Probabilistic Assessment of Multi-mechanism Floods in Inland Watersheds Due to Snowmelt-Influenced Extreme Streamflow Events

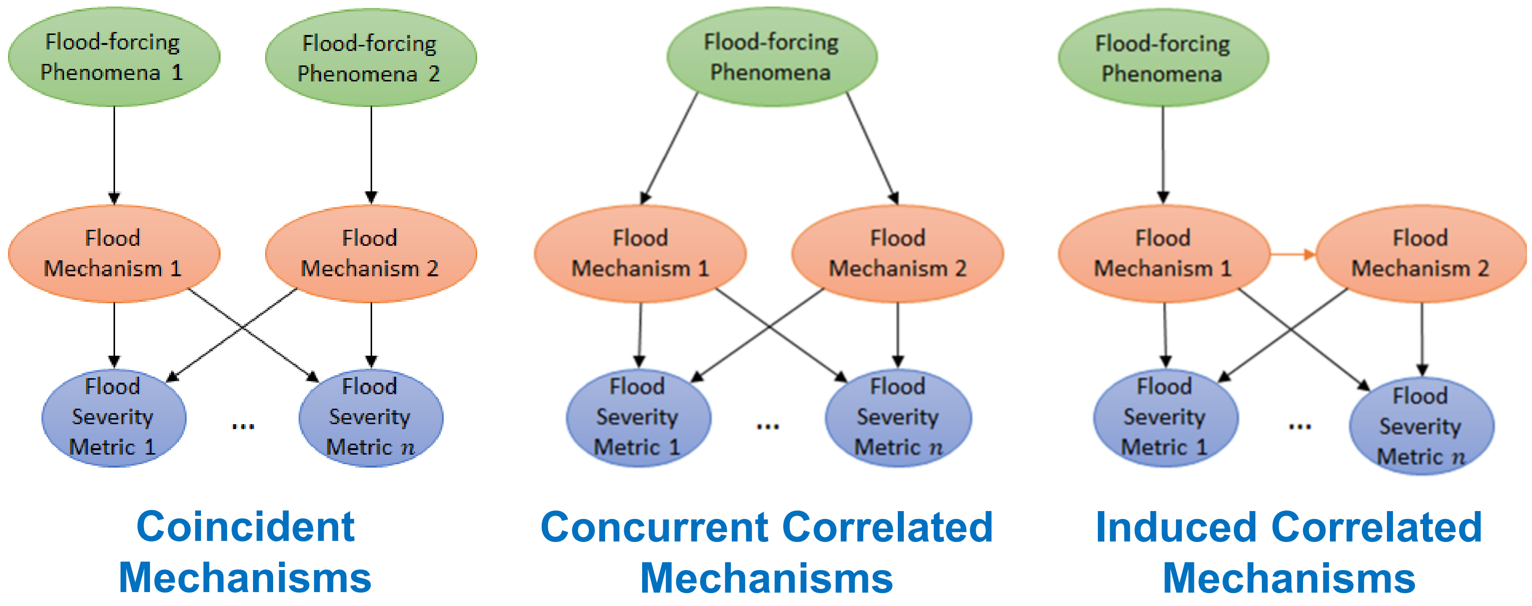
6th Annual NRC PFHA Workshop
February 22 – 25, 2021

Shih-Chieh Kao,[1] Scott T. DeNeale,[1] Michelle (Shelby) Bensi,[2] Somayeh Mohammadi,[2] Elena Yegorova,[3] Joseph Kanney,[3] and Meredith Carr[4]
[1] Oak Ridge National Lab; [2] University of Maryland; [3] US Nuclear Regulatory Commission; [4] US Army Engineer Research and Development Center

Presented by: **Shih-Chieh Kao** (kaos@ornl.gov)

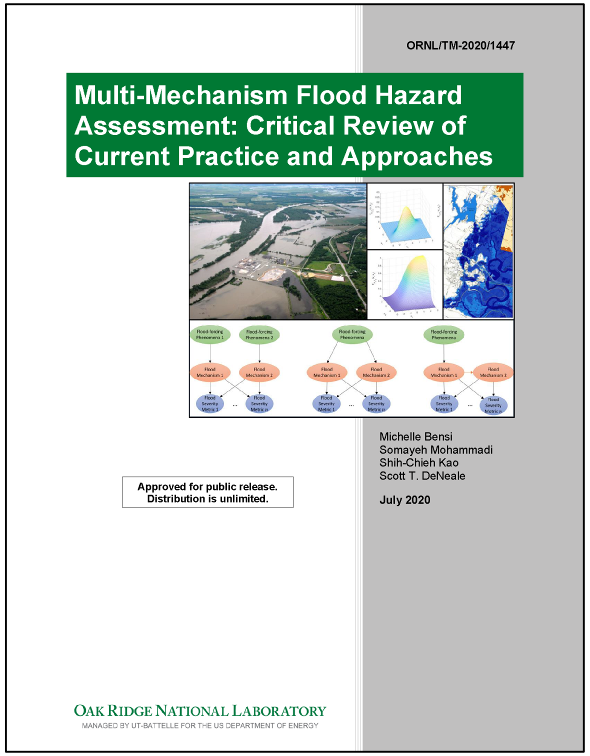
Multi-Mechanism Flood (MMF)

- **MMFs are flood events caused by more than one flooding mechanism.**
  - Also known as compound extreme events
  - Severe MMF may be formed by a combination of mechanisms that themselves are not extreme



MMF for Probabilistic Flood Hazard Assessment (PFHA)

- **A two-stage study to help better understand MMF for PFHA.**
  - Stage 1: Review of current concepts and methods
    - Bensi et al. (2020), Multi-Mechanism Flood Hazard Assessment: Critical Review of Current Practice and Approaches, doi:10.2172/1637939.
  - Stage 2: Development of example case studies
    - **Case 1: Inland study focusing on snowmelt-influenced events using copulas**
    - Case 2: Coastal study focusing on hurricane-induced storm surge and precipitation-induced river discharge using a Bayesian-motivated approach



https://www.osti.gov/biblio/1637939-multi-mechanism-flood-hazard-assessment-critical-review-current-practice-approaches

Objectives of This Case Study

- **We plan to demonstrate:**
  - General procedures to construct multivariate joint distributions using copulas
  - Identification of extreme samples for multivariate frequency analysis
  - Selection of suitable marginal distributions and copula functions
  - Applications of copula-derived joint distributions in PFHA
  - Strengths and limitations of the copula-based MMF assessment approach
- **Focus on inland snowmelt-influenced peak streamflow events**

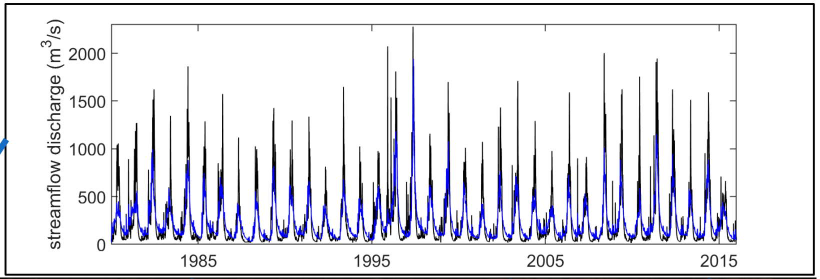
Study Areas

- **Selection considerations**
  - Long-term historic streamflow observations should be available at the watershed outlet to support model validation and frequency analysis.
  - Existing hydrologic model with acceptable performance should be available to simulate snow and other hydrologic processes.
  - The watershed should be large but not under strong flow regulation (e.g., presence of major dams). A headwater basin is preferred.
  - Significant snowpack should be presented to enable the assessment of snowmelt-influenced events.
- **Selected sites**
  - 3 sites meeting the above criteria from the NCAR CAMELS Dataset

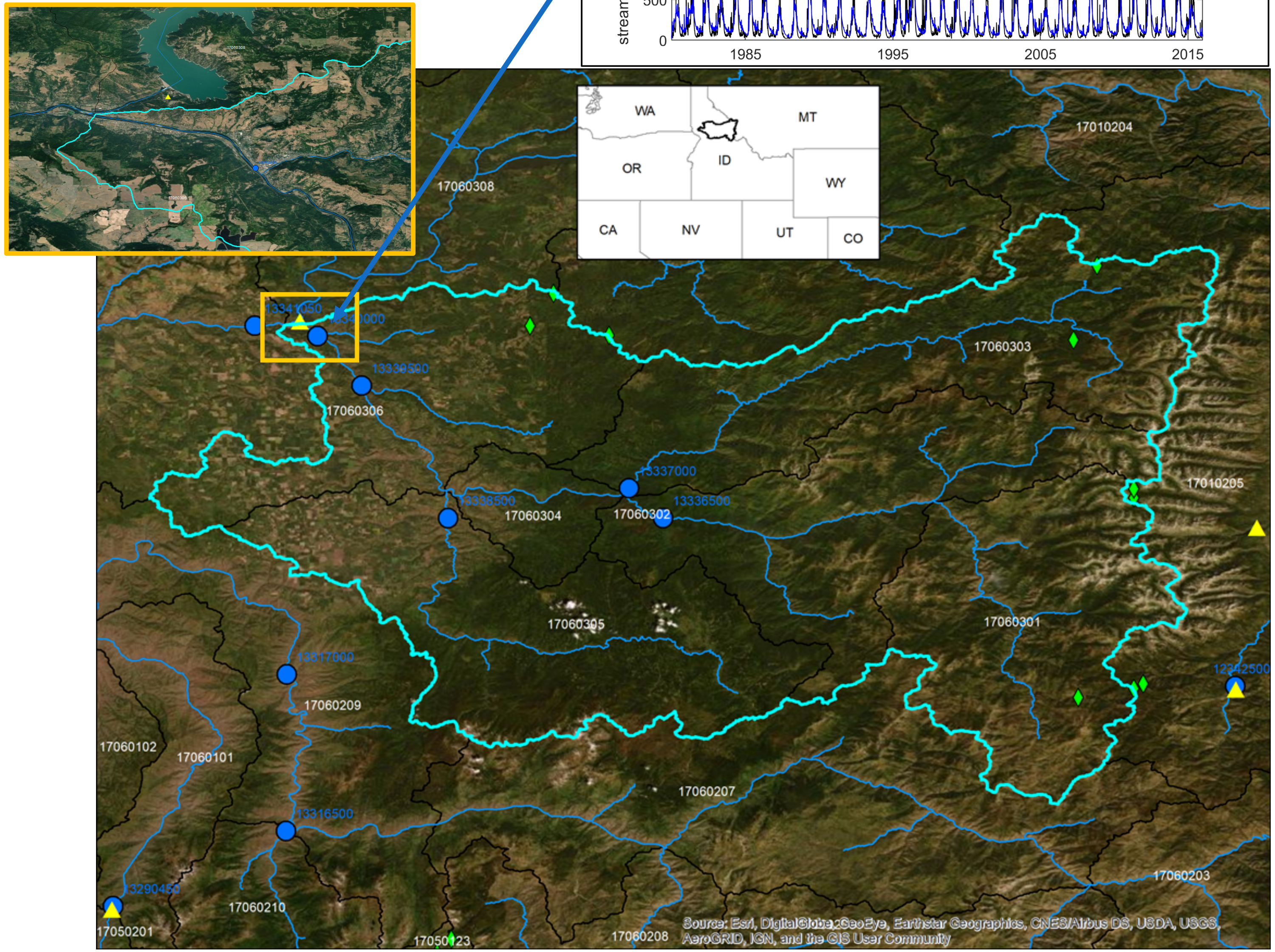
S1: Clearwater River at Orofino, ID

- **HUC08**
  - 17060301–17060306
- **Outlet USGS Station**
  - 13340000
  - 1965–present
  - 247.2 m^3/s
  - 14263 km^2
- **Elevation (m)**
  - min: 338
  - mean: 1451
  - max: 2602

ORNL VIC R^2 (daily) = 0.74



Next to Dworkshak Reservoir

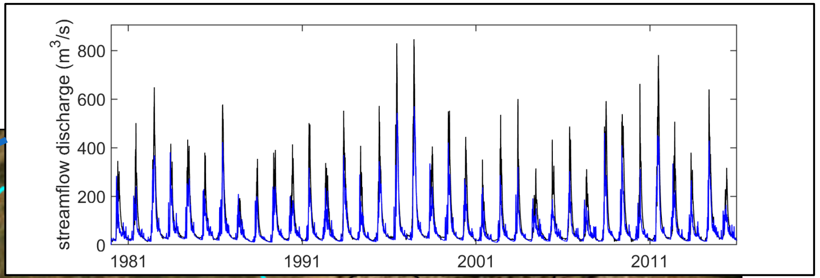


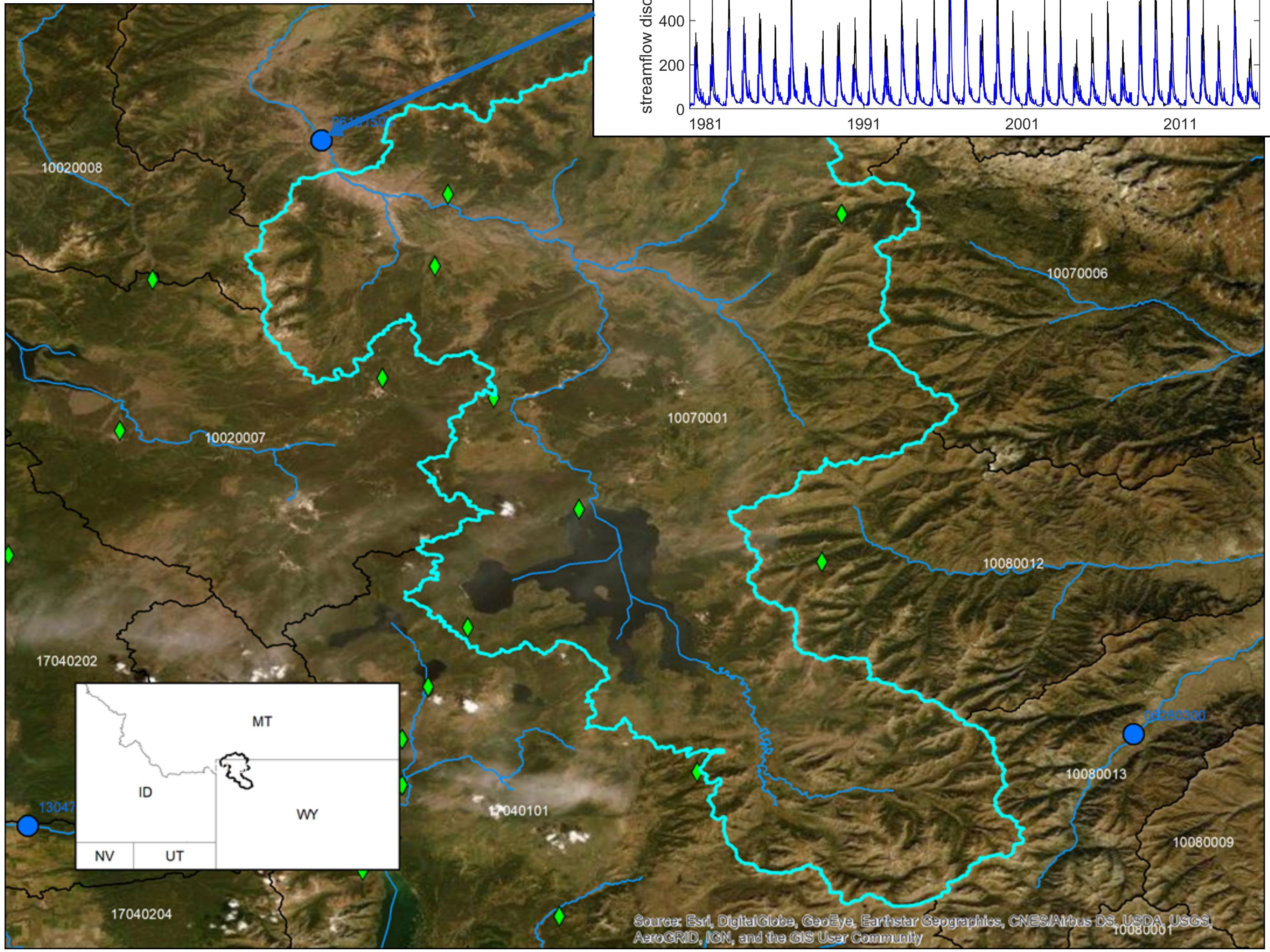
OAK RIDGE National Laboratory

S2: Yellowstone River at Corwin Springs, MT

ORNL VIC R^2 (daily) = 0.80

- **HUC08**
  - 10070001
- **Outlet USGS Station**
  - 06191500
  - 1911–present
  - 88.6 m^3/s
  - 3706.27 km^2
- **Elevation (m)**
  - min: 1560
  - mean: 2542
  - max: 3473

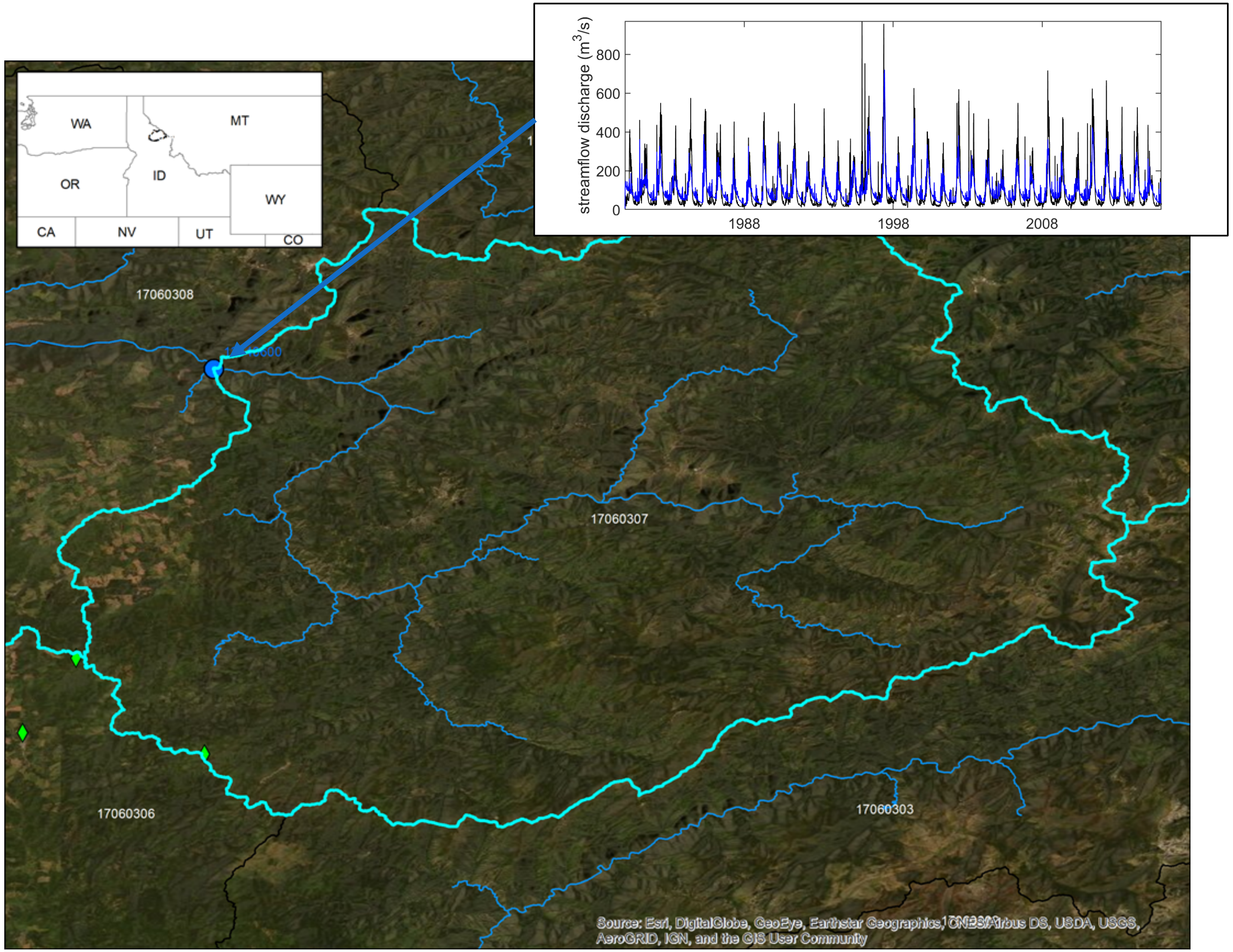




S3: NF Clearwater River NR Canyon Ranger Station, ID

- **HUC08**
  - 17060307
- **Outlet USGS Station**
  - 13340600
  - 1967–present
  - 98.2 m^3/s
  - 3356.6 km^2
- **Elevation (m)**
  - min: 1448
  - mean: 569
  - max: 2241

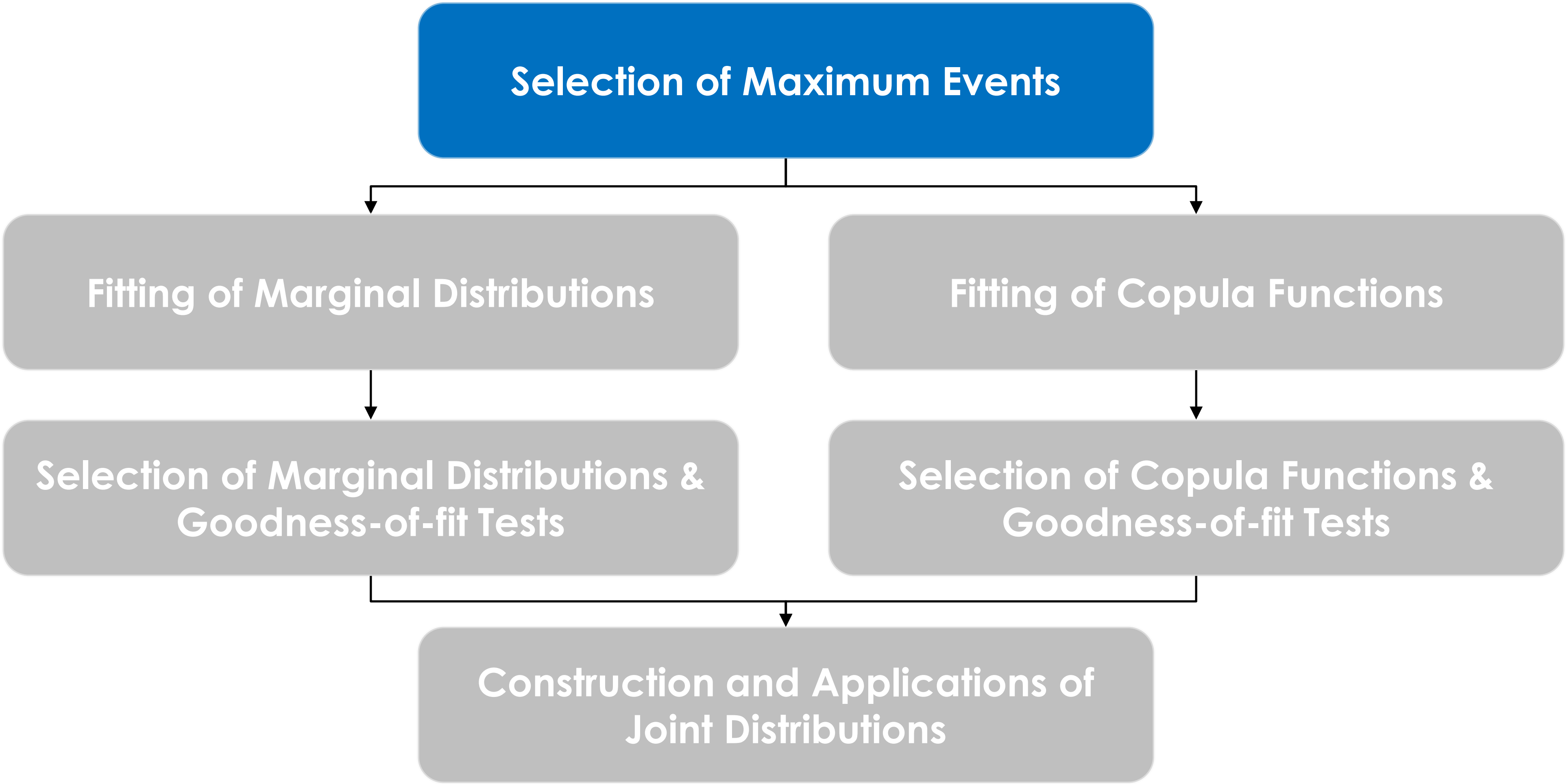
ORNL VIC R^2 (daily) = 0.71



Summary of Available Data

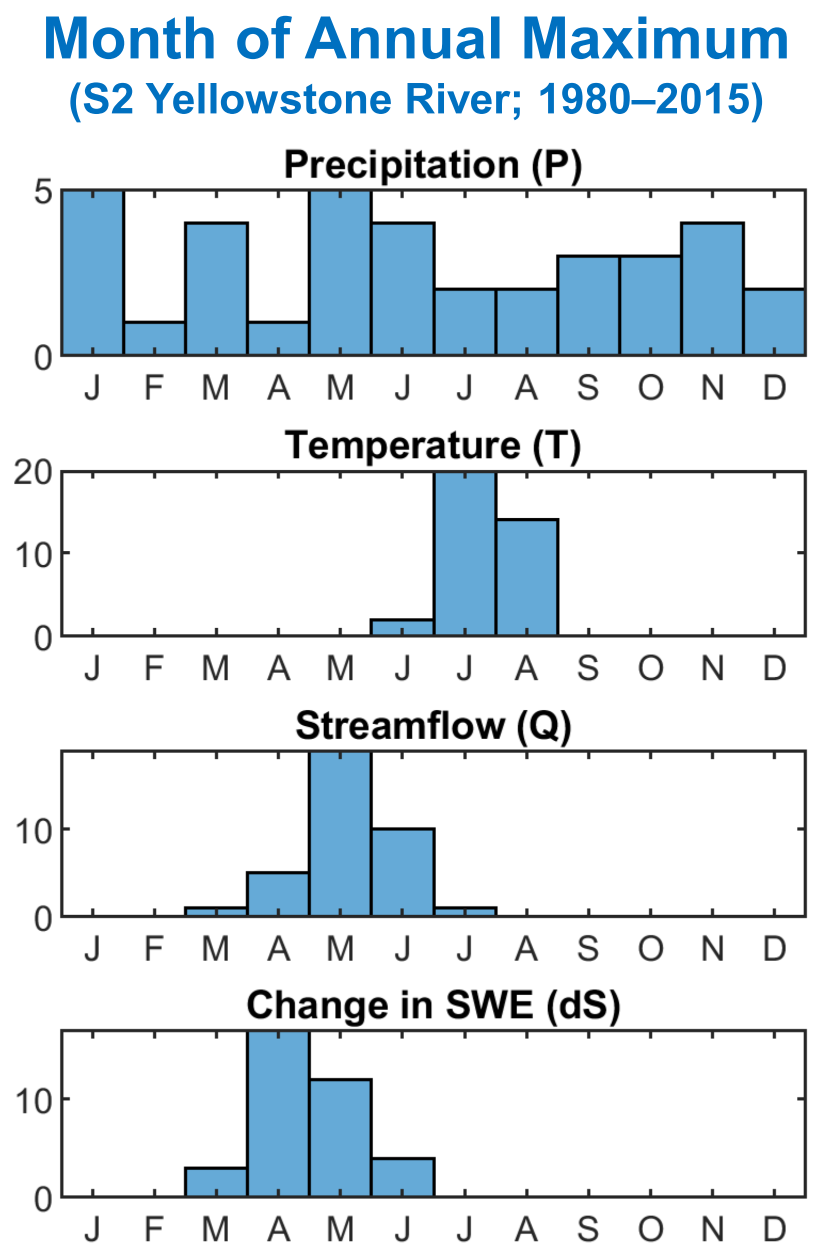
- **Observation**
  - USGS daily streamflow (Q_{obs})
  - 1980–2015 Daymet precipitation (P_{obs}) and temperature (T_{obs})
- **Observation-driven hydrologic model outputs**
  - 1980–2015 streamflow (Q_{VIC}) and change in snow water equivalent ($dSWE_{VIC}$)
- **We focus on the following variables in this case study**
  - Q_{dy}: Daily streamflow (m^3/s)
  - P_{3d}: 3-day precipitation (mm)
  - T_{3d}: 3-day temperature (°C)
  - dS_{3d}: 3-day change in SWE (mm)

Assessment Procedures

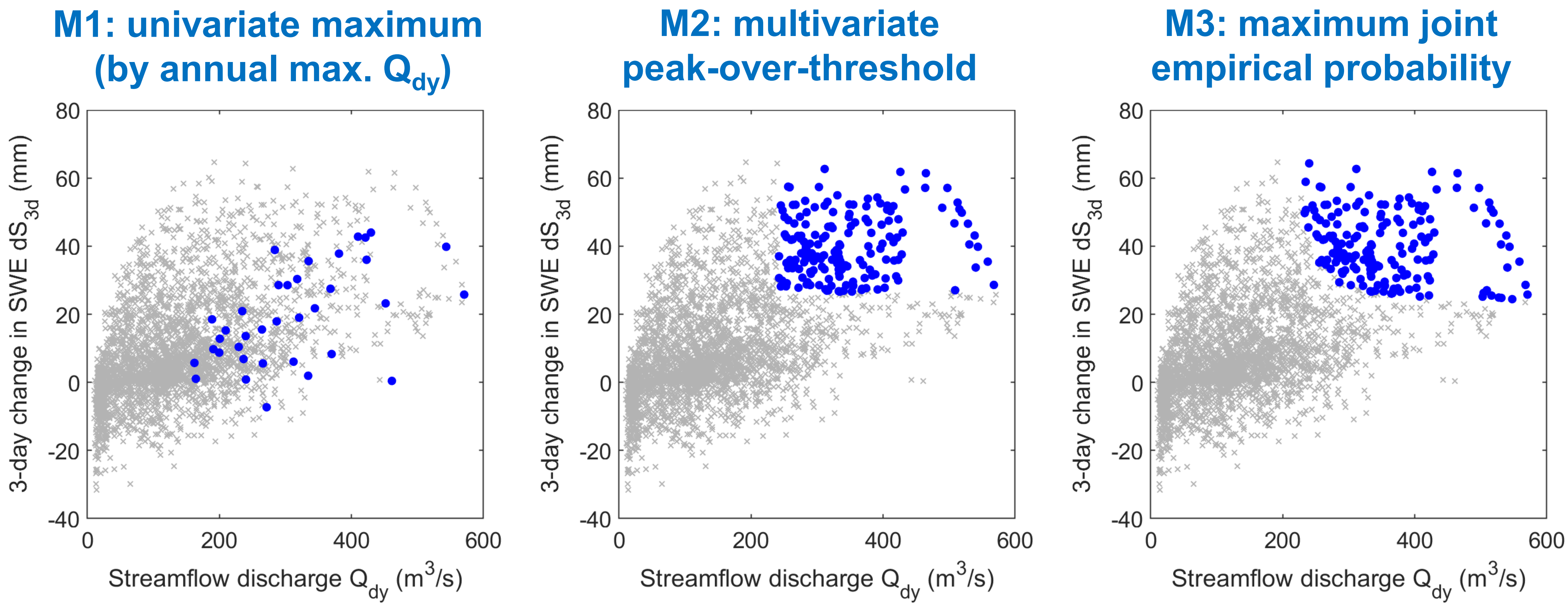


Selection of Maximum Events

- **Samples for multivariate frequency analysis should have consistent timing**
  - (P_{t1}, T_{t1}, Q_{t1}, dS_{t1})
- **The conventional annual maximum approach cannot work directly**
  - (P_{t1}, T_{t2}, Q_{t3}, dS_{t4})
- **In this study, we**
  - select maximum events during April–June
  - compare three different ways in selecting maximum events

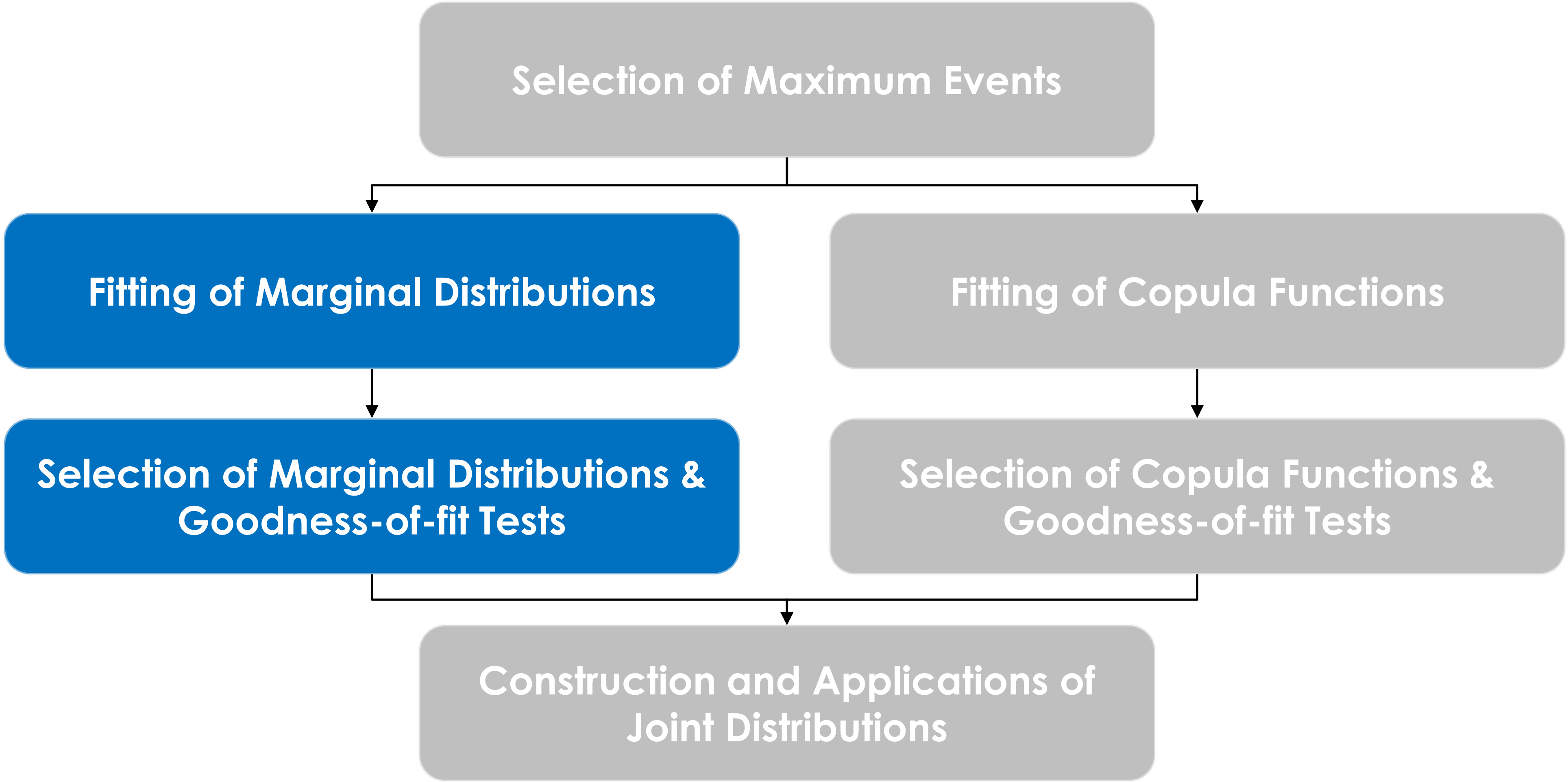


Compare Three Ways to Select Maximum Events



Bivariate example – S2 (Yellowstone River)

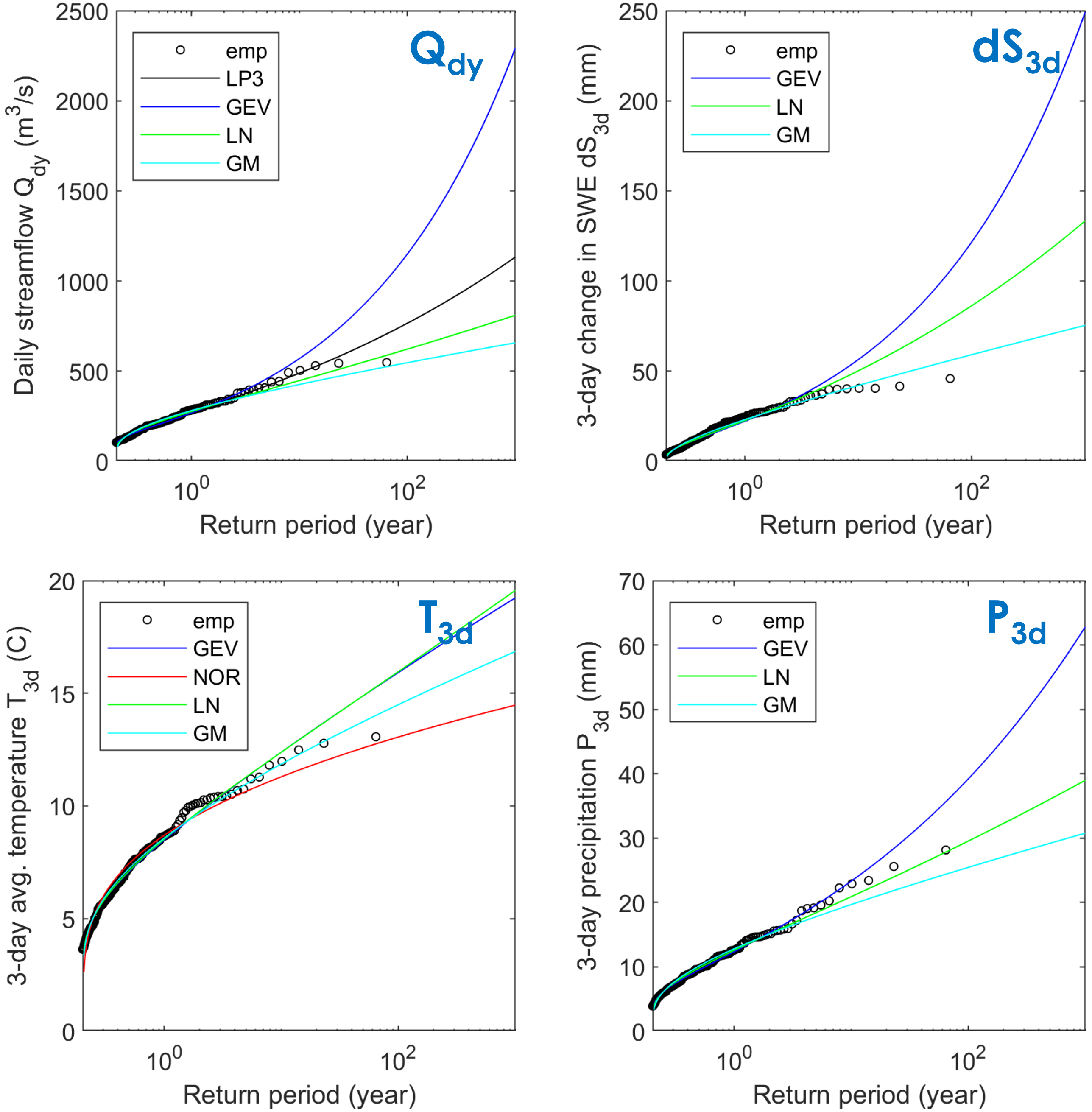
Assessment Procedures



Marginal Distribution

- **Tested distributions**
  - Log Pearson Type III (LP3)
  - General. Extreme Value (GEV)
  - Log-normal (LN)
  - Gamma (GM)
  - Normal (NOR)
- **Parameter estimation**
  - Maximum likelihood
- **Goodness-of-fit**
  - Kolmogorov-Smirnov (KS)
  - Cramer-Von Mises (CM)
  - Akaike Information Criterion (AIC)
  - Bayesian Information Criterion (BIC)

S2M2: multivariate peak-over-threshold



Distribution Selection Using AIC & BIC

S2M2 – Yellowstone river with multivariate peak-over-threshold

- **Select distribution with smaller AIC & BIC values**
  - AIC = −2*(*log-likelihood*)+2*(*numParam*).
  - BIC = −2*(*log-likelihood*)+*numParam*∗log(*numObs*)

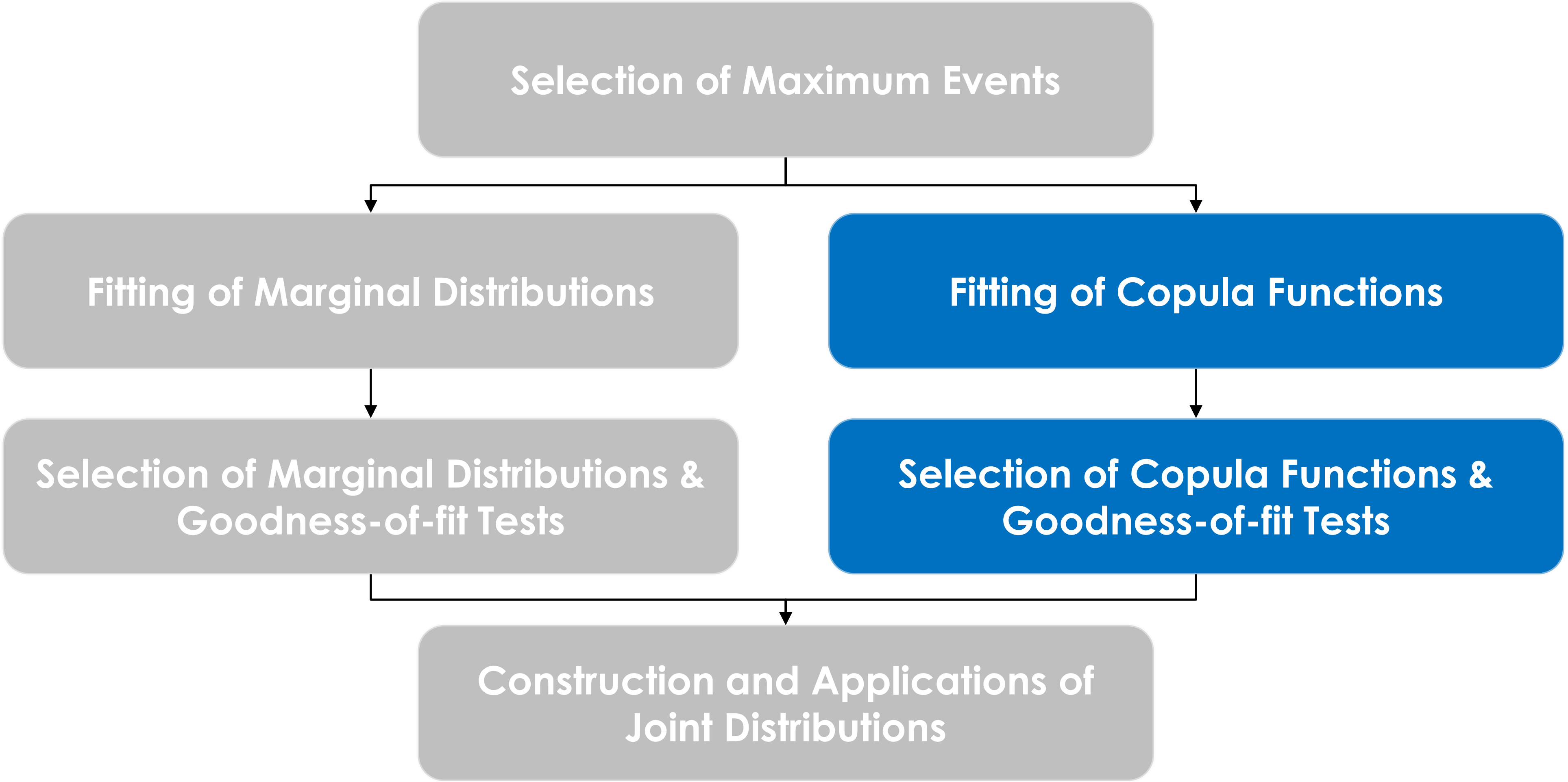
	Streamflow (Q_{dy})		Change in SWE (dS_{3d})		Temperature (T_{3d})		Precipitation (P_{3d})	
	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
LP3	3362.2	3371.8	--	--	--	--	--	--
GEV	3362.0	3371.6	1298.4	1308.0	770.9	780.5	985.7	995.3
NOR	--	--	--	--	787.6	794.0	--	--
LN	3368.9	3375.2	1286.9	1293.3	767.9	774.3	985.4	991.8
GM	3382.2	3388.6	1289.2	1295.6	770.3	776.7	996.6	1003.0

Selected Marginal Distributions

Variables	S1M2 Clearwater, Multivariate Peak-over-threshold	S2M2 Yellowstone, Multivariate Peak-over-threshold	S2M3 Yellowstone, Max. Joint Empirical Probability	S3M2 NF Clearwater, Multivariate Peak-over-threshold
Streamflow (Q_{dy})	LP3	LP3	LP3	LP3
Precipitation (P_{3d})	LN	LN	LN	LN
Temperature (T_{3d})	LN	LN	GEV	GEV
SWE Change (dS_{3d})	LN	GM	GM	GM

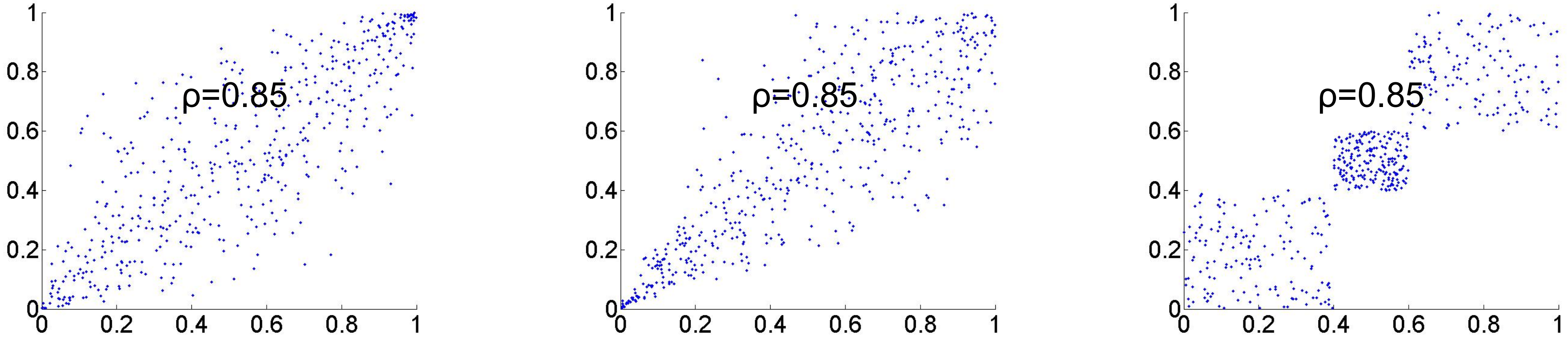
- **The best distributions varied by variables, maximum events, and sites.**
- **Maximum event searching approaches seem to have more profound impacts than different sites.**
  - M1 (univariate maximum) is more different than M2 and M3.
  - **M1 (univariate maximum) is NOT considered further due to the limited sample size.**

Assessment Procedures



Correlation and Dependence Structure

- **Can Pearson's linear correlation coefficient ρ fully characterize the relationship between variables?**
  - $\rho_{XY} = E[(X - \bar{x})(Y - \bar{y})]/Std[X]Std[Y]$

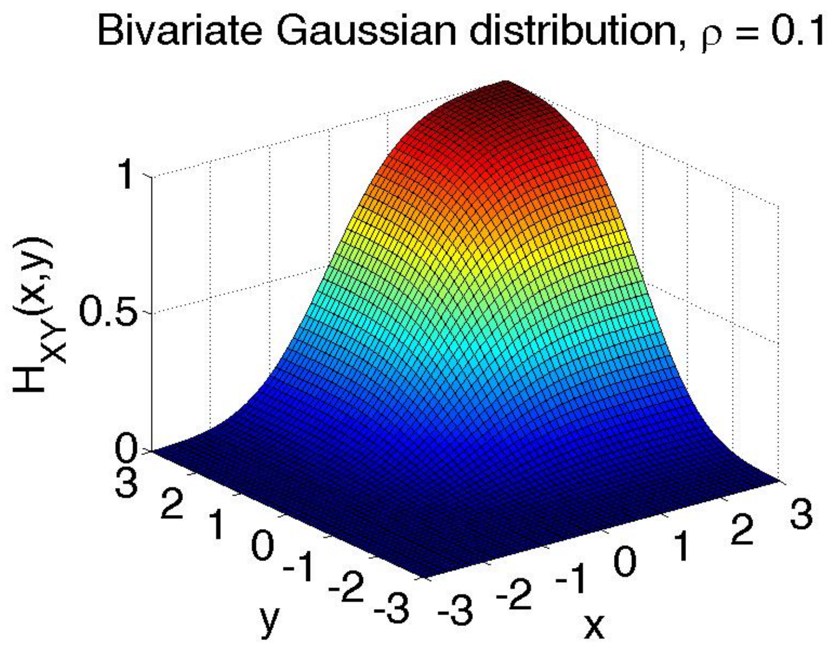


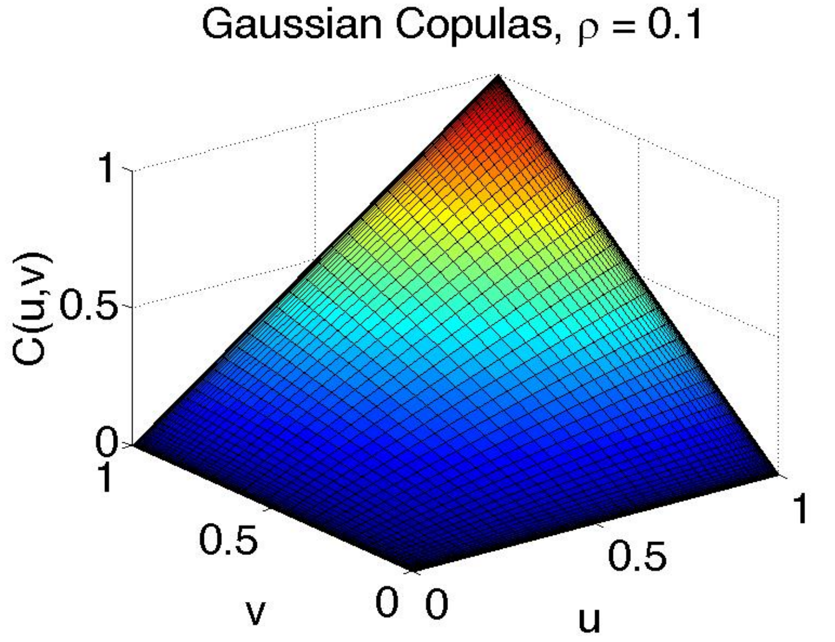
  - Only valid for Gaussian (or elliptic) distributions

- **Dependence structure provides a more comprehensive characterization than the correlation coefficient**

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**

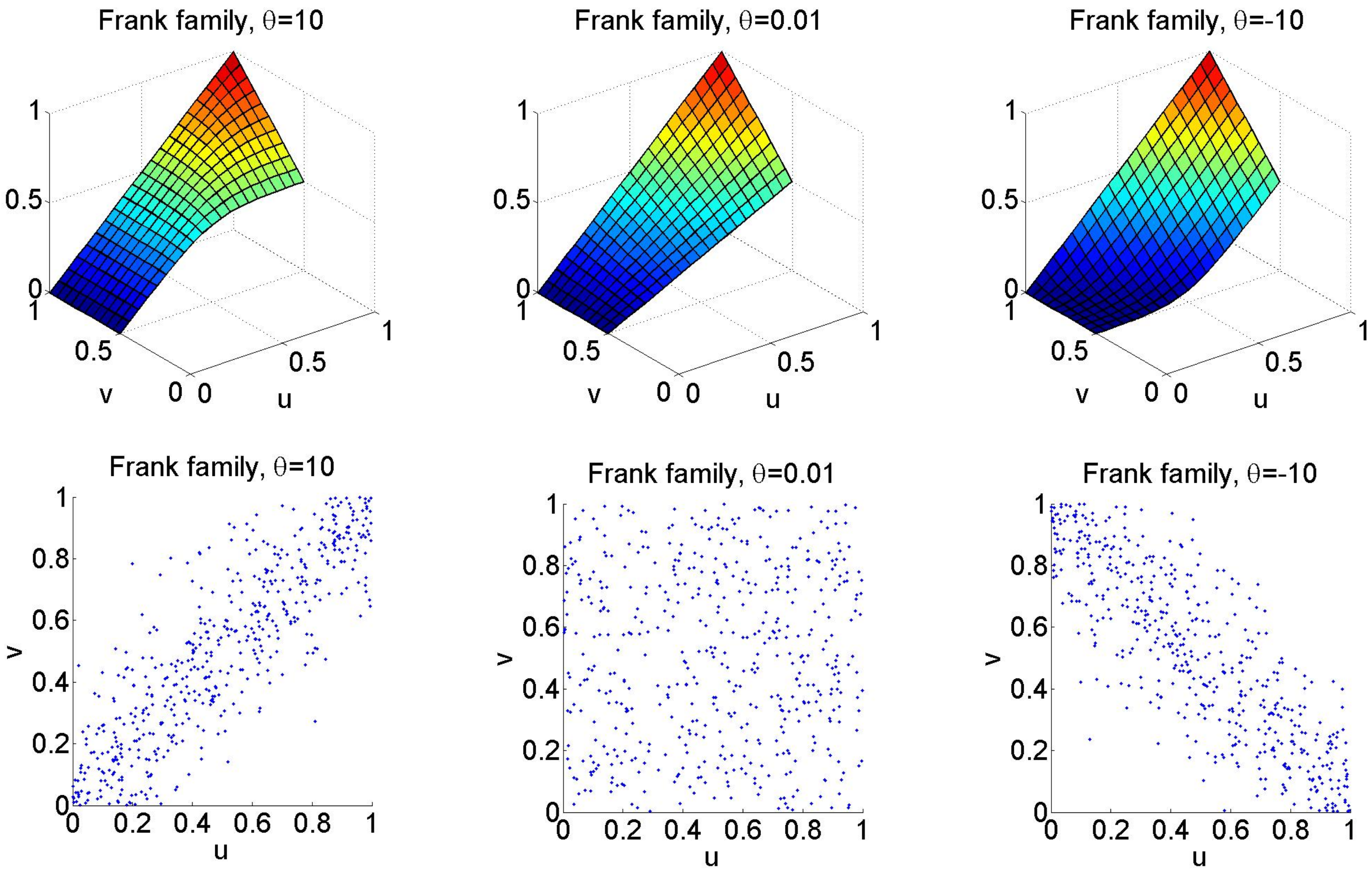




Example of Copulas – Frank Family

- **Frank family of Archimedean copulas**

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



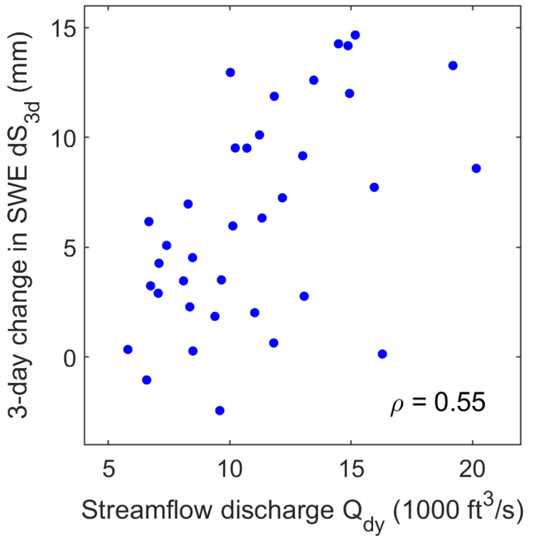
OAK RIDGE National Laboratory

Copula Functions

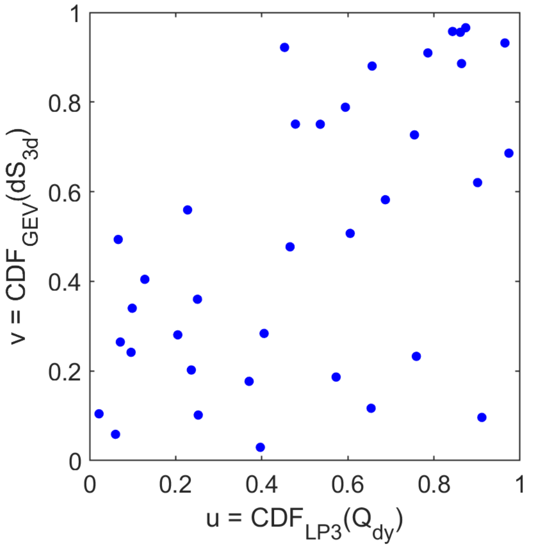
- **Tested copulas**
  - Gaussian (GAU)
  - T, degree of freedom = 2 (TD2)
  - Frank (FRK)
  - Clayton (CLT)
  - Gumbel (GUM)
- **Parameter estimation**
  - Inference Functions for Margins (IFM), using fitted marginals
  - Canonical Maximum Likelihood (CML), using empirical marginals
- **Goodness-of-fit**
  - Multivariate Kolmogorov-Smirnov (KS)
  - Akaike Information Criterion (AIC)
  - Bayesian Information Criterion (BIC)

Example of Copula Function Fitting

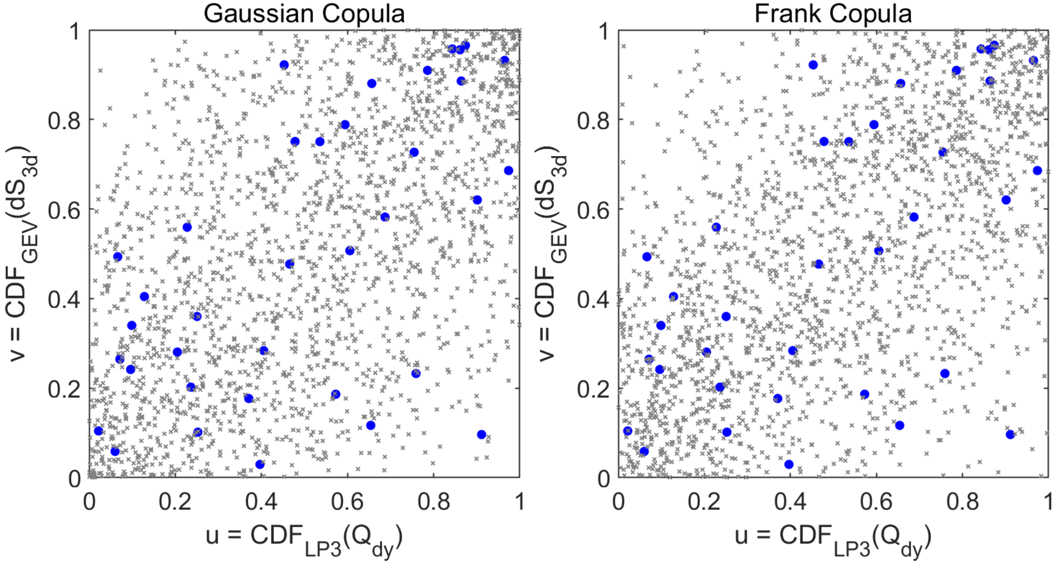
Original Domain

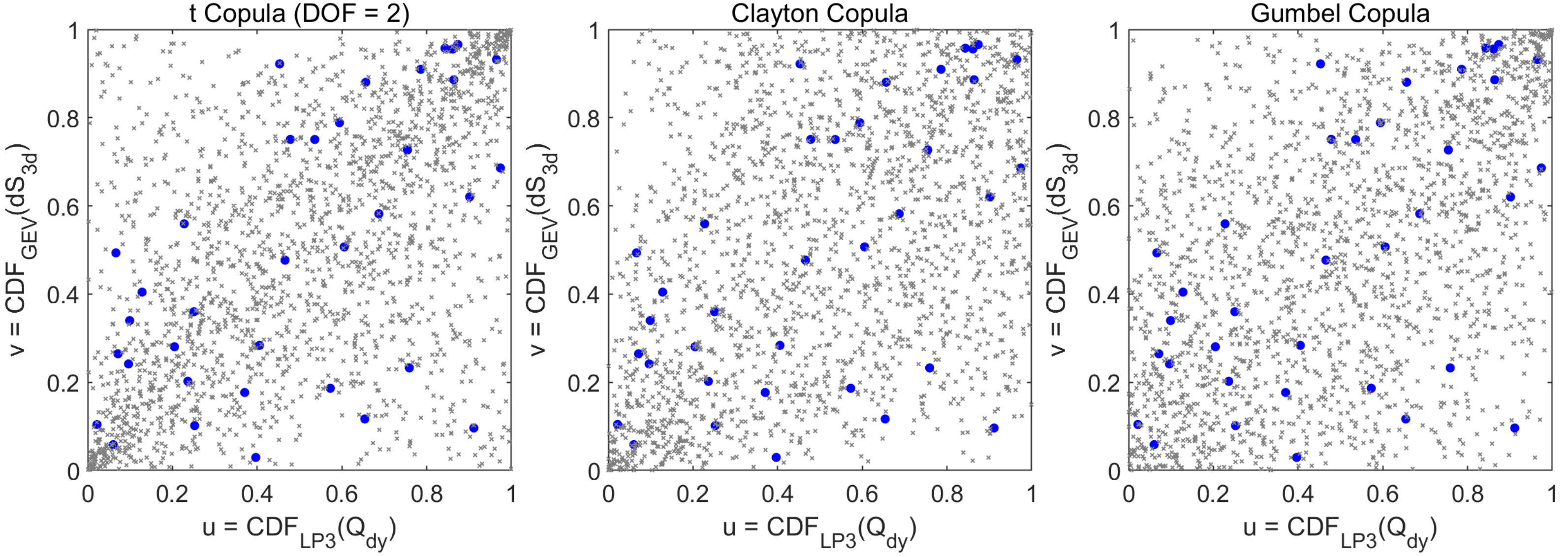


After Transformation

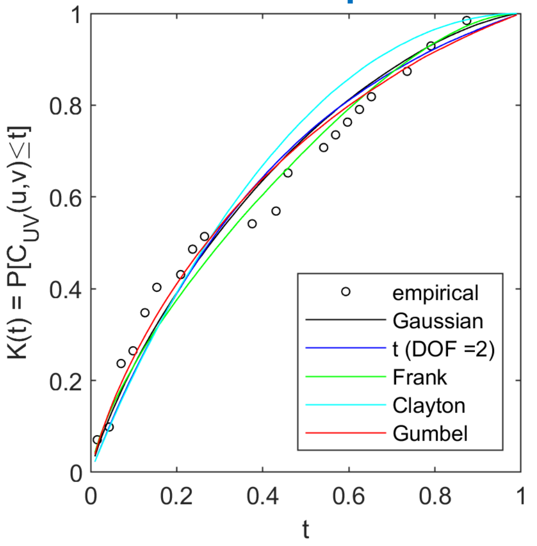


Copula Function Fitting



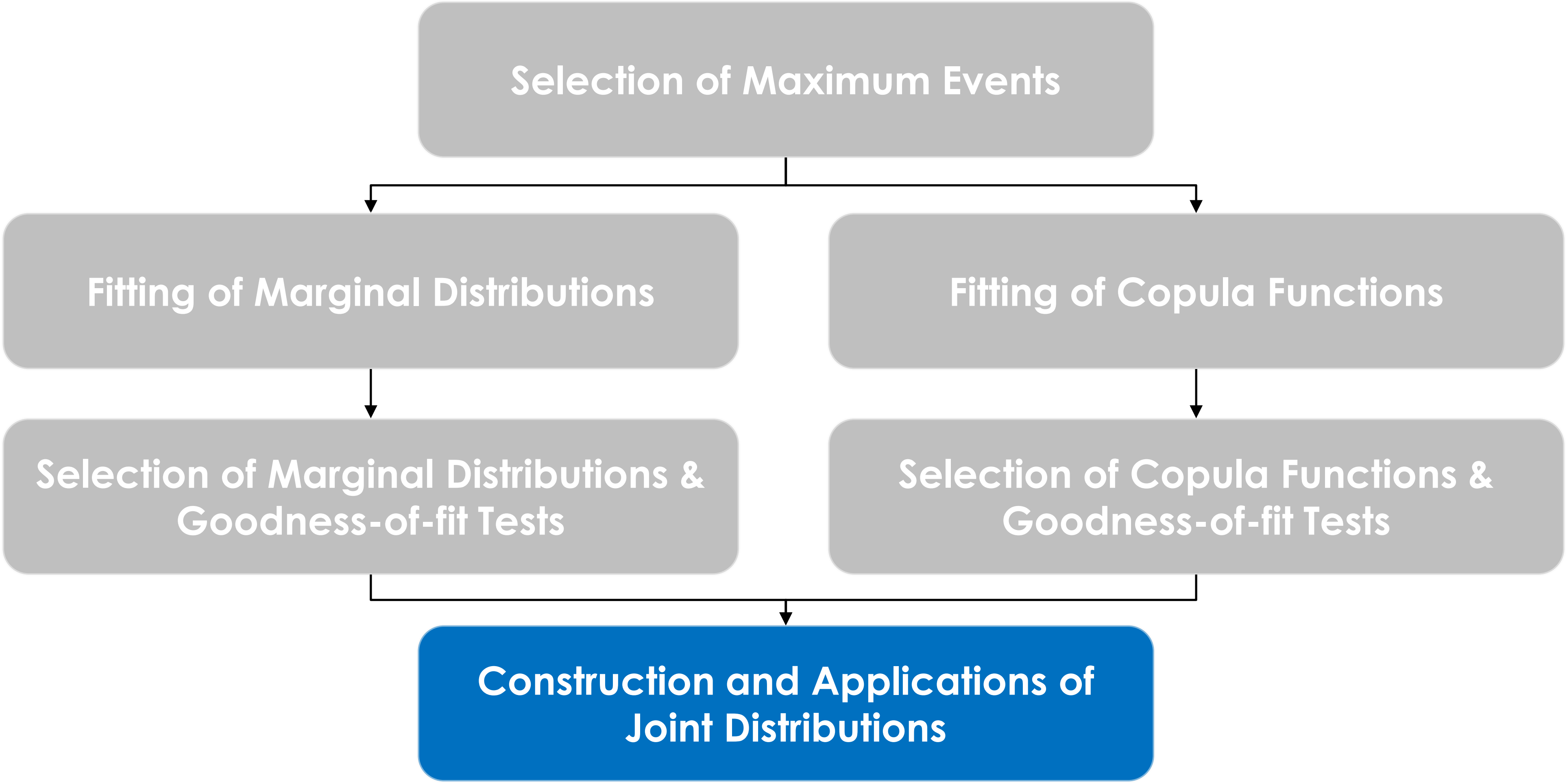


Overall Comparison



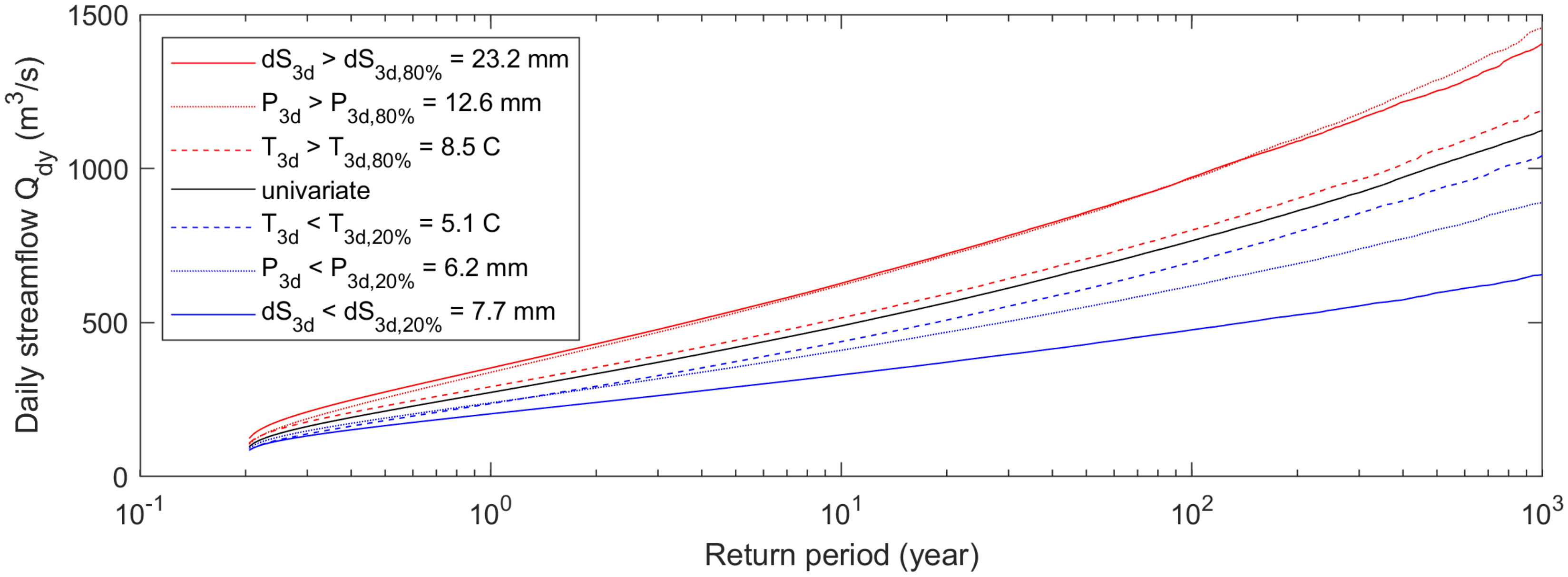
Variables	Correlation coefficient (ρ)	Kendall's τ	Selected copula function	Copula parameter
S1M2 (Clearwater River at Orofino, ID; Multivariate Peak-over-threshold)				
Q_{dy} & P_{3d}	0.1537	0.0606	GUM	1.0697
Q_{dy} & T_{3d}	-0.0455	-0.0077	FRK	-0.0933
Q_{dy} & dS_{3d}	0.3828	0.4045	FRK	4.0447
S2M2 (Yellowstone River at Corwin Springs, MT; Multivariate Peak-over-threshold)				
Q_{dy} & P_{3d}	0.2381	0.1387	GUM	1.1793
Q_{dy} & T_{3d}	0.2311	0.1762	CLT	0.3145
Q_{dy} & dS_{3d}	0.4555	0.3112	GAU	0.4461
S2M3 (Yellowstone River at Corwin Springs, MT; Maximum Joint Empirical Probability)				
Q_{dy} & P_{3d}	-0.1256	-0.0906	GAU	-0.1637
Q_{dy} & T_{3d}	0.2428	0.1351	GUM	1.1398
Q_{dy} & dS_{3d}	0.4287	0.2852	GAU	0.4247
S3M2 (NF Clearwater River NR Canyon Ranger Station, ID; Multivariate Peak-over-threshold)				
Q_{dy} & P_{3d}	0.1071	0.0673	FRK	0.6349
Q_{dy} & T_{3d}	0.0049	0.0026	CLT	0.0368
Q_{dy} & dS_{3d}	0.5273	0.4438	GAU	0.6124

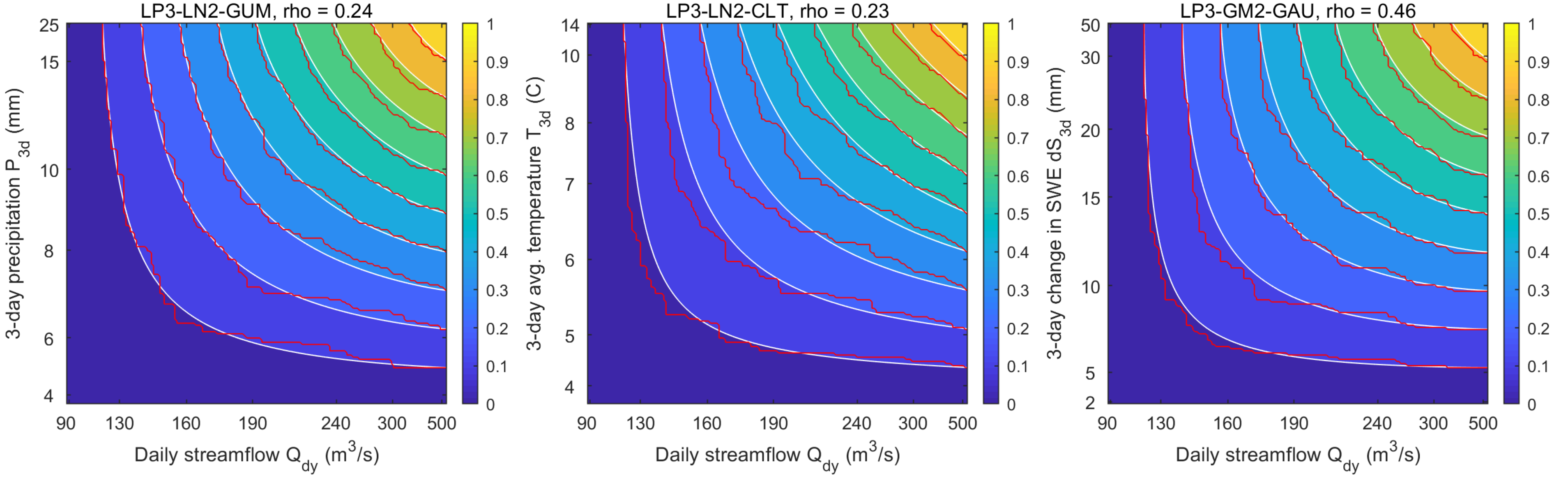
Assessment Procedures



Bivariate Joint Distributions

S2M2 – Yellowstone River with multivariate peak-over-threshold

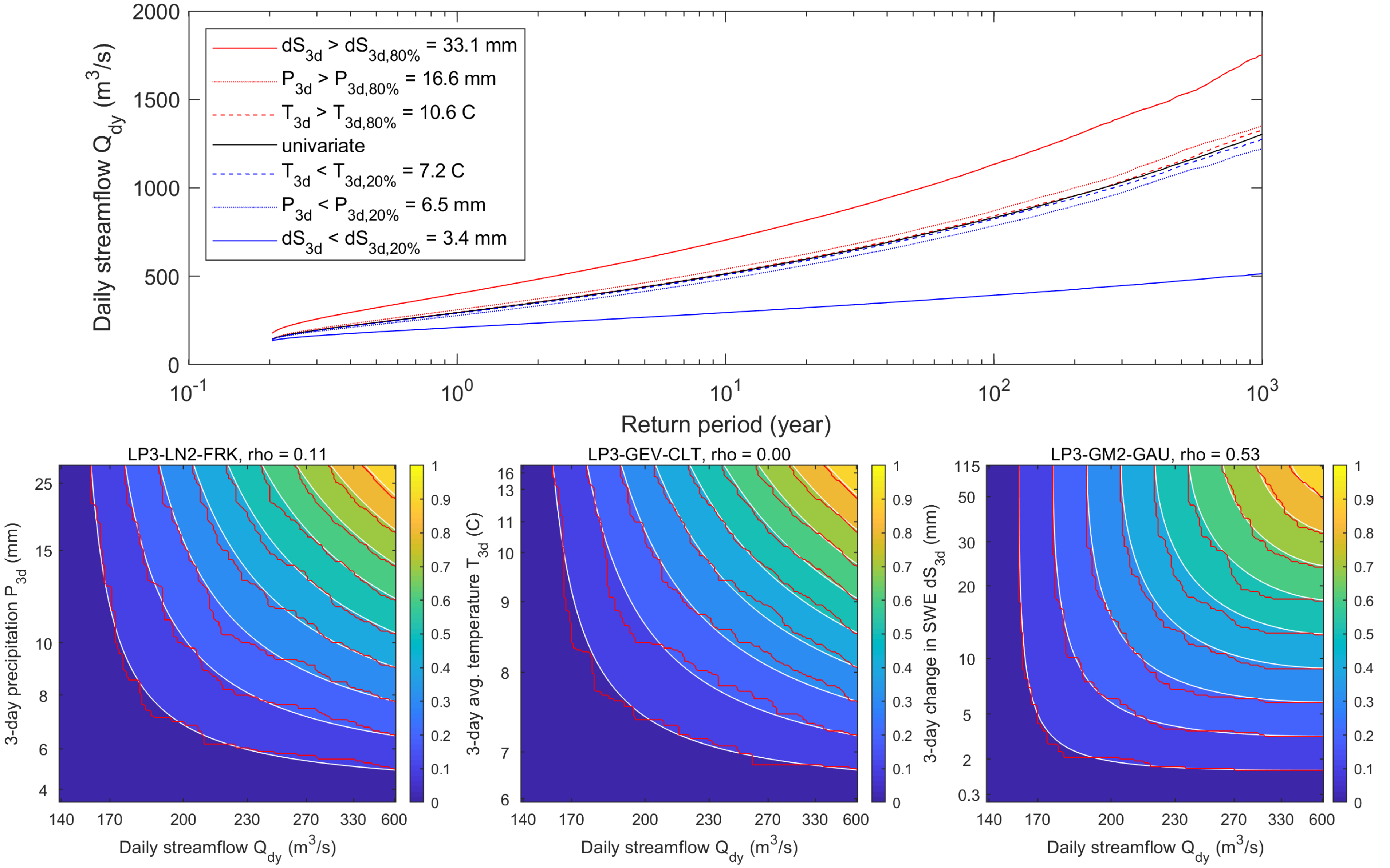




Oak Ridge National Laboratory

Bivariate Joint Distributions

S3M2 – NF Clearwater River with multivariate peak-over-threshold



Summary and Next Steps

- **Copulas offer a natural way to extend our conventional univariate frequency analysis to multiple dimensions**
  - Can be applied to a variety of different MMF applications for PFHA
- **However, there are new issues to be considered**
  - Definition of maximum events
  - Data availability
  - Challenges in higher dimensions
- **Further exploration is needed to identify the best practice of applications**

Thank you!

- Shih-Chieh Kao (kaos@ornl.gov)