Reliability Data Analysis Experiences

William Q. Meeker
Department of Statistics
Center for Nondestructive Evaluation
Iowa State University
wqmeeker@iastate.edu

Conference on Rigorous Test and Evaluation for Defense, Aerospace, and National Security

Crystal City, VA
11 April 2016

Overview

- Examples of Reliability Disasters
- Analysis of Field/Warranty Data
  - Solving a reliability problem with statistics: jet engine bleed system failure
  - Using prior information to strengthen data: bearing cage field failure data
  - Correct analysis of data with multiple failure modes: Device-G field-tracking and Shock Absorber data
  - Other examples
- Accelerated Testing
  - Accelerated life tests (Insulating structure)
  - Accelerated repeated-measures degradation tests (LEDs)
  - Accelerated destructive degradation tests (adhesive bond)
  - Other examples

Reliability Disasters

- The general problem
- Some well known classic examples
- Less well known examples
- Causes

Reliability

- ***Probability*** that a system, vehicle, machine, device, and so on, will perform its intended function under ***encountered*** operating conditions, for a specified period of time (Meeker and Escobar 1998)
- Quality over time (Condra 1993)
- A highly quantitative engineering discipline, often requiring complicated statistical and probabilistic analyses
- Customers expect high reliability

A Common Problem

- A product has been introduced in the field.
- A small number of failures are detected. The “failure mode” is usually unanticipated.
- Do we have a serious problem? What fraction of the product population is susceptible to the failure mode? How fast will they fail?
- What is the risk (cost of future failures)?
- How can we fix the problem?
- How can we avoid similar problems in the future?
- Common element: multiple mistakes

Examples of Reliability Disasters

- These are the kind of things all companies want to avoid
- Some Well Known Classic Examples:
  - Challenger and Columbia Space Shuttle failures
  - United Airlines 232 (aka Sioux City accident)
  - Automobile clear-coat failures (early 1990's)
  - Ford Explorer/Firestone tires
  - GE refrigerator compressor
  - Sony laptop battery design defect
  - Toyota gas pedal
  - BP Blowout-protector

Some Lesser Known Examples

- Failing electrolytic capacitor PC mother boards
- ATT Bell Cells
- Appliance and other environmental problems
- Medical device recalls

Sony Battery Problem

Chilling Tale

This is the kind of thing we are trying to avoid

WALL STREET JOURNAL MONDAY, MAY 7, 1990

1

Chilling Tale

GE Refrigerator Woes Illustrate the Hazards In Changing a Product

Firm Pushed Development Of Compressor Too Fast, Failed to Test Adequately

Missing: the 'Magical Balance'

GE Refrigerator Compressor

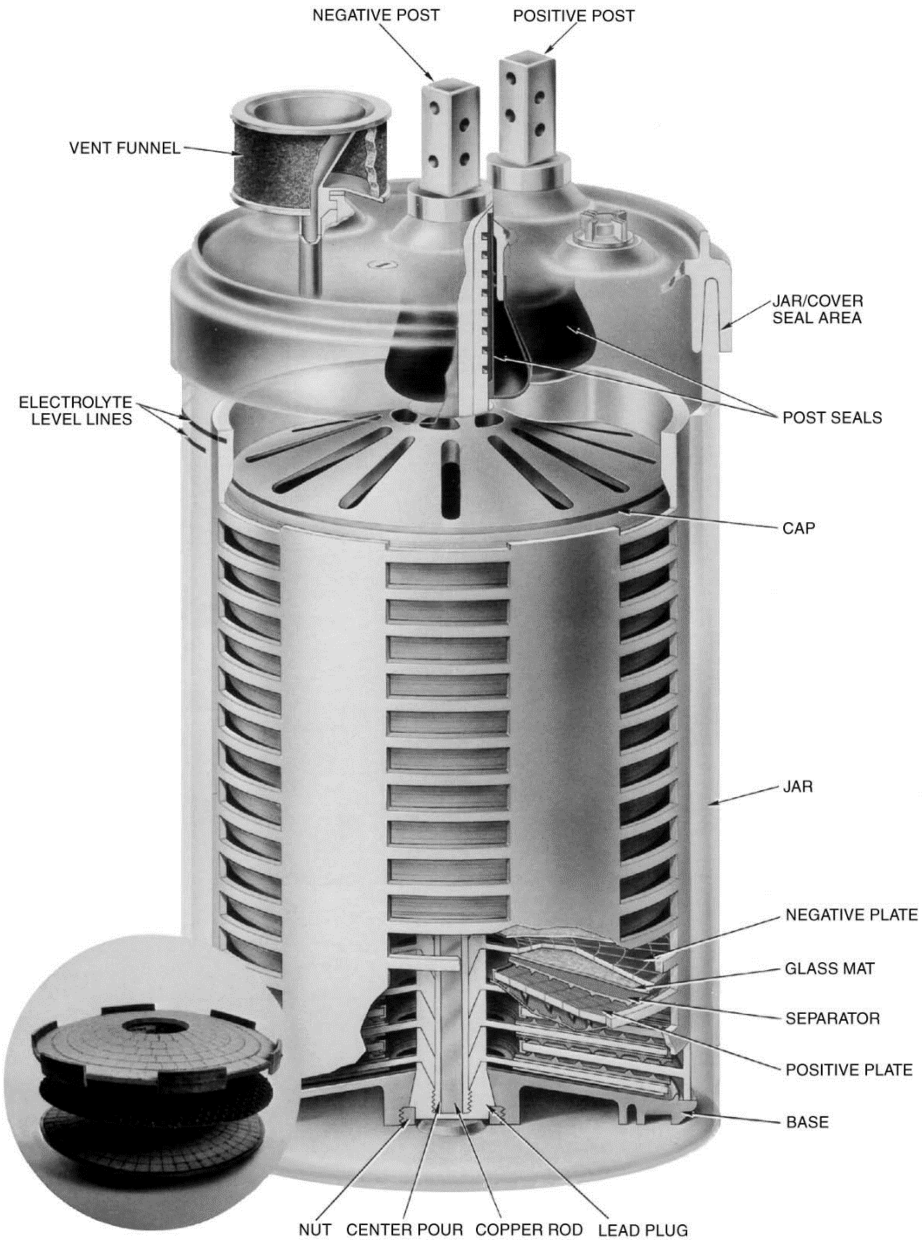
- Early 1980s, GE was losing market share to competitors---Jack Welch was unhappy
- 1983-1986 GE designed, tested, and began to produce a new lower cost “rotary” compressor
- Stopped accelerated testing after one year and no failures
- One million + in service by 1987
- First failure after 1.5 years; virtually all would have eventually failed prematurely
- GE replaced all compressors in refrigerators that it could find. Total cost was more than $450 Million

Failing and Failed PC Electrolytic Capacitors

AT&T Round Cells (aka "Bell Cells")

- Radical new design of a lead-acid battery in 1972
- Longer life predicted (70 years based on accelerated test)
- Lower maintenance cost (epoxy post seal)
- Hundreds of thousands installed during the 1970s
- Problems started in 1978
- A large proportion of the batteries had to be replaced.

Illustration: *Figure 1-1 shows a cut-away view of the Round Cell Battery.*



F-22 Reliability Problems

Premier U.S. Fighter Jet Has Major Shortcomings
F-22's Maintenance Demands Growing

By R. Jeffrey Smith
Washington Post Staff Writer
Friday, July 10, 2009

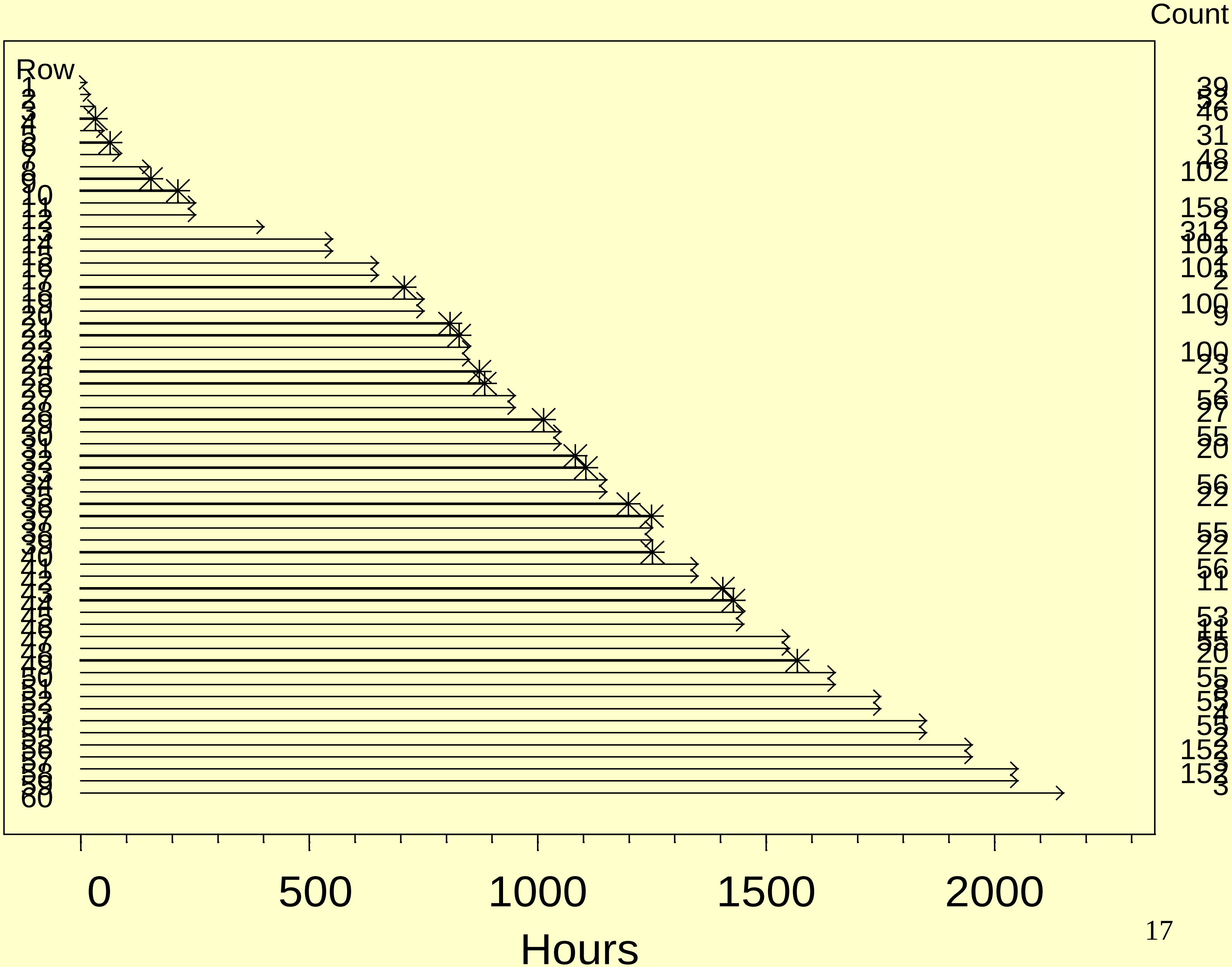
> The United States' top fighter jet, the Lockheed Martin F-22, has recently required more than 30 hours of maintenance for every hour in the skies, pushing its hourly cost of flying to more than $44,000,….

Analysis of Field Data

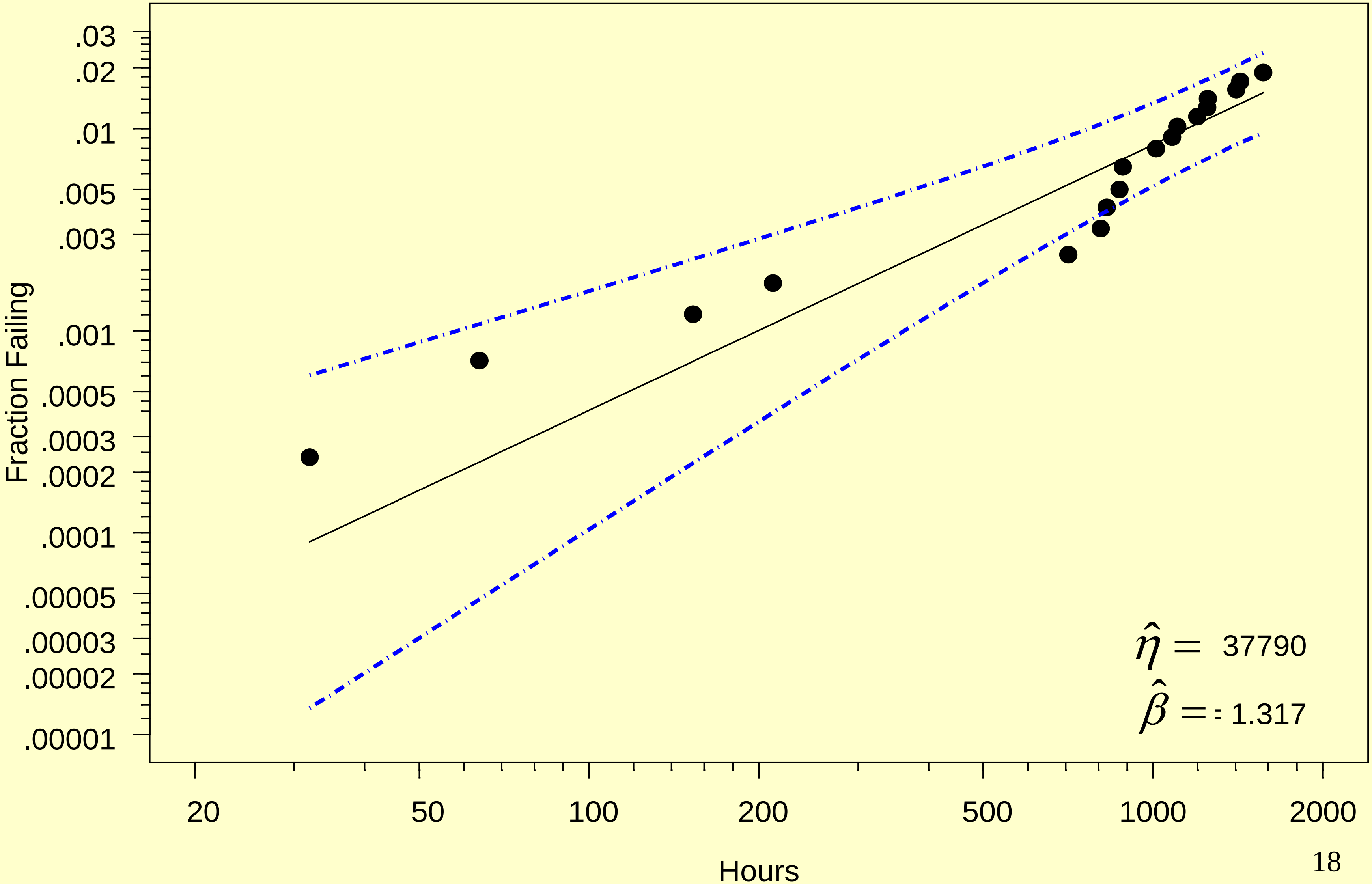
- Warranty data
- Maintenance data
- Field-tracking data

Jet Engine Bleed System Failure

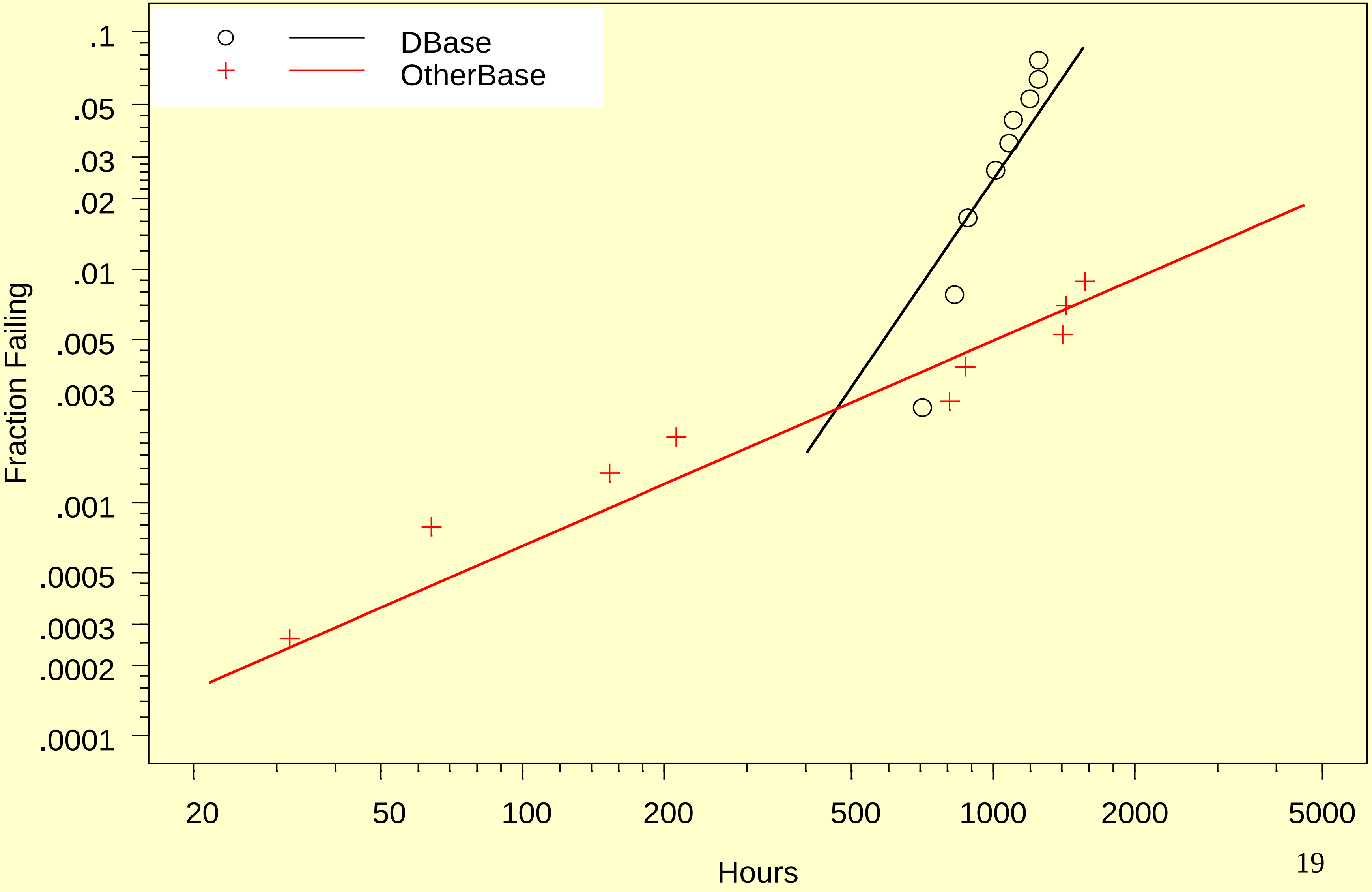
- Data from the *Weibull Handbook (1984)*
- Field data from 2256 systems in the field; staggered entry--multiple censoring
- Unexpected failures
- What is going on??


Count
Row
0
500
1000
1500
2000
Hours

Bleed System Failure Data
with Weibull ML Estimate and Pointwise 95% Confidence Intervals
Weibull Probability Plot



Bleed System Failure Data
With Individual Weibull Distribution ML Estimates
Weibull Probability Plot

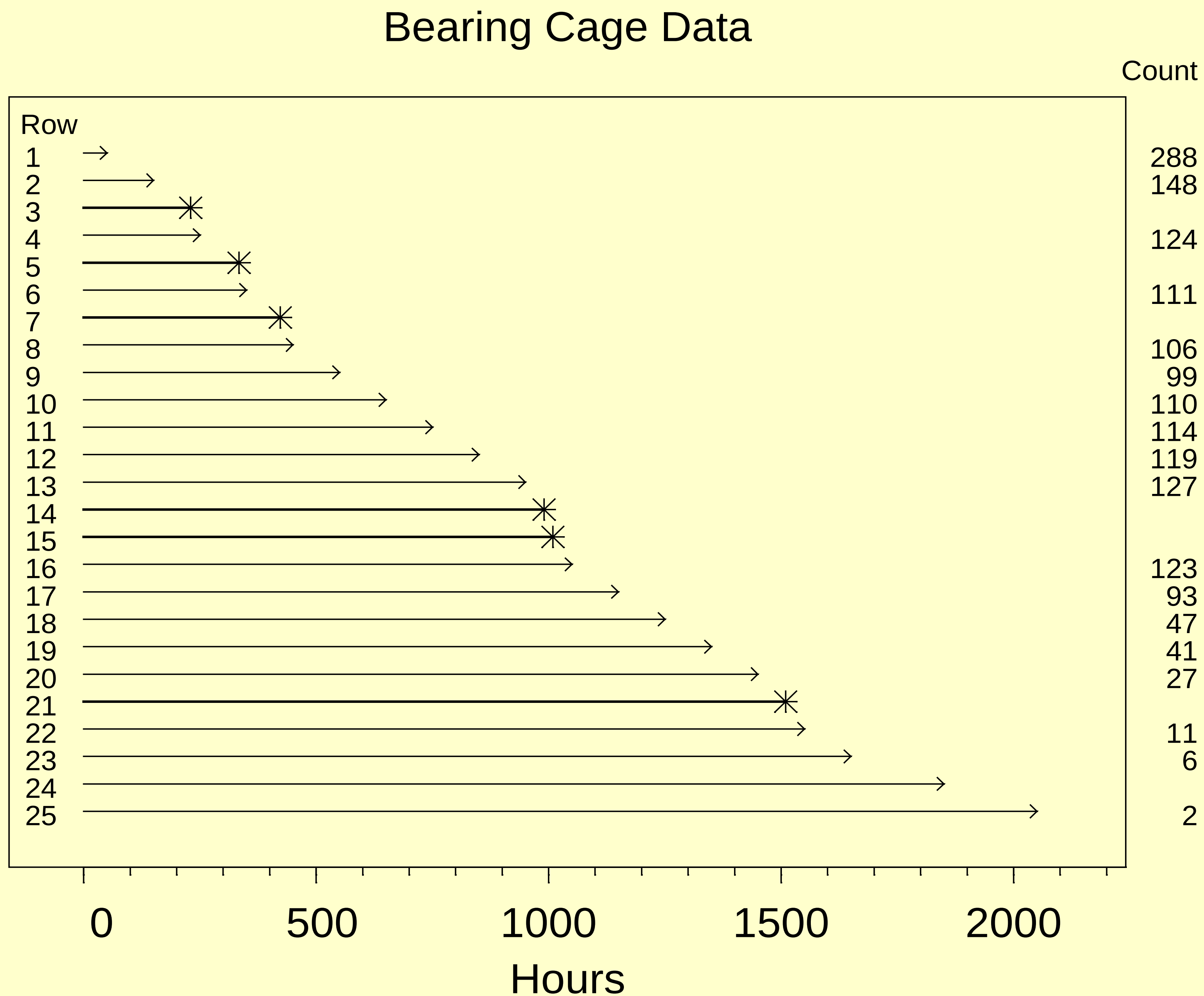


Lessons Learned

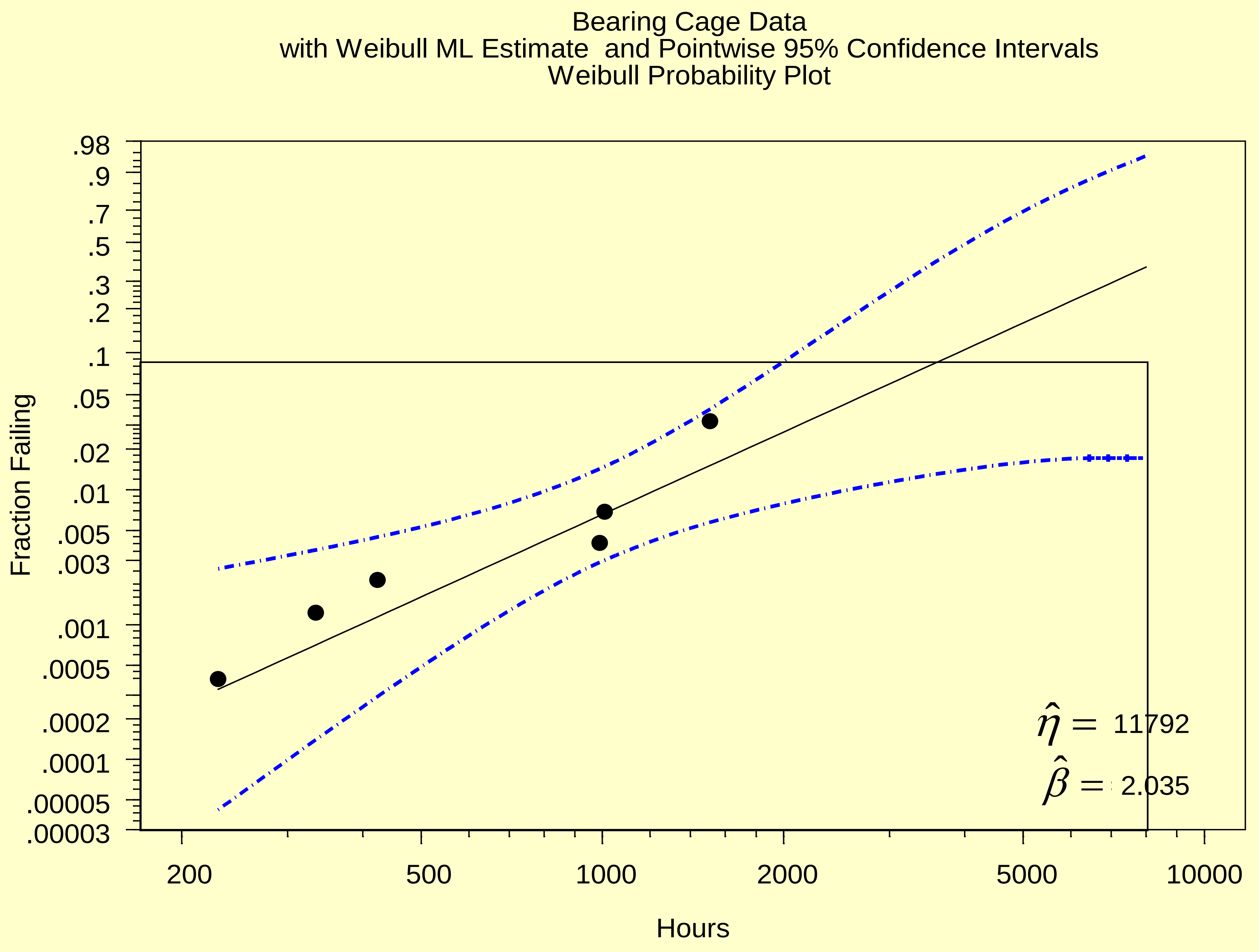
- A shift in slope of a probability plot often indicates a different failure mode
- Look for explanatory variables to help better understand data sources

Bearing Cage Field Failure Data

- Data from the *Weibull Handbook (1984)*
- 1703 units had been introduced into the field over time; oldest unit at 2220 hours of operation.
- 6 units had failed
- Design life specification was B10 = 8000 hours of operation
- Do we have a serious problem? Re-design needed?
- How many spares needed?

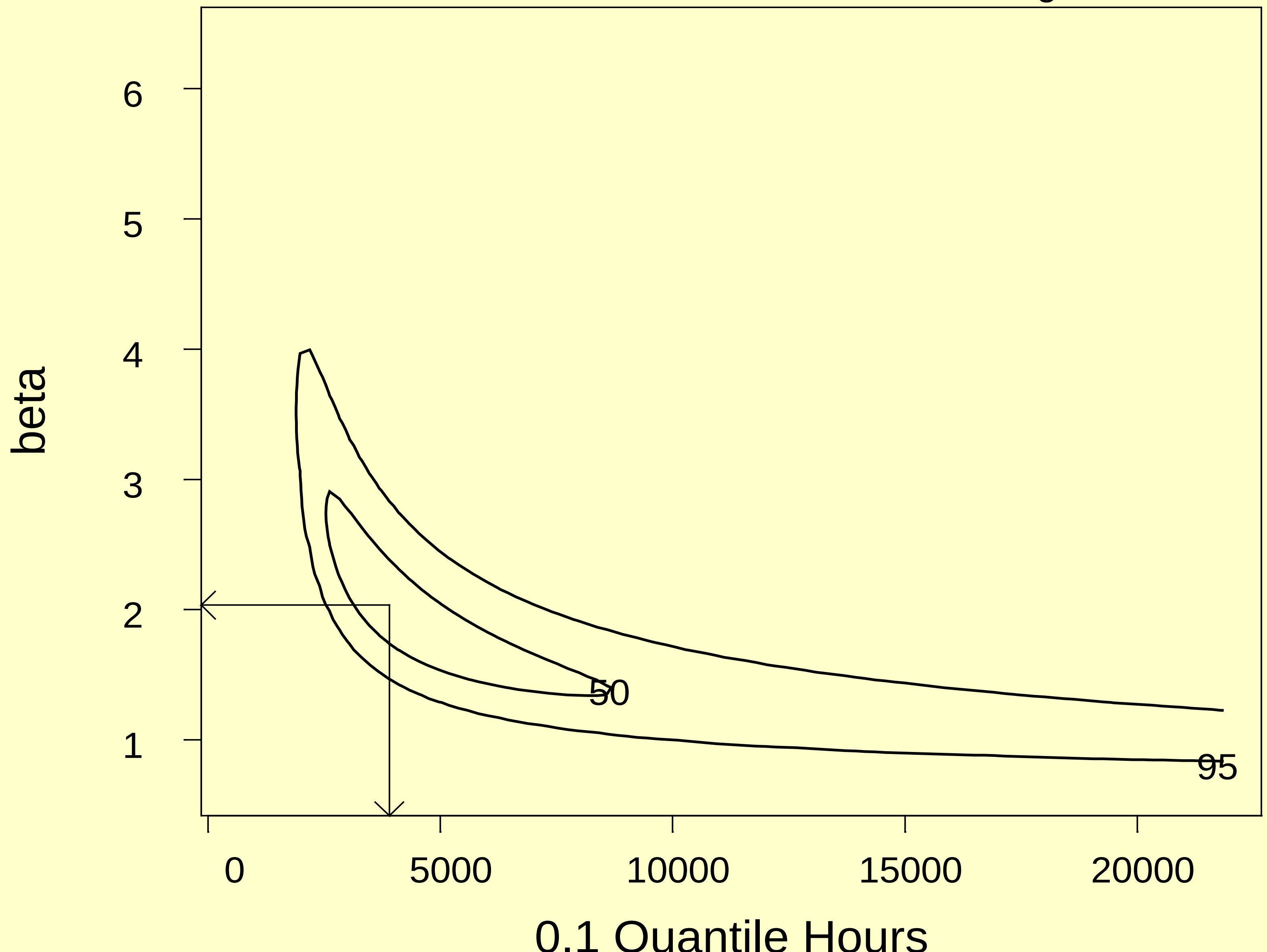


Bearing Cage Data
Count
Row
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
288
148
124
111
106
99
110
114
119
127
123
93
47
41
27
11
6
2
0
500
1000
1500
2000
Hours

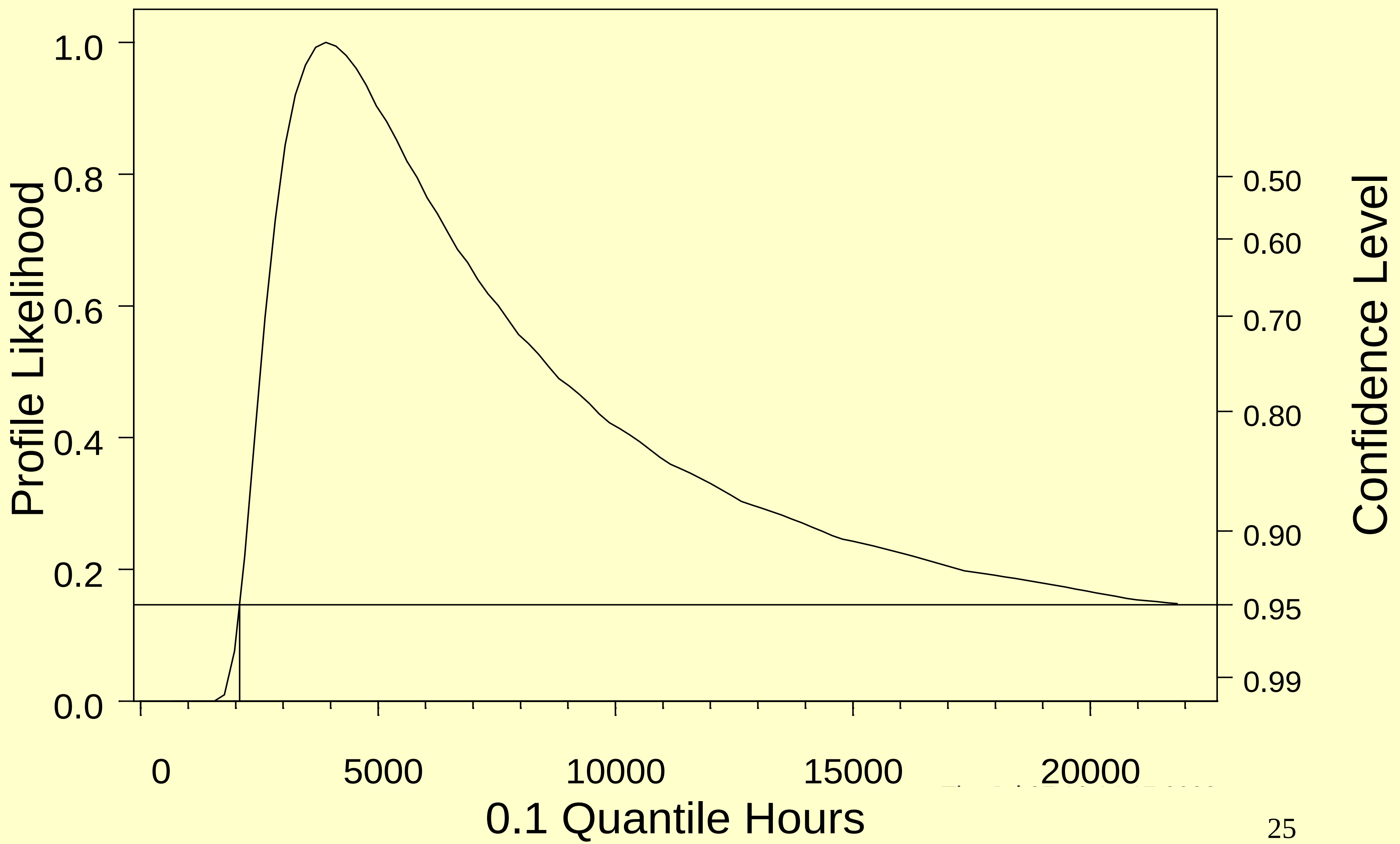


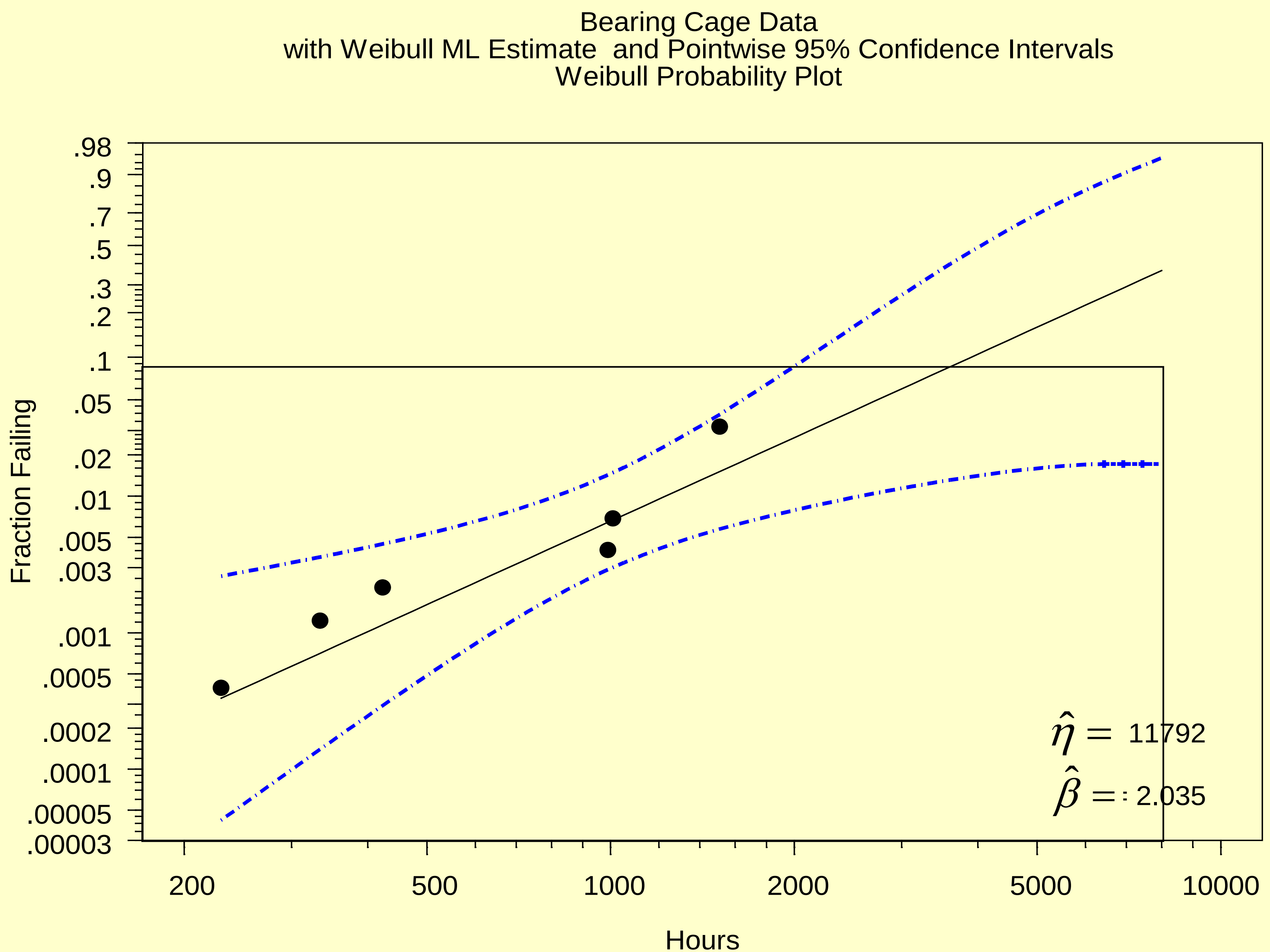
Bearing Cage Data
with Weibull ML Estimate and Pointwise 95% Confidence Intervals
Weibull Probability Plot
Fraction Failing
.98
.9
.7
.5
.3
.2
.1
.05
.02
.01
.005
.003
.001
.0005
.0002
.0001
.00005
.00003
200
500
1000
2000
5000
10000
Hours
$\hat{\eta} =$ 11792
$\hat{\beta} =$ 2.035

Bearing Cage Data
Weibull Distribution Joint Confidence Region

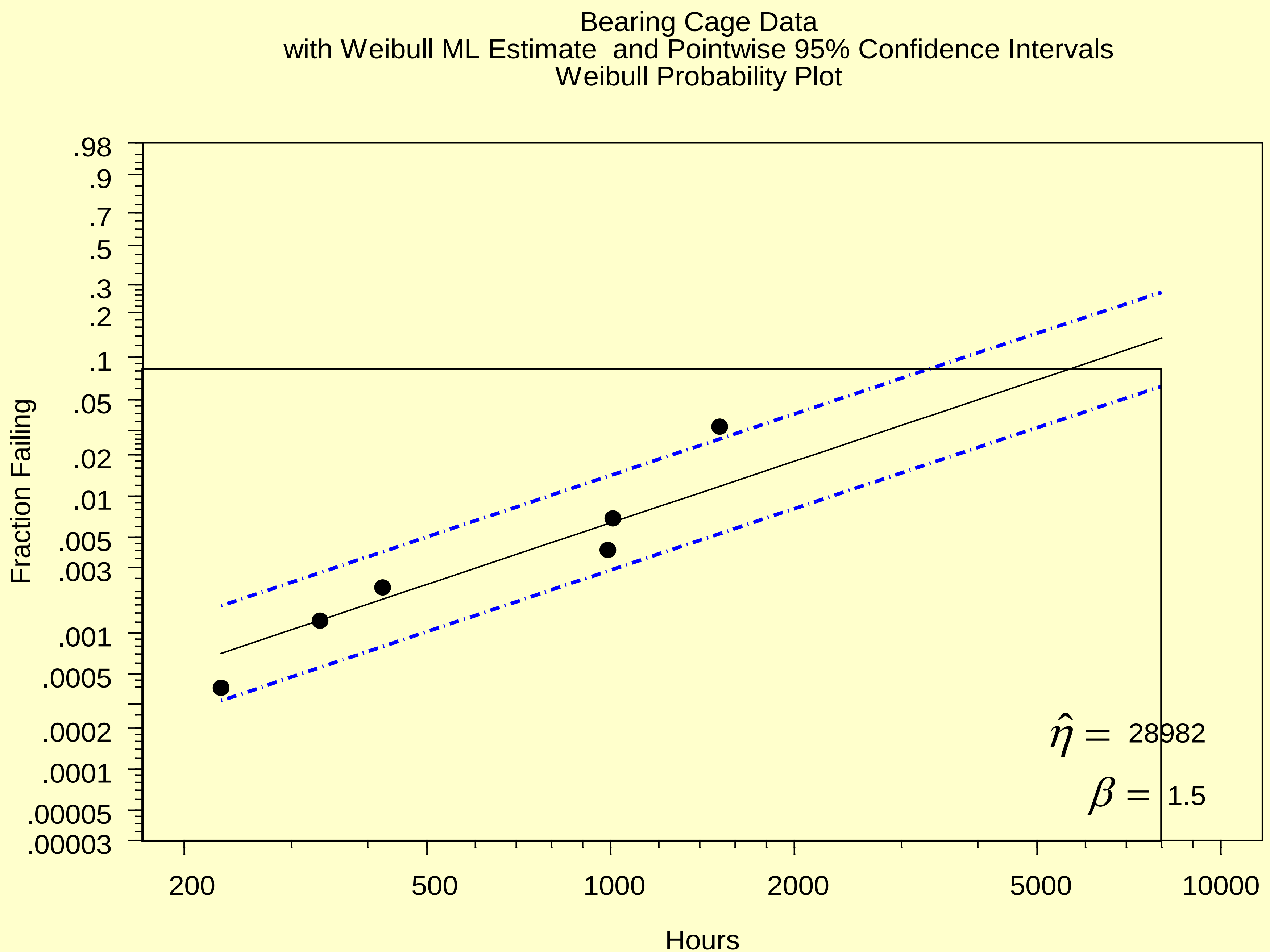


Bearing Cage Data
Profile Likelihood and 95% Confidence Interval
for 0.1 Quantile Hours from the Weibull Distribution



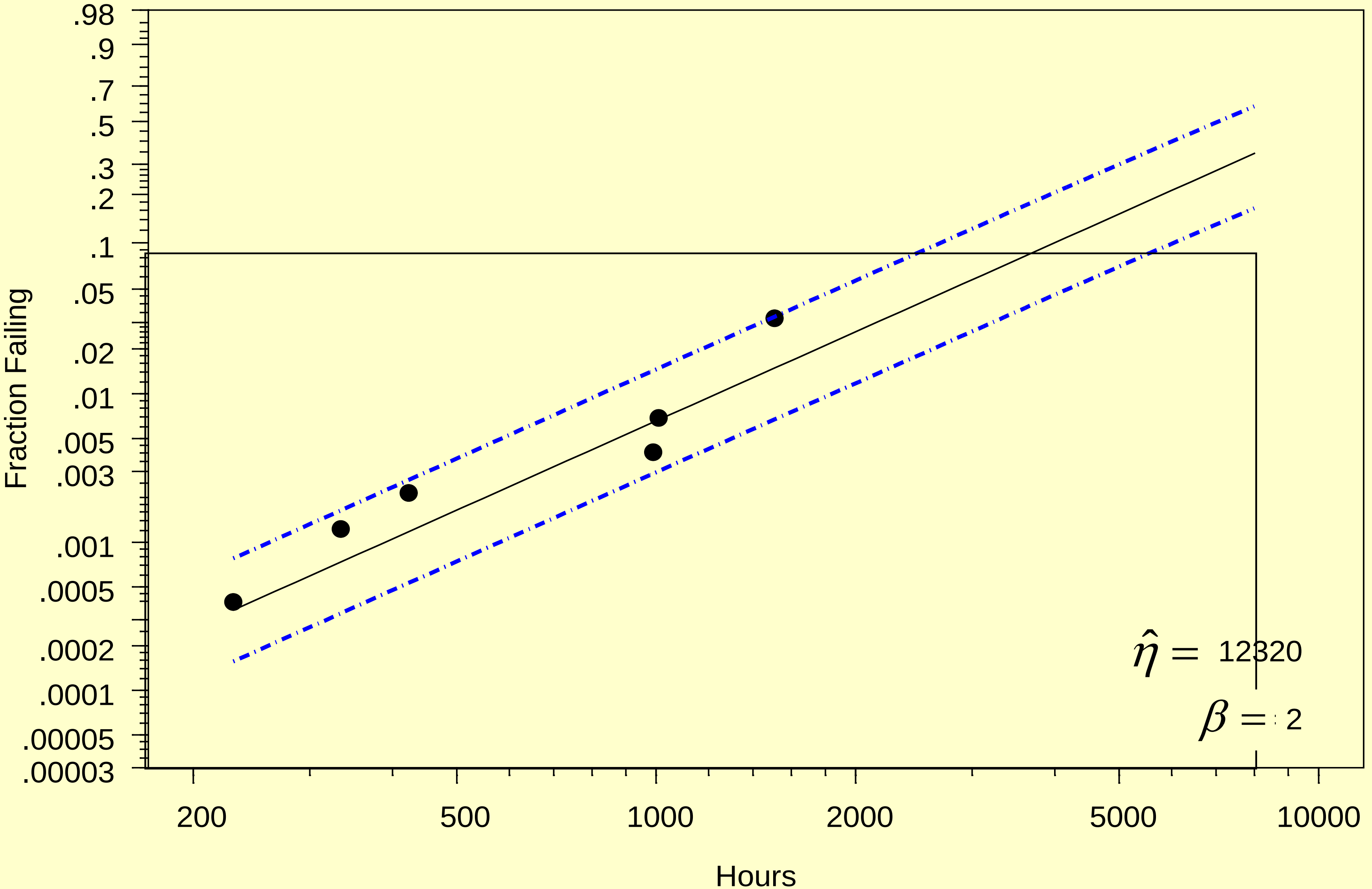


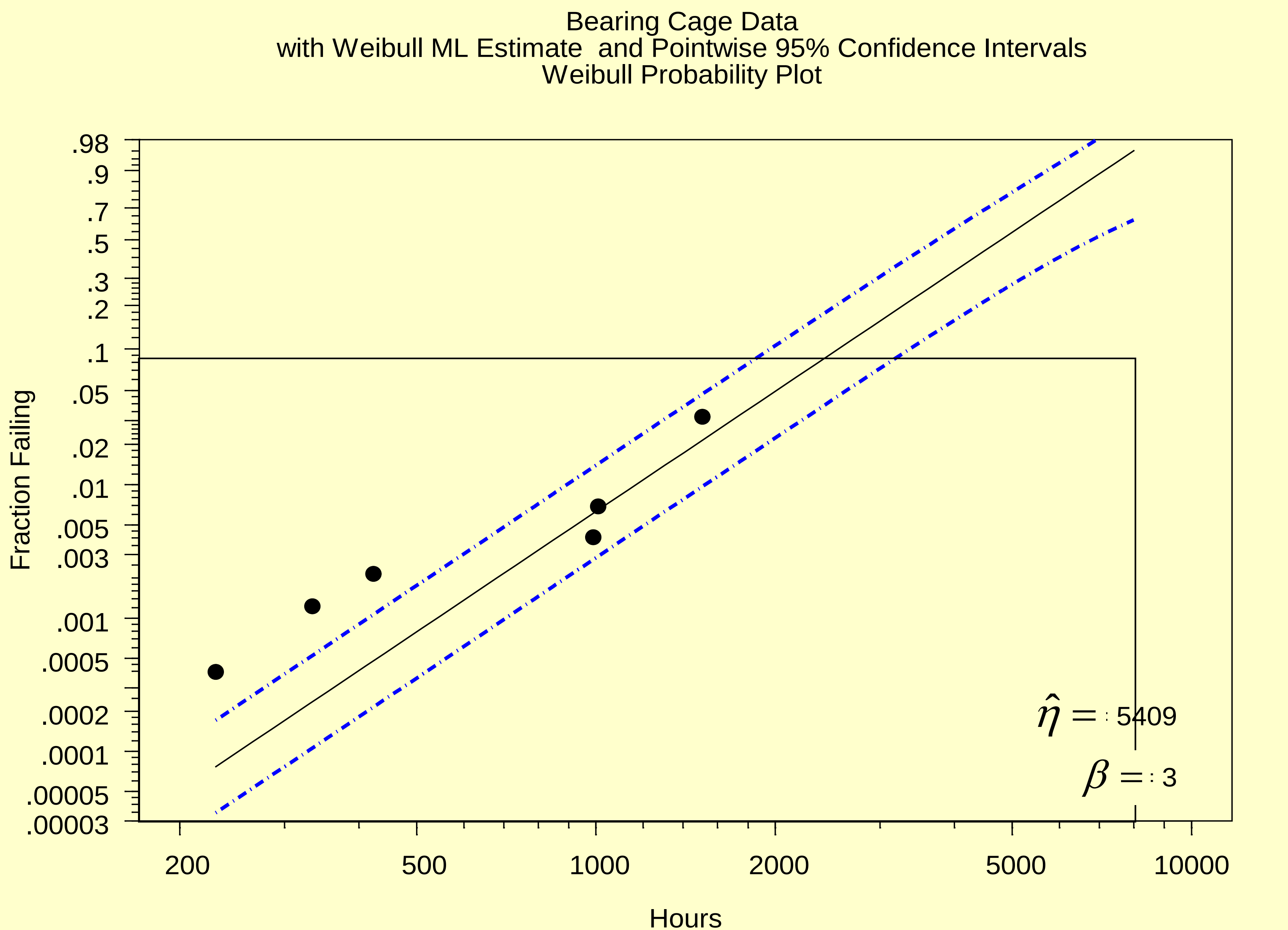
Bearing Cage Data
with Weibull ML Estimate and Pointwise 95% Confidence Intervals
Weibull Probability Plot
Fraction Failing
.98
.9
.7
.5
.3
.2
.1
.05
.02
.01
.005
.003
.001
.0005
.0002
.0001
.00005
.00003
200
500
1000
2000
5000
10000
Hours
$\hat{\eta} =$ 11792
$\hat{\beta} =$ 2.035



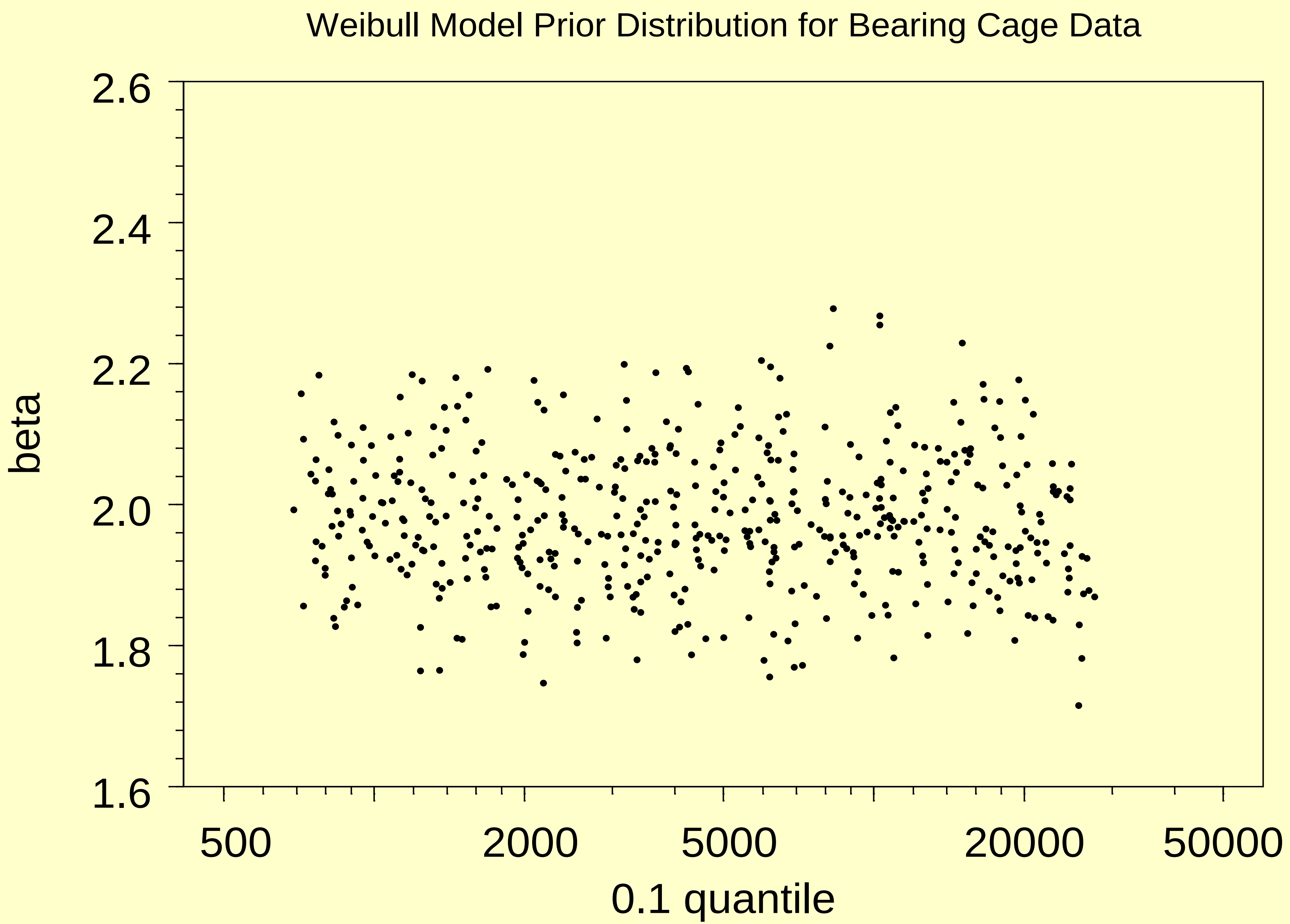
Bearing Cage Data
with Weibull ML Estimate and Pointwise 95% Confidence Intervals
Weibull Probability Plot
Fraction Failing
.98
.9
.7
.5
.3
.2
.1
.05
.02
.01
.005
.003
.001
.0005
.0002
.0001
.00005
.00003
200
500
1000
2000
5000
10000
Hours
$\hat{\eta} =$ 28982
$\beta =$ 1.5

Bearing Cage Data
with Weibull ML Estimate and Pointwise 95% Confidence Intervals
Weibull Probability Plot



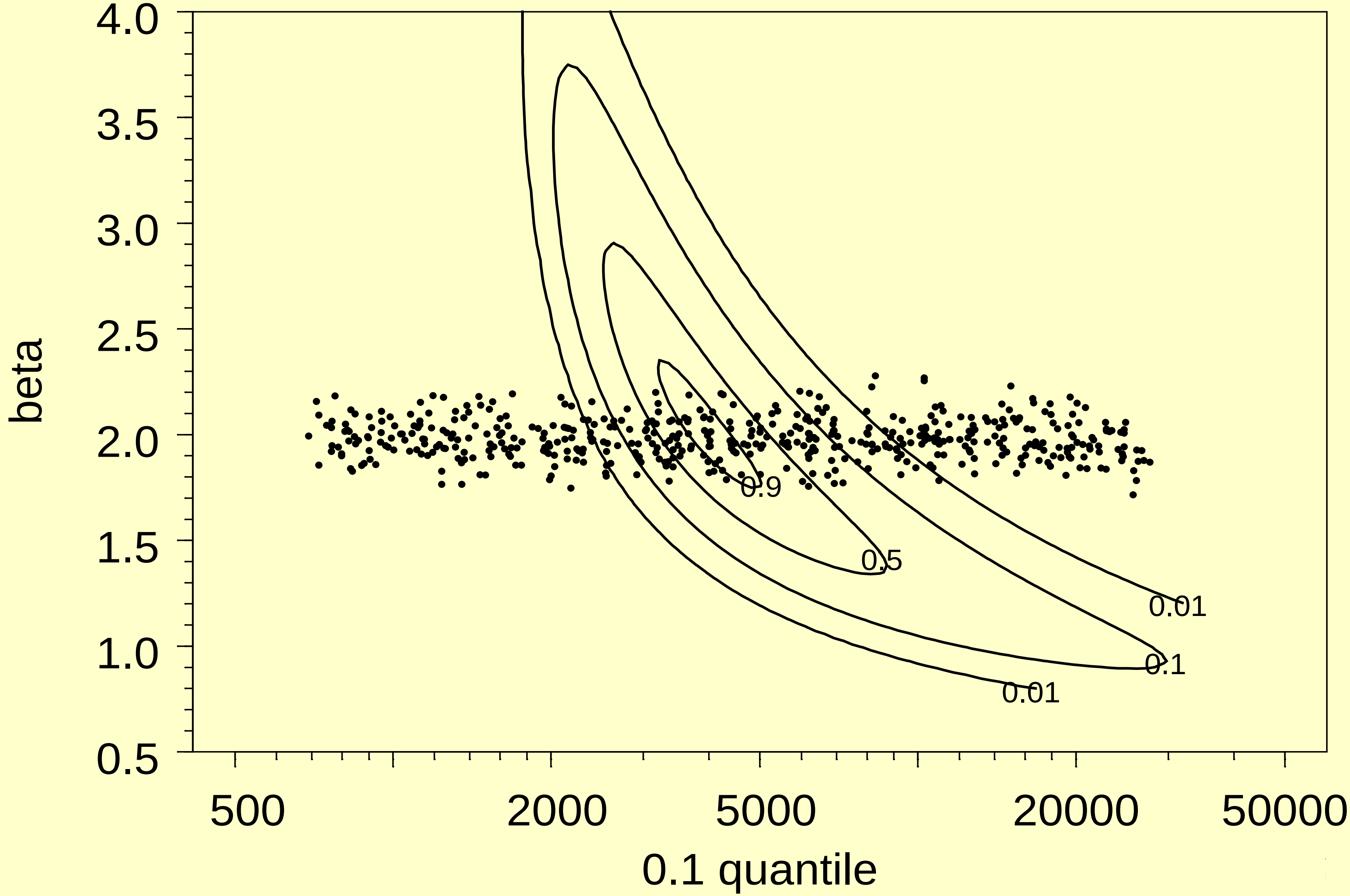


Bearing Cage Data
with Weibull ML Estimate and Pointwise 95% Confidence Intervals
Weibull Probability Plot
Fraction Failing
.98
.9
.7
.5
.3
.2
.1
.05
.02
.01
.005
.003
.001
.0005
.0002
.0001
.00005
.00003
200
500
1000
2000
5000
10000
Hours
$\hat{\eta} = 5409$
$\beta = 3$

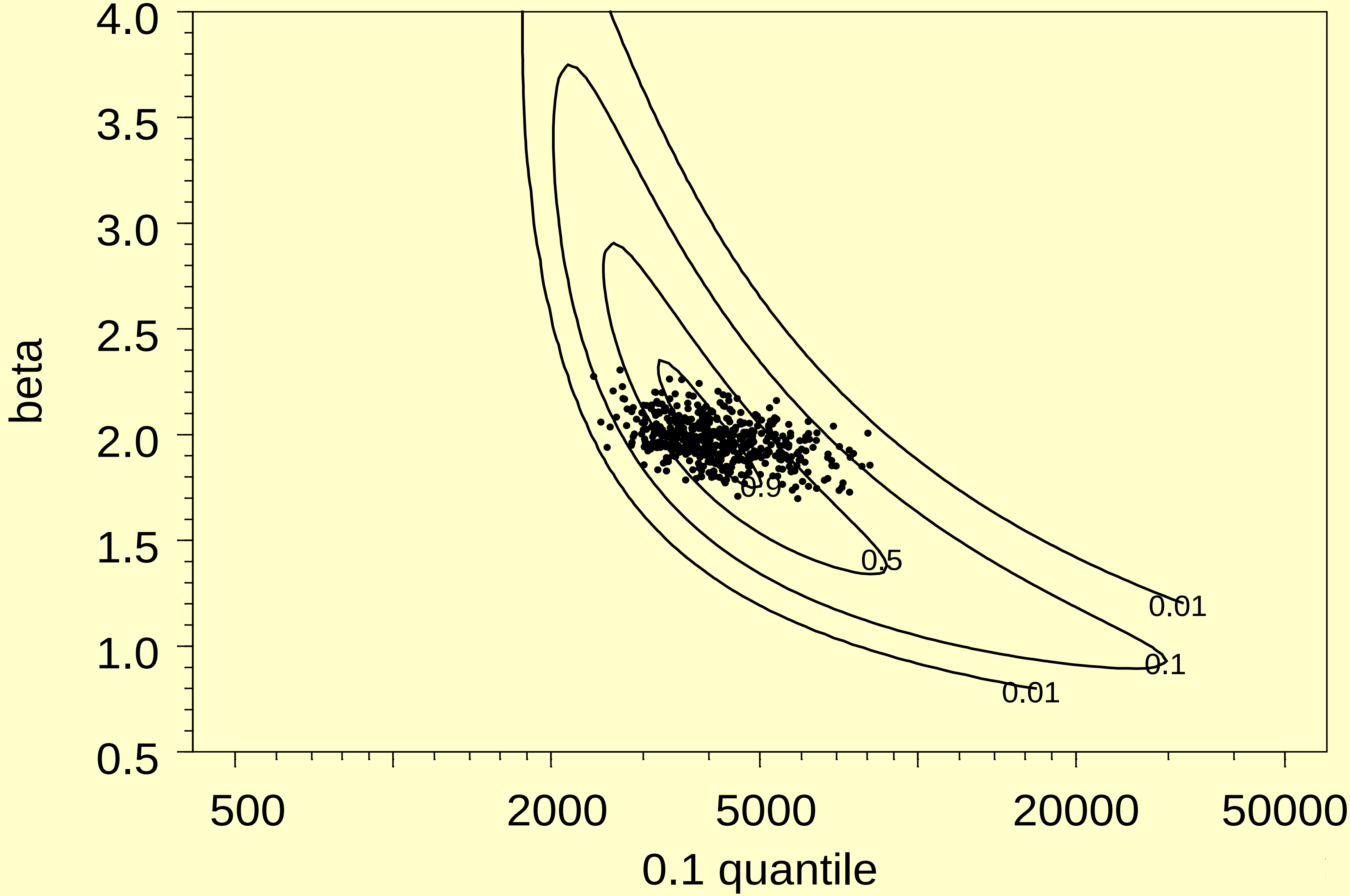


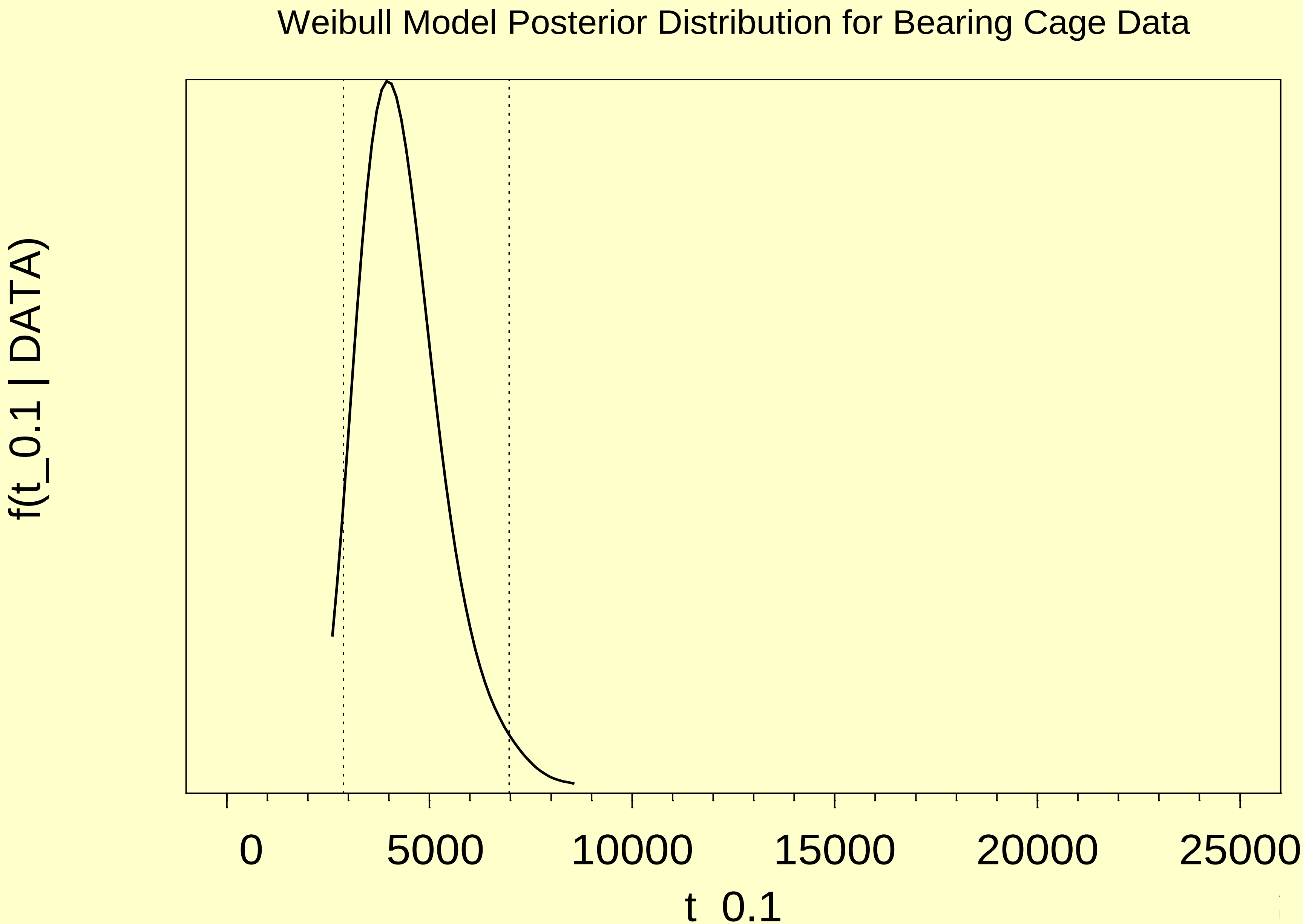
Weibull Model Prior Distribution for Bearing Cage Data
2.6
2.4
2.2
2.0
1.8
1.6
beta
500
2000
5000
20000
50000
0.1 quantile

Weibull Model Prior Distribution for Bearing Cage Data

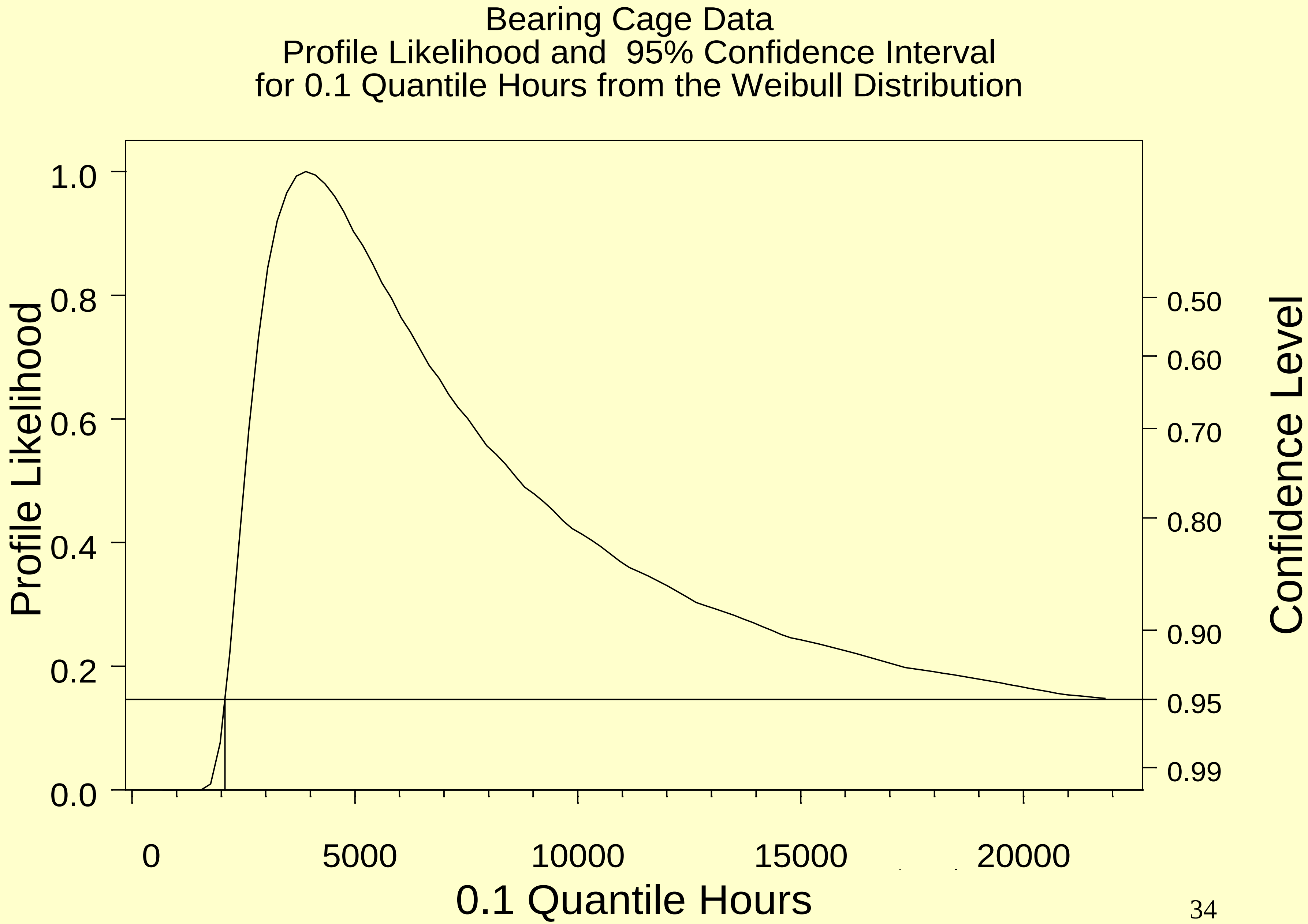


Weibull Model Posterior Distribution for Bearing Cage Data





Weibull Model Posterior Distribution for Bearing Cage Data
f(t_0.1 | DATA)
0
5000
10000
15000
20000
25000
t_0.1


Bearing Cage Data
Profile Likelihood and 95% Confidence Interval
for 0.1 Quantile Hours from the Weibull Distribution
Profile Likelihood
1.0
0.8
0.6
0.4
0.2
0.0
Confidence Level
0.50
0.60
0.70
0.80
0.90
0.95
0.99
0
5000
10000
15000
20000
0.1 Quantile Hours

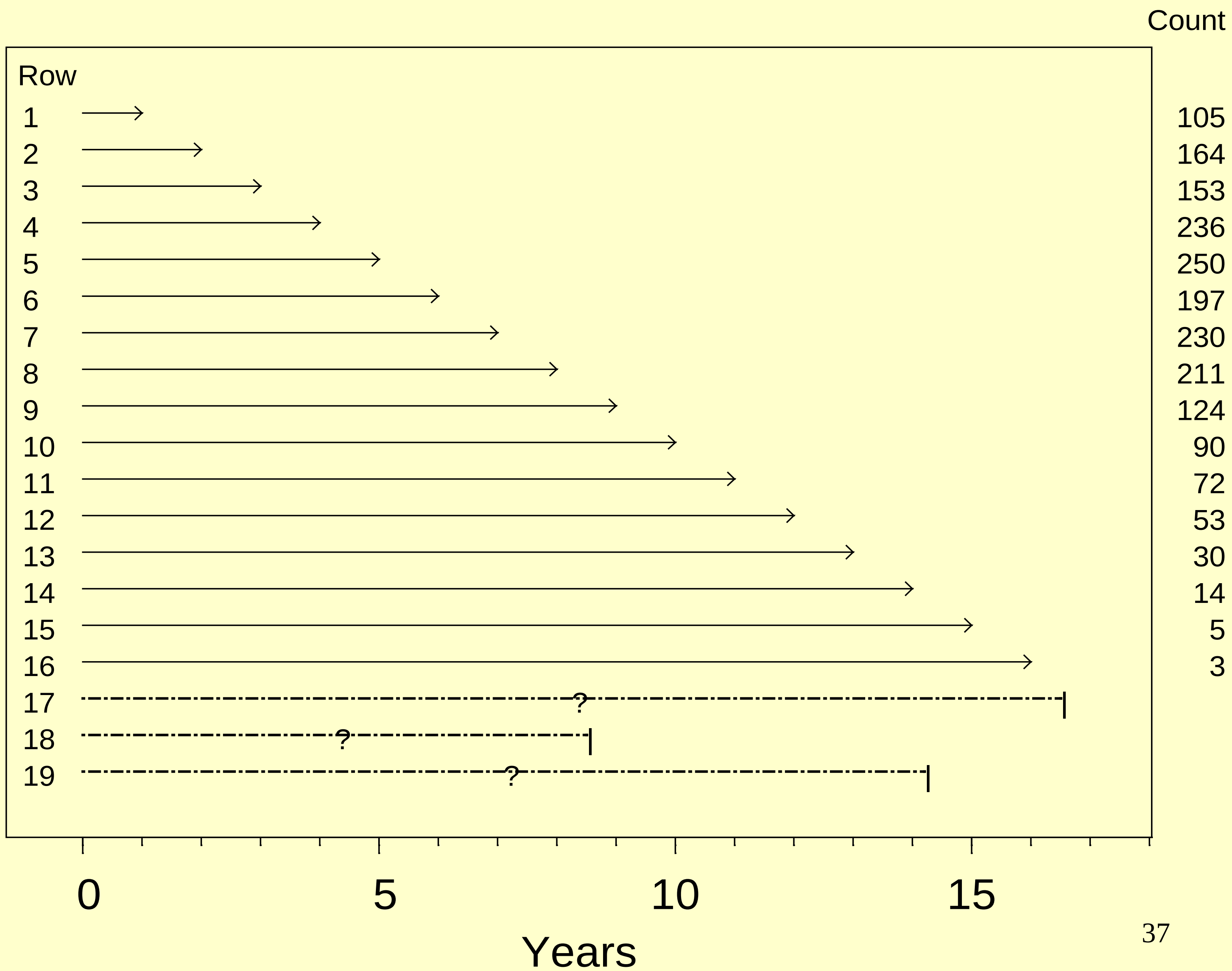
Lessons Learned

- When data are not sufficient to answer the question, it is important to seek external information from past experience or engineering/physical/chemical knowledge about failure mechanisms
- Sensitivity analysis provides insight and assessment of uncertainty
- Bayesian methods provide a convenient summarization of the effects of using “prior” information

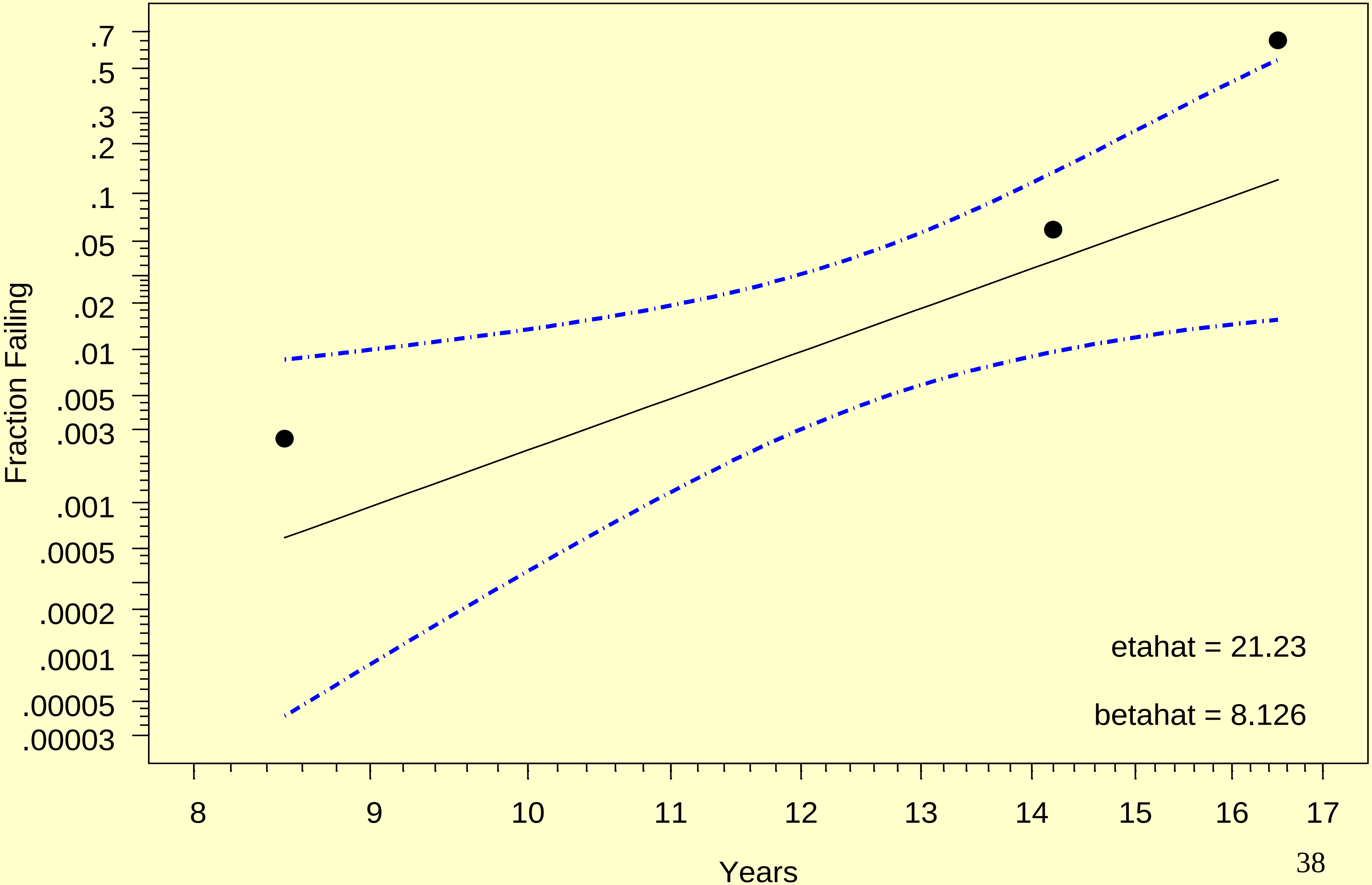
Rocket Motor Field Data Analysis

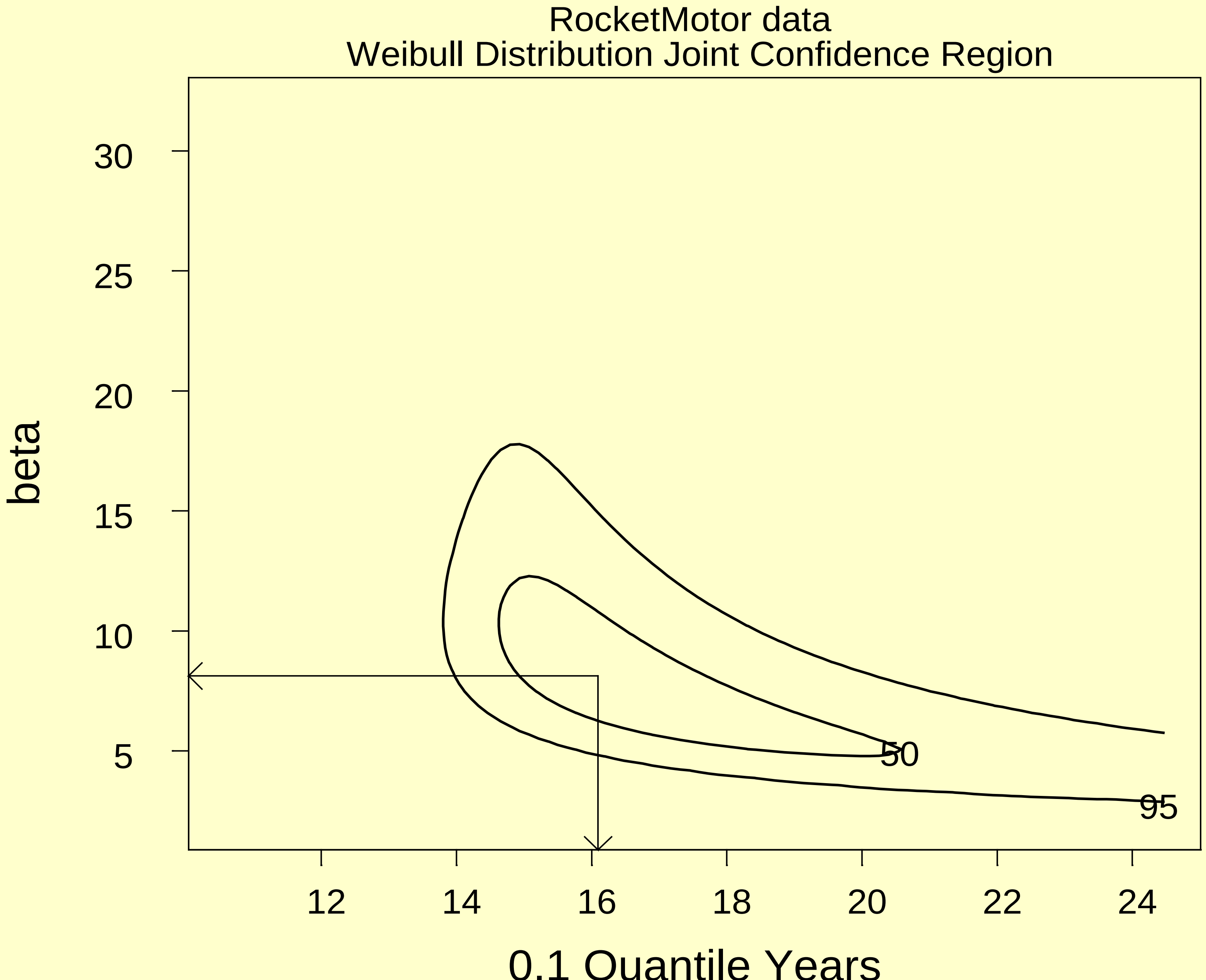
- Data from Olwell and Sorell (2001) RAMS proceedings
- Rocket motor is one of five critical missile components
- 20,000 missiles in inventory
- 2,000 firings over the life of the missile; catastrophic motor failures for three older missiles
- Concern about a possible wearout failure mode

RocketMotor data



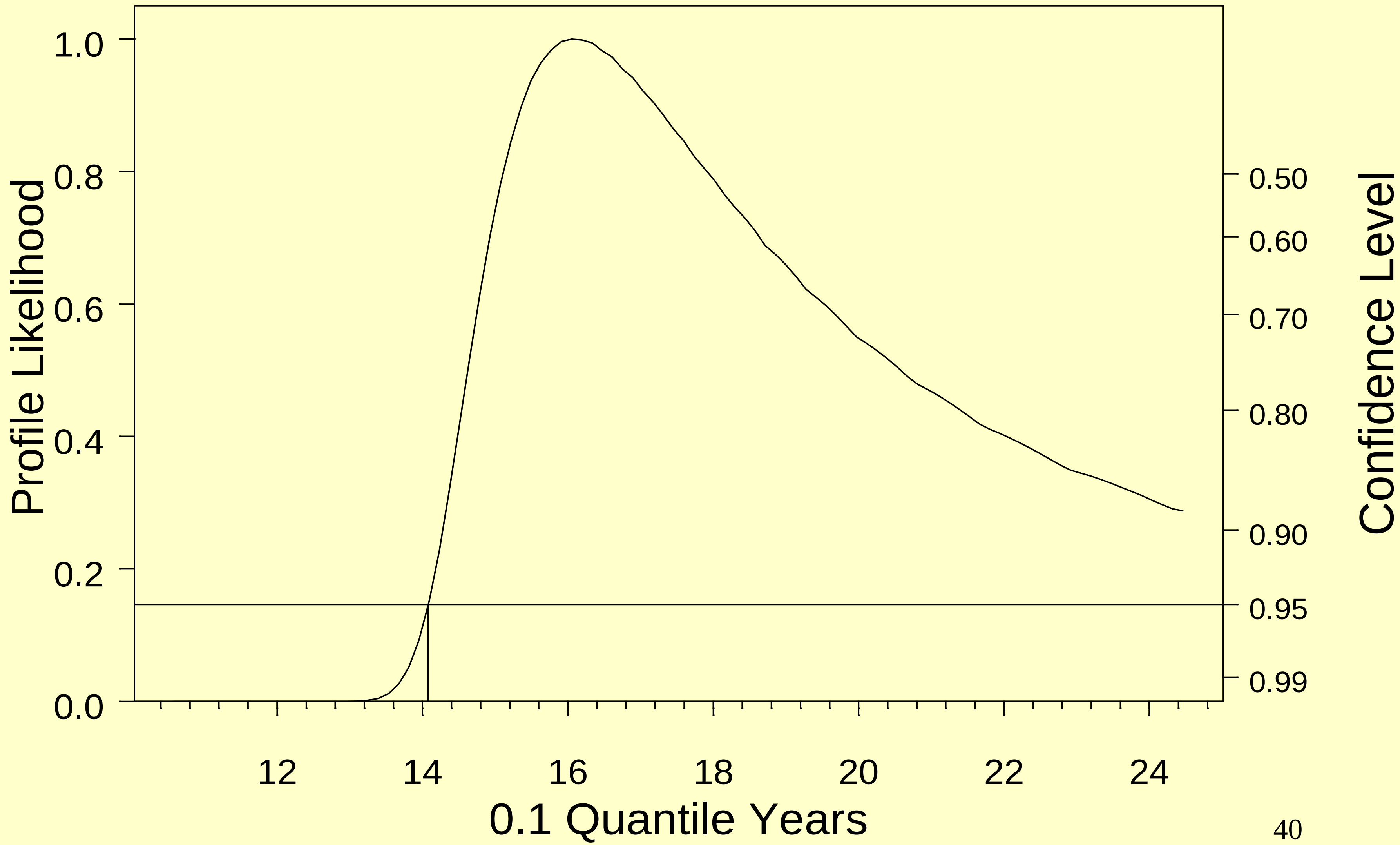
RocketMotor data
with Weibull ML Estimate and Pointwise 95% Confidence Intervals
Weibull Probability Plot



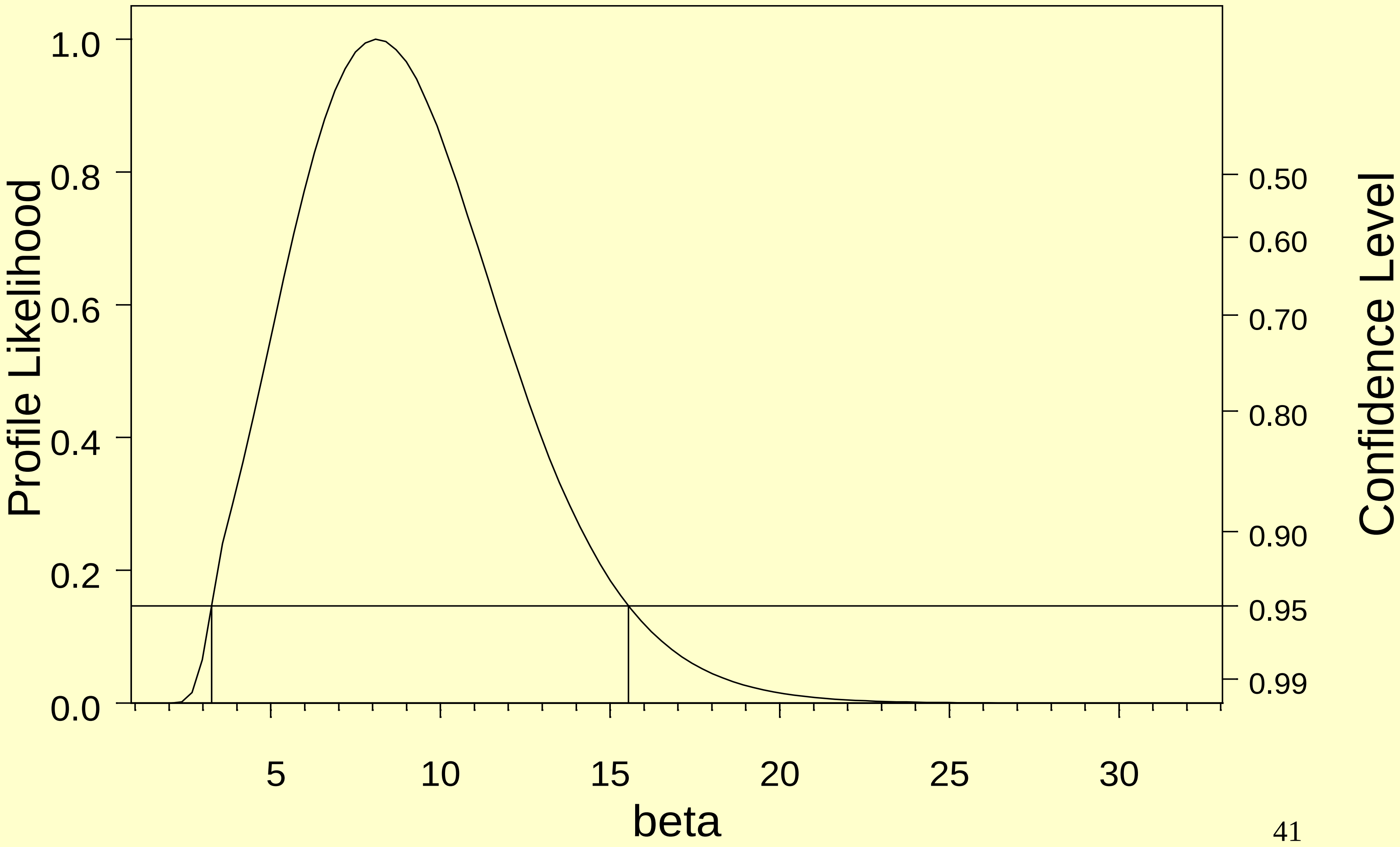


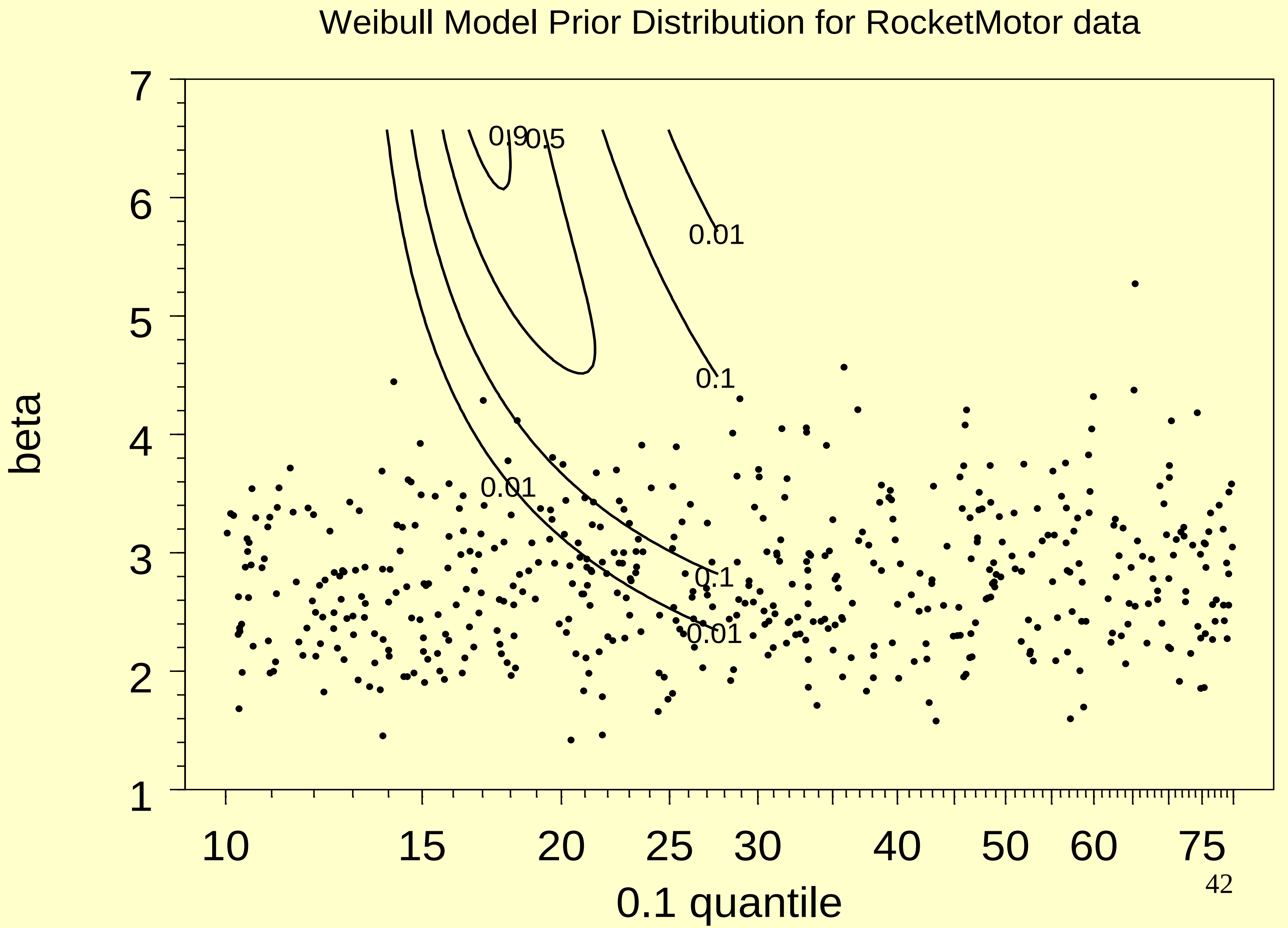
RocketMotor data
Weibull Distribution Joint Confidence Region
beta
30
25
20
15
10
5
12
14
16
18
20
22
24
0.1 Quantile Years
50
95

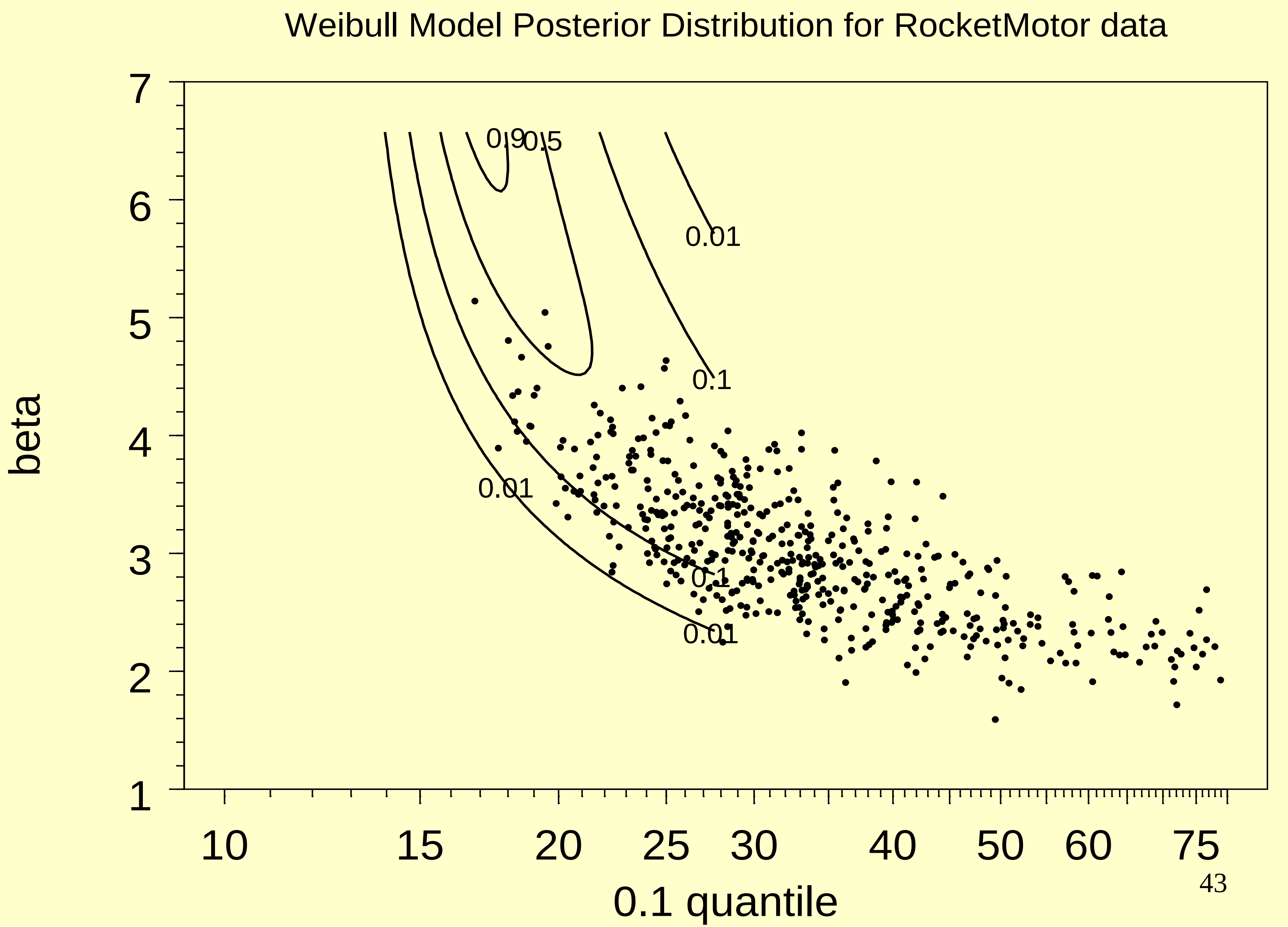
RocketMotor data
Profile Likelihood and 95% Confidence Interval
for 0.1 Quantile Years from the Weibull Distribution



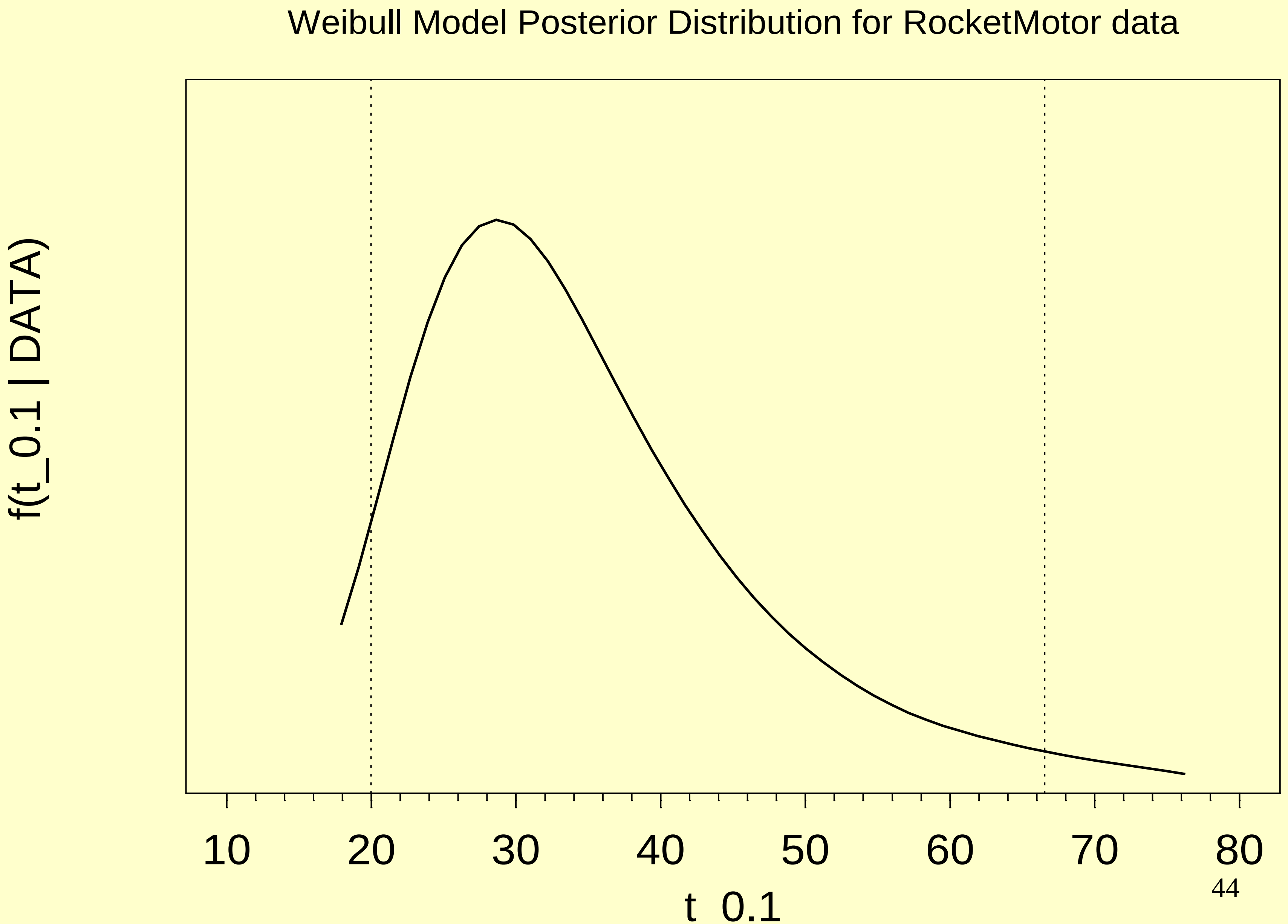
RocketMotor data
Profile Likelihood and 95% Confidence Interval
for beta from the Weibull Distribution



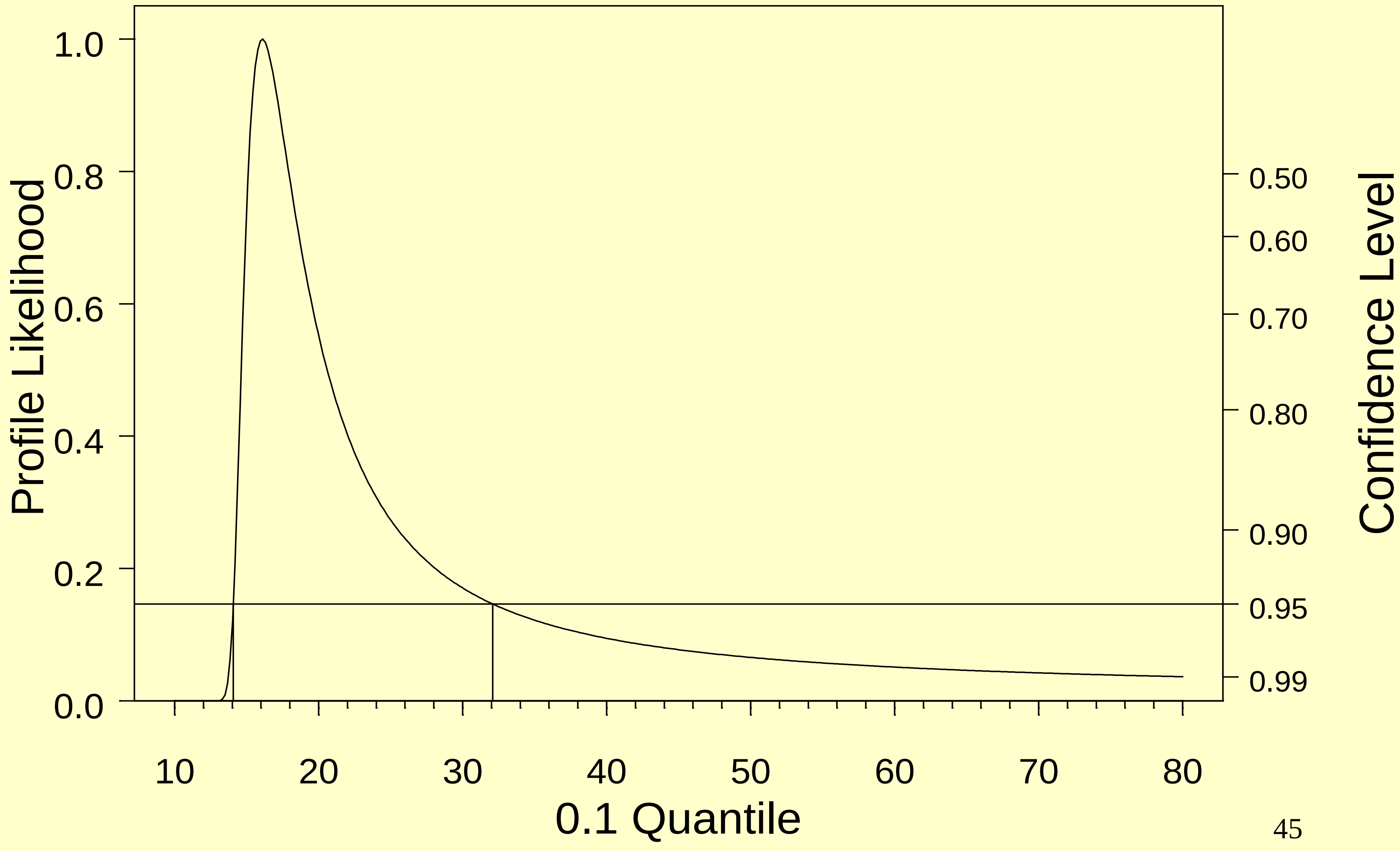

Weibull Model Prior Distribution for RocketMotor data
beta
0.1 quantile
7
6
5
4
3
2
1
10
15
20
25
30
40
50
60
75
0.9
0.5
0.01
0.1
0.01
0.1
0.01



Weibull Model Posterior Distribution for RocketMotor data
beta
0.1 quantile
0.9
0.5
0.1
0.01
1
2
3
4
5
6
7
10
15
20
25
30
40
50
60
75


Weibull Model Posterior Distribution for RocketMotor data
f(t_0.1 | DATA)
10
20
30
40
50
60
70
80
t_0.1

RocketMotor data
Profile Likelihood and 95% Confidence Interval
for 0.1 Quantile from the Weibull Distribution



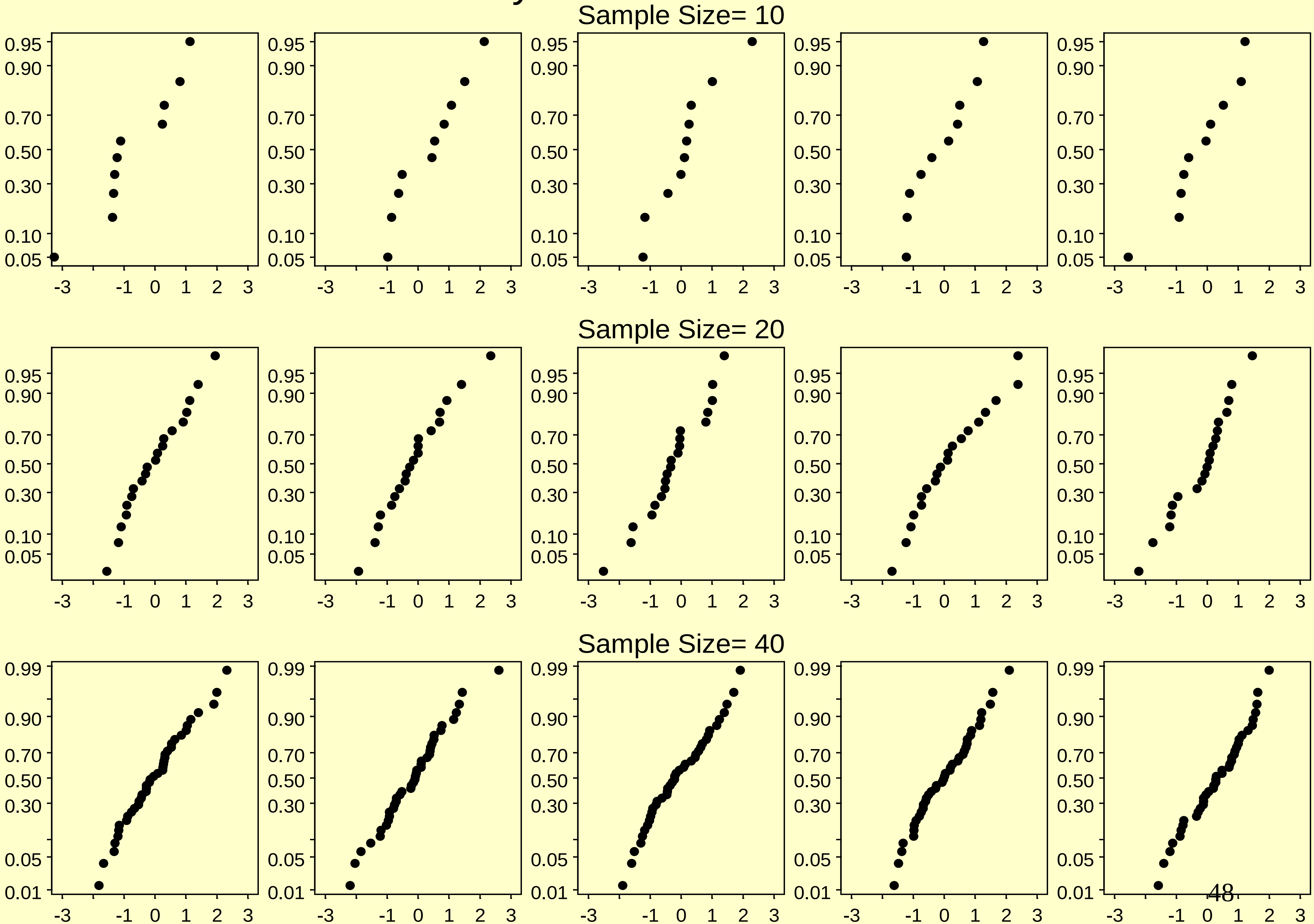
Lessons Learned

- Even with very limited data, useful inferences on reliability may be possible.
- It is important to use external (prior) information, when available
- The shapes of the likelihood (and posterior distribution) are important

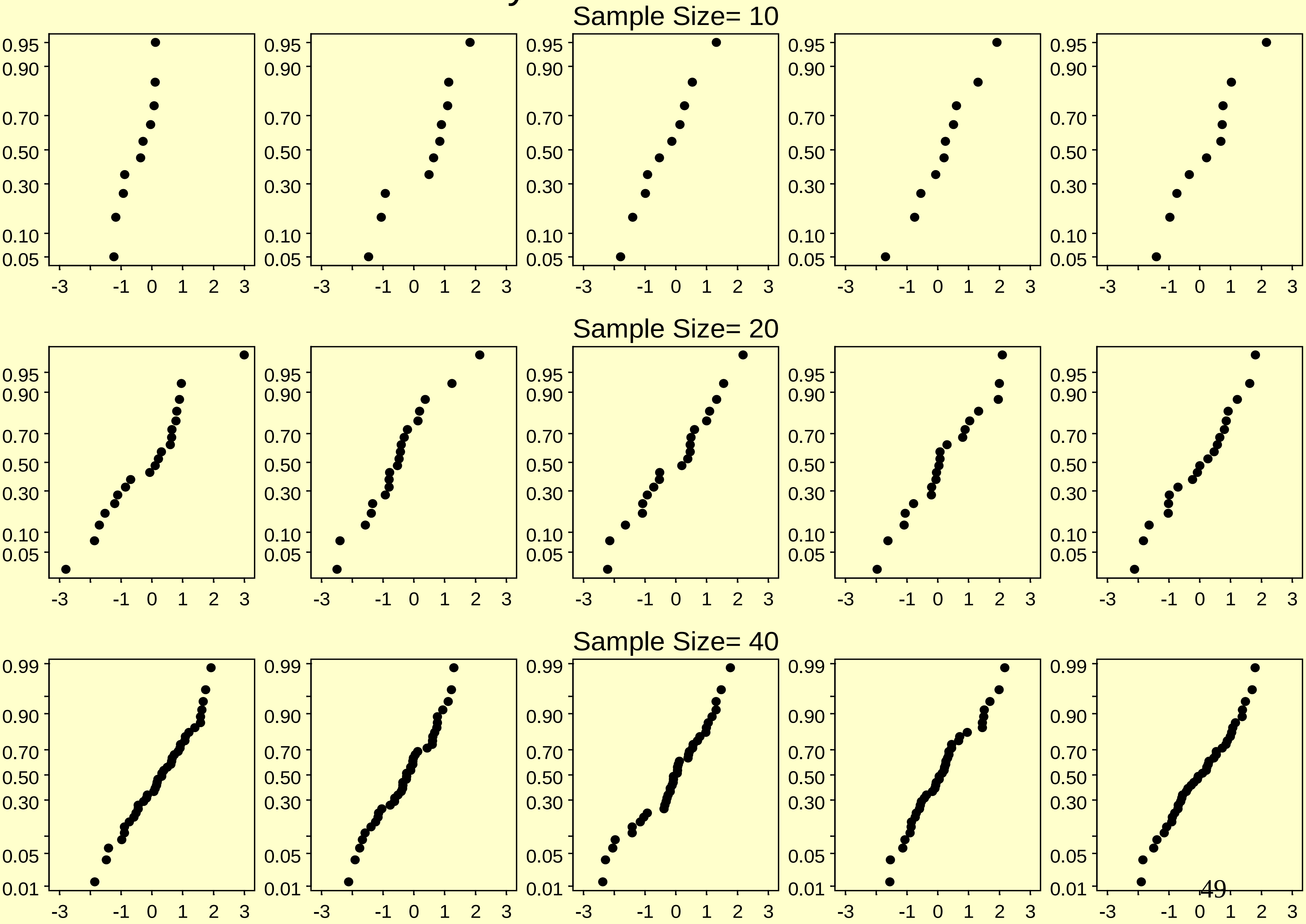
Using Simulation and Graphics to Gain Experience in Assessing Probability Plots

- Motivated by Hahn and Shapiro (1968)
- Dangers of over interpretation of random variability in graphics
- Useful tool for separating signal from noise

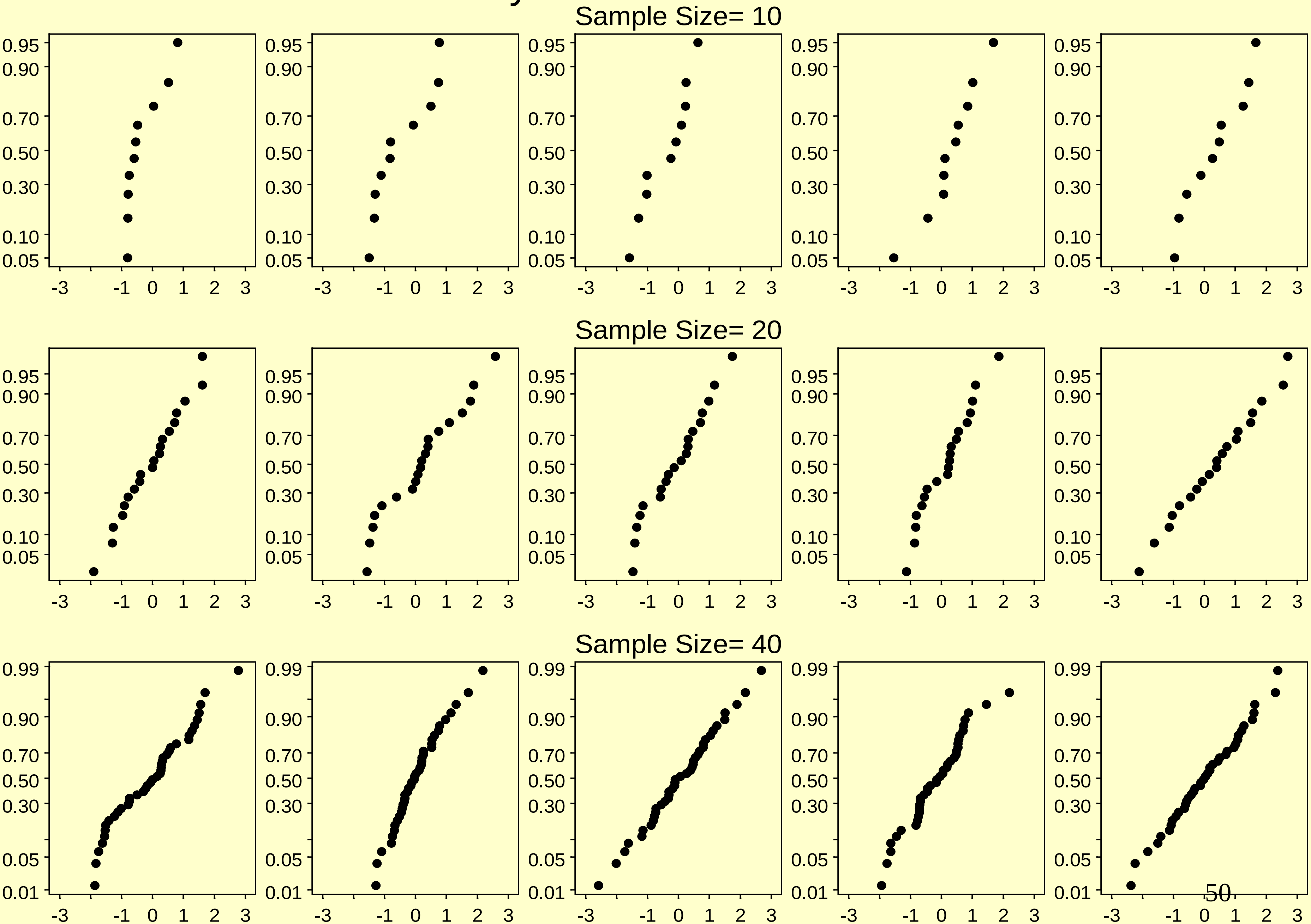
Normal Probability Plot with Simulated Normal Data

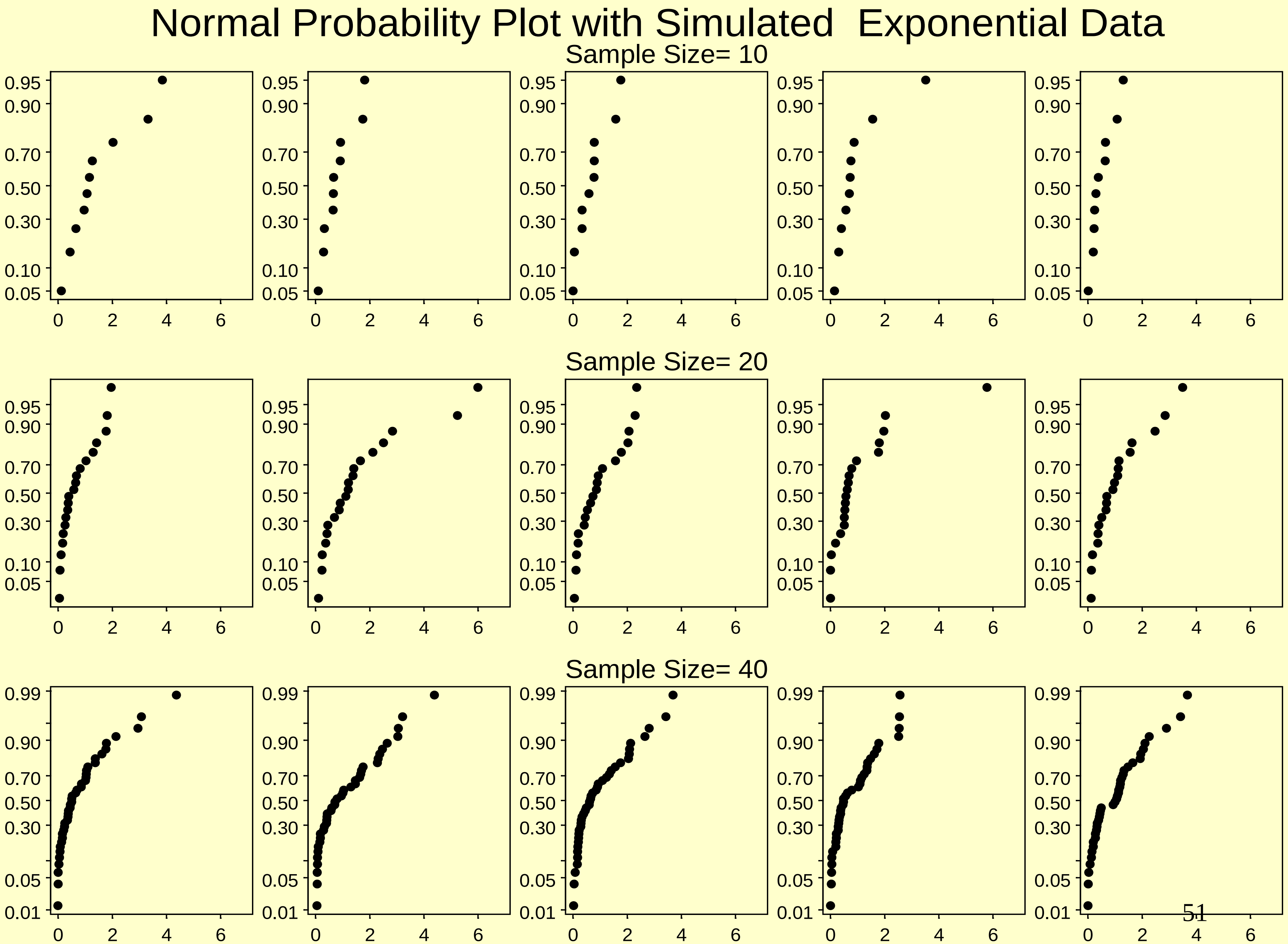


Normal Probability Plot with Simulated Normal Data



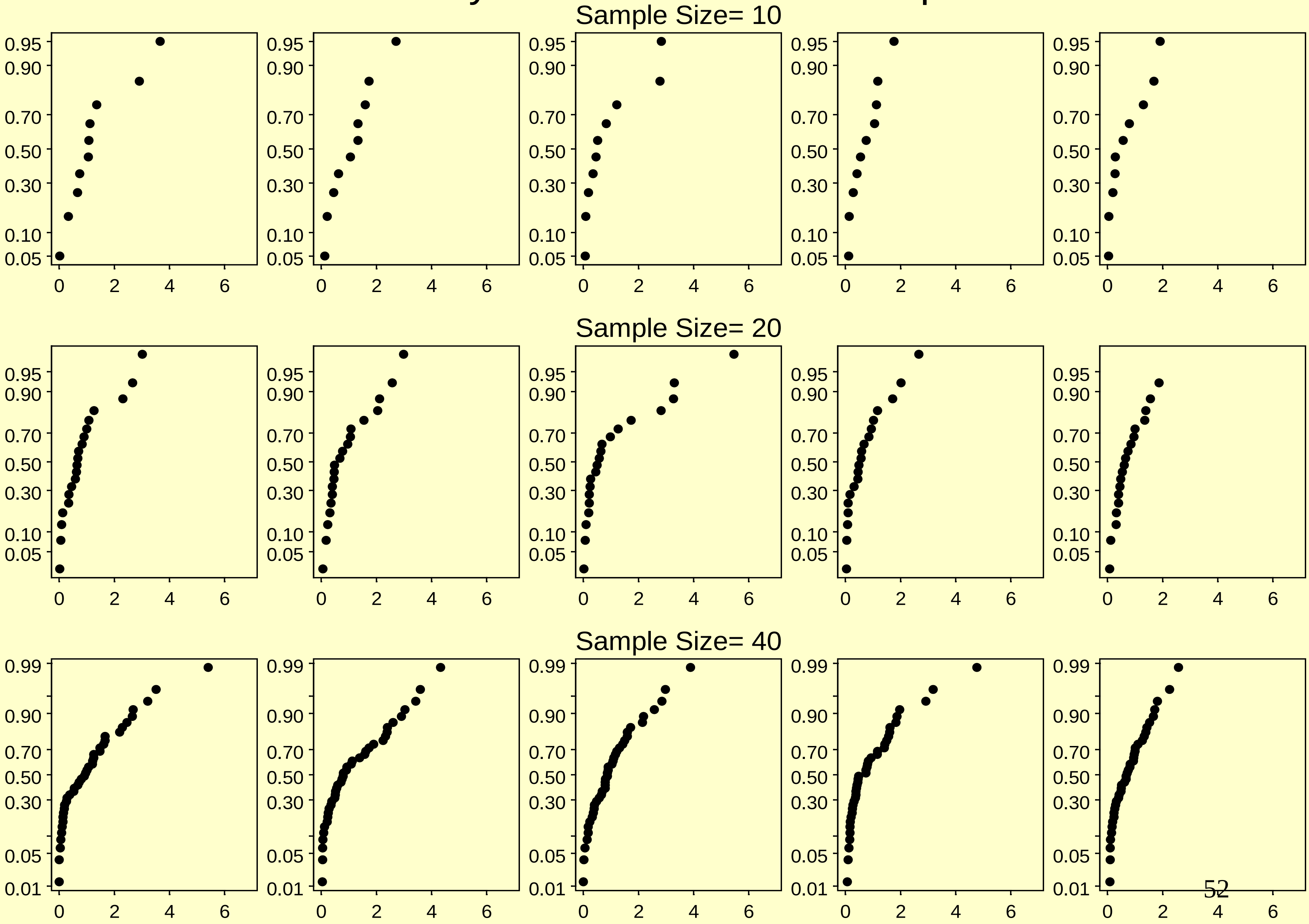
Normal Probability Plot with Simulated Normal Data



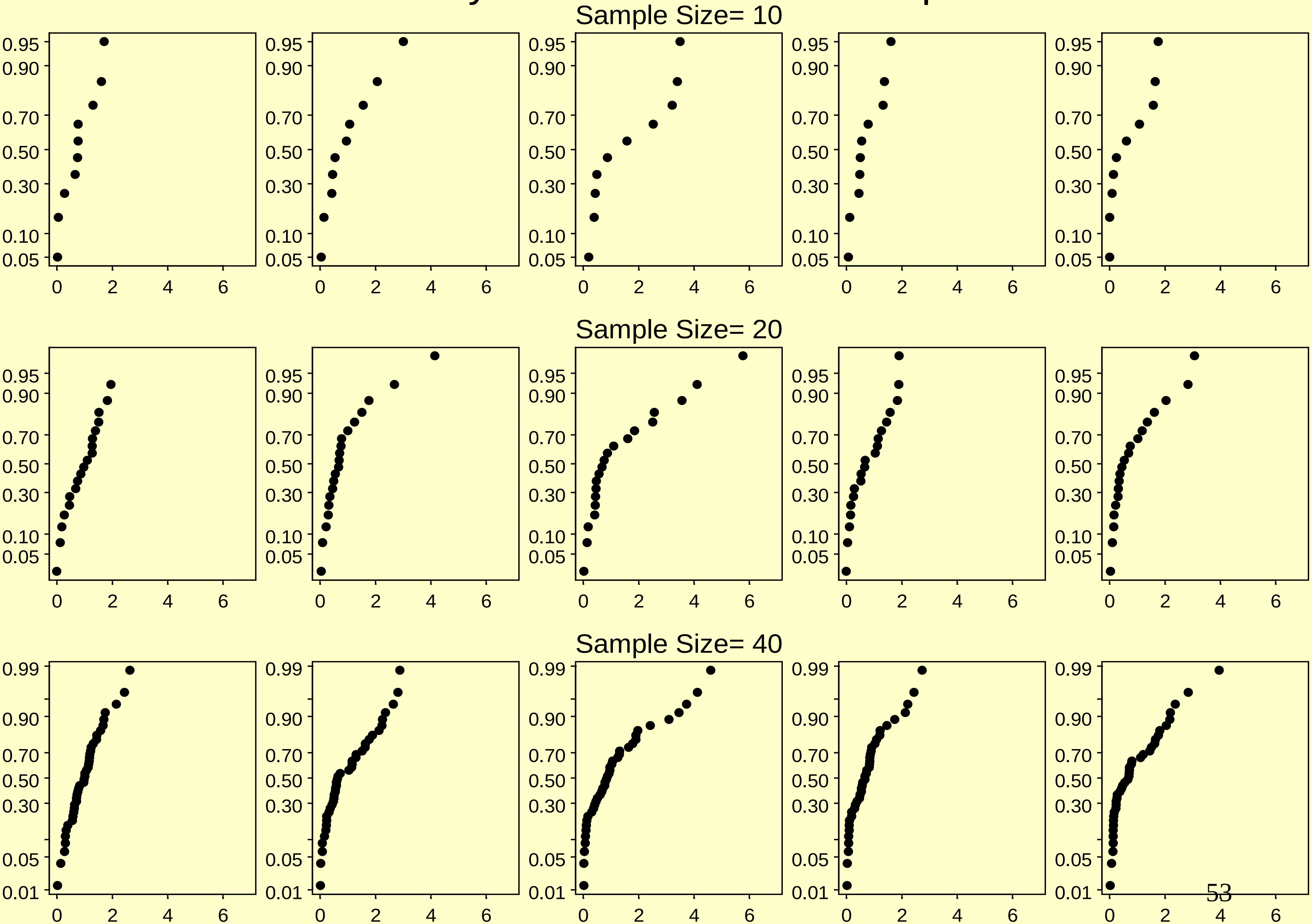


Normal Probability Plot with Simulated Exponential Data
Sample Size= 10
Sample Size= 20
Sample Size= 40
0.95
0.90
0.70
0.50
0.30
0.10
0.05
0.99
0.01
0
2
4
6

Normal Probability Plot with Simulated Exponential Data



Normal Probability Plot with Simulated Exponential Data

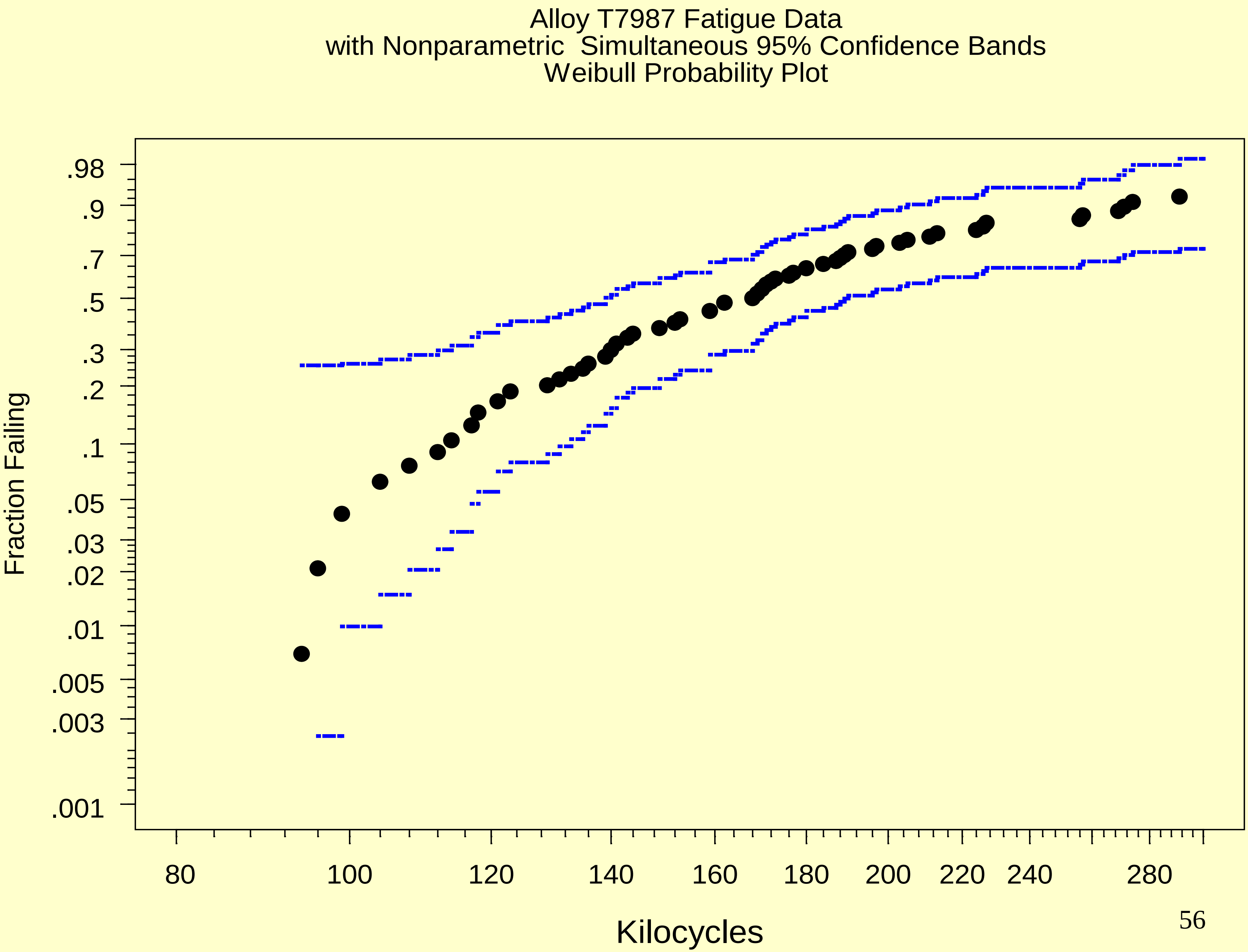


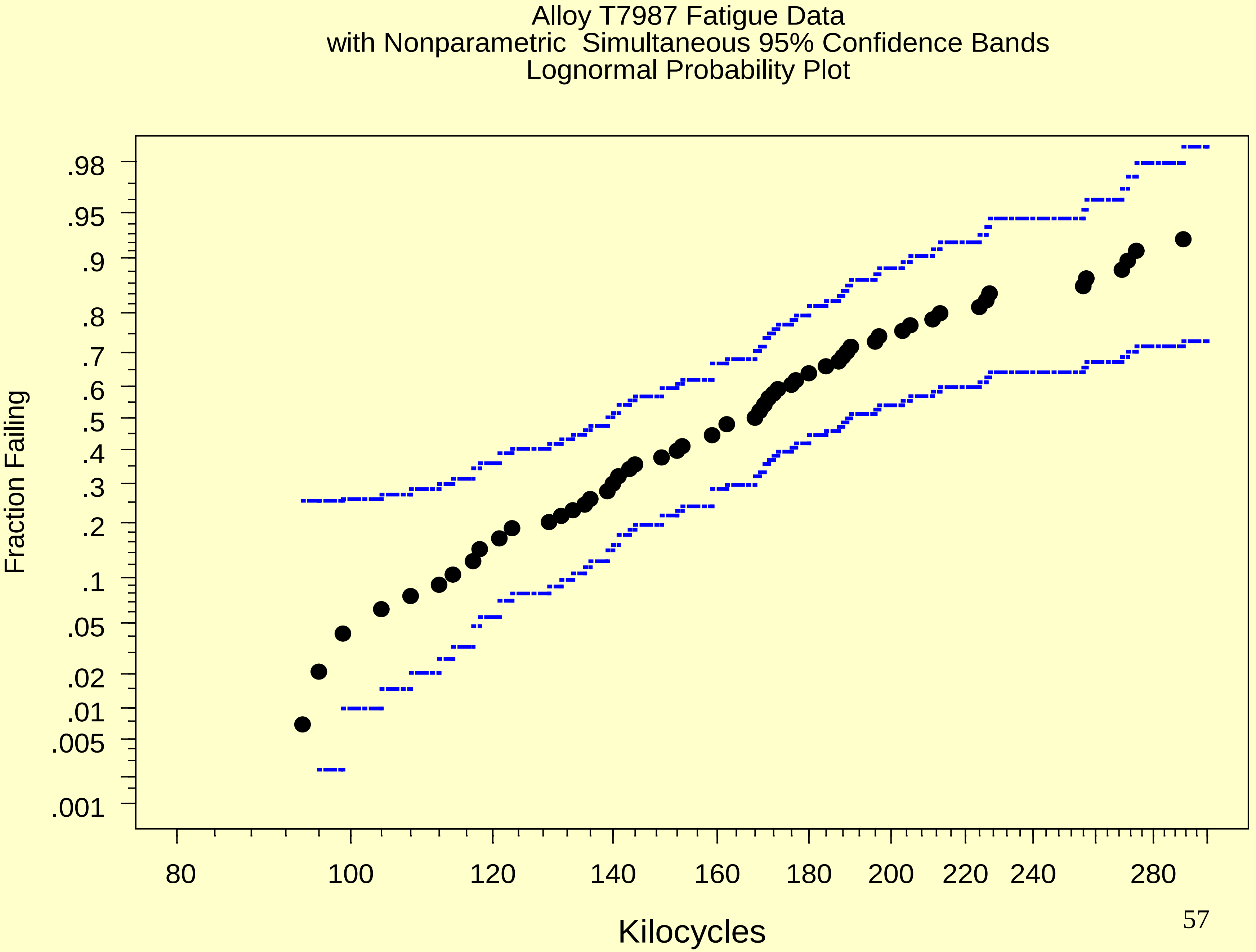
Lessons Learned

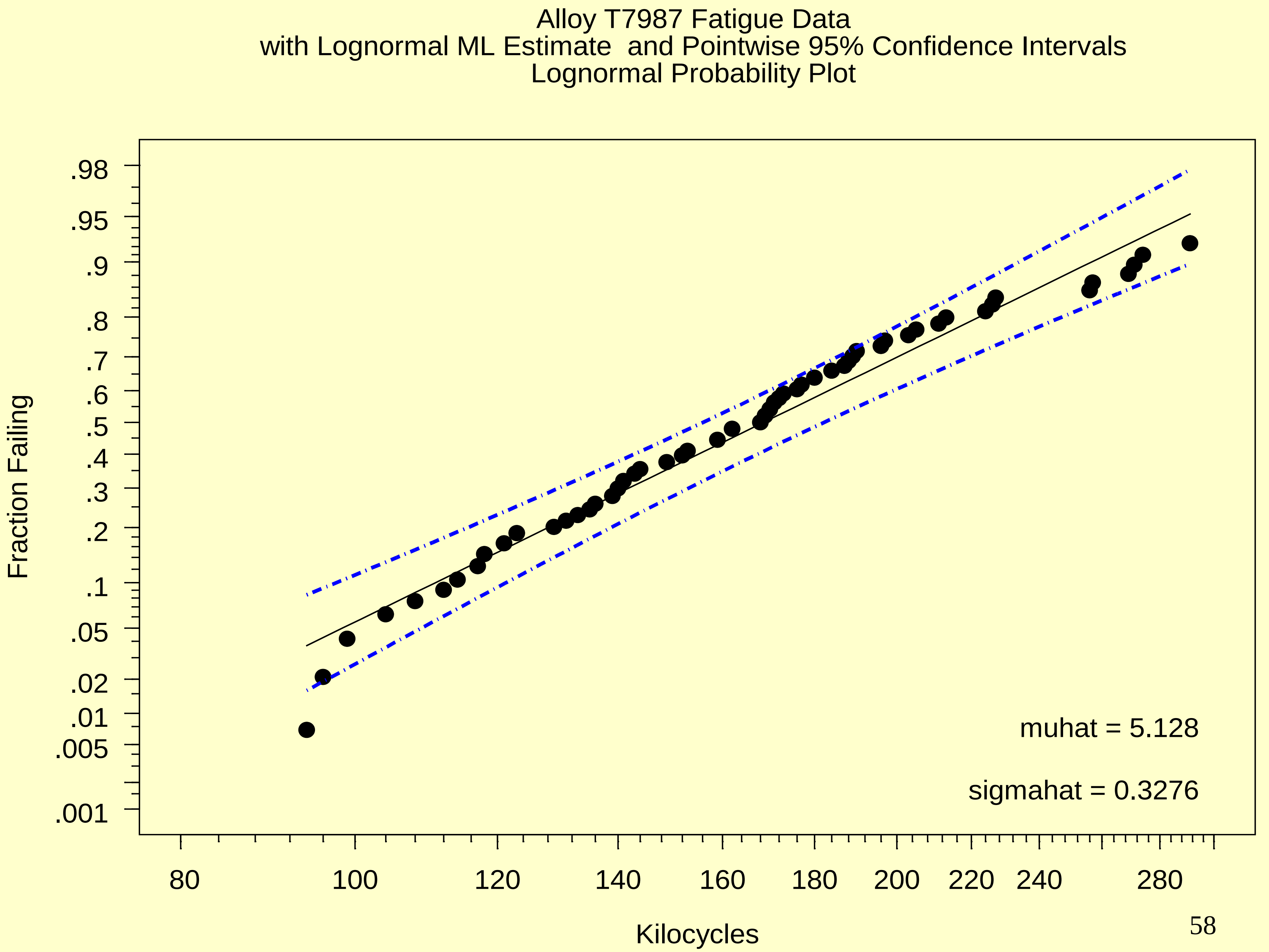
- Probability plots are an important took for (reliability) data analysis
- Simulation can rapidly provide useful experiences and insights for dealing with the interpretation of statistical variability

Alloy T7987 Fatigue Data

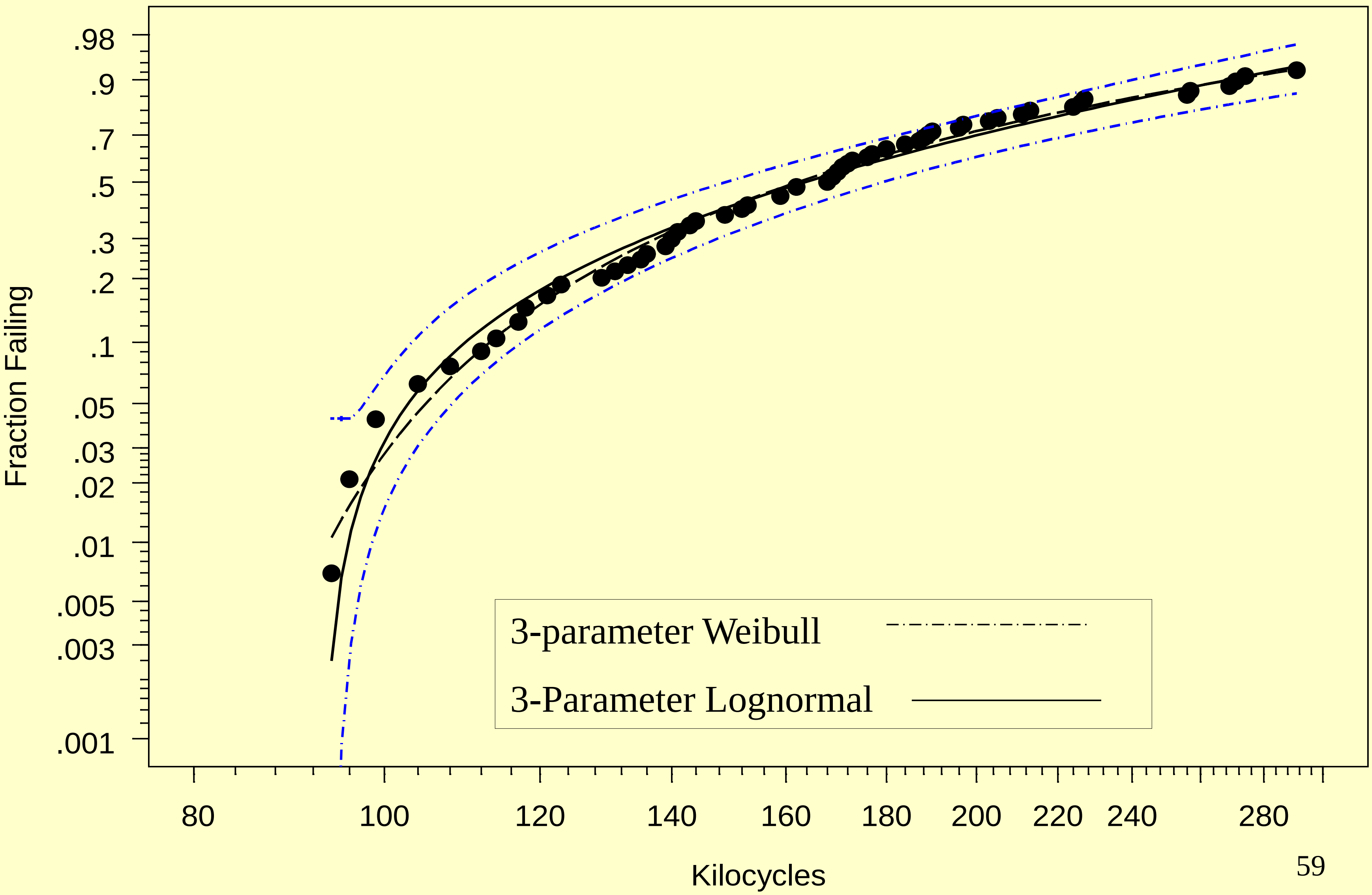
- Data from Meeker and Escobar (1998)
- Dogbone shaped specimens
- Test run until 300 thousand cycles.
  - 67 failures;
  - 5 right-censored observations
- Need to estimate cycles-to-failure distribution.
- Primary interest in the lower tail of the distribution.


Alloy T7987 Fatigue Data
with Nonparametric Simultaneous 95% Confidence Bands
Weibull Probability Plot
.98
.9
.7
.5
.3
.2
.1
.05
.03
.02
.01
.005
.003
.001
Fraction Failing
80
100
120
140
160
180
200
220
240
280
Kilocycles


Alloy T7987 Fatigue Data
with Nonparametric Simultaneous 95% Confidence Bands
Lognormal Probability Plot
Fraction Failing
.98
.95
.9
.8
.7
.6
.5
.4
.3
.2
.1
.05
.02
.01
.005
.001
80
100
120
140
160
180
200
220
240
280
Kilocycles


Alloy T7987 Fatigue Data
with Lognormal ML Estimate and Pointwise 95% Confidence Intervals
Lognormal Probability Plot
Fraction Failing
.98
.95
.9
.8
.7
.6
.5
.4
.3
.2
.1
.05
.02
.01
.005
.001
80
100
120
140
160
180
200
220
240
280
Kilocycles
muhat = 5.128
sigmahat = 0.3276

Alloy T7987 Fatigue Data with sevgets ML Estimate
and Pointwise 95% Confidence Intervals
Weibull Probability Plot

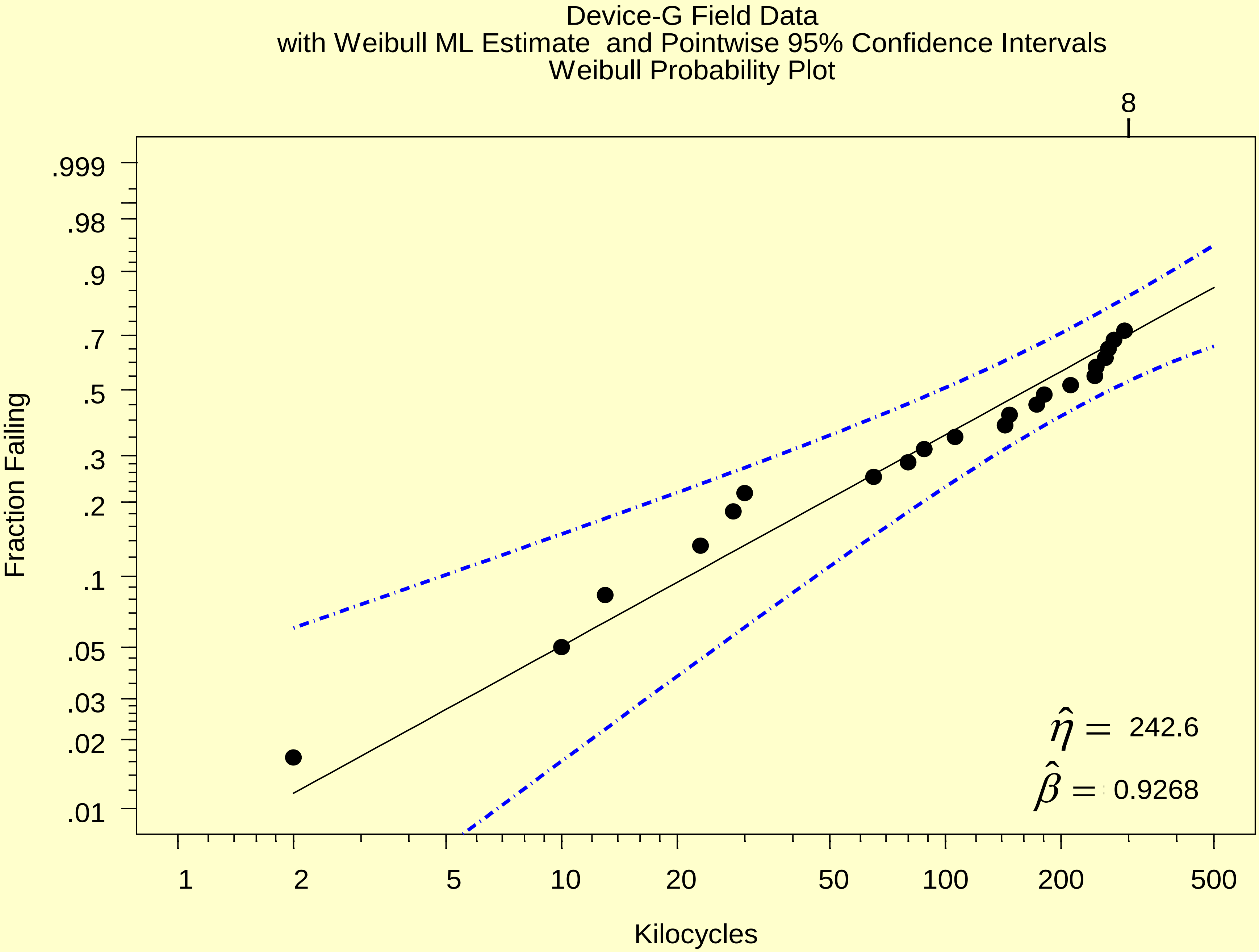


Lessons Learned

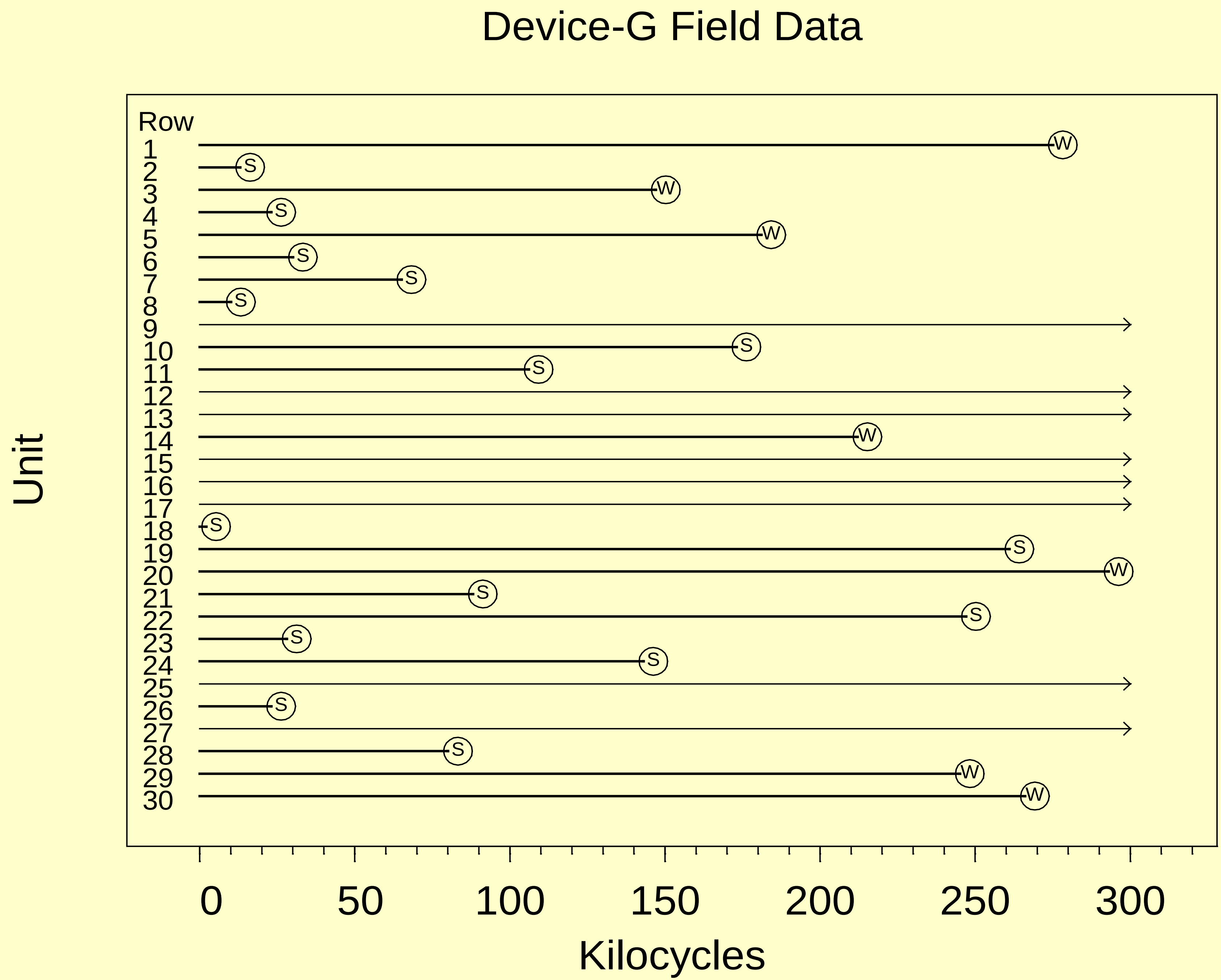
- Fitting a three-parameter Weibull or Lognormal distribution might provide a better fit
- In some cases the use of a three-parameter distribution can be justified because it is known that there is an initial period where the probability of a failure is 0
- Fitting a three-parameter distribution can lead to anti-conservative inferences

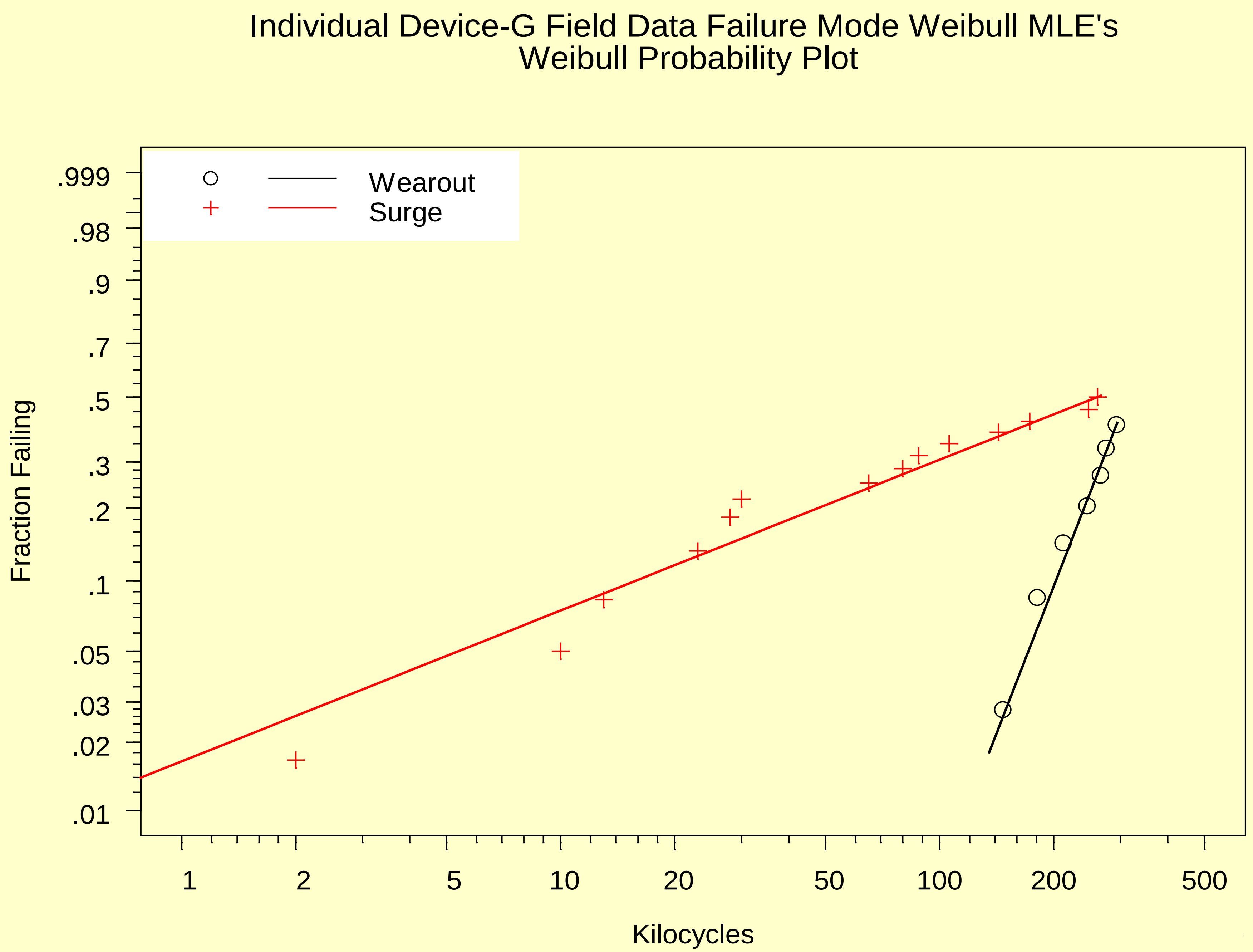
Device-G Field-Tracking Data

- Data from Meeker and Escobar (1998)
- Design life of 300 thousand cycles
- Units were failing in the field more rapidly than had been expected.
- Needs: information on how to improve reliability and an estimate of device MTTF.
- Two failure modes: surge and wear

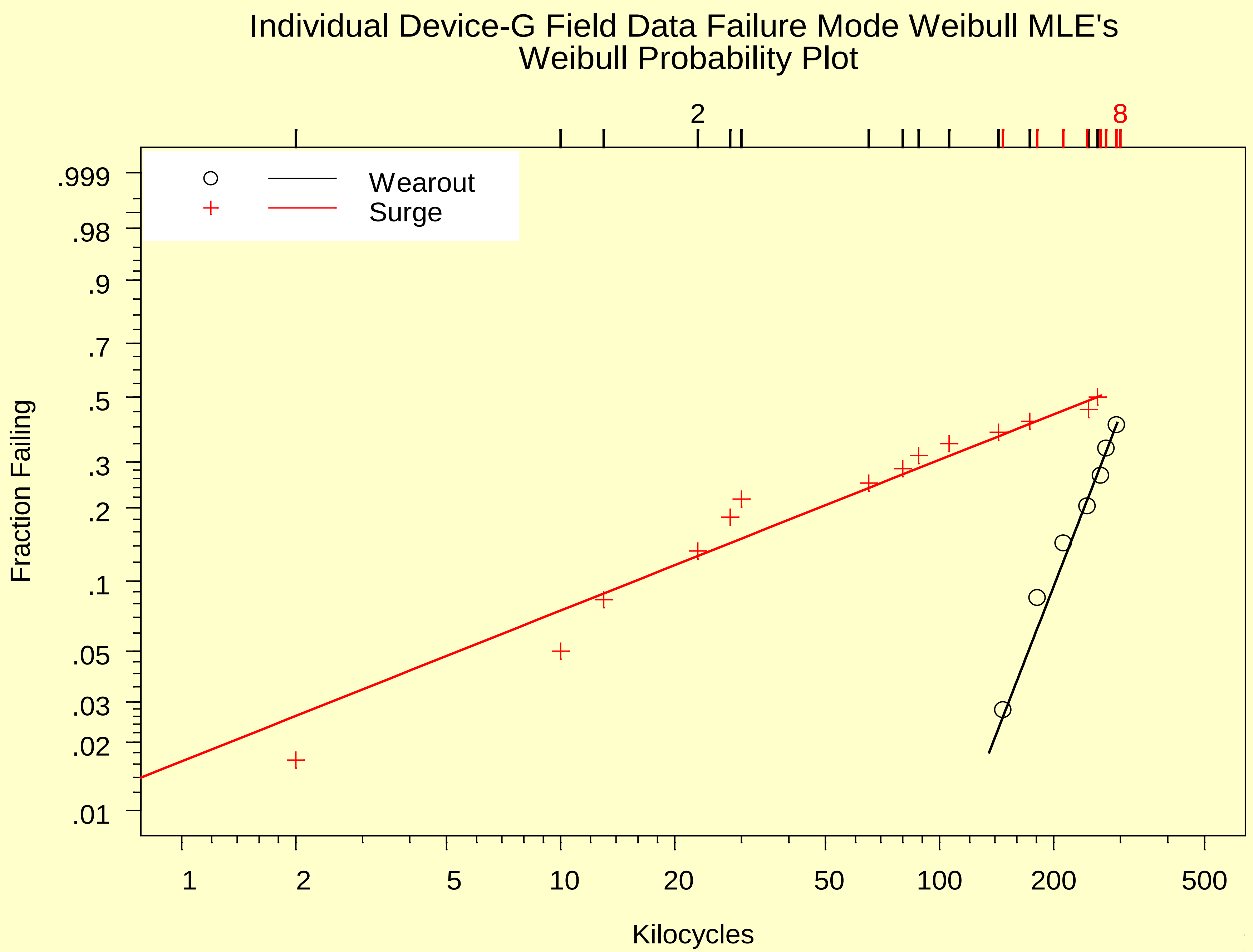


Device-G Field Data
with Weibull ML Estimate and Pointwise 95% Confidence Intervals
Weibull Probability Plot
8
.999
.98
.9
.7
.5
.3
.2
.1
.05
.03
.02
.01
Fraction Failing
1
2
5
10
20
50
100
200
500
Kilocycles
$\hat{\eta} =$ 242.6
$\hat{\beta} =$ 0.9268

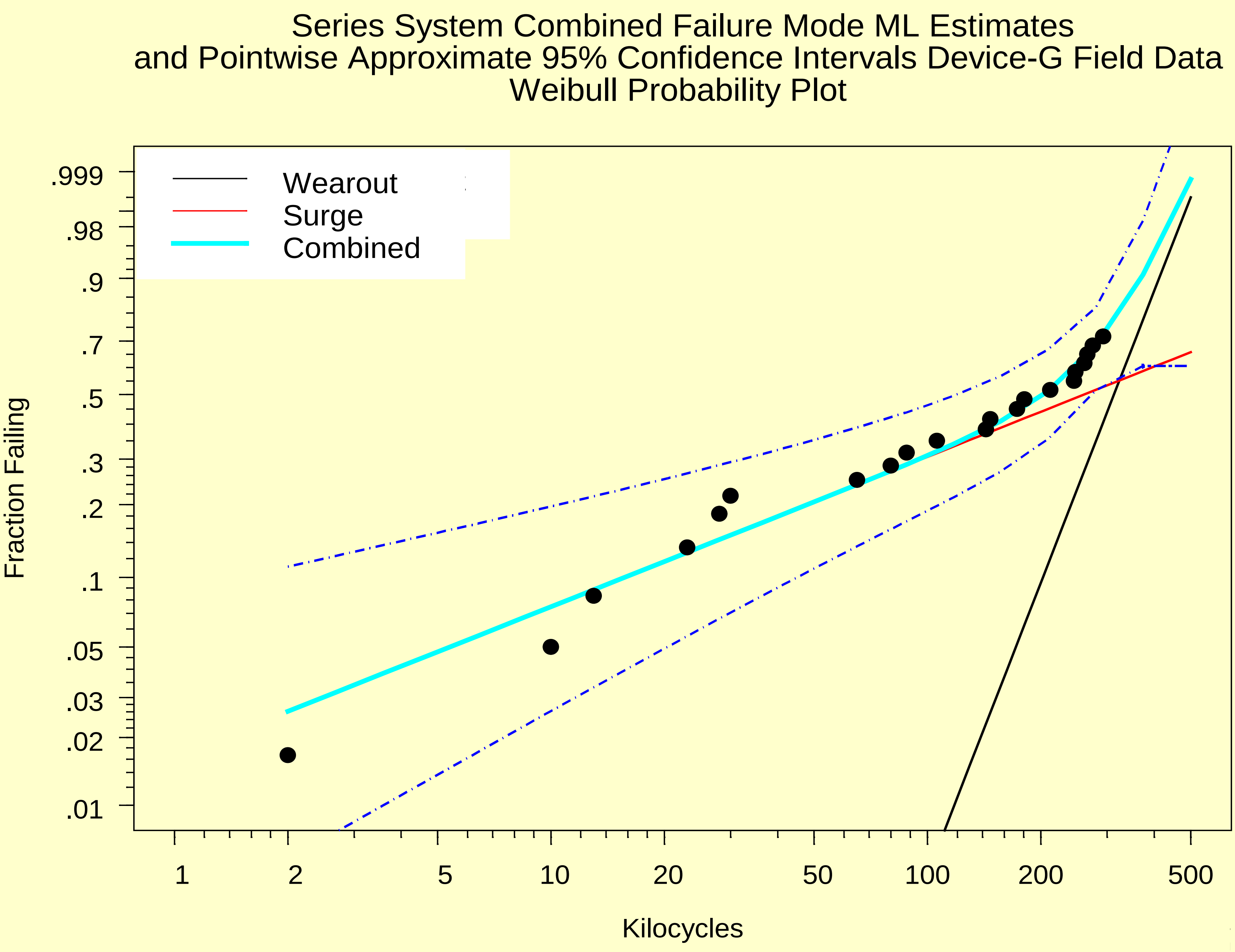

Device-G Field Data
Row
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
W
S
W
S
W
S
S
S
S
S
W
S
S
W
S
S
S
S
S
W
W
Unit
0
50
100
150
200
250
300
Kilocycles



Individual Device-G Field Data Failure Mode Weibull MLE's
Weibull Probability Plot
Wearout
Surge
Fraction Failing
.999
.98
.9
.7
.5
.3
.2
.1
.05
.03
.02
.01
1
2
5
10
20
50
100
200
500
Kilocycles



Individual Device-G Field Data Failure Mode Weibull MLE's
Weibull Probability Plot
2
8
Wearout
Surge
.999
.98
.9
.7
.5
.3
.2
.1
.05
.03
.02
.01
Fraction Failing
1
2
5
10
20
50
100
200
500
Kilocycles

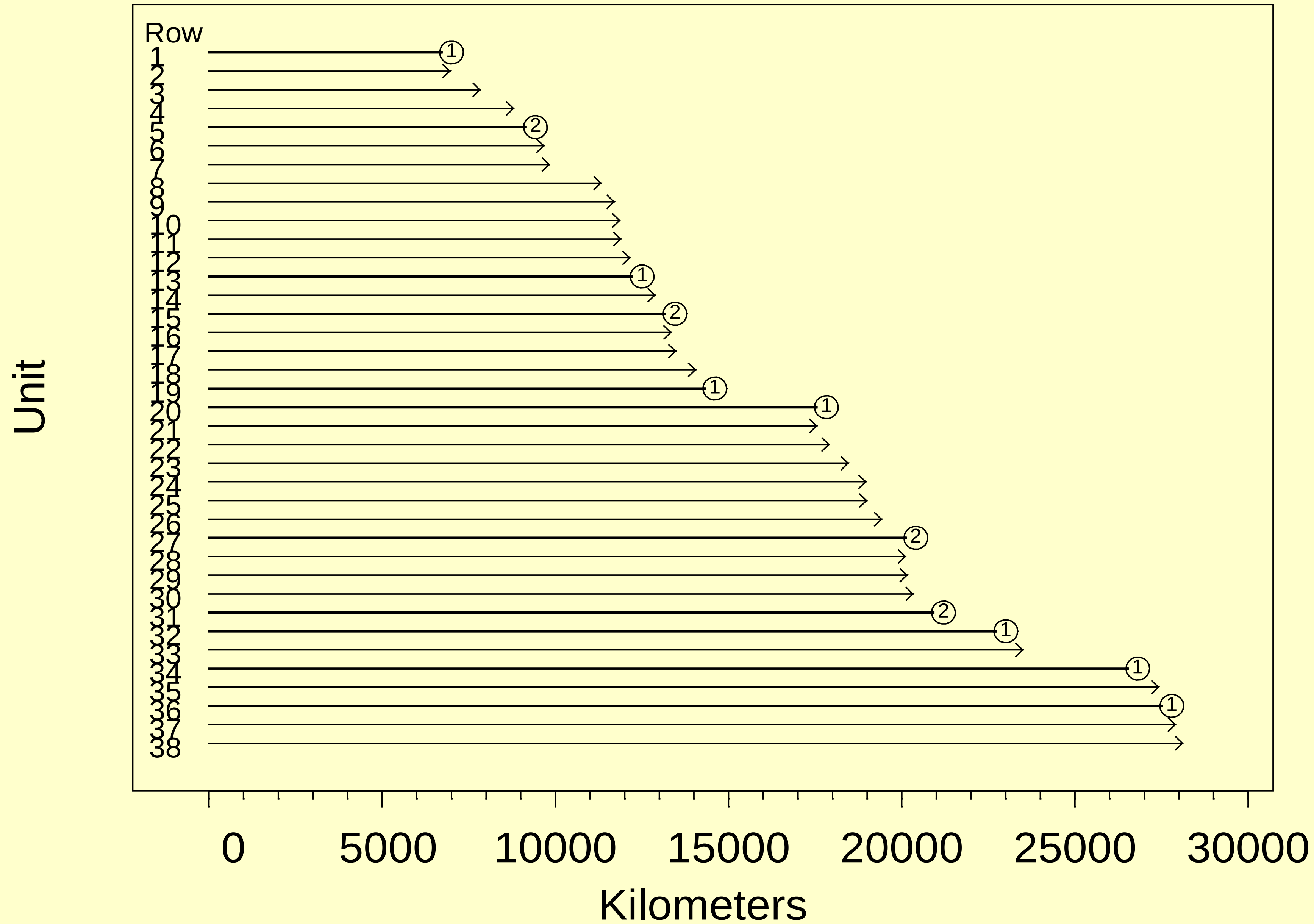


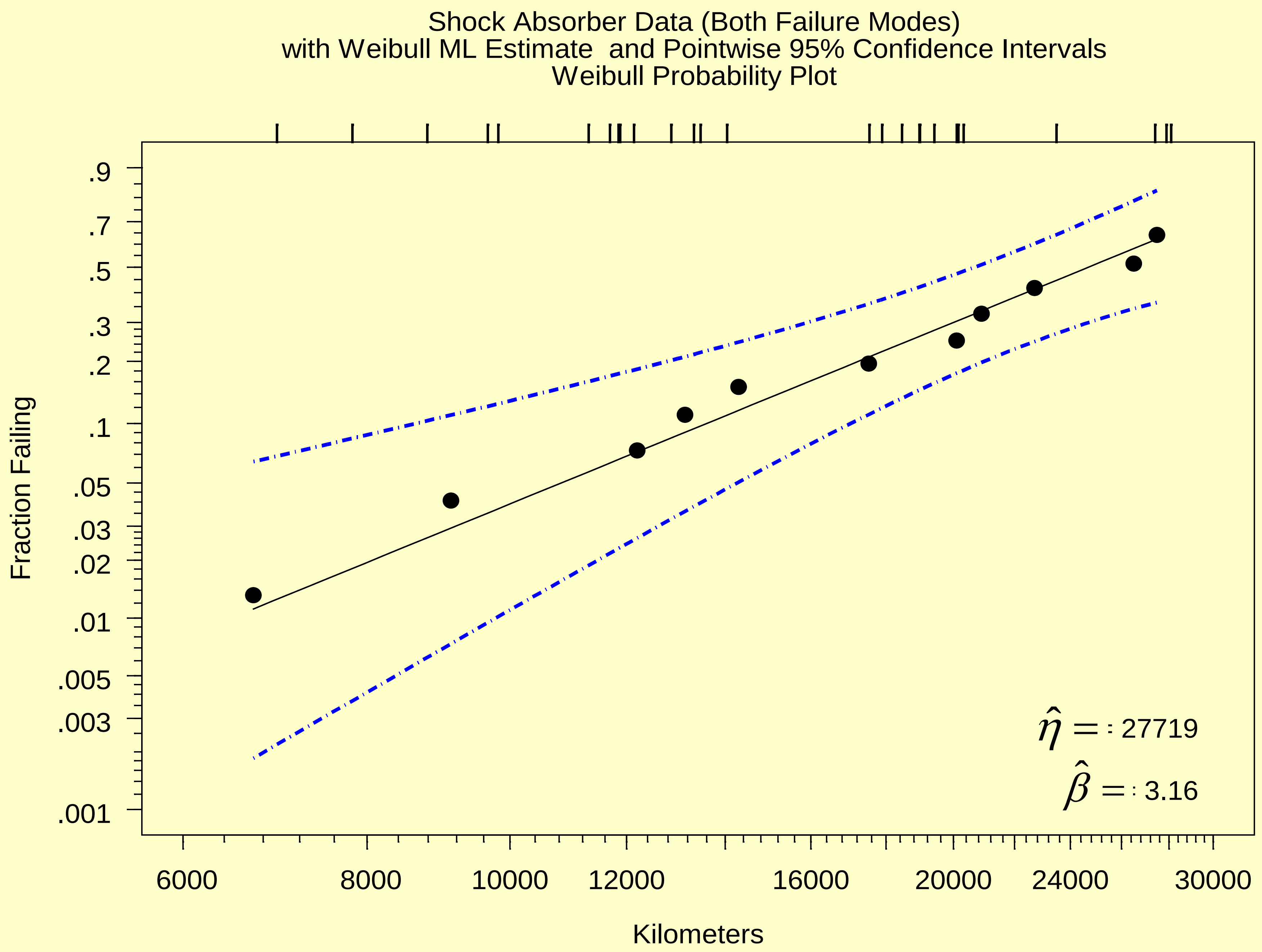
Series System Combined Failure Mode ML Estimates
and Pointwise Approximate 95% Confidence Intervals Device-G Field Data
Weibull Probability Plot
Wearout
Surge
Combined
Fraction Failing
.999
.98
.9
.7
.5
.3
.2
.1
.05
.03
.02
.01
1
2
5
10
20
50
100
200
500
Kilocycles

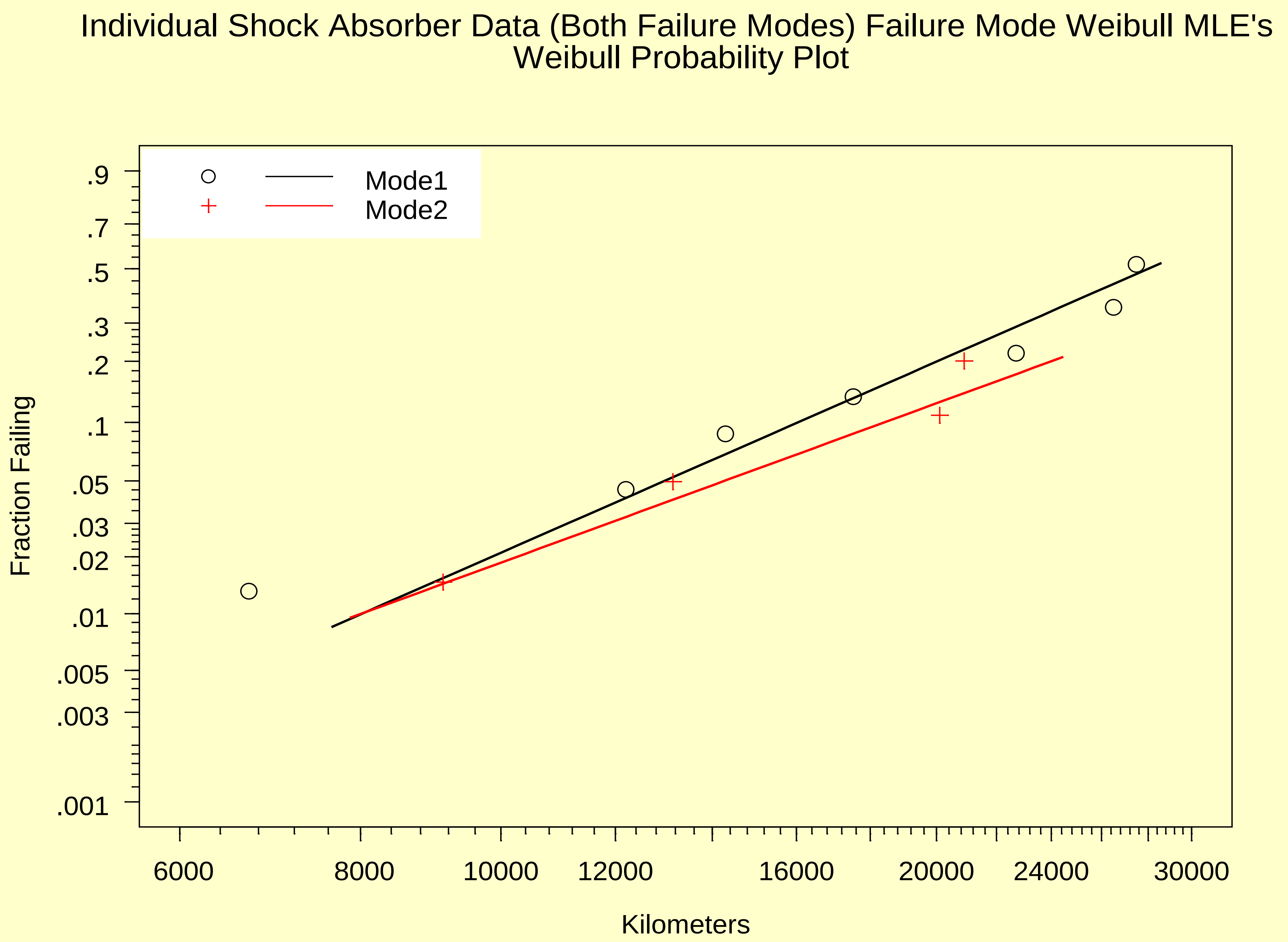
Shock Absorber Field Data

- Data from O'Connor (1985)
- Two failure modes (identified as Mode 1 and Mode 2)
- Need to estimate the failure-time distribution for the shock absorbers

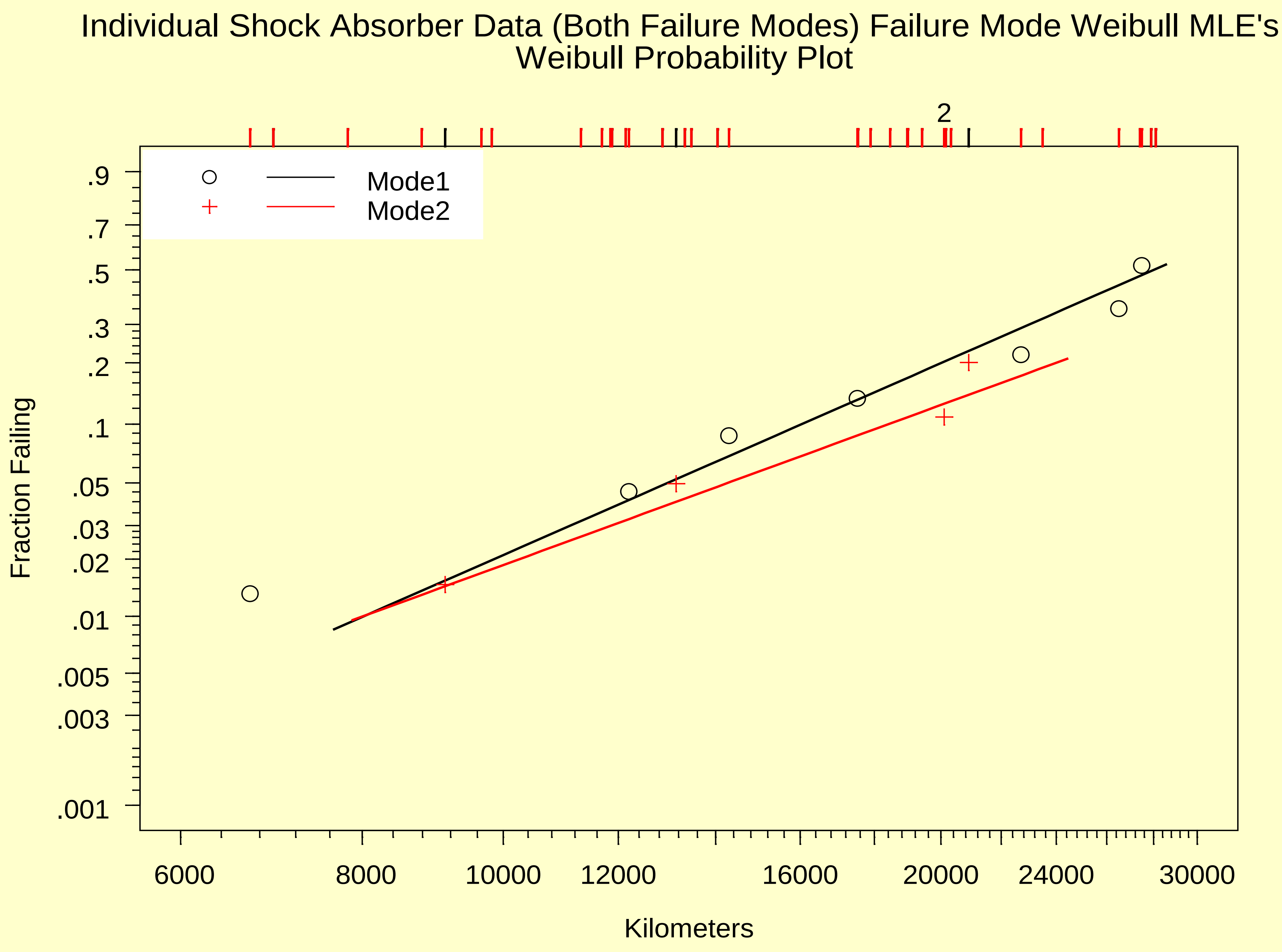
Shock Absorber Data (Both Failure Modes)

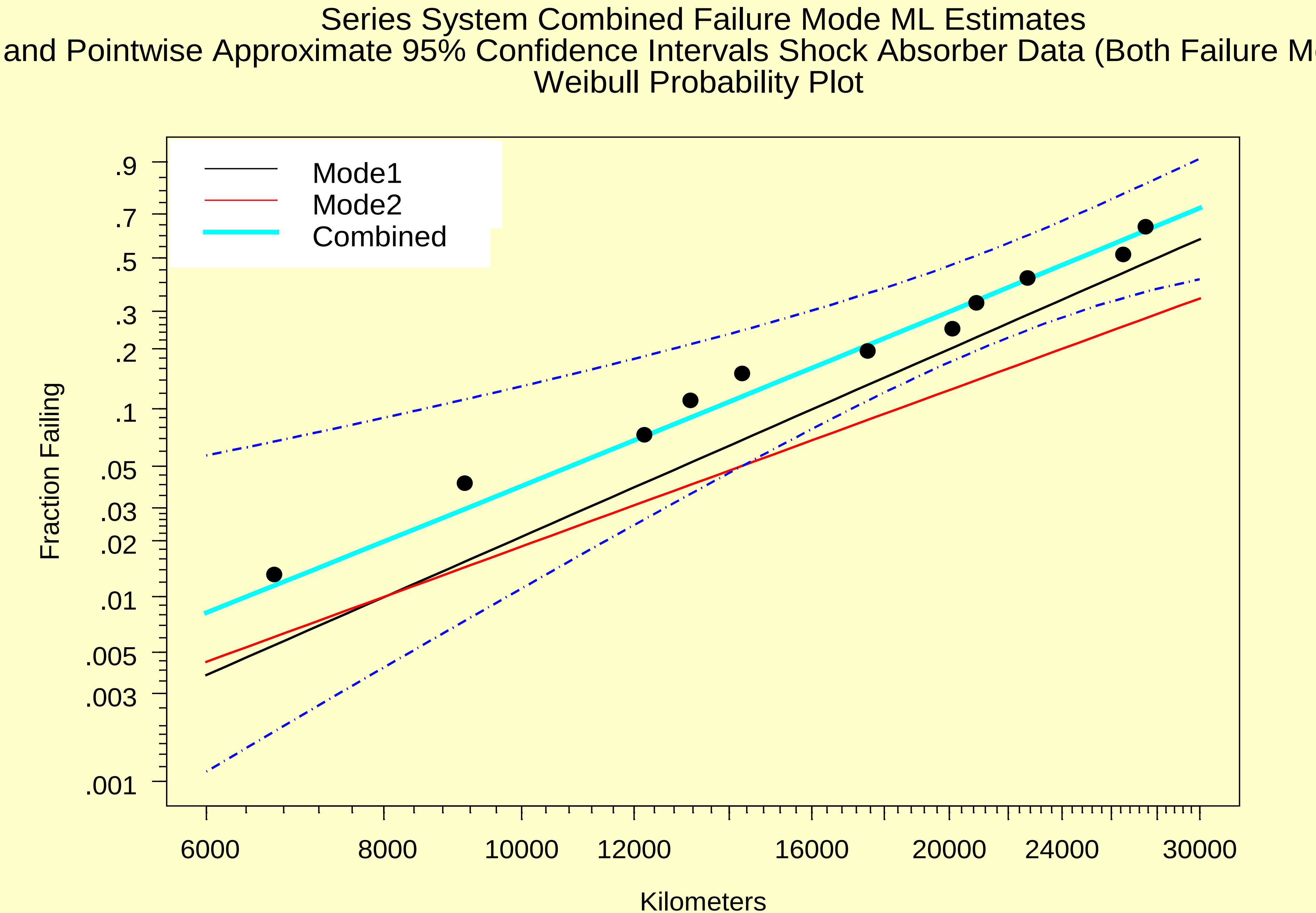



Shock Absorber Data (Both Failure Modes)
with Weibull ML Estimate and Pointwise 95% Confidence Intervals
Weibull Probability Plot
Fraction Failing
.9
.7
.5
.3
.2
.1
.05
.03
.02
.01
.005
.003
.001
6000
8000
10000
12000
16000
20000
24000
30000
Kilometers
$\hat{\eta} = 27719$
$\hat{\beta} = 3.16$

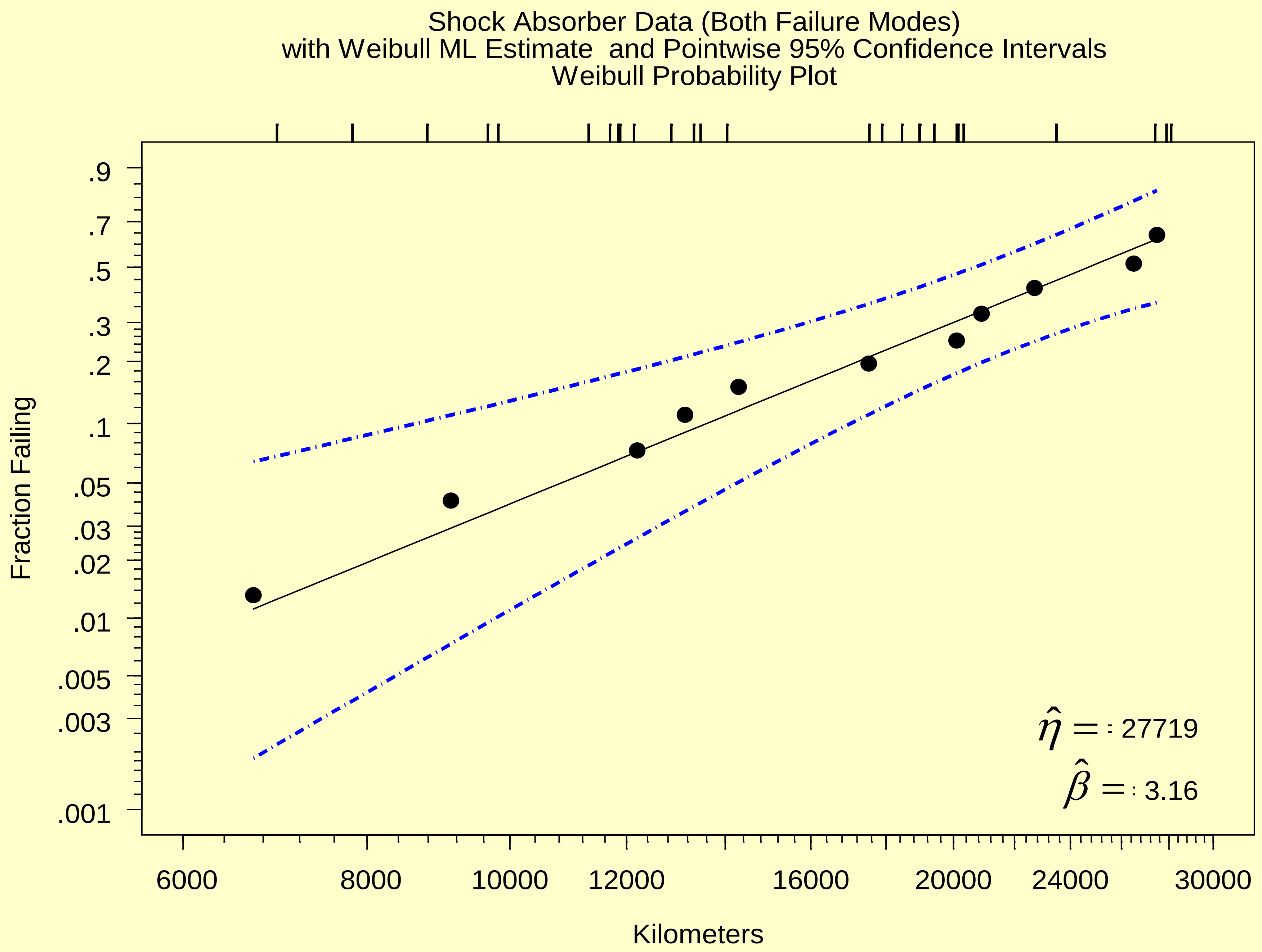


Individual Shock Absorber Data (Both Failure Modes) Failure Mode Weibull MLE's
Weibull Probability Plot
Mode1
Mode2
.9
.7
.5
.3
.2
.1
.05
.03
.02
.01
.005
.003
.001
Fraction Failing
6000
8000
10000
12000
16000
20000
24000
30000
Kilometers


Individual Shock Absorber Data (Both Failure Modes) Failure Mode Weibull MLE's
Weibull Probability Plot
2
Mode1
Mode2
.9
.7
.5
.3
.2
.1
.05
.03
.02
.01
.005
.003
.001
Fraction Failing
6000
8000
10000
12000
16000
20000
24000
30000
Kilometers



Series System Combined Failure Mode ML Estimates
and Pointwise Approximate 95% Confidence Intervals Shock Absorber Data (Both Failure M
Weibull Probability Plot
Mode1
Mode2
Combined
Fraction Failing
.9
.7
.5
.3
.2
.1
.05
.03
.02
.01
.005
.003
.001
6000
8000
10000
12000
16000
20000
24000
30000
Kilometers



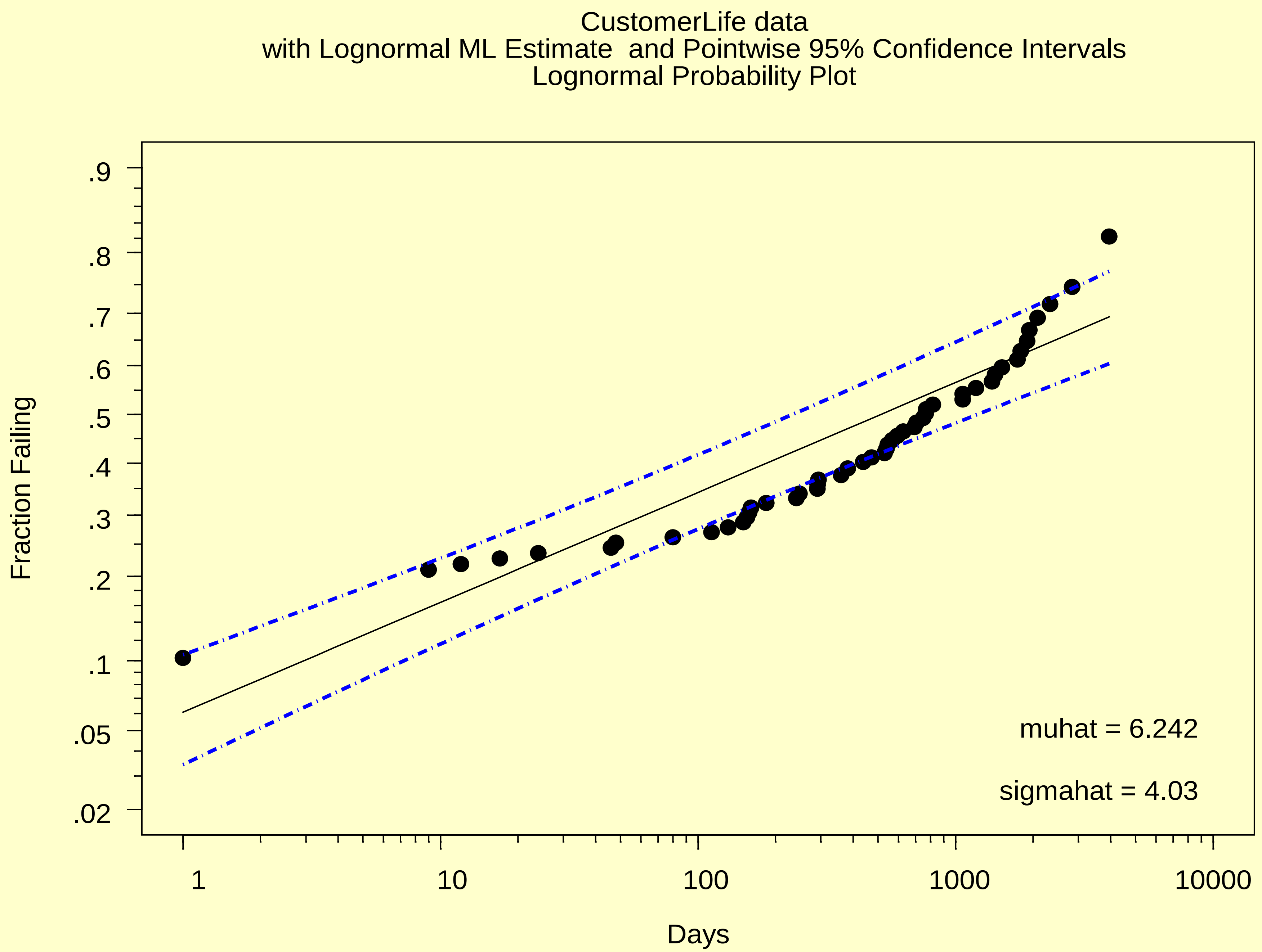
Shock Absorber Data (Both Failure Modes)
with Weibull ML Estimate and Pointwise 95% Confidence Intervals
Weibull Probability Plot
Fraction Failing
.9
.7
.5
.3
.2
.1
.05
.03
.02
.01
.005
.003
.001
6000
8000
10000
12000
16000
20000
24000
30000
Kilometers
$\hat{\eta} =$ 27719
$\hat{\beta} =$ 3.16

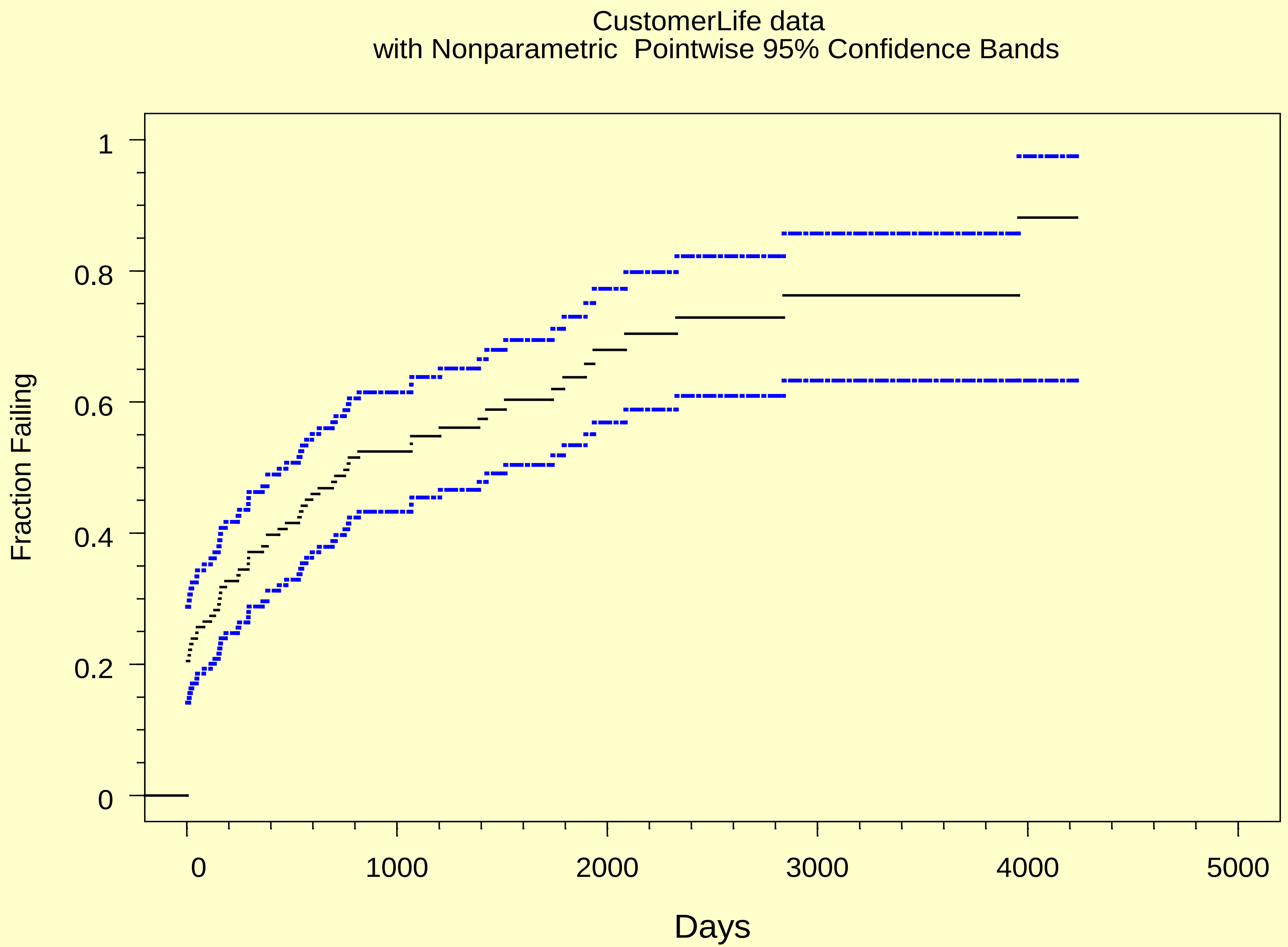
Lessons Learned

- If there is more than one failure mode, it is important to separate and analyze failure modes separately when
  - Shape parameters are very different
  - There is need to assess the impact of eliminating a failure mode

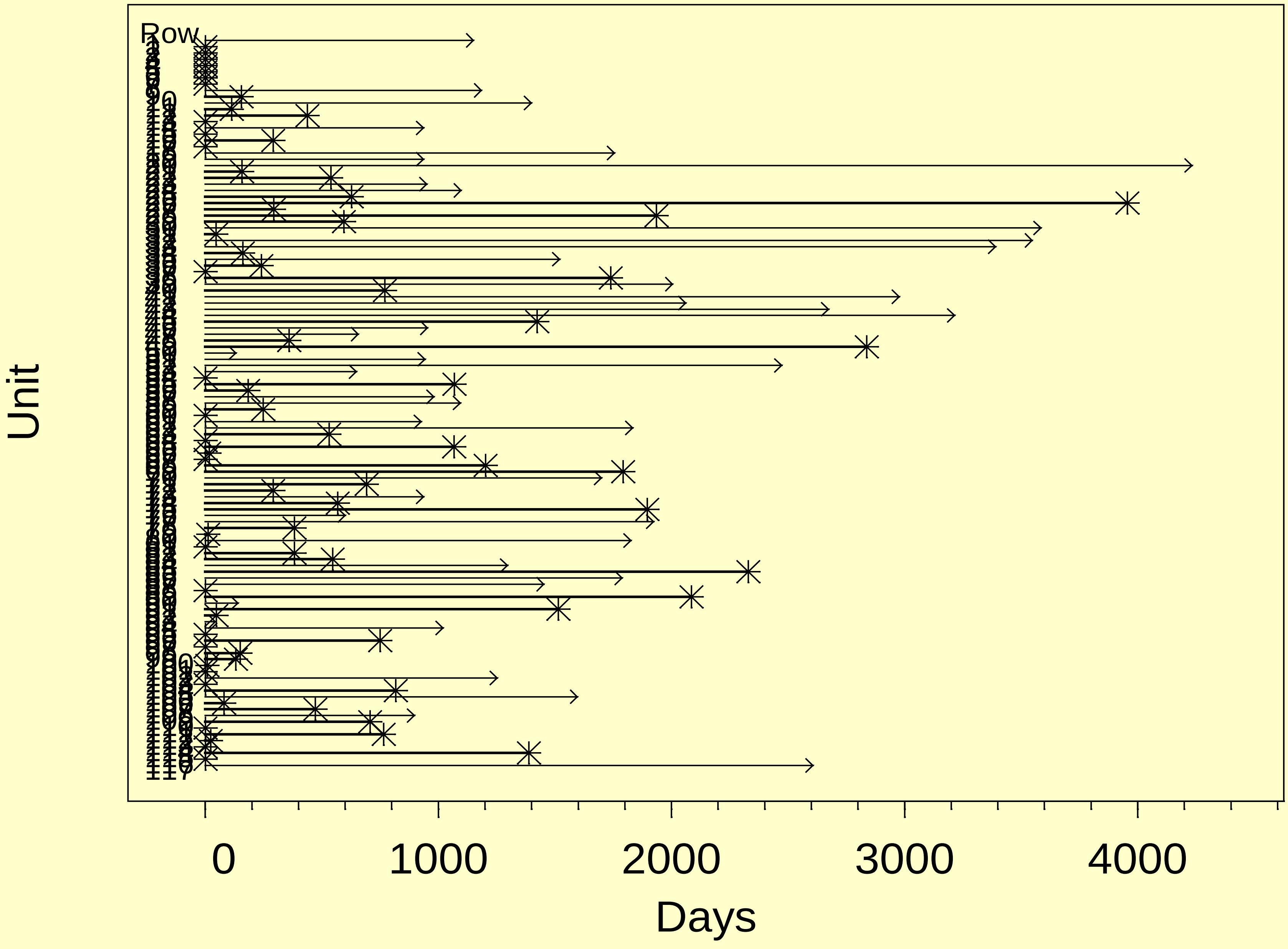
Estimating the Distribution of "Customer Life"

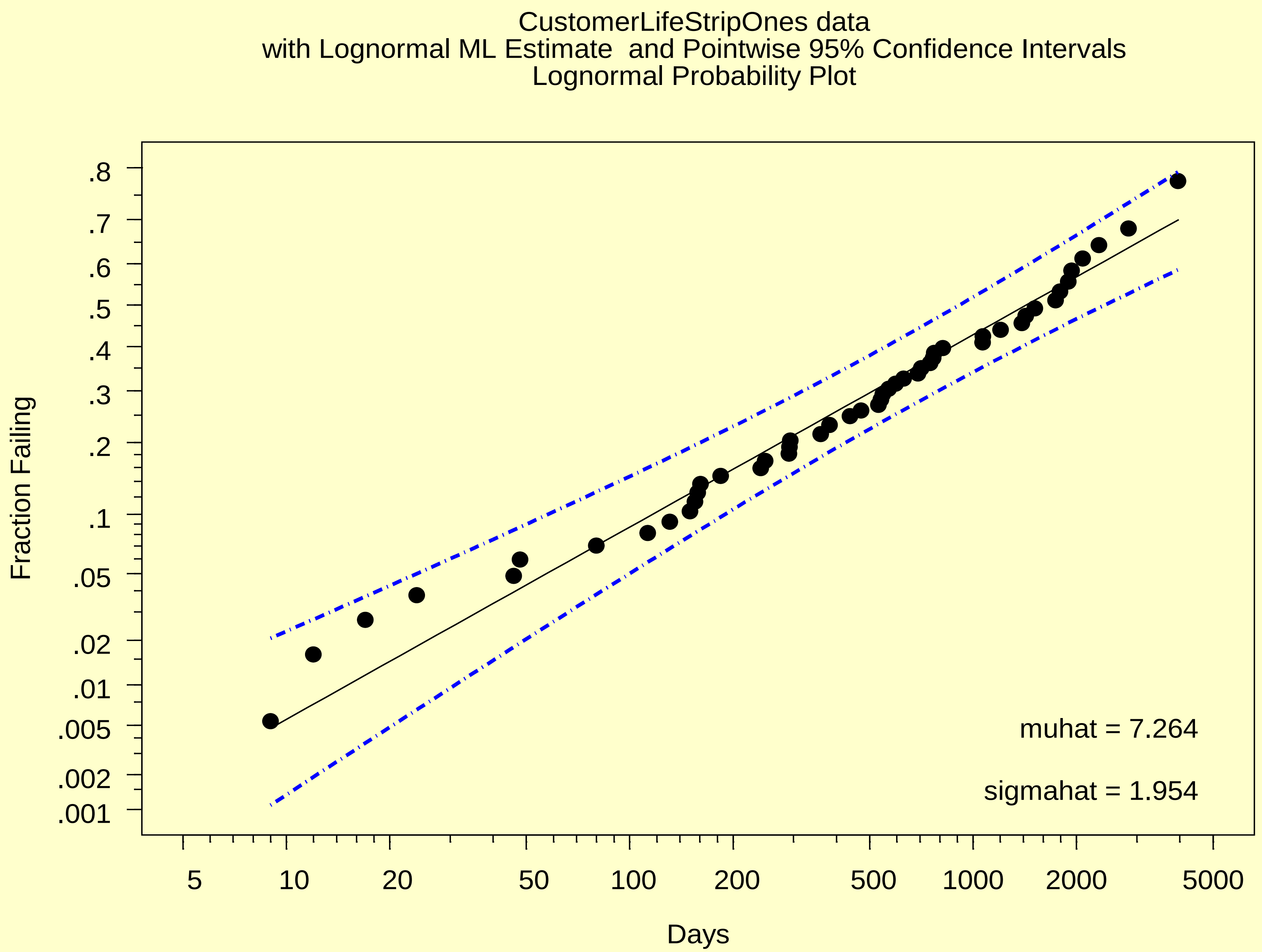
- Data available on all customer transactions for 12 years
- Need to estimate the distribution of remaining life for all current customers on January 1, 2001
- Beginning of "customer life" is the date of the first transaction.
- End of "customer life" is defined at the date of the last transaction preceded by a two-year gap.


CustomerLife data
with Lognormal ML Estimate and Pointwise 95% Confidence Intervals
Lognormal Probability Plot
Fraction Failing
.9
.8
.7
.6
.5
.4
.3
.2
.1
.05
.02
1
10
100
1000
10000
Days
muhat = 6.242
sigmahat = 4.03

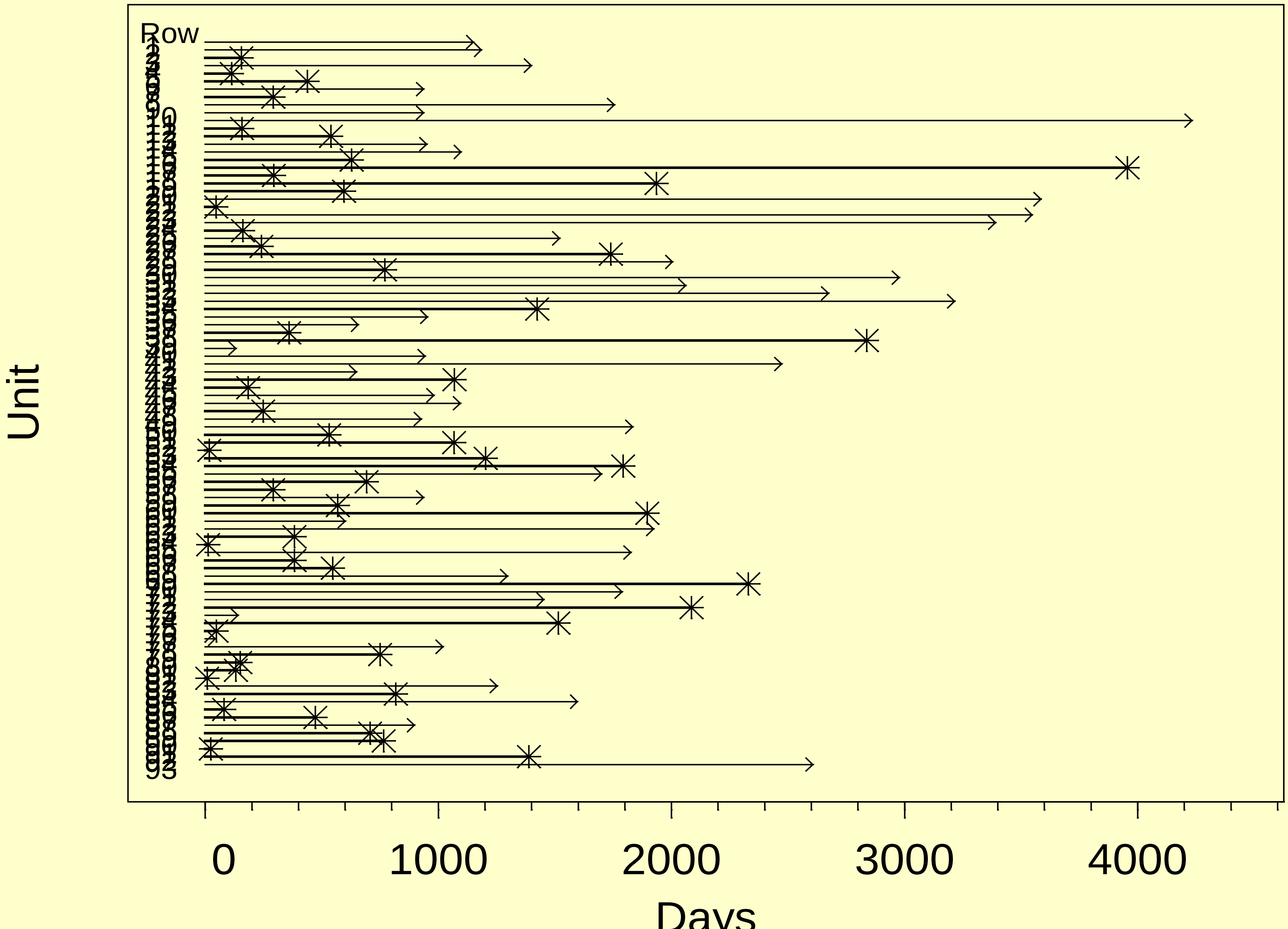

CustomerLife data
with Nonparametric Pointwise 95% Confidence Bands
Fraction Failing
Days
0
0.2
0.4
0.6
0.8
1
0
1000
2000
3000
4000
5000

CustomerLife data




CustomerLifeStripOnes data
with Lognormal ML Estimate and Pointwise 95% Confidence Intervals
Lognormal Probability Plot
Fraction Failing
.8
.7
.6
.5
.4
.3
.2
.1
.05
.02
.01
.005
.002
.001
5
10
20
50
100
200
500
1000
2000
5000
Days
muhat = 7.264
sigmahat = 1.954

CustomerLifeStripOnes data

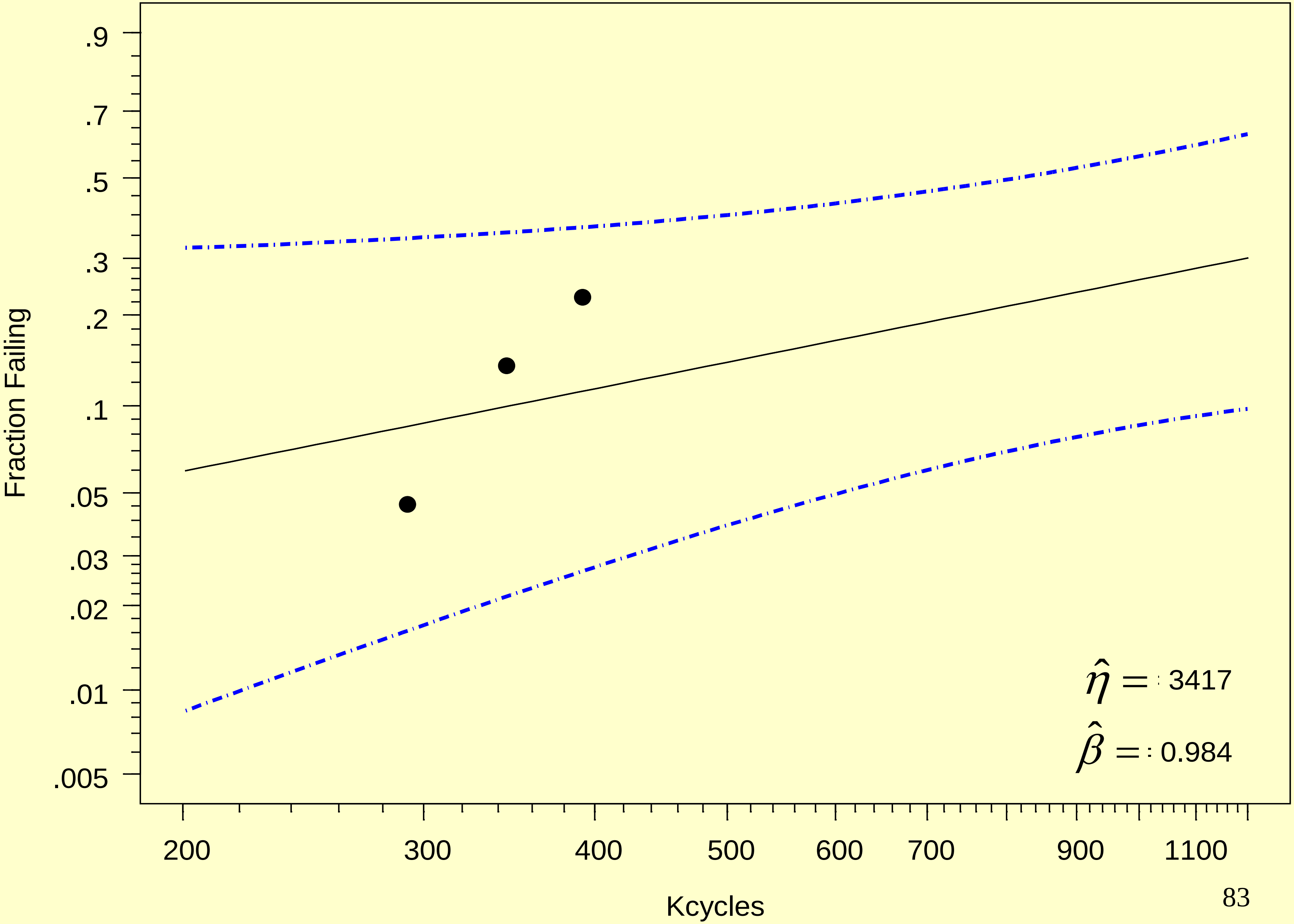


Lessons Learned

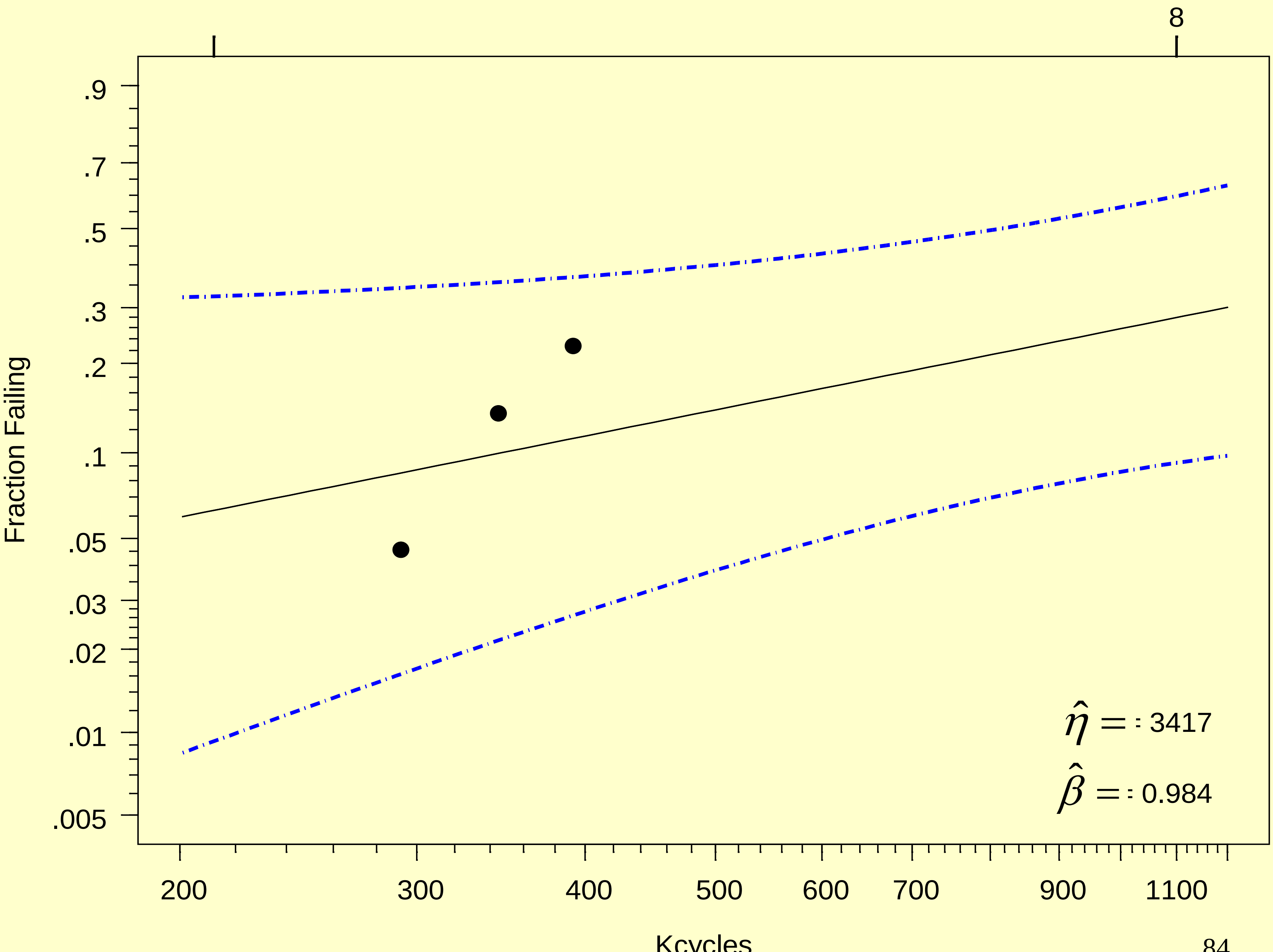
- It will not always be easy to find a distributions that “fits” the data
- A parametric model is generally needed to extrapolate in time
- Properly segmenting data in an appropriate manner may provide a path toward a good model

Bearing-A Bench Test Results

- Data from student in 1999 FTC short course
- Continuous-run test for a newly designed bearing
- Sample of 12 units put on test; one early removal; 3 failures.
- Test terminated at 1100 thousand cycles

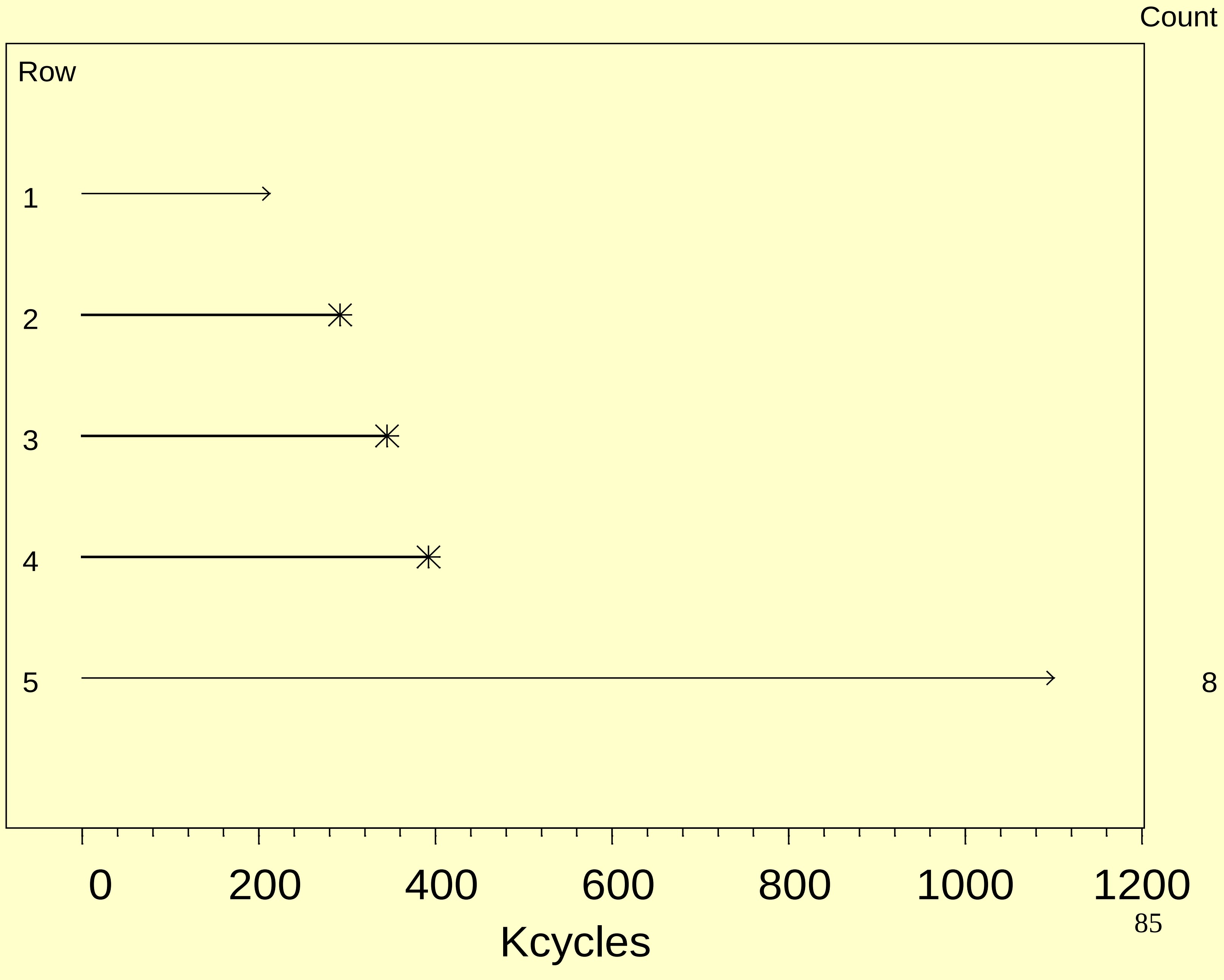


.9
.7
.5
.3
.2
.1
.05
.03
.02
.01
.005
Fraction Failing
200
300
400
500
600
700
900
1100
Kcycles
$\hat{\eta} =$ 3417
$\hat{\beta} =$ 0.984



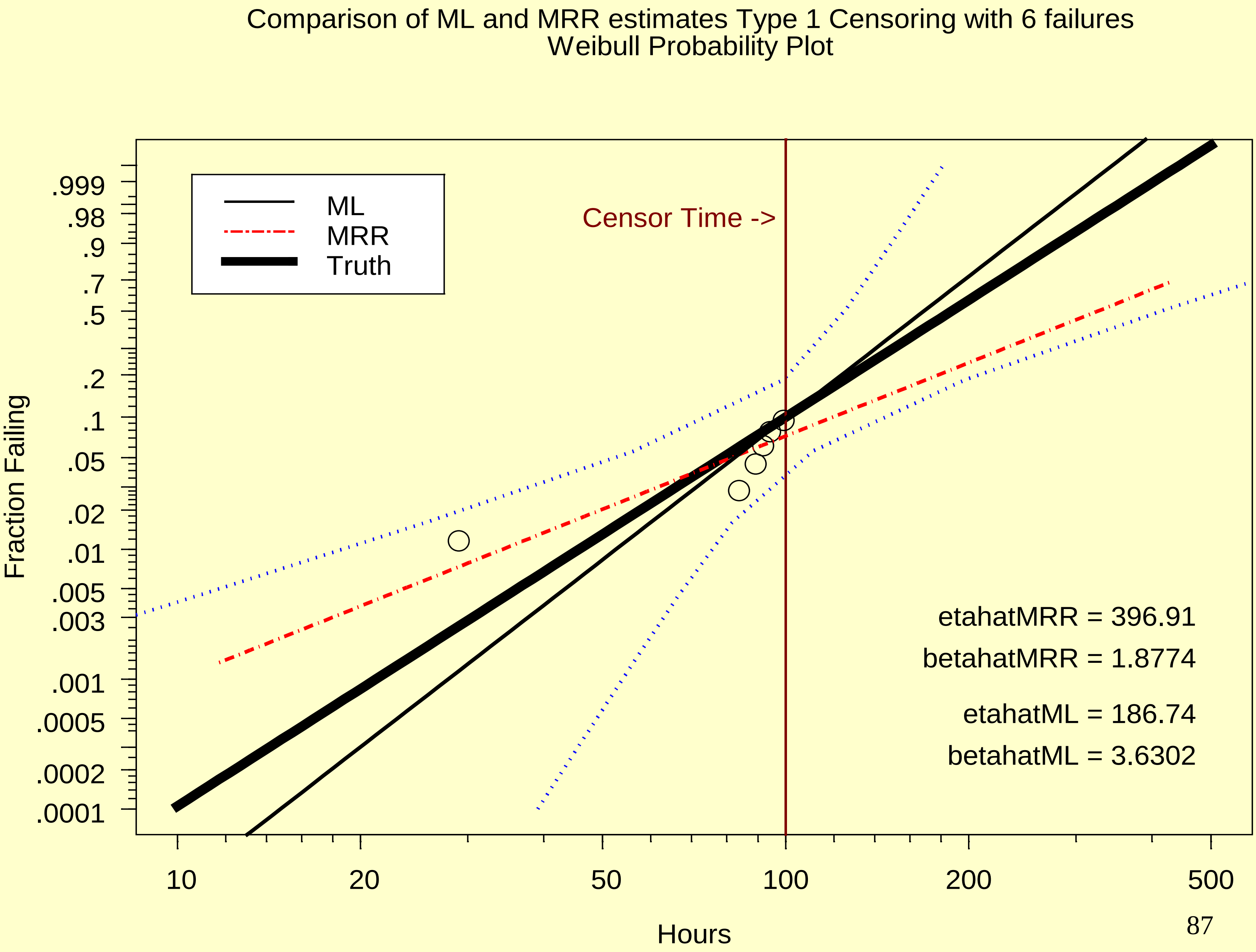
8
.9
.7
.5
.3
.2
.1
.05
.03
.02
.01
.005
Fraction Failing
$\hat{\eta} = 3417$
$\hat{\beta} = 0.984$
200
300
400
500
600
700
900
1100
Kcycles

BearingA data

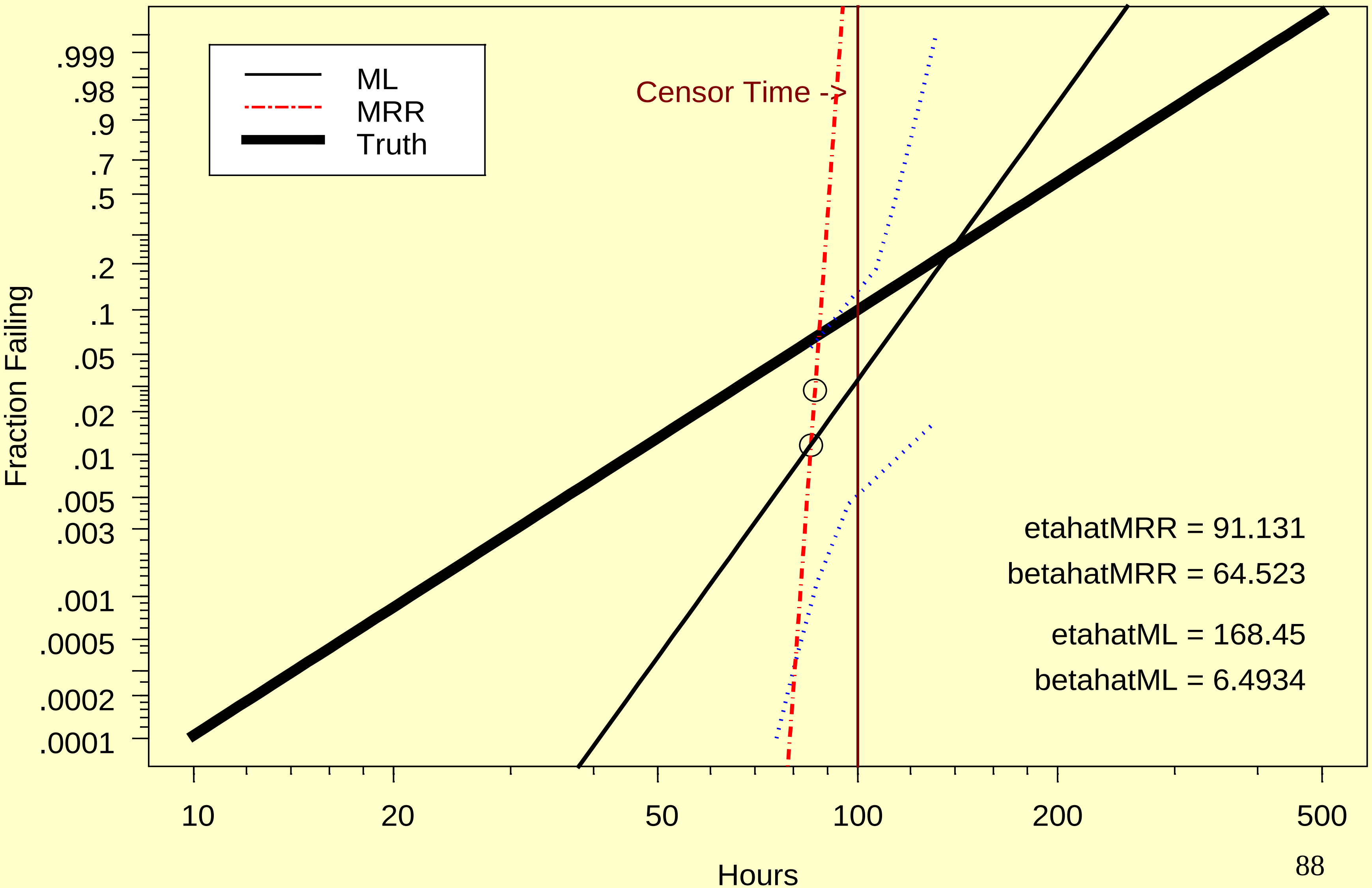


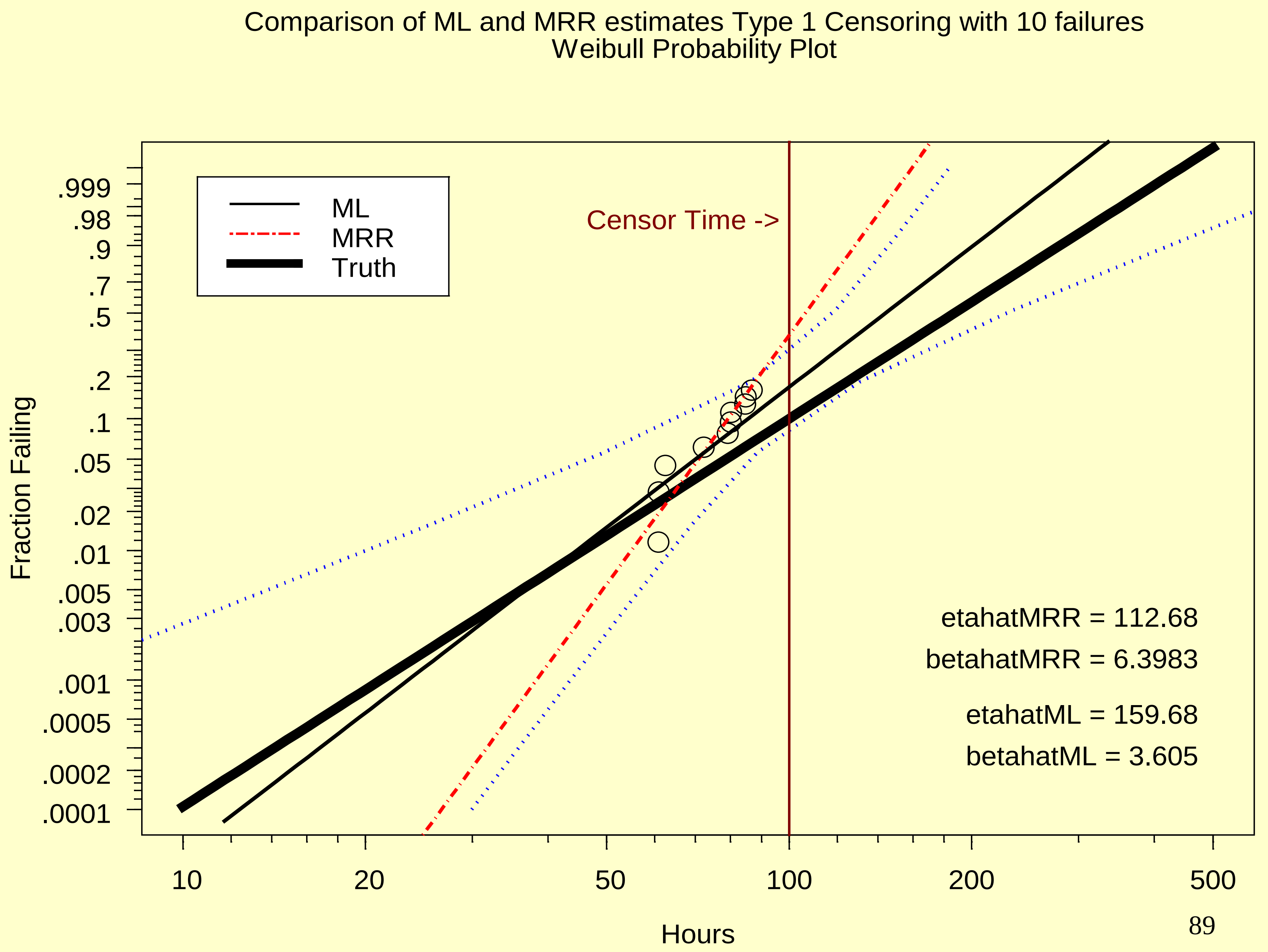
Simulation to Compare ML and MRR

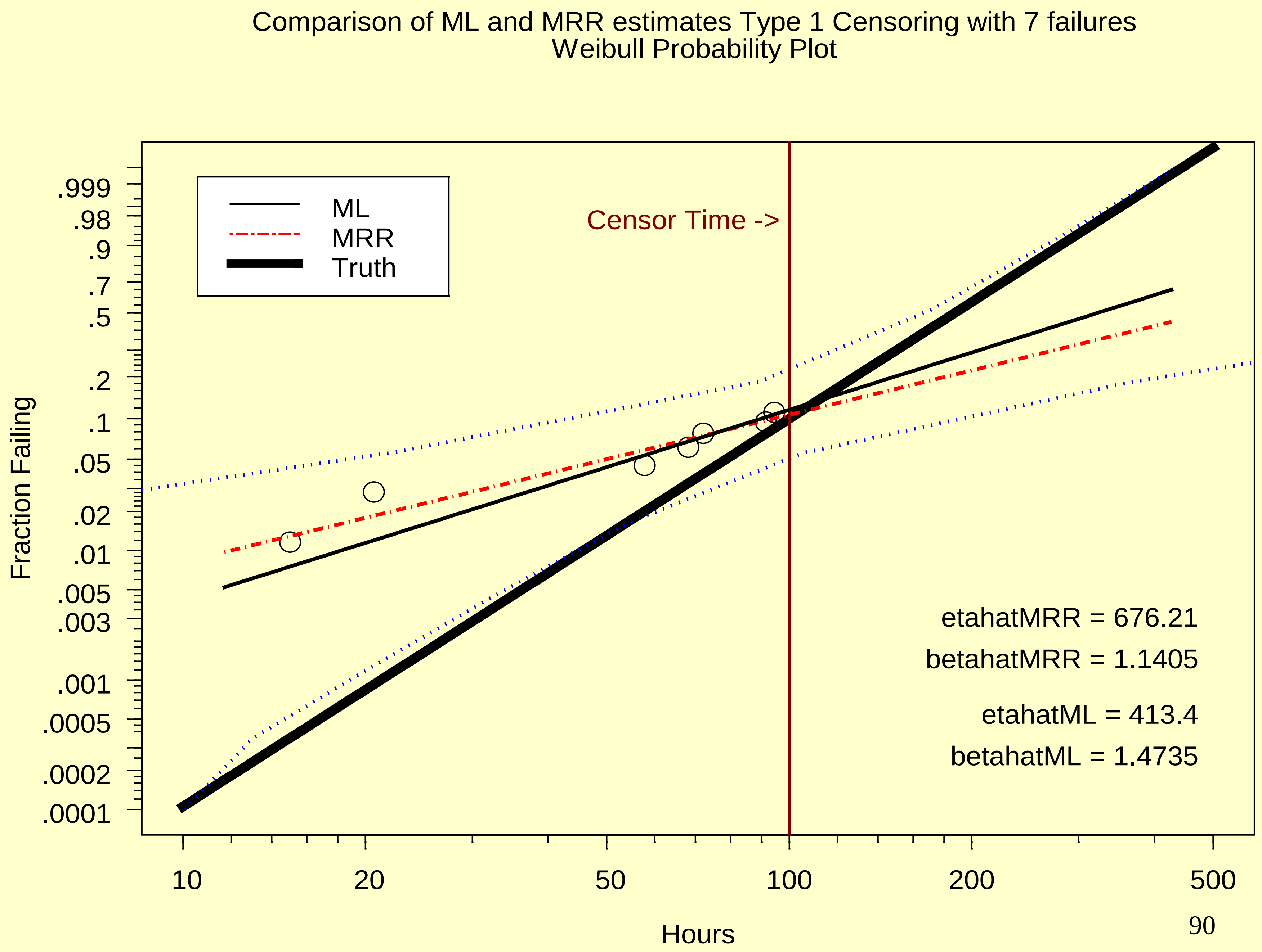
- Subset of results from Genschel and Meeker (2010) in *Quality Engineering*
- Simulation mimicking staggered entry and analysis at a fixed point in time (Type 1 censoring is a special case)
- Evaluated relative efficiency for various quantile estimators


Comparison of ML and MRR estimates Type 1 Censoring with 6 failures
Weibull Probability Plot
ML
MRR
Truth
Censor Time ->
Fraction Failing
.999
.98
.9
.7
.5
.2
.1
.05
.02
.01
.005
.003
.001
.0005
.0002
.0001
etahatMRR = 396.91
betahatMRR = 1.8774
etahatML = 186.74
betahatML = 3.6302
10
20
50
100
200
500
Hours

Comparison of ML and MRR estimates Type 1 Censoring with 2 failures
Weibull Probability Plot

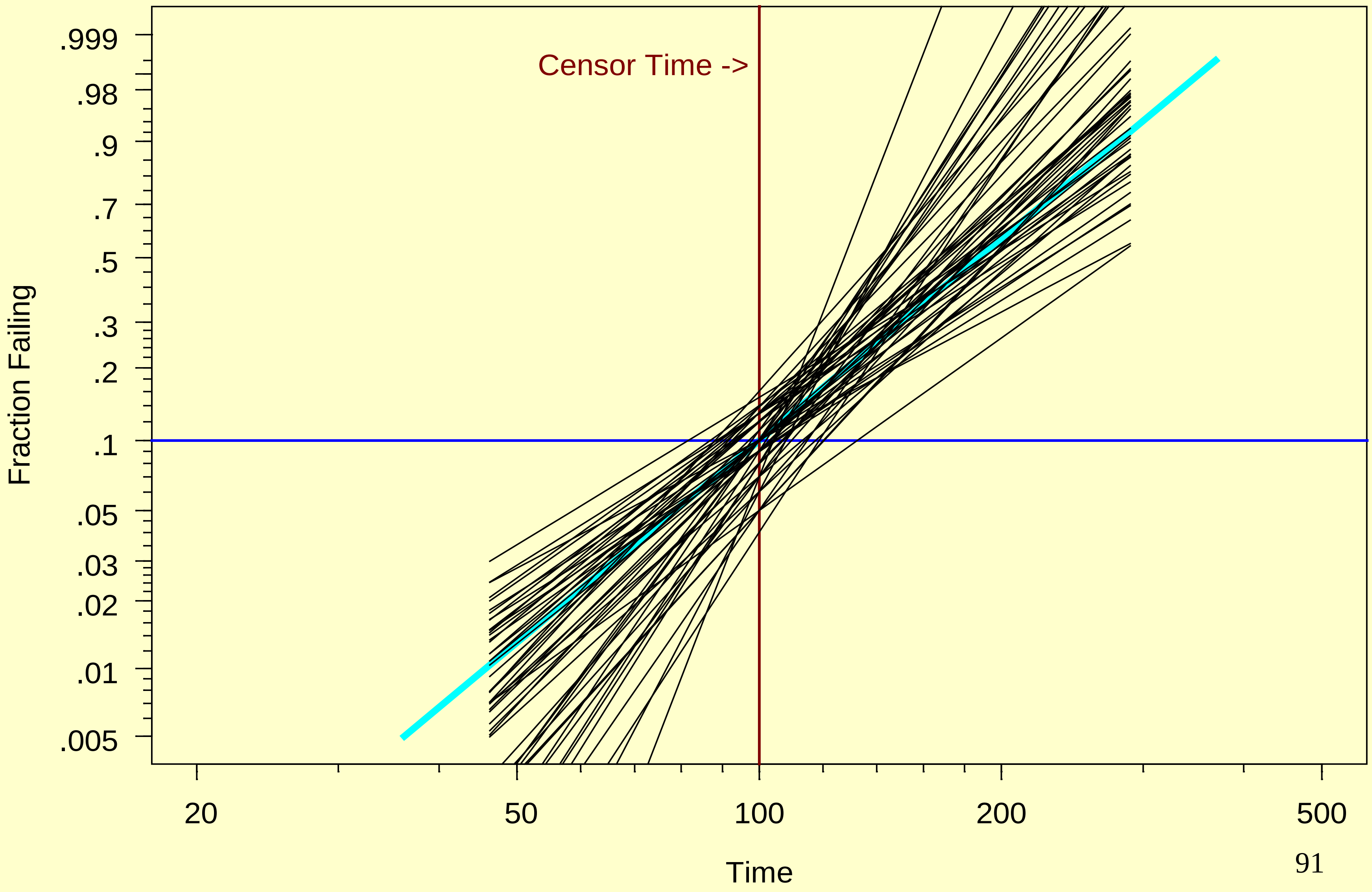


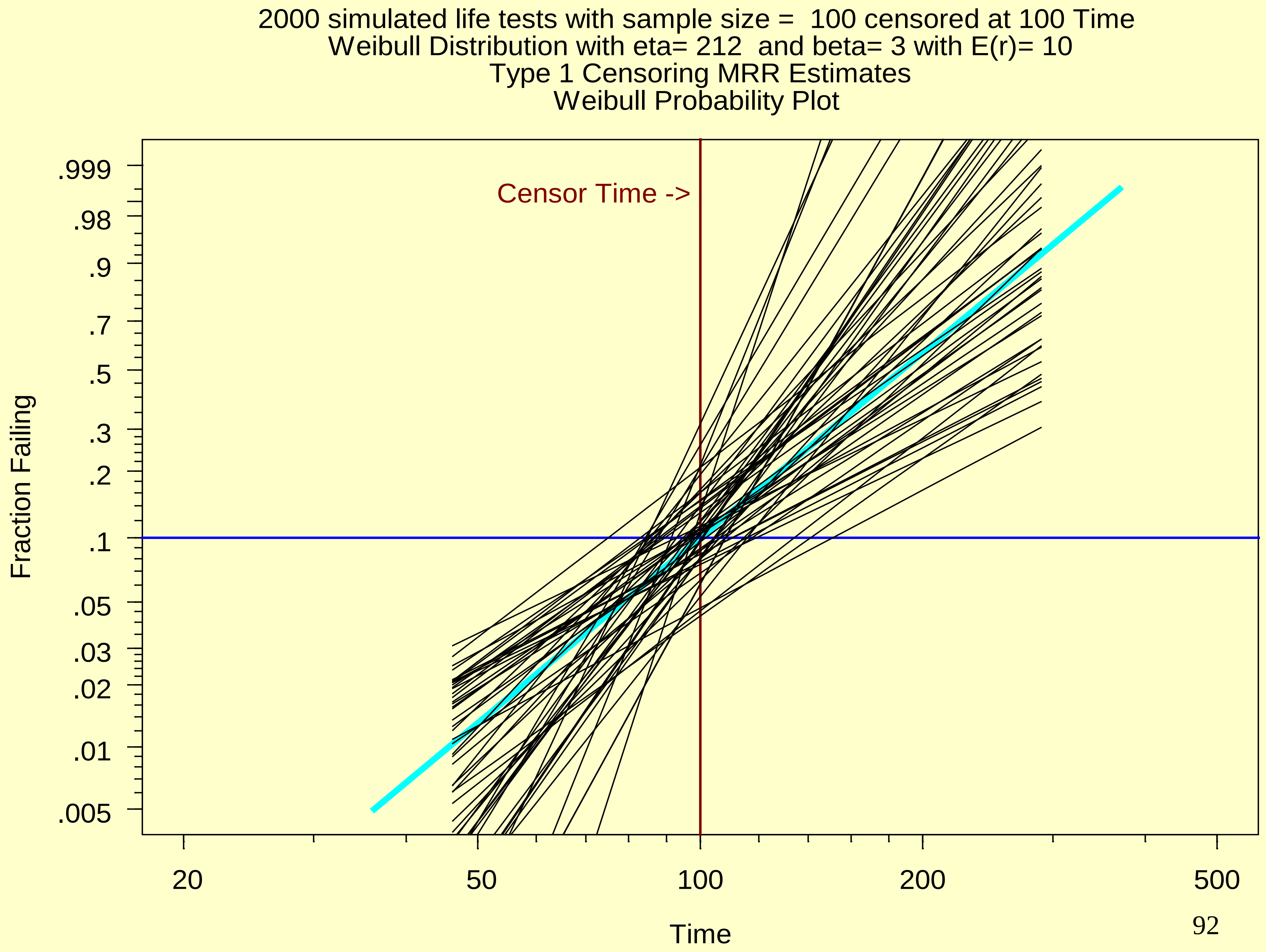

Comparison of ML and MRR estimates Type 1 Censoring with 10 failures
Weibull Probability Plot
ML
MRR
Truth
Censor Time ->
etahatMRR = 112.68
betahatMRR = 6.3983
etahatML = 159.68
betahatML = 3.605
Fraction Failing
.999
.98
.9
.7
.5
.2
.1
.05
.02
.01
.005
.003
.001
.0005
.0002
.0001
10
20
50
100
200
500
Hours



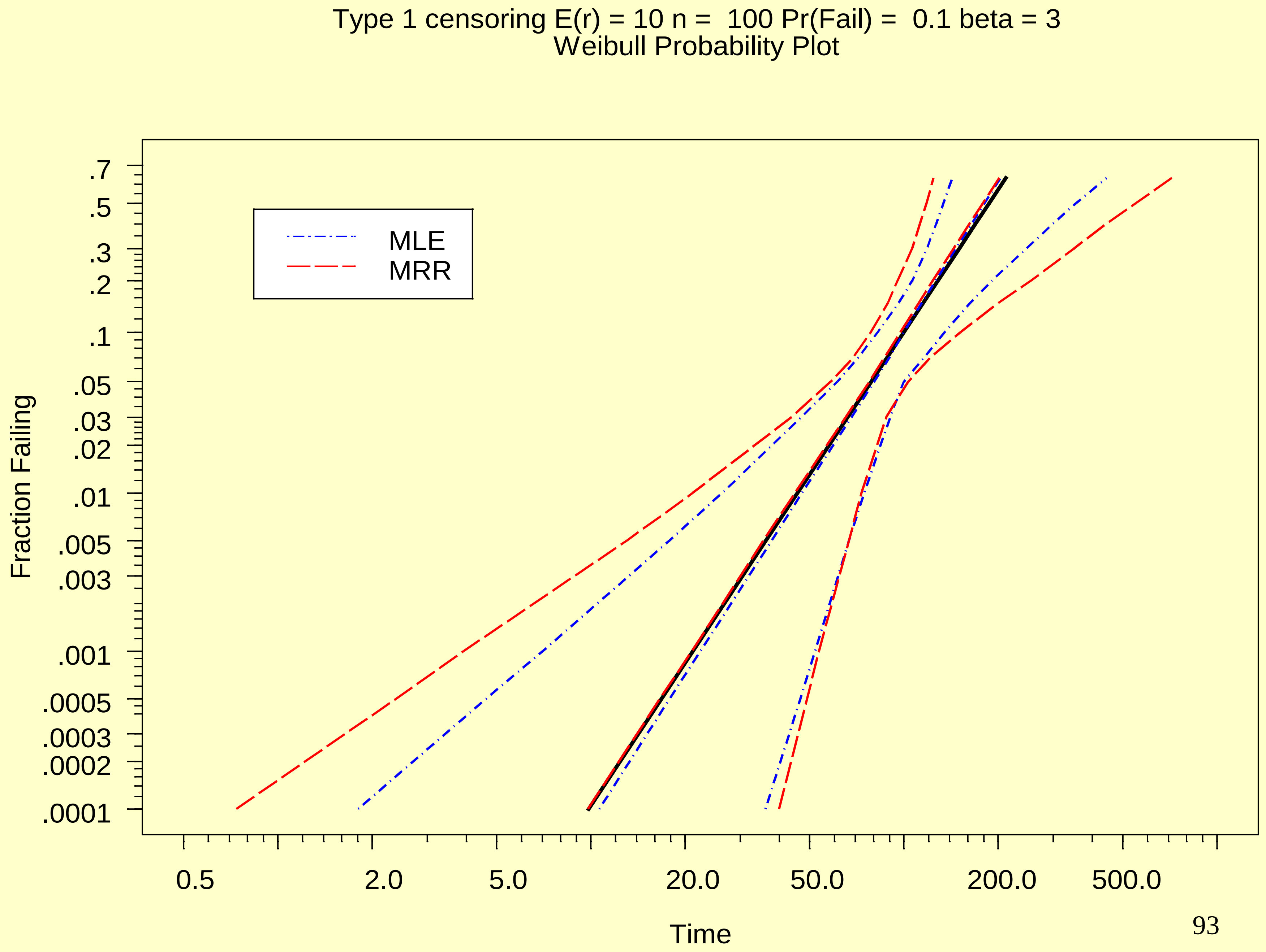
Comparison of ML and MRR estimates Type 1 Censoring with 7 failures
Weibull Probability Plot
ML
MRR
Truth
Censor Time ->
etahatMRR = 676.21
betahatMRR = 1.1405
etahatML = 413.4
betahatML = 1.4735
Fraction Failing
.999
.98
.9
.7
.5
.2
.1
.05
.02
.01
.005
.003
.001
.0005
.0002
.0001
10
20
50
100
200
500
Hours

2000 simulated life tests with sample size = 100 censored at 100 Time
Weibull Distribution with eta= 212 and beta= 3 with E(r)= 10
Type 1 Censoring ML Estimates
Weibull Probability Plot





2000 simulated life tests with sample size = 100 censored at 100 Time
Weibull Distribution with eta= 212 and beta= 3 with E(r)= 10
Type 1 Censoring MRR Estimates
Weibull Probability Plot
Censor Time ->
Fraction Failing
.999
.98
.9
.7
.5
.3
.2
.1
.05
.03
.02
.01
.005
20
50
100
200
500
Time


Type 1 censoring E(r) = 10 n = 100 Pr(Fail) = 0.1 beta = 3
Weibull Probability Plot
MLE
MRR
Fraction Failing
.7
.5
.3
.2
.1
.05
.03
.02
.01
.005
.003
.001
.0005
.0003
.0002
.0001
0.5
2.0
5.0
20.0
50.0
200.0
500.0
Time

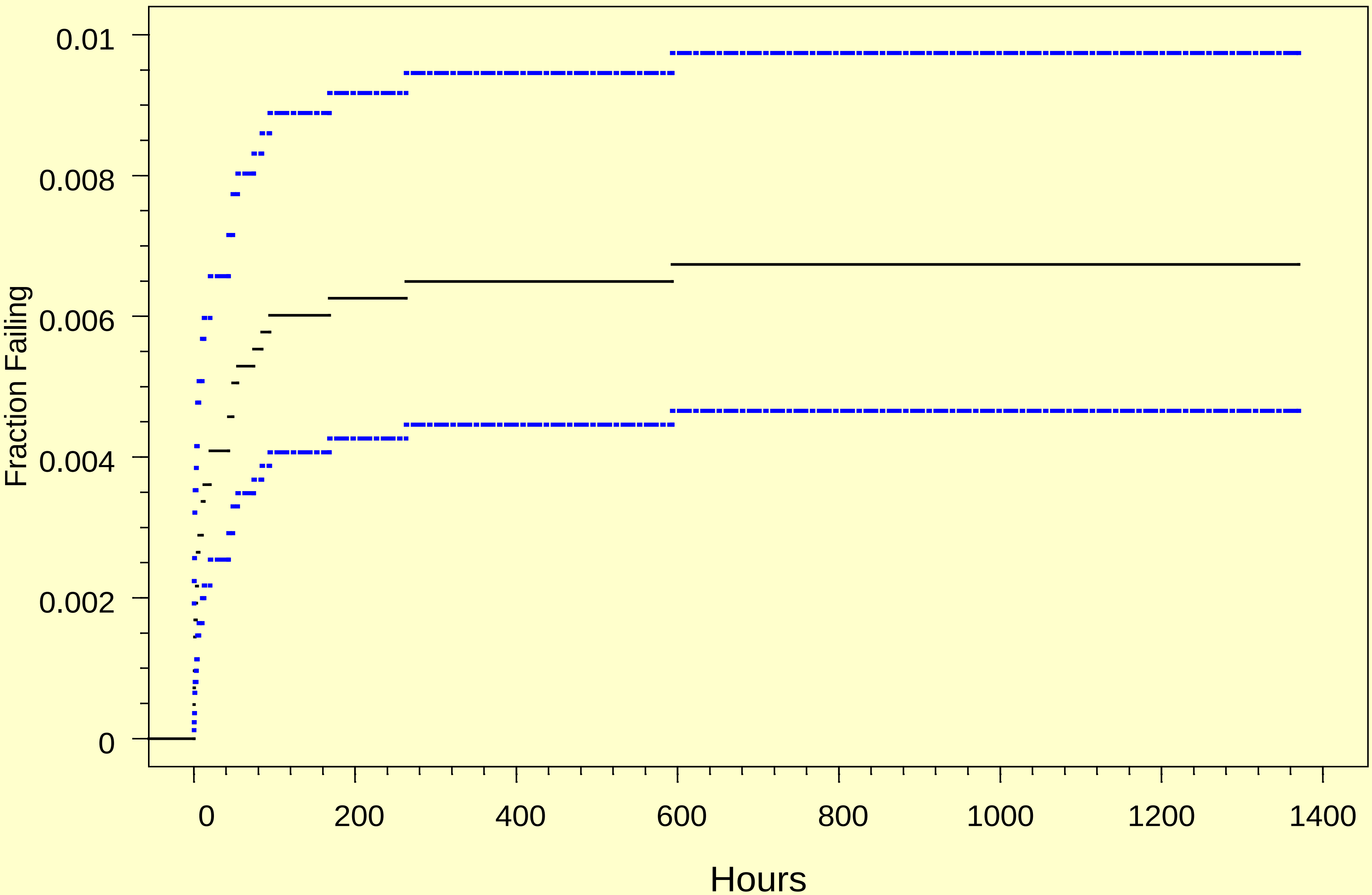
Lessons Learned

- In the analysis of censored data, maximum likelihood is more trustworthy than ordinary least squares
- Understand your failure modes
- ML and MR estimators are, in general, biased, but bias is insignificant relative to spread in the data
- ML estimators are more efficient than MRR estimators

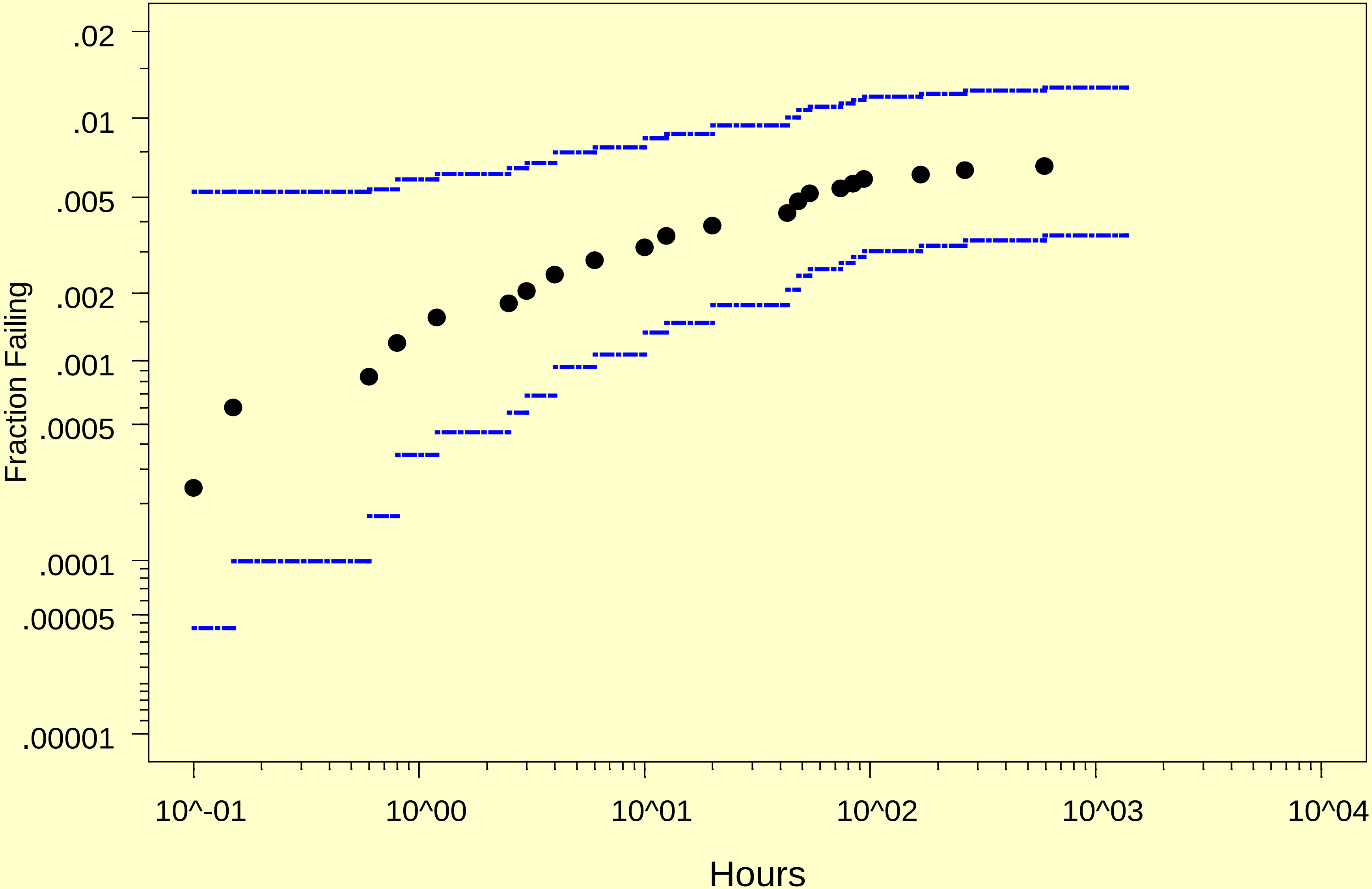
IC Device Accelerated Life Test

- Data from Meeker (1987)
- 4156 units tested for 1370 hours
- 28 failures; last failure at 593 hours
- Needed information: fraction failing at 20,000 hours

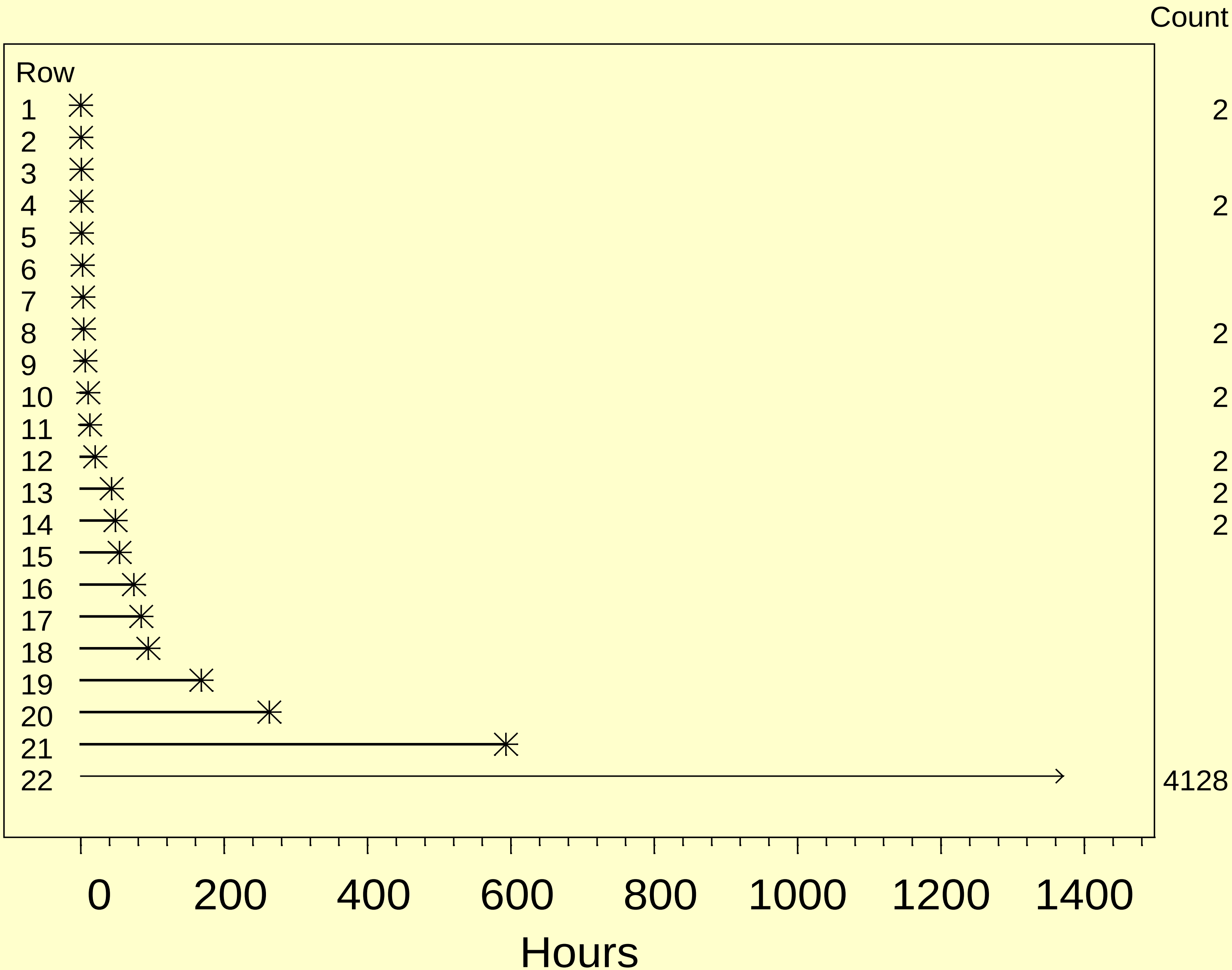
Integrated Circuit Failure Data After 1370 Hours
with Nonparametric Pointwise 95% Confidence Bands

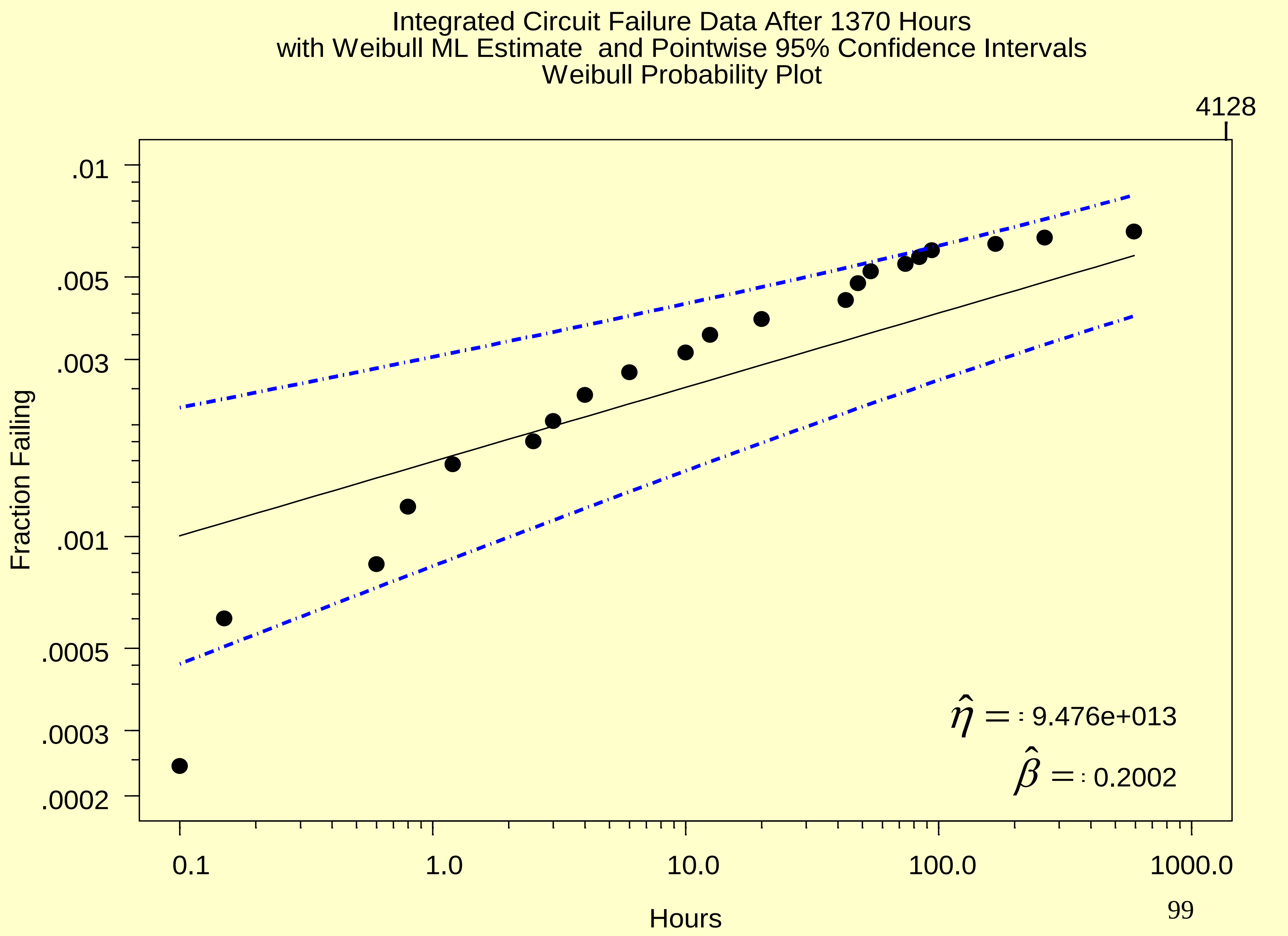


Integrated Circuit Failure Data After 1370 Hours
with Nonparametric Simultaneous 95% Confidence Bands
Lognormal Probability Plot

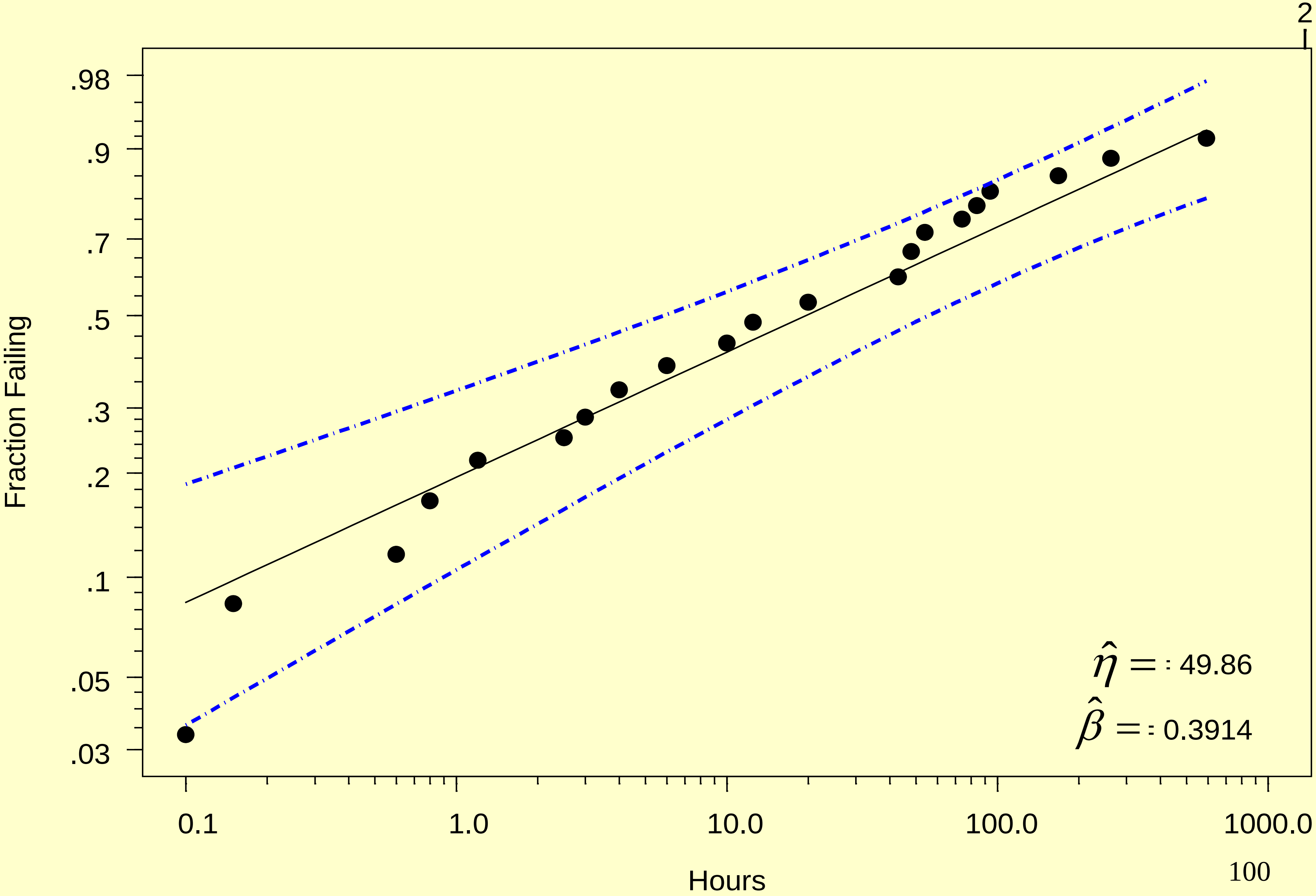


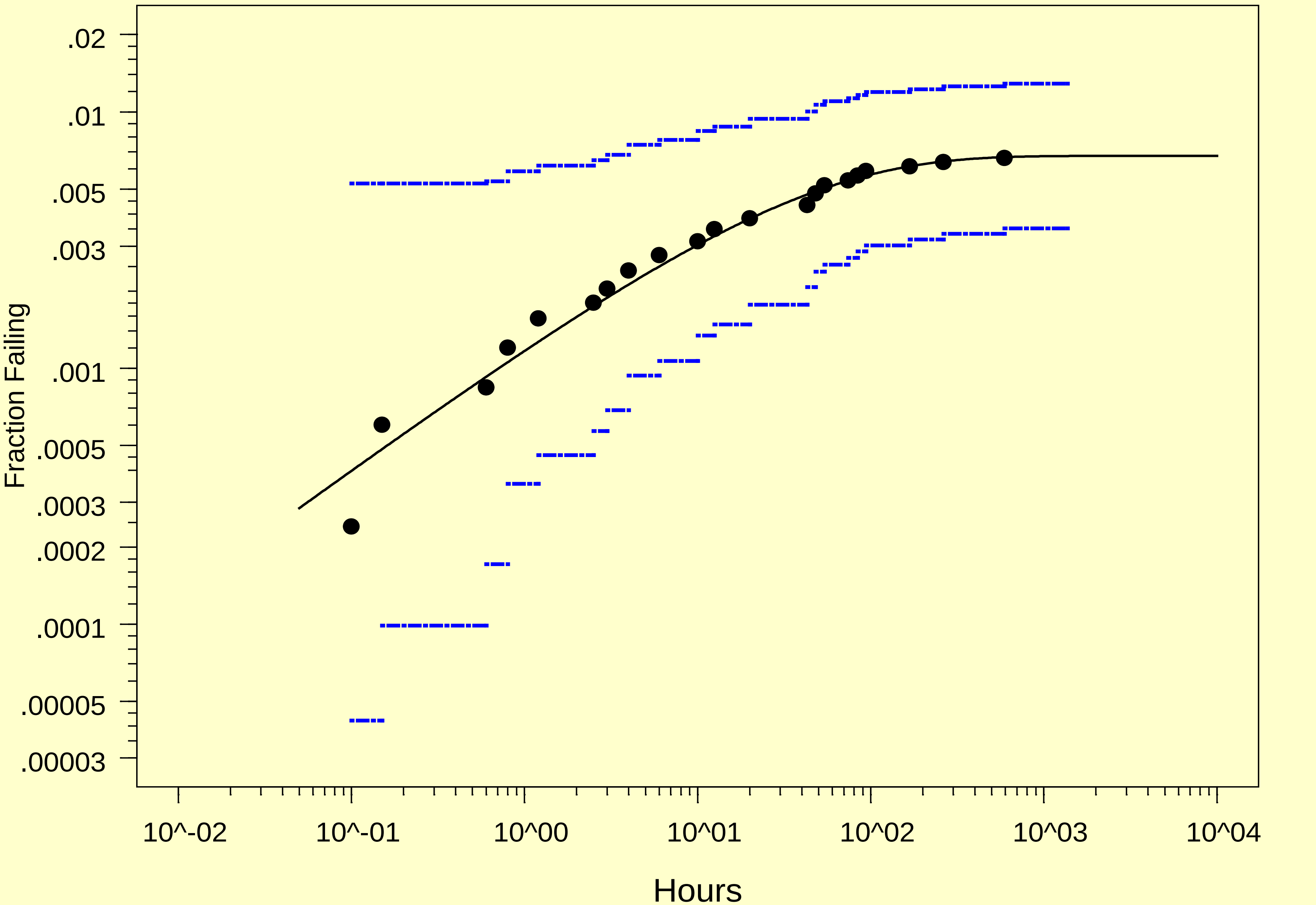
Integrated Circuit Failure Data After 1370 Hours




Integrated Circuit Failure Data After 1370 Hours
with Weibull ML Estimate and Pointwise 95% Confidence Intervals
Weibull Probability Plot
4128
Fraction Failing
.01
.005
.003
.001
.0005
.0003
.0002
Hours
0.1
1.0
10.0
100.0
1000.0
$\hat{\eta} =$ 9.476e+013
$\hat{\beta} =$ 0.2002

lfp1370Fix2 data
with Weibull ML Estimate and Pointwise 95% Confidence Intervals
Weibull Probability Plot





.02
.01
.005
.003
.001
.0005
.0003
.0002
.0001
.00005
.00003
Fraction Failing
10^-02
10^-01
10^00
10^01
10^02
10^03
10^04
Hours

Lessons Learned

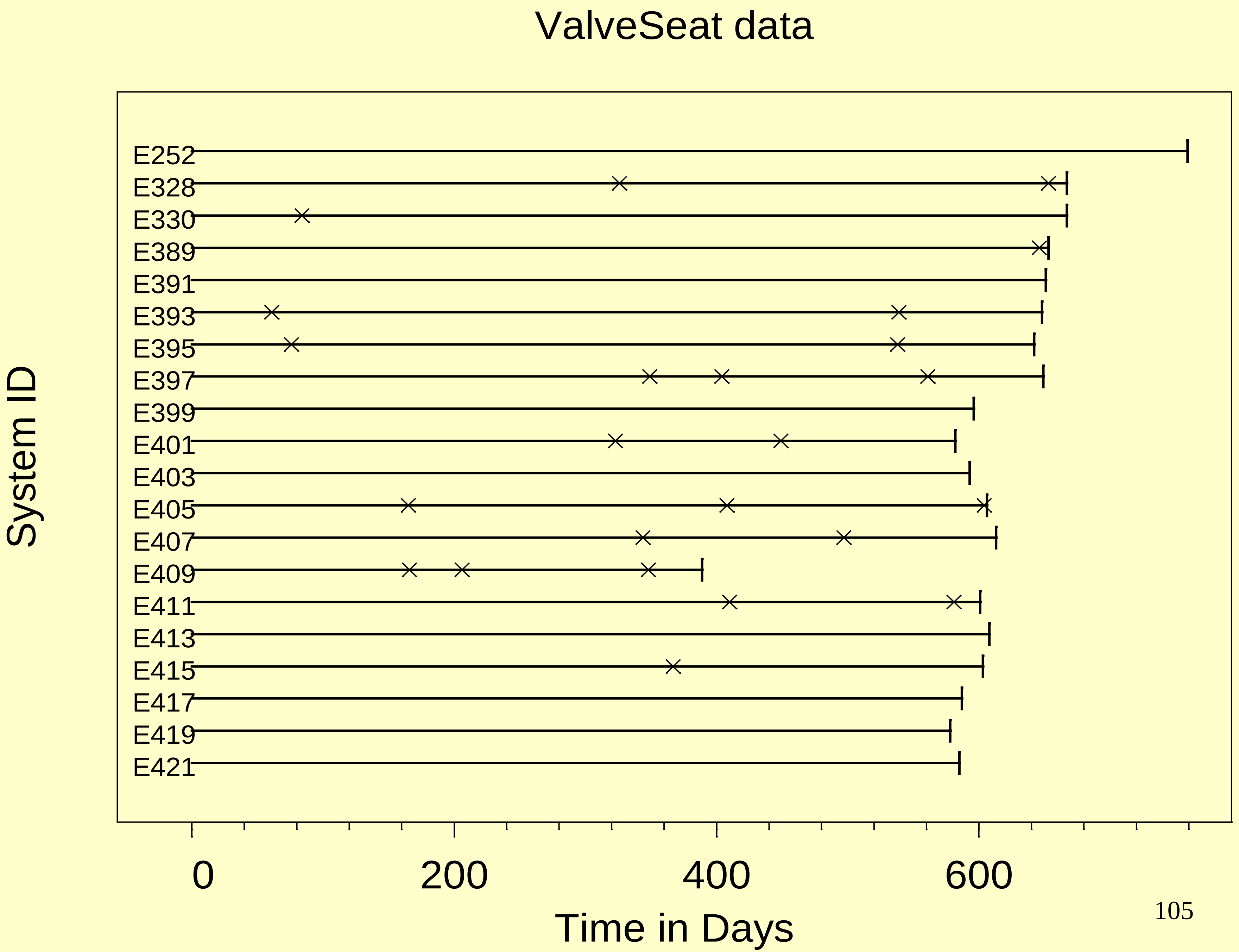
- Leveling of a probability plot usually implies a “limited failure population” (a.k.a “defective subpopulation model.”
- If you do not wait long enough
  - You cannot decide if you have many units failing slowly or a small fraction failing early.
  - Conclusions will depend strongly on the assumed distribution model

Analysis of Recurrence Data

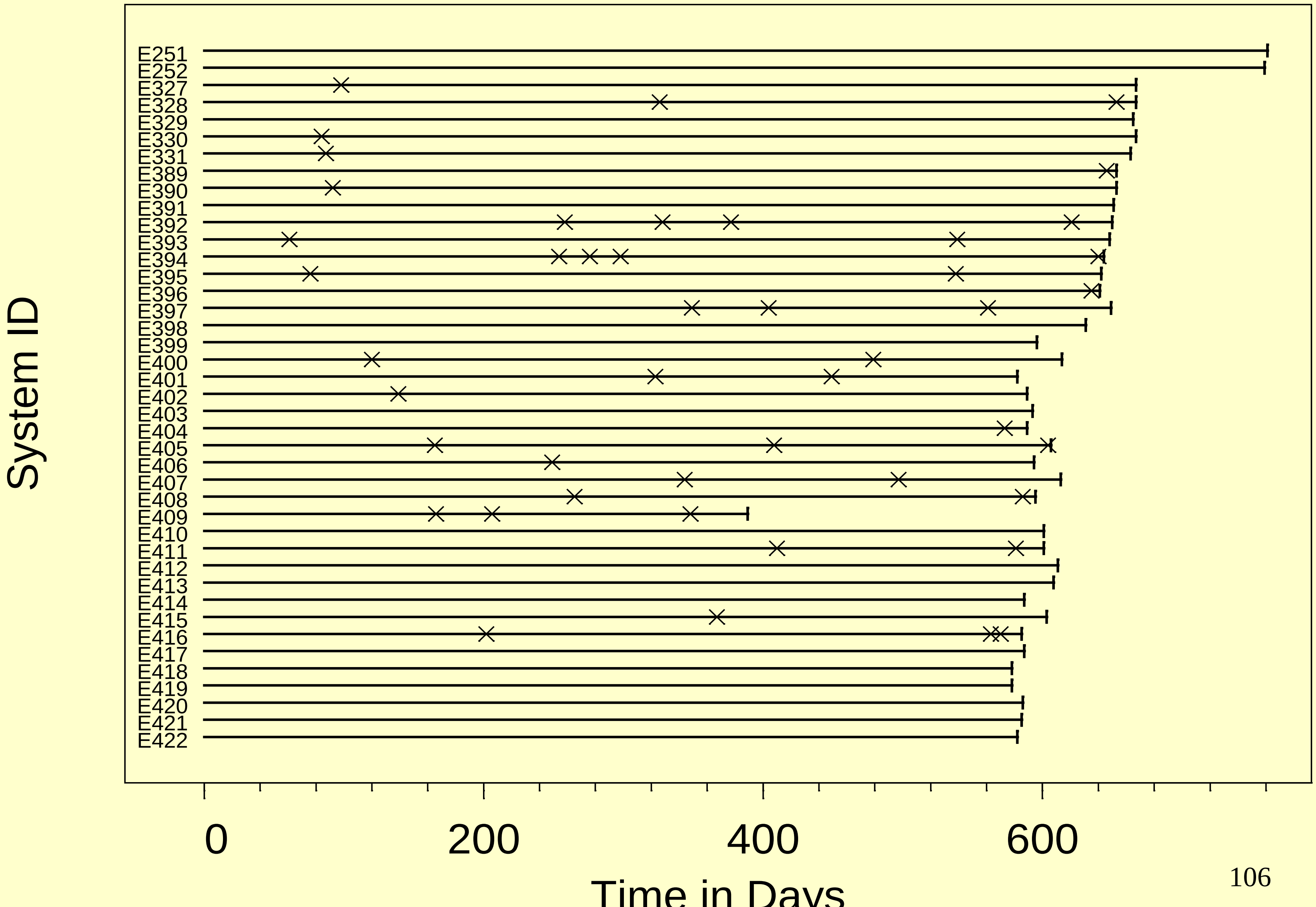
- Recurrence data arise in many applications involving systems that are repaired over time (fleets of automobiles, aircraft, machinery, etc.)
- Events of interest could include failures, maintenance actions, replacements, etc.
- Interest centers on such metrics as
  - Mean cumulative cost
  - Rate of cost accumulation
- See Nelson (2003) and Cook and Lawless (2007)

Valve Seat Replacement Times (Nelson and Doganaksoy 1989)

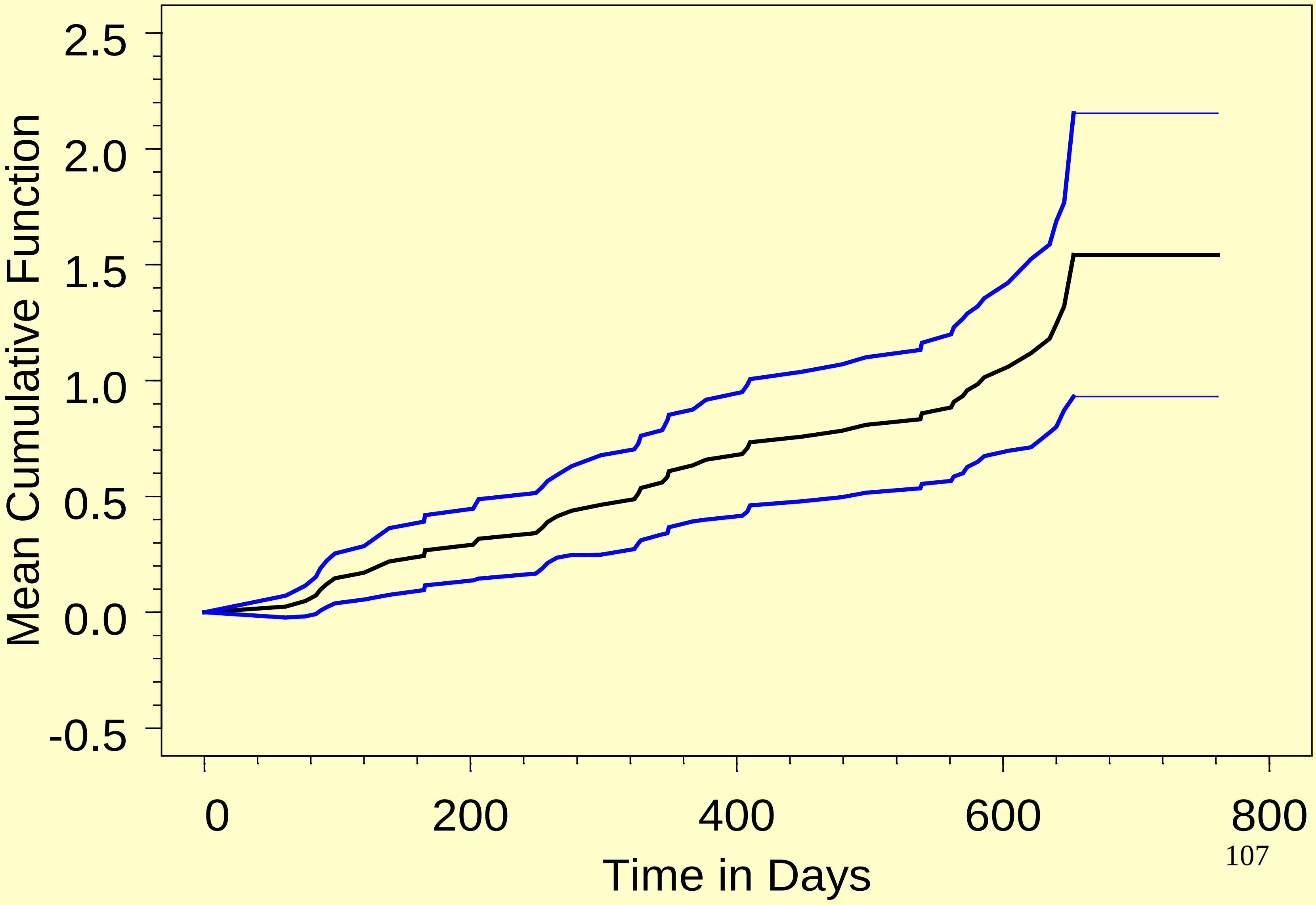
- Data collected from valve seats from a fleet of 41 diesel engines operated in and around Beijing, China (days of operation).
- Each engine has 16 valves.
- Most failures caused by operating in a dusty environment.
- Does the replacement rate increase with age?
- How many replacement valves will be needed in the future?
- Can valve life in these systems be modeled as a renewal process?


ValveSeat data
System ID
E252
E328
E330
E389
E391
E393
E395
E397
E399
E401
E403
E405
E407
E409
E411
E413
E415
E417
E419
E421
0
200
400
600
Time in Days

ValveSeat data

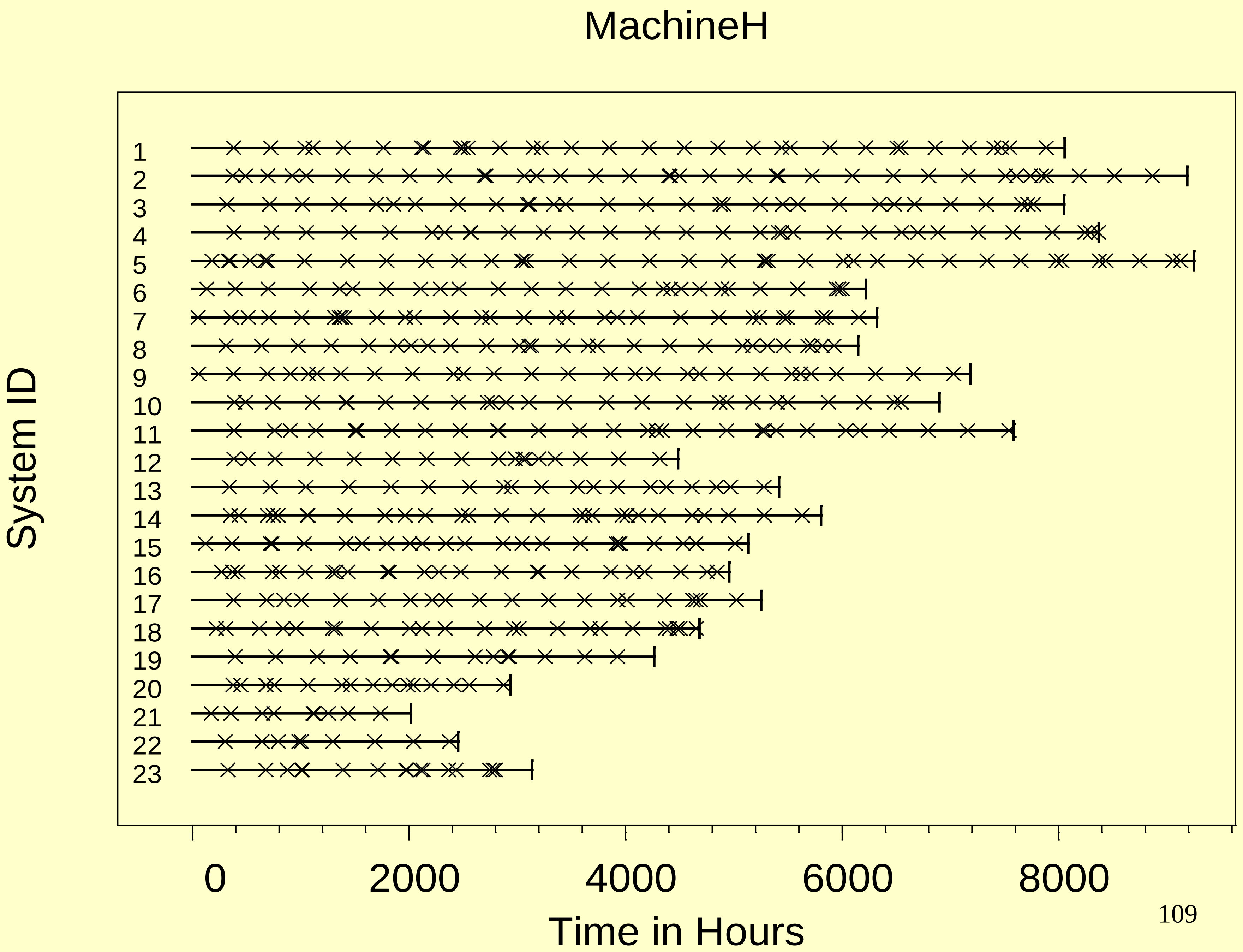


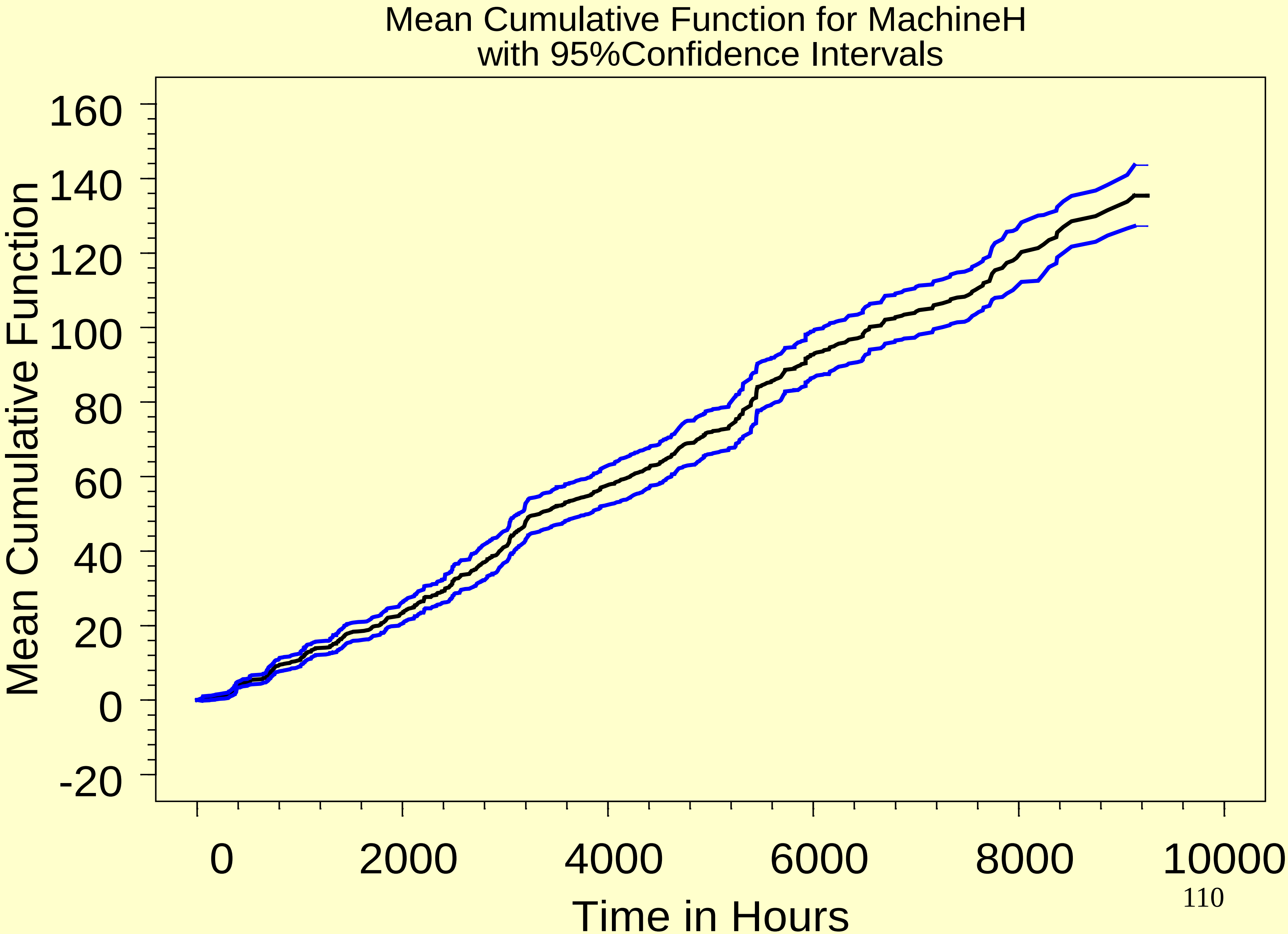
Mean Cumulative Function for ValveSeat data
with 95%Confidence Intervals

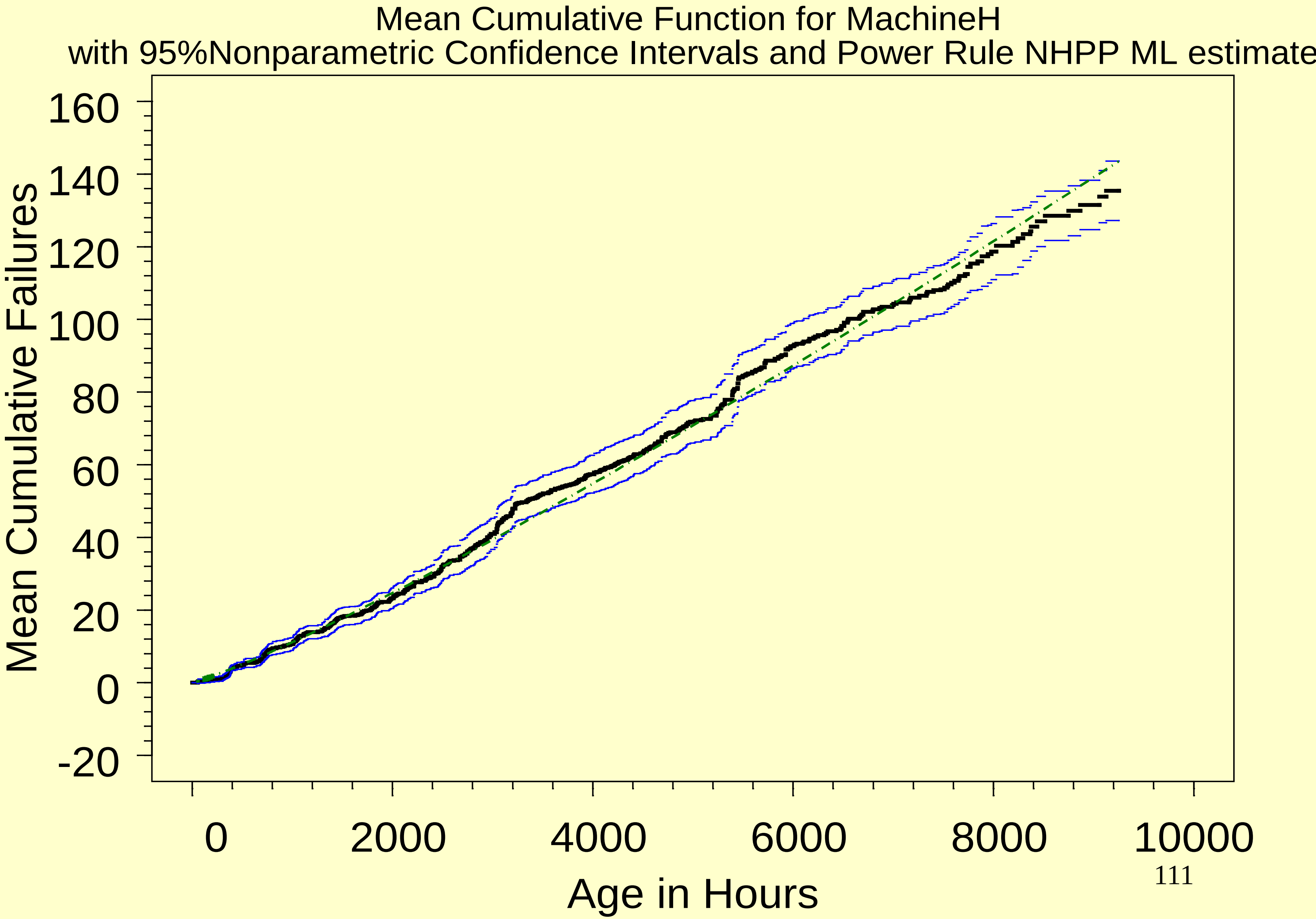


Earth Moving Machine Maintenance Actions (Meeker and Escobar 1998)

- Fleet of 23 earth-moving machines put into service over time
- Preventive maintenance ever 300-400 hours of operation
- Major overhaul every 2000-3000 hours of operation
- Many unscheduled maintenance actions


MachineH
System ID
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
0
2000
4000
6000
8000
Time in Hours


Mean Cumulative Function for MachineH
with 95%Confidence Intervals
Mean Cumulative Function
160
140
120
100
80
60
40
20
0
-20
0
2000
4000
6000
8000
10000
Time in Hours

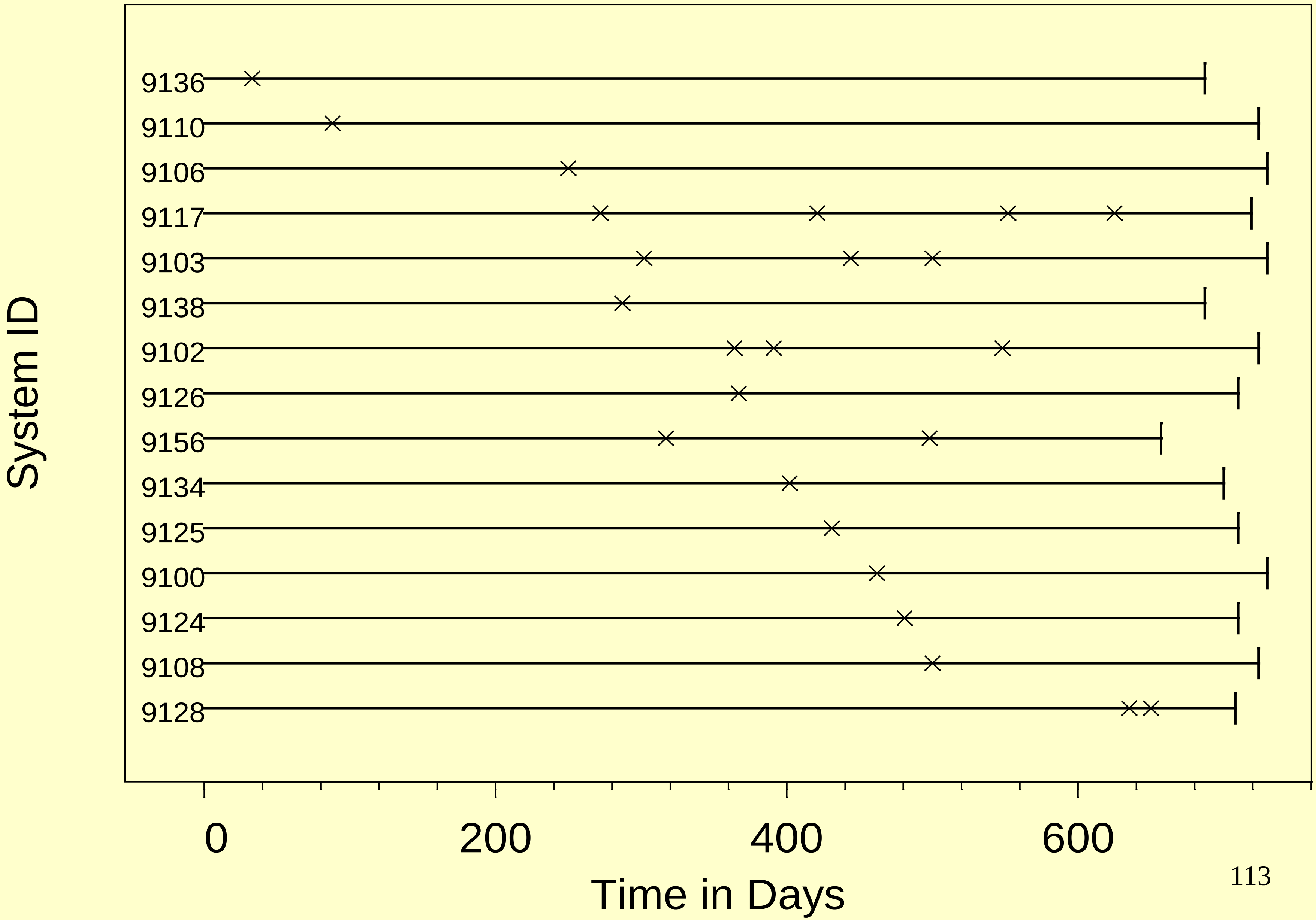


Mean Cumulative Function for MachineH
with 95%Nonparametric Confidence Intervals and Power Rule NHPP ML estimate
Mean Cumulative Failures
160
140
120
100
80
60
40
20
0
-20
0
2000
4000
6000
8000
10000
Age in Hours

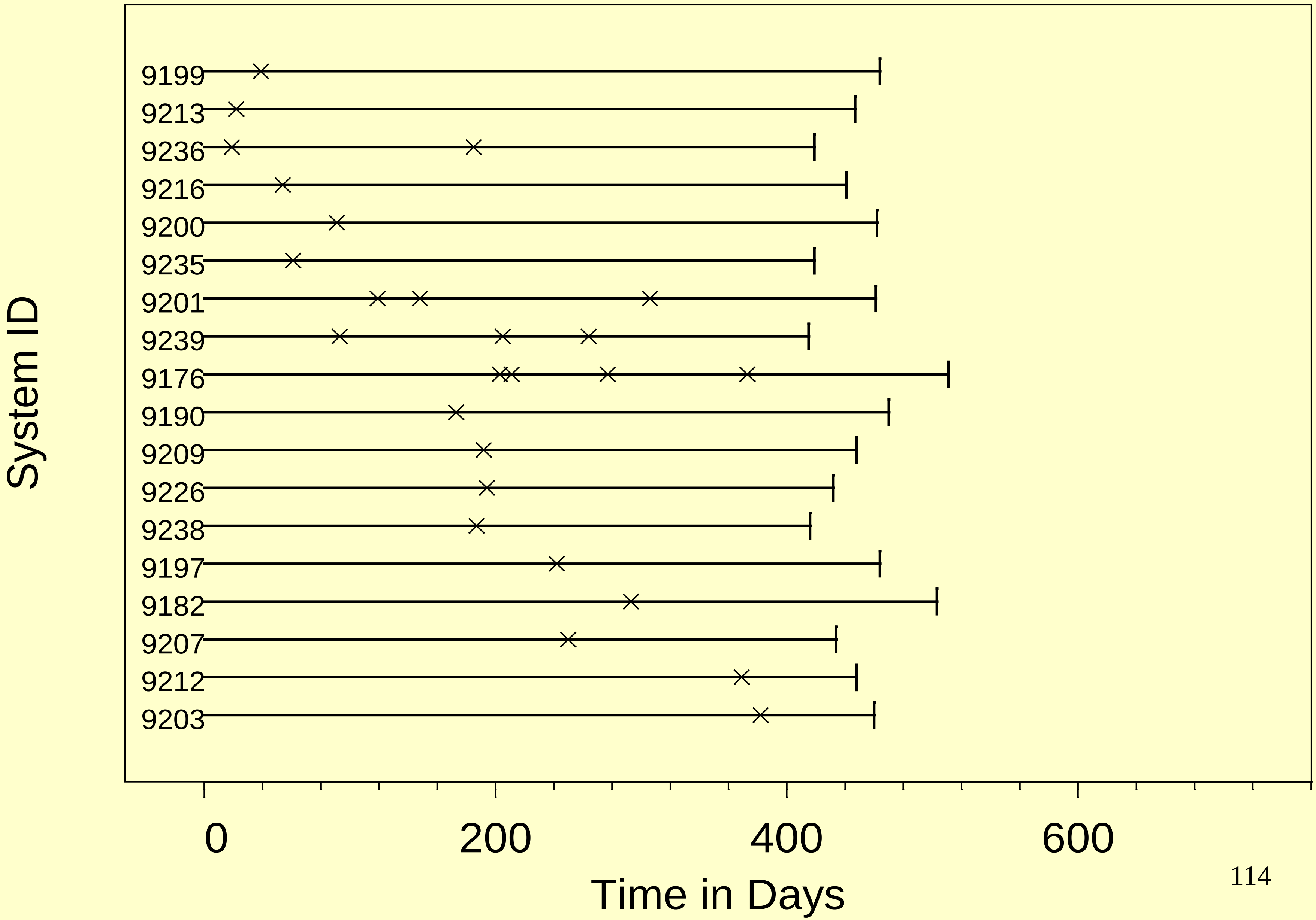
Locomotive Braking Grid Replacement Frequency Comparison (Doganaksoy and Nelson 1991)

- A particular type of locomotive has six braking grids.
- Data available on locomotive age when a braking grid is replaced and the age at the end of the observation period.
- There seemed to be a problem with Batch 2.
- A statistical comparison between two different production batches of braking grids is desired.

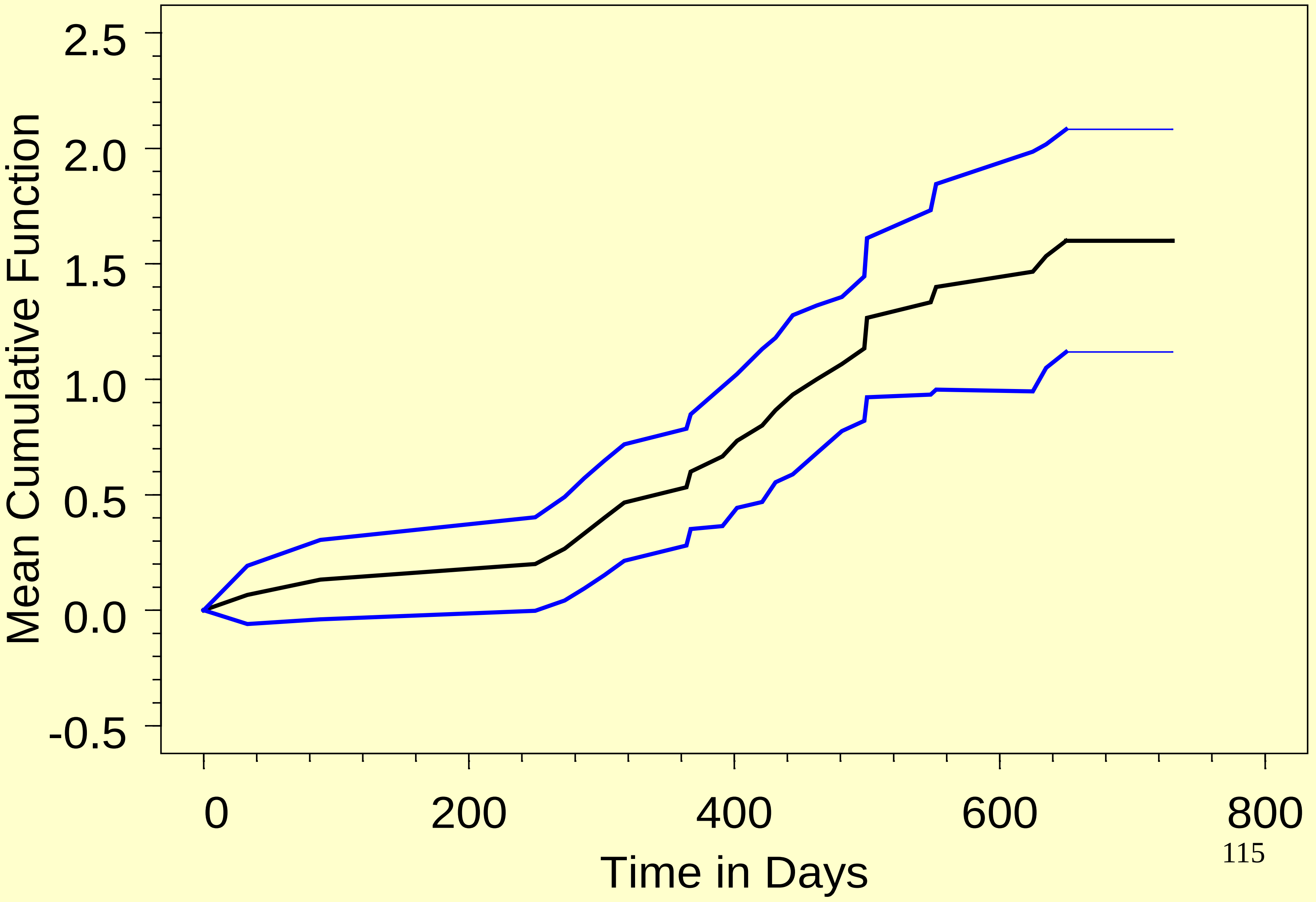
Grids1 Replacement Data

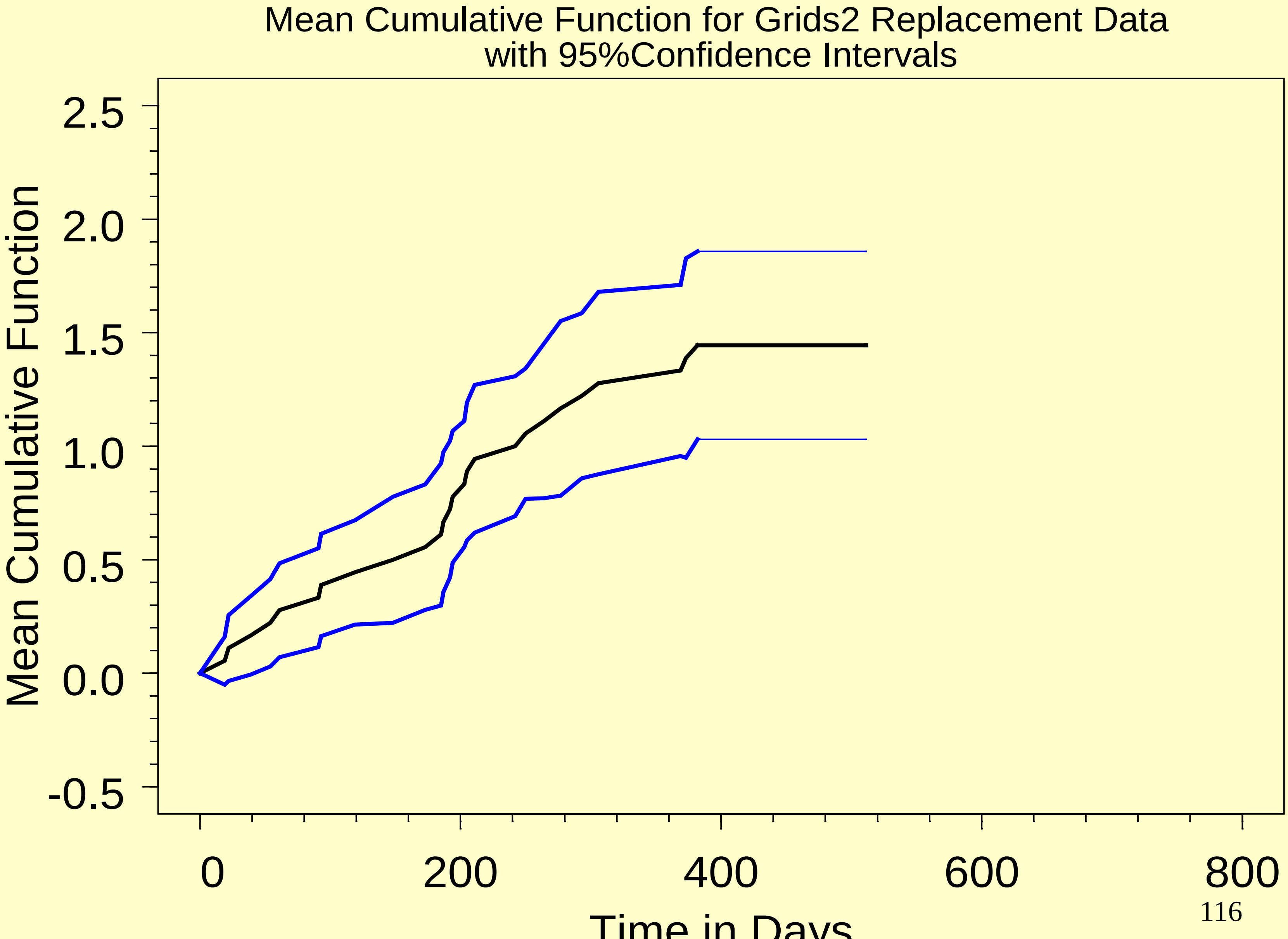


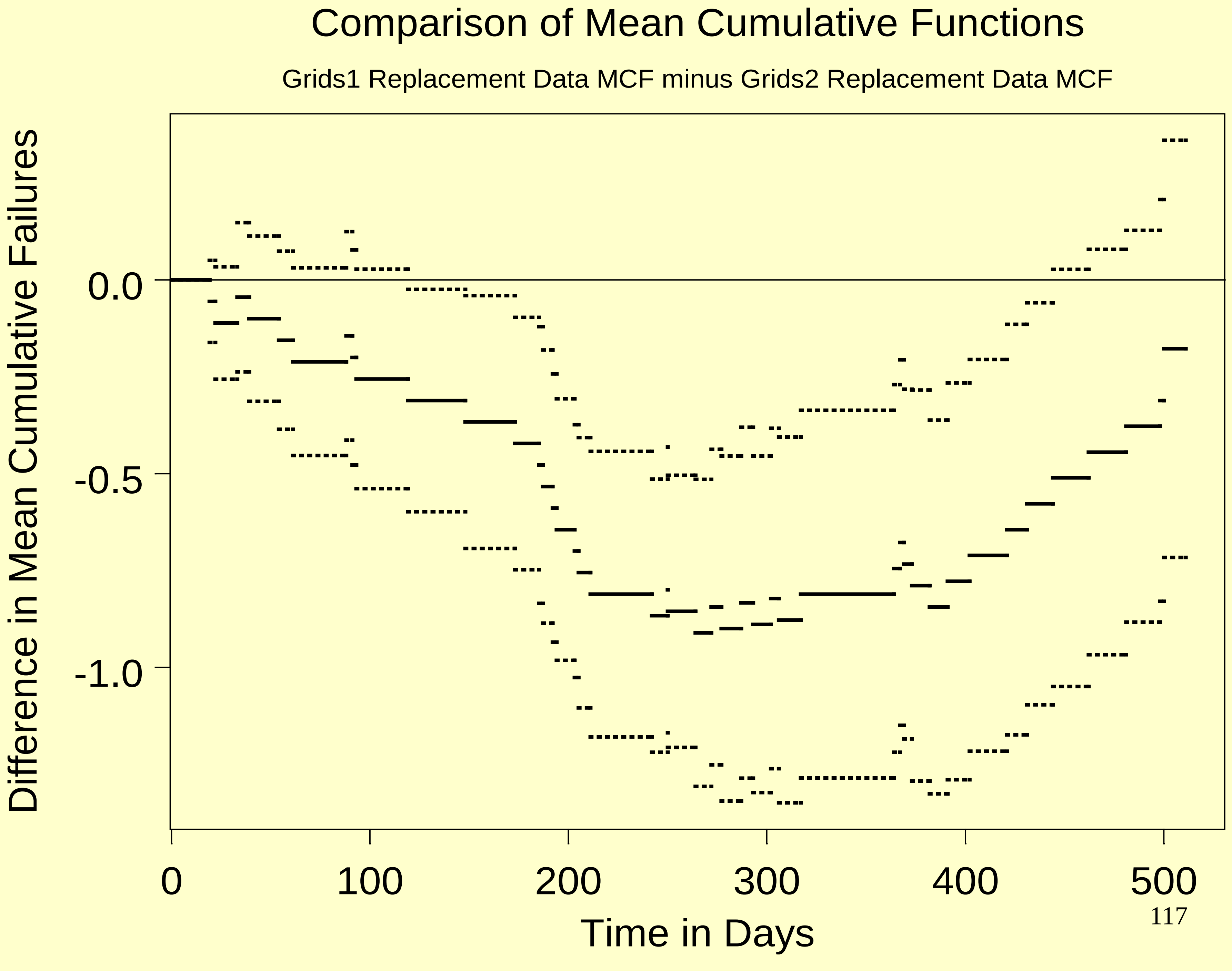
Grids2 Replacement Data



Mean Cumulative Function for Grids1 Replacement Data
with 95%Confidence Intervals




Mean Cumulative Function for Grids2 Replacement Data
with 95%Confidence Intervals
Mean Cumulative Function
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
0
200
400
600
800
Time in Days


Comparison of Mean Cumulative Functions
Grids1 Replacement Data MCF minus Grids2 Replacement Data MCF
Difference in Mean Cumulative Failures
0.0
-0.5
-1.0
0
100
200
300
400
500
Time in Days

Lessons Learned

- Simple nonparametric statistical methods provide powerful tool for drawing conclusions from complicated data with minimum assumptions
- Parametric methods and methods for censored recurrence data are also available (but more complicated to implement)
- It is often a mistake to try to revert to “failure-time analysis” when data are recurrence data.

Motivation for Accelerated Testing

Today manufactures need to develop newer, higher technology products in record time while improving productivity, reliability, and quality.

- Rapid product development.
- Changing technologies/new materials
- More complicated products with more components
- Higher customer expectations for better reliability

Levels of Accelerated Testing

- Materials
- Components
- Subsystem
- Full system

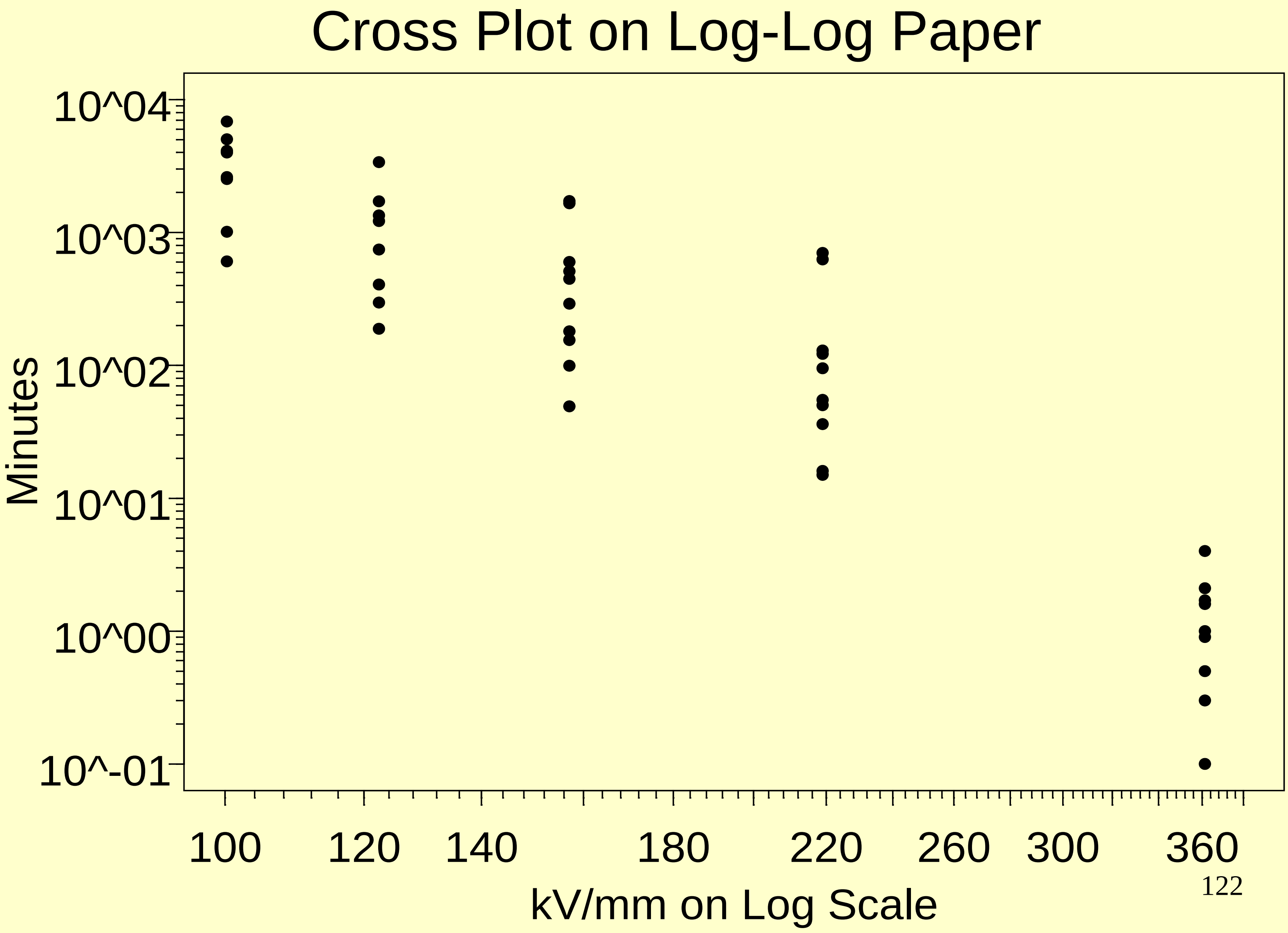
Usually testing at higher levels of integration will result in less acceleration

Most accelerated tests are run at lower levels of a product

Voltage-Accelerated Life Test of a Mylar-Polyurethane Laminated Direct Current High Voltage Insulating Structure

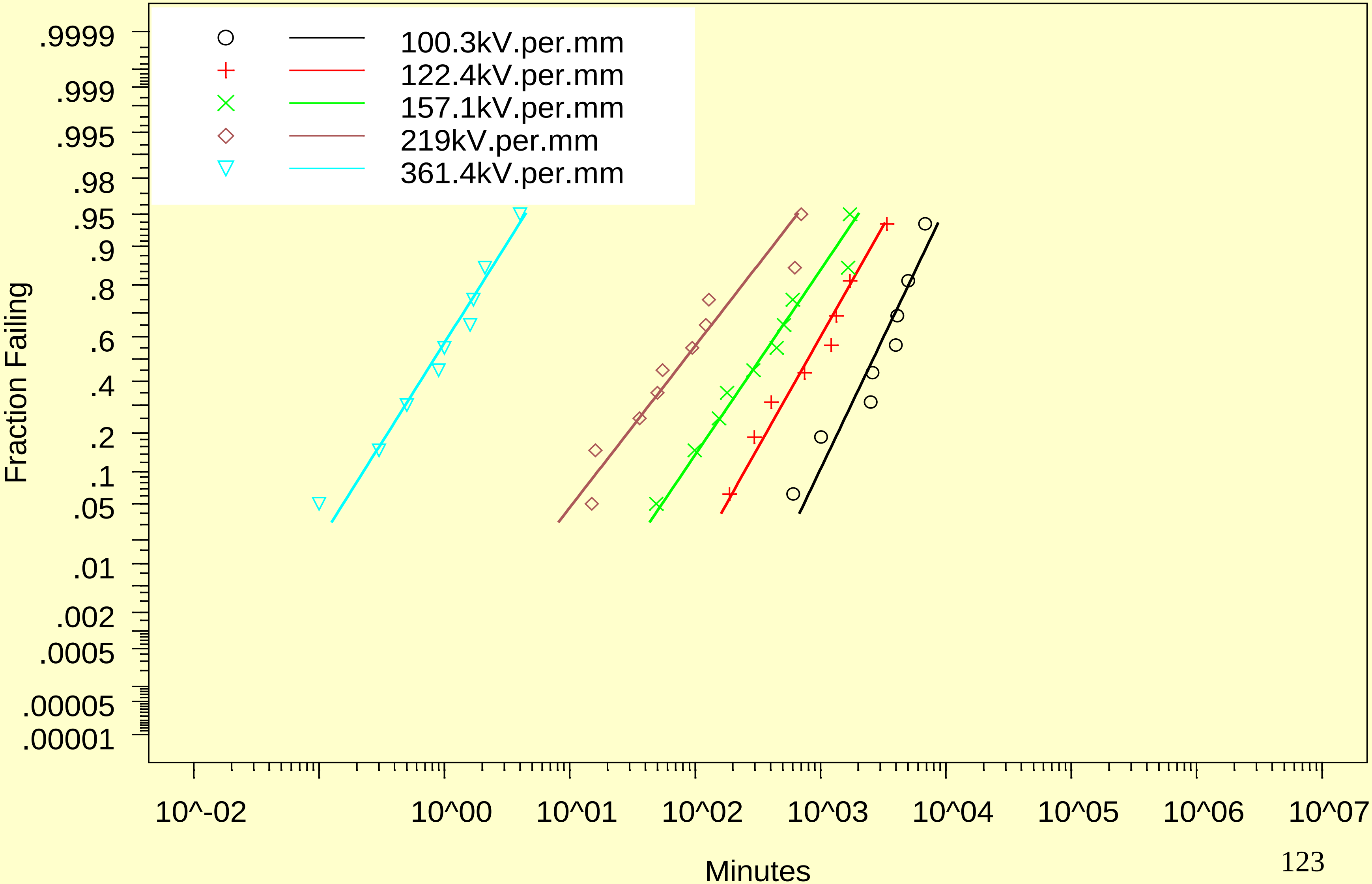
- Data from Kalkanis and Rosso (1989)
- Time to dielectric breakdown of units tested at 100.3, 122.4,157.1, 219.0, and 361.4 kV/mm.
- Needed to evaluate the reliability of the insulating structure and to estimate the life distribution at system design voltages (e.g. 50 kV/mm).

Mylar-Polyurethane Laminated Direct Current High Voltage Insulating Structure



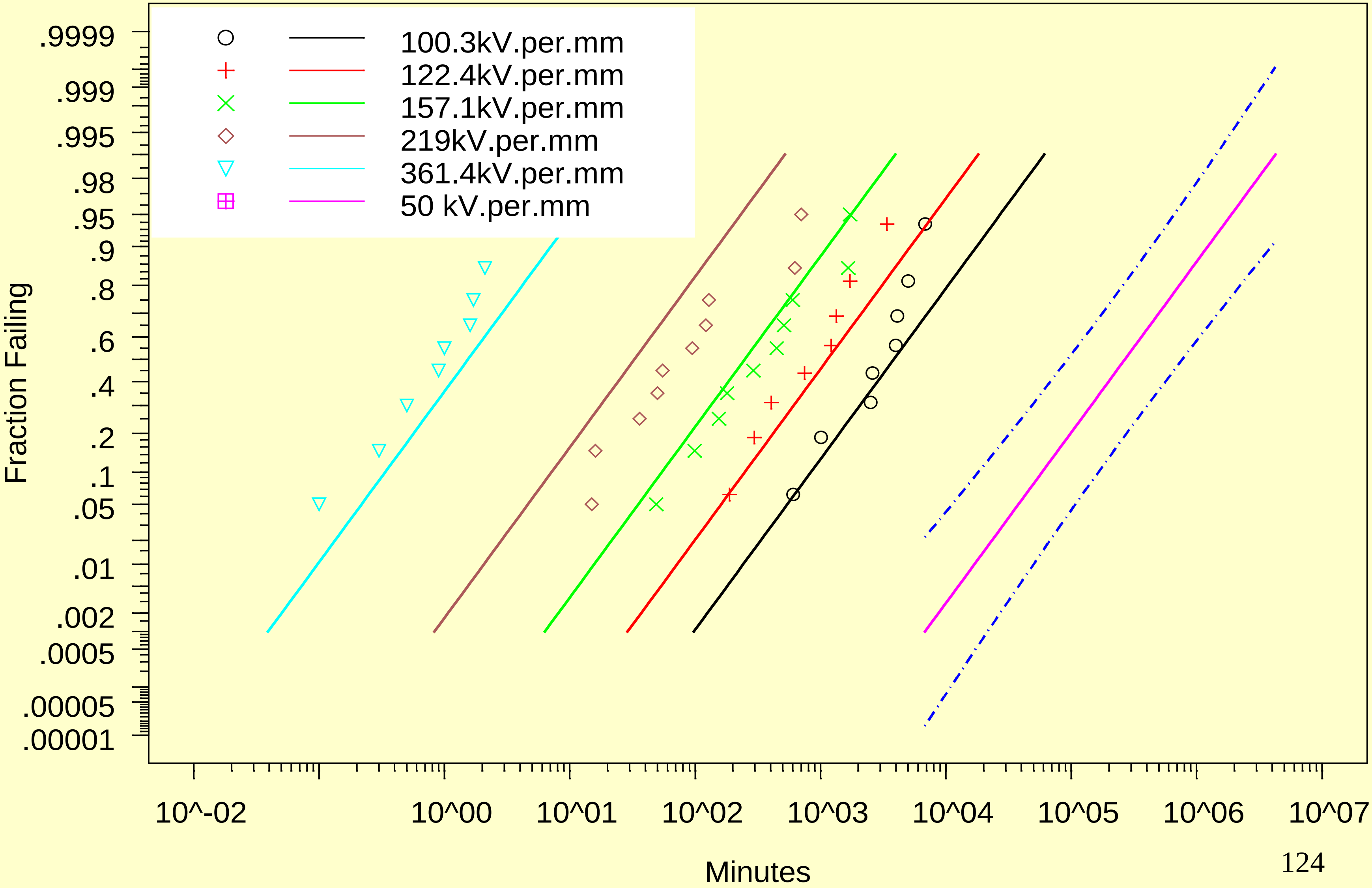
Mylar Polyurethane Insulating Structure

Data Analysis at Individual Conditions



Mylar Polyurethane Insulating Structure Data

Inverse Power Regression Model All Data



The Inverse Power Law/Lognormal Model for Insulation Lifetimes

- $\text{Life} = CV^{\beta_1}$
- The probability of failure as a function of time is

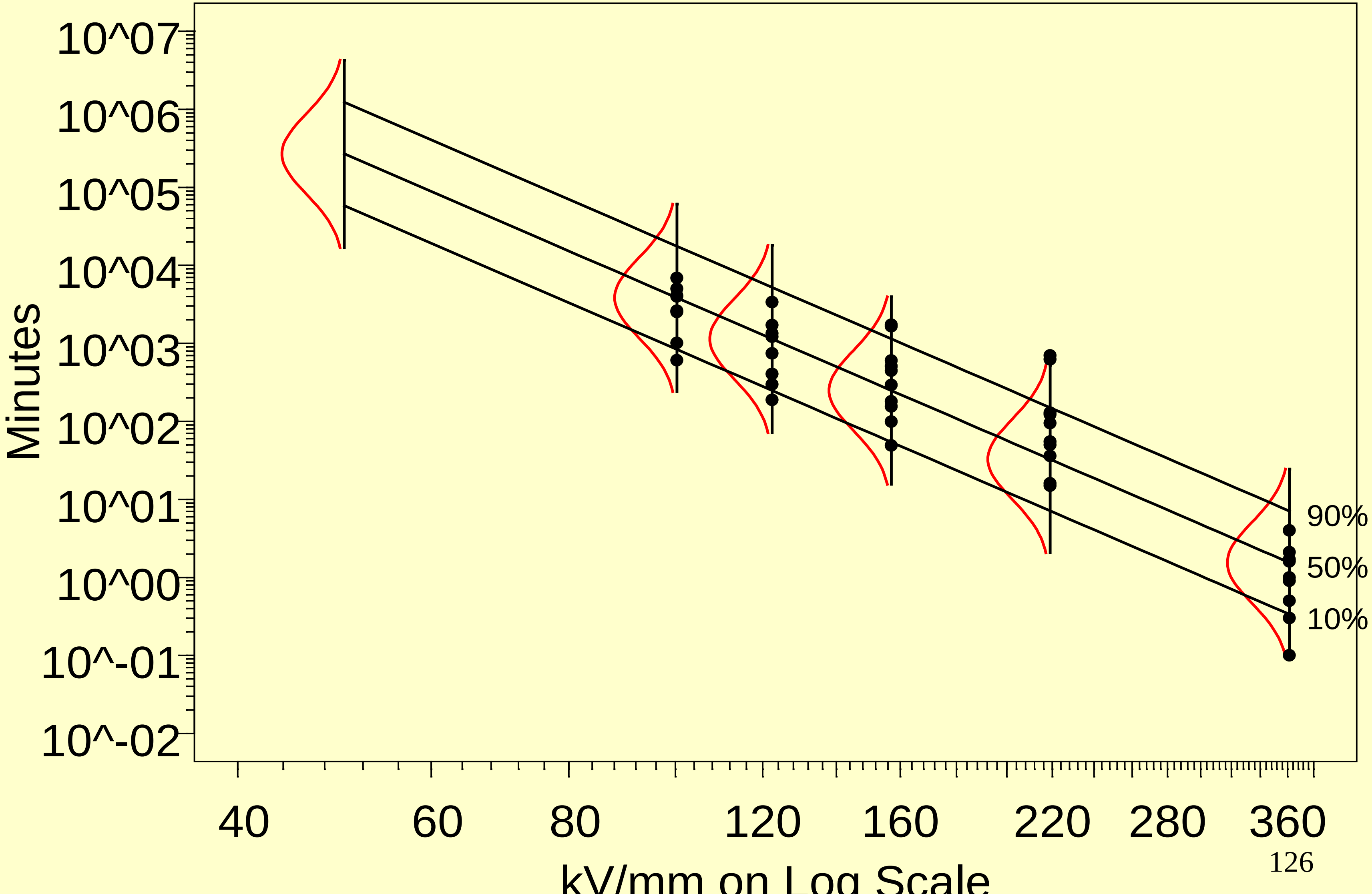
$$\Pr[T(\text{temp}) \le t] = \Phi_{\text{NOR}}\left[\frac{\log(\text{t})\text{-}\mu(x)}{\sigma}\right]$$

$$\mu(x) = \beta_0 + \beta_1 x$$

- x = log(Voltage Stress)
- β_1 is negative because life is shorter at higher levels of voltage stress
- σ is assumed not to depend on voltage stress
- Similar relationships used for pressure and cycling rate

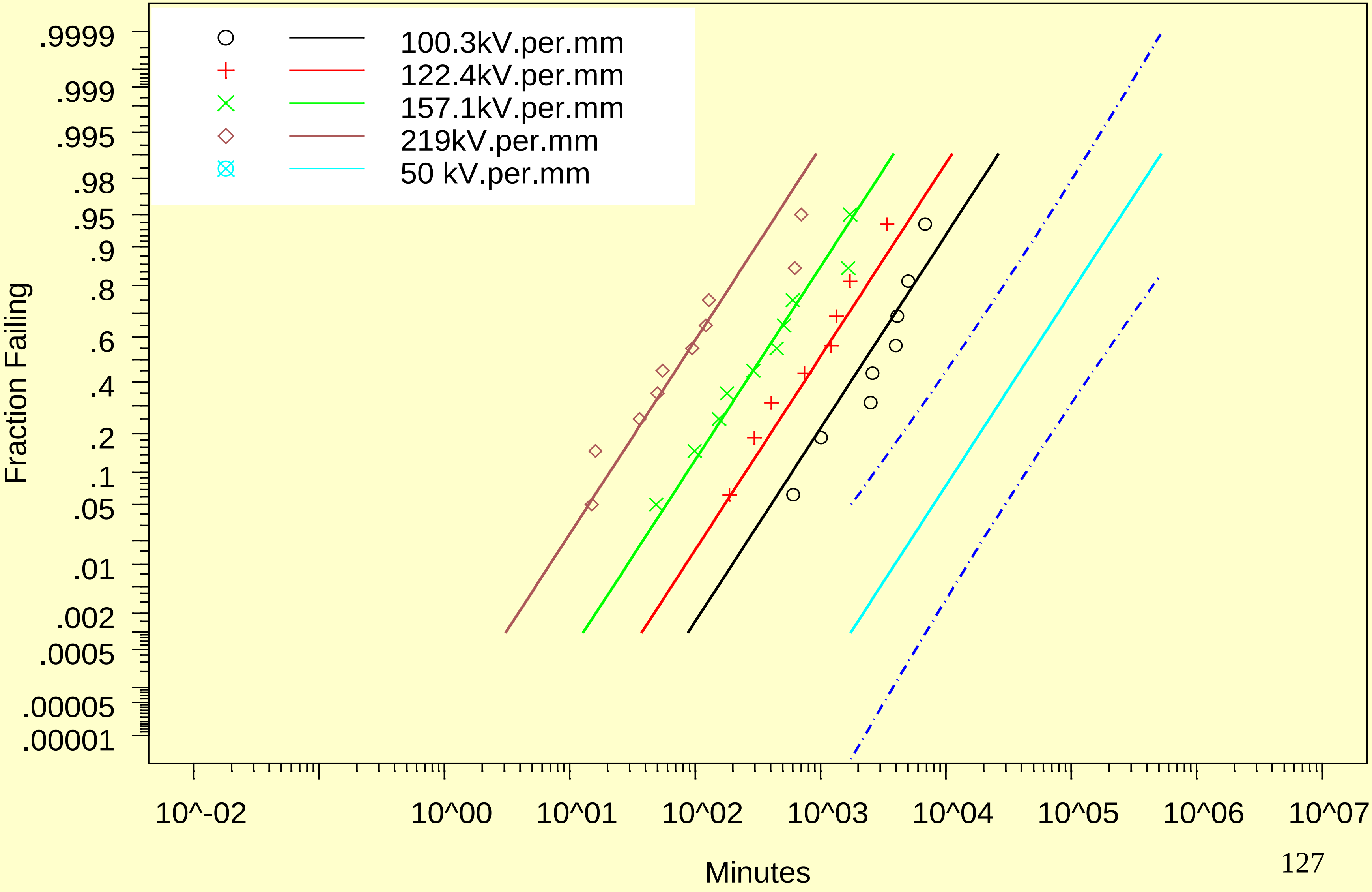
Mylar Polyurethane Insulating Structure Data

Inverse Power Regression Model All Data



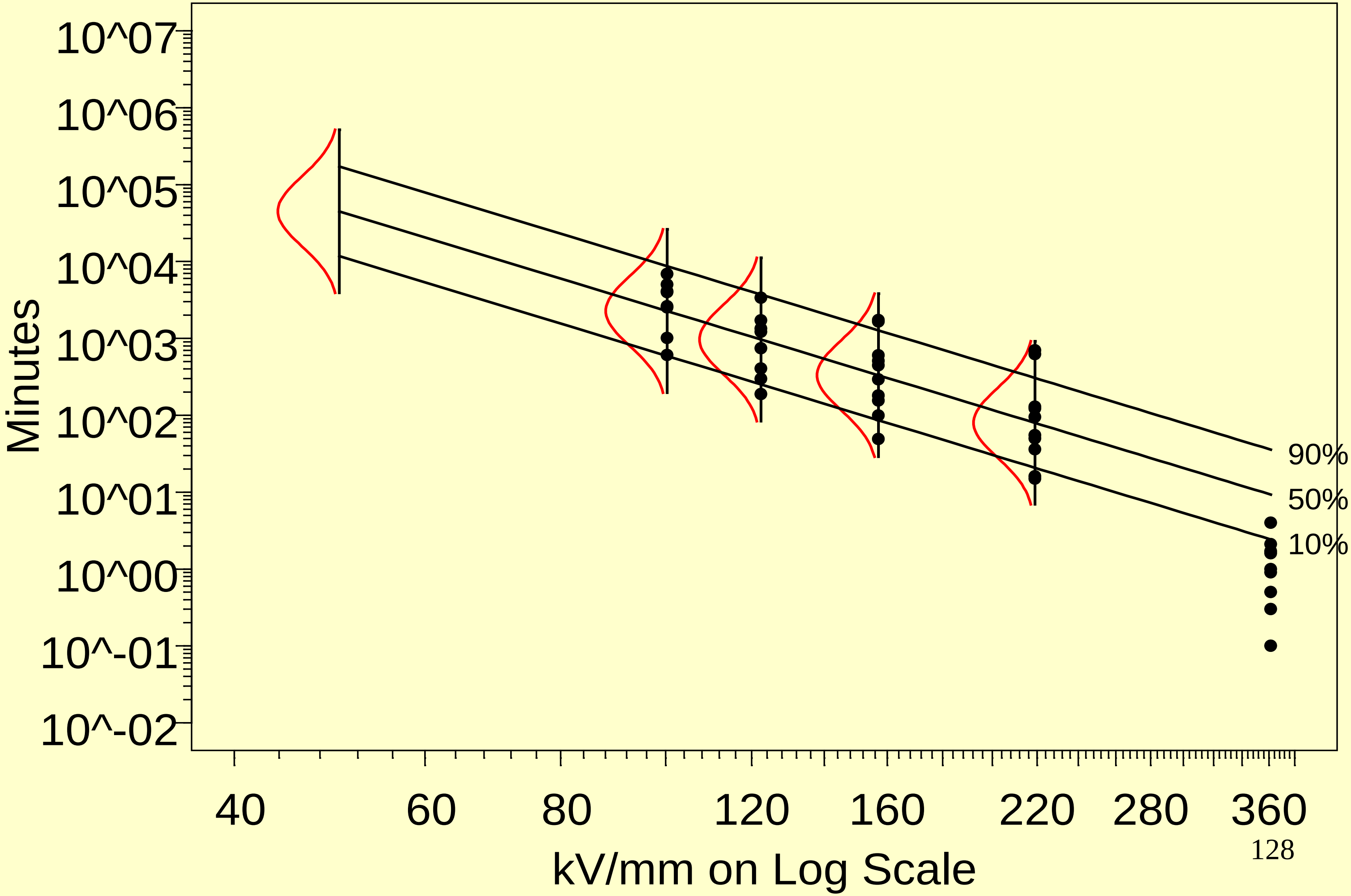
Mylar Polyurethane Insulating Structure Data

Inverse Power Regression Model Good Data



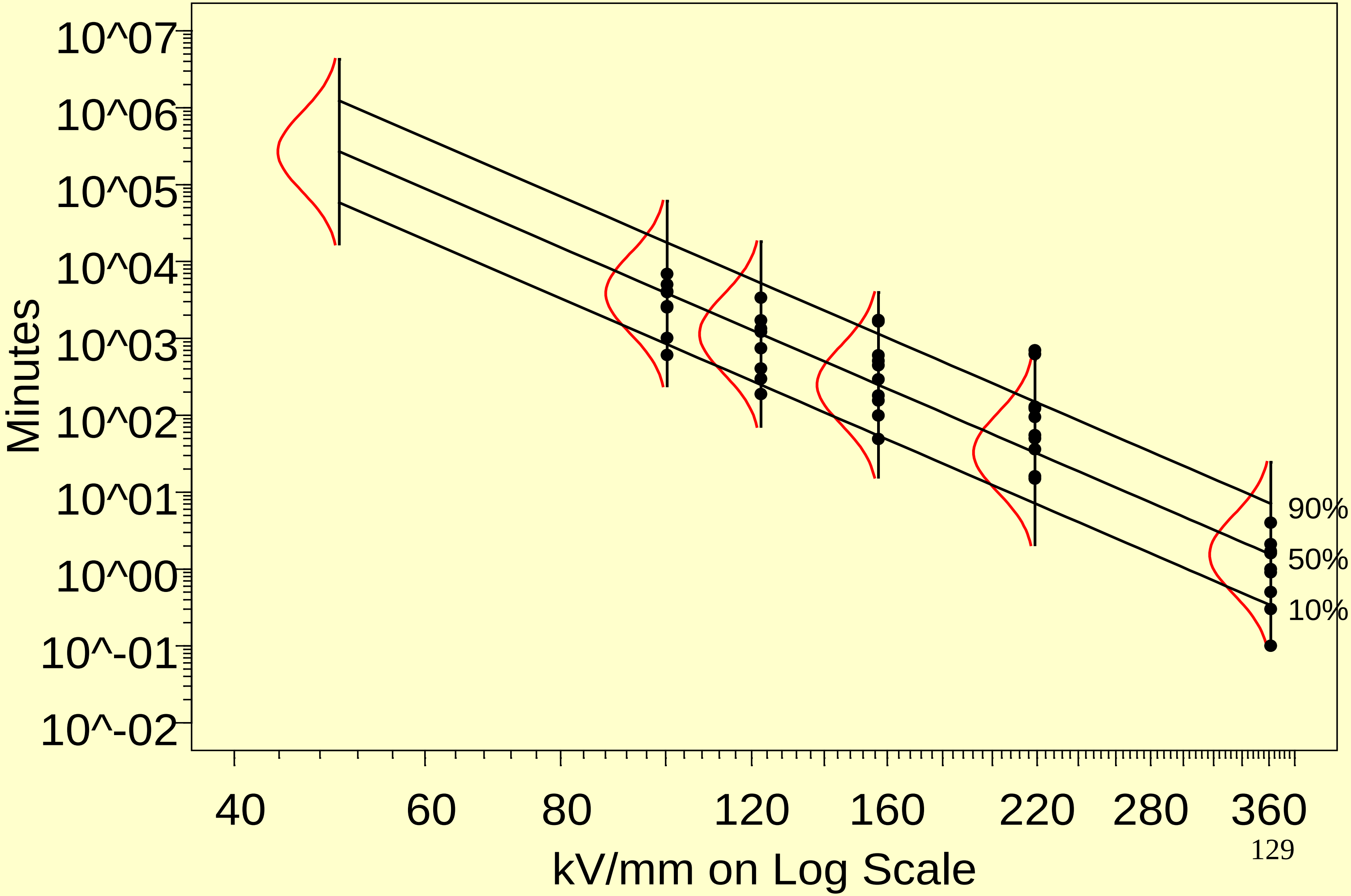
Mylar Polyurethane Insulating Structure Data

Inverse Power Regression Model Good Data



Mylar Polyurethane Insulating Structure Data

Inverse Power Regression Model All Data

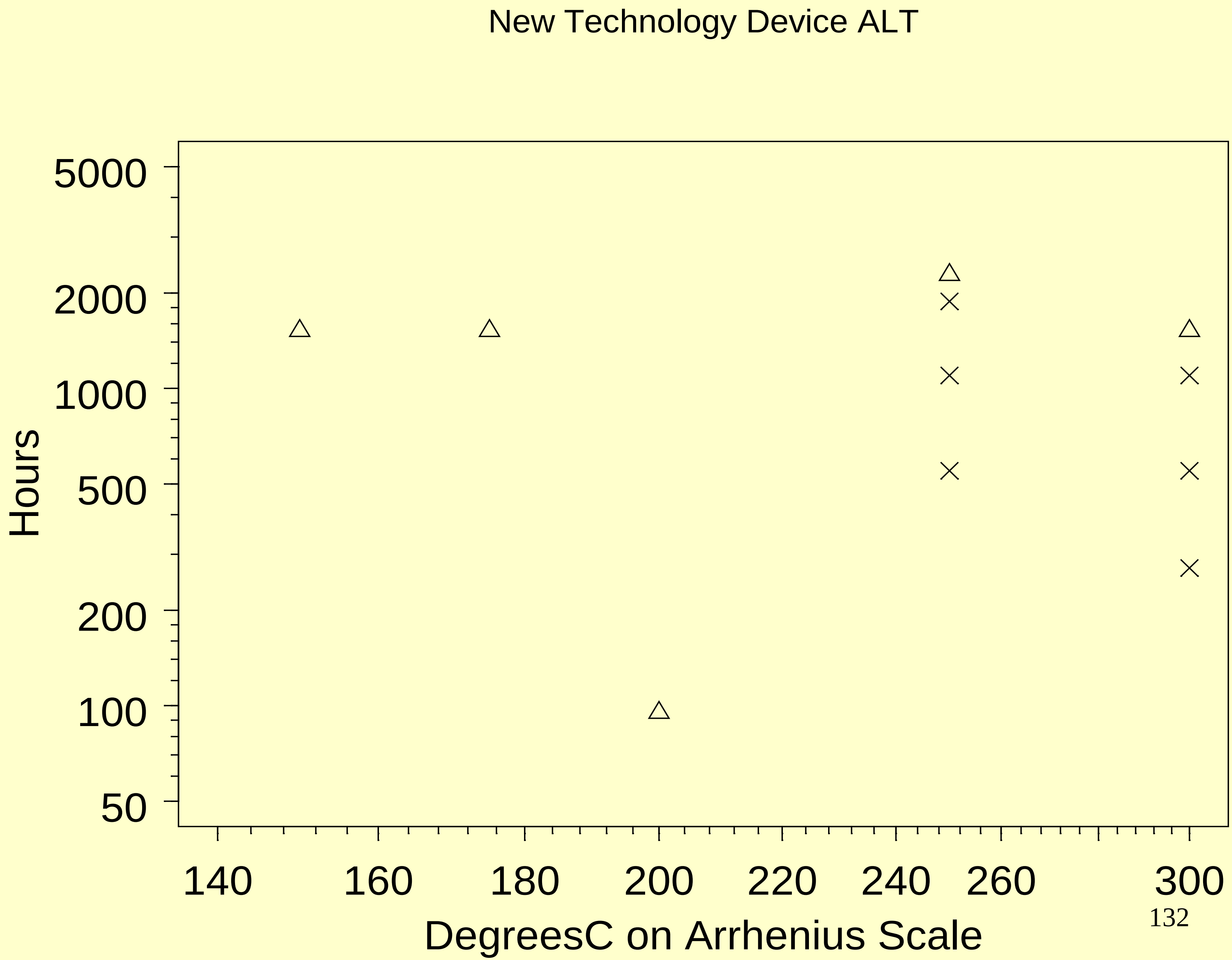


Lessons Learned

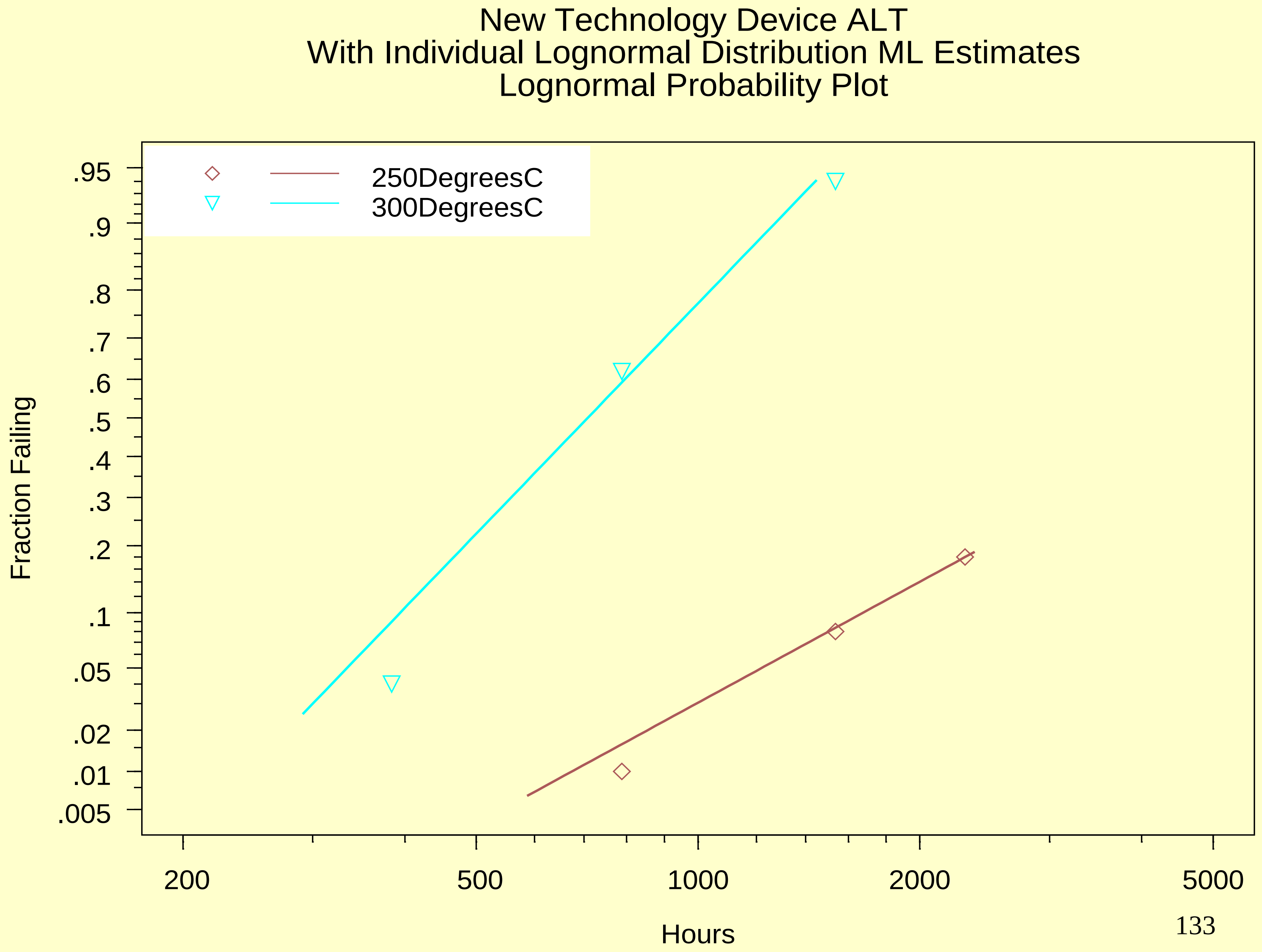
- Transformation of data can simplify modeling
- The inverse power relationship (log life is linear in log(Voltage Stress) may be useful for modeling dielectric life
- Testing at a voltage stress that is too high can cause new failure modes
- New failure modes at the higher levels of stress, if unrecognized, leads to incorrectly optimistic conclusions
- Structure on a data plot can hide important information

New Technology IC Device ALT

- Data from Meeker and Escobar (1998)
- 50 devices tested at each of 150, 175, 200, 250, and 300 degrees C
- Interval censoring (inspection of devices approximately every four days)
- Failures seen only at 250 and 300 degrees C
- Need to estimate the life distribution at 100 degrees C



New Technology Device ALT
Hours
5000
2000
1000
500
200
100
50
140
160
180
200
220
240
260
300
DegreesC on Arrhenius Scale



New Technology Device ALT
With Individual Lognormal Distribution ML Estimates
Lognormal Probability Plot
250DegreesC
300DegreesC
Fraction Failing
.95
.9
.8
.7
.6
.5
.4
.3
.2
.1
.05
.02
.01
.005
200
500
1000
2000
5000
Hours

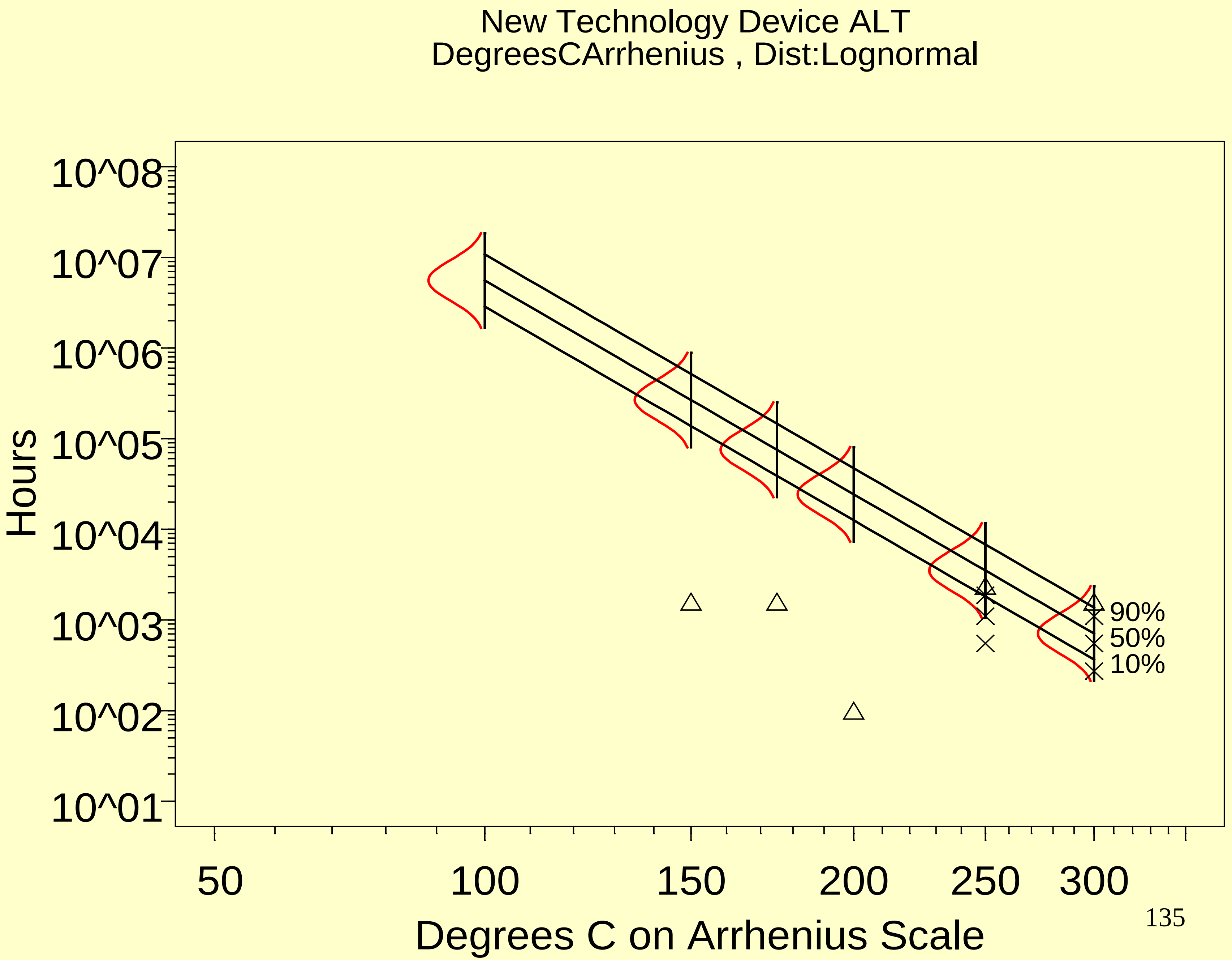
The Arrhenius/Lognormal Model for Temperature Acceleration

- Log life is inversely proportional to reciprocal Kelvin temperature
- The probability of failure as a function of time is

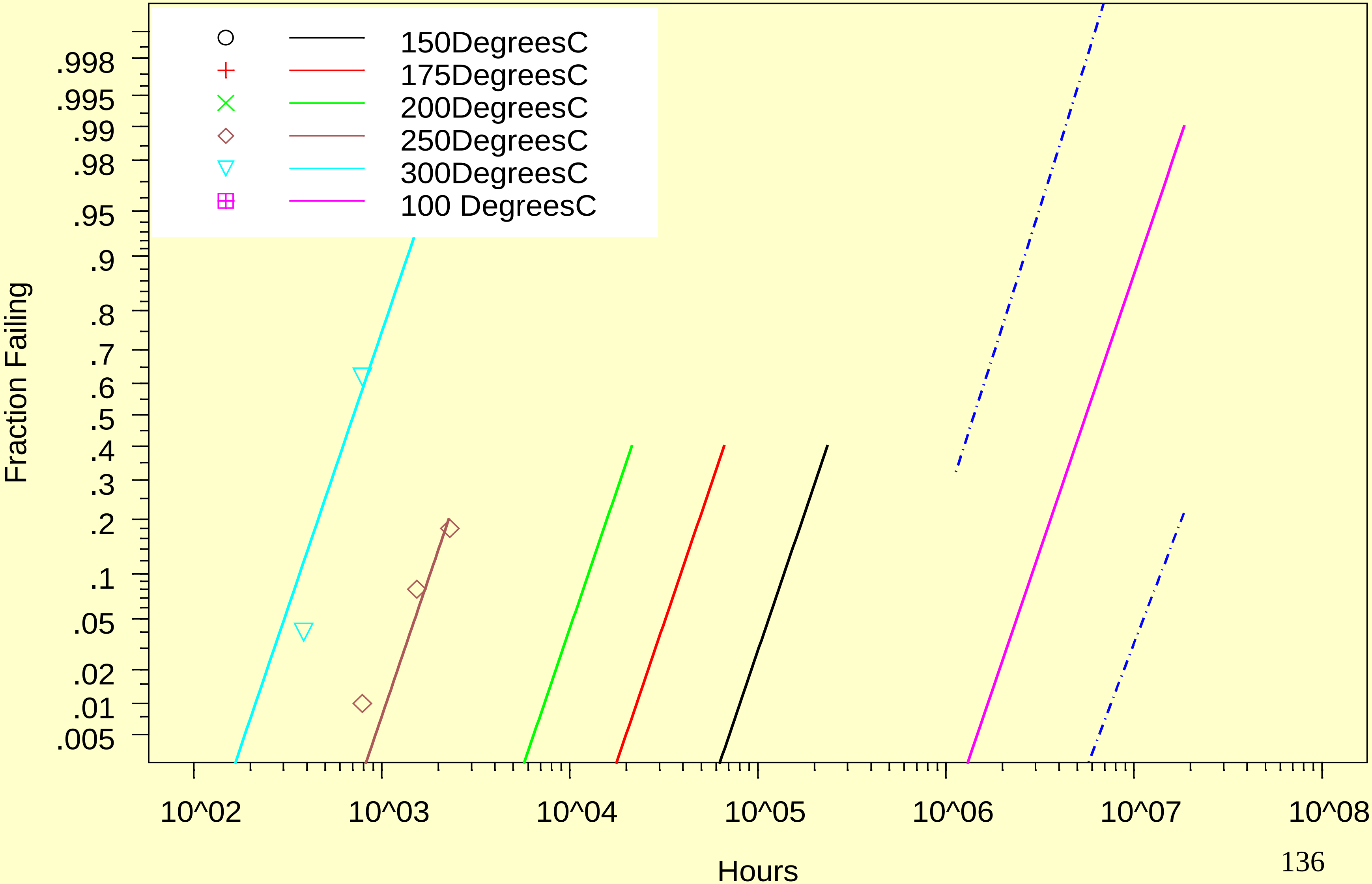
$$\Pr[T(\text{temp}) \le t] = \Phi_{\text{NOR}}\left[\frac{\log(\text{t})\text{-}\mu(x)}{\sigma}\right]$$

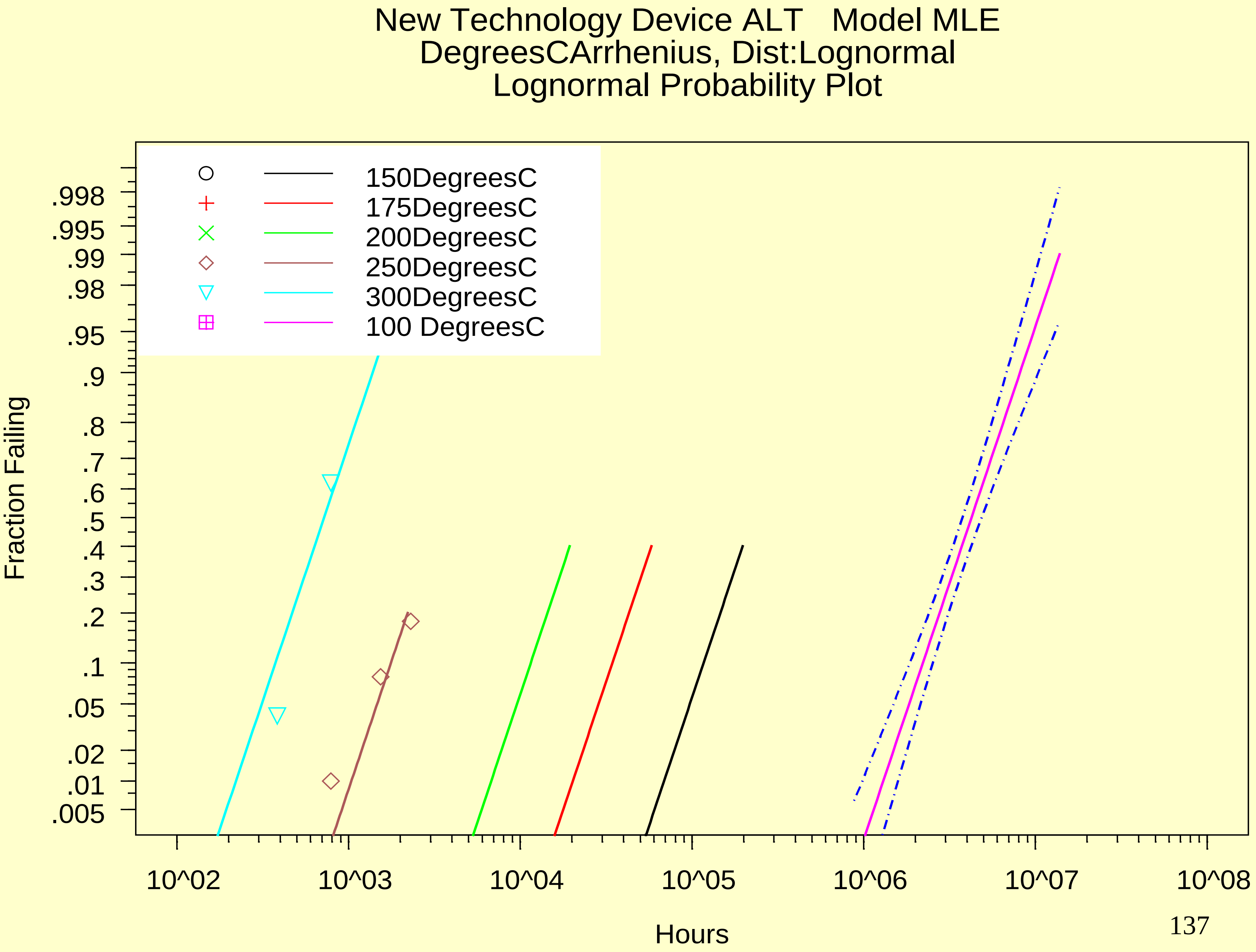
$$\mu(x) = \beta_0 + \beta_1 x$$

- x = 11605/(Degrees C + 273.15)
- β_1 is the “effective activation energy”
- σ does not depend on temperature
- Widely used in the evaluation of electronic components


New Technology Device ALT
DegreesCArrhenius , Dist:Lognormal
Hours
10^08
10^07
10^06
10^05
10^04
10^03
10^02
10^01
90%
50%
10%
50
100
150
200
250
300
Degrees C on Arrhenius Scale

New Technology Device ALT Model MLE
DegreesCArrhenius, Dist:Lognormal
Lognormal Probability Plot



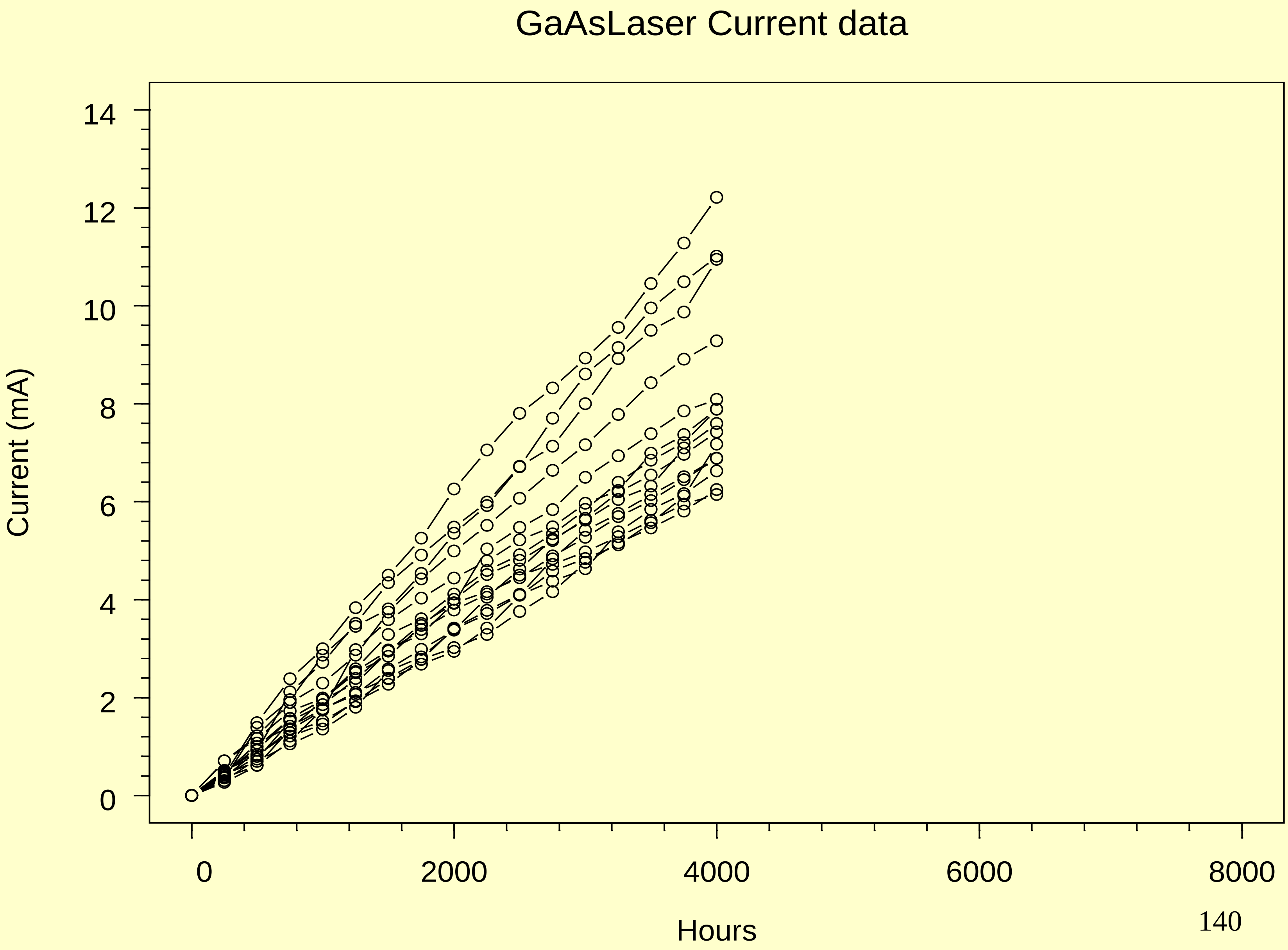

New Technology Device ALT Model MLE
DegreesCArrhenius, Dist:Lognormal
Lognormal Probability Plot
150DegreesC
175DegreesC
200DegreesC
250DegreesC
300DegreesC
100 DegreesC
Fraction Failing
.998
.995
.99
.98
.95
.9
.8
.7
.6
.5
.4
.3
.2
.1
.05
.02
.01
.005
10^02
10^03
10^04
10^05
10^06
10^07
10^08
Hours

Lessons Learned

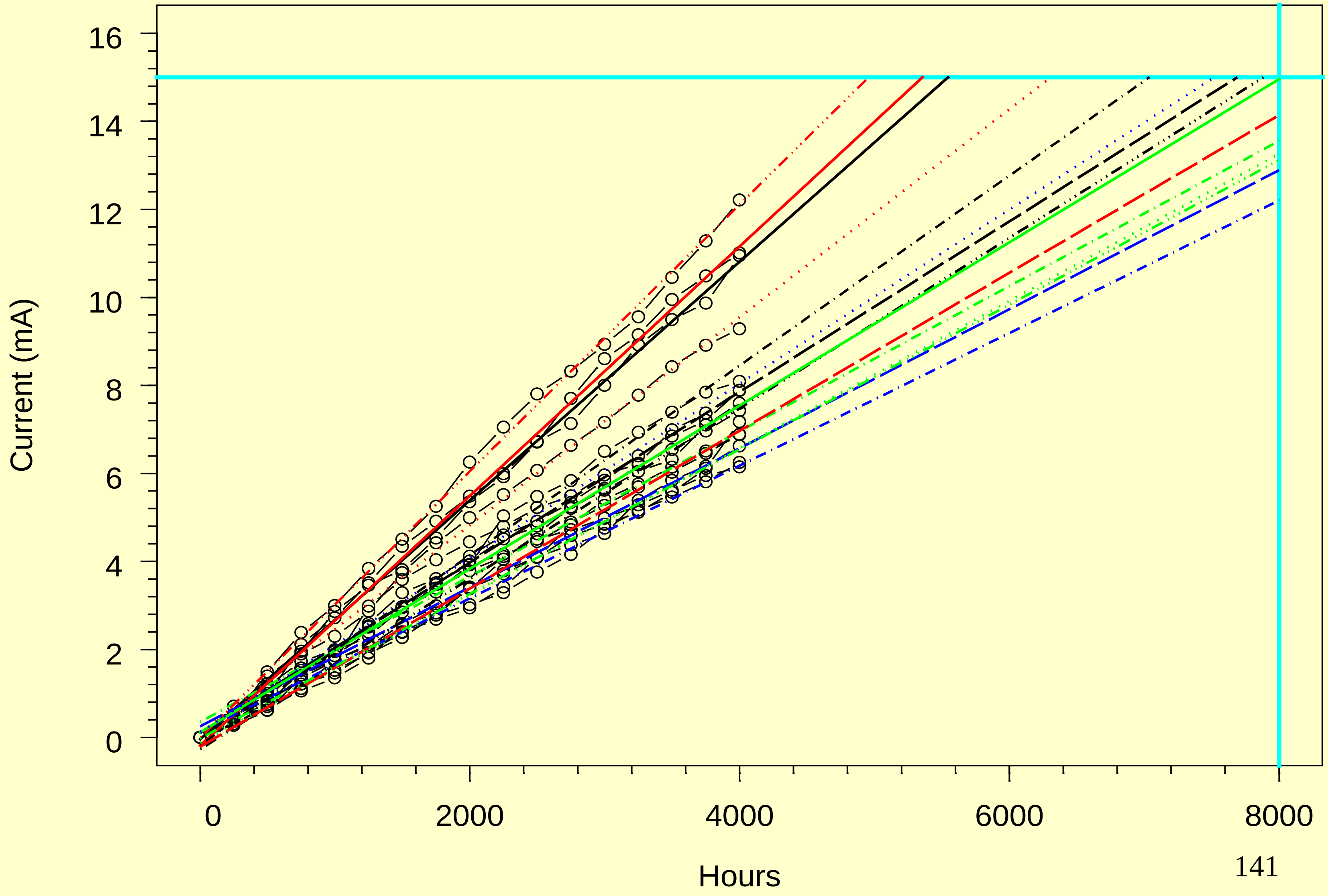
- Maximum likelihood can easily deal with interval-censored data
- A change in the distribution slope parameter implies that simple acceleration models are not appropriate
- Too much stress can cause new failure modes
- Statistical error is greatly amplified when extrapolating
- Engineering information is important in accelerated life testing

GaAs Laser Repeated Measures Degradation Data Analysis

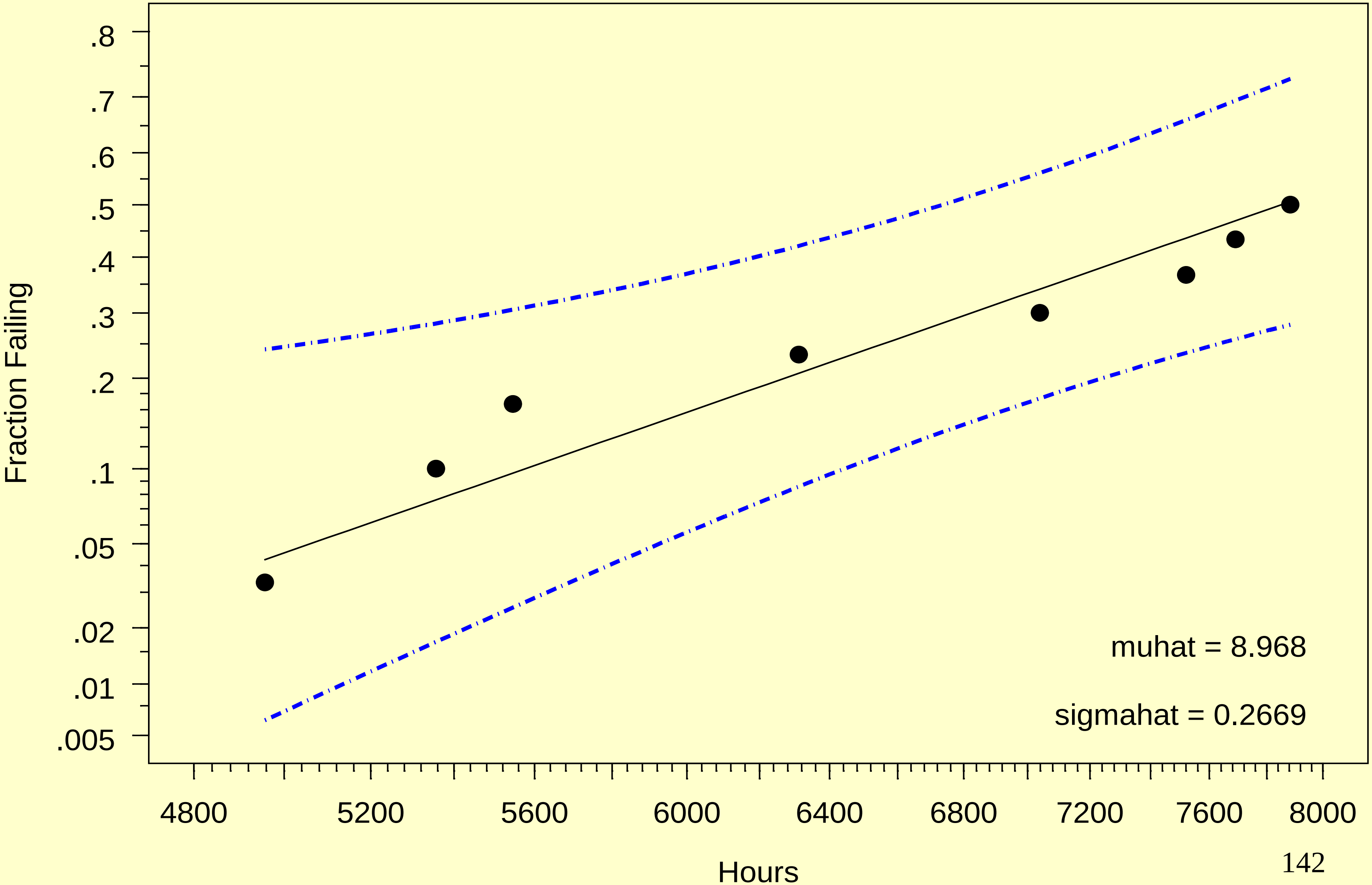
- Lasers used in telecommunications applications
- Lasers tested at 80 degrees C to accelerate life
- Few failures expected even in the accelerated life test


GaAsLaser Current data
Current (mA)
Hours
0
2
4
6
8
10
12
14
0
2000
4000
6000
8000

GaAsLaser.CurrentF15C8000.XLinear.YLinear.ld



GaAsLaser.CurrentF15C8000.XLinear.YLinear.ld
with Lognormal ML Estimate and Pointwise 95% Confidence Intervals
Lognormal Probability Plot

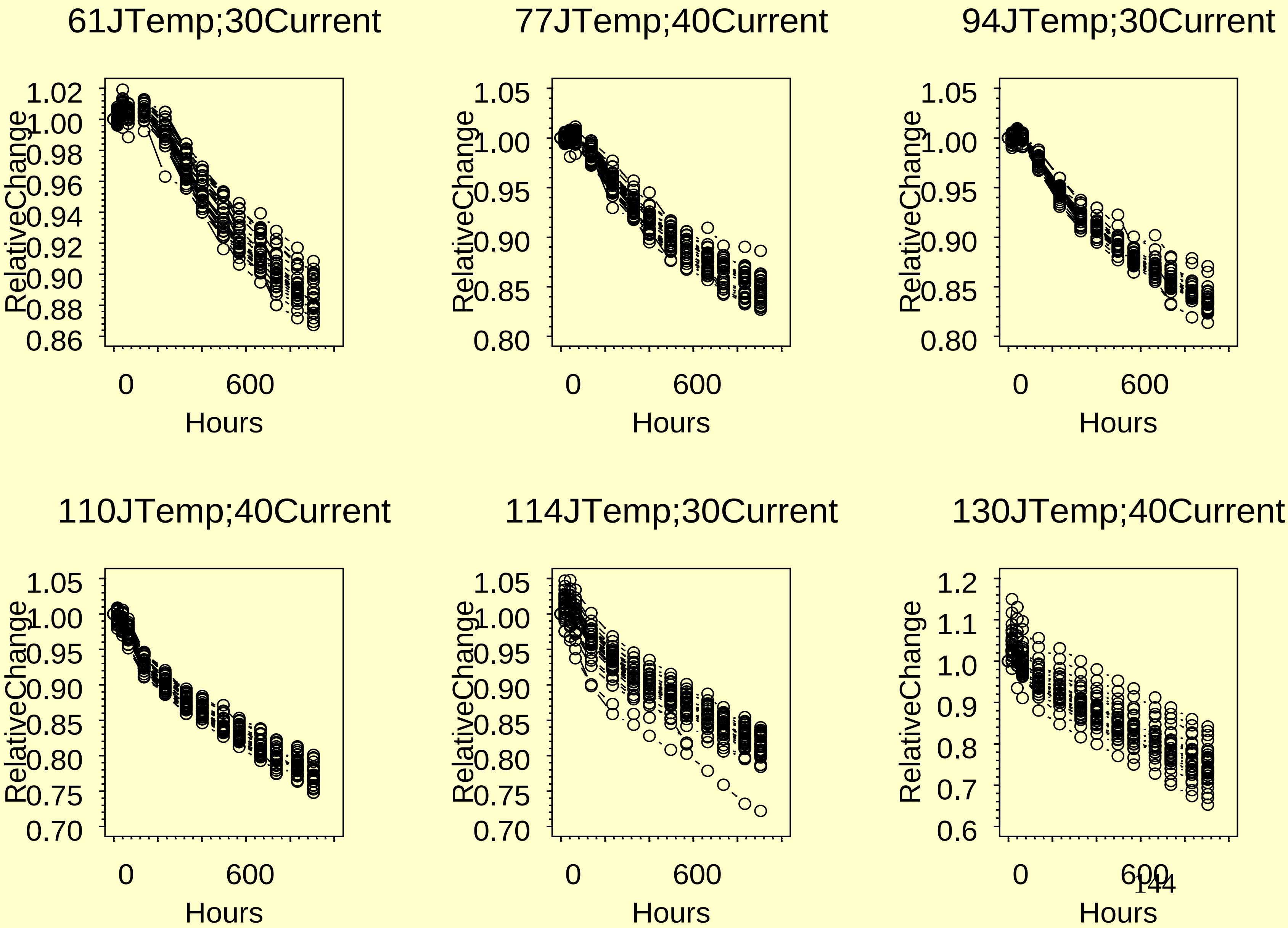


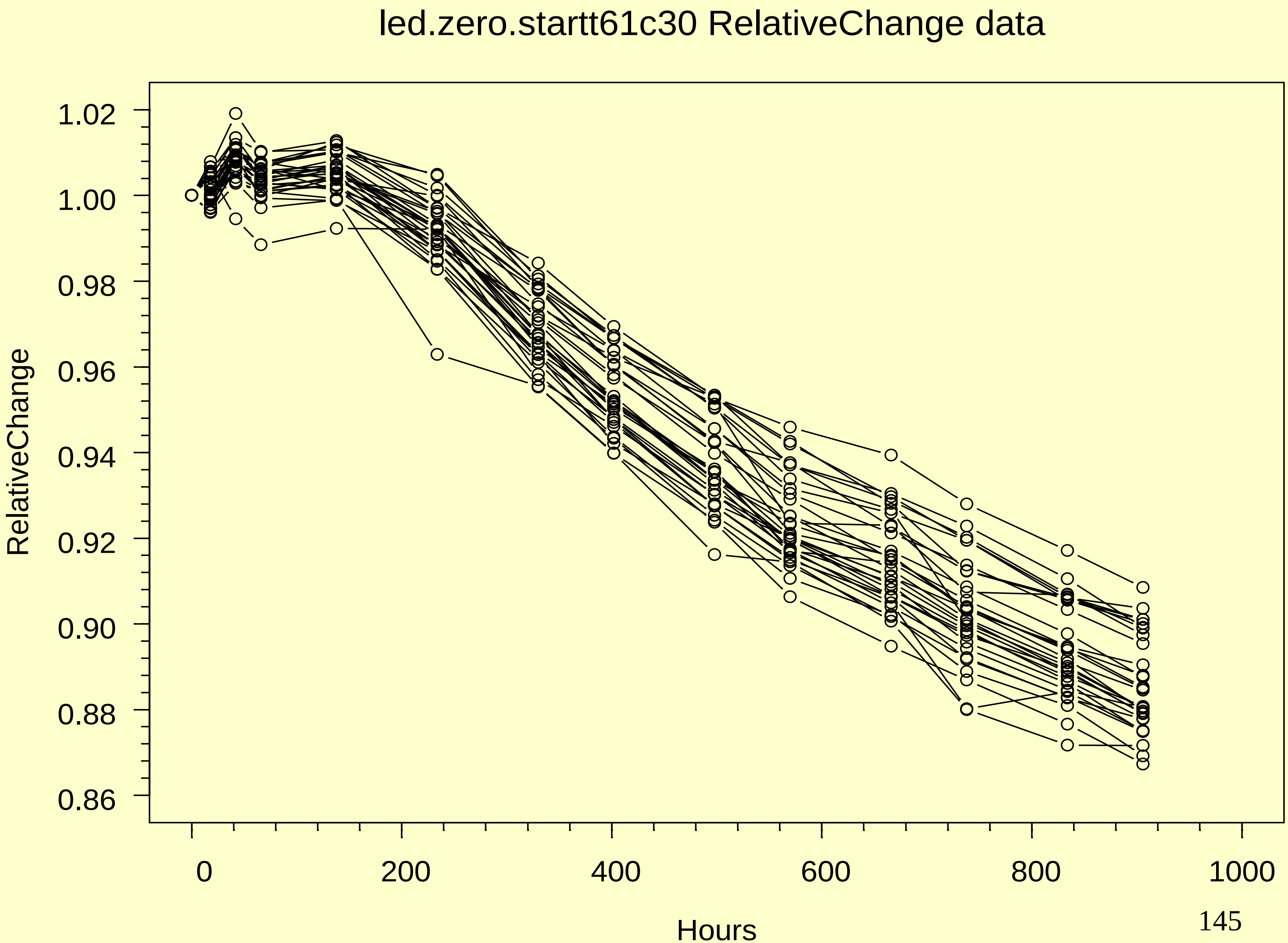
Accelerated Repeated Measures Degradation Test of LEDs

- Tests run with 50 LEDs at each of six combinations of temperature and current (two accelerating variables)
- Measured light output on each unit periodically
- Data are messy

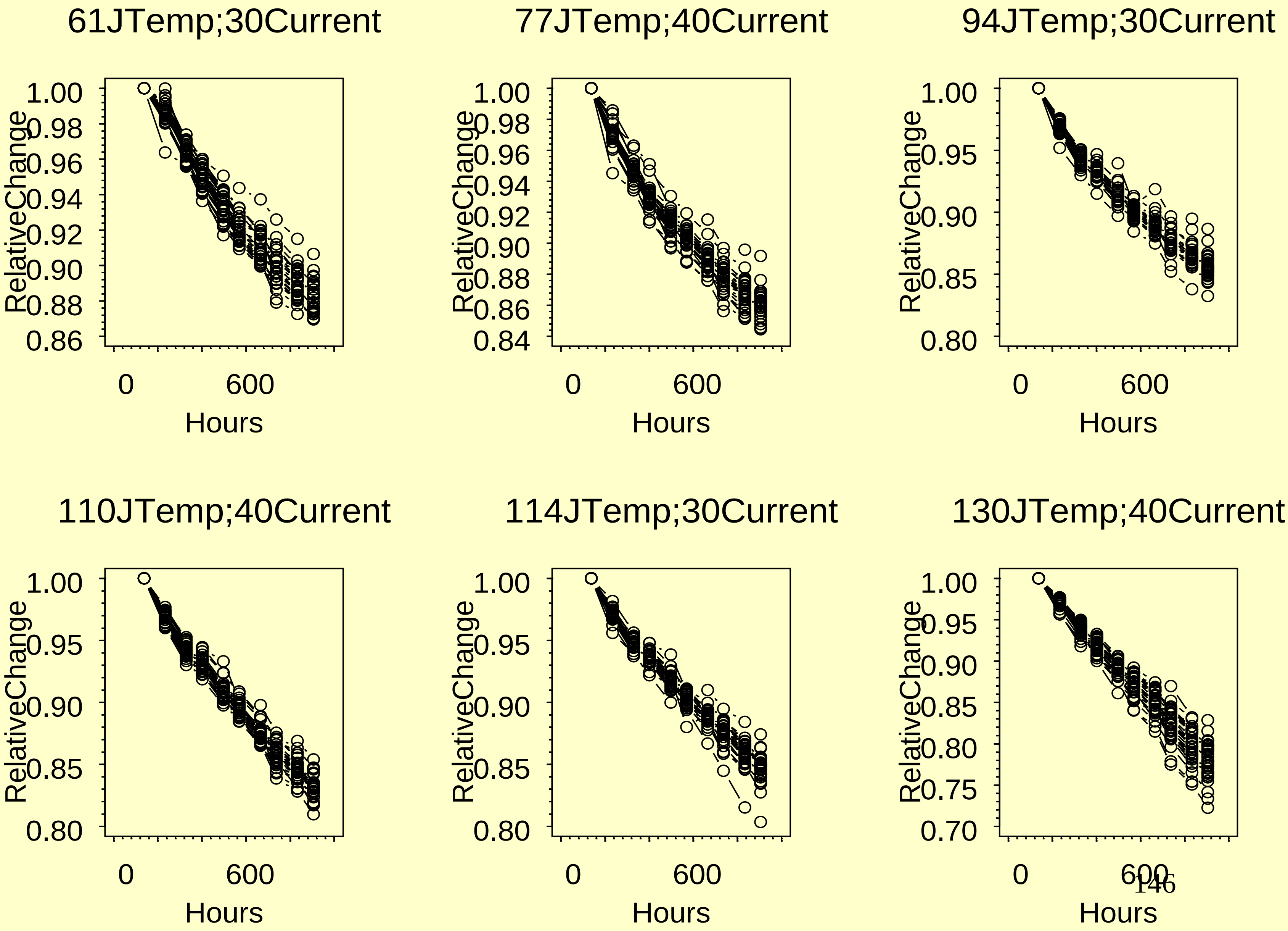
LED Relative Change in Light Output (Zero Start)

x axis: linear y axis: linear

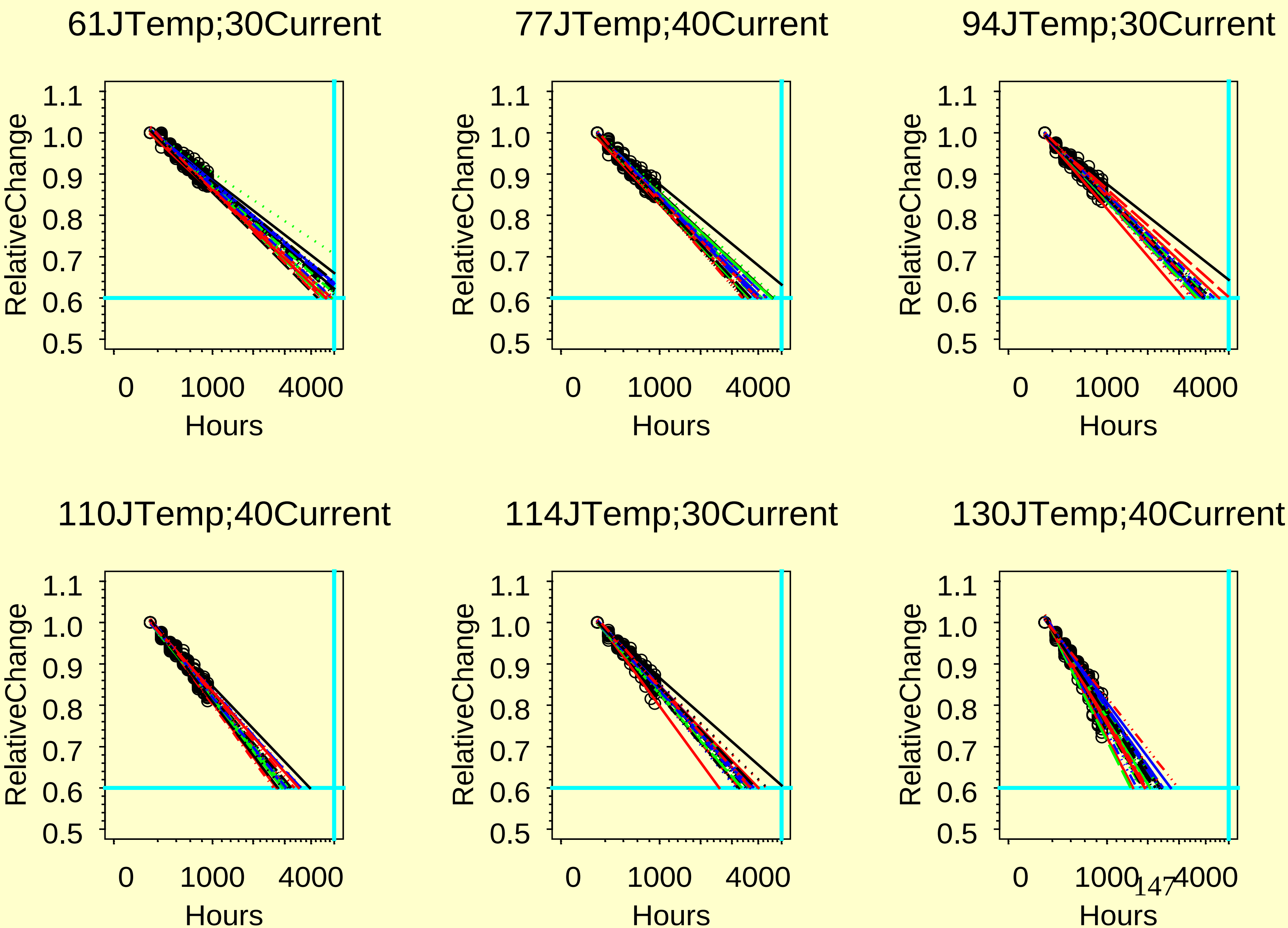


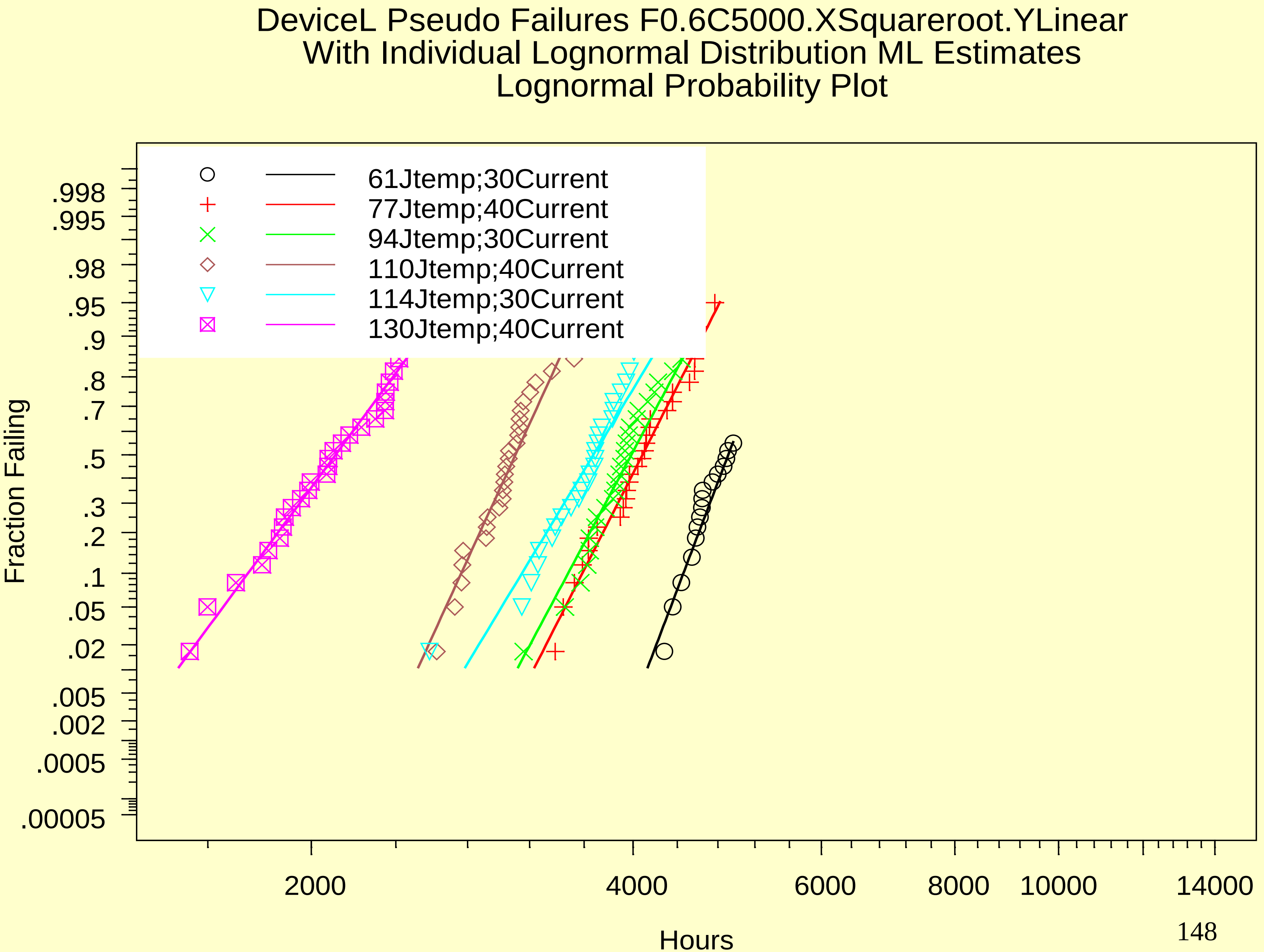

led.zero.startt61c30 RelativeChange data
RelativeChange
1.02
1.00
0.98
0.96
0.94
0.92
0.90
0.88
0.86
0
200
400
600
800
1000
Hours

Relative Change in Light Output from 138 Hours
x axis: linear   y axis: linear

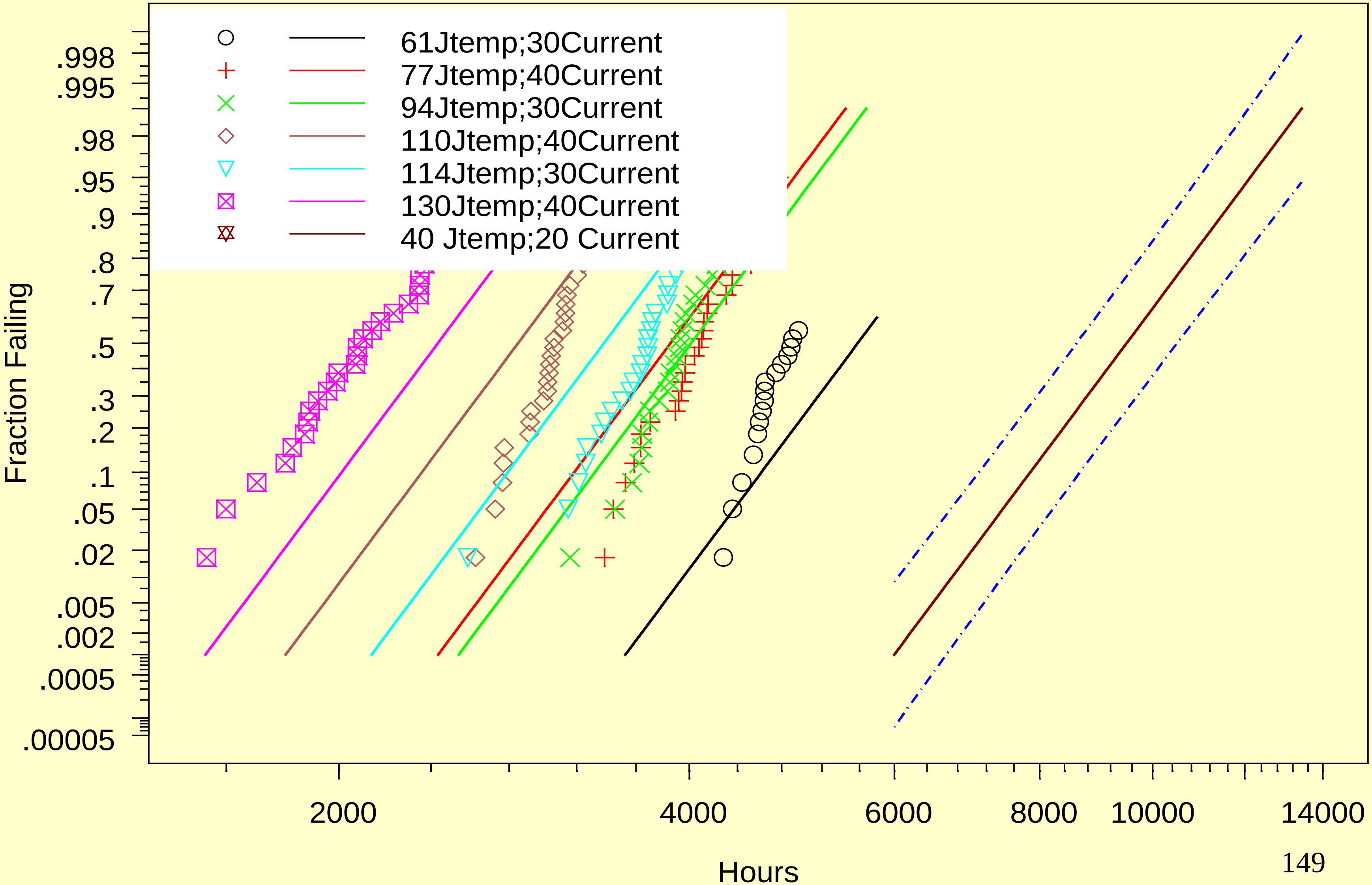


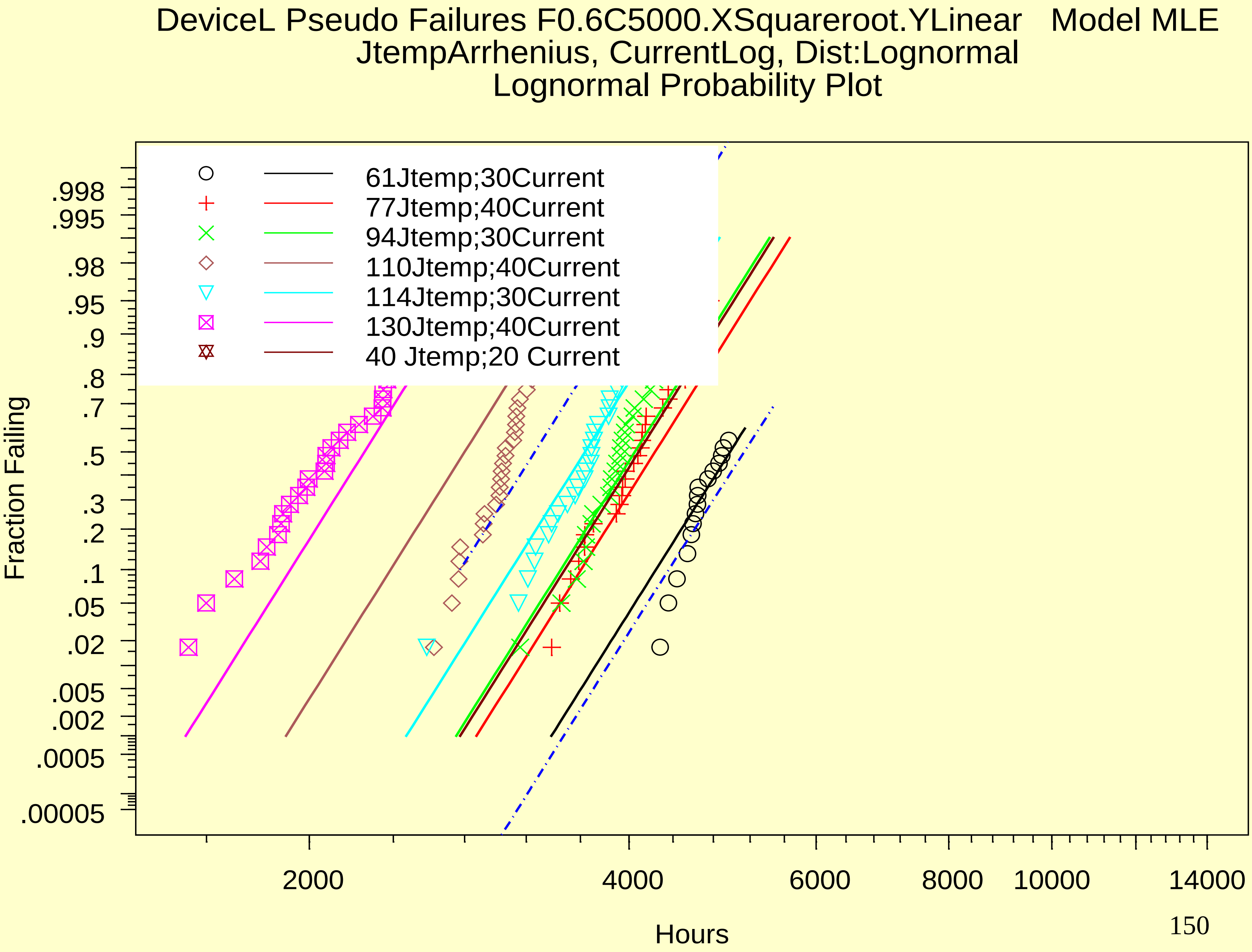
LED Pseudo Failures F0.6C5000.XSquareroot.YLinear



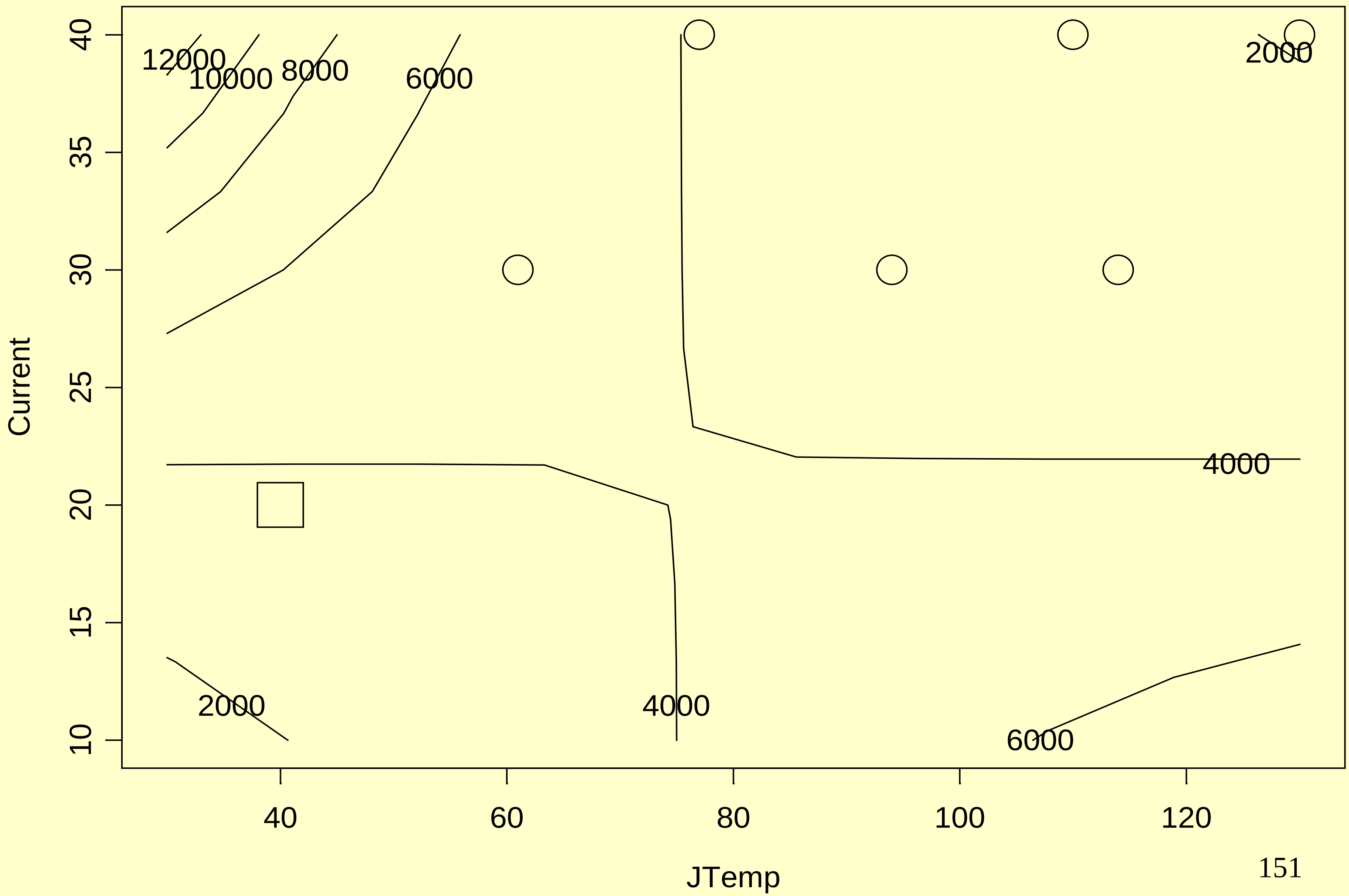

DeviceL Pseudo Failures F0.6C5000.XSquareroot.YLinear
With Individual Lognormal Distribution ML Estimates
Lognormal Probability Plot
61Jtemp;30Current
77Jtemp;40Current
94Jtemp;30Current
110Jtemp;40Current
114Jtemp;30Current
130Jtemp;40Current
Fraction Failing
.998
.995
.98
.95
.9
.8
.7
.5
.3
.2
.1
.05
.02
.005
.002
.0005
.00005
2000
4000
6000
8000
10000
14000
Hours

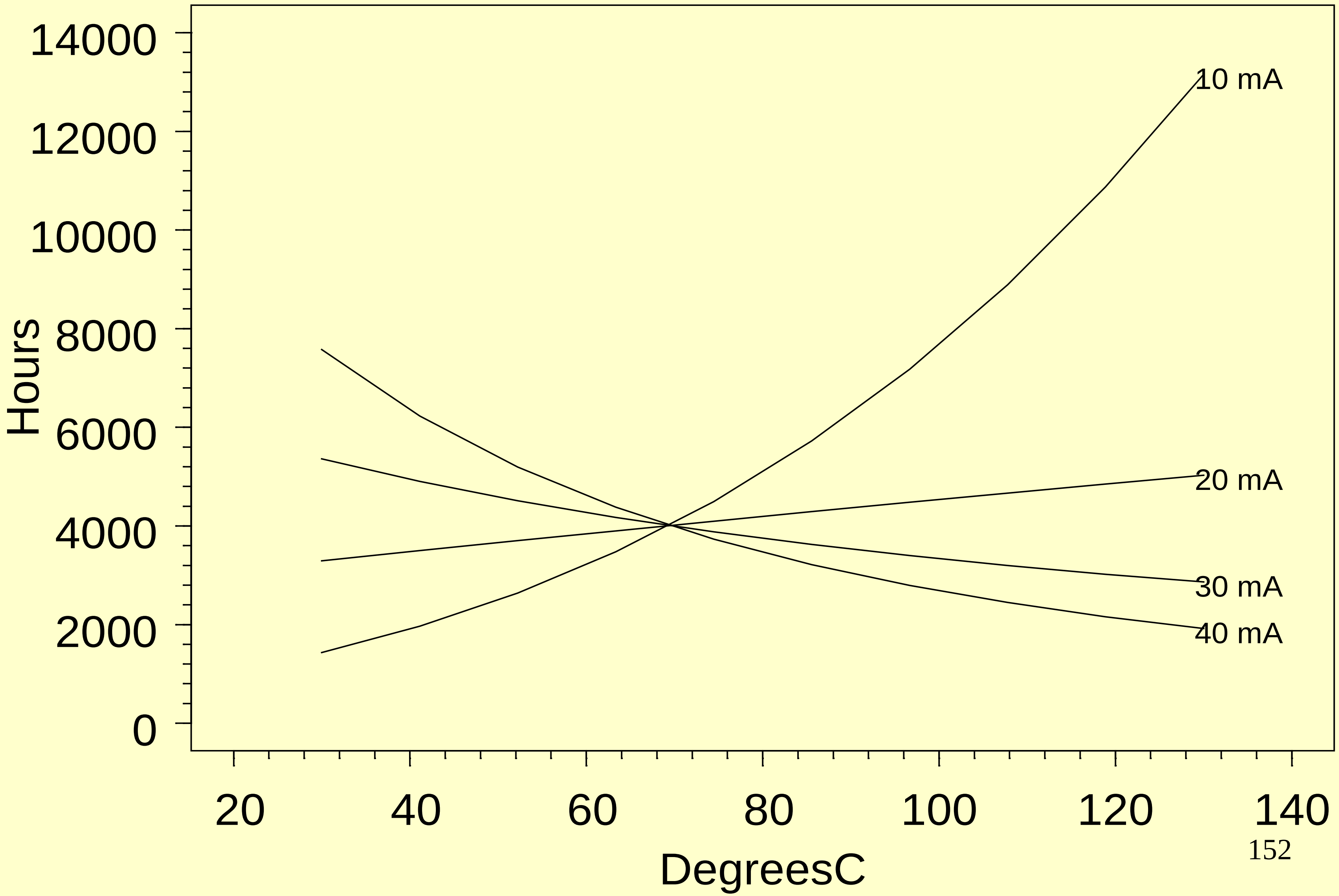
DeviceL Pseudo Failures F0.6C5000.XSquareroot.YLinear Model MLE
JtempArrhenius, CurrentLog, Dist:Lognormal
Lognormal Probability Plot



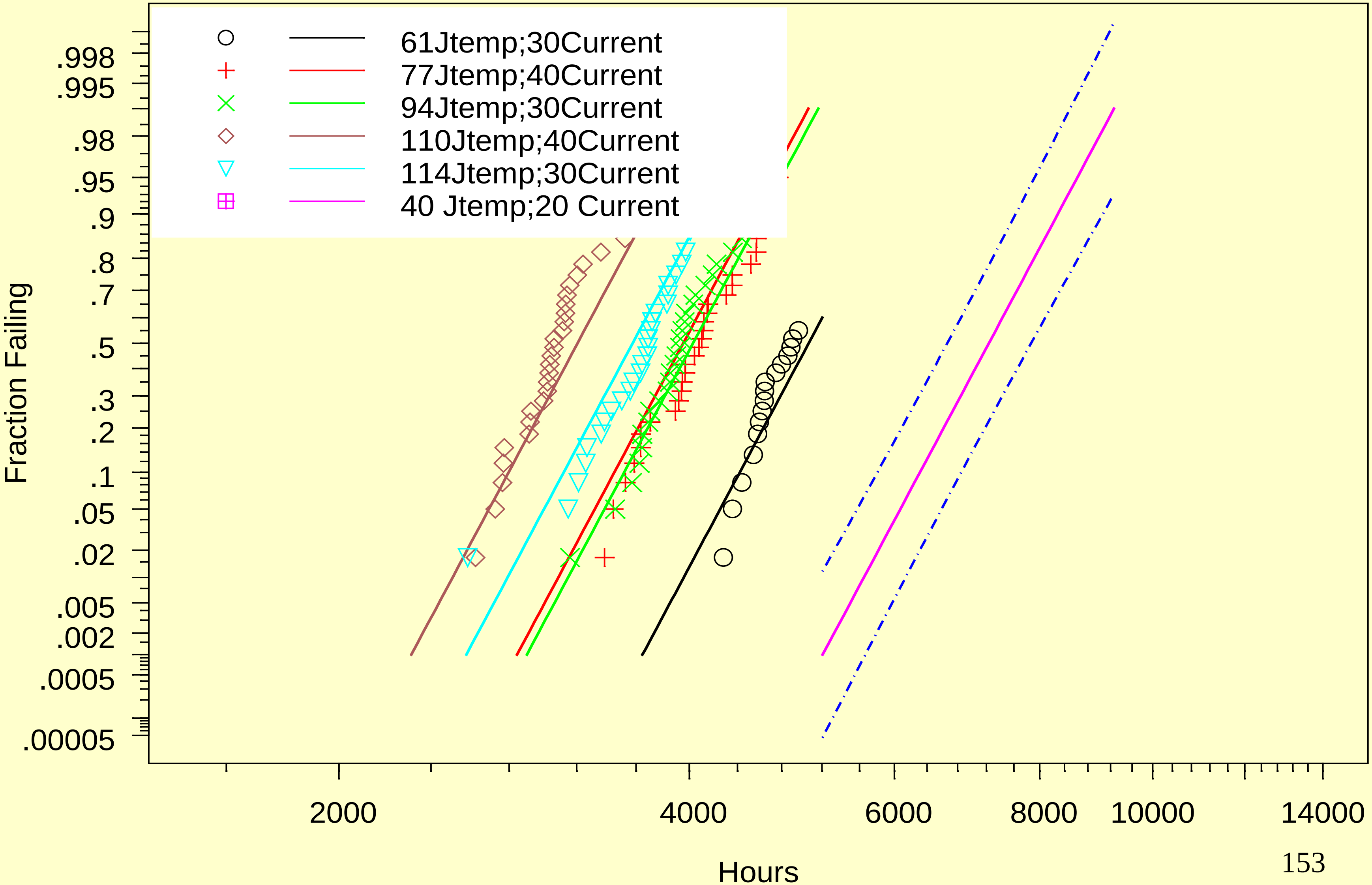

DeviceL Pseudo Failures F0.6C5000.XSquareroot.YLinear Model MLE
JtempArrhenius, CurrentLog, Dist:Lognormal
Lognormal Probability Plot
61Jtemp;30Current
77Jtemp;40Current
94Jtemp;30Current
110Jtemp;40Current
114Jtemp;30Current
130Jtemp;40Current
40 Jtemp;20 Current
Fraction Failing
.998
.995
.98
.95
.9
.8
.7
.5
.3
.2
.1
.05
.02
.005
.002
.0005
.00005
2000
4000
6000
8000
10000
14000
Hours

Linear model with interaction

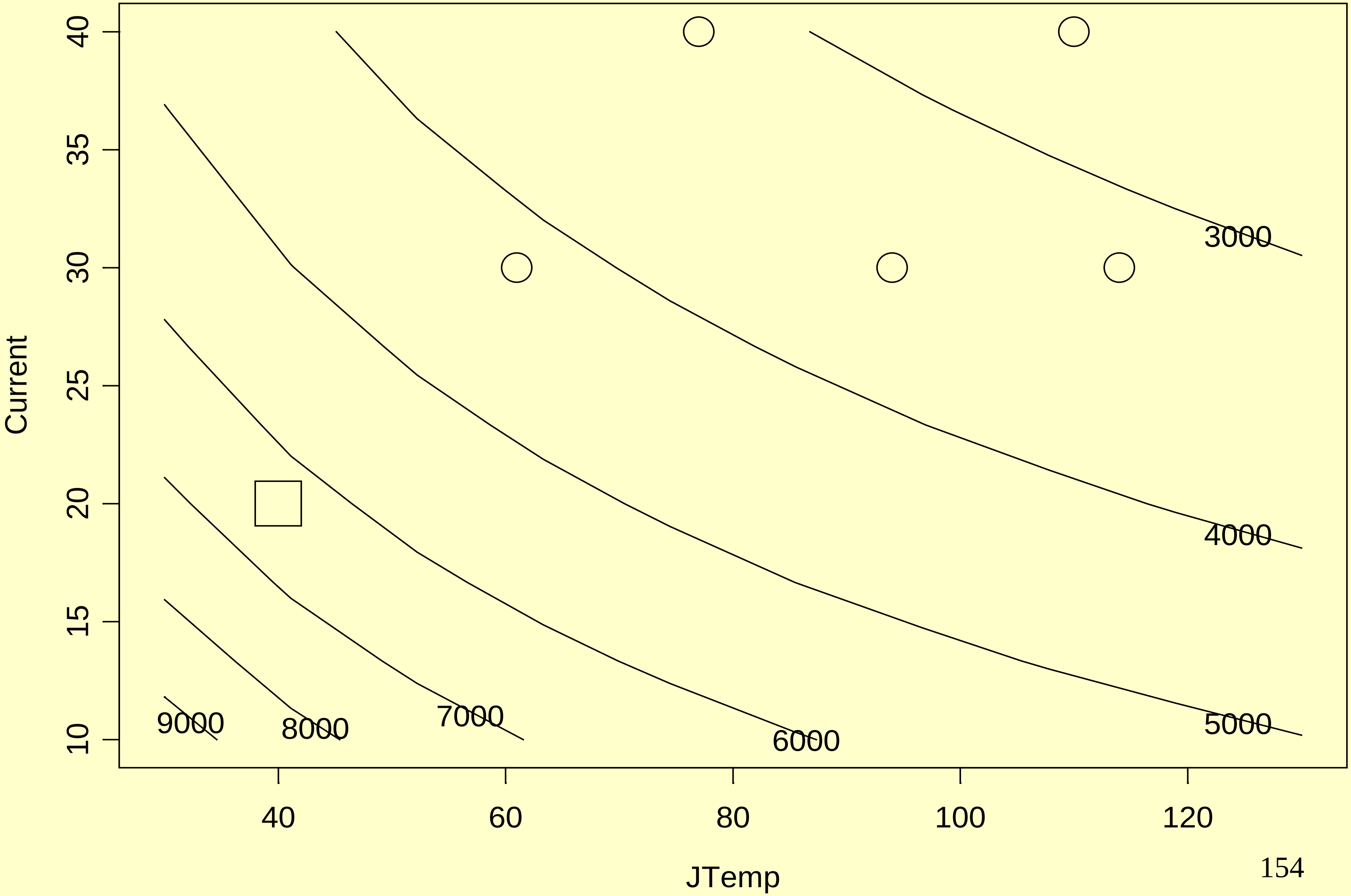


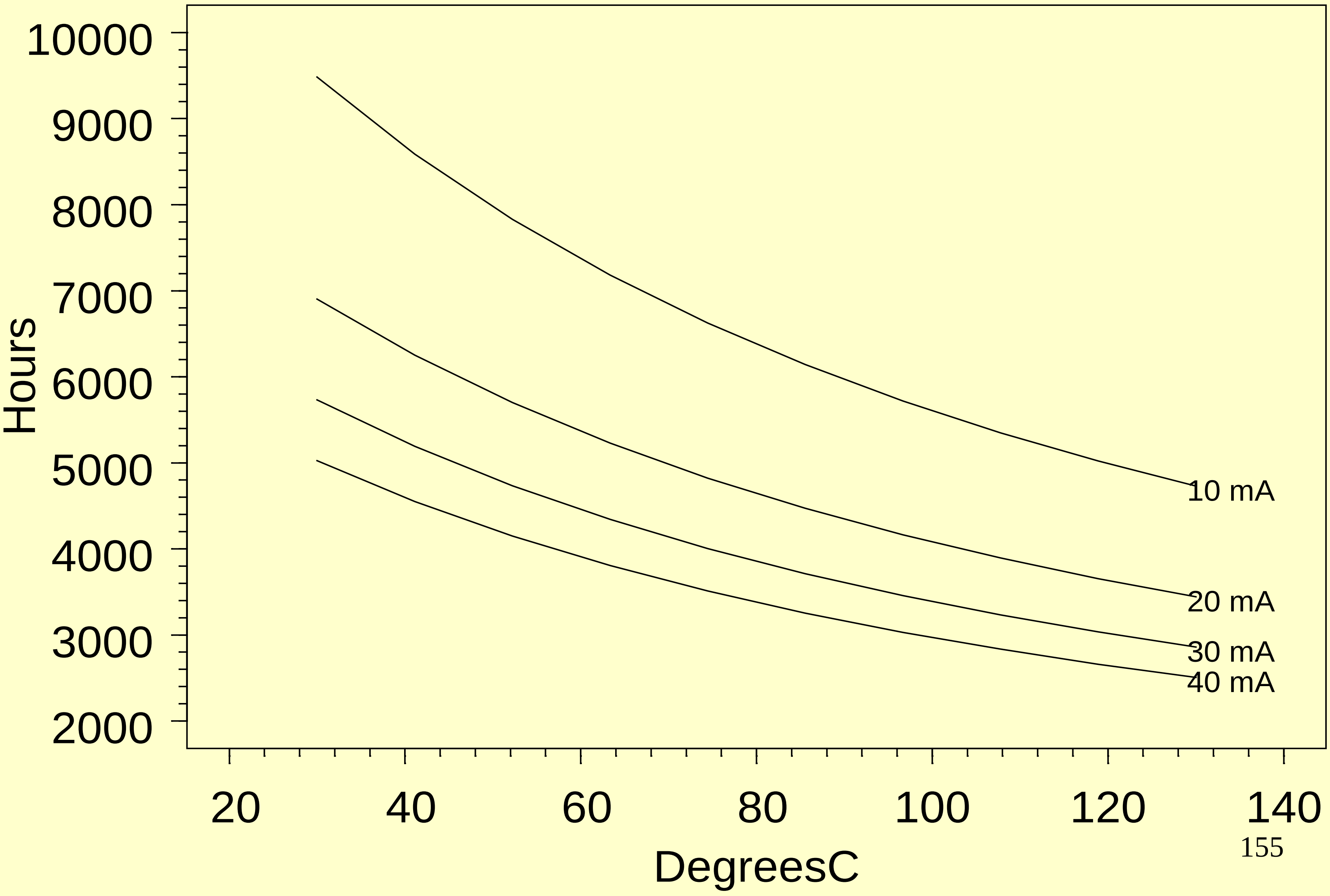

14000
12000
10000
8000
6000
4000
2000
0
Hours
20
40
60
80
100
120
140
DegreesC
10 mA
20 mA
30 mA
40 mA

Fixed DeviceL Pseudo Failures F0.6C5000.XSquareroot.YLinear Model MLE
JtempArrhenius, CurrentLog, Dist:Lognormal
Lognormal Probability Plot



Linear model without JTemp130,Current40 and no interaction




10000
9000
8000
7000
6000
5000
4000
3000
2000
Hours
10 mA
20 mA
30 mA
40 mA
20
40
60
80
100
120
140
DegreesC

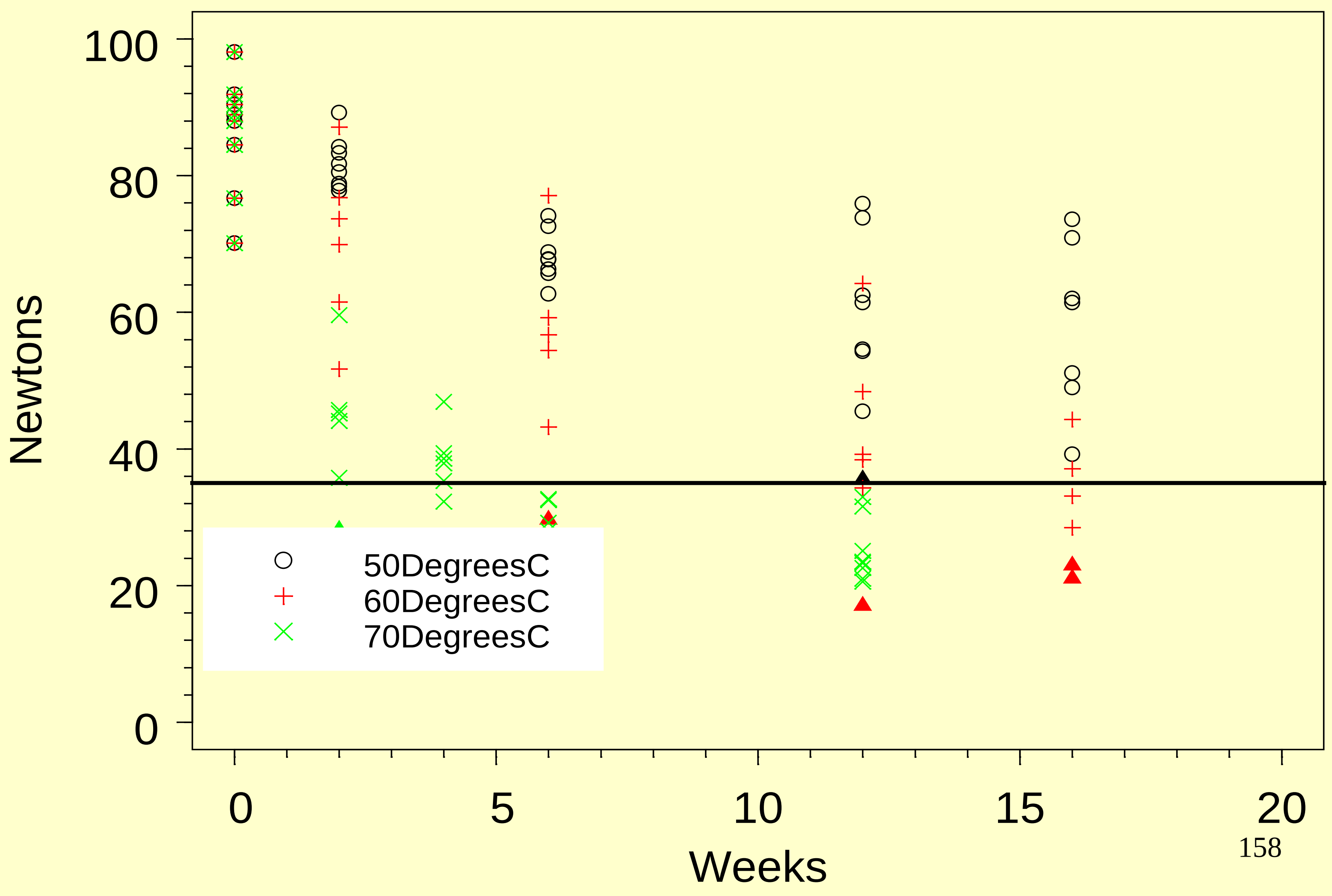
Lessons Learned

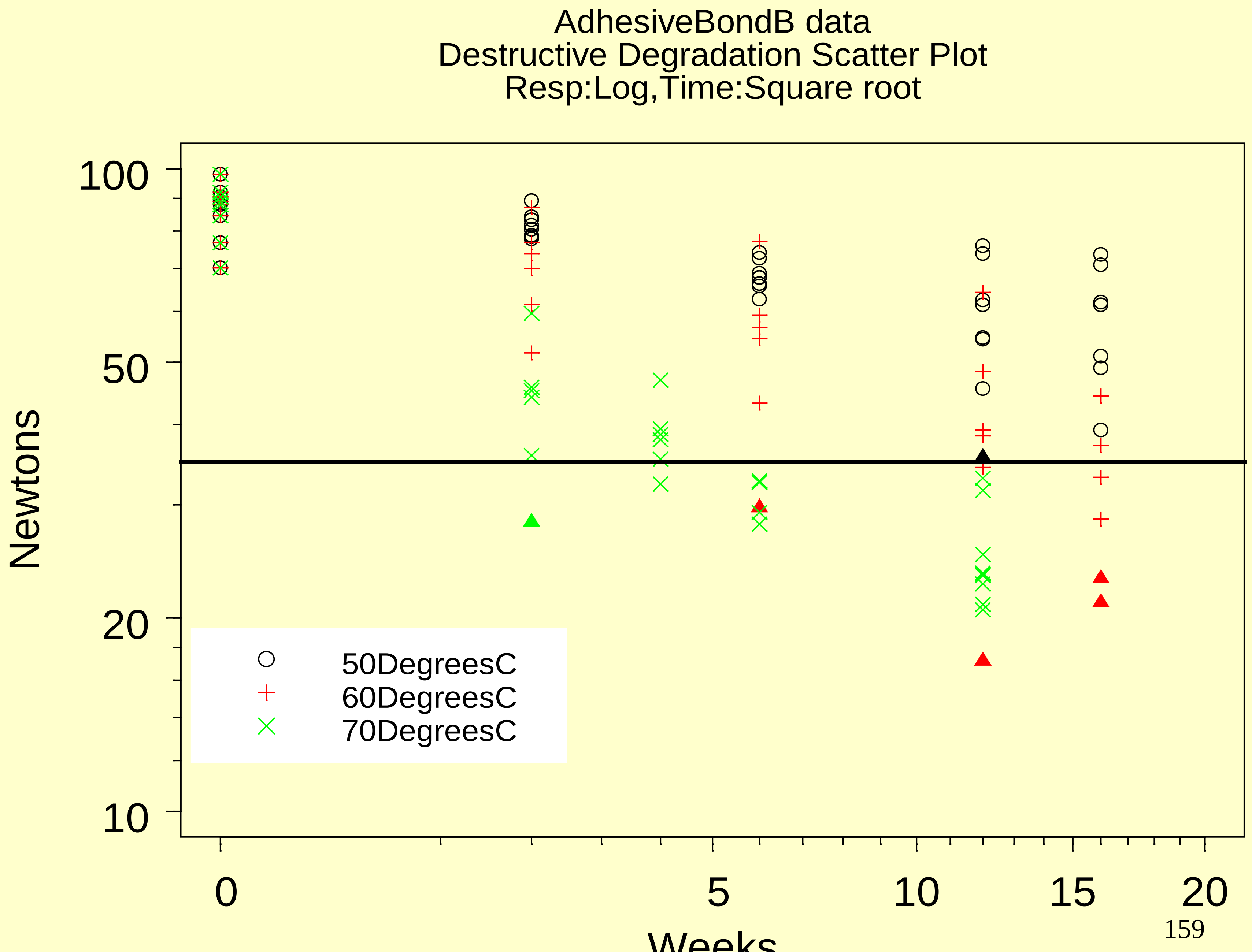
- Repeated measures degradation data can provide much more information that failure-time data
- Early part of degradation path may be complicated, but can be ignored.
- Testing at stress levels that are too high can cause new failure modes
- Modeling and extrapolation in two dimensions (two accelerating variables) can be more complicated and more risky

Accelerated Destructive Degradation Test to Evaluate the Lifetime of an Adhesive Bond

- Some baseline units tested without aging
- Samples of units placed into chambers operating at 50, 60, and 70 degrees C
- Some units removed at specific times and destructively evaluated for strength
- Units with adhesive strength below 35 Newtons are considered to be failures

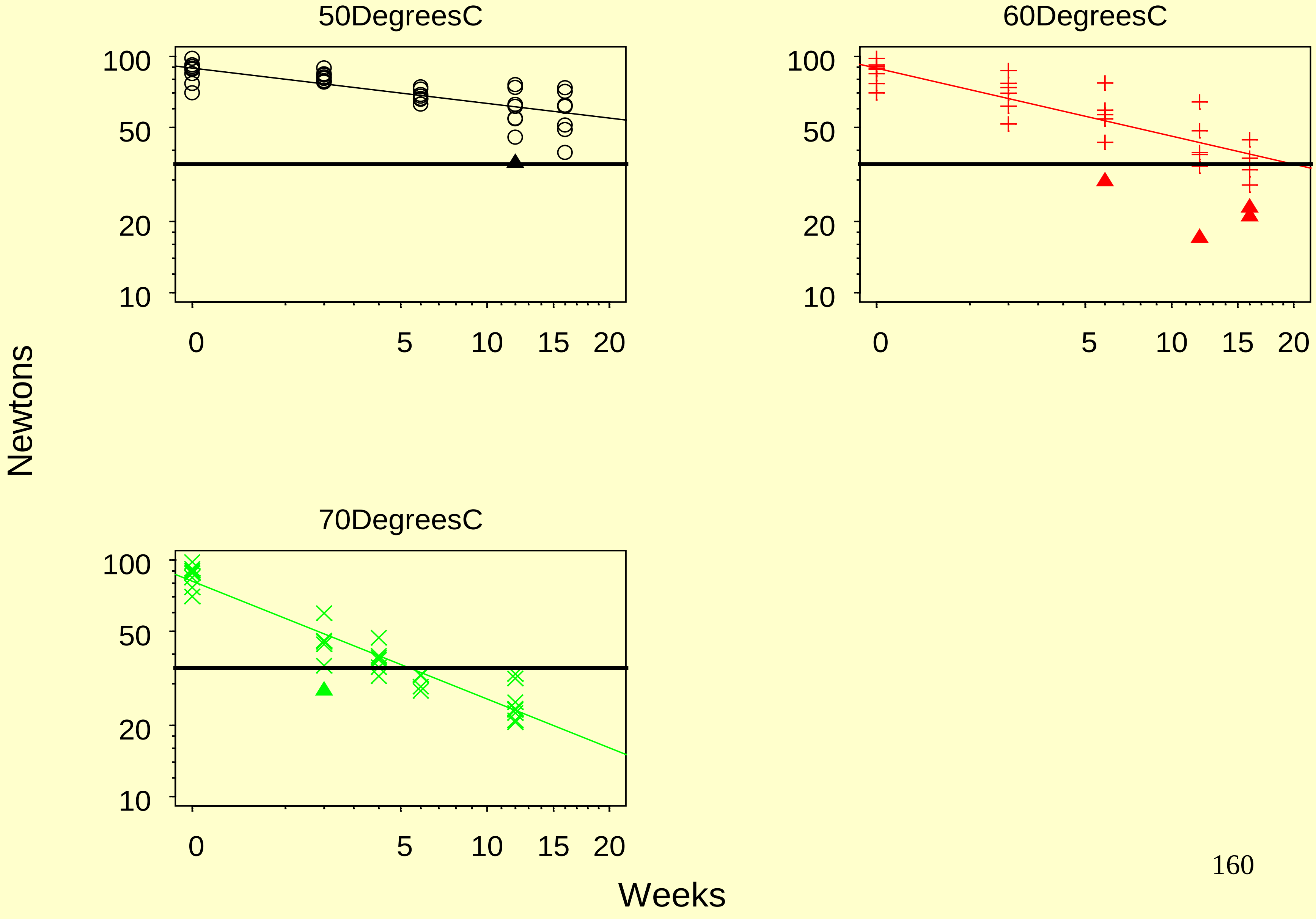
AdhesiveBondB data
Destructive Degradation Scatter Plot
Resp:Linear,Time:Linear

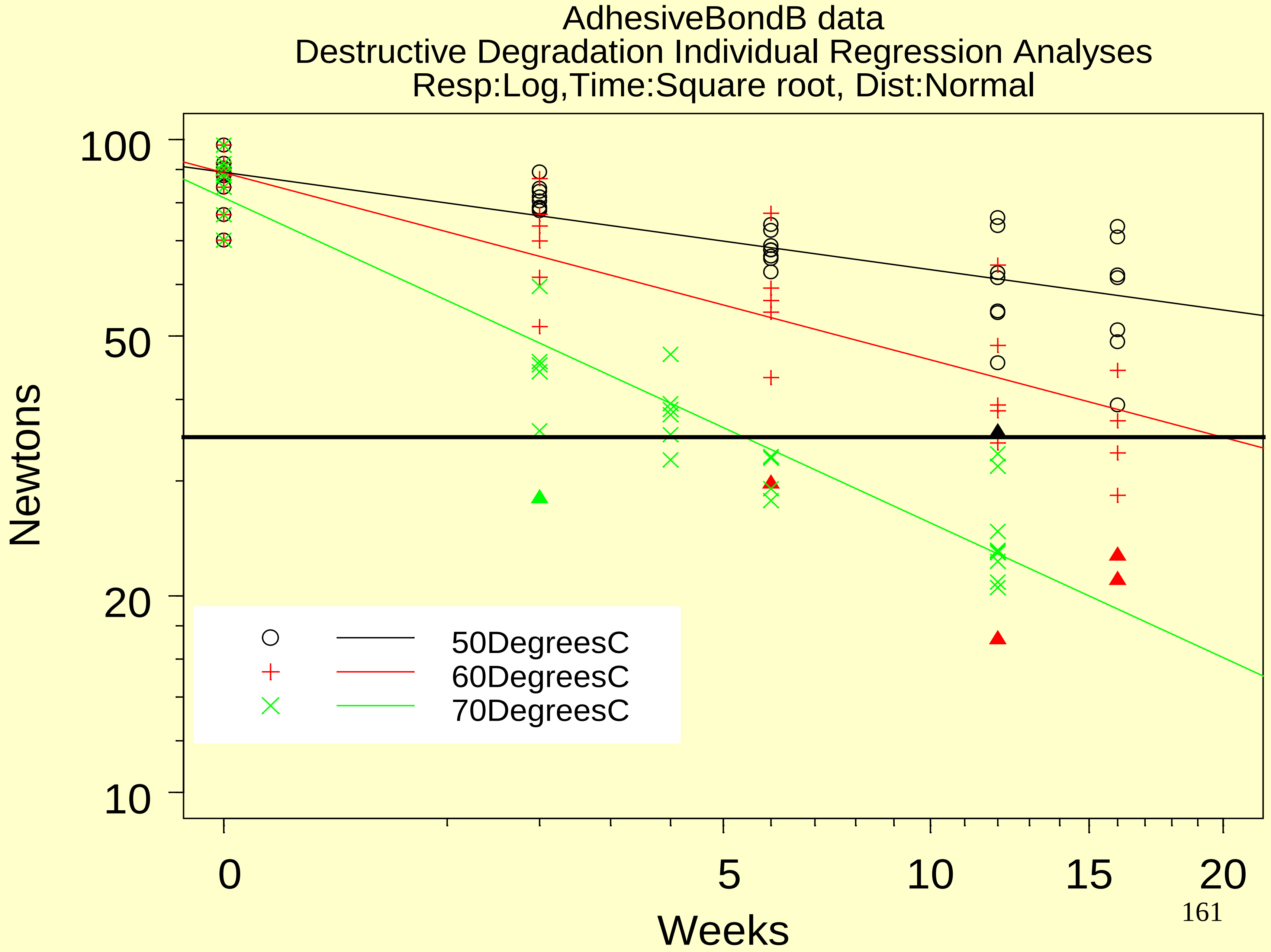


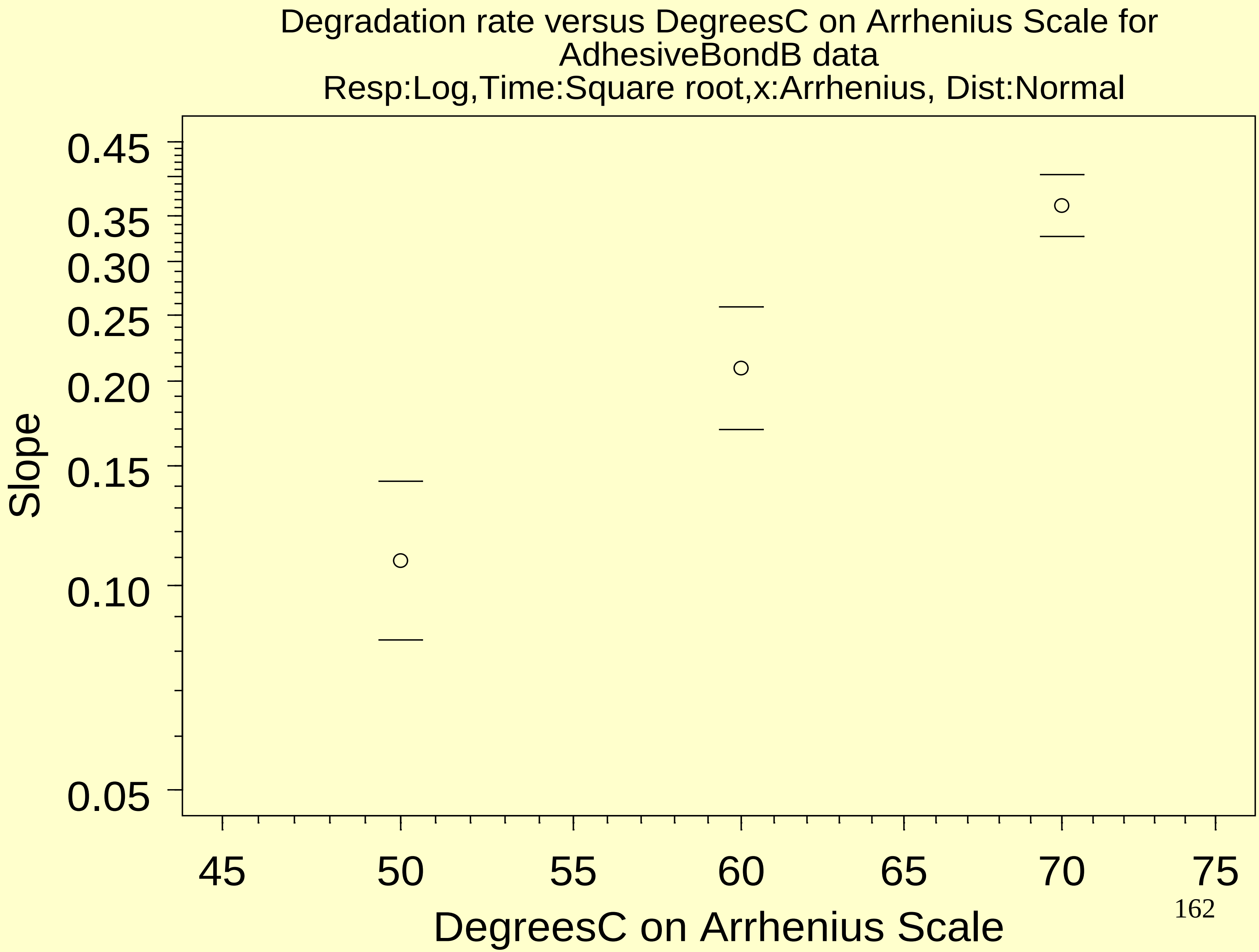


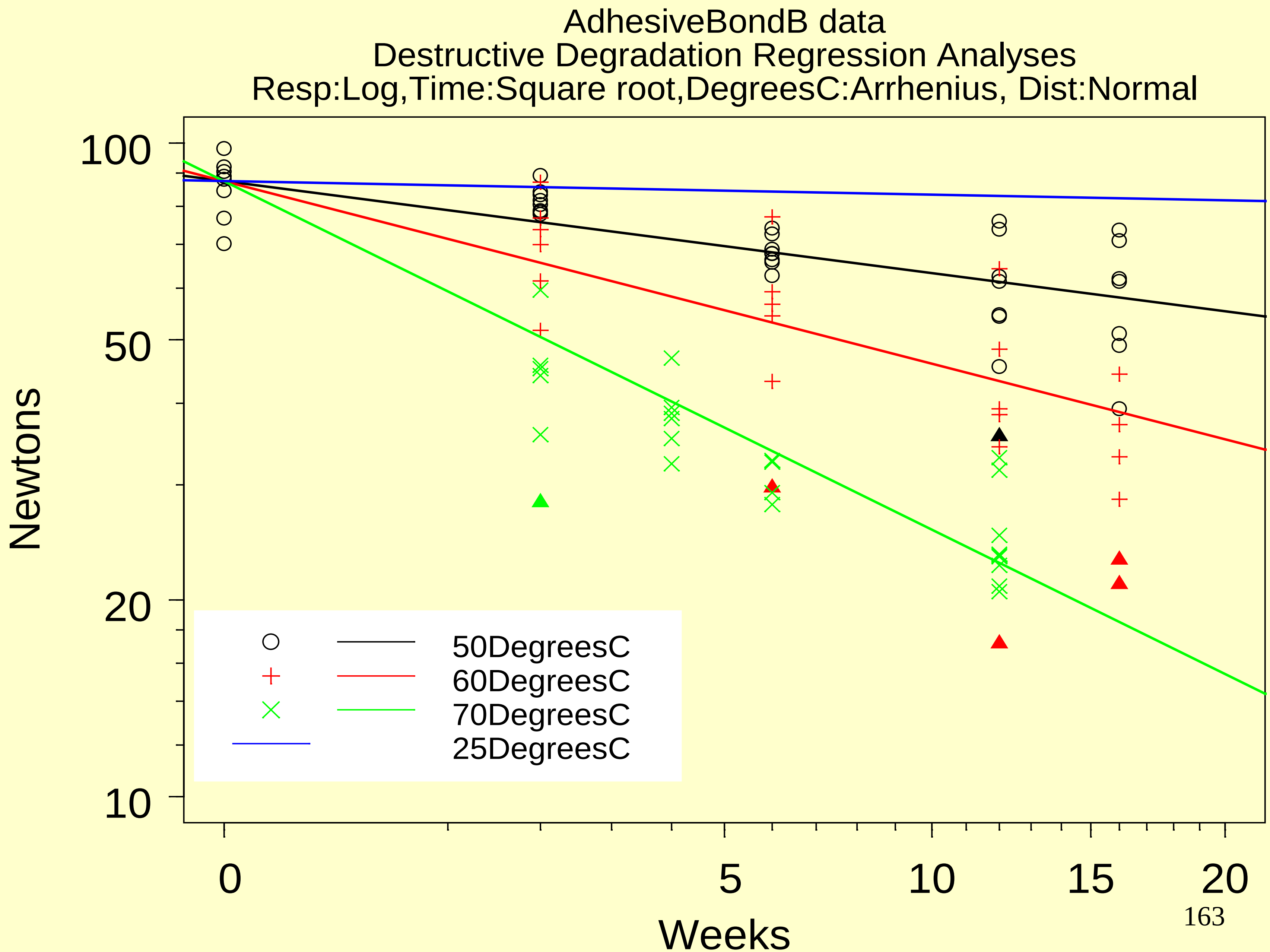
AdhesiveBondB data
Destructive Degradation Scatter Plot
Resp:Log,Time:Square root
Newtons
100
50
20
10
50DegreesC
60DegreesC
70DegreesC
0
5
10
15
20
Weeks

AdhesiveBondB data
Resp:Log,Time:Square root, Dist:Normal

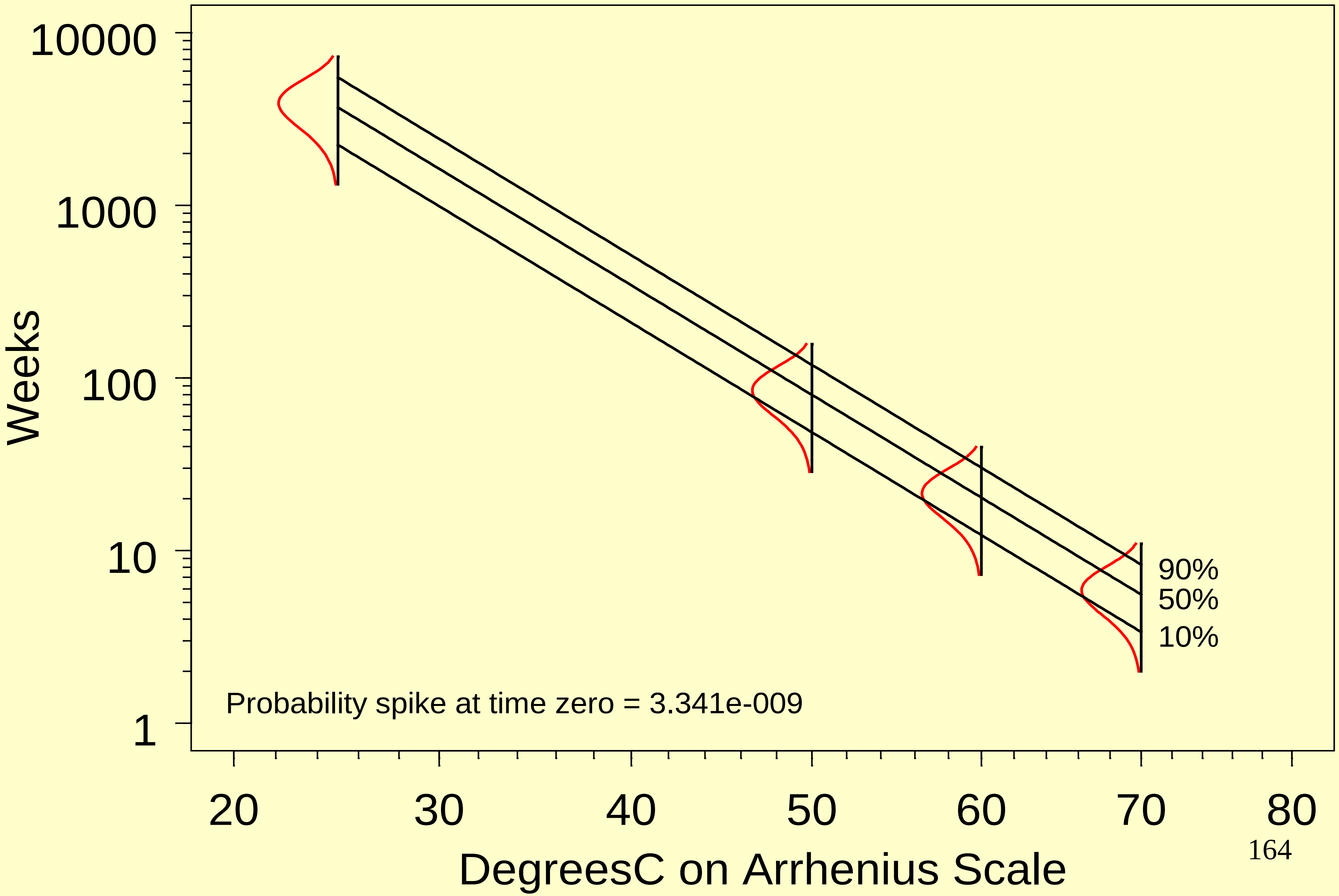


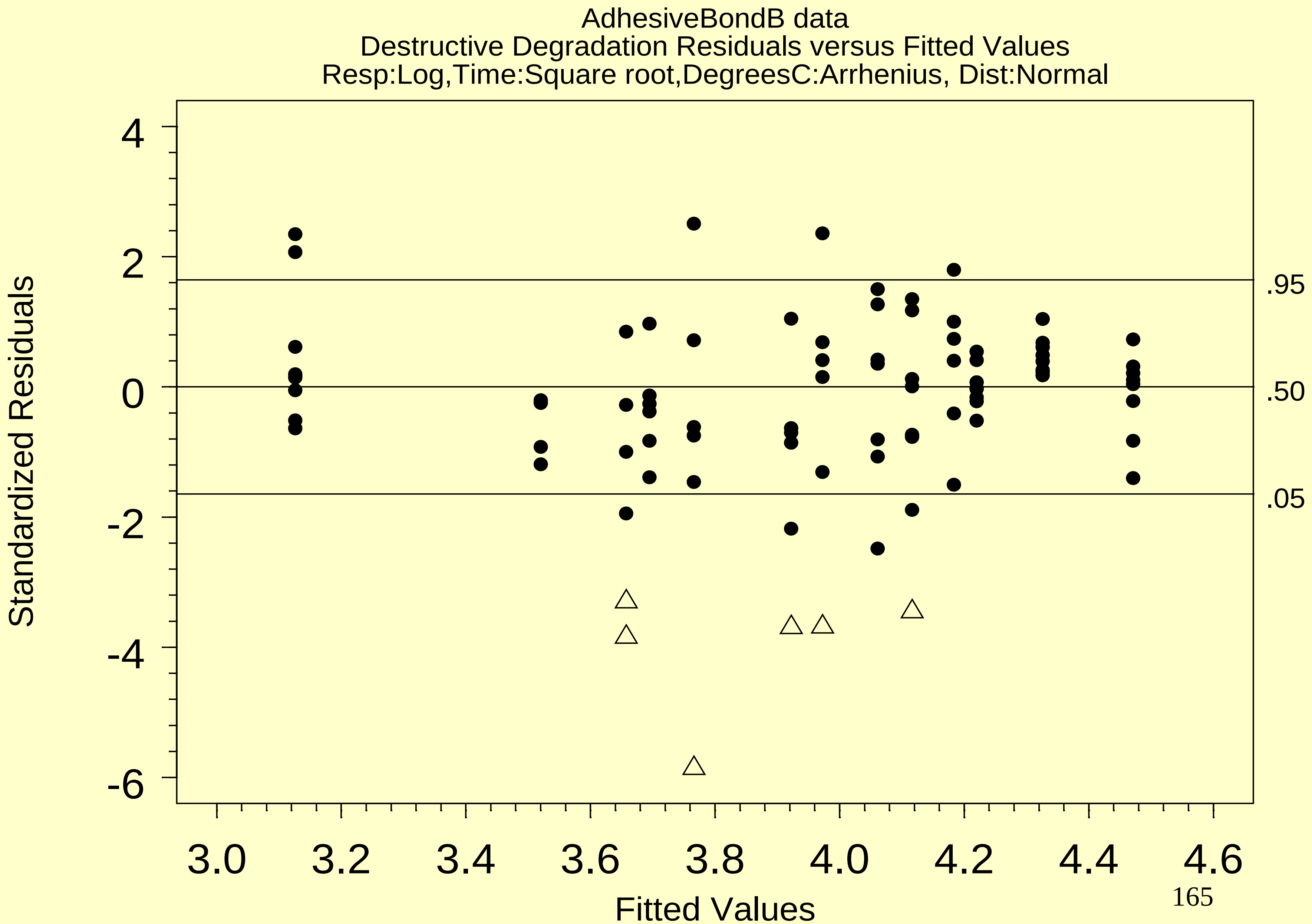

AdhesiveBondB data
Destructive Degradation Individual Regression Analyses
Resp:Log,Time:Square root, Dist:Normal
Newtons
100
50
20
10
0
5
10
15
20
Weeks
50DegreesC
60DegreesC
70DegreesC


Degradation rate versus DegreesC on Arrhenius Scale for
AdhesiveBondB data
Resp:Log,Time:Square root,x:Arrhenius, Dist:Normal
Slope
0.45
0.35
0.30
0.25
0.20
0.15
0.10
0.05
45
50
55
60
65
70
75
DegreesC on Arrhenius Scale


AdhesiveBondB data
Destructive Degradation Regression Analyses
Resp:Log,Time:Square root,DegreesC:Arrhenius, Dist:Normal
Newtons
100
50
20
10
Weeks
0
5
10
15
20
50DegreesC
60DegreesC
70DegreesC
25DegreesC

Model plot for AdhesiveBondB data
Resp:Log,Time:Square root,DegreesC:Arrhenius, Dist:Normal
Failure-time distribution for degradation failure level of 35 Newtons




AdhesiveBondB data
Destructive Degradation Residuals versus Fitted Values
Resp:Log,Time:Square root,DegreesC:Arrhenius, Dist:Normal
Standardized Residuals
4
2
0
-2
-4
-6
.95
.50
.05
3.0
3.2
3.4
3.6
3.8
4.0
4.2
4.4
4.6
Fitted Values

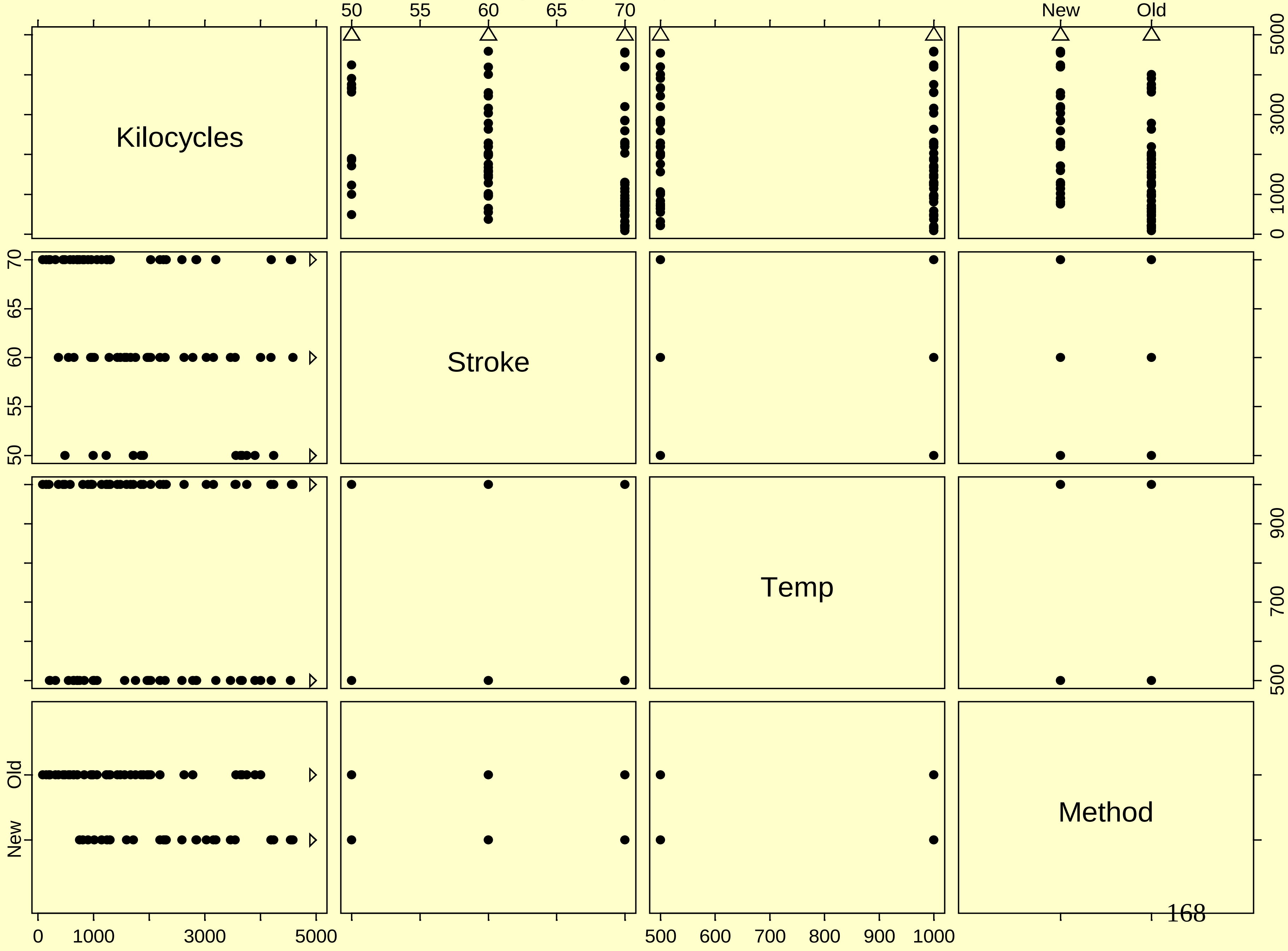
Lessons Learned

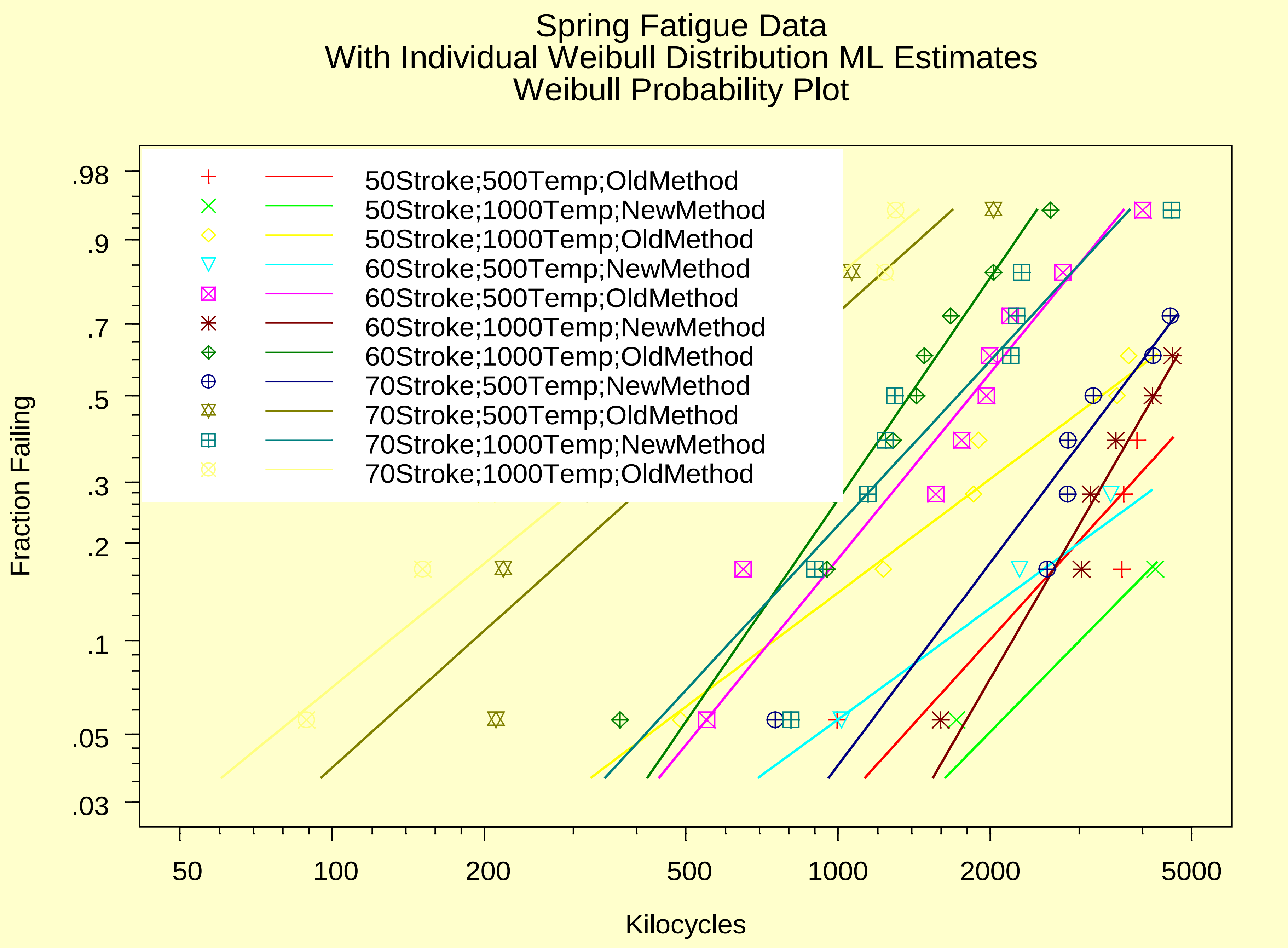
- Destructive degradation data are also useful, but more units need to be tested
- Transformations of variables can sometimes simplify model assumptions
- One need to understand the reason for anomalous observations before acting on them

Experiment to Estimate the Fatigue Life of a New Spring

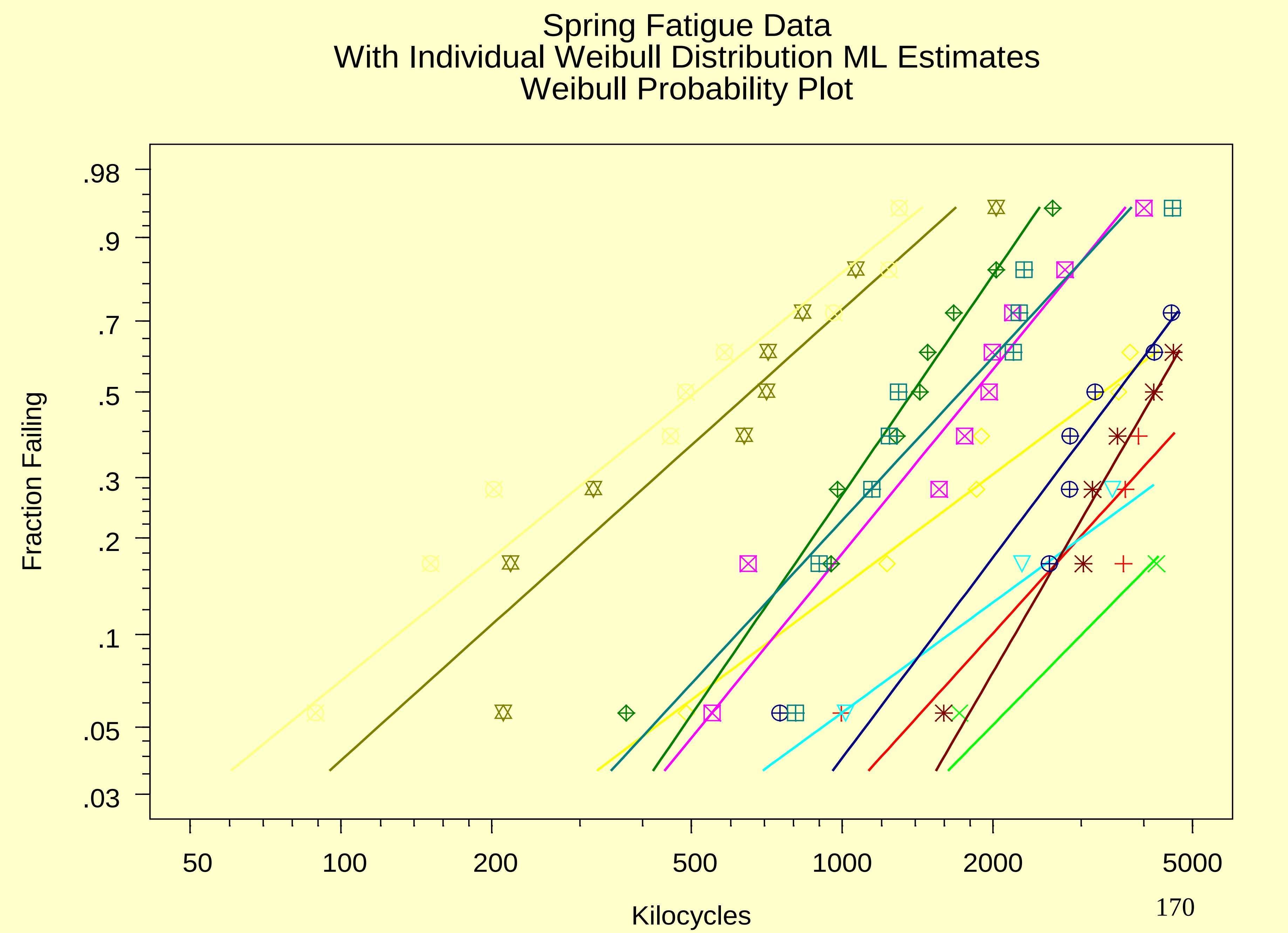
- Data from Meeker, Escobar, and Zayac (2003)
- Large factorial accelerated test experiment. Three factors, 12 combinations of levels, 9 reps at each combination.
- 108 springs tested until 5000 k-cycles
- Goals:
  - Compare New and Old processing methods
  - B01 life > 5000 k-cycles?

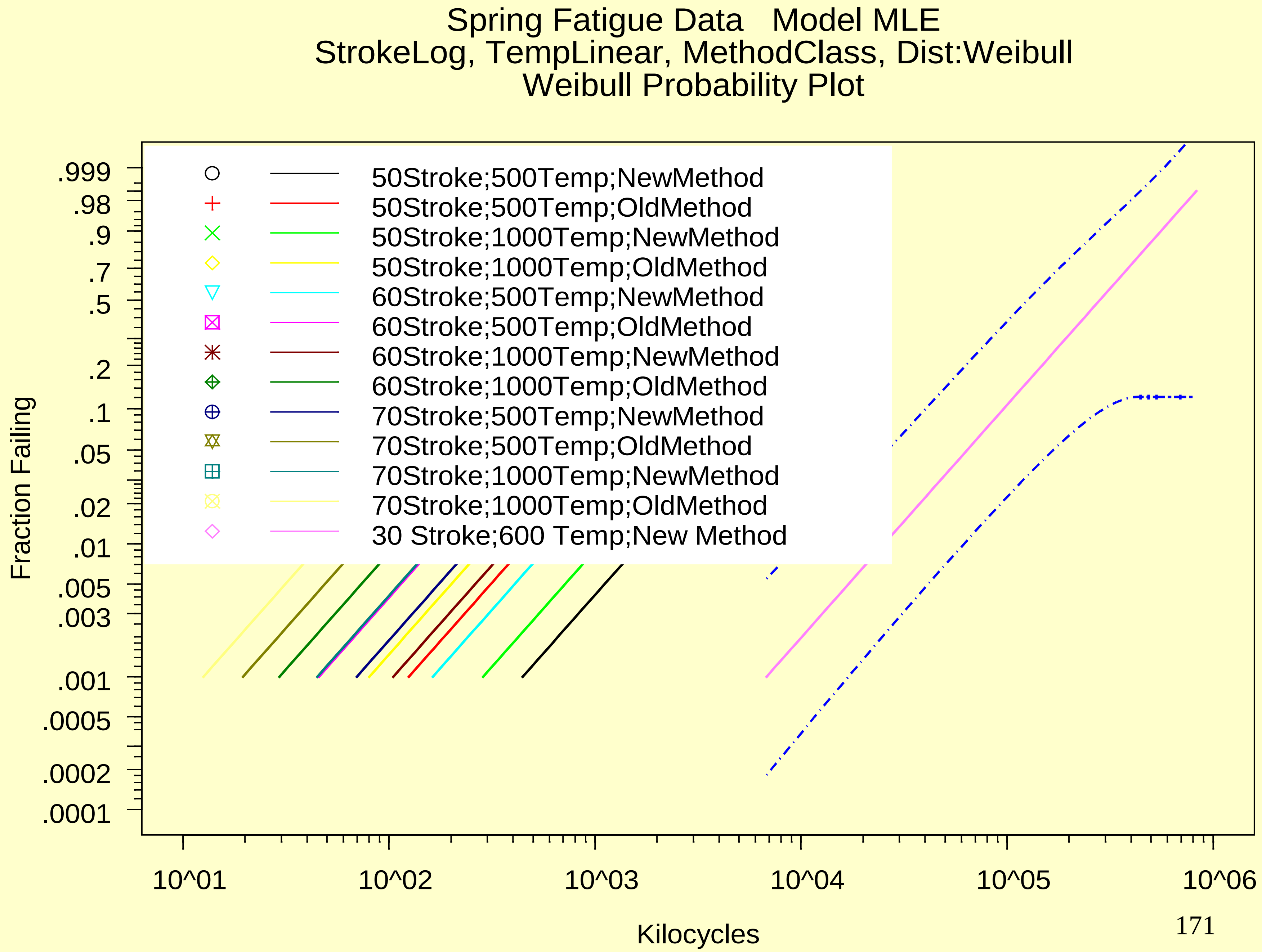
Spring Fatigue Data

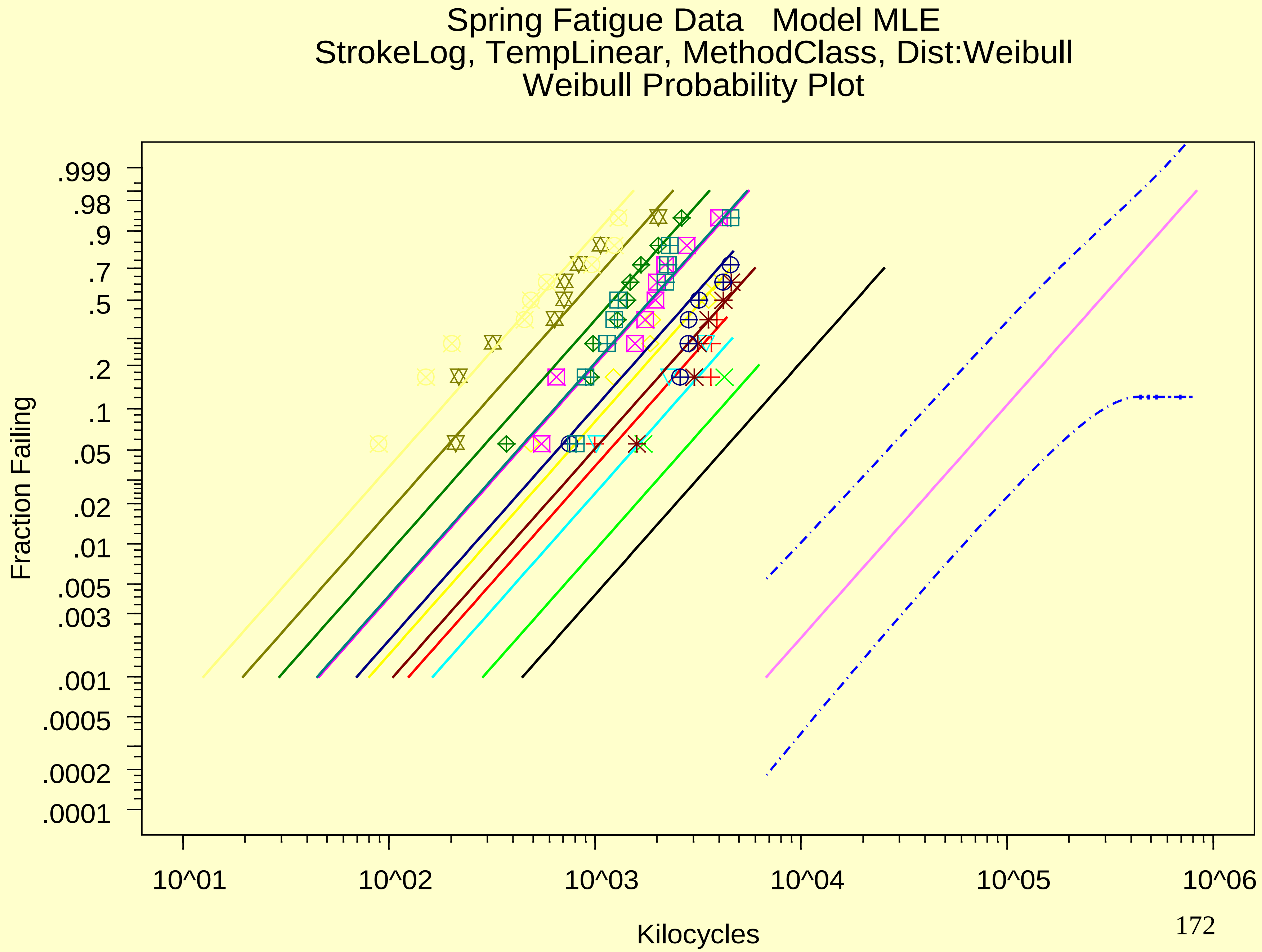




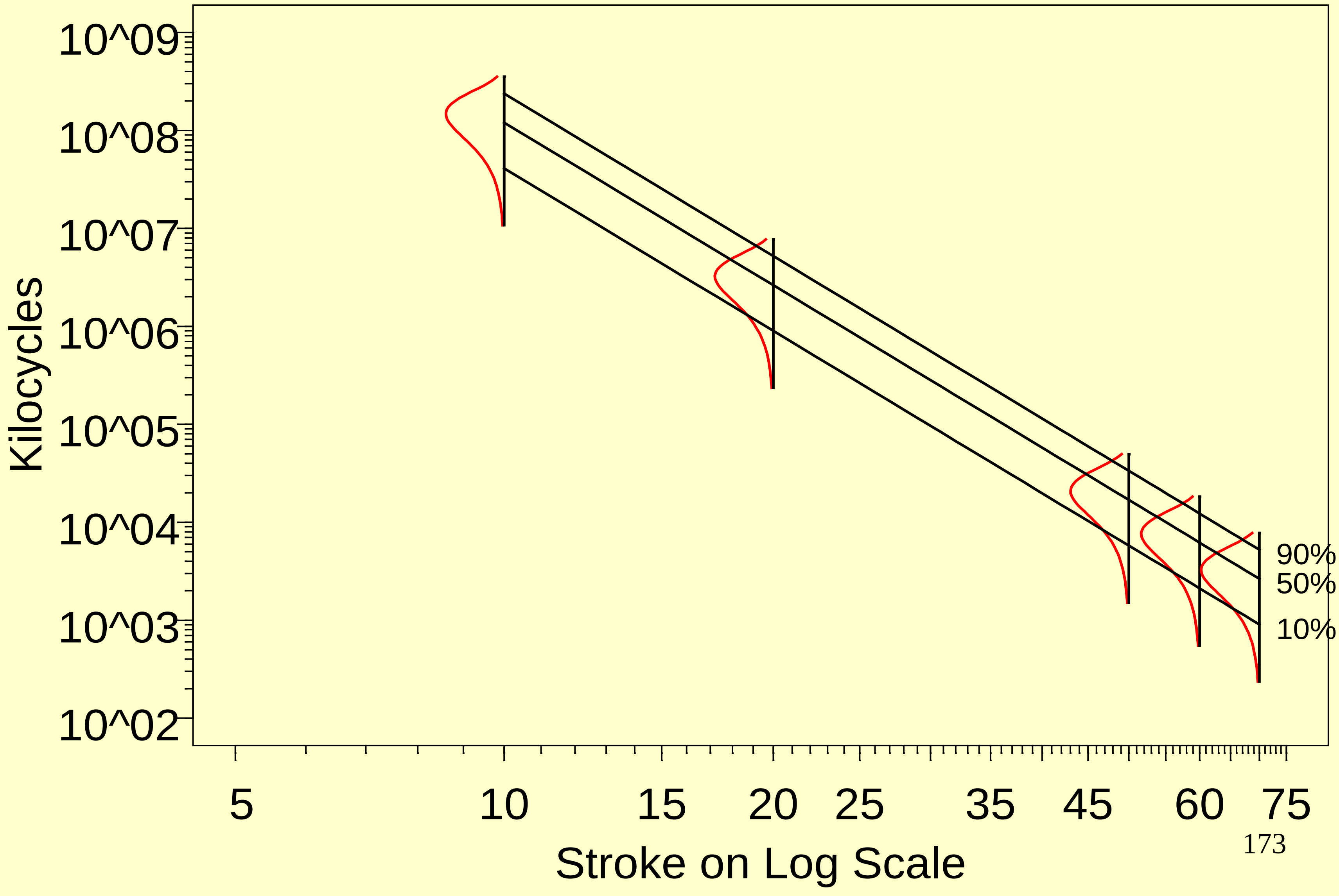
Spring Fatigue Data
With Individual Weibull Distribution ML Estimates
Weibull Probability Plot
50Stroke;500Temp;OldMethod
50Stroke;1000Temp;NewMethod
50Stroke;1000Temp;OldMethod
60Stroke;500Temp;NewMethod
60Stroke;500Temp;OldMethod
60Stroke;1000Temp;NewMethod
60Stroke;1000Temp;OldMethod
70Stroke;500Temp;NewMethod
70Stroke;500Temp;OldMethod
70Stroke;1000Temp;NewMethod
70Stroke;1000Temp;OldMethod
Fraction Failing
.98
.9
.7
.5
.3
.2
.1
.05
.03
50
100
200
500
1000
2000
5000
Kilocycles

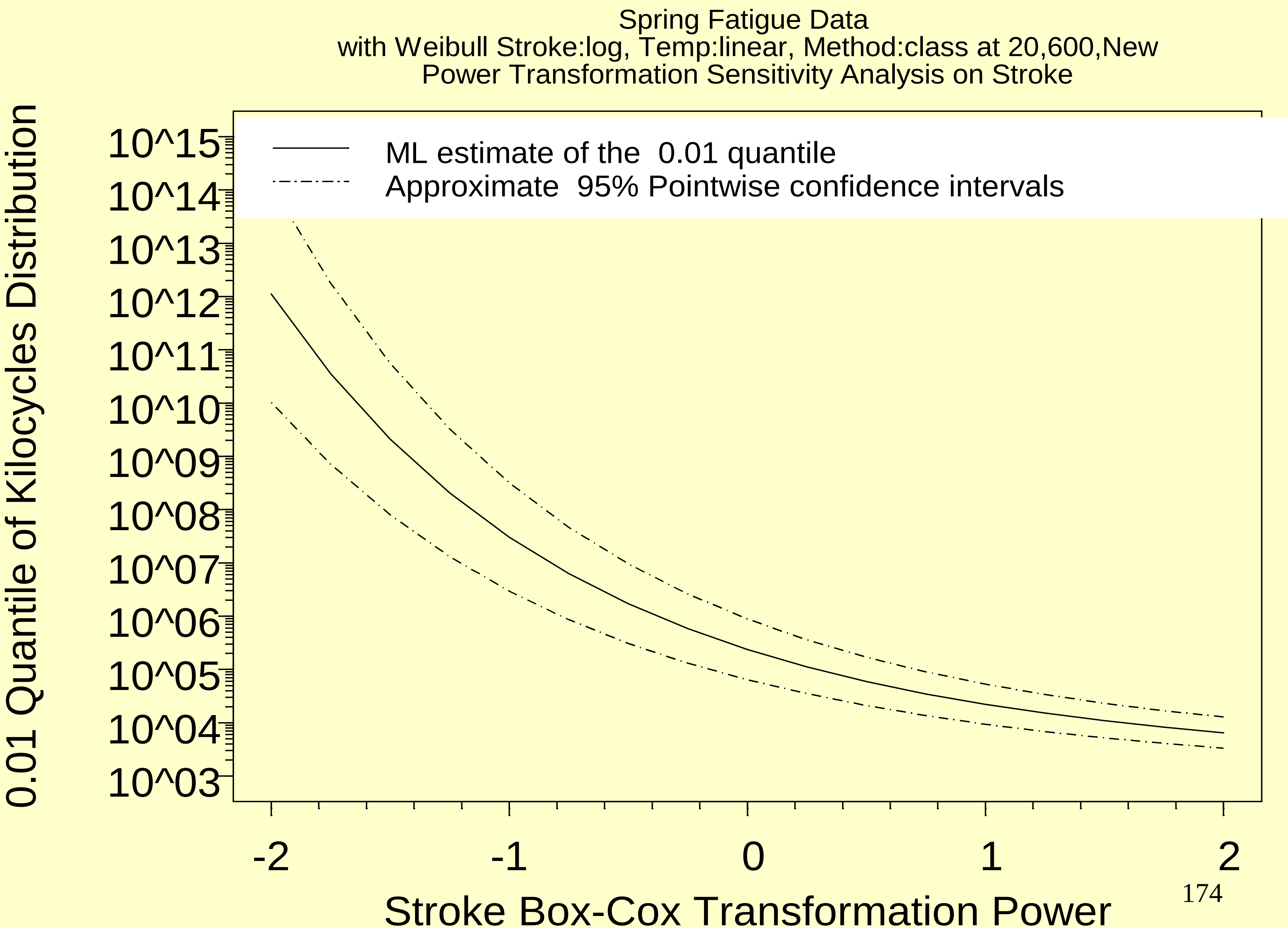

Spring Fatigue Data
With Individual Weibull Distribution ML Estimates
Weibull Probability Plot
Fraction Failing
.98
.9
.7
.5
.3
.2
.1
.05
.03
50
100
200
500
1000
2000
5000
Kilocycles


Spring Fatigue Data Model MLE
StrokeLog, TempLinear, MethodClass, Dist:Weibull
Weibull Probability Plot
50Stroke;500Temp;NewMethod
50Stroke;500Temp;OldMethod
50Stroke;1000Temp;NewMethod
50Stroke;1000Temp;OldMethod
60Stroke;500Temp;NewMethod
60Stroke;500Temp;OldMethod
60Stroke;1000Temp;NewMethod
60Stroke;1000Temp;OldMethod
70Stroke;500Temp;NewMethod
70Stroke;500Temp;OldMethod
70Stroke;1000Temp;NewMethod
70Stroke;1000Temp;OldMethod
30 Stroke;600 Temp;New Method
Fraction Failing
.999
.98
.9
.7
.5
.2
.1
.05
.02
.01
.005
.003
.001
.0005
.0002
.0001
10^01
10^02
10^03
10^04
10^05
10^06
Kilocycles

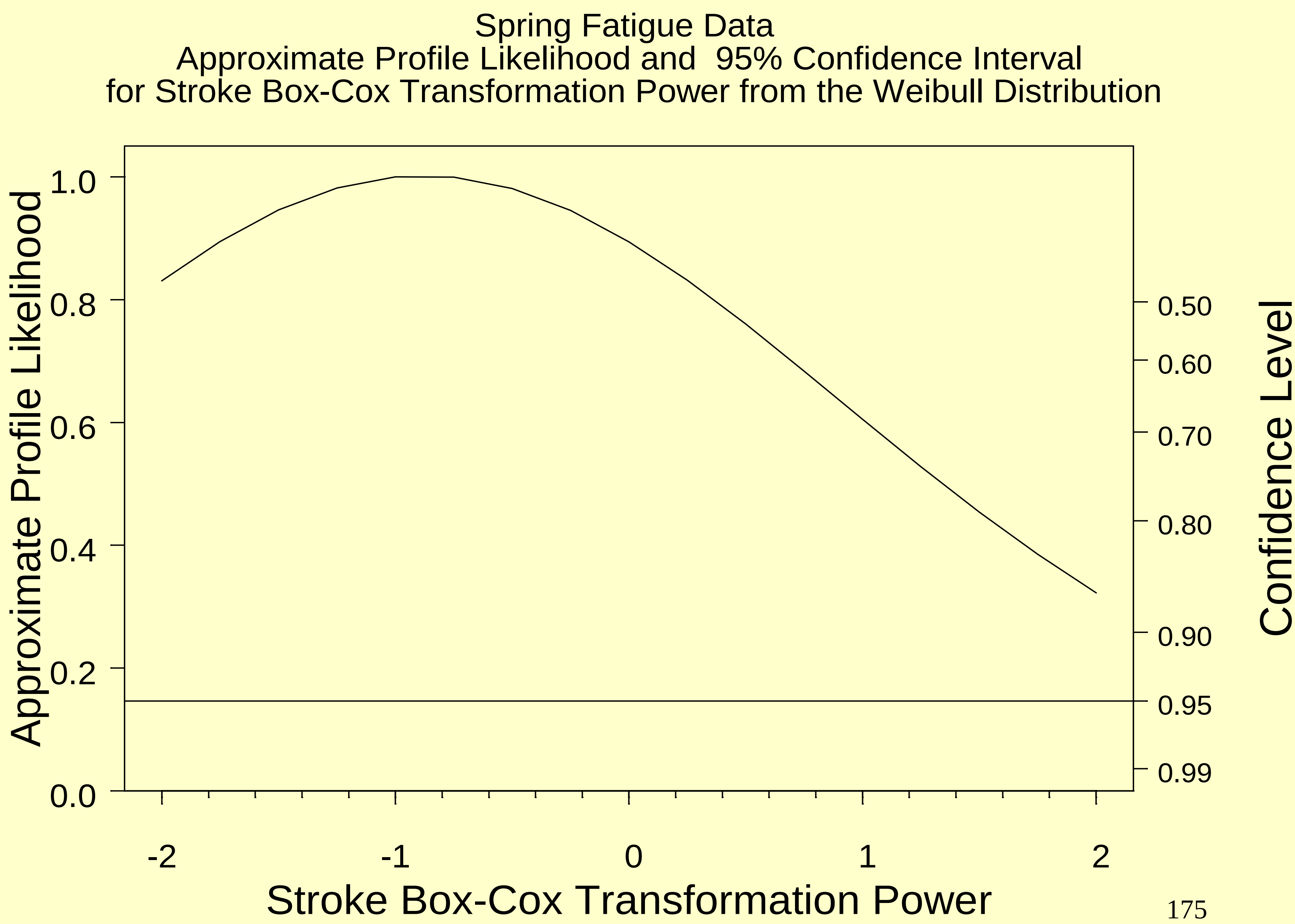

Spring Fatigue Data Model MLE
StrokeLog, TempLinear, MethodClass, Dist:Weibull
Weibull Probability Plot
Fraction Failing
.999
.98
.9
.7
.5
.2
.1
.05
.02
.01
.005
.003
.001
.0005
.0002
.0001
10^01
10^02
10^03
10^04
10^05
10^06
Kilocycles

Spring Fatigue Data StrokeLog, TempLinear, MethodClass, Dist:Weibull
Fixed values of Temp=600, Method=New





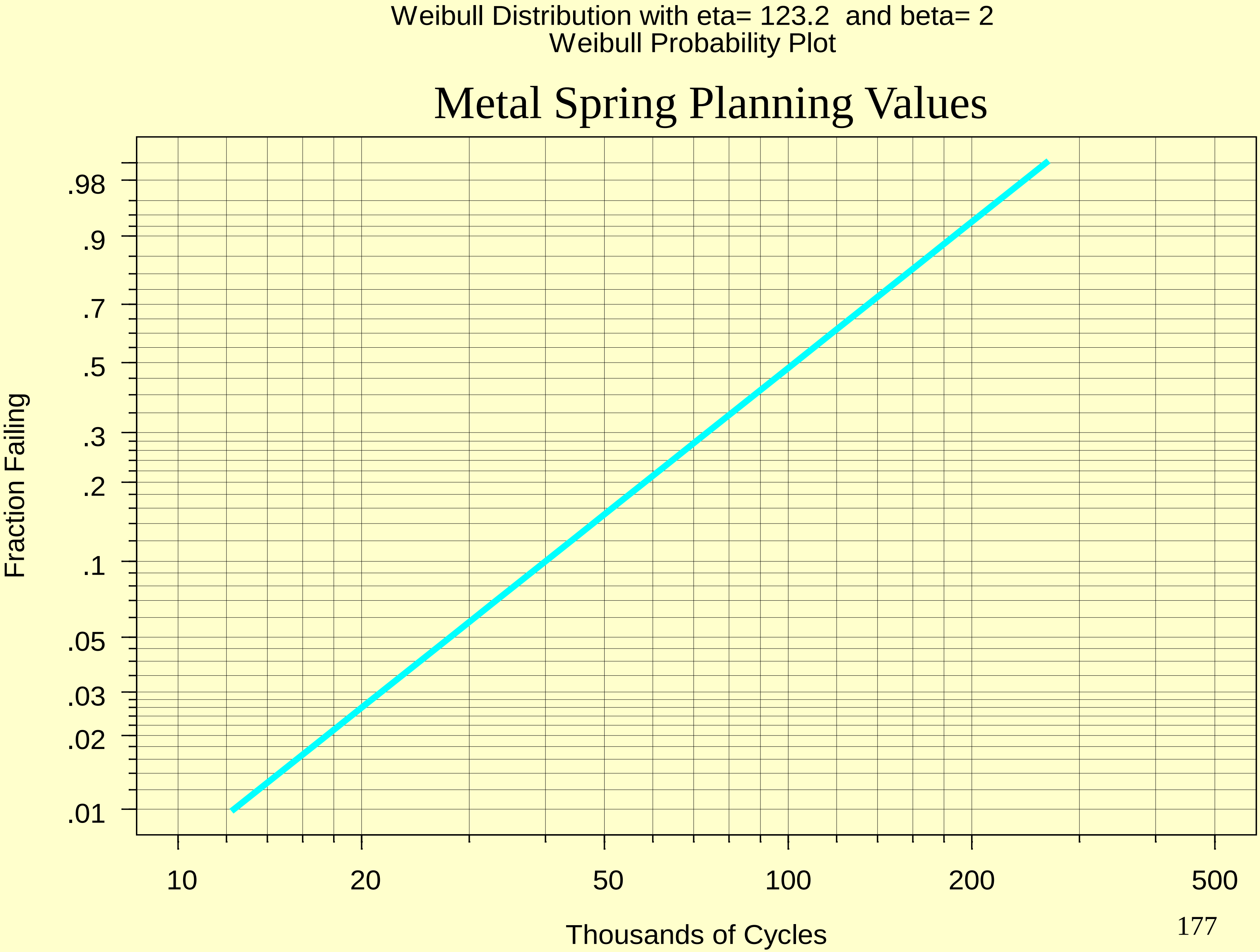
Spring Fatigue Data
with Weibull Stroke:log, Temp:linear, Method:class at 20,600,New
Power Transformation Sensitivity Analysis on Stroke
ML estimate of the 0.01 quantile
Approximate 95% Pointwise confidence intervals
0.01 Quantile of Kilocycles Distribution
10^15
10^14
10^13
10^12
10^11
10^10
10^09
10^08
10^07
10^06
10^05
10^04
10^03
-2
-1
0
1
2
Stroke Box-Cox Transformation Power


Spring Fatigue Data
Approximate Profile Likelihood and 95% Confidence Interval
for Stroke Box-Cox Transformation Power from the Weibull Distribution
Approximate Profile Likelihood
1.0
0.8
0.6
0.4
0.2
0.0
Confidence Level
0.50
0.60
0.70
0.80
0.90
0.95
0.99
-2
-1
0
1
2
Stroke Box-Cox Transformation Power

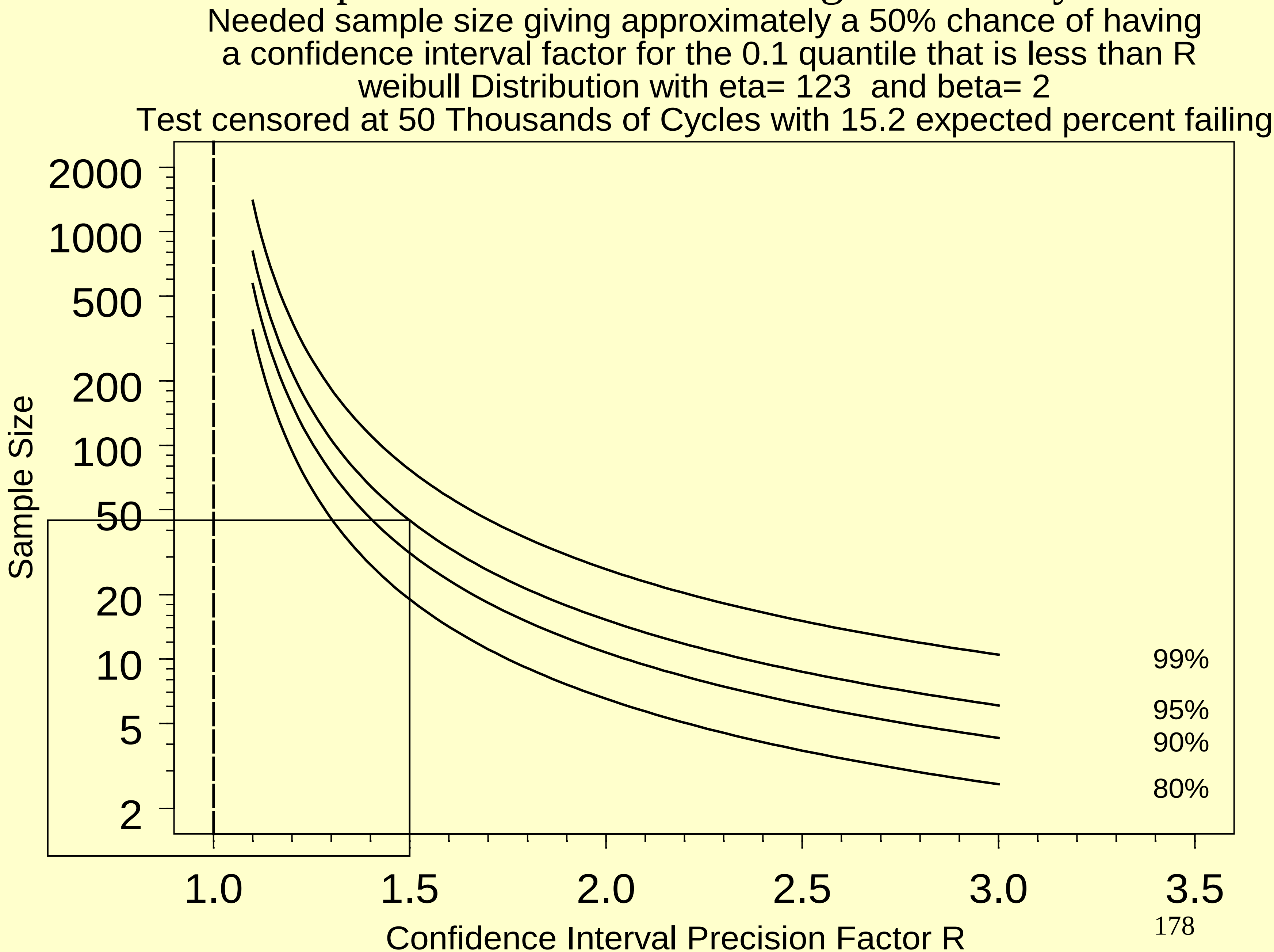
Planning a Life Test for a Metal Spring

- Sample size tool is a generalization of Stat 101 formula
- Simulation and graphics provide more insight
- Illustrated here for single distribution---can be applied to any other kind of model
- Goal: Decide how many units to test and how long to test them

Metal Spring Planning Values

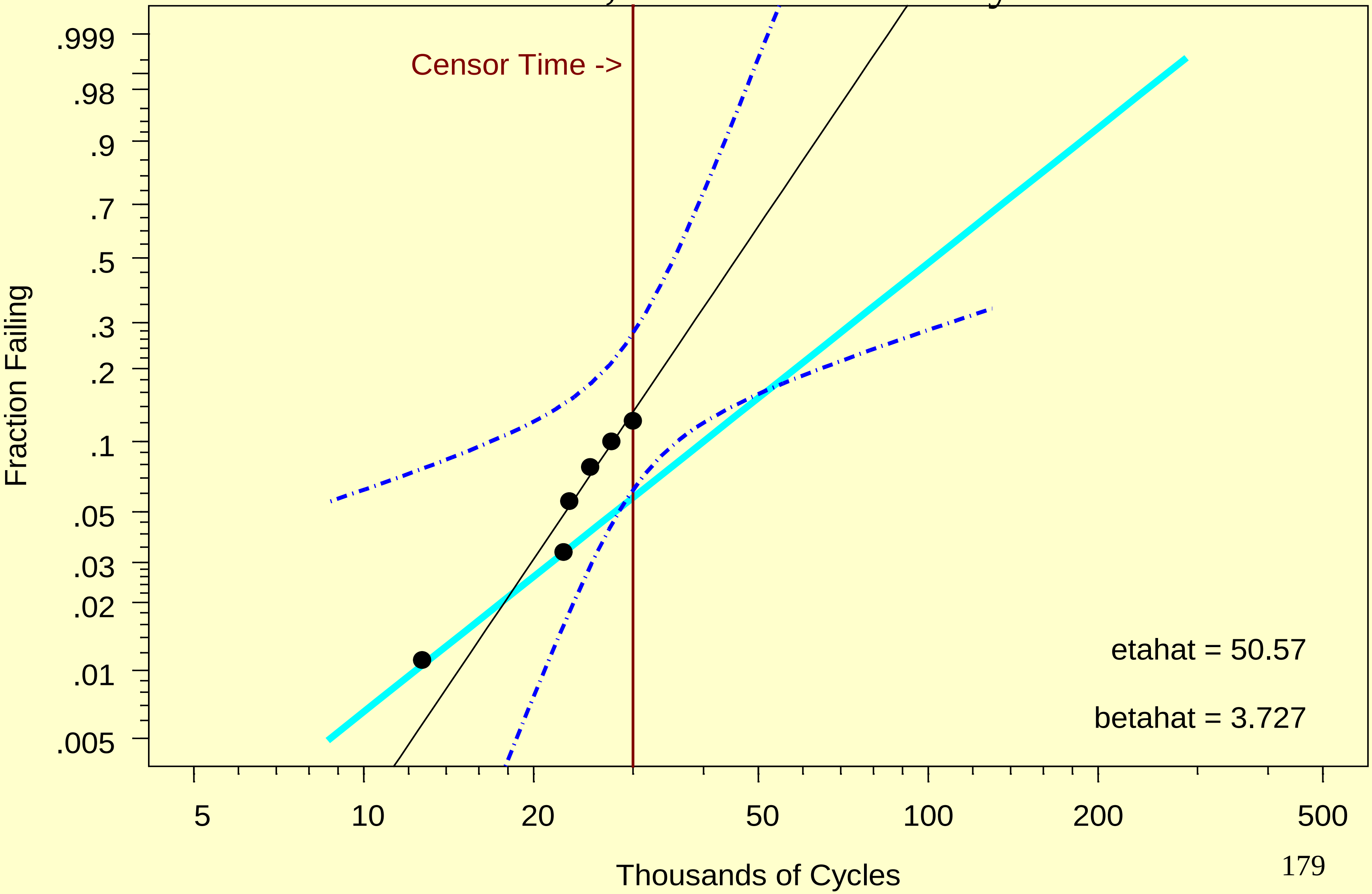


Sample Size Tool --- Censoring at 50 kilocycles



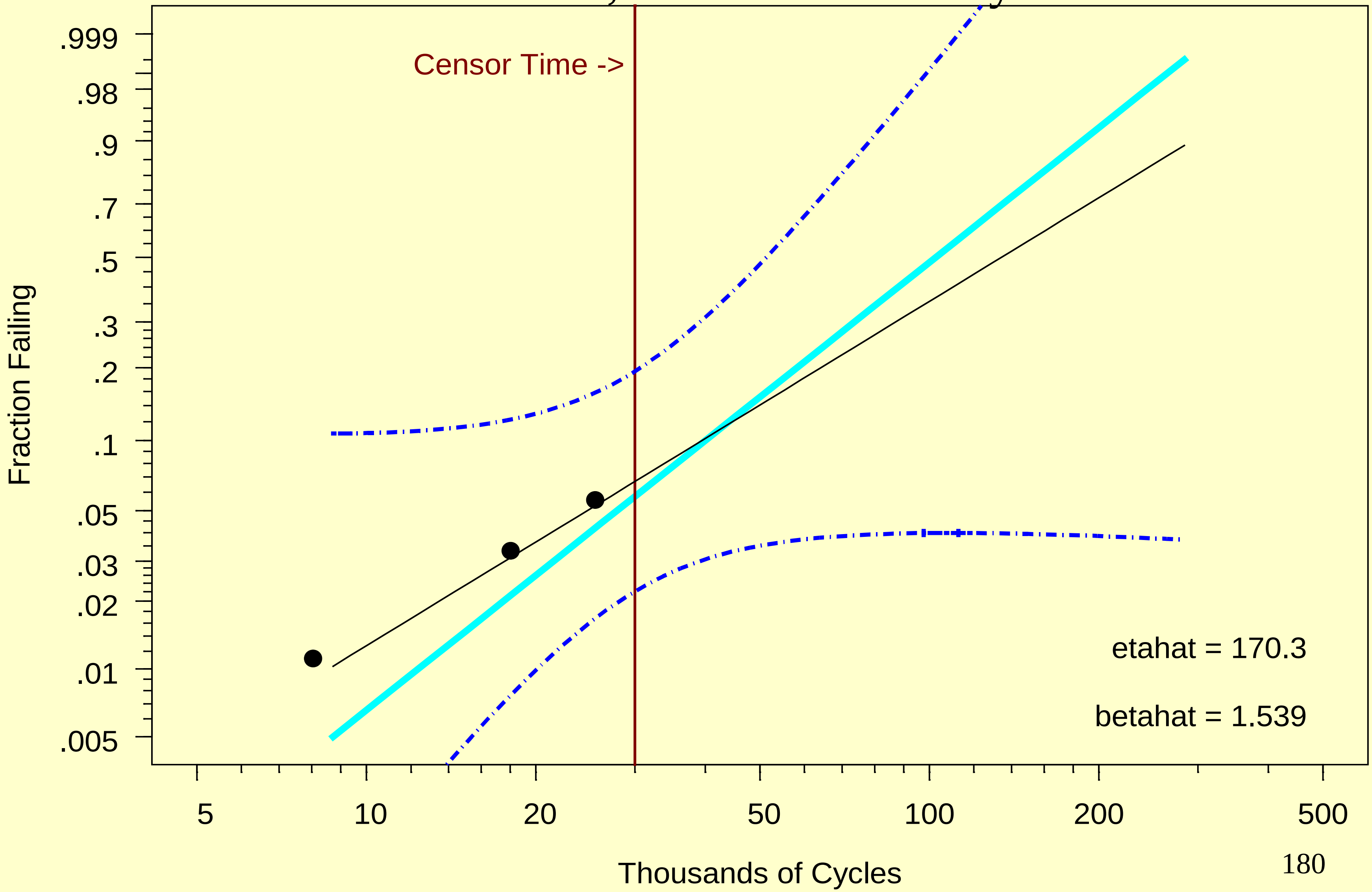
Simulated life test with sample size = 45
Weibull Distribution with eta= 123 and beta= 2
Weibull Probability Plot

n = 45, censor at 30 kilocycles

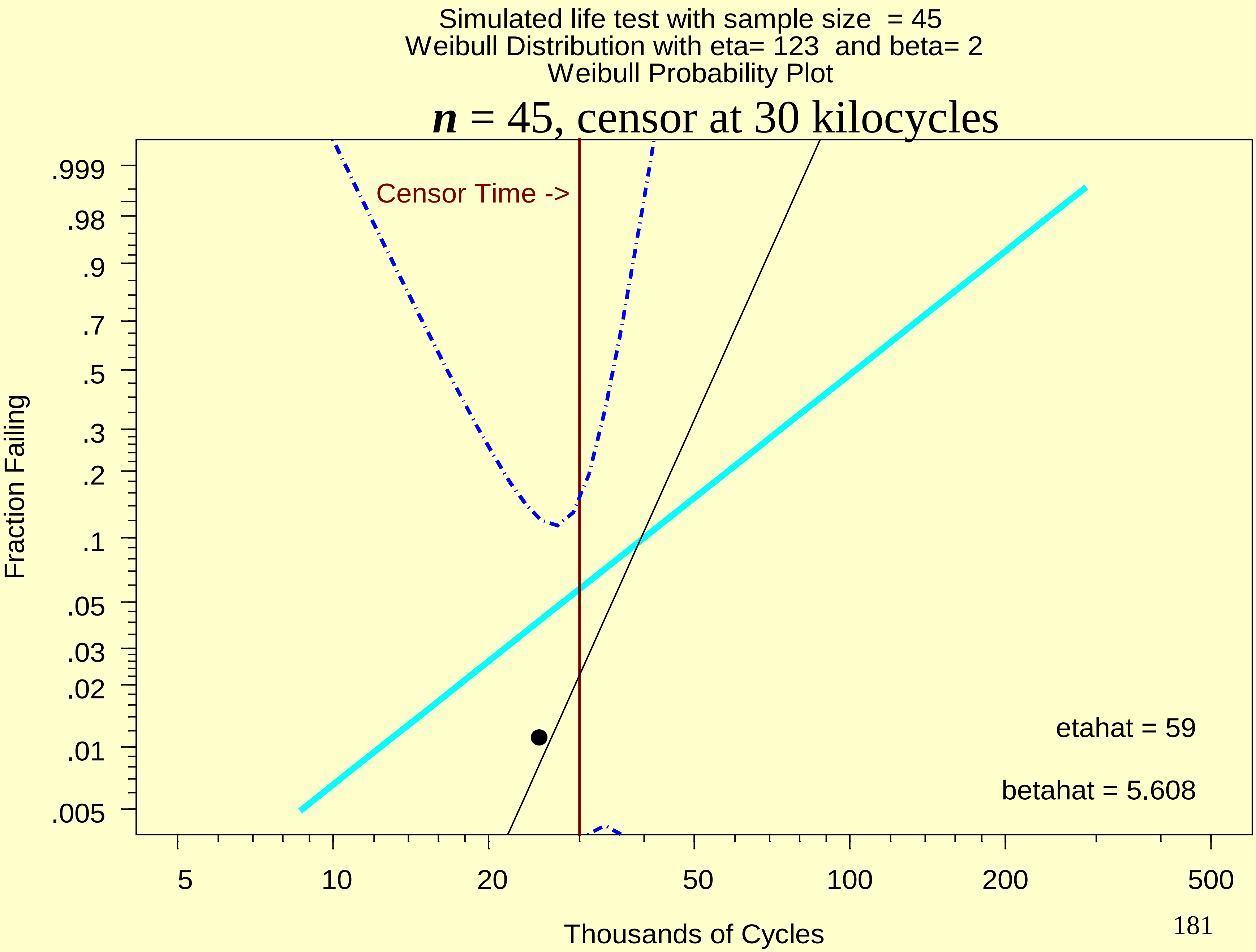


Simulated life test with sample size = 45
Weibull Distribution with eta= 123 and beta= 2
Weibull Probability Plot

n = 45, censor at 30 kilocycles

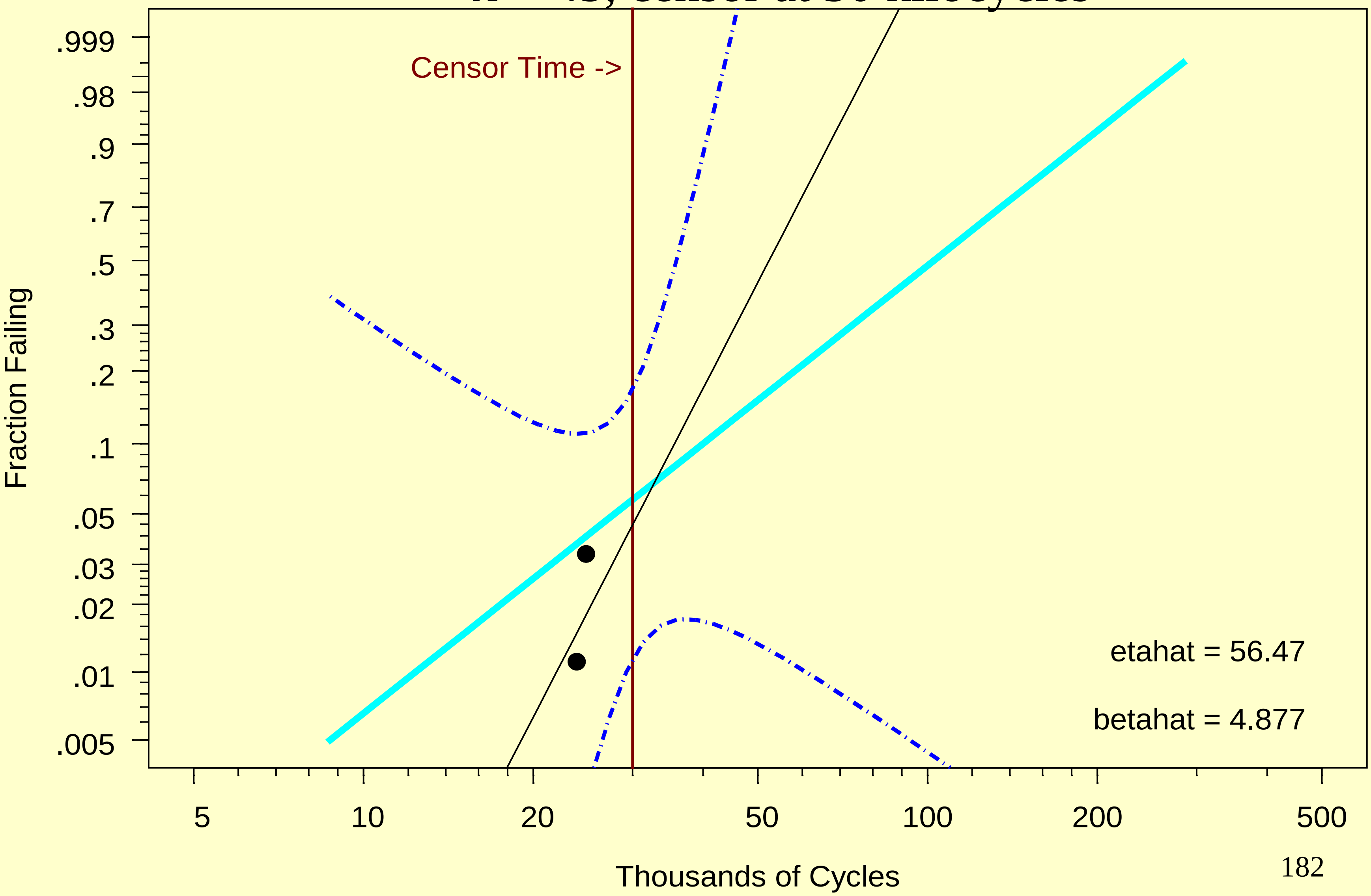


n = 45, censor at 30 kilocycles

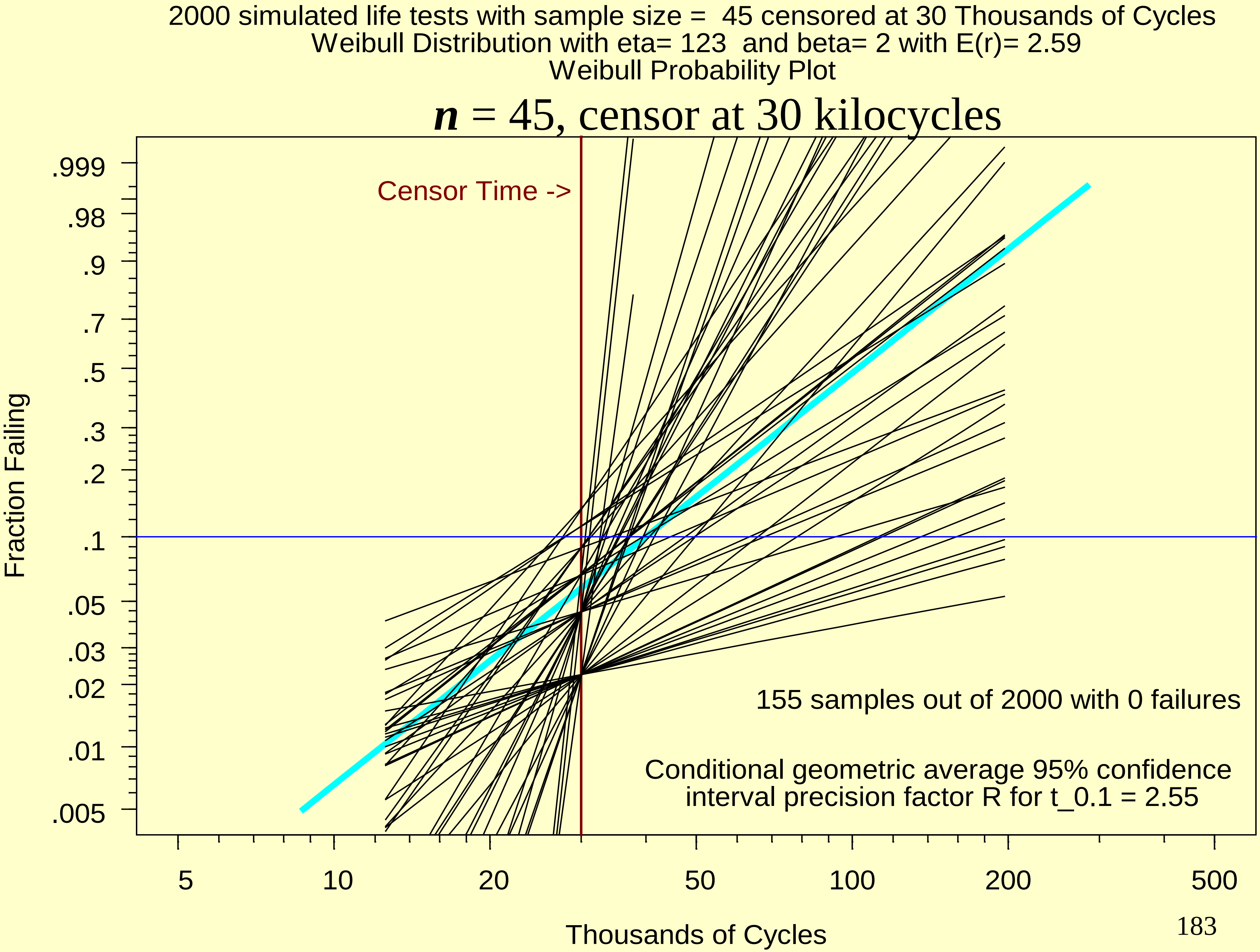


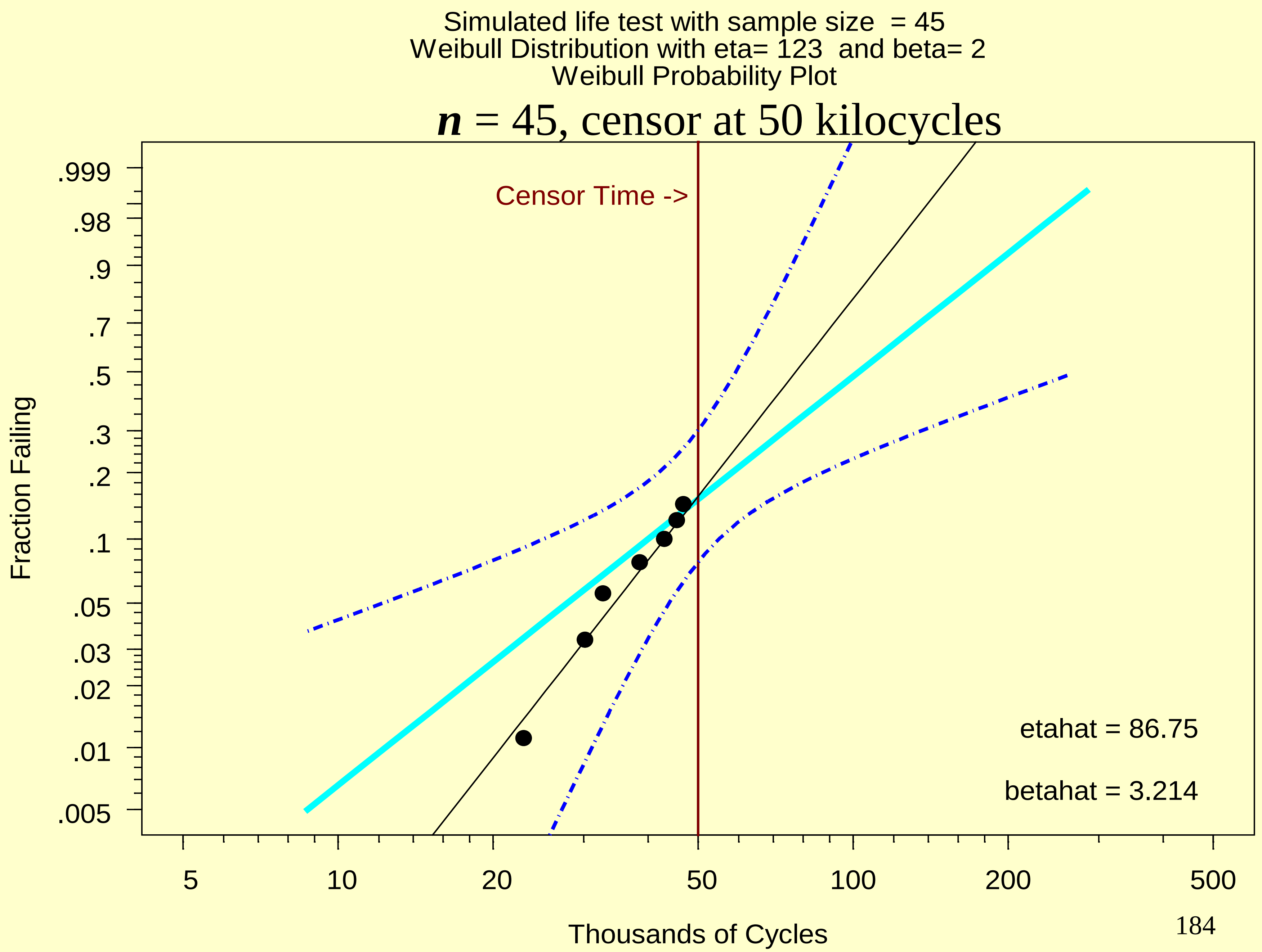
Simulated life test with sample size = 45
Weibull Distribution with eta= 123 and beta= 2
Weibull Probability Plot

n = 45, censor at 30 kilocycles

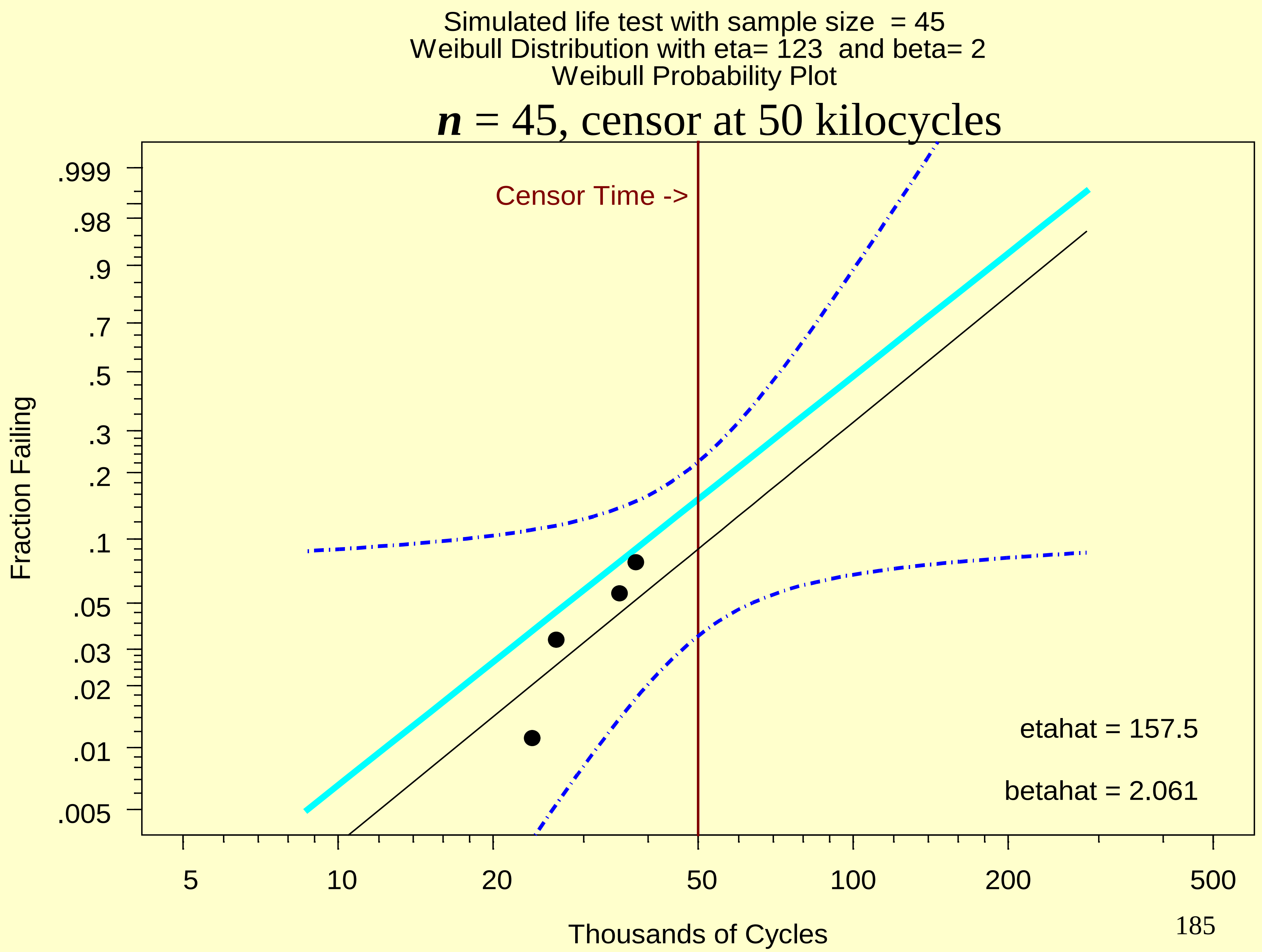


n = 45, censor at 30 kilocycles

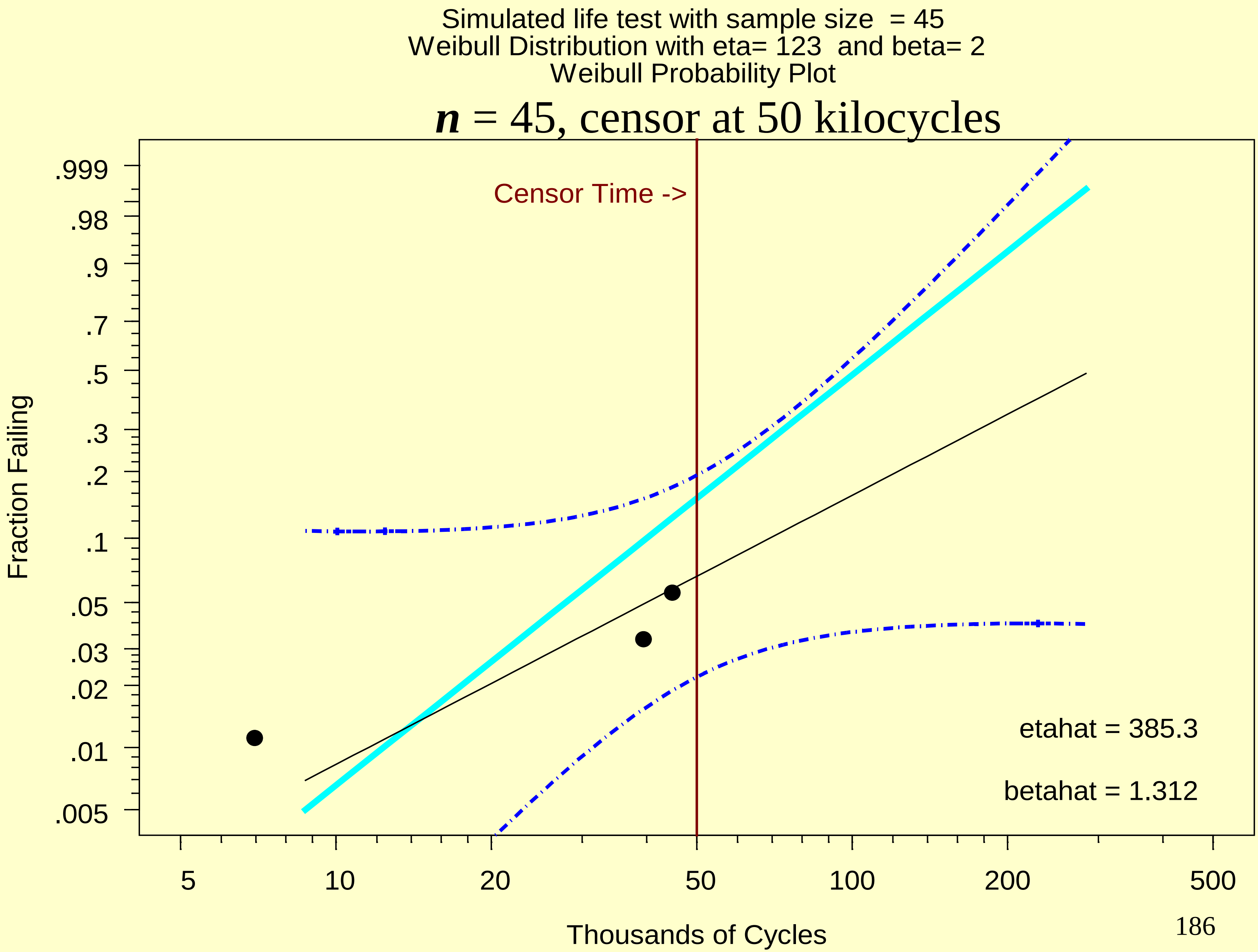




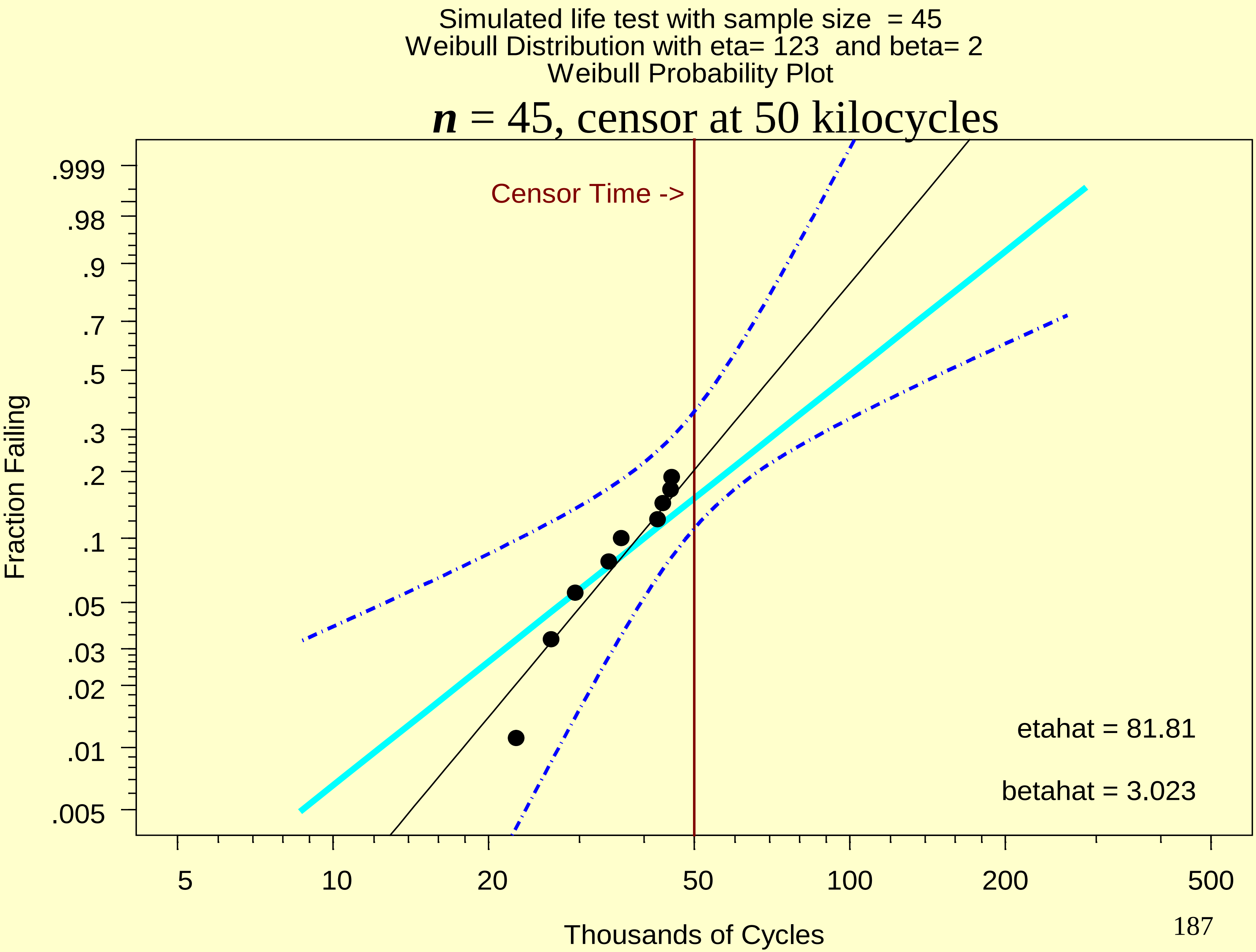
Simulated life test with sample size = 45
Weibull Distribution with eta= 123 and beta= 2
Weibull Probability Plot
n = 45, censor at 50 kilocycles
Censor Time ->
etahat = 86.75
betahat = 3.214
Fraction Failing
.999
.98
.9
.7
.5
.3
.2
.1
.05
.03
.02
.01
.005
5
10
20
50
100
200
500
Thousands of Cycles

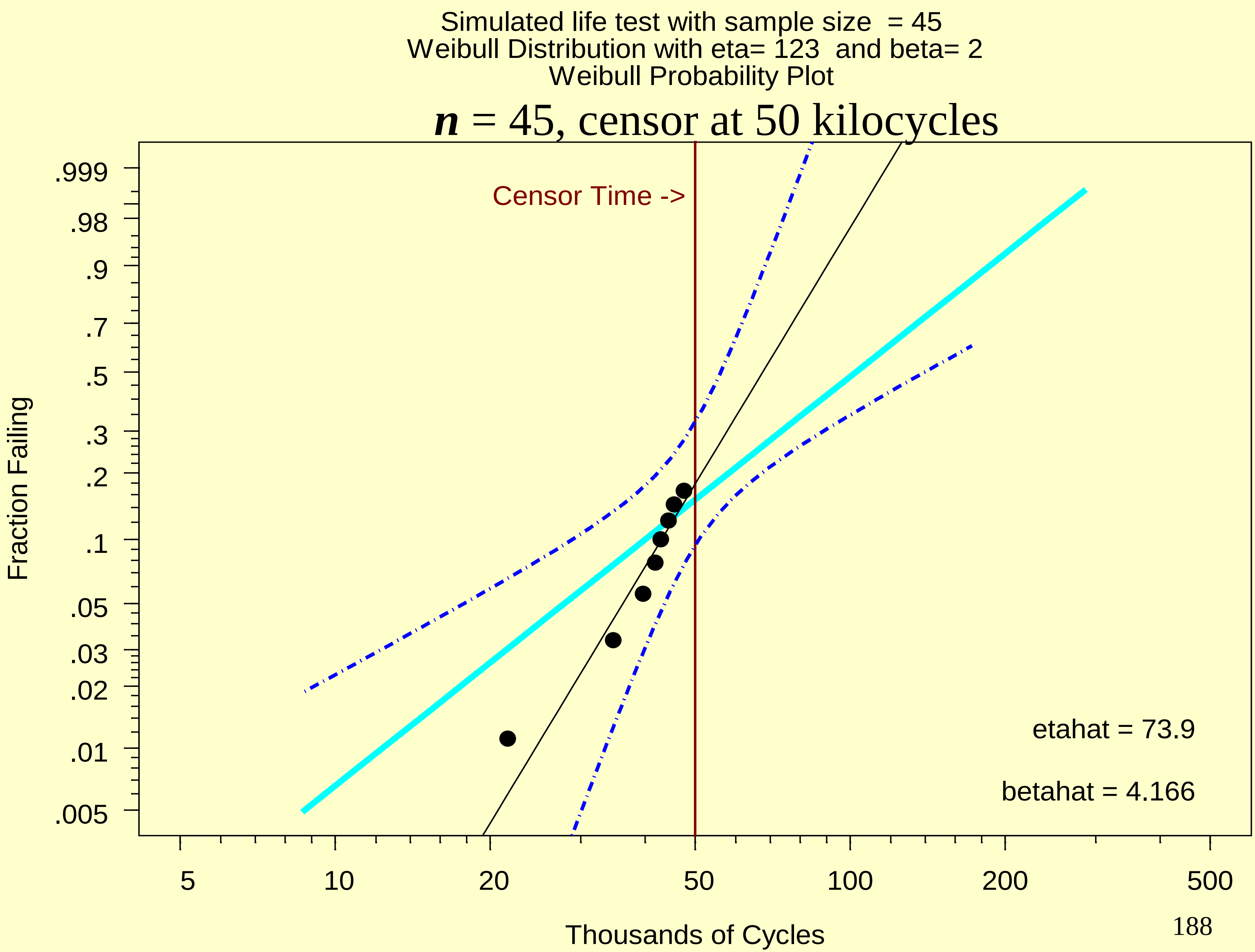

Simulated life test with sample size = 45
Weibull Distribution with eta= 123 and beta= 2
Weibull Probability Plot
n = 45, censor at 50 kilocycles
Censor Time ->
Fraction Failing
.999
.98
.9
.7
.5
.3
.2
.1
.05
.03
.02
.01
.005
5
10
20
50
100
200
500
Thousands of Cycles
etahat = 157.5
betahat = 2.061

n = 45, censor at 50 kilocycles



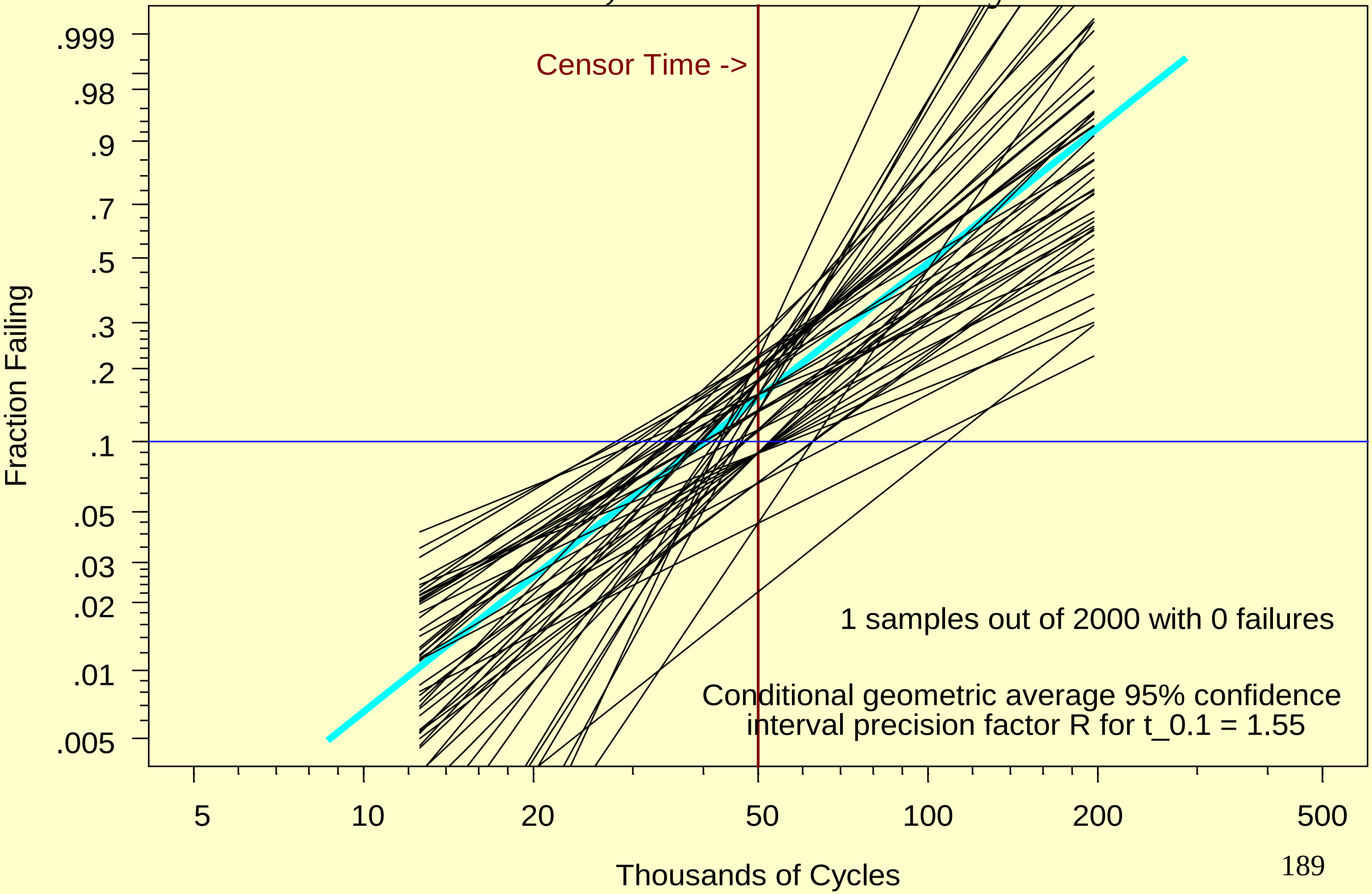
n = 45, censor at 50 kilocycles




Simulated life test with sample size = 45
Weibull Distribution with eta= 123 and beta= 2
Weibull Probability Plot
n = 45, censor at 50 kilocycles
Censor Time ->
.999
.98
.9
.7
.5
.3
.2
.1
.05
.03
.02
.01
.005
Fraction Failing
5
10
20
50
100
200
500
Thousands of Cycles
etahat = 73.9
betahat = 4.166

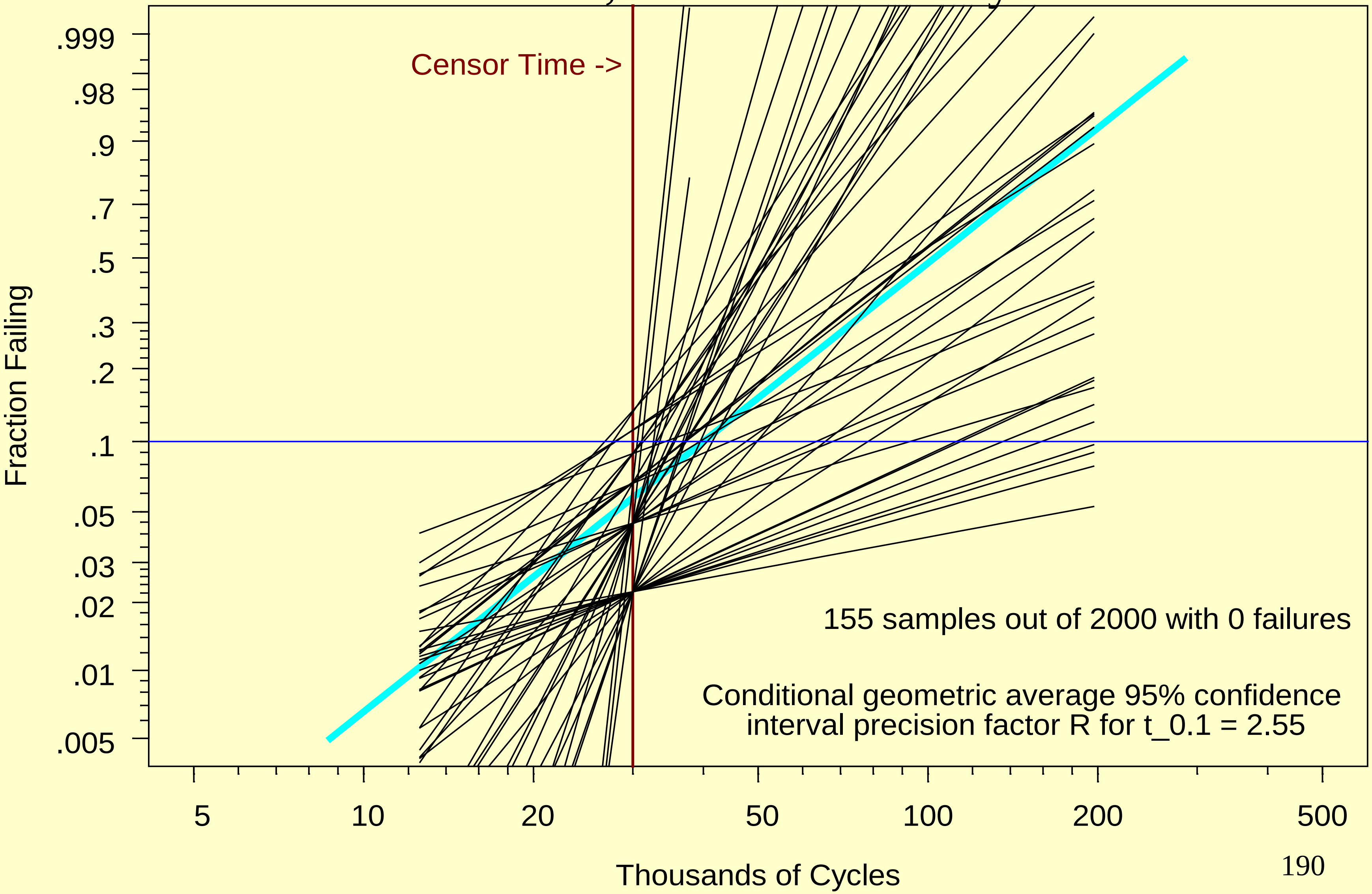
2000 simulated life tests with sample size = 45 censored at 50 Thousands of Cycles
Weibull Distribution with eta= 123 and beta= 2 with E(r)= 6.83
Weibull Probability Plot

n = 45, censor at 50 kilocycles



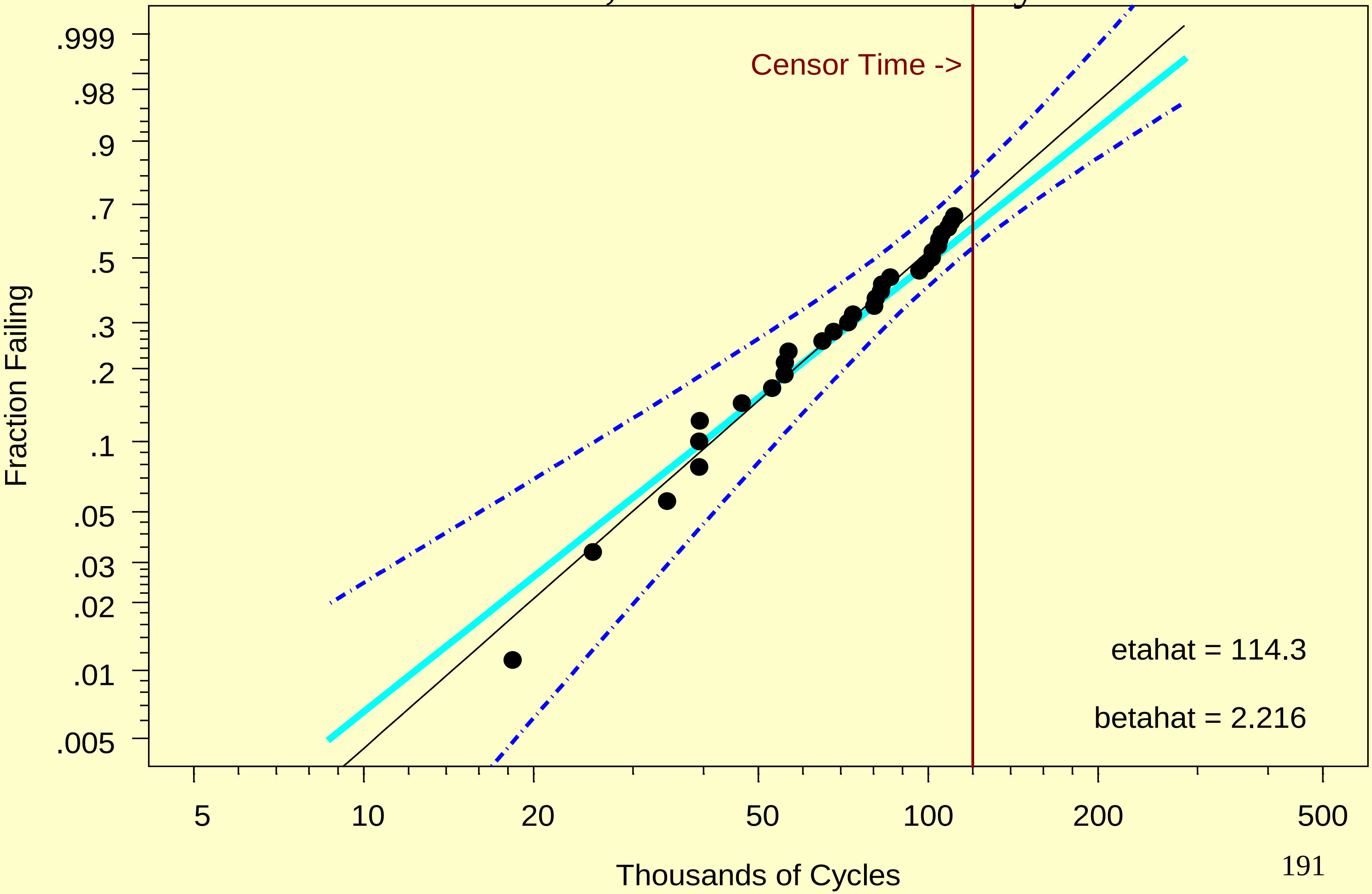
2000 simulated life tests with sample size = 45 censored at 30 Thousands of Cycles
Weibull Distribution with eta= 123 and beta= 2 with E(r)= 2.59
Weibull Probability Plot

n = 45, censor at 30 kilocycles



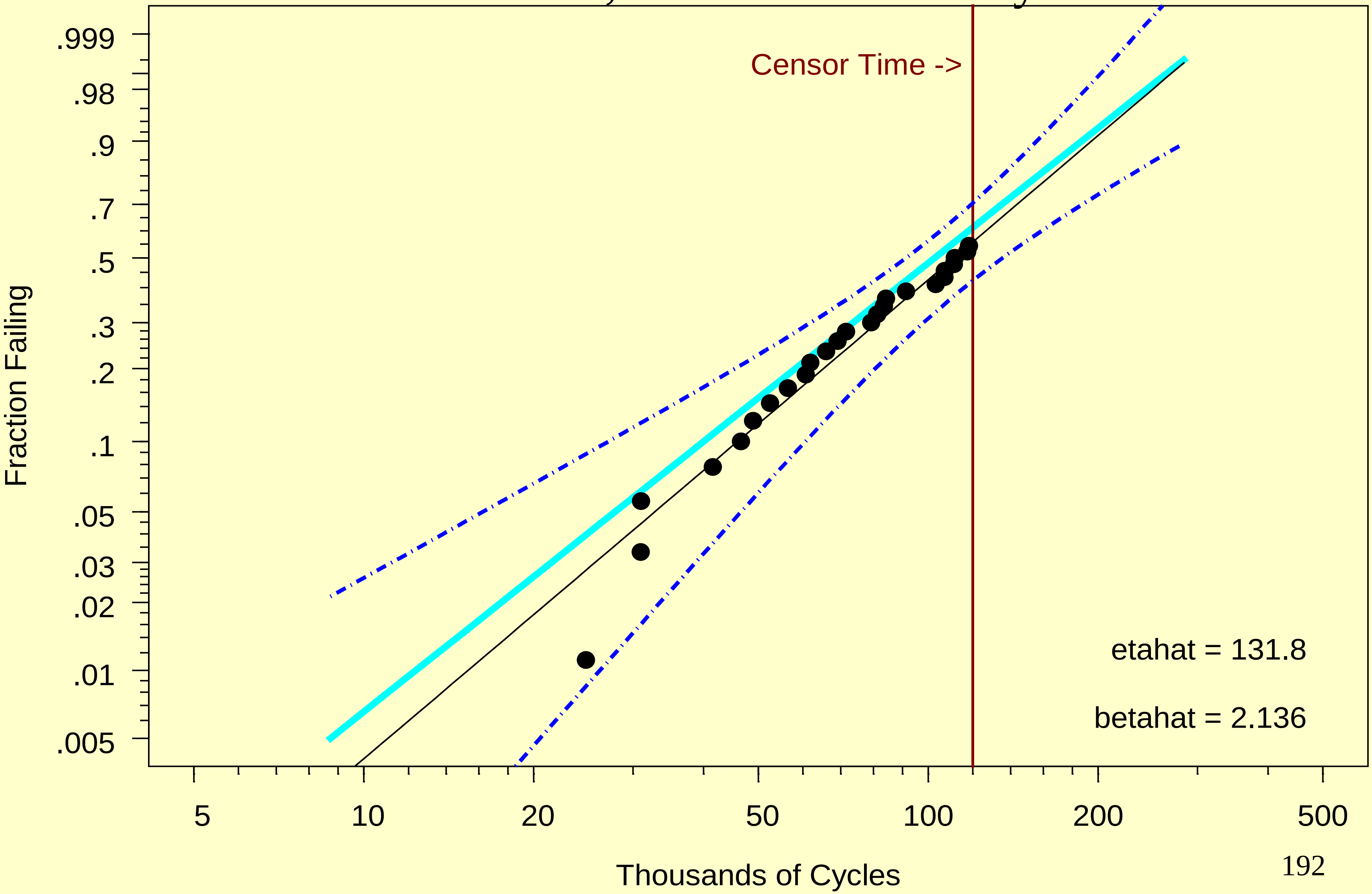
Simulated life test with sample size = 45
Weibull Distribution with eta= 123 and beta= 2
Weibull Probability Plot

n = 45, censor at 120 kilocycles

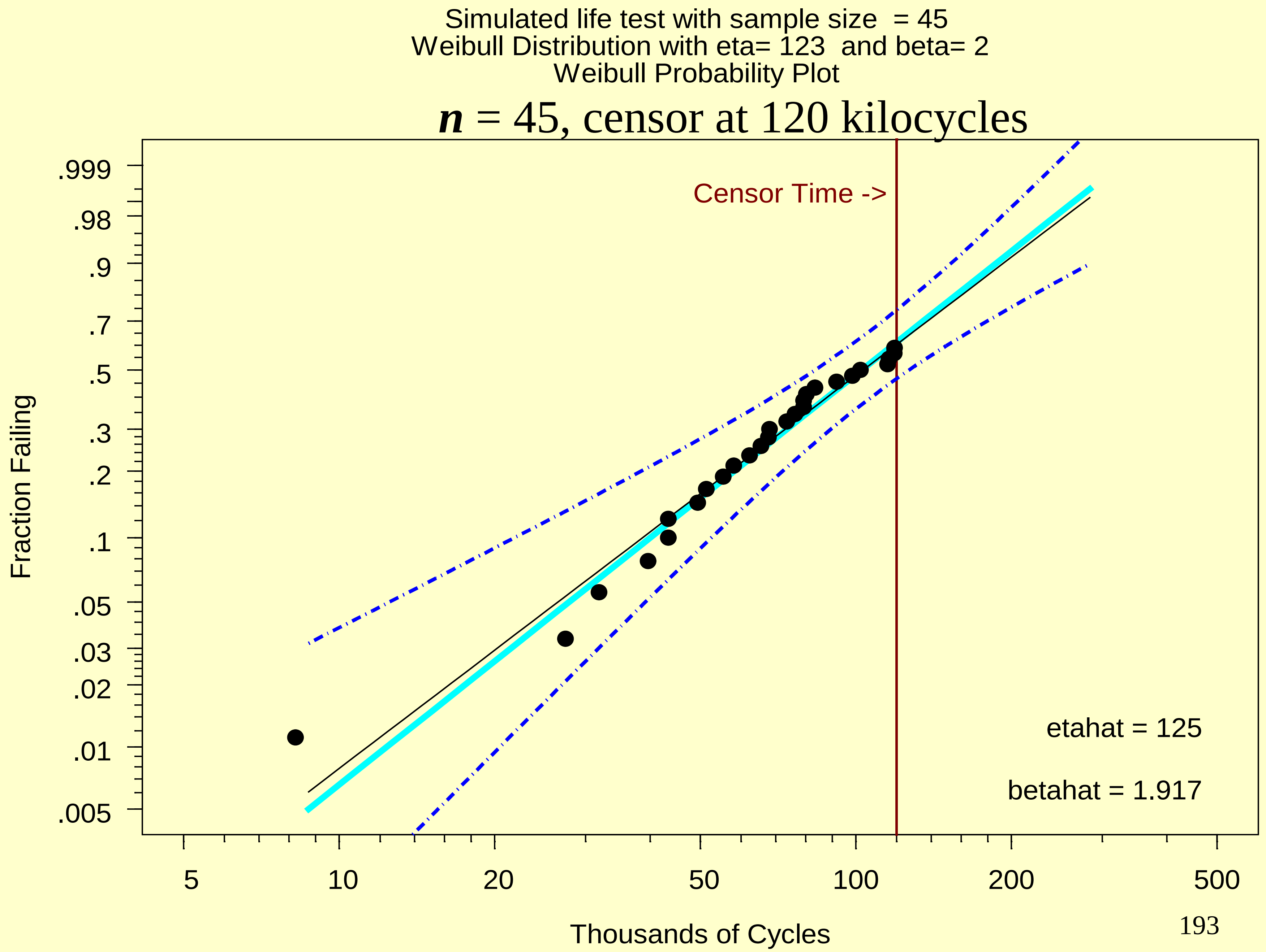


Simulated life test with sample size = 45
Weibull Distribution with eta= 123 and beta= 2
Weibull Probability Plot

n = 45, censor at 120 kilocycles

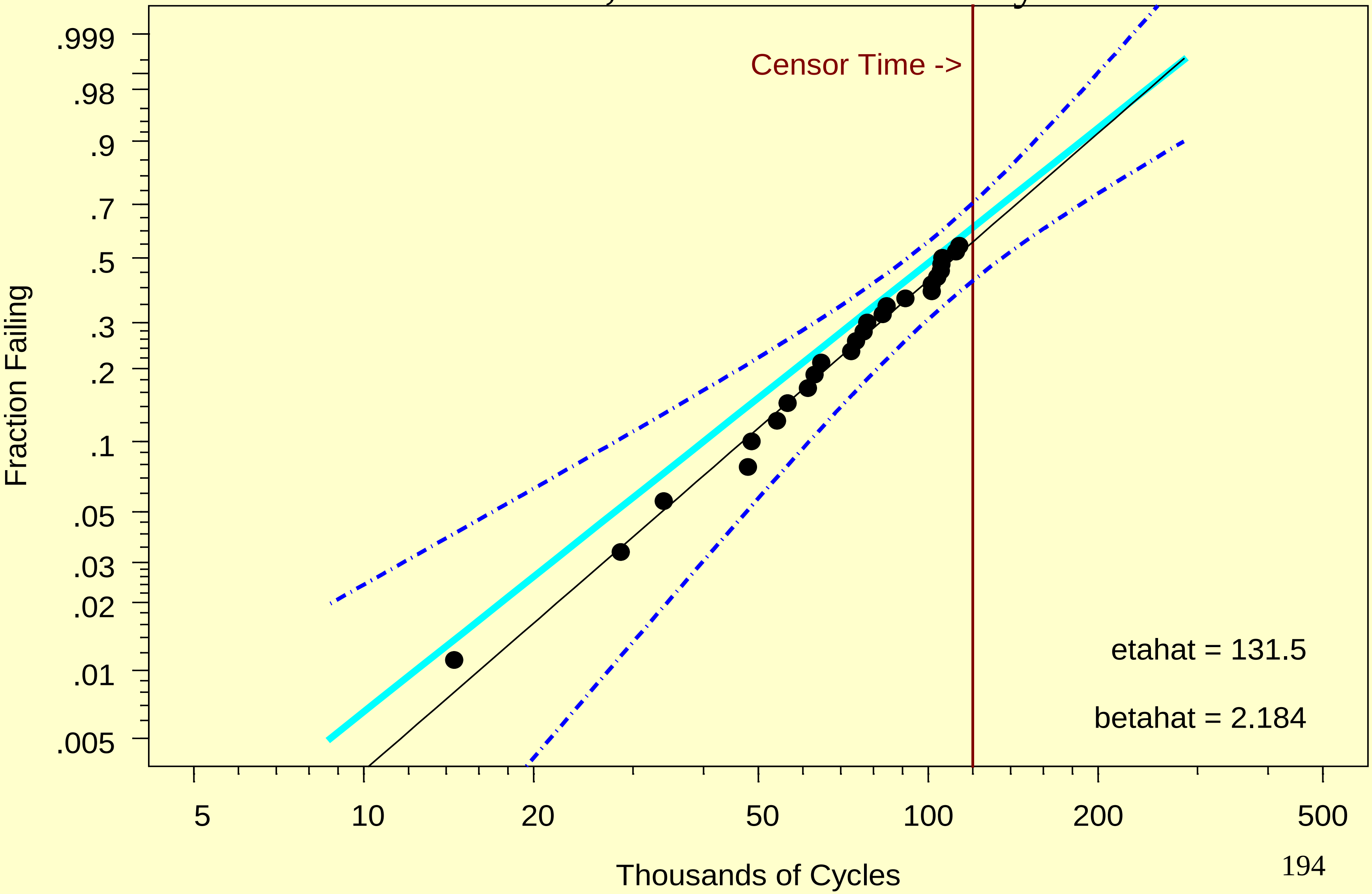


$\boldsymbol{n}$ = 45, censor at 120 kilocycles

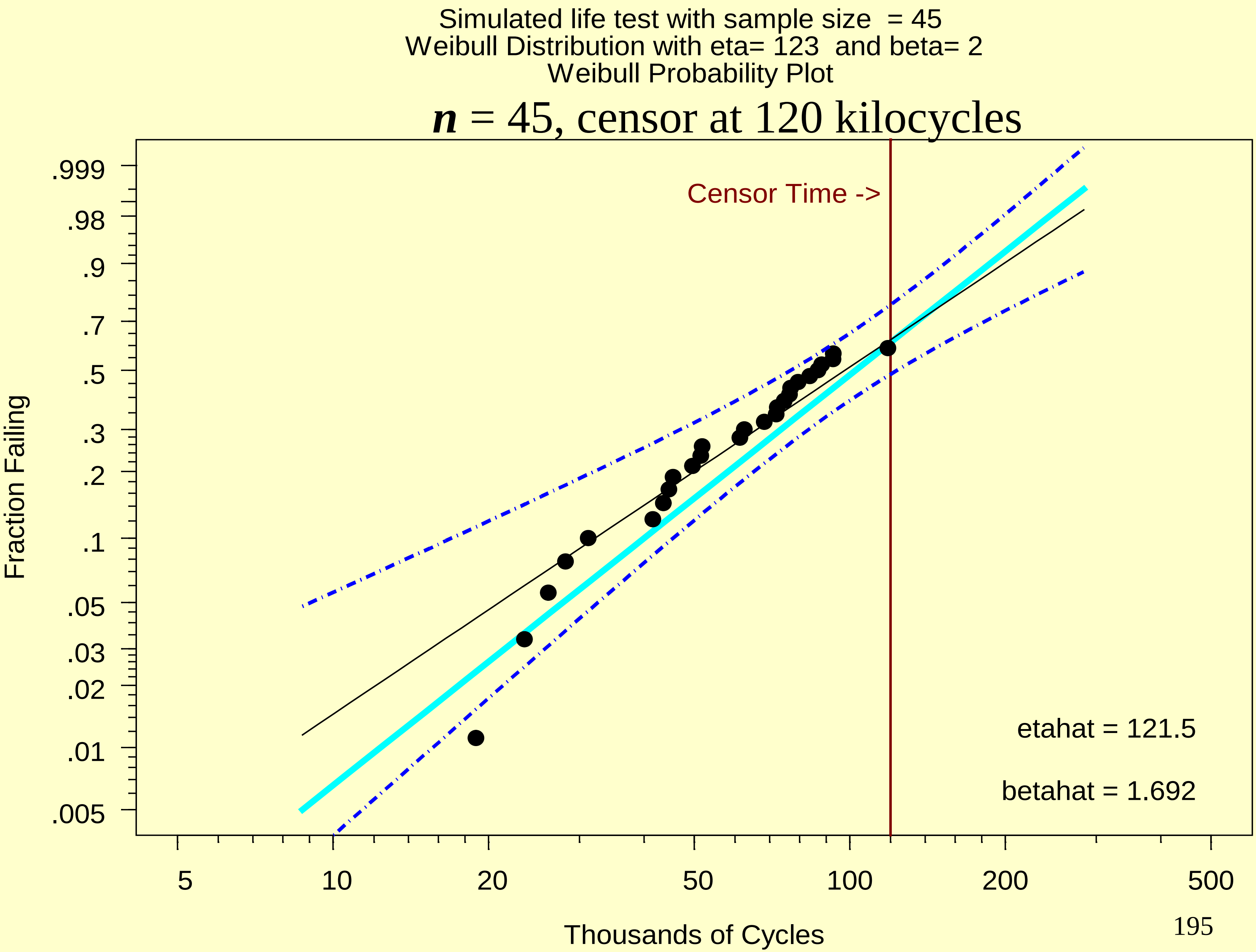


Simulated life test with sample size = 45
Weibull Distribution with eta= 123 and beta= 2
Weibull Probability Plot

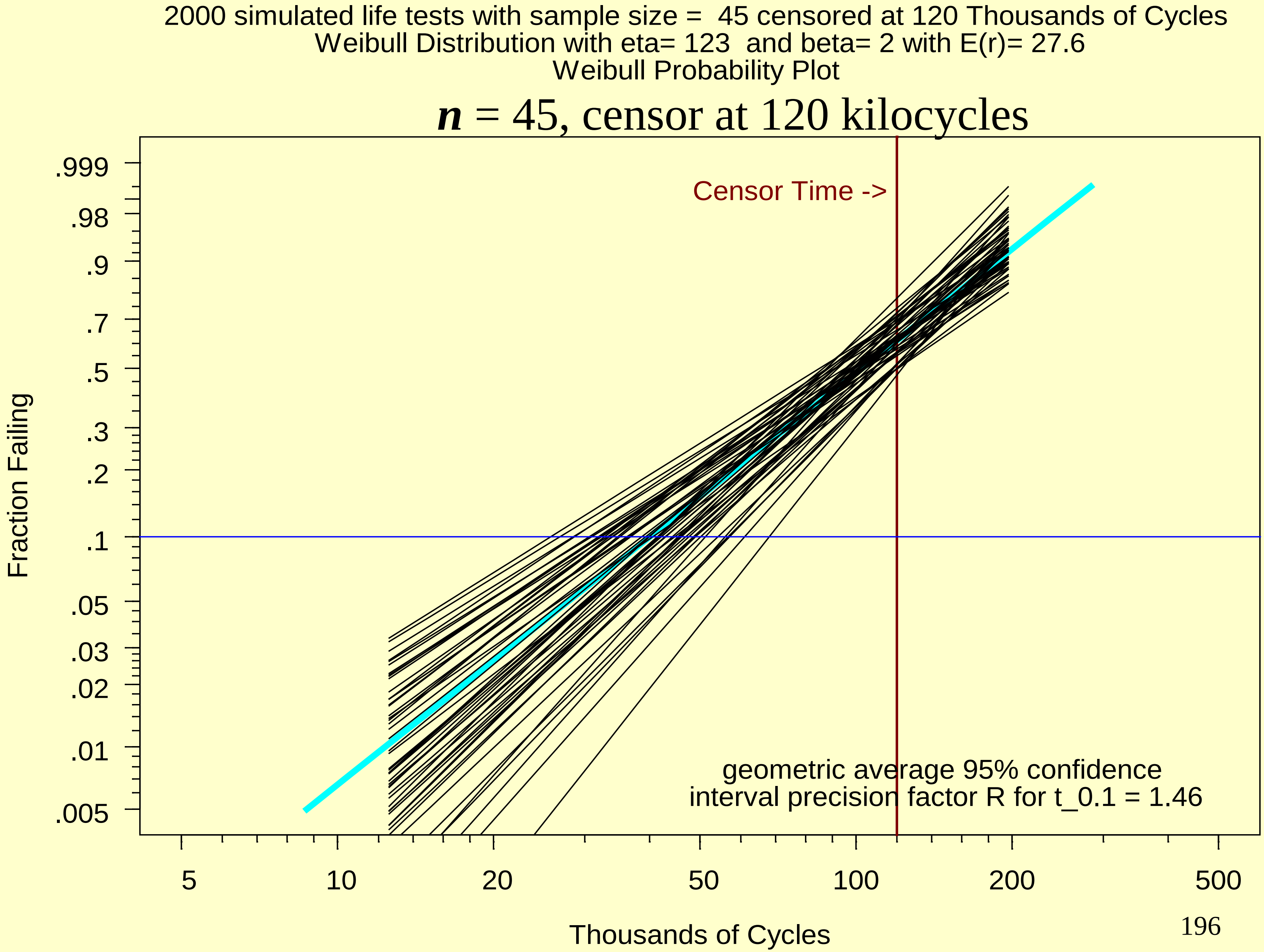
n = 45, censor at 120 kilocycles



$\boldsymbol{n}$ = 45, censor at 120 kilocycles

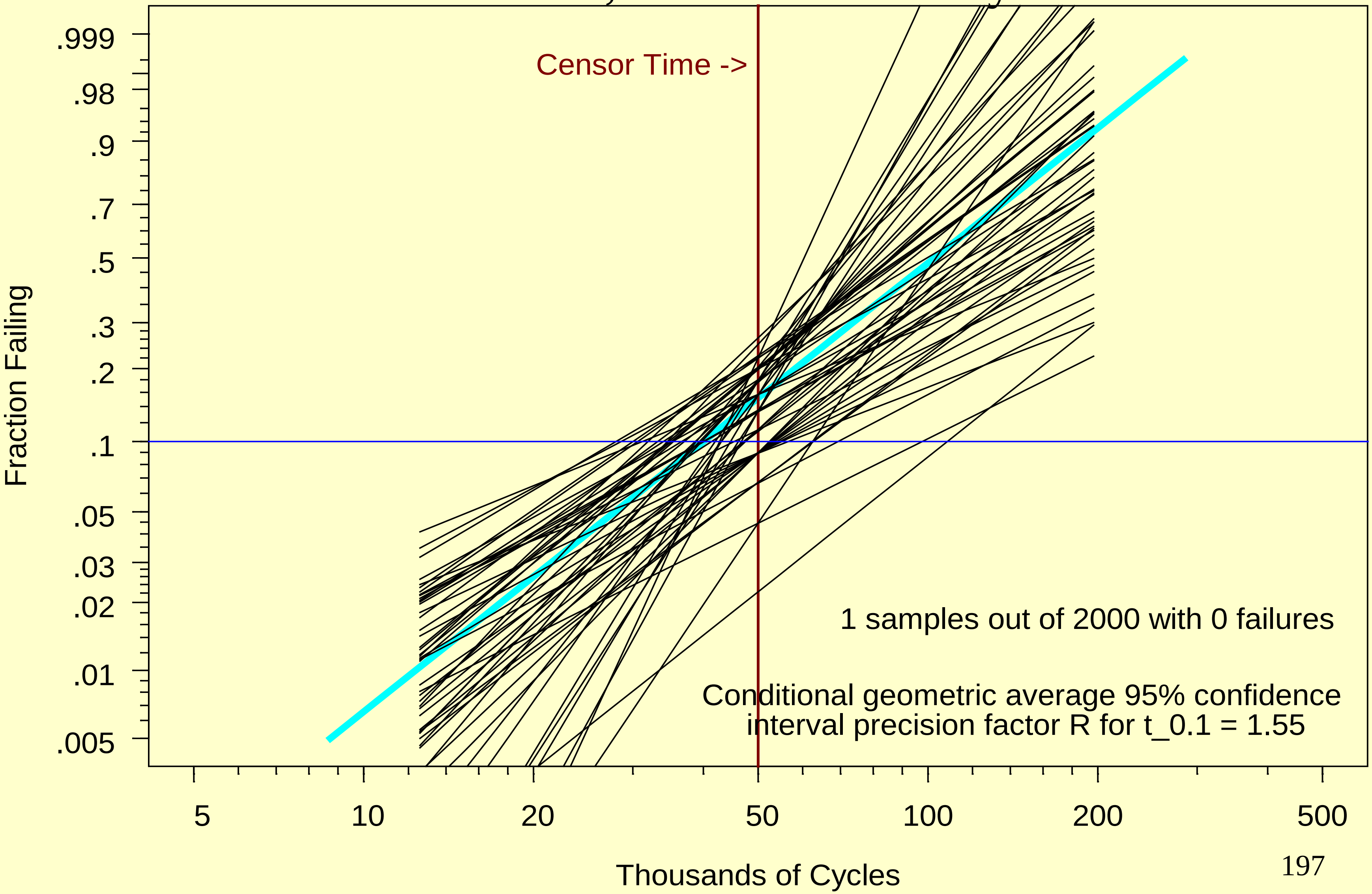


n = 45, censor at 120 kilocycles



2000 simulated life tests with sample size = 45 censored at 50 Thousands of Cycles
Weibull Distribution with eta= 123 and beta= 2 with E(r)= 6.83
Weibull Probability Plot

n = 45, censor at 50 kilocycles



Concluding Remarks

- There are many kinds of reliability data: life, degradation, recurrence
- We often need to use available information outside of your data, but recognize uncertainty in that information.
- Divide and analyze (e.g., failure cause, stratification).
- Do not ignore model uncertainty, especially in extrapolation.
- There is no magic in accelerated testing or statistics
- Finding appropriate accelerating variables and a model adequate for extrapolation are critical concerns.
- Statistical methods allow one to plan efficient tests, fit given models to data, and to quantify statistical uncertainty
- Modern software can provide useful test-planning tools