Comparing the results of Truncated Generalized Fréchet Weibull distributions found in literature with those obtained by the free CumFreq program

R.J. Oosterbaan

October 2022   website: www.waterlog.info

Abstract

Ramadan AbdelHamid Zeineldin et al. have given data to which they have fitted Truncated Generalized Fréchet Weibull 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. Ramadan AbdelHamid Zeineldin et al. (2021) *Generalized Truncated Frechet Generated Family Distributions and Their Applications*. Computer Modeling in Engineering and Sciences 126(1):1-29. January 2021. DOI: https://doi.org/10.32604/cmes.2021.012169
2. R.J.Oosterbaan, CumFreq
   https://www.waterlog.info/cumfreq.htm

Contents

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.

CumFreq uses plotting positions: $F_c = r / (N+1)$ where r is the rank number of the data arranged in ascending order: $r = 1, 2. 3, ... N$ with N being the total number of data to. Here F_c is the observe cumulative frequency or probability (Makkonen 2006). With the help of F_c it is fairly simple to find the parameters of a large number of well known probability distributions.

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. Minutes waiting time before a client receives the desired service in a bank.

According to CumFreq the best of all distributions for the minutes waiting time is the Burr (generalized Pareto-Lomax) cumulative probability distribution function (CDF or Fc):

$$CDF = Fc = 1 - [\, B\ /\ (X^P + B)\,]^E$$

The Fc values can be found by the plotting positions: $Fc = r/(N+1)$ where r is the rank number of the data arranged in ascending order: $r = 1, 2. 3, \ldots\ldots N$ with N being the total number of data (Makkonen 2006, Oosterbaan 2022).

Using the transformed values $Ft = \ln\{1 - Fc\}$ and $Xt = \text{Ln}\{B\ /\ (X^A + B)\}$ one can write: $Ft = E*Xt$, where E can be found from a simple ratio method, while the exponent P and the term B are to be numerically optimized. The distribution has three parameters.

In the case of waiting time in the bank, CumFreq gave the following result:

$P = 1.62E+000 \quad B = 1.15E+002 \quad E = 3.12E+000$

The CDF function is shown in figure 1 while the histogram together with the probability density function is shown in the figure 2 hereunder.
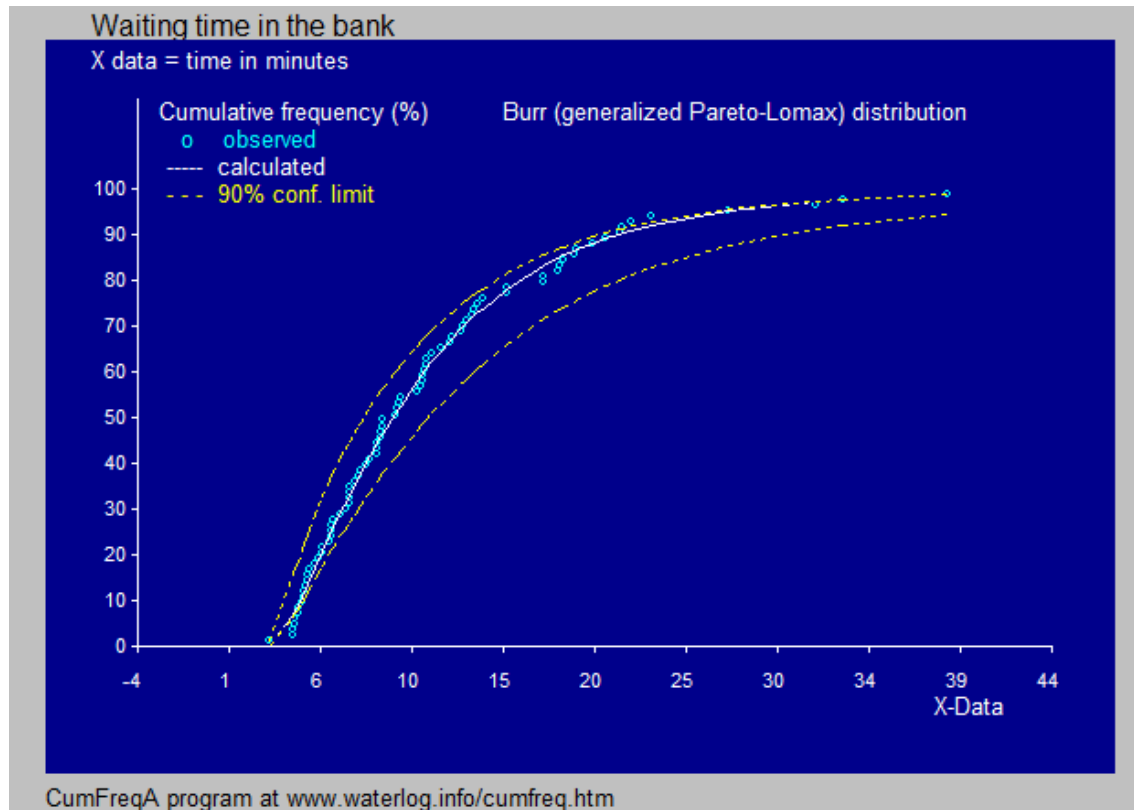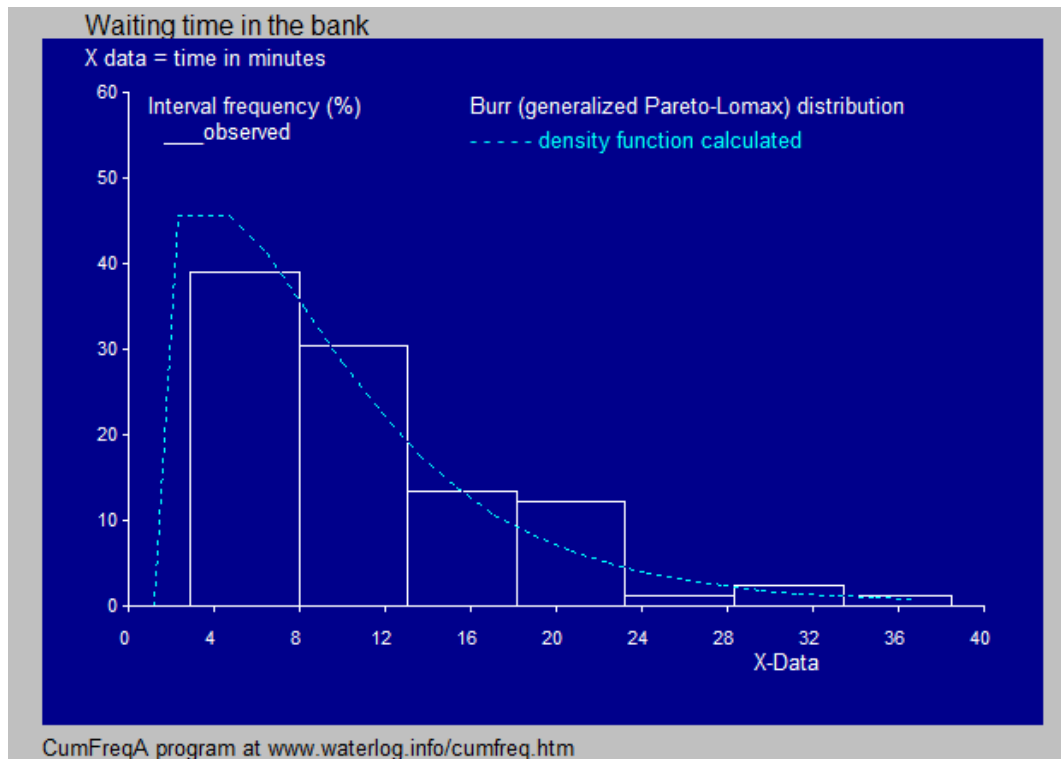

Figure 1.

Figure 2.

R. A. Zeineldin et al., using the Extended Truncated Generalized Fréchet Weibull (TGFrW) distribution, working with three parameters, have reached the following results:
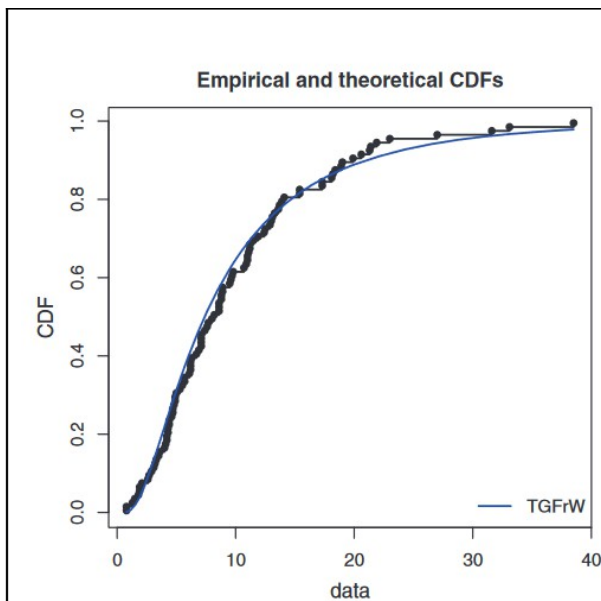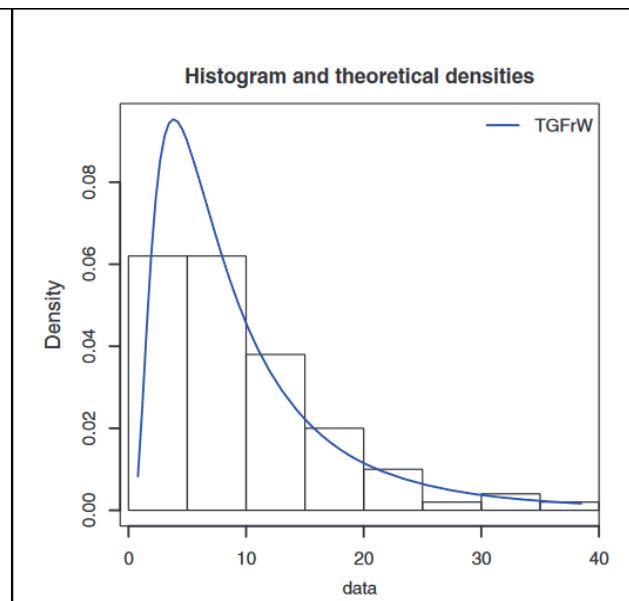

Figure 3.

Figure 4.

The CDF curve in figure 1 fits better to the CDF curve in figure 3 and the density curve in figure 2 fits better to the observed data represented by the histogram in figure 4. Hence the Burr distribution found with CunFreq deserves preference over the TGFrW distribution.

The option in CumFreq to detect the best fitting of all distributions results in a ranking according to goodness of fit. The next table gives the first nine CDF's for waiting times in the bank:

Distribution name *)	Average of absolute differences between calculated and observed cumulative frequencies #)
1. Burr generalized	1.22 %
2. Exponential distribution generalized	1.30 %
3. Dagum generalized	1.33 %
4. Log-normal distribution optimized	1.38 %
5. Generalized extreme value (GEV)	1.44 %
6. Gumbel generalized	1.54 %
7. Frechet (Fisher-Tippet Type 2) distribution	1.76 %
8. Gamma (Erlang) distribution	2.23 %
9. Fisk distribution	2.25 %

*) Generalization means raising X to the power P.
    Optimization means adjusting mean and standard deviation
#) 0–1% excellent, 1–2% very good, 2–3% good, 3 – 4% medium, > 4 % poor

R. A. Zeineldin et al. use different indexes for the goodness of fit so that here a comparison cannot be made.

3. Monthly precipitation (in inches) in Minneapolis.

According to CumFreq the best of all distributions for the monthly precipitation is the Dagum (inverted Burr) cumulative probability distribution function (CDF or Fc):

$$CDF = Fc = [B / (X^P + B)]^E$$

The Fc values can be found by the plotting positions: $Fc = r / (N+1)$ where r is the rank number of the data arranged in ascending order: r = 1, 2. 3, ... N with N being the total number of data (Makkonen 2006, Oosterbaan 2022).

Using the transformed values $Ft = \ln (Fc)$ and $Xt = Ln \{ B / (X^A + B) \}$ one can write:
$Ft = E*Xt$, where A can be found from a simple ratio method while P and B need to be optimized numerically. The distribution has three parameters.
In the case of monthly precipitation, CumFreq gave the following result:
P = - 2.94E+000   B = 1.56E-001   E = 6.45E-001

The CDF function is shown in figure 5 while the histogram together with the probability density function is shown in the figure 6 hereunder.
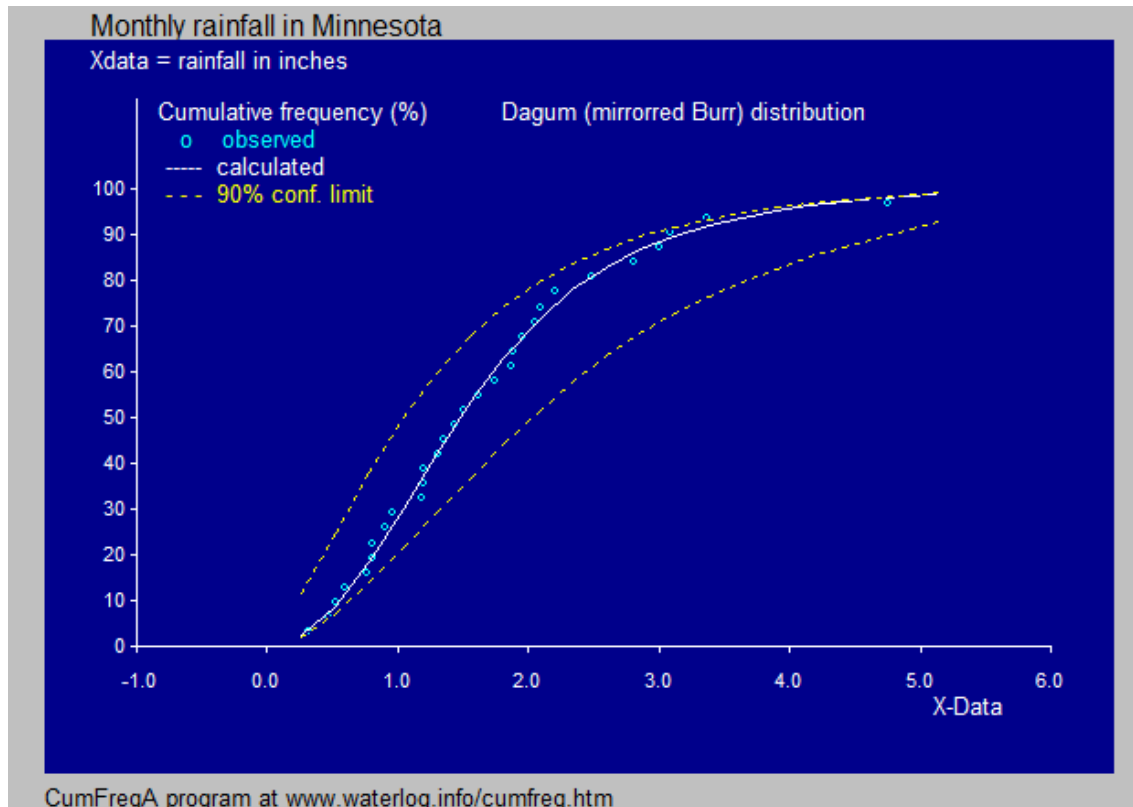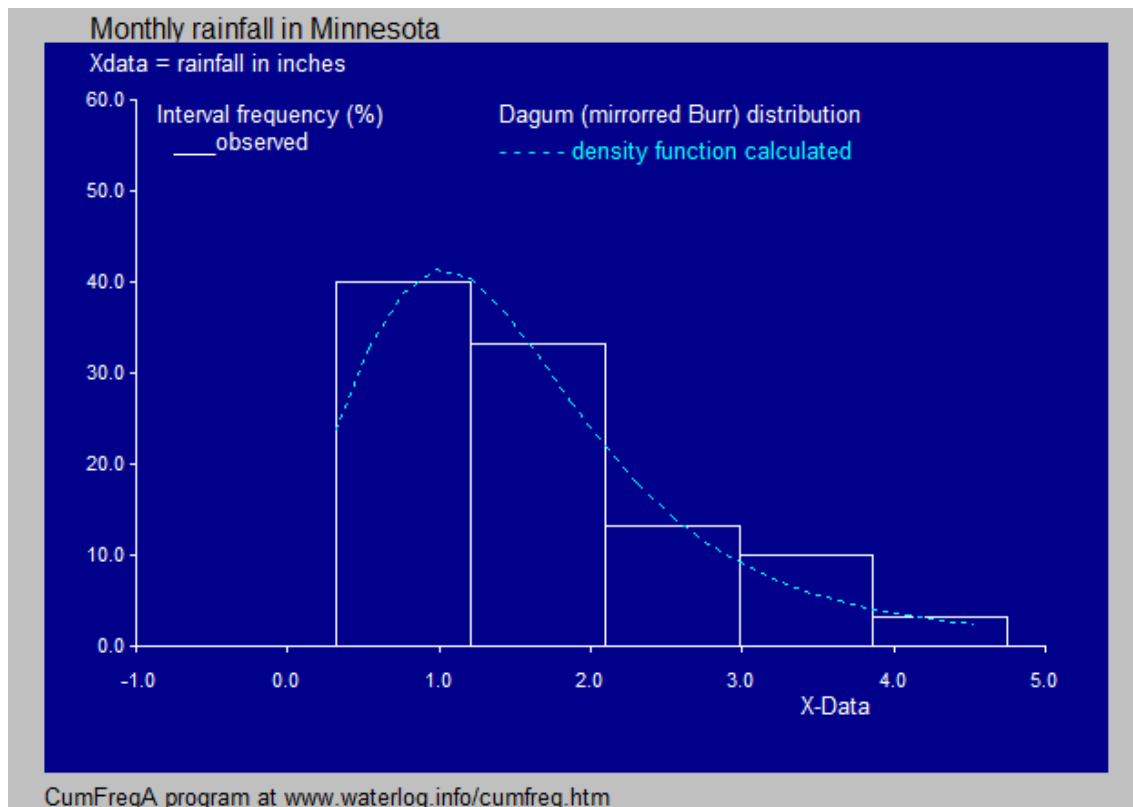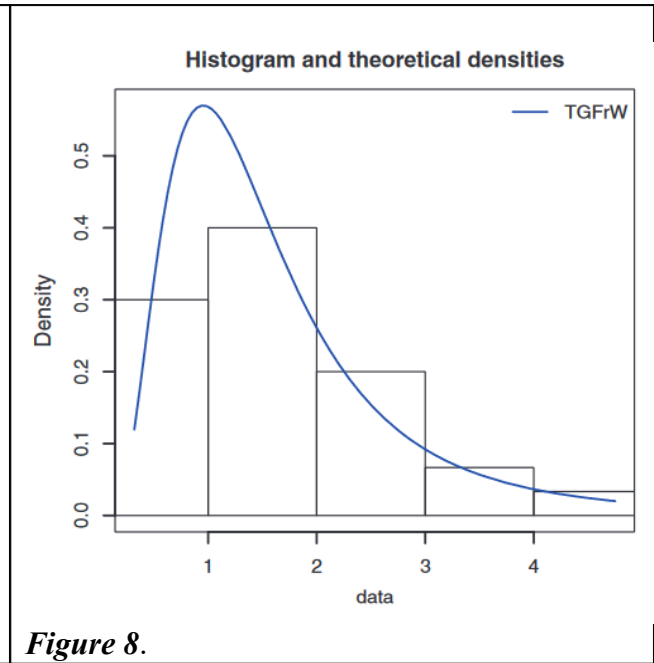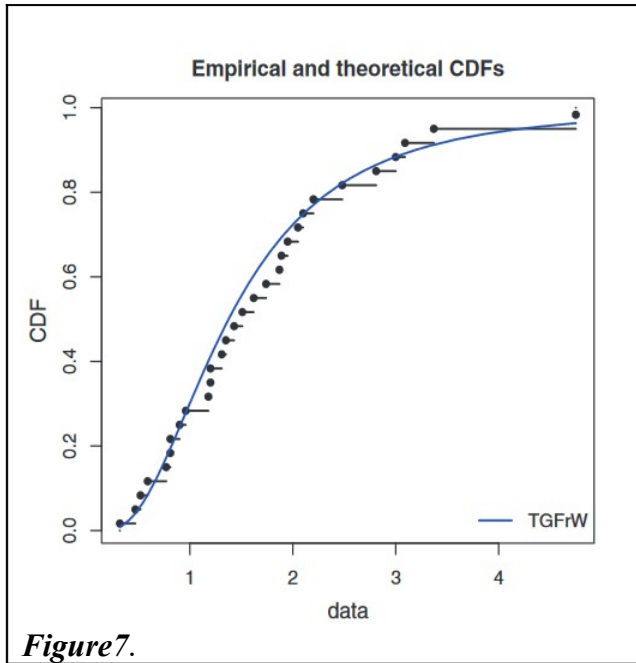
Figure 5.

Figure 6.

R. A. Zeineldin et al., using the Extended Truncated Generalized Fréchet Weibull (TGFrW) distribution, working with three parameters, have reached the following results:

Figure7.

Figure 8.

The CDF curve in figure 5 fits better to the CDF curve in figure 7 and the density curve in figure 6 fits better to the observed data represented by the histogram in figure 8. Hence the Dagum distribution found with CunFreq deserves preference over the TGFrW distribution.

The option in CumFreq to detect the best fitting of all distributions results in a ranking according to goodness of fit. The next table gives the first eight CDF's for monthly precipitation:

Distribution name *)	Average of absolute differences between calculated and observed cumulative frequencies #)
1. Dagum generalized	1.49 %
2. Burr generalized	1.50 %
3. Logistic distribution generalized	1.51 %
4. Gumbel generalized	1.53 %
5. Frechet (Fisher-Tippet Type 2) distribution	1.54 %
6. Generalized extreme value (GEV)	1.55 %
7. Frechet (Fisher-Tippet Type 2) mirrored	1.64 %
8. Rayleigh generalized	1.67 %
9. Weibull distribution	1.97 %

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

In the table, the first five distributions give all practically the same good results.
while the others do not fall much behind.

R. A. Zeineldin et al. use different indexes for the goodness of fit so that here a comparison cannot be made.

4. Conclusion

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

5. References

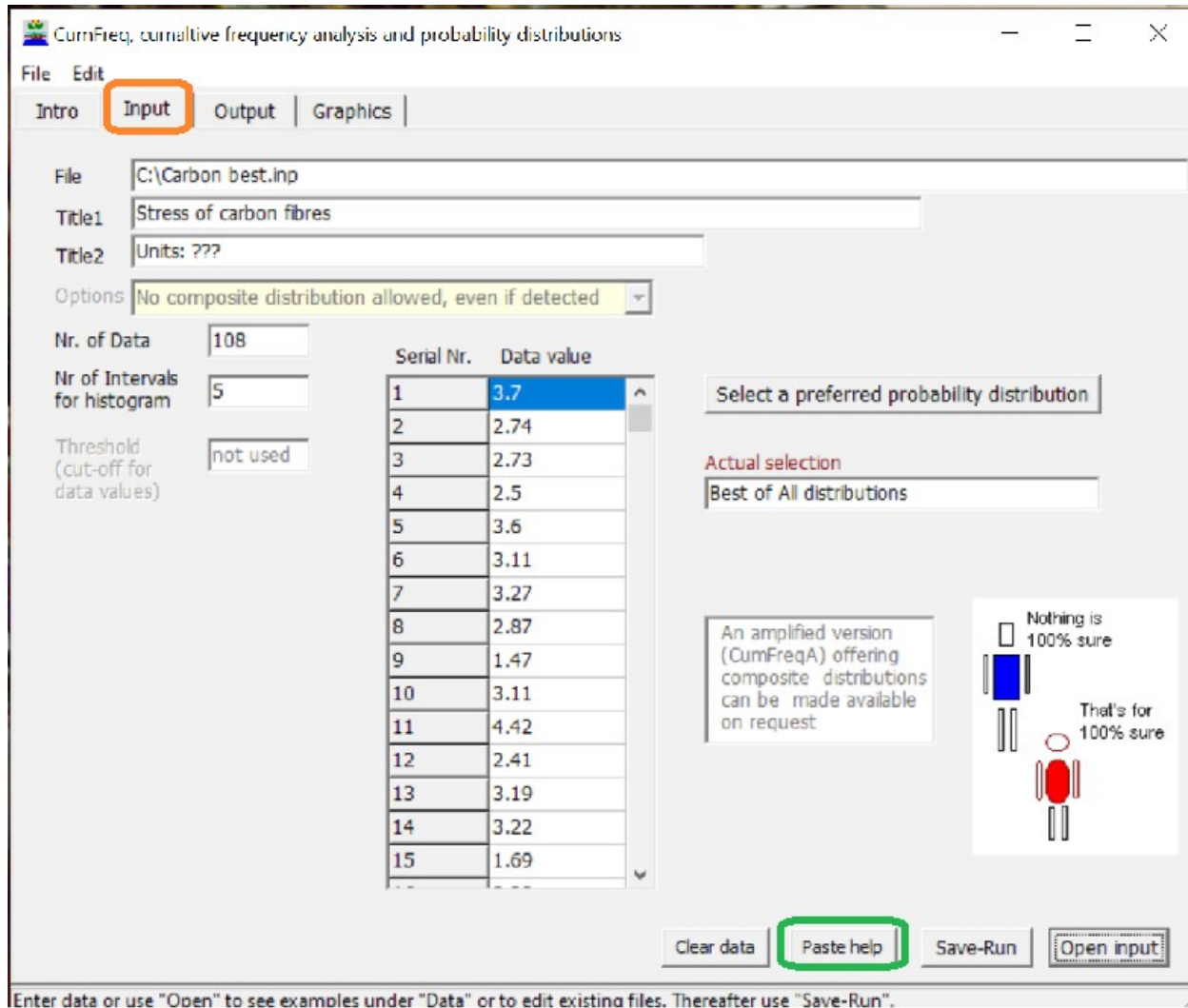L. Makkonen, 2006. Plotting Positions in Extreme Value Analysis. Journal of applied meteorology an climatology, 01 Feb 2006, pp. 334-340.. https://doi.org/10.1175/JAM2349.1

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

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

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

9. Appendices

Appendix A1. CumFreq input menu

Appendix A2. Selection options in CumFreq

CumFreq select distributions

Make a selection then click "Confirm"

- ⦿ **Best of All Distributions**
- ○ Normal (standard)
- ○ Normal (optimized)
- ○ Log-normal (standard)
- ○ Log-normal (optimized)
- ○ Root-normal
- ○ Square-normal
- ○ Generalized normal
- ○ Logistic (standard)
- ○ Log-logistic
- ○ Logistic generalized
- ○ Burr distribution
- ○ Cauchy (standard)
- ○ Cauchy generalized
- ○ Dagum distribution
- ○ Exponential (standard)
- ○ Exponential generalized
- ○ Exponential general. mirr.
- ○ Gamma (Erlang)

- ○ Fisk (adjusted)
- ○ Fisher-Tippet III
- ○ F-T III mirrored
- ○ Frechet (typical)
- ○ Frechet mirrored
- ○ GEV (Gen. Extreme. Value)
- ○ Gompertz (generalized)
- ○ Gumbel (standard)
- ○ Gumbel generalized
- ○ Gumbel mirrored
- ○ Gumbel mirrored generalized
- ○ Kumaraswamy generalized
- ○ Laplace (standard)
- ○ Laplace generalized
- ○ Rayleigh generalized
- ○ Student (1 or 2 d.f.)
- ○ Weibull (standard)
- ○ Weibull mirrored

Confirm