Lecture 2

APPLICATION OF EXREME VALUE THEORY TO CLIMATE CHANGE

Rick Katz

Institute for Mathematics Applied to Geosciences
National Center for Atmospheric Research
Boulder, CO USA

email: `rwk@ucar.edu`

Home page: `www.isse.ucar.edu/staff/katz/`

Lecture: `.../staff/katz/docs/pdf/ubalect2.pdf`

Quote

"In order to apply any theory we have to suppose that the data are homogeneous, i.e. that no systematical change of climate and no important change in the basin have occurred within the observation period and that no such changes will take place in the period for which extrapolations are made."

Emil Gumbel

(*Ann. Math. Stat.*, 1941)

Outline

(1) Non-Stationarity

(2) Trends in Extremes

(3) Other Forms of Covariates

(4) Risk Communication (Under Non-Stationarity)

(5) Economic Impacts / Random Sums

(6) Extreme Weather Spells

(7) Origin of Bounded and Heavy Tails

(8) Multivariate Extremes

(9) Spatial Extremes

(1) Non-Stationarity

- **Sources**

-- Trends

Global climate change:

Trends in frequency & intensity of extreme weather events

-- Cycles

Annual & diurnal natural aspect of climate

-- Physically-based

Increased precision, improved realism

- **Theory**

-- No general extreme value theory under non-stationarity
Only limited results under restrictive conditions

- **Methods**

-- Introduction of covariates resembles generalized linear models
Only limited connection (e. g., Weibull case)

-- Straightforward to extend MLE

- **Issues**

-- Nature of relationship between extremes & covariates
Resembles that for overall / center of data?

(2) Trends in Extremes

- **Trends**

-- Example (Urban heat island)

Trend in summer minimum of daily minimum temperature at Phoenix, AZ (i. e., block minima)

$$\min\{X_1, X_2, \ldots, X_T\} = -\max\{-X_1, -X_2, \ldots, -X_T\}$$

Assume summer minimum temperature in year *t* has GEV distribution with location & scale parameters:

$$\mu(t) = \mu_0 + \mu_1 t, \quad \log \sigma(t) = \sigma_0 + \sigma_1 t, \quad \xi(t) = \xi, \quad t = 1, 2, \ldots$$

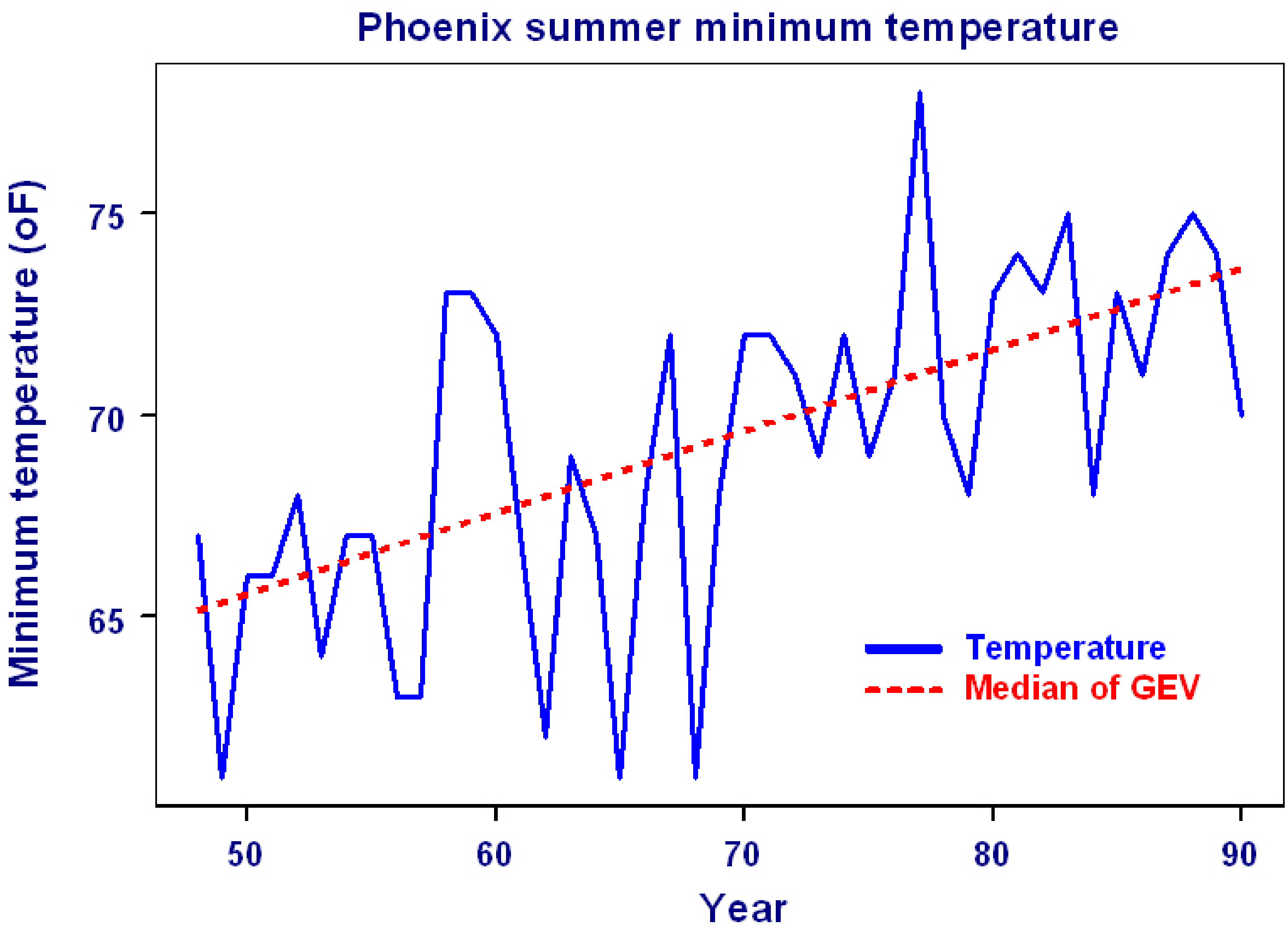
- **Parameter estimates and standard errors**

Parameter		Estimate	(Std. Error)
Location:	μ_0	66.17*	
	μ_1	0.196*	(0.041)
Scale:	σ_0	1.338	
	σ_1	−0.009	(0.010)
Shape:	ξ	−0.211	

***Sign of location parameters reversed to convert back to minima**

-- Likelihood ratio test for $\mu_1 = 0$ (*P*-value < 10^{-5})

-- Likelihood ratio test for $\sigma_1 = 0$ (*P*-value ≈ 0.366)


Phoenix summer minimum temperature
Minimum temperature (oF)
75
70
65
50
60
70
80
90
Year
Temperature
Median of GEV

- **Q-Q plots under non-stationarity**

-- Transform to common distribution

Non-stationary GEV [$\mu(t)$, $\sigma(t)$, $\xi(t)$]

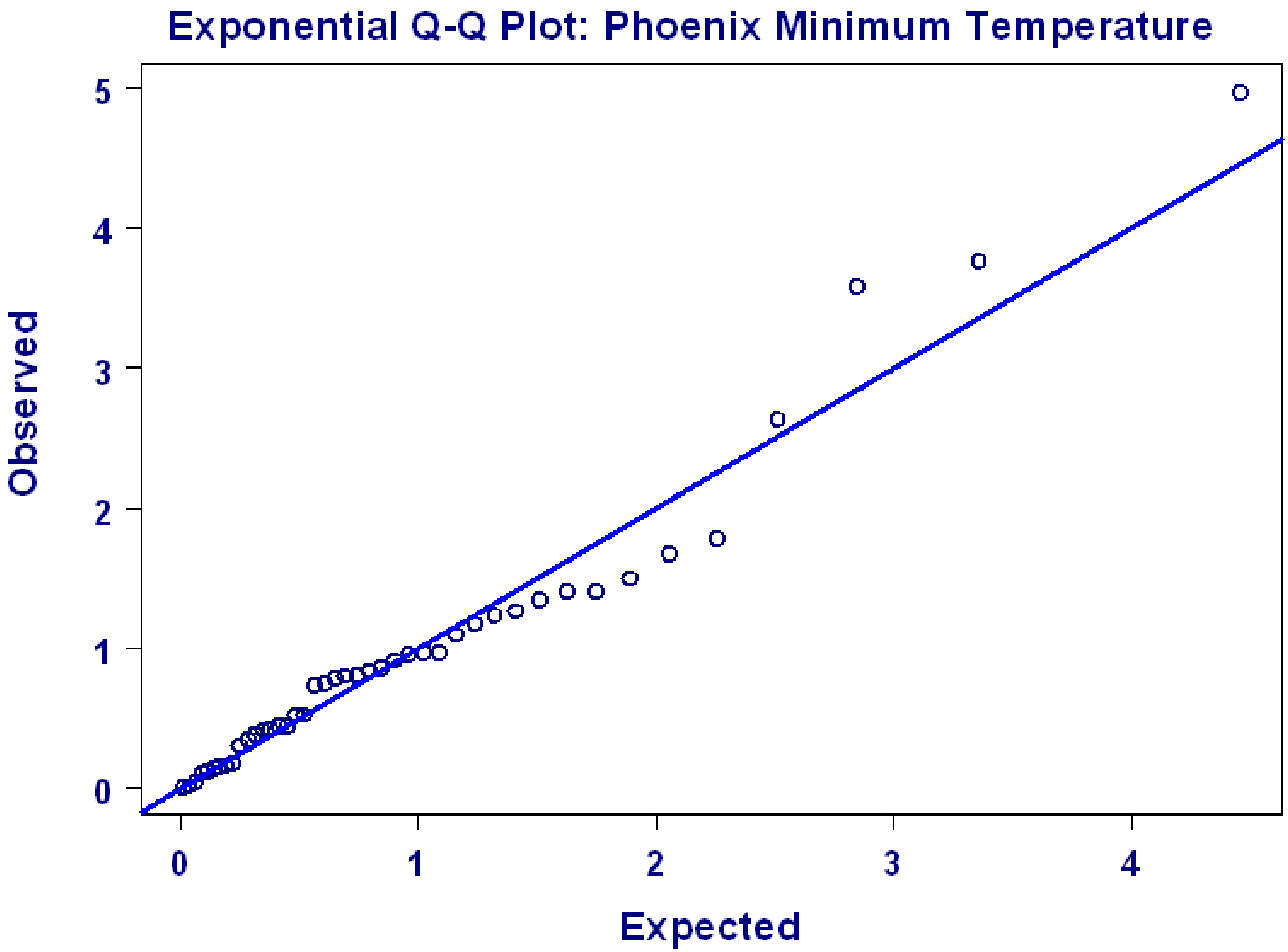
Transform to Gumbel or exponential distribution

(i) Non-stationary GEV to exponential

$$\varepsilon_t = \{1 + \xi(t)\,[M_t - \mu(t)] \,/\, \sigma(t)\}^{-1/\xi(t)}$$

(ii) Non-stationary GEV to Gumbel (used by `extRemes`)

$$\varepsilon_t = [1/\xi(t)] \log \{1 + \xi(t)\,[M_t - \mu(t)] \,/\, \sigma(t)\}$$


Exponential Q-Q Plot: Phoenix Minimum Temperature
Observed
Expected
0
1
2
3
4
5
0
1
2
3
4

(3) Other Forms of Covariates

- **Physically-based covariates**

-- Example [Arctic Oscillation (AO)]

Winter maximum of daily maximum temperature at Port Jervis, NY (i. e., block maxima)

Z denotes winter index of AO

Given $Z = z$, assume conditional distribution of winter max. temp. is GEV distribution with parameters:

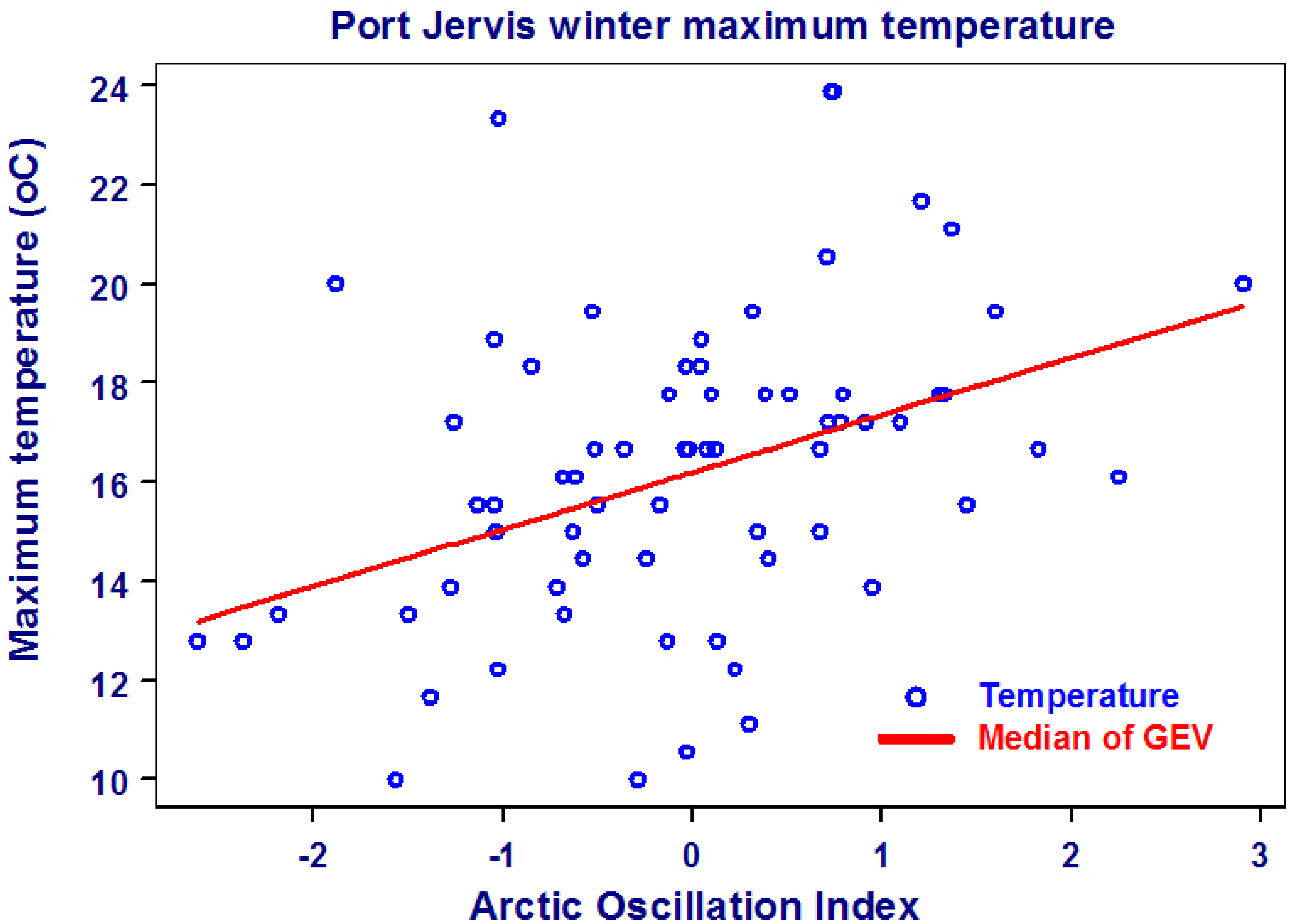
$$\mu(z) = \mu_0 + \mu_1 z, \quad \log \sigma(z) = \sigma_0 + \sigma_1 z, \qquad \xi(z) = \xi$$

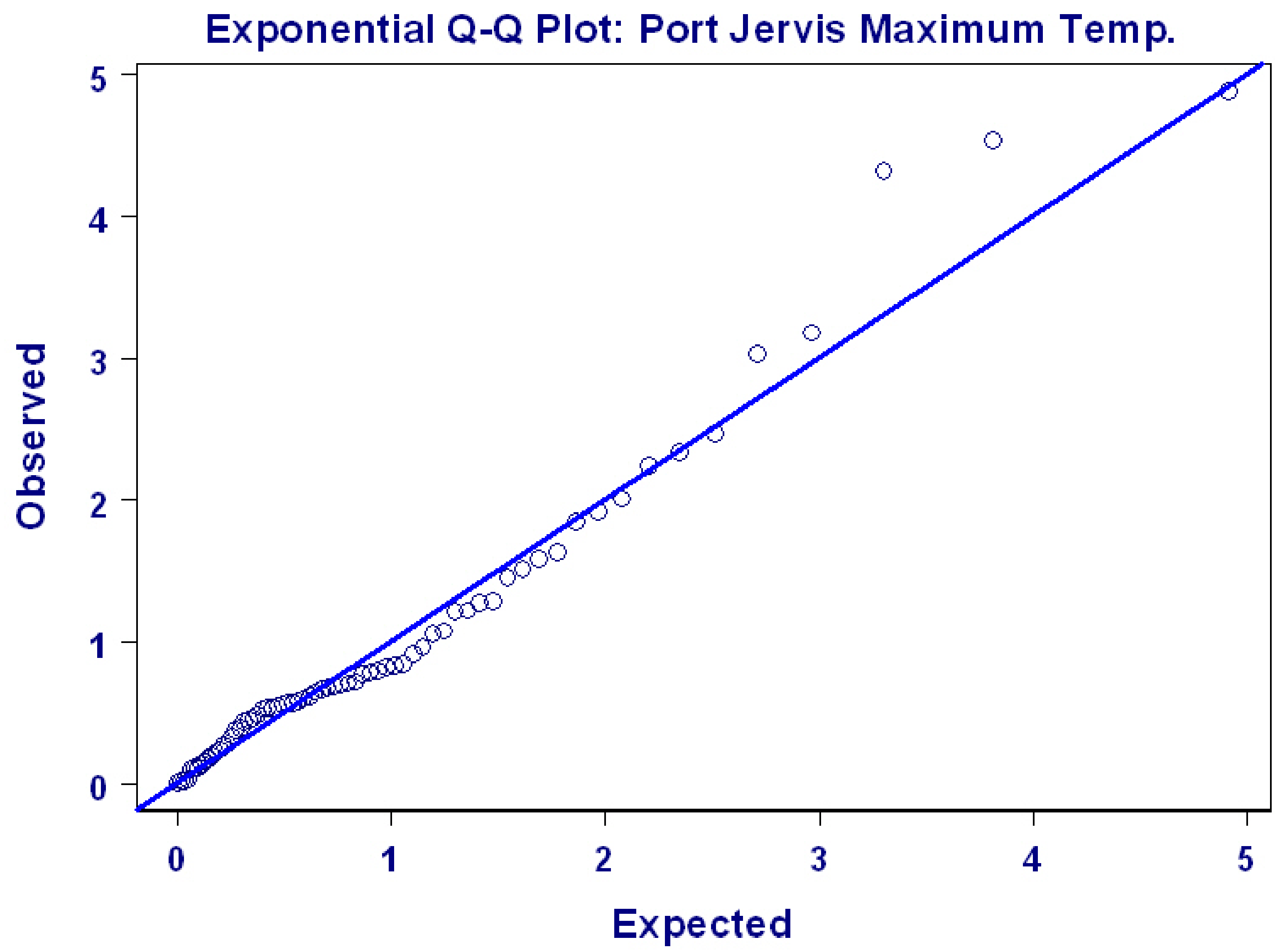
- **Parameter estimates and standard errors**

Parameter		**Estimate**	**(Std. Error)**
Location:	μ_0	**15.26**	
	μ_1	**1.175**	**(0.319)**
Scale:	σ_0	**0.984**	
	σ_1	**−0.044**	**(0.092)**
Shape:	ξ	**−0.186**	

-- Likelihood ratio test for $\mu_1 = 0$ (*P*-value < 0.001)

-- Likelihood ratio test for $\sigma_1 = 0$ (*P*-value ≈ 0.635)


Port Jervis winter maximum temperature
Maximum temperature (oC)
Arctic Oscillation Index
24
22
20
18
16
14
12
10
-2
-1
0
1
2
3
Temperature
Median of GEV


Exponential Q-Q Plot: Port Jervis Maximum Temp.
Observed
Expected
0
1
2
3
4
5
0
1
2
3
4
5

- **Fort Collins precipitation example**

-- Threshold u = 0.395 in (could be time varying as well)

Length of year $T \approx 365.25$ days

-- Orthogonal approach

(i) Annual cycle in Poisson rate parameter

$$\log \lambda(t) = \lambda_0 + \lambda_1 \sin(2\pi t / T) + \lambda_2 \cos(2\pi t / T)$$

Parameter		Estimate	(Std. Error)
Rate:	λ_0	−3.721	
	λ_1	0.221	(0.045)
	λ_2	−0.846	(0.049)

Likelihood ratio test for $\lambda_1 = \lambda_2 = 0$ (P-value ≈ 0)

(ii) Annual cycle in scale parameter of GP distribution

$$\log \sigma^*(t) = \sigma_0^* + \sigma_1^* \sin(2\pi t / T) + \sigma_2^* \cos(2\pi t / T)$$

Parameter		Estimate	(Std. Error)
Scale:	σ_0^*	−1.238	
	σ_1^*	0.088	(0.048)
	σ_2^*	−0.303	(0.069)
Shape	ξ	0.181	

Likelihood ratio test for $\sigma_1^* = \sigma_2^* = 0$ (P-value < 10^{-5})

Q-Q plot: Transform non-stationary GP to exponential

$$\varepsilon_t = [1/\xi(t)] \log \{1 + \xi(t) [Y_t / \sigma^*(t)]\}$$

-- Point process approach (u = 0.395 in, h = 1/365.25)

Annual cycles in location & scale parameters of GEV distribution:

$$\mu(t) = \mu_0 + \mu_1 \sin(2\pi t / T) + \mu_2 \cos(2\pi t / T)$$

$$\log \sigma(t) = \sigma_0 + \sigma_1 \sin(2\pi t / T) + \sigma_2 \cos(2\pi t / T)$$

Parameter		Estimate	(Std. Error)	LR Test
Location:	μ_0	1.281		
	μ_1	−0.085	(0.031)	$\mu_1 = \mu_2 = 0$
	μ_2	−0.806	(0.043)	(P-value ≈ 0)
Scale:	σ_0	−0.847		
	σ_1	−0.123	(0.028)	$\sigma_1 = \sigma_2 = 0$
	σ_2	−0.602	(0.034)	(P-value ≈ 0)
Shape	ξ	0.182		

(4) Risk Communication (Under Non-Stationarity)

- **Interpretation of return level $x(p)$ (under stationarity)**

-- Stationarity implies identical distributions
(not necessarily independence)

(i) Expected waiting time (under temporal independence)
Waiting time W has geometric distribution:

$$\Pr\{W = k\} = (1 - p)^{k-1} p, \quad k = 1, 2, \ldots, \quad E(W) = 1/p$$

(ii) Length of time T_p for which expected number of events = 1

$$1 = \text{Expected no. events} = T_p\, p, \quad \text{so } T_p = 1/p$$

- **Options (under non-stationarity)**

-- Retain one of these two interpretations

Not clear which one is preferable:

Property (ii) is easier to work with (like average probability)

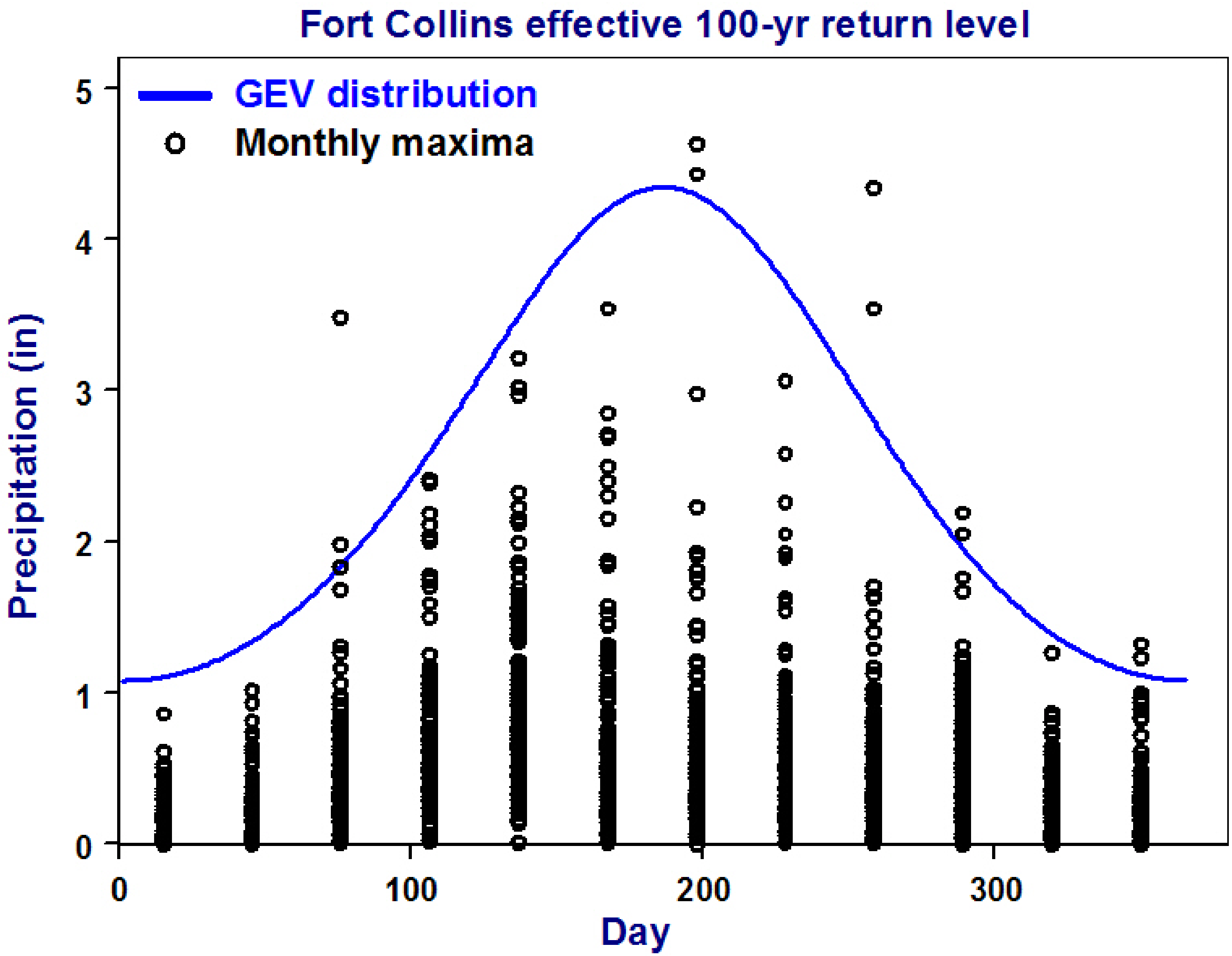
Property (i) may be more meaningful for risk analysis

-- Switch to “effective” return period and “effective” return level

(i. e., varying probabilities over time)

Fort Collins example:

Rescale parameters of GEV distribution for annual maxima to obtain those for monthly maxima



Fort Collins effective 100-yr return level
GEV distribution
Monthly maxima
Precipitation (in)
Day
0
1
2
3
4
5
0
100
200
300

-- Return level with annual cycles

Could still determine single annual return period 1/*p* or return value *x*(*p*) (assuming temporal independence):

$$1 - p = \Pr\{M_T \leq x(p)\} \approx \prod_t p_t \qquad (t = 1, 2, \ldots, T)$$

where p_t denotes probability do not exceed *x*(*p*) at time *t*
(Obtained from fitted point process, see Chapter 7 in Coles 2001)

Fort Collins example (July 1997 flood): Observed value of 4.63 in

Corresponding return period for point process model (with annual cycles in location and scale parameters of GEV dist.): 50.8 yr*

***562.3 yr if assume Gumbel distribution instead**

(5) Economic Impacts / Random Sums

- **Economic impacts from extreme weather events**

-- Both frequency & intensity vary from year to year

-- Damage total (e. g., annual) can be viewed as "random sum"

- **Hurricane damage data**

-- Already analyzed upper tail of damage from individual storms using GP distribution

-- Now consider rate of storms, covariates (e. g., ENSO phenomenon, trends)

- **Statistics of random sums**

-- Notation

$N(t)$ number of events in tth yr

X_k damage from kth event in tth yr, $k = 1, 2, \ldots, N(t)$

X_k's independent and identically distributed with common distribution function F

X_k's independent of $N(t)$

Total damage in tth yr (random sum):

$$S(t) = X_1 + X_2 + \cdots + X_{N(t)}, \quad N(t) \geq 1$$

-- Mean of total annual damage [assume $E(X_k) < \infty$]

$$E[S(t)] = \mathrm{E}[N(t)]\, E(X_k)$$

-- Variance of total annual damage [assume $\mathrm{Var}(X_k) < \infty$]

$$\mathrm{Var}[S(t)] = \mathrm{E}[N(t)]\, \mathrm{Var}(X_k) + \mathrm{Var}[N(t)]\, [E(X_k)]^2$$

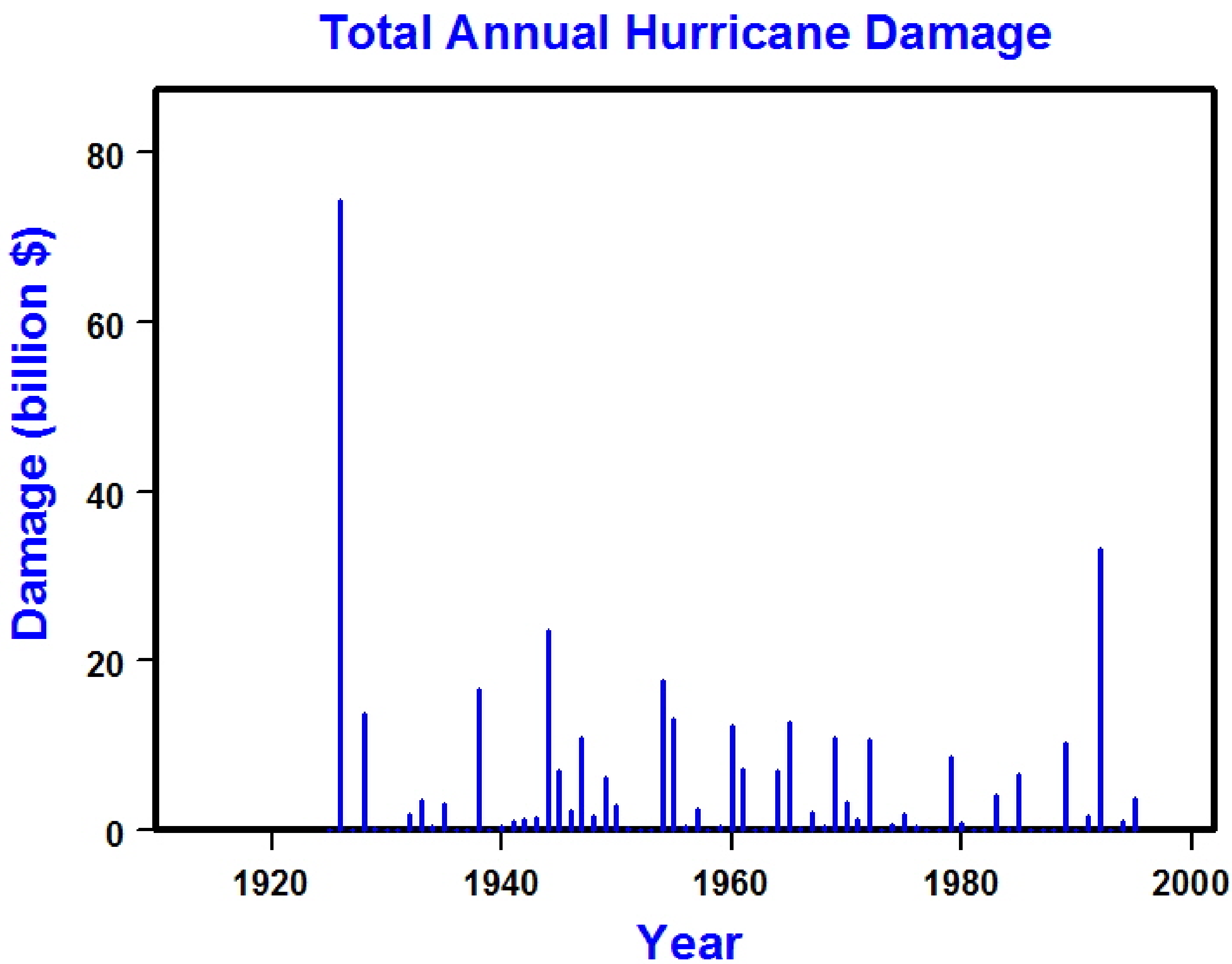
-- Maximum damage from individual events in tth yr, $M_{N(t)}$

(i. e., maximum of *random* number of random variables)

Assume N(t) has exact Poisson distribution (rate parameter λ):

$$\Pr\{M_{N(t)} \le x\} = \exp\{ -\lambda[1 - F(x)]\}$$

(exact result, not just asymptotic)


Total Annual Hurricane Damage
Damage (billion $)
80
60
40
20
0
1920
1940
1960
1980
2000
Year

(i) Occurrence component $N(t)$

-- Assume $N(t)$ has Poisson distribution with rate parameter $\lambda(t)$

-- Trend

$$\log \lambda(t) = \lambda_0 + \lambda_1 t, \quad t = 1, 2, \ldots$$

Estimate of $\lambda_1 \approx 0.0030$ (std. error ≈ 0.0043)

Likelihood ratio test for $\lambda_1 = 0$ (P-value ≈ 0.492)

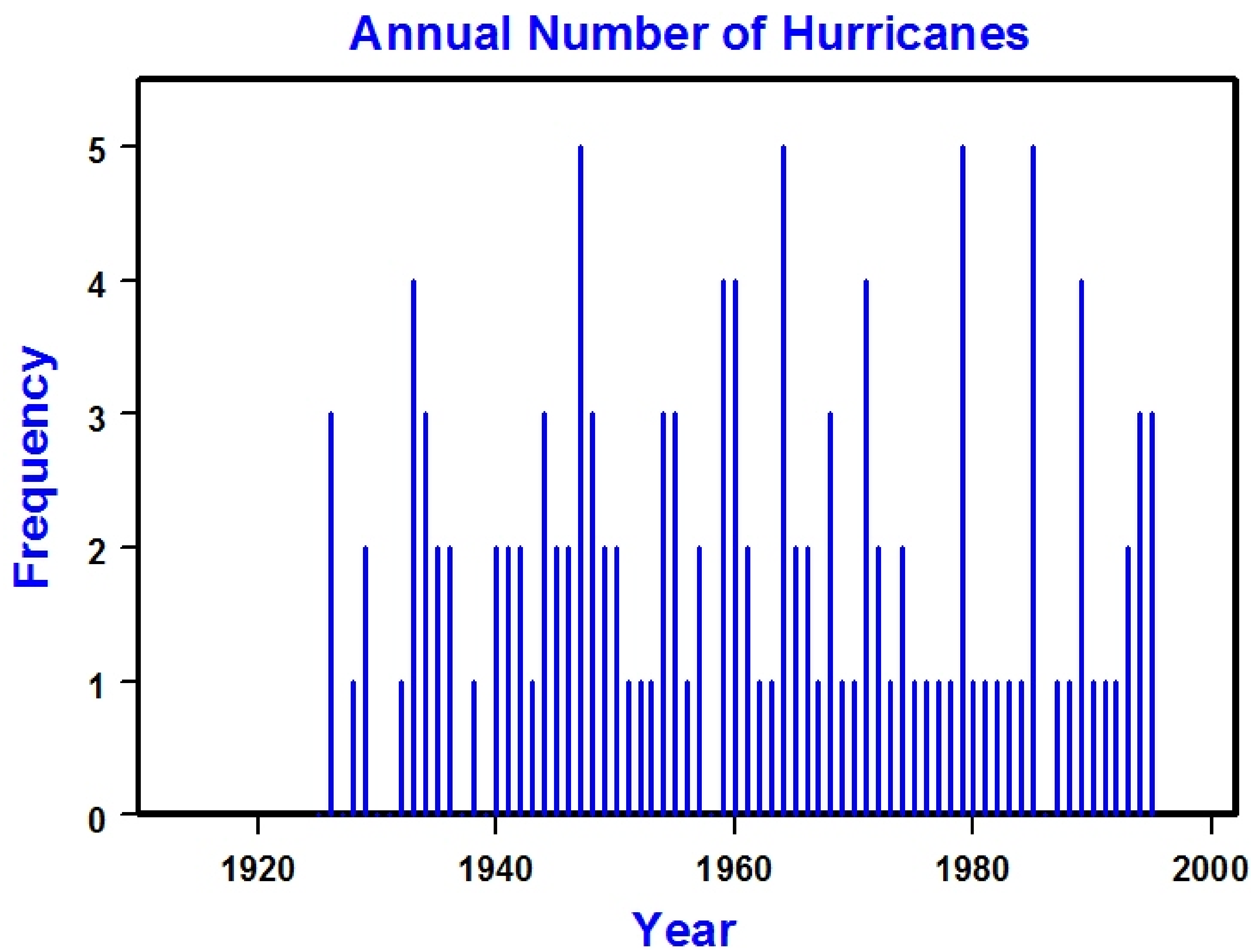
-- ENSO phenomenon as covariate

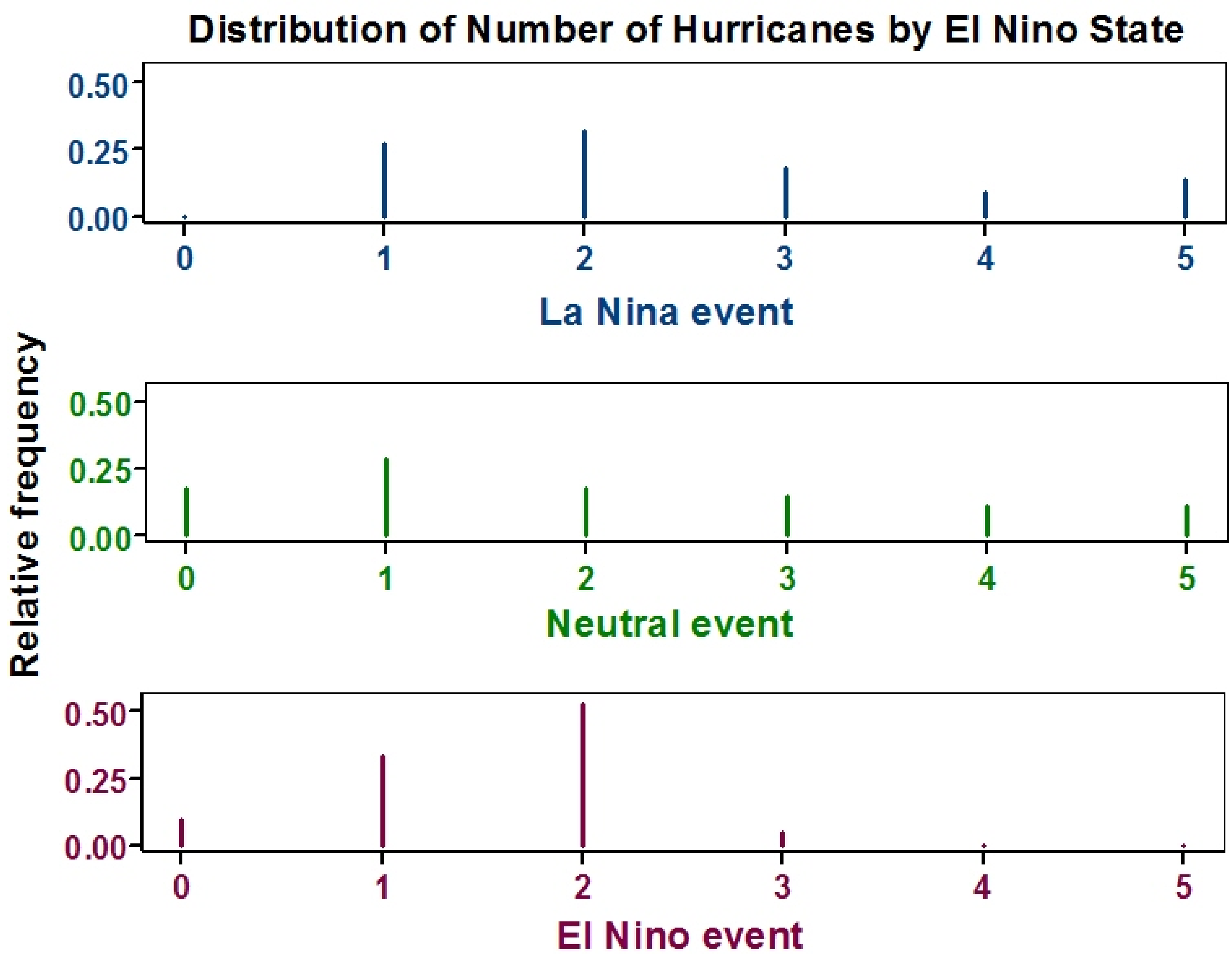
$$\log \lambda(z) = \lambda_0 + \lambda_1 z$$

where $z = -1, 0, 1$ indicate La Niña, neutral, or El Niño event

Estimate of $\lambda_1 \approx -0.248$ (std. error ≈ 0.115)

Likelihood ratio test for $\lambda_1 = 0$ (P-value ≈ 0.029)


Annual Number of Hurricanes
Frequency
0
1
2
3
4
5
1920
1940
1960
1980
2000
Year


Distribution of Number of Hurricanes by El Nino State
Relative frequency
0.50
0.25
0.00
0
1
2
3
4
5
La Nina event
Neutral event
El Nino event

(ii) Damage component X_k

-- Model overall distribution *F* as lognormal

No trend in mean of log-transformed damage

Dependence of mean of log-transformed damage on ENSO state (lower mean log-transformed damage during El Niño event)

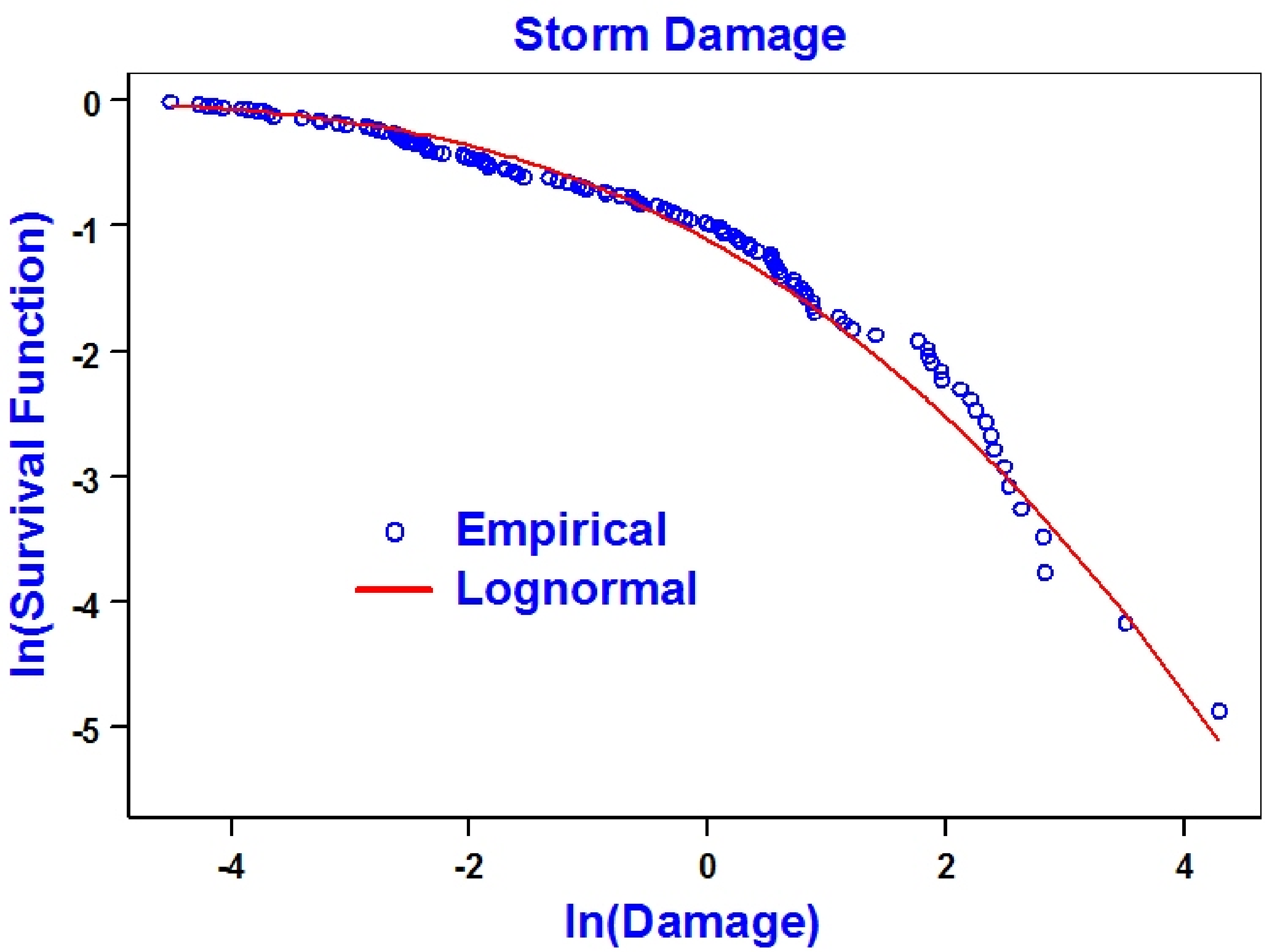
-- GP distribution for upper tail (*Y* excess in damage over $6 billion)

Dependence on ENSO state $Z = z$:

$$\log \sigma^*(z) = \sigma_0^* + \sigma_1^* z$$

Estimate of $\sigma_1^* \approx 0.048$

Likelihood ratio test for $\sigma_1^* = 0$: *P*-value ≈ 0.275



Storm Damage
ln(Survival Function)
0
-1
-2
-3
-4
-5
-4
-2
0
2
4
ln(Damage)
Empirical
Lognormal

(6) Extreme Weather Spells

- **Definition of heat wave / hot spell**

-- Recall "runs de-clustering" algorithm

-- More complex definition (multiple thresholds)

- **Lack of use of extreme value theory in modeling weather spells**

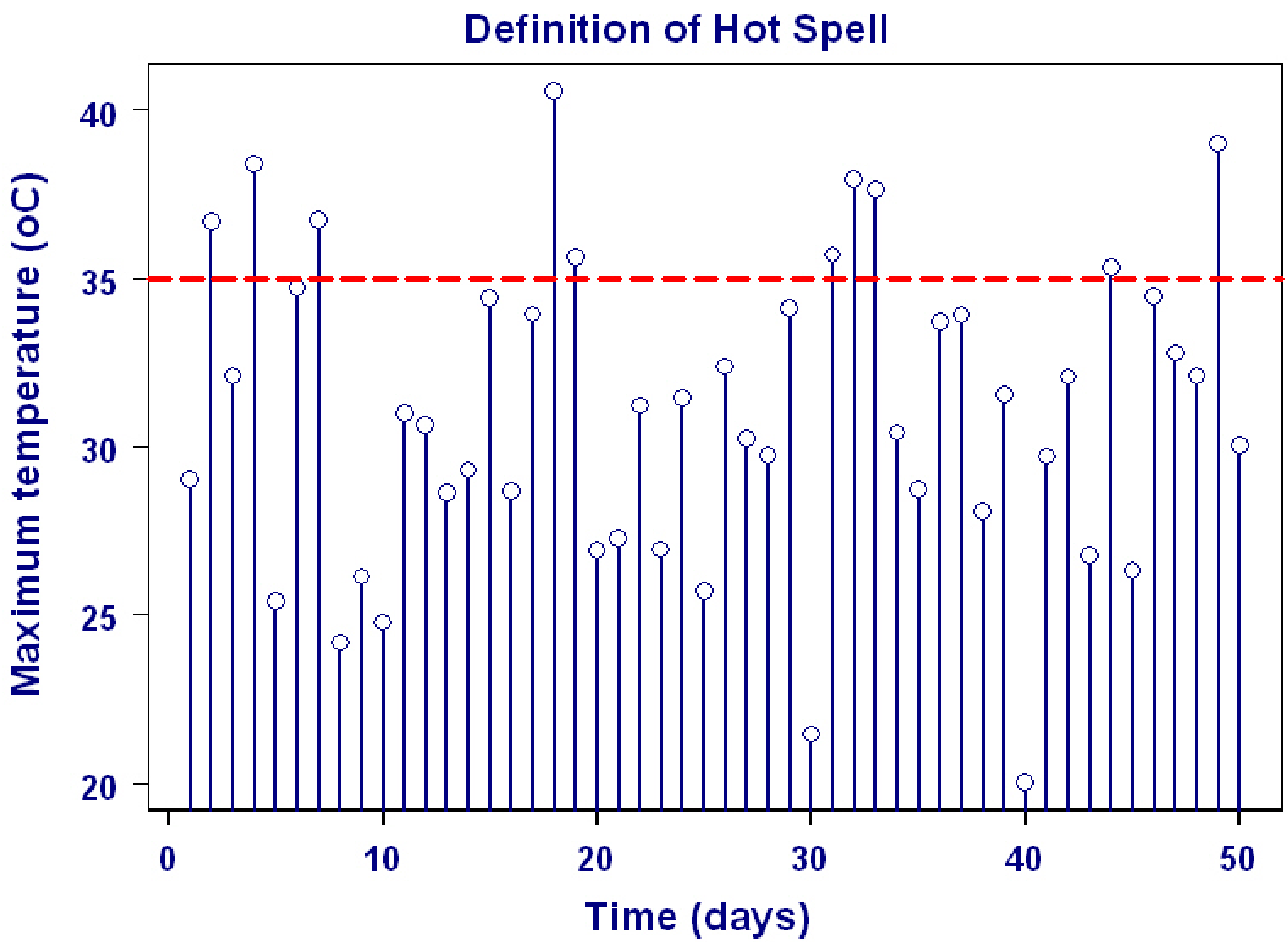
-- Extremal dependence at high levels

- **Cluster statistics**

-- Duration

-- Maxima

-- Other measures of intensity


Definition of Hot Spell
Maximum temperature (oC)
40
35
30
25
20
0
10
20
30
40
50
Time (days)

- **Covariate approach**

-- Advantage

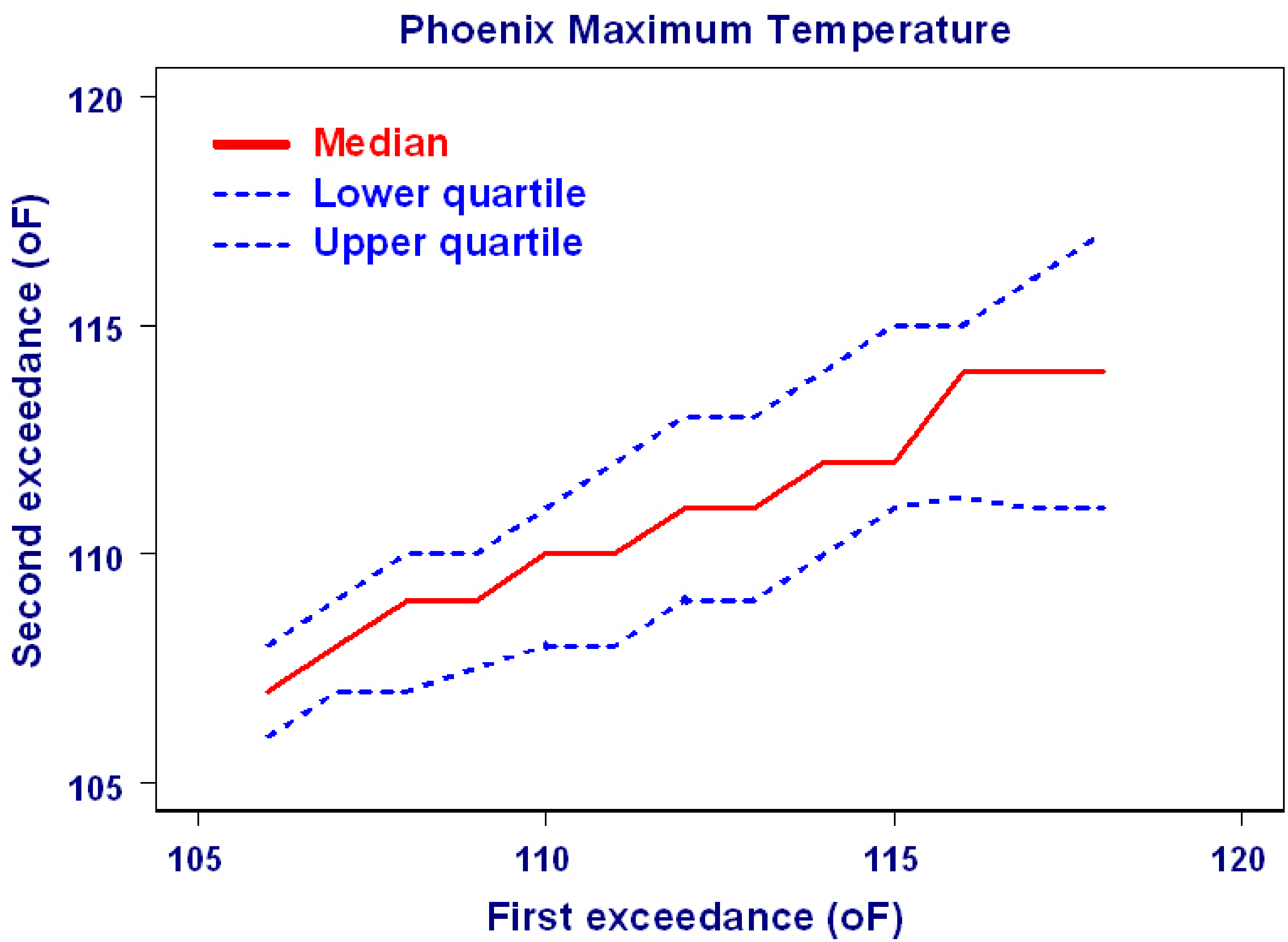
Requires only univariate extreme value theory (not multivariate)

-- Let Y_1, Y_2, . . ., Y_k denote excesses over threshold within given cluster / spell

Model conditional distribution of Y_2 given Y_1 as GP distribution with scale parameter depending on Y_1: e. g.,

$$\sigma^*(y) = \sigma_0^* + \sigma_1^*\, y, \text{ given } Y_1 = y$$

Similar model for conditional distribution of Y_3 given Y_2 (etc.)


Phoenix Maximum Temperature
Median
Lower quartile
Upper quartile
120
115
110
105
Second exceedance (oF)
105
110
115
120
First exceedance (oF)

(7) Origin of Bounded and Heavy Tails

- **Upper Bounds / Penultimate approximation**

-- Weibull type of GEV (i. e., $\xi < 0$) provides better approximation than asymptotic Gumbel when parent distribution is (e. g.) normal (with $\xi \uparrow 0$ as block size $T \to \infty$)

-- "Thermostat hypothesis"

Upper bound on temperature in tropical oceans

Implications for impact of global warming on coral reefs

-- Intensification of hurricanes

Upper bound on maximum wind speed (Trend?)

- **Heavy tails / Penultimate approximation**

-- "Stretched exponential" distribution (traditional form of Weibull)

$$\Pr\{Y > y\} = \exp(-y^c), \quad y > 0, \quad c > 0$$

with shape parameter c (for simplicity, unit scale parameter)

(i) $c < 1$

Fréchet type of GEV (i. e., $\xi > 0$) better approximation than asymptotic Gumbel (explanation for heavy tail of precipitation?)

(ii) $c > 1$

Weibull type of GEV (i. e., $\xi < 0$) better approximation than asymptotic Gumbel (explanation for bounded upper tail of wind speed?)

-- Simulation experiment

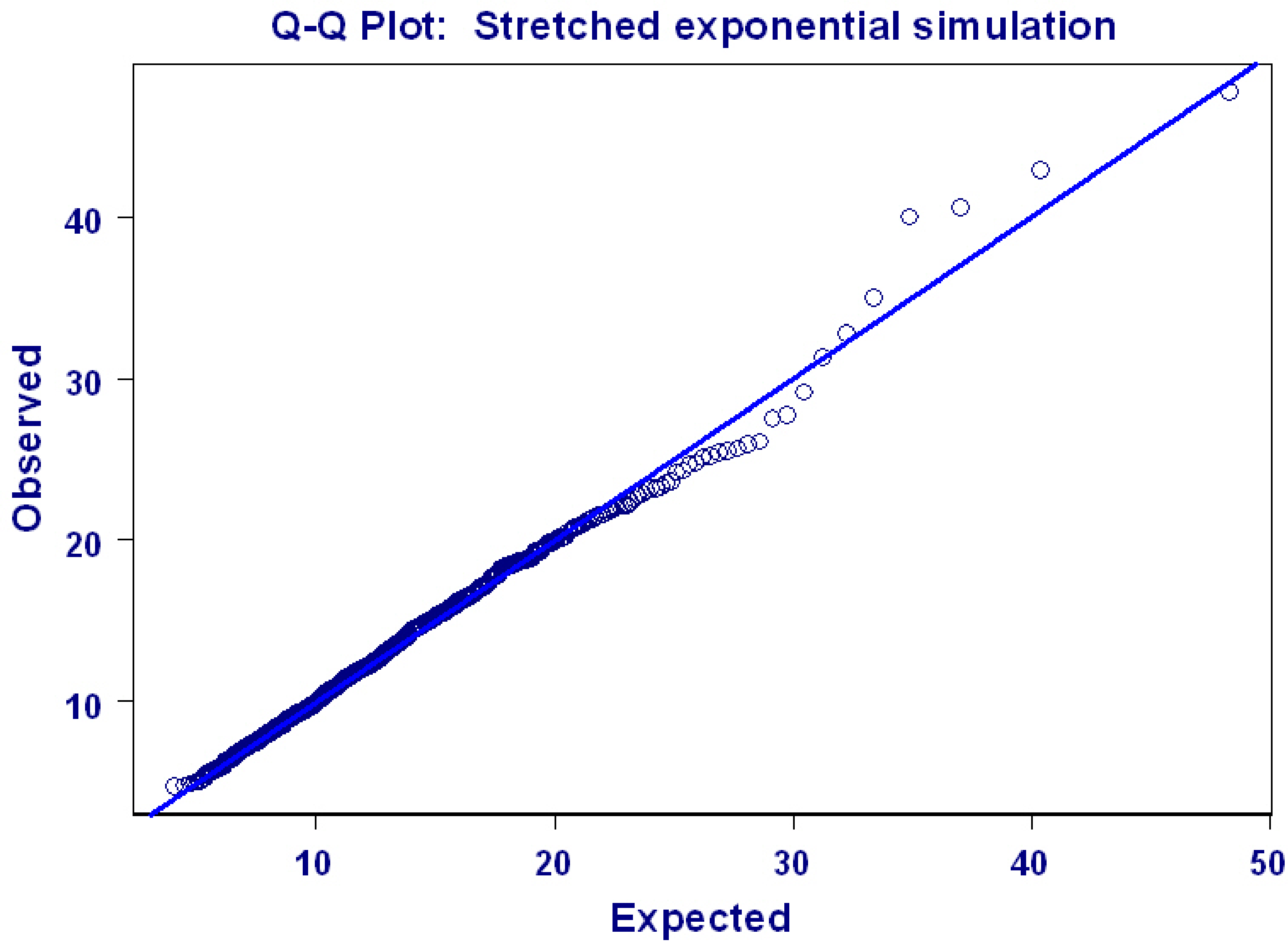
Generated observations with stretched exponential distribution (shape parameter c = 2/3: physical argument, Wilson & Toumi 2005)

Determine maximum of sequence of length T = 100, M_{100} (daily precipitation occurrence rate about 27%)

Should produce GEV shape parameter of just over 0.1 (Heuristic argument gives $\xi \approx 0.11$ for this block size)

Fitted GEV distribution (sample size = 1000):

Estimate of $\xi \approx 0.12$


Q-Q Plot: Stretched exponential simulation
Observed
40
30
20
10
10
20
30
40
50
Expected

- **Heavy Tails / Chance mechanism**

-- Mixture of exponential distributions

Suppose Y has exponential distribution with scale parameter σ^*:

$$\Pr\{Y > y \mid \sigma^*\} = \exp[\, -(y/\sigma^*)]$$

Further assume that the rate parameter $v = 1/\sigma^*$ varies according to a gamma distribution with shape parameter α (unit scale):

$$f_V(v;\, \alpha) = [\Gamma(\alpha)]^{-1}\, v^{\alpha-1} \exp(-v), \quad \alpha > 0$$

The unconditional distribution of Y is heavy-tailed:

$$\Pr\{Y > y\} = (1 + y)^{-\alpha}$$

(i.e., exact GP distribution with shape parameter $\xi = 1/\alpha$)

-- Simulation experiment

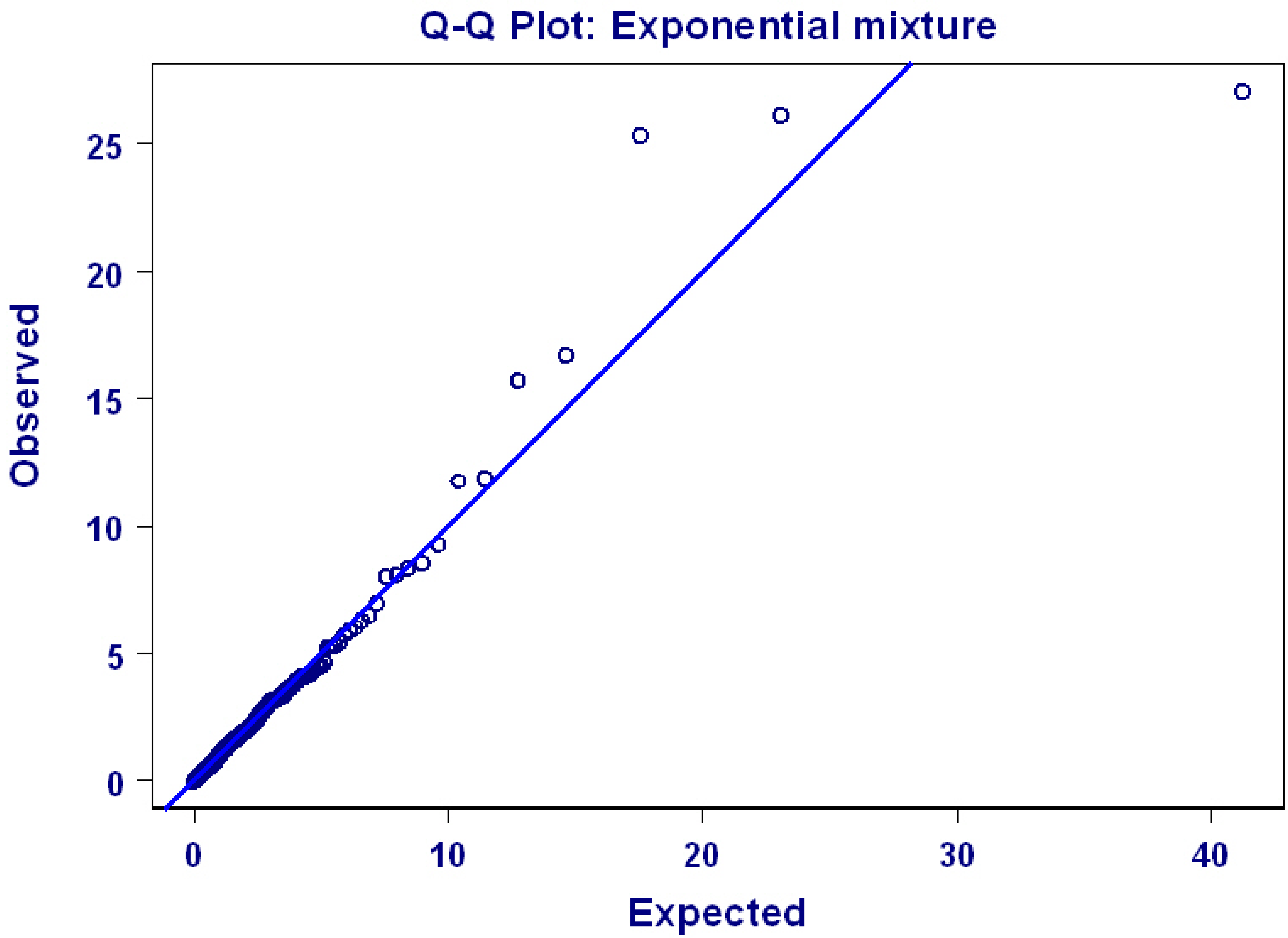
Induce heavy tail from conditional light tails

Let rate parameter of exponential distribution have gamma distribution with shape parameter $\alpha = 2$

Then mixture distribution is GP with shape parameter $\xi = 0.5$

Fit GP distribution to simulated exponential mixture (sample size = 1000):

Estimate of $\xi \approx 0.51$


Q-Q Plot: Exponential mixture
Observed
Expected
0
5
10
15
20
25
0
10
20
30
40

(8) Multivariate Extremes

- **Extreme value theory**

-- Well developed, but complex

- **Bivariate normal distribution (correlation coefficient ρ)**

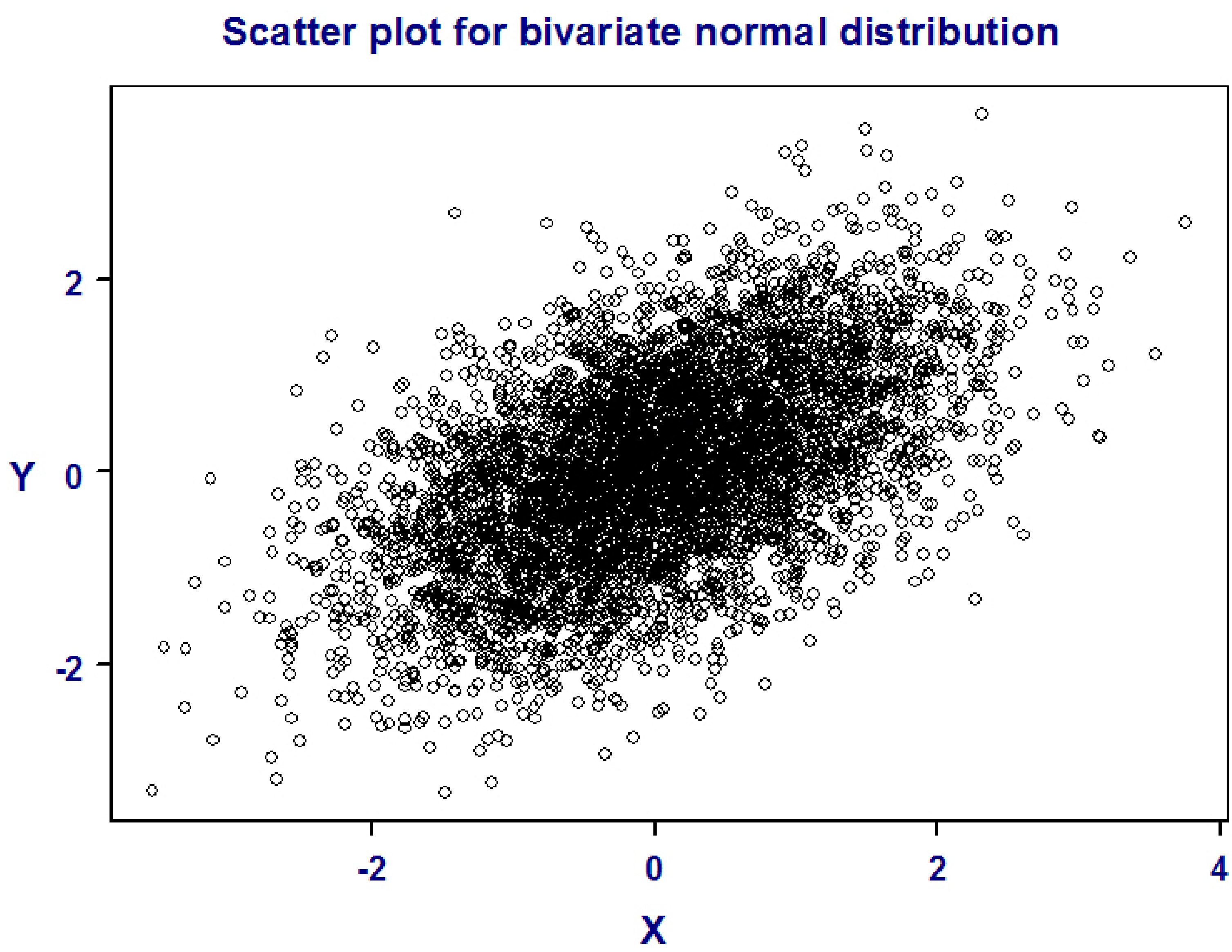
-- Joint distribution of two random variables (X, Y)

-- Clustering in overall sense for $\rho > 0$

(but no clustering at high levels in asymptotic sense)

$$\Pr\{Y > u \mid X > u\} \to 0 \text{ as } u \to \infty$$

[Recall $\Pr\{X > x\} \approx \varphi(x) / x$ if X has standard normal distribution]


Scatter plot for bivariate normal distribution
2
Y 0
-2
-2
0
2
4
X

- **Try univariate extreme value theory with covariates instead**

-- Example (Fruit-frost problem Yakima, WA, USA)

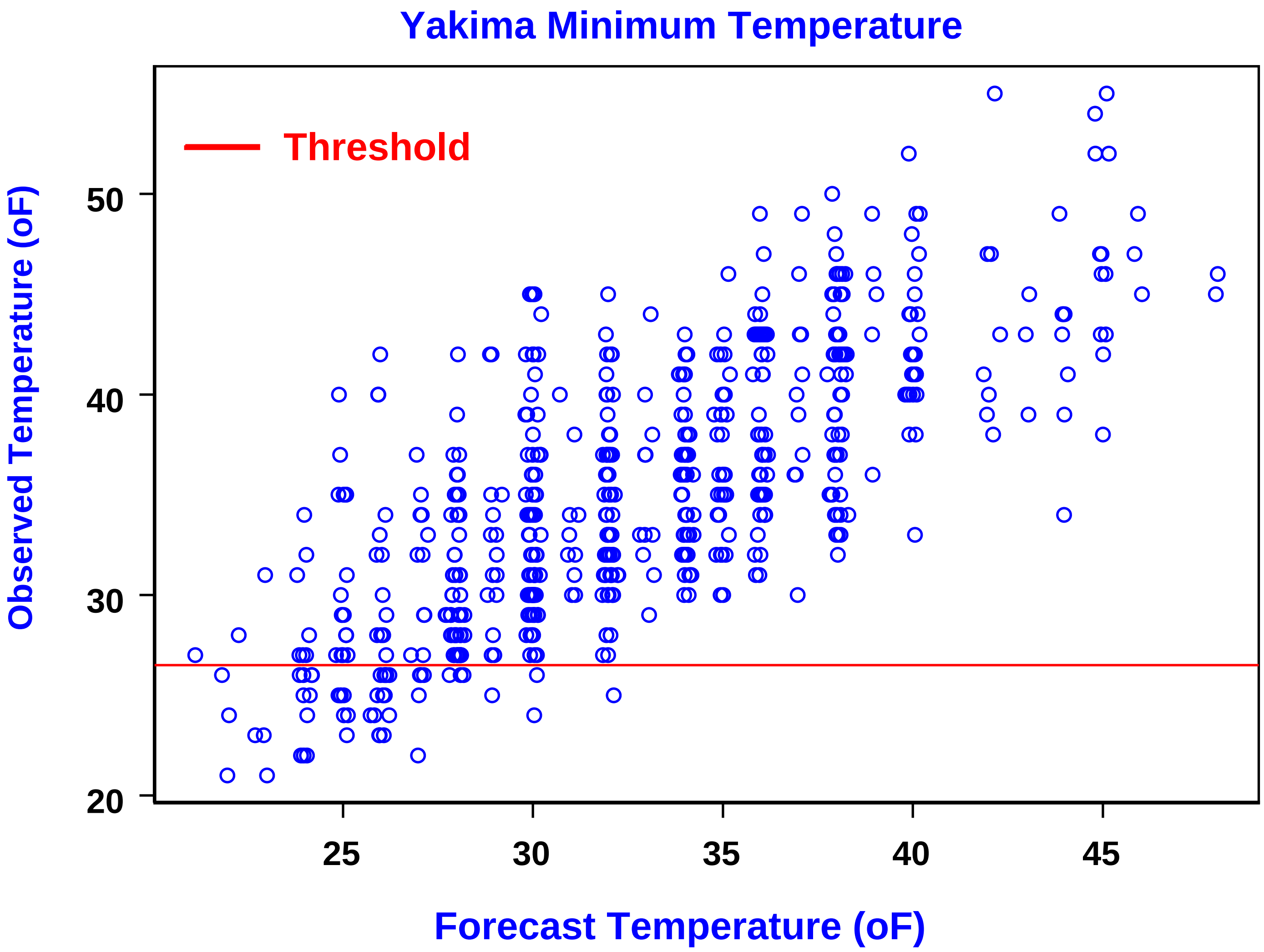
Conditional distribution of observed daily minimum temperature X (for $X < u$ = 26.5 ºF) given forecast Z

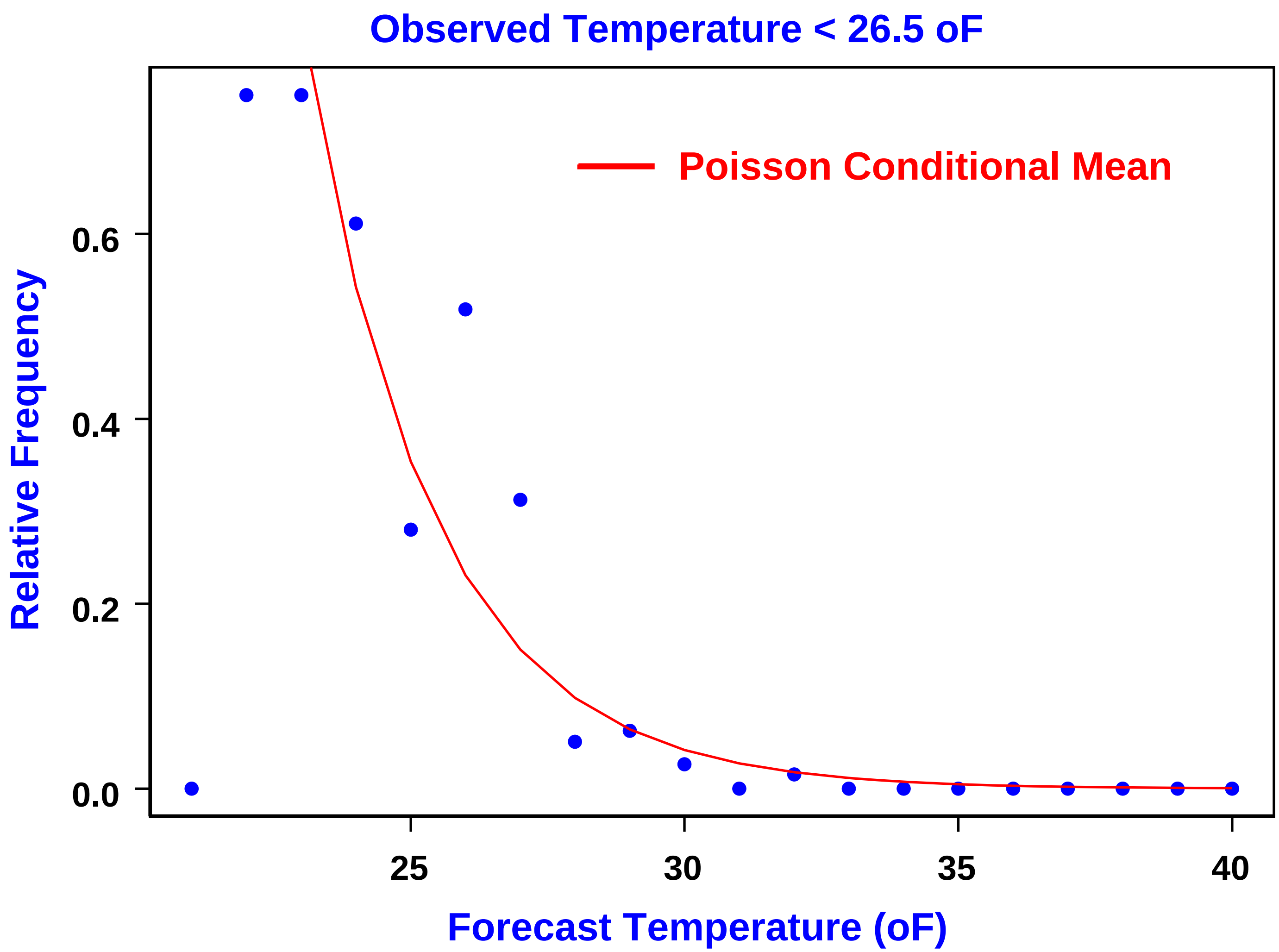
(i) Occurrence of $X < u$ given forecast $Z = z$

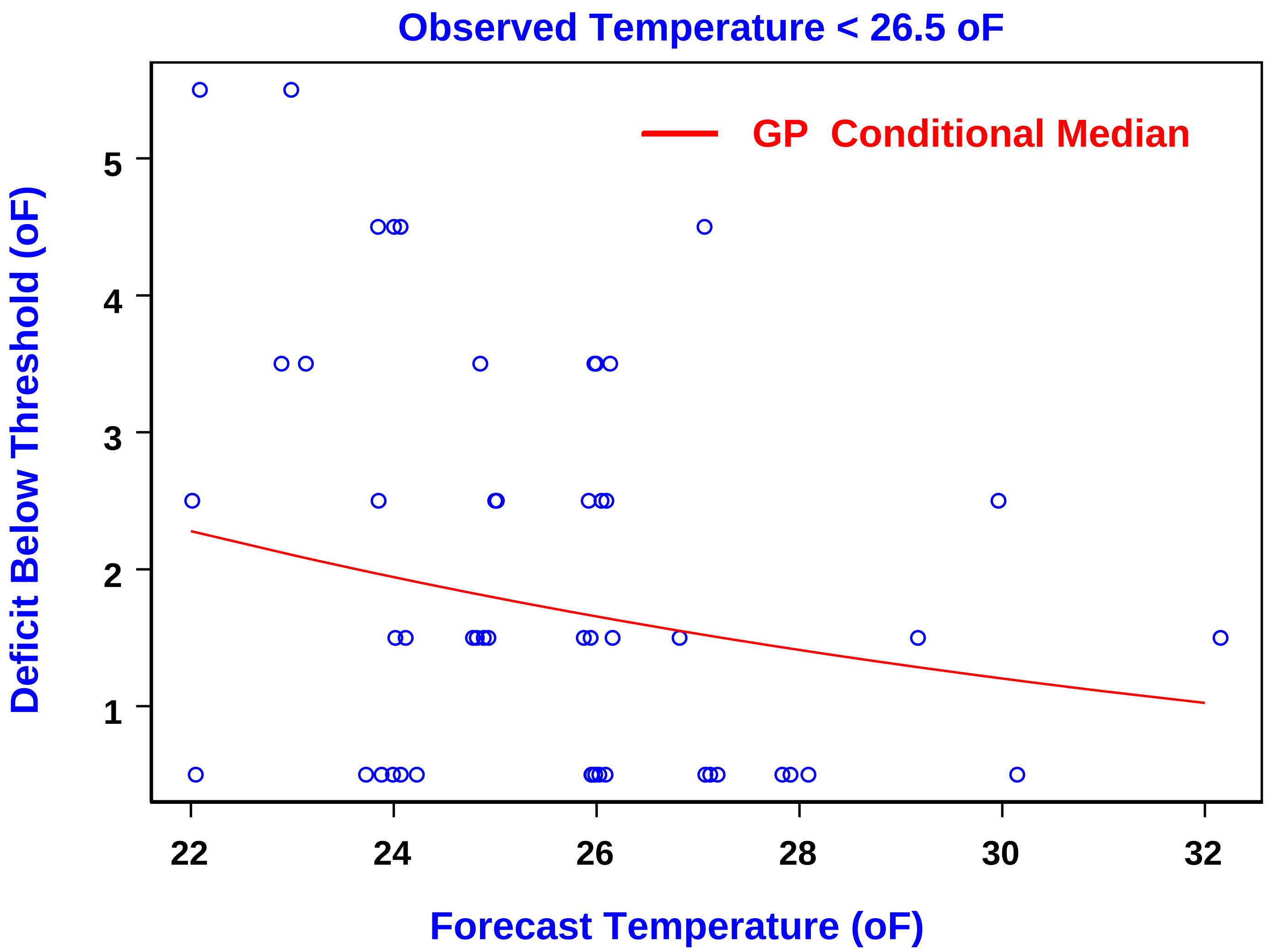
Poisson distribution with rate: $\log \lambda(z) = \lambda_0 + \lambda_1 z$

(ii) Deficit $Y = u - X$ given forecast $Z = z$

GP distribution with scale: $\log \sigma^*(z) = \sigma_0^* + \sigma_1^* z$


Yakima Minimum Temperature
Threshold
Observed Temperature (oF)
Forecast Temperature (oF)
20
30
40
50
25
30
35
40
45


Observed Temperature < 26.5 oF
Poisson Conditional Mean
Relative Frequency
0.0
0.2
0.4
0.6
25
30
35
40
Forecast Temperature (oF)



Observed Temperature < 26.5 oF
GP Conditional Median
Deficit Below Threshold (oF)
1
2
3
4
5
22
24
26
28
30
32
Forecast Temperature (oF)

(9) Spatial Extremes

- **Approaches**

-- Extensions of extreme value theory

Max stable processes (as in univariate case)

Beautiful theory (e. g., model for how parameters depend on spatial support), but lack of realistic applications

-- Spatial smoothing (Soon to be added to `extRemes`?)

First model each site separately

Then smooth parameter estimates or return levels

-- Bayesian hierarchical modeling / Markov Chain Monte Carlo

Attempts, but still under development

(e. g., conditional independence does not induce enough unconditional spatial dependence)

-- Detection of trends in extremes

"Borrow strength"

Akin to "regional analysis" in hydrology for flood estimation

Assume common slope of trend across region (also common shape parameter?)

How to account for spatial dependence?

-- Crude approaches

Assume spatial independence for fitting:

But estimate standard errors by resampling

(to take into account spatial dependence)

Adjust tests of field / local significance:

Use "false discovery rate"

(more powerful than Bonferroni procedure)

Need to adjust for effect of spatial dependence?