Using Extreme Value Theory to Improve Knowledge and Decision Making of Low Probability Events.

Gerald Van Tassell, University of Nebraska-Lincoln, gvantassell2@unl.edu

Dr. Cory Walters, University of Nebraska-Lincoln, cwalters7@unl.edu

Dr. Richard Preston, rkpreston@windstream.net

Dr. Mindy Mallory, Purdue University, mlmallor@purdue.edu

Selected Poster prepared for presentation at the 2023 Agricultural & Applied Economics Association Annual Meeting, Washington DC: July 23- 25, 2023

Using Extreme Value Theory to Improve Knowledge and Decision Making of Low Probability, Financially Devastating Crop Yield Events

Gerald Van Tassell[A], Richard Preston[B], Cory Walters[A], Mindy Mallory[C]

A - Department of Agricultural Economics, University of Nebraska-Lincoln, Lincoln, NE

B- Preston Farms, Hardin County, KY

C- Department of Agricultural Economics, Purdue University, West Lafayette, IN

Background and Objective

- Few data points exist in the low yield area of the yield distribution; therefore, we do not know much about the tail of the yield distribution.
- Other, more devastating events likely exist that have yet to be seen.
  - **Absence of Evidence ≠ Evidence of Absence**
  - Improving our knowledge of these events can lead to improved decision-making.
- We use Extreme Value Theory (EVT) to identify events that have yet to be observed.
- Improving the tails of the yield distribution will allow for better analysis of the impacts of crop insurance and marketing on farm

Methods

Detrending

- Detrending the data represents the foundation for our analysis.
- Our detrending approach must be flexible for use in a diversity of locations and crops while also providing a stationary distribution of residual ratios. Two methods were found to satisfy both criteria.

LOESS (red)- Adapted from Lu *et al*. 2017.

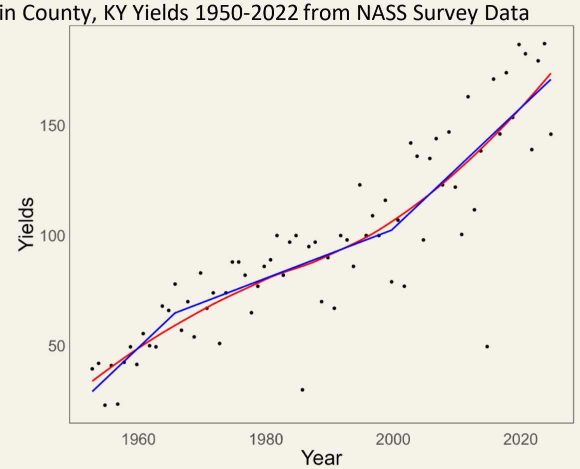
- Weighted quadratic least squares regression with a tri-cube weight function and smoothing parameter fitted using 10-folds cross validation based on the RMSE.
- No global assumptions, locally fitted quadratic curves allow for functional form flexibility

Piecewise Linear Splines (blue)- Introduced by Skees *et al.* 1997

- Divides the data range into distinct intervals and fits separate linear regressions to each interval
- Knot(s) chosen to minimize RMSE

Data

- Hardin County, KY Yields 1950-2022 from NASS Survey Data



Fisher-Tippet-Gnedenko Theorem

- EVT analogue to the Central Limit Theorem.
- Provides the distribution of the maximum of a sample of i.i.d. random variables.
- States that the probability density function of X-u conditional on X>u follows the generalized Pareto distribution regardless of the distribution of X(Coles 2003).
- These conditional densities are referred to as pareto tails and provide an empirical estimate of the tails of the yield distribution without assuming the underlying distribution of crop yields.

Pareto Tail Estimation

- Threshold selection is done according to Coles 2003 using mean excess plot
- Threshold selection is a subjective process especially due to the limited data in the tails of the distribution
- Threshold selected such that the shape parameter is relatively stable regardless of threshold selection

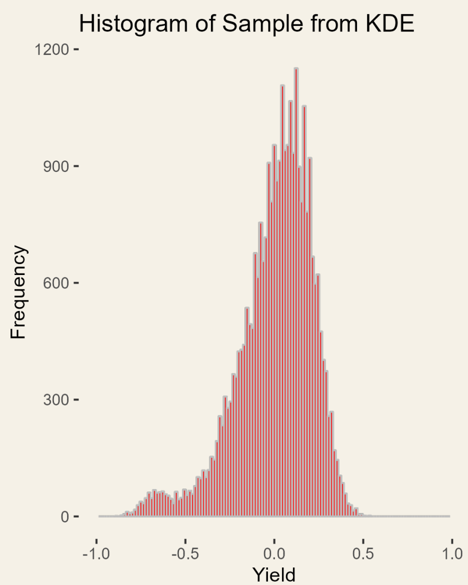
Kernel Density Estimate (KDE)

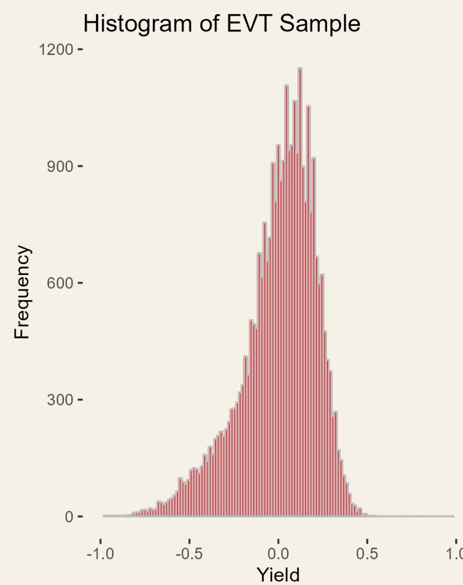
- Estimate of the PDF of yields using a gaussian kernel
- Provides a smooth and continuous estimate of the PDF of yields
- Bandwidth determined by cross-validation
- Does not require an assumption on the underlying distribution

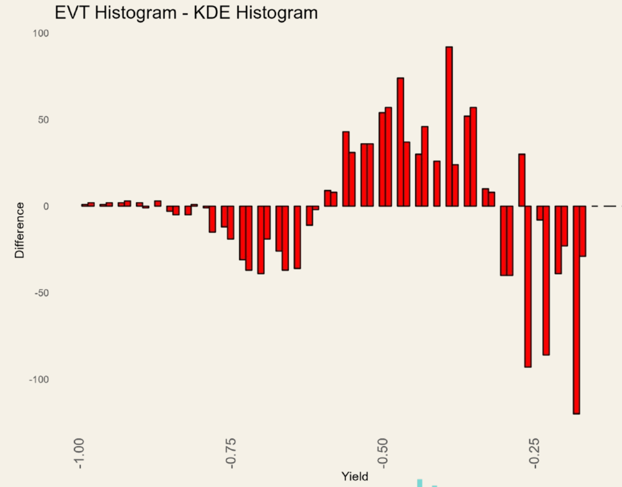
Sampling Method

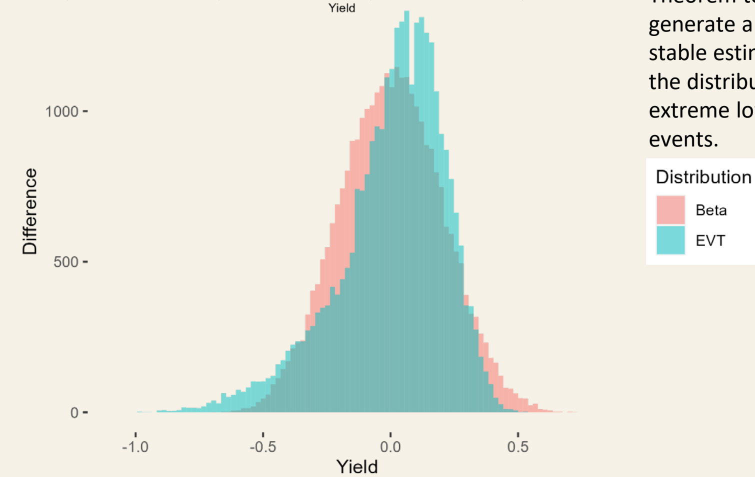
1. Generate a sample from the KDE
2. Replace the observations below the threshold of the pareto tail by sampling from the parameterized GPD to create what we will be calling the "EVT Distribution"

Results









- The use of the pareto tail removes the "hump" in the tail of the distribution.
- The hump is caused by two points that happen to be nearly identical.
- Instead of allowing just two observations to dictate the tail, we employ the Fisher-Tippet-Gnedenko Theorem to generate a more stable estimate of the distribution of extreme low yield events.

Future Directions

- A sample from this yield distribution will be combined with a sample from a price distribution using rank-correlation to create a revenue distribution
- The revenue distribution will be used as an input to a net income function that also takes decisions (crop insurance/marketing) as inputs and outputs a net income distribution
- The improvement to the tails of the yield distribution will improve the accuracy of VaR and ES calculations on the final net income distribution

References

Coles, S., Bawa, J., Trenner, L., & Dorazio, P. (2001). An Introduction to Statistical Modeling of Extreme Values (Vol. 208, p. 208). London: Springer.

Lu, J., Carbone, G. J., & Gao, P. (2017). Detrending crop yield data for spatial visualization of drought impacts in the United States, 1895–2014. Agricultural and Forest Meteorology, 237, 196-208.

Skees, J. R., Black, J. R., & Barnett, B. J. (1997). Designing and rating an area yield crop insurance contract. American Journal of Agricultural Economics, 79(2), 430-438.