Title:	Example for Approximating European loop distribution using Rayleigh Distribution
Source:	TIA TR30.3 Committee
Author:	Jack Douglass
Contact:	Jack Douglass Chairman TIA TR30.3 Voice: 949-233-3558 Email: jackdouglass@cox.net
Status:	For Information

Abstract

This contribution proposes a example method for developing a family of statistically based loop models using Rayleigh Distribution for approximating loop probability. A family of PDF and CDF Loop Distribution curves can be used to represent the range European Loops lengths, without being country specific. This model can be applied to PD TM6 (02)07, Access Transmission Network Model for Evaluating xDSL Modem Performance.

1. Introduction

Creating a Network model using the Network Model Coverage (NMC) methodology requires that the Likelihoods-of-Occurrence (LOO) for different loop lengths be known. It would be desirable to have a general crosstalk model (not country specific) that represents the range of loop lengths that can be experienced in Europe. The range of loop lengths could be represented by a Family of equally spaced PDF and CDF curves that spans the range of probabilities. Having such a model would allow the user of the model to evaluate DSL modem performance over the full range of European loops without having to know the loop length distribution for specific country. But, on the other hand, an operating company, which knows the loop distribution for their country can select an appropriate PDF and CDF curve that closely matches loop distribution for that country. A non-country specific loop model for Europe is needed for the main body of PD TM6 (02)07, Access Transmission Network Model for Evaluating xDSL Modem Performance.

This contribution is not intended to be final solution for modelling the European loop distribution. It is intended to generate discussion on methods for creating a loop distribution model.

2. Procedure

1. PDF and CDF distributions are calculated for a fictitious European Country, where 80% of the loops are ≤3km.
2. Calculated PDF and CDF distributions are compared to Rayleigh distribution (Weibull with Shape=2)
3. PDF and CDF distributions are plotted for North American data. A family or range of CDF curves can be observed.
4. PDFs and CDFs are plotted for a Family of Example European Distributions with a shape=2 and mean loop lengths between 2.5km and 3.75km. Knowing the mean loop length for a particular country the approximate PDF and CDF can be determined from the curves and associated Weibull distribution data.

Note: It is understood that PDF and CDF curves are derived from discrete values and are typically plotted as bar charts. However, in this contribution all PDF and CDF distribution are plotted as line graphs so that data can be easily seen.

2.1. Selection of Distribution Function

This contribution proposes an example method for developing a family of statistically based loop models using Rayleigh Distribution for approximating loop probability. Rayleigh Distribution is selected as an example distribution function that could be used for modelling loop distribution. There may be other distributions function that more suited for modelling European loop distribution. We welcome any input on a distribution functions that would be more suited.

2.1.1. Rayleigh distribution:

A Rayleigh distribution is typically used when both random processes x(t) and y(t) are Gaussian distributed with the same variance and zero mean, assuming x and y are statistically independent. Examples where Rayleigh distributions are used: Modem constellations, scattering of radiation and wind speeds.

2.1.2. Weibull Distribution

The Weibull distribution has a relatively simple distributional form. However, the shape parameter allows the Weibull to assume a wide variety of shapes. The shapes include an Exponential distribution (Shape =1), a Rayleigh Distribution (Shape=2), and a relatively symmetric distribution (Shape=5). A location parameter allows the graph to be shifted left or right on the horizontal axis. This combination of simplicity and flexibility in the shape make it suitable for many distribution applications.

Probability Density Function for Weibull Distribution
f(x,a,b) = a/b(x^(a-1)e^(-(x/b)^a)

Cumulative Distribution function for Weibull Distribution
F(x,a,b)=1-e^((x/b)^a)

2.2. Calculated an example for Europe Loop Distribution where 80% of loops are ≤3km and curve

PDF (Figures 8) and CDF (Figure 9) curves were created for a fictitious European Country, where 80% of the loops are ≤3km, using the following procedure.

1. Created X-Y coordinates for post-codes
2. The number of residences in each post-code equals the number of subscribers in each area
3. Determined the X-Y coordinated of central offices
4. Calculated the straight line distance between the CO and the subscriber. Then multiplied by a factor of 1.5, because cables are not connected by shortest method.

PDF and CDF distributions from this procedure are plotted and compared to the Rayleigh Distribution (Weibull distribution with shape = 2 and location=65% value on the calculated CDF curve). See Figures 1 and 2. The Rayleigh Distribution closely matches the calculated CDF and PDF. It would be nice to validate this calculated PDF with anonymous data from European operating companies. We welcome this kind of input from the operating companies.

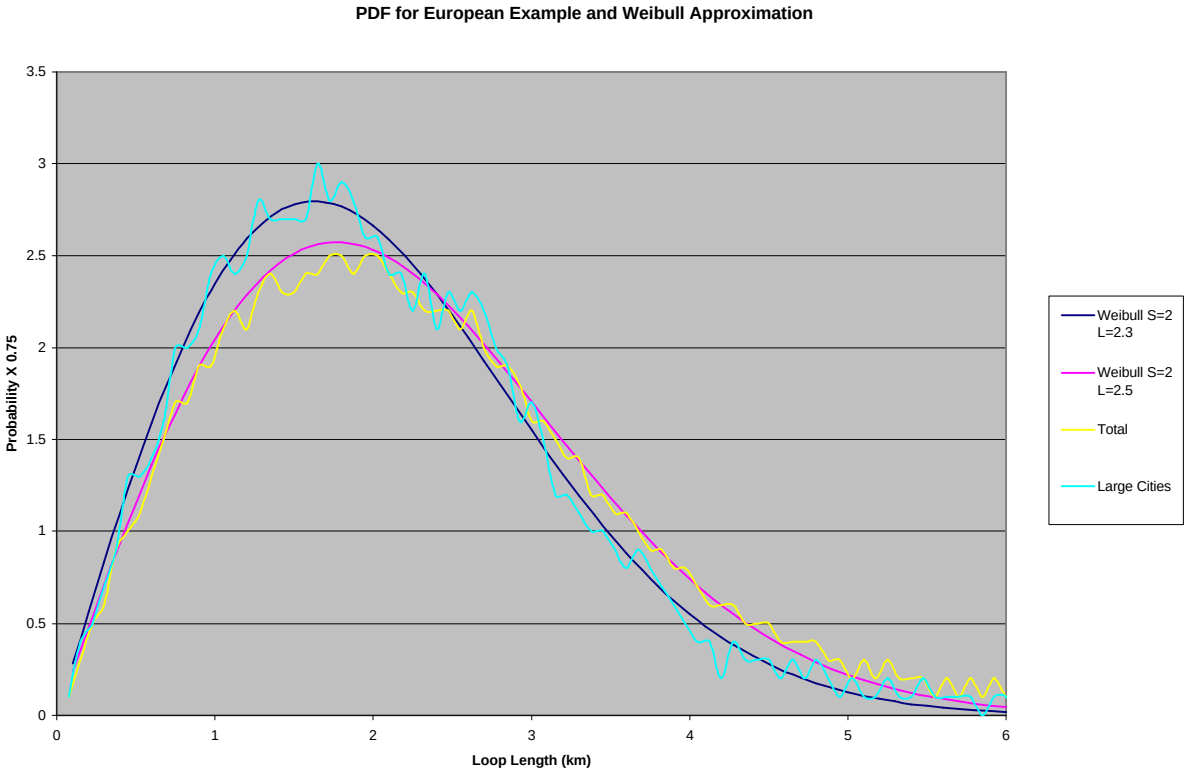

Figure 1 –Example of European Loop PDF with 80% of loop ≤ 3km and Weibull Distribution

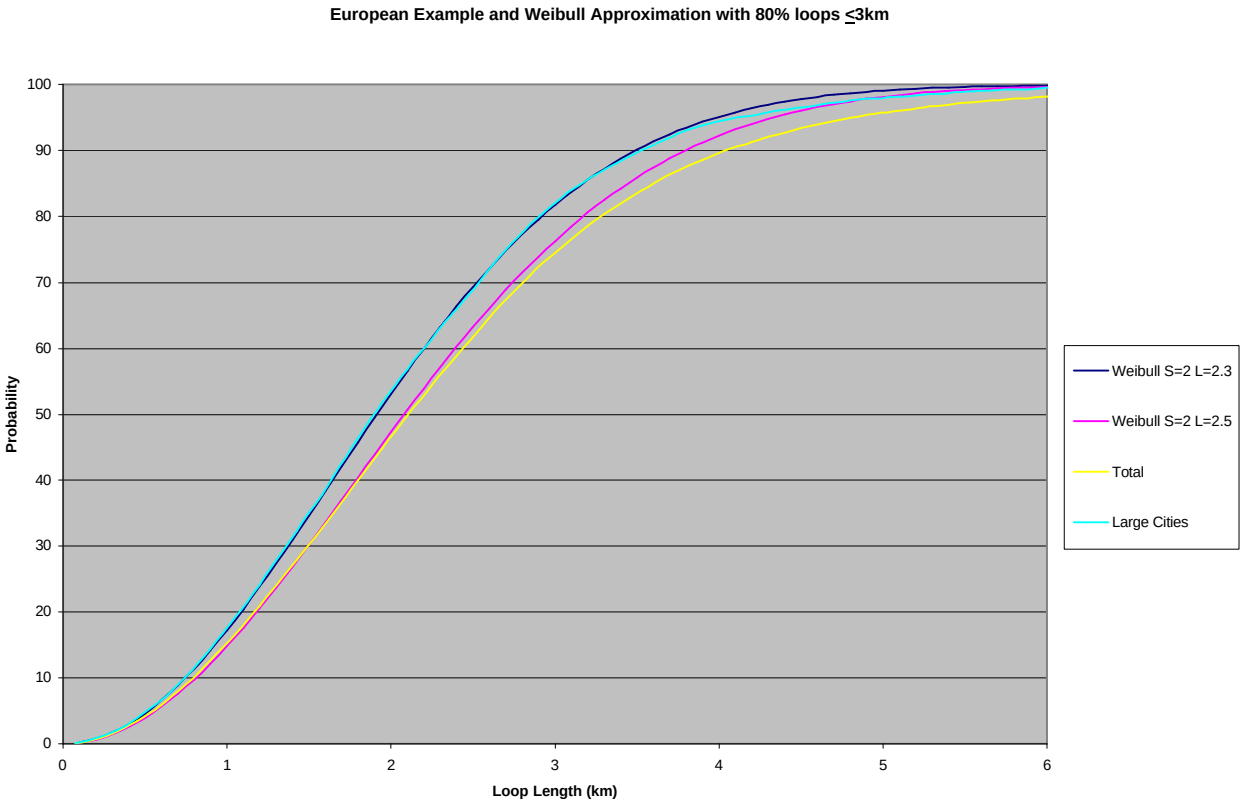

Figure 2 – Example of European Loop CDF with 80% of loops < 3km

2.3. North American Loop Distribution

Loop distribution data for over 14 million North American lines was received from multiple anonymous operating companies and Telecodia. PDF and CDF distributions for loop lengths less than 5.5km are given in Figure 3 and Figure 4. A family or range of CDF Distributions for North America can be observed in Figure 4.

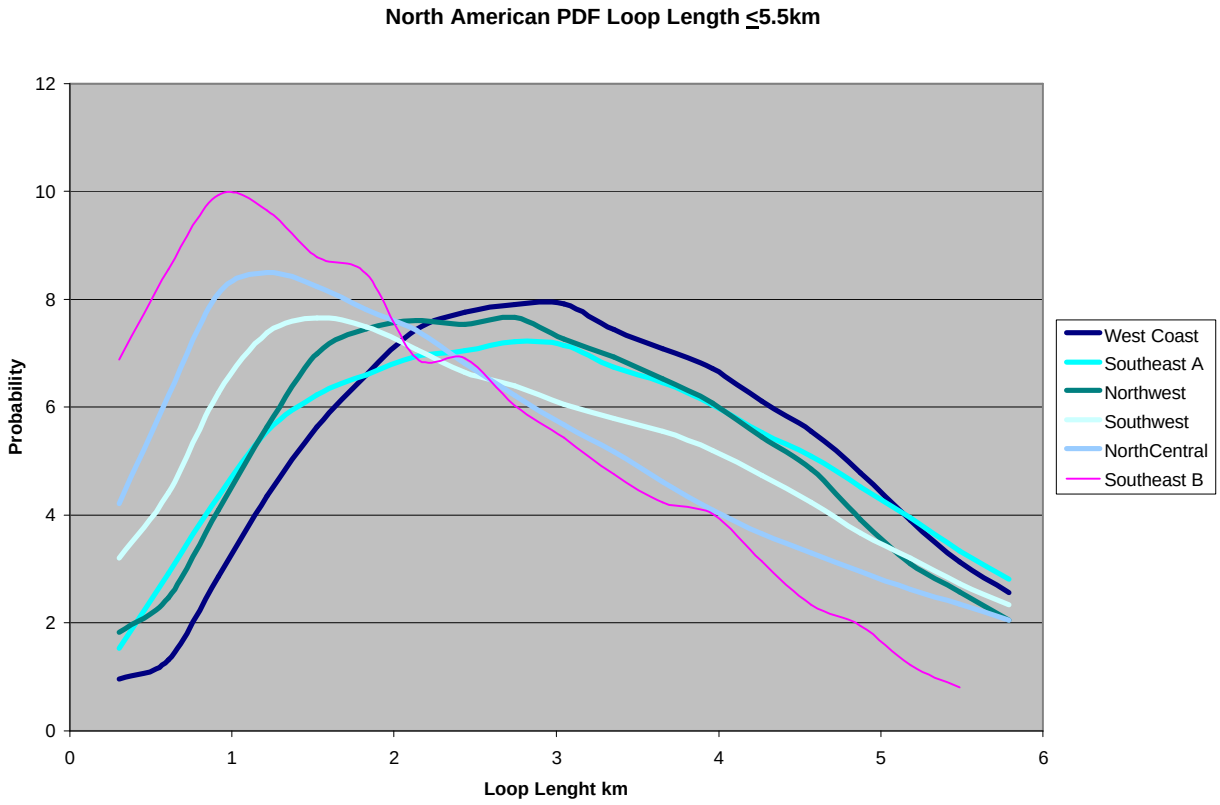

Figure 3 –PDF of North American Loop Data (Truncated to 5.5km)

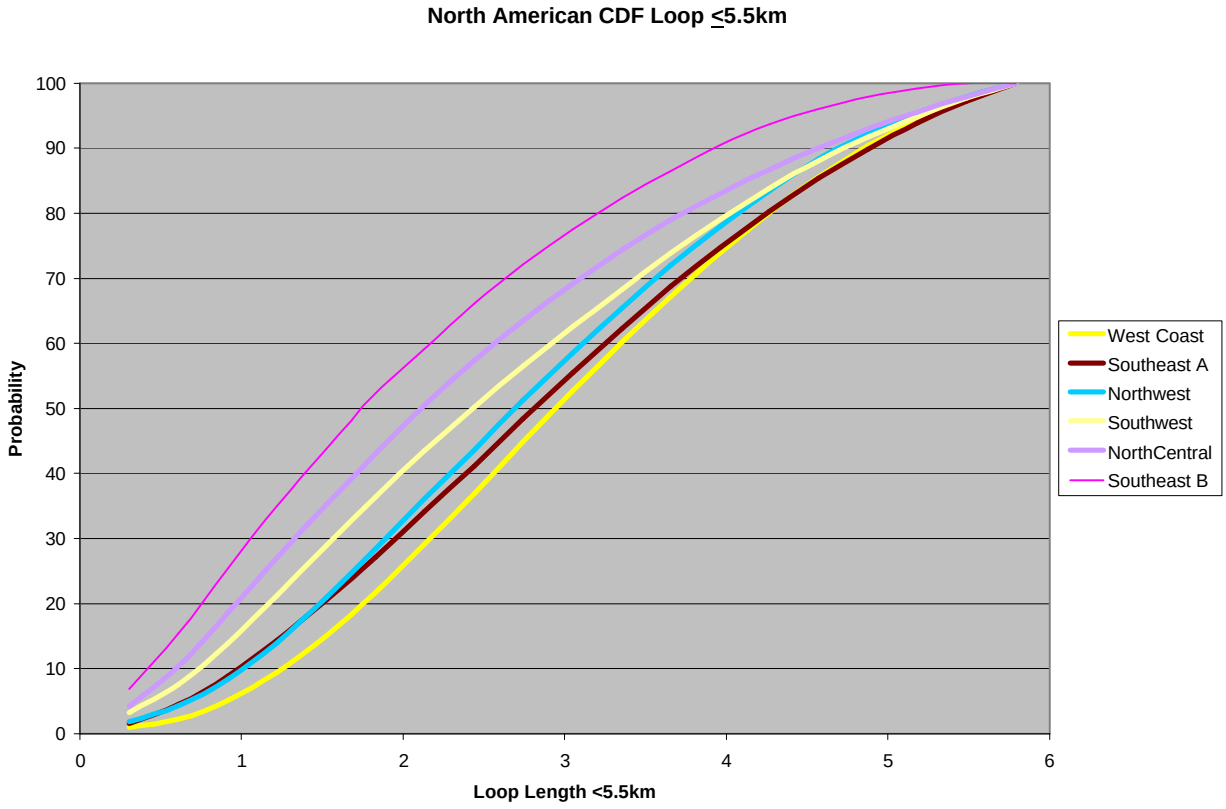

Figure 4 -- North American loop length CDF (Truncated to 5.5km) – A Family or Range of distributions can be observed.

2.4. Family or range of PDFs and CDFs based on example for Europe Loop Distribution

A Family or range of PDF (Figure 5) and CDF (Figure 6) curves can be created by setting the shape=2 and plotting curves for length values between 2.5km and 3.75km. Knowing the mean loop length for a particular country the approximate PDF and CDF can be determined from the curves and associated Weibull distribution data. It would be nice to validate this calculated PDF with anonymous data from European operating companies. We welcome this kind of input from the operating companies. Likelihoods of occurrence values that are needed in PD TM6 (02)07 can be selected 50, 70, 85 and 95 values (Black Line) points on CDF curve.

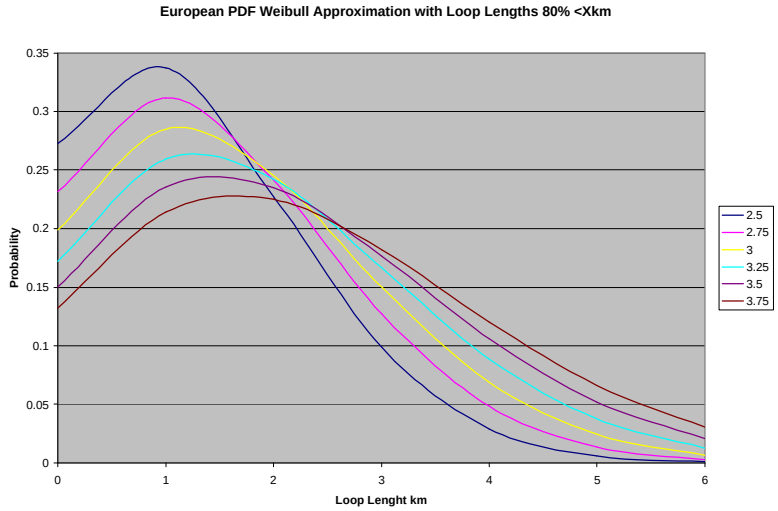

Figure 5 -- Family of European Loop PDF curves for loop lengths between 2.5km to 3.5km (80% loop length) using Weibull Approximation

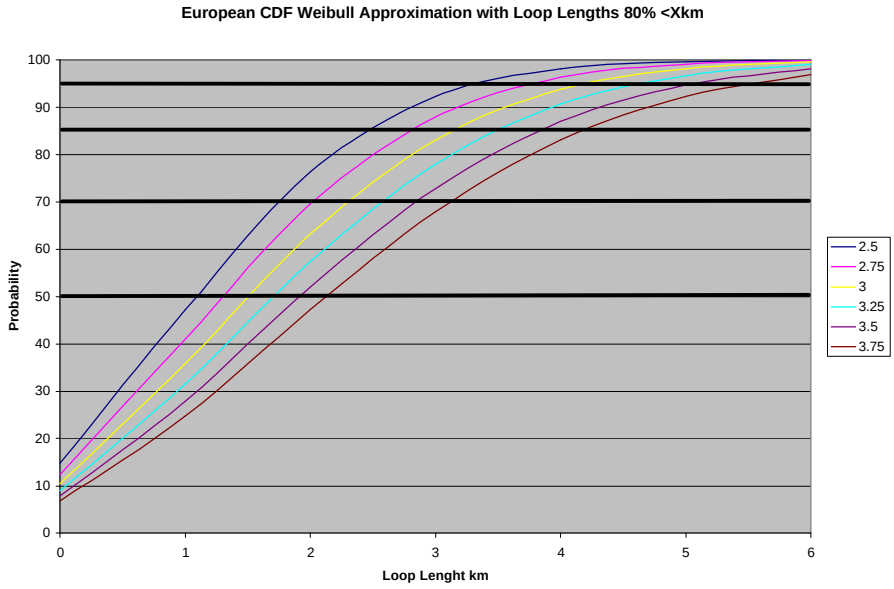

Figure 6 – Family of European Loop CDF curves for loop lengths between 2.5km to 3.75km using Weibull Approximation. CDF for the Network Model can be selected from the graph for the 50%, 70%, 85% and 90% Bold line

3. Open questions

- How closely do the PDF and CDF curves match the actual loop distributions in Europe?

- Is there a better distribution function that can be used to approximate loop distribution?
- Can a family of curves be used to represent the range of loop distributions in Europe without specifying loop distribution information for specific countries.
- Are there any additional factors that should be considered when approximating loop distributions?

4. Summary

This contribution is not intended to be final solution for modelling the European loop distribution. It is intended to generate discussion on methods for creating a loop distribution model. If this approach is acceptable an equation can be used to represent the loop distributions. A statistically based non-country specific European Loop Model (Family of CDF Curves) can be created using the procedure outlined in this document as more anonymous loop data is submitted. We propose that operators compare the family of CDF curves with actual loop distribution data for their country. We welcome any input on making the PDF and CDF curves more accurate. Rayleigh Distributions was selected as an example distribution function that could be used for modelling loop distribution. There may be other distributions function that are more suited for modelling European loop distribution. If a better distribution function is know we welcome contributions in this area.

References:

NIST/SEMATECH e-Handbook of Statistical Methods, http://www.itl.nist.gov/div898/handbook/index.htm

Simple Interactive Statistical Analysis, http://home.clara.net/sisa/weibhlp.htm

The Mathworks Statistic Tool box, http://www.mathworks.com/access/helpdesk/help/toolbox/stats/prob_d24.shtml

Swinburen Astronomy on line, http://astronomy.swin.edu.au/~pbourke/analysis/distributions/#rayleigh