An Analysis of Amarillo's Yearly Rainfall Totals

By Robert Ashcraft
Retired Program Scientist - Pantex

Amarillo's local weather station has been measuring and reporting yearly rainfall totals since 1892. The driest year so far was 2011, with a total of 7.01 inches. The wettest year was 1923, with a total of 39.75 inches. The rainfall data, consisting of 129 values, can be found at the NOAA Online Weather Data website.

Amarillo's rainfall for 2020 was only 12.53 inches. Only seven of the 129 years on record have been drier.

The purposes of this study were to (1) determine if the rainfall totals behaved like a random sample from a statistical distribution, (2) determine the type of the distribution, and (3) determine the distribution parameters that would provide the best approximation to the data.

The analysis began by generating a 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,

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

where q_i is the i^{th} quantile, i ranges from 1 to n, 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 Figure 1.

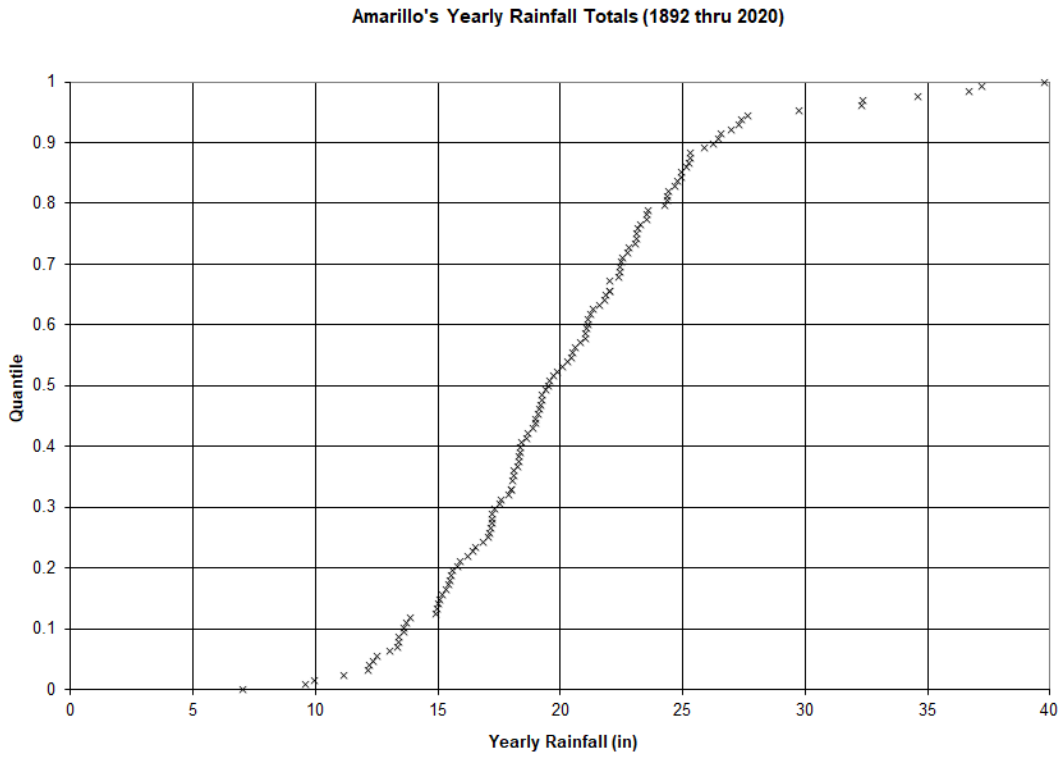


Figure 1. The cumulative distribution of Amarillo's yearly rainfall totals

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 proportion 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 3.80118, 0.25048, and 0.44743, respectively. The resulting curve is shown with the data in Figure 2, and the curve is an excellent approximation to the data.

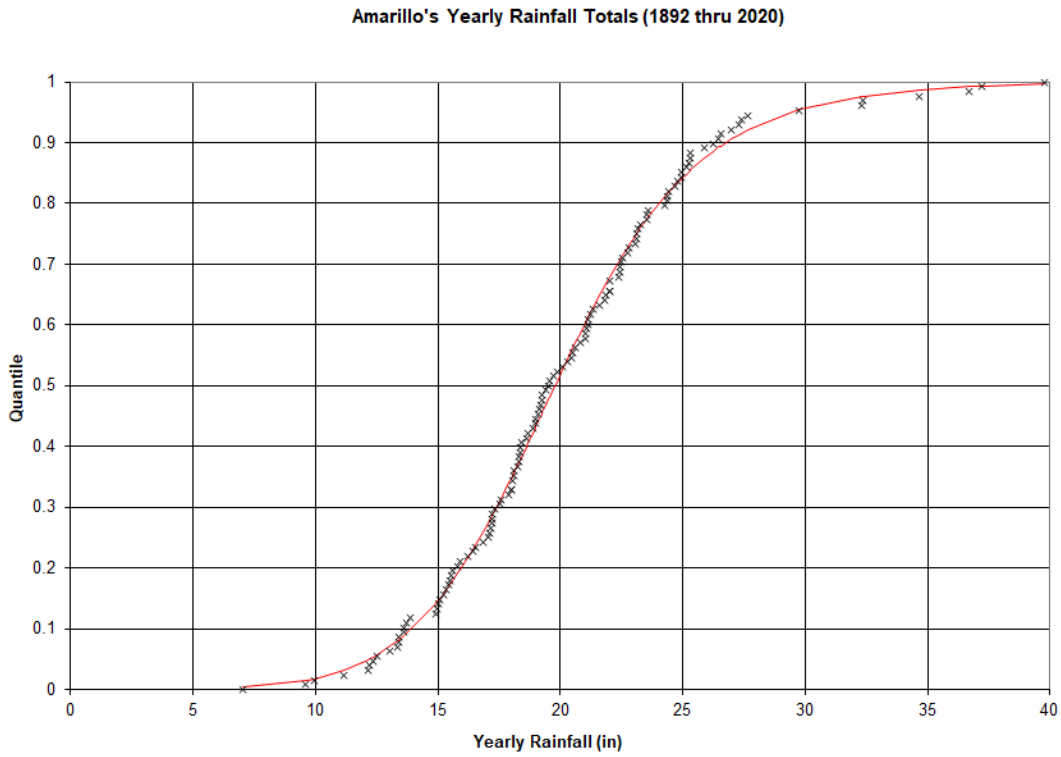
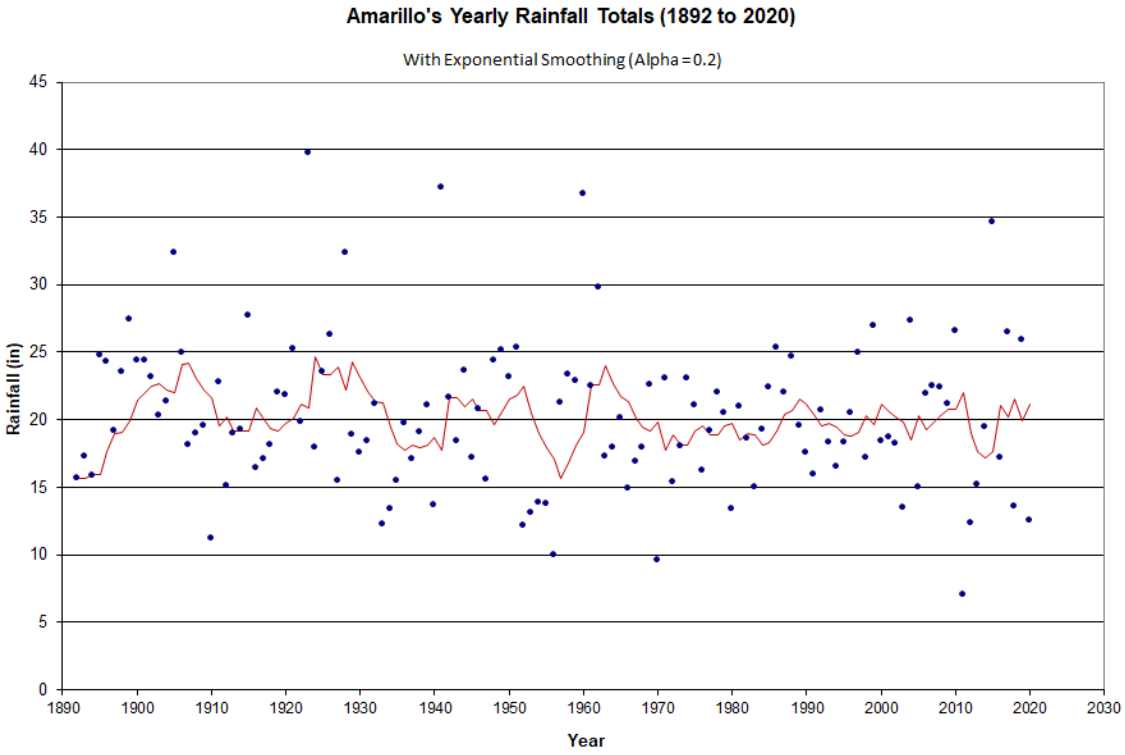


Figure 2. Amarillo's yearly rainfall totals with the best fit sigmoid curve

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.


Amarillo's Yearly Rainfall Totals (1892 to 2020)
With Exponential Smoothing (Alpha = 0.2)
Rainfall (in)
Year
0
5
10
15
20
25
30
35
40
45
1890
1900
1910
1920
1930
1940
1950
1960
1970
1980
1990
2000
2010
2020
2030