,j})/n$ and the correction to field components is:

$B_{i,k} = B_{i,k} – linA_{i,k}$ and the data quality flag for the square wave step element (9^{th} from right) should be set to reflect partial correction of more than one component for samples between (j-7) and (j+n+7). The dqf value for this offset should be set to 2 (all components partially corrected) if the step was identified in all 3 components. The dqf value should be set to 3 (partially corrected in more than one component) if the step was only identified in two components.

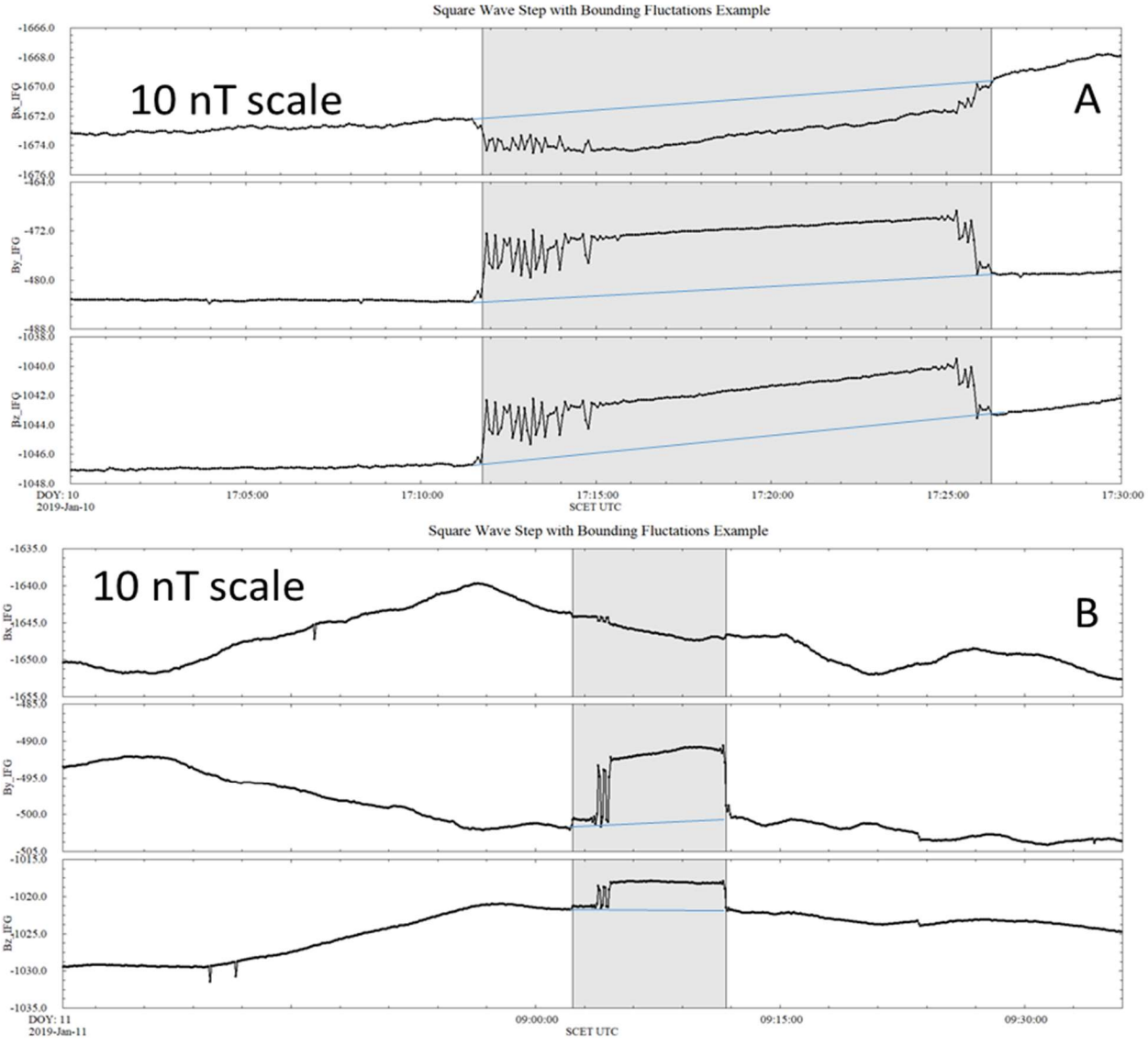
If a compound square wave step is detected, the mitigation follows the same steps as for a single step except that the linear variation in the step amplitude from initial onset to termination needs to be distributed across the total event duration. For simplicity, we will only describe the mitigation of the event shown in Figure 15, panel B. We use the sample subscripts (j, j+n, j+m) to describe the steps at 15:46, 15:53, and 16:48 respectively and call the amplitudes at the step A1, A2, and A3. Amplitudes A1 and A3 are computed as described above for a single step. Amplitude A2 is the difference between A3 and A1 ($A2_i = A3_i – A1_i$). In this example, the step in Bx at 15:46 is zero ($A1_1 = 0$). The linear drift rate ($dA_{i,k}$) for components By and Bz is computed from the time of the initial onset to the termination (j, j+m) while for Bx, it's computed between (j+n,j+m). By and Bz samples between the first and second onsets (j+1 to j+n) are corrected using $A1_i$ and $dA_{i,k}$ (k=j+1, j+n) and those between the second onset and the termination are corrected by using $A2_i$ and $dA_{i,k}$ (k=j+n+1, j+m). In this case, the dqf value is set to 3 for sampled (j-7, j+n-7) and to 2 for samples (j+n-6, j+m+7). The values of Bx are unchanged until sample (j+n-7) so only two components have been corrected prior to this time. Bx values between (j+n-7, j+n) are corrected during the computation of $A2_1$ so the dqf for these samples should reflect partial corrections to all 3 samples.

4.3 Square Wave Steps with Bounding Fluctuations

One of the more complicated forms of data artifact is referred to as a square wave step with bounding fluctuations. As its name implies, these artifact appear similar to the simple or

compound square wave but the determination of the step size at any or all steps is complicated by fluctuations near the steps. Figure 17 shows examples of some events. In panel A, sharp steps occur in the By and Bz components that are either followed by or preceded by field fluctuations that are not sinusoidal (shaded region). At the same time, there are drops and fluctuations in the Bx component that are not abrupt at onset or termination. In this example, the fluctuations are positive and negative excursions from the shifted trend lines. In panel B, there is a small, simple square wave onset in By and Bz, followed by a second step that begins with fluctuations. In this example, the fluctuations appear to be variations between the initial and shifted trend lines. In both panels, the data plots have been annotated with a blue line that shows the likely field trend in the absence of these artifacts. However, the IFG team has not yet determined an algorithm for the detection or correction of these events even though they are easy enough to detect by visual inspection. Any correction algorithm will likely involve removing of the steps and fluctuations as separate actions. If the fluctuations are removed from the data, extreme caution will need to be exercised to prevent the unintentional removal of fluctuations of geophysical interest.

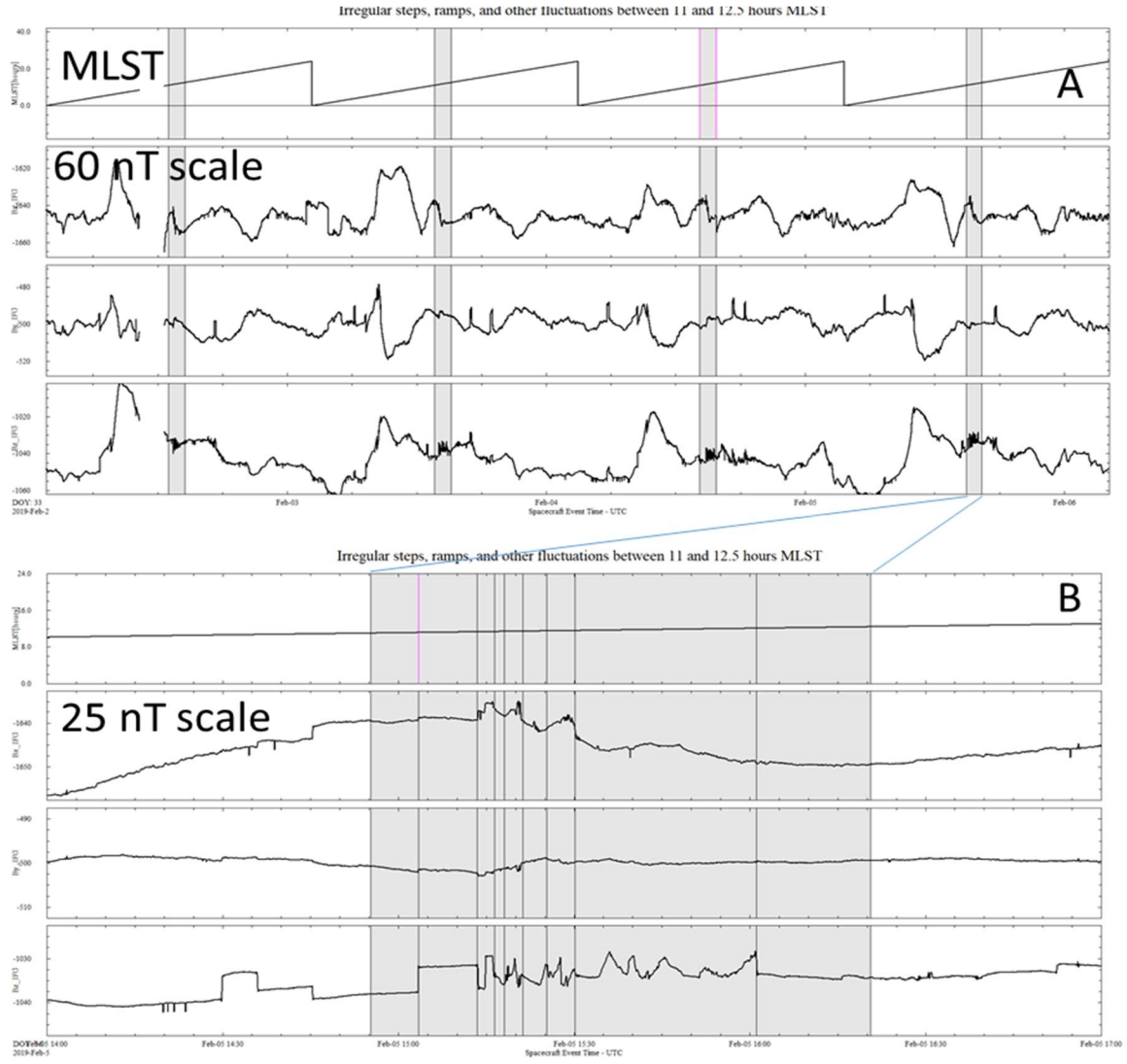
Figure 17: Square wave steps with bounding fluctuations



4.4 Irregular steps, ramps, etc. between 11 and 12.5 hours TLST

The artifacts discussed thus far can occur nearly any time of day on Mars, although nighttime is generally much less impacted than the daytime when the other payload systems are active. This next type of perturbation occurs regularly near mid-day TLST. Figure 18, panel A shows four Mars days of data with the TLST time period between 11 and 12.5 hours shaded in each day. In each of these shaded intervals, you can see a disturbance that is strongest in the Bz component. Panel B expands the time interval on February 5, 2019 so that the types of variations can be shown in more detail. The TLST interval from 11-12.5 hours is shaded and there are some vertical lines are drawn to guide the reader's eyes between panels. On this day, the interference begins at 11.2 hours and ends at 12.1 hours (first/last vertical lines). Over the entire interval,

Figure 18: Irregular steps, ramps, and other fluctuations between 11 and 12.5 hours TLST



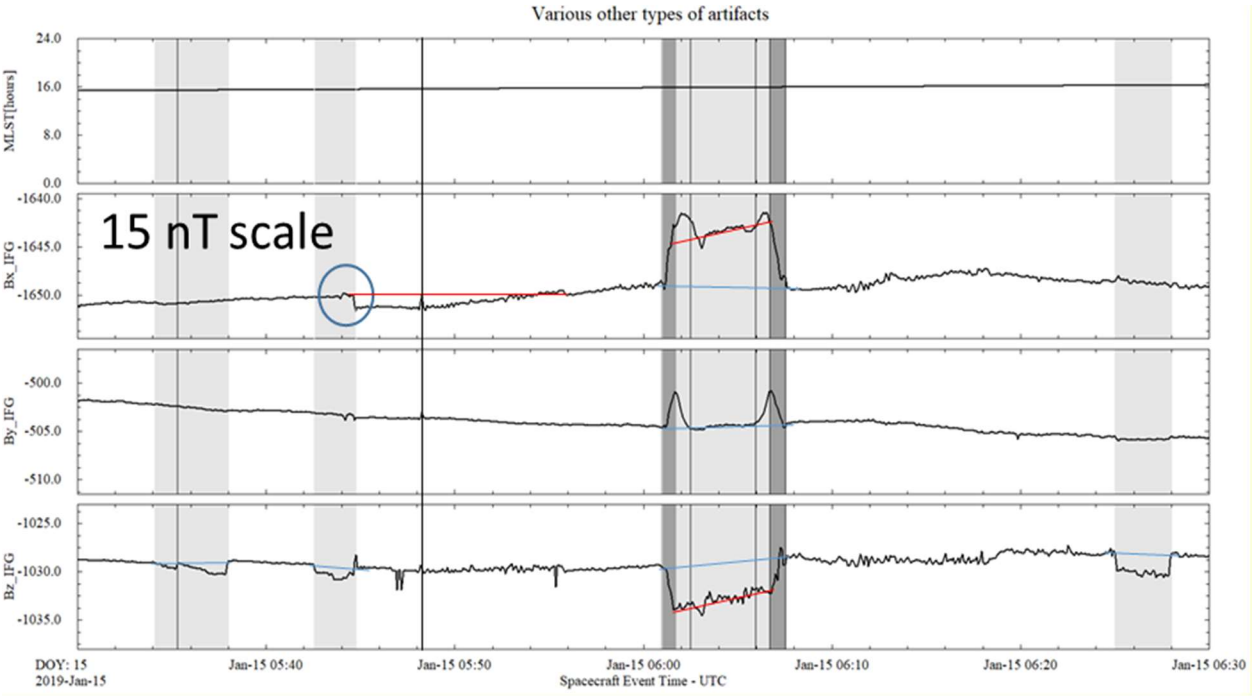
there is very little disturbance in the By (IFG) component, and a lack of participation in these events is characteristic of this type of disturbance. There are fluctuations in the Bx componentat the times of some of the changes in the Bz component, but not for all of them and the shapes of

the fluctuations are not necessarily the same. This is also common for disturbances in this TLST time interval. Many of the Bz fluctuations look more like ramps or humps than steps or spikes. It is unlikely that these types of fluctuation will ever be removed from the data unless a current system on the spacecraft can be found to have shapes to these similar to these signals and the parameter is available in the spacecraft engineering and ancillary data (SCEA) at a cadence that is sufficient to resolve and remove the magnetic signatures. Typically the SCEA data have both low cadence and low continuity on this spacecraft making them very difficult to use directly for the removal of spacecraft artifacts from the magnetic field data.

4.5 Other artifacts

Many if not most of the artifacts that remain in the filter 0.2 Hz data have been described above. However, there are other types of artifacts that are found occasionally throughout the data set.

Figure 19: Various other types of artifacts



Several different artifacts are identified by shading and probably trend lines are shown with blue lines, although other markings are also present. Starting from the left side of Figure 19, the first shaded region shows a set of downward ramps in the Bz component with no corresponding variations in the Bx or By components. The vertical line in this shaded region shows a return to trend followed by another ramp down the sharply returns to the trend after a minute or two. Moving to the right, the next shaded region shows a sharp step in the Bz component that returns to the trend after a few minute. However, during the disturbed interval, Bz does not follow the trend so this can't be described as a square wave step. There are no deviations from the trends in the Bx or By components at the onset of the artifact but both components so variations at its

termination. The Bx component has a step (circled) at this time with no clear step back to the previous level at a later time. It may be that the red line shown represents the trend in Bx and that there is a slow, ramping back to the trend over the next 8-10 minutes, however this is purely speculative. During the middle of this event, there is a brief excursion of the Bx component back to the putative trend line marked with a vertical line. This excursion involves 5 or six data samples, unlike the single point spikes that occur between 15:45 and 15:55 in the Bz component. In addition, there are small fluctuations in the Bz and By components at the time marked by this vertical line. Moving to the next shaded region, there is a fairly sharp drop in the Bz component over several data points, followed by a return to the trend 5 or 6 minutes later. However, the onset and termination steps are different sizes and the slope of data within the disturbed region (red line) is greater than that of the trend line. The By component shows a field increase and decrease of about the same duration as the drops/rises in the Bz component (dark shaded regions) but the field returns to trend in the middle of the Bz disturbance. The disturbance in the Bx component has features that are correlated in time with those in the Bz and By components, including the slope of central trend line (red – parallel to red line in Bz), but are otherwise different from the signatures in the other two components. Moving on to the last shaded region, the appears to a square wave step that only affects the Bz component. While there might be very slight variations in the other two components at the onset and termination events, these would not be identified by the previously stated identification criteria. Clearly there are many different current systems active on the spacecraft during this time period, each of which has its own magnetic signature and the observed field variations are the sum of the contributions of the individual currents. It is unlikely that any automated identification criteria or mitigation algorithm can be developed to flag and/fix any of these types of events in the data. We can only hope that once the spacecraft is fully commissioned, disturbances such as these will no longer be observed in the IFG data.

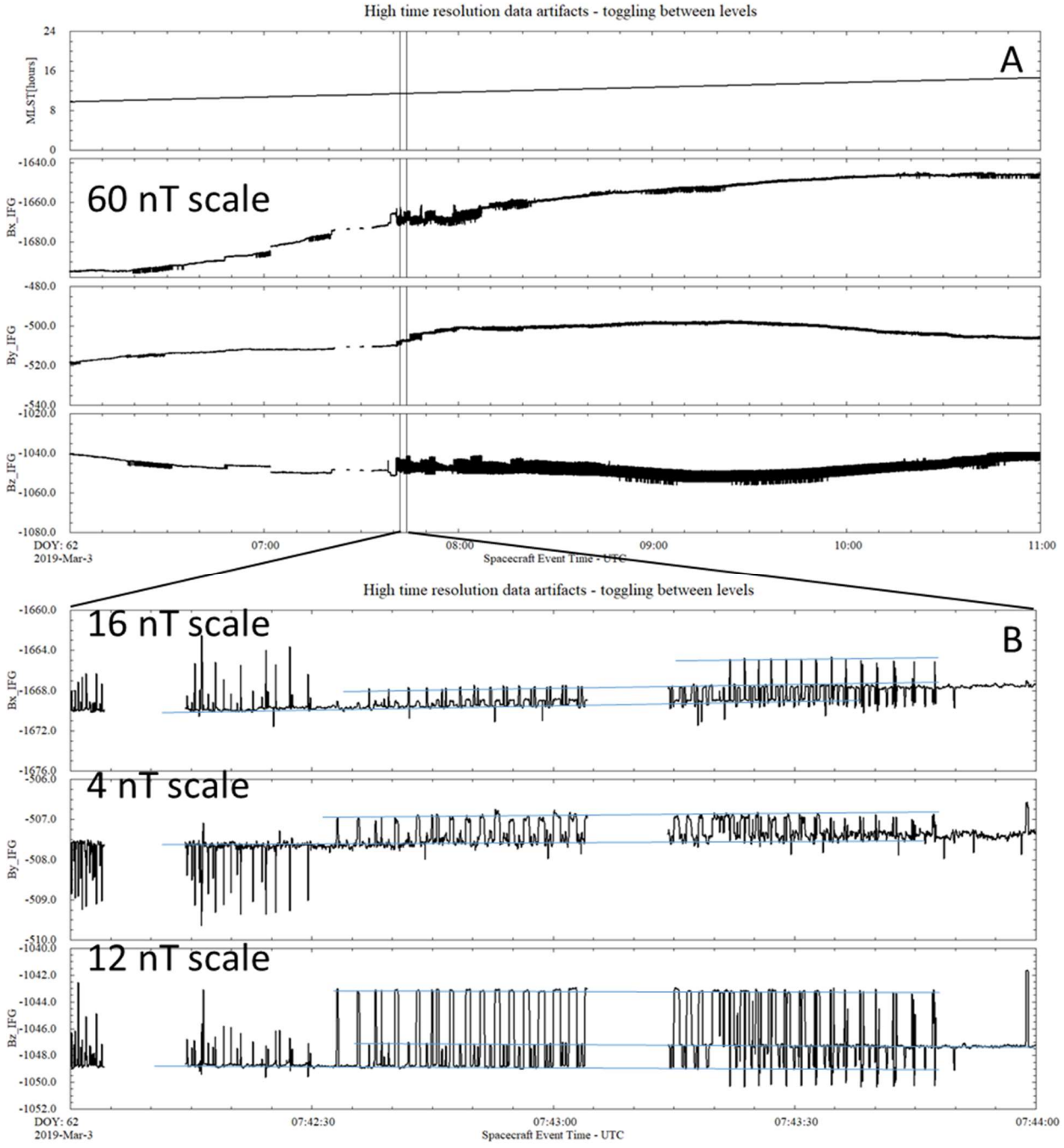
4.6 Artifacts in the high time resolution data

Thus far, the artifacts that have been described have been restricted to those that are identified in the low time resolution (0.2 Hz) continuous data. Data are available at higher data rates for selected time intervals when the downlink bandwidth allowed. All of the artifacts described previously are present in the high time resolution data, although the signatures may be slightly different in terms of onset or termination durations. One phenomenon that is present in nearly all of the high time resolution data is the toggling of the samples between multiple, parallel, baseline values, as shown in Figure 20. The top panel (A) shows five hours of 20 Hz resolution data at a fixed 60 nT scale while the bottom panel (B) zooms in on two minutes of data near the onset of persistent toggling. Panel (A) shows that the onset of the toggling is very rapid, and that Bz is the component that is most severely impacted. The data in panel (B) are displayed at different vertical scales in order to show that similar phenomenology is occurring in all of sensors at different amplitudes. Blue trend lines have been overlaid on the data in panel B to guide the reader's eyes. Focusing on panel (B), and starting on the left side of the figure (07:42:00), the magnetic field appears to have a baseline of about (-1670, -507.5, -1050) nT with spikes of different amplitudes on the order of a few nT around this value. At approximately 07:42:35, the field begins to toggle between values at the existing baseline and a new baseline of (-1668, -507.0, -1043) nT. A few seconds later, a 3rd baseline value (-1047) appears in the Bz component. A third baseline appears briefly in the Bx component (-1666 nT) at around 07:42:50. Eventually the field gradually begins to stabilize around a new baseline value of (-1668, -507.5, -1047) at around 07:43:50 leaving By unchanged. As panel (A) clearly indicates, this is only a brief hiatus

in the toggling, while panel (B) demonstrates that beneath the hash, baseline changes are occurring.

If these fluctuations occurred with step sizes that were integer powers of two in the raw data, they would look like bit flips between two states in one of the middle bits (2^7 - 2^{11}) in the 24 bit analog to digital conversion, of the type that are always present in the lowest order bits in digital data. However, this is not what is happening. The sensor scale factors are all near 140 so that 140

Figure 20: High time resolution toggling near a baseline step



DN (data numbers) approximately equals 1 nT, The observed step sizes require several bits to toggle in the middle registers simultaneously. While suspicious in appearance, this is not an instrumental artifact.

It is unlikely that the IFG team will be able to remove the toggling fluctuations from the high rate data.

5 Data versions and Documentation updates

The version 1 IFG data set provided in the first InSight data release does not identify or correct any of the artifact types presented above. The calibrated data collection has been processed in the same manner as the partially calibrated data. The only differences between the two collections are the reference frames (the calibrated data replace the IFG frame with the local level, local north frame), and a data quality flag (DQF) column has been added. The elements of the DQF that describe data quality are all set to the value 5 (not evaluated), except for the element that describes the irregular fluctuations between 11 and 12.5 TLST, which is set to zero then the TLST is not in the impacted interval.

Version 2 data, which included in data release 2, had single point spikes flagged and removed, and also had some of the simple and compound square wave steps flagged and partially removed. In addition, this version of the data corrected an error in the timing of the sensor temperature data that reduced the overall quality of the calibration.

Version 3 data include all of the version 2 corrections but also corrected an error in the local time column. Data originally labeled as MLST actually contained TLST values. This error explains why many of the figures in this document are labeled as MLST while the text refers to TLST. Version 3 data now include both MLST and TLST data columns, each with the correct contents.

Many steps and artifacts remain in the calibrated IFG data. Many steps appear that look as though they should have been removed by the destepping algorithm remain. Most of the remaining steps occur because either the slope of the step was too shallow to be identified as a true step, or that the background level from which the step occurs could not be clearly identified.

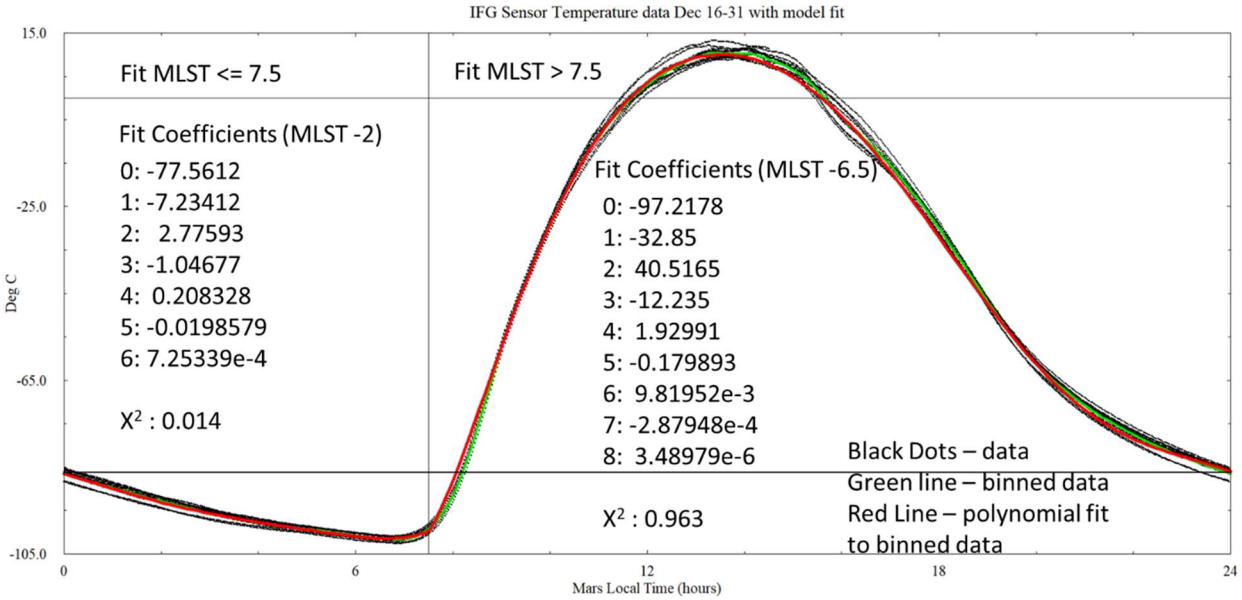
If criteria can be developed to identify and remove these or some of the other artifacts describe here, or discovered in the future, these will be corrected in later versions of the data. This document will be updated and released with each delivery to the PDS that includes new data calibration or processing.

Appendix 1: IFG Sensor Temperature

In general, when the IFG is on and acquiring data, the sensor temperature data are available. The exception occurs immediately after a power on or PAE reset. The FIR filter applied to the temperature data is much longer and it takes more than two hours for the first sensor temperature sample to be output. In order to be able to process the IFG data in the absence of temperature data, a model of the temperature versus MLST has been developed. To develop the model, sensor temperature data from December 16-31, 2018 were binned into 0.1 hour MLST bins. The binned data were then fit using a polynomial function of MLST. In order to make sure that the model was continuous across the 24 to 0 hours MLST boundaries, data from the 22-24 hour bins were duplicated as hours -2 to 0, and similarly, the data from hours 0-2 were replicated as hours 24-26. Binned data from the MLST range -2 to 26 were then fit. The shape of the temperature variation over the Mars day could not be fit with an acceptable error by using a single polynomial function. In order to achieve a reasonable fit, the data are fit in two segments, with one function covering the MLST range -2 to 8 hours, and the second covering hours 6.5 to 26. The two fits agree well at a MLST values of 0 and 7.5 hours so these are used as the transition points between the two functions.

Figure 21 shows the data values as individual black dots. The green line traces the binned values and the red line shows the model fit. There is a horizontal line drawn between the model values at 0 and 24 hours to show that the values are nearly, but not exactly equal. There is also a vertical line drawn at MLST = 7.5 hours to show that the morning transition between the polynomial functions is smooth. The fit coefficients are given, along with the chi-squared values of the two fits. The IFG sensor temperature profile has been fairly stable since landing and there has been no need to update the fit parameters using data acquired in 2019. The data are routinely checked against the model and the model will be updated in the future if required.

Figure 21: Polynomial fit to the IFG sensor temperature data.



The DQF value in the calibrated data set is used to identify the source of the sensor temperature data used in the calibration process. The 2^2 place is where this information is stored and the value

is set to 0 if actual data are used and is set to 3 for samples where the model temperature was used.

Appendix 2: IFG Electronics Temperature

The IFG calibration is also a function of the electronics temperature. Unfortunately, only temperature sample is returned to the spacecraft ground per data processing session so the actual data can't be used for either the data processing or even to develop a model like has been done for the sensor temperature data. Fortunately, there is a temperature measurement for the PAE electronics box in the spacecraft engineering and ancillary data (T-0014) which serves as a reasonably good proxy for the missing data. Unfortunately, this channel is not sampled and returned continuously each Mars day. In general, the channel is frequently sampled when the solar arrays are generating power and is infrequently sampled when the spacecraft is running on its batteries. Since the value is needed continuously for the IFG data processing, these data also need to be fit with a continuous function. The electronics temperature profile is much more variable from day to day than the IFG sensor temperature profile so it is not a good candidate for modeling. Instead, the data points from each week are fit to a continuous function, a running polynomial fit is computes and this fit value (modelET) is stored and used in the data processing.

Figure 22: IFG Electronics temperature data and fits

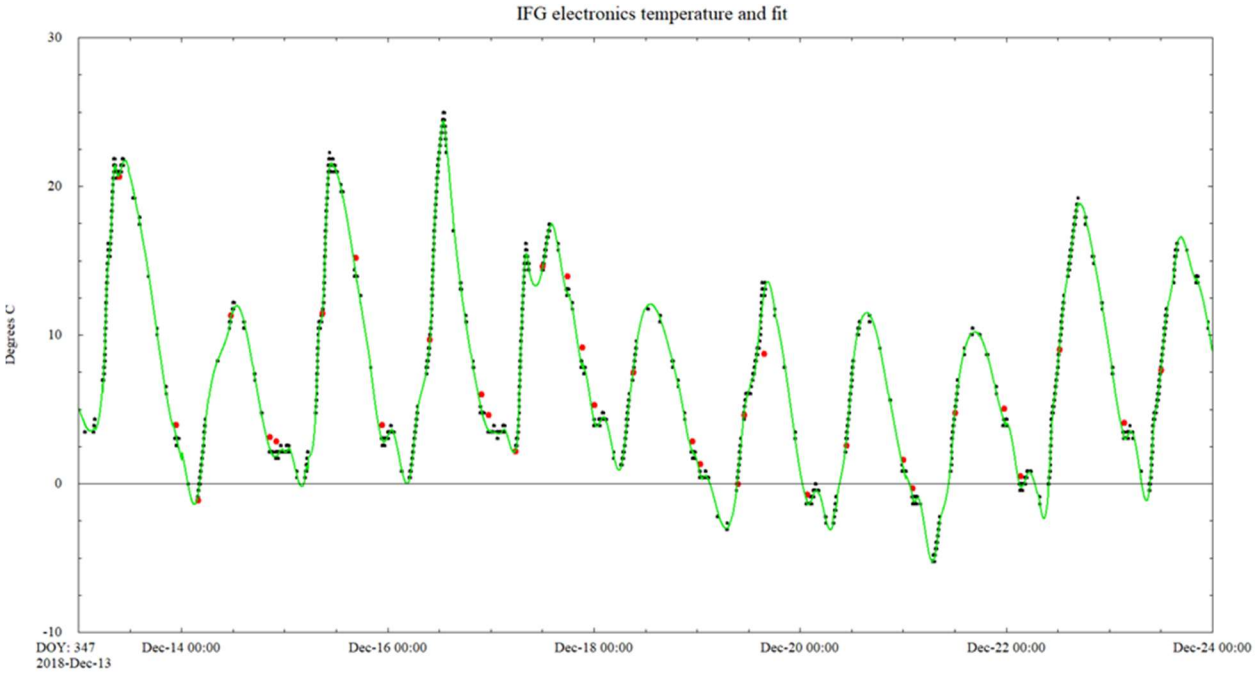


Figure 22 shows the few IFG electronics temperature samples as large red dots, the PAE box temperature (T-0014) values as smaller black dots, and the running fit to the PAE temperature values as the green trace for 11 days in December 2018. The green trace fits the black dots well when the temperature is varying smoothly but it smooths out high frequency fluctuations. The red dots lie along the green trace when the temperature is rising steadily for hours. However, when the temperature is cooling, the actual electronics temperature values are generally

displaced to the right of the green trace indicating that the IFG board is cooling more slowing than the electronics box where it is housed. Note that Figures 5 and 6 in the ground calibration section demonstrate that a few degrees of error in the electronics temperature estimate would only produce a small fraction of a nanoTesla error in the calibrated field for a field of 1000 to 2000 nT. The 2^3 place in the DQF value is always set to a value of 2 to indicate that the electronics temperature is approximated by a fit to the data.

Appendix 3: Fixed Solar Array Currents

The IFG data were shown to be correlated with the fixed solar array current values reported in SCEA channels E-0771 and E-0791. Unfortunately, like the PAE temperature data, these parameters are not returned continuously so the actual data values can't easily be used in the data processing pipeline. The shape of the fix solar array current versus MLST function has been found to be fairly stable in the time since the spacecraft landed so this parameter lends itself to modeling rather than fitting. There were some shape variations with time that occurred in early and mid-January 2019 so the data processing pipeline does use a few slightly different shape models for this time period. However, since Jan 16, 2019 the shape of the daily variation has been constant. There are amplitude fluctuations that occur periodically that require a scaling of the underlying shape model to improve the fit to the measured currents.

Figure 23: Fixed solar array current shape model, Dec 14-31, 2018

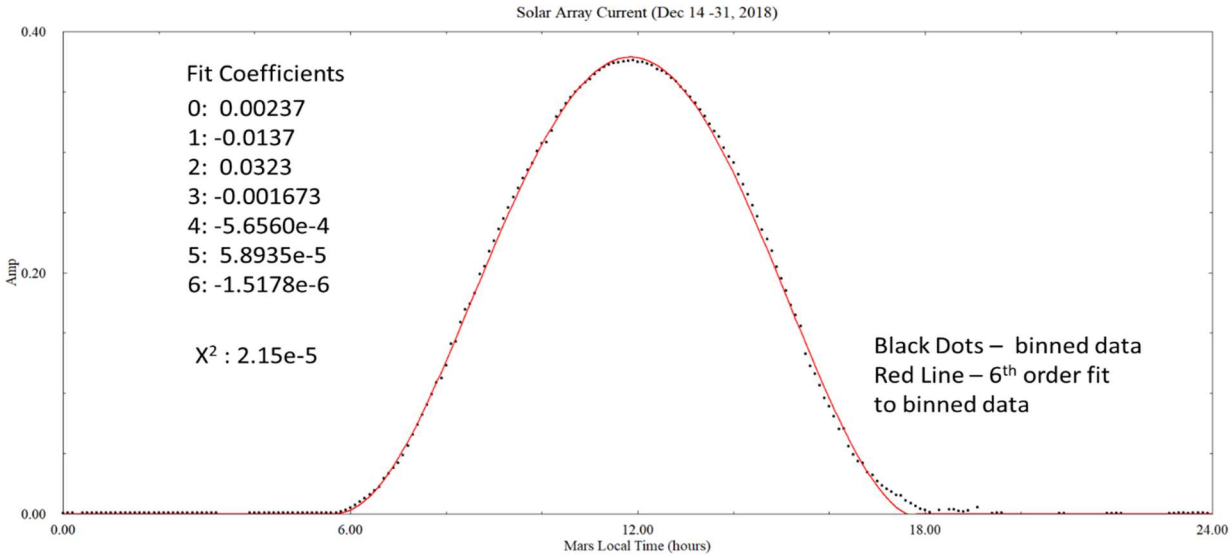
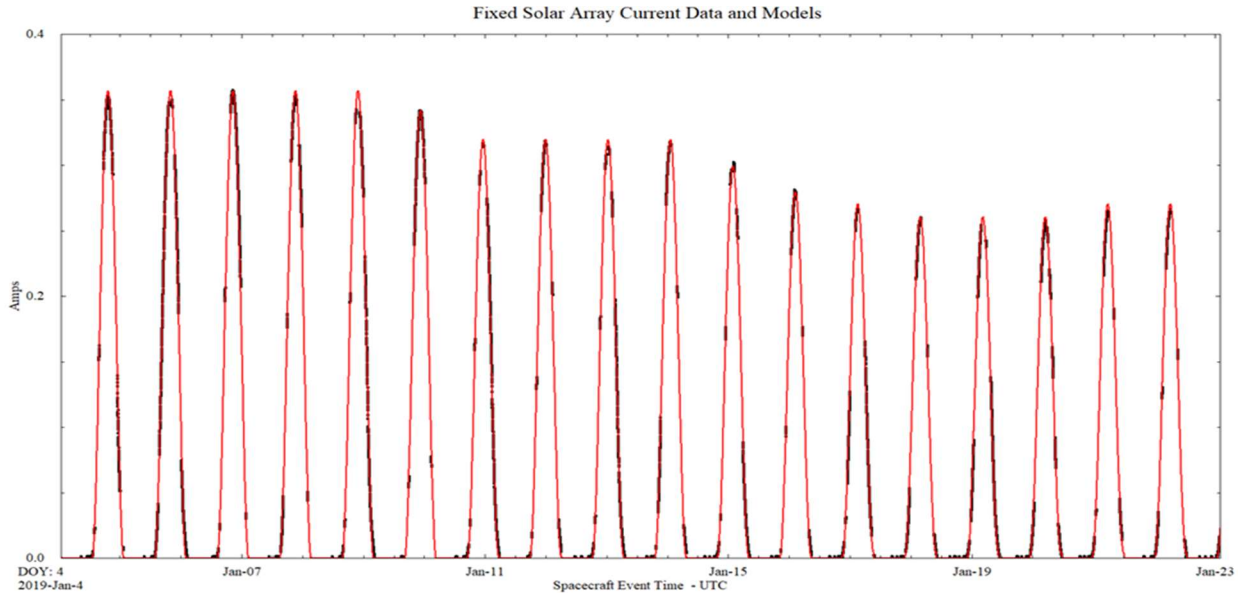


Figure 23 shows a 6th order polynomial fit to the data in channel E-0771 for the time period between December 14 and 31, 2018. The data were binned (black dots) and fit (red trace). The fit coefficients and chi-square value are given in the figure. Figure 24 shows the data (black dots) and model fit (red trace) for the time period in mid-January 2019, when the models were changing.

Figure 24: Fixed solar array current and models, January 2019



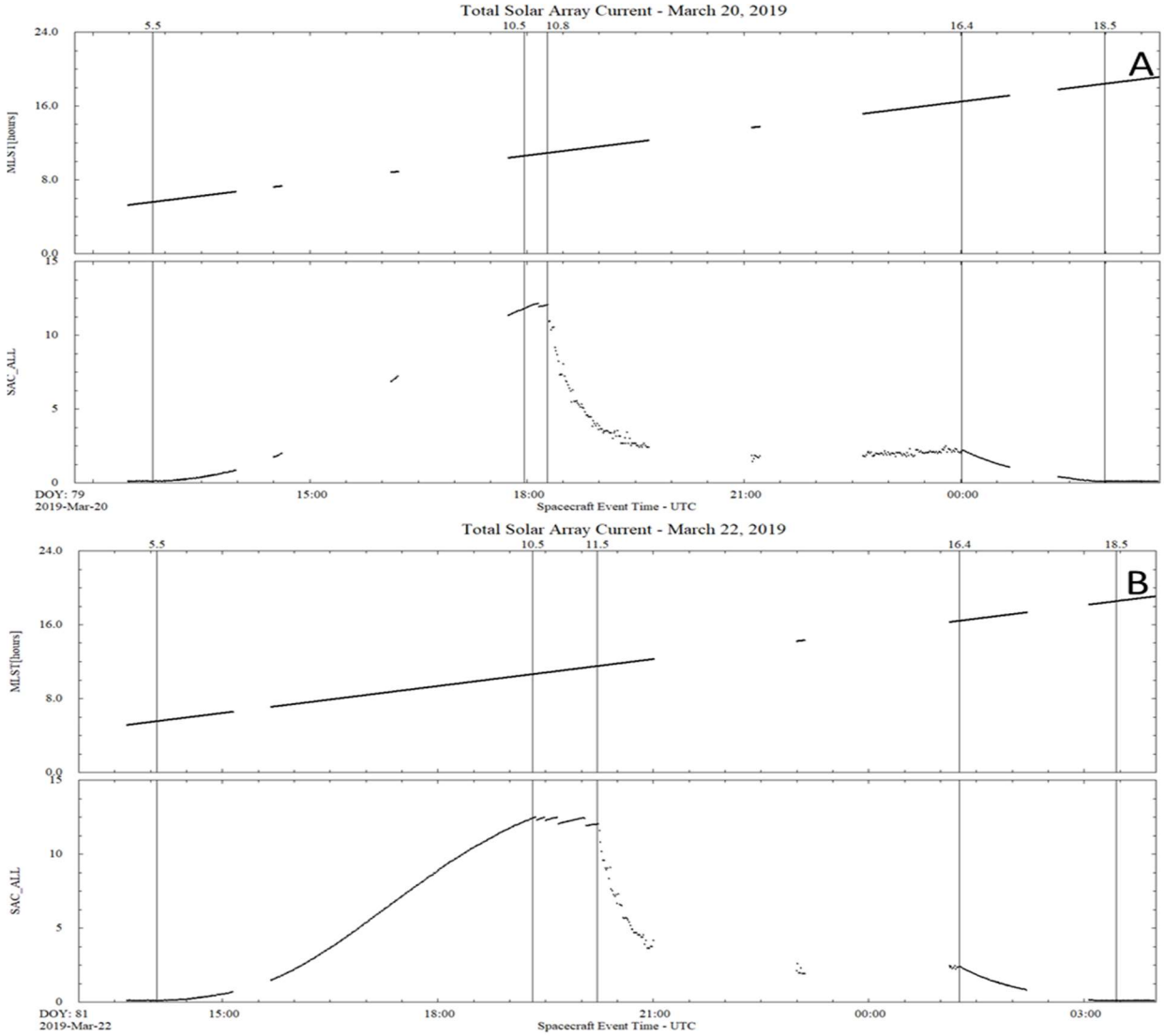
The 2^4 place in the DQF value is always set to a value of 3 to indicate that the fixed solar array current correction is computed using a model fit to the data.

Appendix 4: Total Solar Array Current

The term total solar array current is slightly misleading because the value actually computed and modeled (modSACT) is the total solar array current minus the fixed solar array current, which is a mouthful. Since the fixed solar array currents are modeled separately, it's the remained current that needs to be characterized. When the terms total solar array current or total current are used here, please understand that this is just a shortened way of describing what is a being computed which is the sum of the SCEA channels E-0772 and E-0792. Like the PAE box temperature, the total current data are not returned continuously, nor are the values sampled frequently enough to use them directly in the data processing pipeline.

Figure 25 shows total solar array currents for two days in March 2019 when the data were more continuous than they are normally. In each panel, the top trace is the MLST value and the bottom trace shows the total current data points. Both panels are marked with vertical lines at various values of MLST. The top panel (A) shows data from March 20 and the bottom panel (B) shows March 22, 2019. In both panels, currents start being measure a little before 5.5 hours MLST (first vertical line) and they begin to rise smoothly shortly thereafter. Both panels show the current reaching a maximum value near 10.5 hours MLST, although there is some variation between the two days. On the right side of the figure, both pans show a smooth decay of the current beginning at about 16.4 hours MLST and the current is near zero by 18.5 hours. However, the behavior of the current between about 10.5 and 16.4 hours MLST varies significantly between the two days. Both days show a lot of noise in the data in this time period and both show a rapid

Figure 25: Total Solar Array Current on two days in March, 2019



drop to an intermediate current level before decaying to zero at the end of the day. This behavior is normal for this parameter. There is a smooth rise in the early to mid-morning and an evening fall off that are consistent from day to day. However, the mid-day behavior differs in the details and timing but always shows a rapid decline to an intermediate level that is maintained until the evening. In order to model this behavior, the Mars day is split into 4 segments, morning (5.5-10.5), mid-day (11-16.4), evening (16.4 – 18.5), and night (18.5 to 5.5) hours MLST. The night time values are set near zero, the morning and evening values are binned and fit to functions of MLST that are fixed, and the mid-day period is binned and set to another function that is allowed to shift in MLST by as much as ± 0.7 hours MLST. Morning and evening models take precedence over the mid-day model and gaps in MLST coverage are spanned by linearly interpolating between functions. The timing and slope of the morning rise and evening fall-off changes occasionally so data need to be monitored regularly to see if a new base model needs to be computed.

Figure 26:Model of the total solar array current

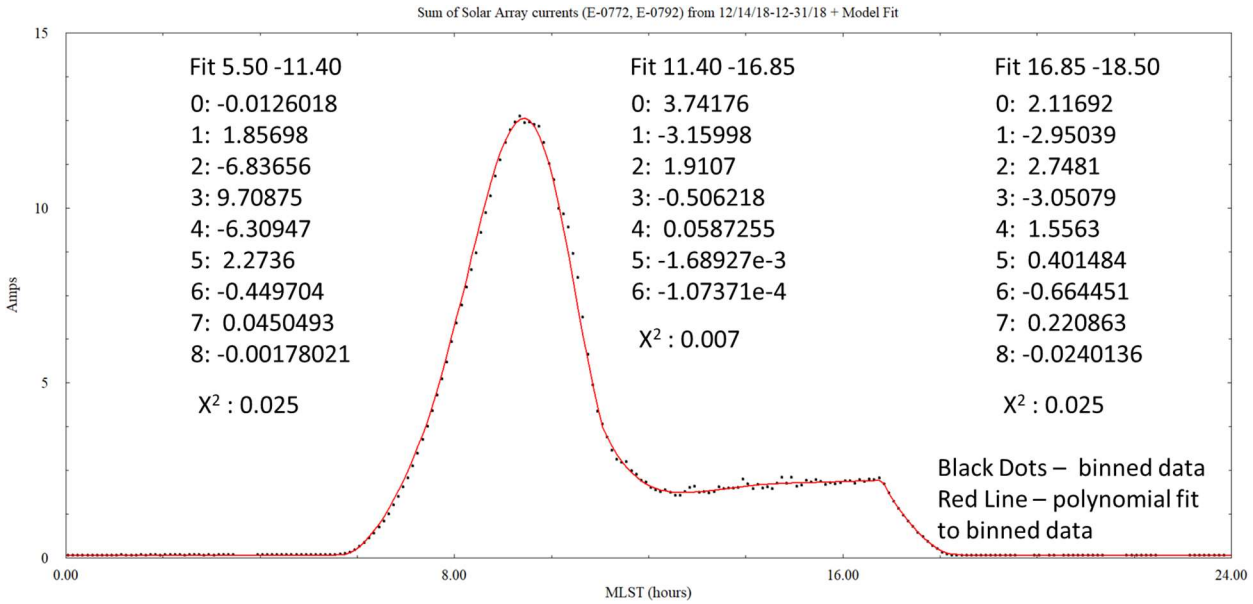


Figure 27: Total solar array current data and model

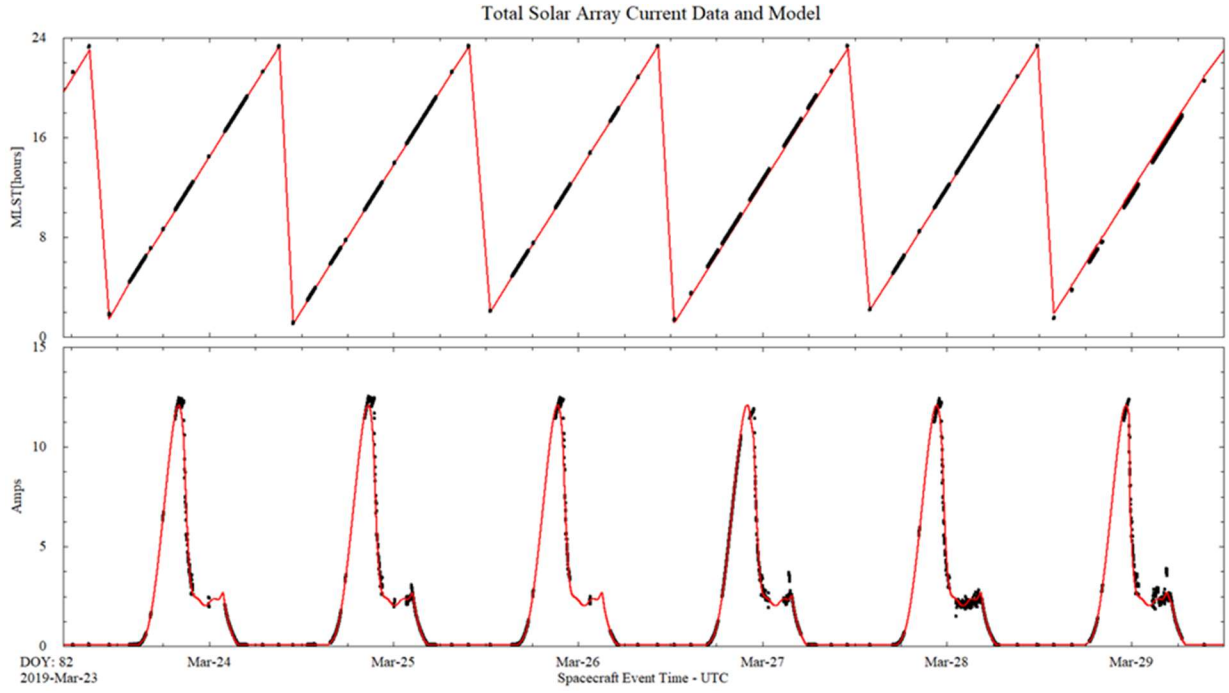


Figure 26 shows total current binned data (black dots) from Dec 14 to 31, 2018 and the morning, mid-day, and evening polynomial fits to the data. Note that mid-day and evening segments begin later in this time period that is does in March 2016. To date, there have been 5 different models, each covering a time interval between two and eight weeks. Once the base model has been

computed for each time interval, the daily MLST shift is computed. The shift that best fits the mid-day drop is selected and then the model total solar array current is recomputed using the shifted MLST times for the mid-day segments.

Figure 27 shows a bit more than five days of solar array current data (black dots, bottom panel) as well as the model current used to correct the IFG data (bottom panel, red trace). The top panel shows the actual MLST values as black dots and the red trace shows the shifted values used for computing the mid-day model segments. When the mid-day segment has been shifted relative to the morning and evening segments, the missing MLST coverage is computed by linearly interpolating across the gap.

Finally, the 2^5 place in the DQF value is always set to a value of 3 to indicate that the total solar array current correction is computed using a model fit to the data.