FORECAST VERIFICATION OF EXTREMES:

USE OF EXTREME VALUE THEORY

Rick Katz

Institute for Study of Society and Environment
National Center for Atmospheric Research
Boulder, CO USA

Email: rwk@ucar.edu

Web site: www.isse.ucar.edu/HP_rick.html

Talk: www.isse.ucar.edu/HP_rick/pdf/atl06.pdf

QUOTE ON EXTREMES

- **Emil Gumbel (1891 – 1966)**

-- Pioneer in application of statistics of extremes

(Germany, France, U. S.)

"*Il est impossible que l'improbable n'arrive jamais*."

OUTLINE

(1) Motivation

(2) Conventional Methods

(3) Extreme Value Theory (EVT)

(4) Application of EVT to Verification

(5) Frost Forecasting Example

(6) Remaining Issues

(7) Resources

(1) MOTIVATION

IMPORTANCE OF EXTREMES

- **Societal Impacts of Extreme Events**
- **Skill in Predicting Extremes?**

VERIFICATION METHODS

- **Not Designed for Extremes Per Se**
- **Lack of Use of Statistical Theory of Extreme Values**

KEY REFERENCE

- **David Stephenson (2004)**

-- "***Verification of Rare Extreme Events***"

International Verification Methods Workshop

`www.bom.gov.au/bmrc/wefor/staff/eee/verif/Workshop2004/presentations/2.1_Stephenson.pdf`

-- **Inadequacy of existing verification measures**

-- **Proposes diagnostic measure (taken from bivariate EVT)**

(2) CONVENTIONAL METHODS

THRESHOLDING

- Dependence on Choice of Threshold

-- Ignore intensity of event ("excess" over threshold)

DISTRIBUTIONS-ORIENTED VERIFICATION

- Joint Distribution of Forecasts & Observations

-- Binning (sparse entries for rare/extreme events)

-- Parametric distributions (tail behavior?)

(3) EXTREME VALUE THEORY

THRESHOLD EXCEEDANCES

- **Poisson Distribution**

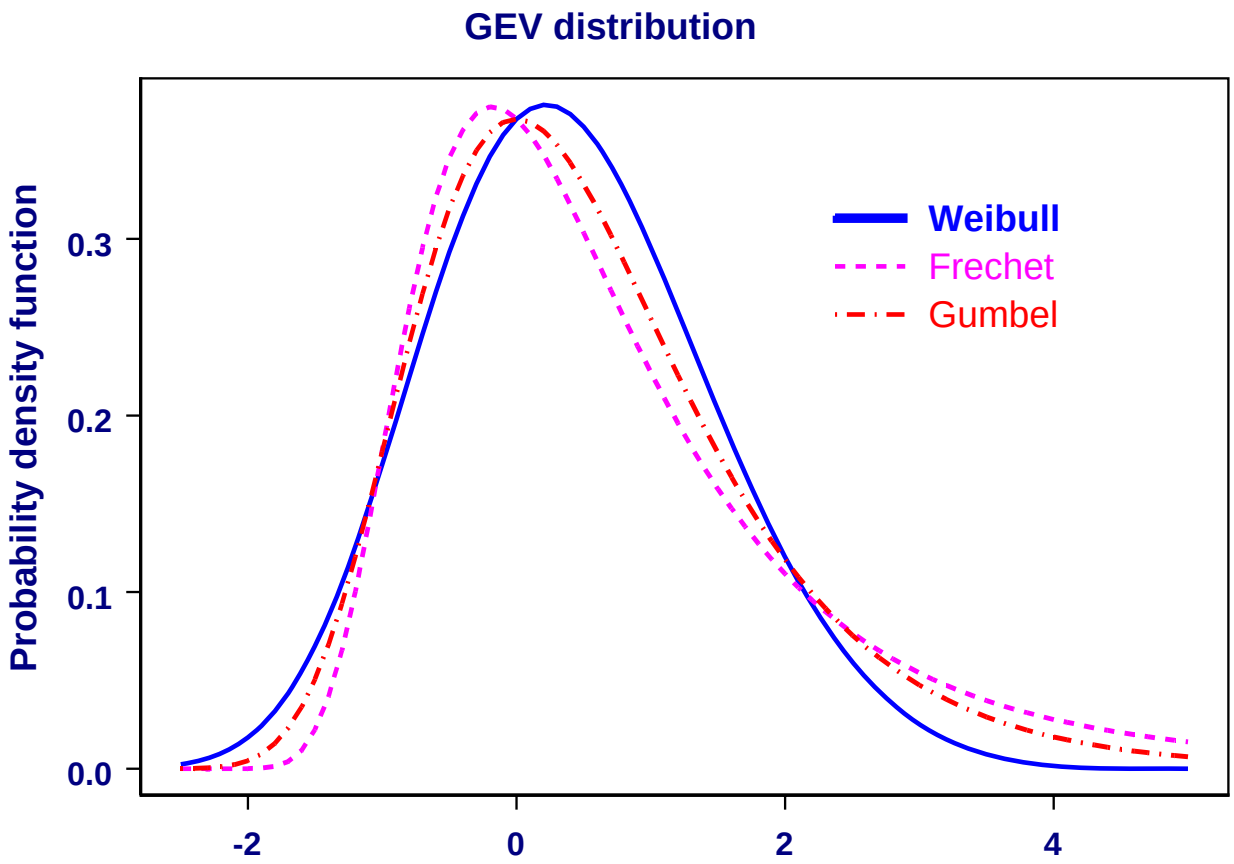
-- **Approximation for frequency of occurrence of rare events (*N* say)**

$$\Pr\{N = n\} = \lambda^n e^{-\lambda} / n! \quad n = 0, 1, 2, \ldots$$

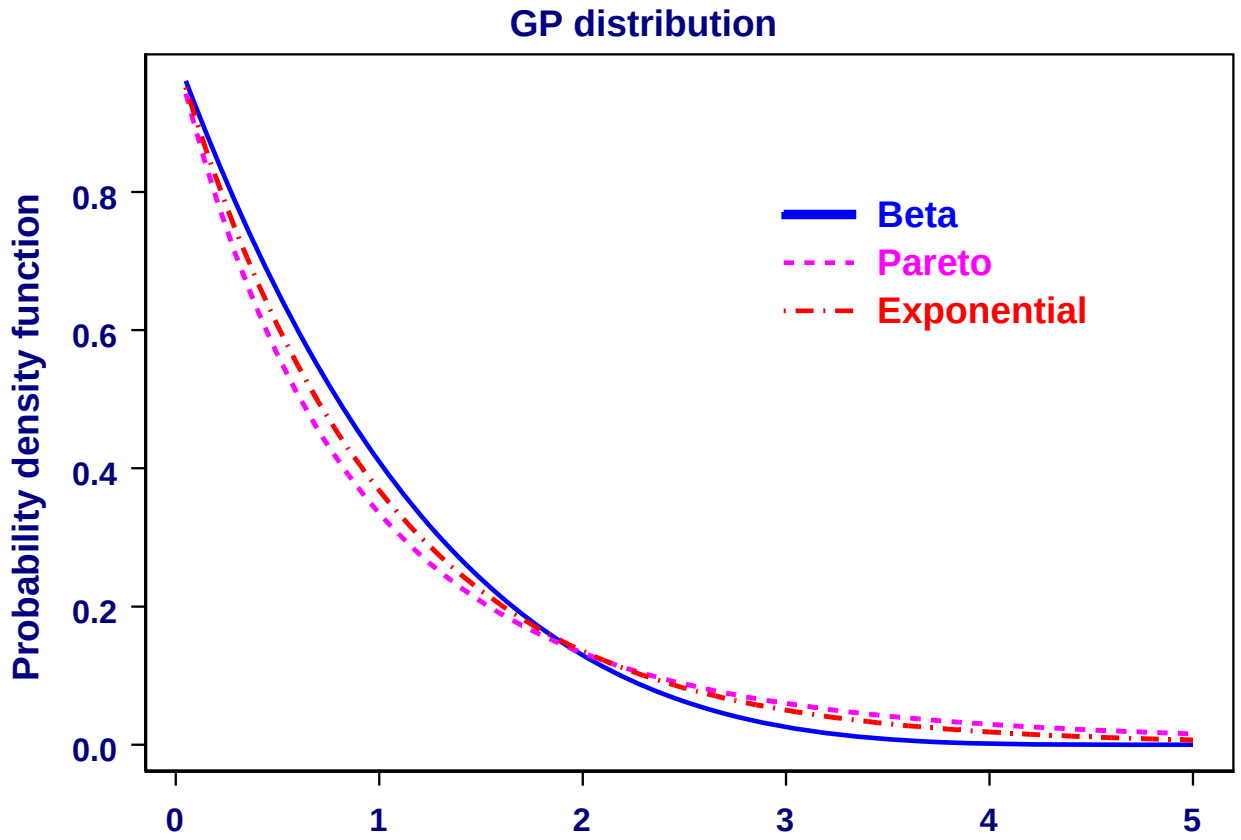
$\lambda > 0$ **rate parameter**

(recall Poisson approximation to Binomial Distribution)

GENERALIZED EXTREME VALUE DISTRIBUTION: MAXIMA



GENERALIZED PARETO DISTRIBUTION: EXCESS OVER THRESHOLD



(4) APPLICATION OF EVT TO VERIFICATION

- Y Observed Weather
- $\hat{Y}$ Forecast Weather

CALIBRATION–REFINEMENT FACTORIZATION

- Factor Joint Distribution of Y & $\hat{Y}$

$$f(y, \hat{y}) = f(y \mid \hat{y})\, f(\hat{y})$$

- Model Conditional Distribution of Y given $\hat{Y}$ using EVT (for y large)

- **Threshold Exceedances**

-- $Y > u$ extreme high threshold (u *fixed*)

$$\Pr\{Y > u \mid \hat{Y} = \hat{y}\}$$

Model as Poisson distribution with rate parameter

λ depending on $\hat{y}$:

$$\ln \lambda = \lambda_0 + \lambda_1 \hat{y}$$

$\lambda_1 \neq 0$ corresponds to *skill* relative to climatology

-- Estimate parameters using maximum likelihood (MLE)

-- Test for $\lambda_1 = 0$ (i. e., *no skill*) using likelihood ratio test (LRT)

- **Excess Over Threshold**

-- $Y - u$ given $Y > u$

$$\Pr\{Y > y \mid \hat{Y} = \hat{y}\}$$

GP distribution with scale parameter $\sigma > 0$ depending on $\hat{y}$:

$$\ln \sigma = \sigma_0 + \sigma_1 \hat{y}$$

$\sigma_1 \neq 0$ corresponds to *skill* relative to climatology

-- Estimate parameters using maximum likelihood (MLE)

-- Test for $\sigma_1 = 0$ (i. e., *no skill*) using Likelihood Ratio Test (LRT)

(5) FROST FORECASTING EXAMPLE

FRUIT–FROST PROBLEM

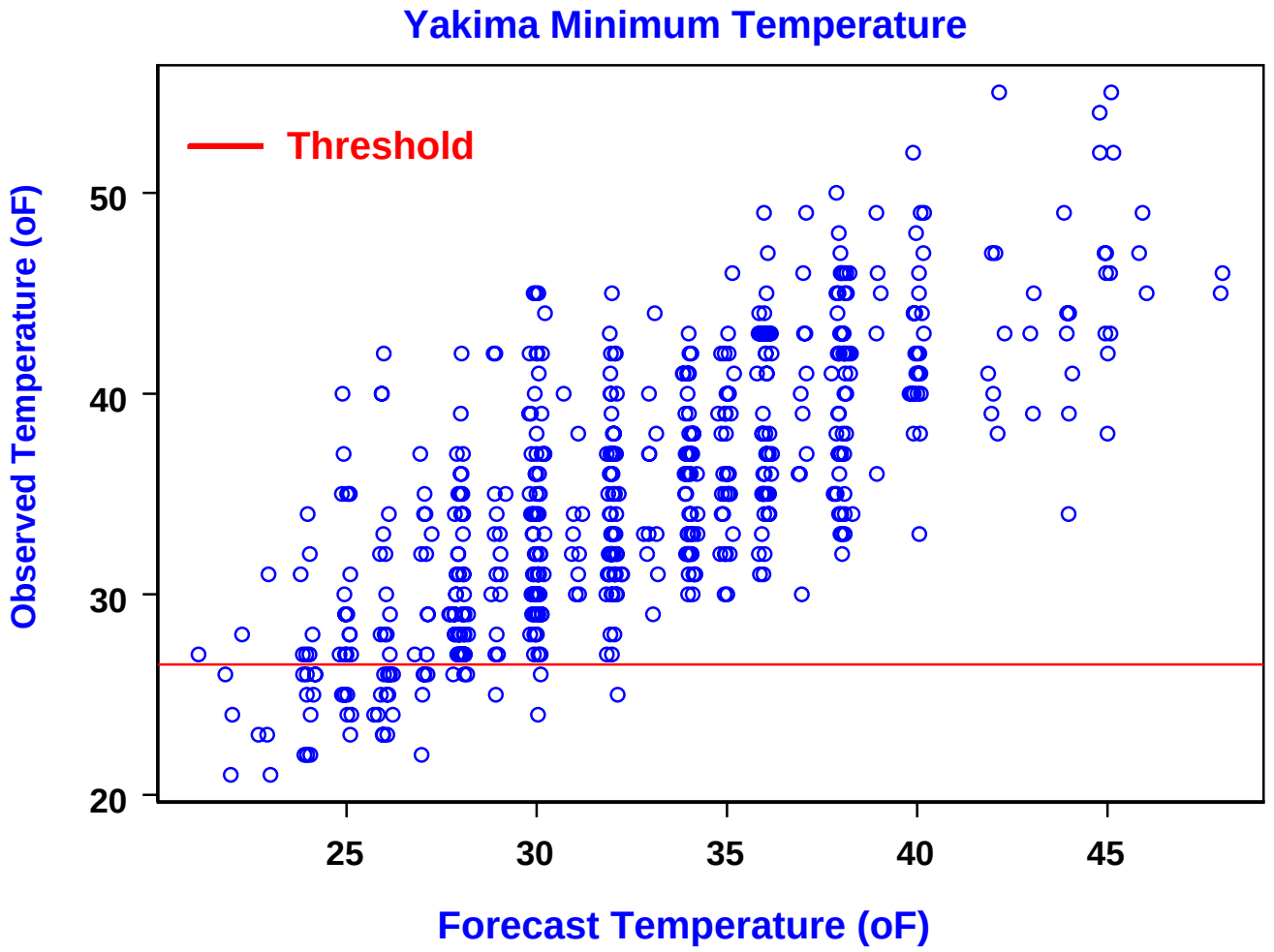
- **Protection of Fruit Buds from Freezing Temperatures**

-- **Katz et al. (*JAM*, 1982)**

- **Specialized Minimum Temperature Forecasts**

-- **Yakima, WA Case Study**

Data: April, 1957 – 1976 (20 yrs.)


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

RESULTS

- **Technical Issue**

-- **Essentially same EVT applies to low extremes (by reflection)**

- **Threshold "Exceedances"**

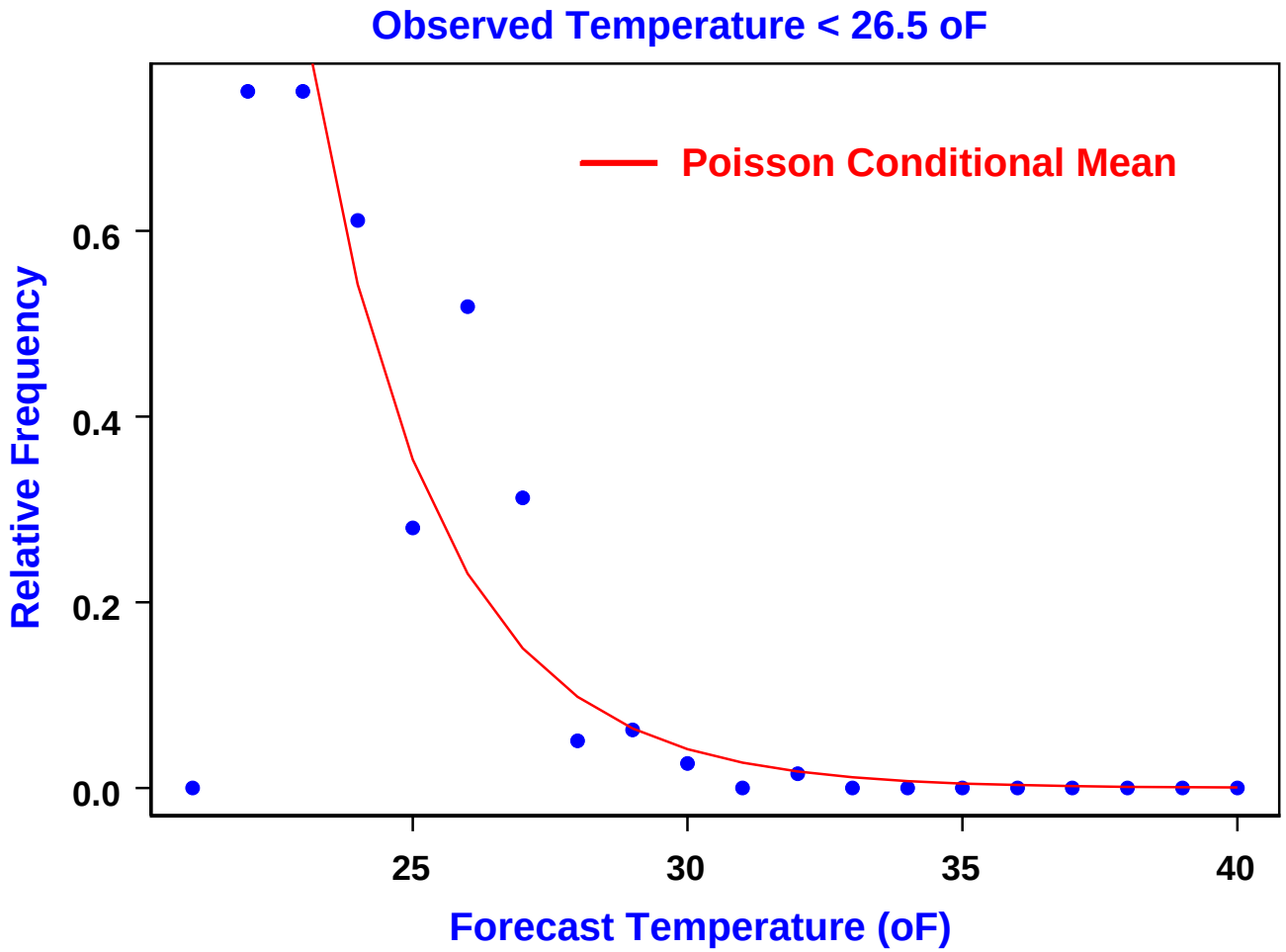
-- **Event of form:** $Y < u$

-- **Used threshold:** $u = 26.5\ °F$

-- **Forecast (as covariate)**

Poisson dist.: λ_1 **MLE of** $\lambda_1 = -0.427$

Likelihood-Ratio Test for $\lambda_1 = 0$**: P-value ≈ 0**


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)

- **Deficit Below Threshold**

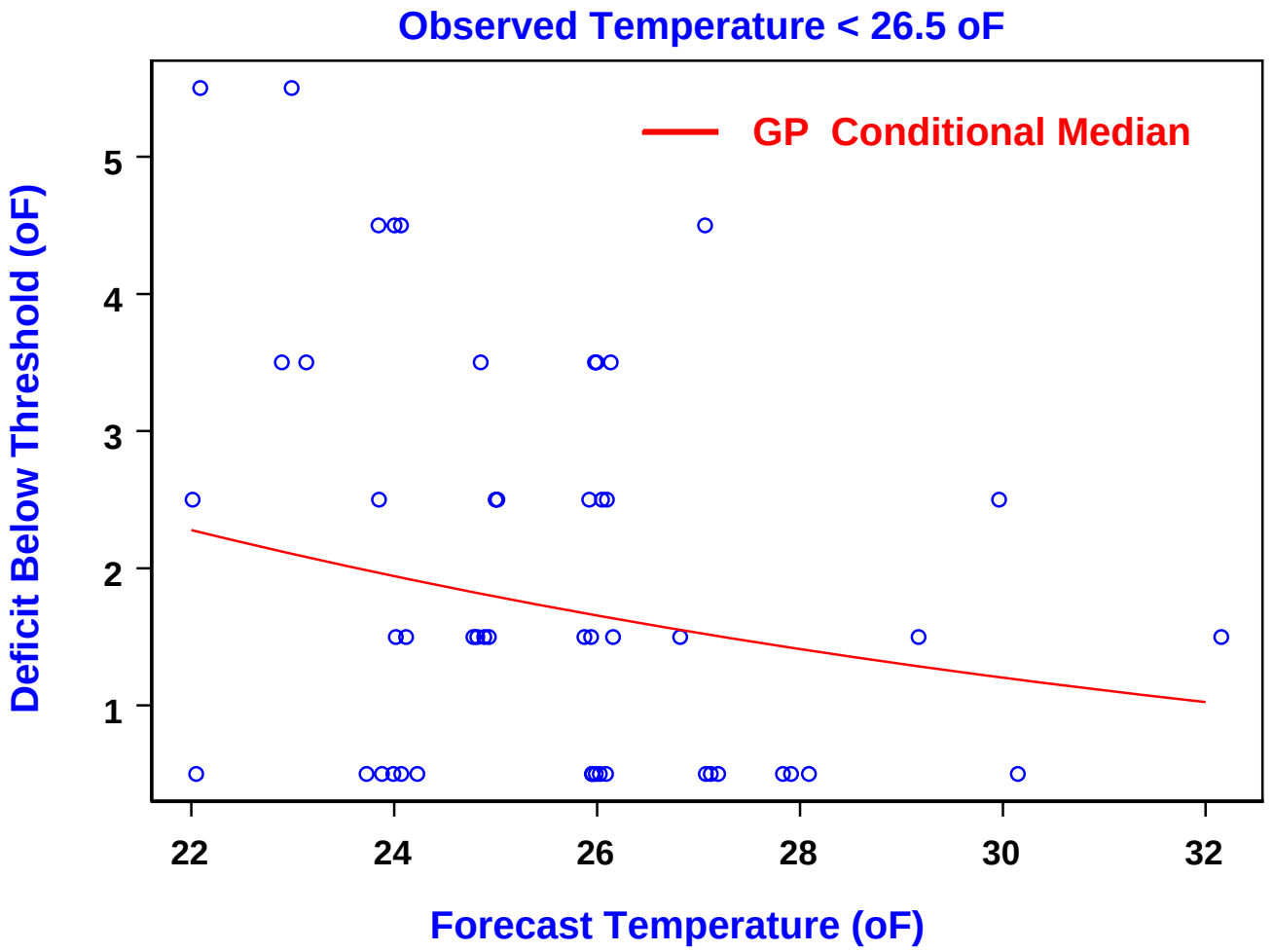
-- Deficit $u - Y$

Threshold u = 26.5 °F

-- Forecast (as covariate)

GP dist.: MLE of $\sigma_1 = -0.0798$

LRT for $\sigma_1 = 0$: P-value < 0.01


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)

(6) REMAINING ISSUES

- **Choice of Threshold**

-- **Vary with forecast?**

- **De-Clustering**

-- **Adjustment for temporal dependence**

- **Parsimony**

-- **Combine Poisson & GP components into single model**

(7) RESOURCES

- **Statistics of Weather and Climate Extremes**

-- **Web site**

`www.isse.ucar.edu/extremevalues/extreme.html`

- **Extremes Toolkit**

-- **Open source software**

`www.isse.ucar.edu/extremevalues/evtk.html`