BODIE TECHNOLOGY, INC.
The Noisy Data Experts®

Kornucopia's Ability to Synthesize Realistic Oscillatory Transient Shock Signals, Including Statistical Considerations

Ted Diehl, PhD

Bodie Technology, Inc

95th Shock & Vibration Symposium, 2025

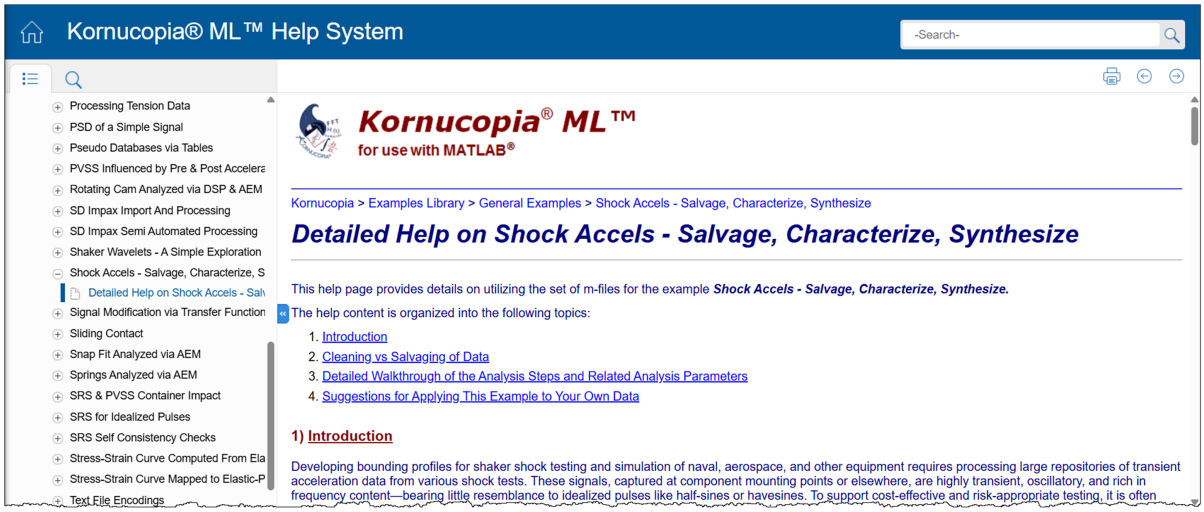
Overview

- Very brief review of a typical synthesis work-flow
  - From a large library of problematic raw data to a credible synthesized transient oscillatory shock signal.
- The method relies on building a statistically-based bounding PVSS curve from a library of curves.
- Current approaches ASSUME Gaussian or Lognormal behavior across PVSS curves.
  - Is that justifiable?
  - Can we do better?
    - YES – by extending on CDF methods.
- Conclusions / Status.

2024 S&V Presentation:

Synthesizing a Realistic Transient Acceleration Shock Signal to Represent and Bound Diverse Oscillatory Shock Time-Histories

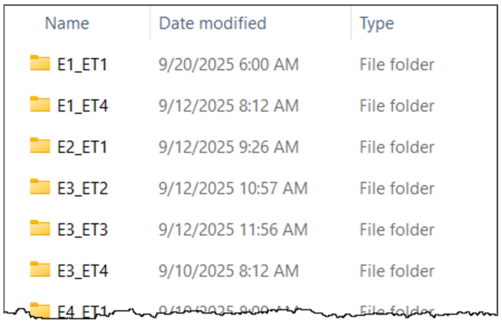
Tools are in Kornucopia now, currently utilized by the US Navy in design and shock test specifications.



BODIE TECHNOLOGY, INC.
The Noisy Data Experts®

Quick Re-cap of Current Salvaging & Synthesis Workflow

Library of Raw Acceleration Measurements



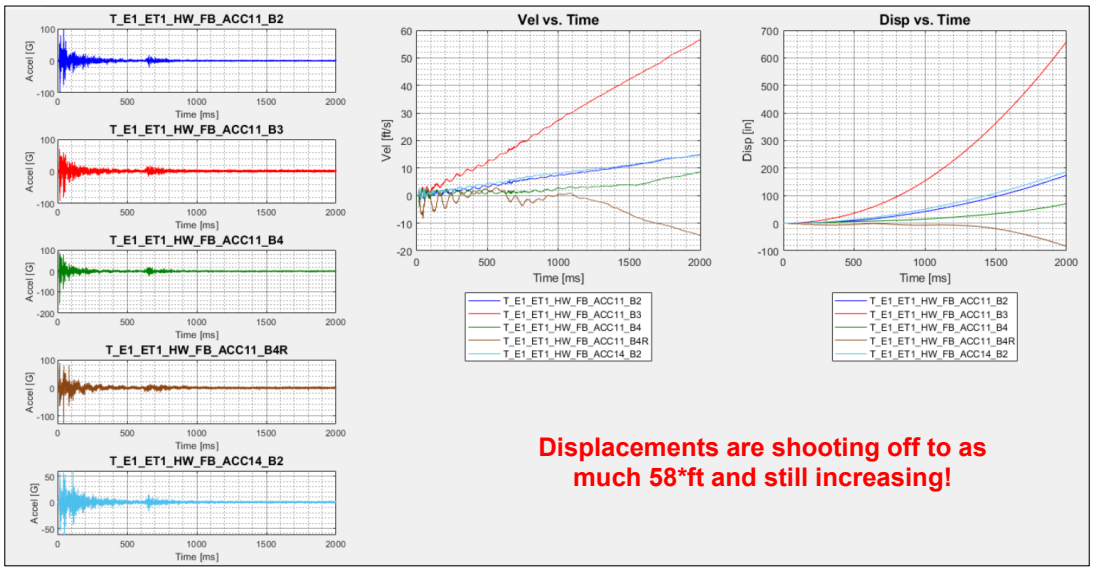
Name	Date modified	Type
E1_ET1	9/20/2025 6:00 AM	File folder
E1_ET4	9/12/2025 8:12 AM	File folder
E2_ET1	9/12/2025 9:26 AM	File folder
E3_ET2	9/12/2025 10:57 AM	File folder
E3_ET3	9/12/2025 11:56 AM	File folder
E3_ET4	9/10/2025 8:12 AM	File folder
E4_ET1	[illegible]	File folder

- **14 folders**
- **Over 2,500 files (signals)**
- **Different test environments**
- **Nearly all have issues**

For algorithm details, see last year's talk or Kornucopia documentation.

Raw Measurement Data is NOT Credible Relative to Velocity & Displacement

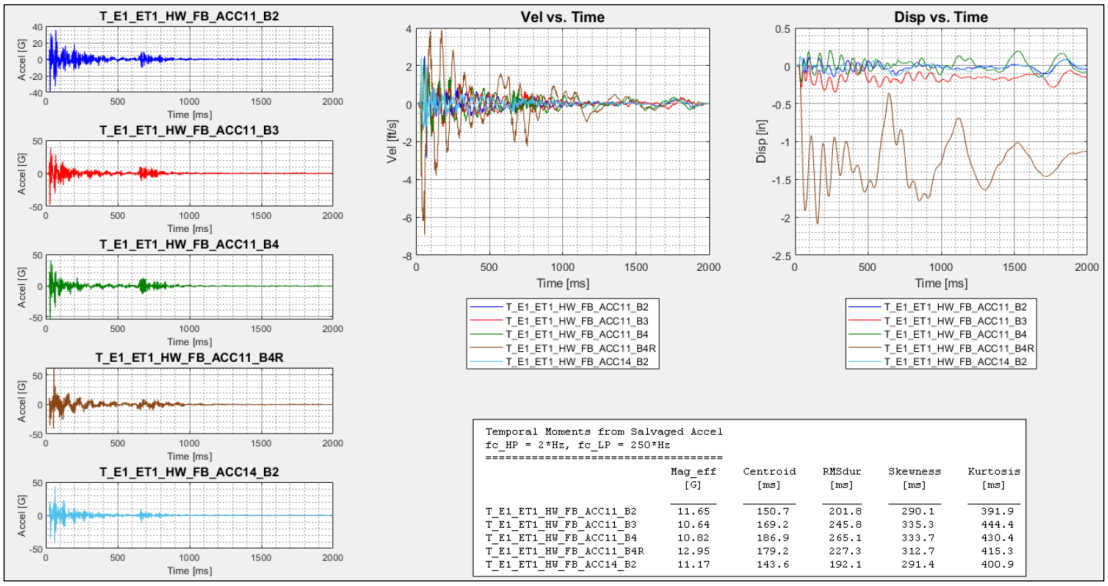
- Raw Heavyweight data.
- **FB** Response Dir.
- Acceleration seems OK.
- Time-integrating to Velocity and Displacement yield implausible results.
- Even if this data was somehow considered plausible, it could NOT be used to drive a mechanical shaker.



Displacements are shooting off to as much 58*ft and still increasing!

Kornucopia-Salvaged Data is Credible Relative to All 3 Kinematic Measures

- Salvaged Heavyweight data.
- **FB** Response Dir.
- Data filtered with fc_HP = 2*Hz, fc_LP = 250*Hz.
- All three kinematic measures are plausible.
- Data could drive a mechanical shaker that has sufficient capabilities.

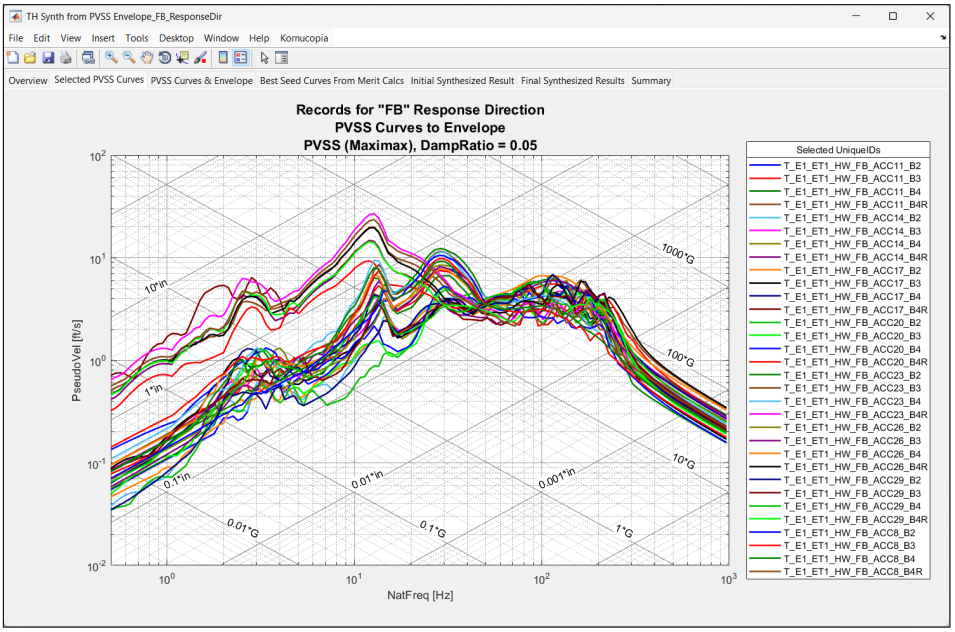


Temporal Moments from Salvaged Accel
fc_HP = 2*Hz, fc_LP = 250*Hz

	Mag_eff [G]	Centroid [ms]	RMSdur [ms]	Skewness [ms]	Kurtosis [ms]
T_E1_ET1_HW_FB_ACC11_B2	11.65	150.7	201.8	290.1	391.9
T_E1_ET1_HW_FB_ACC11_B3	10.64	169.2	245.8	335.3	444.4
T_E1_ET1_HW_FB_ACC11_B4	10.82	186.9	265.1	333.7	430.4
T_E1_ET1_HW_FB_ACC11_B4R	12.95	179.2	227.3	312.7	415.3
T_E1_ET1_HW_FB_ACC14_B2	11.17	143.6	192.1	291.4	400.9

PVSS Plot of All the FB Response Direction Data in the Study

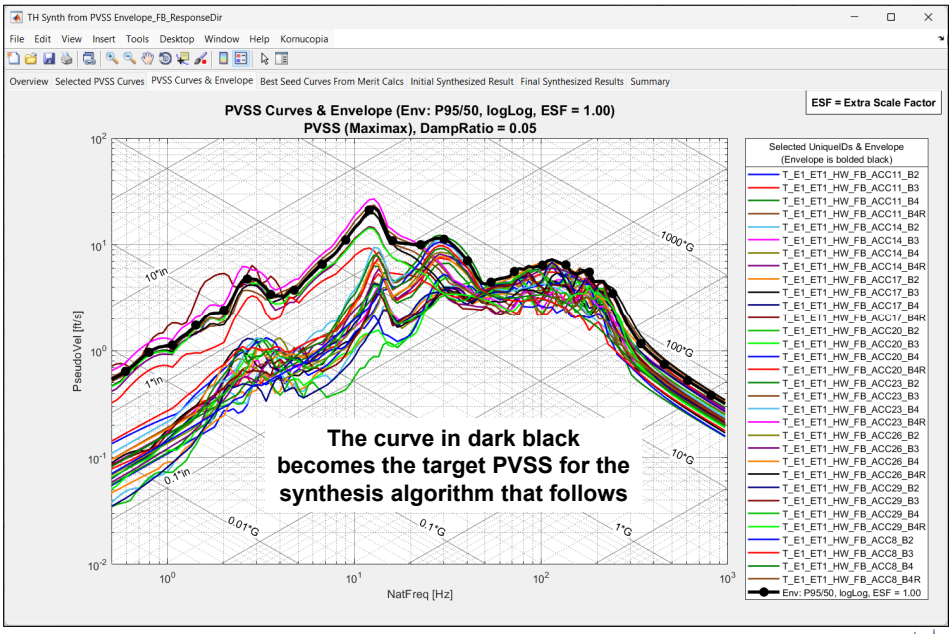
- Maximax PVSS curves for all 32 FB Response Direction signals.
  - PVSS curves all computed from salvaged signals.
- Quick Observations.
  - Below 20*Hz, the curves separate into two clear groupings.
  - Above 50*Hz the curves converge into one group.



Statistically Enveloping All the FB Response Direction PVSS Curves

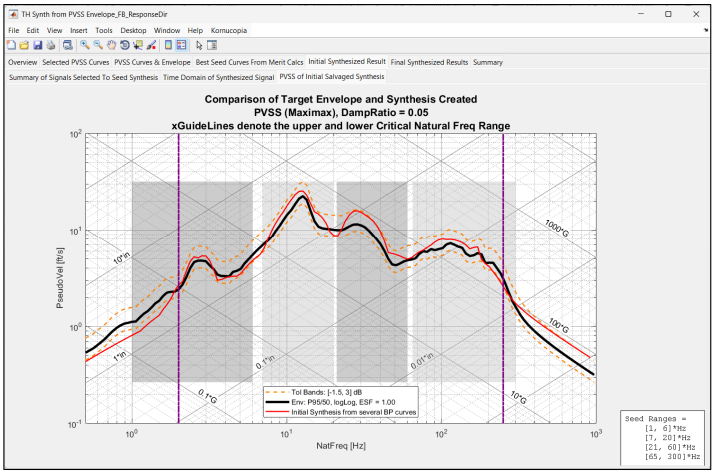
- Several statistically-based enveloping methods could be applied. In this example a **P95/50 envelope** is computed.
  - Based on a one-sided normal tolerance limit approach to meet a 95-percentile limit with 50% confidence.
  - Statistical bounding calculations done in the log-log space as this is traditionally more likely to satisfy normality requirements of statistics.
    - But is this ASSUMPTION of lognormal true???

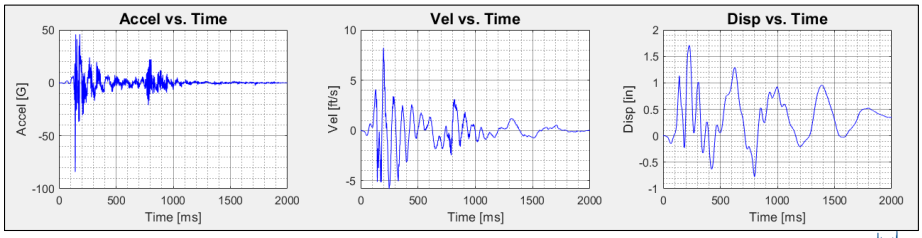
This is applying Gaussian statistics in the log10 space.



Creating Representative Time History from Target PVSS Curve

- User picks key natural frequency ranges in PVSS.
- Best seed signals chosen from salvaged library acceleration time-histories.
- Seed signals time histories are bandpass-filtered, scaled, and combined → Synthesized.
- Synthesized PVSS checked against target (with tolerance).
  - If needed: adjust BP ranges, scaling, or add Shaker Wavelets.
- **Result – realistic oscillatory transient signal, looks like the underlying test data.**





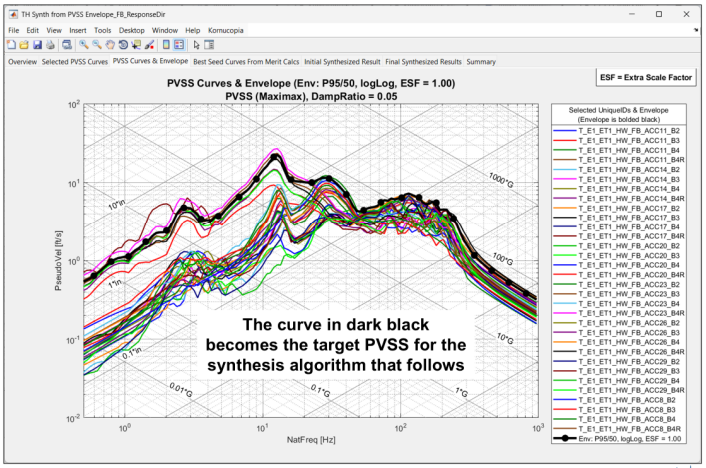
But Was Our Statistical Bounding Justified?

- Let's look a little deeper at our **Gaussian / Lognormal** ASSUMPTIONS.

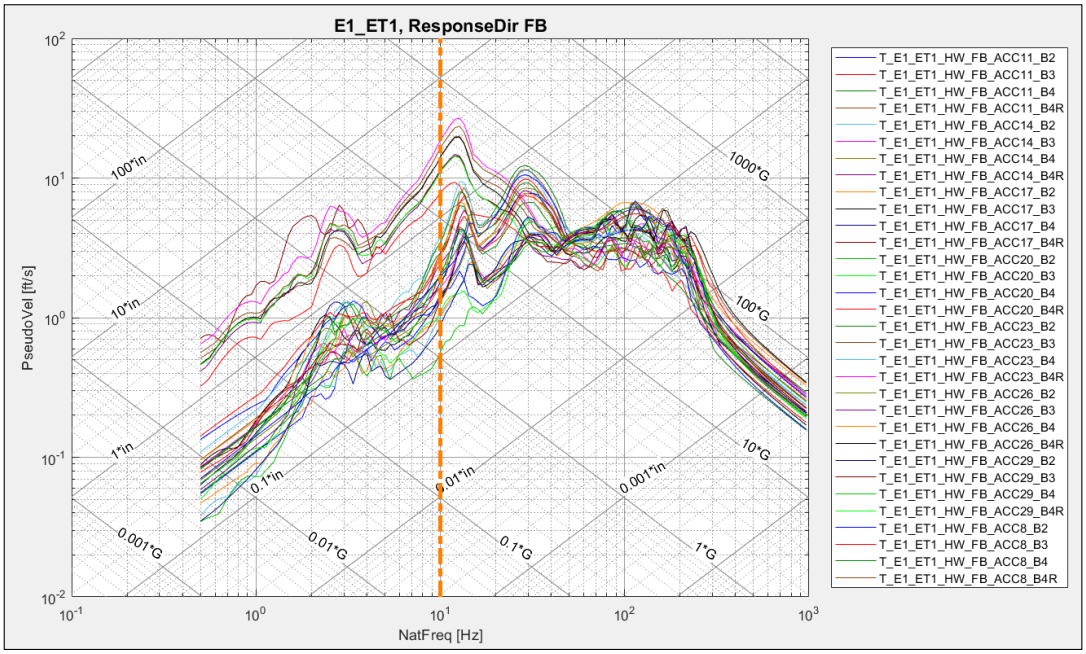
Statistically Enveloping All the FB Response Direction PVSS Curves

- Several statistically-based enveloping methods could be applied. In this example a **P95/50 envelope** is computed.
  - Based on a one-sided normal tolerance limit approach to meet a 95-percentile limit with 50% confidence.
  - Statistical bounding calculations done in the log-log space as this is traditionally more likely to satisfy normality requirements of statistics.
    - But is this ASSUMPTION of lognormal true???

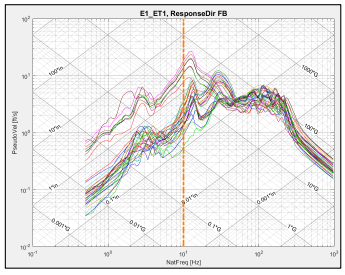
This is applying Gaussian statistics in the log10 space.



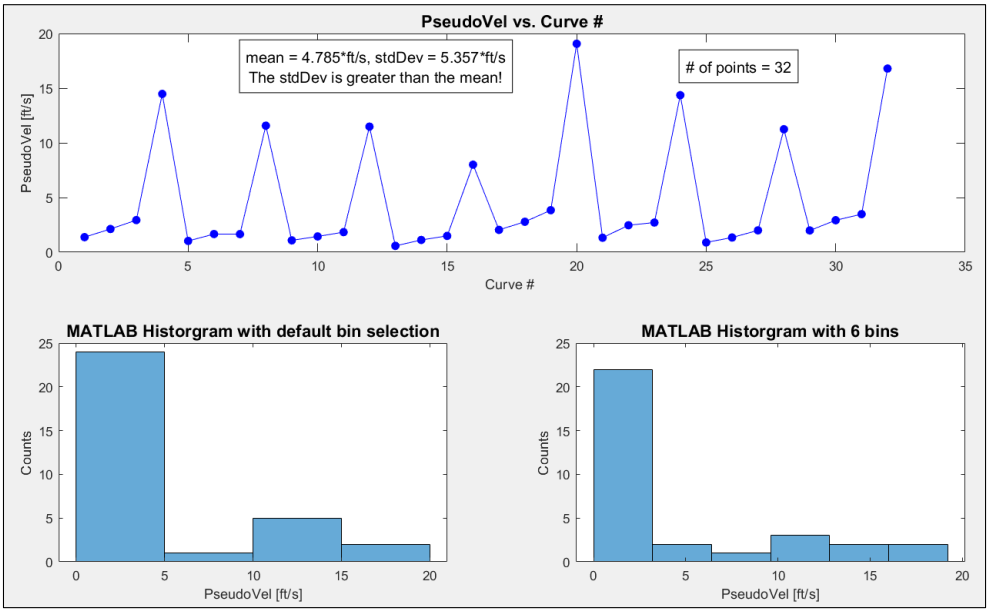
Consider 1 Slice @ 10*Hz



Consider 1 Slice @ 10*Hz



- Clearly some repetition, 3 tests with slight increase, then 4th with large increase.
- Taken all together, the slice is heavily skewed.
- This is not Gaussian. It fails:
  - Shapiro-Wilk (SW)
  - Kolmogorov-Smirnoff (KS)
  - Anderson-Darling (AD)



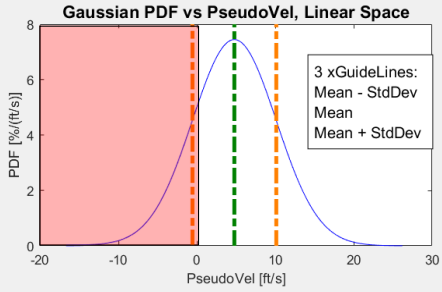
Analytical PDF (Probability Distribution Functions) for Gaussian & Lognormal Assumptions

Gaussian:

- PDF analytically computed based solely on mean & std.
- Roughly one-third of the PDF is physically implausible for this slice!

Lognormal

- Raw PV values first transformed:
  X = log10(PV / xUnits)
  xUnits = ft/s.
- Then compute Gaussian statistics on X but must include the Jacobian of PV due to chain rule.



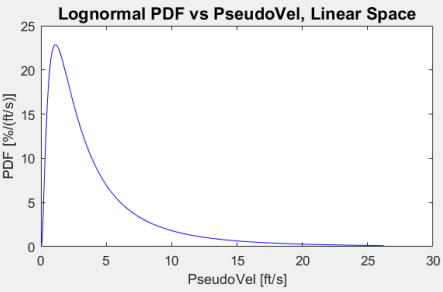
Descriptive Stats (linear space):
mean = 4.785*ft/s
std = 5.357*ft/s

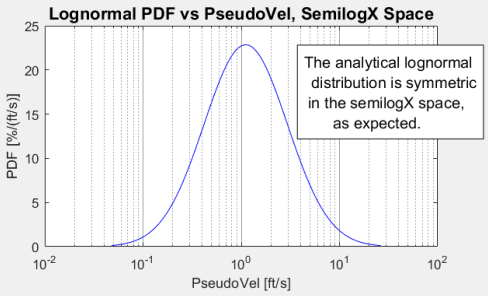
o The std is greater than the mean!
o (mean - 1*std) yields negative PV!

Classical stats checks on **Gaussian** usage:
- SW - failed
- KS - failed
- AD - failed

Classical stats checks on **Lognormal** usage:
- SW - failed
- KS - passed
- AD - passed

Who do you believe?





Curves are Gaussian/Lognormal models of the data, not the data's true PDF.

Cumulative Distribution Functions: CDF and eCDF

- CDFs are computed by integrating PDF.
  - For Gaussian this leads to the equations shown below and the plots to the right.
  - f = PDF
  - F = CDF
  - x = PrimaryData
    For PVSS analysis
    x = PseudoVel (PV)
  - t = dummy integration variable.
  - erf = error function, has no closed form solution. It must be evaluated numerically.

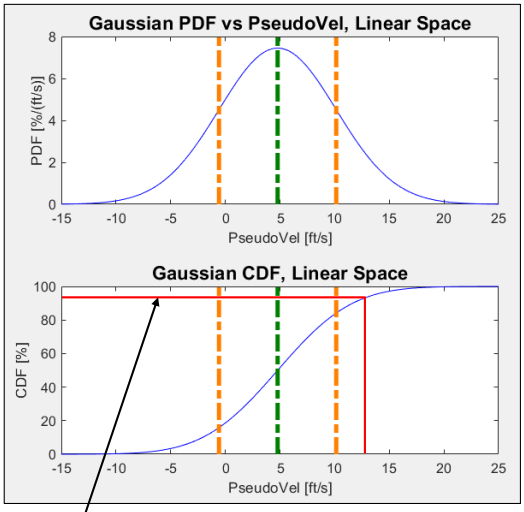
$$F(x) = \int_{-\infty}^{x} f(t)\,dt$$

$$f(t\,|\,\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right)$$

$$F(x\,|\,\mu,\sigma) = \int_{-\infty}^{x}\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right)dt$$

$$F(x\,|\,\mu,\sigma) = \tfrac{1}{2}\left[1+\operatorname{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right]$$

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}}\int_{0}^{x} e^{-t^2}\,dt$$

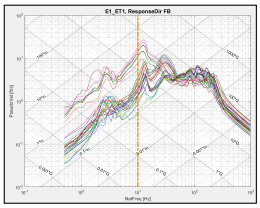


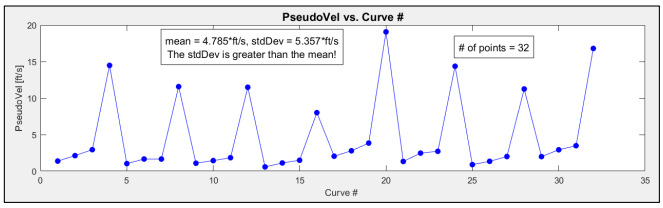
- **Histograms give poor PDF estimates.**
- **CDFs can be estimated (called eCDF) much more robustly using established formulas such as:**
  - **Median Ranks, Bernard, Gringorten, …**

For estimating a P95/50 bound, the CDF tells use what we want, the PDF does not.

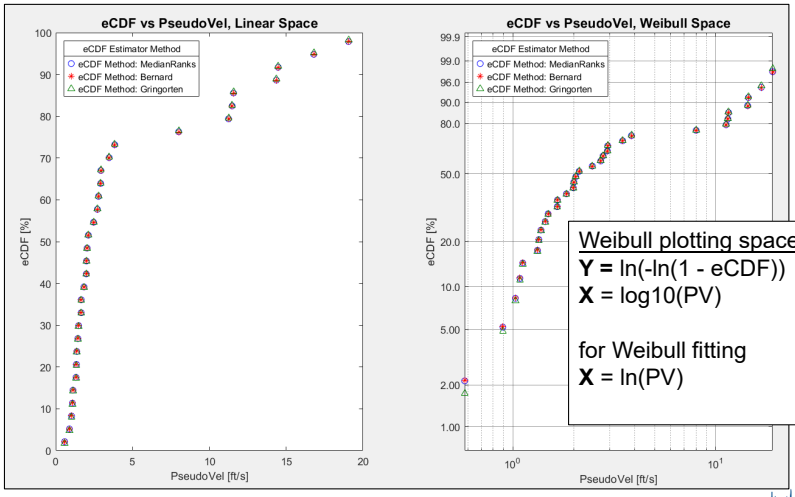
3 PV guidelines
Mean – 1*std
Mean
Mean + 1*std

eCDF Computed on Our Slice of Pseudo Velocity Data





- An **eCDF** is a direct estimate of the **CDF** based on the actual data.
  - This tells us directly the Probability information we desire.
  - *Bootstrapping* to get confidence level.
- Three estimate methods shown:
  - Median Ranks
  - Bernard ✅
  - Gringorten
- Median Ranks and Bernard virtually indistinguishable. Bernard formulation more computationally efficient (important for bootstrapping).
  - Gringorten distorts the tails in a non-conservative manner.

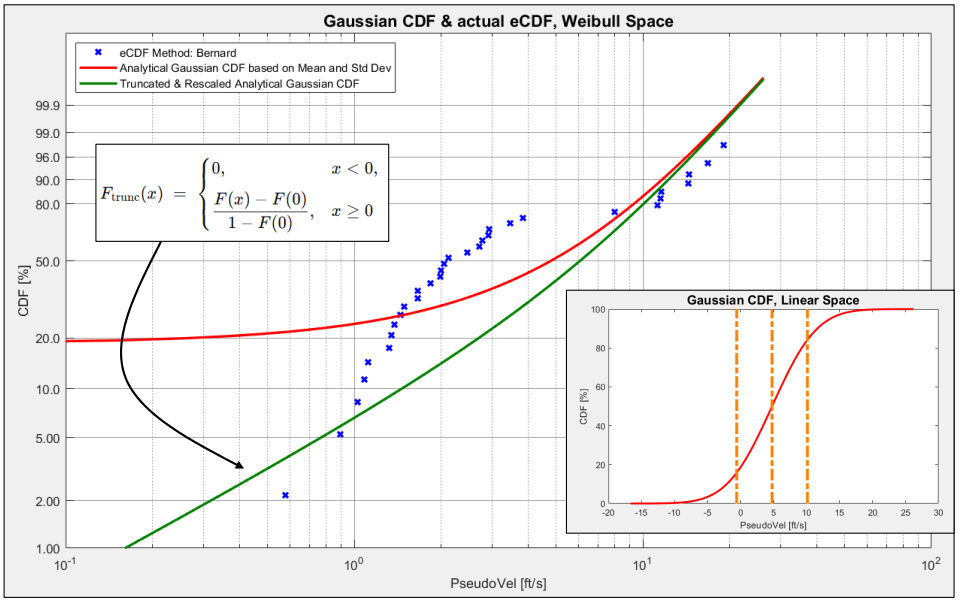


Weibull plotting space
Y = ln(-ln(1 - eCDF))
X = log10(PV)

for Weibull fitting
X = ln(PV)

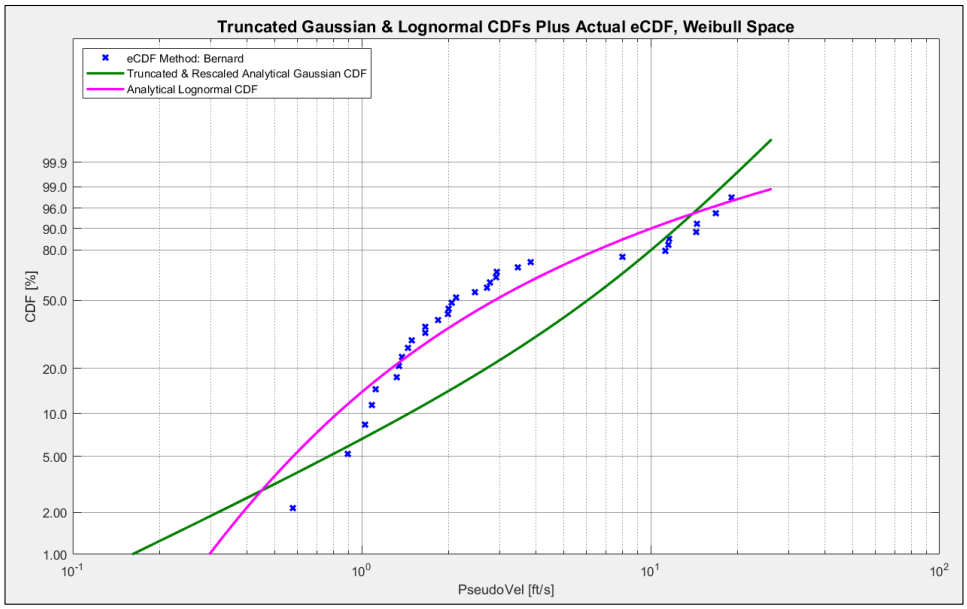
CDF / eCDF Assessment Further Shows Gaussian is Not Valid for This Slice

- **Blue** symbols are eCDF values computed directly from the slice via Bernard's estimate.
- **Red** & **green** curves are Gaussian model of the data based on mean and std.
  - Inset is Gaussian analytical CDF in linear space.
  - Red curve in Weibull space is simply truncating the analytical CDF to plot only positive PV values.
  - Green curve applied classic *truncate and rescale* for CDF formulation.
- **None of these Gaussian "fits" represent the eCDF well.**



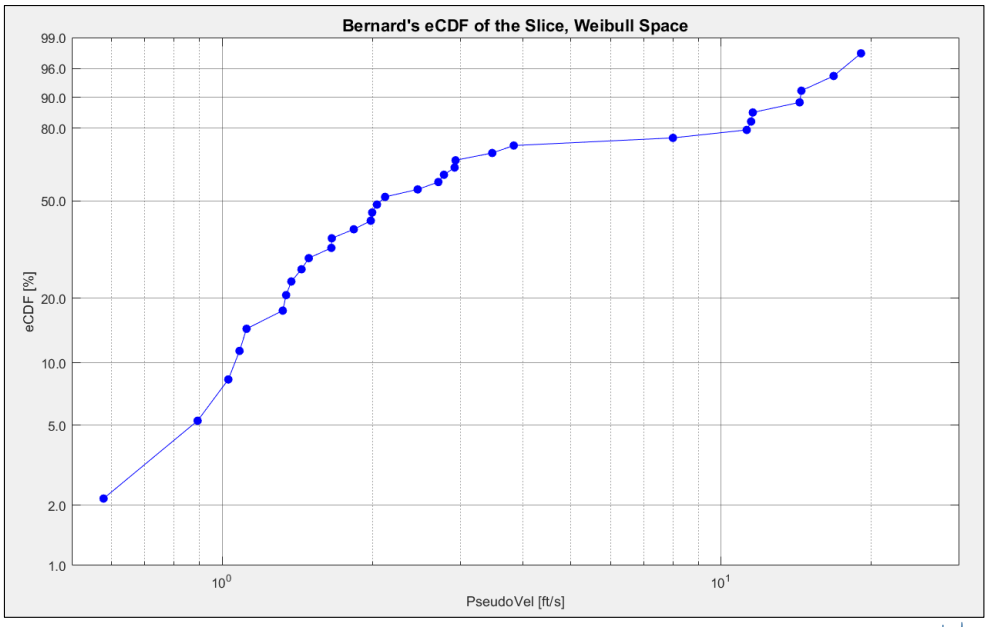
CDF / eCDF Assessment Further Shows Lognormal is Better But Not Great for This Slice

- **Blue** symbols are eCDF values computed directly from the slice via Bernard's estimate.
- **Green** curve is Gaussian model of the data based on mean and std using classic truncate and rescale for CDF formulation.
- **Magenta** curve is Lognormal CDF.
- **Lognormal is an improvement over simple Gaussian for this dataset, but the fit is still not that great.**



Can We Just use the eCDF Directly for Predicting P95/50 or Similar Requests?

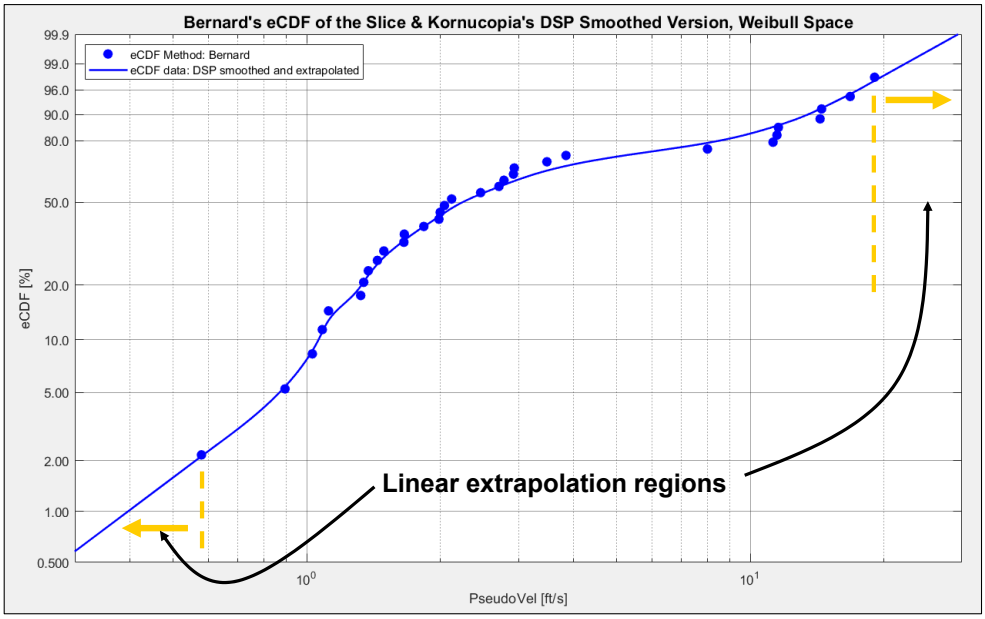
- eCDF vs PV data is strictly monotonic and functional but bumpy (not desirable)
- We would like to smooth the eCDF vs PV data somehow without enforcing an assumed model.
- Polynomial fitting is tricky to apply robustly in this context.
- **DSP-based smoothing shows great promise!**
  - Kornucopia already has all the needed algorithms to achieve this robustly.
  - Then use linear or PCHIP interpolation to create a computable function that is easily invertible (PV vs eCDF).



… and use bootstrapping to yield the confidence level.

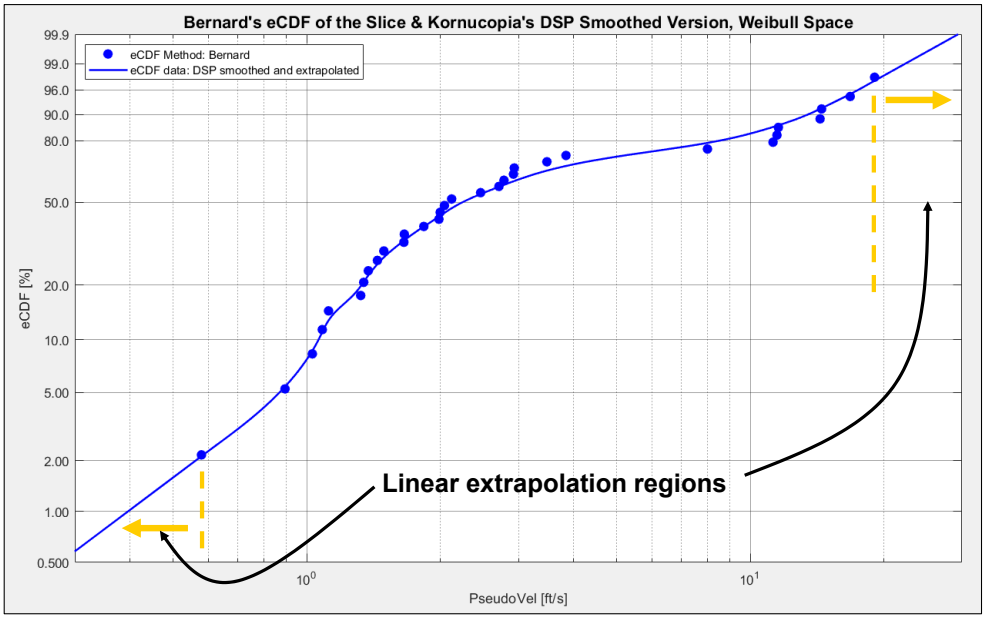
FlexCDF™ – DSP Smoothing of eCDF Data – Robust and Powerful

- **FlexCDF™** – a framework for robust statistical analysis directly from your eCDF vs PrimaryData.
  - Advanced DSP to smooth underlying data.
  - Invertible statistical models without assuming a specific distribution.
- **FlexWeibull™** - specialized positive-only mode based on the Weibull space (for smoothing and interpolation function).
  - Ideal for Shock and Failure Analysis.



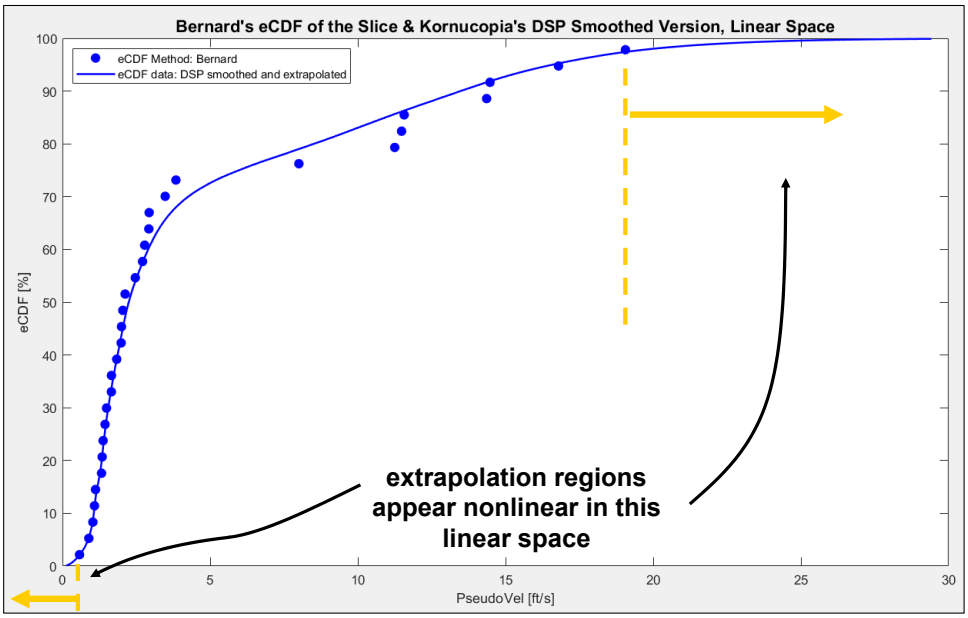
FlexWeibull™ – Some Details

- Computed in Weibull space using specially regularized eCDF data.
- Apply DSP lowpass bidirectional filtering with advanced end-effect minimization technology from Kornucopia.
  - 1st order Butterworth
  - Linear end-projections.
  - Filtered in Weibull space and then mapped back to linear space as needed.
  - User can select cutoff "frequency" (actually a duration factor) for filter. Default setting of (1/4 duration) yields good overall results.



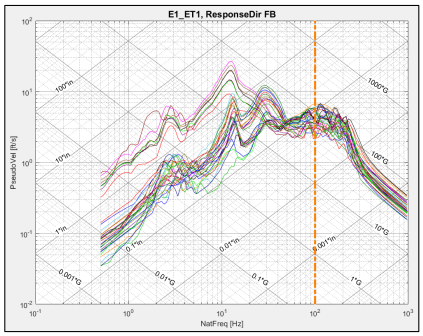
The eCDF Computed Via FlexWeibull Maps Back to Linear Space Well

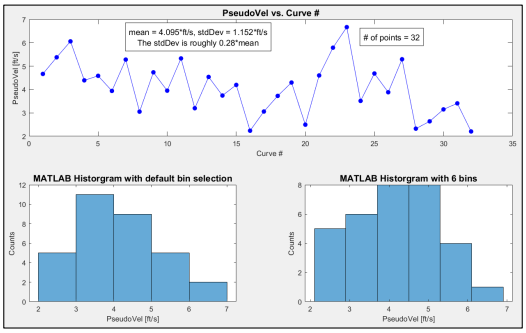
- The entire range appears smooth and plausible, including the extensions.

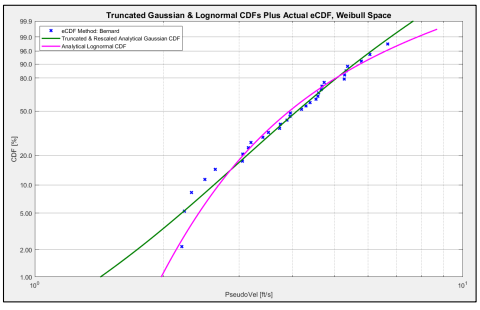


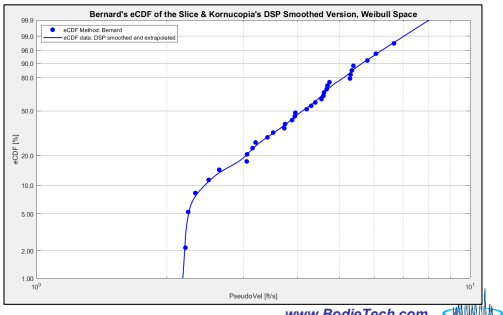
FlexWeibull™ Results From a Different Slice Demonstrates Algorithm's Broad Applicability

- At 100*Hz, the slice has a very different nature relative to the 10*Hz slice.
  - The 100*Hz slice shows that Gaussian is maybe as good as Lognormal.
- **FlexWeibull** run with the same filtering parameters as the other slice equally handles this slice as well as the other slice.
- **Conclusion, FlexWeibull should work well for all the slices with little to no user interventions required!**









Conclusions and Status

- Kornucopia® ML™ offers many advanced tools to help engineers working with shock data.
- This presentation provided a brief overview of Kornucopia's ability to synthesize realistic oscillatory transient shock signals, including statistical considerations.
- As released today, statistical bounding of a set of curves is based on Gaussian or lognormal methods.
- **FlexCDF / FlexWeibull** – Under development, designed to better handle the statistical variations seen in large PVSS datasets and other types of data too.

FlexCDF™ – A general framework for robustly analyzing eCDF vs. PrimaryData using DSP-based smoothing.

FlexWeibull™ – Next-gen statistical bounding for shock data.

Both available in early 2026.