War-Winning Capabilities … On Time, On Cost

Title: Parametric TLE

Date: 11 May 2017

- Author: James Brownlow
- Office 812 TSS/ENTR
- Phone: 661 277-4843

U.S. AIR FORCE

Preamble:

- Target Location Error, in many instances, varies with range to the target
- Todd Remund, Greg Hutto and Jeff Beekman (Beekster) demonstrated the use of a parametric survival model to estimate targeting system CEP as a function of range to the target
- This presentation details how the parametric survival model was adapted to estimation of TLE.

Introduction

- Precision target location is germane to navigation, weapon delivery and target tracking.
- This presentation considers TLE as a function of range to the target, specifically when TLE is expressed in terms of ‘CExx’
  - CEP: 50^{th} percentile. Radius of a circle about the target that has the property that 50% of the time estimated target location is inside the circle, 50% of the time it is outside the circle
  - CE10 and CE90 the 10^{th} and 90^{th} percentile values
- Answer the question, “At range R what is the CEP?” or “what is the CE90?”

public release: 412TW-PA-17-0114

TLE: cumulative probability

- General idea: Probability(TLE ≤ d; r), the probability that TLE is less than some distance, d, at range r, to the target-cumulative distribution function:

$$\Pr(TLE \leq d; r) = F(d;r) = F(d)$$

- Models: *F(d)* for log-normal, Weibull:

z = (log(d)-μ(d))/σ, μ a function of range

- Simplest case: μ is a linear function

$$\mu_i = \beta_0 + \beta_1 r_i$$

- Fit a log-normal and Weibull distribution function

public release: 412TW-PA-17-0114

General Idea

- Also could model the probability that TLE is less than some quantity, d, as a function of a number of explanatory variables- range, software load, system, threat type, and the like
- More elaborate regression models are possible; quadratic regression for μ, and/or relating σ to range, or to other explanatory variables.
- Likelihood methods used to fit the model

Presentation overview

- Show how parametric reliability/failure models can be adapted to TLE analyses.
- Use log-normal and Weibull cumulative distribution functions estimate probability distributions of TLE at each range, r
  - Describe the parameters of the model
  - Set up maximum likelihood estimates of the model
  - Hypothesis tests for significance of parameters
  - Display results graphically

Aside

- Todd Remund, Jeff Beekman and Greg Hutto developed this approach using JMP© and R
- TLE analysis was derived from parametric survival models described in from Meeker and Escobar
- Remund *et al*. found that Weibull and log-normal probability functions appear to have the most utility in TLE analysis

Simplest case

For $\Phi_{\ln}$, log-normal distribution,

$$\Pr(Y \le y) = F(y;\mu,\sigma) = F(y;\beta_0,\beta_1,\sigma) = \Phi_{\ln}((y-\mu)/\sigma)$$

The quantile function for this model is

$$y_p(r) = \mu + \Phi_{\ln}^{-1}(p)\sigma = \beta_0 + \beta_1 r + \Phi_{\ln}^{-1}(p)\sigma$$

The quantile function then gives us probability curves for TLE, just select 'p'

Estimation of parameters, β_0, β_1, and σ is based on the likelihood function

$$likelihood(\beta_0,\beta_1,\sigma) = \prod_{i=1}^{n} \frac{1}{\sigma}\phi(\frac{y_i-\mu_i}{\sigma})$$

Example:

- Start with a data set of TLE values taken from one or more flights
- Thin the data as necessary to eliminate 'point-to-point' correlation of observations
- Model the cumulative distribution of the errors as a log-normal cumulative distribution function: ($\Phi_{\ln}$ the log-normal distribution function)

$$\Pr(D \le d) = \Phi_{\ln}((d - \mu(d))/\sigma)$$

  - μ is a function of d; $\mu_i = \beta_0 + \beta_1 d_i$
- Use maximum likelihood to get estimates of β_0, β_1, and σ

Maximum likelihood

- Find values of b_0, b_1, and σ so that the probability

$$\log(likelihood) = p(b_o, b_1, \sigma \mid d) = \prod_{i=1}^{n} \frac{1}{\sigma} \phi\left(\frac{d_i - \mu_i}{\sigma}\right)$$

is a maximum

- Use logs to find the maximum, i.e. apply Python function ‘minimize’ to –log(likelihood)
- Python function *minimize* from scipy.optimize
- Use the MLE estmates b_0*, b_1*, and σ* to get TLE as a function or range to the target

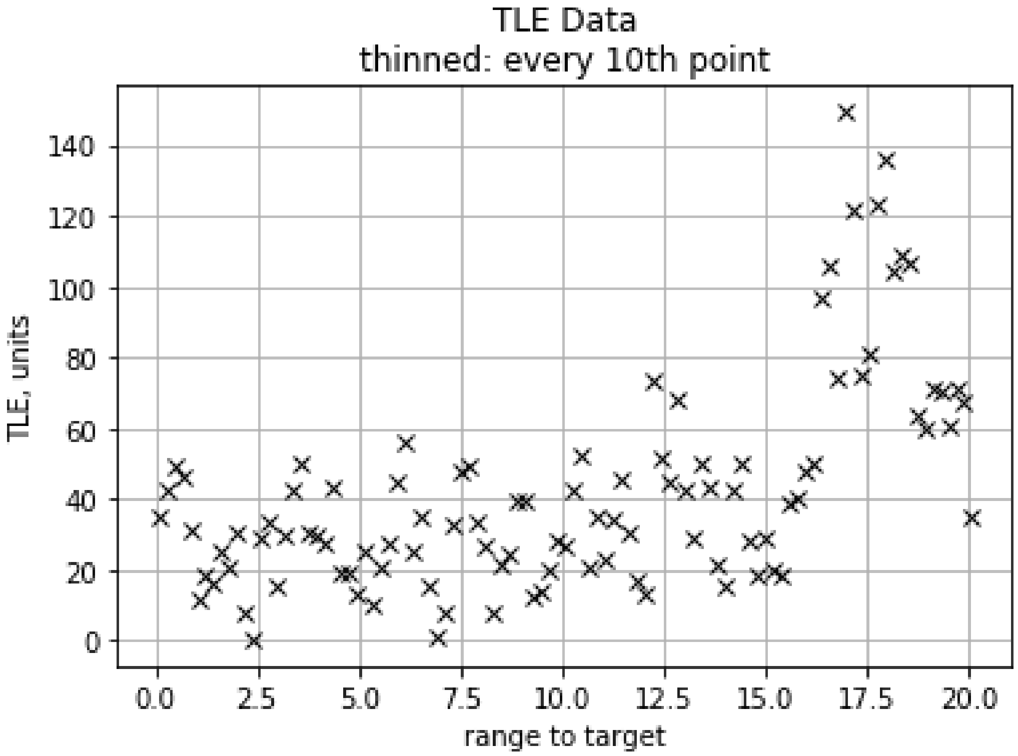
What it looks like in Python

```
def logLikelihood(b, density, distribution, r, X, delta):
    sig = b[0]
    logLike = 0
    mu = np.dot(X, b[1:])
    for i in range(len(r)):
        z=(np.log(r[i])-theta[i])/sig
        if delta[i]: logLike -= float( np.log(density(z) )-
np.log(sig)-np.log(r[i])) # delta == 0 => observed value
                              # delta != 0 => truncated value
        else: logLike -= float(np.log(1- distribution(z)))
    return logLike
```

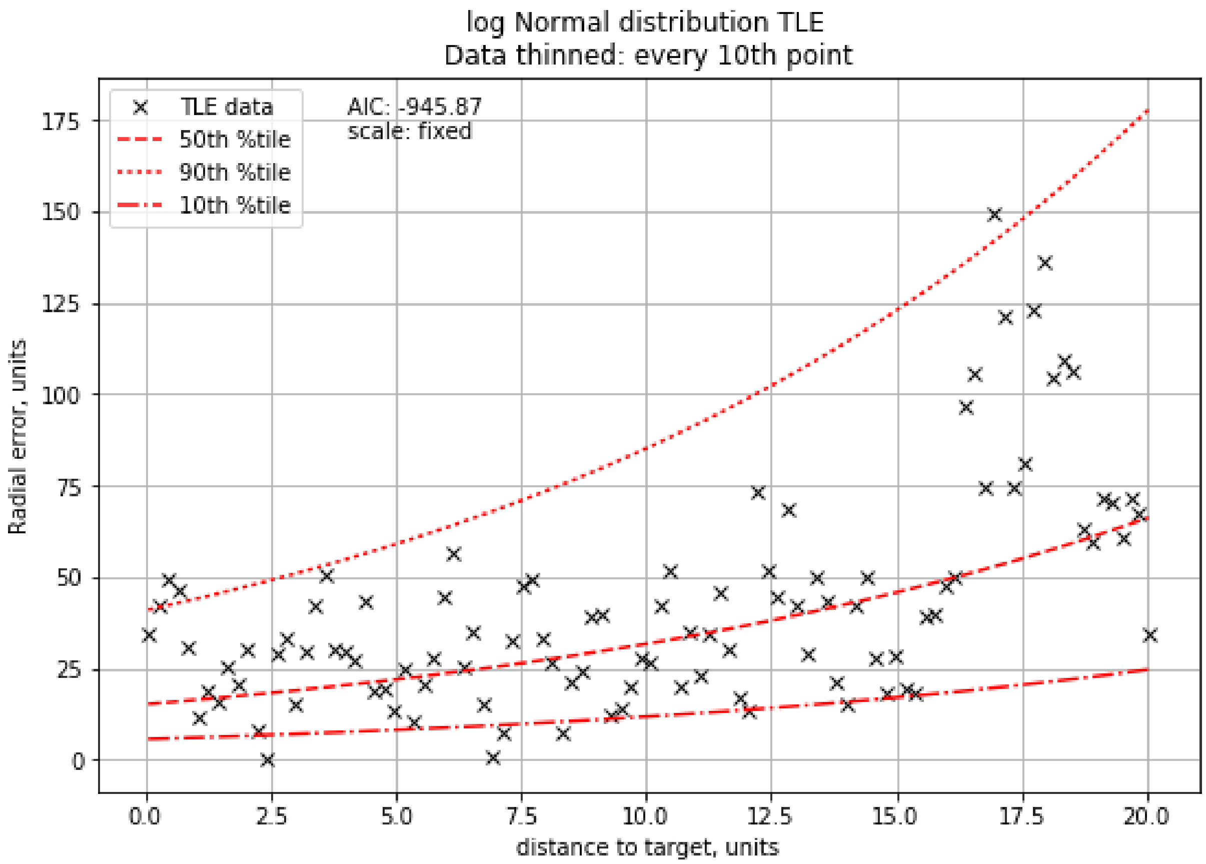
Plot original data, TLEs

- Plot(distance, r)
- Get the median value from the fit distribution, Φ:
  - Model: $mu = b_0 + b_1*d$
  - $CEP = exp(mu+\Phi^{-1}(0.5))*scale$
  - Similar for CE90, CE10
- Plot CE10, CEP, CE90 curves with the data

Original data, thinned



Fit to Log normal distribution



public release: 412TW-PA-17-0114

Estimates!

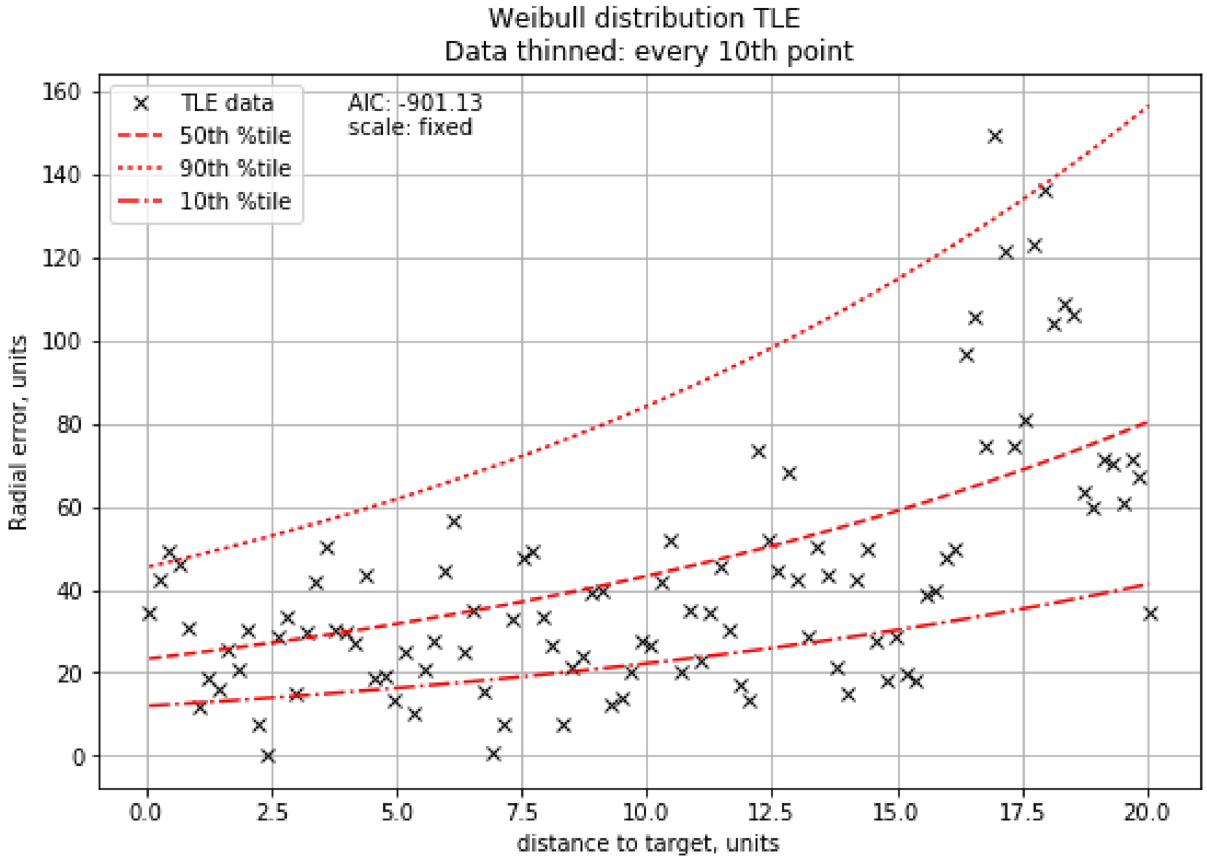
- In the previous slide, the x's represent the (thinned) data
  - MLE assumed observations were independent
  - Used partial autocorrelation function to find the thinning factor
- The red dashed lines show the 90^{th}, 50^{th} (CEP) and 10^{th} percentile estimates at each range.
- The AIC is given; different cumulative distribution functions can be used: select as 'best' among them the one with lowest AIC

Log Normal Fit

parameter		std error	t-value	p-value
Sigma:	0.77	0.023	33.6	0.0
b_0:	2.72	0.044	62.4	0.0
b_1:	0.07	0.008	9.38	0.0

AIC: -945.9

Weibull distribution



public release: 412TW-PA-17-0114

Weibull fit

parameter		std error	t-value	p-value
Sigma:	0.52	0.040	13.00	0.0
b_0:	3.15	0.11	28.58	0.0
b_1:	0.06	0.01	7.39	0.0

AIC: -901.13

AIC: log-normal a better fit than Weibull

What about non-constant scale, σ?

- Model both μ and σ as dependent on one (or more) explanatory variable(s)
- Again, $p(R \le r) = \Phi_{\ln}((\log(r) - \mu)/\sigma)$
- where

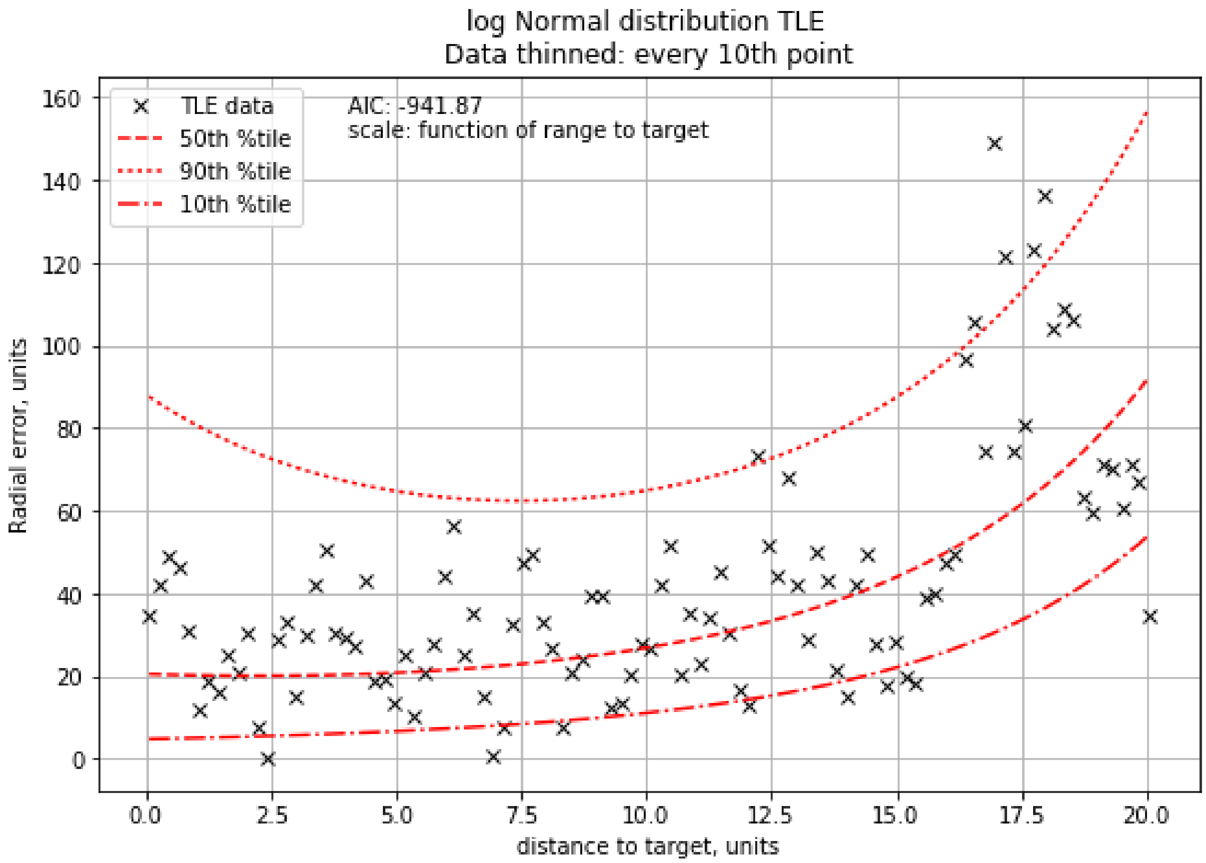
$$\mu = \beta_0^{[\mu]} + \beta_1^{[\mu]} x + \beta_2^{[\mu]} x^2, \ and$$

$$\log(\sigma) = \beta_0^{[\sigma]} + \beta_1^{[\sigma]} x$$

- Log quantile, to get CEP, CE90, is then

$$r_p(x) = \beta_0^{[\mu]} + \beta_1^{[\mu]} x + \beta_2^{[\mu]} x^2 + \Phi^{-1}(p)\exp(\beta_0^{[\sigma]} + \beta_1^{[\sigma]} x)$$

Log normal, quadratic fit



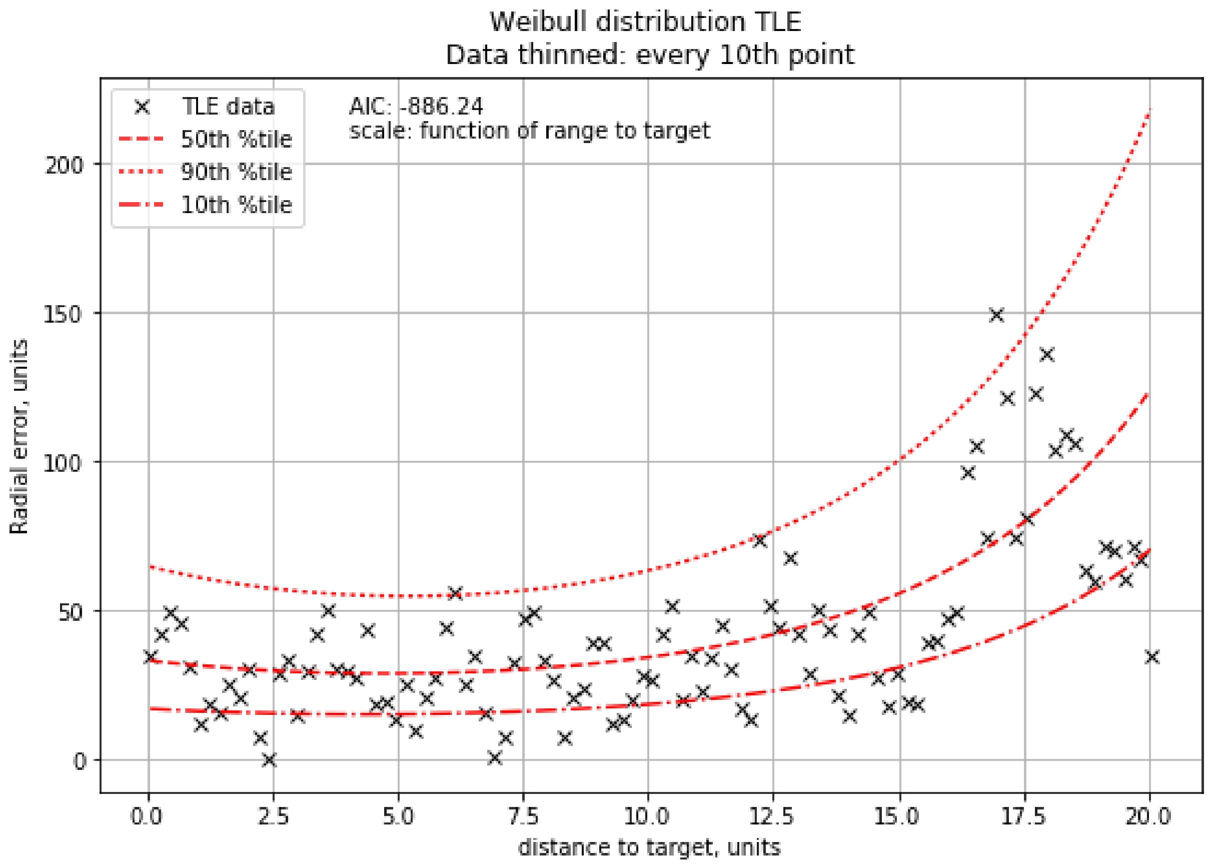
public release: 412TW-PA-17-0114

Log-normal, quadratic fit

parameter		std error	t-value	p-value
$b_0^{[\mu]}$:	3.02	0.28	10.69	0.00
$b_1^{[\mu]}$:	-0.02	0.05	-0.40	0.66
$b_2^{[\mu]}$:	0.005	0.002	2.16	0.02
$b_0^{[\sigma]}$:	0.13	0.14	0.92	0.18
$b_1^{[\sigma]}$:	-0.05	0.012	-4.26	0.00

AIC: -941.9

Weibull, quadratic fit



public release: 412TW-PA-17-0114

Weibull, quadratic fit

parameter	std error	t-value	p-value
$b_0^{[\mu]}$: 3.51	0.14	24.795	0.0
$b_1^{[\mu]}$: -0.059	0.032	-1.85	0.97
$b_2^{[\mu]}$: 0.0062	0.002	4.04	0.00
$b_0^{[\sigma]}$: -0.65	0.15	-4.29	0.0
$b_1^{[\sigma]}$: -0.0084	0.013	-0.661	0.74

AIC: -886.24

Conclusions?

- AIC values identify the log-normal fit, with fixed sigma as the best model
- For quadratic fit, the linear term (in both log-normal and Weibull) is not significantly different from zero
- Any reason to add other distributions?
  - Log-logistic, Gumbel first asymptote (largest extreme value),
  - Use AIC to decide, or to weight model results for a final estimate of CE values

Questions??

- Slides will be posted on sharepoint!

- References:
- Meeker and Escobar, *Statistical Methods for Reliability*, John Wiley & Sons, New York, 1998

- Leonhard Held, *Methoden der staistischen Inference-Likelihood und Bayes,* Spektrum Akademischer Verlag Heidelberg, 2008