GPS FOM Chimney Analysis using Generalized Extreme Value Distribution

Rick Ott, Joe Frisbee and Kanan Saha

Abstract. Many a time an objective of a statistical analysis is to estimate a limit value like 3-sigma 95% confidence upper limit from a data sample. The generalized Extreme Value Distribution method can be profitably employed in many situations for such an estimate.
It is well known that according to the Central Limit theorem the mean value of a large data set is normally distributed irrespective of the distribution of the data from which the mean value is derived. In a somewhat similar fashion it is observed that many times the extreme value of a data set has a distribution that can be formulated with a Generalized Distribution.

For a set of data $x_1, x_2, \ldots, x_n$, let the mean value be denoted by X_m where

$$X_m = (x_1+x_2+..+...+..+x_n)/n$$

Let $Z = (X_m - \mu) / \sigma/\sqrt{n})$ where μ, σ^2 are population mean and variance.

Central limit theorem states that for large n, Z is normally distributed with zero mean and unit variance. Similarly in many situations the extreme value of a large data set has a Generalized Extreme value distribution given by

$$G(z) = \exp\{-[1+\xi((z-\mu)/\sigma)]^{-1/\xi}\},$$

where ξ is a parameter that together with μ and σ specifies distribution characteristics.
In space shuttle entry with 3-string GPS navigation the Figure Of Merit (FOM) value gives a measure of GPS navigated state accuracy. A GPS navigated state with FOM of 6 or higher is deemed unacceptable and is said to form a FOM 6 or higher chimney. A FOM chimney is a period of time during which the FOM value stays higher than 5. A longer period of FOM of value 6 or higher causes navigated state to accumulate more error for a lack of state update. For an acceptable landing it is imperative that the state error remains low and hence at low altitude during entry GPS data of FOM greater than 5 must not last more than 138 seconds.
To test the GPS performance many entry test cases were simulated at the Avionics Development Laboratory. Only high value FOM chimneys are consequential. The extreme value statistical technique is applied to analyze high value FOM chimneys. The Maximum likelihood method is used to determine parameters that characterize the GEV distribution. And then the limit value statistics are estimated.

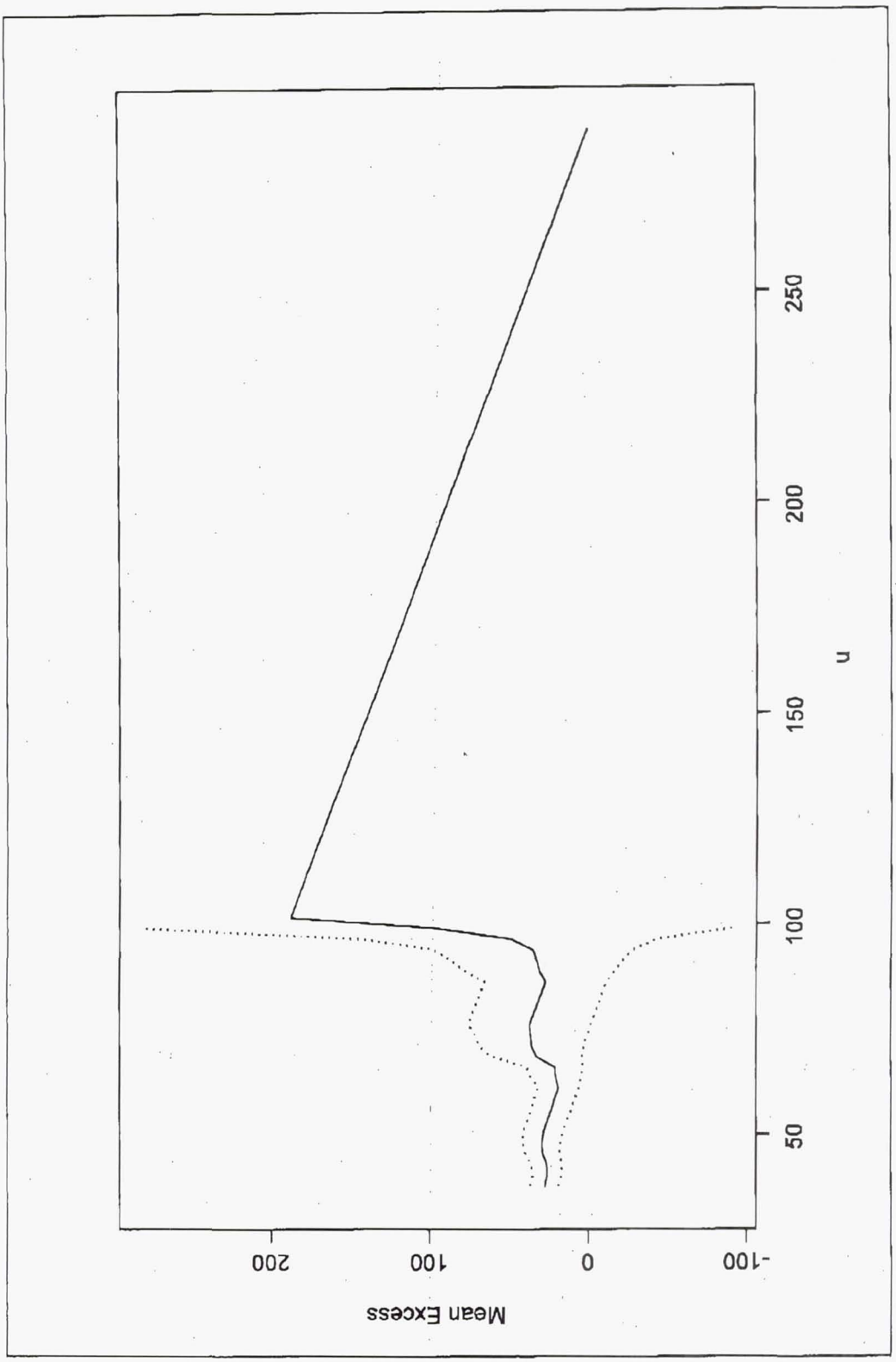
GPS FOM Chimney Analysis using Generalized Extreme Value Distribution

- Rick Ott, Joe Frisbee and Kanan Saha

STPOR Recovery Data

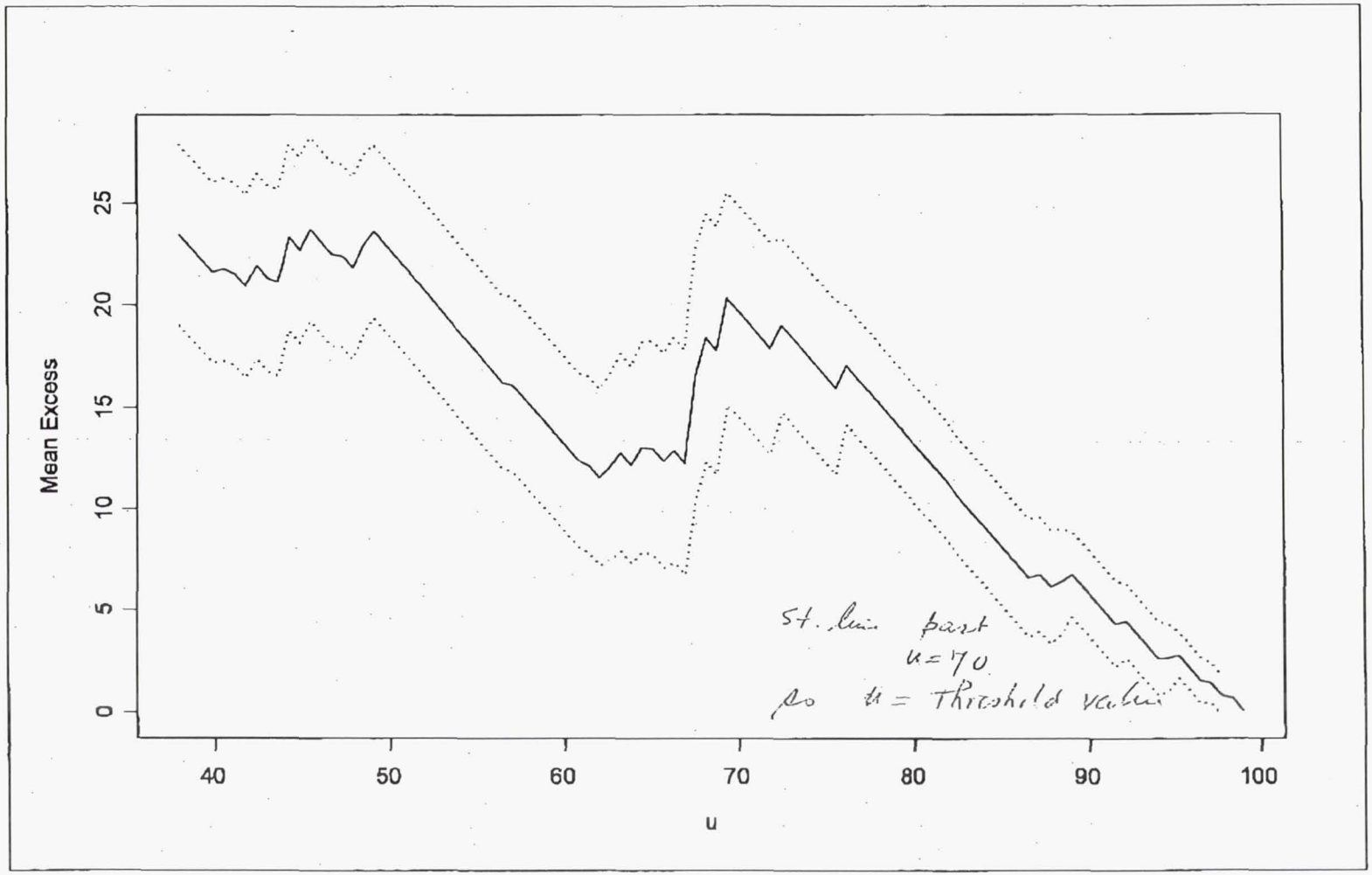
- Data: One entire small set, n = 60 (59)
- Data does **NOT** pass Normal QQ Regression Test r = .949377, p-value < 1% (See Advanced Statistics: Monte Carlo, Bootstrapping, Multivariate Methods, 1991 C. Hallum & R. Chhikara)
- Since complete set is given, the first method chosen is a Generalized Pareto Threshold Model
- Threshold Chosen from Mean Residual Life Plot and Parameter Threshold Plots (if possible)
- Generalized Pareto Threshold Model is Nonparametric, Very Conservative, No Assumptions besides Max of F(x) follows a Generalized Extreme Value Distribution!

STPOR Recovery Mean Residual Life Plot (w/ 289)



STPOR Recovery Mean Residual Life Plot (wo/ 289)

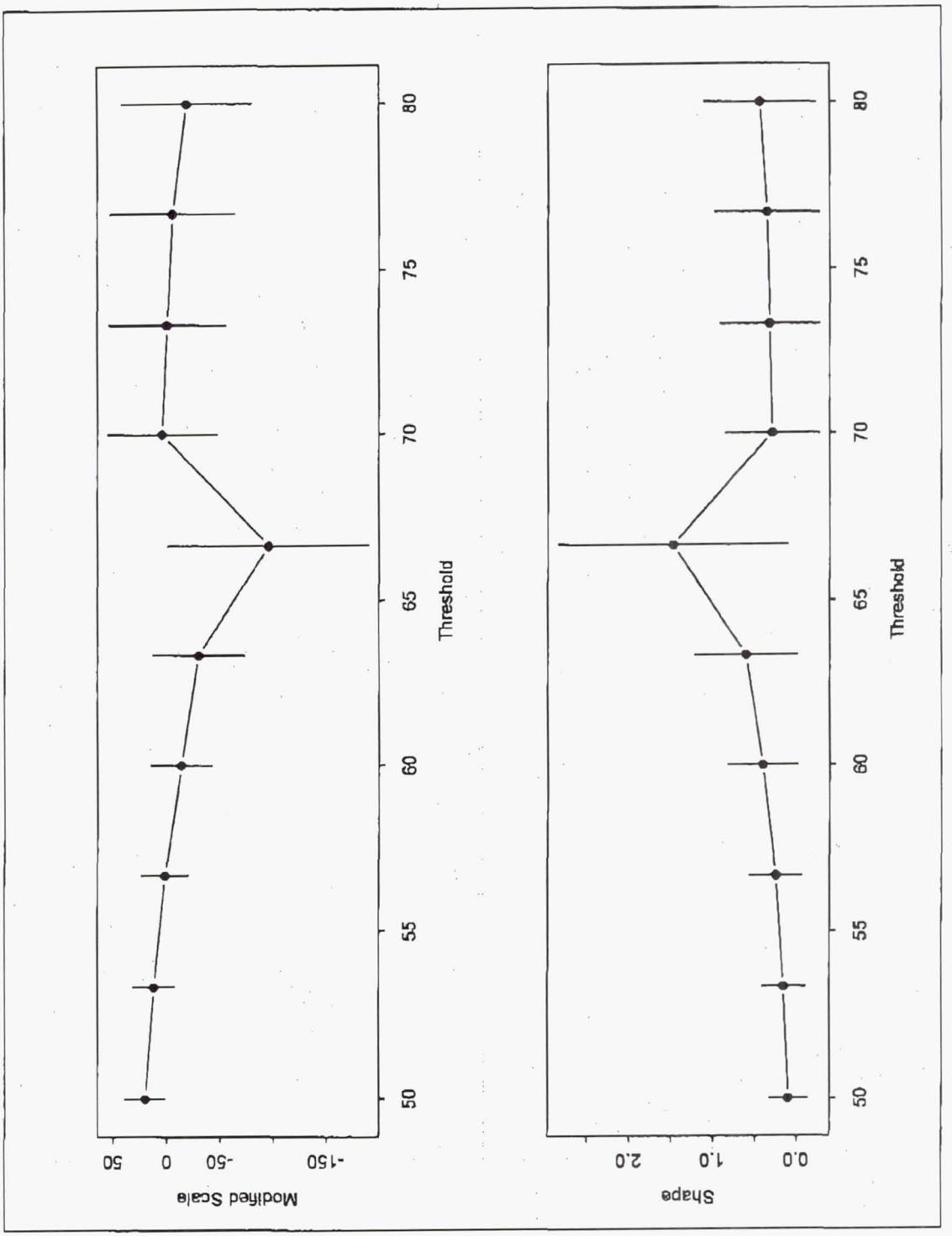
without outlier point of value 289



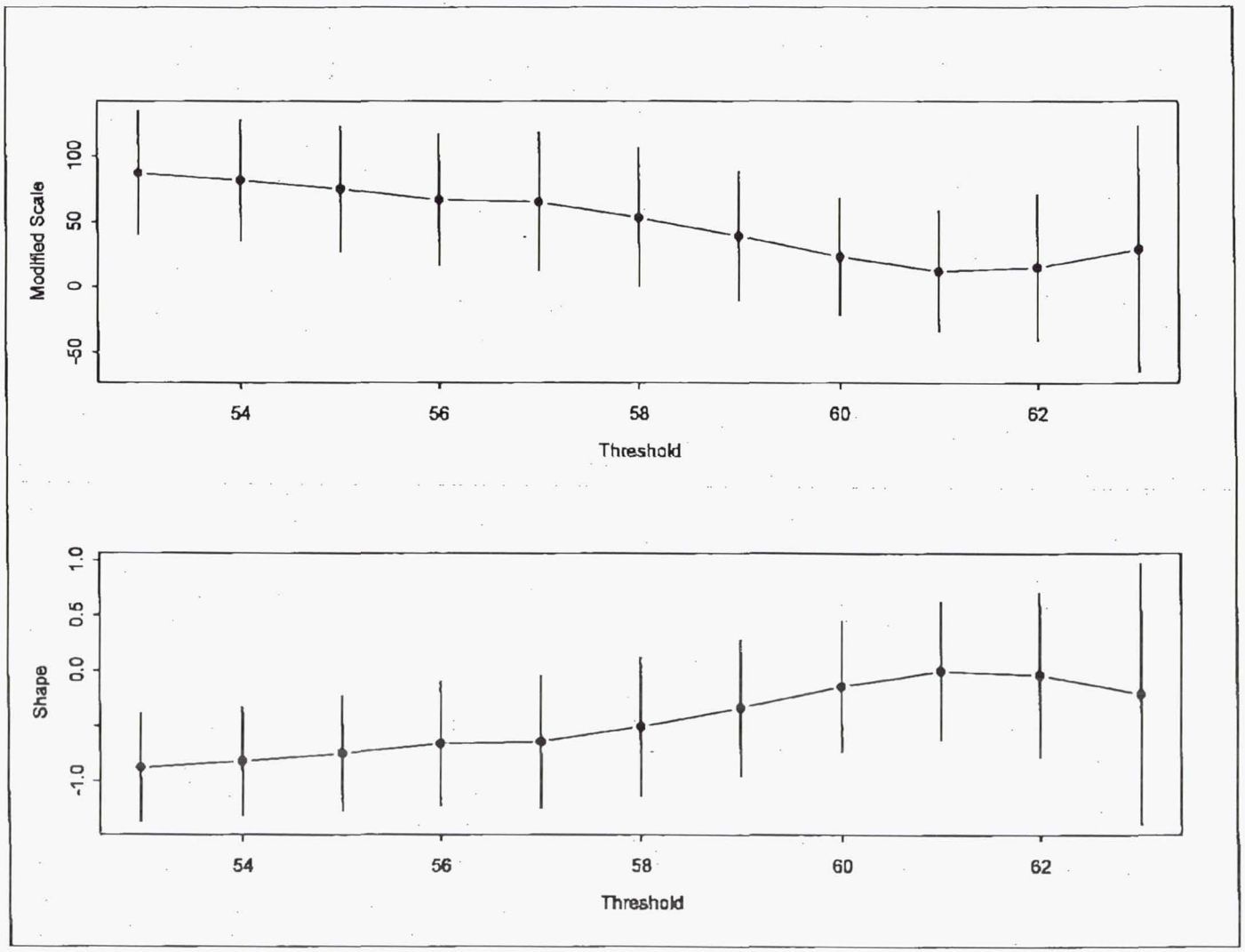
Mean Residual Life Plots

- Without 289 MRLP suggest threshold u ≈ 70
- With 289 MRLP plot severely influenced by suspected 289 outlier – No threshold suggested
- PROBLEM: Even if 70 is used as threshold only 12 (11) points used for parameter estimation. Without 289 Data Set, software does not converge to give estimates if threshold above 63.

Parameter Threshold Plot w 289



Parameter Threshold Plot wo 289



STPOR Recovery GPD
(Generalized Pareto Distribution)
Estimates and Standard Errors (w/ 289)
Threshold **63**

Number of points used in estimation: 28 out of 60

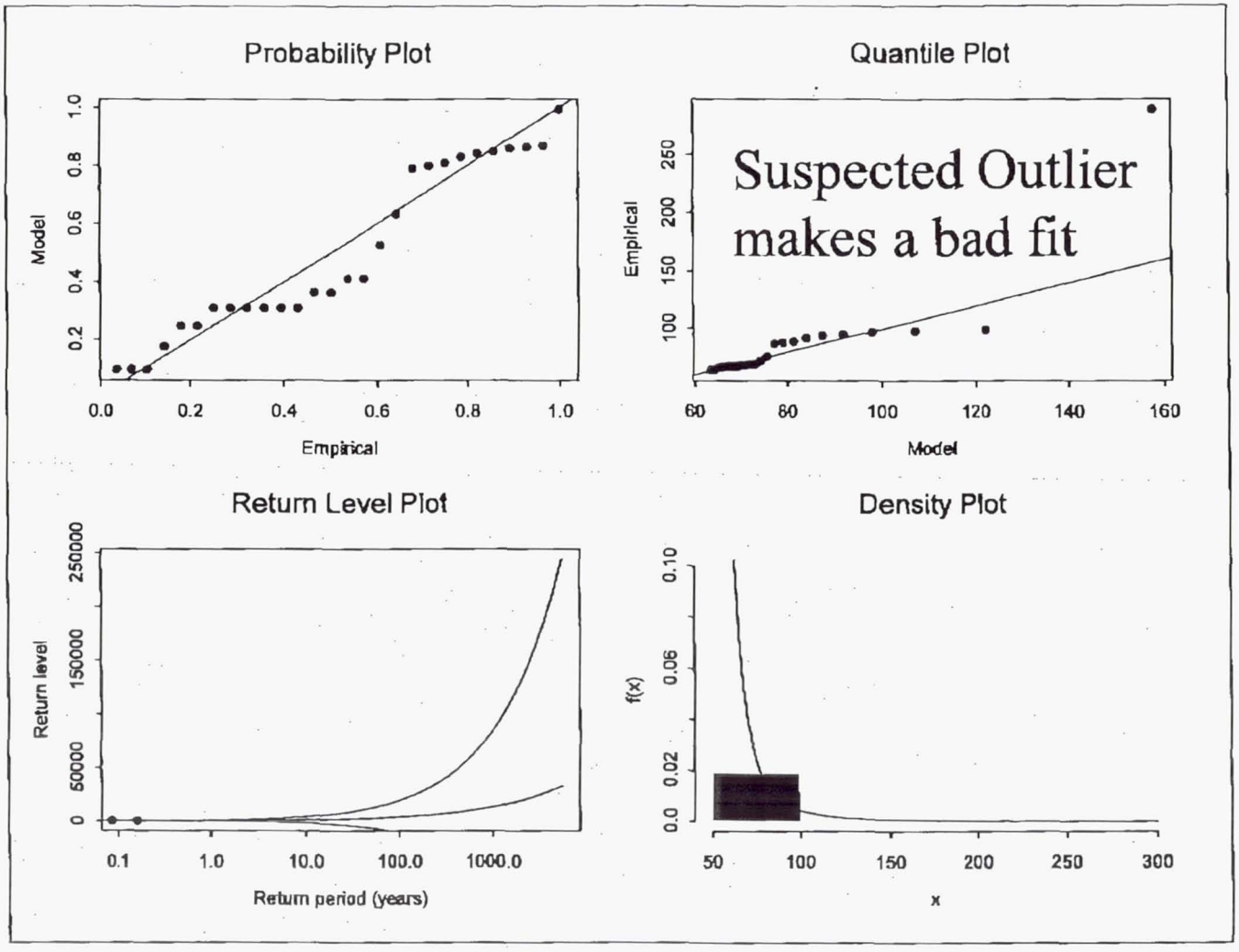
Max Negative log-likelihood:107.126

maximum likelihood estimate

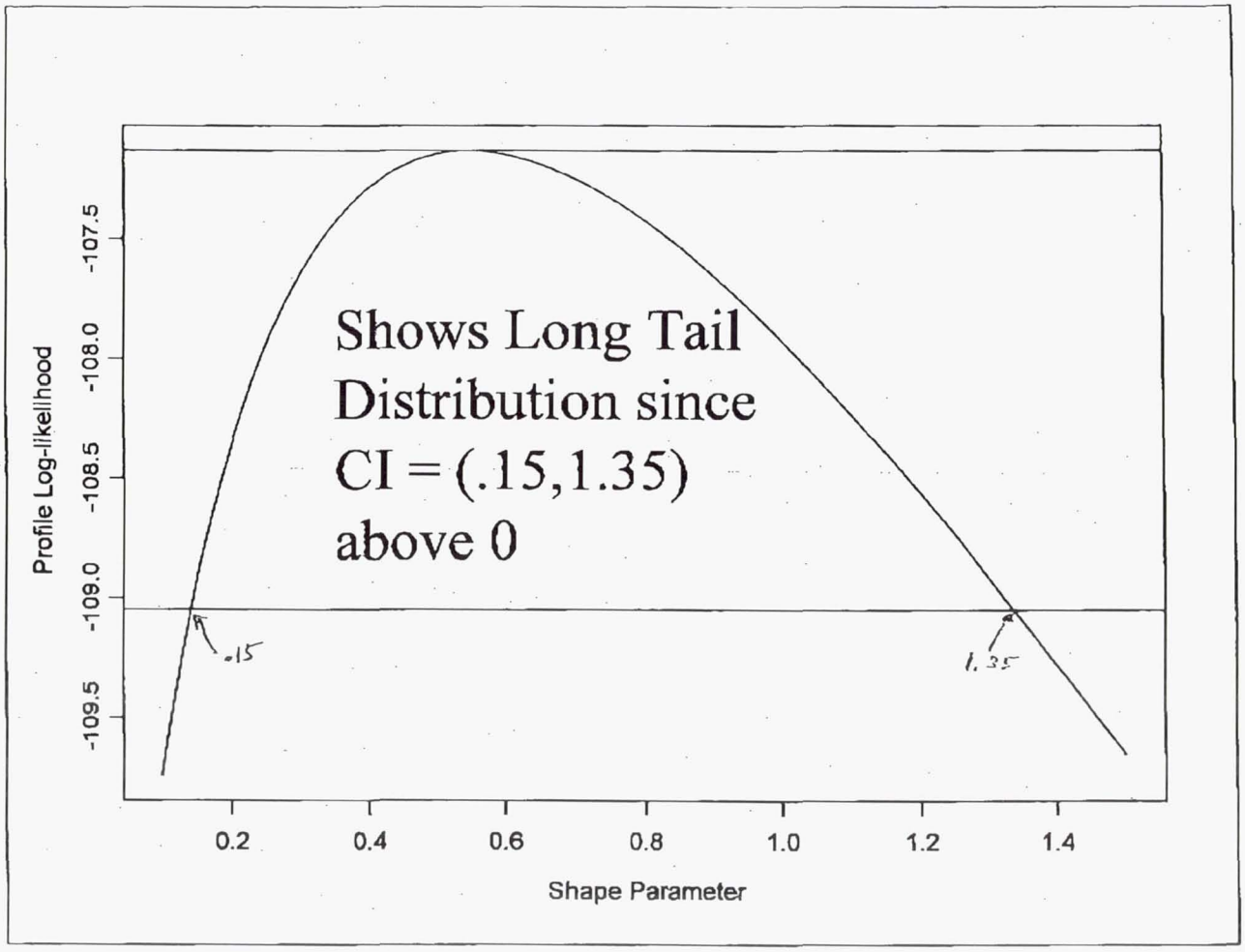
mles: $\sigma^* = 9.7694690$ $\xi = 0.5466667$

$se_{\sigma^*} = 3.2344973$ $se_{\xi} = 0.2908163$

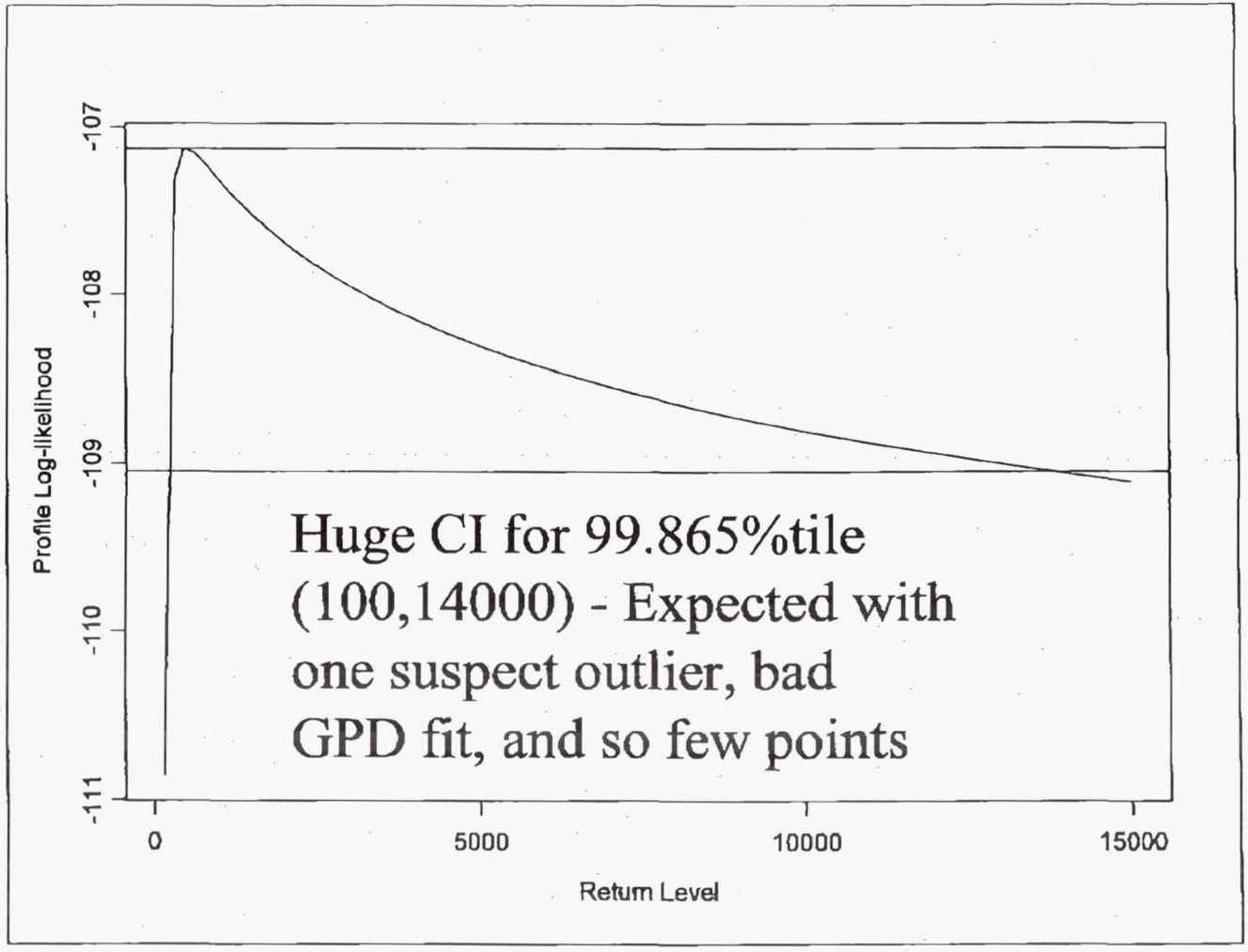
Threshold 63 w 289 Diagnostic Plots



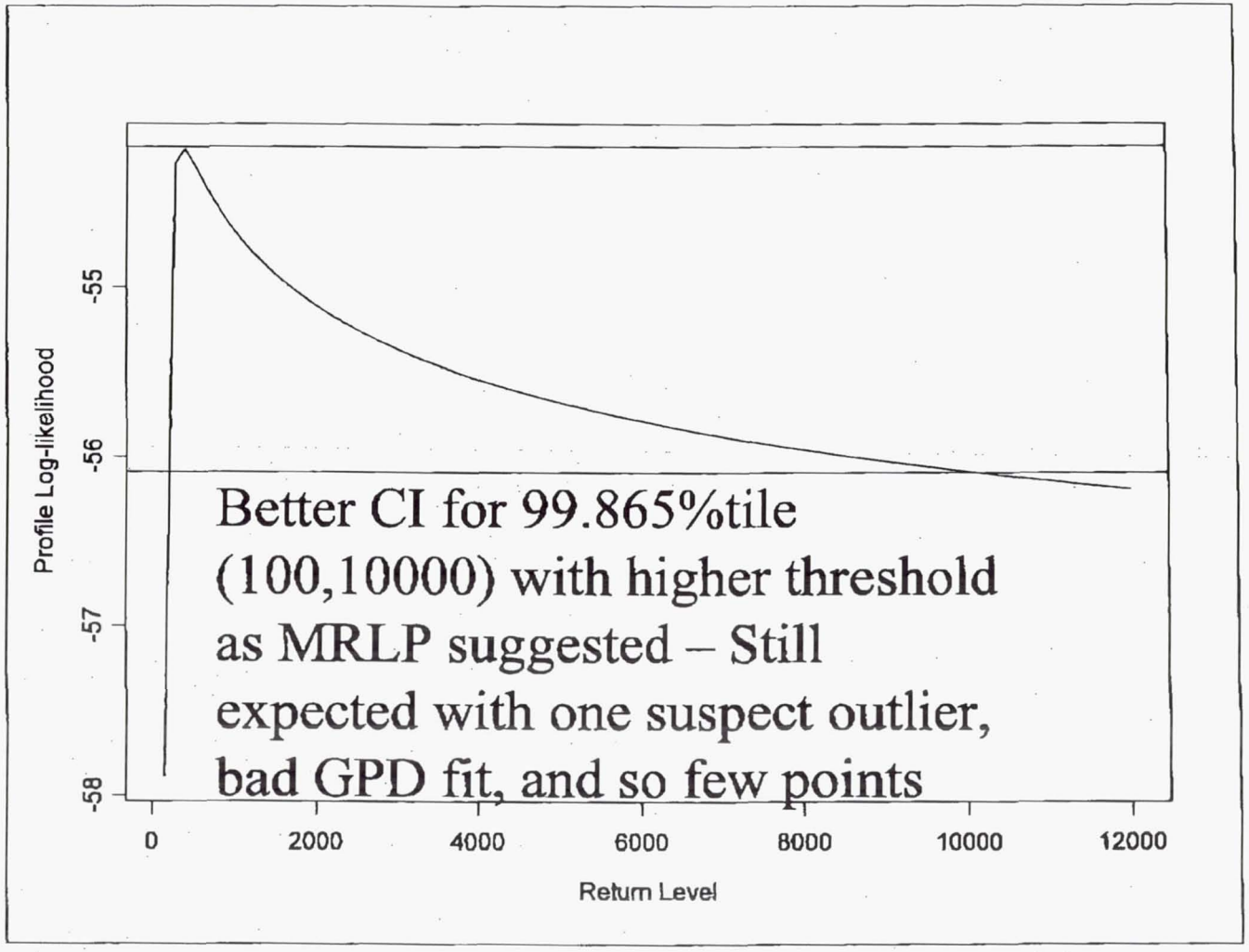
STPOR Recovery GPD
ξ Profile Log-Likelihood (w/ 289, u = 63)



STPOR Recovery GPD 99.865%tile Profile Log-Likelihood (w/ 289, u = 63)



STPOR Recovery GPD 99.865%tile Profile Log-Likelihood (w/ 289, u = 70)



STPOR Recovery GEV
Estimates and Standard Errors (wo/ 289)
Threshold **63**

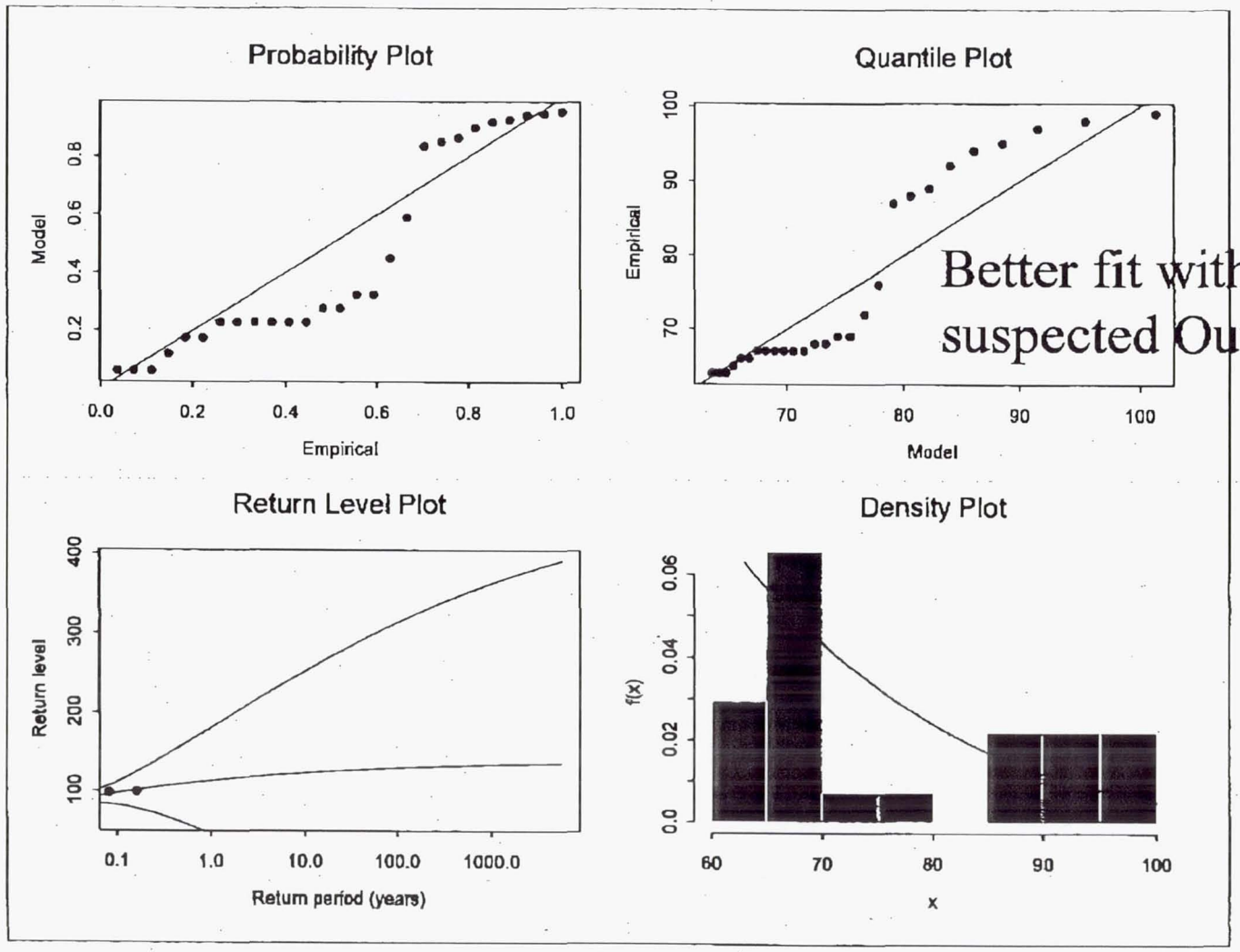
Number of points used in estimation: 27 out of 60

Max Negative log-likelihood:96.18224

mles: $\sigma^* = 15.9403808$ $\xi = -0.2065505$

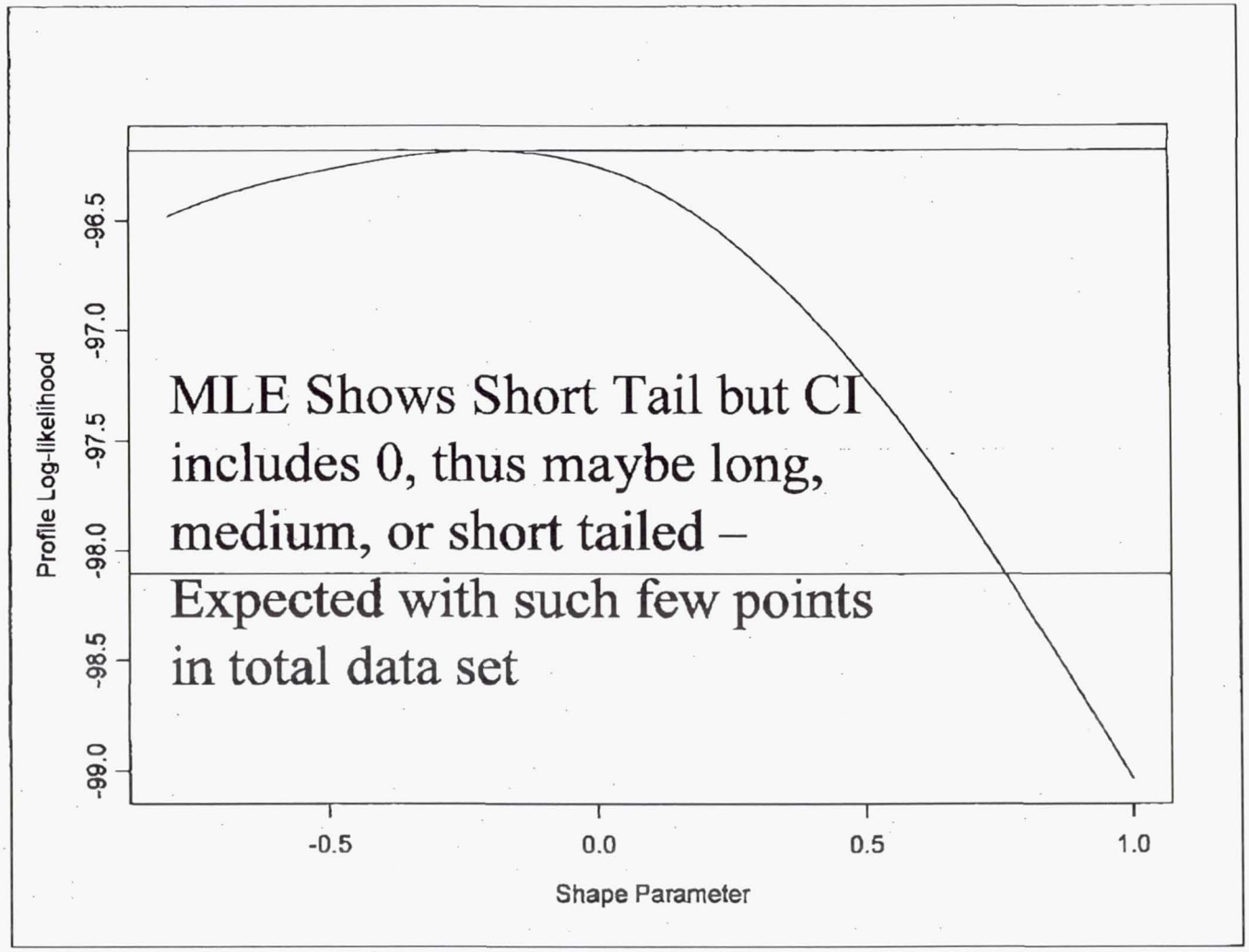
$se_{\sigma^*} = 10.0720236$ $se_{\xi} = 0.6031458$

Threshold 63 wo 289 Diagnostic Plots

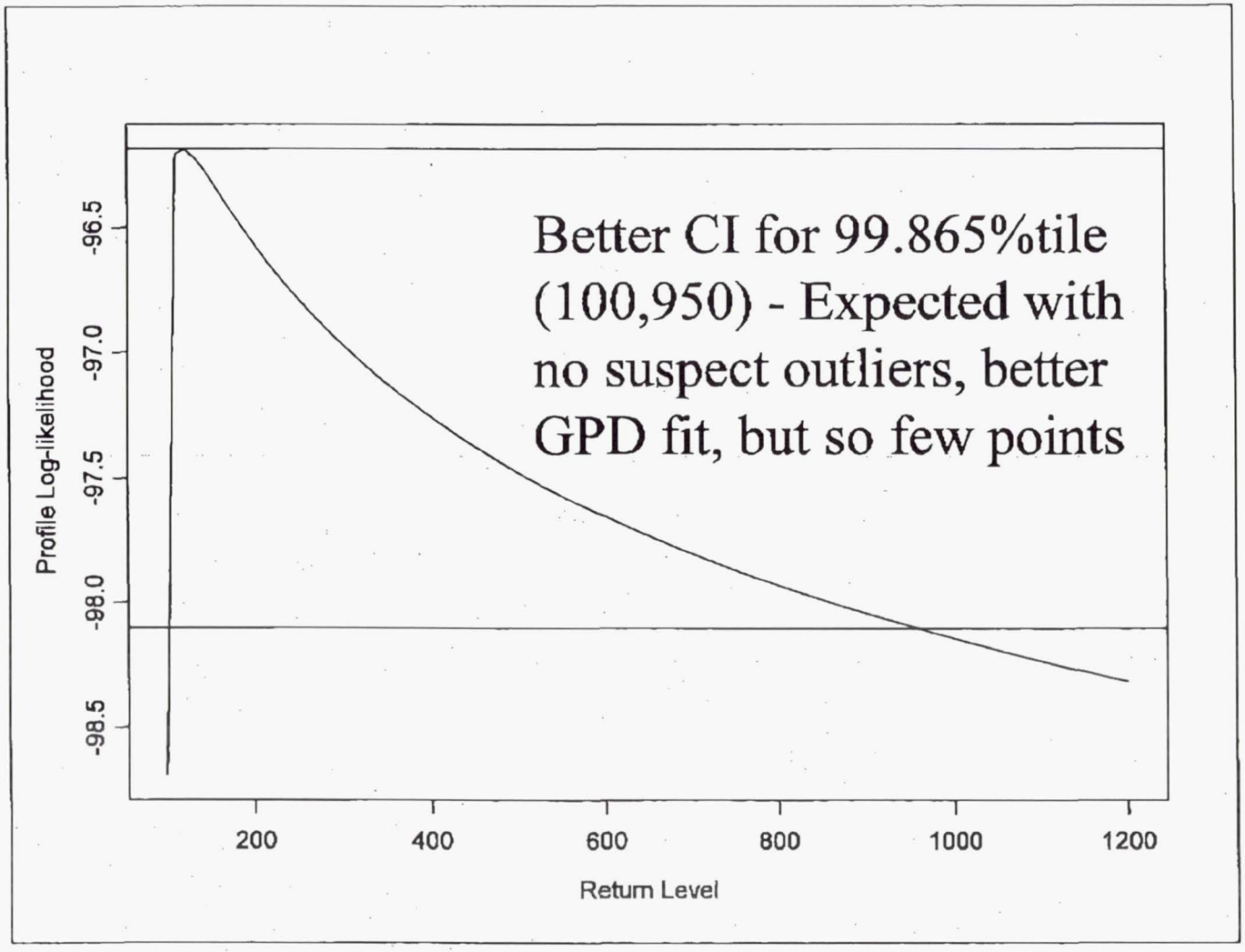


Better fit without suspected Outlier

STPOR Recovery GPD
ξ Profile Log-Likelihood (wo/ 289, u = 63)



STPOR Recovery GPD 99.865%tile Profile Log-Likelihood (wo/ 289, u = 63)



Conclusions

- Data does NOT pass Normality Test
- Large CI for 99.865%tile due to lack of data and nonparametric GPD method

- Next to Come:
  - Weissman Estimates based on assuming a specific GEV family for k largest statistics
  - Extreme Spacing Test to Reject Medium Tail Distribution in Favor of Large or Short Tail