Statistical Techniques for Analyzing Process or “Similarity” Data in TID Hardness Assurance

R. Ladbury

Abstract—We investigate techniques for estimating the contributions to TID hardness variability for families of linear bipolar technologies, determining how part-to-part and lot-to-lot variability change for different part types in the process.

Index Terms—radiation effects, reliability estimation, quality assurance

INTRODUCTION

Radiation Hardness Assurance (RHA) methodologies against Total Ionizing Dose (TID) degradation impose rigorous statistical treatments for data from a part’s Radiation Lot Acceptance Test (RLAT)[1] and/or its historical performance.[2],[3],[4],[5] However, no similar methods exist for rigorously statistical treatments for data from a part’s Radiation Hardness Assurance (RHA) methodologies—e.g. —radiation effects, reliability estimation, quality assurance against Total Ionizing Dose (TID) degradation imposing rigorous statistical treatments for data from a part’s Radiation Lot Acceptance Test (RLAT)[1] and/or its historical performance.[2],[3],[4],[5] However, no similar methods exist for determining how part-to-part and lot-to-lot variability change for different part types in the process. We then develop analysis techniques and apply them to the different data sets.

TABLE I: ADI BIPOLAR OP AMP FAILURE DOSES—KRAD(SI)

<table>
<thead>
<tr>
<th>Part #</th>
<th># lots</th>
<th>Mean Failure Dose</th>
<th>Avg. Lot Stand. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>OP177</td>
<td>1</td>
<td>2.5</td>
<td>1.7</td>
</tr>
<tr>
<td>OP400</td>
<td>4</td>
<td>5.32</td>
<td>1.2</td>
</tr>
<tr>
<td>OP497</td>
<td>1</td>
<td>5.7</td>
<td>4.5</td>
</tr>
<tr>
<td>OP467</td>
<td>2</td>
<td>9.2</td>
<td>1.9</td>
</tr>
<tr>
<td>OP07</td>
<td>1</td>
<td>11.1</td>
<td>2</td>
</tr>
<tr>
<td>OP470</td>
<td>1</td>
<td>12.5</td>
<td>2</td>
</tr>
<tr>
<td>OP470</td>
<td>1</td>
<td>12.5</td>
<td>2</td>
</tr>
<tr>
<td>OP11</td>
<td>1</td>
<td>12.6</td>
<td>2</td>
</tr>
<tr>
<td>OP77</td>
<td>3</td>
<td>16.4</td>
<td>2</td>
</tr>
<tr>
<td>OP200</td>
<td>1</td>
<td>22</td>
<td>1.7</td>
</tr>
<tr>
<td>OP27</td>
<td>3</td>
<td>29.5</td>
<td>4.3</td>
</tr>
</tbody>
</table>

TABLE III: ADI OP AMP FAILURE DOSSES (KRAD(SI)) FOR MULTI-LOT SAMPLES

<table>
<thead>
<tr>
<th>Part #</th>
<th>Lot Mean Fail. Level</th>
<th>Lot St. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>OP27(1)</td>
<td>39.00</td>
<td>2.80</td>
</tr>
<tr>
<td>OP27(2)</td>
<td>21.00</td>
<td>3.60</td>
</tr>
<tr>
<td>OP27(3)</td>
<td>28.50</td>
<td>5.07</td>
</tr>
</tbody>
</table>

Mean | 29.50 | 4.29
St. Dev. | 9.04 | 1.29

TABLE II: AFLB FOR LTC RH SERIES PARTS (a@50 KRAD(SI))

<table>
<thead>
<tr>
<th>Part #</th>
<th>Function</th>
<th>Mean AFLB</th>
<th>St. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>RH1499</td>
<td>Op Amp</td>
<td>5%</td>
<td>0.02</td>
</tr>
<tr>
<td>RH1119</td>
<td>V. Comp.</td>
<td>45%</td>
<td>0.04</td>
</tr>
<tr>
<td>RH1011</td>
<td>V. Comp.</td>
<td>56%</td>
<td>0.03</td>
</tr>
<tr>
<td>RH108</td>
<td>Op Amp</td>
<td>56%</td>
<td>0.10</td>
</tr>
<tr>
<td>RH1014</td>
<td>Op Amp</td>
<td>26%</td>
<td>0.06</td>
</tr>
<tr>
<td>RH1078</td>
<td>Op Amp</td>
<td>151%</td>
<td>0.04</td>
</tr>
<tr>
<td>RH27</td>
<td>Op Amp</td>
<td>215%</td>
<td>0.76</td>
</tr>
</tbody>
</table>

Mean | 8.75 | 7.70
St. Dev. | 9.20 | 4.50

Mean | 21.19 | 4.08
St. Dev. | 17.50 | 2.20

Mean | 10.50 | 1.60
St. Dev. | 16.40 | 2.63

For all datasets, we chose either a normal distribution (if the quantity being fit was distributed from $-\infty$ to $+\infty$) or a lognormal fit (for quantities with values from 0 to $+\infty$).
failure dose). However, since fit parameters for the lognormal distribution can be found by fitting the logarithms of the data to the desired distribution, we will represent the resulting means with the symbol $\mu$ and the standard deviations by $\sigma$.

Because similarity data must consider TID degradation in many different part types, a first challenge is developing a meaningful criterion for comparing degradation across such different part types. For the parts in Tables I and III, we defined failure dose as that where the first parameter (usually input bias current, $I_b$) goes out of specification for the device.

Such a failure criterion will not work for parts in Table II, since none of the parts failed parametrically or functionally at the highest dose of 50 krad(Si). Here, we compare the parts’ parametric degradation—with input bias current change ($\Delta I_b$) as a proxy for degradation. Because pre-rad specifications of $I_b$ varied widely from part to part, we normalized input bias current changes to pre-rad values (to arrive at $\%\Delta I_b$). Although the LTC site has only one lot of data for each part type, previous data show series performance to be quite stable from lot to lot. For instance, 38 lots of RH1014 op amps showed that mean lot $\Delta I_b$ varying by less than 2-3x from 60-200 krad(Si).[5]

Exploratory data analysis[8] revealed neither systematic correlation between mean failure level and standard deviation about the mean, nor any systematic time dependence of TID results that could indicate a process change. As such, we determined distributions for mean failure doses and standard deviations independently, assuming all parts in a given process are representative of a single distribution.

II. INFERENCE WITH LIMITED DATA

Our goal is to quantify types of variation that affect TID response for parts fabricated in a given process. These include part-to-part variation within the flight lot and variation in both mean hardness and part-to-part variability from lot to lot. We can estimate part-to-part variation within a lot using RLAT data.[1] We can also bound it to a desired confidence level if lot-to-lot variation is well behaved and we have enough representative historical data.[2]-[5] Likewise, unless the flight parts are exceptional, we can bound lot-to-lot variation with a large enough dataset of data for similar parts. However, the dataset must include data for enough part types to be a representative sample for the process. That is, we must be able to assess variability of mean hardness and part-to-part variation across lots for several (>3) part types in the process.

Because many of the parts in Table I include data only for a single lot, it is not possible to disentangle lot-to-lot variation from the part-type-to-part-type contribution. Rather, the rank plot in Fig. 1 (upper) shows the probability (abscissa) that the mean failure dose of a random lot of a random part type drawn from the process will exceed a given failure dose. Assuming Weibull statistics (which give the best fit), with 90% confidence 90% of lots in the process will not exhibit first failure below 3.3 krad(Si), and a “typical” lot of a typical part will be hard to ~13 krad(Si). The lower plot in figure 1 is a similar rank plot for the standard deviation of failures in a lot.

Likewise, for part types in Table II, we have data only for a single lot. However, while we have data for fewer part types, the RH series is a radiation hardened process, so we expect lot-to-lot and part-to-part variation to be moderate.

As such, the larger part-to-part variation exhibited by the RH27—sufficient to distort the lognormal fit to the other parts (Fig. 2a, b)—is surprising. The part also exhibited the highest overall $\%\Delta I_b$, although this was less out of family (Fig. 2c). A query to LTC[9] revealed that the RH27 uses the same design as the commercial OP27 with no additional design hardening. As such, it likely represents a worst case for parts fabricated in the RH process. In contrast, the RH1499 exhibits very little proportional increase in $I_b$, as well as excellent part-to-part consistency because it was specifically designed to capitalize on the hardness of the RH process. This part probably approaches the limit of what can be achieved with process and design hardening.
Even minimally restrictive data can constrain failure distributions. For the data in Table IV on ADI’s eXtra-Fast Complementary Bipolar (XFCB) process,[10] we know only that all parts performed within specifications at the highest dose (column III). Such suspension data (that is where the test is suspended before failure is seen) provide poor constraints on the failure distribution, since it can be fit equally well by a narrow distribution centered just above the highest test levels or a broad distribution centered far above the failure levels. These two extremes give very different answers as we look at behavior in their tails. Even so, a fit of this data to a lognormal distribution (Fig. 3) shows that >99% of parts in the XFCB process will survive 45 krad(Si) with 90% confidence unless the failure distribution is extremely broad (e.g. even worse than a random failure distribution).

**TABLE IV: SUSPENSION DATA FOR ADI XFCB PARTS**

<table>
<thead>
<tr>
<th>Part #</th>
<th>Type</th>
<th>Dose krad(Si) (Dose Rate)</th>
<th># Parts</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD6640</td>
<td>ADC</td>
<td>&gt;100 (50 rad(Si)/s)</td>
<td>5</td>
</tr>
<tr>
<td>AD6640</td>
<td>ADC</td>
<td>&gt;1000</td>
<td>5</td>
</tr>
<tr>
<td>AD9042</td>
<td>ADC</td>
<td>&gt;1000</td>
<td>5</td>
</tr>
<tr>
<td>AD8184</td>
<td>MUX</td>
<td>&gt;103 (0.06 rad(Si)/s)</td>
<td>10</td>
</tr>
<tr>
<td>AD8063</td>
<td>Op Amp</td>
<td>&gt;103 (0.06 rad(Si)/s)</td>
<td>10</td>
</tr>
<tr>
<td>AD8001</td>
<td>Op Amp</td>
<td>&gt;100 (1.5 rad(Si)/s)</td>
<td>5</td>
</tr>
<tr>
<td>AD8001</td>
<td>Op Amp</td>
<td>&gt;1000</td>
<td>5</td>
</tr>
<tr>
<td>AD8009</td>
<td>Op Amp</td>
<td>&gt;1000</td>
<td>5</td>
</tr>
</tbody>
</table>

**Figs. 2 and 3**

- **Fig. 2**: Rank plots of mean and standard deviation for \(\%\Delta I_b\). a) The RH27’s standard deviation is an outlier, distorting the fit to the other part types (b). c) A Weibull form fits the mean increase in \(I_b\), with the RH27 distorting the fit only slightly.
- **Fig. 3**: Estimated lower-limit doses where 99% of XFCB parts would pass based on lognormal fits of Table IV’s data. Parameters (\(\mu, \sigma\)) within the 90% confidence interval are in unshaded cells. The analysis shows that the 99/90 lower limit for the process is 45 krad(Si) unless the failure distribution is exceptionally broad (\(\sigma>0.9\)), which is very unlikely given prior experience.

**III. DISENTANGLING VARIABILITY**

The data for multi-lot samples of ADI op amps in Table III allow us to estimate the contributions of part-to-part, lot-to-lot and part-type-to-part-type variation to the overall process variability. However, because our dataset is still small, we want a process that is robust and relatively insensitive to fluctuations so that it evolves smoothly as we add more data. Our basic philosophy can be summarized as: “When in doubt, average.” Averages tend to be robust against minor sampling fluctuations and other random errors. As a measure of variability, we use the standard deviation (normal or lognormal), which, as the square root of the variance, or second moment, is also robust.
We describe the method in figure 4. For each lot of each part type, we start with the mean failure dose and the standard deviation of failures about that mean. Then we fit these means and standard deviations (both positive definite quantities) to their own lognormal distributions for each part type. This results in 4 parameters as in reference 5:

1) $\mu_i$—the expected mean failure dose for an average lot of a given part type
2) $\sigma_i$—the expected part-to-part variability (standard deviation) in an average lot for a given part type
3) $\sigma_p$—the expected lot-to-lot standard deviation of expected mean failure dose for a given part type
4) $\sigma_o$—the expected lot-to-lot standard deviation of part-to-part variability for a given part type.

Then we fit each of these 4 parameters to appropriate distributions across part types, resulting in likelihoods for normal or lognormal parameters (one mean and one standard deviation) for each of the above variables (1-4):

1) $P(\mu_i, \sigma_i)$ describes the distribution of $\mu_i$ across part types as a lognormal distribution with lognormal mean $\mu_{i0}$ and lognormal standard deviation $\sigma_{i0}$.
2) $P(\mu_o, \sigma_o)$ describes the distribution of $\mu_o$ across part types as a lognormal distribution with lognormal mean $\mu_{o0}$ and lognormal standard deviation $\sigma_{o0}$.
3) $P(\sigma_p, \sigma_o)$ describes the distribution of $\sigma_p$ across part types as a lognormal distribution with lognormal mean $\mu_{p0}$ and lognormal standard deviation $\sigma_{p0}$.
4) $P(\sigma_o)$ describes the distribution of $\sigma_o$ across part types as a lognormal distribution with lognormal mean $\mu_{o0}$ and lognormal standard deviation $\sigma_{o0}$.

The parametric combination with the highest likelihood is the maximum likelihood distribution describing how the given variable is distributed across part types in the process under study. However, because our dataset is small, it is unlikely to produce a sharp maximum in the likelihood. As such, we use a likelihood-weighted model averaging approach similar to that in reference 11 and used later in 5, with weights given by:

$$w(\mu_{xy}, \sigma_{xy}) = \frac{1, \mu_{xy}, \sigma_{xy})}{\mu(x, \mu_{xy}, \sigma_{xy})}$$

(1)

where $z$ represents the data.

Thus, the probability distribution for expected hardness is

$$P(\mu) = \int P(\mu, \mu_{xy}, \sigma_{xy}) w(\mu_{xy}, \sigma_{xy}) d\mu_{xy} d\sigma_{xy}$$

(2)

This method takes the contribution to $P(\mu)$ for each parametric combination $(\mu_{xy}, \sigma_{xy})$ weighted according to the likelihood of that combination based on our data. In many cases such model-averaging approaches can outperform maximum likelihood, because we preserve the contribution of every parametric combination, merely changing their weights as we add data, rather than selecting different parameters with each new addition.[11] As such, the averaged probability distribution $P(\mu)$ tends to evolve more continuously. We calculate similar averaged distributions over the other parameters $\mu_o$, $\sigma_o$, and $\sigma_i$.

These distributions allow us to bound expected failure dose, part-to-part and lot-to-lot variability to any desired confidence level for parts in the process. While the limited dataset we have here likely results in a broad distribution, this merely means that the bounds we estimate are likely conservative. The bounds will improve as we add data for more parts to constrain the model. Moreover, we can use the distributions to spot parts that may perform out of family for the process.

![Fig. 4 (Top)](image)

Generate likelihood for data $z$ and use likelihood-weighted average to estimate probability distributions

$$w(\mu, \sigma) = \frac{1, \mu, \sigma) }{1, \mu, \sigma) }$$

The resulting distributions describe part-to-part, lot-to-lot variability over the range of similar part types in the process.

Applying this method to the data in Table III, the OP400 appears to exhibit greater variability than other parts—especially in its part-to-part standard deviation—from one lot to the next. As we did for the RH27, we initially perform the procedure outlined above excluding the OP400 data. Figures 5-8 summarize the various contributions to variability within the process (omitting the OP400). These curves show that while the OP400 is somewhat soft and variable in its lot mean hardness for the process (86% WC in expected mean hardness and 91% WC in lot-to-lot variation of mean hardness), lot-to-lot changes in part-to-part standard deviation are at the 95% WC level for the process. This indicates that the part could be out of family for the process. Certainly, inclusion of the OP400 results in much broader distributions except for that of mean hardness.
Fig. 5 The expected mean failure level for arbitrary part types in the ADI bipolar (>2.5 μm feature size) process.

Fig. 6 The expected lot-to-lot variation of mean failure dose for arbitrary part types in ADI’s bipolar (>2.5 μm feature size) process.

Fig. 7 The distribution of expected part-to-part standard deviations of failure doses for arbitrary part types in the ADI bipolar (>2.5 μm) process.

Fig. 8 The distribution of variation of part-to-part standard deviations of from lot to lot for arbitrary part types in the ADI bipolar (>2.5 μm) process.

Another part we would like to investigate is the OP484 quad op amp, which has been seen to exhibit significant lot-to-lot and part-to-part variability.[4],[5] For example in one sample of 9 lots, the hardest lot had a mean failure level about 9× harder than the softest lot. This would place this part roughly at 97% worst-case for the process based on figure 6, and it places at about the 90th percentile for variability in standard deviation (based on figure 8).

Although we cannot say definitively based on our small sample size that the OP400 and OP484 are out of family for the ADI bipolar (>2.5 μm) family, these parts exhibit much more lot-to-lot variability that other parts we have investigated in the family. It is important to be able to identify such out-of-family parts produced in a process, because they preclude qualification of parts in the process based solely on similarity data—unless we understand why the parts are out of family. That we could spot the OP400 and OP484 even with the limited dataset available illustrates the model’s robustness.

IV. INFERENCE WITHOUT MULTI-LOT DATA

The lack of multi-lot data for the LTC parts in Table II precludes disentangling lot-to-lot and part-type-to-part-type contributions to variability for the process. However, we can look at performance of a "typical" lot for a "typical" part type in the RH process. Figure 9 shows the distribution of lot mean ∆Ib after 50 krad(Si) of low-dose-rate irradiation. The curves were generated using a fitting and averaging procedure similar to that used in section IV. Based on these curves, the RH27 does indeed appear out of family, falling at the 81st worst case (WC) percentile for mean increased bias current (∆Ib) and at the 99.8th WC percentile for part-to-part standard deviation σ_{ptp}. Even so, the degradation and variability are mild compared to the commercial processes like the ADI process investigated in section IV.
We have developed a statistical method that allows use of similarity data to bound likely radiation behavior of other parts in the process for which we do not yet have data or to supplement limited data for a part fabricated in the same process. The method is sufficiently robust that it can provide meaningful constraints on radiation behavior even with small, imperfect data sets drawn from public sources. Although the methods developed do not necessarily require multi-lot samples of data for a given part, we can begin to disentangle the contributions to variability for the process if we have at least three part types with data for three lots. The multi-lot analysis provides estimates for expected mean performance and part-to-part variability, as well as how these quantities vary from lot to lot and part type to part type.

Even if we lack data for multiple lots, we can still draw useful conclusions about how an “average lot” will perform—especially for radiation hardened part families like the LTC RH series of parts. Indeed, if the process is sufficiently hard, even suspension data like that for the ADI XFCB process in Table IV can place useful constraints on possible failure distributions.

Although similarity data has traditionally been used mainly for qualification decisions, the methods developed here have several other uses. Perhaps most important, they indicate when similarity data are inappropriate for qualification. Thus, the “out of family” performance of the OP484 and OP400 relative to other parts in the ADI bipolar (>2.5 µm) process is problematic because it is not understood, while the RH27’s greater degradation relative to other parts in the RH process can be understood based on its reliance exclusively on process hardening. Also, because the method of necessity looks at all available data for parts in a process, it is useful for spotting process changes that may affect other parts in the process as well. Because the method provides separate estimates of part-to-part, lot-to-lot and part-type variation, it will likely also provide useful input for physics-based modeling for TID and for hardening efforts.

While the above plots do not separate the different contributions to variability, we noted above that the lot-to-lot consistency of the RH1014 quad op amp suggests that lot-to-lot variation should be mild for the RH process. As an example, mean ∆Ib at 100 krad(Si) (standard dose rates) varied by only a factor of 2.6 over 38 lots.[4],[5] We assumed that ∆Ib was lognormally distributed and found that this lot-to-lot variation equates to about half as much variability as we saw for the parts in Table II. If the RH1014 is typical for the process, this suggests that lot-to-lot variability and variation of response over part types are roughly commensurate in the LTC RH process. While this represents only a first guess that could change as we add data for more part types to refine our model, it demonstrates that even limited data can provide a useful working hypothesis to guide part selection while LTC continues to add to its list of parts tested at ELDRS rates.

V. CONCLUSIONS AND RECOMMENDATIONS

We have developed a statistical method that allows use of similarity data to bound likely radiation behavior of other parts in the process for which we do not yet have data or to supplement limited data for a part fabricated in the same process. The method is sufficiently robust that it can provide meaningful constraints on radiation behavior even with small, imperfect data sets drawn from public sources. Although the methods developed do not necessarily require multi-lot samples of data for a given part, we can begin to disentangle the contributions to variability for the process if we have at least three part types with data for three lots. The multi-lot analysis provides estimates for expected mean performance and part-to-part variability, as well as how these quantities vary from lot to lot and part type to part type.

Even if we lack data for multiple lots, we can still draw useful conclusions about how an “average lot” will perform—especially for radiation hardened part families like the LTC RH series of parts. Indeed, if the process is sufficiently hard, even suspension data like that for the ADI XFCB process in Table IV can place useful constraints on possible failure distributions.

Although similarity data has traditionally been used mainly for qualification decisions, the methods developed here have several other uses. Perhaps most important, they indicate when similarity data are inappropriate for qualification. Thus, the “out of family” performance of the OP484 and OP400 relative to other parts in the ADI bipolar (>2.5 µm) process is problematic because it is not understood, while the RH27’s greater degradation relative to other parts in the RH process can be understood based on its reliance exclusively on process hardening. Also, because the method of necessity looks at all available data for parts in a process, it is useful for spotting process changes that may affect other parts in the process as well. Because the method provides separate estimates of part-to-part, lot-to-lot and part-type variation, it will likely also provide useful input for physics-based modeling for TID and for process and circuit hardening efforts.

Although our goal in this study has been to develop a method that is sufficiently robust to yield useful results even with small, imperfect datasets, the model’s utility will improve significantly as we increase the dataset size or quality. In particular, if data are drawn from a long time series of lot qualification efforts with consistent test procedures and conditions, much greater precision is possible. Moreover, while here we have concentrated on a single parameter (Ibias), one can apply it across the board to all parameters or to different definitions of failure (e.g. functional).

The robustness of the method should also allow it to be used with data from different sources, provided that test conditions are known to be consistent. As a result, it increases the value of data sharing and of implementing consistent test methods that facilitate data interpretation as well as data sharing.

REFERENCES


