Interpreting Archival Cross-Section vs. LET Fit Parameters Based On Data Quality

Ray Ladbury

Radiation Effects and Analysis Group

NASA Goddard Space Flight Center

To be presented by Ray Ladbury at the Single Event Effects (SEE) Symposium/Military and Aerospace Programmable Logic Devices (MAPLD) Workshop, La Jolla, CA, May 18 2022.

Abbreviations

GLM—Generalized Linear Model

LET—Linear Energy Transfer

LET_0—Onset LET

s—Weibull shape parameter

SEE—Single-Event Effect

SEL—Single-Event Latchup

σ_s—Saturated SEE cross section (= σ_{sat})

w—Weibull width parameter

Agenda

- Introduction: Radiation Databases—The Latest Old Idea
- Uses of Radiation Databases
  - Qualification of single part type based on historical radiation performance
  - Look for evidence of variable response in historical/similar data
  - "How bad can it be?" studies based on very large, broad datasets*
    - % failure, parametric studies**
- How do you compare fit parameters across parts/experiments?
  - Different data qualities, # points in cross section (σ) vs. LET, error bars…
    - Similar to modelling acceptance, other parameters in nuclear/particle physics
    - Monte Carlo approach
- How do such studies modify results?
- Conclusion: More steps toward bigger data

*Effectively, these are extreme-value studies—the more data, the more likely to be valid; and the next entry could invalidate the analysis

**Comparative SEE rates; SEE σ vs. LET fit parameters …

Radiation Databases Are in Fashion

- Radiation databases are receiving new emphasis
  - Receives repeated mention in National Academies study: "Testing at the Speed of Light: The State of U.S. Electronic Parts Space Radiation Testing Infrastructure"
  - Increased efforts to update material in and increase access to NASA and ESA databases
  - New databases—e.g. PMPedia.space and other efforts to consolidate data
- More data is a good thing
  - Reduces duplication of efforts—or if efforts are duplicated, can look for evidence of variability
  - Provides basis of knowledge for radiation response in different part types—guiding test development
  - Availability of large amounts of data in one place facilitates new and different types of analysis
- Unfortunately, current efforts have the same vulnerabilities as past databases
  - No dedicated source of funding to keep database current
    - Value of data may decrease over time
    - Data may require validation, correction or updating;
    - Data field definitions needed to facilitate analysts' queries (especially for complex, "big-data" studies)
  - Obtaining access to data remains challenging—e.g. proprietary datasets, etc.
  - Standards to ensure and evaluate data quality remain elusive
    - In part because they depend on the purpose for which data are to be used.

What Kind of Data Is in a Database? What Is It Good For?

More information in the database entries means more detailed analysis to assist design and reliability

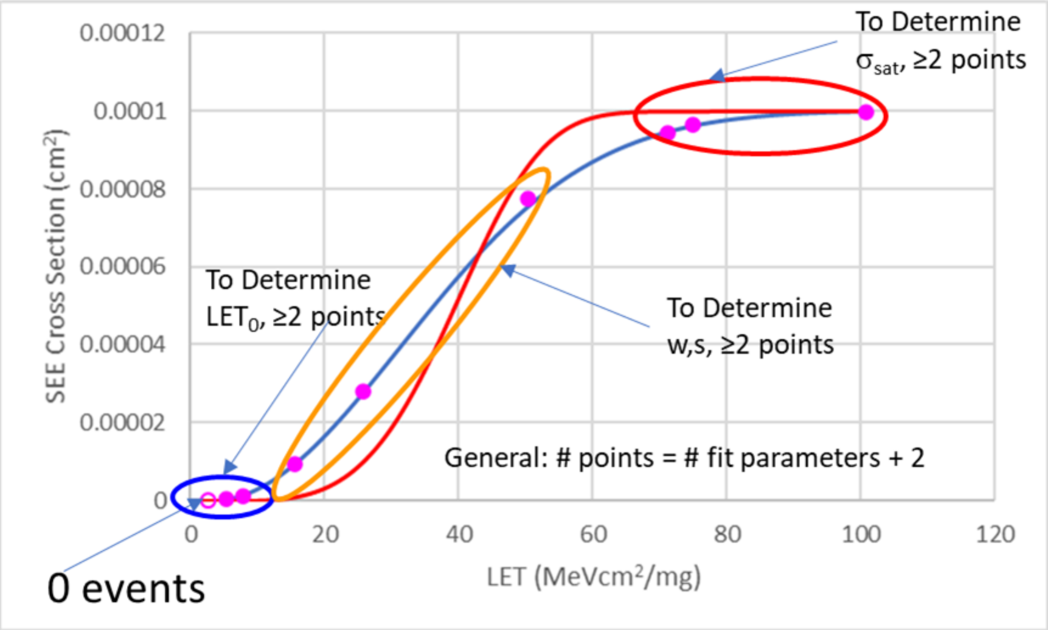
Go/No-Go Test Results
"LET_0>75 MeVcm2/mg"
Useful for risk avoidance, determining consequences

Partial Results
"LET_0=20 MeVcm^2/mg"
"σ_s=2×10^{-4} cm^2"
Useful for rough rate limit

Rate in Single Environment
Useful for that environment;
Rough rate in other orbits

Weibull Fit to σ vs. LET
Rate estimate for any environment

σ vs. LET + Event counts
Best-fit/bounding rates
Fit parameter estimates
Data quality



WC Consequences for Each Mode
Criticality, mitigation, etc.

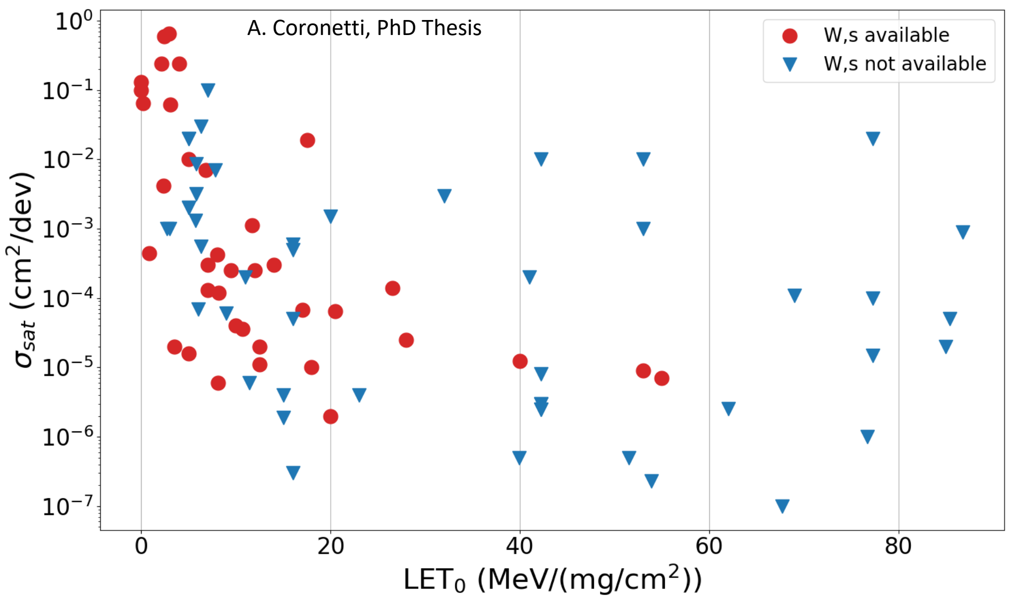
Range of Consequences vs. Rate
Detailed Risk Analysis

Consequences/Detection/Recovery
Risk Analysis
System-Level Availability
System-Level Recovery Design

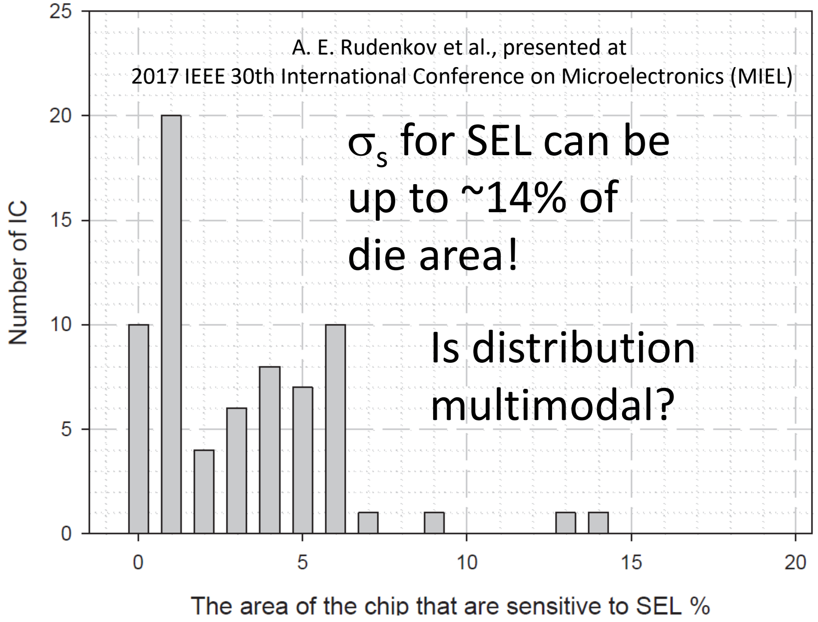
- Rate estimates derived from SEE databases likely "good enough" as long as
  - Part-to-part and lot-to-lot variation in SEE response is negligible
  - Test conditions are representative or bounding of application conditions
- Rate estimation/bounding is a very forgiving task
  - σ vs. LET with as few as 6 LET points and ~4 events per LET can come with 2-3x
  - E. Petersen: rate mainly depends on rapidly rising portion of σ vs. LET (10% to ~90% of σ_s), so it is possible to get Weibull fit parameters wrong and still get close to the right rate

But Large Databases Very Useful For Other Analyses

A large dataset over a broad range of Commercial CMOS parts can elucidate the *a priori* risks of such parts by showing the range of SEL behaviors within the population. Also, more specific data is not necessarily more predictive.



Is high SEL σ_s correlated with low onset LET?
Is correlation lost at higher onset LET?



- For some radiation threats (e.g. SEL), increased restriction of database does not increase its representativeness
  - Assemble as large a dataset of susceptible parts as possible to look for trends
    - Example: Low onset LET seems to correlate with high saturated cross section for susceptible parts…but only at low LET, but correlation is less convincing in the larger dataset, especially when Weibull width and shape not known
    - But how well do we know Onset LET (LET$_0$) and saturated cross section (σ_s)? Is it different if we know Weibull w and s or not?
    - How were these parameters determined? Least-Squares? Generalized Linear Model? By eye? And are they best-fit or bounding?
    - Do different data qualities (e.g. # LETs in σ vs. LET, # events per point) result in different errors for different LET$_0$/σ_s estimates?

Challenges of Fitting σ vs. LET Data to a Weibull Form

- Data characteristics
  - Data span several orders of magnitude
    - Determining LET0 requires fitting well at low LET
    - Determining σ_s requires fitting well at high LET
    - Rate is driven by behavior at intermediate LET(10%-90% of σ_s)
- Fitting strategies
  - SEE data inconsistent with Linear Regression
  - "Eyeball" fitting is still widely done, but
    - Method subjective (different analysts=different fit)
    - Fits may be done for different purposes (WC vs. best fit)
- Does it matter?
  - Rate estimates forgiving of small errors in fit parameters
    - True for both best-fit and bounding rate estimates
    - Rate determined largely by behavior of rapidly rising portion of σ vs. LET (e.g. from 10% to 90% of σ_s.
  - Parametric studies need accurate determination of fit parameters—or at least good error estimates.
- If parameters to be used in parametric studies, objective, self-consistent fitting highly desirable
  - Candidate: Generalized Linear Model SEE (GLM)

Generalized Linear Model for Fitting SEE σ vs. LET data

-Assumes errors on event counts Poisson and uses maximum likelihood to determine best fit to data.

Inputs to Generalized Linear Model (GLM)

n Test LETs={LET_i},
Fluences@each LETi={F_i}
Observed SEE counts @each LETi={O_i}

Generalized Linear Model (GLM)

Predicted SEE Counts @ each LETi={Pi}
$P_i = F_i * \sigma(\sigma_s, LET_i - LET_0, w, s)$

Likelihood, $\Lambda = \prod_{i=1}^{n} Poisson(O_i, mean = P_i)$
Maximize Λ (Λ_{MAX}) to define most likely LET_0, σ_s,w, s

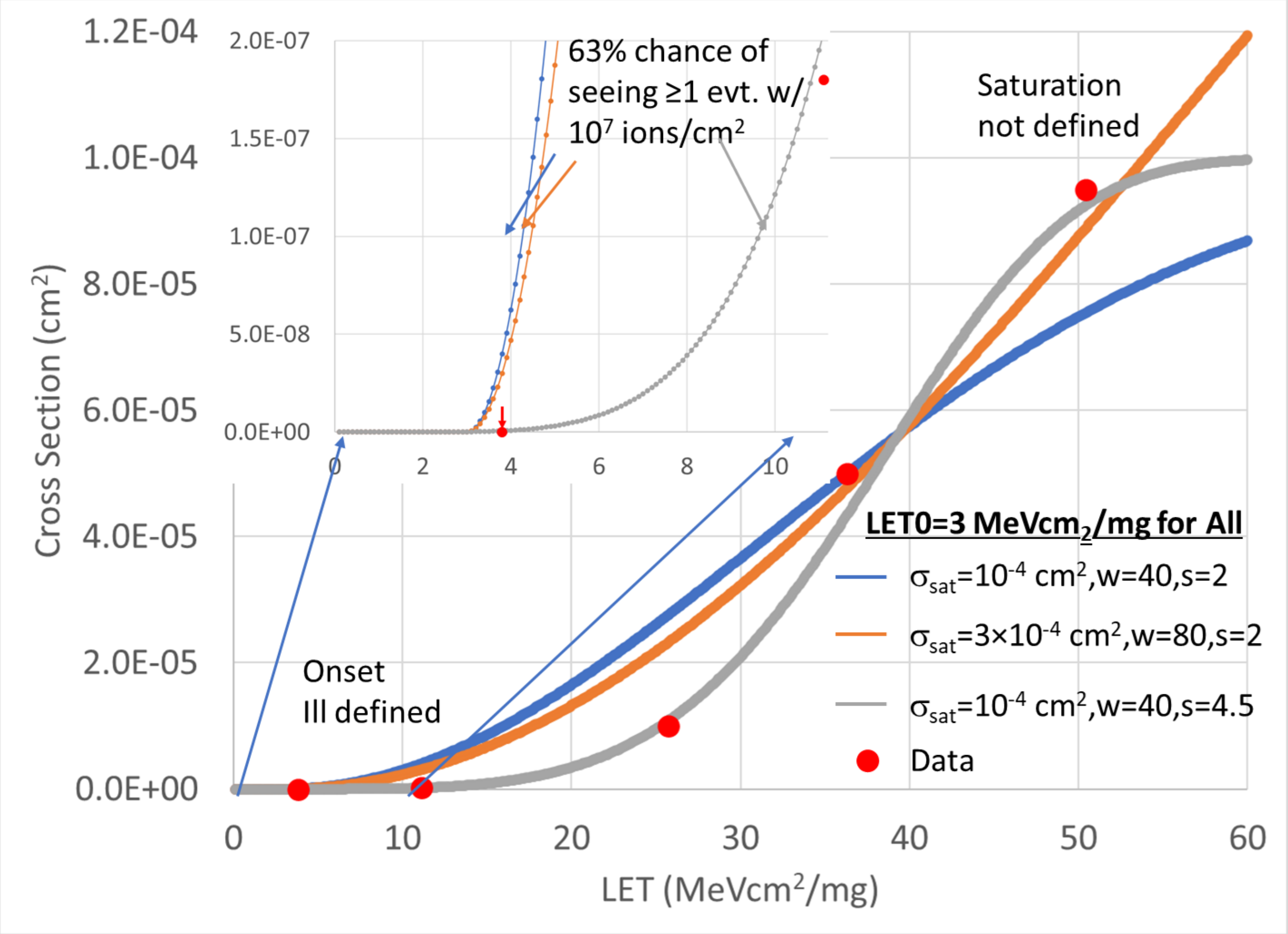
$\Lambda(CL)/\Lambda_{MAX} \sim \exp(-0.5*\chi 2(1-CL, \#\ \text{fit parameters})$,
Contours in parameter space for Confidence CL

Output of Generalized Linear Model (GLM)

Best Fit Rate and Estimates for LET_0, σ_s,w, s
WC Rate and LET_0, σ_s, w, s for any Confidence CL

Systematic Errors and Parametric Correlation

- Systematic data errors also affect fits
  - If σ not saturated at highest test LET, small event count fluctuations can affect fit
    - Can result in significant errors on rate
    - Even more significant errors on parametric values
  - Test LET too far apart near onset can result in significant errors in onset LET, s
- Weibull parameters tend to be correlated
- σ_s correlates with Weibull w (coefficient~0.66)
  - For any best-fit combination of (σ_s,w), there will be one that is nearly as good where both σ_s and w are larger.
  - Especially true if saturation not found
- LET_0 inversely correlated to shape s
  - Correlation especially evident if s>2.5, because σ near threshold <<10^{-7} cm^2, so unlikely to see events
- These correlations exist regardless of fitting method



Given these uncertainties, what do parameter measurements mean?

Interpreting fit parameter measurements using Monte Carlo

Monte Carlo (MC) Studies often used to determine uncertainties on nuclear/particle physics results

Assume fit to σ vs. LET returns
$LET_0^{obs} LET_0$=3 MeVcm^2/mg, w^{obs}=80,
s^{obs}=3.7 σ_s^{obs}=5E-4 cm^2
Assume event counts or error bars know for each LET

Assume "real" fit parameters are:
LET'$_0$, σ'$_s$, w' and s'

Want probability distribution of σ'$_s$ given value of w^{obs}=80

Must loop over w' and σ'$_s$ because they are correlated

Similar procedure can generate probability of actual LET0 given s^{obs}=3.7

Generate MC events for values of σ'$_s$, w' that can produce observed σ_s^{obs} and w^{obs}

Let event counts fluctuate Poisson-wise about mean = observed counts @ each LET

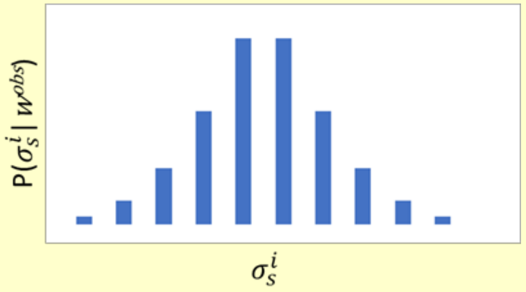
Let C_j^i be # of MC events that yield σ_s^{obs} and w^{obs} as best fit when generated with σ_s^i and w_j

w' \ σ'$_s$	w^1	•••	w^m
σ_s^1	C_1^1		C_m^1
⋮			
σ_s^n	C_1^n		C_m^n

Let $C^i=\sum_{j=1}^{m} C_j^i$

Let $C=\sum_{i=1}^{m} C^i$, total evts w/ σ_s^{obs} and w^{obs}

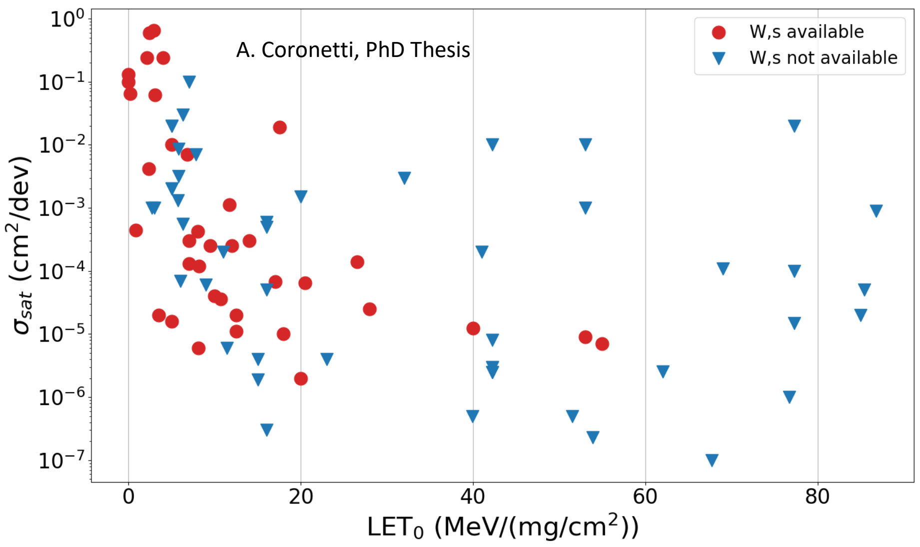
$P(\sigma_s^i \mid w^{obs})$=$C_i$/C

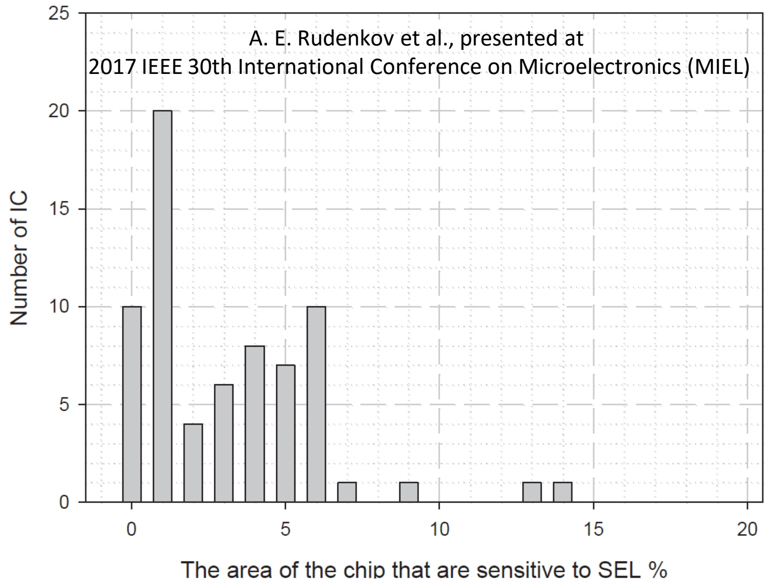


Point value of σ_s^{obs} gets replaced by distribution reflecting Data quality and Observed results.

If w^{obs} not known, C_j^i now include all MC events producing best-fit σ_s= σ_s^{obs} regardless of best-fit w→ broader distribution

What Does This Mean for Our Parametric Studies





- (σ_s, LET$_0$) points replaced by distributions reflecting data quality, systematic errors (e.g. σ not really saturated)
  - Effect on distribution difficult to discern without knowing more about each entry, but
    - Correlation between LET0 and σ_s at low LET likely persists
  - Even without doing the analysis, note departures from correlation are mostly at high LET and w, s not known
    - w, s not known imply broader distributions for true σ_s, LET$_0$
    - At higher LET$_0$, have fewer ions let to find saturation
  - Correlation may persist even at higher onset LET!

- For this study w and s were not included
  - Bins are actually much fuzzier than they would appear
  - Highest σ_s values are based on very few parts, so they are less certain
    - We get paid to be pessimistic—they could be higher!
  - Given correlation between LET$_0$ and σ_s per Coronetti, parts in the high σ_s tail likely dominate risk
- Is there a glimmer of hope for SEL screening?
  - 40 MeV/u Ion ranges at Texas A&M: Kr-0.59 mm, Ar-1.04 mm
  - Just sayin'

Conclusions

- Large databases hold out hope of bringing "Big Data" to bear on intractable SEE issues
  - Example: How do we reduce SEL risk when using commercial parts
- Unfortunately, available data entries represent highly variable data quality
  - Need to develop techniques that weight entries based on underlying data quality
    - # of LET values on σ vs. LET; # of events per σ point
    - Need to be cognizant of potential systematic errors:
      - Cross section not saturated at highest test LET in σ vs. LET
      - σ too small to see events w/ 10^7 ions/cm^2 near onset LET when shape parameter >3
- Monte Carlo studies can be very useful in determining data-quality weights for database entries
  - Useful for determining conditions that could have led to observed results—and from these probability distributions for observables
  - Self-consistent, objective, repeatable σ vs. LET fitting routine (e.g. GLM strategy) highly desirable
- Proper weighting ensures we can be confident about important results of an analysis
- Results here also argue for including more rather than less data in SEE databases
  - Event counts/error bars on cross sections allow proper analysis reflecting random and systematic errors
  - Where data are incomplete, analysts will have to get a lot more creative (and conservative)