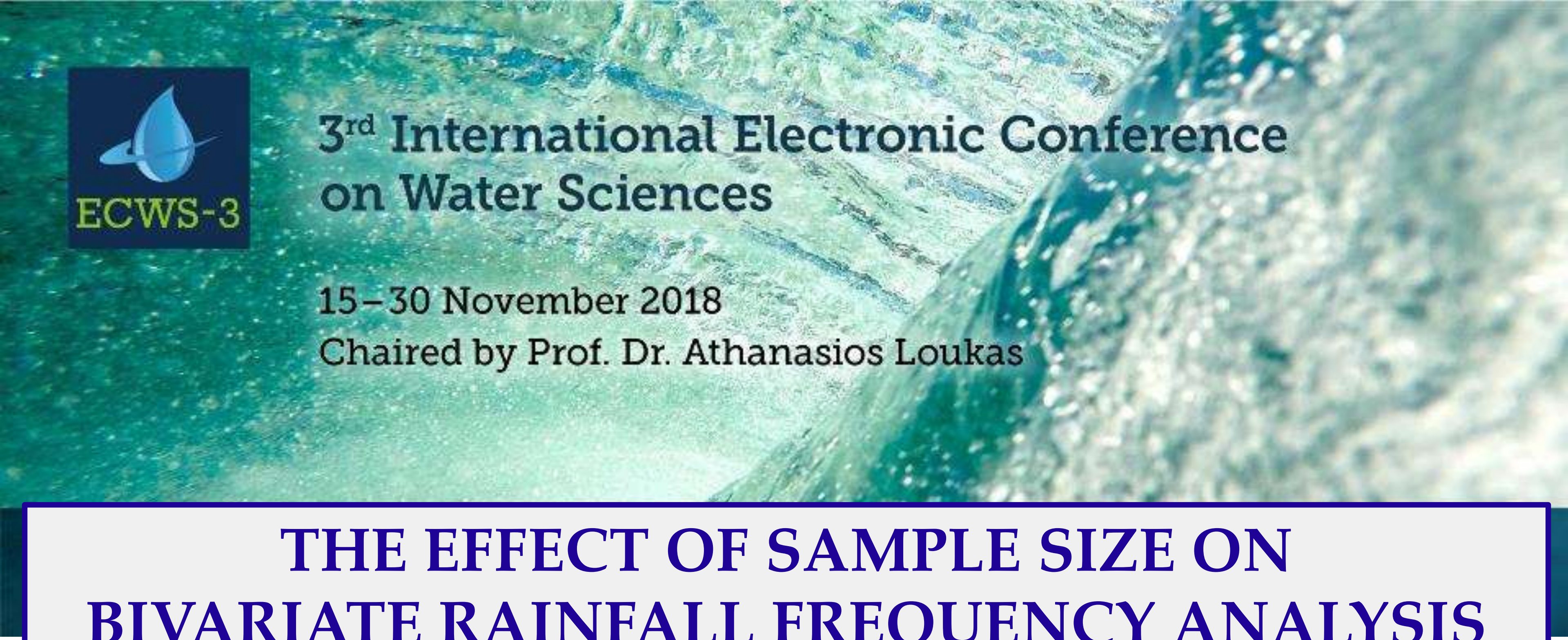


ECWS-3

3rd International Electronic Conference on Water Sciences

15–30 November 2018

Chaired by Prof. Dr. Athanasios Loukas

THE EFFECT OF SAMPLE SIZE ON BIVARIATE RAINFALL FREQUENCY ANALYSIS OF EXTREME PRECIPITATION

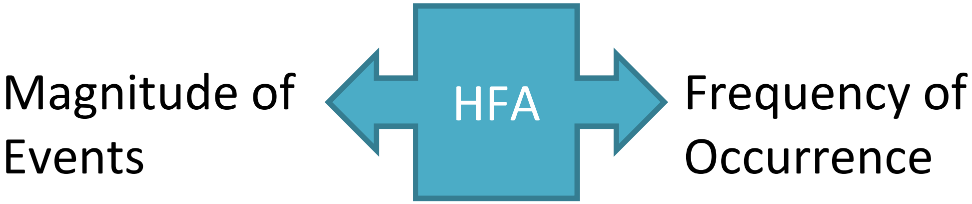
Nikoletta Stamatatou, Lampros Vasiliades and Athanasios Loukas
Laboratory of Hydrology and Aquatic Systems Analysis,
Department of Civil Engineering, University of Thessaly, Volos, Greece
E-mail: lvassil@civ.uth.gr

OUTLINE

- Introduction
- Study Area and Database
- Univariate Rainfall Frequency Analysis
- Bivariate Rainfall Frequency Analysis
- Modelling Dependence
- Copulas
- Results – Concluding Remarks

HYDROLOGICAL FREQUENCY ANALYSIS

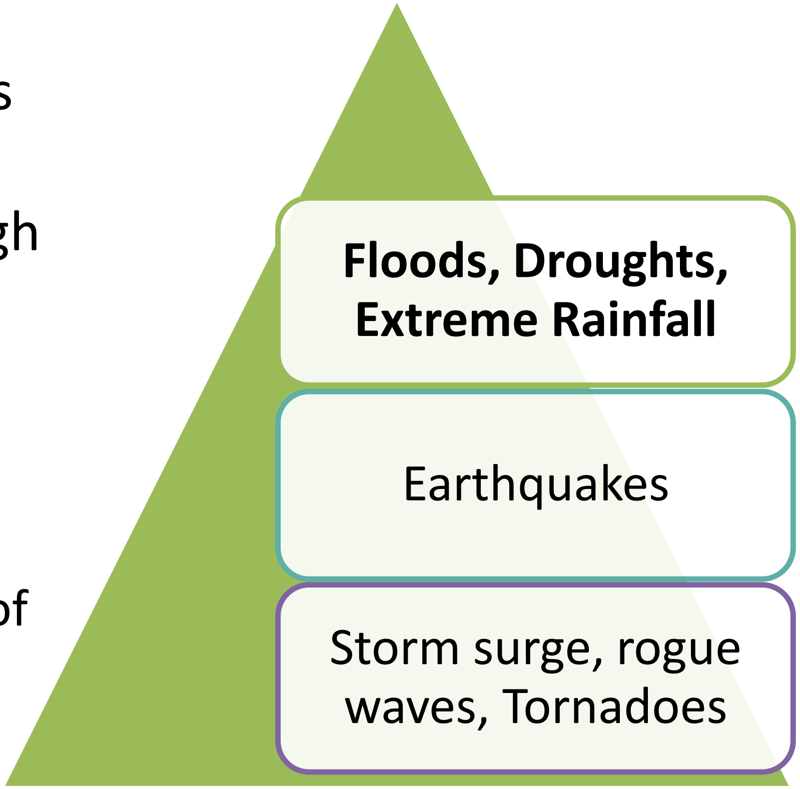
The objective of frequency analysis is to relate the magnitude of events to their frequency of occurrence through probability distributions.



Magnitude of Events

HFA

Frequency of Occurrence



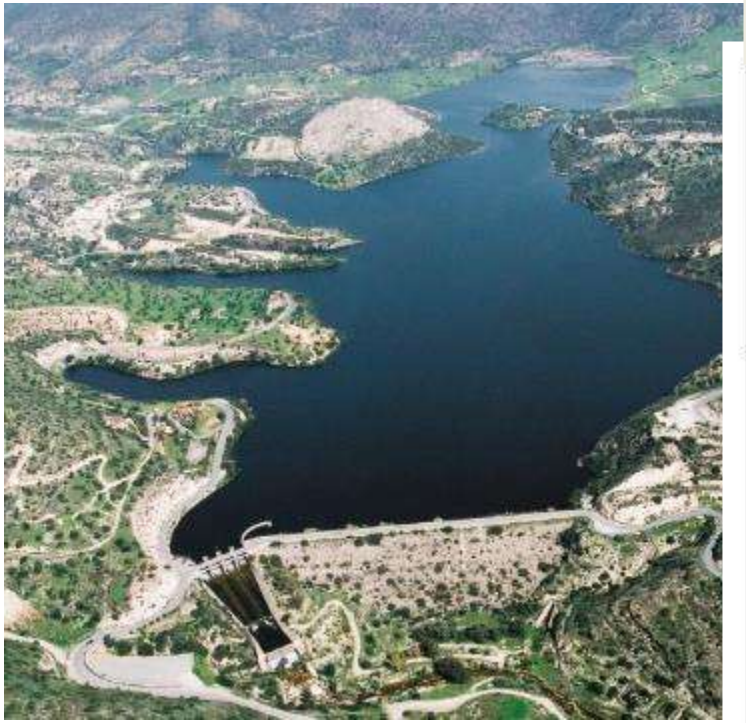
Floods, Droughts, Extreme Rainfall

Earthquakes

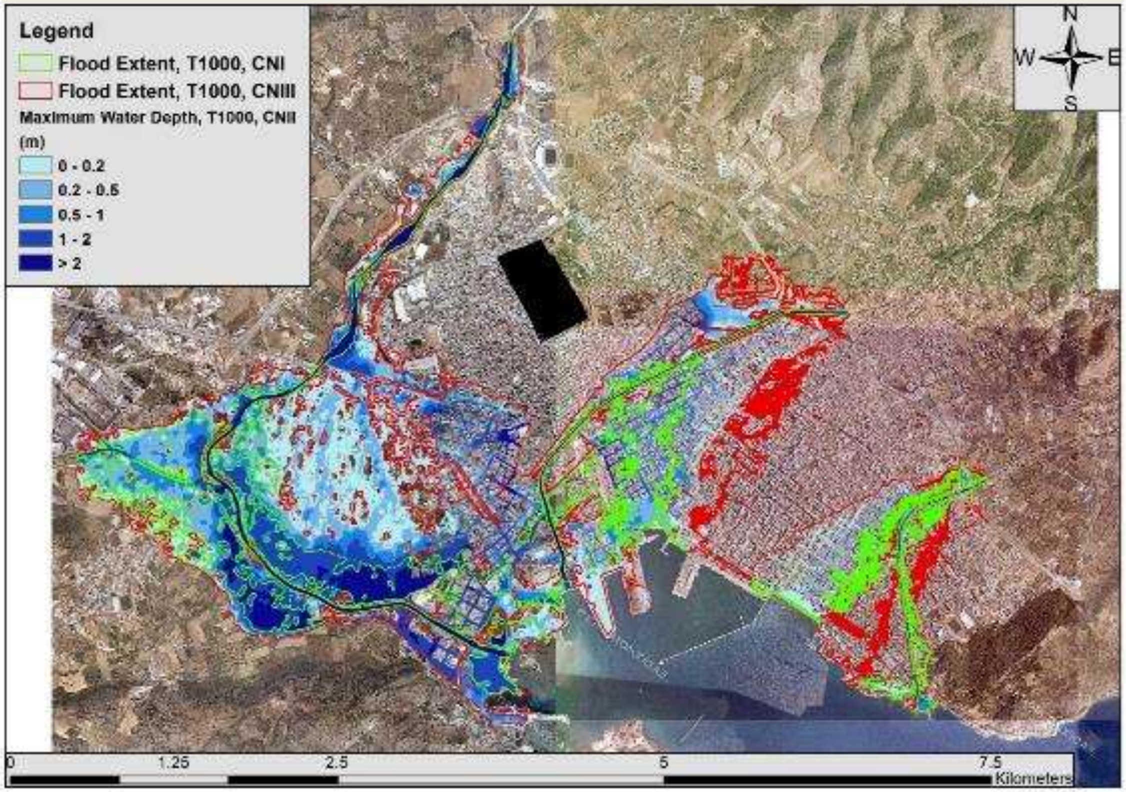
Storm surge, rogue waves, Tornadoes

DESIGN FLOOD
HYDROLOGICAL FREQUENCY ANALYSIS

Dam Design values:
Yermasoyia Dam



Flood risk map:
Flood extent and water depths of return period T = 1000 years for Volos city, Greece



RAINFALL FREQUENCY ANALYSIS

- **Objective: Multivariate approach** on **RFA** using **copulas**

- **Design variables of the hydraulic structures**
  - Rainfall frequency estimation in a multivariate framework
    - **Dependence** of rainfall characteristics
      - peak rainfall
      - Volume of extreme rainfall
      - Storm duration
  - Design of copulas for various hydrologic (-meteorological) applications (variables)
  - Calculation and comparison of univariate and joint bivariate return periods
    - **Conditional** return period
    - **Joint *OR*** and ***AND*** return periods

$$T(u|v) = \mu \frac{1-v}{1-u-v+C(u,v)} \text{ and}$$

$$T(v|u) = \mu \frac{1-u}{1-u-v+C(u,v)},$$

Rainfall peak and **storm duration**:

$$T_{u,v}^{OR} = \frac{\mu}{1-C_{u,v}(u,v)},$$

either peak and duration exceed their threshold (**cooperative risk**)

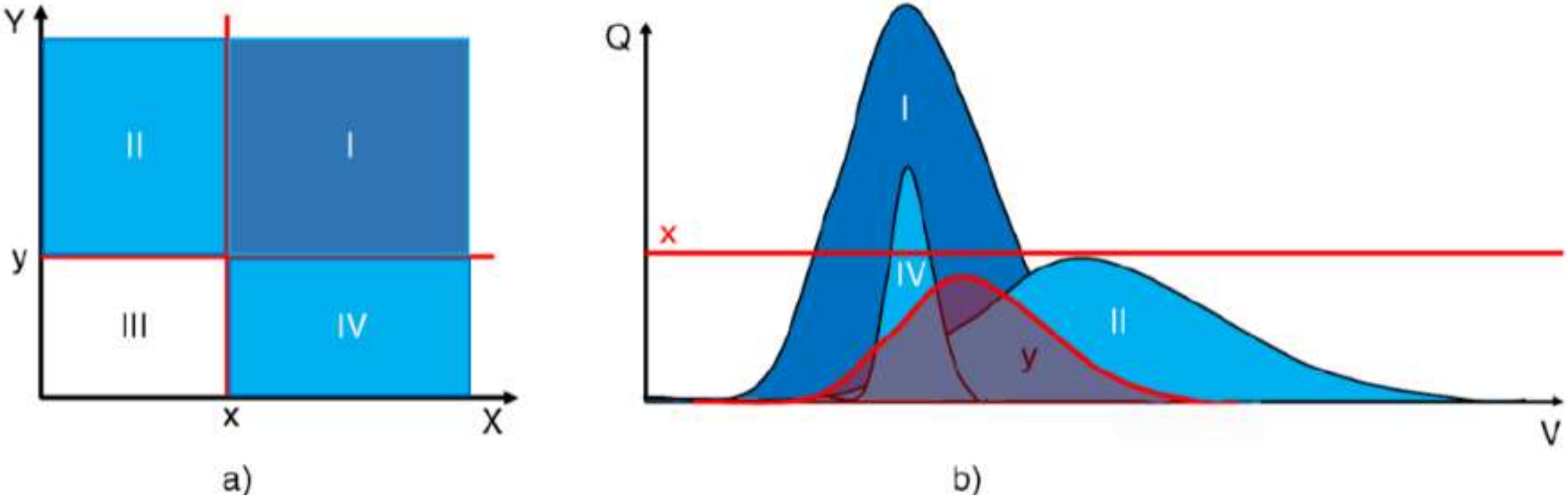
$$T_{u,v}^{AND} = \frac{\mu}{1-u-v+C_{u,v}(u,v)},$$

both peak and duration exceed their threshold simultaneously (**dual risk**)

BIVARIATE
RAINFALL FREQUENCY ANALYSIS

Illustration of joint probabilities

(from Brunner et al., 2016)



a) b)

Quadrant I: $\Pr\left[X > x, Y > y\right] = 1 - F_X(x) - F_Y(y) + F_{XY}(x, y) = S_{XY}(x, y)$

Quadrant II: $\Pr\left[X \leq x, Y > y\right]$

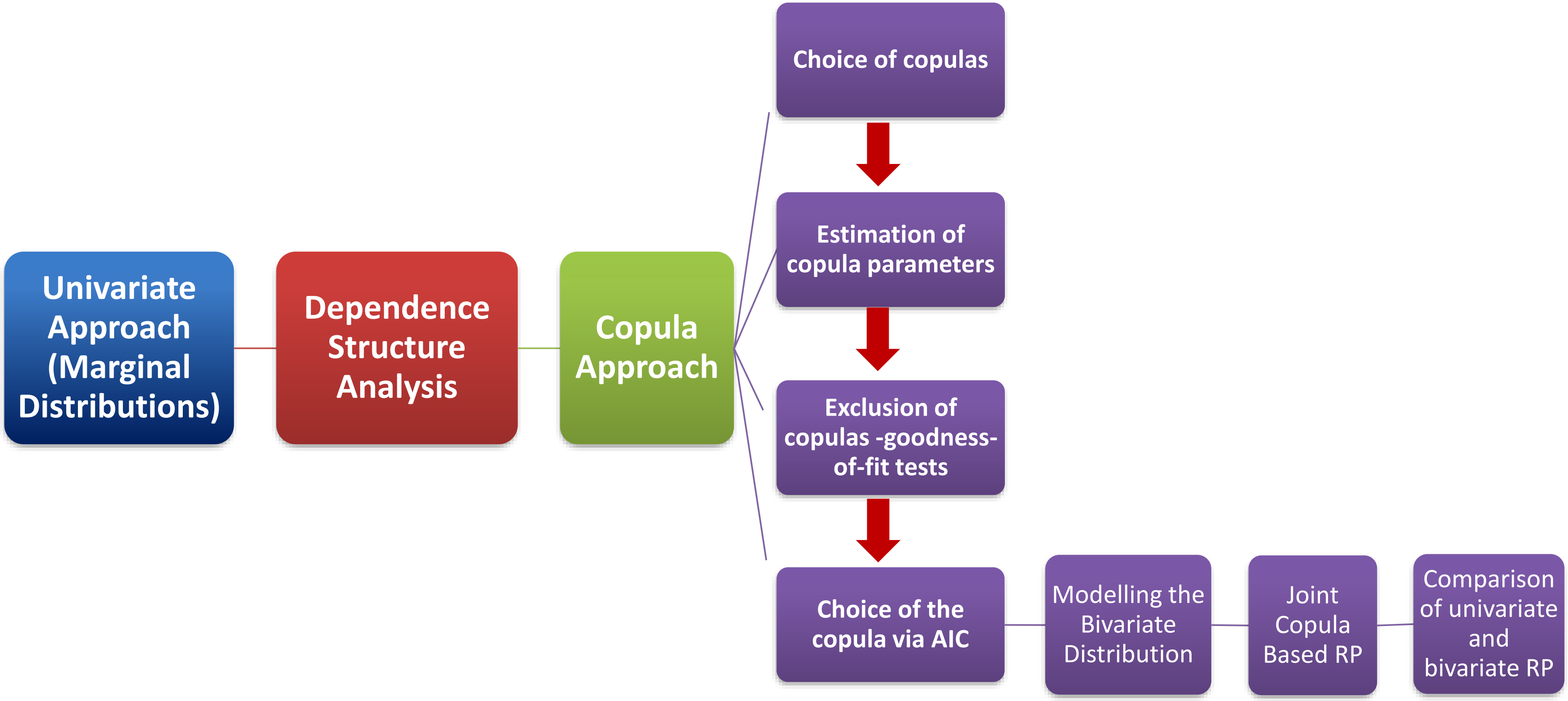
Quadrant III : $\Pr\left[X \leq x, Y \leq y\right] = F_{XY}(x, y)$

Quadrant IV : $\Pr\left[X > x, Y \leq y\right]$

Brunner M.I, Favre A.C., Seibert J. 2016. Bivariate return periods and their importance for flood peak and volume estimation. *WIREs Water* 2016. doi: 10.1002/wat2.1173

BIVARIATE RAINFALL FREQUENCY ANALYSIS

FLOW DIAGRAM OF THE METHOD



Univariate Approach (Marginal Distributions)

Dependence Structure Analysis

Copula Approach

Choice of copulas

Estimation of copula parameters

Exclusion of copulas -goodness-of-fit tests

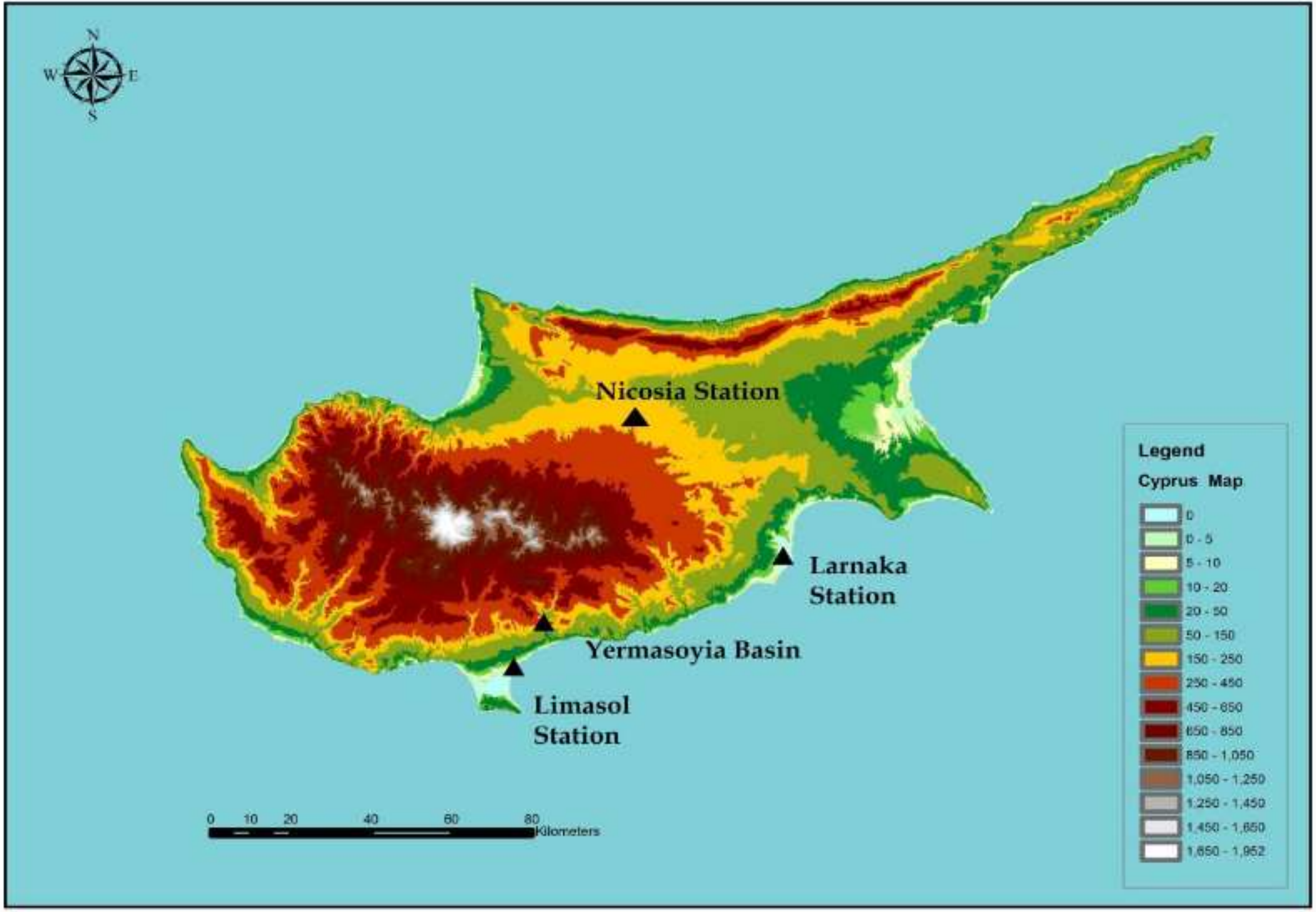
Choice of the copula via AIC

Modelling the Bivariate Distribution

Joint Copula Based RP

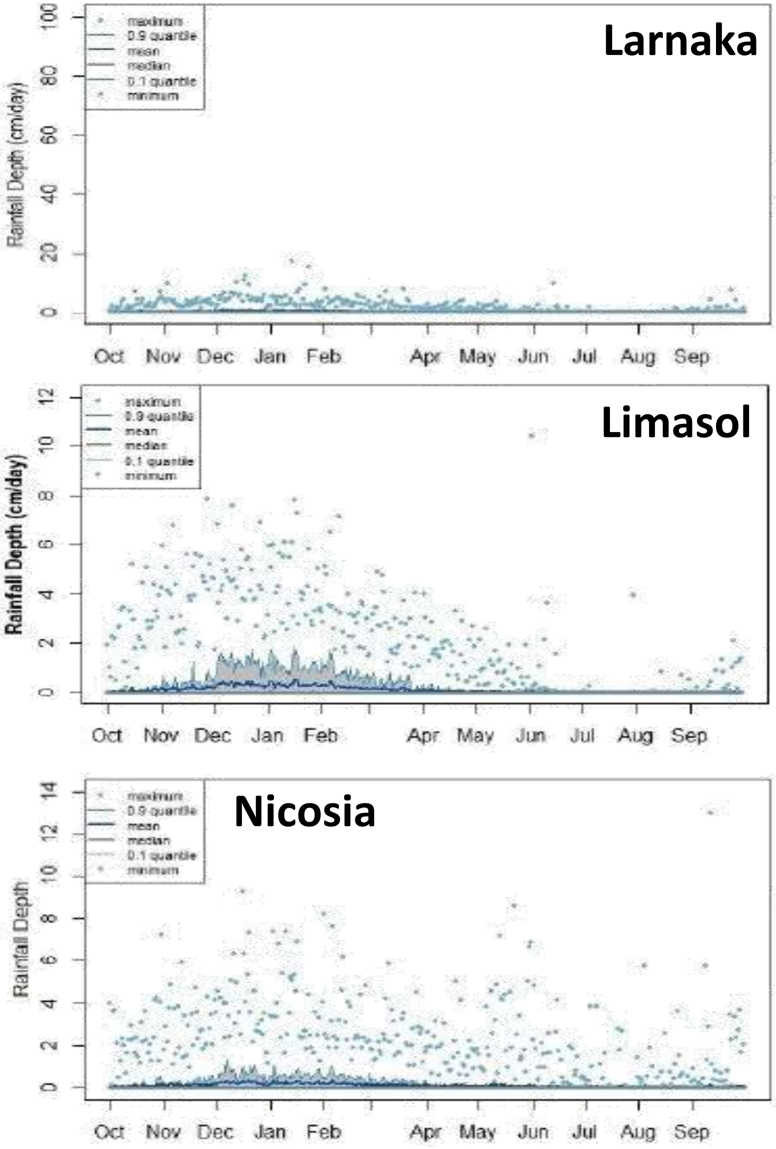
Comparison of univariate and bivariate RP

LOCATION OF METEOROLOGICAL STATIONS, CYPRUS



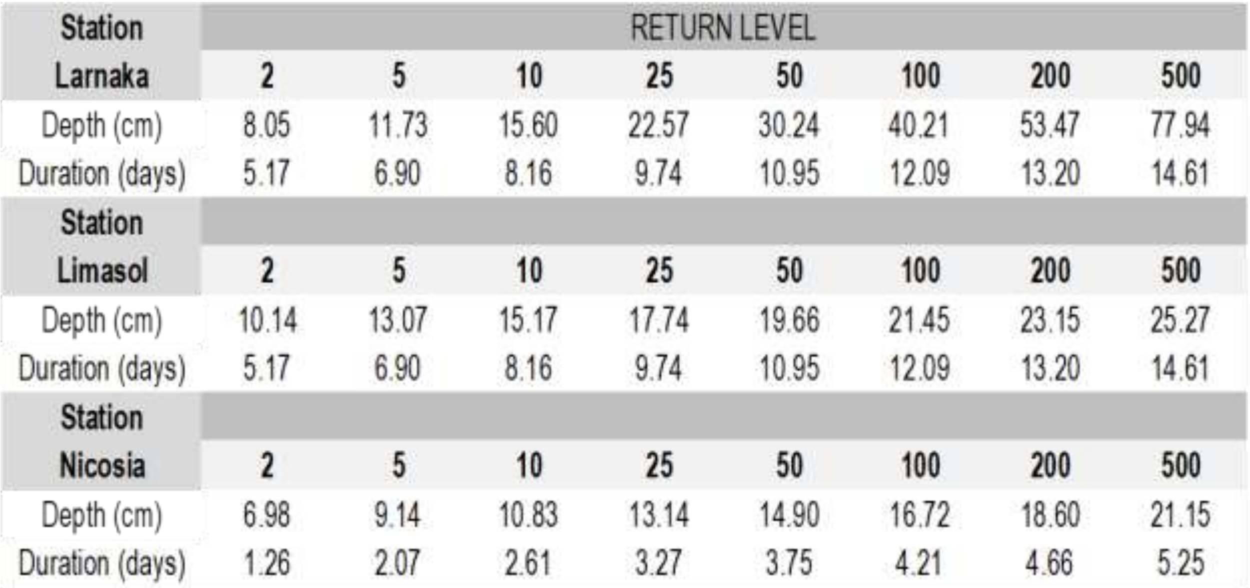
RAINFALL DATA, CYPRUS

- **Three Meteorological Stations** (Larnaka, Limassol and Nicosia)
  - 90 year daily rainfall data
    - **Historical period:** 1920 - 2100
  - **Rainfall Depth** (in cm) and **Rainfall Duration** (in days) were extracted.
  - **Annual Maxima Series** for both variables



UNIVARIATE RAINFALL FREQUENCY ANALYSIS

Determination of marginal distributions



Station	RETURN LEVEL							
Larnaka	2	5	10	25	50	100	200	500
Depth (cm)	8.05	11.73	15.60	22.57	30.24	40.21	53.47	77.94
Duration (days)	5.17	6.90	8.16	9.74	10.95	12.09	13.20	14.61
Station								
Limasol	2	5	10	25	50	100	200	500
Depth (cm)	10.14	13.07	15.17	17.74	19.66	21.45	23.15	25.27
Duration (days)	5.17	6.90	8.16	9.74	10.95	12.09	13.20	14.61
Station								
Nicosia	2	5	10	25	50	100	200	500
Depth (cm)	6.98	9.14	10.83	13.14	14.90	16.72	18.60	21.15
Duration (days)	1.26	2.07	2.61	3.27	3.75	4.21	4.66	5.25

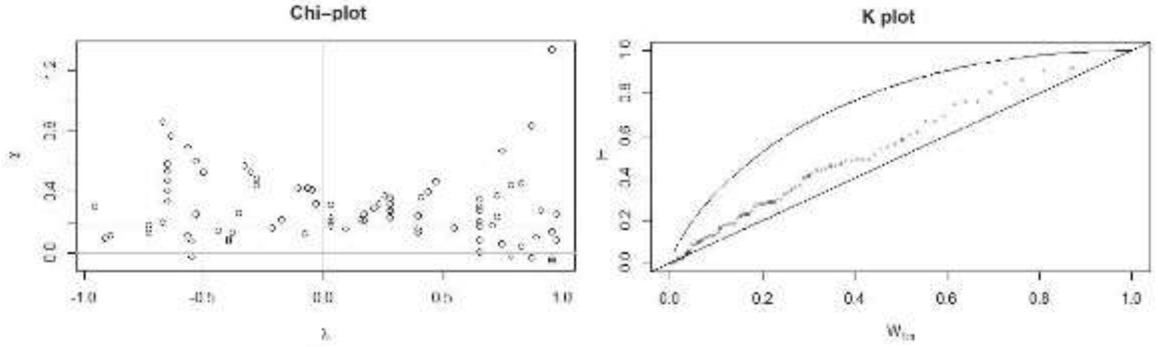
GEV

GEV

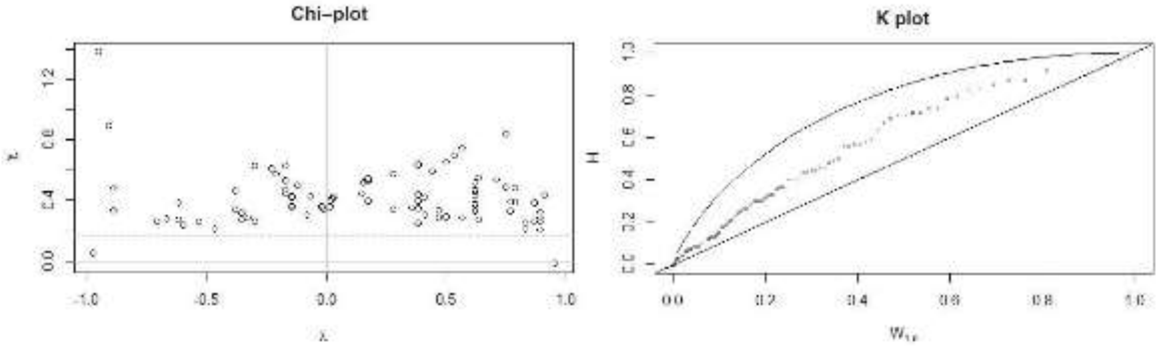
Gamma
GEV

- **GEV** distribution was the optimal probability model for both rainfall depth and duration for **Larnaka** and **Limassol** Stations.
- For **Nicosia** Station, **Gamma** distribution had a better fit for rainfall depth and **GEV** distribution for the hydrograph duration respectively.
- Finally, with the help of the marginal distributions, the **univariate return periods** are estimated for design return periods

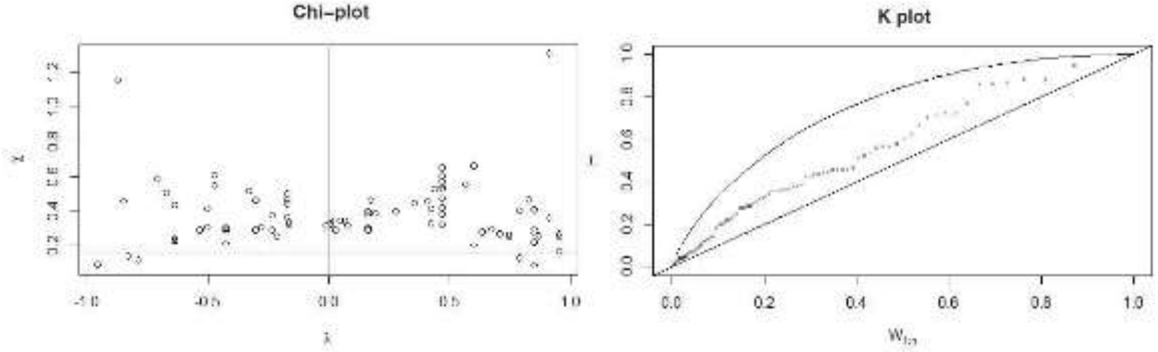
Dependence: Use of Chi and K plots



Larnaca Station



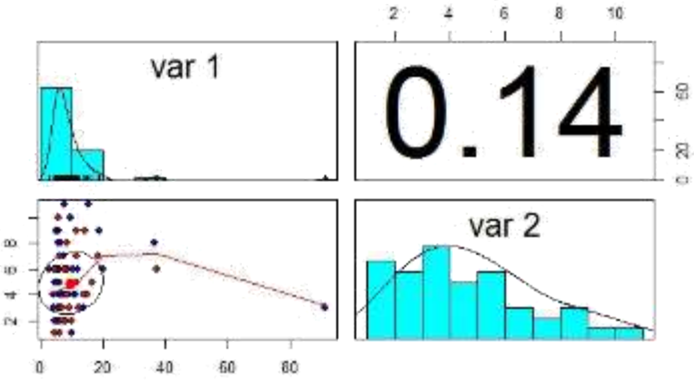
Limasol Station



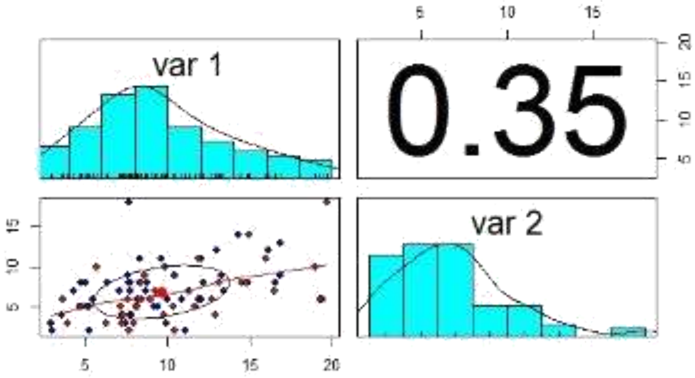
Nicosia Station

Dependence: Use of Kendall's tau coefficient

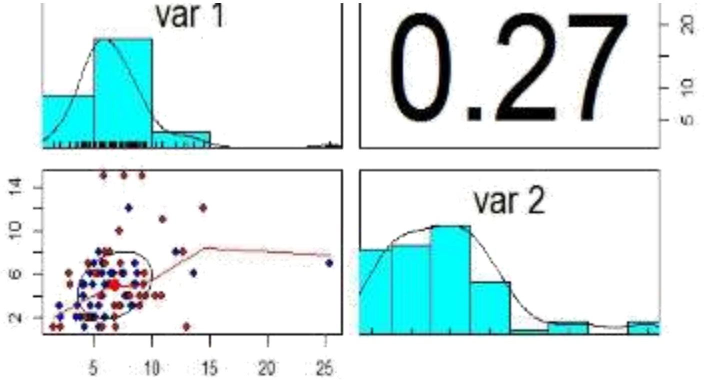
Larnaka



Limasol

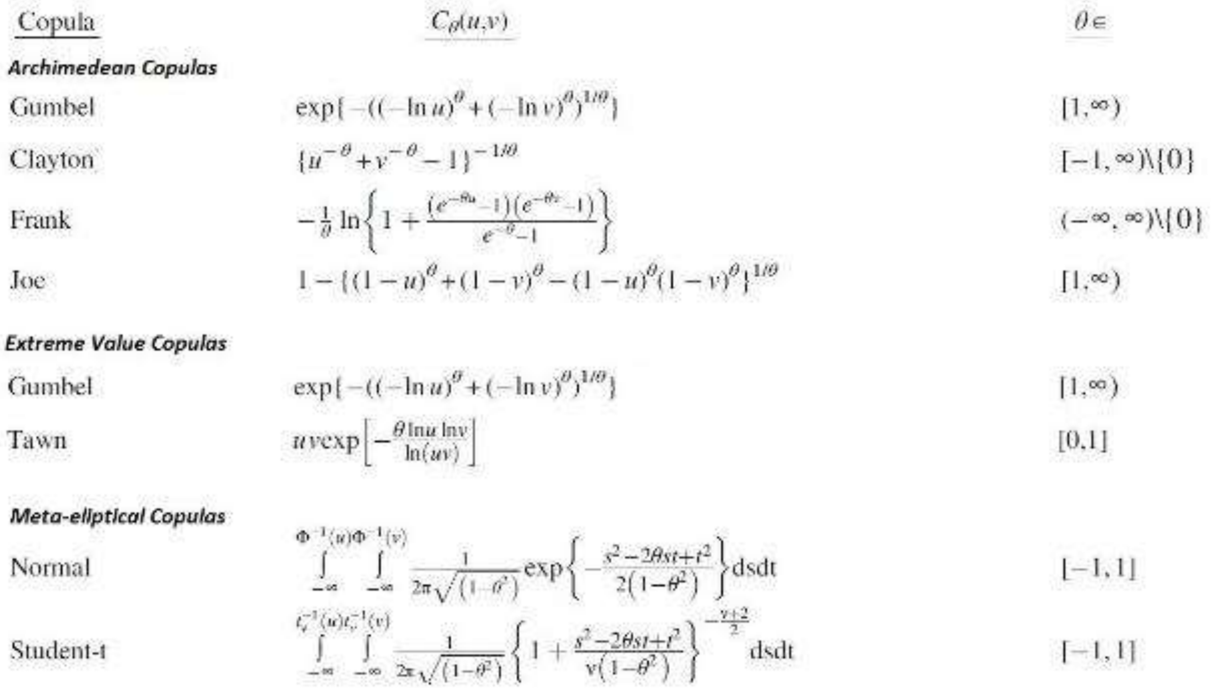


Nicosia



Fitting of a copula model

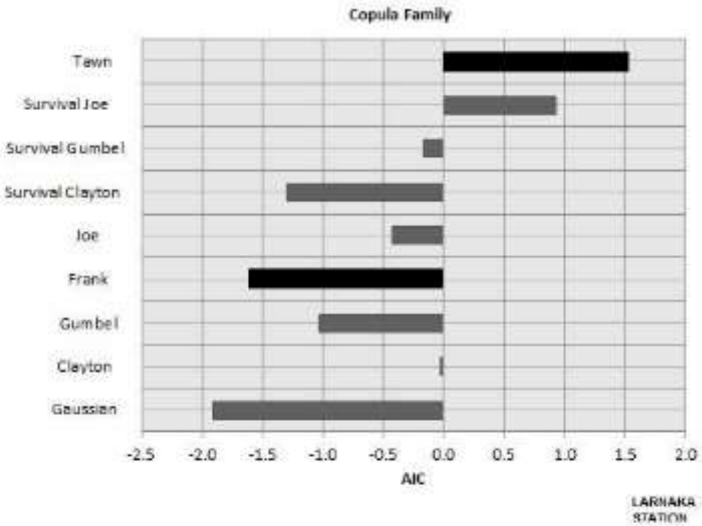
- **Copulas** from **Archimedean, Elliptical and Extreme Value families** are fitted using a pseudo-likelihood estimation method
  - **Evaluation procedure**
    - Graphical approaches and a goodness-of-fit test based on the Cramér von Mises statistic
    - **AICc**

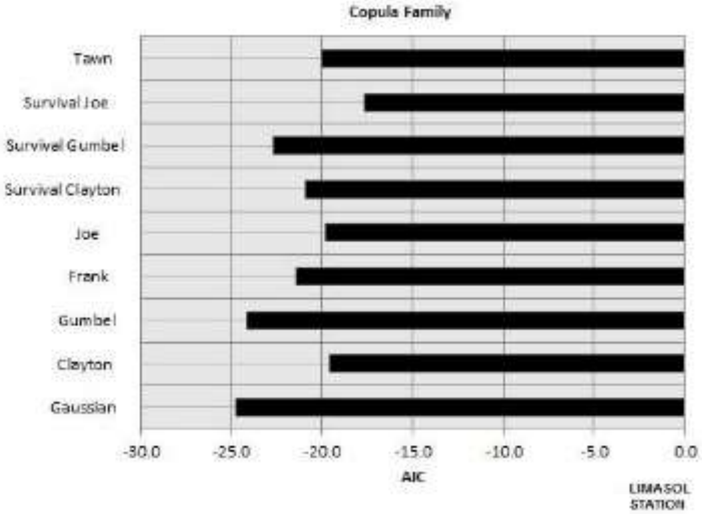


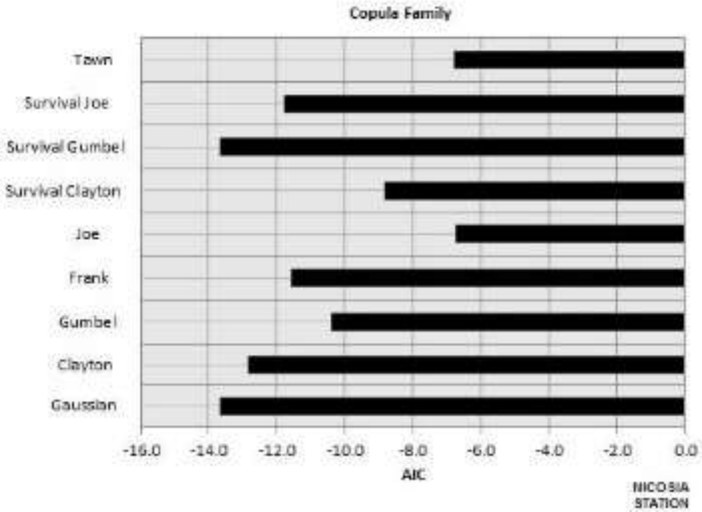
Copula	$C_\theta(u,v)$	$\theta\in$
Archimedean Copulas		
Gumbel	$\exp\{-((-\ln u)^\theta+(-\ln v)^\theta)^{1/\theta}\}$	$[1,\infty)$
Clayton	$\{u^{-\theta}+v^{-\theta}-1\}^{-1/\theta}$	$[-1,\infty)\setminus\{0\}$
Frank	$-\frac{1}{\theta}\ln\left\{1+\frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}\right\}$	$(-\infty,\infty)\setminus\{0\}$
Joe	$1-\{(1-u)^\theta+(1-v)^\theta-(1-u)^\theta(1-v)^\theta\}^{1/\theta}$	$[1,\infty)$
Extreme Value Copulas		
Gumbel	$\exp\{-((-\ln u)^\theta+(-\ln v)^\theta)^{1/\theta}\}$	$[1,\infty)$
Tawn	$uv\exp\left[-\frac{\theta\ln u\ln v}{\ln(uv)}\right]$	$[0,1]$
Meta-elliptical Copulas		
Normal	$\int_{-\infty}^{\Phi^{-1}(u)}\int_{-\infty}^{\Phi^{-1}(v)}\frac{1}{2\pi\sqrt{(1-\theta^2)}}\exp\left\{-\frac{s^2-2\theta st+t^2}{2(1-\theta^2)}\right\}\mathrm{dsdt}$	$[-1,1]$
Student-t	$\int_{-\infty}^{t_v^{-1}(u)}\int_{-\infty}^{t_v^{-1}(v)}\frac{1}{2\pi\sqrt{(1-\theta^2)}}\left\{1+\frac{s^2-2\theta st+t^2}{v(1-\theta^2)}\right\}^{-\frac{v+2}{2}}\mathrm{dsdt}$	$[-1,1]$

BIVARIATE RAINFALL FREQUENCY ANALYSIS

Fitting of a copula model







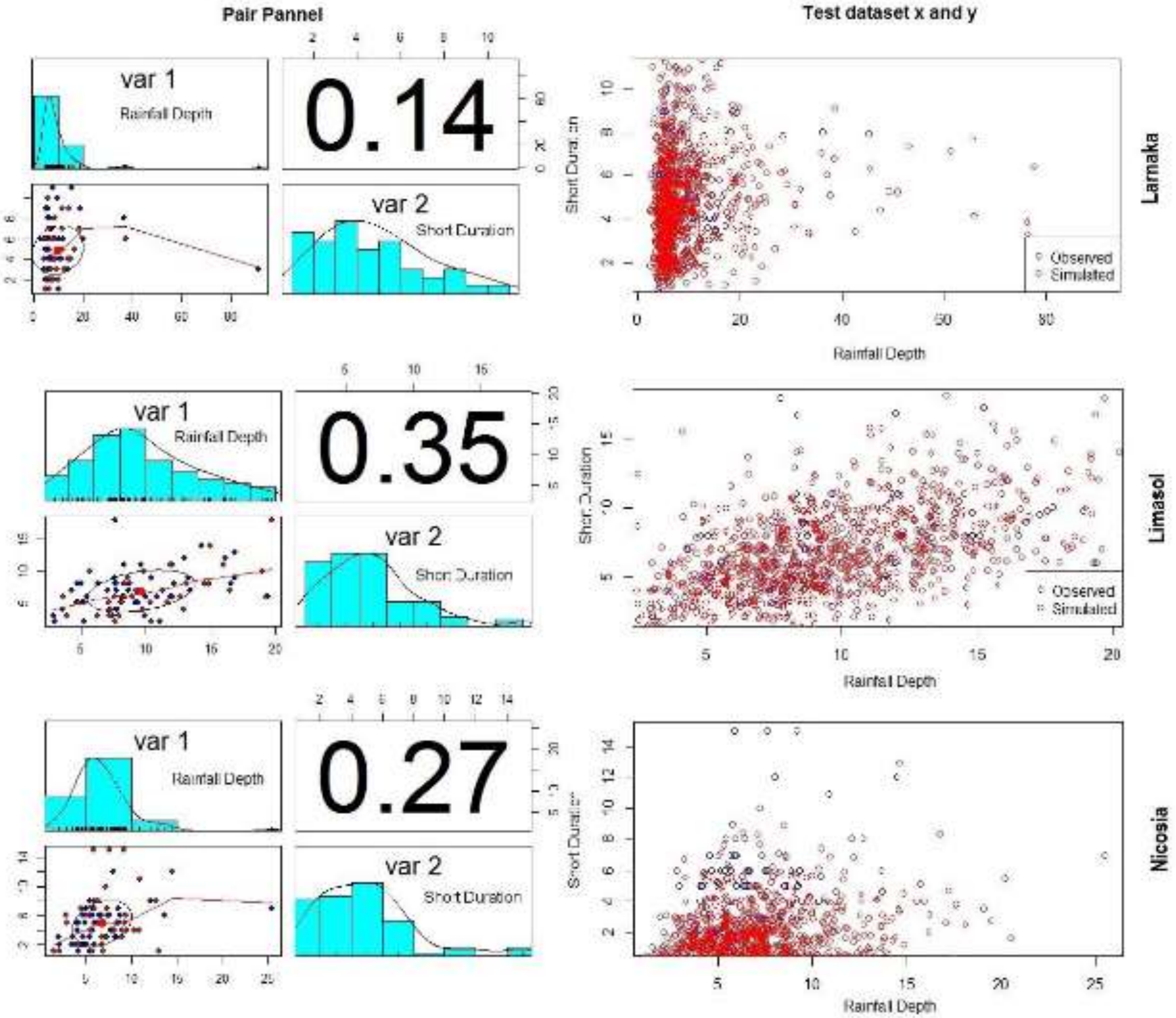
AIC values and accepted copulas for the three rainfall stations

■ Accepted Copula

▨ Rejected Copula

BIVARIATE
RAINFALL FREQUENCY ANALYSIS

Modelling dependence and fitting of copulas

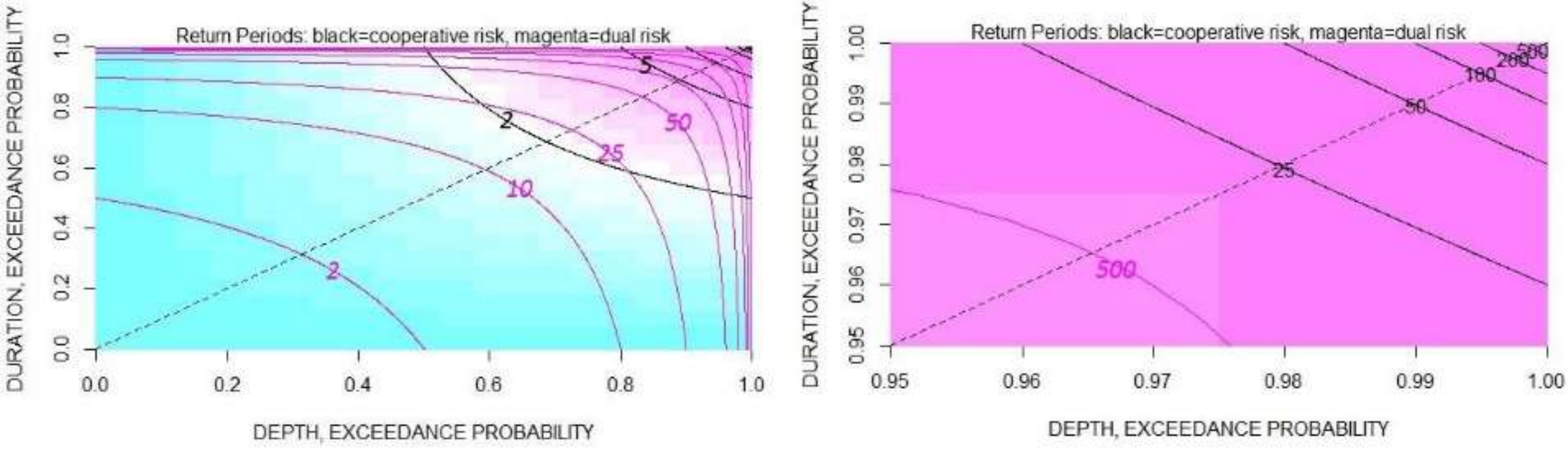


BIVARIATE RAINFALL FREQUENCY ANALYSIS

Results: AMS Joint Return Periods

$T^{OR} < T^{UNI} < T^{AND}$

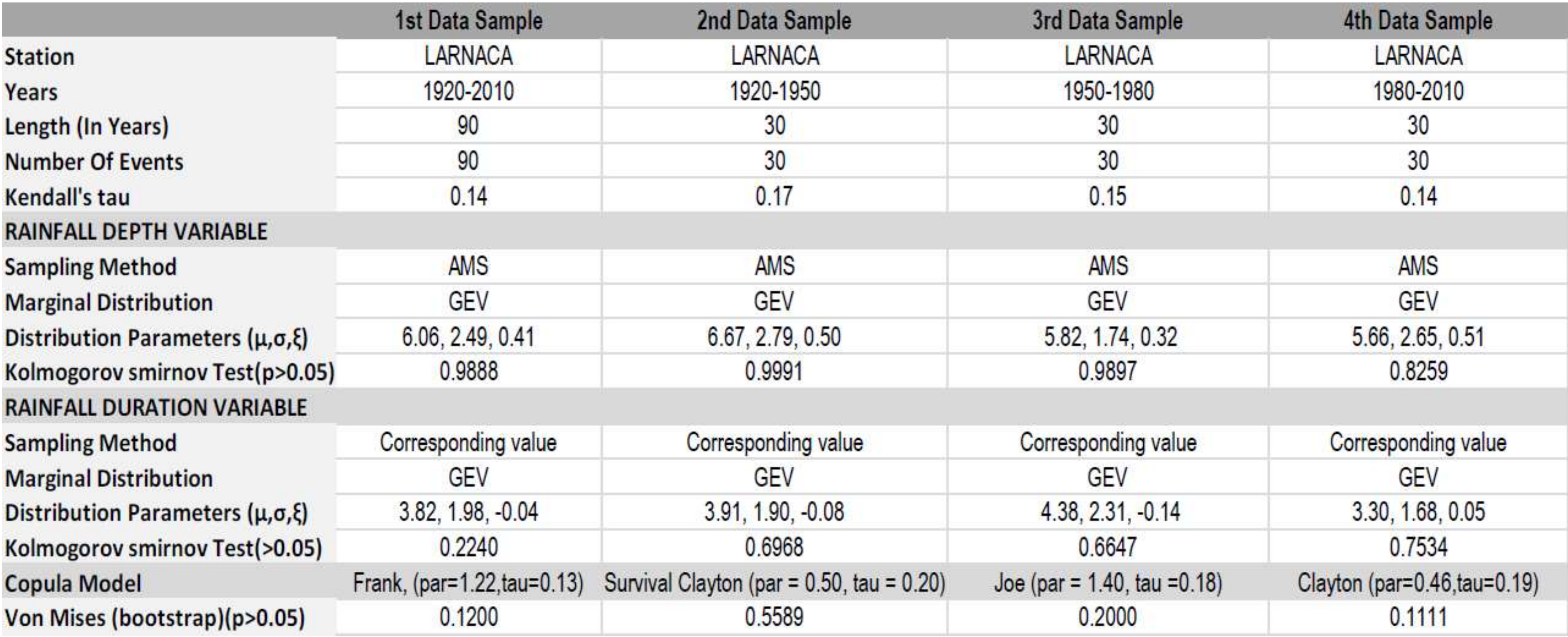
Station	RETURN LEVEL							
Larnaka	**2**	**5**	**10**	**25**	**50**	**100**	**200**	**500**
Depth (cm)/ dual	5.33	6.69	7.36	7.88	8.10	8.21	8.27	8.30
Depth (cm)/ cooperative	8.55	13.75	18.80	27.85	37.22	49.66	66.09	78.07
Duration (d)/ dual	3.80	5.75	7.03	8.65	9.84	11.00	12.14	13.59
Duration (d)/ cooperative	6.03	8.25	9.62	11.25	12.41	13.52	14.59	24.44



STATION LARNACA

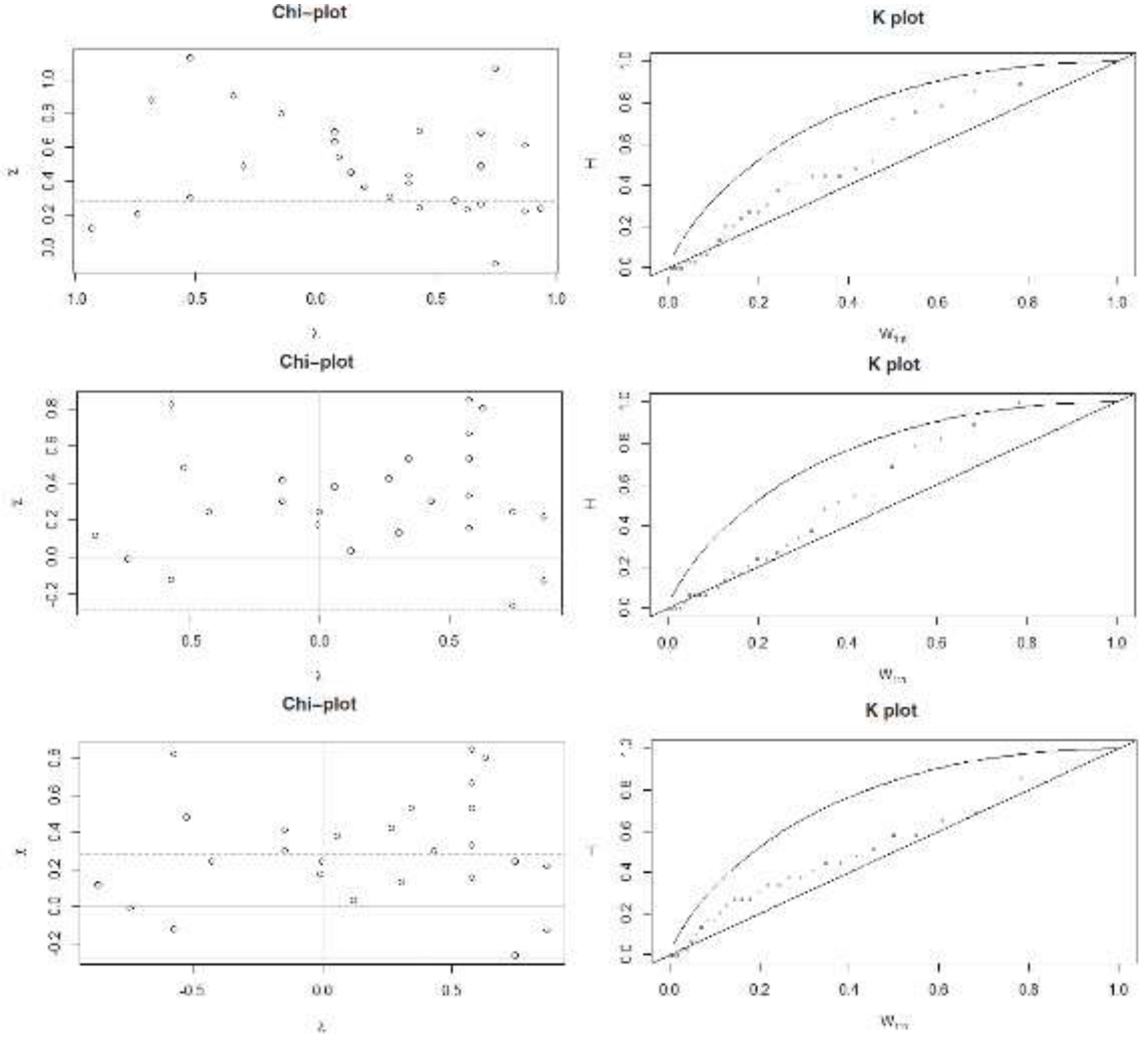
BIVARIATE
RAINFALL FREQUENCY ANALYSIS

Effect of Sample Size: Larnaca Station



	1st Data Sample	2nd Data Sample	3rd Data Sample	4th Data Sample
Station	LARNACA	LARNACA	LARNACA	LARNACA
Years	1920-2010	1920-1950	1950-1980	1980-2010
Length (In Years)	90	30	30	30
Number Of Events	90	30	30	30
Kendall's tau	0.14	0.17	0.15	0.14
RAINFALL DEPTH VARIABLE				
Sampling Method	AMS	AMS	AMS	AMS
Marginal Distribution	GEV	GEV	GEV	GEV
Distribution Parameters (μ,σ,ξ)	6.06, 2.49, 0.41	6.67, 2.79, 0.50	5.82, 1.74, 0.32	5.66, 2.65, 0.51
Kolmogorov smirnov Test(p>0.05)	0.9888	0.9991	0.9897	0.8259
RAINFALL DURATION VARIABLE				
Sampling Method	Corresponding value	Corresponding value	Corresponding value	Corresponding value
Marginal Distribution	GEV	GEV	GEV	GEV
Distribution Parameters (μ,σ,ξ)	3.82, 1.98, -0.04	3.91, 1.90, -0.08	4.38, 2.31, -0.14	3.30, 1.68, 0.05
Kolmogorov smirnov Test(>0.05)	0.2240	0.6968	0.6647	0.7534
Copula Model	Frank, (par=1.22,tau=0.13)	Survival Clayton (par = 0.50, tau = 0.20)	Joe (par = 1.40, tau =0.18)	Clayton (par=0.46,tau=0.19)
Von Mises (bootstrap)(p>0.05)	0.1200	0.5589	0.2000	0.1111

Effect of Sample Size on Dependence: Larnaca Station



Years: 1920-1950

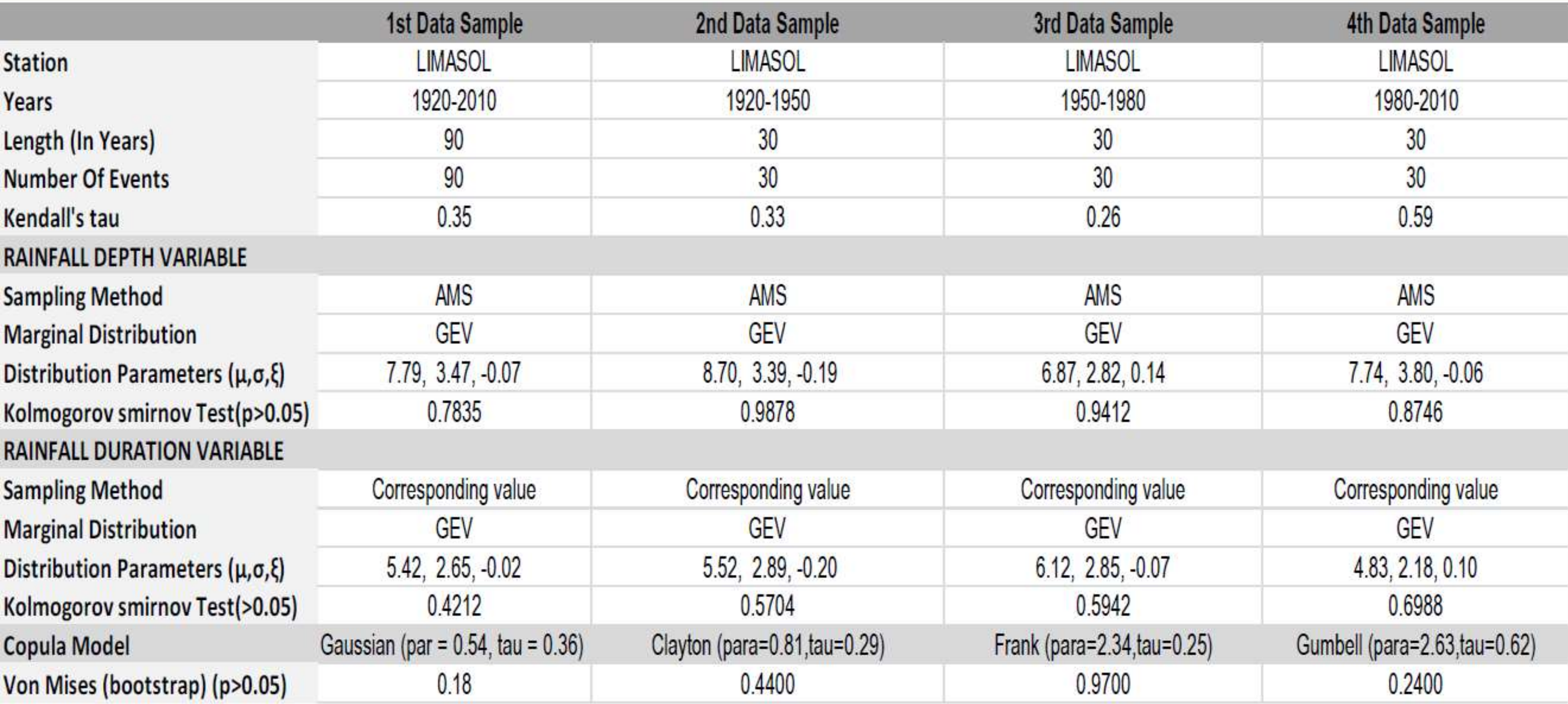
Larnaka Station

Years: 1950-1980

Years:1980-2010

BIVARIATE
RAINFALL FREQUENCY ANALYSIS

Effect of Sample Size: Limasol Station



	1st Data Sample	2nd Data Sample	3rd Data Sample	4th Data Sample
Station	LIMASOL	LIMASOL	LIMASOL	LIMASOL
Years	1920-2010	1920-1950	1950-1980	1980-2010
Length (In Years)	90	30	30	30
Number Of Events	90	30	30	30
Kendall's tau	0.35	0.33	0.26	0.59
RAINFALL DEPTH VARIABLE				
Sampling Method	AMS	AMS	AMS	AMS
Marginal Distribution	GEV	GEV	GEV	GEV
Distribution Parameters (μ,σ,ξ)	7.79, 3.47, -0.07	8.70, 3.39, -0.19	6.87, 2.82, 0.14	7.74, 3.80, -0.06
Kolmogorov smirnov Test(p>0.05)	0.7835	0.9878	0.9412	0.8746
RAINFALL DURATION VARIABLE				
Sampling Method	Corresponding value	Corresponding value	Corresponding value	Corresponding value
Marginal Distribution	GEV	GEV	GEV	GEV
Distribution Parameters (μ,σ,ξ)	5.42, 2.65, -0.02	5.52, 2.89, -0.20	6.12, 2.85, -0.07	4.83, 2.18, 0.10
Kolmogorov smirnov Test(>0.05)	0.4212	0.5704	0.5942	0.6988
Copula Model	Gaussian (par = 0.54, tau = 0.36)	Clayton (para=0.81,tau=0.29)	Frank (para=2.34,tau=0.25)	Gumbell (para=2.63,tau=0.62)
Von Mises (bootstrap) (p>0.05)	0.18	0.4400	0.9700	0.2400

Effect of Sample Size: Nicosia Station

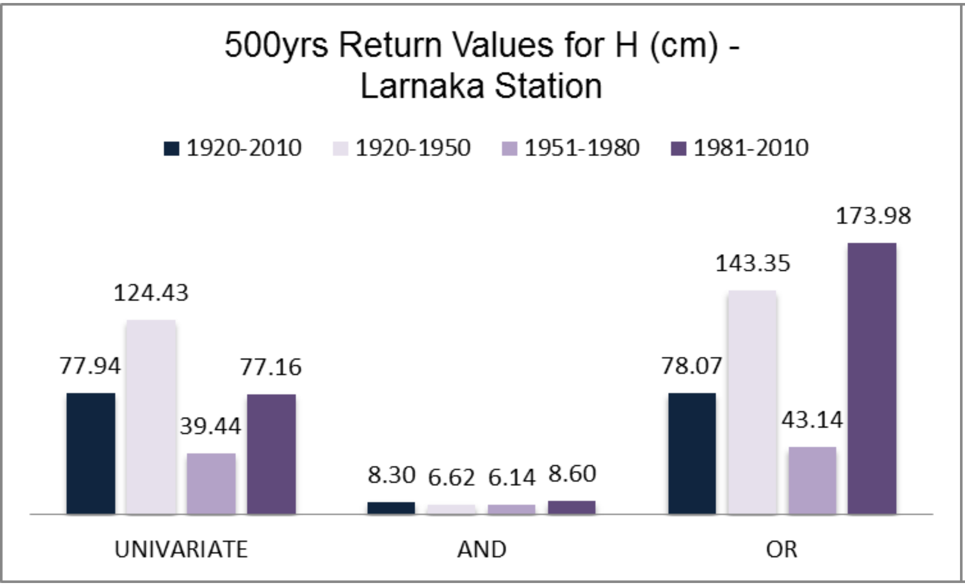
	1st Data Sample	2nd Data Sample	3rd Data Sample	4th Data Sample
Station	NICOSIA	NICOSIA	NICOSIA	NICOSIA
Years	1920-2010	1920-1950	1950-1980	1980-2010
Length (In Years)	90	30	30	30
Number Of Events	90	30	30	30
Kendall's tau	0.27	0.33	0.19	0.18
RAINFALL DEPTH VARIABLE				
Sampling Method	AMS	AMS	AMS	AMS
Marginal Distribution	GEV	GEV	GEV	GEV
Distribution Parameters (μ,σ,ξ)	5.39, 2.26, 0.04	6.04, 2.54, 0.12	5.09, 2.16, -0.02	5.57, 2.13, -0.44
Kolmogorov smirnov Test(p>0.05)	0.9071	0.9909	0.8721	0.4083
RAINFALL DURATION VARIABLE				
Sampling Method	Corresponding value	Corresponding value	Corresponding value	Corresponding value
Marginal Distribution	GAMMA	GAMMA	GEV	GEV
Distribution Parameters (μ,σ,ξ)	shape= 2.66,rate= 0.54	shape= 2.68,rate= 0.48	2.81, 1.95, 0.31	3.54, 1.72, 0.06
Kolmogorov smirnov Test(>0.05)	0.1094	0.4000	0.4600	0.5469
Copula Model	Survival Gumbel (par = 1.37, tau = 0.27)	Clayton (para=1.18,tau=0.37)	Joe (para=1.43,tau=0.19)	Survival Joe (par = 1.62, tau = 0.26)
Von Mises (bootstrap) (p>0.05)	0.1425	0.1000	0.2000	0.0600

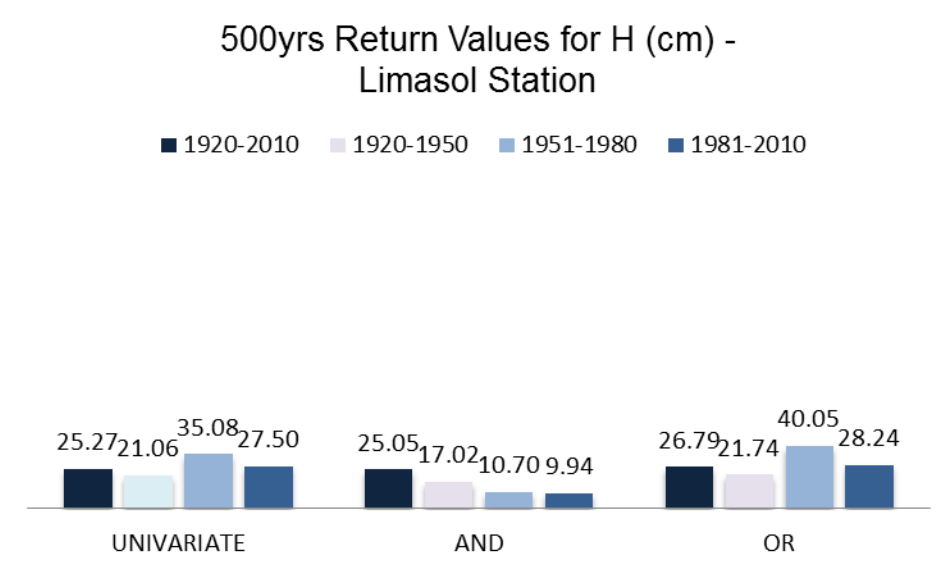
BIVARIATE RAINFALL FREQUENCY ANALYSIS

Effect of Sample Size: Results

Kendall's tau constant in all applications

Kendall's tau indicated much stronger correlation during the 30 last years

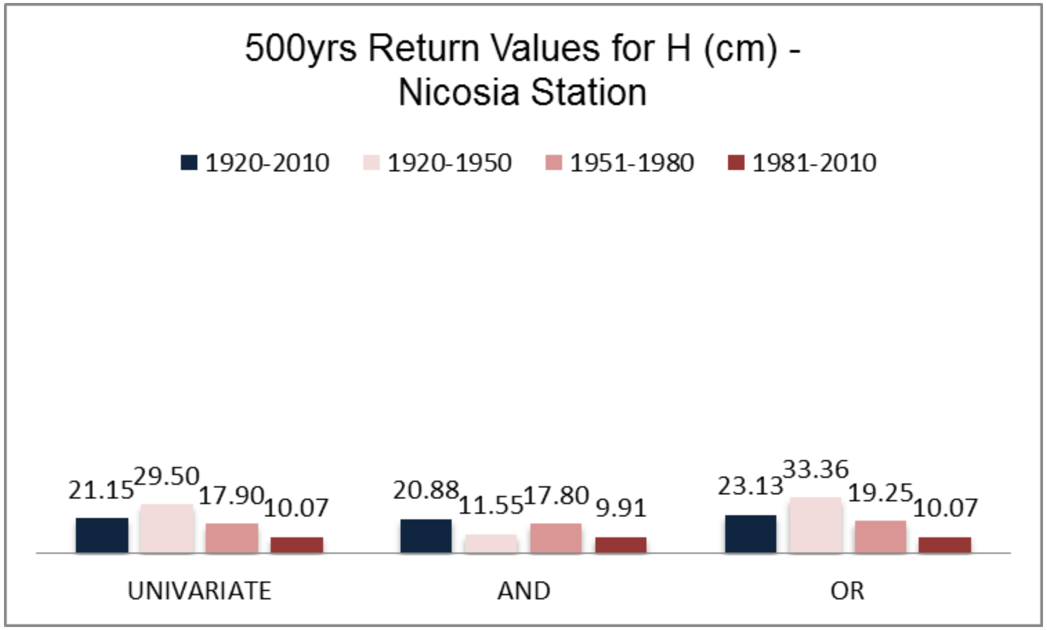


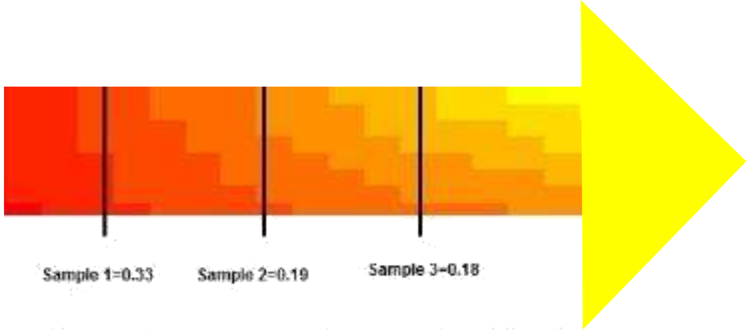


$T^{OR} < T^{UNI} < T^{AND}$ relationship

Significant inconsistences in AND and OR cases especially in Larnaka Station

Effect of Sample Size: Results





Kendal's tau reduces with time

$T^{OR} < T^{UNI} < T^{AND}$ relationship

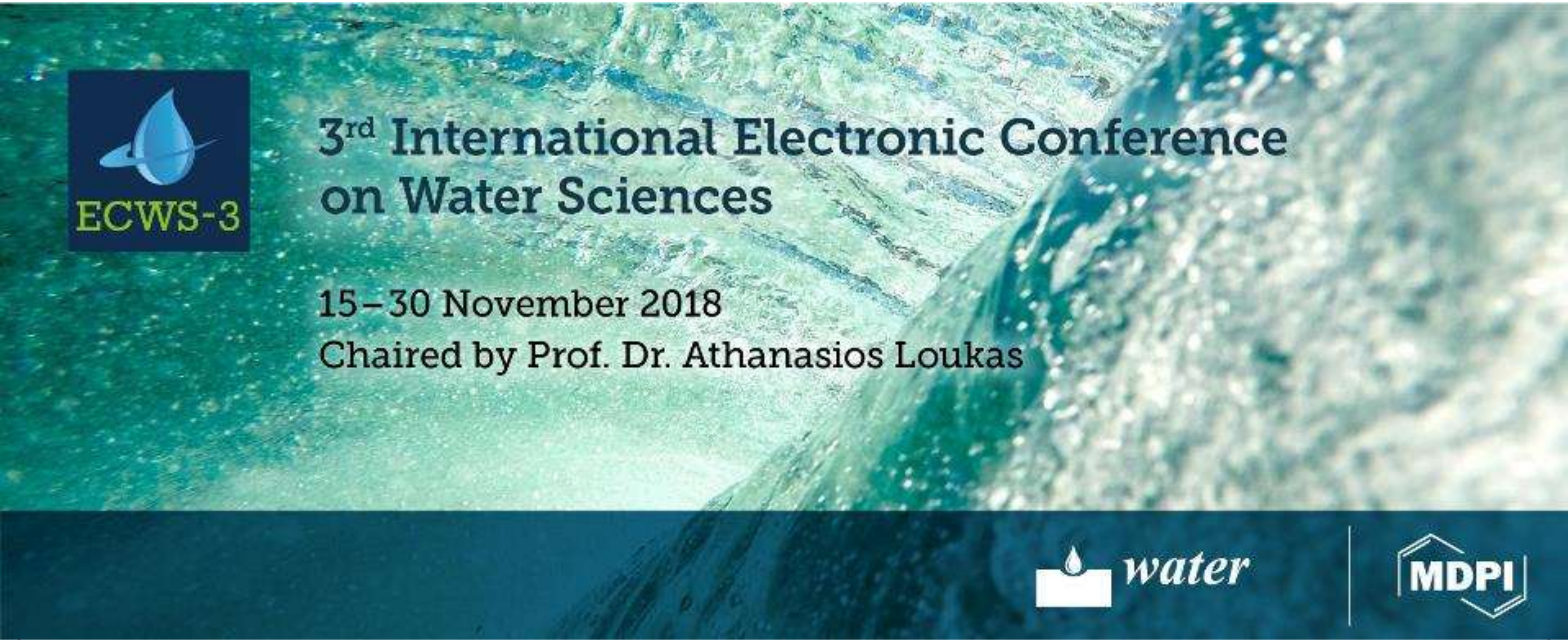
CONCUDING REMARKS

- Results show that u**nivariate analysis** can't provide a complete assessment of the probability of occurrence of extreme rainfall if two or more dependent variables are significant in the design process
  - **univariate** approaches might lead to an inadequate estimation of the risk associated with a given event
- **Minor dependence** between **rainfall peaks** and **storm duration**
  - **bivariate analysis** could be considered in the estimation of **design values**
- **Sample size** has large impacts on the derived results
  - Further investigation is needed for variable data lengths (small and large samples)
- **Design values** at the study **return periods** are in consensus with Salvadori et al., (2007) following the equation $T^{OR} < T^{UNI} < T^{AND}$

Salvadori G, De Michele C, Kottegoda NT, Rosso R. Extremes in Nature. An Approach Using Copulas, vol. 56. Dordrecht: Springer; 2007, 292 p.

THANK YOU FOR YOUR ATTENTION!

Dr. Lampros Vasiliades
Laboratory of Hydrology and Aquatic Systems Analysis
Department of Civil Engineering, School of Engineering, University of Thessaly
Pedion Areos, 38334 Volos, Greece
E-mail: lvassil@civ.uth.gr


ECWS-3

3rd International Electronic Conference on Water Sciences

15–30 November 2018
Chaired by Prof. Dr. Athanasios Loukas