Probit Regression for Causality

- Reference:
  - Tuttle, S., and G. Salvucci, 2016: Empirical evidence of contrasting soil moisture–precipitation feedbacks across the United States. *Science*, **352**, 825-828, doi: 10.1126/science.aaa7185.

- Principle:
  - A generalized linear model of precipitation occurrence with independent predictors of lagged soil moisture (one day), atmospheric pressure (1-4 days) and precipitation (1-4 days, plus sinusoidal terms to represent persistent seasonal and interannual variations) is optimized using the Akaike Information Criterion (AIC) to optimize the best fit with minimum parameters. The same is done with soil moisture as the predictand to isolate the direct influence of soil moisture on precipitation.
  - The probit regression indicates daily probability of precipitation. Because of the endogenous relationship between precipitation and soil moisture, the correlation due to the forward relationship between the two (that soil moisture integrates past precipitation regardless of any other sources of lagged autocorrelation of precipitation) is estimated by a bootstrap technique and removed.
  - The Granger Causality paradigm is used to determine significant impact of past soil moisture on precipitation.

- Data needs:
  - Daily soil moisture, precipitation, surface pressure (indicator of synoptic/dynamical sources of precipitation), preferably over several decades. Easily applied to Earth system model output and reanalyses.

- Observational data sources:
  - Well suited to meteorological station data.

- Caveats:
  - Indicates whether soil moisture is a contributor or not to increased predictability, but not what sign of anomaly contributes to what sign of response without further processing of the data/method (e.g., compositing by soil moisture anomaly, cf. Tuttle and Salvucci 2016).
  - The "false correlation" element is a key factor – one can demonstrate this by calculating lagged correlations between precipitation and lagged soil moisture using a simple soil moisture model where precipitation is specified as a "noisy" input without possibility of feedback (e.g., Entekhabi & Rodriguez-Iturbe, 1994):

$$\frac{dS(t)}{dt} = P(t) - \eta S(t)$$

Supported by NASA/MAP Grant: NNX13AQ21G