RECLAMATION

Managing Water in the West

Developing Precipitation Frequency Estimates in Regions of Complex Terrain

Kathleen D. Holman[1], Andrew P. Verdin[1], David P. Keeney[1], and Joseph Kanney[2]

[1] Bureau of Reclamation Technical Service Center
[2] U.S. Nuclear Regulatory Commission

U.S. Department of the Interior
Bureau of Reclamation

Outline

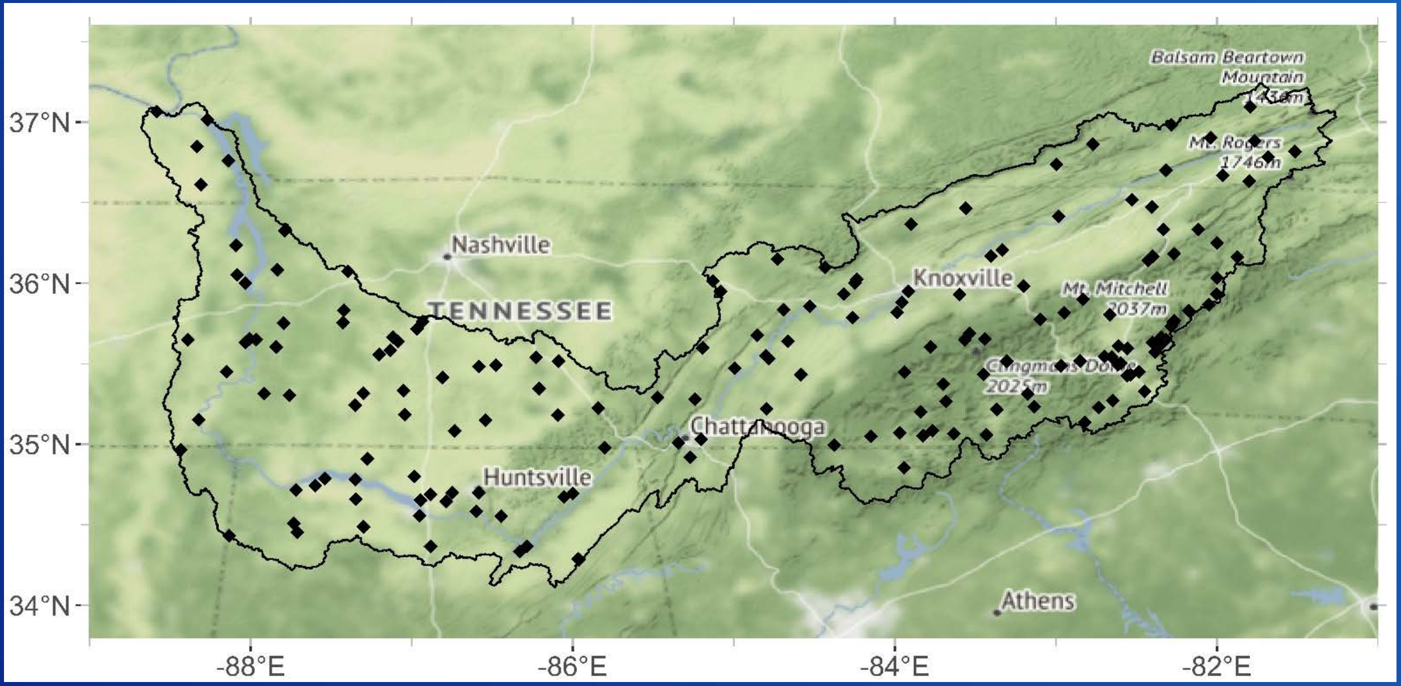
1. Motivation
2. Case study in Tennessee River Valley
   A. Data
   B. Precipitation-Frequency Methods
   C. Results
3. Summary & Conclusions

Motivation

I. Extreme precipitation events vital for engineering design and planning
II. Deterministic methods widely used for design projects
    a) Singe value with no associated probability
III. Precipitation-frequency analyses provide range of magnitudes and probabilities

Case Study

Tennessee River Valley Watershed



Observations from GHCN-Daily stations with 85% data availability for 10+ years period of record

RECLAMATION

Precipitation-Frequency Methods

L-Moments:

– Summarized in Hosking and Wallis (1997)
– Deterministic system for describing PDFs
– Based on linear combinations of moments
  – λ_1=L-location (mean)
  – λ_2=L-scale (variability or dispersion)
  – λ_3=L-skewness (asymmetry)
  – λ_4=L-kurtosis (thickness of tail)
– Can estimate parameters for 6 distributions; focus on GEV

Bayesian Inference:

– Application of Bayes Rule; assume observations ($\boldsymbol{Y}$) are fixed and estimate PDF parameters (θ)

$$p(\theta|\boldsymbol{Y}) = \frac{p(\boldsymbol{Y}|\theta)p(\theta)}{p(\boldsymbol{Y})} \propto p(\boldsymbol{Y}|\theta)p(\theta)$$

– Consider the GEV likelihood function
– Define prior distributions for GEV params, apply Monte Carlo sampling, acceptance criteria, build posterior distributions of θ

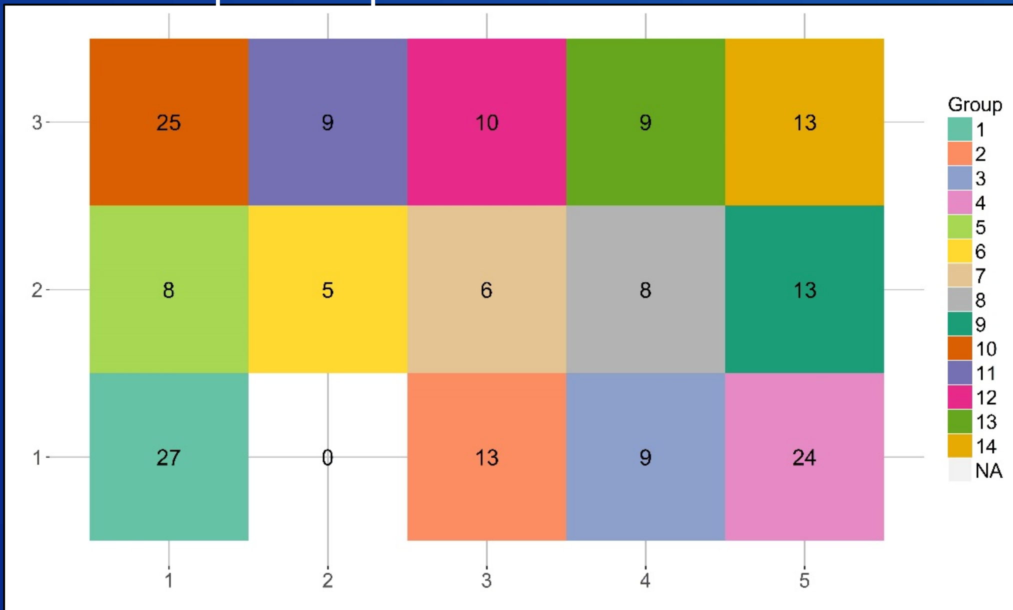
RECLAMATION

Regional Frequency Approach

1. Identify homogeneous region(s)
2. Screen observations for false/erroneous records
3. Identify annual (or seasonal) maxima
4. Estimate GEV distribution parameters
   - L-moments
   - Bayesian inference
5. Compute point precipitation frequency results

Self-Organizing Map

SOM Output Map



Clustering algorithm used to "group" stations with similar attributes

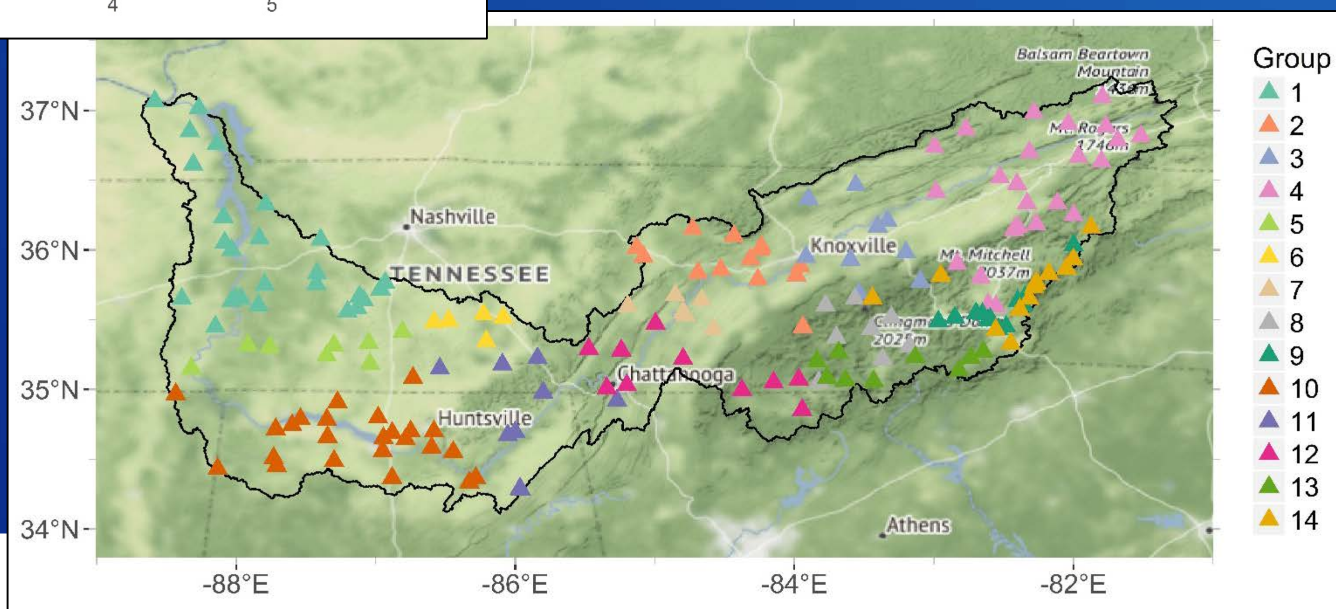
Apply SOM algorithm to :

- Latitude
- Longitude
- Elevation
- *Avg annual precipitation*
- *Avg annual max one-day precipitation*

Each station maps to a single SOM node

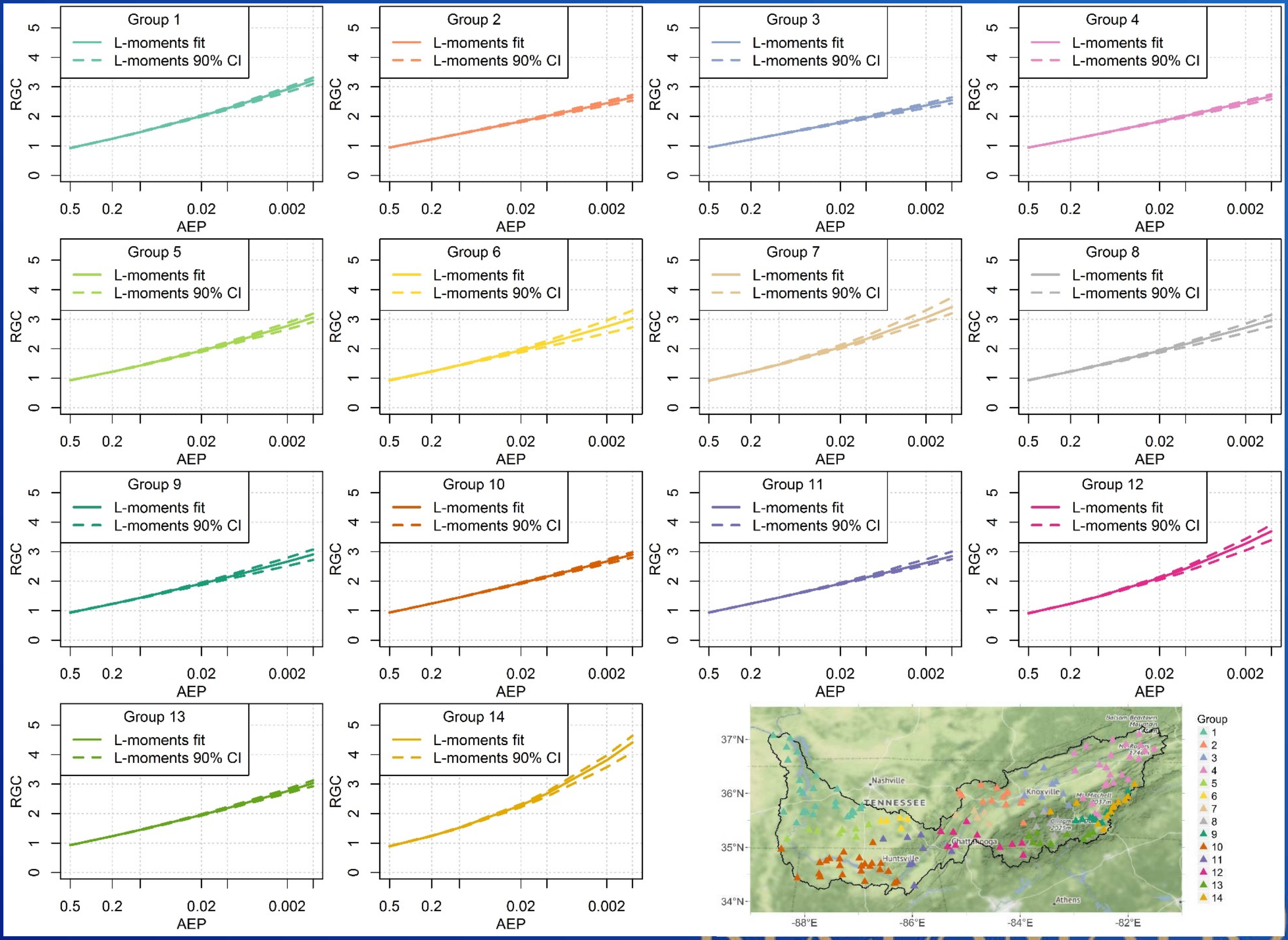
Gauges mapped to same node define homogeneous regions

Homogeneous regions need not be contiguous

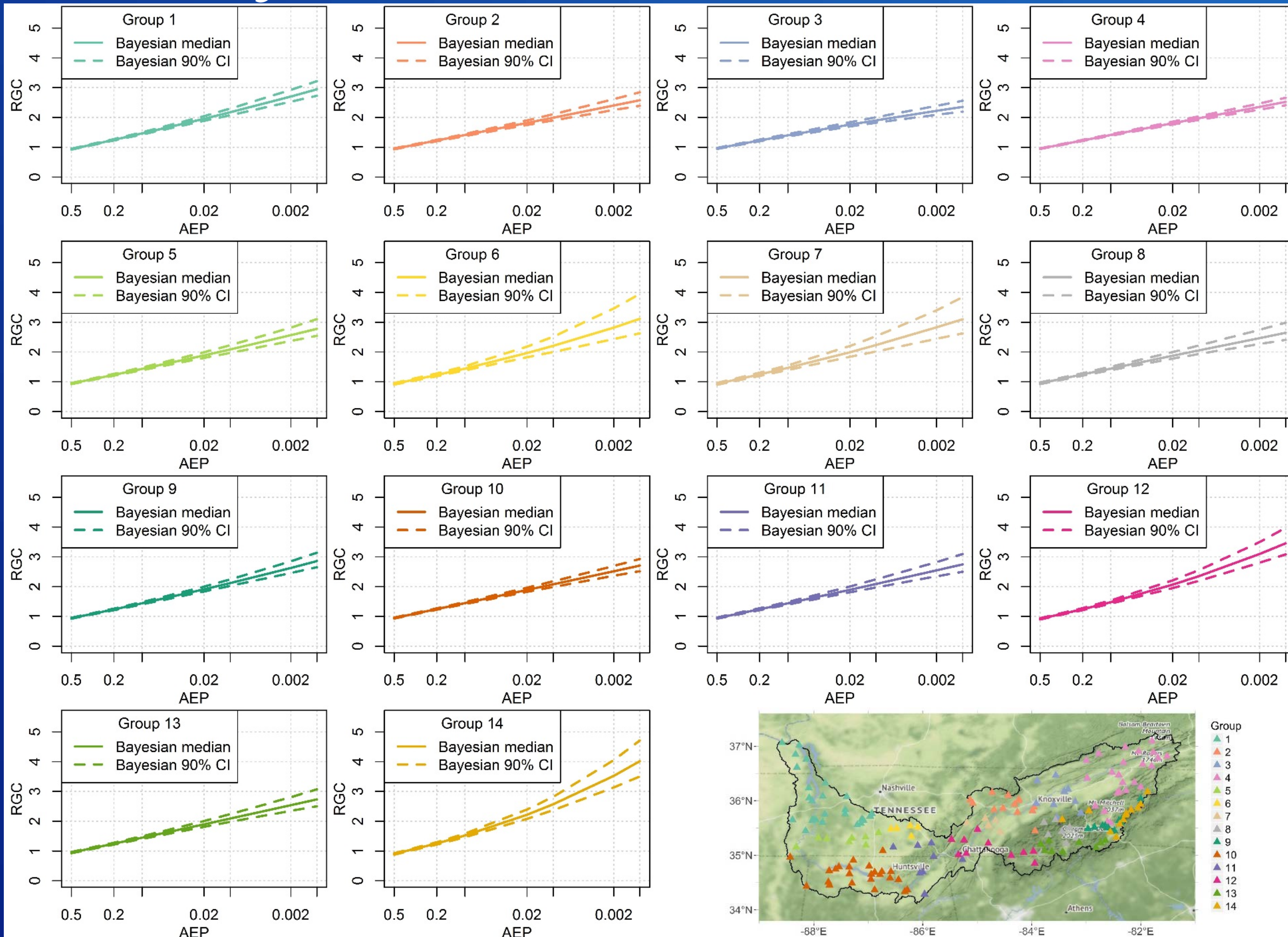


Available R packages for SOM analysis:
library("som")
library("kohonen")

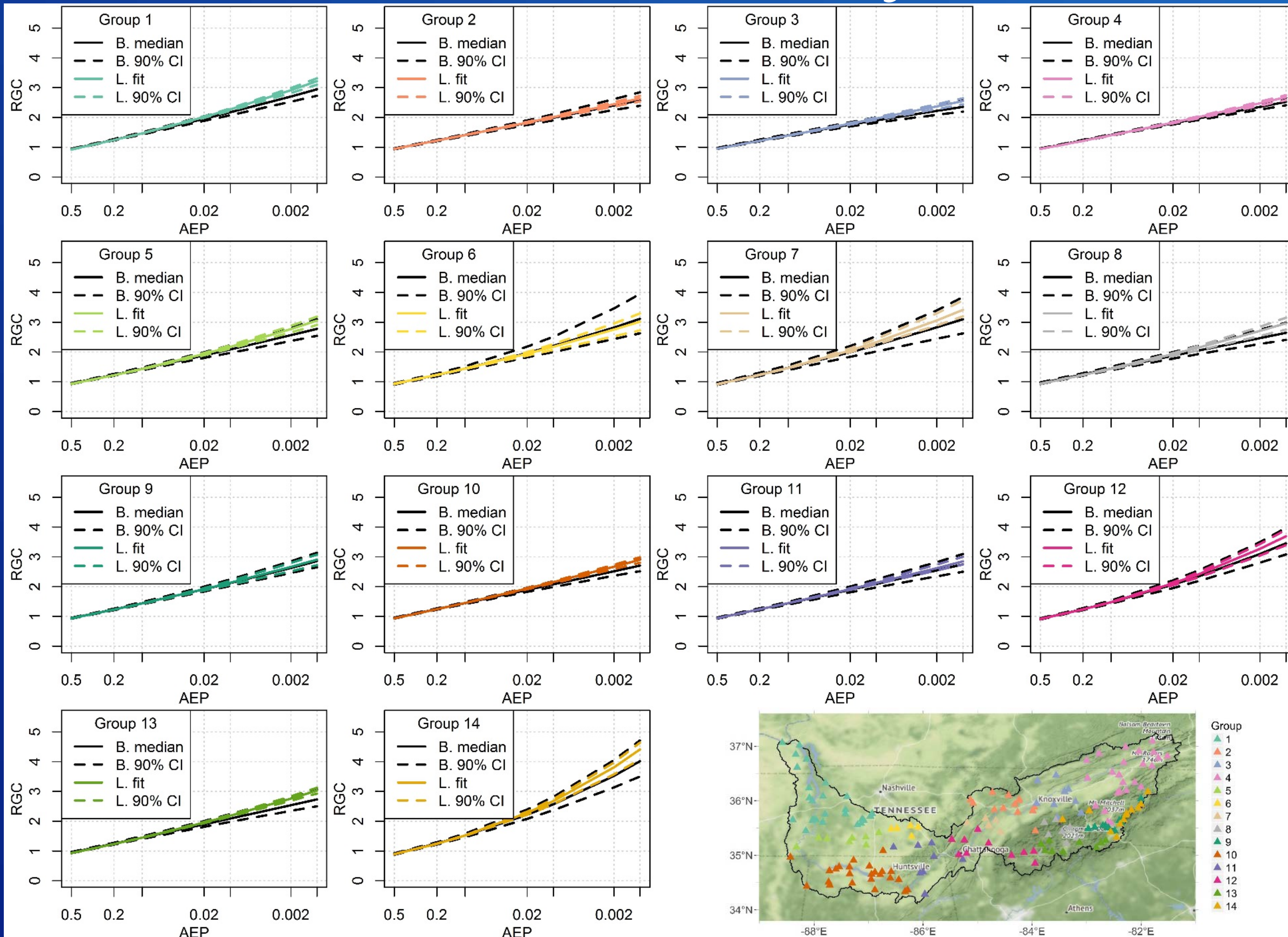
L-Moments RGCs



Bayesian Inference RGCs



L-Moments vs. Bayesian



Summary

Precipitation-frequency analyses provide users with expected return periods of heavy events

Frequency estimates vary based on estimation method

- In all but one region, L-moments best-estimates exceed Bayesian best-estimates
- Uncertainty bounds from Bayesian inference always exceed L-moments

Additional testing is needed to understand operational benefits of Bayesian inference over L-moments

RECLAMATION

Questions?

Katie Holman
kholman@usbr.gov
+1.303.445.2571

Andrew Verdin
averdin@usbr.gov
+1.303.445.3647

Joseph Kanney
joseph.kanney@nrc.gov
+1.301.415.1920

This work was funded by the U.S. Nuclear Regulatory Commission under contract NRC-HQ-60-11-I-006

RECLAMATION

Question

How do precipitation-frequency estimates developed using the L-moments algorithm compare with estimates developed using Bayesian inference?

L-Statistics

System for describing probability distribution functions based on linear combinations of moments

<u>L-moments:</u>

λ_1=L-location (mean)

λ_2=L-scale (variability or dispersion)

λ_3=L-skewness (asymmetry)

λ_4=L-kurtosis (thickness of tail)

<u>L-moment ratios (dimensionless):</u>

$T_r=\lambda_r/\lambda_2$

T=L-CV= λ_2/λ_1 (variability)

Hosking and Wallis (1997)

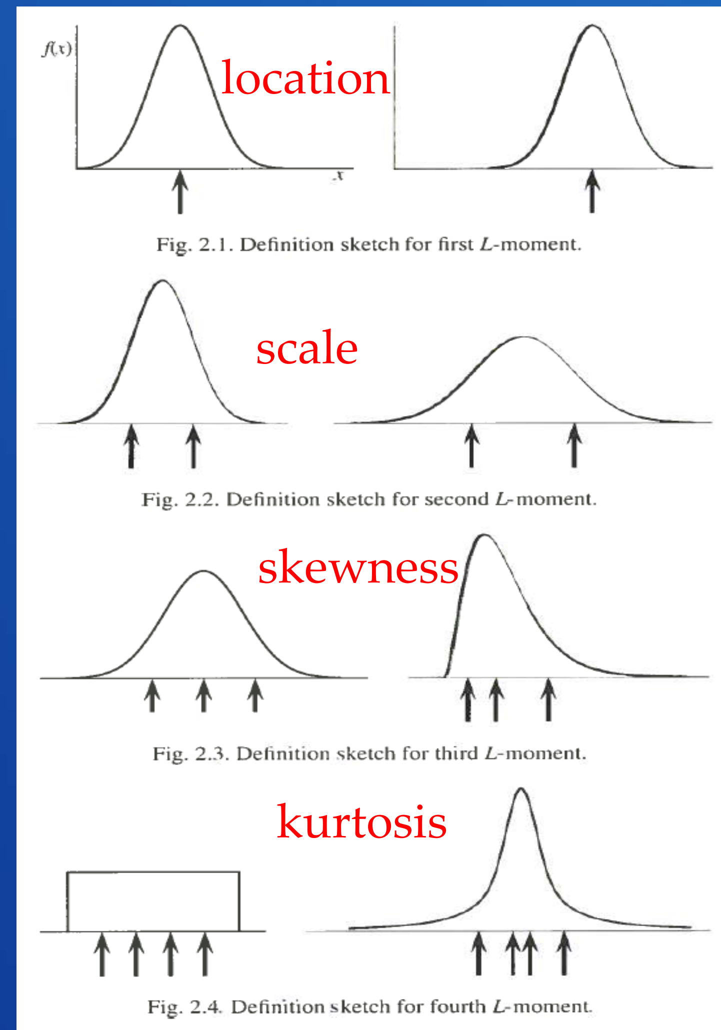


Fig. 2.1. Definition sketch for first *L*-moment.

Fig. 2.2. Definition sketch for second *L*-moment.

Fig. 2.3. Definition sketch for third *L*-moment.

Fig. 2.4. Definition sketch for fourth *L*-moment.

Available R packages for L-moments:
library("lmom")
library("lmomRFA")

Bayesian Inference

Bayes' Rule in a modeling framework:

$$p(\theta|\boldsymbol{Y}) = \frac{p(\boldsymbol{Y}|\theta)p(\theta)}{p(\boldsymbol{Y})} \propto p(\boldsymbol{Y}|\theta)p(\theta)$$

where $\boldsymbol{Y} = (y_1, y_2, \dots, y_n)$ and $\theta = (\mu, \sigma, \xi)$

- Define *prior distributions* for model parameters θ (a priori knowledge)
- Consider GEV likelihood function (can consider GNO, GLO, etc.)
- Monte Carlo, acceptance criteria, builds *posterior distributions* of θ

Bayesian inference derives the *posterior probability* as a consequence of a *prior probability* and a *likelihood function*

Available R packages for Bayesian inference:
library("rstan")
library("spBayes")

Bayesian inference

Prior $p(\boldsymbol{\theta})$: the strength of our belief in θ without the data Y

Posterior $p(\boldsymbol{\theta}|\boldsymbol{Y})$: the strength of our belief in θ when the data Y are taken into account

Likelihood $p(\boldsymbol{Y}|\boldsymbol{\theta})$: the probability that the data Y could have been generated by the model with parameter values θ

Evidence p(Y): the probability of the data according to the model, determined by summing across all possible parameter values weighted by the strength of belief in those parameter values

→ typically unknown, can be ignored with proportionality

→ essentially a normalizing constant

→ does not enter into determining relative probabilities (models)

Available R packages for Bayesian inference:
library(“rstan”)
library(“spBayes”)

SOM Results – At-Site Means