An Analysis of Amarillo's Rainfall Totals

By Robert Ashcraft
Retired Program Scientist - Pantex

Amarillo's local weather station has been measuring and reporting rainfall totals since 1892. In this report we will look at cumulative distributions for both annual and monthly rainfall totals and develop mathematical models for those distributions.

Part 1. Annual rainfall totals

Annual rainfall totals in Amarillo have ranged from 7.01 inches (2011) to 39.75 inches (1923). The annual rainfall data, consisting of 131 values, can be found at the NOAA Online Weather Data website.

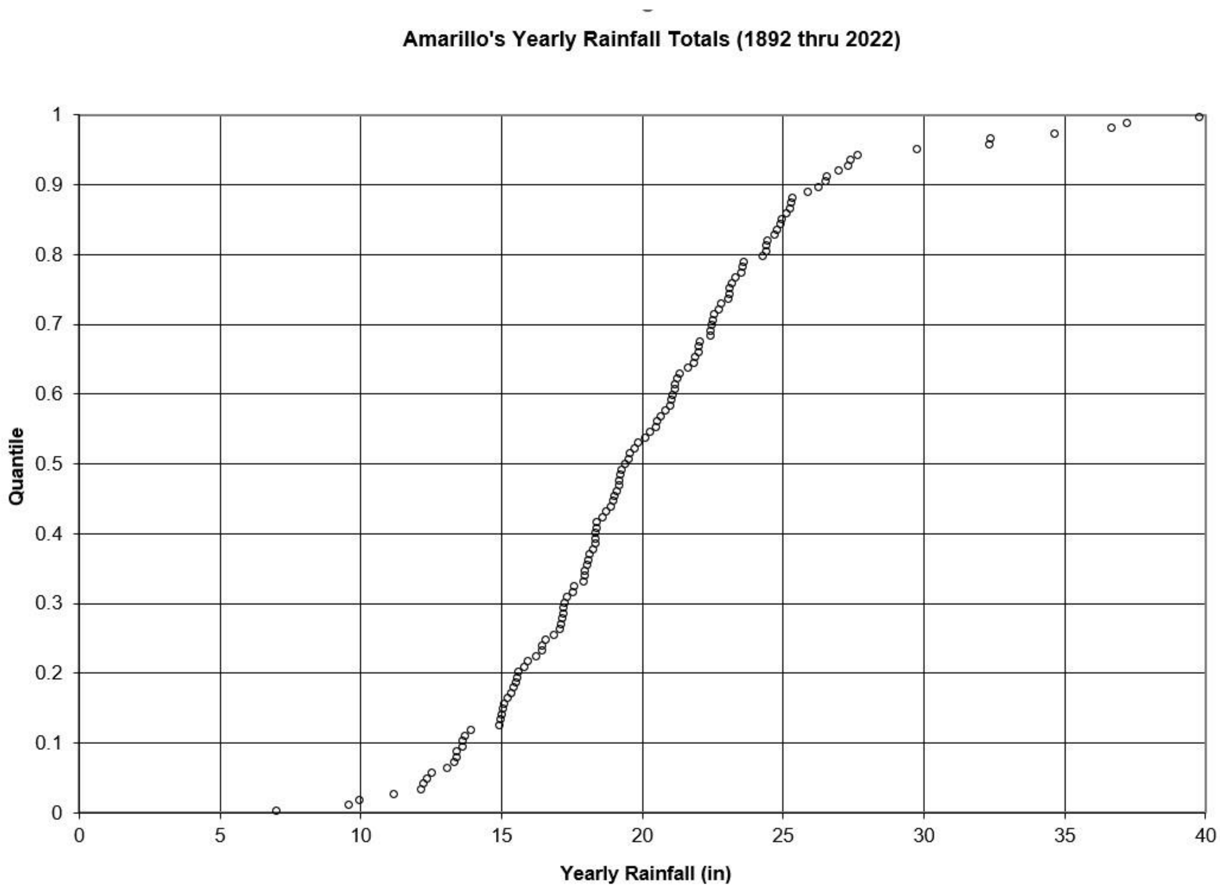
Amarillo's rainfall for 2022 was 16.43 inches. Only 30 of the 131 years on record have been drier.

The purposes of this study were to (1) generate a cumulative distribution for the rainfall totals, (2) determine the nature of the distribution, and (3) determine the distribution parameters that would provide the best approximation to the data.

The analysis began by generating the cumulative distribution for the data. To accomplish this, the data were sorted in ascending order. Then a quantile was computed for each data value using Hazen's formula,

$$q_i = \left(\frac{i-0.5}{n}\right) \quad i = 1, 2, \dots, n$$

In this formula q_i is the i^{th} quantile and n is the number of data points. A plot was then made with the quantiles on the y-axis and the sorted rainfall totals on the x-axis, as shown in the following graph.

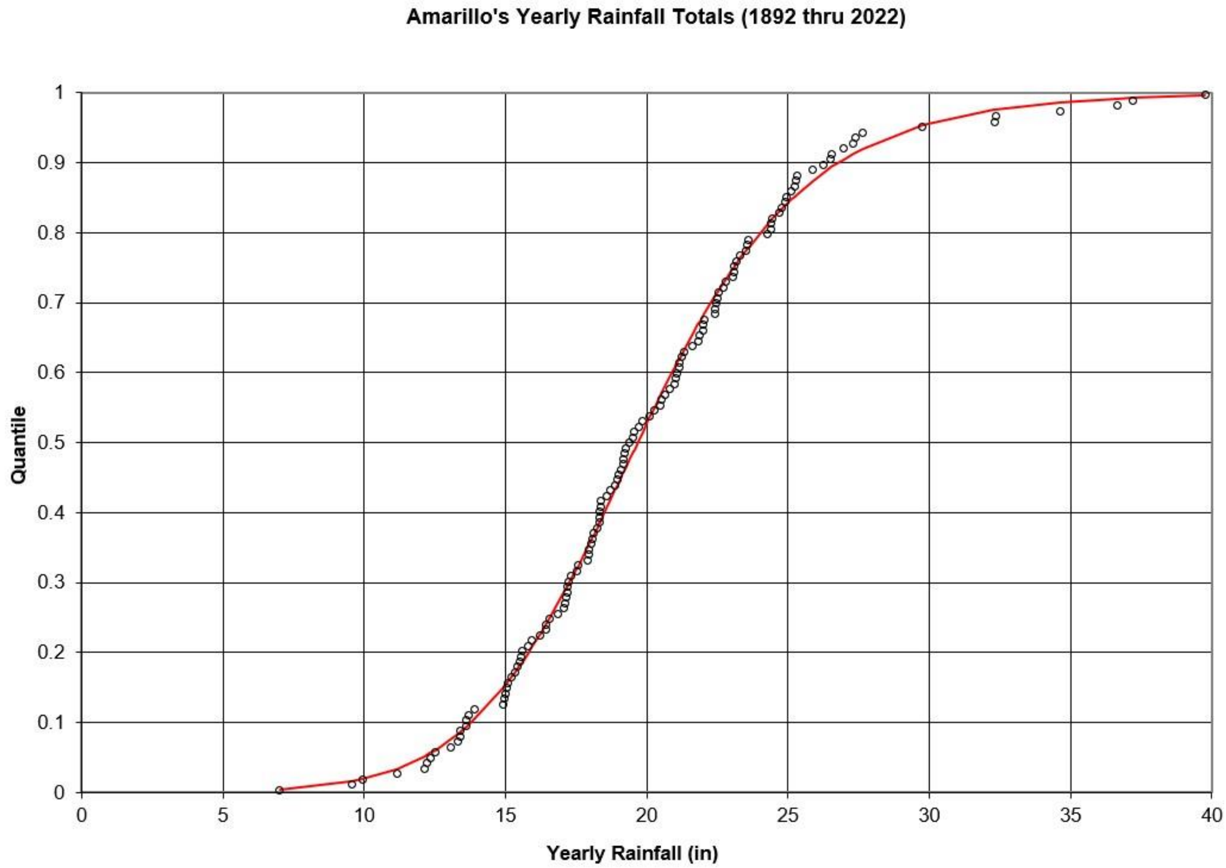


The data appear to be roughly "S" shaped, but are not symmetric since the right tail is longer than the left tail. Thus, we can rule out a normal distribution. Several asymmetric distributions were considered, such as log normal, square root normal, etc., but the distribution that fit the data the best was a modified sigmoid curve. Its equation is

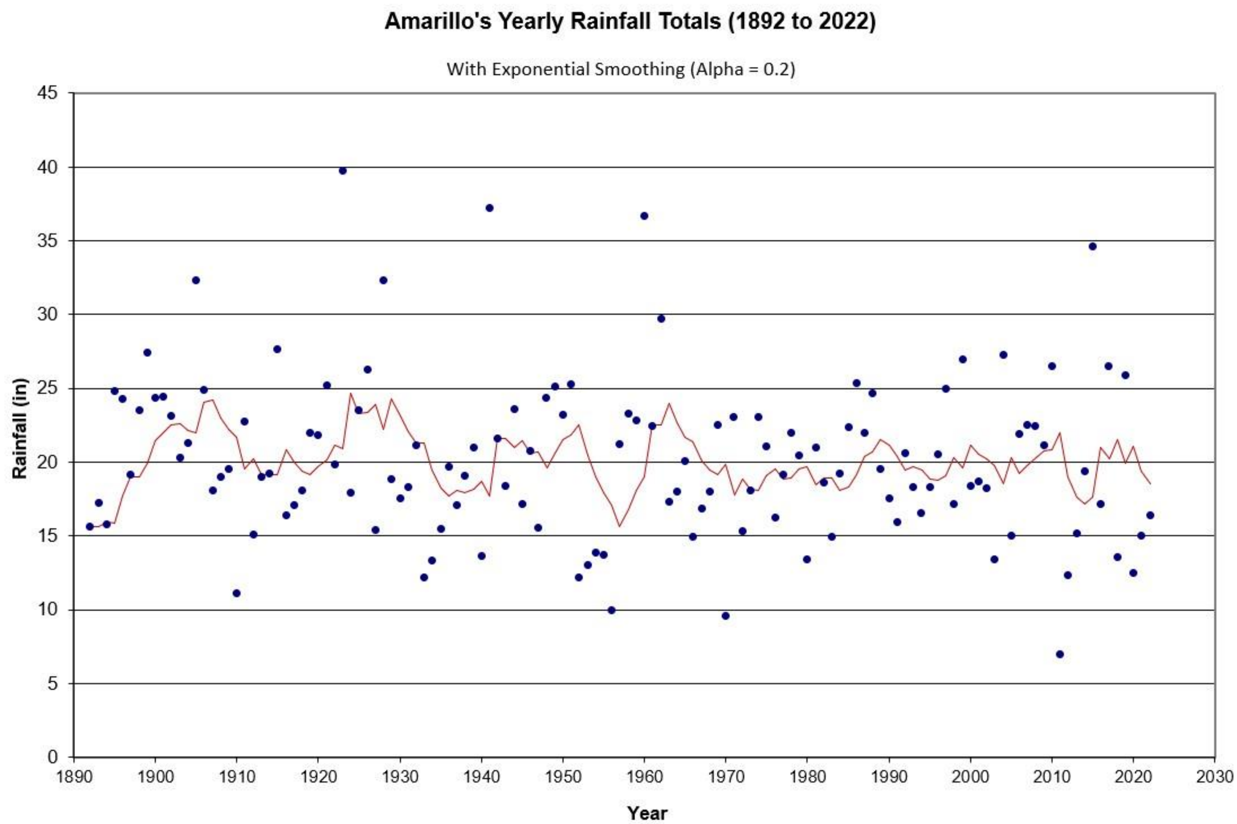
$$F(x) = \frac{1}{1 + exp\left(-\frac{x^{\gamma} - \alpha}{\beta}\right)}$$

where *F(x)* is the fraction of yearly rainfall totals that will be less than or equal to *x* inches, and α, β, and γ are parameters to be determined by a least squares regression.

The parameters α, β, and γ were found to be 3.60752, 0.23316, and 0.43064, respectively. The resulting curve is shown with the data in the following plot, and the curve is an excellent approximation to the data.



Finally, here is a time series plot of the yearly rainfall totals (blue circles) with exponentially smoothed values (red line segments). The smoothed values indicate no apparent trend in the data.



Part 2. Monthly rainfall totals

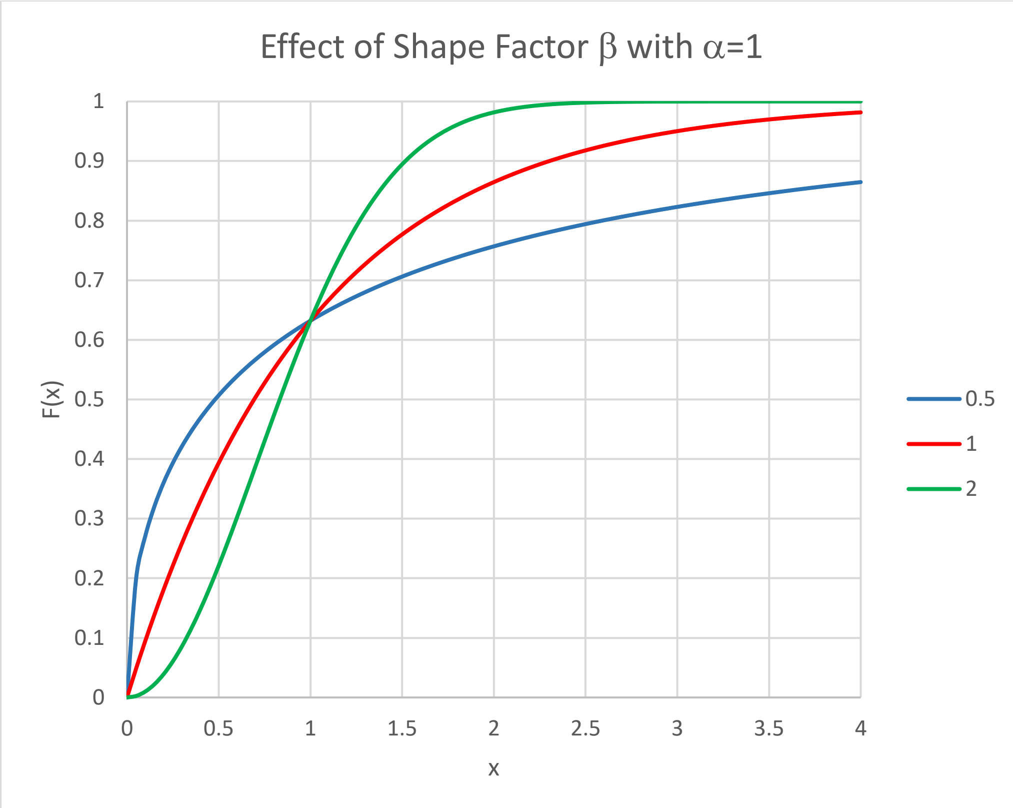
Amarillo's monthly rainfall totals required a different model than the one used for the annual totals. The monthly totals were found to be well approximated by a cumulative Weibull distribution function of the form

$F(z) = 1 - e^{-z}$ where $z = \left(\frac{x}{\alpha}\right)^{\beta}$

In this equation, *F(z)* is the quantile, *x* is the rainfall total in inches, α is the scale factor, and β is the shape factor. More information about the Weibull distribution can be found at this website: https://en.wikipedia.org/wiki/Weibull_distribution

As before, the rainfall totals for each month were sorted in ascending order and became the x-coordinates. A quantile q_i was calculated for each of *n* rainfall totals and the quantiles became the y-coordinates. The parameters α and β for each month were found by least squares regression.

The following plot shows the effect of the shape factor β. For each curve, $\alpha = 1$ while $\beta = 0.5, 1, and\ 2$. For $\beta > 1$, the curve has an S shape. The curve is a smooth exponential for $\beta \leq 1$.

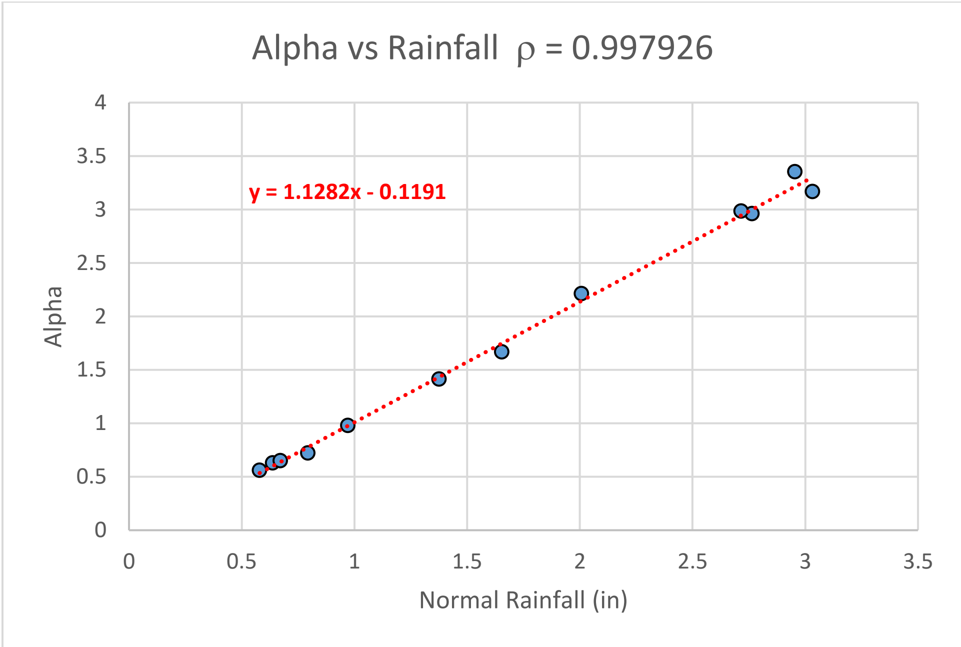


Least squares regression with the cumulative Weibull distribution function was done for each of the 12 months. Plots of the data and resulting curve fit for each month are shown in the Appendix. Results of the curve fits are shown in the following table:

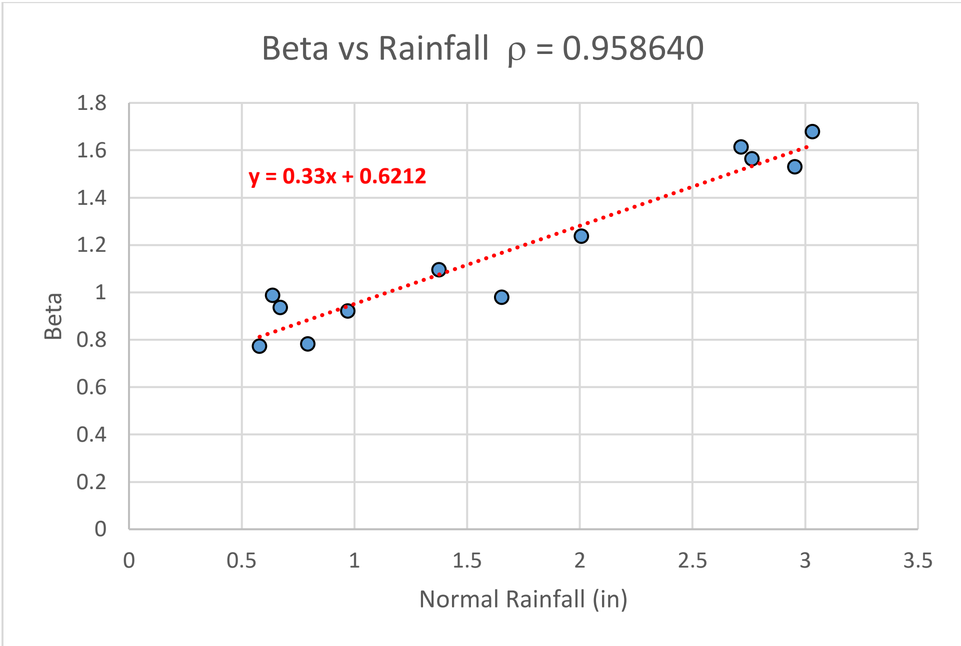
Month	Normal Rainfall (in)	Minimum (in)	Maximum (in)	α	β	R^2
Jan	0.579313	0	3.17	0.561629	0.773422	0.991
Feb	0.637099	0	2.93	0.628965	0.988155	0.995705
Mar	0.970382	0	4.14	0.978857	0.921028	0.995747
Apr	1.375191	0	0	1.413209	1.095165	0.997444
May	2.76313	0.04	9.81	2.960761	1.564436	0.996491
Jun	3.032061	0.01	10.73	3.16845	1.679149	0.995263
Jul	2.714962	0.04	8.02	2.984173	1.614753	0.994954
Aug	2.953817	0.15	8.07	3.352816	1.530101	0.997227
Sep	2.006794	0.03	6.42	2.21221	1.238098	0.998274
Oct	1.653206	0	7.64	1.668541	0.980026	0.994503
Nov	0.792366	0	5.09	0.722406	0.782879	0.985148
Dec	0.671221	0	4.52	0.650772	0.937076	0.995035

The table shows, for each month, the normal rainfall for that month, the minimum and maximum observed values, α and β from the curve fit, and R^2, a measure of how well the curve fits the data. An R^2 value of 1.0 indicates a perfect fit.

The scale factor α is a good approximation to the normal rainfall. Note how the normal rainfall totals for each month are close to the corresponding α values. These results suggest that there is a good correlation between the values of α and the normal rainfalls. Here is the plot with the trendline:



Hence, the relationship is highly linear with a correlation coefficient of 0.997926. What about the shape factor β? Does it also correlate with normal rainfall? Here is the plot with the trendline:



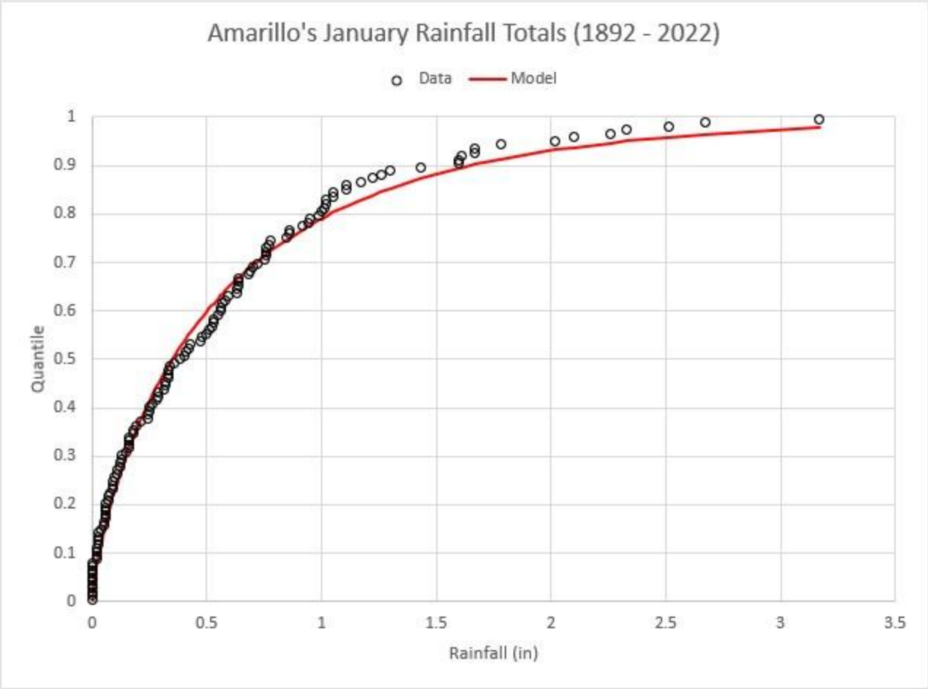
There is a tendency for β to increase with rainfall, but the effect is not as linear as it is with α. The correlation coefficient has dropped to 0.958640.

The best fit lines for α and β as functions of normal rainfall (x) are:

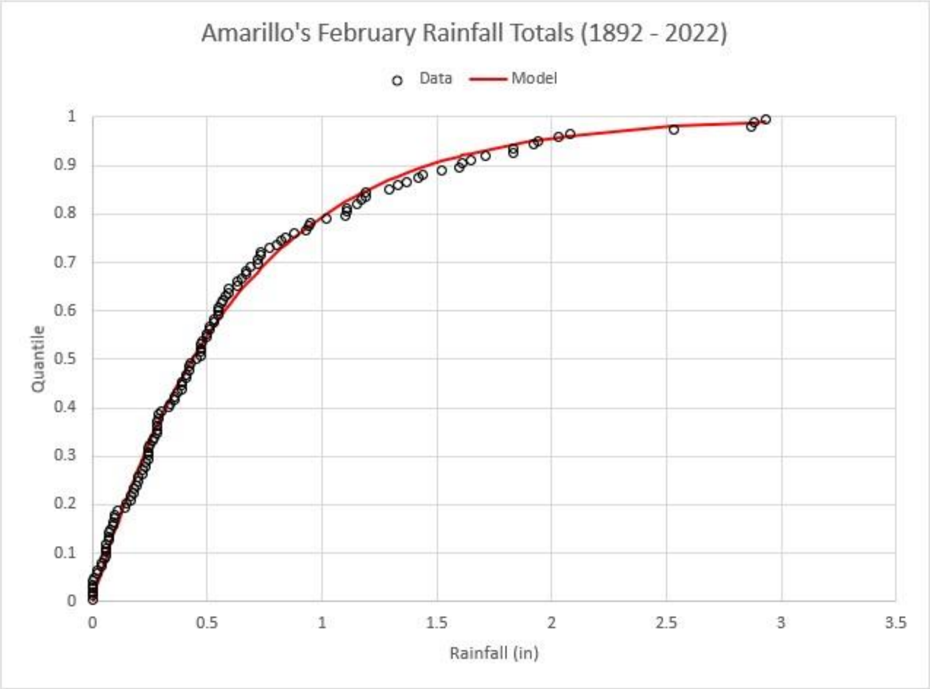
$$\alpha = 1.1282x - 0.1191$$
$$\beta = 0.330x + 0.6212$$

In conclusion, Amarillo's monthly rainfall totals from the past 131 years can be approximated by cumulative Weibull distribution functions with parameters α and β. The parameters have been shown to correlate with monthly normal rainfalls. Both parameters tend to increase as monthly rainfall totals increase.

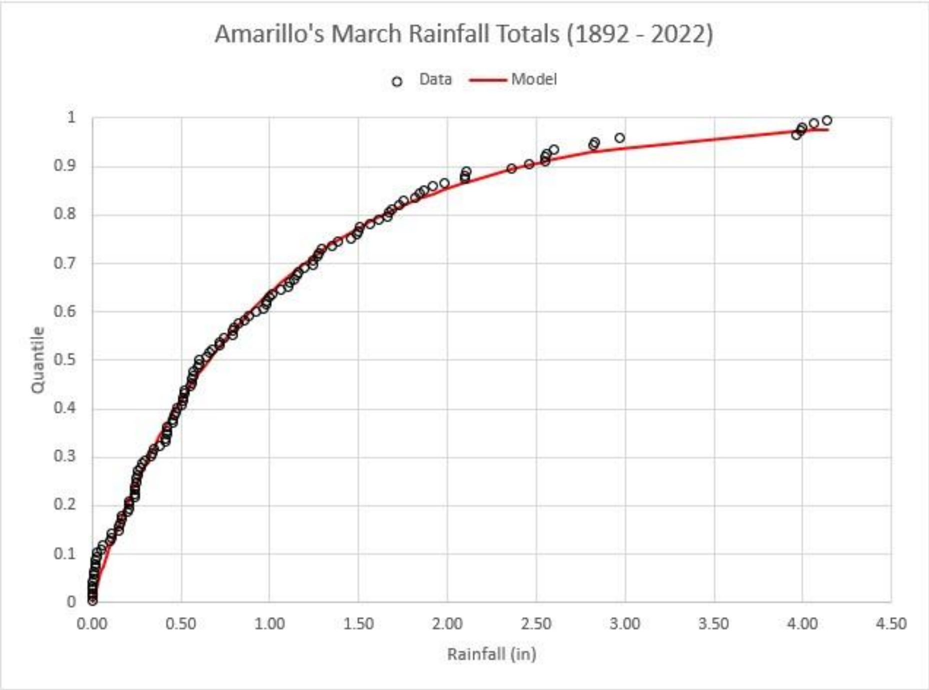
APPENDIX

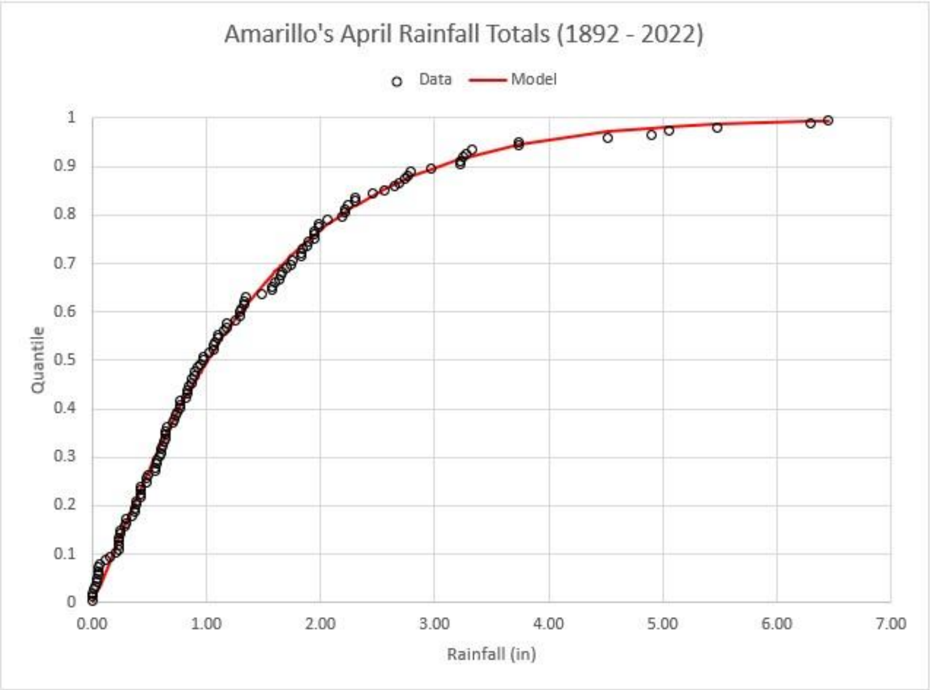


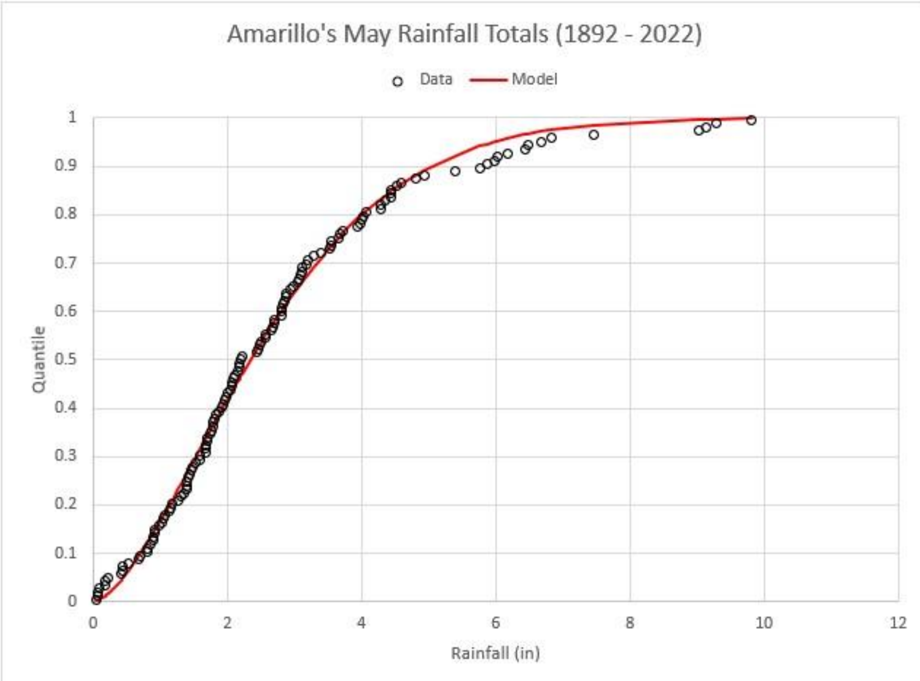
Amarillo's January Rainfall Totals (1892 - 2022)
Data
Model
Quantile
Rainfall (in)

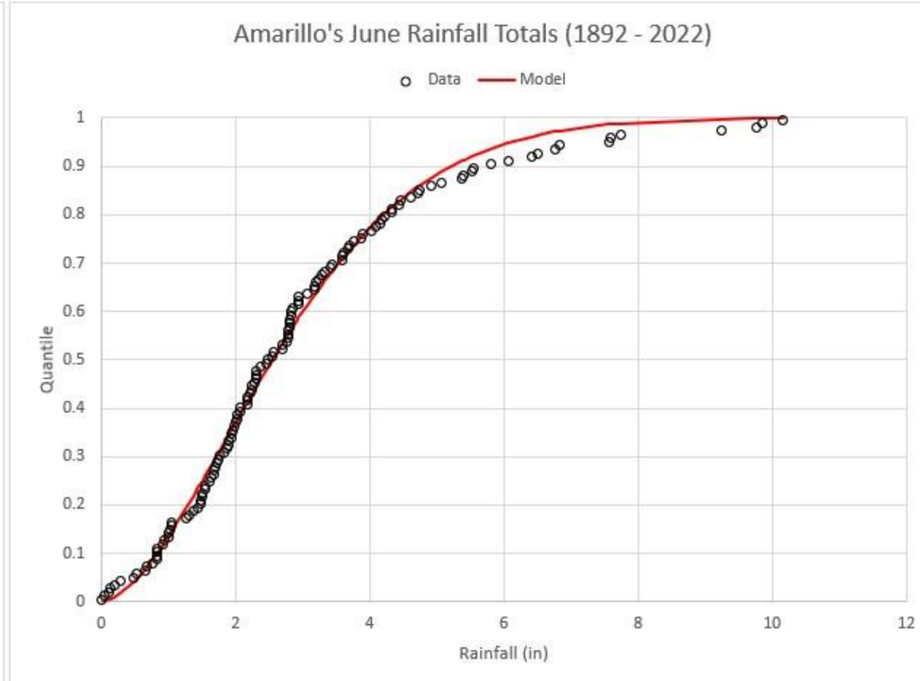


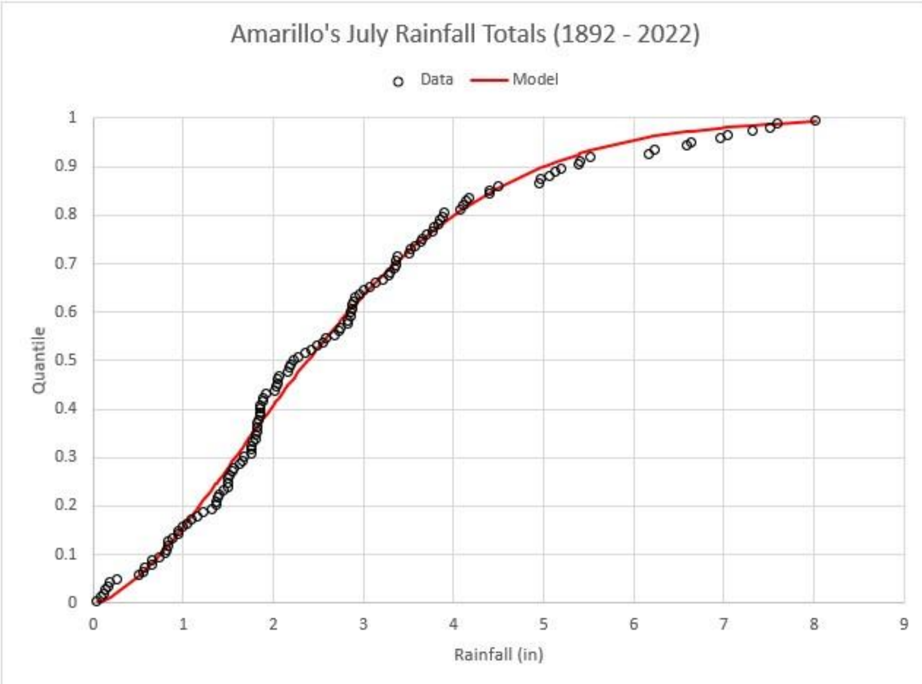
Amarillo's February Rainfall Totals (1892 - 2022)
Data
Model
Quantile
Rainfall (in)

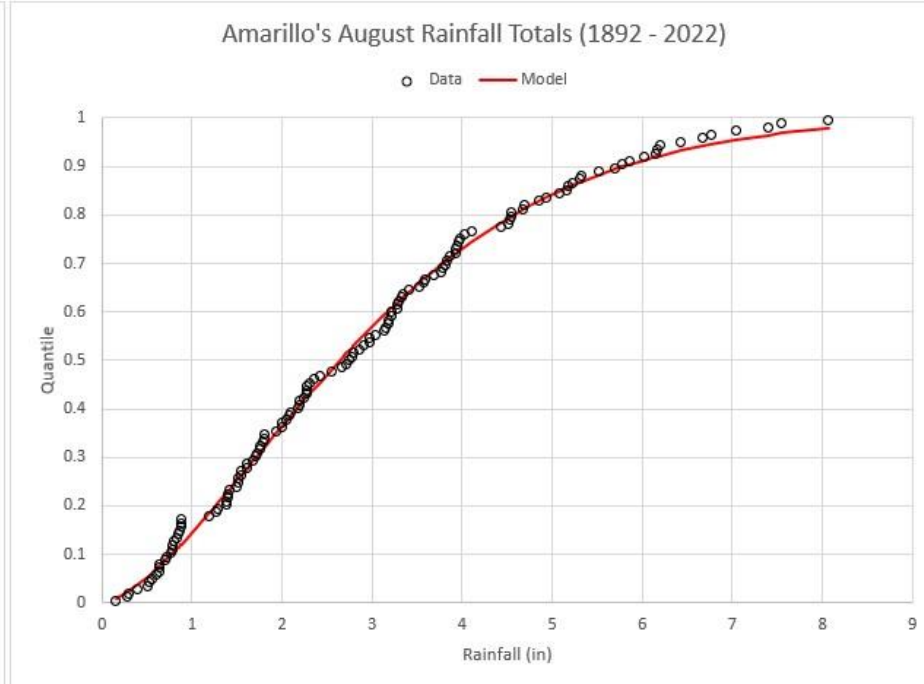

Amarillo's March Rainfall Totals (1892 - 2022)
Data
Model
Quantile
Rainfall (in)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0.00
0.50
1.00
1.50
2.00
2.50
3.00
3.50
4.00
4.50

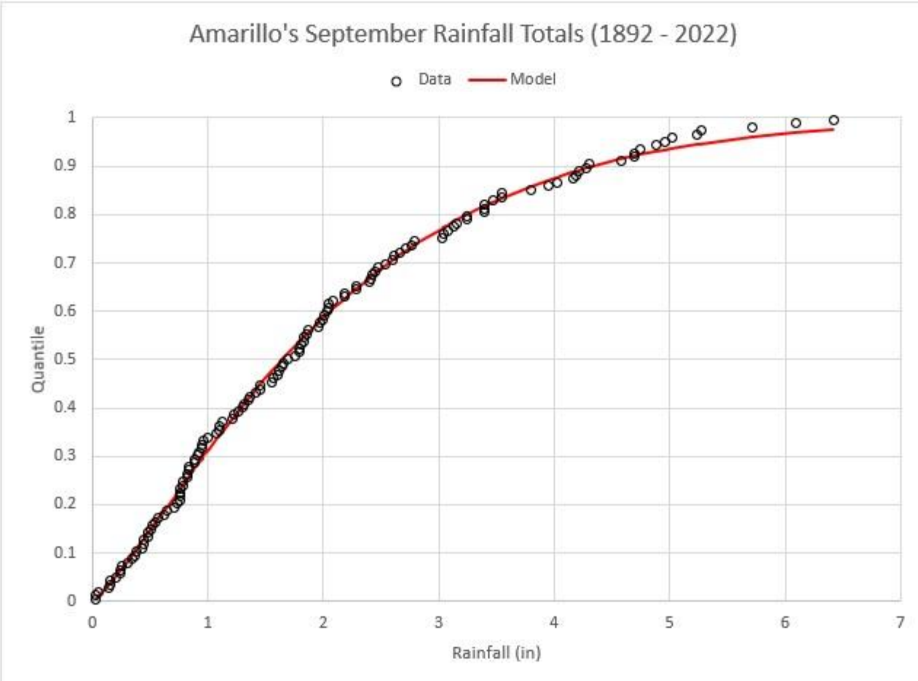

Amarillo's April Rainfall Totals (1892 - 2022)
Data
Model
Quantile
Rainfall (in)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0.00
1.00
2.00
3.00
4.00
5.00
6.00
7.00

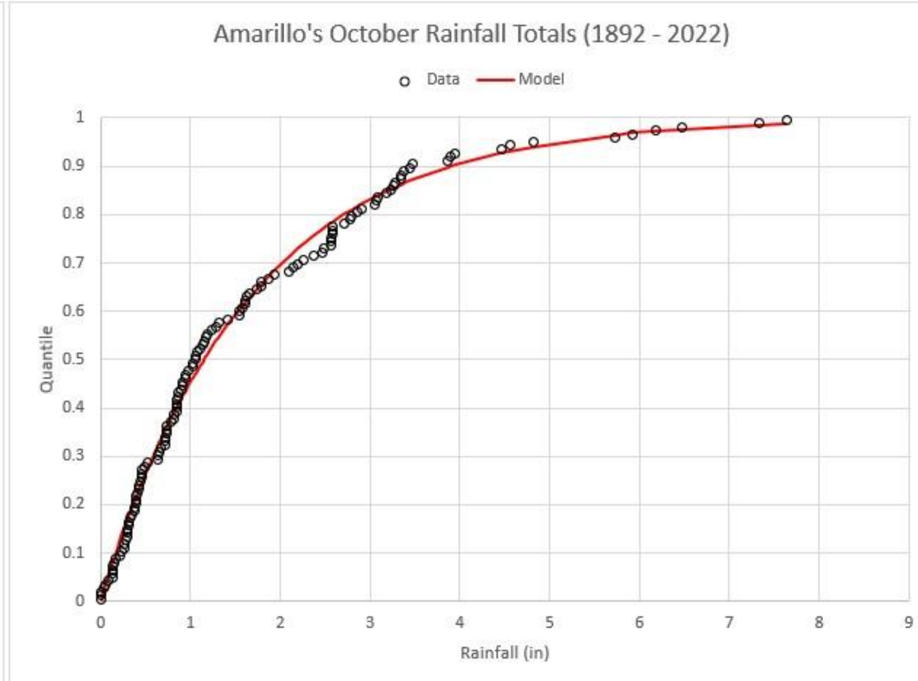

Amarillo's May Rainfall Totals (1892 - 2022)
Data
Model
Quantile
Rainfall (in)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
2
4
6
8
10
12

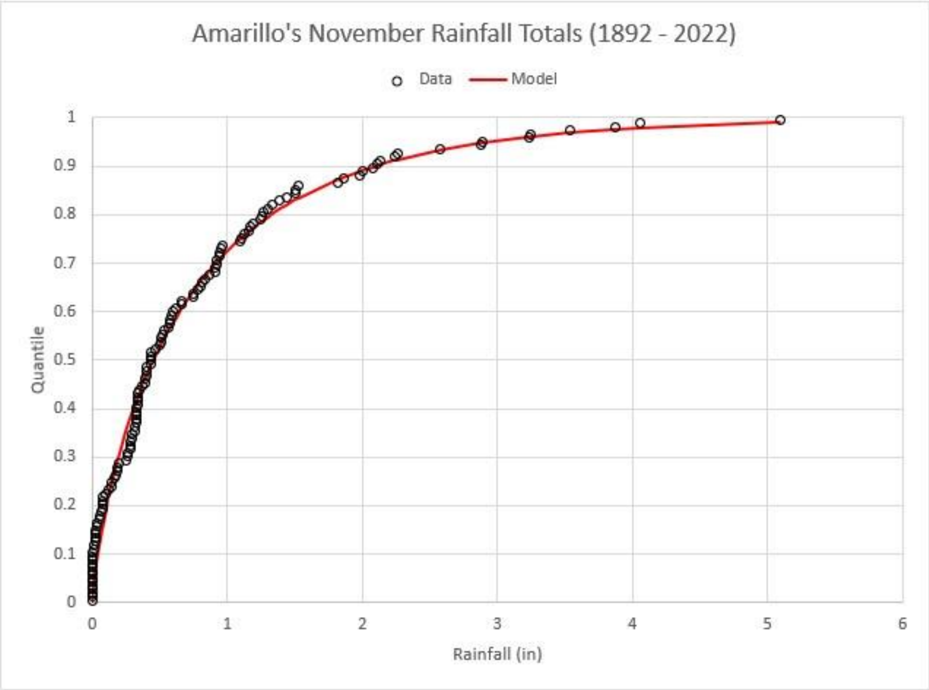

Amarillo's June Rainfall Totals (1892 - 2022)
Data
Model
Quantile
Rainfall (in)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
2
4
6
8
10
12

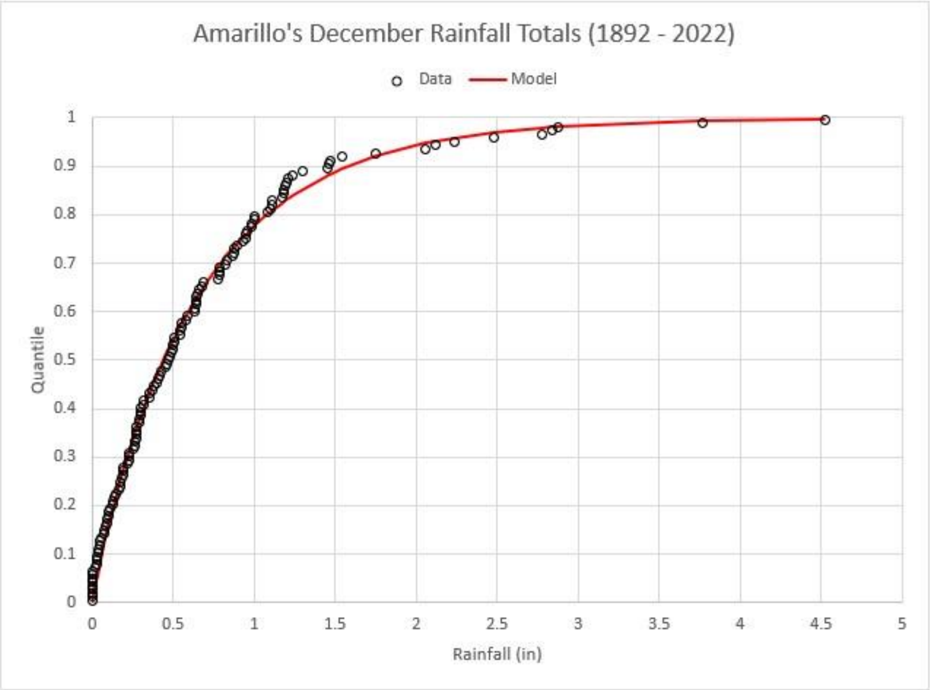

Amarillo's July Rainfall Totals (1892 - 2022)
Data
Model
Quantile
Rainfall (in)


Amarillo's August Rainfall Totals (1892 - 2022)
Data
Model
Quantile
Rainfall (in)


Amarillo's September Rainfall Totals (1892 - 2022)
Data
Model
Quantile
Rainfall (in)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
1
2
3
4
5
6
7


Amarillo's October Rainfall Totals (1892 - 2022)
Data
Model
Quantile
Rainfall (in)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
1
2
3
4
5
6
7
8
9


Amarillo's November Rainfall Totals (1892 - 2022)
Data
Model
Quantile
Rainfall (in)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
1
2
3
4
5
6


Amarillo's December Rainfall Totals (1892 - 2022)
Data
Model
Quantile
Rainfall (in)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5