# Fidelity of Climate Extremes in High Resolution Global Climate Model (GCM) Simulations

Salil Mahajan[1], Katherine J. Evans[1], Marcia L. Branstetter[1], Valentine Anantharaj[1] and Juliann Leifeld[2]

[1]Oak Ridge National Laboratory

[2]University of Minnesota

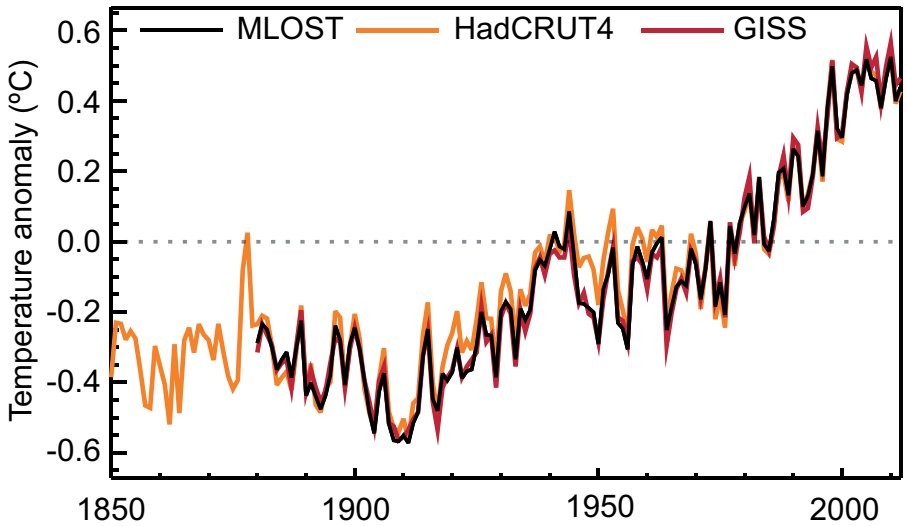
Oak Ridge National Laboratory

Climate Change Science Institute
Oak Ridge National Laboratory

# Outline

- Motivation
- Precipitation
- Use an objective approach to quantify precipitation extremes:
  - Generalized Extreme Value theory
- Use a regionalization framework to improve sampling
- Compare stationary and non-stationary extremes

Global Mean Temperature Anomalies

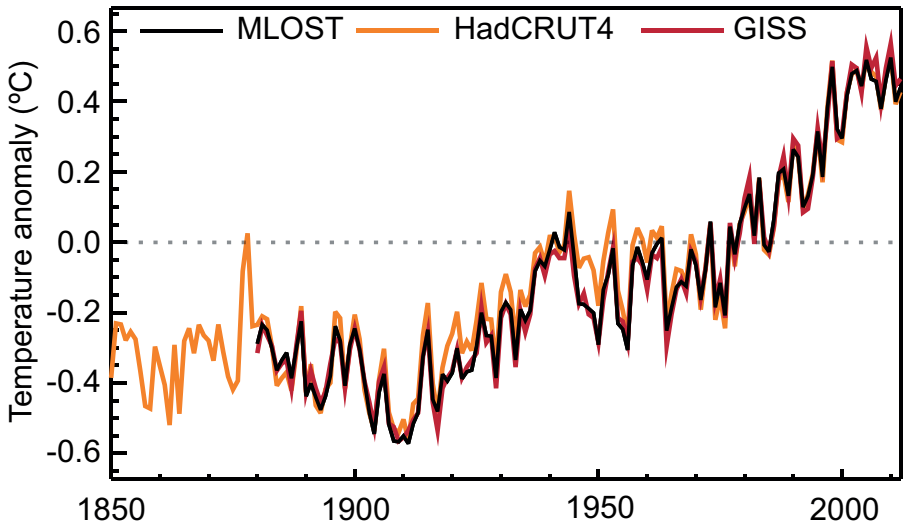



*Source: IPCC AR5 (Climate Change: The Physical Basis)*

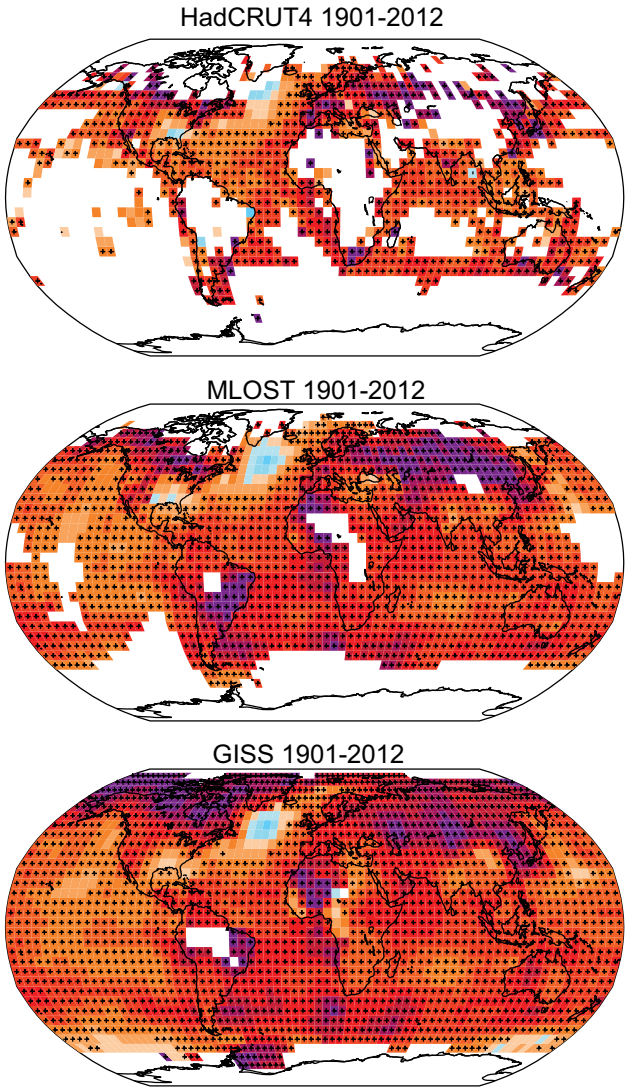
# Motivation

- Trends

Global Mean Temperature Anomalies




Trends in Mean

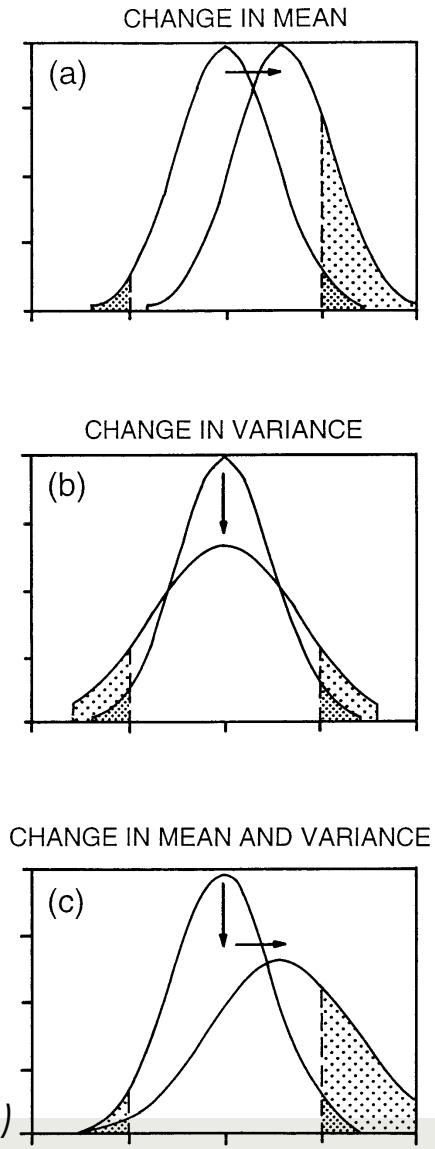



*Source: IPCC AR5 (Climate Change: The Physical Basis)*

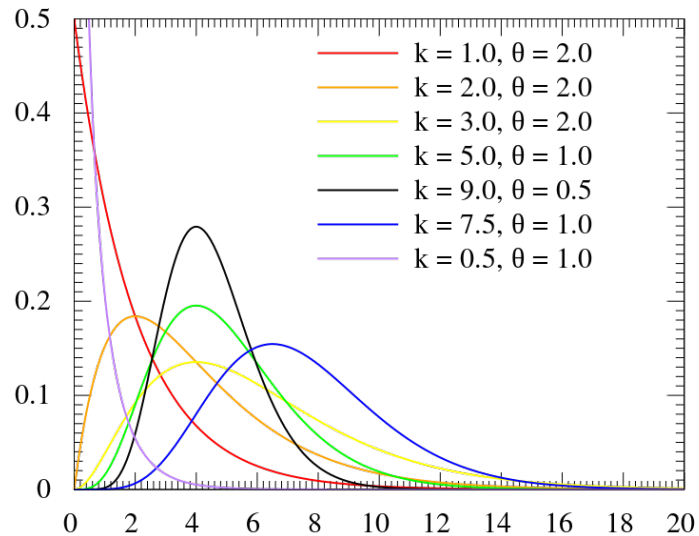
# Extremes

- Change in PDF

Temperature




Precipitation



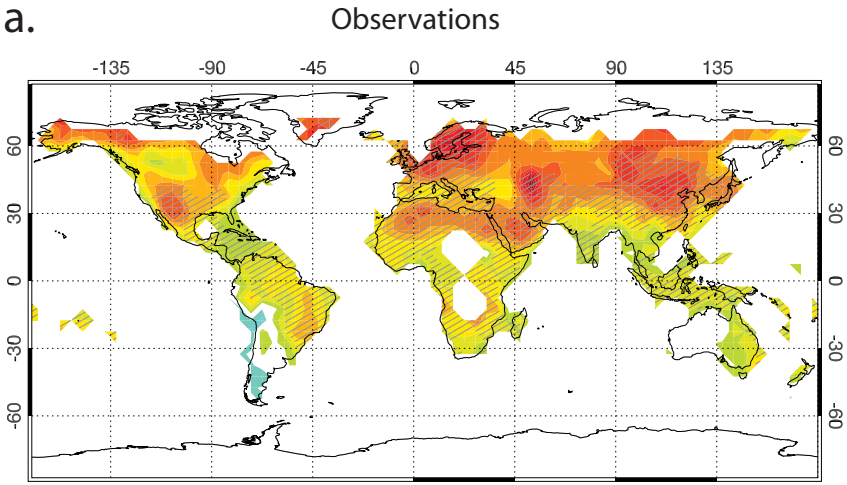

*Source: Wikipedia (Gamma Distribution)*

*Source: Meehl et al. (2000)*

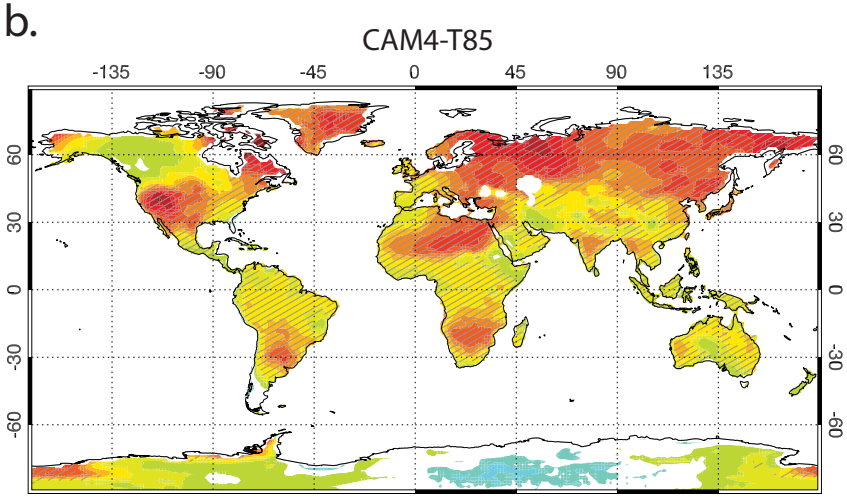
# Extremes: Trends in Temperature

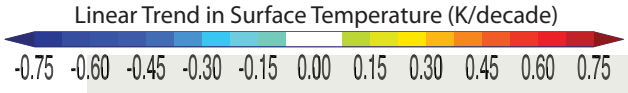
## Trends in Mean Temperature: (1979-2005)

a.



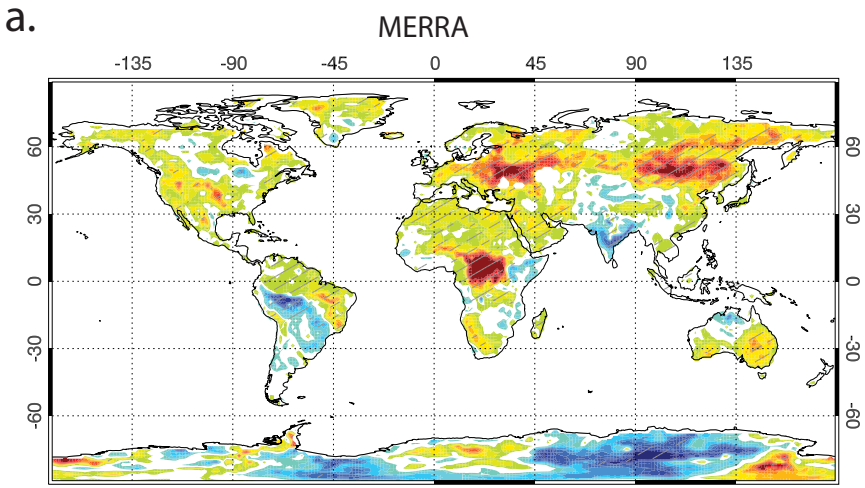

b.



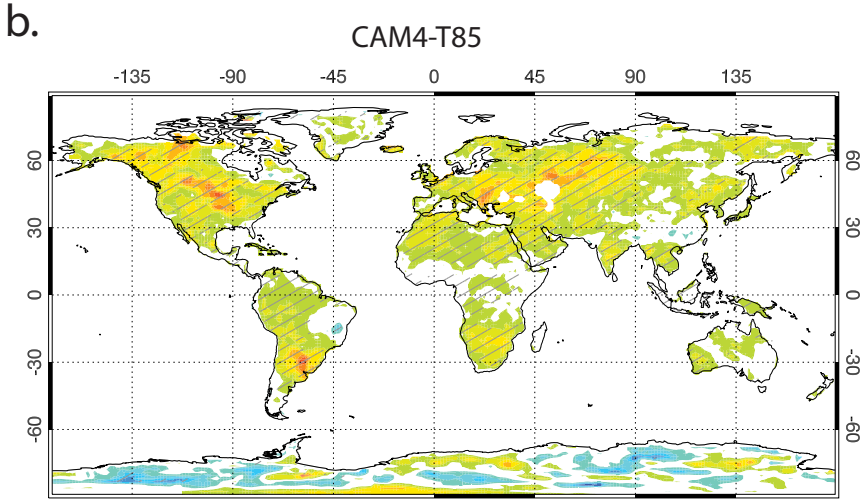



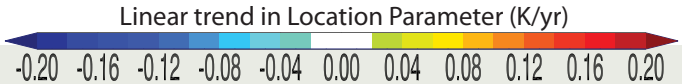

## Trends in Extremes: (1979-2005)

a.




b.



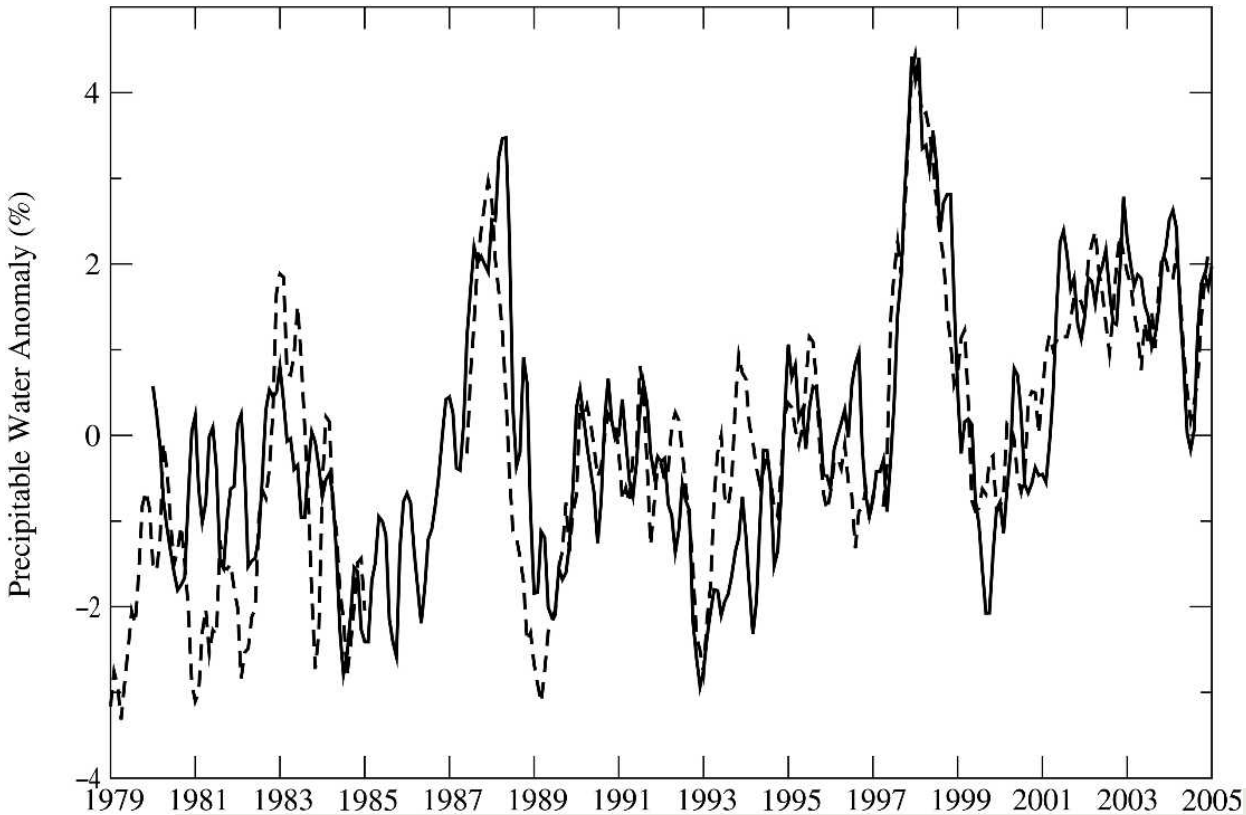

*Evans et al. (2014)*

# Precipitation Response to Warming

- Clausius-Clapeyron Equation:

$$\frac{d\ \ln e_s}{dT} = \frac{L_v}{RT^2}$$

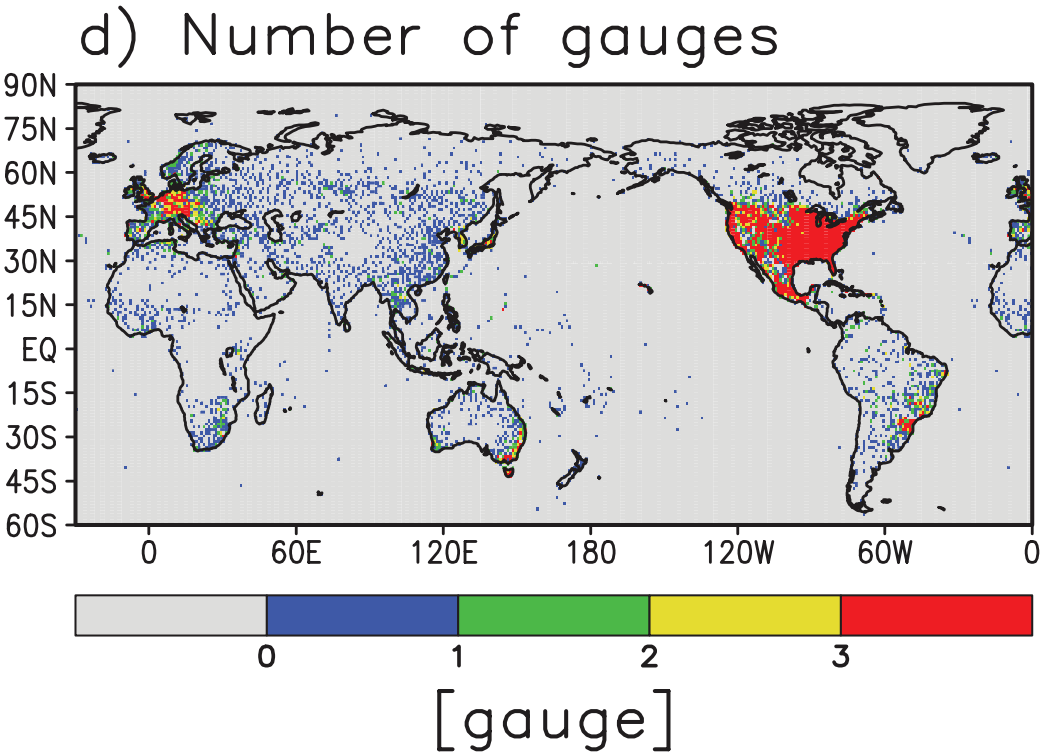
- Increase in atmospheric water vapor: 7% per K




*Time Series of globally averaged ocean-only column integrated water vapor anomalies: Held and Soden (2006)*

# Precipitation Data

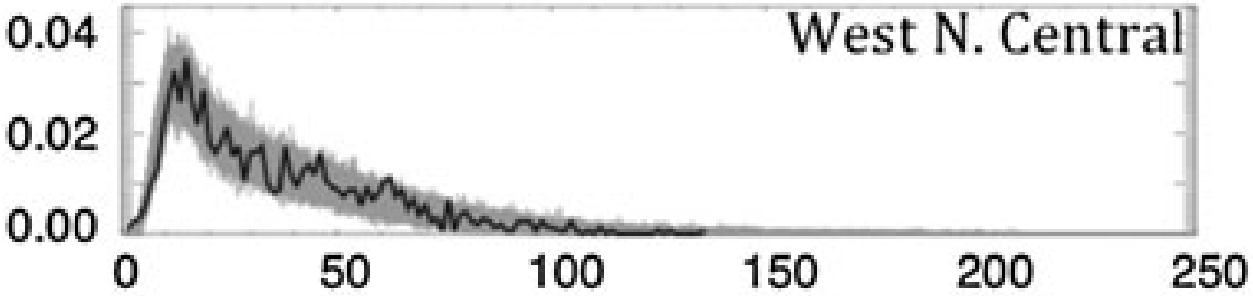
- Long-term high density observations exists for only parts of the world
- Long Historical Data only reliable after large spatial aggregation:
  - to reduce error due to poor sampling




*Gauge density in 1 degree grid boxes in 2005. Source: Chen et al. (2007)*

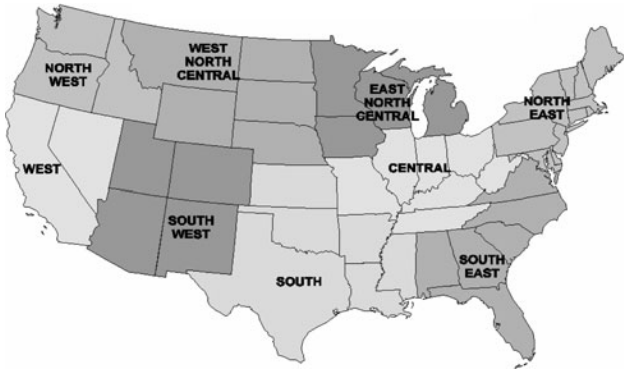
# Precipitation

- Delta function at zero
- Skewed
- Typically modeled by exponential, gamma or log-normal distribution.
- Spatial correlation
- Non-stationary?




*PDF of monthly Precipitation (Mahajan et al. 2011)*

# Non-stationary Precipitation Extremes



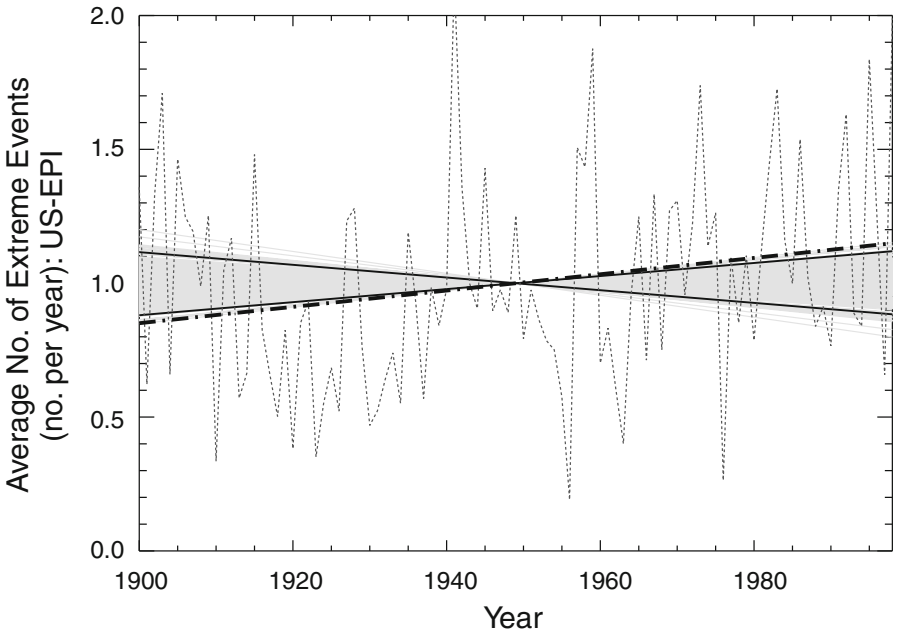

*Climate Regions of US. (Source: NOAA CPC)*

- Long term trends (~100 yrs)
- Fixed regional domains
- Parametric:
  - Monthly precipitation
  - Allow for spatial correlation
  - Assume log-normal distribution of precipitation
  - Threshold for extremes: 90 percentile
  - Bootstrapping to establish confidence intervals:
    - Random sampling from the distribution
- Non-parametric:
  - Random sampling from the observed data

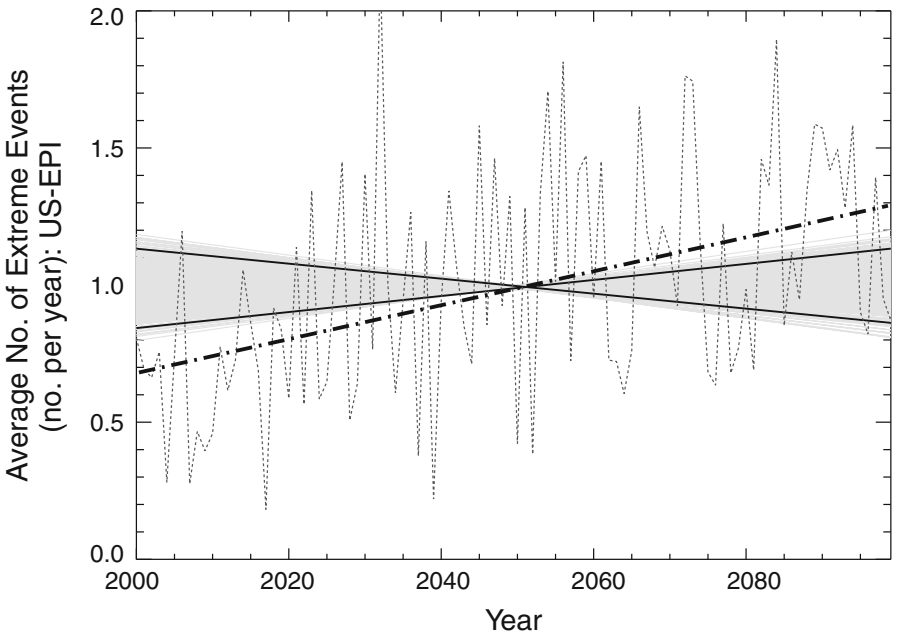
# Non-stationary Precipitation Extremes

- Positive trends observed over the past century
- Significant trends in climate model projections of the 21st century

Observations (1900-2000)



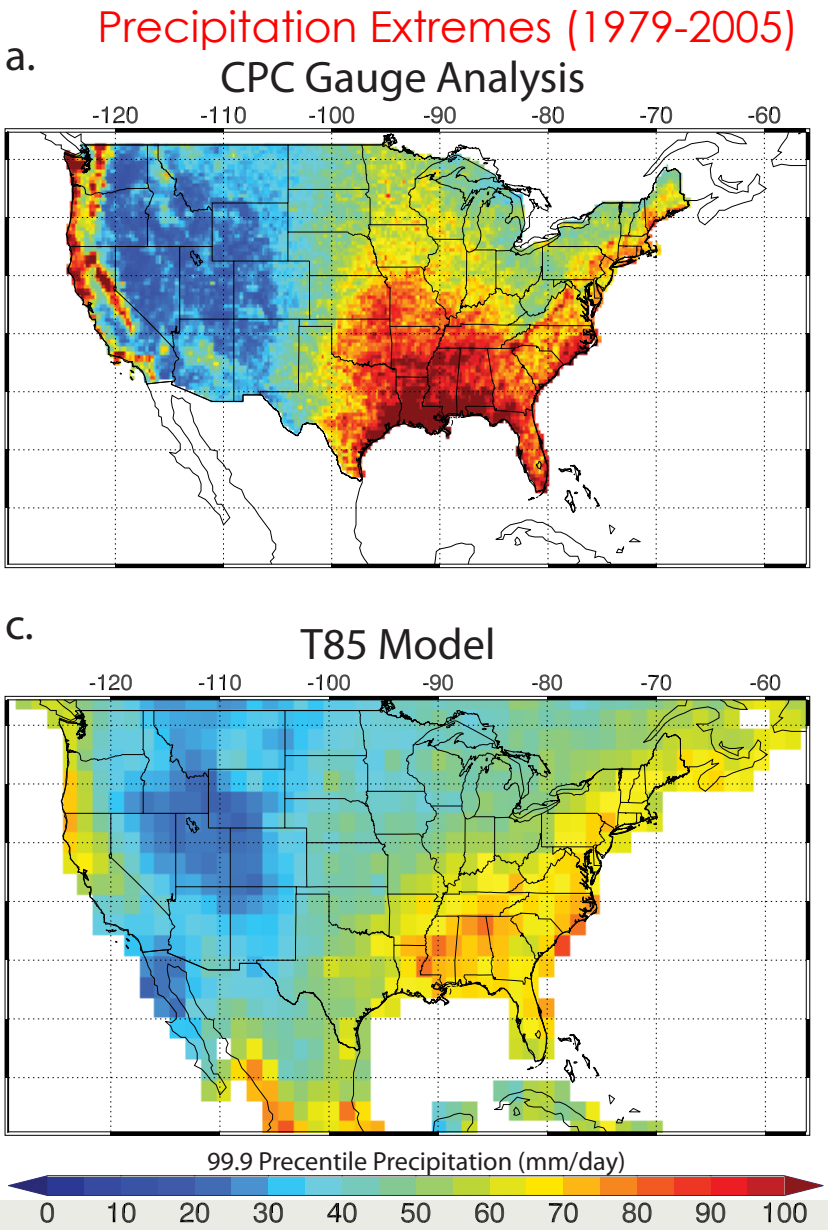

Model Projections (2000-2100)




*Source: (Mahajan et al. 2011)*

# Biases in CMIP5 models

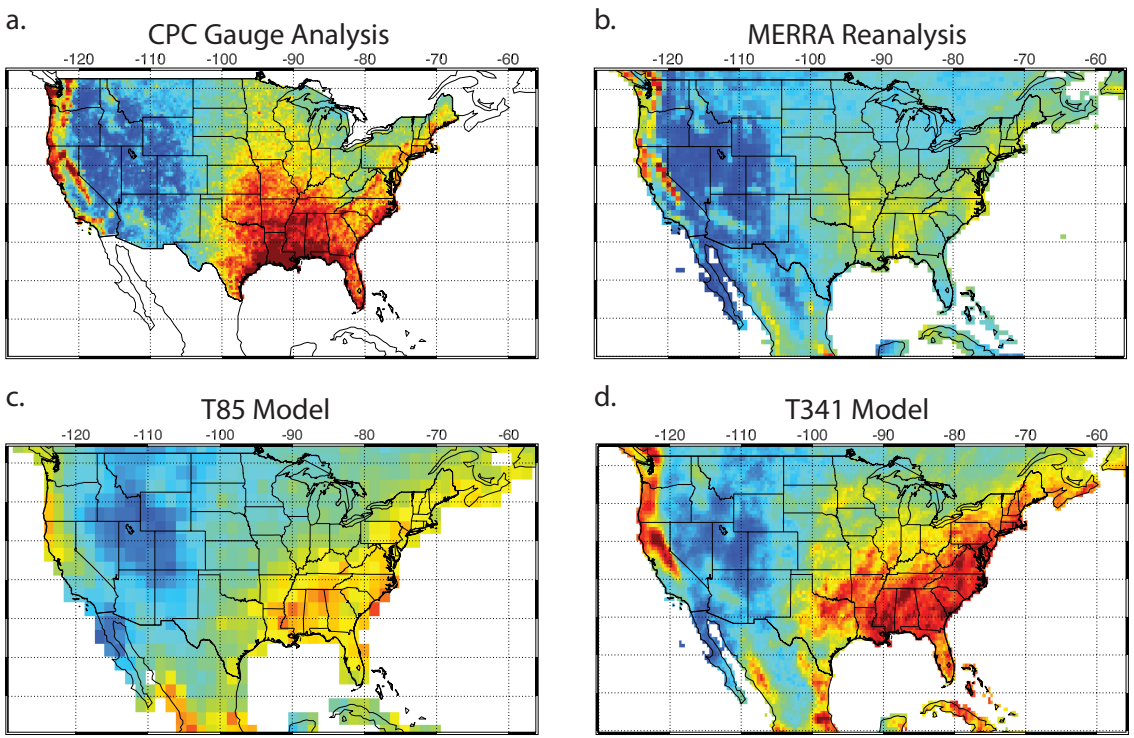
- Long standing mean biases remain
- Biases in precipitation extremes
- Create Better Models:
  - Improve and add processes
  - Resolve finer features: High Resolution




*Mahajan et al. 2014*

# High Resolution Models

- High resolution (~ 0.25 degrees) climate simulations are becoming common.
  - Dedicated projects
  - Increase in computational resources
  - Long standing mean biases remain, more realism, improvements over orography
- Explore methods to evaluate climate extremes statistics



# Generalized Extreme Value (GEV) Distribution

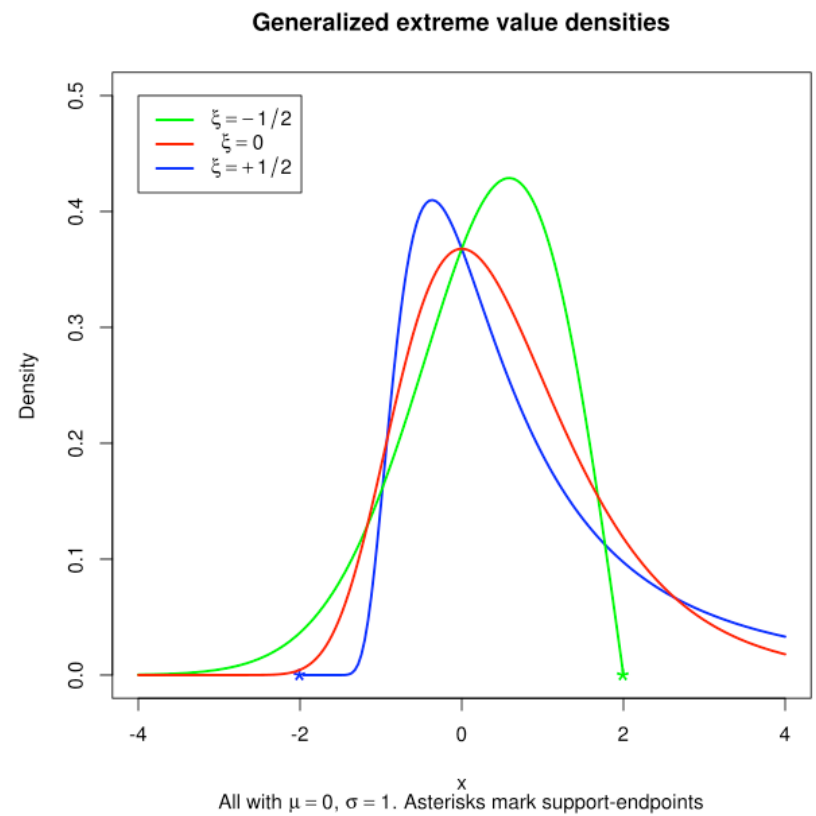
$$G(z) = \exp\left\{-[1+\xi(\frac{z-\mu}{\sigma})]^{-1/\xi}\right\}$$

where $\mu$, $\sigma$ and $\xi$ represent the location, scale and shape parameter respectively.

defined on:

$$z : 1 + \xi(z-\mu)/\sigma > 0$$

- Extreme Value Theory: Linearly normalized values of the max./min. of a process belong to GEV irrespective of the distribution of the population
- Analogous to central limit theorem: sample means belong to normal distribution




*Source: Wikipedia - http://en.wikipedia.org/wiki/File:GevDensity.svg#/media/File:GevDensity.svg*

# Stationary and Non-stationary Extremes

$$G(z) = \exp\left\{-[1+\xi(\frac{z-\mu}{\sigma})]^{-1/\xi}\right\}$$

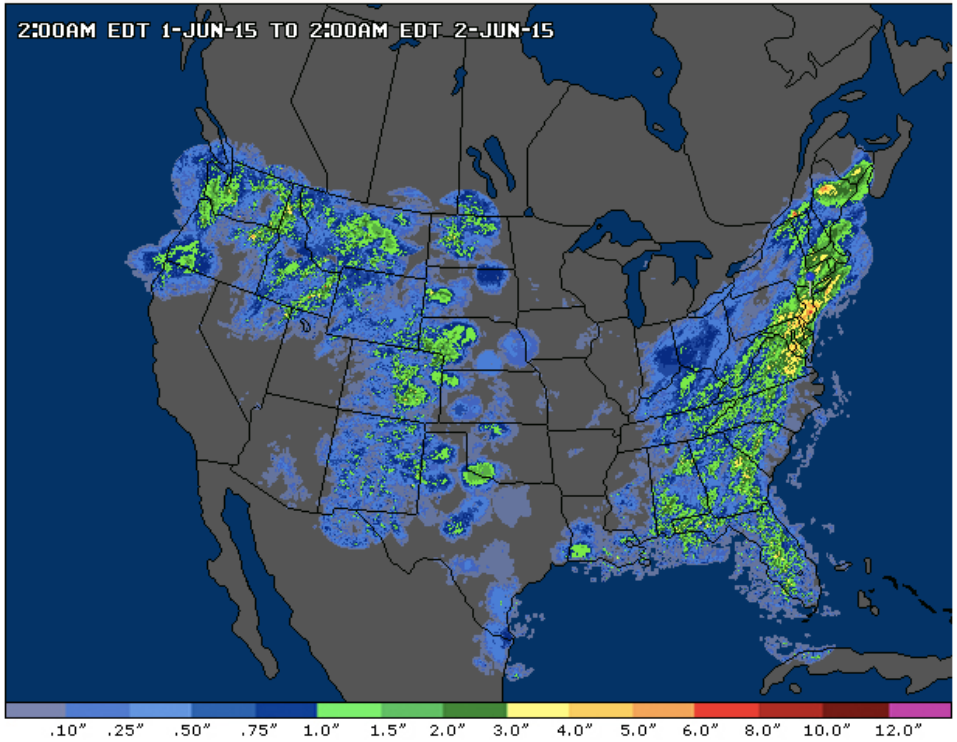
- Here, Annual Maximum of Daily Precipitation
- Parameters estimated using Maximum Likelihood:
  - Maximizes the probability of the occurrence of the fitting data
- Parameters ~ multivariate normal distribution
- Return periods can be easily derived
- Non-stationarity in parameters can be introduced, e.g.:

$$\mu = \mu_0 + \alpha t$$

where, *t*, is a time index

# Regionalization Framework

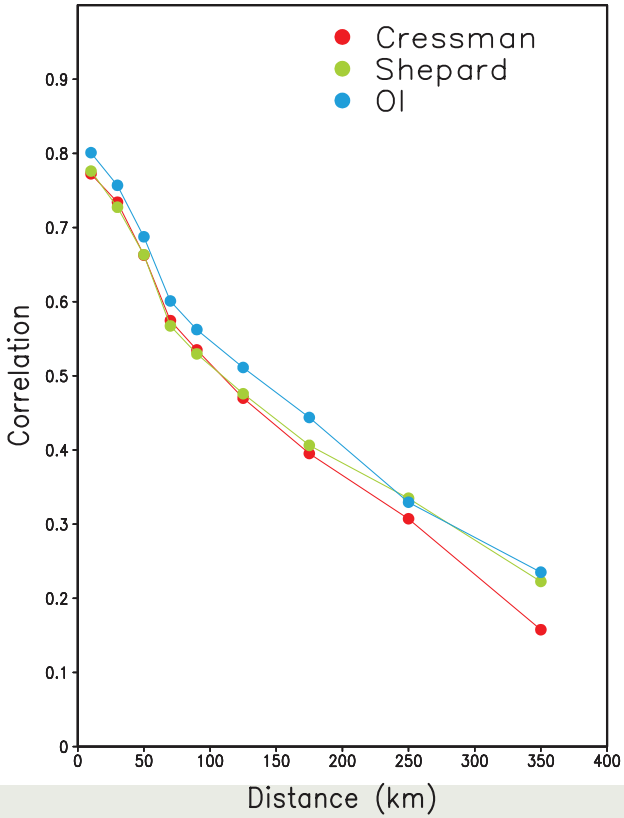
- Quantifying extremes suffers from poor sampling and small datasets.
- Data from surrounding regions can improve sampling, if they have a homogeneous climate
- Use a flexible region of influence approach, used in flood assessment models




*Daily Precipitation Radar Map for June 1 2015 (Source: Accuweather.com)*

# Regionalization Framework

- Two grid boxes have a homogeneous climate if:
  - The distance between them is within the length scale of daily precipitation (300km here)
  - They have statistically equal means
  - They exhibit statistically significant correlation in the daily anomalous precipitation time-series




*Serial Correlation in daily precipitation between a station and nearby stations. Source: Chen et al. (2007)*

# Simulations and Data

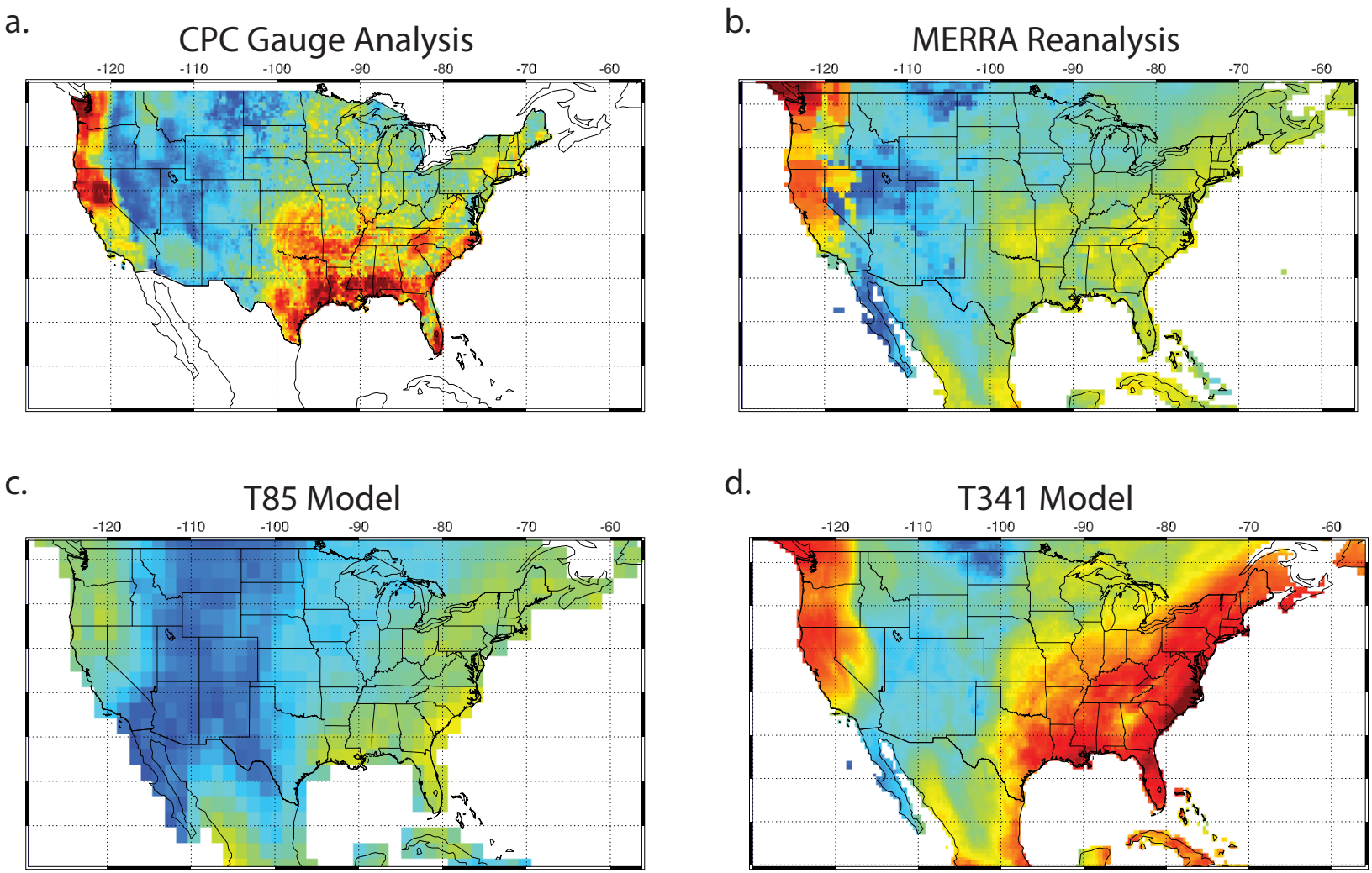
- NOAA Climate Prediction Center (CPC) Gauge-based Unified Daily Precipitation Data:
  - Optimally Interpolated to 0.5° resolution globally (0.25° over US).
- NASA MERRA Reanalysis: 0.5°x0.67°
- DOE/NCAR Community Earth System Model (CESM1):
  - Community Atmosphere Model (CAM4)
  - Low resolution: T85 (~1°)
  - High resolution: T341 (~0.25°)
  - 5 member Ensemble for the period 1979-2005
  - Forced with observed ocean and sea-ice conditions
  - Forced with observed greenhouse gases, aerosols, ozone and solar forcings.

# Workflow

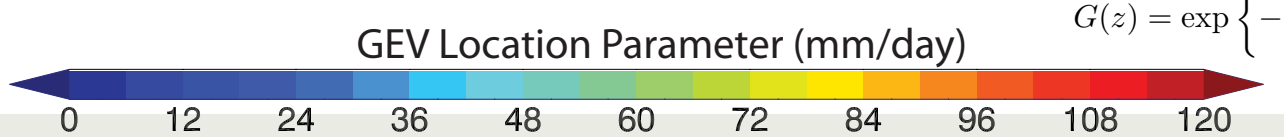
- Establish Region of Influence (ROI) for each grid point
  - Compare means
  - Correlations
- Pool data from ROI
- Compute Annual Max. from the data pool
- Fit GEV to the the annual max. values for each ensemble member for each grid point
- Compute Ensemble average of parameters and their standard errors
- Conduct Statistical tests

- T341 Data: >20 Gb for one 25-yr simulation
- Parallel Implementation: Python with mpi4py

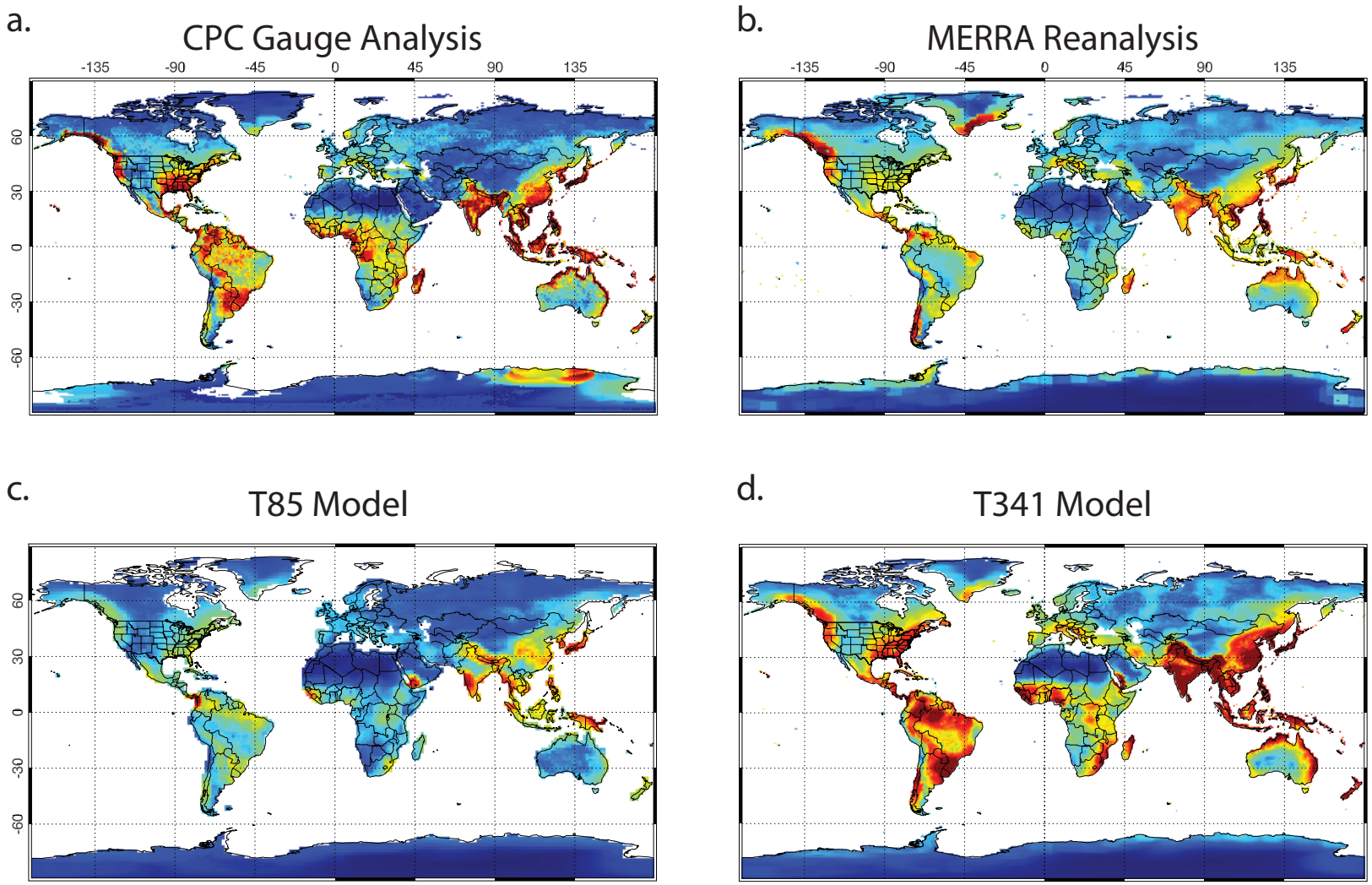
# Results: Stationary Extremes




$$G(z) = \exp\left\{-[1+\xi(\frac{z-\mu}{\sigma})]^{-1/\xi}\right\}$$



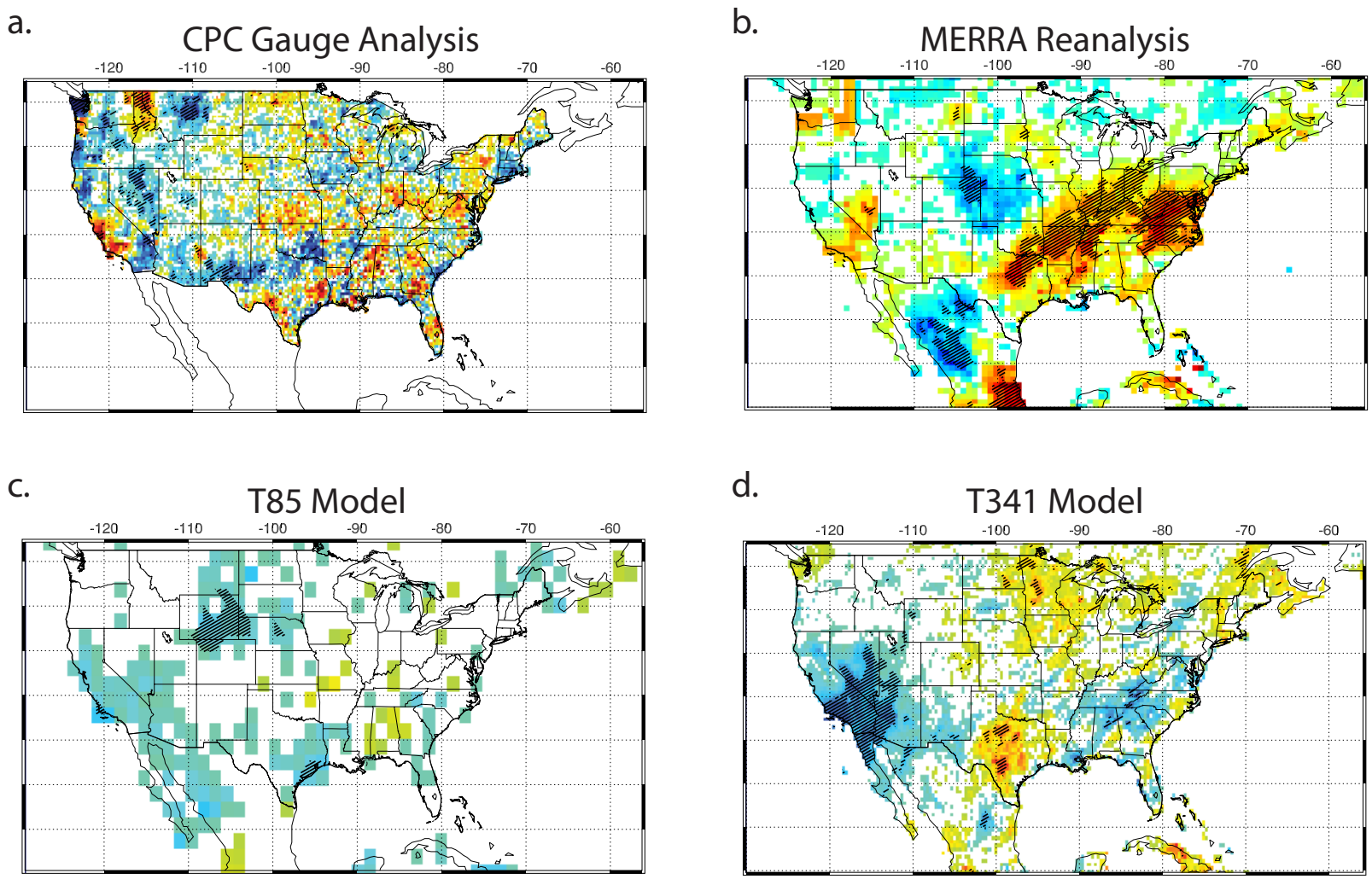
# Results: Stationary Extremes



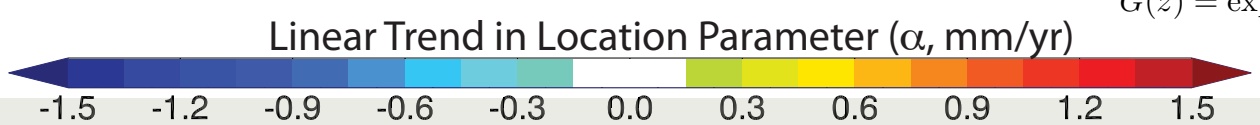

$$G(z) = \exp\left\{-[1+\xi(\frac{z-\mu}{\sigma})]^{-1/\xi}\right\}$$

GEV Location Parameter (mm/day)

# Results: Non-Stationary Extremes

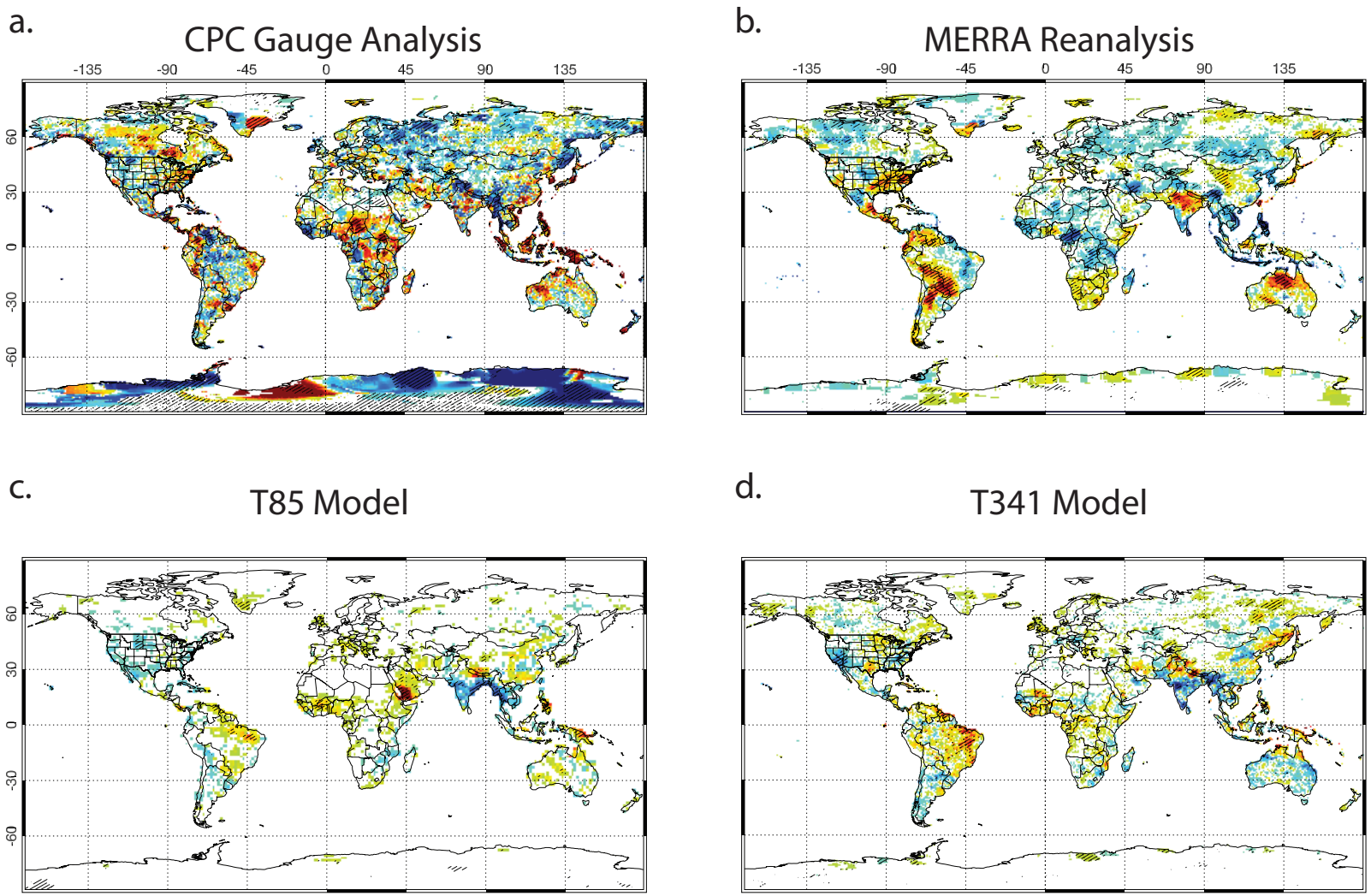



$$G(z) = \exp\left\{-[1+\xi(\frac{z-\mu}{\sigma})]^{-1/\xi}\right\}$$

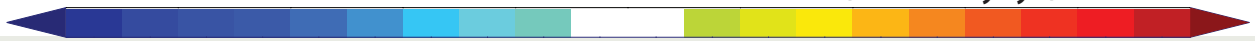



$$\mu = \mu_0 + \alpha t$$

# Results: Non-Stationary Extremes



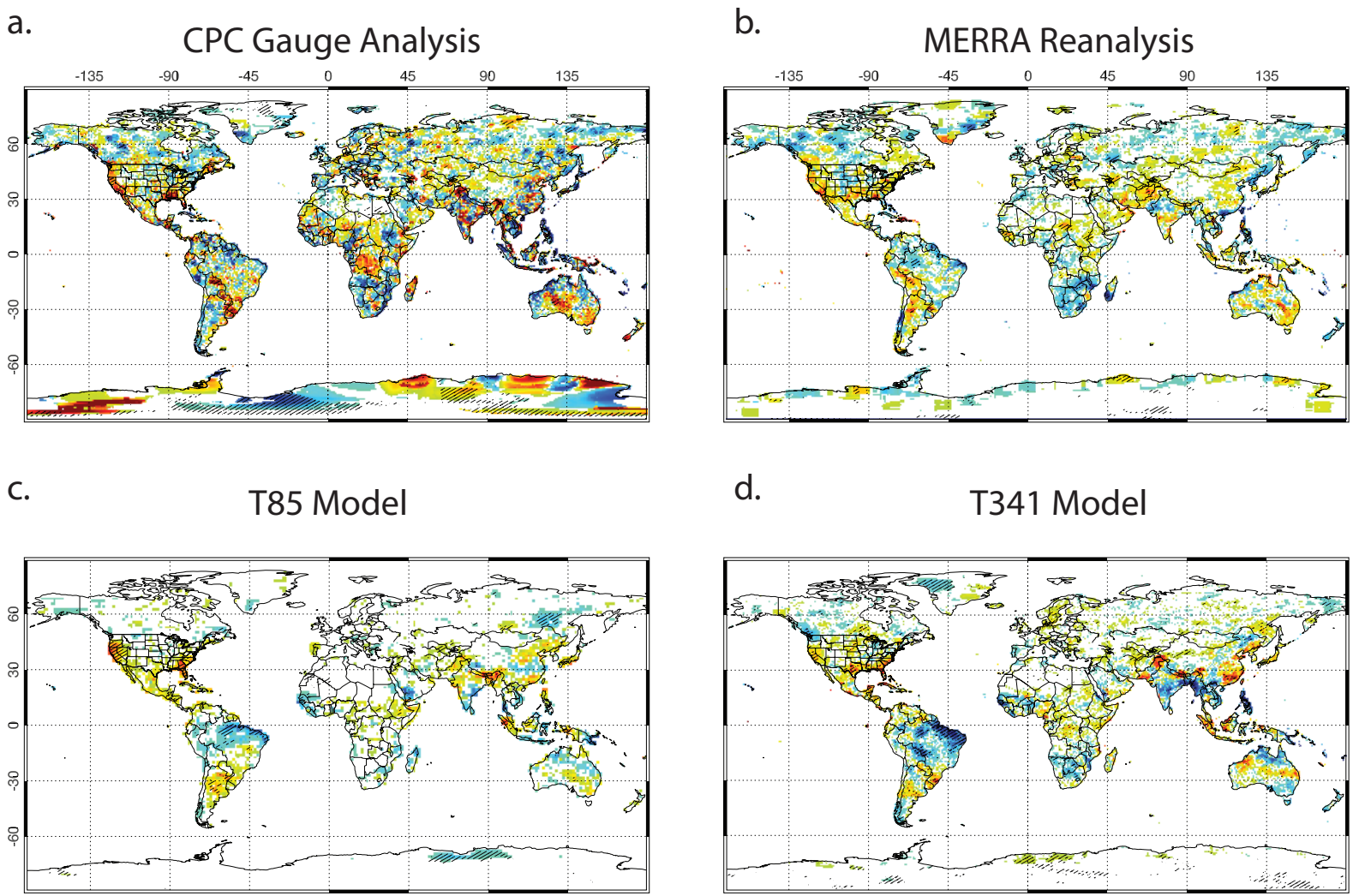

GEV Location Parameter Time Trend (mm/day/yr)




$$G(z) = \exp\left\{-[1+\xi(\frac{z-\mu}{\sigma})]^{-1/\xi}\right\}$$

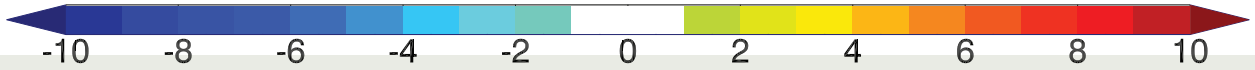
$$\mu = \mu_0 + \alpha t$$

# Results: Tele-connections (ENSO)




$$G(z) = \exp\left\{-[1+\xi(\frac{z-\mu}{\sigma})]^{-1/\xi}\right\}$$

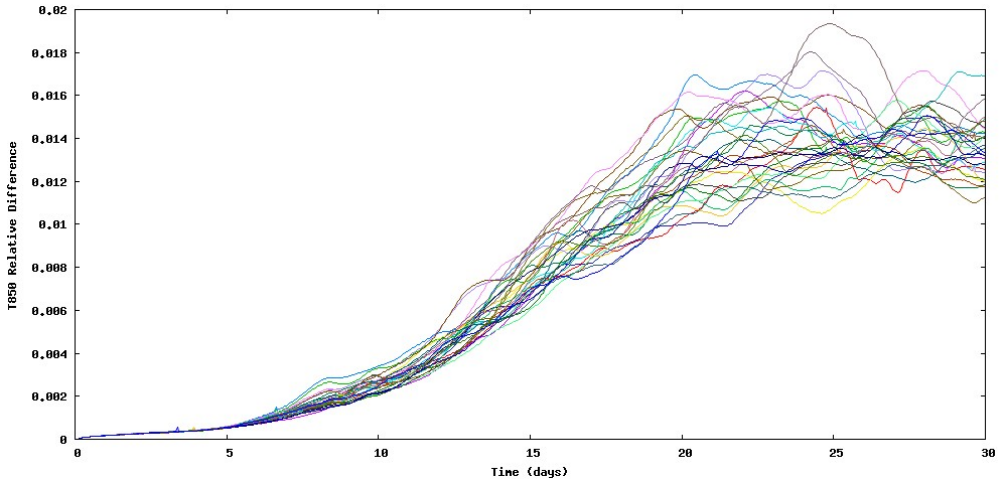
GEV Location Parameter ENSO component (mm/day/K)




$$\mu = \mu_0 + \alpha t$$

# Application: Ensemble Testing

- Ensembles provide better throughput on HPC machines.
- Traditionally, long simulations (>50 yrs) are integrated to quantify atmospheric noise (natural variability).
- Atmosphere – time scales of a few weeks.
- Is a long simulation equivalent to an ensemble of several short runs?
- Perturb atmospheric initial conditions within numerical error
- Go beyond climate means for evaluation: Extremes.



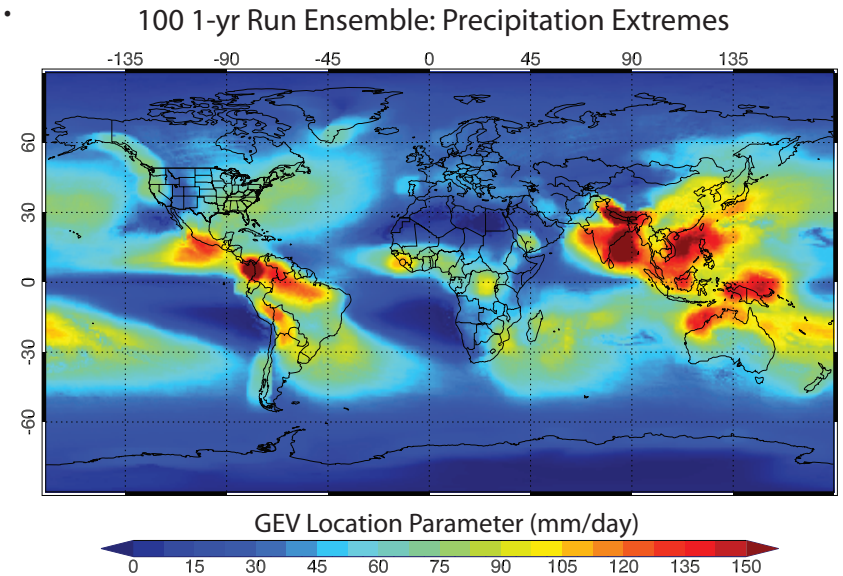

*Evolution of Temperature (Courtesy: Matt Norman)*

# Application: Ensemble Testing

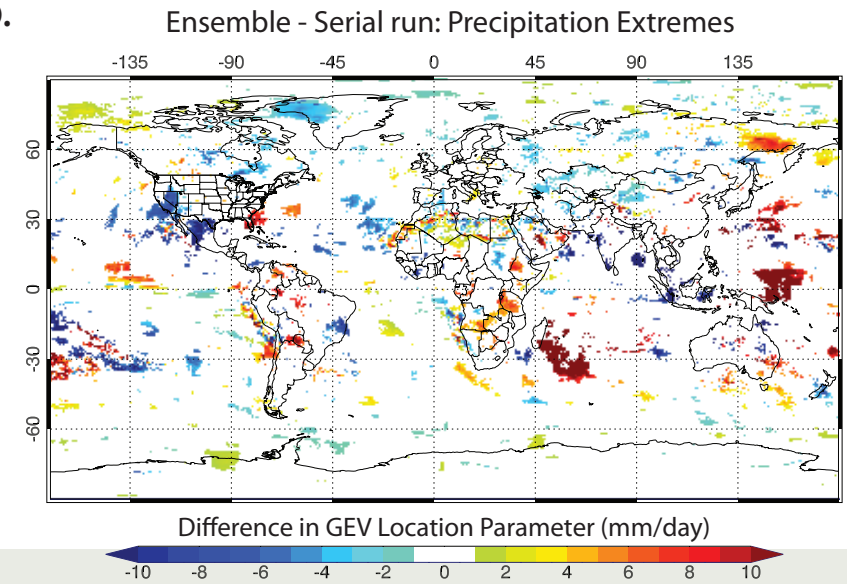
- Short run ensembles do not exhibit the variability of precipitation extremes in the long run
  - 11% grid points differ
  - Should be about 5%
    - 95% confidence intervals
    - Also from bootstrap (conf. interval: 3.5-9.5%)

- Hypothesis: Ensembles not really independent - Land initial conditions add memory.

a.




b.



# Summary

- High-resolution models better simulate stationary precipitation extremes.
- Developed a regionalization framework for evaluating climate model simulations of climate extremes.
- Parallel implementation for large high-resolution data.
- Observational data do not display non-stationarity (time trends) in precipitation extremes over the past 25 years.
- Models replicate the observed lack of non-stationarity.
- Future work: include other indices to evaluate dependence on other climate phenomenon like NAO.

# Acknowledgements


- Accelerated Climate Modeling Project (DOE BER)
- ORNL Leadership Computing Facility (OLCF)
- mpi4py module in Python by Lisandro Dalcin
- ‘ismev’ and ‘evd’ R package by Alec Stephenson

# Questions?

- Thanks!