Comparing the results of transmuted Gumbel probability distributions found in literature with those obtained by the free CumFreq program

R.J. Oosterbaan
September 2022 website: www.waterlog.info

Abstract
Farwah Willayat and Doaa Elhertaniy have given data to which they have fitted different upgraded Gumbel probability distributions like the Marshall-Olkin extended Gumbel type II and the quartic rank transmuted Gumbel distributions. In this paper these data are used in the free CumFreq program with the option to find the best fitting of all distribution dealt with in this software and compare the results. One reason for this operation is to demonstrate that CumFreq is easy to operate and all computations are done automatically. Another reason is to try to find even better fitting distributions than those presented by the authors mentioned before while the number of parameters is less, which enhances the robustness.

Specifications of references in abstract

1. Farwah Willayat et al. (2022)
   Marshall–Olkin Extended Gumbel Type-II Distribution: Properties and Applications
2. Doaa ELhertaniy (2022)
   Quartic Rank Transmuted Gumbel Distribution
3. R.J.Oosterbaan, CumFreq
   https://www.waterlog.info/cumfreq.htm

Contents

1. Brief introduction to CumFreq
2. Breaking stress of carbon fibers
3. Service times of aircraft windshields
4. Fracture toughness of alumina
5. Glass strength
6. Maximum monthly wind speed
7. Conclusion
8. References
9. Appendices
   A1. CumFreq input menu
   A2. Selection options in CumFreq
   A3. Overview of advanced probability distributions applied to glass strength data

1. Brief introduction to CumFreq

After entering the data in the input menu (for example by copying from excel and pasting, see Appendix A1)), one can select a preferred probability distribution (see Appendix A2) or choose "*best of all*". Clicking on "Save-Run" CumFreq will then perform the probability distribution fitting automatically, show the results, and provide the opportunity to see graphs.

For more detail consult R.J. Oosterbaan (2022) How to derive a probability distribution from a data set using the simple method of plotting positions and the free CumFreq model

2. Breaking stress of carbon fibers (data from Farwah Willayat et al.)

According to CumFreq the best of all distributions for the carbon data is the Rayleigh cumulative probability distribution function (CDF):

CDF = 1 – exp (X^P/A) with P = 2.76 and A = 20.19

This is a simple expression with only 2 parameters, yet very efficient.

Graphics are shown on the next page.

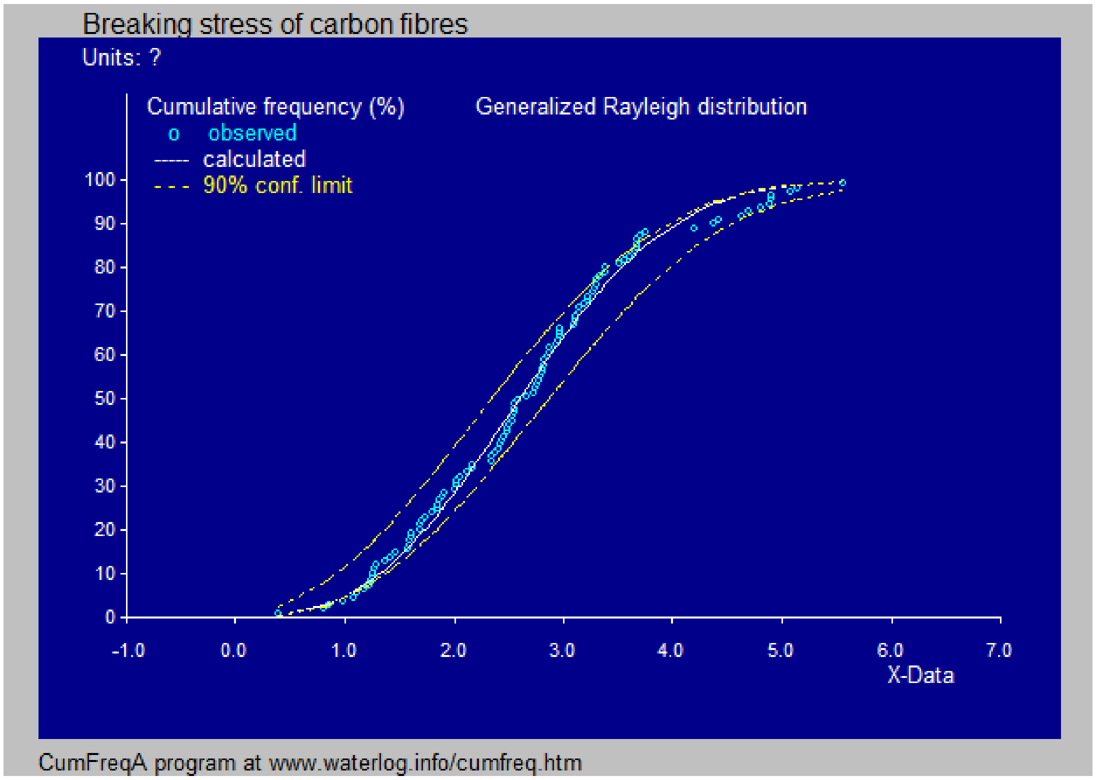


Figure 1.

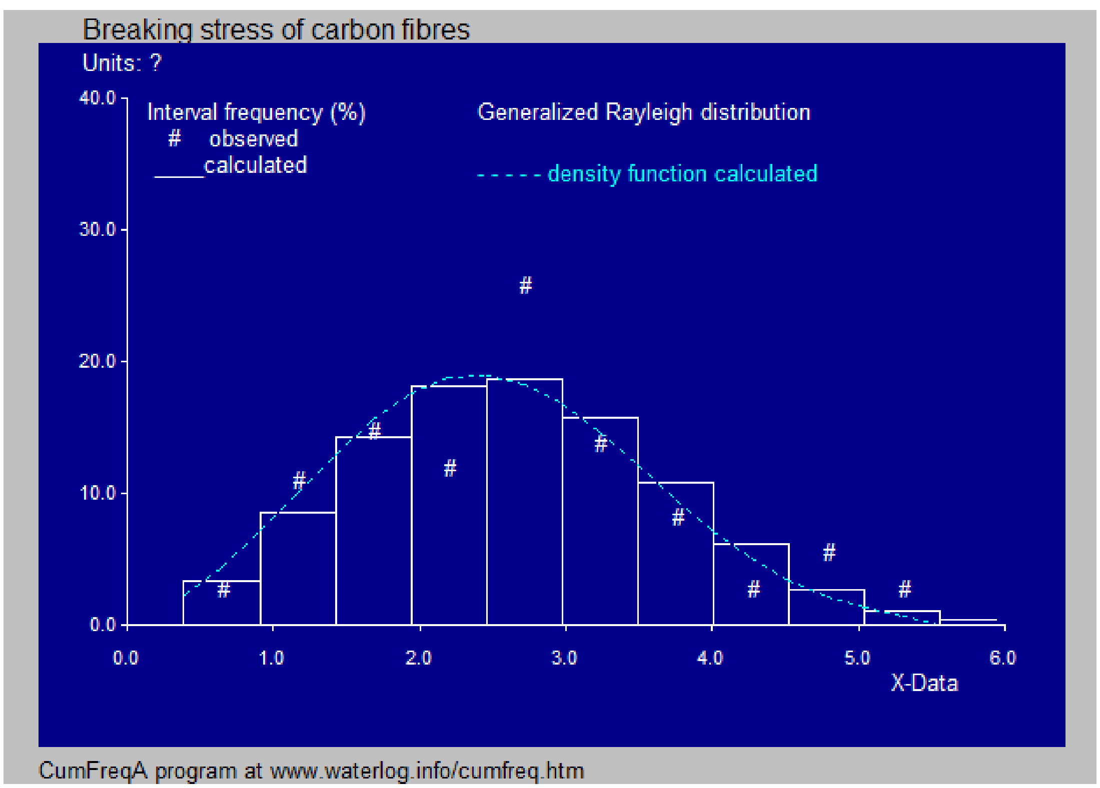
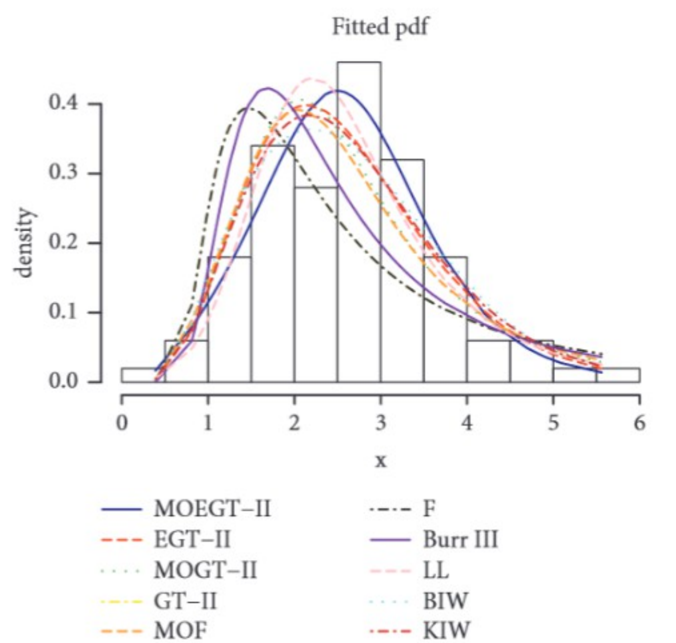


Figure 2.



__Comparison__

The name of the principle model used by Farwah Willayat et al. is abbreviated to MOEGT-II. It uses 4 parameters.

Their probability density function (pdf) is represented by the blue curve in the figure to the left. It looks quite similar to the density curve for the generalized Rayleigh distribution in the previous figure.

However, the Rayleigh model uses only 2 parameters. Therefore it might be given preference

Figure 3.

The CumFreq ranking according to goodness of fit of the first nine CDF's for carbon stress is:

Distribution name *)	Average of absolute differences between calculated and observed cumulative frequencies #)
1. Rayleigh generalized	1.85 %
2. Logistic generalized	1.89 %
3. Burr generalized	1.92 %
4. Gumbel generalized	1.94 %
5. Fisher-Tippett type 3	1.98 %
6. Weibull	2.07 %
7. Mirrored Frechet (Fisher-Tippet Type 2)	2.08 %
8. Kumaraswamy	2.09 %
9. Generalized exponential (Poisson-type)	2.22 %

*) Generalization means raising X to the power P
#) 0–1% excellent, 1–2% very good, 2–3% good, 3 – 4% medium, > 4 % poor

Farwah Willayat et al. use different indexes for goodness of fit, so it is difficult to compare them.

3. Service times of aircraft windshields (data from Farwah Willayat et al.)

According to CumFreq the best of all for the windshield data is the generalized logistic probability distribution function (CDF):

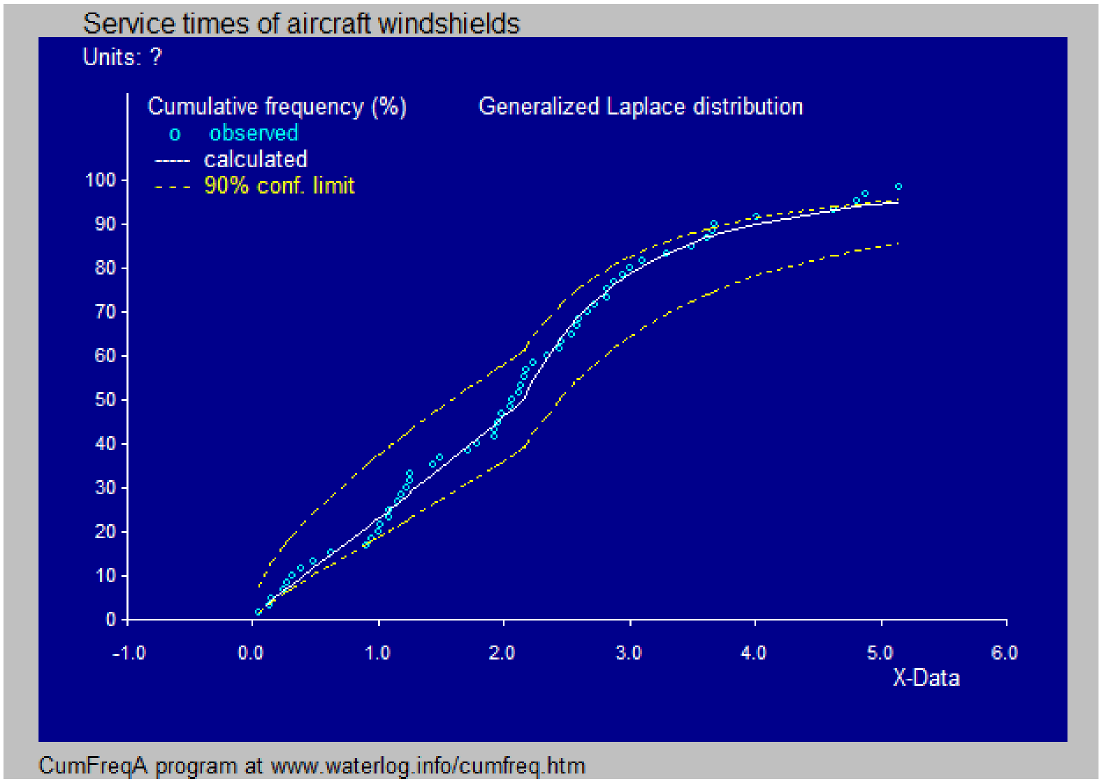
The cumulative frequency function is generalized logistic :

CDF = 1 / { 1+exp(A*X^P+B) }
The exponent P = 8.30E-001,   A   = -1.76E+000,   B   = 3.14E+000

Cumulative frequency of the generalized Laplace distribution :

X<B Freq = 0.5*exp{A1*(X^E1-B)}
X>B Freq = 1-0.5*exp{A2*(B-X^E2)}
B = 1.08E+000 A1 = 9.65E+000 A2 = 2.34E+001
The exponents are: E1 = 1.00E-001 and E2 = 1.00E-001

Graphics are shown on the next page.


Service times of aircraft windshields
Units: ?
Cumulative frequency (%)
Generalized Laplace distribution
o observed
----- calculated
- - - 90% conf. limit
100
90
80
70
60
50
40
30
20
10
0
-1.0
0.0
1.0
2.0
3.0
4.0
5.0
6.0
X-Data
CumFreqA program at www.waterlog.info/cumfreq.htm
Figure 4.

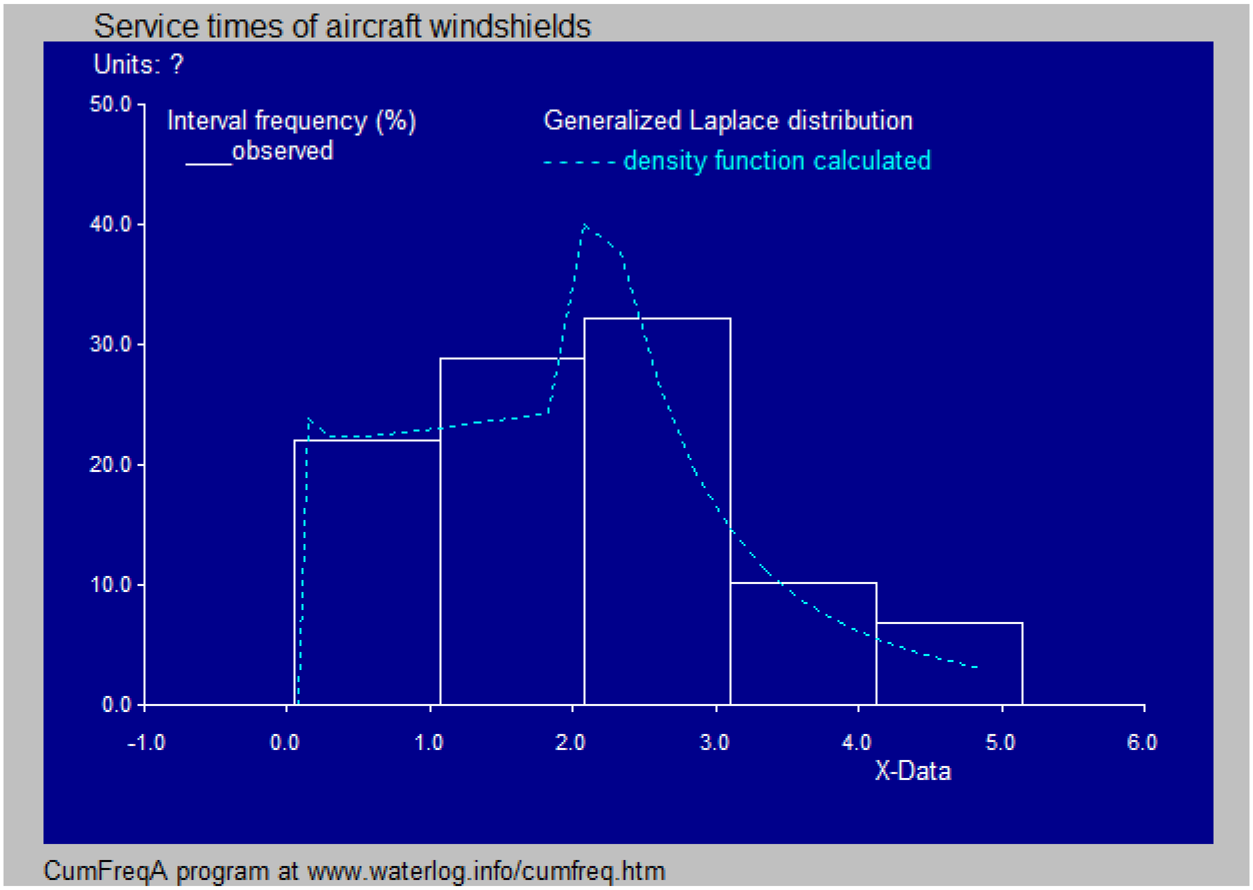
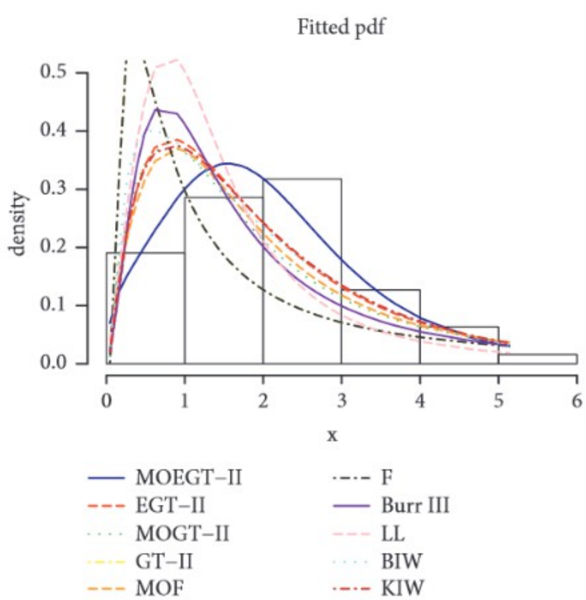


Figure 5.



Comparison

The name of the principle model used by Farwah Willayat et al. is abbreviated to MOEGT-II. It uses 4 parameters.

Their probability density function (pdf) is represented by the blue curve in the figure to the left. The Laplace distribution in the previous figure shows a peak that corresponds better with the peak of the histogram.

Hence, the Laplace distribution is preferable.

Figure 6a.

Noor A. Ibrahim et al. (2022) used the same windshield data for the application of the Weibull Burr type X (WBX) distribution with 4 parameters. Their result is shown in the next figure 6b (red CDF curve). This curve does not have such a good fit as the curve of the generalized Laplace distribution in figure 5 using only 3 parameters.

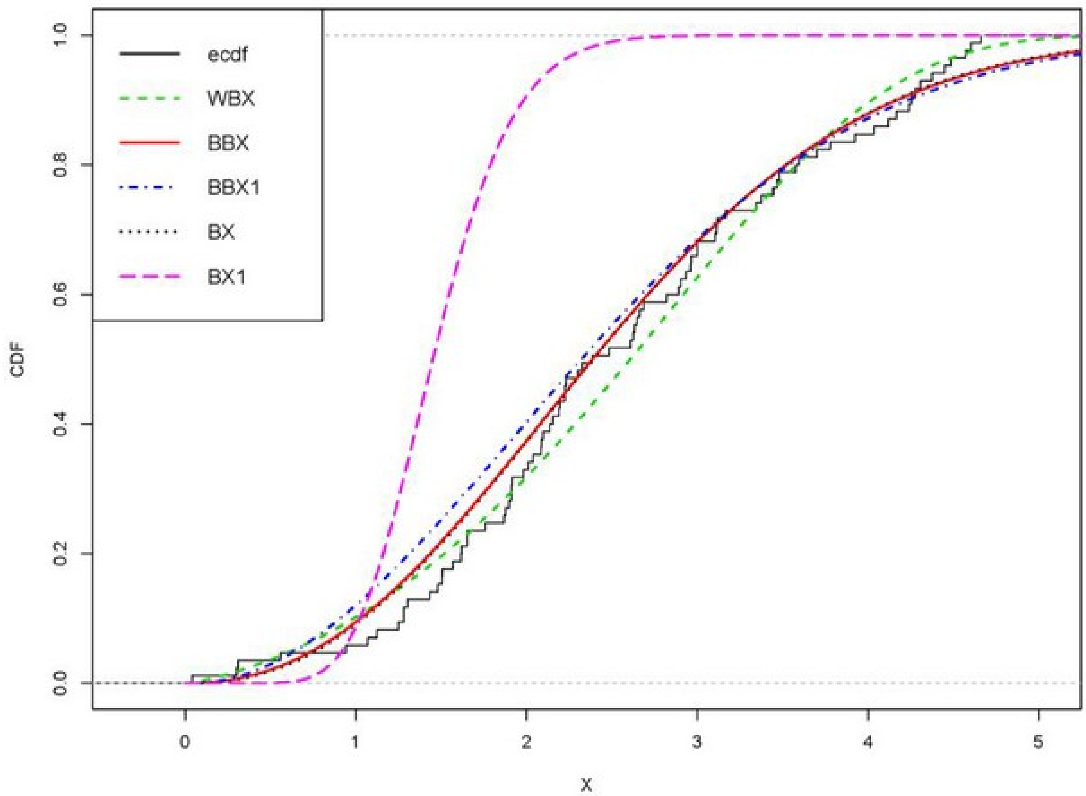


Figure 6b. ecdf = expected (observed) CDF

The CumFreq ranking according to goodness of fit of the first nine CDF's for service times is:

Distribution name *)	Average of absolute differences between calculated and observed cumulative frequencies #)
1. Laplace generalized	1.63 %
2. Logistic generalized	1.87 %
3. GEV (generalized extreme value)	1.91 %
4. Mirrored Frechet (Fisher-Tippet Type 2)	1.92 %
5. Fisher-Tippett type 3	1.94 %
6. Gumbel generalized	1.95 %
7. Mirrored Gumbel generalized	2.02 %
8. Generalized exponential (Poisson-type)	2.04 %
9. Burr generalized	2.08 %

*) Generalization means raising X to the power P

#) 0–1% excellent, 1–2% very good, 2–3% good, 3 – 4% medium, > 4 % poor

4. Fracture toughness of alumina (data from Farwah Willayat et al.)

According to CumFreq the best of all for the fracture toughness data is the generalized logistic cumulative probability distribution function (CDF):

$$CDF = 1 / \{1 + \exp(A*X^{P} + B)\}$$

The exponent P = 1.81E+000
A = –2.93E-001
B = 4.33E+000

Three parameters, more is not required.

Graphics are shown on the next page.

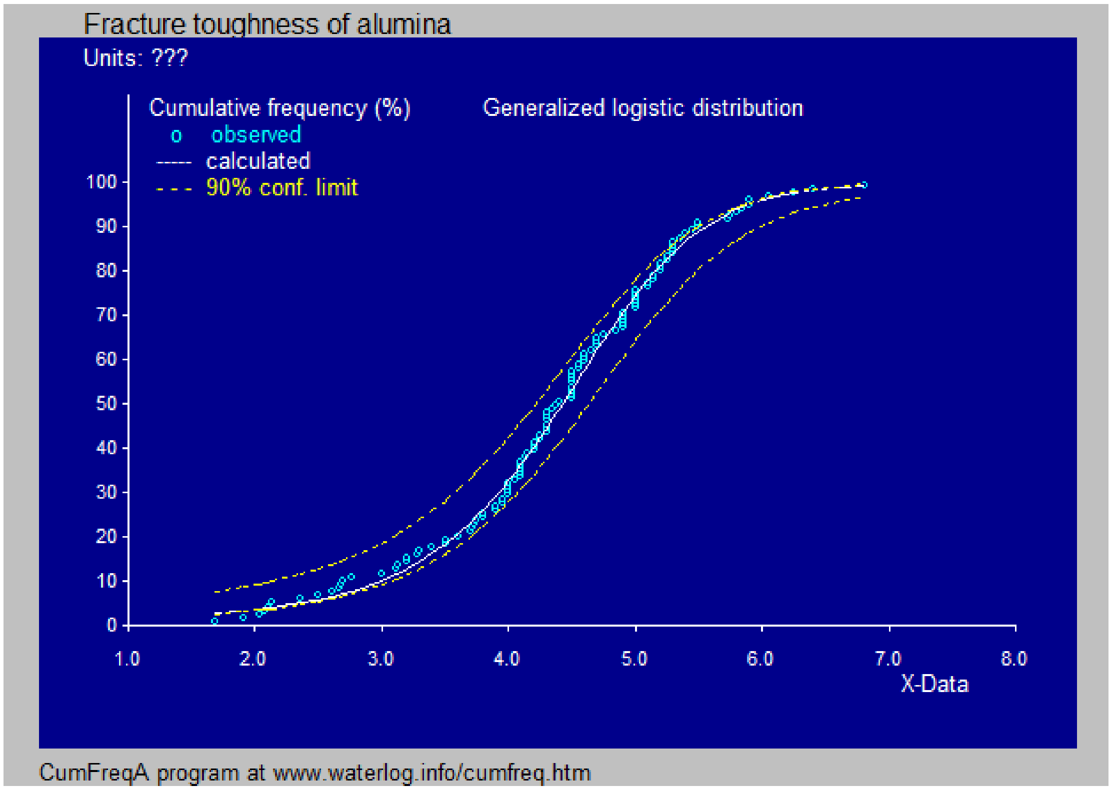


Figure 7.

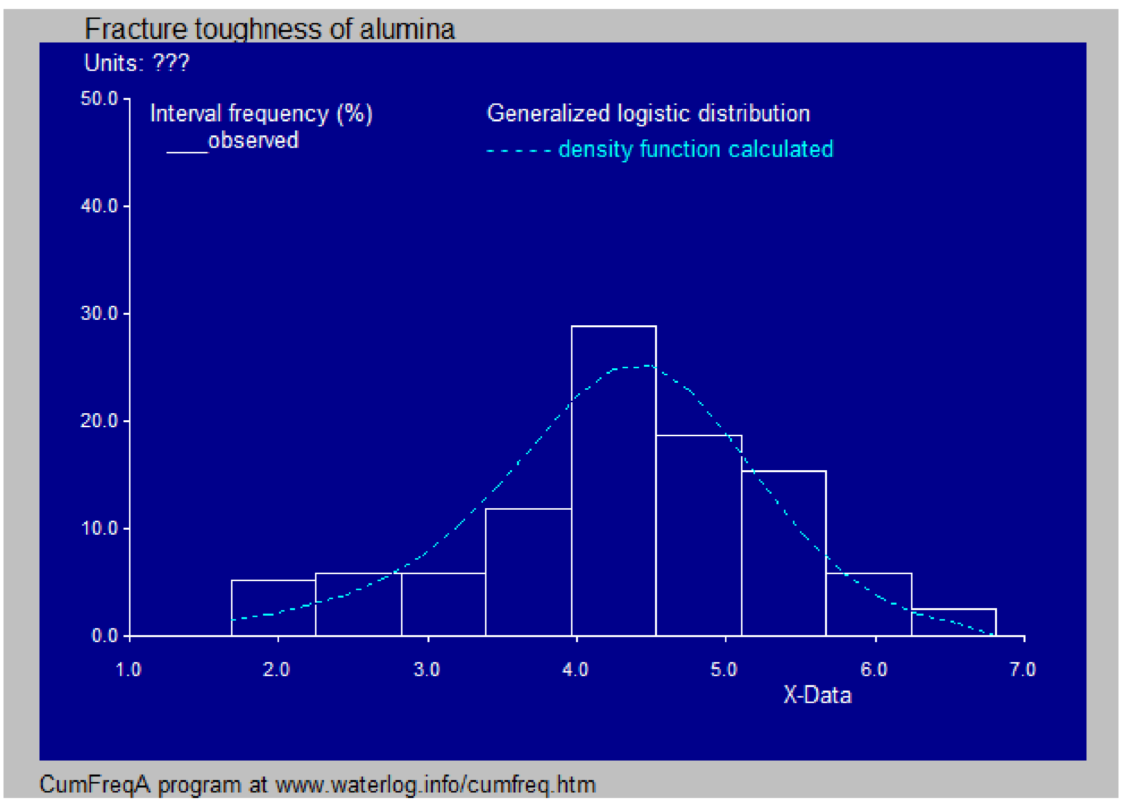
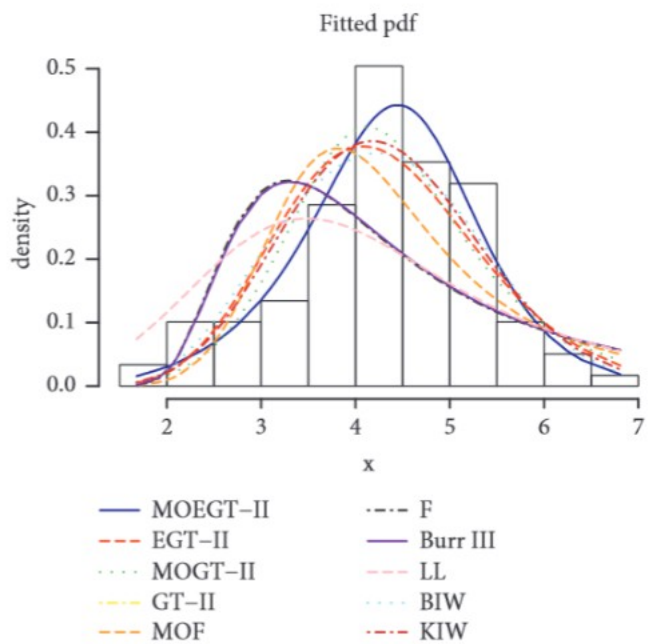


Figure 8.

.



Comparison

The name of the principle model used by Farwah Willayat et al. is abbreviated to MOEGT-II. It uses 4 parameters.

Their probability density function (pdf) is represented by the blue curve in the figure to the left. It looks quite similar to the density curve for the generalized logistic distribution in the previous figure.

However, the logistic model uses only 3 parameters. Therefore it might be given preference

Figure 9.

The CumFreq ranking according to goodness of fit of the first nine CDF's for the fracture toughness is:

Distribution name *)	Average of absolute differences between calculated and observed cumulative frequencies #)
1. Logistic generalized	1.42 %
2. Laplace generalized	1.85 %
3. Normal generalized	1.88 %
4. Mirrored Frechet (Fisher-Tippet Type 2)	1.90 %
5. Laplace standard	1.91 %
6. Mirrored Gumbel generalized	1.94 %
7. GEV (generalized extreme value)	1.99 %
8. Mirrored Gumbel standard	2.04 %
9. Mirrored Fisher-Tippett type 3	2.14 %

*) Generalization means raising X to the power P, mirrored means CDF = 1 – standard CDF
#) 0–1% excellent, 1–2% very good, 2–3% good, 3 – 4% medium, > 4 % poor

Farwah Willayat et al. use different indexes for goodness of fit, so it is difficult to compare them.

For the versatility of the generalized logistic distribution consult R.J. Oosterbaan (2022) [FITTING THE VERSATILE LINEARIZED, COMPOSITE, AND GENERALIZED LOGISTIC PROBABILITY DISTRIBUTION TO A DATA SET]

5. Glass strength (data from Doaa ELhertaniy)

According to CumFreq the best of all for the glass strength data is the standard (composite) Laplace cumulative probability distribution function (CDF):

Cumulative frequency of the composite Laplace distribution (standard) :

$X<B$ CDF = $0.5*\exp\{ A_1*(X-B) \}$
$X>B$ CDF = $1-0.5*\exp\{ A_2*(B-X) \}$

where:

B = 1.59E+000
A_1 = 2.91E+000
A_2 = 6.60E+000

Three parameters are sufficient.

Graphics are shown on the next page.

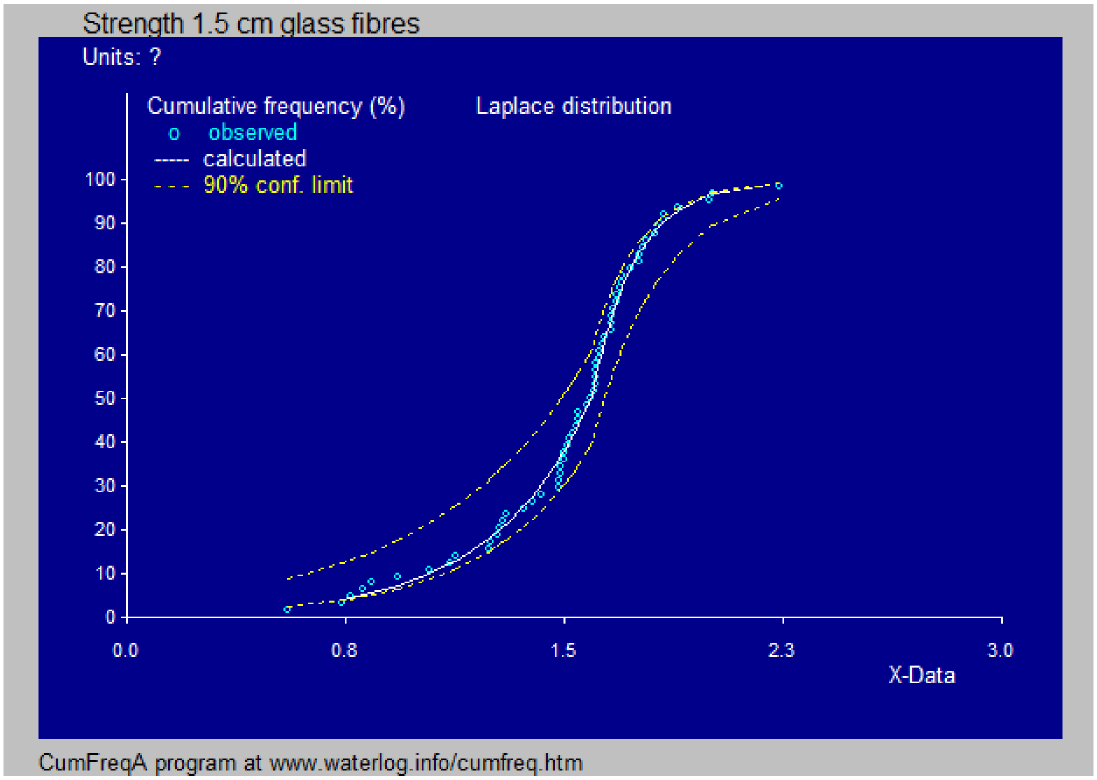


Figure 10.

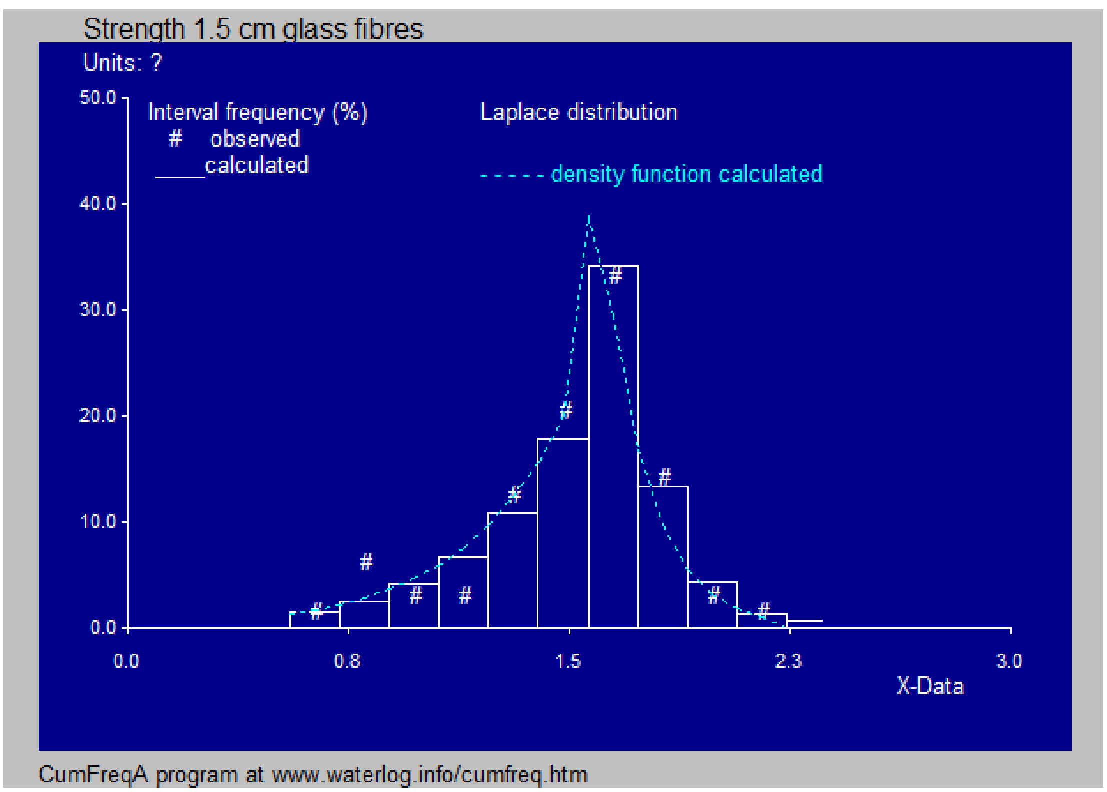


Figure 11.

The CumFreq ranking according to goodness of fit of the first two CDF's for the glass strength is:

Distribution name *)	Average of absolute differences between calculated and observed cumulative frequencies #)
1. Laplace standard	1.30 %
2. Logistic generalized	2.94 %

*) Generalization means raising X to the power P
#) 0–1% excellent, 1–2% very good, 2–3% good, 3 – 4% medium, > 4 % poor

Here, only one of the selectable distributions is good. For glass strength Doaa ELhertaniy (2022) presented the following figure 12 on the probability density function in which G = standard Gumbel distribution with 2 parameters, TG = quadratic rank transmuted Gumbel distribution with 3 parameters, CTG = cubic transmuted Gumbel distribution with 4 parameters, and QTG = quartic transmuted Gumbel distribution with 5 parameters. None of these, even the one with 5 parameters, shows as good a fit as the previous figure made with the standard Laplace distribution employing 3 parameters. This proves that it is difficult to find an applicable probability distribution for the glass strength data without a composition of two different distributions as the Laplace distribution actually provides.

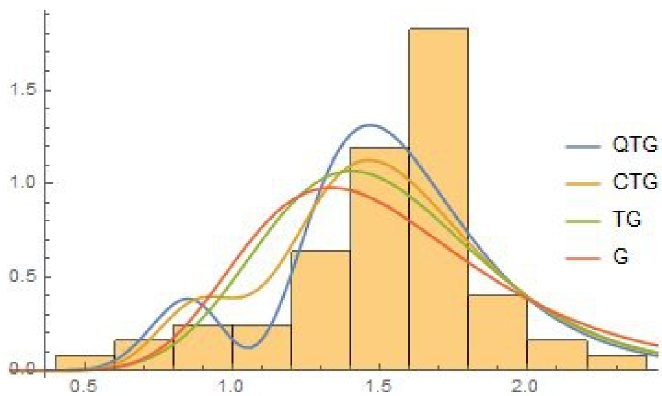


Figure 12.

There are various other cases in literature where the same glass strength data have been used for different advanced probability distributions beyond the standard ones. An overview of these is given in *Appendix 1C*.

6. Maximum monthly wind speed (data from Doaa ELhertaniy)

According to CumFreq the best of all for the wind speed data is the GEV cumulative probability distribution function (CDF):

Cumulative frequency of the generalized extreme value (GEV) type :

CDF = exp **[**–**{** 1+C(X–A)/B **}** **^** (–1/C) **]**
A = 2.30E+001   B = 4.89E+000   C = 1.90E-001

Three parameters, more is not required.
Graphics are shown hereunder and on the next page.

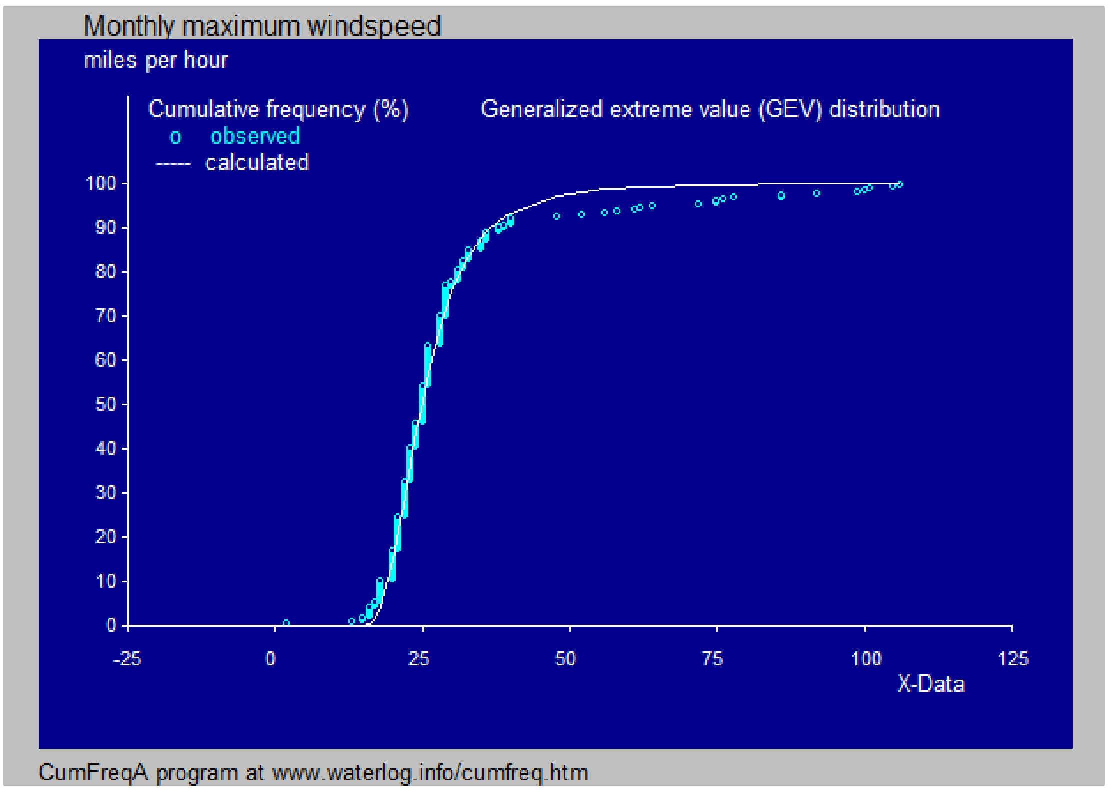


Figure 13.

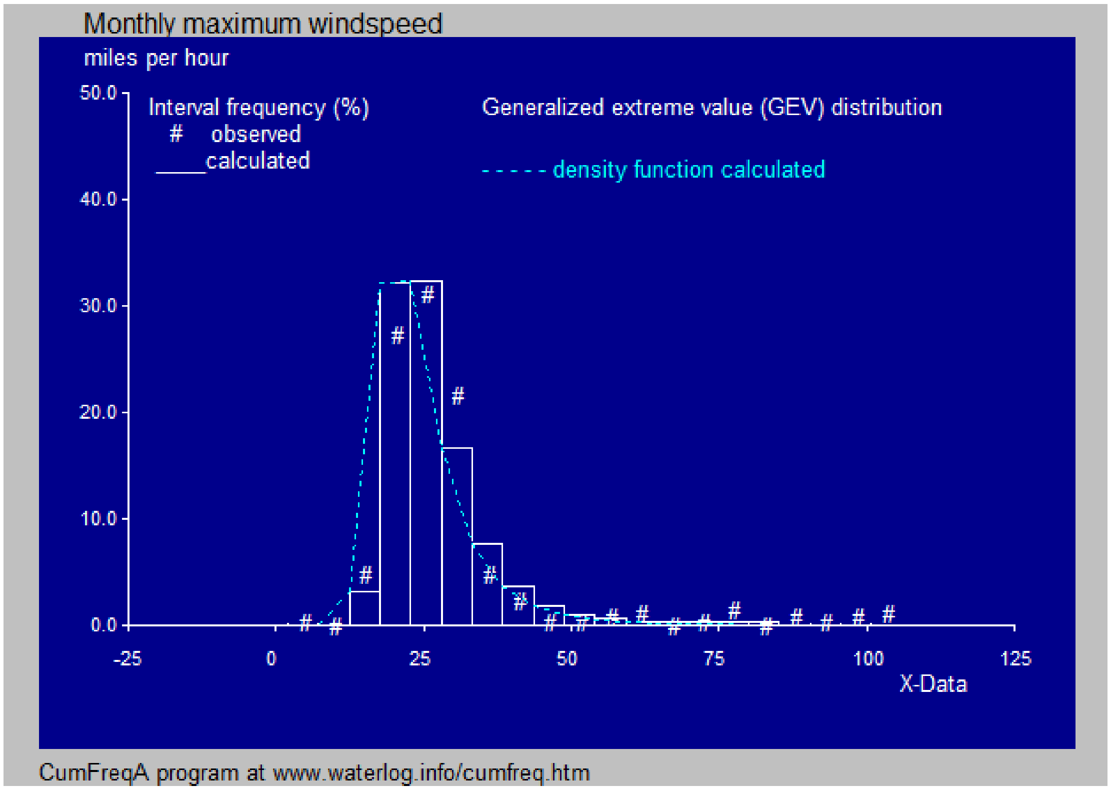


Figure 14.

For wind speed Doaa ELhertaniy (2022) presented the following figure on the probability density function in which G = standard Gumbel distribution with 2 parameters, TG = quadratic rank transmuted Gumbel distribution with 3 parameters, CTG = cubic transmuted Gumbel distribution with 4 parameters, and QTG = quartic transmuted Gumbel distribution with 5 parameters. None of these, even the one with 5 parameters, shows as good a fit as the previous figure made with the GEV distribution employing only 3 parameters. The GEV distribution deserves preference.

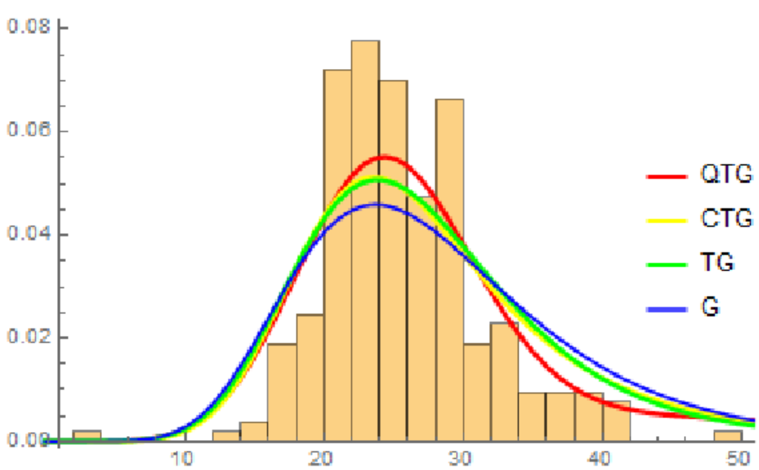


Figure 15.

The CumFreq ranking according to goodness of fit of the first five CDF's for wind speed is:

Distribution name *)	Average of absolute differences between calculated and observed cumulative frequencies #)
1. GEV (generalized extreme value)	2.08 %
2. Log-Normal optimized	2.59 %
3. Laplace generalized	2.77 %
4. Gamma (Erlang)	2.96 %
5. Laplace standard	3.09 %

*) Generalization means raising X to the power P, Optimization signifies that mean and standard deviation are considered parameters to be optimized

#) 0–1% excellent, 1–2% very good, 2–3% good, 3 – 4% medium, > 4 % poor

7. Conclusion

For practical purposes CumFreq is easy to use, versatile, and recommendable (Oosterbaan, 2019)

8. References

1. Farwah Willayat et al. (2022)
   Marshall–Olkin Extended Gumbel Type-II Distribution: Properties and Applications

2. Doaa ELhertaniy (2022)
   Quartic Rank Transmuted Gumbel Distribution

3. R.J.Oosterbaan, CumFreq
   https://www.waterlog.info/cumfreq.htm

4. R.J. Oosterbaan (2022)
   How to derive a probability distribution from a data set using the simple method of plotting positions and the free CumFreq model

5. R.J. Oosterbaan (2022)
   FITTING THE VERSATILE LINEARIZED, COMPOSITE, AND GENERALIZED LOGISTIC PROBABILITY DISTRIBUTION TO A DATA SET

6. R.J. Oosterbaan (2019). Software for generalized and composite probability distributions. International Journal of Mathematical and Computational Methods, Volume 4, 2019, 1-9.

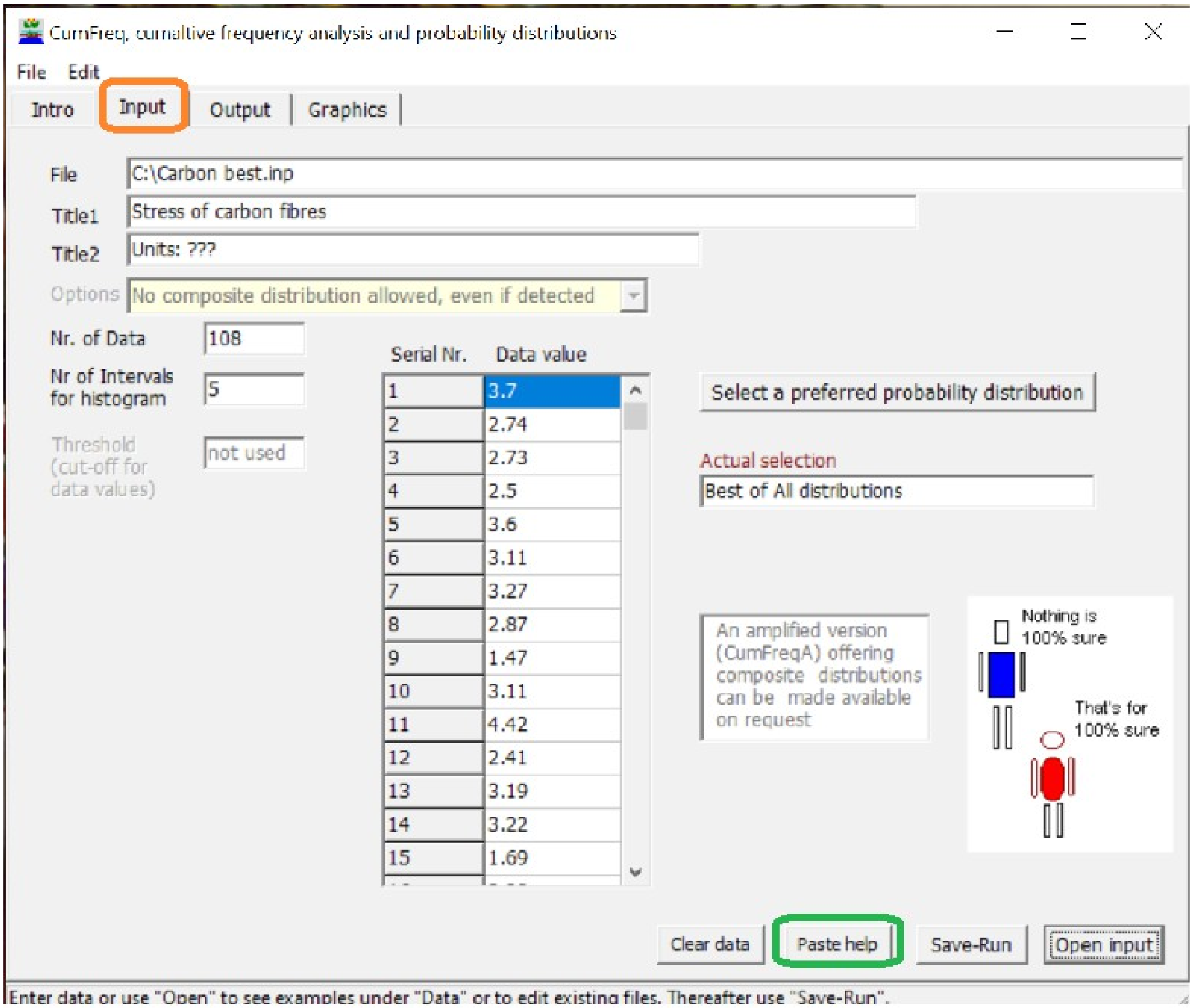
7. Mundher Abdullah Khaleel et al. (2022). New extension of Burr type X distribution properties with Application. Journal of King Saud University - Science (2017). Doi: http://dx.doi.org/10.1016/j.jksus.2017.05.007

8. Noor A. Ibrahim et al, (2022). Weibull Burr X distribution properties and application. Available from: https://www.researchgate.net/publication/319504741_Weibull_Burr_X_distribution_properties_and_application

9. Pelumi E. Oguntunde et al (2022). A New Generalization of the Lomax Distribution with Increasing, Decreasing and Constant Failure Rate. Available from: https://www.researchgate.net/publication/317329320_A_New_Generalization_of_the_Lomax_Distribution_with_Increasing_Decreasing_and_Constant_Failure_Rate

9. Appendices

Appendix A1. CumFreq input menu



Appendix A2. Selection options in CumFreq

Appendix A3. Overview of advanced probability distributions applied to glass strength data

Contents:

A3.1 The EGBX distribution
A3.2 The GoL distribution
A3.3 The BetaBurX distribution
A3.4 The GoFW distribution
A3.5 The TEMO-R distribution
A3.6 The TEMO–Go distribution
A3.7 The OFWE distribution

A3.1 The EGBX distribution

Mundher Abdullah Khaleel et al. (2017) used the same glass strength data for the application of the Exponentiated Generalized Burr type X (EGBX) distribution with 4 parameters. Their result is shown in the next figure A3.1 (red density curve). This curve has a lower peak than the QTG curve in figure 12 (Section 5) which already has a relatively low peak, but it is at a somewhat higher X value which is closer to the modal value of the observed data in the histogram than the QTG curve. The goodness of fit between the to solutions is therefore difficult to compare. Anyhow, the curve does not show as good a fit as the curve in figure 11 (Section 5) made with the standard Laplace distribution using the CumFreq software employing only 3 parameters.

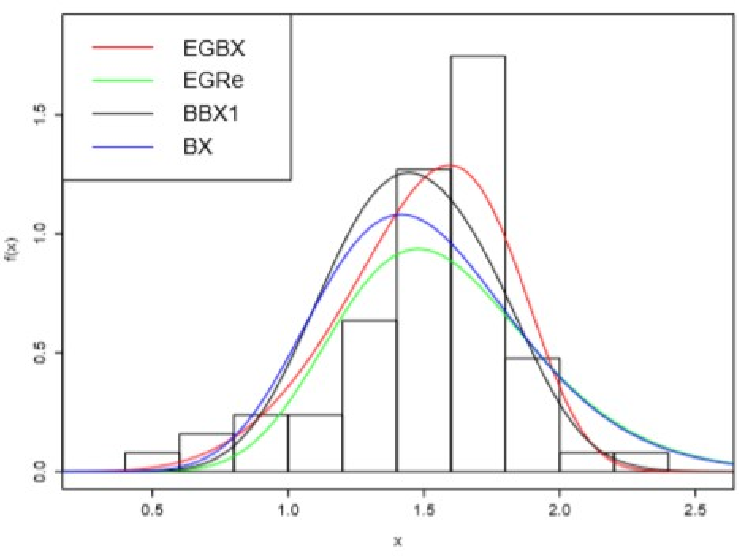


Figure A3.1.

Reference:

Mundher Abdullah Khaleel, Noor Akma Ibrahim, Mahendran Shitan, and Faton Merovci (2017). New extension of Burr type X distribution properties with application. Article in Journal of King Saud University - Science · May 2017. DOI: http://dx.doi.org/10.1016/j.jksus.2017.05.007

__A3.2 The GoL distribution__

Pelumi E. Oguntunde et al. (2017) used the same glass strength data for the application of the New Generalization of the Lomax Distribution using the Gompertz-Lomax (GoL) distribution with 4 parameters. Their result is shown in the next figure A3.2 (red density curve). This curve looks much the same as the EGBX curve in figure A3.1. The comments made under figure A3.1 also hold for the present GoL distribution.

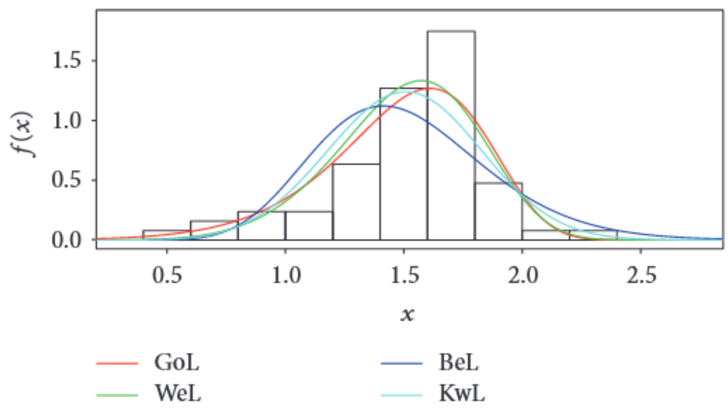


Figure A3.2.

Reference:

Pelumi E. Oguntunde, Mundher A. Khaleel, Mohammed T. Ahmed, Adebowale O. Adejumo and Oluwole A. Odetunmibi (2017). A New Generalization of the Lomax Distribution with Increasing, Decreasing and Constant Failure Rate. Hindawi Modelling and Simulation in EngineeringVolume 2017, Article ID 6043169, 6 pages. https://doi.org/10.1155/2017/6043169

__A3.3 The BetaBurX distribution__

Faton Merovci et al. (2016) used the same glass strength data for the application of the BetaBurX Distribution with 4 parameters. Their result is shown in the next figure A3.3 (red density curve). This curve looks much the same as the EGBX curve in figure A3.1 and the GoL distribution in figure A3.2. The comments made under figure A3.1 also hold for the present BetaBurX distribution.

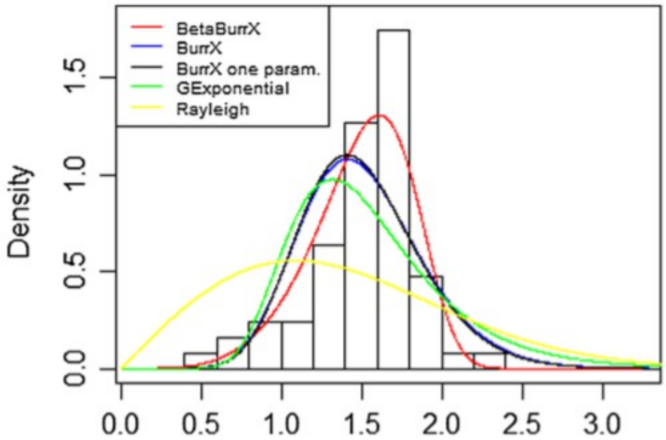


Figure A3.3.

Reference:
Faton Merovci, Mundher Abdullah Khaleel, Noor Akma Ibrahim and Mahendran Shitan (2016). The beta Burr type X distribution properties with application Springer Plus volume 5, Article number: 697 (2016). DOI: https://doi.org/10.1186/s40064-016-2271-9

A3.4 The GoFW distribution

Khaleel, M. A. (2020) used the same glass strength data for the application of the GoFW Distribution with 4 parameters. Their result is shown in the next figure A3.4 (green density curve).

This curve looks much the same as the EGBX curve in figure A3.1 and the GoL distribution in figure A3.2 and the BetaBurX distribution in figure A3.3. The comments made under figure A3.1 also hold for the present GoFW distribution.

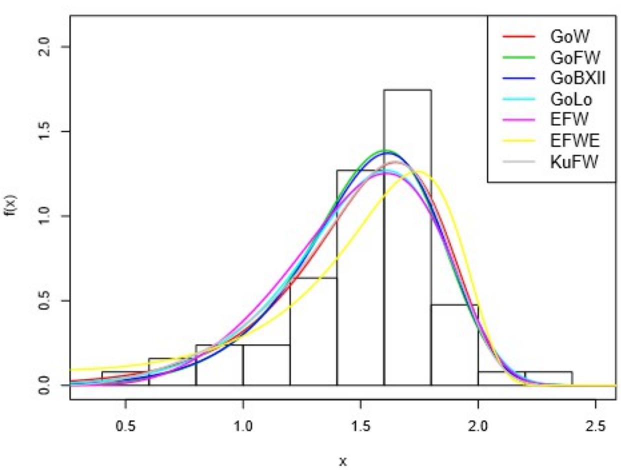


Figure A3.4.

Reference:
Khaleel, M. A., Oguntunde, P.E, Ahmed, M. T., Ibrahim, N. A., and Loh, Y. F. (2020). The Gompertz Flexible Weibull Distribution and its Applications . Malaysian Journal of Mathematical Sciences 14(1): 169-190 (2020).
http://119.40.117.179/fullpaper/2020-January-14-1/Khaleel,%20M.%20A.-169-190.pdf

A3.5 The TEMOR distribution

Ali Mahmood Munef and Mundher A. Khaleel (2021) used the same glass strength data for the application of the TEMO-R Distribution with 3 parameters. Their result is shown in the next figure A3.5 (black density curve). This curve looks much the same as the EGBX curve in figure A3.1, the GoL distribution in figure A3.2 and the BetaBurX curve in figure A3.3, with the exception that the peak is somewhat higher and closer to the peak of the histogram of the observed data. Yet is does reach as high as the peak of the Laplace distribution with 3 parameters in figure 11 (Section 5) obtained with the CumFreq software, so that the latter deserves preference.

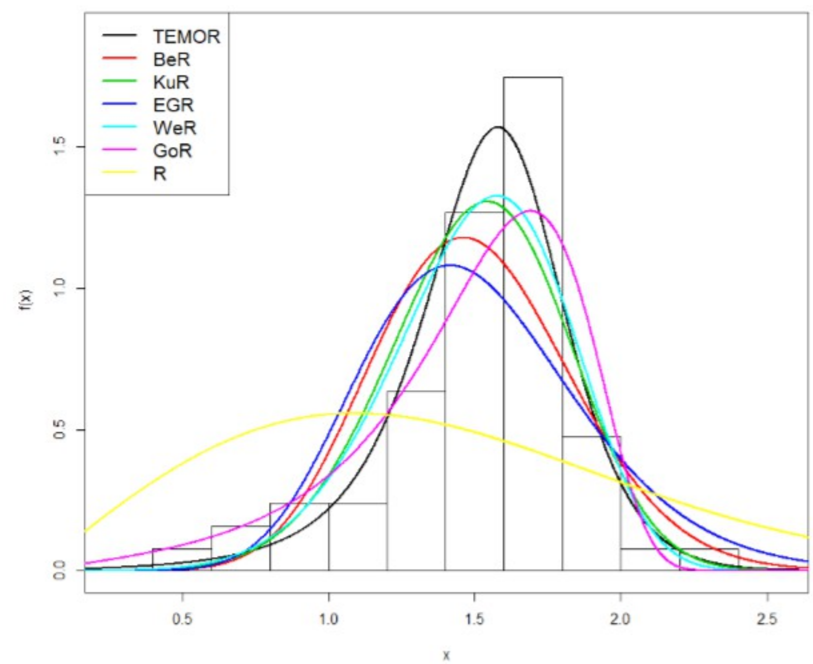


Figure A3.5.

Reference:
Ali Mahmood Munef and Mundher A. Khaleel (2021). Truncated Exponential Marshall Olkin Rayleigh distribution Properties and Applications.
file:///D:/Downloads/db77bb0d0983b46a-3.pdf

A3.6 The TEMO-Go distribution

Alaa Abdulrahman Khalaf and Mundher A. Khaleel (2020) used the same glass strength data for the application of the TEMO-Go Distribution with 4 parameters. Their result is shown in the next figure A3.6 (black density curve). This curve looks much the same as the TEMO-R curve in figure A3.4 Yet is does reach as high as the peak of the Laplace distribution in figure 11 (Section 5) employing only 3 parameters, so that the latter deserves preference.

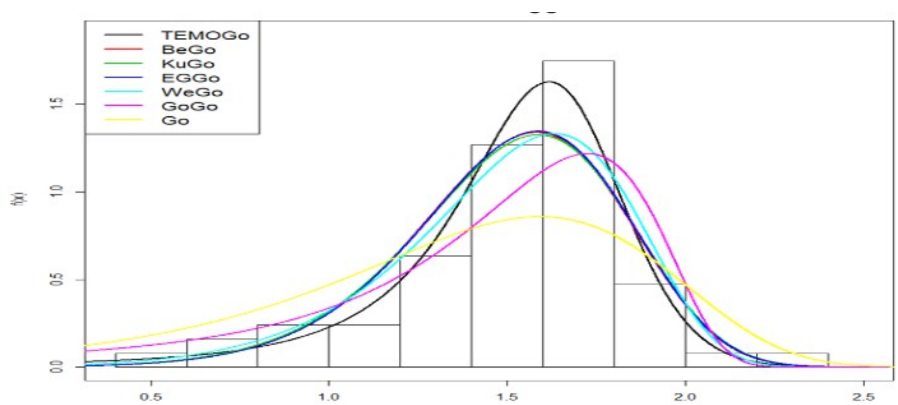


Figure A3.6.

Reference:

Alaa Abdulrahman Khalaf and Mundher A. Khaleel (2020). Truncated Exponential Marshall-Olkin-Gompertz Distribution: Properties and Applications. file:///D:/Downloads/2020.pdf

A3.7 The OFWE distribution

M. El-Morshedy and M. S. Eliwa (2020) used the same glass strength data for the application of the OFWE Distribution with 3 parameters. Their result is shown in the next figure A3.7 (red density curve). This curve looks much the same as the (EGBX in figure A3.1, the GoL distribution in figure A3.2, the BetaBurX distribution in figure A3.3, and the TEMO-R curve in figure A3.4

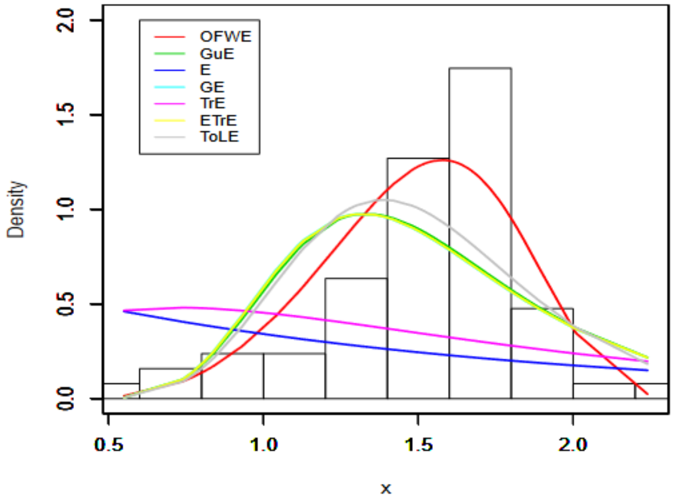


Figure A3.7.

Reference:
M. El-Morshedy, M. S. Eliwa (2019). The Odd Flexible Weibull-H Family of Distributions: Properties and Estimation with Applications to Complete and Upper Record Data. Filomat 33:9 (2019), 2635–2652. https://doi.org/10.2298/FIL1909635E