Data, What Is It Good For?

Ray Ladbury
NASA GSFC Radiation Effects and Analysis Group

Abbreviations and Symbols

- ADI-Analog Devices, Inc.
- ADC—Analog-to-Digital Converter
- CAD—Computer-Aided Design
- CMOS—Complementary Metal-Oxide-Semiconductor
- COTS—Commercial Off The Shelf
- DAC—Digital-to-Analog Converter
- EDAC—Error Detection And Correction
- Figure of Merit (FOM) parameters
  - C_{ENV}—proportionality constant to turn Figure of merit into rate estimate—specific to radiation environment
  - σ_{sat}—saturated cross section
  - $LET_{0.25}$—LET where cross section reaches 25% of its saturated value=$LET_0+w*(0.288)^{1/s}$
- GEO—Geosynchronous Equatorial Orbit
- LDC—Lot Date Code
- LET—Linear Energy Transfer
- LTC-Linear Technologies Corp., now part of ADI
- MBU—Multi-Bit Upset
- MRED—Monte Carlo Radiative Energy Deposition
- Nuclear and Space Radiation Effects Conference
- PO/R—Power-On/Reset
- REDW—Radiation Effects Data Workshop
- SEE—Single-Event Effect
- SEECA—SEE Criticality Analysis
- SEFI—Single-Event Functional Interrupt
- SEL—Single-Event Latchup
- SV—Sensitive Volume
- TNS—Transactions on Nuclear Science
- WC—Worst Case
- Weibull Fit parameters
  - LET_0—onset LET, σ_{sat}—saturated cross section, w—Weibull width, s—Weibull shape

Databases Are In Fashion

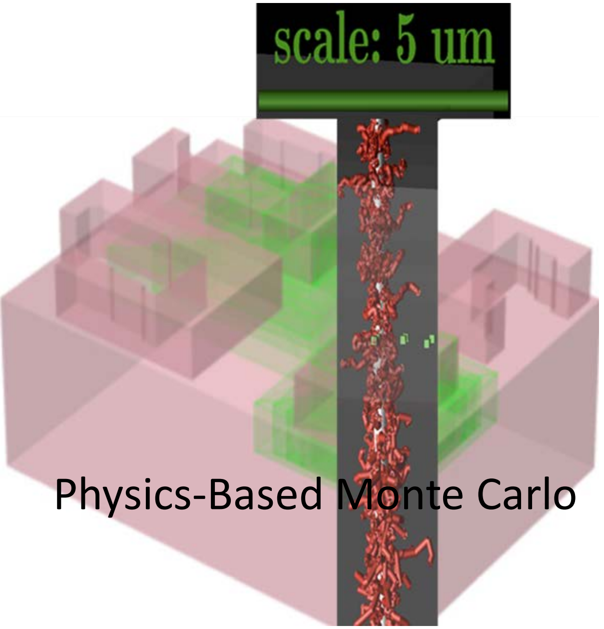
- Radiation databases are increasingly acknowledged as important to radiation hardness assurance
  - Importance discussed in recent National Academies Report,* along with caveats and limitations
  - Emphasized in recent NASA workshops and reports
  - Increasing numbers of new players in Commercial and Smallsat community increases importance
- Unfortunately, access is becoming more difficult and data more difficult to use
  - Fewer radiation databases accessible to general public
  - More data is application specific or pertains to older or obsolete parts
- Value of radiation databases is increasing as new techniques for using data are developed
  - New techniques combine data from multiple tests and part types to assess risk
  - Different types of data may be important for combining datasets
    - How are parts similar or different
    - What are limitations of the data that may preclude its inclusion in a larger analysis
- How good a database is depends on the analysis being done
  - Maximizing database utility requires understanding the types of analyses that may use the data
    - This is changing rapidly

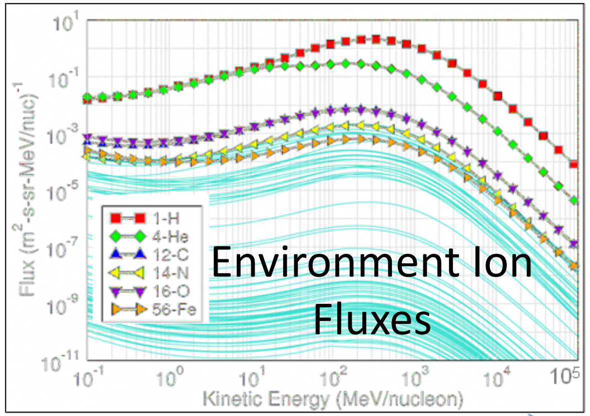
*https://www.nap.edu/catalog/24993/testing-at-the-speed-of-light-the-state-of-us

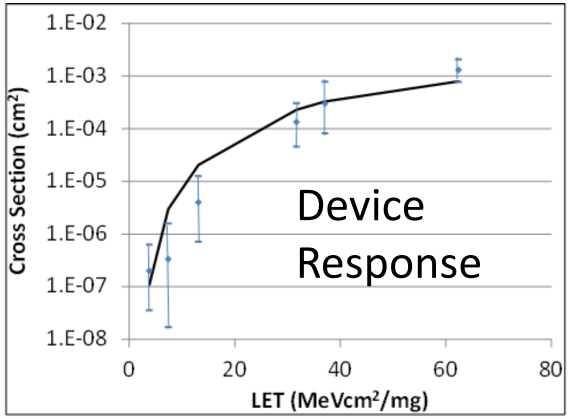
Why We Need Data: The Basics for SEE

- Quality of a database depends on what it is to be used for
  - Limiting risk of using a particular part– Risk=Probability(failure) $\times$ Consequences(Failure) $\rightarrow$ 2 types of data
    - Modes to which part is susceptible and their consequences
    - Data to estimate/limit rate for each mode
    - SEE testing is destructive—Limiting risk requires parameterizing risk model with data from test and/or database
  - Other analyses likely require combining/comparing data for multiple tests/parts
    - Historical analyses combine data from multiple tests of same part to assess lot-to-lot, part-to-part and application dependence
    - Similarity analyses combine data for multiple part types to answer question: What is risk of using a part like these parts?
    - Trend analyses combine data for parts that differ in one or more critical criteria to assess impact of such criteria
      - Principle Component Analysis (PCA) will even tell you which linear combinations of criteria are most interesting
    - System Analyses combine results from above analyses and assess how they impact a particular system or range of systems
    - Others—limited only by imagination, available data and available statistical techniques
- Risk-Limiting analyses may treat each part individually, because what matters is:
  - Understanding consequences of SEE modes well enough to assess system/mission impact
  - Understanding rate to assess how often one has to deal with each mode
- Other analyses require sufficient information on parts to assess similarities and differences between them
  - Technology, vendor, fabrication facility, process, feature size, traceability, part function, SEE modes, application conditions

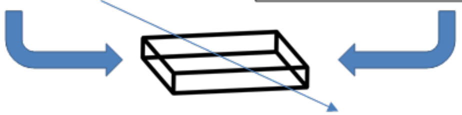
All Models Are Wrong; Some Models Are Useful







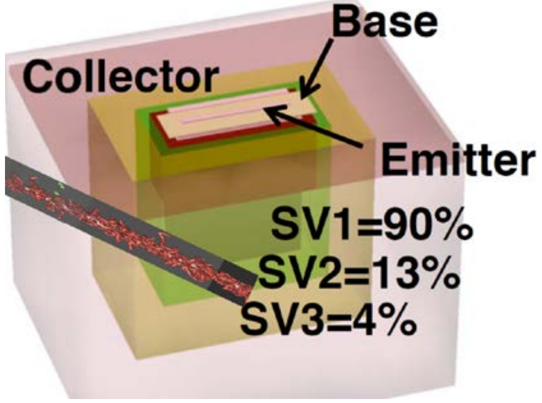
- Rectangular Parallelepiped
- Assumes constant ion LET
- 100% of charge deposited along ion paths (chords) through SV collected
- Device response fit w/ **4-parameter** Weibull (LET_0, σ_{sat}, Weibull w and s)



Rate Prediction Model

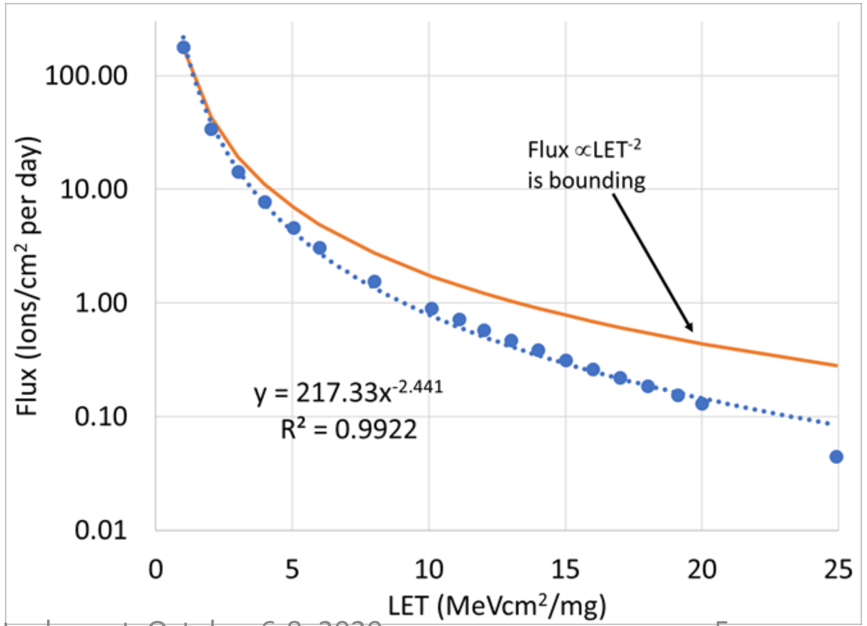
Single-Event Rate

- Monte Carlo models can be arbitrarily accurate AND arbitrarily complicated
- Complexity found in both SV model and physics of ion transport, charge collection, etc.
- Device model from CAD model (MRED) or derived from test data (CRÈME-MC Nested Volume) or both



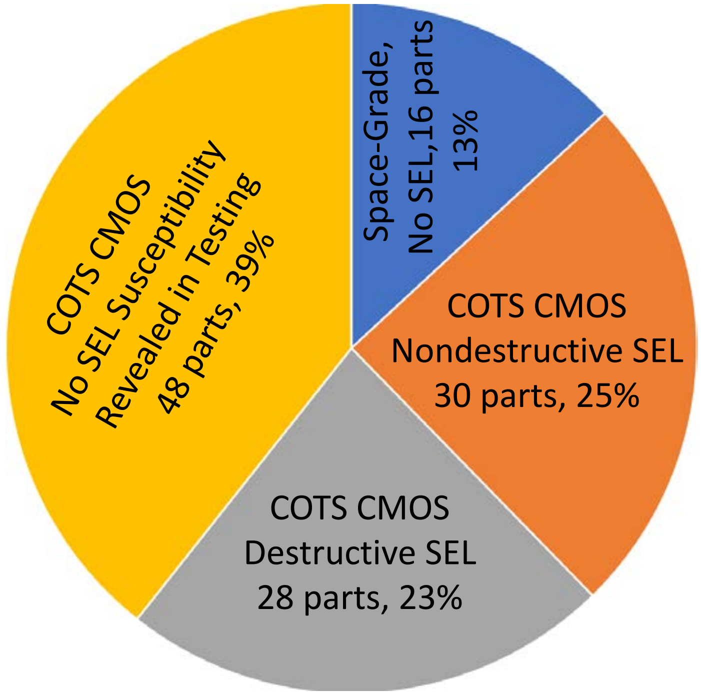
Nested-Volume Monte Carlo

- Figure of Merit gives approximate, analytical expression for SEE rate
- $FOM=C_{ENV}* \sigma_{sat}/(LET_{0.25})^2$
- C_{ENV} depends environment
- Algebraic form useful
- **2-parameters**



How Much Information Does A Database Need?

Adapted from G. Allen et al. 2017 REDW, pp. 1-14



- Even a single-chart summary of testing on a large # of parts is very useful
    - 58 of 106 (~55%) of commercial CMOS parts SEL susceptible
    - ~26% susceptible to destructive SEL
    - Provides proper caveat regarding risks of using COTS CMOS parts and likely magnitude of test campaign to find SEL immune parts
    - Tells us nothing about how any one part will perform
- Most databases provide more information
    - REDW summary table format has won Best Data Workshop papers for JPL, GSFC
    - Summary of SEE modes and onset LET, along with perhaps limiting cross section
    - Sufficient to say whether mode common or rare, not enough to estimate rate
    - Device information allows comparison w/ similar parts, vendors, technologies…

Adapted from M. Obryan et al. REDW Data Summary

Part Number	Manufacturer	LDC or Wafer#, REAG ID#	Device Function	Technology	Particle: (Facility/Year/ Month) P.I.	Test Results: LET in MeV•cm^2/mg, σ in cm^2/device, unless otherwise specified	Supply Voltage	Sample Size (Number Tested)
Identifier	Vendor	Traceability	Device Type	Defines SEE modes	Type Test By Whom	Short Summary of results; Modes, LET_0, σ_{sat}.	Applic. Info	Devices Tested
MT29F128G0 8CBECBH6	Micron	201448, 14-088	Flash Memory	16 nm CMOS	H: (LBNL2015 Aug; 15Dec) DC	SEU LET_{th} < 0.9 MeV·cm2/mg, SEU σ = 1.7x10^{-10} at LET of 58; SEFI: Part is vulnerable to SEFI in static biased and dynamic test modes. SEFI LET_{th} < 0.9; No device functional failure up to LET of 118. However observed block erase failure at LET of 58.	3.3V	2

Limiting Risk=**Consequences** × Probability of Event I

Consequences

Increasing Information

Range of consequences + actions needed for recovery (e.g. SEFI requires PO/R or Soft reset OK; MBU vs. EDAC; fraction of transients causing effect) → SEECA based on SEE consequences and recovery—emphasizes availability, reliability and survivability

Range of consequences → (e.g. proportion of destructive events, distribution of transient widths & amplitudes, #MBU, range of SEFI effects) → SEECA based on both likely and worst-case effects—emphasizes reliability and survivability

WC Consequences → (e.g. destructive or not, WC transient width/amplitute, WC MBUs, etc.) → SEE Criticality Analysis (SEECA) based on worst-case effects—emphasizes survivability

Summary of Part's SEE Modes—consequences may be inferred from past experience → Mainly useful for parts selection, although preliminary SEECA may be done

Limiting Risk=Consequences × Probability of Event II

SEE Probability/Rate

Increasing Information

σ vs. LET +error bars and/or event counts + test report/summary, full dataset	Analysis below + can assess degree to which results representative for application, consistent w/ models
σ vs. LET +error bars and/or event counts	Can estimate RPP rate for any desired confidence level
σ vs. LET curve	In addition to RPP Rate, can assess goodness of fit (GOF) to Weibull and RPP models
σ vs. LET Weibull fit parameters LET_0, σ_{sat}, Weibull w and s	RPP rate in any environment, comparisons and correlations of rate, and Weibull fit parameters
$LET_{0.25}$ and σ_{sat}	FOM rate for any environment + comparison to similar parts + some idea of how σ increases w/ LET
SEE rate for a given environment, e.g. GEO	Compare rates for similar parts estimate rate in other environment using ratios of FOM constants
LET_0 and/or σ_{sat}?	Is mode Frequent or Rare?

What Does More Data Buy You?

- Increasing information—whether regarding SEE consequences or SEE rates increases fidelity of SEE analyses as well as types of analyses we can do.
  - Increased data for SEE consequences and recovery from SEE modes allows a better fidelity for SEECA
- Improved device response data (SEE modes and their σ vs. LET) likely produces better rate estimate/bound.
  - σ vs. LET fit parameters (LET_0, σ_{sat}, w, s) or FOM parameters ($LET_{0.25}$ and σ_{sat}) allow rate estimation for any environment and studies of correlations between parameters vs. rate
    - Parameters fitting σ vs. LET are not always independent
    - If error bars/event counts available, can examine correlations between likelihood + fit parameters
      - Useful for understanding weaknesses/biases in the data
- Largest increment in storage results from archiving test reports/summaries and datasets for each part
  - What does it buy?
    - May reveal dependence of SEE rate/consequences on test/application conditions
    - Full dataset can also reveal departures from RPP, LET_{EFF} dependence, Weibull form for σ vs. LET
    - Dependence of σ on angle could suggest more sophisticated sensitive volume models for Monte Carlo SEE rate estimation
- What about Similarity analyses and other data that go beyond data for a single part?

Analyses and Data: Similarity

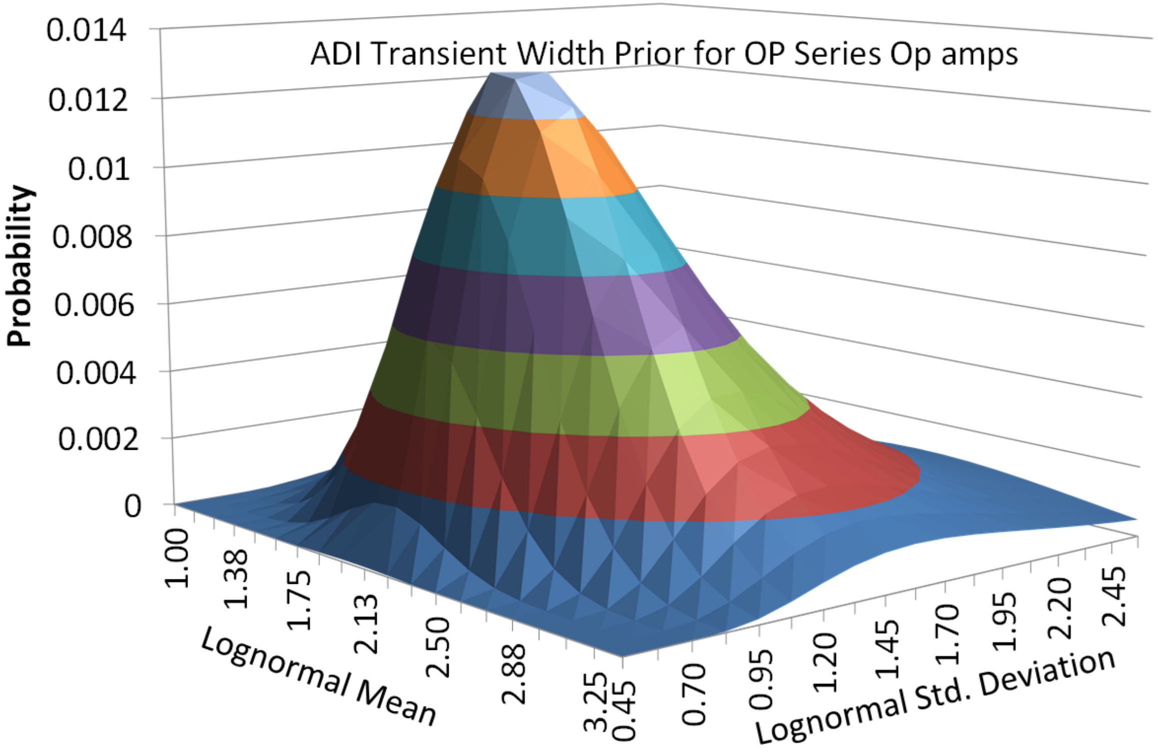
- Similarity is inherently a subjective term
  - Definitions vary not just in terms of degree of similarity acceptable, but also in terms of criteria
- For SEE, we mean “sufficiently similar” in terms of criteria that we expect similar SEE susceptibility
  - Nondestructive SEE—may be sufficient for parts to be fabricated at same foundry w/ same process node
    - May also need to require similar cell type, part type and/or function
  - Destructive SEE—defining similarity may be more challenging
    - May depend on parasitic elements of circuitry (e.g. thyristor in CMOS) or other processes not occurring in normal operations
    - Foundry and process may be poor predictors of performance
- If such analyses are to use a database, data in database must be sufficient to determine “similarity”:
  - Vendor
  - Semiconductor material
  - Technology—at a minimum technology type (e.g. bipolar, CMOS, FLASH…,)
  - Foundry and process node
  - Device type/function,
- Ideally, data would be sufficient in quantity to determine which of these factors was important
  - Principal Component Analysis finds appropriate “eigenvectors” and ranks them in order of influence
  - Unfortunately, radiation analyses are almost always starved for data

When Similarity Works—Even With Limited Data

Table I: SET Data for ADI and LTC Op Amps

	Vendor/ Process	SET Rate (day^{-1})	WC duration	WC amplitude	Supply—Output
RH108[2]	LTC—RH Bipolar	0.011	7 µs	4.5 V	0/12V—6V
RH118[2]	LTC--RH Bipolar	0.046	>5 µs	2 V	-8V/+8V—0V
RH1014[2]	LTC--RH Bipolar	0.029	>35 µs	rail-to-rail	-12V/+12V—6V
RH1499[3]	LTC--RH Bipolar	0.0082	13 µs	4.4 V	-15V/+15V—4V
RH1078[4]	LTC--RH Bipolar	0.04	30 µs	2 V	0/10V—2V
OP27[5]	ADI—Bipolar>2.5 µm	0.036	10 µs	10V	-10V/+10V—varied
OP113[2]	ADI—Bipolar>2.5 µm	0.01	>2 µs	1.7 V	0/5V—2.55V
OP270[2]	ADI—Bipolar>2.5 µm	0.35	1.2 µs	3 V	-5V/+5V—0.42V
OP400[2]	ADI—Bipolar>2.5 µm	0.56	10 µs	3 V	-12V/+12V—±2.4V
OP05[5]	ADI—Bipolar>2.5 µm	--	12 µs	--	--
OP15[5]	ADI—Bipolar>2.5 µm	--	15 µs	--	--

Vendor/ Process	Typical Width/ Typical Part	Typical Width/ 90% WC Part	90% WC Width/ 90% WC Part	90% WC Rate
ADI	15 µs	<35 µs	<218 µs	<0.19/day
LTC	22 µs	<35 µs	<170 µs	<0.13/day



- Bayesian treatment of data yields useful bounds
  - Rate~1-1.4 transients/week@90% confidence
  - Transient widths are relevant
  - Transient amplitude likely to be rail-to-rail for most parts
  - And this is with limited datasets

- Data used for Bayesian analysis over parameters assuming transient widths distributed lognormally over part family
- Can select parameters for best fit or for bounding distributions, depending on analysis requirements

Just Because You Have Data, Doesn't Mean It Makes Sense

SEL for ADI 0.6 μm CMOSADCs/DACs

BNL Data

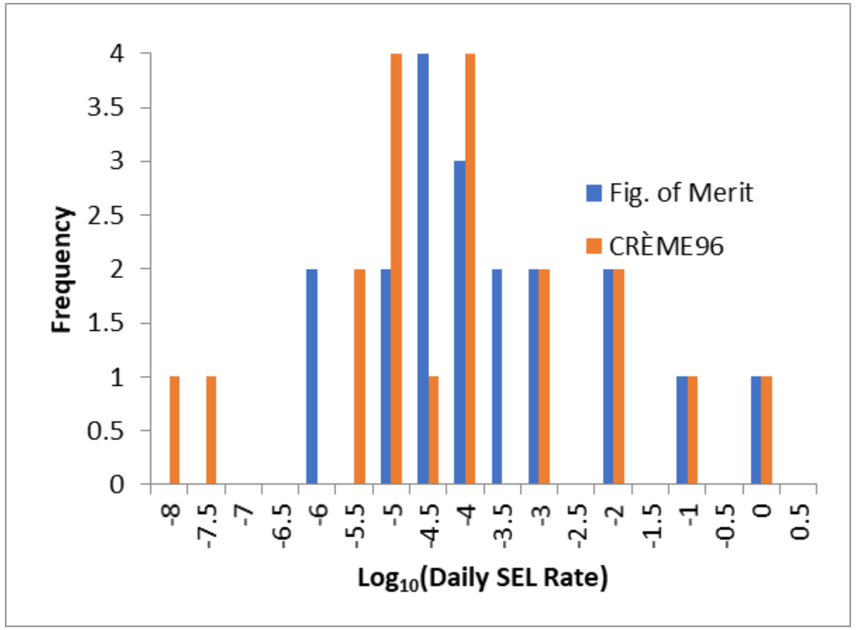
Part #	Onset LET (MeVcm2/mg)	σ_{lim} (cm^2)
AD5305[6]	LET<11	??
AD7472[6]	11<LET<15	5.00E-04
AD7476[6]	LET<60	??
AD7664[6]	6	2.50E-04
AD7714[6]	15<LET<24	6.00E-04
AD9260[6]	8	??

TAMU Data

Part #	Onset LET (MeVcm2/mg)	σ_{lim} (cm^2)
AD5334[6]	5	5.00E-04
AD7664[6]	4	1.00E-03
AD7675[6]	4	3.00E-05
AD7714[6]	20<LET<22	1.00E-04

Adapted from Ladbury and Campola, TNS Vol. 60, pp. 4464-4469

- Little correlation in SEL data even for similar parts
- Onset LET appears bimodal
- σ_{sat} spans nearly 3 orders of magnitude



Representative Recent COTS CMOS

Part # [REF]	Mfg.	Function	σ_{sat}	LET$_0$	Shape	Width
LTC1409 [10]	LTC/ADI	14-Bit ADC	0.000007	55.0	3.1	36
LTC1604 [10]	LTC/ADI	16-Bit ADC	0.000009	53.0	1.5	32
LTC3708 [13]	LTC/ADI	Buck Controller	0.000006	8.1	3.2	18.5
MAX2981 [9]	Maxim	Line Driver	0.000011	12.5	3.4	21
MAIA ROIC [14]	TSMC/ASIC	Detector Read-out IC	0.000065	20.5	1.75	53.3
AD7714 [9]	ADI	24-bit ADC	0.00014	26.5	1.6	90
TPS51200RDC [13]	TI	Regulator	0.000036	10.7	2	31.3
LTC2271 [12]	LTC/ADI	16-Bit ADC	0.000020	12.5	5.2	10
AD7609 [9]	ADI	Data Acq. Sys.	0.00002	3.5	1.5	21
LTC2502 [9]	LTC/ADI	Op. Amp.	0.00013	7.0	1.5	53
ADS8320 [13]	TI	16-Bit ADC	0.000068	17.0	4.8	6
93LC96BT [9]	Microchip	EEPROM	0.000250	9.5	2.1	31
LTC1595 [9]	LTC/ADI	16-Bit ADC	0.00012	8.2	3.3	14.6
SIG60 [9]	YAMAR	XCVR	0.00112	11.7	2.5	25.5
AD5328 [11]	ADI	12-bit DAC	0.00044	0.8	3.8	22
TMS320F2906 [9]	TI	DSP	0.019	17.5	4.6	26
MAX2992 [9]	Maxim	XCVR	0.0070	6.8	3.3	18
dsPIC30F6014A [11]	Microchip	Micro-controller	0.0616	3.1	2.4	23
ADV212 [12]	ADI	Vid. Codec	0.065	0.2	2	16

- Since, similarity yields no predictive value, expand the selection to be "representative" of recent COTS
- Select parts with good Weibull fits
- Try to have a few parts per vendor, a range of part types, a range of process nodes…
- Not easy to find good Weibull fit for SEL, as many parts tested only over limited LET range
- Data extracted from my NSREC talk

- Above histogram is for parts in Representative COTS CMOS table
- SEL rates for recent COTS CMOS span 5 orders of magnitude
- Rate appears multi-modal and does not correlate with vendor, process node…
- No strong correlations between any Weibull fit parameters

Just Because Data Is Nonsensical Doesn't Mean It's Useless

$LET_0 \rightarrow \{L_1 \ldots L_n\}$

$\sigma_{sat} \rightarrow \{\sigma_1 \ldots \sigma_n\}$

$Width \rightarrow \{w_1 \ldots w_n\}$

$Shape \rightarrow \{s_1 \ldots s_n\}$

Bootstrap Parameters

(L_1,σ_1,w_1,s_1)

(L_1,σ_1,w_1,s_2)

⋮

(L_1,σ_1,w_2,s_1)

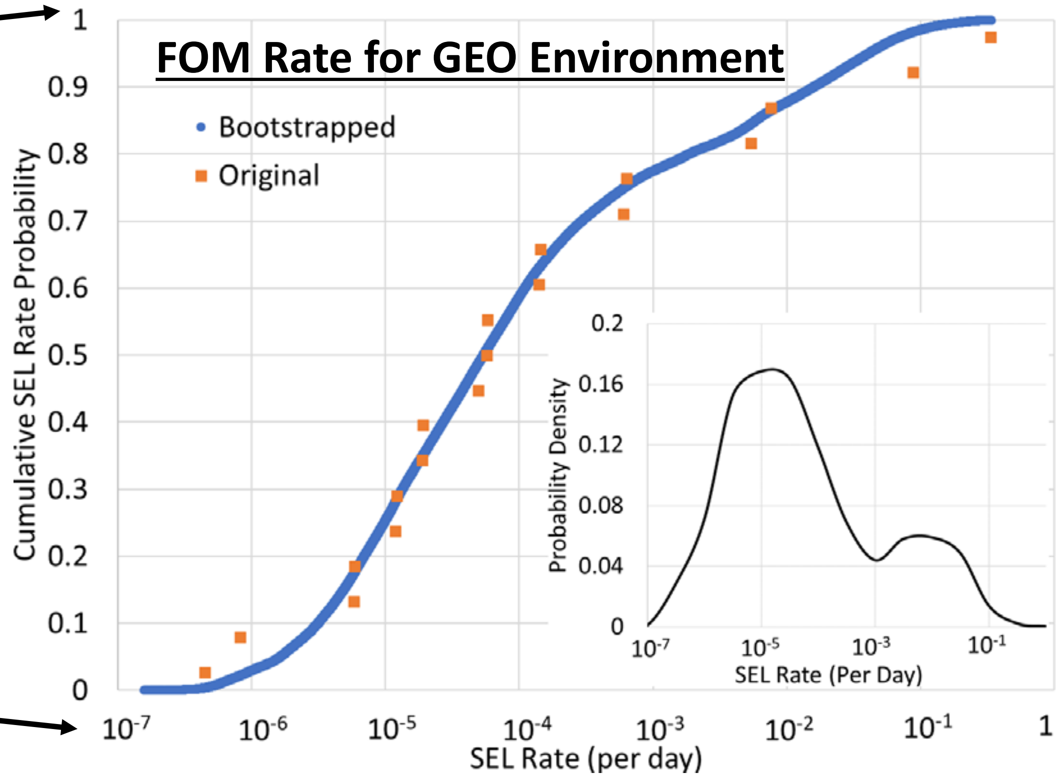
⋮

(L_1,σ_2,w_1,s_1)

⋮

(L_n,σ_n,w_n,s_n)

FOM Rate



- SEL data is sufficiently ill behaved that use of parametric statistics produces no useful *a priori* constraints on SEL rates
- Nonparametric statistics can still be useful
- No correlations: all parametric combination~equally likely
- 19 part types→19^4 combinations=130321, nearly continuous

- Quantile plot useful for summarizing SEL risk
  - Reveals clearly bimodality suggested in histogram
  - Can bound rate for any desired confidence level
  - Risk arises as probability of "bad" part (upper mode) rather than gradual increase from lower mode
- See my NSREC talk for use in a system analysis

Bayesian Analysis: Bringing Type I and Type II Analyses Together

- Bayesian analysis combines two sources of information
  - Prior Probability based on info sources other than data specific to the parts of interest: $P(Hypothesis)$
    - May be based on similarity, trend, expert opinion…
    - Hypothesis is often that the data are distributed according to a certain distribution, e.g. normal with mean μ and standard deviation σ; it may be phrased simply as "Probability of μ and σ"
  - Data—specific to the parts of interest—either test data or from database and is represented in terms of the Likelihood: $P(Data|Hypothesis)$
  - Prior and Data are combined using Bayes Theorem: $P(Hypothesis|Data) = \frac{P(Data|Hypothesis) \times P(Hypothesis)}{P(Data)}$
  - $P(Data)=$probability of observing data regardless of which hypothesis is true, e.g. integrated over all values of μ and σ
- Why use the Prior?
  - Prior can supplement data, allowing conclusions to be stated with higher confidence
  - Or, Prior can alert the analyst that part for which data taken appears to be "out of family" with "similar parts" or contrary to the trend or engineering judgment—and so more data required to reach valid conclusion with high confidence.
  - Questions:

1) Is more information better?
2) Is the analysis done correctly under correct assumptions?
3) If not, do you really understand these parts well enough to draw conclusions?

Conclusions

- Data in databases can be used for multiple analysis types
  - Type 1: What is the risk of using this part in this application?
    - Need data on Consequences of SEE modes—preferably including range of consequences and recovery strategies
    - Need data to constrain rate—using models ranging from FOM (2 parameters) to RPP (4-7 parameters) to physics based Monte Carlo (arbitrarily complicated)
  - Type 2: Risk of using parts like these in applications like mine?
    - These analyses require data on the parts being tested—how are they similar to or different from other parts
    - Data needed may include: Technology, vendor, fabrication facility, process, feature size, traceability, part function, SEE modes, application conditions
- Type 2 analysis techniques progressing rapidly
  - Can extract additional value from test data
  - Types of analyses include
    - Historical Analyses (usually less important for SEE) can verify what type of variability is important
    - Similarity Analyses can provide evidence for risk arising from use of classes of parts
    - Trend Analyses—how bad (or good) can it get
    - System Analyses (see my NSREC talk)
- Bayesian analyses combine information from Type 1 and Type 2 analyses
- Best databases are those that add value to the data—only limitations are the data and creativity