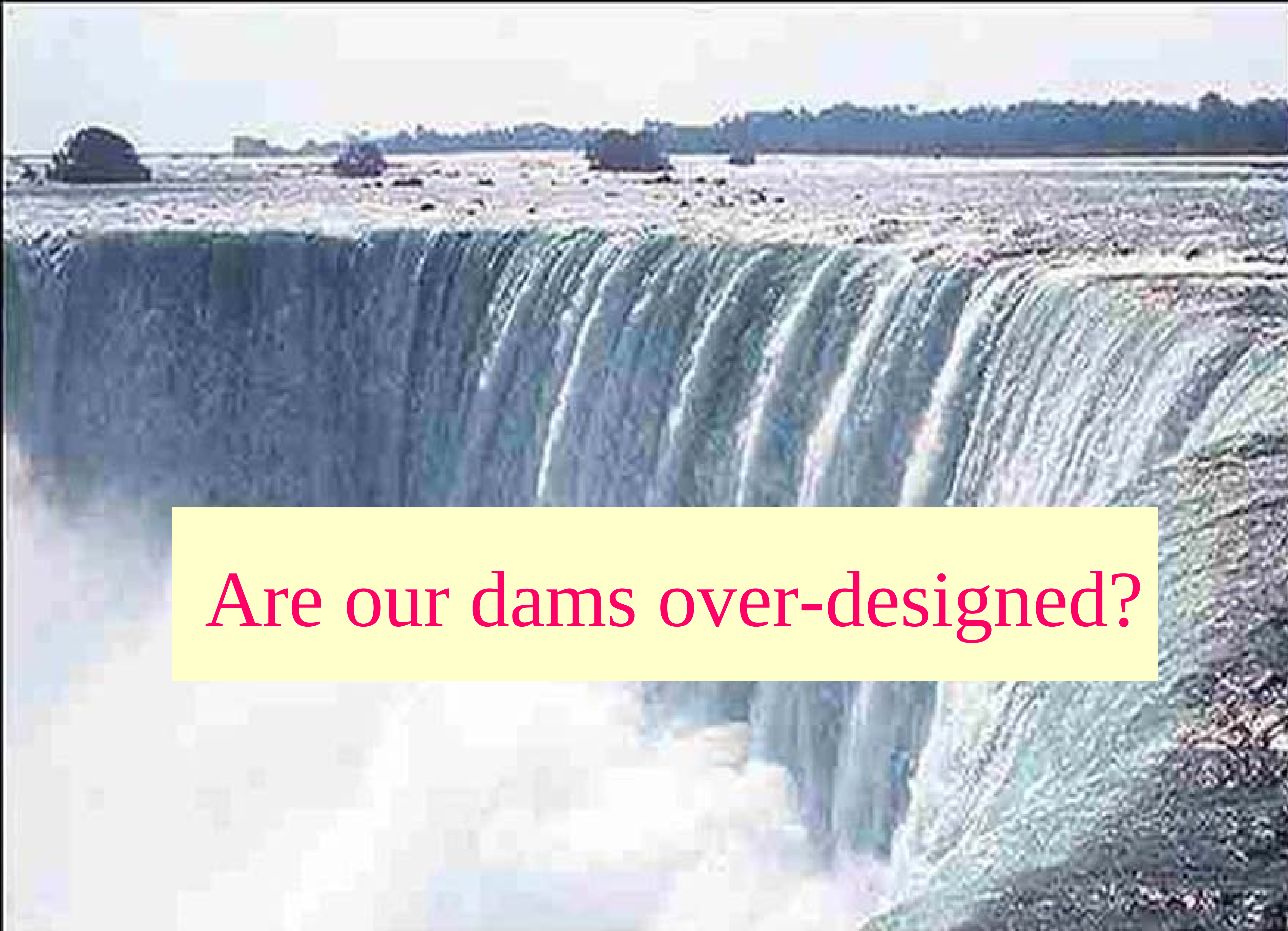


Are our dams over-designed?

Dams

Desirable effects	Undesirable effects
Irrigation	Displacement of people
Electricity	Submergence of valuable forest
	Long gestation period
	Large investment

One factor affecting cost- strength required

A large dam must be able to withstand even a rare but heavy flood

How big a flood to assume in dam design?

PMF- Probable maximum flood

PMP- Probable maximum precipitation

P.C.Mahalnobis

Floods in Orissa

Work half a century ago

Engineers: Build embankments to avoid flood
(caused by rise in riverbed)

PCM: (1923) no noticeable rise in riverbed
Build dams upstream (Hirakud)

North Bengal: build retarding basins to control flood

PCM(1927): Rapid drainage needed
(Lag correlation between rainfall and flood)

Indian Institute of Tropical Meteorology, Pune

1 day PMP atlas for India (1989)

Higher PMP → costlier Dam

Our finding: PMP calculations of IITM - Overestimates

may lead to costlier dams (avoidable)

How to calculate PMP?

X: Max rainfall in a day in one year (at a location)

$P(X > X_T) = 1/T$

Then X_T is called T year return period value of rainfall (convention in hydrology)

X_{100} : value of one day max rainfall exceeded once in 100 years

X_{100} is considered suitable PMP for '**minor dams'.**

Major dams: X_{10000} as PMP

T= 10,000 years.

How to estimate X_T?

Data: daily rainfall records for 90 years.

Y_i = maximum rainfall in a day in year i

$Y_1, Y_2, \ldots Y_{90}$ available

$Est(X_{90}) = \max(Y_1, Y_2, \ldots Y_{90})$

Good enough for minor dams.

What about X_{10000} ?

Purely empirical approach - inadequate.

Model based approach needed.

Model based approach

Extreme value: Gumbel distribution used commonly

$f(x) = 1/ \alpha \, . \exp(- (x-\mu)/ \alpha - \exp((x-\mu)/ \alpha))$

Estimate μ, α by maximum likelihood

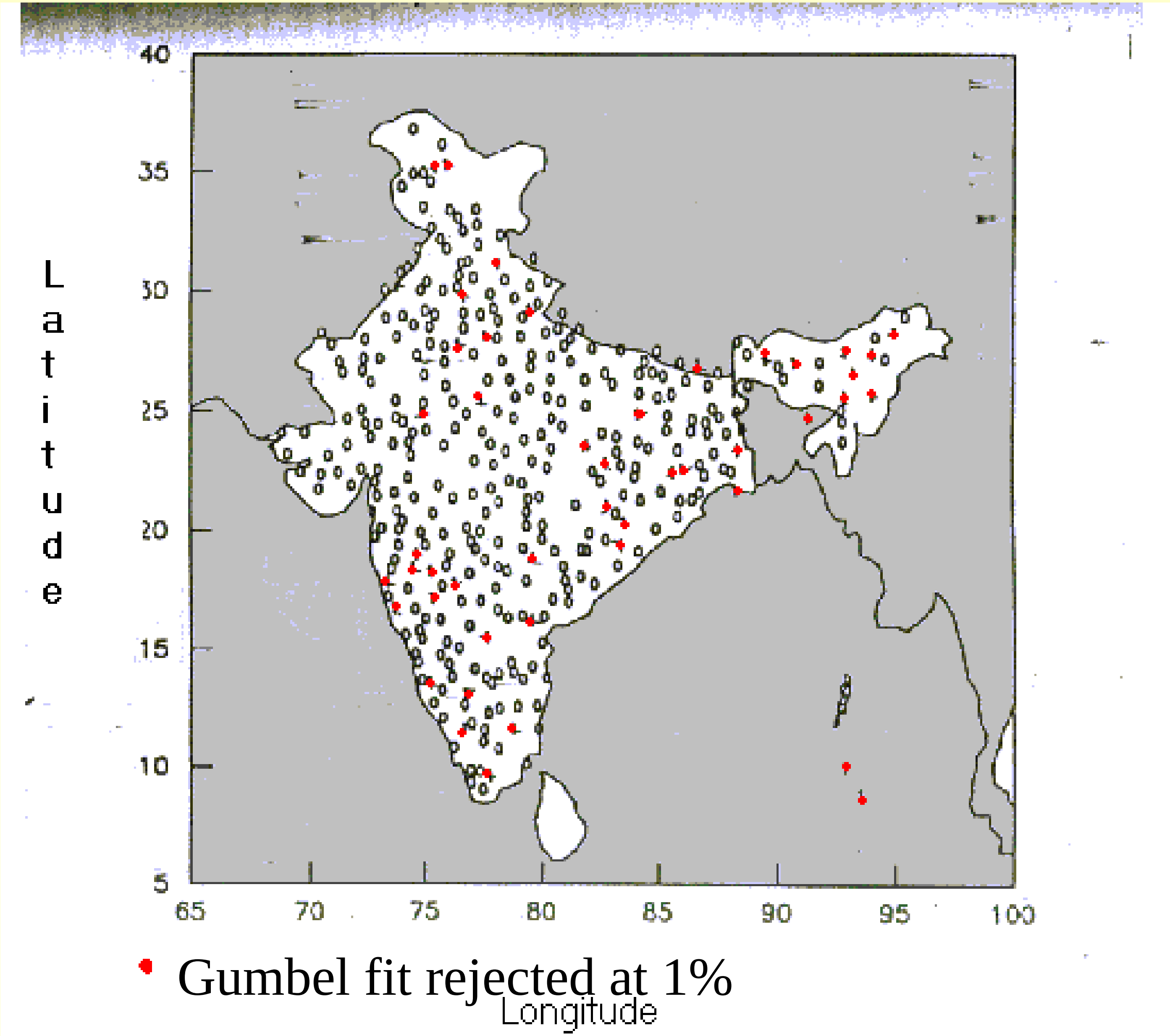
Test goodness of fit by chi-square.

- Data available -358 stations
- Fit good at 86% stations ($\alpha = 0.05$) , 94% stations ($\alpha = 0.01$)

If fit is good

–obtain 10^{-4} upper percentile of the fitted distribution

-- use as estimate of X_{10000}


Latitude
40
35
30
25
20
15
10
5
65
70
75
80
85
90
95
100
Gumbel fit rejected at 1%
Longitude

Hershfield method:

a) $X_1, X_2, \ldots, X_n$ annual one day maxima at a station for n years

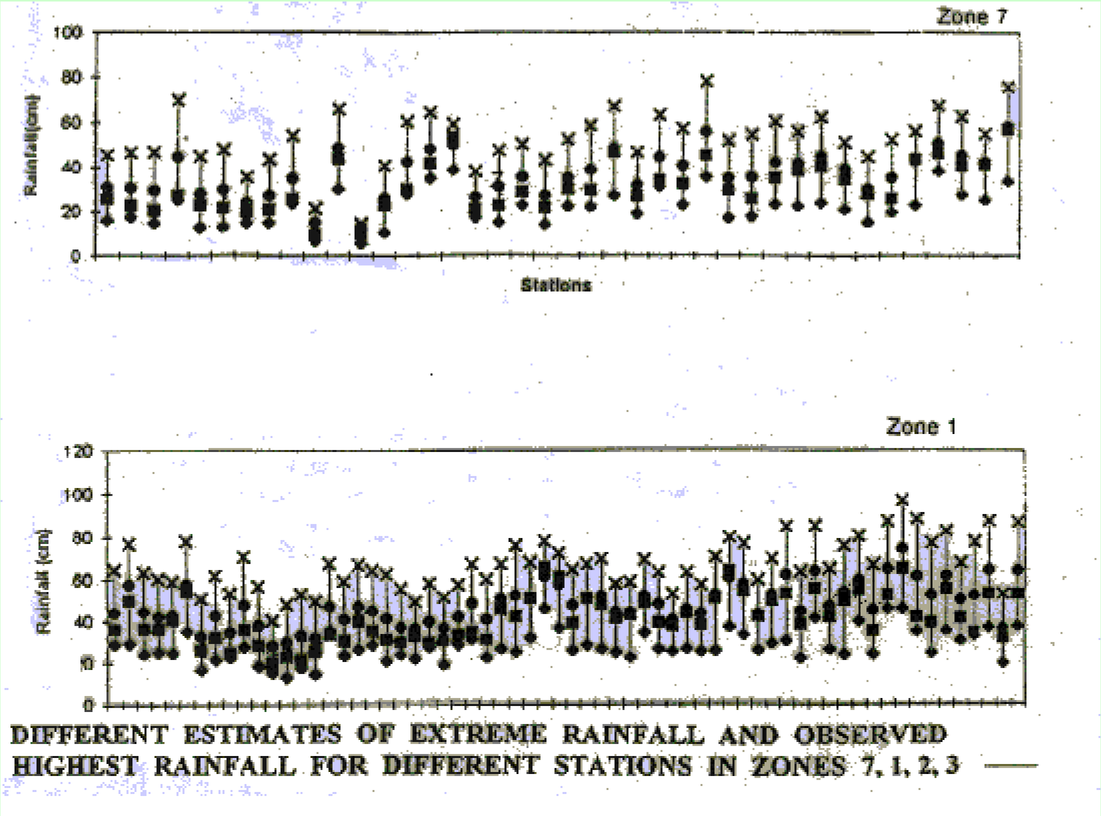
$K = (X_{max} - av(X_{n-1}))/S_{n-1}$

$av(X_{n-1})$: average after dropping X_{max}

b) K_m = largest K over all stations in a locality

$X_{PMP} = av(X_n) + K_m S_n$

How does this method compare with model based method?



X:Hershfield, ◆ Observed highest

■ :10,000 year value(model based)

Hershfield Method overestimates PMP

Stability of model based estimate

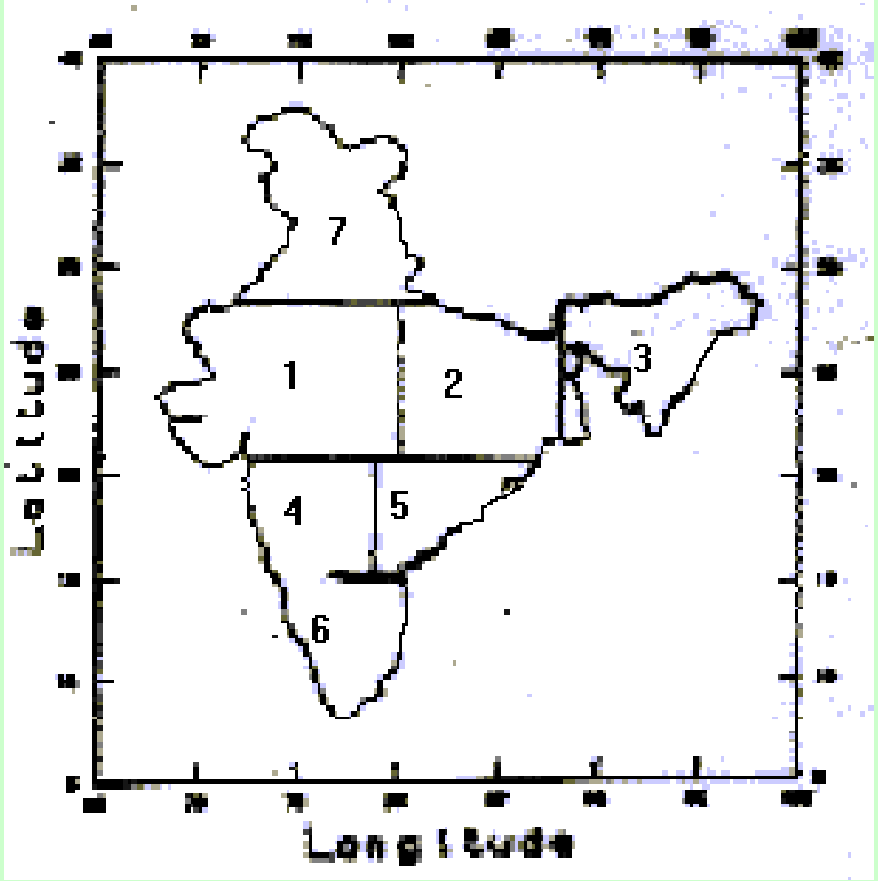
If the new estimator is volatile i.e. has large standard error and is unstable, it may not be usable.

Computing standard errors -analytically intractable

Simulation study carried out

design: for each station generate
100 samples each of size 100
compute competing estimates
empirical mean and sd
Zone-wise comparison

Homogeneous rainfall zones in India



Comparative volatility of estimators

Zone	Model based estimator(cm)		Hershfield estimator(cm)	
	mean	sd	mean	sd
7	30.81	2.54	49.9	4.16
1	41.22	3.59	63.01	5.29
2	43.82	3.65	62.31	5
3	54.16	4.27	64.71	5.92
4	36.90	3.01	46.32	3.93
5	43.07	3.65	55.36	4.50
6	37.15	3.00	34.31	3.12

Proposed estimator more stable

Further work

Gumbel model fitted to
rainfall data from 358 stations

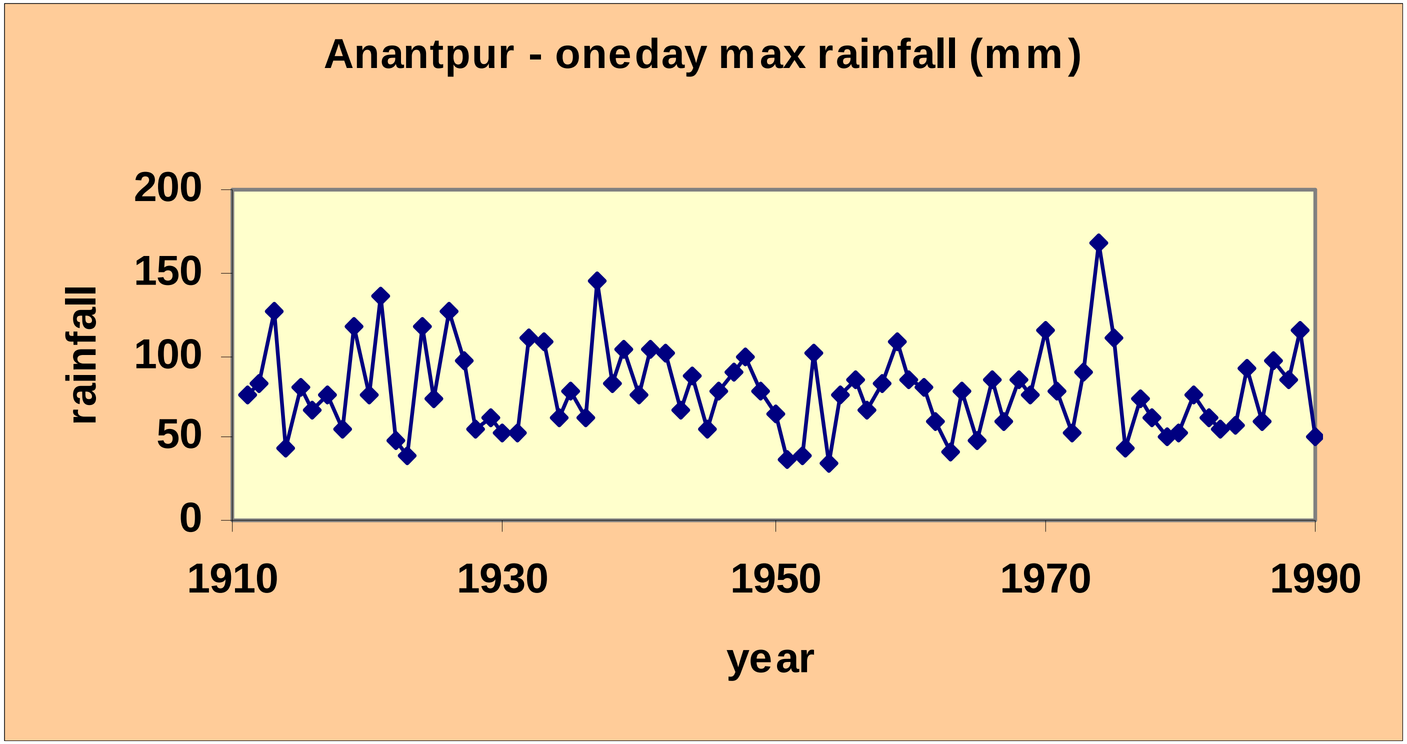
Acceptable fit – 299 stations

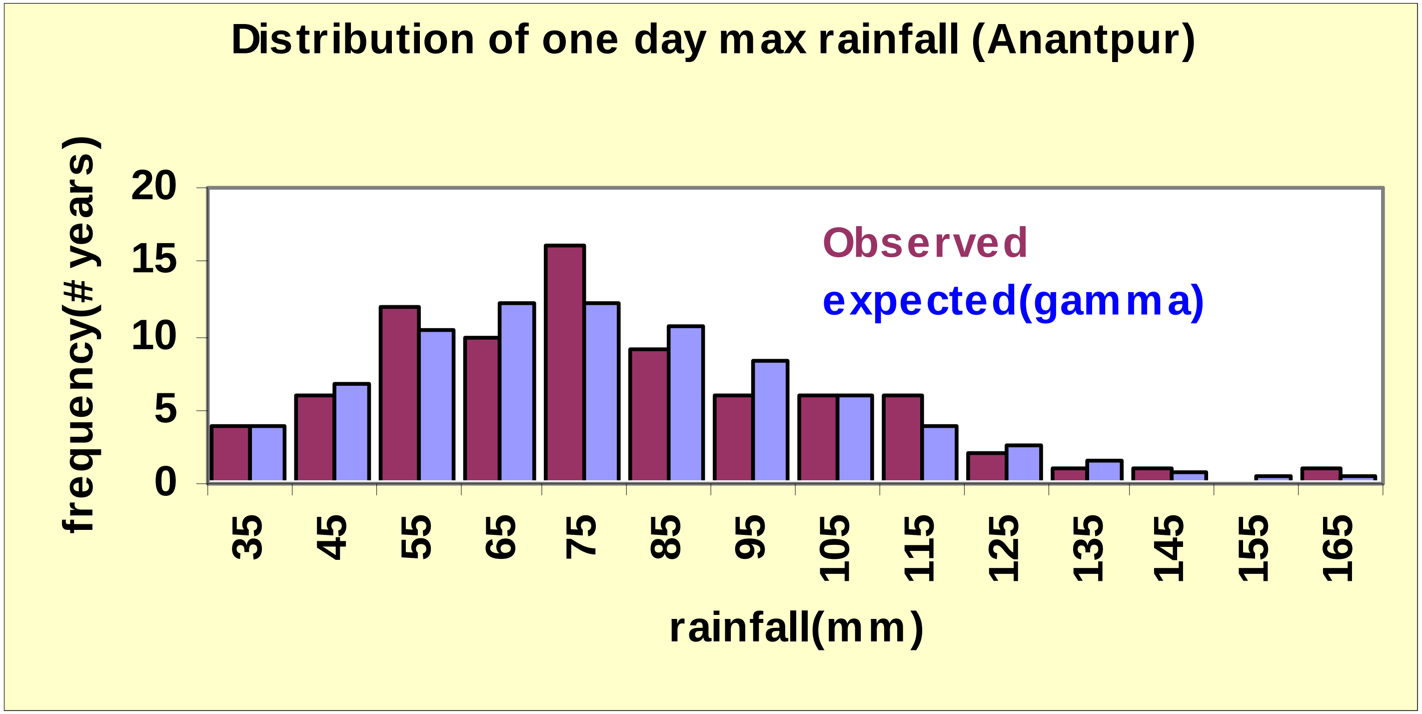
What about remaining stations?

Alternative models: log-normal, gamma, Weibull, Pareto

Which model gives good fit to data?
How robust is the resulting estimate?

Pareto does not fit any data set.


Anantpur - oneday max rainfall (mm)
rainfall
200
150
100
50
0
1910
1930
1950
1970
1990
year


Distribution of one day max rainfall (Anantpur)
frequency(# years)
20
15
10
5
0
Observed
expected(gamma)
35
45
55
65
75
85
95
105
115
125
135
145
155
165
rainfall(mm)

Fitting alternative models

Station	Best fit distribution	Percentile(cm) .001	.0001
Nagpur	Log- normal	20.77	29.88
Bankura	gamma	29.89	36.08
Lucknow	Weibull	22.21	24.36
Baramati	Gamma	17.41	20.41

Estimates based on Gumbel were robust
What about the above?

Simulation study

One station (Pune)
All models

Data 1901- 1990

$Av(X_{(n)}) = 7.09$ cm

$sd(X_{(n)}) = 2.64$ cm

Parameter of each model

Chosen such that

Mean, sd match with

Observed values

Sample size 100

Parameters chosen and true quantiles

Distribution	Parameters		Quantiles(mm) .001	.0001
Gamma	a= 7.21	b=9.84	180.88	213.0
Weibull	a=3.76	b=77.36	129.4	139.7
Log-normal	μ =4.2	σ =0.36	136.6	195.9

a: shape parameter, b: scale parameter

Bias and MSE stabilized at 2000 samples

Results of simulation study (10,000 simulations)

Distribution	Return period(T)	True value	estimate	RMSE
Gamma	1000	180.99	182.07	1.54
	10000	213.01	213.76	1.54
Log-normal	1000	136.62	135.65	8.60
	10000	195.92	194.30	16.63
Weibull	1000	129.40	129.80	5.14
	10000	139.77	140.14	6.20

Estimates are stable in these distributions as well
Gamma model performs better than others.