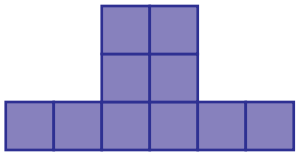
MIT CENTER FOR SPACE RESEARCH



ACIS FLIGHT SOFTWARE
PART # 36-56101

CCD BIAS LEVEL DETERMINATION ALGORITHMS

Rita Somigliana and Peter Ford

Massachusetts Institute of Technology
Center for Space Research

prepared for the ACIS Flight Software Critical Design Review

Version 2.1

June 19, 1995

1.0 Introduction

To distinguish X-ray events within a CCD pixel from background noise, it is necessary to know the pixel value that would be reported by the analog electronics in the absence of any event or background, a quantity known as the pixel's **bias level**. In high-quality CCDs that have not been subjected to appreciable radiation dosage, the i'th measurement of this value is closely approximated by a Gaussian random function $\exp[-(p_1-p_0)^2/\sigma^2]$ whose modal value p_0 and width σ are nearly identical for all pixels except for a small number of damaged pixels termed "hot" or "flickering" according to their anomalous bias level behavior.

During the course of a science run lasting from 10^3 to 10^5 seconds, all bias values are expected to vary with slow changes in the DC level of the analog electronics. These variations are compensated by "overclocking" the CCD, i.e. reading pixels from the frame store that never received charge from imaging pixels. The average value of the overclock pixels will directly measure the change in DC level, and can therefore be used to correct it.

Baseline pixel values may also change slowly across the CCD as a result of light leaks, temperature inhomogeneities, or other changes in the operating environment. It was originally thought that, after DC level changes, this would be the most serious systematic error in determining the bias levels, and it was intended to include special hardware in the ACIS Front End Processors (FEPs) to detect and compensate for these spatially varying changes in bias value. The algorithm divided each exposure frame into groups of 64×64 pixels and used their modal value, corrected for any change in average overclock, as an estimate of bias level for the corresponding pixels in the following exposure. It is described in detail in section 3.2.2.3 of Applicable Document 4.

The situation has changed with the analysis of CCD data from the ASCA mission, which uses devices that are similar in many ways to those to be flown on ACIS. As the report by Rasmussen makes clear (see Applicable Document 1), the pixel bias levels of ASCA CCDs changed significantly, presumably as a result of radiation damage, in a spatially random manner. Bautz has concluded (in Applicable Document 2) that this phenomenon will most likely affect ACIS CCDs, and would constitute the leading source of error in the measurement of X-ray energies, unless corrected by on-board calibration of each pixel. To this end, it has been decided to make significant alterations to the design of the front-end processor hardware (see Applicable Document 3):

- Remove the hardware that accumulates the 64×64 pixel histograms, along with their associated lookup tables.
- Add a radiation-tolerant **bias map buffer** in which to accumulate and store the 12-bit bias estimates for each pixel. Each FEP will therefore contain a 1.5 Mbyte bias map buffer in addition to its existing image frame buffer.
- Add hardware that will compare the corresponding bias map value against each pixel input from the DEA in order to determine threshold crossings.

The bias map is too large to be made of radiation hardened devices, so it is anticipated that it will suffer several SEUs per orbit, especially when AXAF passes through the Earth's radiation belts. The map will therefore be protected by a 1-bit parity plane so that the FEPs can detect SEUs and disregard the possibly false event thresholds. However, once an SEU occurs, that pixel is effectively lost to science until the bias map is recalibrated. We therefore anticipate the need to recompute the bias levels at least once per orbit, perhaps more often, and the remainder of this memorandum details our search for an algorithm or algorithms for computing the bias levels automatically from a series of calibration exposures.

We begin in Section 2.0 by describing the FEP hardware in greater detail, and discussing its relation to the DEA and Back-End Processor (BEP). Since many of the more robust bias determination algorithms require many bytes of temporary RAM per bias value, we pay especial attention to FEP memory usage. We also describe the operation of the redesigned pixel thresholder hardware, and the part played by overclocking in compensating for changes in the average threshold level.

The algorithms themselves are discussed in Section 3.0. They fall naturally into three classes, according to their need for additional RAM to save information about each pixel while they estimate its bias threshold.

- Algorithms that require no additional RAM beyond the 1.5 Mbyte image map buffer and bias map buffer. These algorithms execute in a time proportional to N, the number of calibration exposures they make.
- Algorithms that must store all values of a given pixel from N independent exposures in order to determine its bias threshold. These algorithms execute in a time proportional to N^2.
- Algorithms that need to store a constant amount of additional data for each pixel. They execute in a time proportional to N.

In general, we shall see that the most robust algorithms fall into the second class, i.e. those that are potentially the least efficient. It is therefore crucial to compare their behavior against realistic image data in order to select one or two for inclusion in the FEP flight software. We therefore asked the ACIS science team to provide us with several sets of test data, as described in Section 4.0, and copied them to optical discs (CD-R) with accompanying documentation.

We anticipate that the bias maps will be used by ACIS in two ways: either (a) they will start to be dumped to the ground immediately they are computed, or (b) individual bias values will accompany each event in the photon event lists that are downlinked during the science exposures. When an entire bias map is to be dumped, considerable time can be saved by compressing the telemetry stream. In Section 5.0 we discuss possible compression algorithms.

Section 6.0 reports on a comparison of some of the algorithms introduced in Section 3.0, operating on the test data described in Section 4.0, using several statistical methods to determine the "best" algorithm. At the same time, these tests have also proven to be highly effective at locating anomalous pixels within the CCDs, and we recommend that they be used for that purpose during all ACIS phases—test, calibration and operation.

The two statistical techniques that are used in this document, the “Analysis of Variance” (ANOVA) for the large-scale comparison of bias frames, and the “T Student” test used in Section 7.0 for estimating differences between bias algorithms operating on small groups of pixels, are summarized in Appendix A.

1.1 Applicable Documents

1. ASCA-D SIS Memo #557, “Residual Dark Distribution: Modelling the Corner Pixel Distribution, etc.”, A. Rasmussen, MIT.
2. ACIS Memo #PS-45, “Pixel-by-Pixel Bias Determination for ACIS?”, M. Bautz, MIT.
3. ACIS Memo, “DPA Hardware Specification & System Description”, D. Gordon.
4. ACIS Memo, “Pixel and Bias Bit Maps - Alternative memory Architectures”, D. Gordon, MIT.
5. MIT-CSR 36-01103, version 3, “ACIS Science Instrument Software Requirements Specification”.
6. MIT-CSR 36-02402, version 1, “ACIS Science Instrument Software Preliminary Design Specification”.
7. MIT-CSR Part #TBD, version 3, “ACIS Hardware Specification and System Description”.
8. “Statistics for Experimenters”, G. Box, W. Hunter, and J, Hunter, Wiley, New York, 1978—tabulation of $F_{\alpha}(r,s)$ function values.

2.0 Problem Requirements

It is expected that Science observations will be conducted in two phases:

- *Bias Level Calibration*- the system collects data for several consecutive exposures with the purpose of calculating bias thresholds for all CCD pixels. The calculation uses the values reported for the same pixel in a certain number of frames, which may, or may not come from consecutive exposures. At the end of calculation, new bias threshold levels are stored and used in future event detections.
- *Event detection* - the system collects science data frame by frame. For each frame, the system locates pixels whose data values exceed the established thresholds, from which it generates an event list. In “bright” mode, optimized for event throughput, the system subtracts the thresholds from the raw data, inspects neighboring pixels to identify “split” events, rejecting those events that do not belong to a set of “grade” patterns, sums charge if split among event pixels, and reports the results. In “faint” mode, optimized for energy resolution, the system downloads the detected events along with the thresholds. These thresholds consist of three terms which are added together for each pixel:
  - *Uplink Threshold*, one value for the entire science observation, defined by explicit uplink command as part of the observation parameter block.
  - *Overclock Levels*, four values for each frame, representing the average DC level in each output node of the CCD, calculated from the previous frame.
  - *Bias Levels*, one value for each pixel in the frame, defining the zero-level of that pixel, calculated at the beginning of the observation and possibly recalculated at intervals by ground request.

In practice, the bias levels will be stored as 12-bit *unsigned* integers. To prevent them taking negative values, the overclock levels will be treated as *signed* integers, representing the difference between the DC level of the current exposure and that of the first calibration exposure.

Figure 2-1 represents the flow of data flow between the principal components of the data collection system that are involved with pixel bias levels. Control commands are omitted.

2.1 Focal Plane

This is an array of 6 CCDs that detect X-rays and ionizing particles. Each CCD generates a set of analog signals whose voltages are proportional to the charge deposited by the events. The noise sources are summarized in the following equation taken from Applicable Document 2:

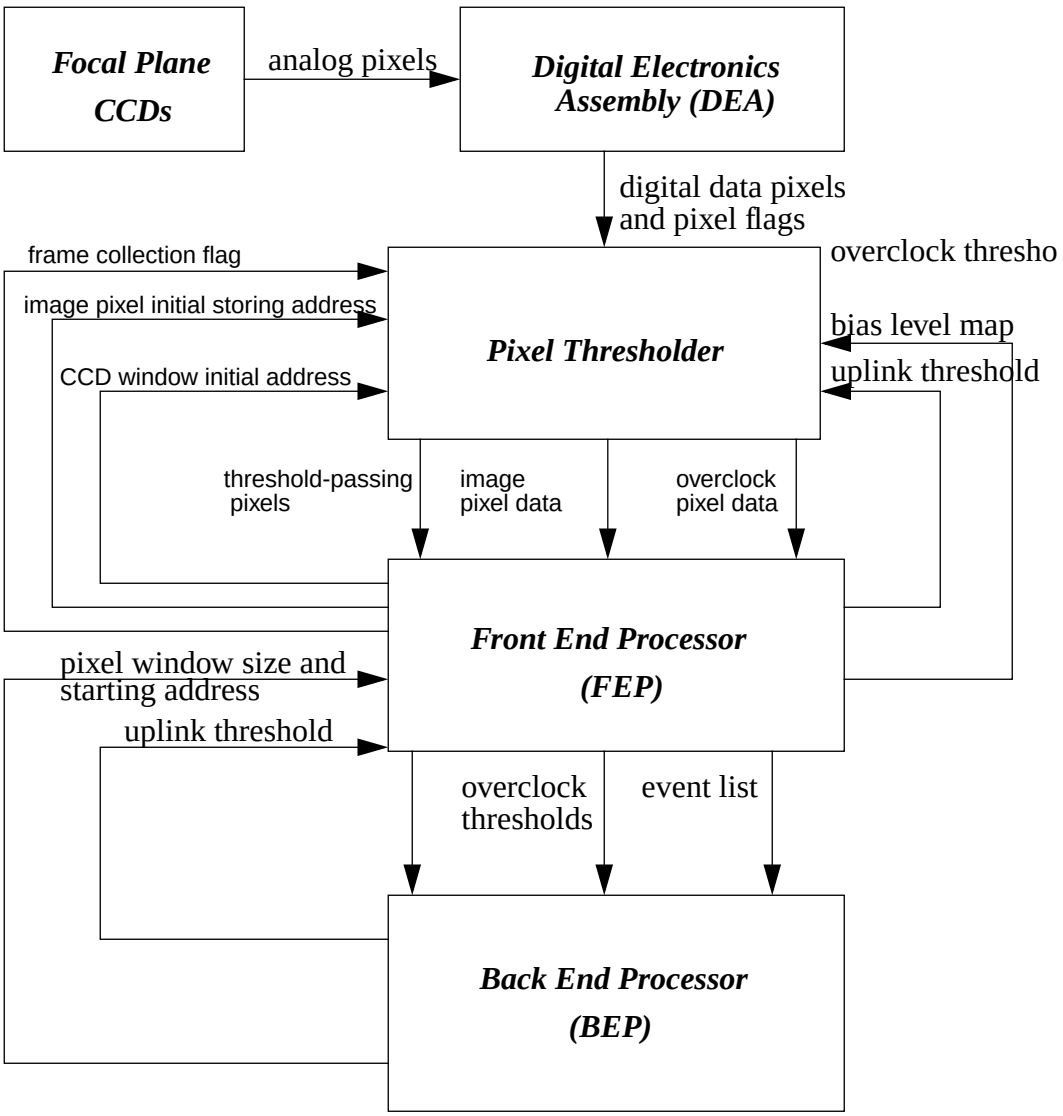
$$b_i(x, y) = dc_i(x, y) + p_i(x, y) + r_i(x, y) + n_i(x, y) \quad (2.1)$$

where x,y is the column, row coordinate, i is the frame index, and

- $b_i(x, y)$ = bias level.

- $dc_i(x, y)$ = dark current, which depends on temperature, integration time, and pixel number, but only varies slowly with time.
- $p_i(x, y)$ = contribution from charged particles and diffuse X-ray events; $p_i(x, y)$ is a random variable with a parent distribution that varies slowly with x and y, but which does not have zero mean.
- $r_i(x, y)$ = other spatial variations in bias level which may be time dependent (e.g. from light leaks).
- $n_i(x, y)$ = DC level which may be approximated by a zero-mean, normally distributed random variable with no correlation from pixel to pixel.

FIGURE 2-1. ACIS Data Flow involving Bias Levels



2.2 DEA (Detector Electronic Assembly)

Each of six DEAs sends “clocking” signals to a CCD to transfer its image array (1024×1026 pixels) to its on-chip frame store (1024×1024 pixels), then drains the frame store into four serial shift registers and reads each register into an analog-to-digital converter. Each converter generates a 256×1024 array of 12-bit digital pixels, and transfers them to the “Pixel Thresholders”, accompanied by flags that indicate the first pixel of a new exposure and the presence of overclock pixels. In addition, each DEA provides CCD clock synchronization, analog pixel amplification, and removal of noise created by the clocking signals.

2.3 Pixel Thresholder

A“Pixel Thresholder” stands between each DEA and its corresponding FEP accepting digital data from the DEA and write them to the FEP’s image buffer. Flags within the data indicate the end of each row and frame, and the base address of the “Image Pixel Map Buffer” can be selected by the FEP software. Overclock pixels are also flagged by the DEA and written to a separate buffer in the FEP. The processing of image and overclock pixels is controlled by a collection flag—if deactivated, the thresholder ignores all pixels until the next start of frame. This property allows the FEPs to perform calculations that take longer than the frame readout time, since it prevents the image map from being overwritten until the FEP has finished examining it.

Digital pixels, starting at the “CCD Frame Initial Row” and continuing to the end of the frame, are transferred to the image buffer, starting at the “Image Buffer Initial Row” and continuing for a pre-determined number of rows (or until the physical end of the Image Buffer). The “CCD Frame Initial Row” and “Image Buffer Initial Row” are determined by the contents of FEP hardware registers that must be set before processing a new frame.

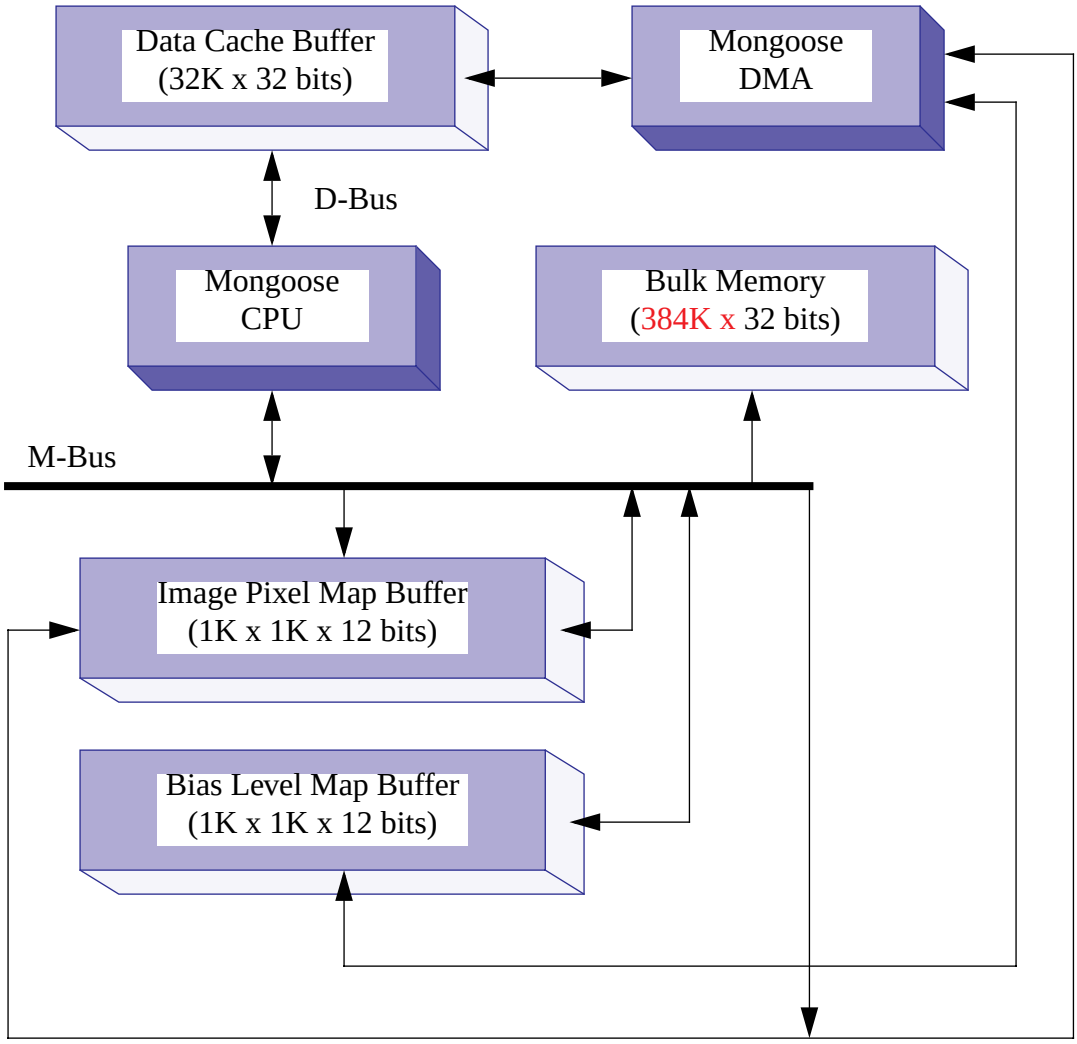
In a similar fashion, the overclock pixels are transferred from the thresholders to the overclock buffer starting at an initial CCD address for a requested number of rows (up to a maximum of 1024) and columns (up to a maximum of 32). The overclock data are stored in an FEP memory buffer starting at a specified address. The CCD starting address, the number of rows and columns to transfer, and the FEP buffer address, are all specified by FEP hardware registers that must be set before processing a new frame (see Applicable Document 5).

The FEP hardware thresholder transfers each DEA pixel to the image pixel buffer and sets the corresponding bit in the T-Plane buffer indicating whether the pixel value is greater than the threshold defined by the sum of the corresponding overclock level, the calibrated bias level, and the uplink threshold. In the event of a parity error in the bias level buffer, the pixel thresholders will set the threshold bit and assert a memory upset flag. It will be the responsibility of the FEP software to detect and correct the condition.

2.4 Front End Processor (FEP)

One FEP is assigned to process data from each DEA and from each CCD device. The major functions of the FEP are to calibrate the bias levels, calculate the average overclock values, and generate the photon event lists. Since each image frame consists of up to 1.5 Mbytes of data, and the frame readout time is approximately 2.65 seconds, FEP operations are time critical. The FEP itself consists of a radiation hardened MIPS “Mongoose” CPU and a variety of random access memory modules, as shown in Figure 2-2.

FIGURE 2-2. Mongoose Memory Organization



Several blocks of memory are available. Some are dedicated memory blocks, others are buffers in a bulk memory. They are the following:

- Image Pixel Map buffer - dedicated radiation-tolerant memory to contain the raw pixel data. Buffer size is 1K x 1K x 12 bits = 1.536 MBytes
- Bias Level Map buffer - dedicated radiation-tolerant memory to contain the bias level thresholds. Since this RAM isn't immune to single event upsets, and the bias levels will depend on exposure time and details of CCD clocking, it will be necessary to recalibrate the bias levels at frequent intervals.
  Buffer size: 1K x 1K x 12 bits = 1.536 MBytes
- Data Cache Buffer - dedicated radiation-hardened memory, which is used for calculations. Buffer size: 128 kBytes.
- Bulk Memory - a contiguous space of radiation-tolerant memory that may be used for multiple purposes, buffer size: 1.5 MBytes. This memory is subdivided into the following areas:
  - Overclock buffer - containing the overclock pixel values selected from the DEA by the pixel thresholders, maximum buffer size = 4 x 32 x 1024 x 16 bits = 256 KBytes, where: 4 = number of CCD shift registers (output nodes); 256 = maximum number of pixels readout per row for each shift register; 32 = number of overclock pixels for each row; 16 = pixel size
  - T-Plane Buffer - contains flags indicating which data pixel values are over the thresholds. Each data pixel is represented by one bit in the T-Plane.
    Buffer size = 1K x 1K x 1 bit = 128 KBytes
  - Parity Buffer - contains a single bit representing the parity, even or odd, of the corresponding bias level value, buffer size = 1K x 1K x 1 bit = 128 KBytes
  - Overclock Level Buffer - containing the results of the overclock level calculation. Buffer size is 4 x 16 bits = 8 bytes.
  - Free Memory Size - is about 1192 KBytes

In addition, the following thresholder registers are mapped into the processor's memory space:

- Uplink Threshold - contains the commanded value to use when computing the T-Plane values for incoming data pixels, size is 4 32-bit words, one for each DEA output node.
- Pixel Window Information - contains the pixel window size, starting CCD row, and starting Pixel Map row to use in the calculation, size is 12 32-bit words.

Access to 32-bit words in Cache Memory typically occurs within one machine cycle (0.1 µsec). Access to bulk memory takes much longer, 0.4–0.5 µsec. The FEP processors also contain DMA controllers capable of transferring data between bulk memory and cache at about 1 µsec per word plus setup time.

The remainder of this section describes more fully those FEP features that affect the Bias Threshold Algorithms. Their operation is shown in Figure 2-3.

2.4.1 Overclock Level Calculation

This function calculates the average pixel DC level on a frame by frame basis. The overclock levels are stored in the “Overclock Level Buffer”. Their values are updated each time a new frame is received. Four averages are taken, one for each CCD shift register. The results are used by the pixel thresholders to compensate the pixel data for this noise source.

2.4.2 Bias Level Calculation

This function calibrates the zero-event pixel levels and stores them in the “Bias Level Map Buffer”. The calculation is executed for each pixel in the array. At the same time, the parity of each bias level is calculated and stored in the “Bias Parity Buffer”. This parity is used by the pixel thresholders to detect isolated bit flips in the bias buffer. A damaged bias value will be detected by the thresholder and cause the corresponding pixel to be flagged in the T-Plane so that the FEP software can report it.

The data used for the bias level calculation are the raw image pixels generated by the pixel thresholder from a sequence of calibration exposures, corrected for any change in the node-average overclock levels. More than one algorithm may be used for these calculations, depending on the memory and time available, and the accuracy desired.

Memory constraints: The bias calculation will use pixel values from multiple exposures and may require more FEP memory than is available to store them all simultaneously. This problem can be resolved by dividing the CCD into strips and repeating the bias calculation for each strip, storing the bias threshold values into the corresponding locations in the bias buffer.

Calculation time constraints: the calculation may require more CPU time than the interval between calibration exposures. This problem can be resolved by de-activating the pixel collection flag until the current exposure frame has been processed.

It is anticipated that the resulting bias levels will be downloaded via the ACIS science telemetry stream. At 24 Kbits/sec, it would take over 500 seconds to download each FEP bias map, but it is expected that simple data compression algorithms (see section 4, below) will reduce this time very considerably.

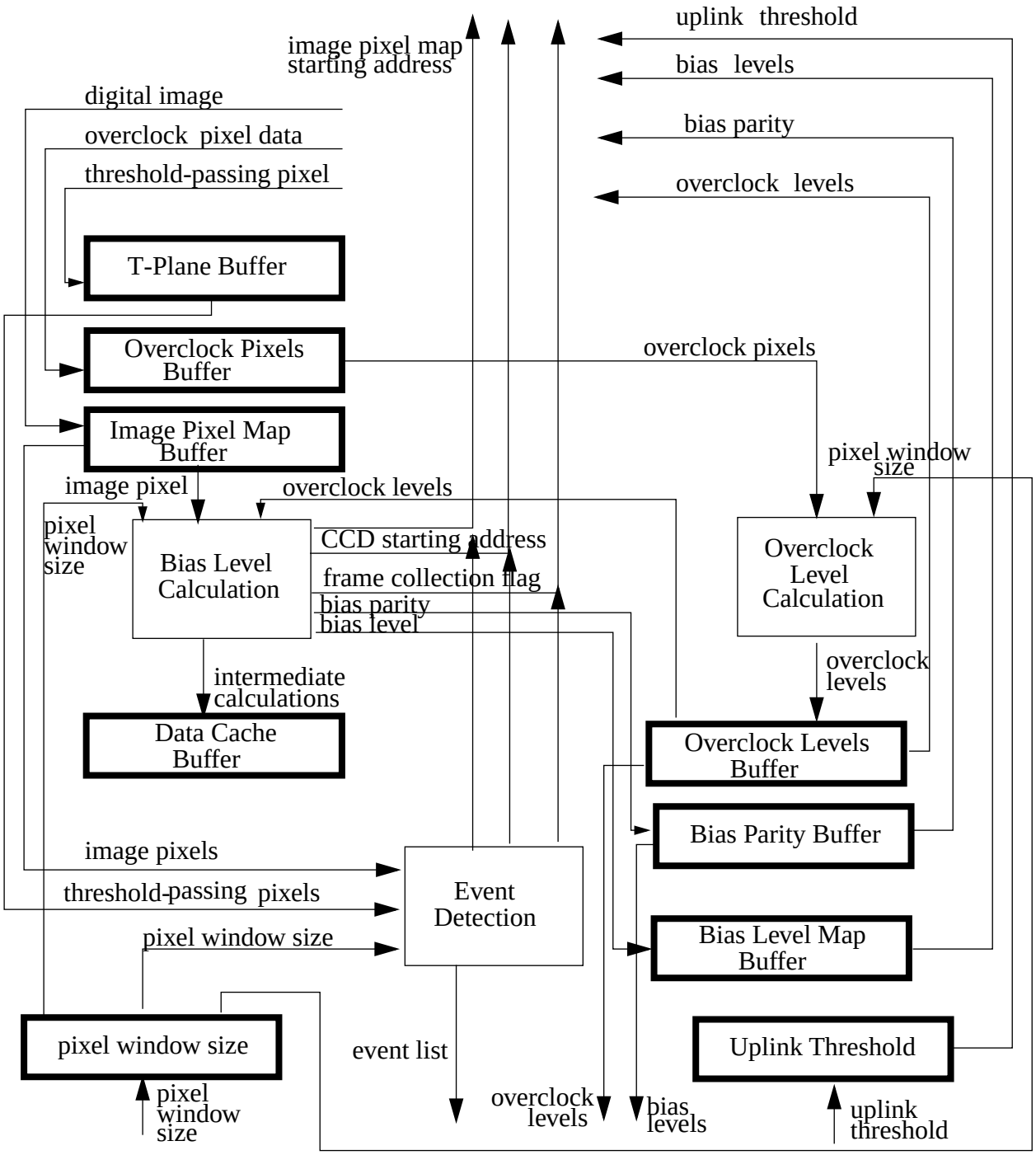
2.4.3 Event Detection

During a science observation, a program executing in each FEP will examine the threshold crossings indicated in the T-Plane, determine whether that pixel represents a local maximum value, and, if so, add it to an event list than is subsequently read by the BEP.

2.5 Back End Processor (BEP)

A single BEP controls the higher-level ACIS functions, i.e. accepting uplink commands, generating code for DEA clocking, loading code into the FEPs, scheduling exposures, and collecting and packaging event and housekeeping data for transmission to the downlink telemetry system.

FIGURE 2-3. Front End Processor Data Flow



3.0 Bias Algorithm Definitions

All the algorithms under consideration assume that the CCDs are to be clocked and exposed in the same manner as in the subsequent science run. The algorithms are applied to every pixel in each CCD. They must operate on the pixel data collected during the calibration exposures and cannot use any per-pixel information transmitted from the ground because the bias threshold levels are assumed to be time dependent and their calibration requires recent pixel data.

All algorithms assume that raw pixel values have been corrected for changes in analog level using the average overclock values from the most recent exposure. This makes the bias levels less sensitive to slow drifting of the DC component of the DEA preamplifier response. However, it is not possible simply to subtract the average overclock from the raw pixel values since this would lead, in some situations, to negative bias levels, which could not be stored in the 12-bit bias map. To avoid this possibility, the overclock corrections δO_i for output node i are defined by the difference between the overclock values saved from the *first* calibration exposure $\overline{O}_i$ and those of the current exposure O_i, i.e.

$$\delta O_i = O_i - \overline{O}_i \tag{3.1}$$

For the remainder of this document, it will be assumed that “raw” pixel values will already have been corrected for drift in average overclock by subtracting δO_i.

All algorithms are defined by a set of input parameters which must be uploaded from the ground:

- Number of calibration exposures, N
- CCD array window size, usually either 1024 or 1024/N, depending on the algorithm
- Method used to characterize the bias level, e.g., average, median, modal, etc.
- Rejection threshold levels and confidence measure for the rejection of outlying pixels, e.g., *n* times the variance, *m* times the inter-quartile range (IQR), etc.
- Termination rules, e.g., all bias values must lie within an acceptance range, maximum number of iterations reached, etc.

All algorithms will generate outputs consisting of 12-bit bias levels which may be compressed and downlinked to the ground. Some may also generate individual rejection levels, pixel distribution widths, etc., which will not be reported.

With these broad features in common, the algorithms can be subdivided into three major groups:

- Those not requiring additional storage—they execute rapidly, but are generally less accurate, particularly when encountering “pathological” pixels.
- Those that must store the value of each calibration pixel—they execute more slowly than the previous group because they cannot operate on the full area of the CCD at once, but they are generally more accurate, especially in their treatment of pathological pixels.
- Algorithms requiring some additional storage, but not for all input values—their execution

speed and accuracy is intermediate between the other two groups.

3.1 Algorithms Requiring No Additional Storage

We have investigated a class of algorithms that possesses the considerable advantage that they use no additional FEP memory beyond the Image and Bias Map Buffers and a modest amount of Data Cache. At any moment during the course of the algorithm, each pixel is represented by only two quantities—the 12-bit value in the Image map which will be overwritten by each fresh exposure, and the 12-bit value in the Bias Map whose value must converge to the desired bias threshold level.

The distribution of values, p_i, of a pixel observed N times ($1 \le i \le N$) may be approximated by a narrow Gaussian, $\exp[-(p_i-p_0)^2/\sigma^2]$, with the addition of a few outlying values, especially those with $p_i >> p_0$ produced by X-rays or ionizing particles. p_0 is, by definition, the modal value and σ ($<< p_0$) is the width, typically a few data units. For new CCDs, not yet subjected to radiation damage, both p_0 and σ are nearly identical from pixel to pixel, except for a few pathological "hot" or "flickering" pixels.

When a CCD is damaged by radiation, it is anticipated that p_0 and σ will increase with time, and at rates that vary randomly from pixel to pixel. Such behavior has been reported in Applicable Document 1 from a study of the CCDs aboard the ASCA observatory. Algorithms that use no additional storage are unable to estimate σ from p_i, so they concentrate on the modal value p_0. Once this has been computed for all pixels, its variance $(<p_0^2>/N)-<p_0>^2$ serves as a good approximation to $<\sigma^2>$.

3.1.1 Pixel Conditioning

The algorithms start by examining a series of exposures to "condition" the p_i values and derive b_i, an estimate of p_0. After the first exposure, the best estimate of b_1 is clearly p_1. After the i'th exposure, two possible algorithms yield improved values of b_i, *viz.*

$$b_i = \min(b_{i-1}, p_i) \tag{3.2}$$

which guarantees that, after a number of exposures, none of the b_i will retain anomalously high values from X-ray or ionizing particle events. The resulting b_i will generally, of course, be less than p_0 by a few σ, and it will be seriously compromised if even a few of the p_i possess anomalously *low* values. A rather more accurate conditioning can be achieved by the following two-step algorithm:

$$b_i = p_i \text{ if } p_i < (b_{i-1} - T_1), \tag{3.3}$$

and then,

$$b_i = C\,p_i + (1-C)\,b_{i-1} \text{ if } p_i < (b_{i-1} + T_2), \tag{3.4}$$

otherwise,

$$b_i = b_{i-1} \tag{3.5}$$

i.e. if the new value p_i is much less than the previously stored value b_i, let it replace the stored value. Otherwise, if the new value isn't much larger than the stored value, use a running average to better approximate the "mean" pixel value, b_i. Optimum choice of the thresholds T_1 and T_2, and of the partitioning factor C, will depend on the expected pixel distribution functions themselves. The effect of anomalously low values can be mitigated by applying a *modal* filter immediately after the conditioning phase, e.g. if any bias value is found to be less than all of its neighbors by more than some constant, the value should be replaced by the median of the 8 surrounding values.

3.1.2 Estimate of the Modal Value

After N conditioning exposures, the Bias Map Buffer contains values b_N that are guaranteed to lie within a few σ's of the modal values p_0. The b_N can now be used as rejection thresholds to identify X-ray and ionizing particle events in subsequent calibration exposures, and thereby improve their own values in the process. This is achieved by making a further set of M exposures. The pixel values p_i , $0 < i < M$, are subjected to the following two-step algorithm:

- For each pixel in the fresh image buffer, set

  $$p_i = 4095 \text{ if } p_i > (b_{i-1} + T_3) \tag{3.6}$$

  where T_3 is a threshold value. In addition, the 8 neighbors of any pixel that meets this criterion may also receive a contribution from the event that caused the central pixel to lie above the threshold. These neighboring pixels are therefore also reset to a value of 4095.

- The Image Map pixels are re-examined and those with values less than 4095 are used to refine the Bias Map values,

  $$b_i = C\, p_i + (1 - C)\, b_{i-1} \tag{3.7}$$

  Although this algorithm guarantees that b_i will converge to the neighborhood of p_0 for all positive $C < 1$, b_i will eventually execute a random walk with width $\sim C\sigma$. In practice, the choice of C must be balanced against the exposure count M to optimize the width of the b_{N+M} distribution versus total calibration time $t \propto M + N$.

3.1.3 Discussion

The principal advantage of this class of algorithms is that they are very simple to implement—just a few lines of computer code—requiring no DMA transfers between Image Map and D-Cache, eliminating the need to reset the Initial Image Store address or the CCD Window Initial Address. Preliminary tests show that the resulting bias maps compare favorably with those generated from "strip–by–strip" algorithms. However, they do have some weaknesses, as follows:

- They converge on p_0, the modal pixel values. This would be satisfactory if the widths σ were identical for each pixel over the lifetime of the CCD since a bias threshold value of, say, $p_0+3\sigma$ would be a robust discriminator of energy deposition. However, this is not even true of undamaged CCDs—the width of flickering pixel values can be many times the average σ,

and the situation is expected to deteriorate as the CCDs absorb high doses of radiation. The choice of bias threshold, p_0+T_b, must be made carefully—if T_b is too low, flickering pixels will appear as events; if too high, low-energy events will go unreported.

- Since all parts of the final Bias Map are calculated after the last calibration exposure, there is no opportunity to copy the bias values to the downlink telemetry system while the bias is being calculated. The time required to compress and transmit the six Bias Maps, ca. 12 minutes assuming 75% compressibility (see Section 5.0), must therefore be added to the bias calibration time.
- These algorithms rely on particular properties of the pixel distributions. For instance, if equation 3.2 or 3.3+3.5 are used to condition the bias values, the presence of anomalously low pixel values, $p_i << p_0$, will start the second phase of the algorithm in a poorly conditioned state and the number of subsequent exposures, M, may be insufficient for equation 3.7 to change b_i so that $|b_i–p_0| <= C\sigma$, as desired. This situation can be alleviated by applying an additional "grading" process at the end of the conditioning phase in which the anomalously low values of b_N are identified and replaced by better estimates, e.g. by the median of the b_N of the surrounding pixels.

3.2 Algorithms Requiring Additional Storage

These algorithms need to store all calibration pixel values before determining the optimum threshold level. They require that the FEP analyze the CCD image in strips, copying several exposures of the same strip into separate portions of the image buffer, calculating the bias threshold levels, transferring them to the equivalent locations within the bias buffer, and then repeating the same procedure for the remaining strips. If s is the strip size and N is the number of exposures required to determine the bias of a single pixel, the storage needed for the samples is sN, and this will be re-used for the remaining strips. The full calibration time T_c is given by rNT_f, where r is the number of strips, and T_f is the duration of a single exposure frame. For example, T_c for N samples collected by dividing the CCD image into N strips and storing the N samples for each strip in portions of the "Image Pixel Map" buffer is proportional to N^2. Note that under usual CCD clocking conditions, the minimum value of exposureTime is about 2.67 seconds—the time taken to transfer the contents of the full frame store through the output nodes.

The following sub-sections examine a number of algorithms in this class. They use the following terms:

- $\text{overclockCorrection}_j$ = the average overclock correction for the j'th output node
- x_i = (raw pixel value) – $\text{overclockCorrection}_j$
- N = number of exposures used to characterize each pixel
- k = a confidence-level measure, typically setting the rejection threshold to be $k\sigma$, where σ is the width of a pixel's distribution.

3.2.4 Truncated Mean - Rejection Based on σ

This algorithm calculates the mean and standard deviation σ, of N values of a pixel. It then rejects values that diverge from the mean by more than a specified σ (e.g. those 3σ higher or lower than the mean, equivalent to a confidence level of 99.73%), and then recalculates the mean and σ.

$$\text{Definition: } \bar{x} = \frac{1}{N}\sum_{j=1,N} x_j \tag{3.8}$$

$$\sigma^2 = \frac{1}{N-1}\sum_{j=1,N} (x_j - \bar{x})^2 \tag{3.9}$$

$$\text{Rejection rule: } |x_j - x| \geq k\sigma \tag{3.10}$$

The algorithm is executed by successive iterations. The mean and σ are calculated, then the outliers are rejected and the mean and σ are recalculated. The algorithm is repeated until the termination rule is satisfied. The value *k* can be thought of as an estimated value derived by Gaussian distributions; e.g. *k* = 2 implies a probability of 5% that good pixel values are being rejected. The termination rules are defined as any of the following:

- All pixel values are within an acceptance range.
- The maximum number of iterations was reached.
- The number of remaining pixels is less or equal to the minimum number accepted.

The algorithm is expected to work best with Gaussian or symmetric heavy-tailed distributions in the presence of outlier points for which the assumption of normality does not hold. These outlier points are rejected and the remaining data are treated as Gaussian.

The calculation is speeded up by saving the intermediate values: $\sum x_j$ and $\sum x_j^2$. During the rejection phase, rejected values are subtracted from $\sum x_j$, $\sum x_j^2$ and the sample number is updated.

3.2.5 Median

The algorithm calculates the median of N pixel values. It sorts them in ascending order, and identifies the central value, thereby rejecting the outliers since they will be sorted to one end of the list or the other. Once the $\{x_n\}$ are sorted, the formula for the median is:

$$x_{med} = x_{\frac{(N+1)}{2}}, \qquad \text{N odd} \tag{3.11}$$

$$= \frac{1}{2}\left(x_{\frac{N}{2}} + x_{\left[\frac{N}{2}+1\right]}\right), \qquad \text{N even} \tag{3.12}$$

A suitable rejection rule is based on the interquartile range (IQR) which indicates the distance between the upper and lower quartile:

$$x \le x_{med} - k\ IQR \quad \text{and} \quad x \ge x_{med} + k\ IQR \tag{3.13}$$

A value of $k = 3$ can be considered a good approximation to reject severe outliers. The termination rule for the median is defined as follows:

- All pixel values are within an acceptance range.
- The maximum number of iterations was reached.

The algorithm is expected to work best with Gaussian or symmetric heavy-tailed distributions with few outlier points which have large σ, but zero mean. It fails if the area in the tails is large[1], and is somewhat influenced by points with very large values, but the rejection of contaminated sources is particularly simple.

3.2.6 Modal (Peak)

Modal algorithms calculate the maximum of the pixel distribution. They collect N exposures, generate a histogram whose bin size is a fraction of the total range of pixel values, and identify the peak bin. They are expected to work best with distributions that possess a single sharp maximum. Outlier points are easy to eliminate, but the number of samples must be large enough to permit a histogram to be constructed, or else the sparse histogram must be smoothed in order to identify the modal value.

A preliminary evaluation of modal algorithms indicates that, for a given level of bias threshold accuracy, this class requires a larger number of exposures than the other algorithms in this group.

3.2.7 M-Estimators

These algorithms perform maximum-likelihood estimates of fitting the distribution to a class of functions derived from a particular regression model. The residuals are approximated by a weighted least squares (WLS), and the model parameters are fitted by regression. These algorithms are applied iteratively until there is a negligible change from one iteration to the next.

M-Estimator algorithms are expected to work best with skewed distributions or distributions in which the outlying values are important. The measurement errors are not distributed normally, but have heavier-than-normal tails. Their probability density depends on their residuals, scaled by weight factors which can be assigned to each point. Several probability density models can be used to estimate the weight, each one reflecting a different behavior among the outliers.

3.2.7.1 Weight Mean Estimator with Tukey Functions

This class of algorithms calculates the mean and σ^2 and then uses Tukey bi-weight functions to estimate the weights. These are defined by:

1. Press, William H., etc. “Numerical Recipes in C” 14.1 “Median and Mode”

$$w(z) = z\left(1 - \frac{z^2}{c^2}\right)^2 \quad , |z| < c \tag{3.14}$$

$$w(z) = 0 \quad , |z| > c \tag{3.15}$$

where: $z(x_i) = (x_i - m)/\sigma$ is the scaled residual, x_i is the observed value. m and σ are the estimated mean and standard deviation, respectively, and c is a suitably chosen threshold. The algorithm is defined by the following steps:

a. Calculate the mean and σ^2

$$\bar{x} = \frac{1}{N}\sum_{i=1,N} x_i \tag{3.16}$$

$$\sigma^2 = \frac{\sum_{i=1,N} (x_i - \bar{x})^2}{N-1} \tag{3.17}$$

b. Calculate the weight mean estimator, scaled residuals, weights, and weighted mean, $\bar{W}$:

$$z(x_i) = \frac{(x_i - x)}{\sigma} \tag{3.18}$$

$$w(x_i) = z(x_i)\left[1 - \left(\frac{z(x_i)}{c}\right)^2\right]^2 \quad , z \le c \tag{3.19}$$

$$w(x_i) = 0 \quad , z > c \tag{3.20}$$

$$\bar{W} = \left(\sum_{i=1,N} w(x_i)\cdot x_i\right) \Big/ \left(\sum_{i=1,N} w(x_i)\right) \tag{3.21}$$

c. Calculate the new σ^2 with the equation:

$$\sigma^2 = \sum_{i=1,N} [w(x_i)\cdot x_i - \bar{W}]^2 \tag{3.22}$$

d. Repeat steps b and c until a pre-determined termination rule is satisfied.

3.2.7.2 Weight Median Estimator

This algorithm calculates the median, *M*, and estimates the MAD (Median Absolute Deviation) by the equation

$$MAD = 1.483\, M[|x_i - M(x_i)|],\ i = 1,N \tag{3.23}$$

This algorithm is similar to the Weight Mean Estimator, with the mean is replaced by the median, and σ replaced by the MAD. It is considered more robust than the weight mean because the calculation of σ is less sensitive to large deviations in the distribution shape.

3.3 Algorithms Requiring Some Additional Storage

These algorithms calculate the bias levels to some precision using one of the algorithms of the preceding section, and then update these values using empirical rules which serve to speed up the calculation. Their implementation requires that the CCD image be subdivided into strips similar to the previously described algorithms, but they require a smaller number of exposures for each strip because the full set of pixel values is only used for the initial estimate of bias level. The subsequent levels are calculated by successive approximations using all pixels from the CCD image at one time.

The storage needed for the initial phase of these algorithms is the same as that for the previous group. The duration of the entire calibration is given by:

$$T_{tot} = [r \cdot (N - Q) + Q]\, T_e \tag{3.24}$$

where:

- N = total number of calibration exposures
- Q = number of exposures for the second phase of the algorithm
- r = number of strips
- T_e= duration of a single exposure
- T_{tot}= total calibration time

For example, T_{tot} is proportional to $(N\text{-}Q)^2 + Q$ for an N exposure calibration obtained by making N-Q exposures for the initial bias threshold estimation and Q exposures for the final calculation.

3.3.8 Truncated Mean - Dynamic Updating

This algorithm is divided into two parts. In the first, the mean and σ are computed using the truncated mean equations (3.8-3.10), but the calculation uses a smaller number of exposures (Q) than necessary to compute reliable bias values. Outliers are rejected based on this σ value in the manner described in Section 3.2.4. At the end of this phase of the calculation, the bias levels are stored in the bias buffer.

In the second phase, the hardware pixel thresholder is turned on before starting a second set of exposures and those pixels above threshold events are ignored. The dynamic mean is calculated for the remaining pixels by a weighted average obtained from the new pixel value, as follows:

$$b_i = \frac{(n-1)}{n} b_{i-1} + \frac{1}{n} p_{i-1} \tag{3.25}$$

where:

- b_i = mean estimation after processing frame i
- p_i = new raw pixel value in frame i
- n = the number of samples used for the bias calculation, a number that will be incremented

each time a new sample is added. At the beginning of the dynamic calculation, it is set equal to N, the number of samples used for the first phase of the calculation.

These algorithms are expected to work best for Gaussian or symmetric heavy-tailed distributions with outlying points for which the assumption of normality does not hold. The outlier points are rejected and the remaining data are treated as Gaussian. The algorithm fails if the mean of the tails is large[2].

3.4 Scenarios

All algorithms except those that use no additional storage will probably transfer the input pixel values to the data cache to speed up subsequent calculations, which will most probably operate on a few rows at a time. The "hybrid" algorithms discussed in Section 3.3 will also use the data cache or bulk memory to store partial sums. Each calibration algorithm will require the combined assistance of the FEP thresholders, the FEP processors, and the BEP.

The *pixel thresholders* will be instructed to copy the pixel data from a particular range of CCD rows into a set of strips in the image pixel buffer. To decrease the calculation time, the opportunity should be taken to copy these values from the image pixel buffer to bulk memory during pixel data collection, provided the read and write operations don't interfere with one another.

The *front-end processors* must control the registers in the pixel thresholders that control the pixel strip collection and storage, then execute the specific algorithm, and regulate the data transfer from image pixel buffer to data cache and from there to bulk memory (possibly via DMA transfer).

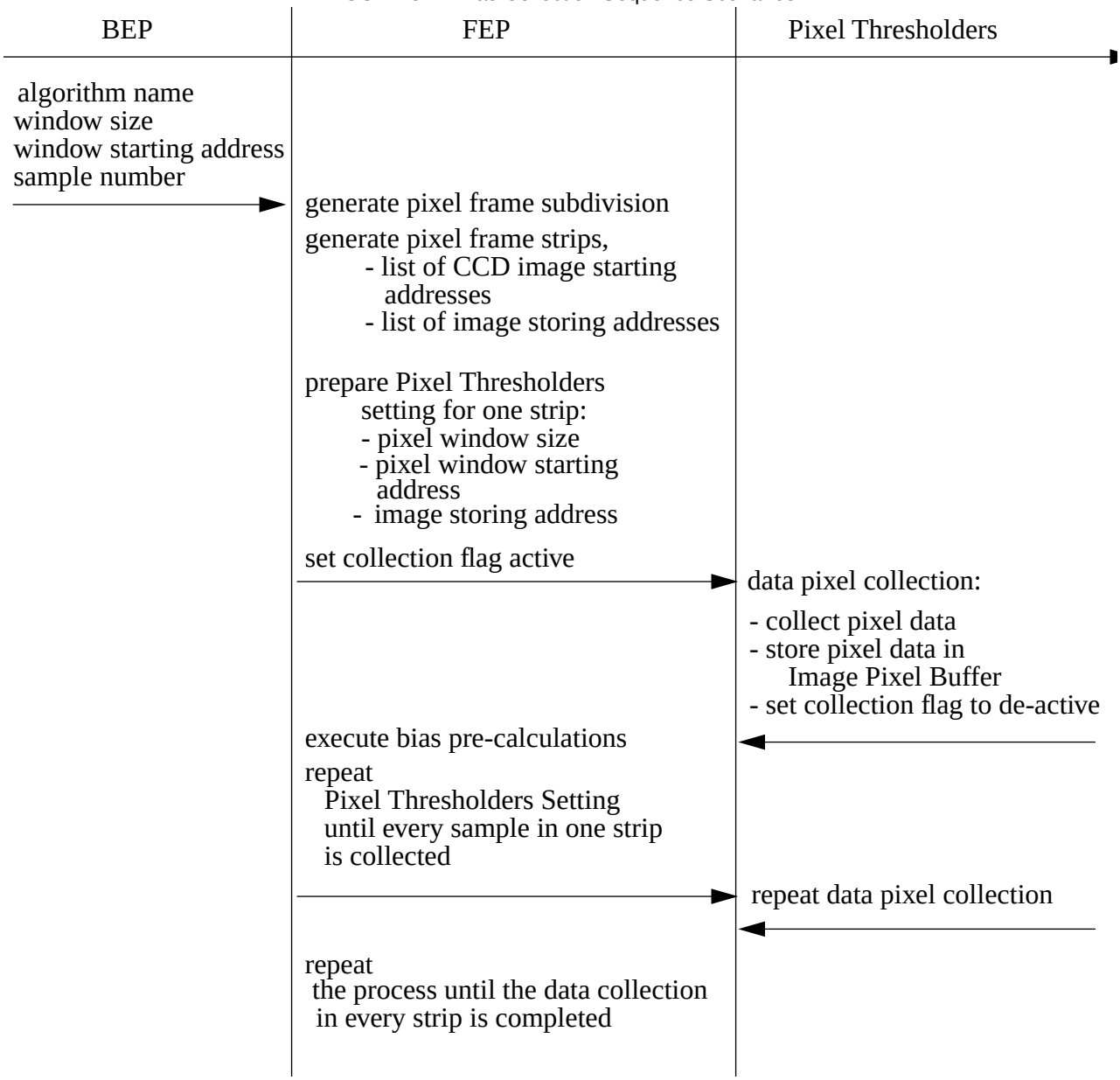
The *back-end processor* must start the bias calculation and initializes its parameters. It may also be required to collaborate with the FEPs in compressing and downlinking the bias maps while they are being created.

Figure 3-1 shows a scenario in which the three hardware units act together to perform a bias calibration that requires that all pixel values be examined together. The steps are as follows:

- Expose the CCD N times, copying the same strip of CCD rows into N separate blocks of Image Buffer.
- For each pixel in the strip, copy its N values to the data cache.
- Calculate the bias and rejection levels for this pixel.
- Apply the rejection rule and iterate the if necessary.
- Calculate the parity of the bias value.
- Transfer the bias and parity levels from the data cache to the bias level map buffer.
- Repeat the process with another strip of CCD pixels until the bias map is complete.

2. Press, William H., etc. "Numerical Recipes in C" 14.1 "Median and Mode"

FIGURE 3-1. Bias Collection Sequence Scenarios



The following storage buffers are used for this process:

- The image pixel map buffer is divided into strips containing N samples of the same strip of CCD image pixels.
- The data cache stores the N sample data (2K x N), sample number indicating the pixel number not rejected (1K), bias level calculated (2K), rejection levels (1K). The algorithm might store intermediate results to speed up this calculation.
- The bias level map buffer stores the results of the bias level calculation. These values are

transferred to the buffer as soon as they are available. At any later time, the BEP may compress them and copy them to downlink telemetry.

A summary of the storage and exposure times for the major algorithms is reported in Table 3-1, where s is the strip size, n is the number of strips, N is the total number of exposures, T_f is the frame readout time,and Q (< N) is the number of “conditioning” exposures used by some algorithms. Since the strips must be stored simultaneously in the Image Buffer, $n \propto N$.

TABLE 3-1.Algorithm Summary

Class of Algorithm	Storage Size	Exposure Time (seconds)
Algorithms requiring additional storage	s N	$n N T_f$
Algorithms not requiring additional storage	one image	$N T_f$
Algorithms requiring some additional storage	s Q	$(n Q + N - Q) T_f$

4.0 Test Data Selection

In order to test the algorithms described in the previous section, a series of data sets must be used. After discussions with the ACIS science team, the following characteristics were considered necessary:

- 100 or more consecutive full frames from the same CCD.
- Data from both front-side and back-side illuminated CCDs.
- The CCDs should be exposed to 4 sources:
  - Ambient (no source)
  - A soft X-ray spectrum, e.g. C_K, Al_K, etc.
  - A hard X-ray spectrum, e.g. Fe^{55}.
  - A hard X-ray spectrum, e.g. Fe^{55}, accompanied by a gamma source, e.g. Co^{60}.
- All data should be paralleled by sets from the same CCDs after subjecting them to the average radiation dose to be anticipated for ACIS after 5 years of on-orbit operation.

In addition, it was stipulated that all images should be in 16-bit FITS format, with an overclock region in each row of each output node, and each data set should be fully documented.

The CCDs contained 1026 rows of 1024 pixels each, and were clocked out in 4 nodes, digitized into 12 bits and recorded in 16 bits, although the preamplifier levels were frequently set sufficiently high that pixel values above 4095 were not uncommon. Apart from these bias level variations, the system gain was kept as constant as possible—typically within ±1% of nominal. The contents of each image line are shown in Table 4-1. Note the presence of underclocks, overclocks, and 3 throw-away pixels at the end of each output node.

TABLE 4-1. Pixel Arrangement in Test Files

Bytes	Contents	Bytes	Contents
0–3	Node A underclocks	680–683	Node C underclocks
4–259	Node A image pixels	684–939	Node C pixels
260–336	Node A overclocks	940–1016	Node C overclocks
340–343	Node B underclocks	1020–1023	Node D underclocks
344–599	Node B image pixels (reverse order)	1024–1279	Node D image pixels (reverse order)
600–676	Node B overclocks	1280–1356	Node D overclocks

The measurements were made on two CCDs, one front-side illuminated (FI), the other back-side (BI), in late 1994, and the resulting FITS images were compressed and copied to a set of CD-R disks, each including documentation. The details are shown in Table 4-2.

TABLE 4-2. Data Sets used for Bias Algorithm Evaluation

CCD ID	Status	Illumination	Frames	Date	Volume	Pathname
ccid17-12-4 (FI)	pre-irradiation	none	200	12/06/94	ac_0002c	ccid17-12-4/bias
		Al_K	200	12/07/94	ac_0001c	ccid17-12-4/aluminum
		O_K	200	12/07/94	ac_0001c	ccid17-12-4/oxygen
		Fe^{55}+Co^{60}	200	12/06/94	ac_0002c	ccid17-12-4/fe55co60
		Fe^{55}	100	12/05/94	ac_0005c	ccid17-12-4/fe55
	post-irradiation	none	150	12/20/94	ac_0008c	ccid17-12-4/bias
		none	50	12/20/94	ac_0009c	ccid17-12-4/bias
		Al_K	200	01/10/95	ac_0008c	ccid17-12-4/aluminum
		O_K	200	01/10/95	ac_0008c	ccid17-12-4/oxygen
		Fe^{55}	199	12/20/94	ac_0009c	ccid17-12-4/fe55
		Fe^{55}+Co^{60}	200	12/20/94	ac_0009c	ccid17-12-4/fe55co60
ccid17-38-3 (BI)	pre-irradiation	none	200	12/09/94	ac_0006c	ccid17-38-3/bias
		Al_K	200	10/20/94	ac_0007c	ccid17-38-3/aluminum
		O_K	200	12/01/94	ac_0003c	ccid17-38-3/oxygen
		C_K	200	12/01/94	ac_0006c	ccid17-38-3/carbon
		Fe^{55}+Co^{60}	200	12/09/94	ac_0003c	ccid17-38-3/fe55co60
		Fe^{55} (-100C)	12	12/09/94	ac_0005c	ccid17-38-3/fe55c100
		Fe^{55} (-110C)	100	12/09/94	ac_0005c	ccid17-38-3/fe55c110
		Fe^{55} (-115C)	100	12/09/94	ac_0005c	ccid17-38-3/fe55c115
		Fe^{55} (-120C)	100	12/09/94	ac_0005c	ccid17-38-3/fe55c120
	post-irradiation	none	100	01/12/94	ac_0004c	ccid17-38-3/bias
		Al_K	200	01/11/94	ac_0004c	ccid17-38-3/aluminum
		C_K	200	01/11/94	ac_0004c	ccid17-38-3/carbon
		none	100	01/12/94	ac_0010c	ccid17-38-3/bias
		O_K	200	01/11/94	ac_0007c	ccid17-38-3/oxygen
		Fe^{55}	200	01/12/95	ac_0010c	ccid17-38-3/fe55
		Fe^{55}+Co^{60}	200	01/12/95	ac_0010c	ccid17-38-3/fe55co60

5.0 Bias Map Compression

Since bias maps will contain many pixels with values lying in a narrow range, it will be possible to apply a simple non-destructive compression algorithm to reduce the number of bits that must be downlinked. Unless the variance of the bias values changes significantly during the mission, the recommended compression scheme is the Huffman First Difference. It is applied as follows:

- A set of 8192 Huffman strings is prepared. These are varying-length bit strings with the special property that their length can be determined from their parity properties, i.e. any two Huffman strings can be distinguished even when they have been concatenated.
- The shortest Huffman string is associated with data value 0, the next shortest with data value +1, the next with –1, the next with +2, and so forth.
- Each row of Bias Map data is separately compressed and communicated to the telemetry system as a single data record. It begins with the row index 0–1023, followed by the 12–bit value of the first Bias Map value in that row. Subsequent 12-bit values are first subtracted from the preceding 12-bit value, and the associated Huffman string $H(p_{n+1}-p_n)$ is written to the telemetry record.
- Two special pixel values—denoting bad pixels and those whose bias values have suffered a parity error and are therefore unreliable—are given special treatment: they are exempted from differencing and compressed to special values. After transmitting such a value, e.g. $H(P_n)$, the next field will be $H(p_{n+1}-p_{n-1})$, provided that p_{n+1} is not itself either of the two special values.

When the telemetry records are received on the ground, the row number and first 12-bit value are extracted, followed by the individual Huffman strings. These latter are used as indices into an inverse of the original Huffman table to recover the Bias Table differences, and thence to the complete Bias Table.

6.0 Data Analysis

6.1 Overview

The raw data that are used for bias calibration must be collected by the CCDs in the presence of several types of radiation and by various kinds of pixels, termed “good”, “hot”, “flickering”, and just plain “bad”. The calibration algorithms must distinguish between the following event signatures:

- X-rays—isolated pixels, or small groups of pixels, with high ADUs.
- Ionizing Particles—typically larger groups of pixels with high ADUs.

The pixel values are also expected to vary for other reasons:

- Random variations from pixel to pixel that may be compensated by bias calibration.
- Random time-dependence of the value of a single pixel.
- Small systematic variations due to temperature, DEA gain, etc., which are partially compensated by averaging the overclocks from each DEA output node.
- Defective or “hot” pixels with anomalously high data values that are generally independent of exposure time and are present in several consecutive exposures.
- Pathological Pixels—those that possess an abnormal distribution of values in the absence of events, e.g., rectangular or double peaked distributions.

The purpose of analyzing the test data of Section 4.0 is to quantify the variation of real pixel values in the presence of events, and to see how they impact the bias level calculation. To observe the greatest possible variety of behaviors, it has been necessary to use full CCD frames and to examine entire images. On the other hand, the algorithms themselves may be affected both by the number of events occurring in a given pixel, and by their energy, and if we wish to compare one algorithm against another we must use a limited number of observations with precisely known characteristics. The statistical techniques used for analyzing these data sets, which are formally presented in Appendix 8.0, must therefore be able to compare bias level values from large images with very many pixels, as well as from small data sets. They must identify systematic bias level variations from image to image (*i.e.*, bad pixels) in particular pixel locations, and estimate the statistical significance of variations in the bias level values.

6.2 Bias Calculation

The data used for the bias calculations are first corrected for changes in the average output node overclock values. The following corrections were applied:

$$p'_{ij} = p_{ij} - \overline{o_{i,n}} + \kappa \qquad (6.1)$$

where

- p'_{ij} indicates the pixel data at CCD row *i* and column *j*, corrected for overclock variations.
- p_{ij} indicates the raw pixel data at CCD row *i* and column *j*.
- $\overline{o_{i,n}}$ represents the mean of the overclocks in the *i*th row. *n* is the number of overclocks used in the calculation.
- the constant κ was added to the right hand side of equation 6.1 to avoid negative numbers. For this analysis, a value of 200 was used.

This overclock correction differs from that to be used in the flight system, and was chosen to counteract the higher noise level in the calibration data, which used prototype data collection hardware. Since each row of pixels from each output node was accompanied by 77 overclock values (see Table 4-1), the average could be calculated for each row and node without significant sampling error.

The bias levels were calculated from sets of ten consecutive image frames. The algorithms themselves were variants of the *mean* and *median* algorithms described in Section 3.2. Three different versions of the mean algorithm were used: the simple mean (MEAN–I1L1), the mean with rejection of pixel values greater than $\pm\sigma$, as described in Section 3.2.4 (MEAN–I2L1), and the mean with rejection of pixel values greater than $\pm 2\sigma$ (MEAN–I2L2). To these was added the 50% median algorithm described in Section 3.2.5 (MEDIAN).

6.3 Objectives

There were three:

- Visualize and quantify the range of bias level variations.
- Identify the existence of areas in the images with anomalous bias level values.
- Determine the dependence of the bias level values on particular factors, using the data content itself.

The visualization of bias level variations necessitates the examination of entire images and was used to quantify the range of these variations. The identification of areas with high bias levels also required the visualization of entire images and was used to locate bad and/or “hot” pixels. To investigate the dependency of bias level values from the data content itself required a knowledge of that content, was therefore limited to small data sets, and was used to compare the performance of the various algorithms.

6.4 Bias Level Variations

The objective of this phase of the analysis was to quantify the range of variations of the bias level calculation across an entire CCD frame. This was accomplished by applying a single algorithm (MEAN–I2L1) to three data sets, and then calculating the maximum, minimum, mean, and σ of each bias pixel’s level, and by counting the number of bias values greater than a specified threshold.

The data sets used were the FITS frames (from 0010 to 0019) collected from the pre-irradiated backside-illuminated CCD 17–18–3, (1) in absence of X-ray sources, and then illuminated by X-ray sources emitting (2) Al_K X-rays and (3) Fe^{55} X-rays with Co^{60} γ-ray contamination. These data sets were named, respectively, `ccid17-38-3/bias`, `ccid17-38-3/aluminum`, and `ccid17-38-3/fe55co60`, and the results of this survey for each output node is reported in Table 6-1.

TABLE 6-1. Summary Of Bias Level Variations

X-Ray Source	Fit Frames	Node	Bias Level				Number of Bias Level Values above Threshold (Threshold)			
			Min[1]	Max[2]	Mean	s	205	210	220	240
none	0010 – 0019	1	186	1092	201.34	4.0	31849	1135	5	3
		2	188	4095	201.34	11.24	23688	362	15	8
		3	188	3666	201.34	8.11	23842	369	11	7
		4	187	4088	201.35	9.0	23952	928	281	14
Al_K	0010 – 0019	1	188	536	201.15	3.19	21547	442	12	1
		2	0	4095	201.22	8.03	11958	89	20	5
		3	189	4095	201.22	8.72	18286	236	17	4
		4	0	671	201.06	3.37	21677	661	19	1
Fe^{55}+Co^{60}	0010 – 0019	1	184	1161	201.35	4.23	31546	1251	73	42
		2	188	4095	201.38	11.27	23886	544	75	38
		3	187	3640	201.37	8.09	24253	462	68	40
		4	187	4085	201.41	9.0	24418	984	312	36

1. a minimum value of 0 indicates a defective pixel.

2. values of 4095 were assigned to defective pixels.

Table 6-1 shows that the variations in the minimum bias values are negligible, whereas the maximum variations change considerably. The mean appears quite stable, while the standard deviation σ varies over a limited range. The numbers of bias levels greater than a given threshold indicate the impact of the high bias level values on the entire image. As shown, the number of these high values is negligible, and this is confirmed by the σ values.

6.5 The Identification of Anomalous Pixels

The analysis also identifies damaged pixel locations, tests whether they remain constant or change from exposure to exposure, and whether their values change over longer time scales. These are best achieved by calculating the bias levels for entire CCDs, and then analyzing the results with the ANOVA techniques. The resulting F factors, defined in Section A.1, can be interpreted as indices of pixel health—those whose F values exceed the values of F_α (tabulated in Applicable Document 8) at a particular level of significance α, corresponding to the number of degrees of freedom of the

experiment, are assumed to be either defective or contaminated by events that were not rejected during the bias calculation.

A more precise way of identifying bad pixels is to calculate the F values from two sets of bias levels—derived from two sets of raw data produced at different times from the same CCD. The F values are compared, and areas in the CCD are identified, where the Fs for both data sets are higher than the tabulated F_{α}. This technique gives more reliable information on areas of bad pixels because the probability is low that two different sets of measurement of the same pixel will be contaminated by events or particles that were not rejected by the bias calculation.

Unfortunately, no information on the performance of the bias algorithms themselves can be extracted from this analysis because, if the algorithm identifies events and removes their effects, the bias values correspond to good pixels and the success of the algorithm is not in question. When the bias algorithm does not identify the event, it yields an anomalous value that is confused with a bad pixel.

6.5.1 Data Sets

The data sets used for this analysis came from CCID 17-38-3, a back-side device that had not been irradiated. The frames were collected (1) in absence of X-ray sources, and in presence of X-ray sources corresponding to (2) Al_K X-rays and (3) Fe^{55} X-rays with Co^{60} γ-ray contamination. The names of these data sets were, respectively, `ccid17-38-3/bias`, `ccid17-38-3/aluminum`, and `ccid17-38-3/fe55co60`.

The bias levels were calculated for each pixel in the CCD, using several algorithms—MEDIAN, MEAN-I2L2 and MEAN-I2L1, as defined in Section 6.2. For statistical analysis, two sets of bias values were generated using two sets of image frames. Each set contained three groups of data (labelled “none”, “Al_K”, and “$Fe^{55}Co^{60}$”), corresponding to no X-Ray sources, Al_K sources, and Fe^{55} X-Rays + Co^{60} γ-rays. This was repeated 5 times for each data set, so a total of 150 image frames was used for each set—10 for each bias calculation, 5 replications of each, and 3 different X-Ray sources.

Each group of CCD frames exposed to different X-ray sources contained a number of bad and hot pixels, along with a number of X-ray and background events. In general, the events do not affect the analysis that follows because the number that are not rejected by any of the bias algorithms is much smaller than the total number of pixels.

6.5.2 Statistical Analysis

The one-way and two-way ANOVA methods used were used to perform the factor analysis.

- One-way classification—the “treatments” are the X-Ray sources (three in this case) and the 5 “replications” are the repeated bias calculations executed for the same pixel from different groups of calibration frames. The total number of bias levels used in this analysis is 15 and

their division into repetitions and treatments is shown in Table 6-2.

TABLE 6-2. One-Way Anova Classification To Identify Bad and Hot Pixels

Treatment	Repetitions				
	1	2	3	4	5
None	Bias level	Bias level	Bias level	Bias level	Bias level
Al_K	Bias level	Bias level	Bias level	Bias level	Bias level
Fe^{55}+Co^{60}	Bias level	Bias level	Bias level	Bias level	Bias level

The same analysis is repeated for each algorithm. The results give information about the general health of each pixel.

- Two-way classification—the “rows” are the X ray sources (3 in this case), the “columns” are the algorithms (also 3), and the (5) replications are the repeated bias calculations executed for the same pixel from different groups of calibration frames. The total number of bias levels used in this analysis is 45 and their division into repetitions and treatments is shown in Table 6-3.

TABLE 6-3. Two-Way Anova Classification to Identify Bad and Hot Pixels

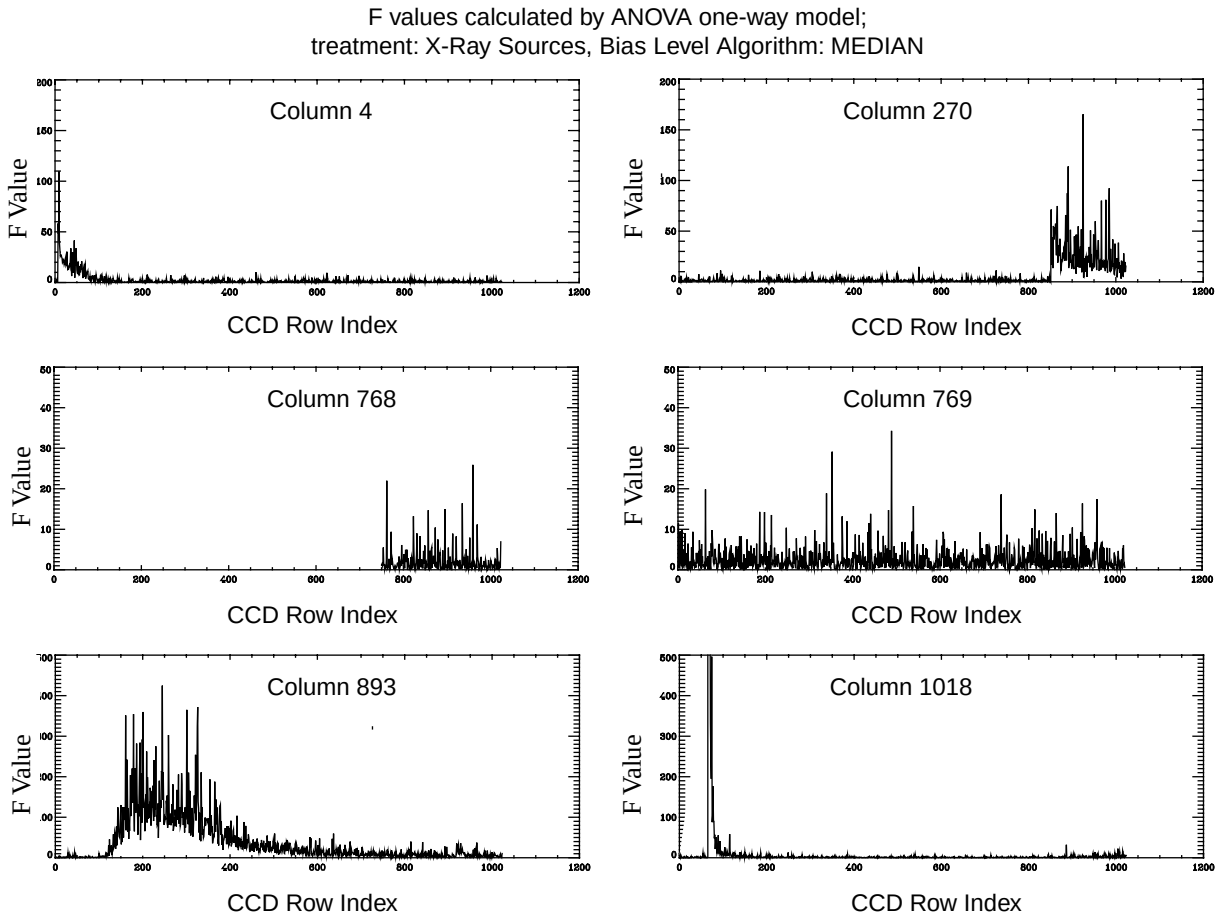
Treatment	Algorithm		
	MEAN I2L2	MEAN I2L1	MEDIAN
None	5 Bias levels	5 Bias levels	5 Bias levels
Al_K	5 Bias levels	5 Bias levels	5 Bias levels
Fe^{55}+Co^{60}	5 Bias levels	5 Bias levels	5 Bias levels

The results tell us about the simultaneous dependency of the bias level on the choice of X-Ray source and bias algorithm.

6.5.2.1 F Distribution within the CCD

Figure 6-1 illustrates the dependence of F on CCD row number for several anomalous columns of CCID 17-38-3. The Fs show considerable variation from column to column—in some there is a relatively sharp peak in the middle of noise, in others the peak is approximately rectangular, while others contain only isolated pixels scattered at random row locations.

FIGURE 6-1. Distribution of Fs in Selected CCD Columns



6.5.2.2 Summary of ANOVA Results

The results from both ANOVA techniques are summarized by identifying the maximum F value, their mean and standard deviation, and the number of pixels with F greater than the tabulated F_α—two values of α were used, one corresponding to a 5% significance level and the other to 1%. The calculation was performed for each CCD output node, for both of the duplicate data sets, and for both ANOVA techniques. The results for the one-way model are summarized in Table 6-4, those for the two-way model in Table 6-5.

The one-way model results reported in Table 6-4 indicate that:

- The maximum F values vary considerably from node to node. However, the mere presence of high values is not a reliable measure of CCD health since they could result from no more than a handful of pixels.

TABLE 6-4. Results of the One-Way ANOVA Model (3 X-Ray Sources, 5 Replications)

Algorithm	Data Set	Node	F Values				
						Number of Pixels with F > F_α	
			Max	Mean	Sigma	5% F_α =4.46	1% F_α = 8.65
MEAN I2L2	1	1	59.38	1.27	1.59	10829	1537
		2	4457.23	1.41	9.17	12923	2198
		3	279437	2.42	546.54	10712	1521
		4	11923533	47.05	23288.0	12048	2386
MEAN I2L1	1	1	110.06	1.27	1.60	10641	1583
		2	4111.55	1.39	8.76	12505	2027
		3	274945.34	2.39	537.58	10603	1534
		4	9484628	37.67	18524.61	11707	2273
MEDIAN	1	1	109.65	1.27	1.62	10810	1630
		2	4963.41	1.39	10.20	12551	2018
		3	309202.94	2.52	604.38	10569	1473
		4	8429955.0	33.72	16464.76	11699	2303
MEAN I2L2	2	1	3815.48	1.34	8.58	11525	1733
		2	13667.26	1.49	27.10	14027	2245
		3	190971.1	2.34	394.02	11055	1555
		4	15454921.0	60.53	30185.29	12768	2503
MEAN I2L1	2	1	3948.88	1.33	8.85	11323	1653
		2	9158.15	1.44	18.37	13359	2168
		3	225103.55	2.33	445.02	10934	1534
		4	15454921.0	60.51	30185.30	12351	2378
MEDIAN	2	1	3927.22	1.33	8.78	11427	1650
		2	12929.29	1.45	25.58	13414	2133
		3	272606.25	2.61	543.96	10793	1607
		4	11590624.0	45.75	22637.87	12282	2365

- The mean F is not statistically significant for three of the nodes of this CCD, but it is significant for the fourth. The mean and its standard deviation give some indication of the CCD health, but this is not a robust index because a small number of very high values can greatly affect both the mean and the standard deviation.
- The total number of pixels with F values greater than the tabulated F_α is almost independent of bias algorithms, and is approximately the same for bias levels calculated from different data sets. It represents the number of pixels that have a bias level variation greater than the

residual error at the α level of significance. The pixels identified by this criterion might be either “bad” or merely “hot”, and further analysis is needed to see whether the high bias level values were not merely the result of unidentified events. The number of pixels with F greater than F_α is about 1% of the total number of pixels in the frame for a 5% level of significance and about 0.2% for 1% level of significance.

TABLE 6-5. Results of the Two-Way ANOVA Model (3 X-Ray Sources, 3 Algorithms, 5 Replications)

Data Set	Node	Factor	F Values				
						Number of Pixels with F > F_α [1]	
			Max	Mean	Sigma	5% F_α=3.28	1% F_α=5.30
1	1	X-Ray Sources	259.35	3.39	4.12	91742	51612
		Algorithms	4.5	.2	.22	10	0
	2	X-Ray Source	13452.44	3.74	27.7	97962	56609
		Algorithms	5.53	.23	.253	25	2
	3	X-Ray Source	861241	6.90	1683.85	91193	50953
		Algorithms	5.89	.22	.24	15	1
	4	X-Ray Source	29227940	115.69	57085.69	92318	52353
		Algorithms	4.90	.22	.242	23	0
2	1	X-Ray Source	11689.43	3.56	26.05	93921	53494
		Algorithms	6.30	.20	.23	21	2
	2	X-Ray Source	34656.30	3.91	68.84	100750	59123
		Algorithms	5.22	.24	.26	26	0
	3	X-Ray Source	674308.94	6.73	1346.80	92237	52050
		Algorithms	5.392	.22	.24	14	1
	4	X-Ray Source	41727608.0	163.27	81498.96	94738	54487
		Algorithms	5.01	.22	.24	21	0

1. The tabulated values of F_α correspond to 2 and 36 degrees of freedom, respectively.

The two-way model analyzes simultaneously the effect of X-Ray sources and bias algorithms. The results for the entire area of a CCD image frame are summarized in Table 6-5. These indicate the general behavior of the CCD but they are not sensitive to some anomalous conditions. To summarize:

- The choice of X-ray source is statistically significant for the bias level calculation, but the choice of algorithm is not. The ANOVA technique separates the effects of sources from those of algorithms and allows the differentiation between the two factors.
- The algorithm factor is not significant in the bias level calculation. When the results with (and without) X-Ray sources are compared, the algorithm factor increases the data variabil-

ity and merely confuses the ANOVA analysis. This does not imply that the bias level is insensitive to the choice of algorithm, merely that it is not possible, by examining an entire CCD, to identify those pixels and events in which the bias calculation depends on the choice of algorithm.

- For output nodes A and B (columns 1-512), the X-Ray source factor is not significant at a 5% level, but it is significant for nodes C and D (columns 513-1024). The results from this model agree with those of Table 6-4 (one-way ANOVA) for nodes A, B, and D, but differ for node C (columns 513-768), where the source factor is significant in two-way ANOVA but not in one-way. This contradictory behavior might be explained by the large variability added by the algorithm factor.
- The behavior of the maximum, mean, and variance of F values follows the same trends as seen in Table 6-4 for the one-way ANOVA model.
- The number of pixels with Fs greater than the tabulated F_α is greater than in the one-way analysis because the number of bias levels used in the two-way analysis is triple that in the one-way case and the algorithm factor therefore increases the variability. About 9% of the pixels have an F value greater than F_α for a 5% level of significance and about 5% of them exceed it at the 1% level.

6.5.2.3 Comparison of the ANOVA Results from the Duplicate Data Sets

A more precise identification of bad and hot pixels can be obtained by comparing the Fs calculated from the two duplicate data sets. The Fs resulting from the ANOVA analysis were represented as pixel arrays, compared, and those pixels with Fs higher than the tabulated F_α were identified. The images contain zero for all pixels with F values less than the tabulated F_α and ten for the others. They readily identify the CCD areas containing anomalous pixels—they appear as long or short sections of columns or unique bad pixels surrounded by good pixels. The bad columns are in every output node, the isolated pixels are distributed evenly, although a relatively large concentration is noticeable at the top of the image.

Figure 6-2 shows those F values higher than the tabulated F_α for both data sets in two areas of the CCD. It used the ANOVA one-way model applied to bias levels calculated by the MEAN-I2L1 algorithm. The upper image (rows 1-256) shows an area of high density of bad or hot pixels centered in the first rows of the CCD, and some segments of bad columns numbered 4, 893, and 1018. The lower image (rows 769-1024) shows segments of bad columns numbered 270, 893, and 1018.

FIGURE 6-2. Bad and Hot Pixel Image

Pixels with F > F_α in each of two sets of data of the same CCD, both calculated by the one-way ANOVA model (treatment: X-ray sources; algorithm: MEAN-I2L1). The anomalous pixels cluster in columns and in rows at the top of the CCD. NOTE: the levels of grey in these images have been chosen to emphasize bad pixel morphology.

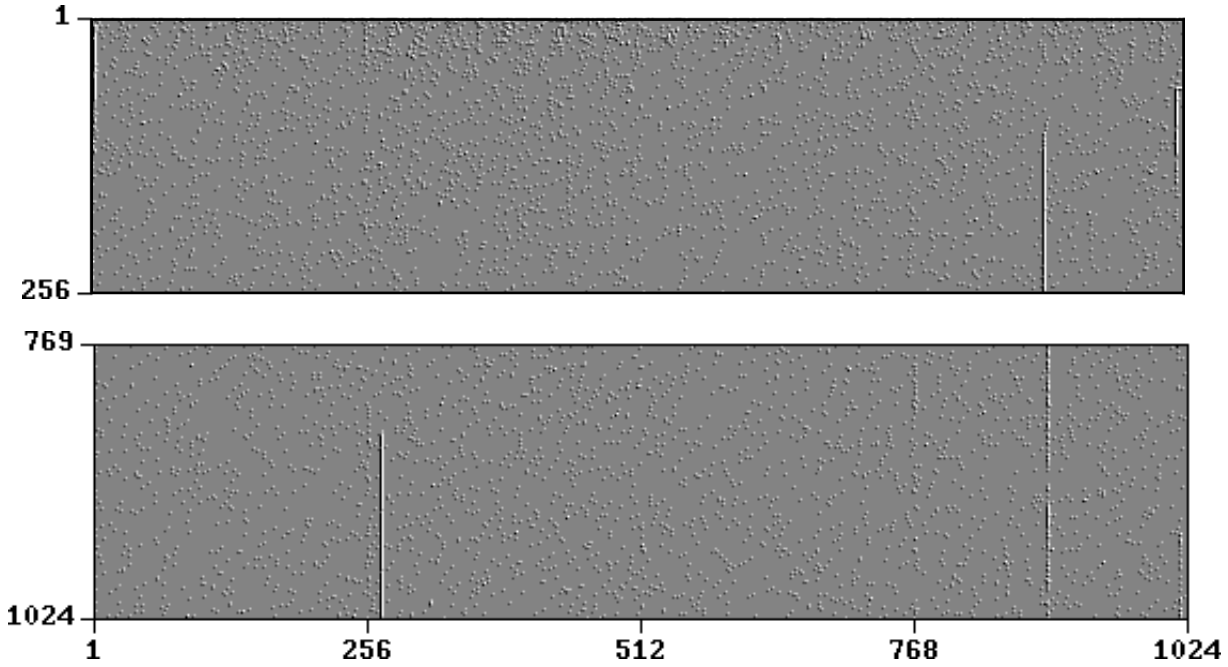


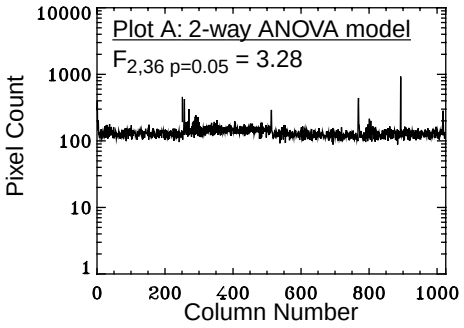
Figure 6-3 shows a series of histograms prepared from the same data, identifying the number of pixels per columns with Fs higher than the tabulated F_α in both data sets. They also show the number of bad and hot pixels in each column. In general, the columns containing an high number of bad pixels are independent of the choice of bias algorithm, and the ANOVA techniques identifies them with few exceptions.

Comparing the results from the ANOVA one-way model as applied to the bias levels calculated by the three algorithms shows that the CCD columns with bad pixels are common to every algorithm except in the case of a single column, which is only seen to be anomalous with the bias levels calculated by the MEAN–I2L2 and MEDIAN algorithms. This might be because MEAN–I2L2 has a larger acceptance range than MEAN–I2L1 and might reject less data. The increased number of bad columns reported by the MEDIAN algorithm relative to MEAN–I2L1 might be attributed to the fact that MEDIAN is data dependent and might generate higher values if the data set contains proportionately higher data values.

The results from the ANOVA two-way model show a higher number of bad pixels than the ANOVA one-way model. This result can be attributed to the greater number of bias level values used in the analysis and to the added variations due to the algorithm factor.

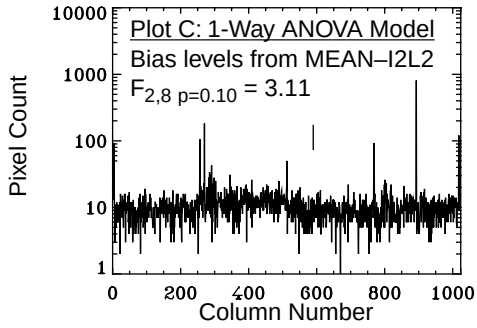
FIGURE 6-3. Distribution of Bad and Hot Pixels as a Function of CCD Column Number

Comparing pixels with Fs > F_α in each of two separate sets of data
from CCID17-38-3, a back-illuminated CCD, before irradiation.



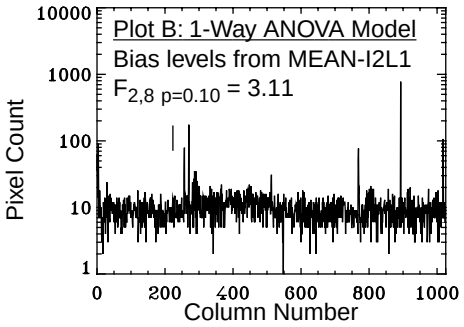
$F_{2,36\ p=0.05}$ values > 250

Columns	0	1	257	270	513	768	769	893	1018
Count	292	395	426	301	292	300	442	927	284



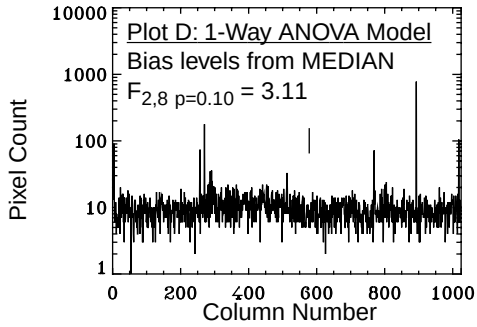
$F_{2,8\ p=0.10}$ values > 50

Columns	1	4	257	270	768	769	893	1018
Count	81	92	105	183	56	95	511	121



$F_{2,8\ p=0.10}$ values > 50

Columns	1	4	257	270	769	893	1018
Count	64	91	79	175	77	777	107



$F_{2,8\ p=0.01}$ values > 50

Columns	1	4	257	270	768	769	893	1018
Count	60	91	74	175	52	72	776	105

6.5.2.4 Conclusions

The results from this analysis may be summarized as follows:

- It is better to examine the bias algorithms one at a time because their simultaneous measurement complicates the factor analysis. In the remainder of this report, bias algorithms will be considered separately.
- The number of bad and hot pixels is consistent from frame to frame.
- The comparison between F images of bias levels generated from duplicate sets of data is a good way of locating bad and hot CCD pixels.
- The identification of anomalous pixels might be obscured by X-Ray events.

7.0 Algorithm Comparison

The data used for the bias calculation have been collected in the presence of external X-rays, high energy particles, and calibration sources, because the events cannot be eliminated during the bias calibration periods. These events vary in energy from about 100 eV to 10 KeV, but are small in number. Their energy can be absorbed by a single pixel or split between two or more neighboring pixels in varying proportions. The present study was limited to energies between 72 and 316eV, because its purpose was to evaluate the minimum energy of X-Ray that could be detected above the noise level.

The bias calculation uses consecutive observations of the same pixel. The number of events that can occur within this pixel during the calibration period depends on the number of observations and duration of each. For this study, the number of events is considered to vary between 0 and 3 events. The goals are as follows:

- To calculate the bias level with different algorithms in the presence and absence of events, and to estimate the average bias values for a fixed number of replications.
- To compare the differences between average bias levels calculated by different algorithms applied to the same data sets.

To rank the algorithms according to their responses to different conditions, two data sets were prepared—*eventless data* are experimental data sets from real CCD frames that were selected so that they not contain any events and correspond either to the same pixel or to different pixels that have similar bias values. *Simulated data* are eventless data sets where “events” of a known energy have been added. Multiple events are simulated by adding the same value to more than one pixel belonging to the set used to calculate a particular bias value.

TABLE 7-1. Bias Level Calculation for Identifying Eventless Data

Pixels	X-Ray Source		
	none	Al_K	Fe^{55}+Co^{60}
pixel 0	10 replications	10 replications	10 replications
pixel 1	10 replications	10 replications	10 replications
...	10 replications	10 replications	10 replications
pixel 19	10 replications	10 replications	10 replications

7.1 Eventless Data

The choice of data to be used was made in two steps: (a) identification of a group of possible pixels from a large number of CCD images, and (b) extraction of a smaller sample group from it. The initial pixel identification was made from the comparison images discussed in Section 6.5 by identifying those pixels with F less than some predetermined value, in this case 1.0. This value is

more restrictive than the expected value, but it was chosen to reduce the number of pixels to examine—twenty were chosen

The values of each selected pixel were then extracted from 100 raw frames for each X-Ray source (three hundred raw frames for each pixel), and the bias levels were calculated using the simple mean algorithm with no rejection (MEAN–I1L1). This is more sensitive to data variations than the other algorithms and can better identify the presence of events in raw data.

The pixel configuration used for choosing the “eventless” group from the pixel list is shown in Table 7-1. The minimum and maximum bias level values were identified and their means and variances were calculated for each group of ten bias level replications. These values are tabulated in Table 7-2.

TABLE 7-2. The Characteristics of “Eventless” Pixel

Pixel Id	X-Ray Source	Bias Levels	Min Bias Level	Max Bias Level	Bias Level Range	Bias Level Mean	Bias Level Variance
pixel 0 s=34804 x y 1012 33	None	201 198 201 198 199 202 203 197 197 198	197	203	6	199.4	4.24
	Al_K	202 197 202 199 199 200 200 202 198 201	197	202	5	200.0	2.8
	Fe^{55}+Co^{60}	196 198 202 204 203 200 199 201 196 199	196	204	8	199.8	6.76
pixel 1 s=70070 x y 438 68	None	202 202 199 202 197 201 200 197 204 199	197	204	7	200.3	4.81
	Al_K	201 204 196 202 202 202 200 198 200 200	196	204	8	200.5	4.65
	Fe^{55}+Co^{60}	199 199 198 202 206 203 197 200 201 199	197	206	9	200.4	6.44
pixel 2 s= 72022 x y 342 70	None	199 195 199 202 201 197 198 201 200 198	195	202	7	199.0	4.0
	Al_K	202 197 200 198 199 197 199 201 205 196	196	205	9	199.4	6.64
	Fe^{55}+Co^{60}	199 201 199 197 200 200 199 200 200 198	197	201	4	199.3	1.21
pixel 3 s=72080 x y 400 70	None	196 199 204 197 195 197 201 201 200 195	195	204	9	198.5	8.05
	Al_K	198 197 198 202 200 203 196 197 199 197	196	203	7	198.7	4.81
	Fe^{55}+Co^{60}	201 199 198 199 195 198 196 200 200 196	195	201	6	198.2	3.56

TABLE 7-2. The Characteristics of "Eventless" Pixel

Pixel Id	X-Ray Source	Bias Levels	Min Bias Level	Max Bias Level	Bias Level Range	Bias Level Mean	Bias Level Variance
pixel 4 s=72383 x y 703 70	None	203 201 202 202 203 200 200 200 208 200	200	208	8	201.9	5.49
	Al_K	204 200 203 199 202 200 202 200 203 204	199	204	5	201.7	3.01
	Fe^{55}+Co^{60}	196 207 202 203 203 201 201 204 200 203	196	207	11	202.2	7.4
pixel 5 s=72504 x y 824 70	None	199 199 197 197 201 200 200 205 199 199	197	205	8	199.60	4.64
	Al_K	197 200 203 200 198 256 201 206 198 205	197	256	59	206.4	281.44
	Fe^{55}+Co^{60}	193 198 196 201 204 205 205 199 196 201	193	205	12	199.8	15.36
pixel 6 s=73312 x y 608 71	None	200 203 201 200 200 197 198 200 200 207	197	207	10	200.6	6.84
	Al_K	204 198 203 201 197 202 197 201 203 198	197	204	7	200.4	6.44
	Fe^{55}+Co^{60}	204 197 198 205 230 199 202 200 204 198	197	230	33	203.7	84.21
pixel 7 s=76029 x y 253 74	None	197 200 207 209 204 205 201 205 199 205	197	209	12	203.2	12.96
	Al_K	203 199 201 206 202 206 201 200 202 203	199	206	7	202.3	4.81
	Fe^{55}+Co^{60}	200 203 203 205 201 199 202 202 211 204	199	211	12	203	10.0
pixel 8 s=77638 x y 838 75	None	201 199 201 201 198 198 196 200 202 201	196	202	6	199.7	3.21
	Al_K	195 201 198 197 259 200 199 199 202 257	195	259	64	210.7	563.01
	Fe^{55}+Co^{60}	200 198 200 200 197 198 202 197 203 202	197	203	6	199.7	4.21
pixel 9 s=79073 x y 225 77	None	202 201 205 205 201 198 204 208 197 204	197	208	11	202.5	10.25
	Al_K	205 219 204 201 200 203 201 201 199 202	199	219	20	203.5	29.65
	Fe^{55}+Co^{60}	204 199 204 202 202 202 203 201 207 197	197	207	10	202.1	6.89

TABLE 7-2. The Characteristics of "Eventless" Pixel

Pixel Id	X-Ray Source	Bias Levels	Min Bias Level	Max Bias Level	Bias Level Range	Bias Level Mean	Bias Level Variance
pixel 10 s=322496 x y 960 314	None	202 199 201 200 203 198 199 201 203 202	193	203	10	199.60	6.84
	Al_K	199 201 203 202 198 198 197 201 204 197	196	204	7	200.0	5.80
	Fe^{55}+Co^{60}	198 202 202 202 197 196 199 201 201 379	197	379	183	217.70	2895.21
pixel 11 s=323313 x y 753 315	None	203 204 201 202 202 196 202 201 204 203	196	204	8	201.80	4.76
	Al_K	202 202 205 202 205 200 197 206 201 204	197	206	9	202.40	6.64
	Fe^{55}+Co^{60}	202 201 201 199 203 201 198 201 202 202	198	203	5	201.0	2.0
pixel12 s=324802 x y 194 317	None	201 202 198 199 195 201 200 197 198 205	195	205	10	199.60	7.24
	Al_K	199 195 203 199 201 205 199 202 201 198	195	205	10	200.20	7.16
	Fe^{55}+Co^{60}	202 197 203 199 202 197 201 204 201 201	197	204	7	200.70	5.01
pixel 13 s=325051 x y 443 317	None	206 204 201 199 205 200 208 201 201 204	199	208	9	202.90	7.69
	Al_K	203 206 204 201 204 202 201 205 204 202	201	206	5	203.2	2.56
	Fe^{55}+Co^{60}	199 203 205 201 208 205 205 199 203 202	199	208	9	203.0	7.40
pixel 14 s=325391 x y 783 317	None	202 202 202 203 203 201 204 201 200 204	200	204	4	202.20	1.56
	Al_K	200 200 215 200 206 197 201 204 201 204	197	215	10	202.80	22.56
	Fe^{55}+Co^{60}	206 207 202 202 195 201 201 202 201 201	195	207	12	201.80	9.36
pixel 15 s=325594 x y 986 317	None	198 199 200 199 205 201 199 203 200 201	198	205	7	200.50	4.05
	Al_K	201 200 202 201 199 201 201 197 199 204	197	204	7	200.50	3.25
	Fe^{55}+Co^{60}	199 200 205 200 196 202 201 198 200 200	196	205	9	200.10	5.09

TABLE 7-2. The Characteristics of "Eventless" Pixel

Pixel Id	X-Ray Source	Bias Levels	Min Bias Level	Max Bias Level	Bias Level Range	Bias Level Mean	Bias Level Variance
pixel 16 s=325840 x y 208 318	None	205 204 197 199 196 199 204 199 200 197	196	205	9	200.00	9.40
	Al_K	200 203 201 196 201 199 199 207 197 201	196	207	11	200.40	8.64
	Fe^{55}+Co^{60}	194 279 205 200 197 202 199 199 200 197	194	279	85	207.20	580.76
pixel 17 s=325236 x y 628 317	None	203 200 204 200 196 199 200 201 201 204	196	204	8	200.80	5.36
	Al_K	203 200 204 203 196 199 200 199 203 200	196	204	8	200.70	5.61
	Fe^{55}+Co^{60}	248 200 205 196 199 197 200 203 200 201	196	248	52	204.90	212.49
pixel 18 s=325488 x y 880 317	None	198 196 202 198 204 203 199 201 200 198	196	204	8	199.90	5.89
	Al_K	198 199 200 200 251 213 195 197 198 204	195	251	56	205.50	252.65
	Fe^{55}+Co^{60}	202 200 198 199 200 202 203 200 201 201	198	203	5	200.60	2.04
pixel 19 s=325634 x y 2 318	None	203 205 199 200 198 198 199 203 197 196	196	205	9	199.80	7.76
	Al_K	201 201 201 196 203 197 201 198 194 194	194	203	9	198.60	9.44
	Fe^{55}+Co^{60}	196 203 204 202 201 198 195 203 199 199	195	204	9	200.0	8.60

It is clear from Table 7-2 that pixels 5, 6, 8, 10, 16, 17 and 18 possess a higher variance than the others. The sets of pixel values that produce these large variances values possess one or more bias values that are considerable higher than the remainder. For example, pixel 8 has a variance of 563.01 and two high bias level values: 259 and 257. The 259 value came from the raw data values: 776, 199, 204, 203, 194, 198, 191, 200, 221, and 207, whereas the 257 value was calculated from the raw values: 195, 201, 202, 197, 209, 180, 214, 191, 217, and 759. Both sets have one or two values higher than 200 and indicate the presence of events. The pixels with the high variances were considered ill-fitted to the purpose of comparing algorithms and were eliminated.

The pixels chosen had a range of less than 15 between maximum and minimum, and a variance of less than 25. Two set of five pixels were chosen, the first containing pixels 0–4, the second of pixels 11–15, and the two-way ANOVA model was applied to both sets. The ANOVA two-way classification for verifying the properties of these two sample groups is shown in Table 7-3.

TABLE 7-3. Two-way ANOVA Classification to Analyze Eventless Data

Pixels	X-Ray Source		
	none	Al_K	Fe^{55}+Co^{60}
pixel 1	10 replications	10 replications	10 replications
pixel 2	10 replications	10 replications	10 replications
...	...	...	...
pixel 10	10 replications	10 replications	10 replications

The two factors in this analysis are (1) the X-ray sources to which the CCD were exposed during the data collection (even if no event was collected by these pixels) and (2) the pixel number corresponding to the chosen pixel. The bias calculation used the simple MEAN-I1L1 algorithm and was then repeated for MEAN-I2L2, MEAN-I2L1, and MEDIAN to provide a better indication of the algorithm sensitivity to data variations and to improve the subsequent comparisons with simulated data.

The results of the two-way ANOVA analysis are reported in Table 7-4, including the tabulated F_α at a significance level of 0.05. The relative size of F versus F_α determines whether the variance of the two factors and their interaction has a larger variation than the residual error at this level of significance. When $F < F_\alpha$, the null hypothesis is assumed. Similar results were found for all four algorithms and both data sets are summarized in Table 7-4. Since there were insufficient data frames available to generate simulated data for a single pixel, it was decided to combine the data from different pixels even if their values are not statistically equal.

TABLE 7-4. Eventless Model Summary Results

Factors	Null Hypothesis	
	Valid	Invalid
Pixel Number		Pixels are different
X-Ray Source	Data from same pixel and different X-Ray sources are equal	
Interaction	There is no interaction between pixels and X-Ray sources	

To overcome this problem, it was decided to compare the values obtained by the eventless and simulated models and to identify their variations; also to repeat the analysis twice—one for each data set—and then compare the results. This comparison would hopefully confirm the common behavior.

7.2 Simulated Data

The simulated data were generated by adding “events” of known energy to the eventless data. Multiple events in the same data set were simulated by adding multiple instances of the same ADU value to different pixels belonging to the same data set. The energy factor was assigned to the pixel number and event number factor to the X-Ray source. The “events” added to the eventless sample have the following specifications:

- Energy: five levels of 20, 40, 60, 80, and 100 in ADU units, corresponding to energies of approximately 72, 144, 216, 288, and 316 eV.
- Event number: three levels corresponding to one, two, and three events in the same set used for calculating the bias level.

For each energy and each event factor, ten replications of the same condition were generated. The experiment configuration is represented in Table 7-5.

TABLE 7-5. Two-way ANOVA Classification To Analyze Simulated Data

Approximate Energies	Events		
	One Event	Two Events	Three Events
72 eV	10 replications	10 replications	10 replications
144 eV	10 replications	10 replications	10 replications
216 eV	10 replications	10 replications	10 replications
288 eV	10 replications	10 replications	10 replications
316 Ev	10 replications	10 replications	10 replications

This configuration is equivalent to that of Table 7-3 used for characterizing the eventless data, but in Table 7-5 the pixel id is replaced by the energy and the X-Ray source by the number of simulated events.

The bias level was calculated for each pixel in the same way as for the eventless model. Four algorithms were used: MEAN–I1L1, MEAN–I2L2, and MEAN–I2L1, and MEDIAN. The two-way ANOVA model was applied to both data sets, and the results are represented in Table 7-6 along with the tabulated F_α at significance level of 0.05.

The conclusions to be drawn from Table 7-6 are summarized in Table 7-7, which indicates that the bias levels are dependent on the event energy at the 0.05 significance level, whereas the dependency on event number depends on the choice of algorithm.

The interaction between the two factors (energy and number of events) is statistically significant for two of the algorithms (MEAN-I1L1 and MEAN-I2L2) and not for the others (MEAN-I2L1 and MEDIAN). It appears that the interaction between the two factors is significant when the variations generated by the factors are themselves significant.

TABLE 7-6. Simulated Model ANOVA Two-Way Model Results

Data Set ID	Algorithm	Factor	Degrees of Freedom	Mean Square	Calculated F	Tabulated F_α (p=.05)
1	MEAN I1L1	Energy	4	2779.547	507.9720	2.45
		Event Number	2	918.360	167.833	3.07
		Interaction	8	95.403	17.435	2.02
		Residual Error	135	5.472		
	MEAN I2L2	Energy	4	1756.080	183.109	2.45
		Event Number	2	1940.790	202.369	3.07
		Interaction	8	239.605	24.984	2.02
		Residual Error	135	9.590		
	MEAN I2L1	Energy	4	204.213	24.997	2.45
		Event Number	2	1.27	0.155	3.07
		Interaction	8	11.457	1.402	2.02
		Residual Error	135	8.17		
	MEDIAN	Energy	4	147.813	16.27	2.45
		Event Number	2	26.823	2.952	3.07
		Interaction	8	2.347	0.258	2.02
		Residual Error	135	9.085		
2	MEAN I1L1	Energy	4	2414.820	338.421	2.45
		Event Number	2	889.803	124.700	3.07
		Interaction	8	98.415	13.792	2.02
		Residual Error	135	7.136		
	MEAN I2L2	Energy	4	1260.053	107.499	2.45
		Event Number	2	1922.823	164.043	3.07
		Interaction	8	246.972	21.070	2.02
		Residual Error	135	11.721		
	MEAN I2L1	Energy	4	178.753	19.374	2.45
		Event Number	2	29.830	3.233	3.07
		Interaction	8	7.977	0.865	2.02
		Residual Error	135	9.227		
	MEDIAN	Energy	4	116.780	12.511	2.45
		Event Number	2	111.323	11.927	3.07
		Interaction	8	5.955	0.638	2.02
		Residual Error	135	9.334		

TABLE 7-7. Simulation Model - ANOVA Two-Way Model Results

Algorithm	Factor	Null Hypothesis	
		Valid	Invalid
Mean I1L1	Energy		Bias Calculation depends on energy
	Events		Bias Calculation depends on event number
	Interaction		Bias Calculation depends on Energy/ Event interaction
Mean I2L2	Energy		Bias Calculation depends on energy
	Events		Bias Calculation depends on event number
	Interaction		Bias Calculation depends on Energy/ Event interaction
Mean I2L1	Energy		Bias Calculation depends on energy
	Events	Data set 1: Bias Calculation is independent on event number	Data set 2: Bias Calculation depends on event number
	Interaction	Bias Calculation is independent on Energy/Event interaction	
Median	Energy		Bias Calculation depends on energy
	Events	Data set 1: Bias Calculation is independent on event number	Data set 2: Bias Calculation depends on event number
	Interaction	Bias Calculation is independent on Energy/Event interaction	

The MEAN-I1L1 algorithm contains no pixel rejection criterion and the resulting bias levels always respond to the presence of a simulated event. The results from this algorithm shown in Table 7-6 should be considered as a control case for rating the performance of the other algorithms.

Table 7-6 shows that the MEAN-I2L2 algorithm is sensitive to energy level, to event number, and to the interaction between them. This can be explained as follows: the presence of events yields a higher σ than their absence. Since the algorithm rejects pixels whose values deviate from the mean by more than $\pm 2\sigma$, the presence of events increases the acceptance range, and therefore the number of rejected events decreases.

The MEAN-I2L1 algorithm is seen to be sensitive to the energy level but less so to the number of events, which can be explained as follows: as in the MEAN-I2L2 case, the presence of events increases the acceptance range and decreases the rejection level, but the acceptance range is smaller and the number of rejected events increases. The weaker dependence on the number of events may be due to rejection level limits—if the data used for the bias level calculation have values in the neighborhood of the rejection level, they will sometimes be rejected, sometimes not, and the result will become sensitive to the presence of additional noise.

The MEDIAN algorithm is also sensitive to the energy level but less so to the number of events. This algorithm sorts the data values, so when the number of higher value pixels increases, so does the bias level. The bias level values depend upon the values of the eventless data—if several sets contain predominately low pixel values, the MEDIAN algorithm will generate a low bias value and will be not affected by the few higher pixel values. Conversely, if several sets contain mostly higher pixel values, the MEDIAN algorithm may pick one of these higher values.

The comparison between the two pixel data sets indicates that the residual errors are less for the first set than for the second one. This variation might be responsible of the rejection level at the border condition because the data used for the bias calculation might have more variations.

7.3 Comparison Between Eventless and Simulated Models

Since the energy factor used data from different pixels that were not statistically equivalent, it was necessary to compare the F values calculated from the two models, eventless and simulated, as shown in Table 7-8. This indicates that each MEAN algorithm shows an increase in F

TABLE 7-8. Eventless and Simulated Models Summary Comparison

Algorithm	Factor	F Values			
		Eventless Model		Simulated Model	
		Data Set 1	Data Set 2	Data Set 1	Data Set 2
MEAN I1L1	Pixel Number/Energy	18.193	12.728	507.972	338.421
	X-Ray Source/Event Number	0.066	0.253	167.833	124.700
MEAN I2L2	Pixel Number/Energy	19.135	18.454	183.109	107.409
	X-Ray Source/Event Number	0.019	0.026	202.369	164.043
MEAN I2L1	Pixel Number/Energy	18.654	17.752	24.997	19.374
	X-Ray Source/Event Number	0.003	0.053	0.155	3.233
MEDIAN	Pixel Number/Energy	16.723	17.515	16.27	12.511
	X-Ray Source/Event Number	0.119	0.276	2.952	11.927

corresponding to the energy factor and therefore that each algorithm depends on the event energy. The F values for the number of events also show an increase and indicate a dependency on that factor.

The comparison of results from eventless and simulated data for the MEDIAN algorithm shows an increase in F corresponding to the number of events, but none for event energy. Presumably the bias values increased with the number of events because the number of low level data values decreased. This algorithm is less sensitive to event energy because it sorts the pixel values and the resulting bias value only depends on the number of pixels containing events and not on the precise value of the event pixel itself.

To better characterize the algorithms, the *factor mean effects* and the *confidence intervals* between mean differences were calculated for both classes of model. The factor mean effect represents the difference between the factor mean and the experimental grand mean, while the confidence interval is the minimum statistically-significant difference between a pair of factor mean effects. The confidence interval is defined by $t\,s\sqrt{2/n}$, where

- t is the t-Student value, tabulated in Applicable Document 8, for the factor's degrees of freedom;
- s is an estimate of the sigma of the factor mean difference. In this case, we use the square root of the residual error, which is a good estimator once the factor dependencies have been eliminated;
- n is the number of replications.

Two factors were considered— X-Ray source and pixel number for the eventless model, and energy and event number for the simulated model. Their mean effect values for a particular set of test pixels are shown in two figures, Figure 7-1 for the eventless model and Figure 7-2 for the simulated-event model. Each graph shows the factor mean effects calculated for four algorithms operating on the same data set.

FIGURE 7-1. EventLess Model: Bias Level Residual for Two-Way ANOVA

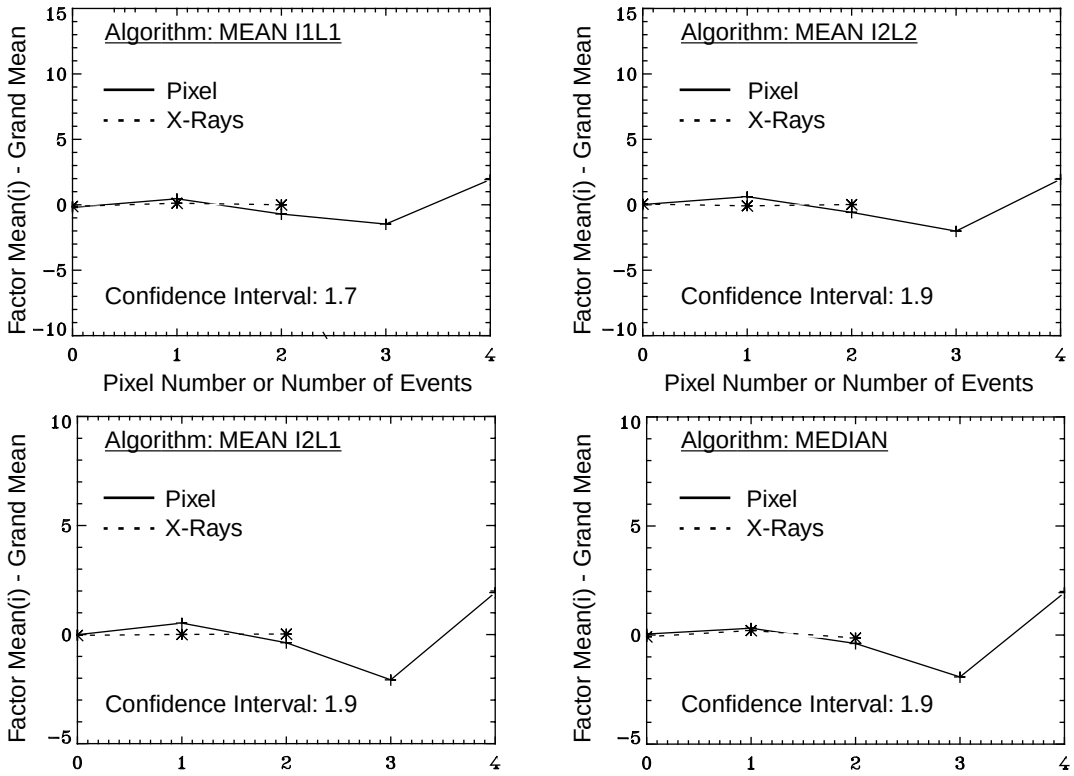
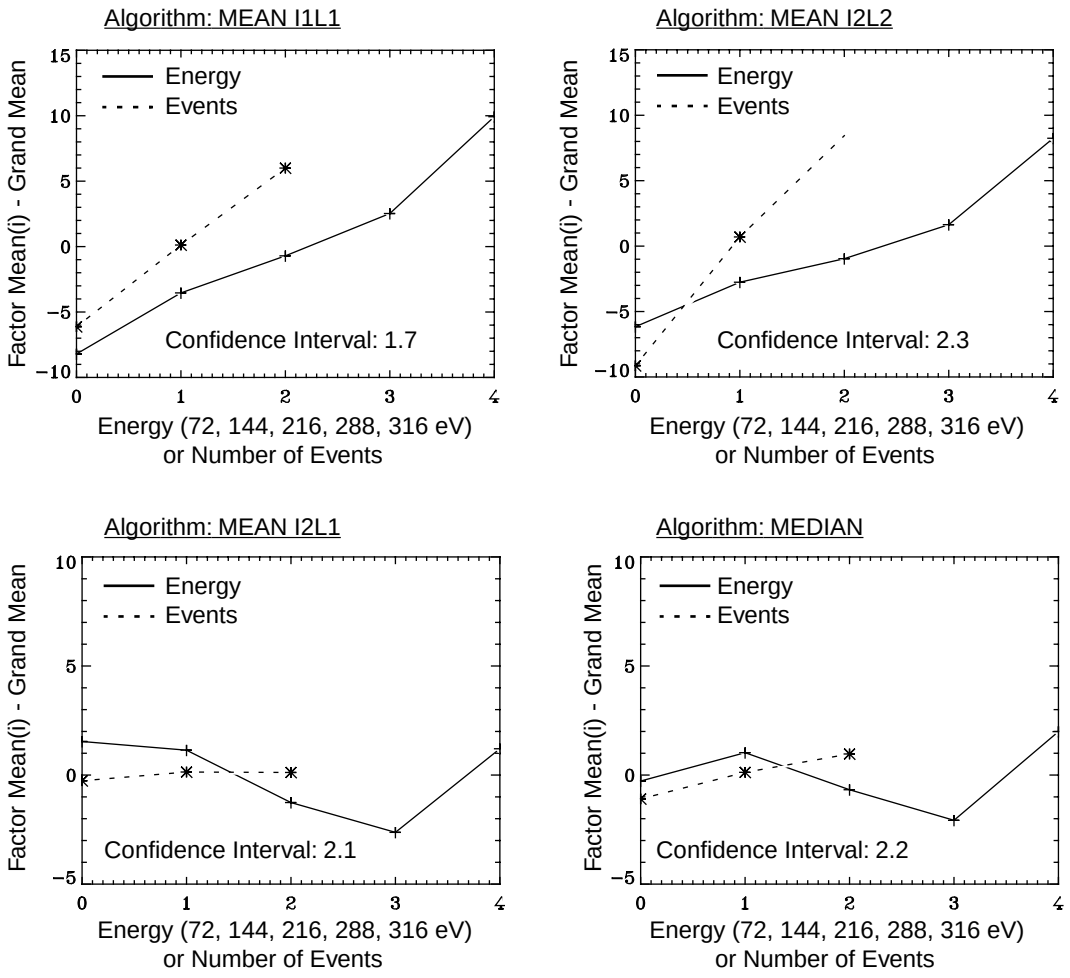


Figure 7-1 indicates that the differences between mean effects for the eventless model are within the confidence interval for the X-Ray factor but not for the pixel factor—the pixel interval confidence is

approximately ±1.9 and some differences are greater than this value. This behavior is observed in each algorithm. Figure 7-2 shows that the differences between energy mean effects for the simulated event model are not within the confidence interval for all algorithms; also that the differences between event number mean effects are outside the confidence interval for some algorithms and dependent on internal pixel noise for others.

FIGURE 7-2. Simulated Model: Bias Level Residual for Two-Way ANOVA



The energy dependence also differs according to the chosen algorithm. This is shown more clearly in Figure 7-3 and Figure 7-4 which represent the expected bias values calculated from the ANOVA model for the same data set.The figures confirm the previous result, and provide more insight into the variation of bias level as a function of the separate factors.

FIGURE 7-3. EventLess Model: Expected Bias Level Values for Pixel, X-Ray Sources

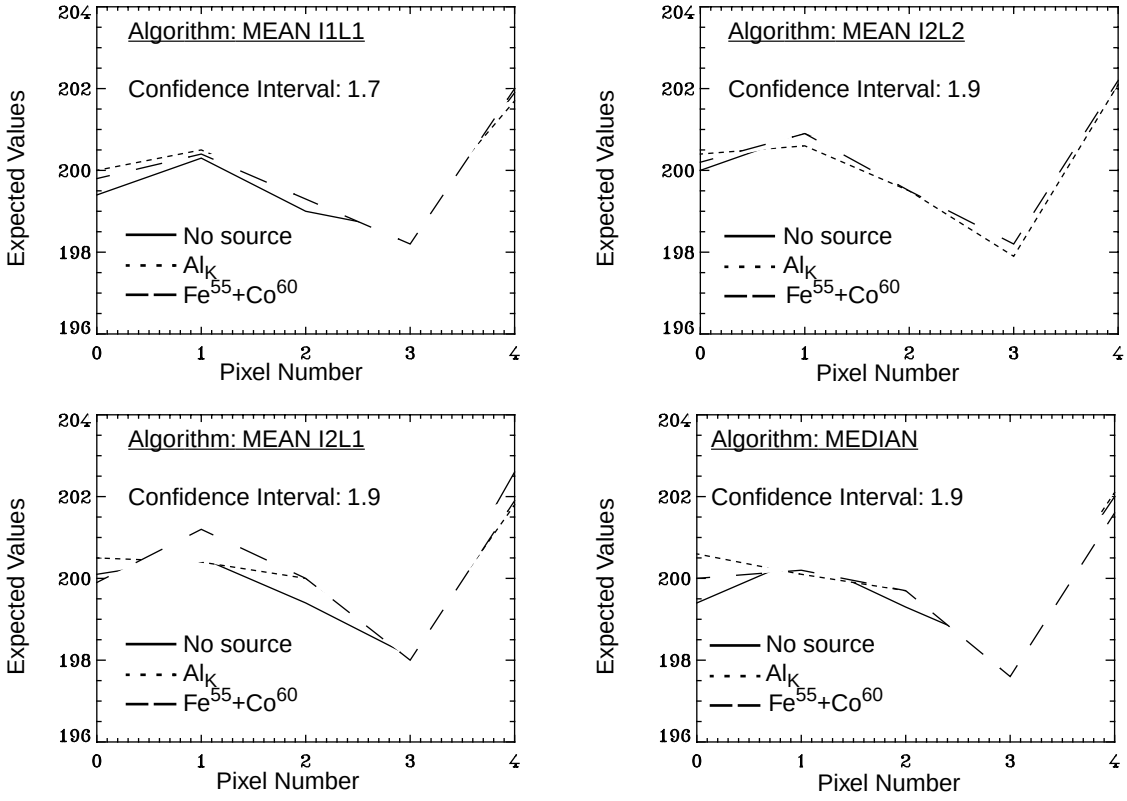


Figure 7-3 indicates that the variations of the expected bias levels with respect to the choice of X-Ray source are within the level of confidence for each algorithm, whereas the variations with respect to the pixel number depend upon the pixel. The graphs do not show any variation of expect value with respect to the choice of algorithm.

Figure 7-4 shows that the variations of the expected values with respect to the number of events and energy depend upon the choice of algorithm. The expected data from the MEAN-I1L1 and MEAN-I2L2 algorithms show a dependency on event energy and on number of events that is considerably larger than the confidence level, whereas the dependency of the MEAN-I2L1 and MEDIAN algorithms is at the border line of the confidence level.

The results from the eventless and simulated event models are summarized in Table 7-9. The study clearly indicates that both the MEAN-I2L1 and MEDIAN algorithms provide “robust” values of pixel bias in the presence of energetic and soft events, with a slight preference in favor of MEAN-I2L1.

FIGURE 7-4. Simulated Model: Expected Bias Level Values for Energy, Event Number

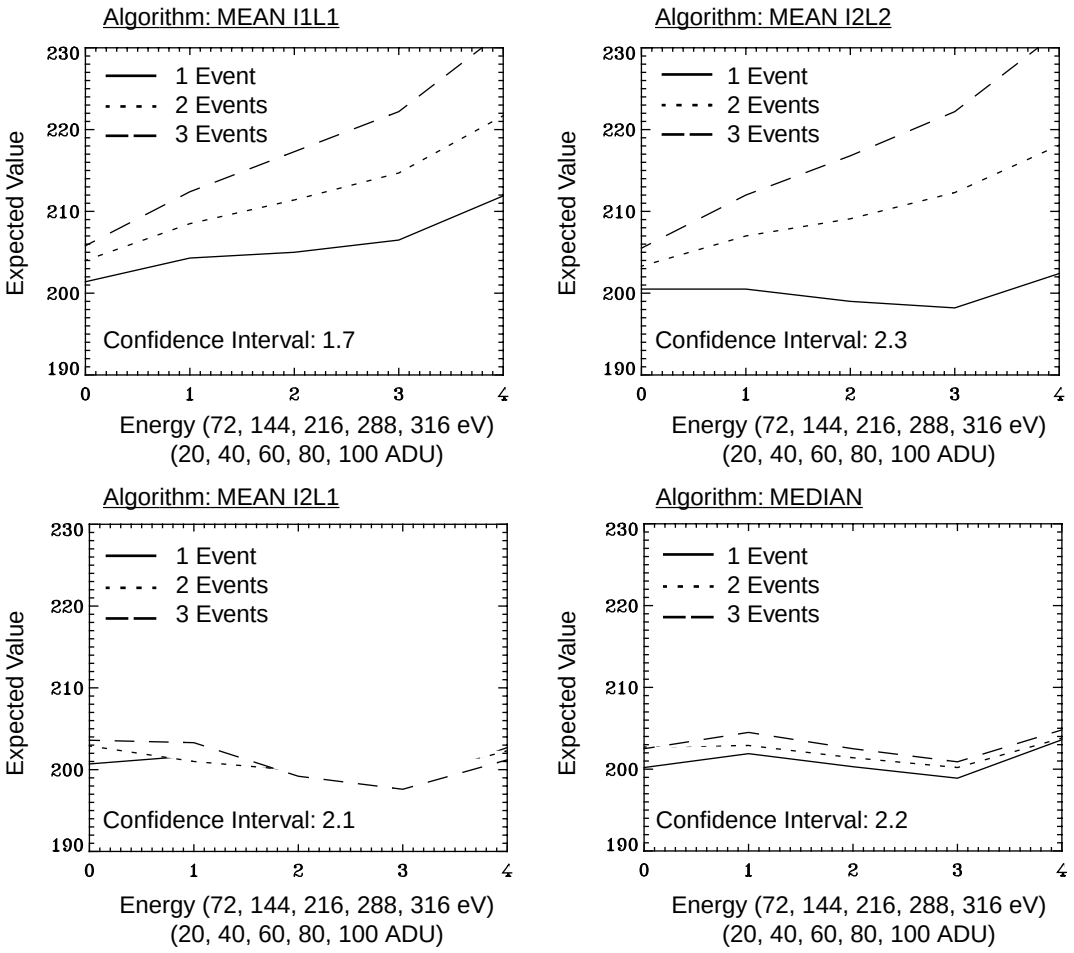


TABLE 7-9. Eventless and Simulated Model Summary Results

Algorithm	Factor	Dependency	
		Eventless Model	Simulated Model
MEAN I1L1	Pixel Number/Energy	YES	YES - heavy dependency
	X-Ray Source/Event Number	NO	YES - heavy dependency
MEAN I2L2	Pixel Number/Energy	YES	YES - heavy dependency
	X-Ray Source/Event Number	NO	YES - heavy dependency
MEAN I2L1	Pixel Number/Energy	YES	YES - slight dependency
	X-Ray Source/Event Number	NO	Depend on the internal noise
MEDIAN	Pixel Number/Energy	YES	YES - slight dependency
	X-Ray Source/Event Number	NO	Depend on the internal noise

8.0 Conclusions

While the factor analysis of Section 6.0 proves to be most effective in identifying anomalous pixels, neither it nor Student's test described in Section 7.0 is capable of pointing to a particular bias algorithm as the best under all possible circumstances of CCD type, illumination source, or history of irradiation. Of those tested, both the MEDIAN and the MEAN–I2L2 (iterated mean with 2–σ outlier rejection) proved to be highly effective in circumstances under which less than 30% of the samples of a given pixel contain events. For a set of 10 bias exposures from a CCD containing 10^6 pixels, experiencing 10,000 "events" per exposure, this will affect less than one pixel.

We recommend that both MEDIAN and MEAN–I2L2 algorithms are implemented in the ACIS FEPs, along with a version of the algorithm described in Section 3.1. The former require at least 100 exposures, the latter at least 40, but none is capable of processing each frame within a nominal 2.6 second readout time. From a preliminary analysis of prototype computer code, the additional calibration time will amount to approximately 120 seconds for MEDIAN and MEAN–I2L2 and 60 seconds for the algorithm involving no additional storage.

Appendix A. Statistical Techniques

A.1 Analysis Of Variance

The analysis of variance (commonly referred to, both in the literature and within this document, by the acronym ANOVA) investigates simultaneously data sets that were obtained from several observations. Is is assumed that some of the sets (termed “replications”) were gathered under identical conditions, whereas others were gathered under changing conditions that involve one or more “factors”. Data sets collected under a change of factors are said to belong to a “level”.

These techniques are applied to random observations and calculate additive means and variances. More precisely, they partition the total sum of squares of deviations from the overall mean into two or more component sums of squares, each of which is either associated with a particular factor or with the experimental measurement error. They subdivide the total number of degrees of freedom between the different sums of squares, and compare the effect on the mean of the factors and their levels, corresponding to the component sums of squares, by using a set of statistical tests (F tests) that indicate whether the observed differences in mean factor effects are real or random.

Compared with less sophisticated statistical techniques, these methods yield conclusions of greater generality because (1) the same hypothesis can be applied simultaneously to several factors and several levels of the same factor, and (2) the interactions between factors can be identified at the same time.

In the present case, these techniques can test the dependency of the choice of bias algorithm on the energy and number of CCD events. Several models of analysis of variance are available. In this document we concentrate on One-Way and Two-Way classifications.

A.1.1 One-Way Classification or One Factor Experiment

In this model the data are classified according to one factor, termed a “treatment”, namely the illumination source (or lack thereof) to which the CCD was exposed while the data were being collected.

In general, the ANOVA technique compares data from *N* different treatments, where each treatment has been replicated *M* times. It separates and estimates the variances among and within treatments, and determines whether the values are dependent on the treatments. In our case, it can be used to determine whether the bias level values calculated by a single algorithm from data collected in presence of different X-ray sources are dependent on those sources.

The naming conventions used by this model are illustrated in Table A-1,where *N* is the total number of treatments, *M* the total number of replications of the same treatment, *n* is the generic treatment and *m* is the generic replication, and $x_{n,m}$ are the bias values corresponding to treatment *n* and replication *m.*

TABLE A-1. ANOVA One-Way Data Configuration

Treatments	Replications		
	1	**m**	**M**
1	$x_{1,1}$	$x_{1,m}$	$x_{1,M}$
2	$x_{2,1}$	$x_{2,m}$	$x_{2,M}$
n	$x_{n,1}$	$x_{n,m}$	$x_{n,M}$
N	$x_{N,1}$	$x_{N,m}$	$x_{N,M}$

We assume (1) that the replications $x_{n,m}$ on the same treatment represent random samples from a normal population, and (2) that their mean values and variances are additive. A mathematical model that might describe such a set of data is:

$$x_{n,m} = \mu + \alpha_n + e_{n,m} \qquad 1 \le n \le N \qquad 1 \le m \le M \tag{A.1}$$

where

- μ = grand mean
- α_n = deviation of the *n*-th treatment mean from the grand mean
- $e_{n,m}$ = residual error of the *m*-th data receiving treatment *n* from the mean treatment *m*-th.

The $e_{n,m}$ have independent normal distribution with mean zero, and the α_n are not independent but satisfy the condition $\sum_{n=1,N} \alpha_n = 0$. Associated with such a model is the decomposition of the observations:

$$x_{n,m} = \bar{X} + (\bar{X}_n - \bar{X}) + (x_{n,m} - \bar{X}_n) \tag{A.2}$$

where:

- $x_{n,m}$ represents the general data from the observations.
- $\bar{X} = \frac{1}{N \times M} \times \sum_{n=1,N \; m=1,M} x_{n,m}$ indicates the grand mean and is identified by μ in the mathematical model.
- $\bar{X}_n = \frac{1}{M} \times \sum_{m=1,M} x_{n,m}$ indicates the *n*-th treatment mean (mean of the *n*-th row).
- $\alpha_n = \bar{X}_n - \bar{X}$ is the deviation of *n*-th treatment mean from the grand mean and is represented in the mathematical model by α_n.
- $e_{n,m} = x_{n,m} - \bar{X}_n$ is called the residual error because it represents what it is left over after the grand mean and treatment differences have been allowed for.

This test model has two independent sources of variations operating at the same time which are:

- analysis error, corresponding to the variance among treatments.
- sampling error, corresponding to the variance within treatments or the residuals.

The resulting variance is the sum of the separate variances and is expressed by the equation:

$$V = \sum_{n=1}^{N}\sum_{m=1}^{M}(x_{n,m} - \bar{X})^2 = M\sum_{n=1}^{N}(\bar{X}_n - \bar{X})^2 + \sum_{n=1}^{N}\sum_{m=1}^{M}(x_{n,m} - \bar{X}_n)^2 \quad \text{(A.3)}$$

where:

$$\bar{X} = \frac{\sum_{n=1}^{N}\sum_{m=1}^{M} x_{n,m}}{N \times M} = \frac{T}{N \times M} \quad \text{(A.4)}$$

$$\bar{X}_n = \frac{\sum_{m=1}^{M} x_{n,m}}{M} = \frac{T_n}{M} \quad \text{(A.5)}$$

While V represents the total variance, the first term on the right hand side of equation A.3 indicates the variance among treatments, and the second the variation within treatments or the residual error. The algebraically equivalent form for numerical work is represented by Table A-2.

TABLE A-2. Analysis of Variance for One Factor Experiments with Replications

Source of Variation	Variation	Mean Square	Degrees of Freedom	F	F Degrees of freedom
Among Treatments V_n	$V_n = \sum_n \frac{T_n^2}{M} - \frac{T^2}{N \times M}$	$S_n^2 = \frac{V_n}{(N-1)}$	N - 1	$\frac{S_n^2}{S_e^2}$	numerator N-1 denominator $N \times (M-1)$
Within treatments V_e	$V_e = V - V_n$	$S_e^2 = \frac{V_e}{N \times (M-1)}$	$N \times (M-1)$		
Total variance V	$V = \sum_{n,m} x_{n,m}^2 - \left(\frac{T^2}{N \times M}\right)$		$N \times M - 1$		

To test the “null” hypothesis—that data values do not depend on treatments—it is necessary to evaluate F, the ratio between the mean square among treatments and the mean square within treatments, and compare it with the tabulated $F_\alpha(r,s)$ function (Applicable Document 8), where *r* is the numerator degrees of freedom and *s* is the denominator degrees of freedom and α indicates the

significance level at which to verify the hypothesis. We accept the null hypothesis at a significance level of α if the calculated F does not exceed the tabulated F_α, whereas, if the null hypothesis is false, the F test will fail and the treatments are shown to be statistically distinct.

For example, let the model represent the bias levels calculated by an algorithm from data collected in the presence of various sets of X-ray sources. If the null hypothesis is true, the choice of source will not affect the bias calculated by that algorithm, whereas, if the F test fails, the bias level calculation is affected by the presence of the X-rays.

The α_n in equation A.1 are indices of the treatment variations with respect to the grand mean, μ, and their representations are a useful indicator of the corresponding treatment *n*-th. The expected values from the grand mean can be calculated by the equation $\hat{x}_n = \mu + \alpha_n$, which represents the mean of each treatment, and these values can be used to evaluate treatment differences.

A.1.2 Two-Way Classification or Two Factor Experiments

In this model the data are classified in two dimensions, described by two factors which vary simultaneously. In general the model compares data sets that have been influenced by two different factors, each of which possesses several levels, and the measurements have been replicated *M* times. The model separates and estimates the variances of the two factors, the interaction between them, and the residual measurement error. It determines whether the data are dependent on the factor levels or on their interactions.

In the present example, the model can determine whether the bias values calculated by a particular algorithm from data collected in the presence of events depends on the number of events in the original data and their energy. The same bias calculation is repeated several times with different data corresponding to the same factor levels.

TABLE A-3. ANOVA Two-Way Data Configuration

Factor 1 (Rows) Levels	Factor 2 (Columns)			
	Level 1		Level P	
	Replication 1	Replication M	Replication 1	Replication M
1	$x_{1,1,1}$	$x_{1,1,M}$	$x_{1,P,1}$	$x_{1,P,M}$
2	$x_{2,1,1}$	$x_{2,1,M}$	$x_{2,P,1}$	$x_{2,P,M}$
.	.	...	...	...
N	$x_{N,1,1}$	$x_{N,1,M}$	$x_{N,P,1}$	$x_{N,P,M}$

The naming conventions for this model are illustrated in Table A-3. In the literature, the two factors are generally called Rows and Columns because they are estimated by sums on rows and columns of the table representing the data configuration. They have nothing whatever to do with rows and columns of an X-ray CCD. In this example, *N* is the number of row levels, *P* the number of column

levels, and *M* the number of replications. We assume that the replications $x_{n,p,m}$ on the same row and column levels represent random samples from a normal population, and possess additive means and variances. The mathematical model that might describe such a set of data is:

$$x_{n,\,p,\,m} = \mu + \alpha_n + \beta_p + \delta_{n,\,p} + e_{n,\,p,\,m},\ 1 < p < P,\ 1 \le n < N, \text{ and } 1 \le m \le M \qquad \text{(A.6)}$$

and where:

- μ = grand mean
- α_n = deviation of the row mean effect at level *n* from the grand mean
- β_p = deviation of the column mean effect at level *p* from the grand mean
- $\delta_{n,\,p}$ = interaction between row mean at level *n* and column mean at level *p* (a joint effect, beyond the total of their individual effects)
- $e_{n,p,m}$ = residual error of the *m*-th data receiving row 1 at level *n* and column 2 at level *p*.

The $e_{i,j,m}$ have an independent normal distribution with mean zero, and the α_n and β_p are not independent but satisfy the conditions $\sum_n \alpha_n = \sum_p \beta_p = 0$. The dependency of the interaction is expressed by the condition: $\sum_{n,\,p} \delta_{n,\,p} = 0$. Associated with such a model is a decomposition of the observations:

$$x_{n,\,p,\,m} = \bar{X} + (\bar{X}_n - \bar{X}) + (\bar{X}_p - \bar{X}) + (\bar{X}_{n,\,p} - \bar{X}_n - \bar{X}_p + \bar{X}) + (x_{n,\,p,\,m} - \bar{X}_{n,\,p}) \qquad \text{(A.7)}$$

where:

- $x_{n,p,m}$ indicates the general data from the observations
- $\bar{X} = \frac{1}{N \times P \times M} \times \sum_{n=1}^{N} \sum_{p=1}^{P} \sum_{m=1}^{M} x_{n,\,p,\,m}$ indicates the grand mean.
- $\bar{X}_n = \frac{1}{P \times M} \times \sum_{p=1}^{P} \sum_{m=1}^{M} x_{n,\,p,\,m}$ indicates the mean of the *n*-th row.
- $\bar{X}_p = \frac{1}{N \times M} \times \sum_{n=1}^{N} \sum_{m=1}^{M} x_{n,\,p,\,m}$ indicates the mean of the *p*-th column.
- $X_n - X$ is the deviation of *n*-th row mean from the grand mean and is represented in the mathematical model by α_n.
- $\bar{X}_p - \bar{X}$ is the deviation of the *p*-th column mean from the grand mean and is represented in the mathematical model by β_p
- $\delta_{n,\,p} = \bar{X}_{n,\,p} - \bar{X}_n - \bar{X}_p + \bar{X}$ indicates the interaction between the row at level *n* and the column at level *p* and is represented in the mathematical model by $\delta_{n,\,p}$
- $e_{n,\,p,\,q} = x_{n,\,p,\,q} - X_{n,\,p}$ is called the error residual because it represents what it is left over after the grand mean and factor differences and their interactions have been allowed for.

This model has two or more independent sources of simultaneous variation, and the resulting variances are the sum of the separate variances. These sources of errors are:

- V_c — error among rows (replications of rows)
- V_r — error among columns, (replications of columns)
- V_i — error due to the interactions between rows and columns
- V_e — sampling error (replications of the same row and column levels)

These error sources operate independently, and the total variation is obtained by summing them:

$$\sum_{n,\,p,\,m} (x_{n,\,p,\,m} - \bar{X})^2 = P \times M \sum_{n=1}^{N} (\bar{X}_n - \bar{X})^2 + N \times M \sum_{p=1}^{P} (\bar{X}_p - \bar{X})^2 M + M \sum_{m=1}^{M} (\bar{X}_{n,\,p} - \bar{X}_n - \bar{X}_p + \bar{X})^2 + \sum_{n=1}^{N} \sum_{p=1}^{P} \sum_{m=1}^{M} (x_{n,\,p,\,q} - \bar{X}_{jk}) \quad \text{(A.8)}$$

where:

$$\bar{X} = \frac{\sum_{n,\,p,\,m} x_{n,\,p,\,m}}{N \times P \times M} = \frac{T}{N \times P \times M} \quad \text{(A.9)}$$

$$\bar{X}_n = \frac{\sum_{p,\,m} x_{n,\,p,\,m}}{P \times M} = \frac{T_n}{P \times M} \quad \text{(A.10)}$$

$$\bar{X}_p = \frac{\sum_{n,\,m} x_{n,\,p,\,m}}{N \times M} = \frac{T_p}{N \times M} \quad \text{(A.11)}$$

$$\bar{X}_{n,\,p} = \frac{\sum_{m} x_{n,\,p,\,m}}{M} = \frac{T_{n,\,p}}{M} \quad \text{(A.12)}$$

The left hand side of equation A.8 indicate the total variance V, the first term on the right the variance between rows V_r, the second term the variance between columns V_c, the third the variance due to the interactions between columns and rows V_i, and the fourth the residual variation V_e.

An equivalent form of this model is represented in Table A-4. In this case, the “null” hypothesis, i.e. that the results are statistically independent of one or more factors, can be tested by calculating the three F functions listed in the 5th column, and comparing their values with tabulated $F_\alpha(r, s)$ (see Applicable Document 8), where *r* are the degrees of freedom of the numerator and *s* the degrees of freedom of the denominator.

α_n, β_p, and $\delta_{n,\,p}$ in equation A.6 are indices of the variation among the two factors, and their interaction at the *n*-th and *p*-th levels with respect to the grand mean, μ. These parameter

TABLE A-4. Analysis of Variance for Two Factor Experiments with Replications

Source of Variation	Variation	Mean Squares	Mean Square Degrees of Freedom	F	F Degrees of freedom
Among Rows V_r	$\sum_n \frac{T_n^2}{P \times M} - \frac{T^2}{N \times P \times M}$	$S_r^2 = \frac{V_r}{(N-1)}$	N-1	$\frac{S_r^2}{S_e^2}$	numerator N-1 denominator NP(M-1)
Among Columns V_c	$\sum_p \frac{T_p^2}{N \times M} - \frac{T^2}{N \times P \times M}$	$S_c^2 = \frac{V_c}{(P-1)}$	P-1	$\frac{S_c^2}{S_e^2}$	numerator P-1 denominator NP(M-1)
Interaction V_i	$\frac{1}{M}\sum_{n,p} T_{n,p}^2 - \frac{T^2}{N \times P \times M} - V_c - V_r$	$S_i^2 = \frac{V_i}{(N-1)(P-1)}$	$(N-1) \times (P-1)$	$\frac{S_i^2}{S_e^2}$	numerator (N-1)(P-1) denominator NP(M-1)
Residual V_e	$V - V_r - V_c - V_i$	$S_e^2 = \frac{V_e}{N \times P(M-1)}$	$N \times P \times (M-1)$		
Total variance V	$V = \sum_{n,p,m} x_{n,p,m}^2 - \frac{T^2}{N \times P \times M}$		$N \times P \times M - 1$		

representations are a useful indicator of the corresponding factors and interactions. Their values are calculated by partitioning equation A.8. The expected values calculated from the grand mean and α_n, β_p, and $\delta_{n,p}$ can be calculated by the equation $\hat{x}_{n,p} = \mu + \alpha_n + \beta_p + \delta_{n,p}$, which represent the mean with respect to rows, columns, and their interactions at level *n* and *p*. The values calculated by this equation can be used to evaluate differences between data corresponding to different factor levels.

A.2 Student's T Test

This test allows us to estimate if there are significant differences between treatments, to calculate confidence intervals for bias level differences, and to estimate the minimum significant differences between bias levels. It permits us to establish a metric with which to compare one bias algorithm against another. The test is valid for any sample size and is a good approximation, especially for large samples, even if the population is significantly non-normal.

In general, Student's "T Score" is calculated by the equations:

$$t = \frac{(\bar{y}_1 - \bar{y}_2)}{\sigma} \qquad \sigma = \sqrt{\frac{(n_1 \times s_1 + n_2 \times s_2)}{(n_1 + n_2 - 2)}} \tag{A.13}$$

where:

- y_1 and y_2 are two values corresponding to the factor 1 and factor 2 at a specified level
- n_1 and n_2 are the number of replications corresponding to the factor 1 and factor 2
- σ is the sigma corresponding to the experiment
- s_1 and s_2 are the estimated sigma for the data corresponding to the factor 1 and factor 2

The test rejects the hypothesis that the two values belong to the same population with a significance level α if their "T Score" value exceeds the value of $t_{\alpha/2}$ (tabulated in Applicable Document 8) with degree of freedom ($n_1 + n_2 - 2$). For multiple comparisons, the confidence interval κ for difference means of the *i*-th and *j*-th treatments, calculated by the analysis of variance, is estimated as follows:

$$\kappa = \bar{y}_i - \bar{y}_j \pm s \times t_{\nu, \frac{\alpha}{2}} \sqrt{\frac{1}{n_i} + \frac{1}{n_j}} \tag{A.14}$$

where:

- $y_i - y_j$ indicates the observed difference between the *p*-th and *q*-th treatments.
- $t_{\alpha, \nu/2}$ indicates the tabulated value of t for ν degrees of freedom within treatments.
- $\sqrt{s \times \left(\frac{1}{n_i} + \frac{1}{n_j}\right)}$ is the estimated variance of treatment differences when σ is estimated by the mean square difference within treatments of s^2.

The minimum significative difference between treatments is estimated by comparing the mean calculation with respect to a standard treatment,[3] and is defined by:

$$t_{\nu, \alpha/2} \cdot s \sqrt{\frac{1}{n_i} + \frac{1}{n_j}} \tag{A.15}$$

3. Special data sets used as benchmarks to compare against specific treatments.