Multivariate Frequency Analysis of Extreme Rainfall Events using Copulas

Shih-Chieh Kao, Ph.D.

Post-doctoral Research Associate
School of Civil Engineering
Purdue University
email: kao@purdue.edu
web: http://web.ics.purdue.edu/~kao

Extreme Events

Feb. 5, 2008
Delphi, Indiana
Flooding of Tippecanoe River

(AP Photo/Journal & Courier, Michael Heinz)

Sept., 2007
George H. Sparks Reservoir
Lithia Springs, Georgia

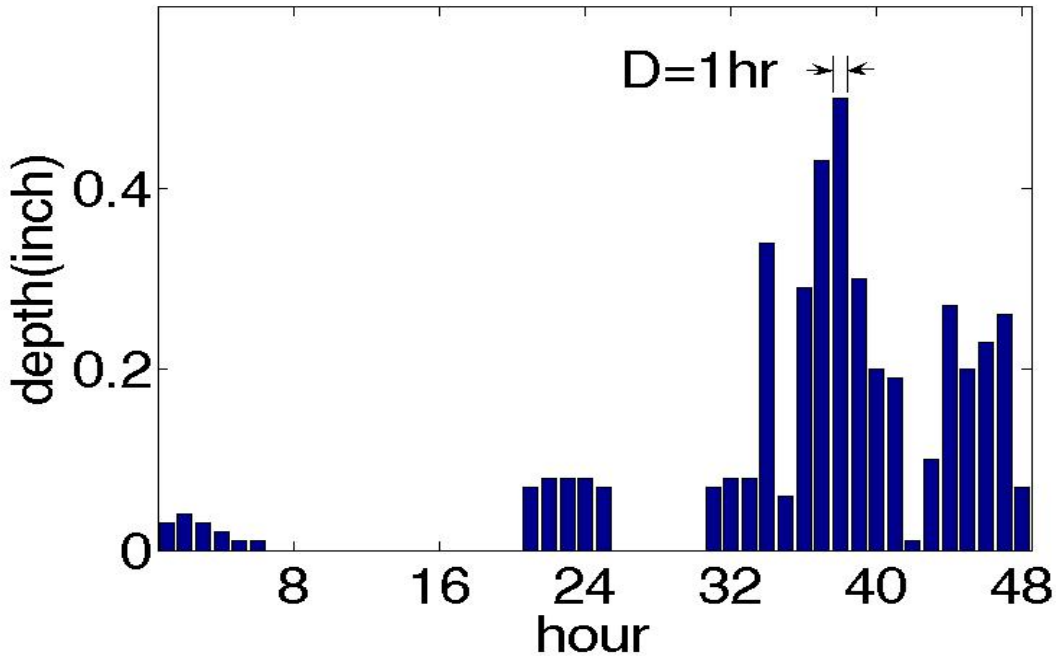
(Barry Gillis, http://www.drought.unl.edu/ gallery/ 2007/Georgia/Sparks1.htm)

Outline

- Background and motivation
  - Limitations in univariate approach
- Introduction to copulas
- Research objectives
  - Topic 1: Probabilistic structure of surface runoff
  - Topic 2: Extreme rainfall frequency analysis
  - Topic 3: Drought frequency analysis
- Summary and concluding remarks

Limitations in Univariate Approach

- Eg: Selection of annual maximum precipitation events
  - *Durations* are not the actual durations of rainfall events
  - Long-term maximum may cover multiple events
  - Short-term maximum encompasses only part of the extreme event

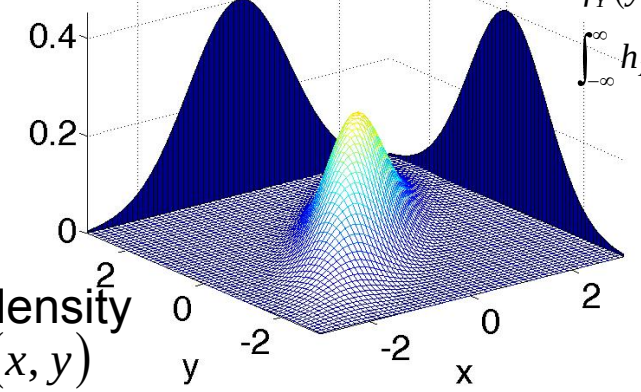


Bivariate Distribution Example

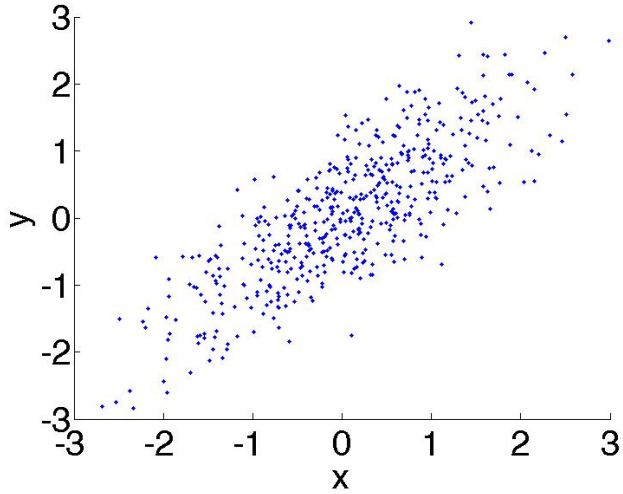
Bivariate Gaussian distribution, $\rho = 0.8$

Marginals $f_X(x) = \int_{-\infty}^{\infty} h_{XY}(x, y)dy$

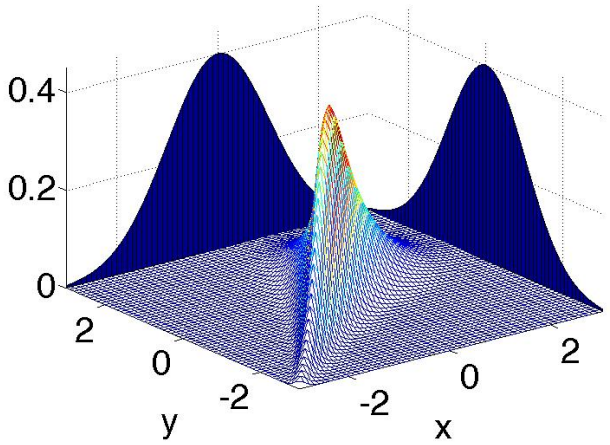
$f_Y(y) = \int_{-\infty}^{\infty} h_{XY}(x, y)dx$

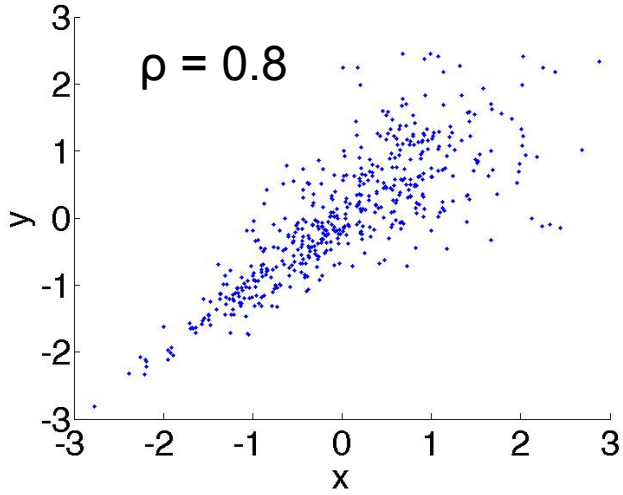


Joint density $h_{XY}(x, y)$



Gaussian marginals with Clayton Copulas



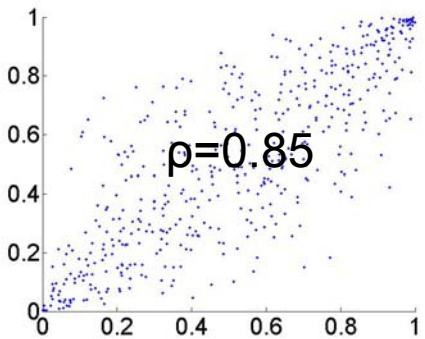


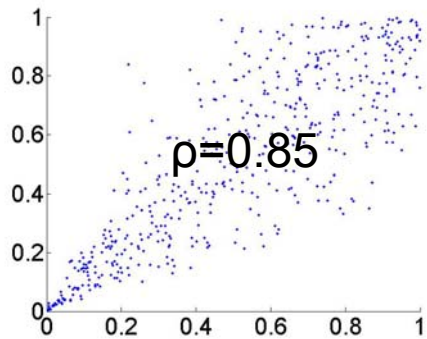
Concept of Dependence Structure

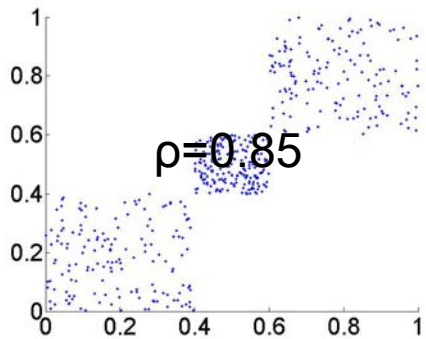
- Conventionally quantified by the linear correlation coefficient ρ

$$\rho_{XY} = \frac{E[(X-\bar{x})(Y-\bar{y})]}{Std[X]Std[Y]}$$

  - Only valid for Gaussian (or some elliptic) distributions



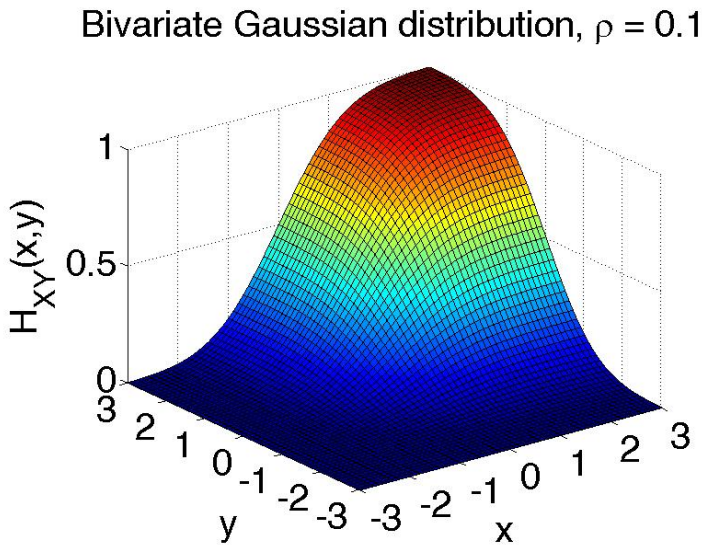


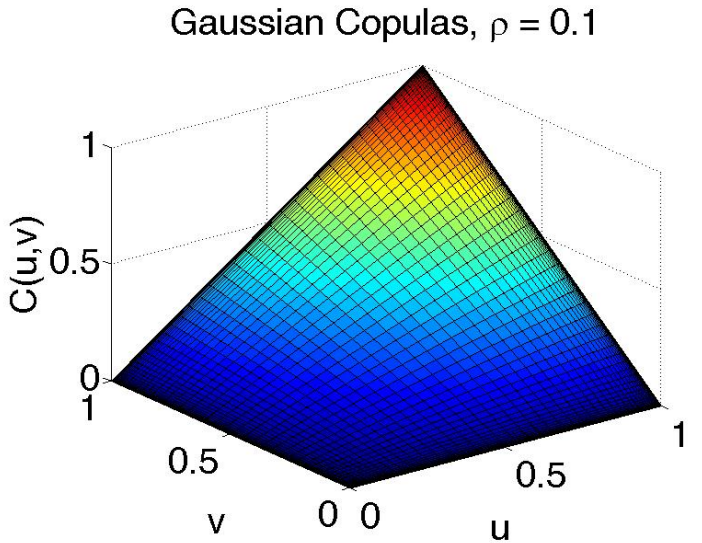


  - Can not correctly describe association between variables
  - A better tool is required to characterize dependence

Copulas

- Transformation of joint cumulative distribution
  - $H_{XY}(x,y) = C_{UV}(u,v)$
    marginals: $u = F_X(x)$, $v = F_Y(y)$
  - Sklar (1959) proved that the transformation is *unique* for continuous r.v.s
- Use copulas to construct joint distributions
  - Marginal distributions => selecting suitable PDFs
  - Dependence structure => selecting suitable copulas
  - Together they form the joint distribution

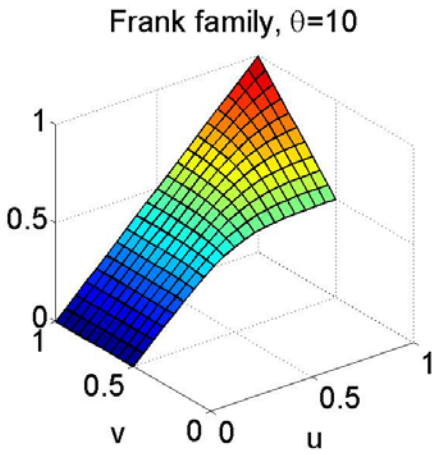


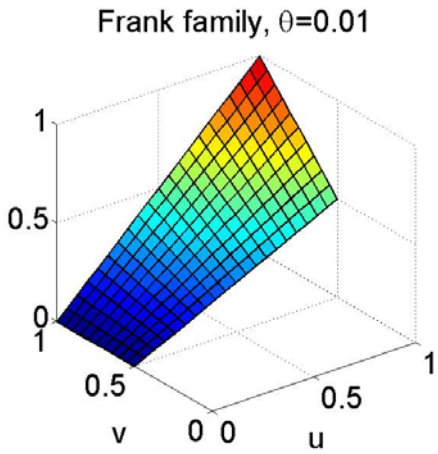


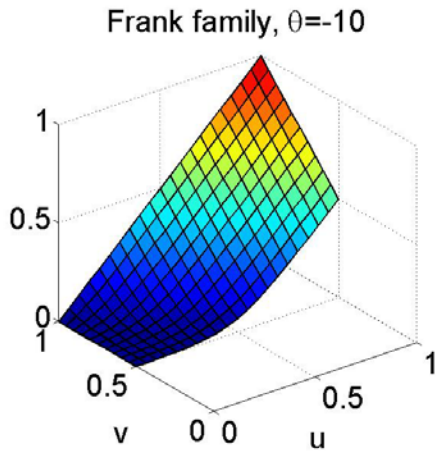
Example of Copulas – Frank Family

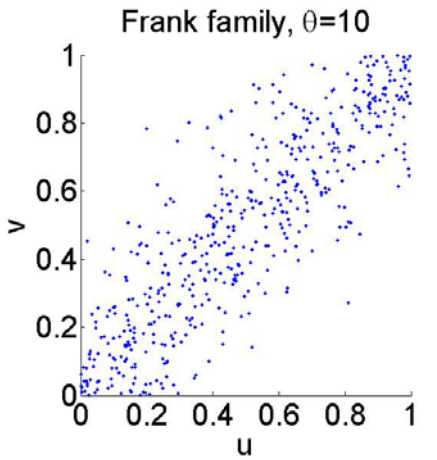
- Frank family of Archimedean copulas

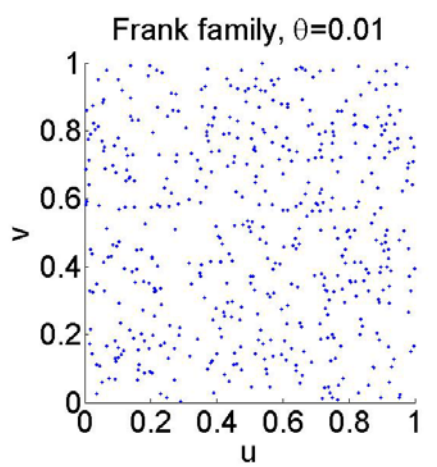
$$C_{Frank}(u,v) = -\frac{1}{\theta}\ln\left(1+\frac{\left(e^{-\theta u}-1\right)\left(e^{-\theta v}-1\right)}{e^{-\theta}-1}\right)$$

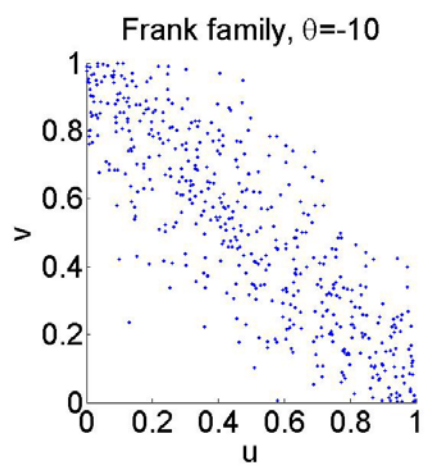












Use of Copulas in Hydrology

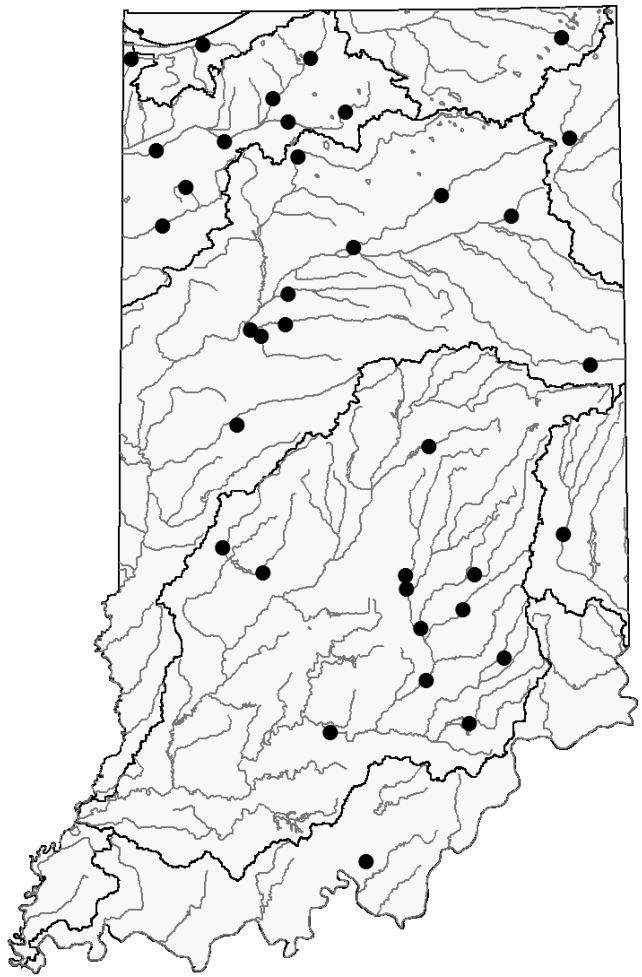
- Since 2003, over 20 research articles have been published in water resources related journals
  - Rainfall and flood frequency analysis, groundwater parameters estimation, sea storms analysis, rainfall IDF curves, and *etc*.
  - Full potential of copulas is yet to be realized (Genest and Favre, 2007)

- For copulas in rainfall frequency analysis:
  - How dependence affects rainfall depth and surface runoff?
  - The definition of extreme events was not clear
  - Few stations were examined

- For copulas in drought frequency analysis:
  - What's the dependence between various drought indices?

Data Sources & Study Area

- Precipitation
  - NCDC hourly precipitation dataset
    - 53 stations with record length greater than 50 years
  - NCDC daily precipitation dataset
    - 73 stations with record length greater than 80 years
- Streamflow
  - USGS unregulated daily mean flow
    - 36 stations with record length greater than 50 years



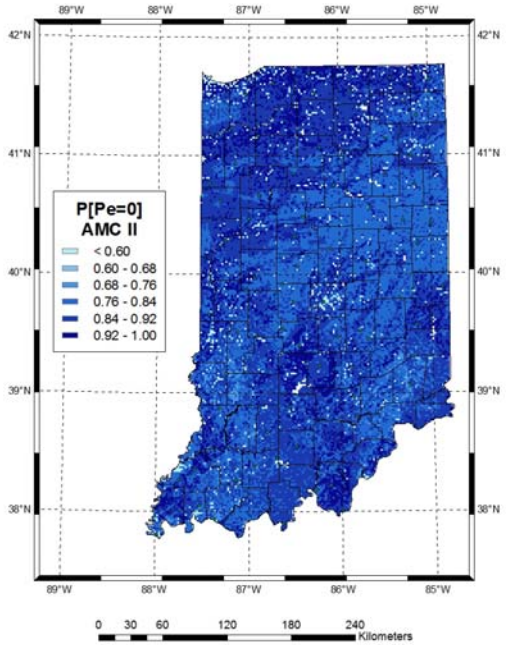
Topic 1

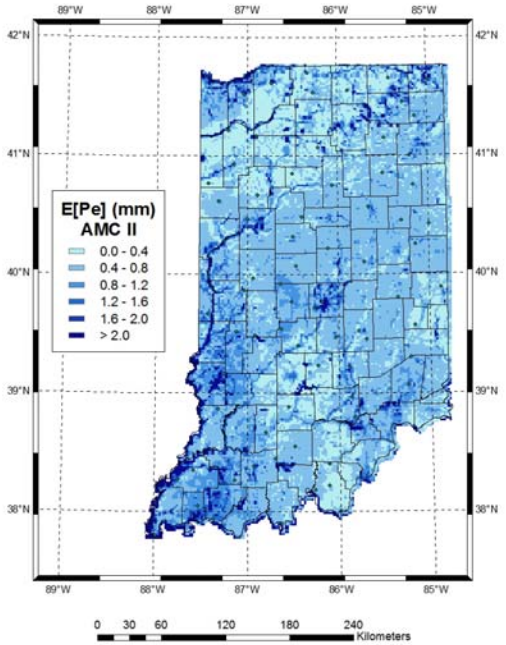
Probabilistic Structure of Surface Runoff

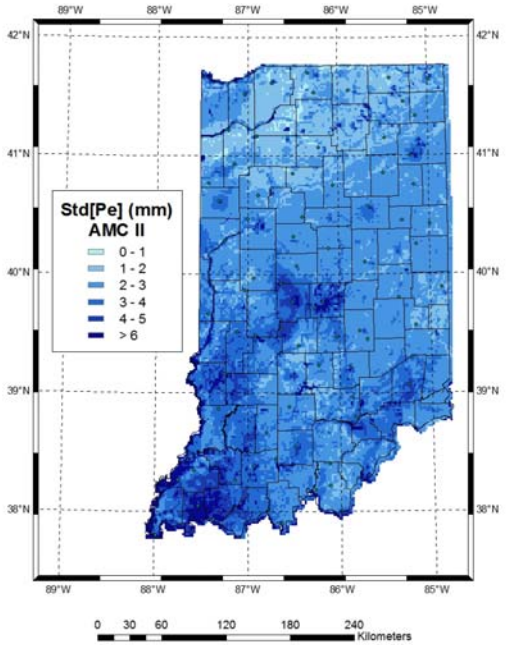
- SCS CN method for excess rainfall depth P_e

$$P_e = \frac{(I \cdot D - 0.2S)^2}{(I \cdot D + 0.8S)}$$

  - A function of both average intensity I and rainfall duration D
  - How does the dependence between I & D affect P_e?
  - GIS maps of $P[P_e=0]$, $E[P_e]$, and $Std[P_e]$ for Indiana



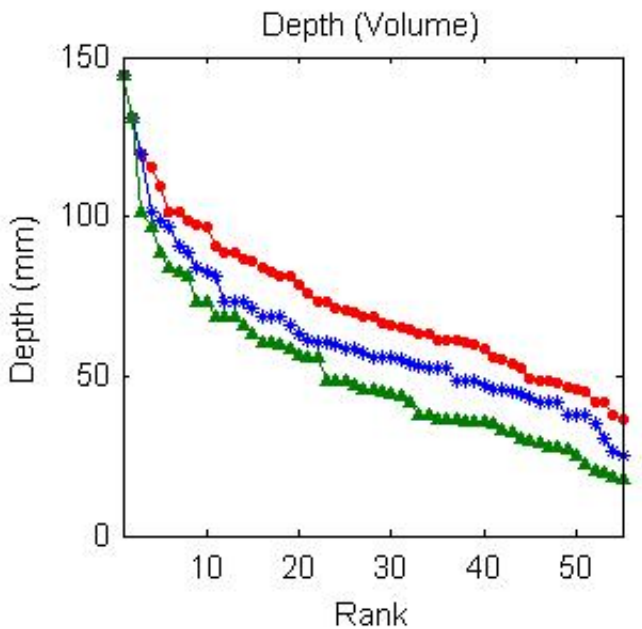


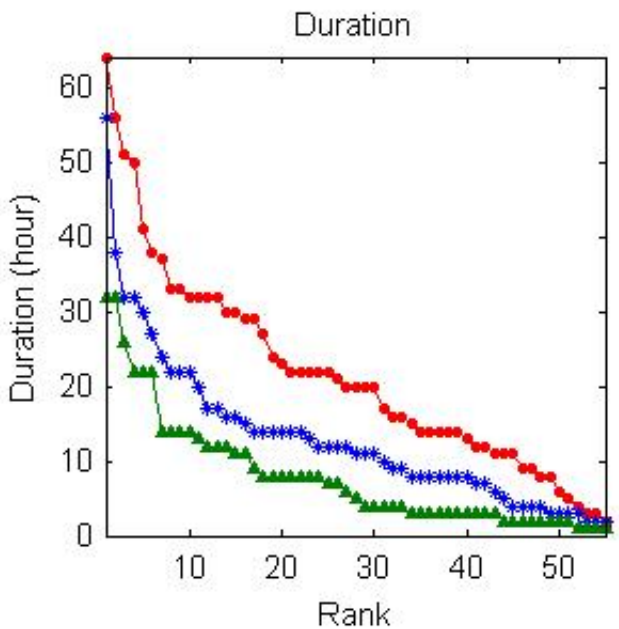


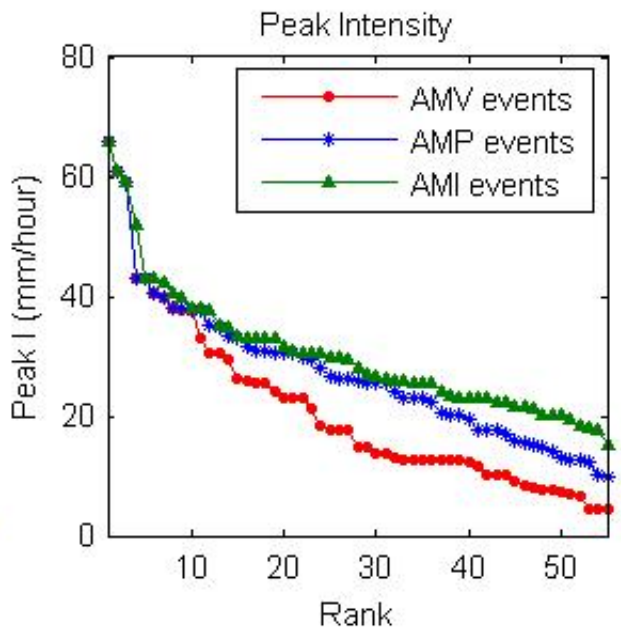
Topic 2

Extreme Rainfall Frequency Analysis (I)

- Definitions of Extreme Rainfall Events
  - Hydrologic designs are usually governed by depth (volume) or peak intensity
  - Annual maximum volume (AMV) events
    - Longer duration
  - Annual maximum peak intensity (AMI) events
    - Shorter duration
  - Annual maximum cumulative probability (AMP) events
    - The use of empirical copulas between volume and peak intensity
    - Wide range of durations







Topic 2

Extreme Rainfall Frequency Analysis (II)

- Bivariate distribution H_{PD}, H_{DI}, H_{PI}
  - Marginal: Extreme Value Type I (EV1), Log Normal (LN)
  - Dependence: Frank Family
- Applications
  - Estimate of depth for known duration

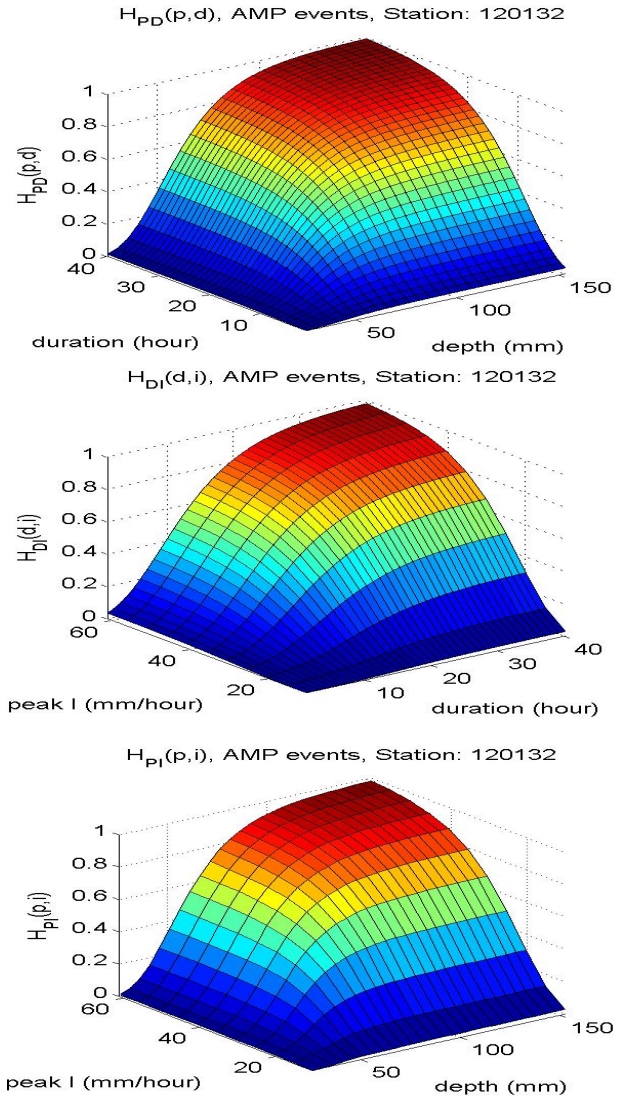
$$F_P\left(p_T \middle| d-1 < D \le d\right) = 1 - 1/T$$

  - Estimate of peak intensity for known duration

$$F_I\left(i_T \middle| d-1 < D \le d\right) = 1 - 1/T$$

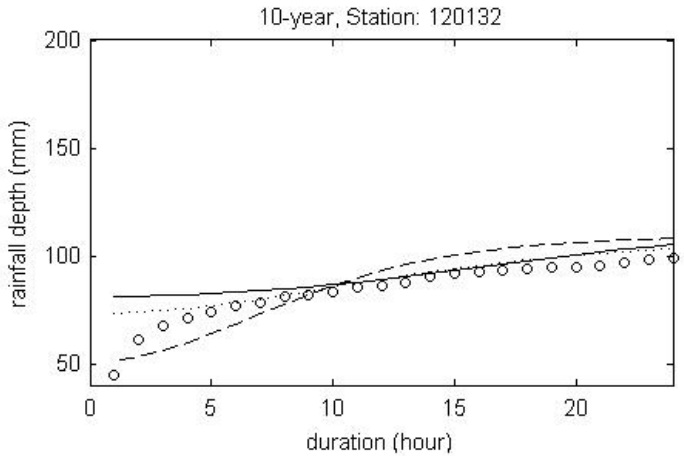
  - Estimate of peak intensity for known depth

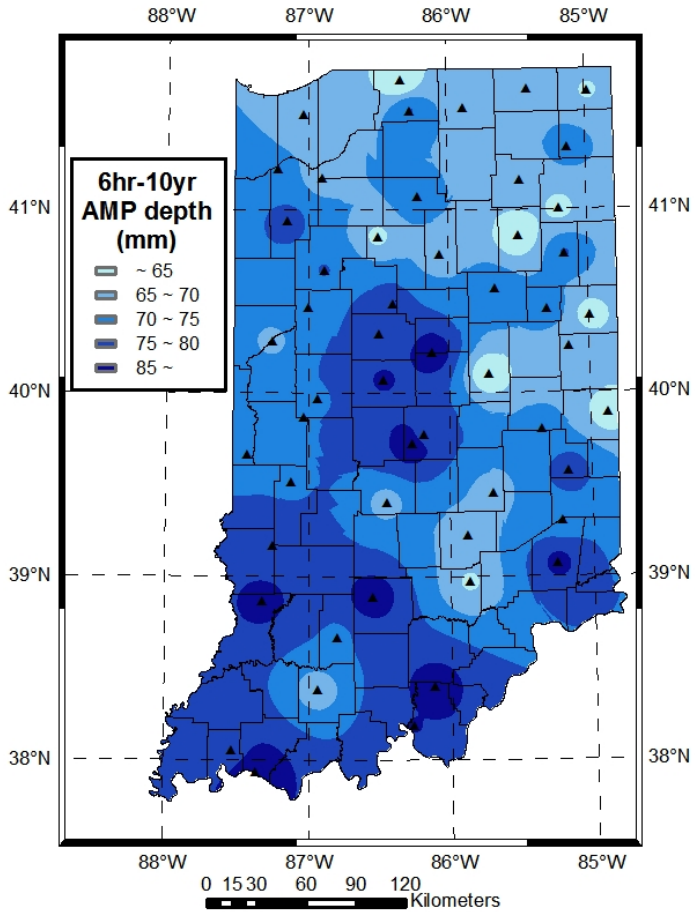
$$E[I \mid P > p]$$



Estimate of depth for known duration

T-year depth p_T given duration d

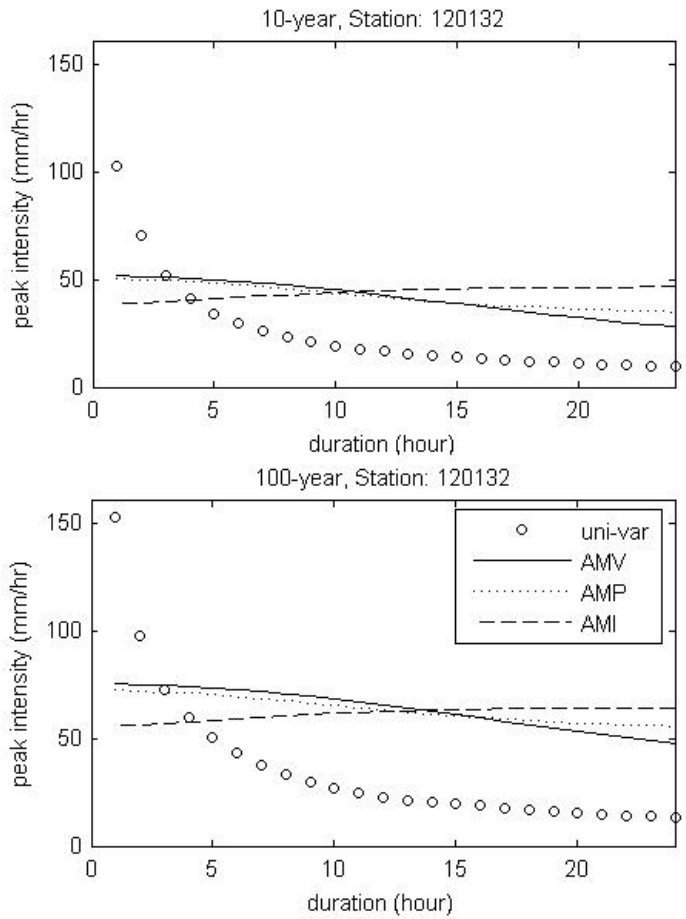


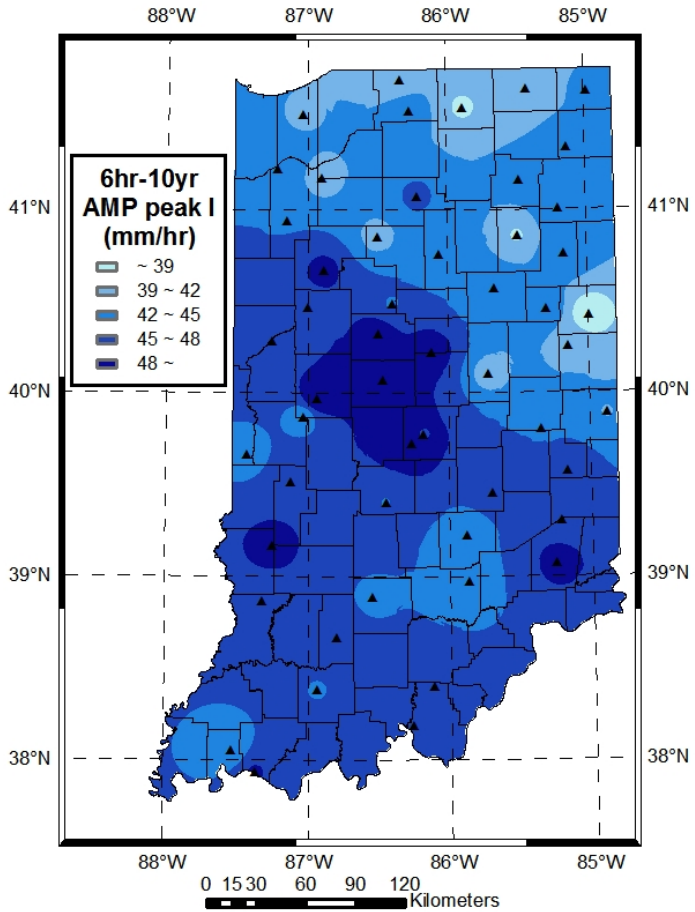


- The estimates of depth are similar for durations larger than 10 hours
- AMP definition seems to be an appropriate indicator for defining extreme events

Estimate of peak intensity for known duration

T-year peak intensity i_T given duration d

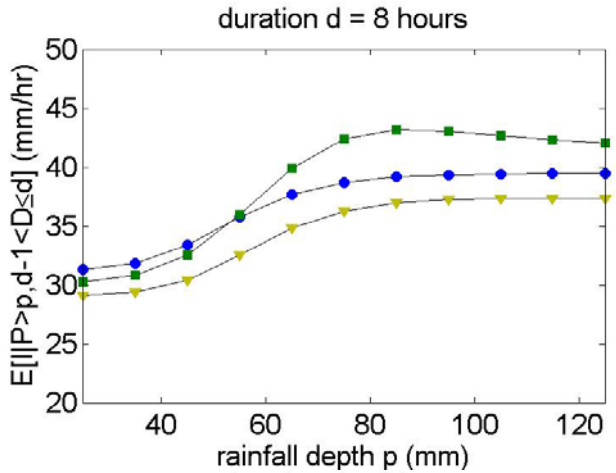


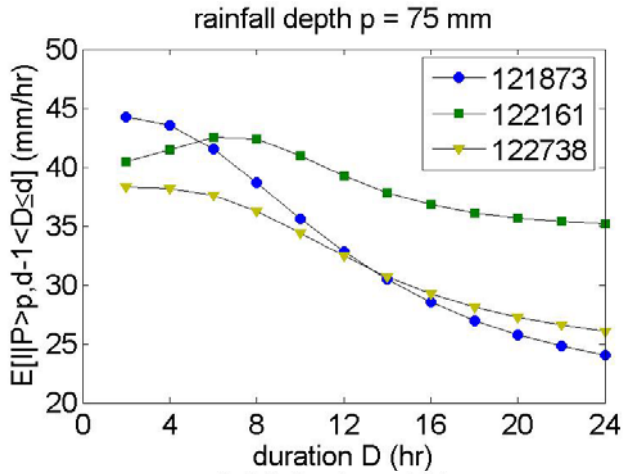


- Conventional approach fails to capture the peak intensity

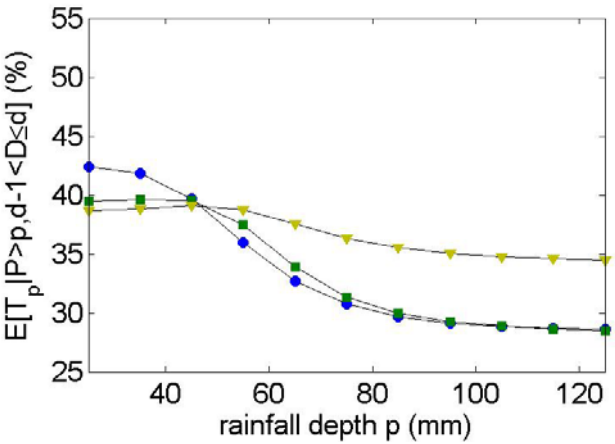
Rainfall Peak Attributes

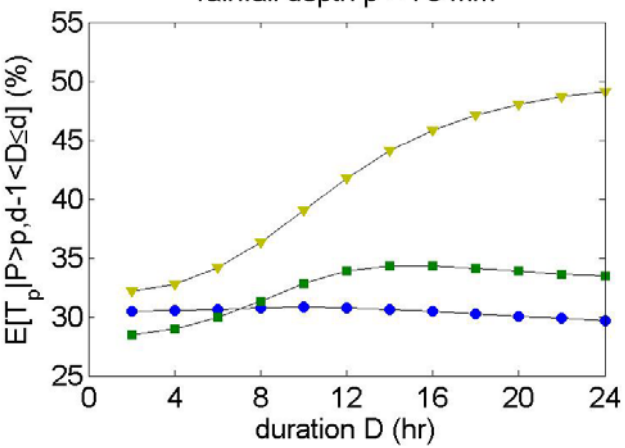
- Given depth (P) and duration (D), compute the conditional expectation of peak intensity (I) and percentage time to peak (T_p)





Expectation of peak intensity given P & D

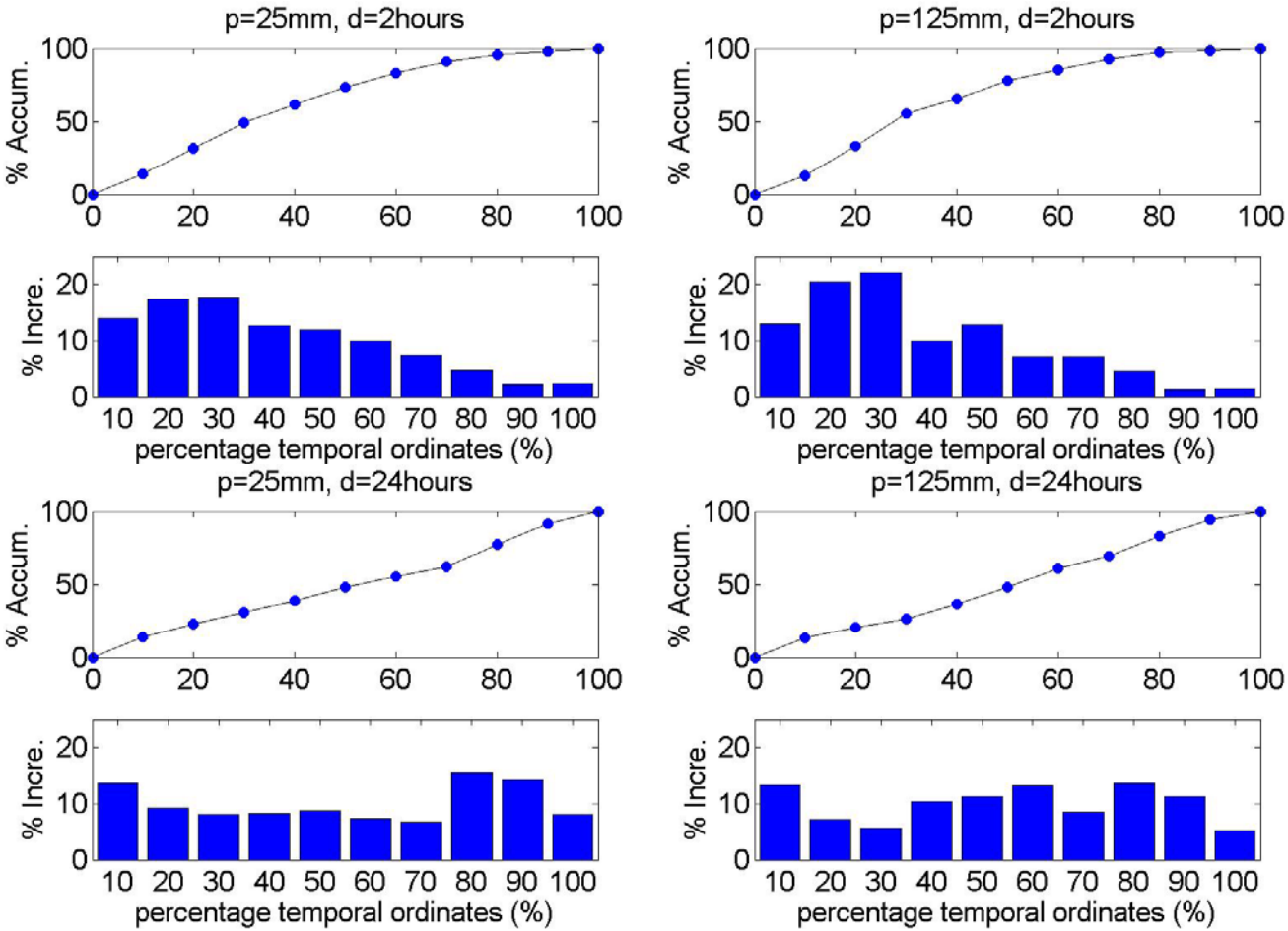




Expectation of time to peak (%) given P & D

Temporal Accumulation Curves

- Given depth (P) and duration (D), compute the conditional expectation of percentage accumulations at each 10% temporal ordinates (A_{10}, A_{20}, …, A_{90})



Expectation of % accumulation given P & D

Topic 3

Drought Frequency Analysis

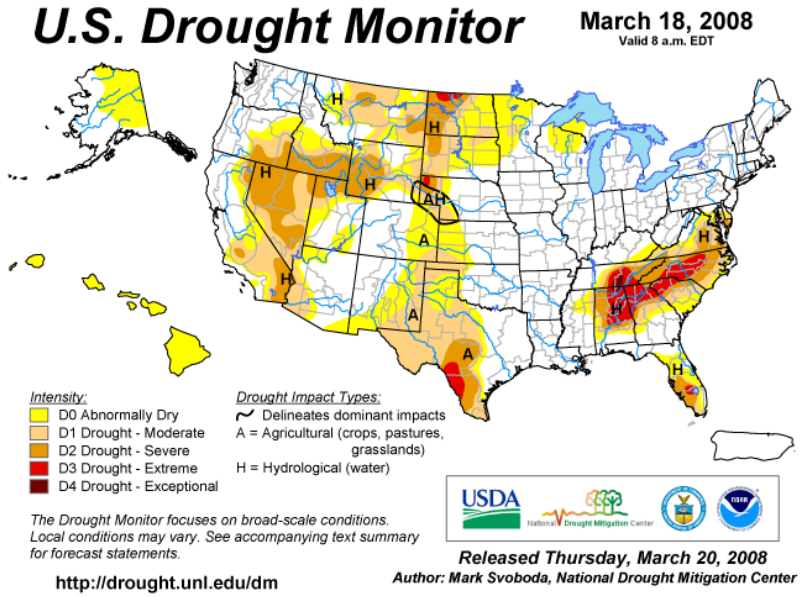
- Challenges in characterizing droughts
  - No clear (scientific) definition: deficit of water for prolonged time
  - Phenomenon dependent in time, space, and between various variables such as precipitation, streamflow, and soil moisture

- Classification of droughts
  - Meteorological drought: precipitation deficit
  - Hydrologic drought: streamflow deficit
  - Agricultural drought: soil moisture deficit

- Various drought indices
  - Palmer Drought Severity Index (PDSI), Crop Moisture Index (CMI), Surface Water Supply Index (SWSI), Vegetation Condition Index (VCI), CPC Soil Moisture, Standardized precipitation index (SPI)

US Drought Monitor

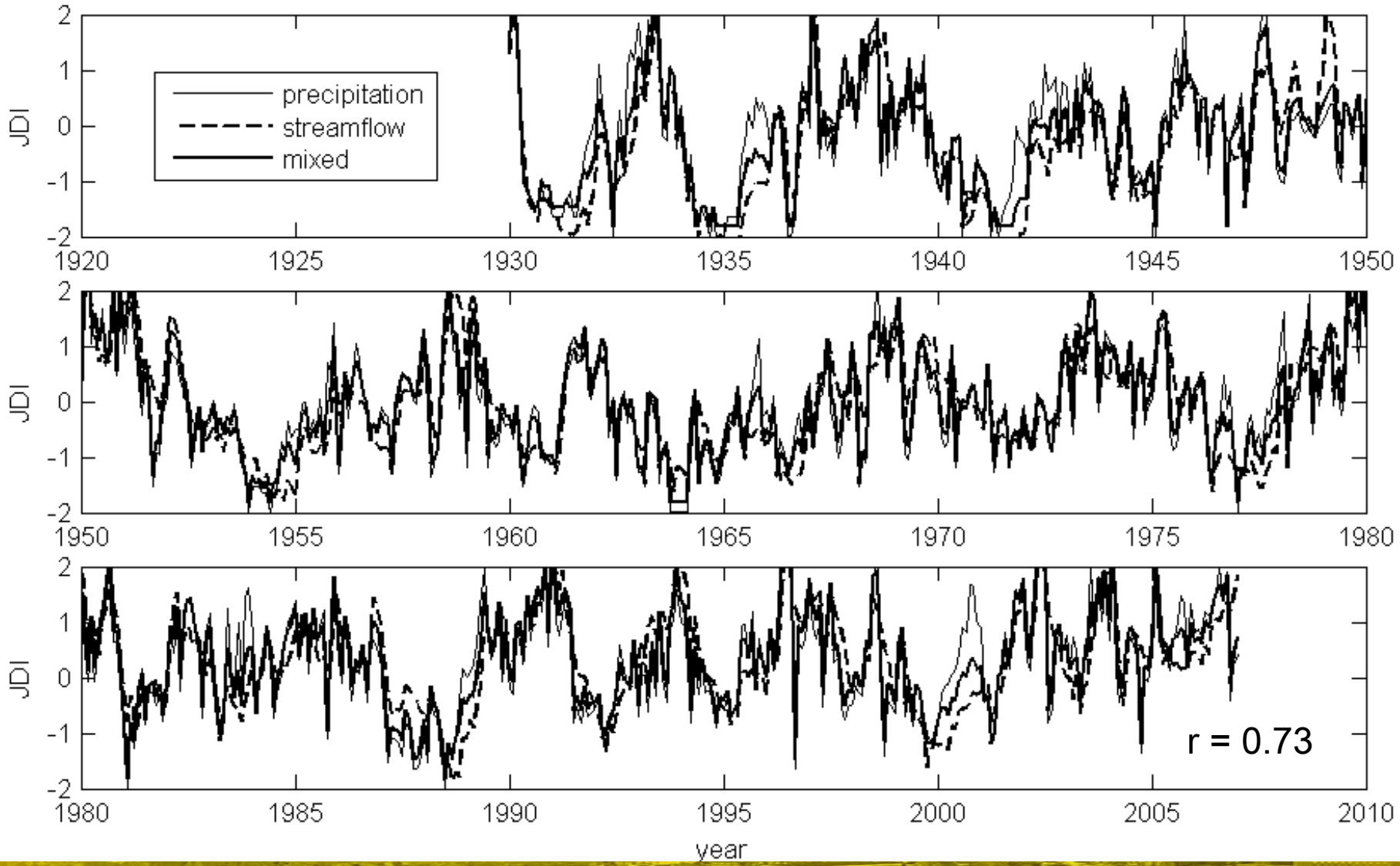
- Overall drought status (D0 ~ D4) determined based on various indices together (Svobada *et al.*, 2002)
  - PDSI
  - CPC Soil moisture
  - USGS weekly
  - Percentage of normal
  - SPI
  - VCI



http://drought.unl.edu/dm/monitor.html

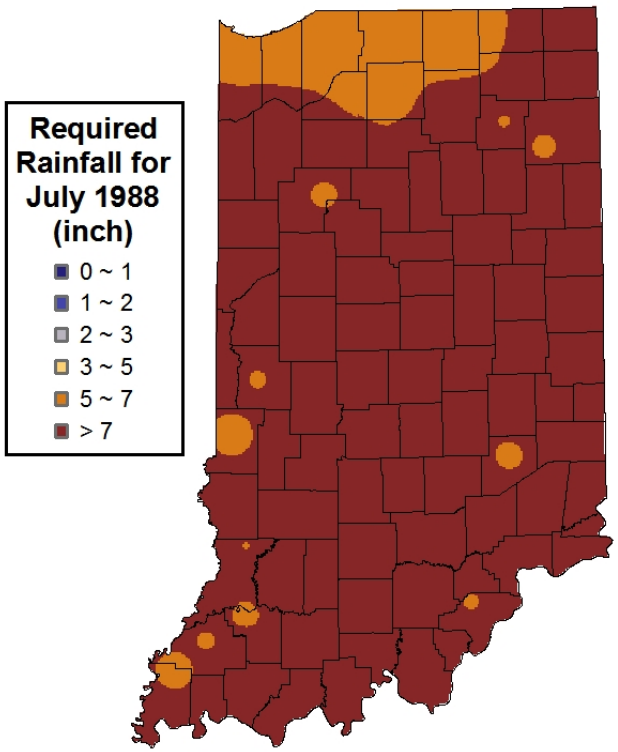
- Linear combination of selected indices (OBDI, objective blend of drought indicator) was adopted as the preliminary overall drought status
- The decision of final drought status relies on subjective judgment
- Copulas can be adopted to compute the joint distribution of drought indices, and then used to compute the joint deficit index (JDI)

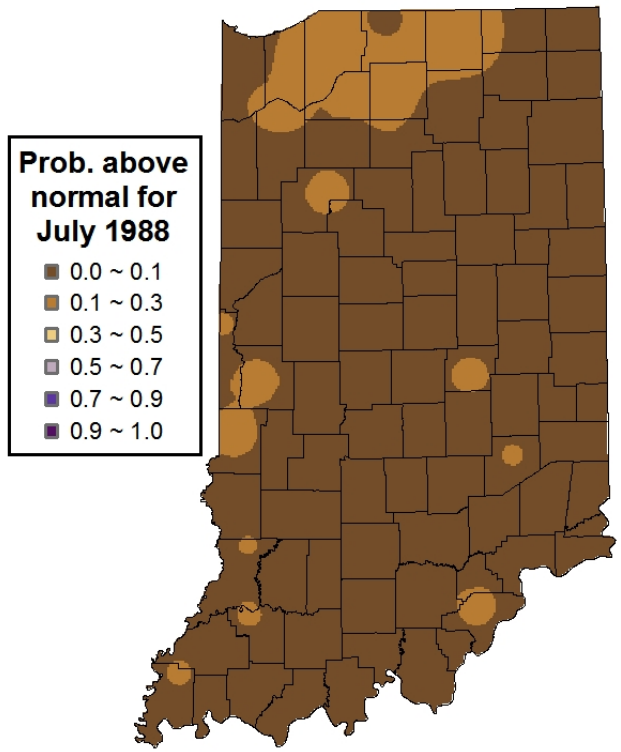
Precipitation and Streamflow JDIs



Potential of Future Droughts

- Required precipitation for reaching joint normal status (cumulative probability to be 0.5) in the future
- Probability of drought recovery





Summary and Concluding Remarks

- Copulas are found to be flexible for constructing joint distributions (no specific marginals are required).
- The dependence structure can be faithfully preserved
- Copula-based multivariate algorithm is potential for general statistic-related engineering problems
- Caution when using copulas
  - Need sufficient historic records
  - Intrinsic assumption of stationary
  - Difficulties arise in higher dimensions
    - Mathematical complexity
    - Hard to preserve all lower level mutual dependencies
    - Compatibility problem

Acknowledgements

- Deepest gratitude to my advisor, Dr. Rao S. Govindaraju

- Thank all of my committee members
  - Dr. Dennis A. Lyn
  - Dr. Devdutta S. Niyogi
  - Dr. Venkatesh M. Merwade
  - Dr. Bo Tao

- Special thanks to
  - Dr. A. Ramachandra Rao
  - Dr. Jacques W. Delleur

- Thank my family and friends

Thank you
Questions?