Basic Statistical Analysis for the Utility Cable Industry

Nigel Hampton
Borealis AB
Sweden

nigel.hampton@borealisgroup.com

Issues

- Where do statistics turn up
- Data collection – the starting point
- Statistical approaches
- Weibull
  - A walk through of a real problem
  - Number of samples – bias & accuracy
  - Scale and scatter
  - Size & length
- Non Parametrics
- Reliability Growth Models
- Further sources of information

Where do statistics turn up

- Quality Control – Routine HV Test, Cleanliness, Smoothness
- Qualification – AWTT, Cenelec
- Endurance Testing – ACLT
- Service Performance – Field Failures

Data collection – the starting point

- Understand the questions your want answered – but be aware there are probably questions that you dont know about yet
- Classification – the more you can separate things the easier and clearer the analysis will be
- Understand your measure of performance
  - Time: elapsed time, time on voltage, # of cycles, length
  - Stress: voltage, maximum or mean etc
  - Continuous or Discrete or bounded: count, concentration, hours vs days, percentage etc

Statistical approches

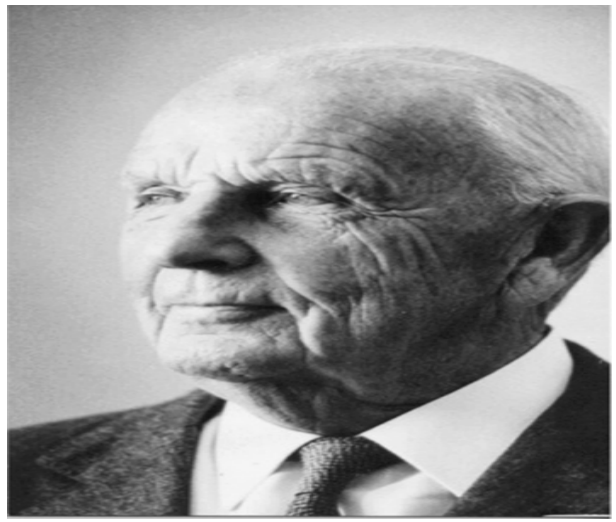
The presentation will not deal with all of the possible approaches but will highlight techniques that have been found to be useful

- Weibull
- Non Parametric
- Reliability Growth – Crow AMSAA

Techniques are primarily graphical in nature – see how the data looks

Weibull

- Comes from Extreme Value Statistics, it is the best approach for a small (<30) number of samples
- Works best for one failure mechanism at a time
- $P=1-\exp(S/a)^{-\beta}$



Data are ordered, smallest to largest.
The Probability of failure (P) is calculated.
The Probabilities are plotted against the Stress.
The data are linearised by plotting on a double log vs log plot

Why use the Weibull Distribution

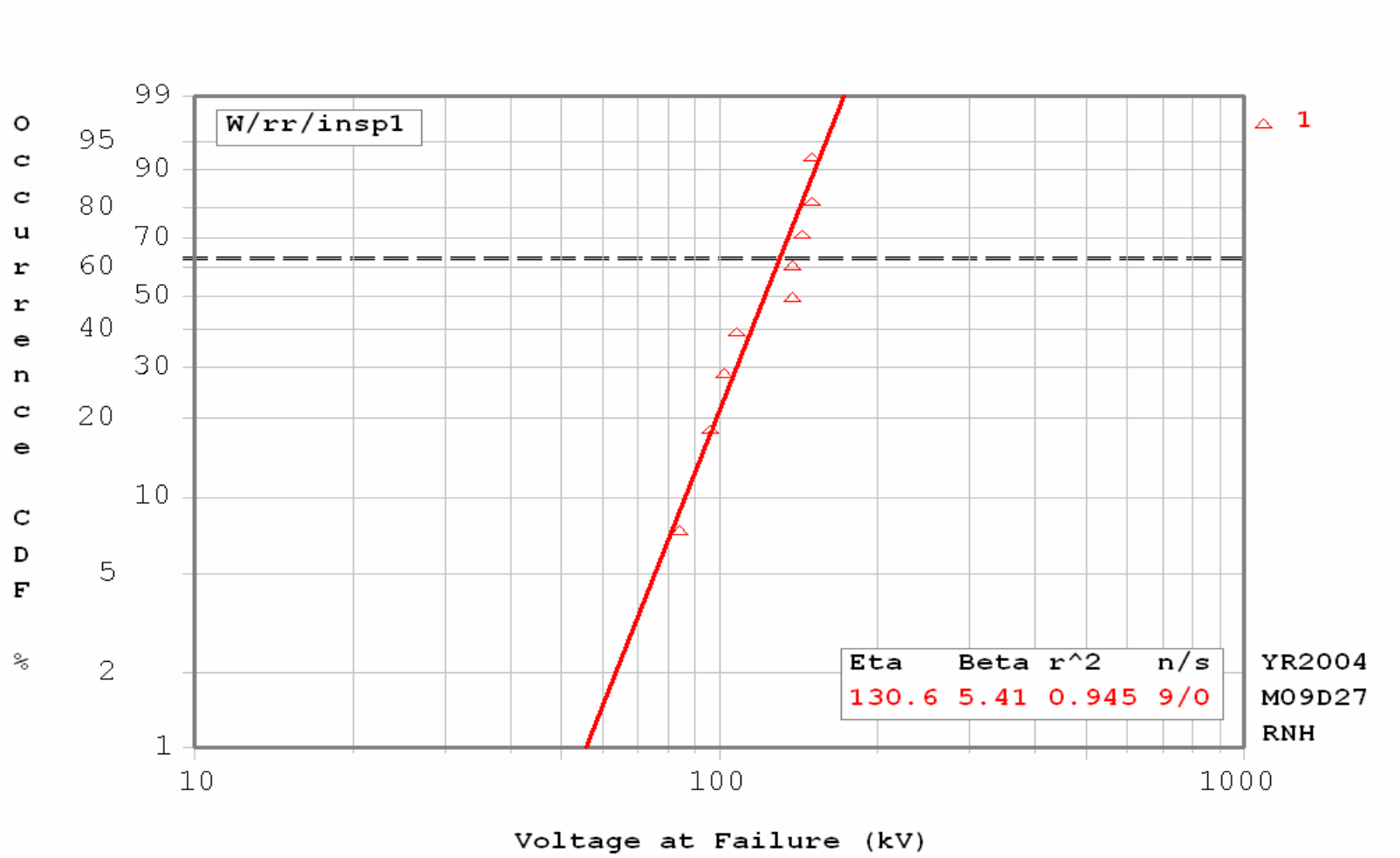
- It has the advantage that it uses the data both to estimate the magnitude (scale) and the shape of the distribution.

- It gives information about the failure mechanisms

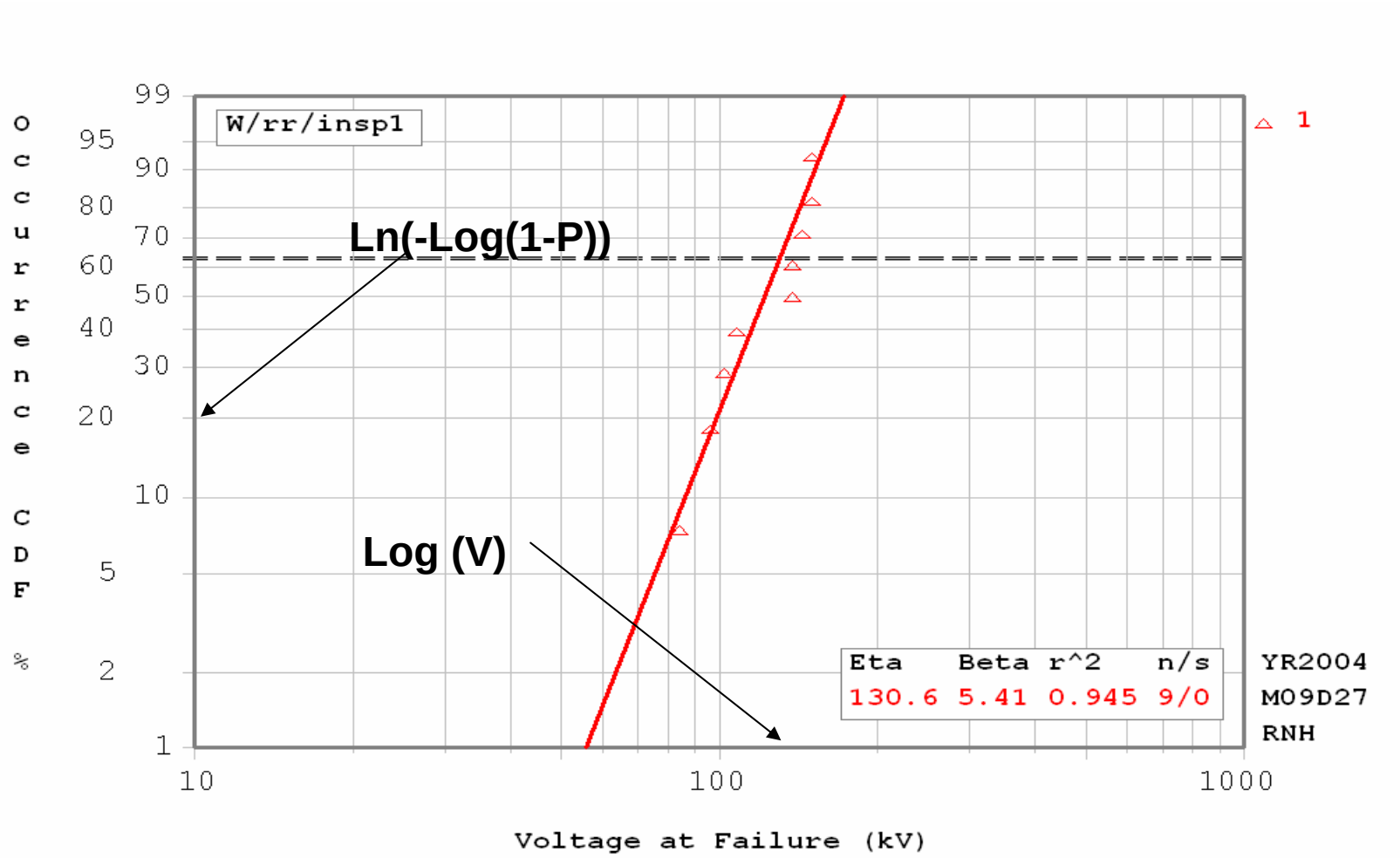
- It works well with ”extreme value” or ”weakest link” type problems

Machine Tool Breakages	Breakdown Strengths	Corrosion
Human Heights	Boiler Tube Failures	Stress Cracking
Times to Failures	Length of Earthworms	Tree Diameters

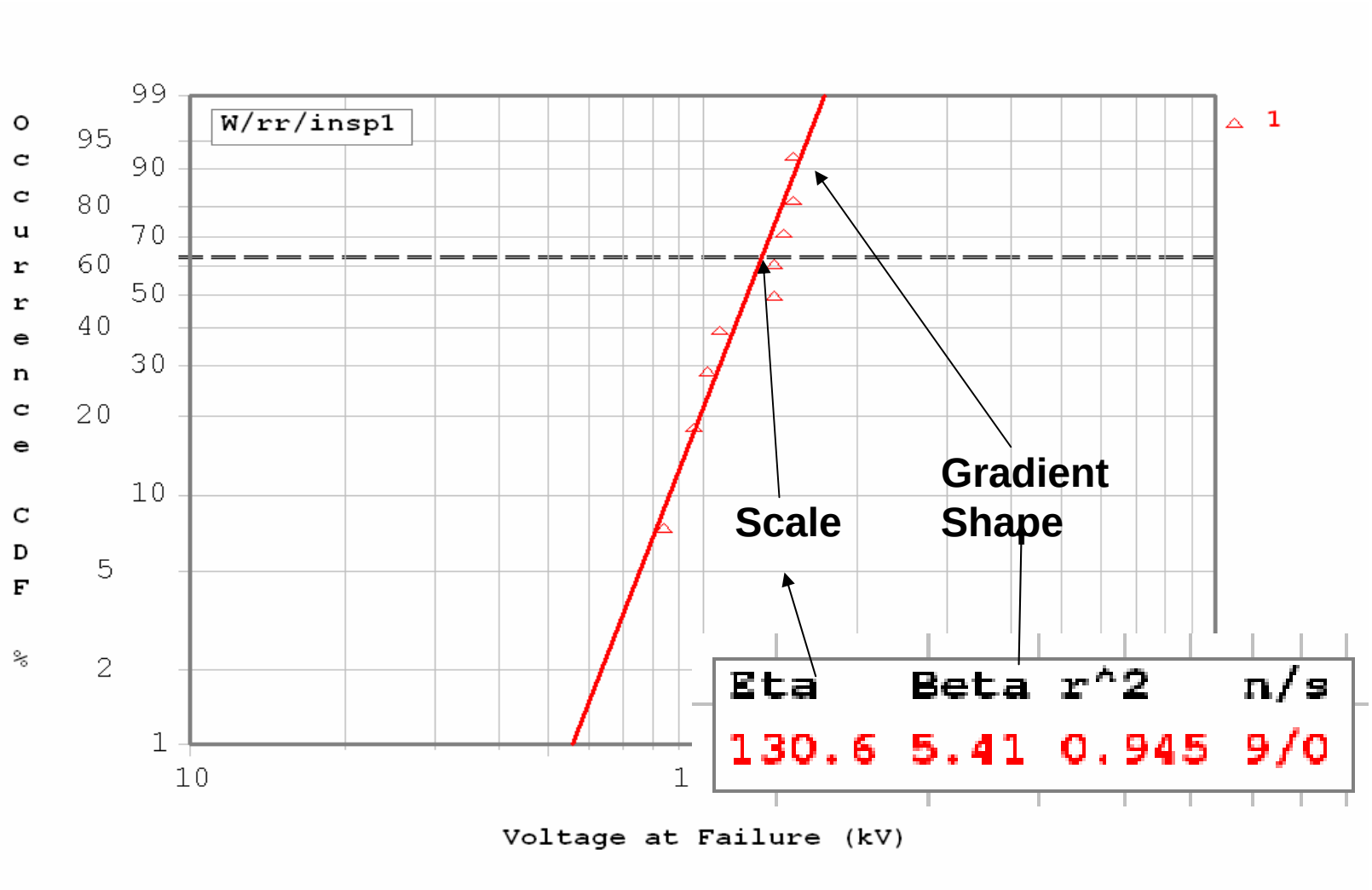
It does not work all of the time but it is always worth a try: the inventor said ”that it may **sometimes** render good service”.



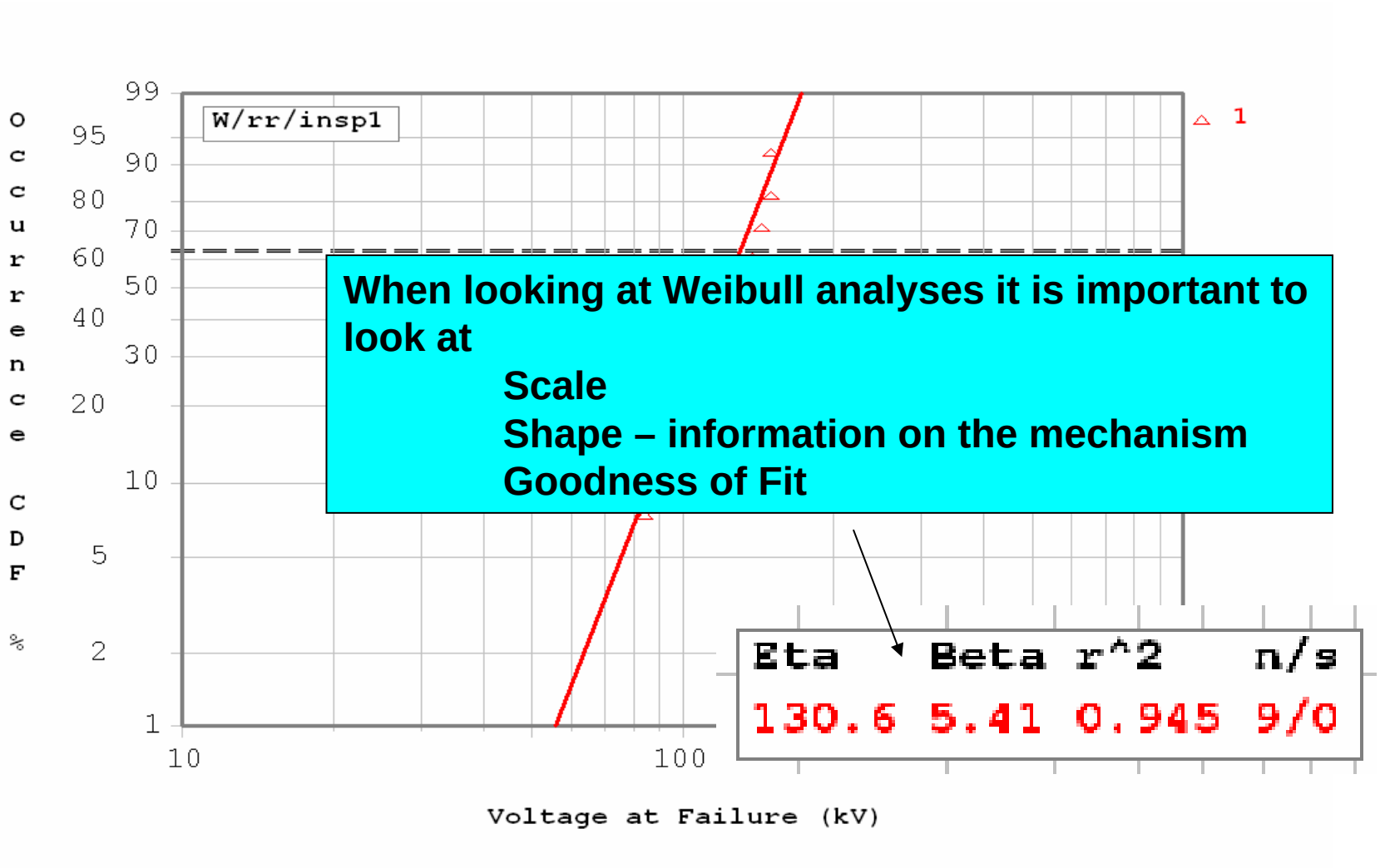
Accelerated long term evaluation of MV XLPE cables; Benjaminsen et al; Jicable03 – cable 10



Accelerated long term evaluation of MV XLPE cables; Benjaminsen et al; Jicable03 – cable 10

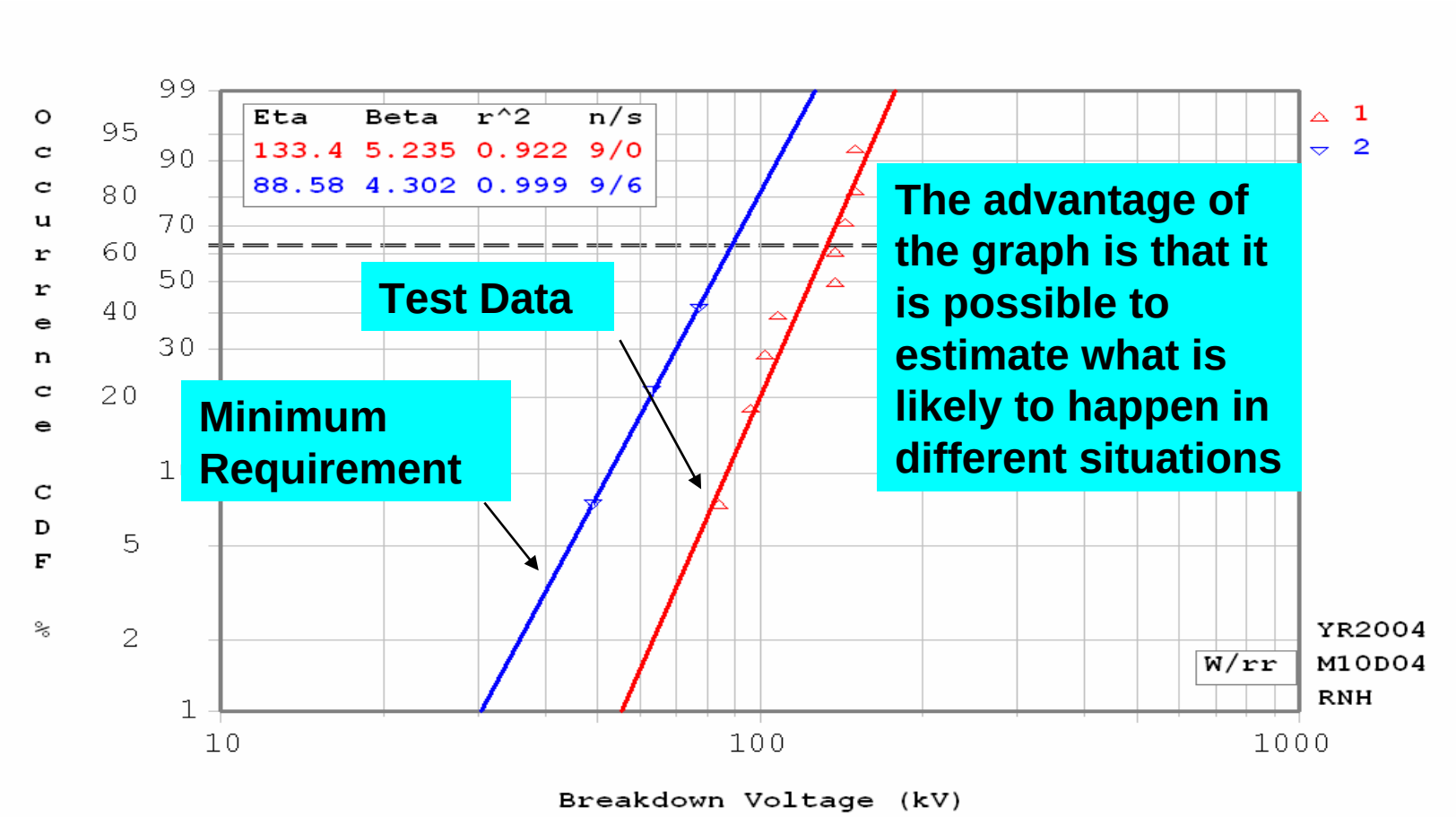


Accelerated long term evaluation of MV XLPE cables; Benjaminsen et al; Jicable03 – cable 10

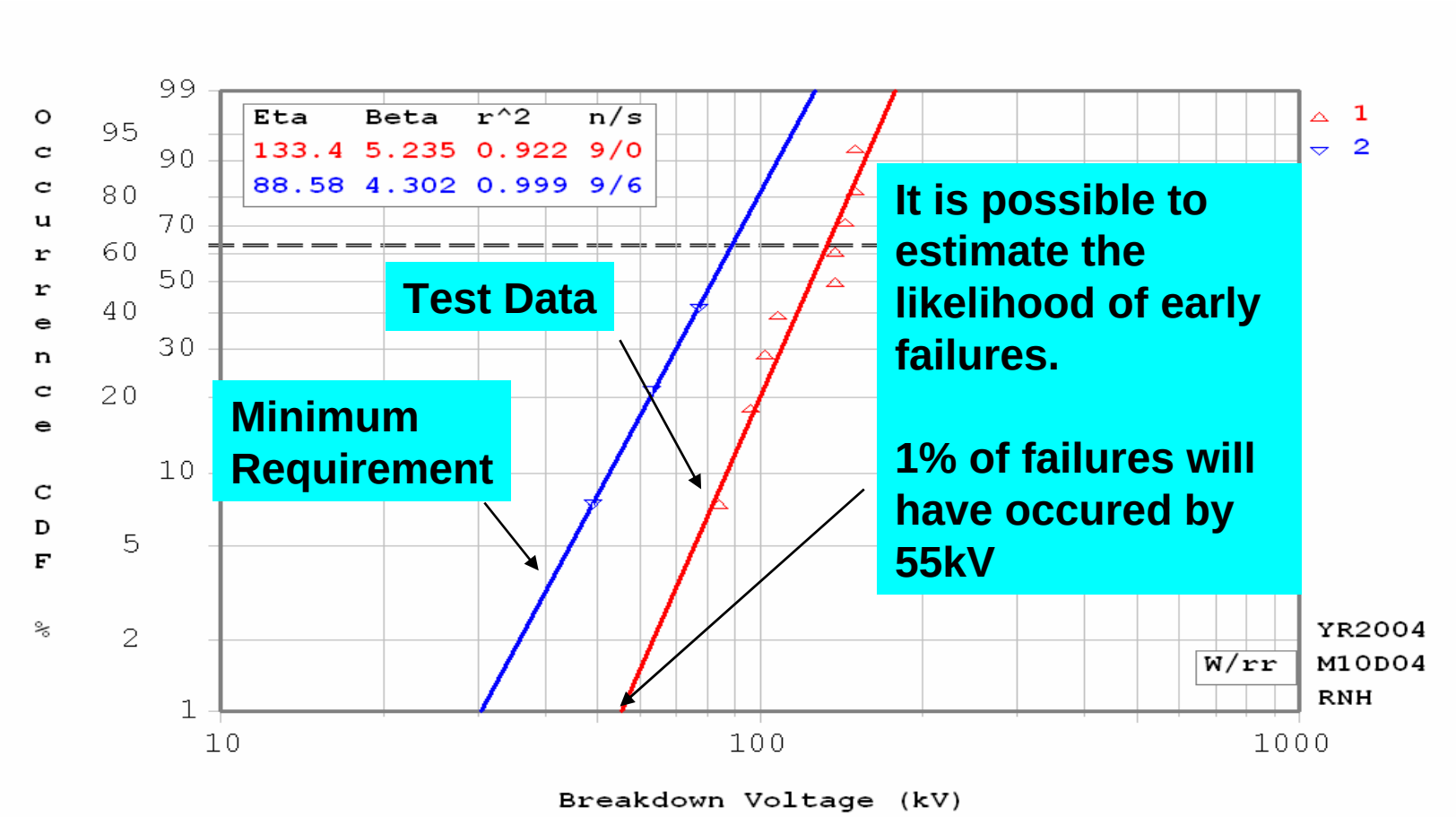


Accelerated long term evaluation of MV XLPE cables; Benjaminsen et al; Jicable03 – cable 10

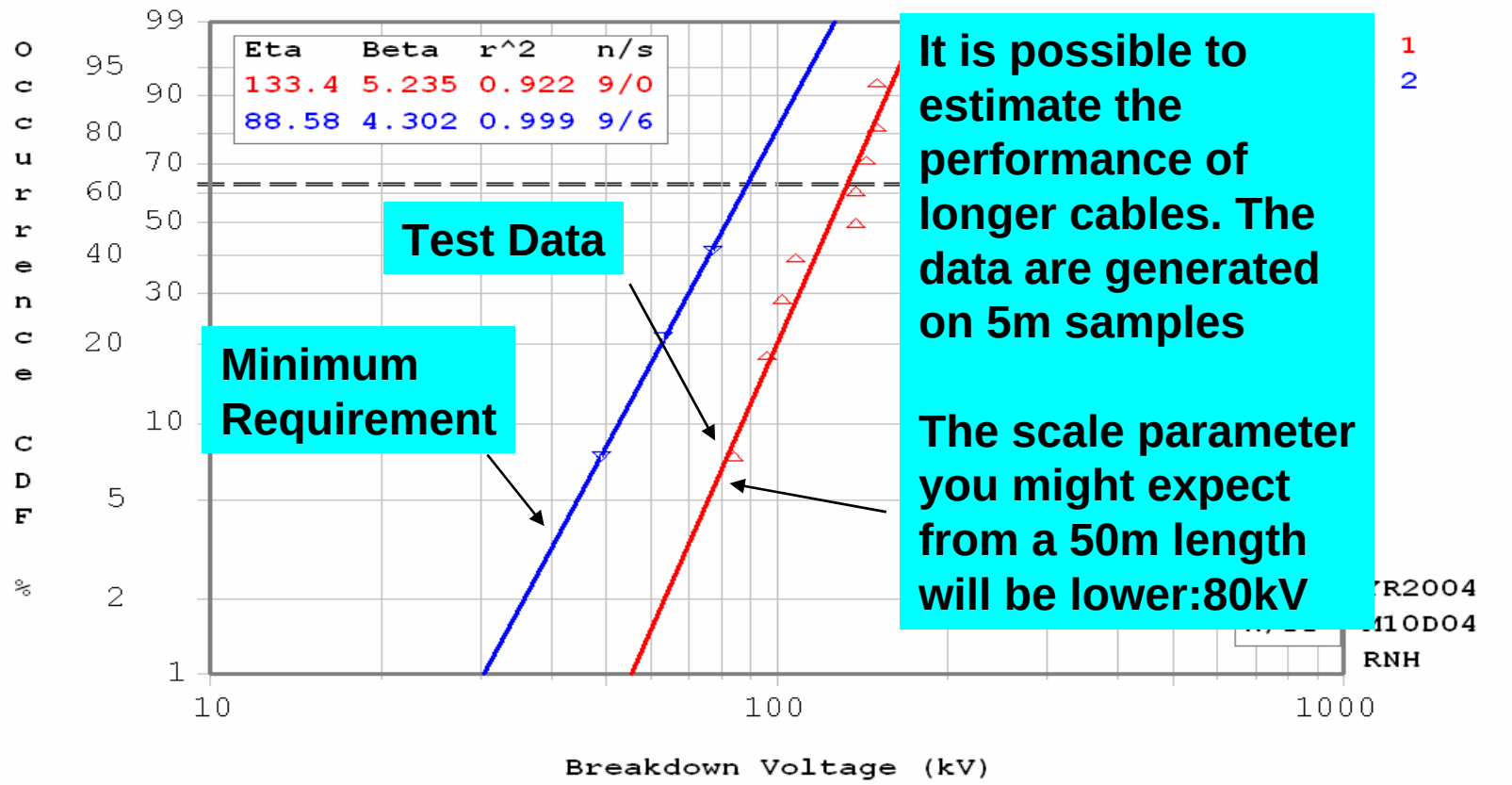
Predictions



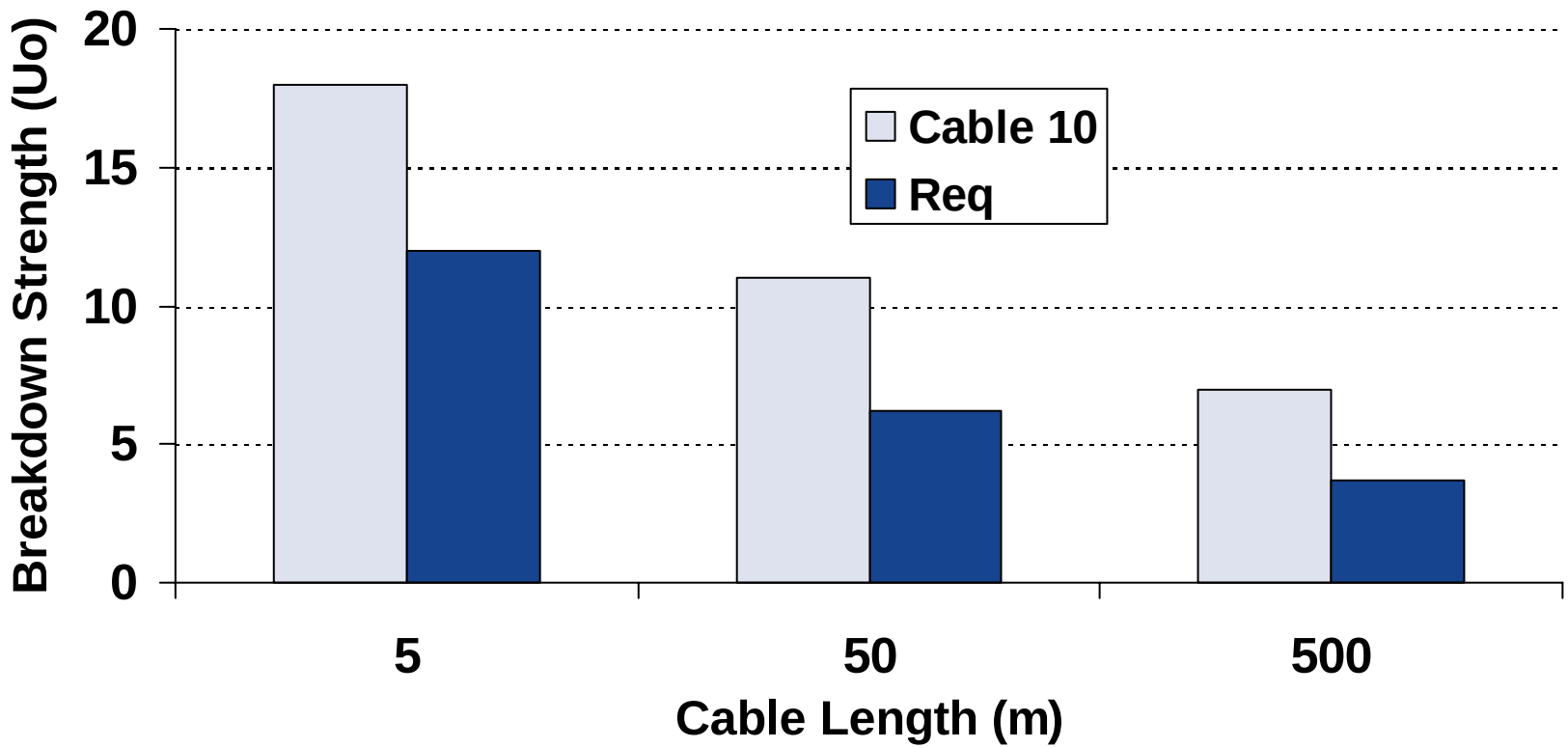
Predictions



Predictions



How good a performance do I need on short lengths to get good performance in service?



50% probability of failure

And now to ”Dirty Data”

- Suspensions
- Mixed Mechanisms
- Visitors
- Missing Data
- Size

Weibull Analysis – a walk through of real problems

- M Abou Dakku, A Bulinski & S Bamji

  Corelation between space charge development and breakdown in polymeric insulations under DC fields; ICSD2004 Toulouse

  Endurance tests under DC for thin samples of different insulations

- J Rodgers – Nodak Electric

  Service failure data for an installed cable system – data segregated by failure cause, year installed

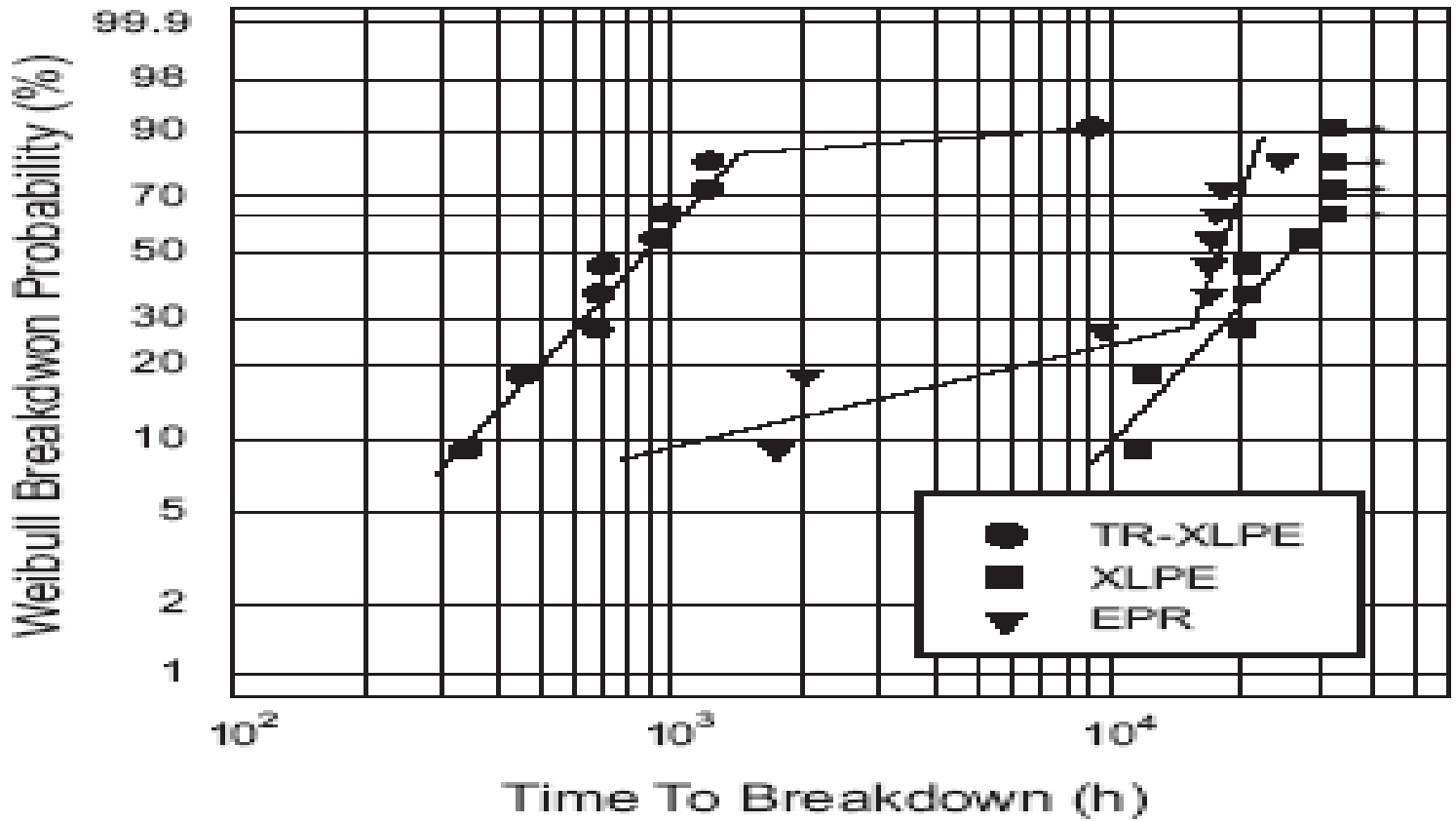
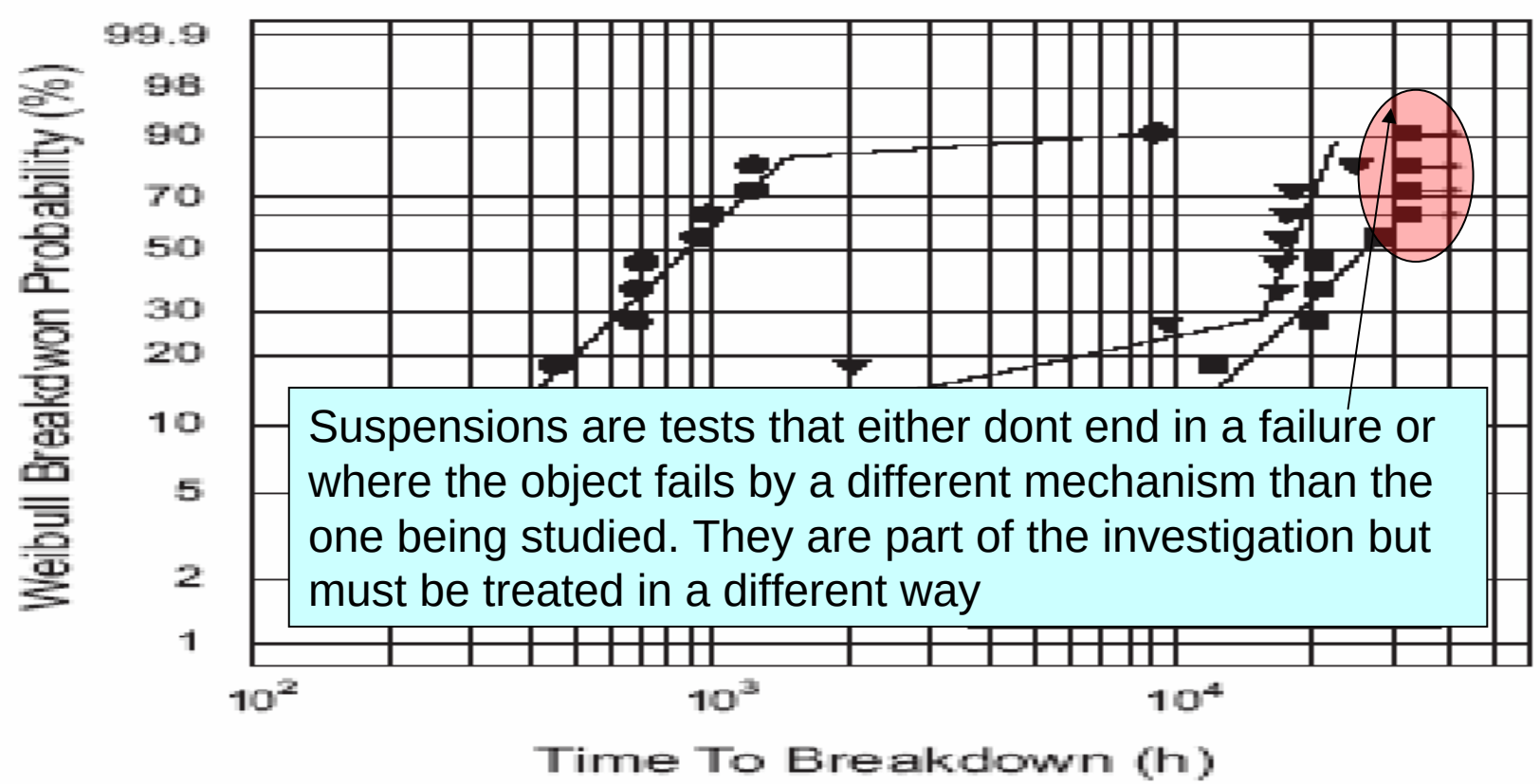
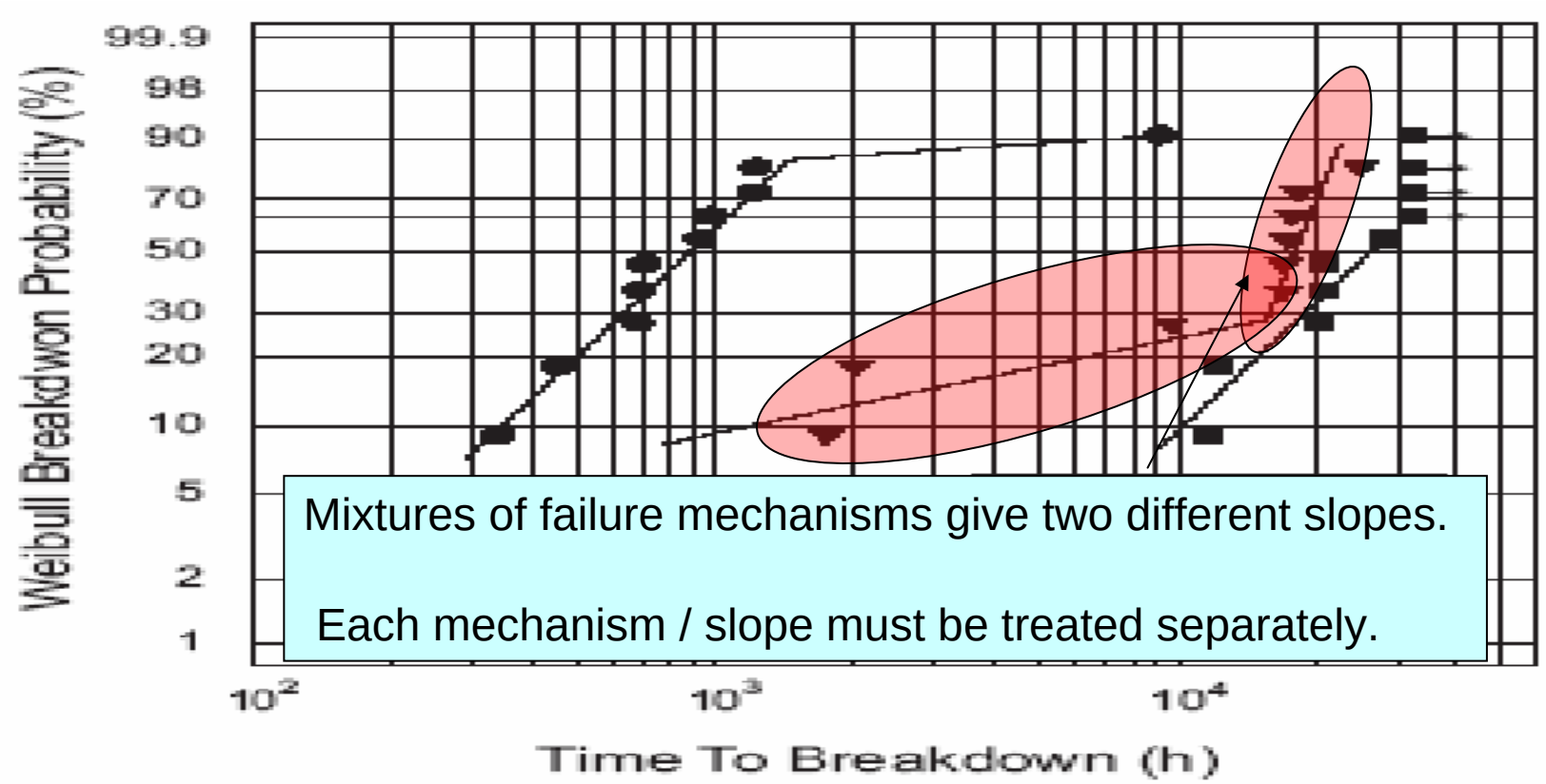


Fig. 6. Times to breakdown plotted in the Weibull plot.

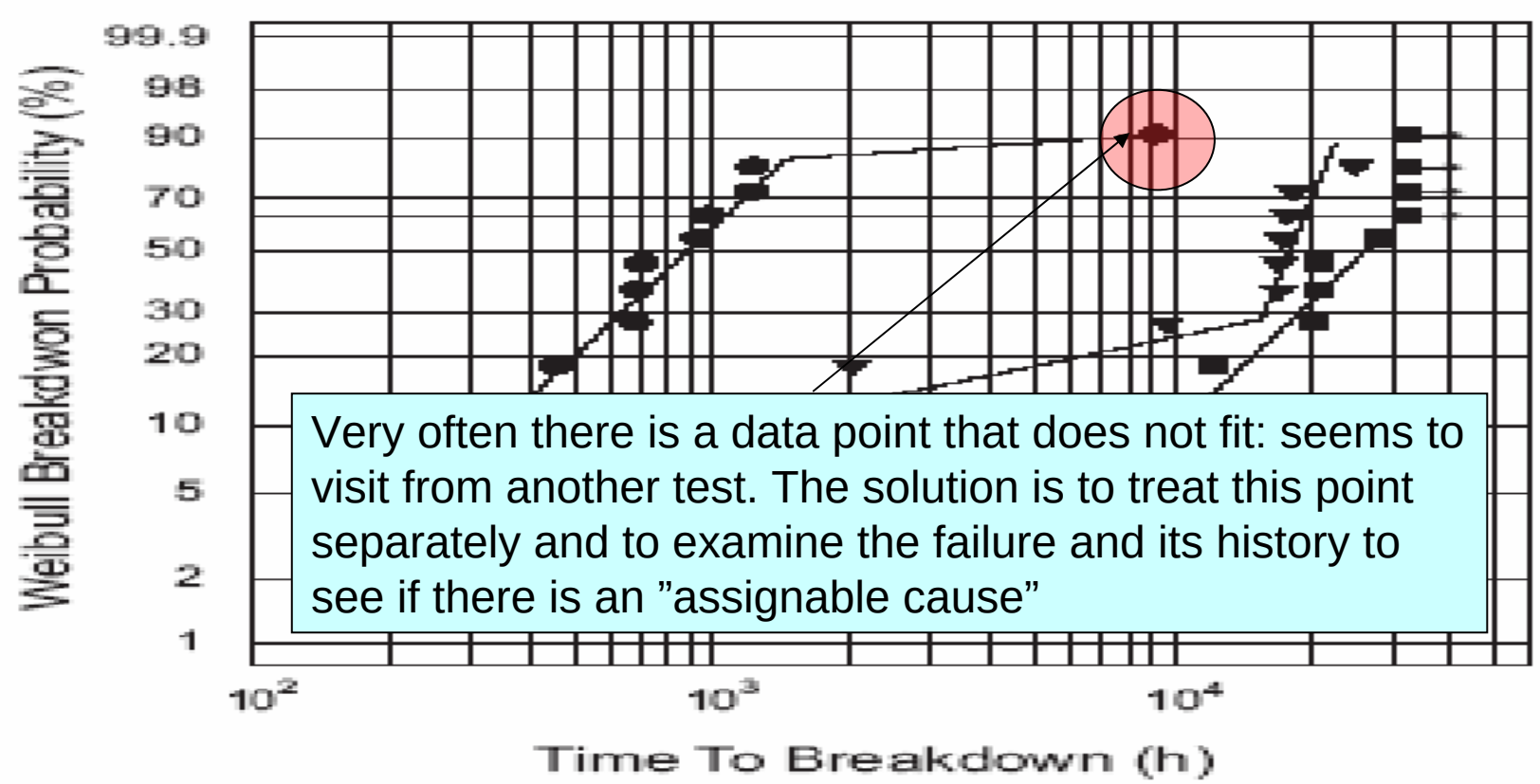
Suspensions



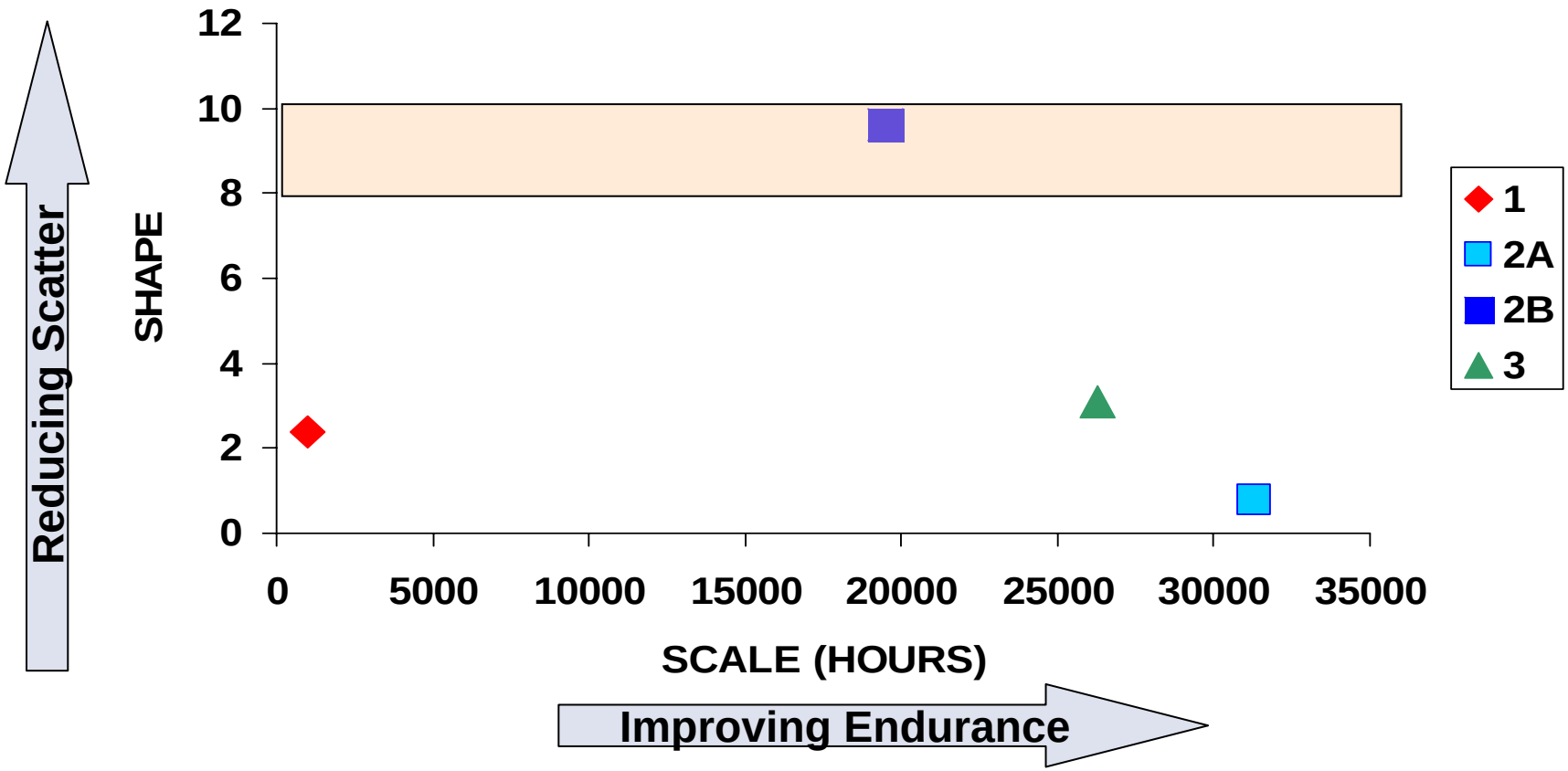
Mixed mechanisms



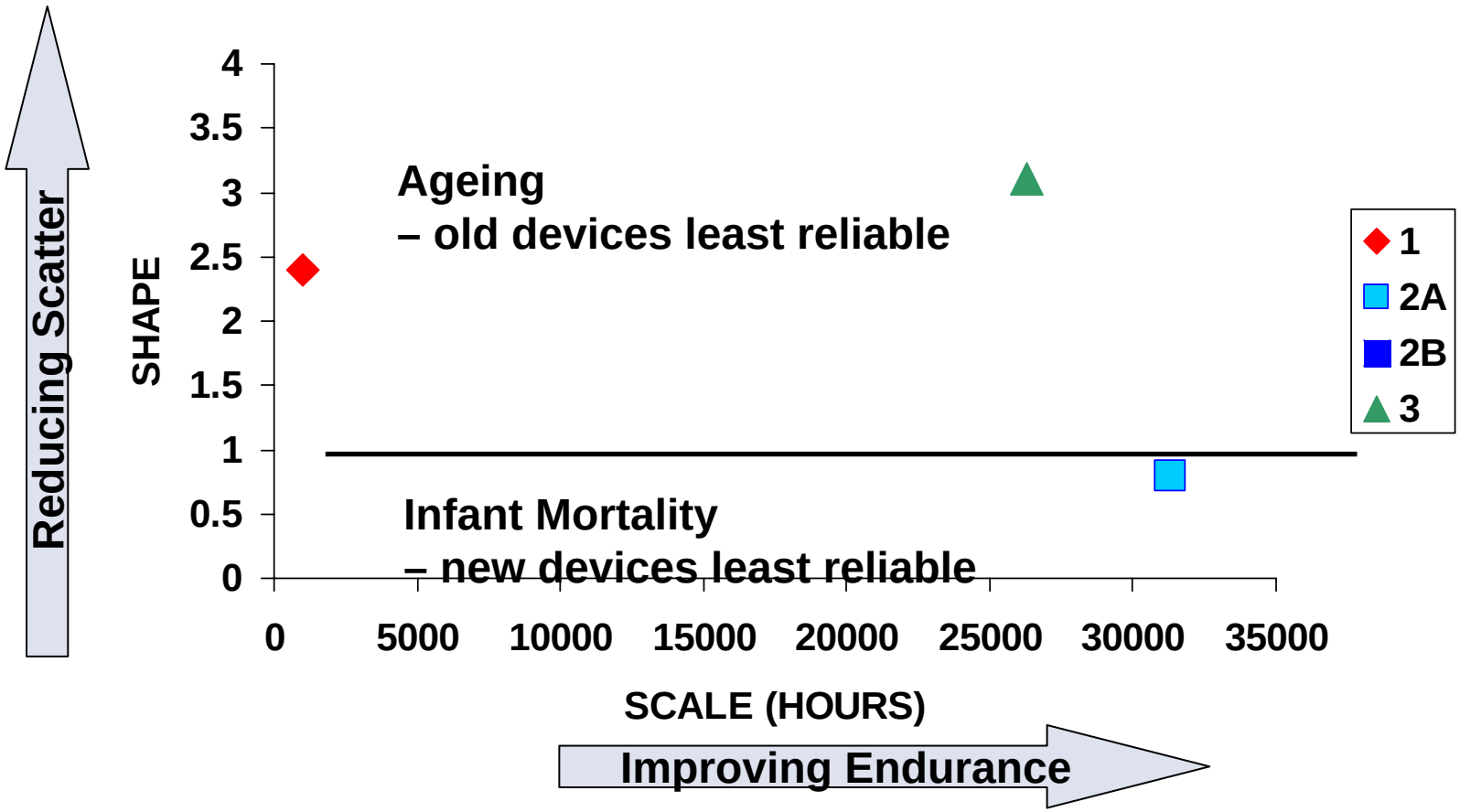
Odd point - ”visitor”



Representing scale & shape

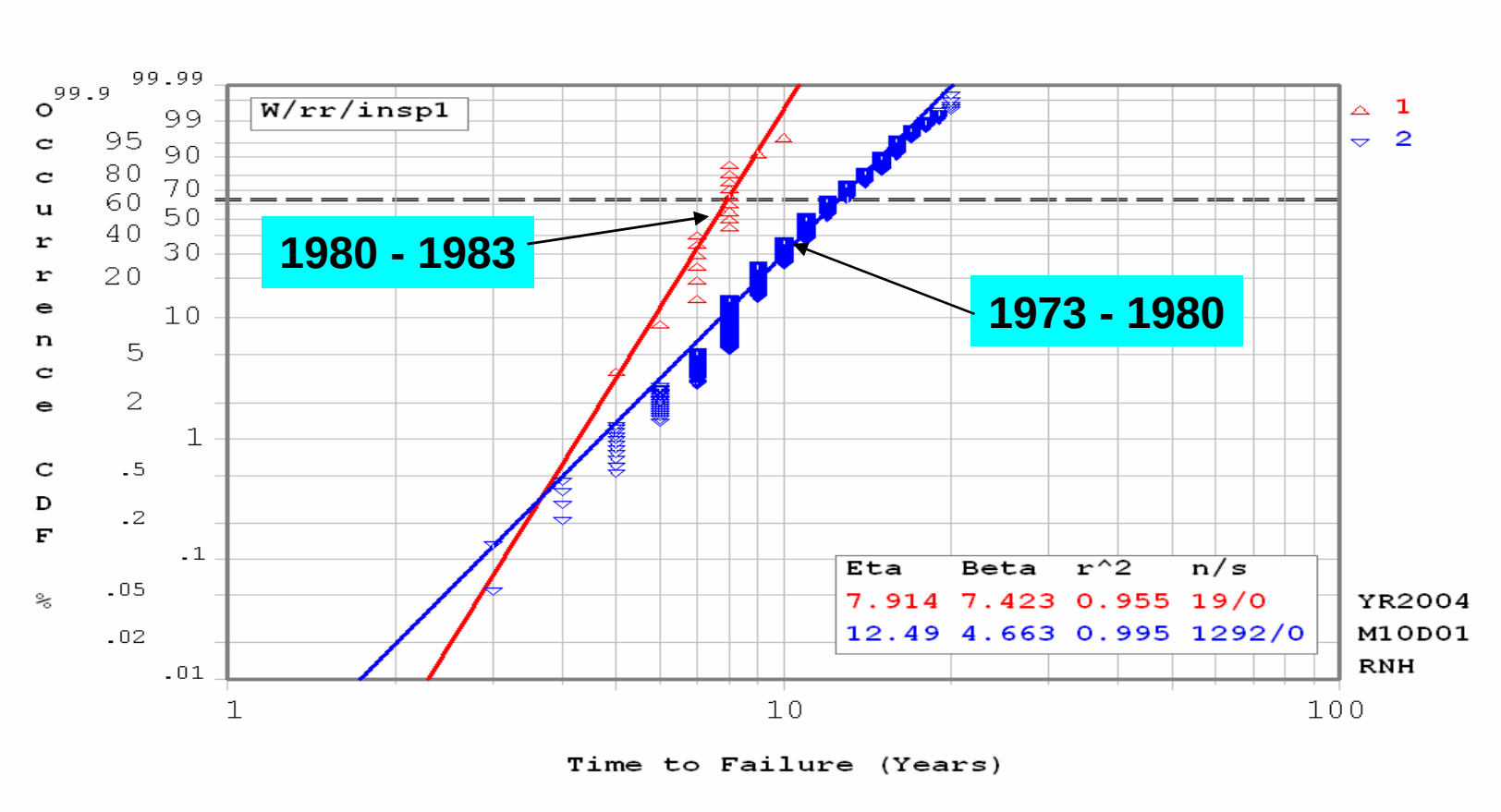


Representing scale & shape



Nodak Electric Service Failure data

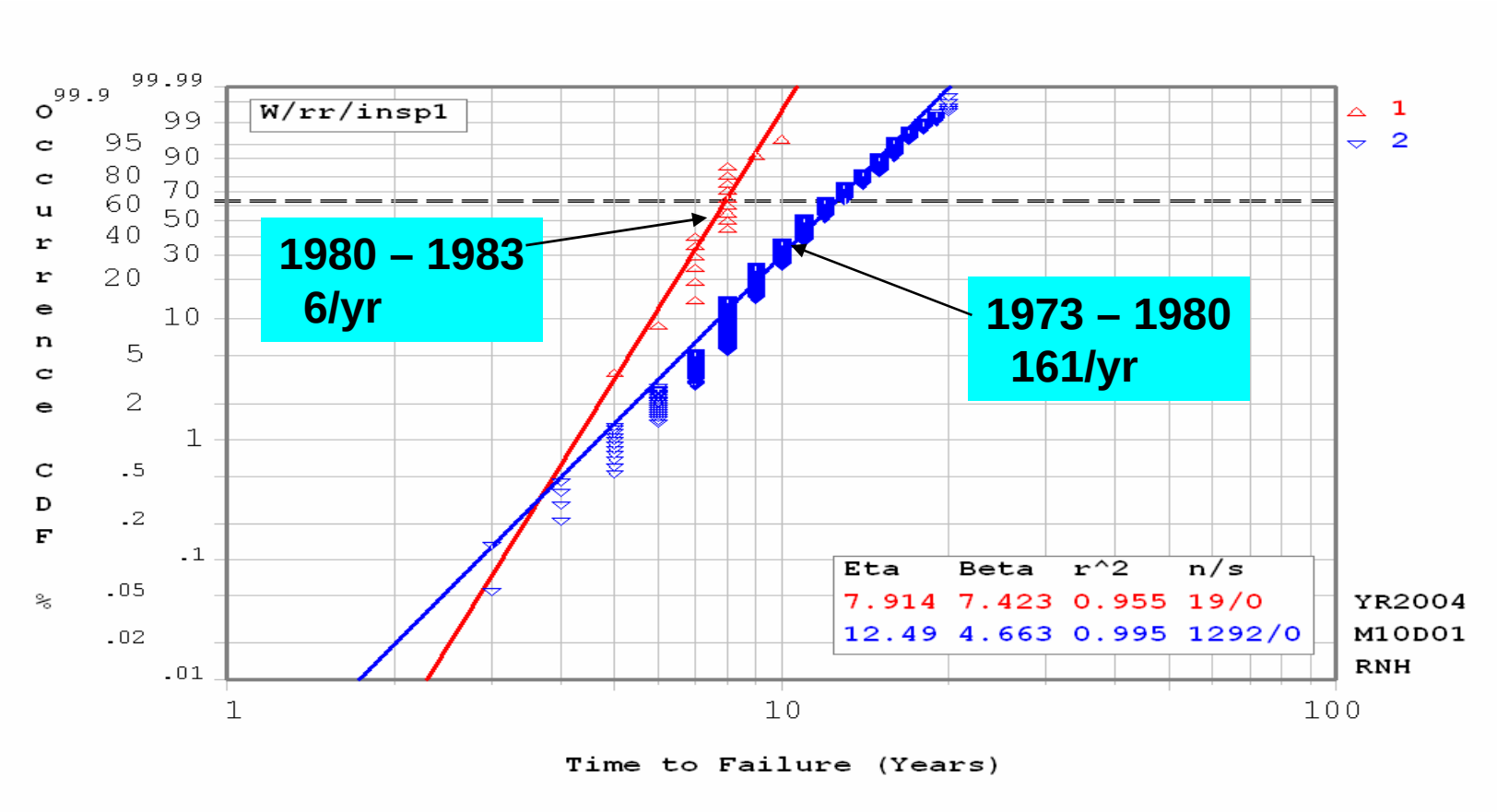
as recorded in database – database only records failures



Data courtesy of John Rodgers

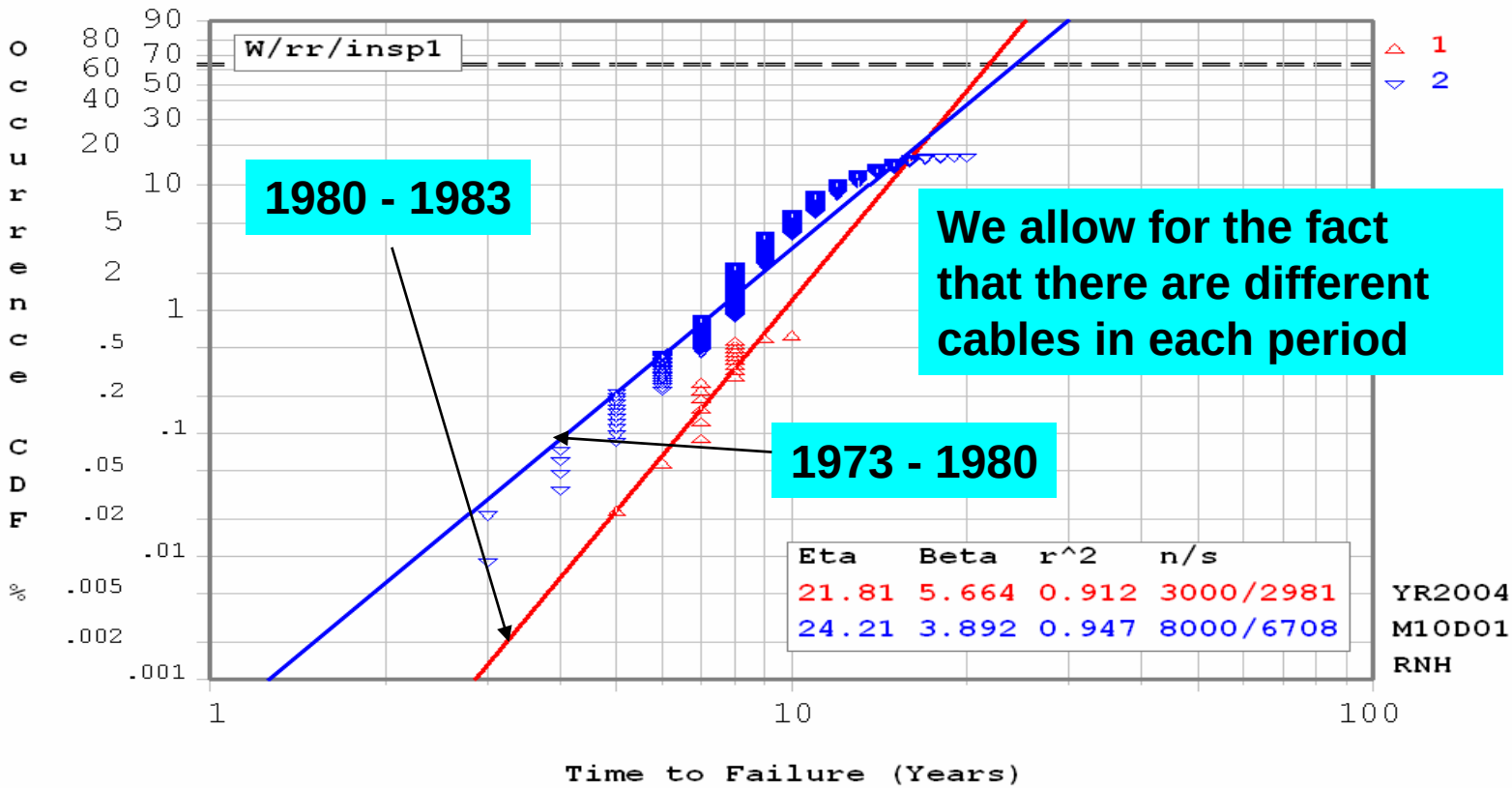
Nodak Electric Service Failure data

as recorded in database – database only records failures



Data courtesy of John Rodgers

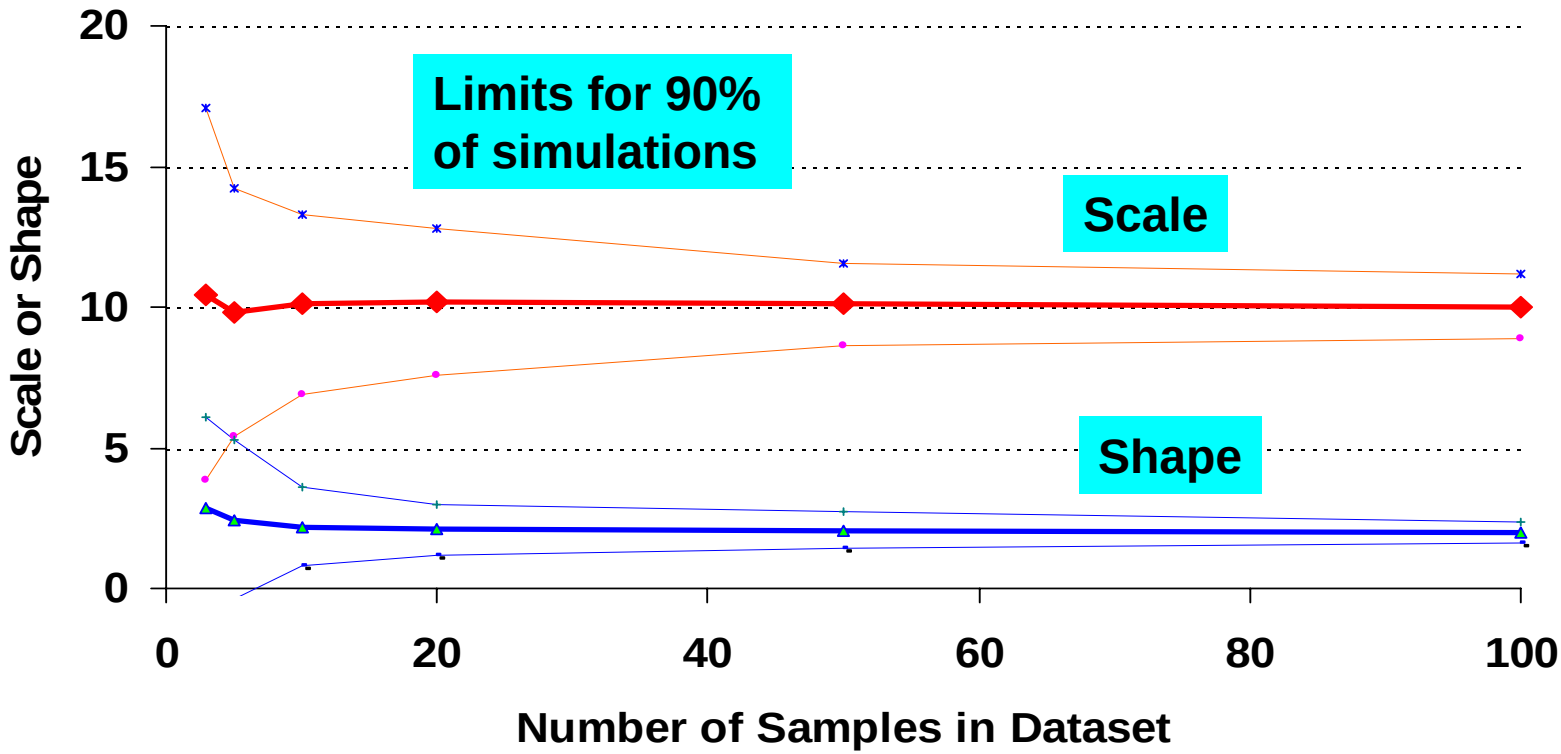
Nodak Electric – looking at the whole population



Number of samples – bias & accuracy

- The larger the number of samples
  - Easier to spot ”Mixed Mechanisms” & ”Visitors”
  - Better accuracy
  - Reduced Bias: a systematic over or under estimate of the parameter

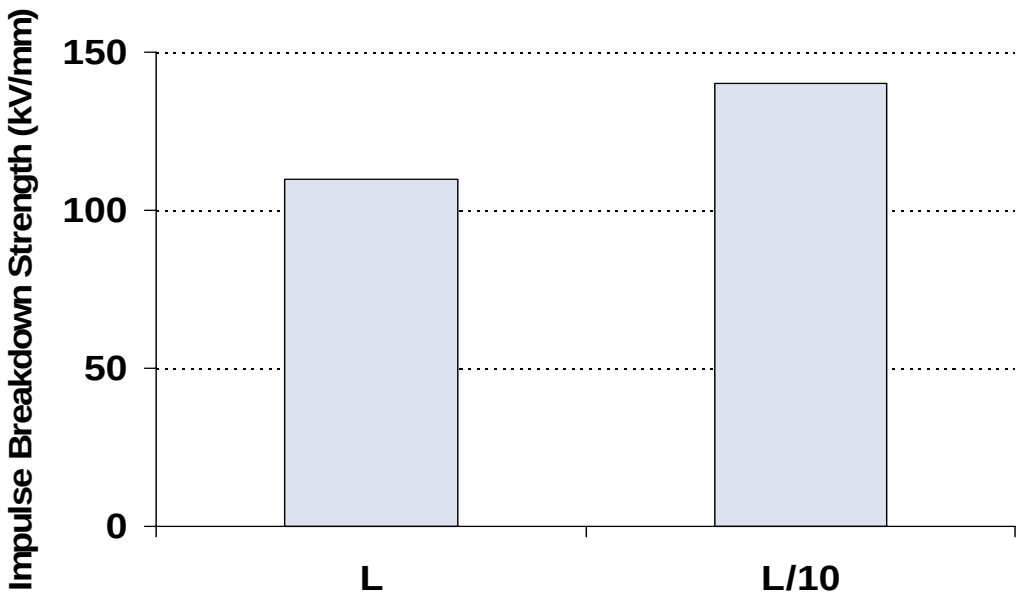
Scatter and sample size



Monte Carlo Simulations – 1000, Scale =10, Shape = 2

Size matters

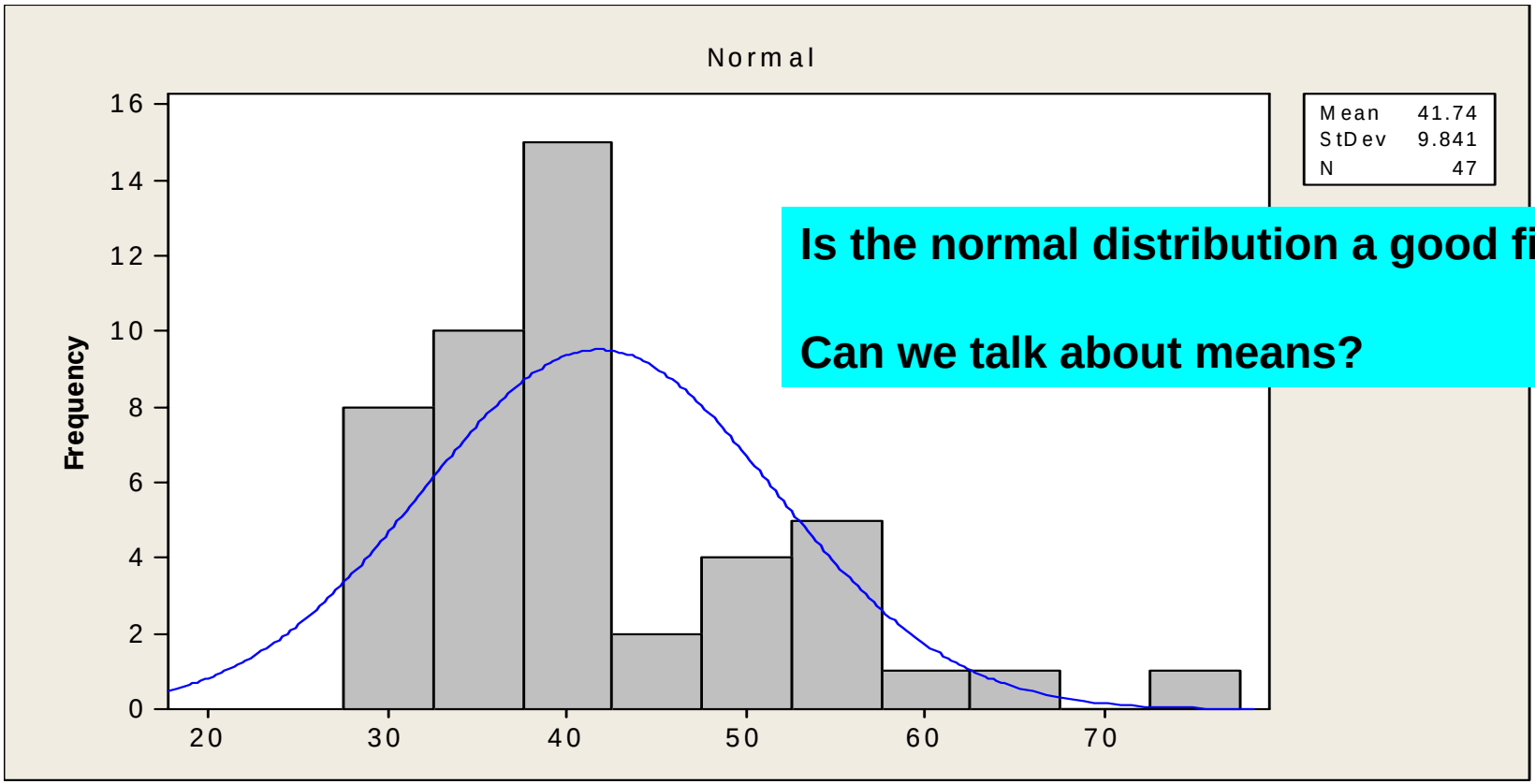
- Longer lengths, larger volumes, more surface will reduce the level of scale – there is a higher chance to find defects



Non parametric methods

- Fast
- Robust
- Simple – can be done by hand or in a spreadsheet
- Work best with segregated data
- No judgement is made at the start about the statistics or the distribution of the problem

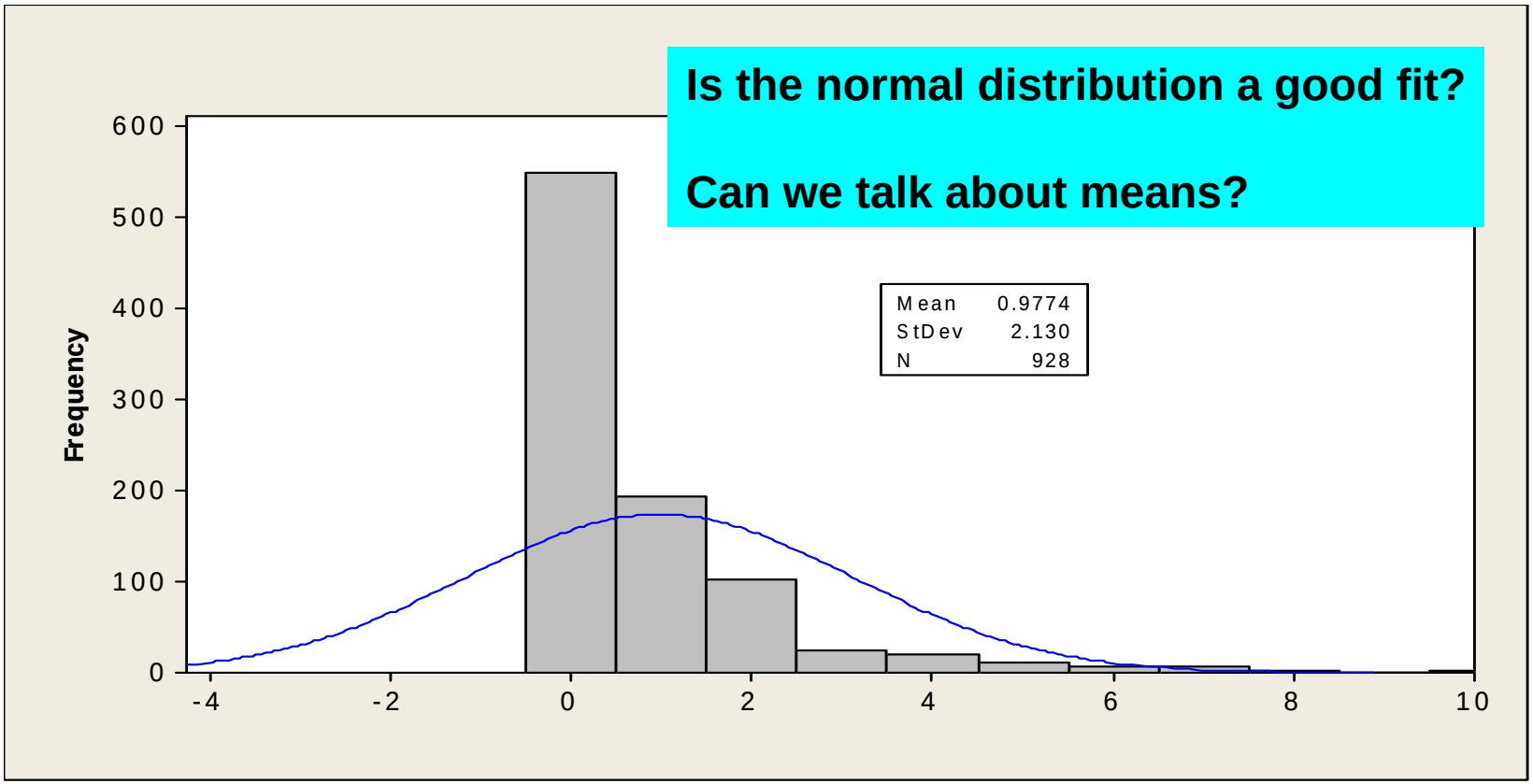
Strip Force Measurements



Is the normal distribution a good fit?

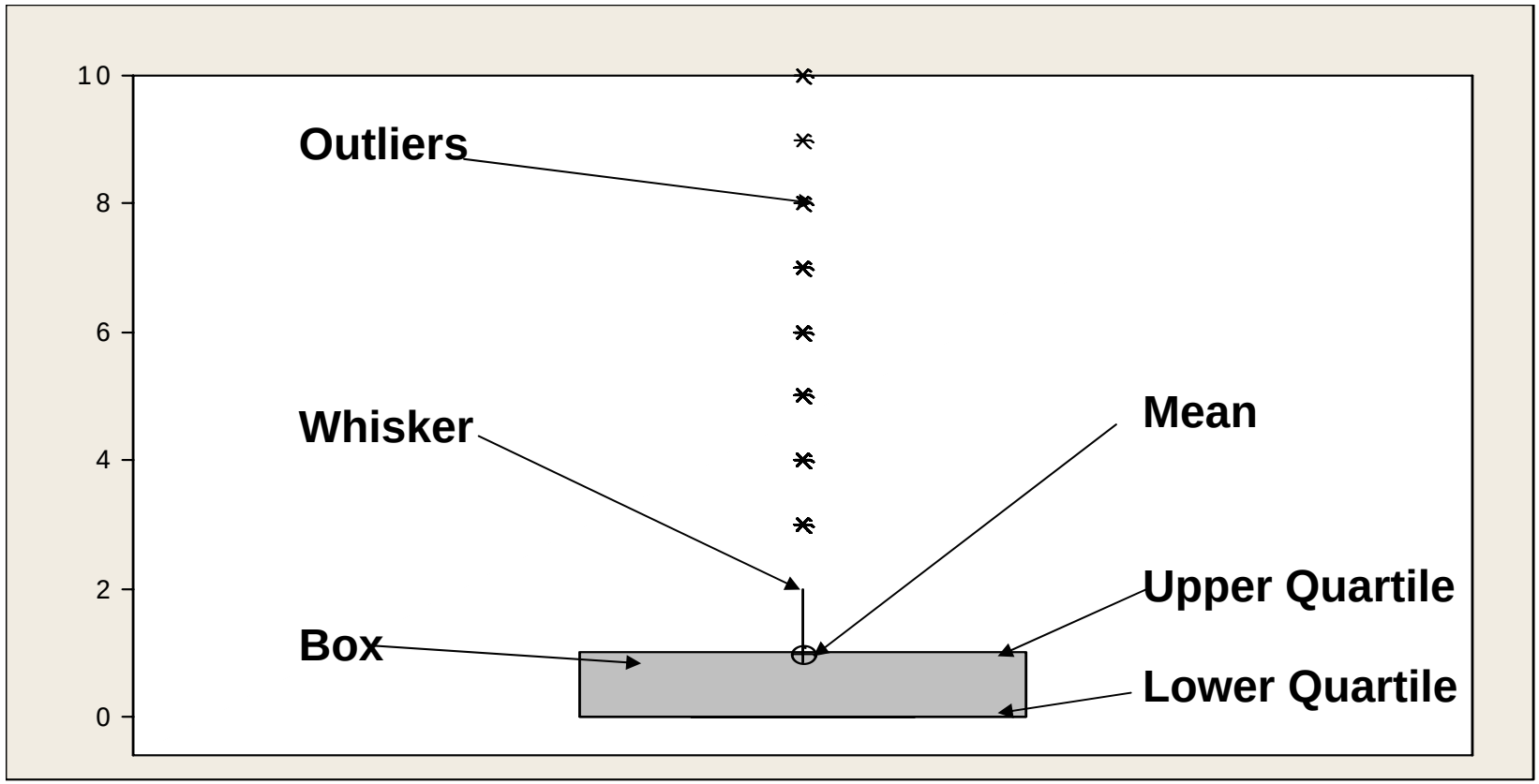
Can we talk about means?

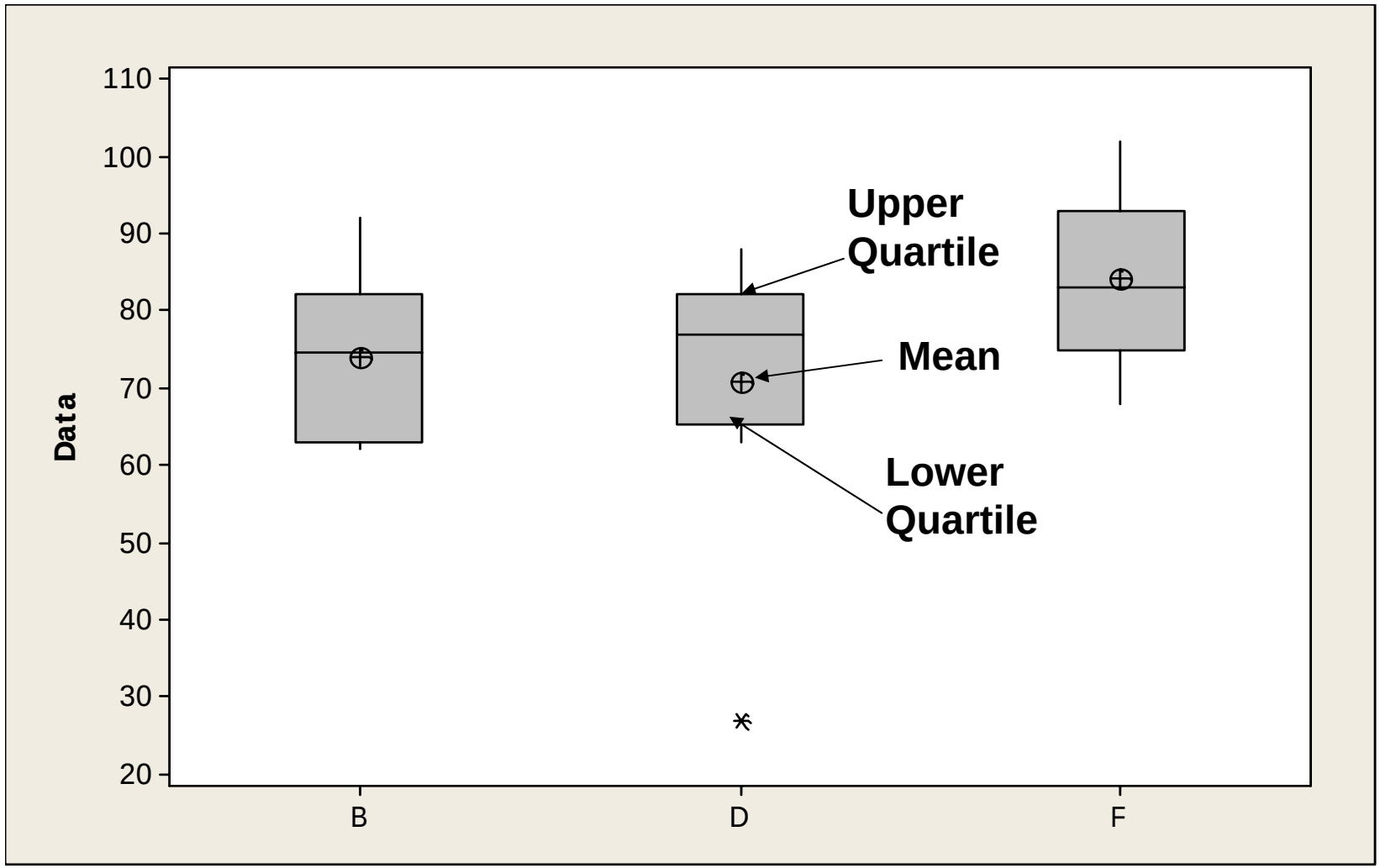
Cleanliness data



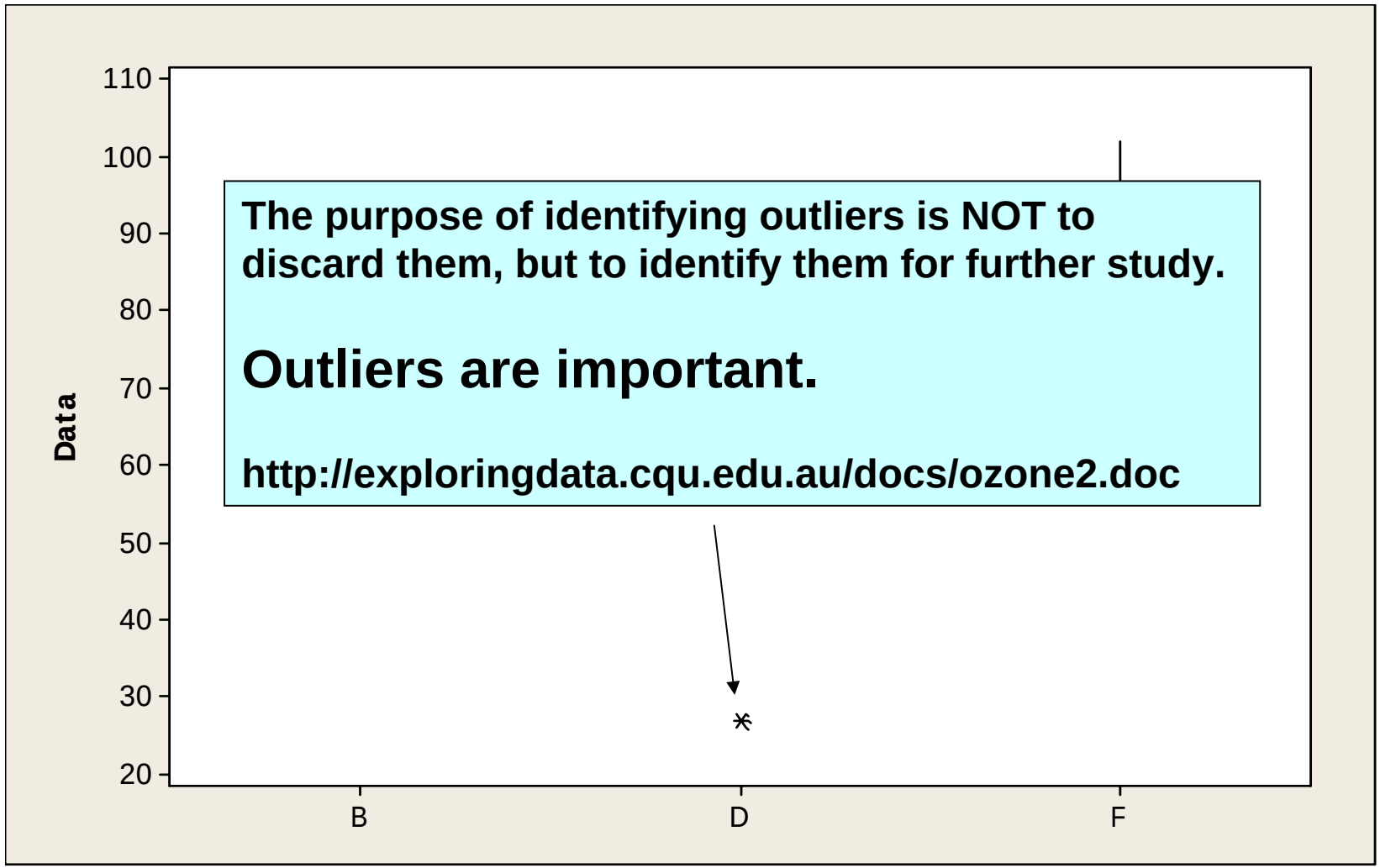
Concentration of contaminants

A Box and Whisker Plot – with outliers





Julian & Hampton; Distribution 2000 Sydney



Julian & Hampton; Distribution 2000 Sydney

Are differences statistically significant?

D	F
27	
63	
66	
66	
75	68
78	68
78	82
82	82
83	84
88	
	91
	92
	96
	102

There is a non parametric test for two data sets of any size. Count the number of values in the first set before the first failure in the second. Count the number of values in the second set after the last failure in the first. If the sum is >6 then there is a significant difference.

Are differences statistically significant?

D	F
27	
63	
66	
66	
75	68
78	68
78	82
82	82
83	84
88	
	91
	92
	96
	102

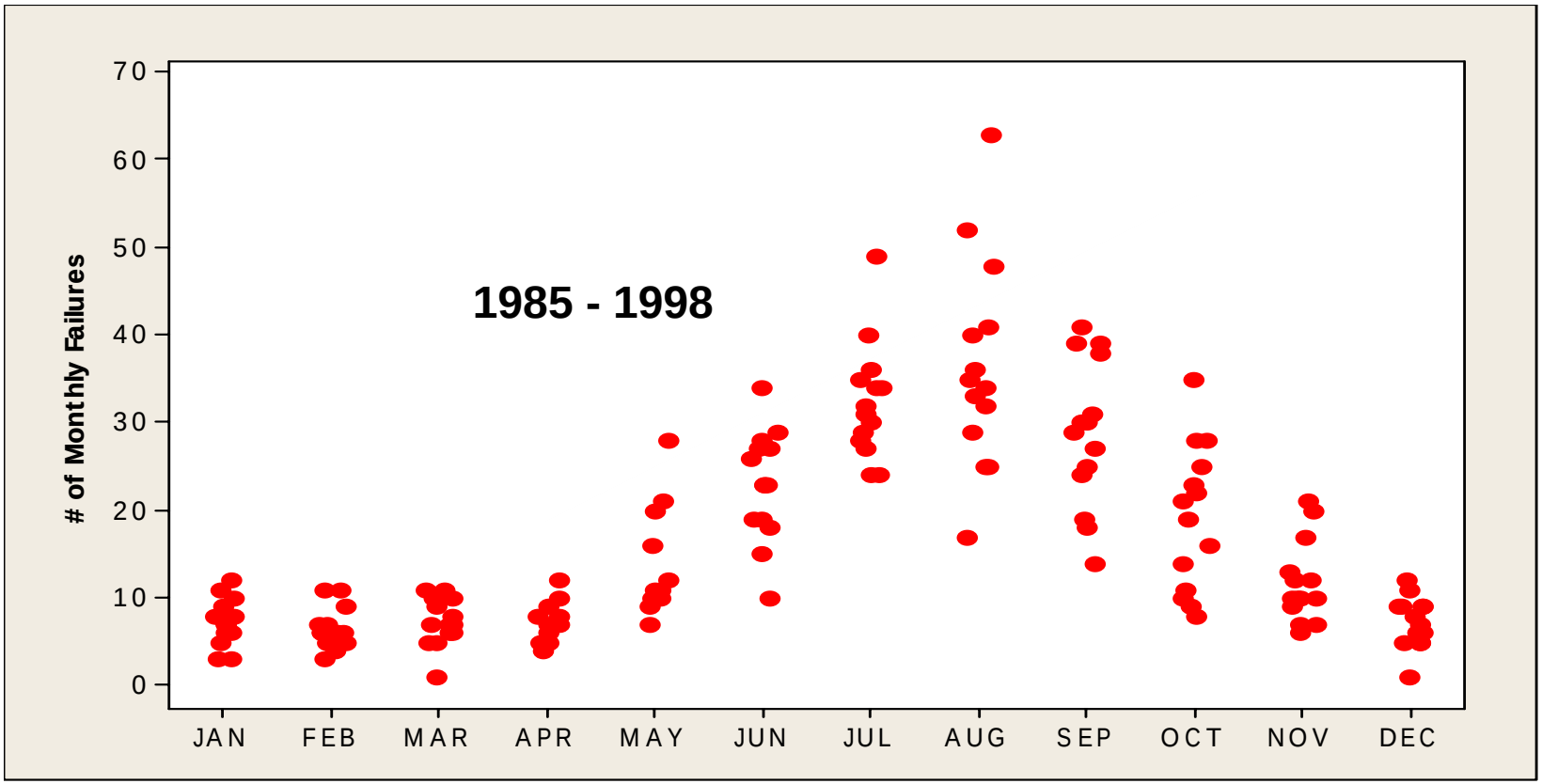
4

4

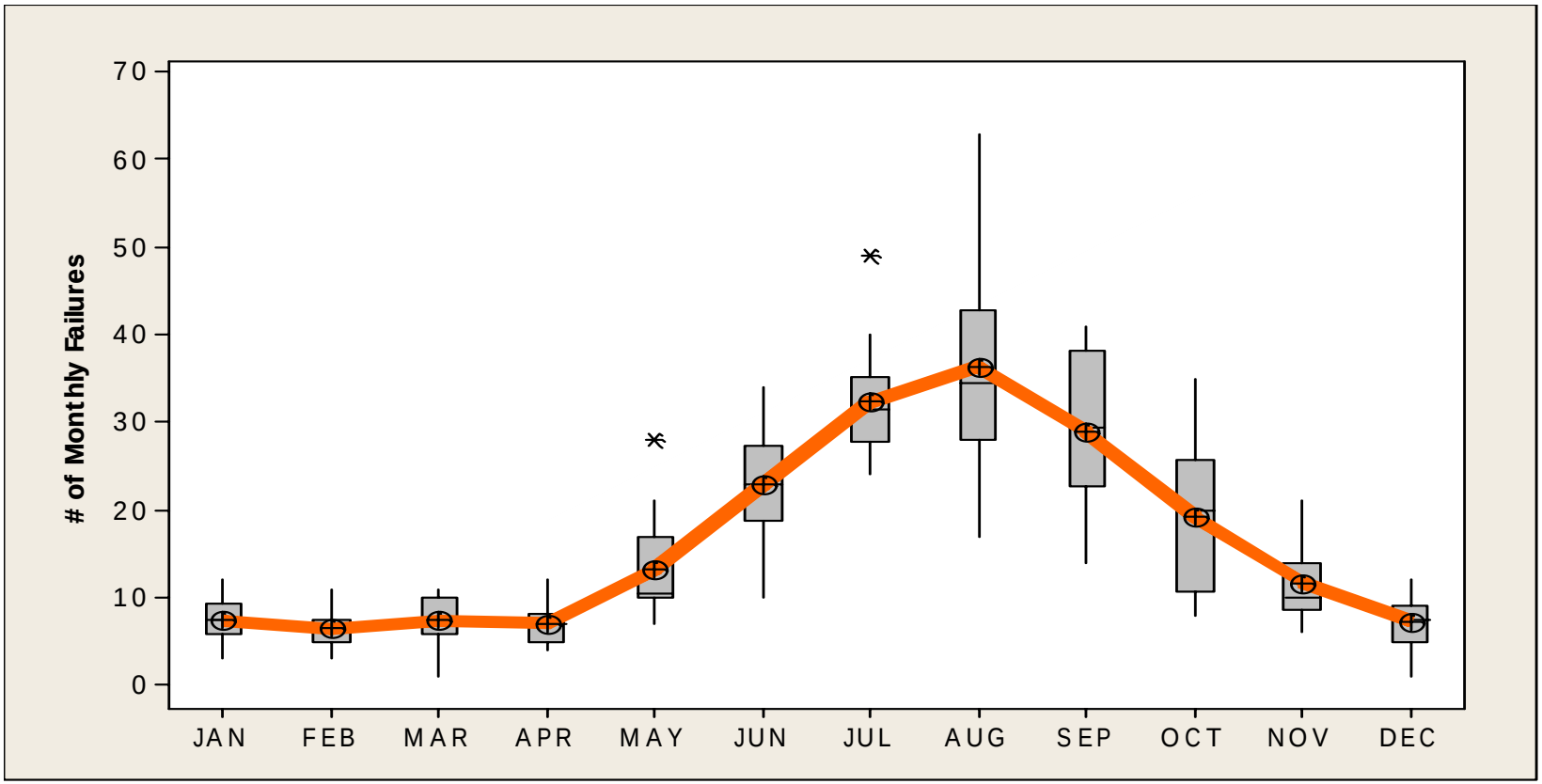
4 + 4 = 8 > 6

Therefore F is significantly better than D

Monthly Failures of Nodak



Monthly Failures of Nodak

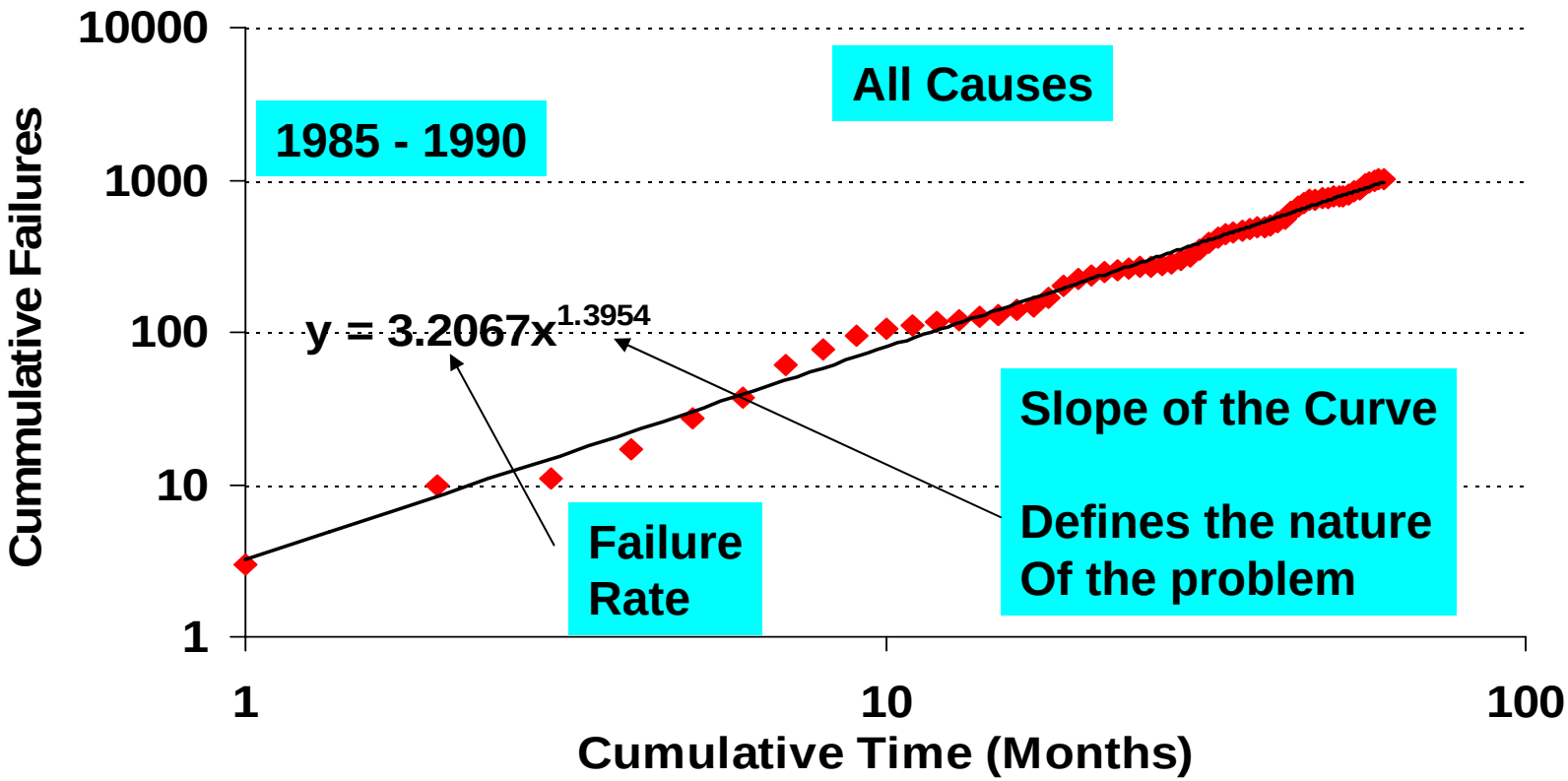


Reliability Growth – Crow AMSAA

- Fast & Simple
- Works when there are a lot of data, can cope with missing information
- Works with mixed failures
- Shows trends very clearly

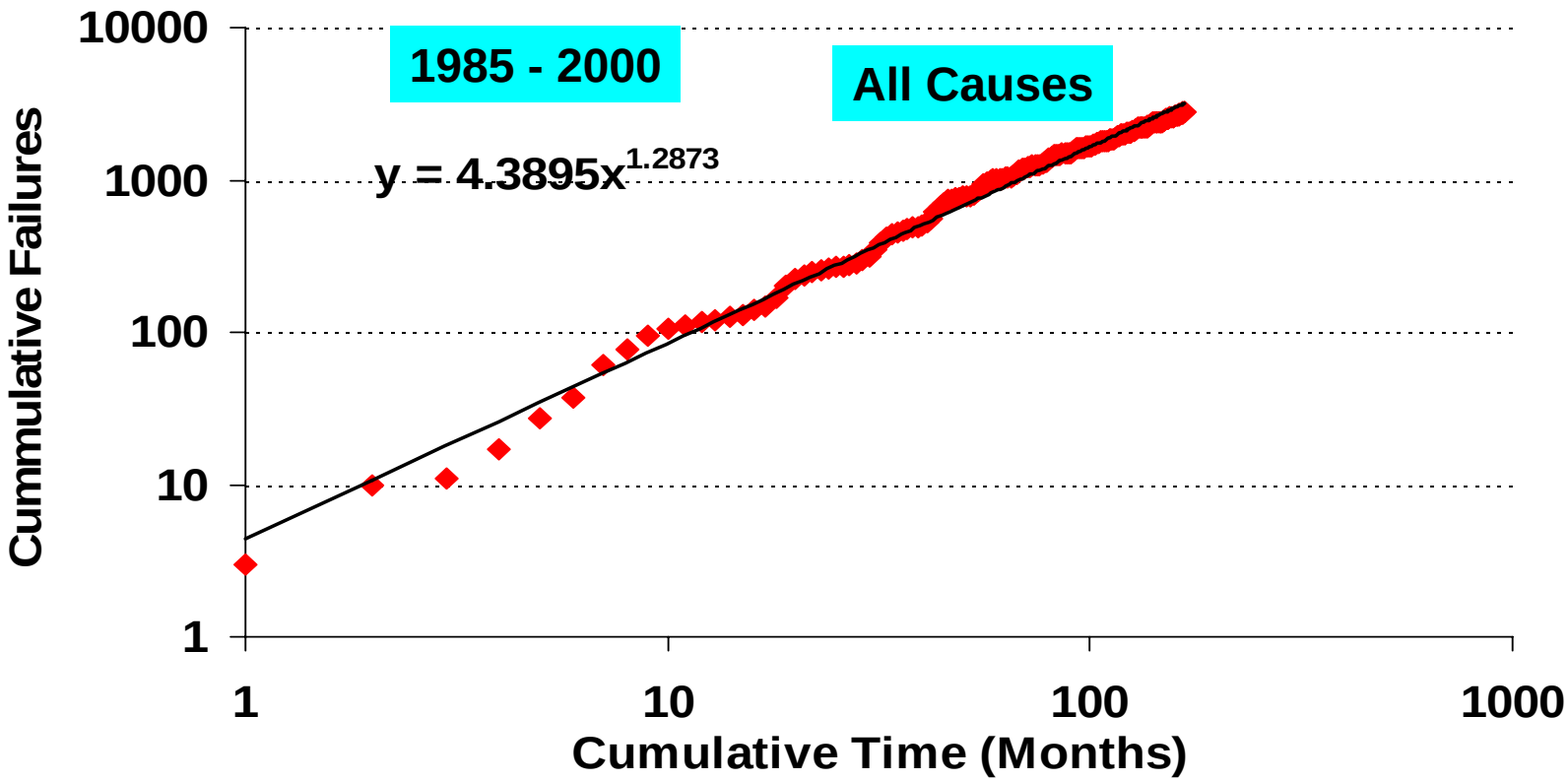
- Graph Cumulative Extent (Time, Number, Length) against Cumulative Occurence of Feature (Failure, Defect). Data are plotted as Log Log plots

Nodak Cable Failure Data

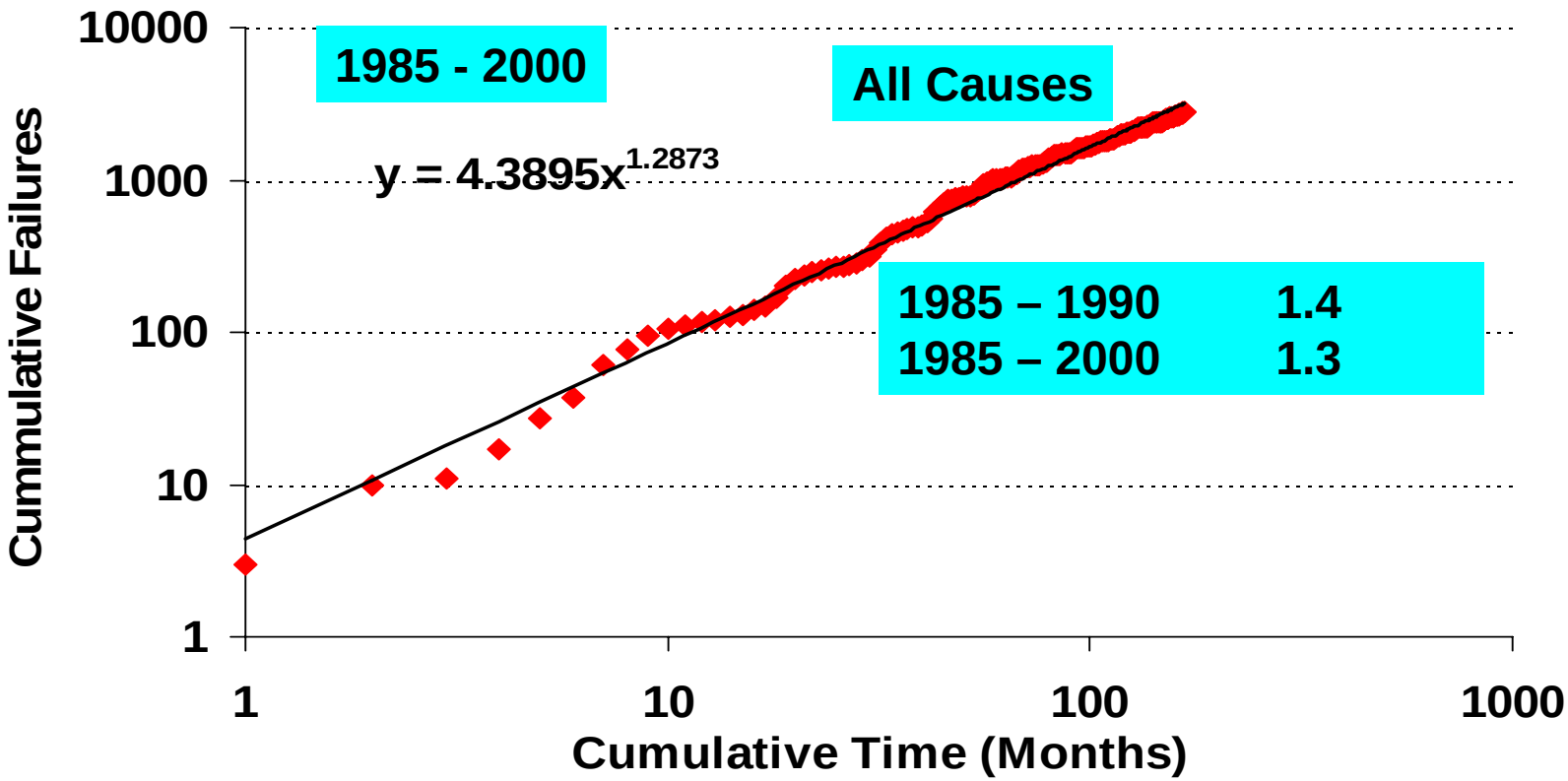


Data courtesy of John Rodgers

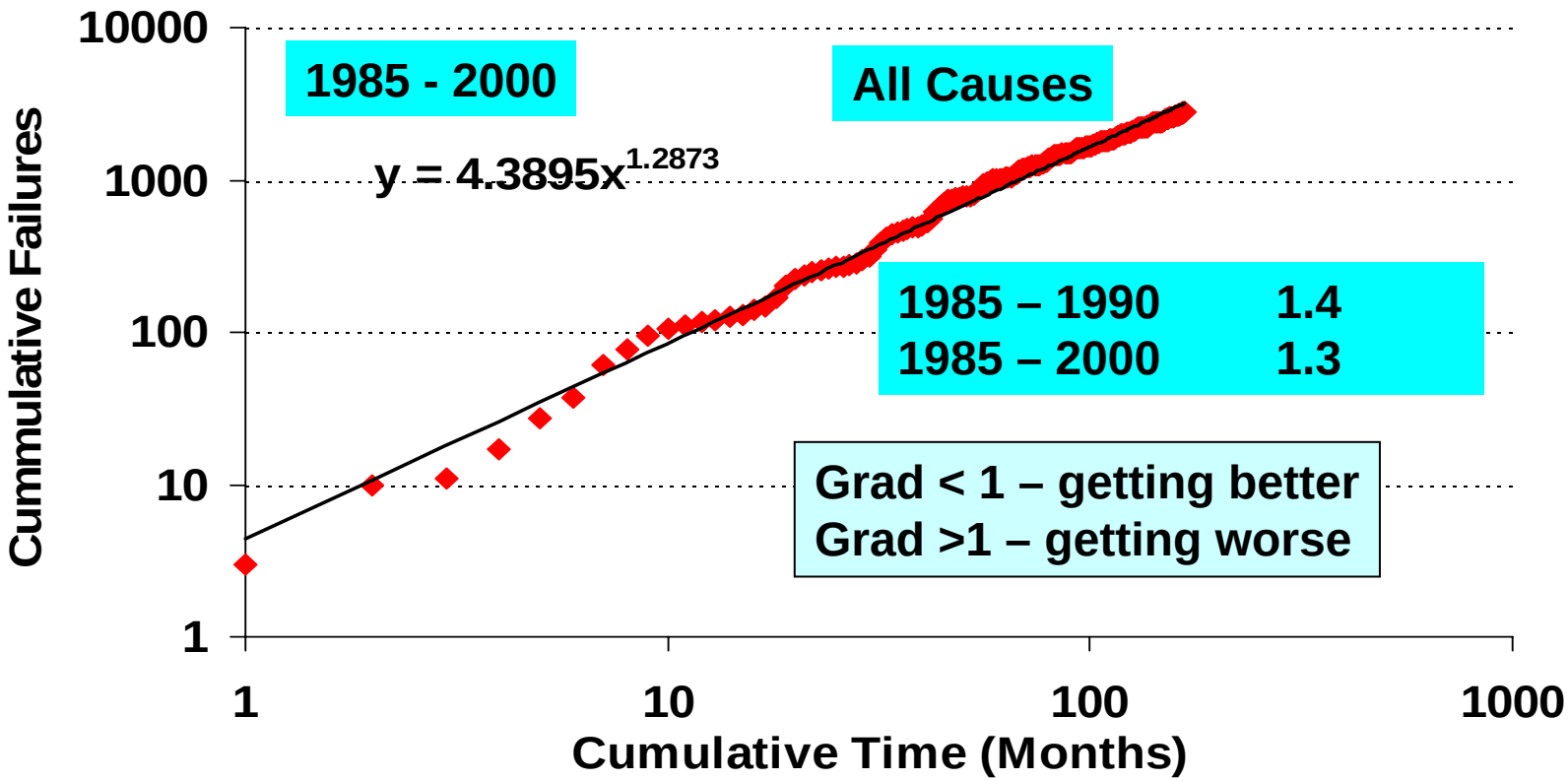
Nodak Cable Failure Data



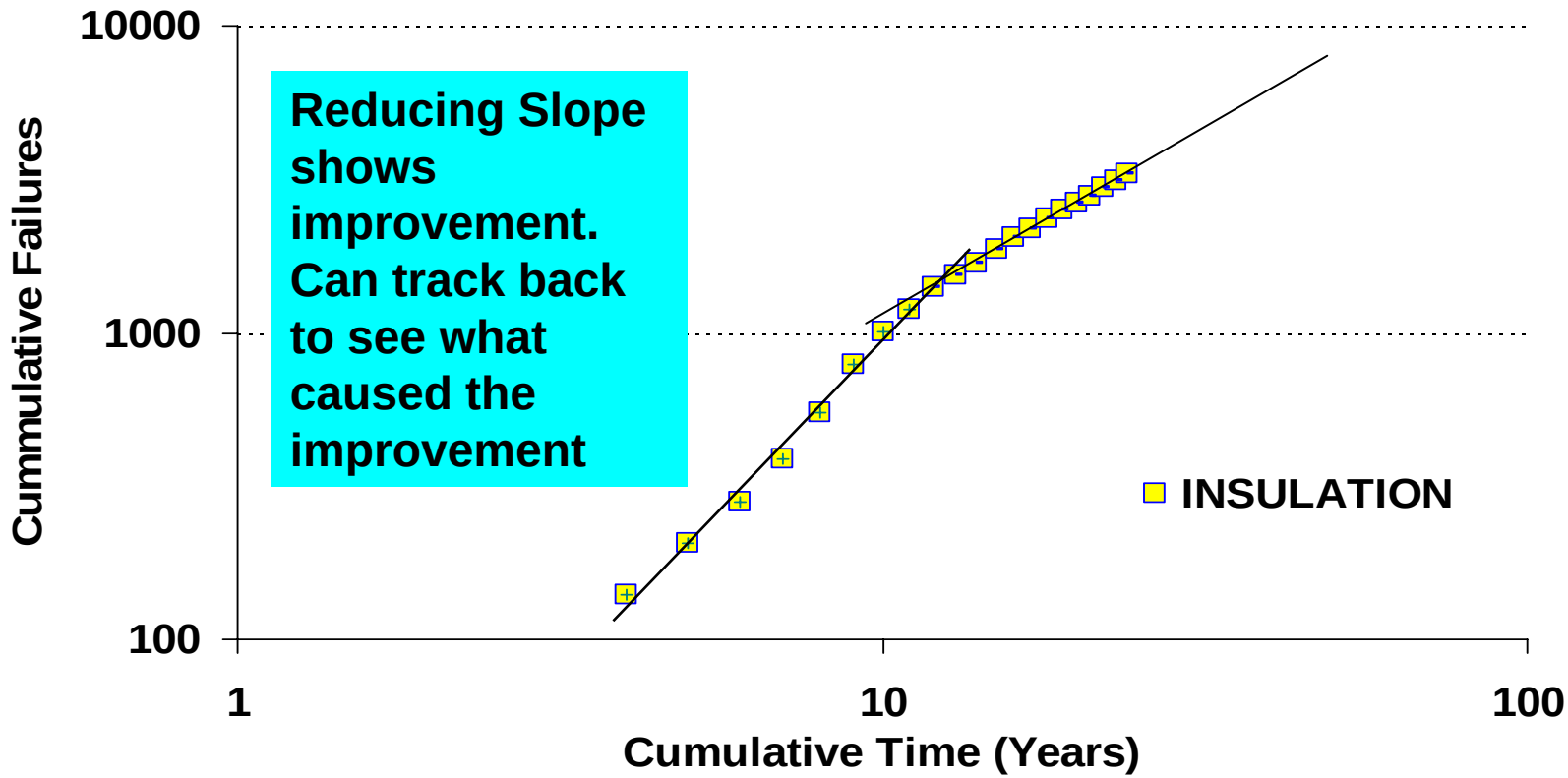
Nodak Cable Failure Data



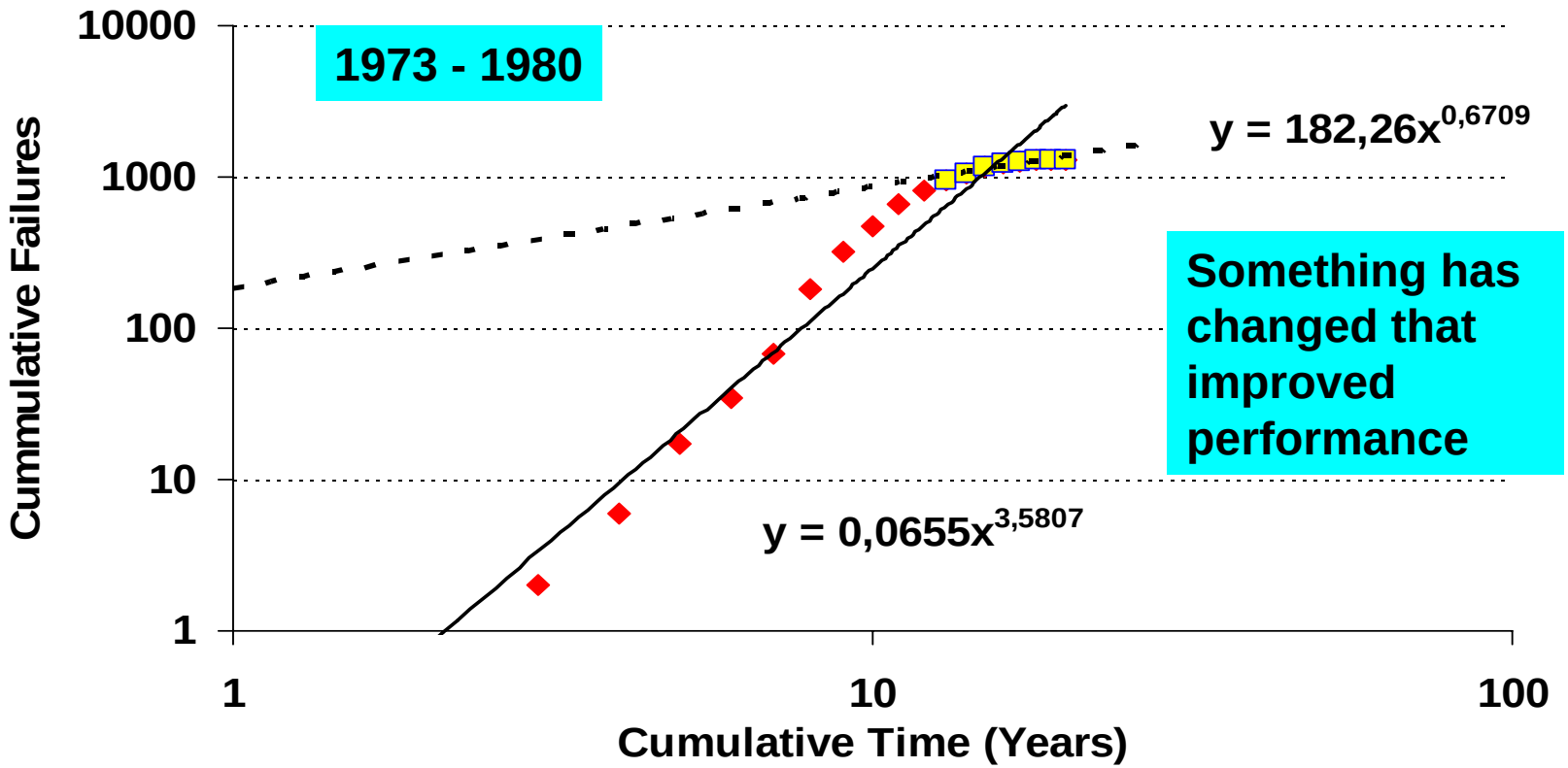
Nodak Cable Failure Data



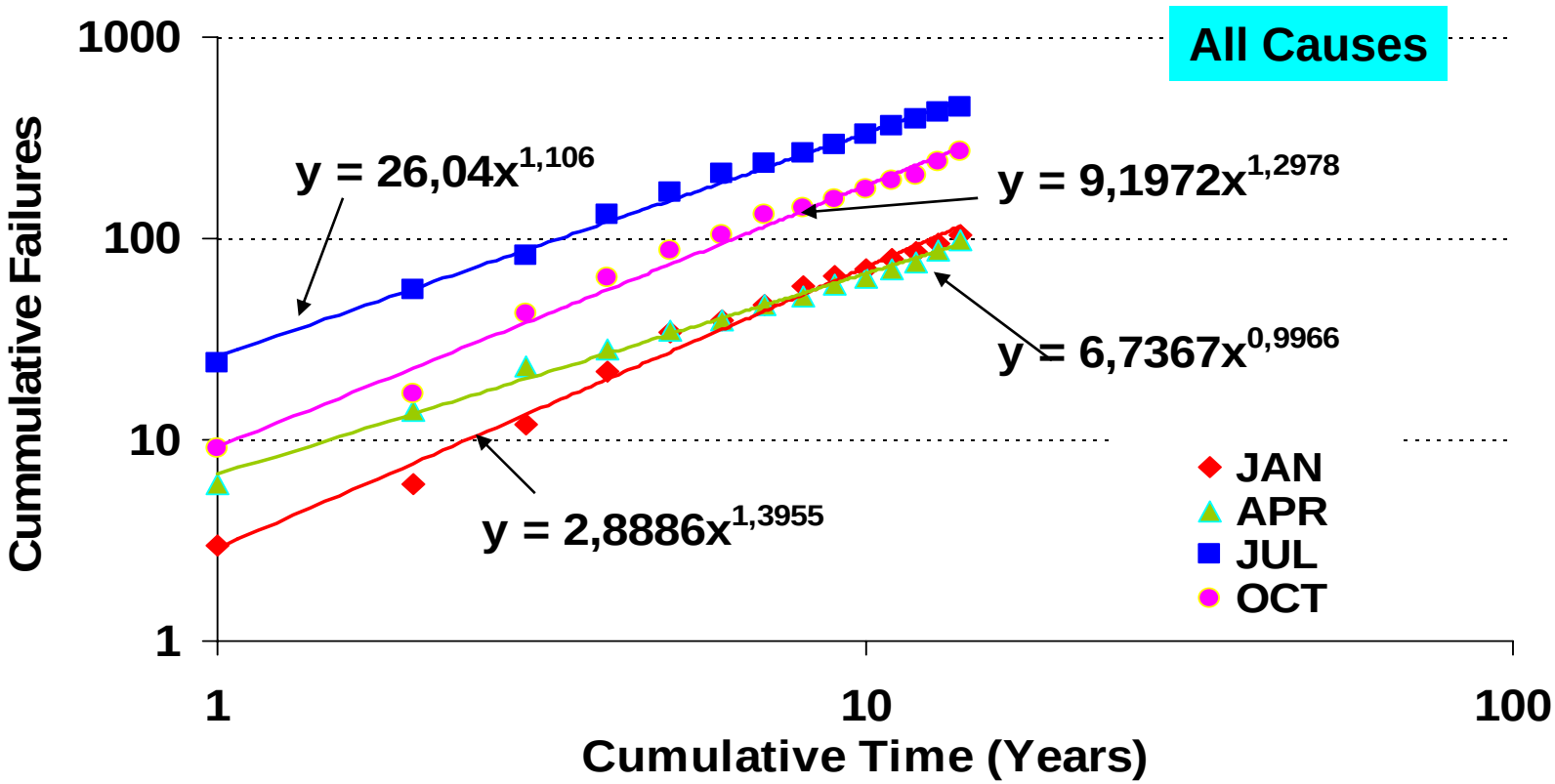
Spotting Changes



Nodak Insulation Data – an alternative to Weibull

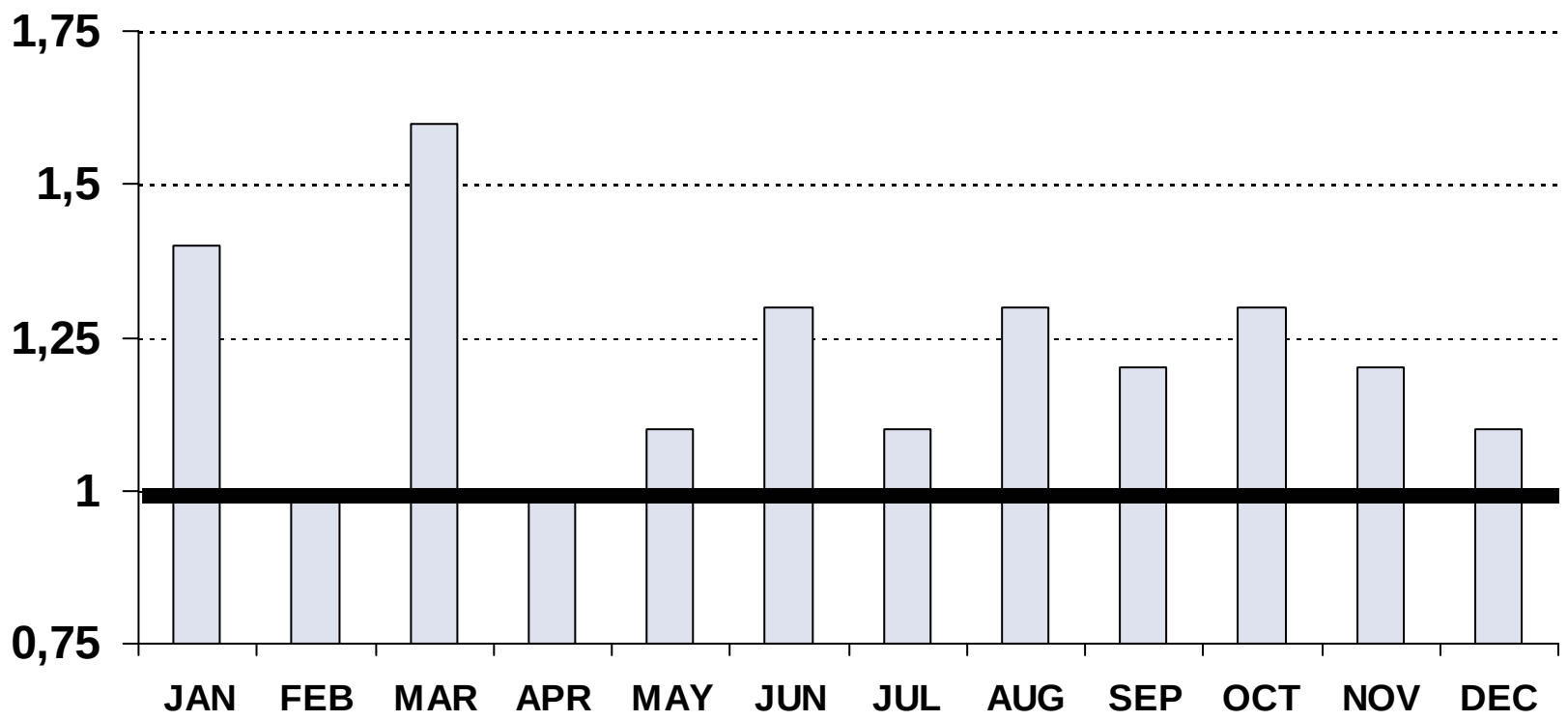


Nodak Cable Failure Data

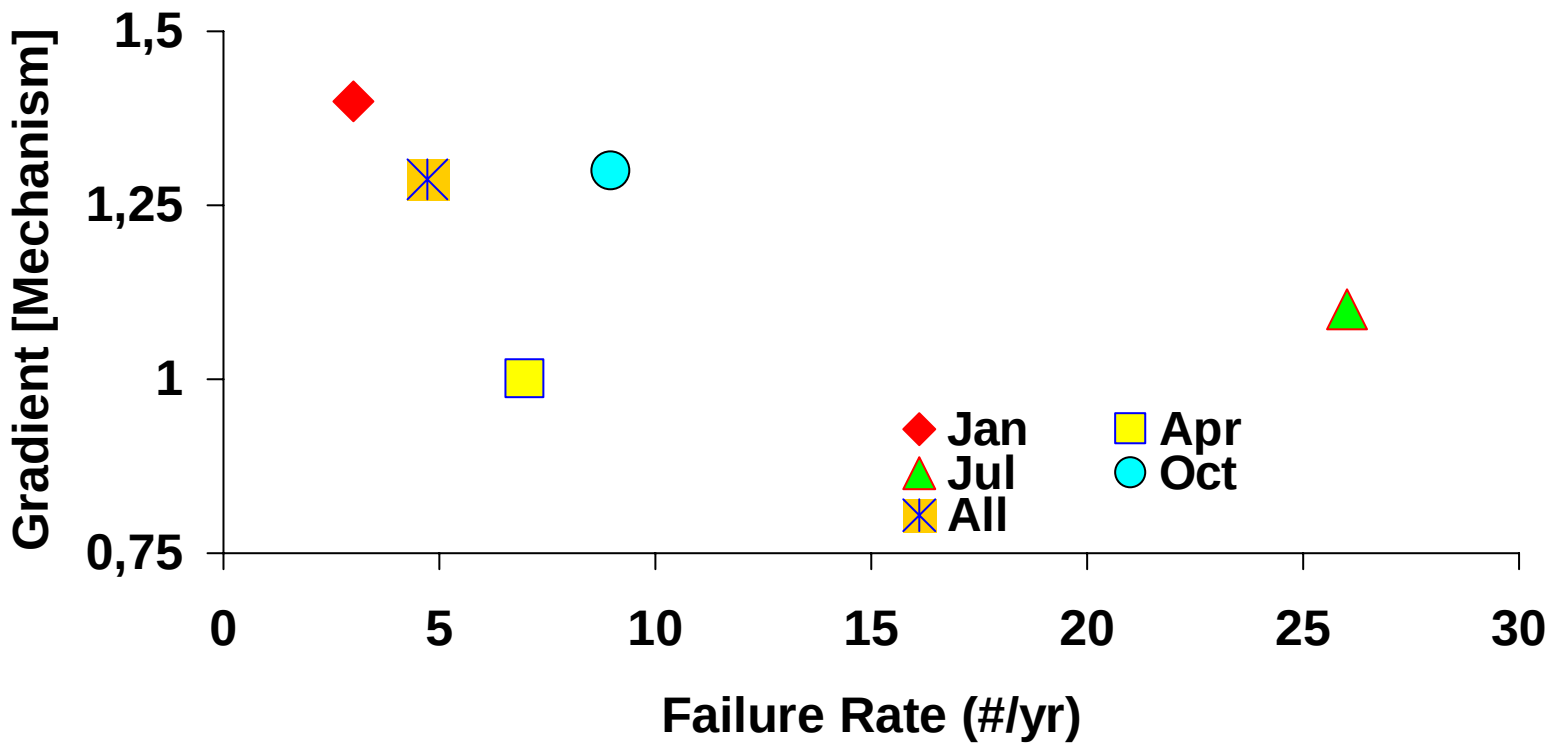


Data courtesy of John Rodgers

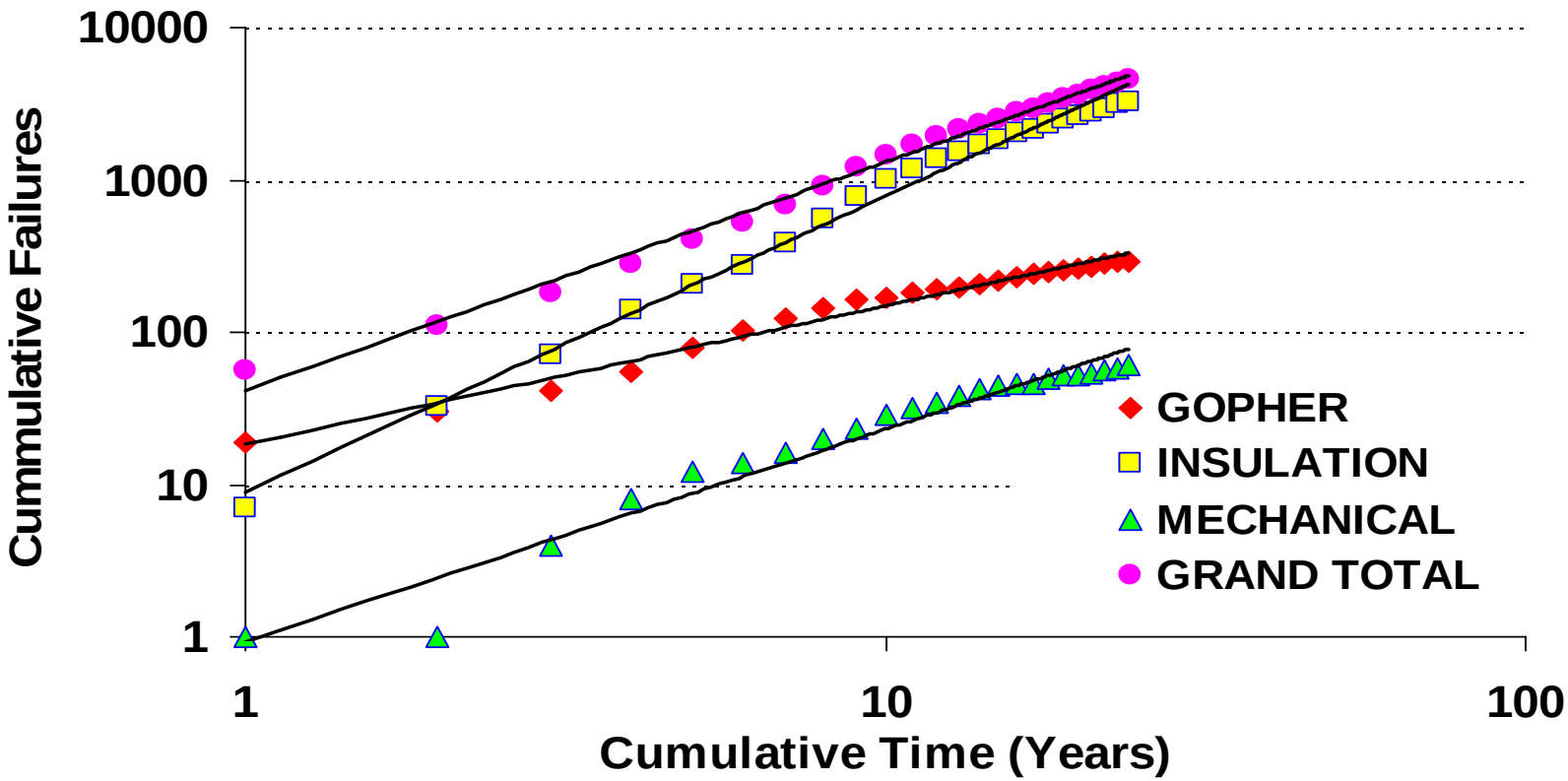
Failure Mechanisms for different months



Failure rate and mechanism

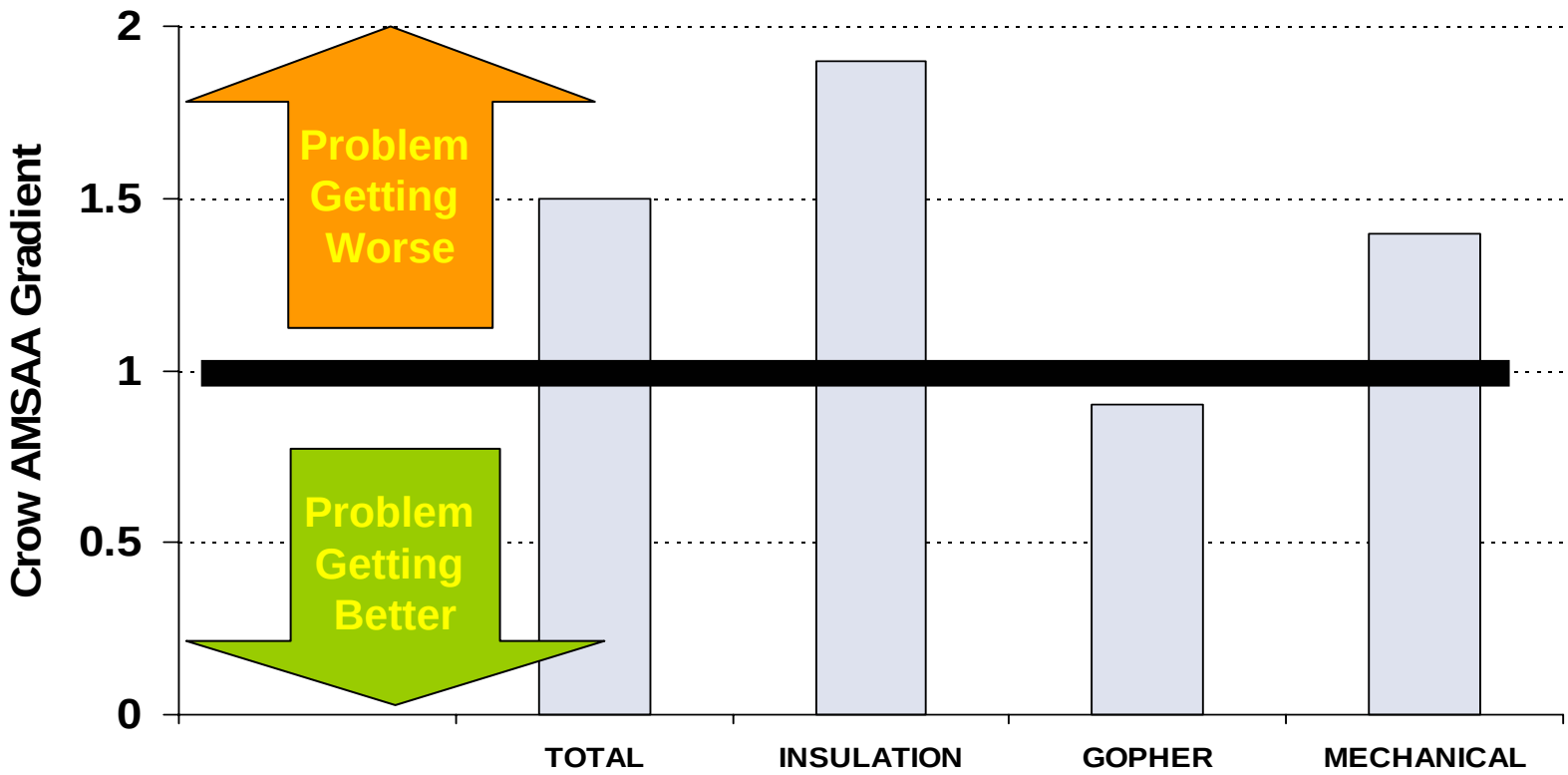


Nodak Cable Failure Data



Data courtesy of John Rodgers

Interpreting the gradients



Conclusion

- Think about the question / problem
- Go and carefully catch your data
- When testing cables: test as many of the longest that you can afford
- Look at the data from a variety of viewpoints
- Investigate all of the clues that the data give you

Further sources of information

Weibull

- http://www.barringer1.com/pdf/Chpt1.pdf
- Accelerated long term evaluation of MV XLPE cables; Benjaminsen et al; Jicable03 – cable 10
- http://www.qualitydigest.com/jan99/html/body_weibull.html
- M Abou Dakku, A Bulinski & S Bamji; Corelation between space charge & breakdown in polymer insulations under DC; ICSD2004 Toulouse
- RN Hampton; Weibull analysis – dealing with ”real” data; Fall ICC 1997

Reliability Growth

- http://www.barringer1.com/nov02prb.htm
- RN Hampton; Estimating system reliability from utility fault data; Fall ICC 1999

Non Parametrics

- http://exploringdata.cqu.edu.au/boxplots.htm
- Julian & Hampton; Distribution 2000 Sydney

Nodak data courtesy of John Rodgers

Thanks for your attention

Any questions??