# Statistics, Numerical Models and Ensembles

Douglas Nychka,
*Reinhard Furrer, , Dan Cooley  Claudia Tebaldi,*
Linda Mearns, Jerry Meehl and Richard Smith (UNC).

- Spatial prediction and data assimilation
- Precipitation extremes
- Combining climate models


NCAR

*Supported by the National Science Foundation DMS*

# Statistical Science

What do you want to know? e.g. $\theta$

# Statistical Science

What do you want to know? e.g. $\theta$

What have you measured? e.g $\boldsymbol{Y}$

# Statistical Science

What do you want to know? e.g. $\theta$

What have you measured? e.g $\boldsymbol{Y}$

Relate them using a probability distribution.

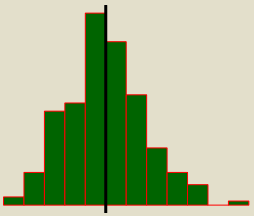
# Statistical Science

What do you want to know? e.g. $\theta$

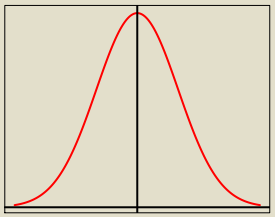
What have you measured? e.g $\boldsymbol{Y}$

Relate them using a probability distribution.

Data = parameter + error




$= \theta +$



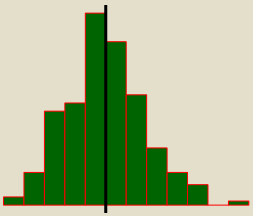
# Statistical Science

What do you want to know? e.g. $\theta$

What have you measured? e.g $\boldsymbol{Y}$

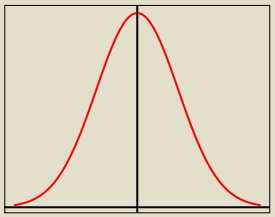
Relate them using a probability distribution.

Data = parameter + error



θ

$= \theta +$



0

*Characterize reasonable values for $\theta$ given the data*

# The statistical method

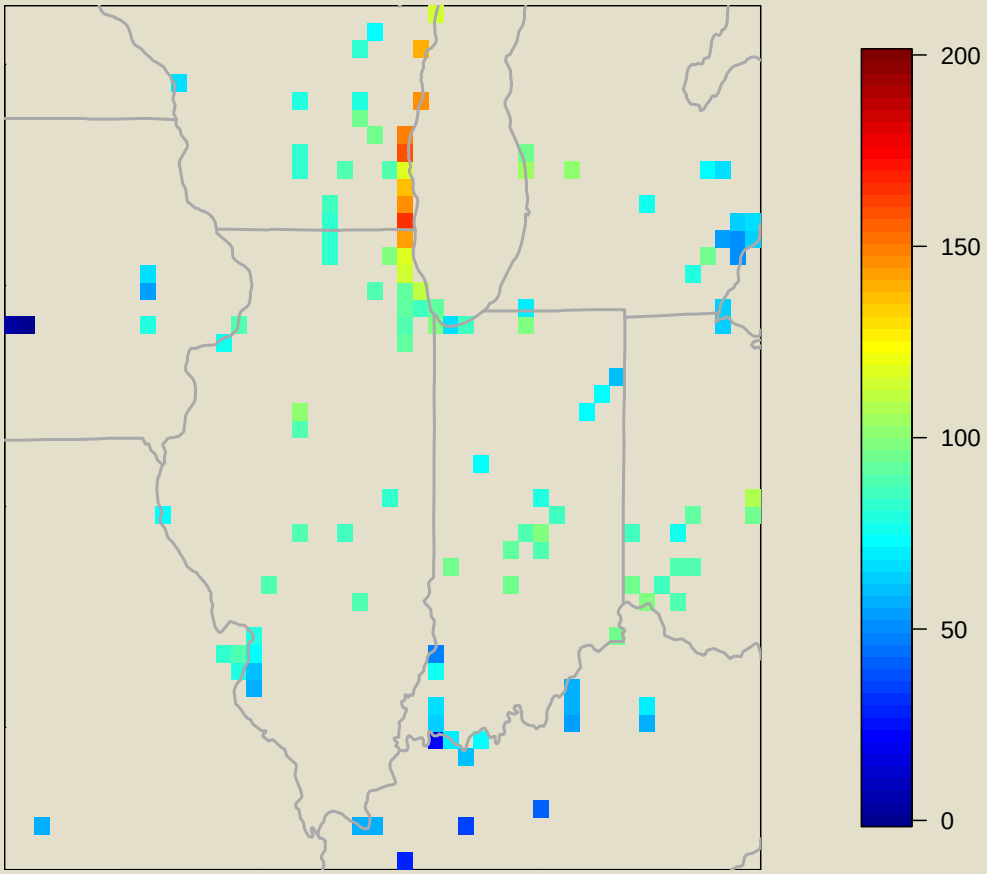
- Use Bayesian models and Monte Carlo methods to generate a statistical ensemble for $\theta$ .
- The ensemble mean is a good estimate for $\theta$.
- The spread is a good measure of uncertainty for $\theta$.

# *Part 1: Observations are in the wrong place!*

# Spatial Prediction

Predict surface ozone where it is not monitored.

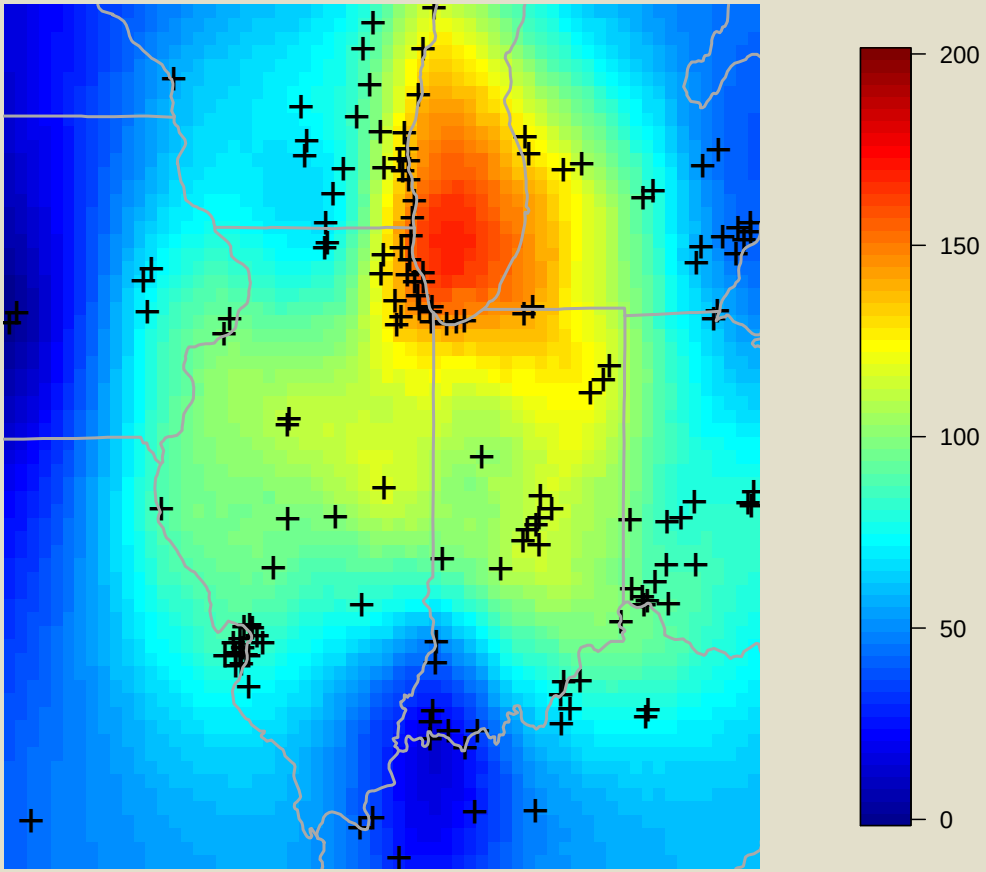
Ambient daily ozone in PPB June 16, 1987, US Midwestern Region.



# Spatial Prediction

## Typical Spatial Interpolation

Ambient daily ozone in PPB June 16, 1987, US Midwestern Region. Best estimate of surface (Kriging, Bayes, OI,, BLUE)



# An ensemble approach

Start with a ensemble of fields that are distributed according to ones best guess or forecast – without consulting the data.

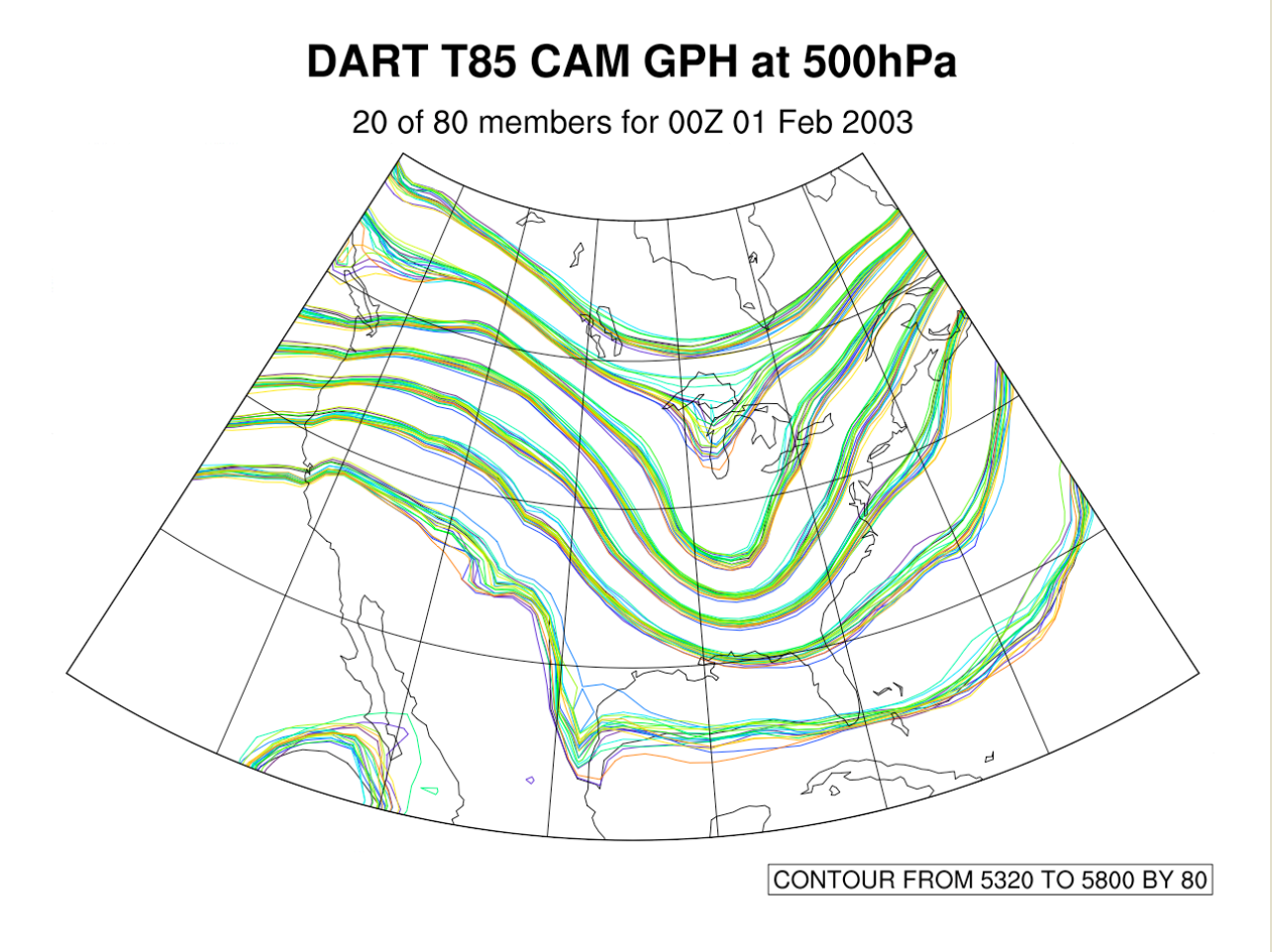
Update each member of the ensemble using the observed data.

*The sample mean and covariances among the ensemble members are used for the update calculations.*

*Animations*

# An ensemble forecast.

Updates done just as in ozone example



# *Part 2 Observations do not measure what we want!*

# Precipitation extremes

How will climate change effect extreme precipitation?

# Precipitation extremes

## How will climate change effect extreme precipitation?

Extremes in precipitation are used to determine flood potential for urban areas, for dam and roadway specifications and also have extensive ecological importance.

- How does one estimate extremes where no observations are made?
- How does one determine a possible “25 year event” from 20 years worth of data?

# Precipitation extremes

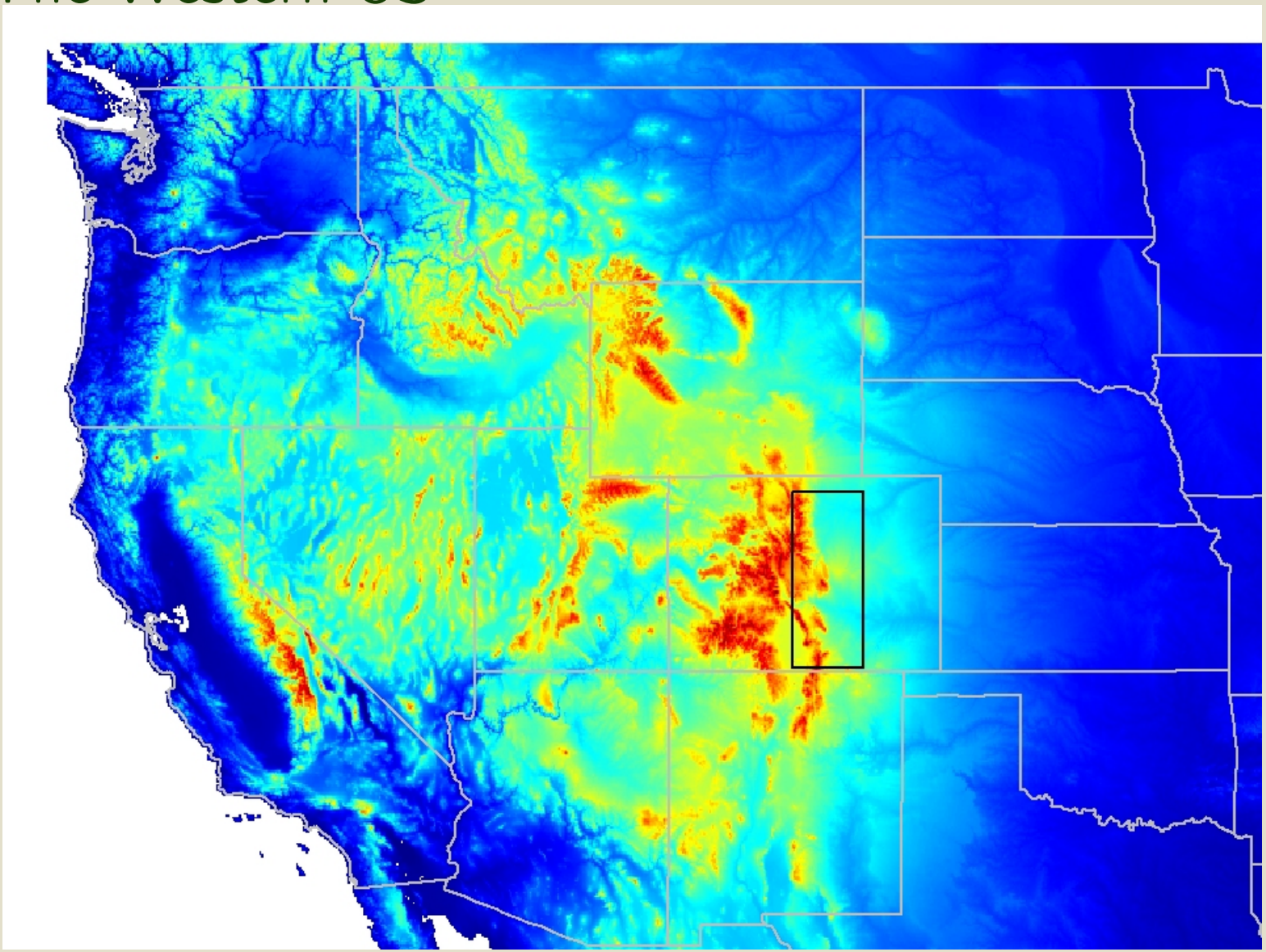
## How will climate change effect extreme precipitation?

Extremes in precipitation are used to determine flood potential for urban areas, for dam and roadway specifications and also have extensive ecological importance.

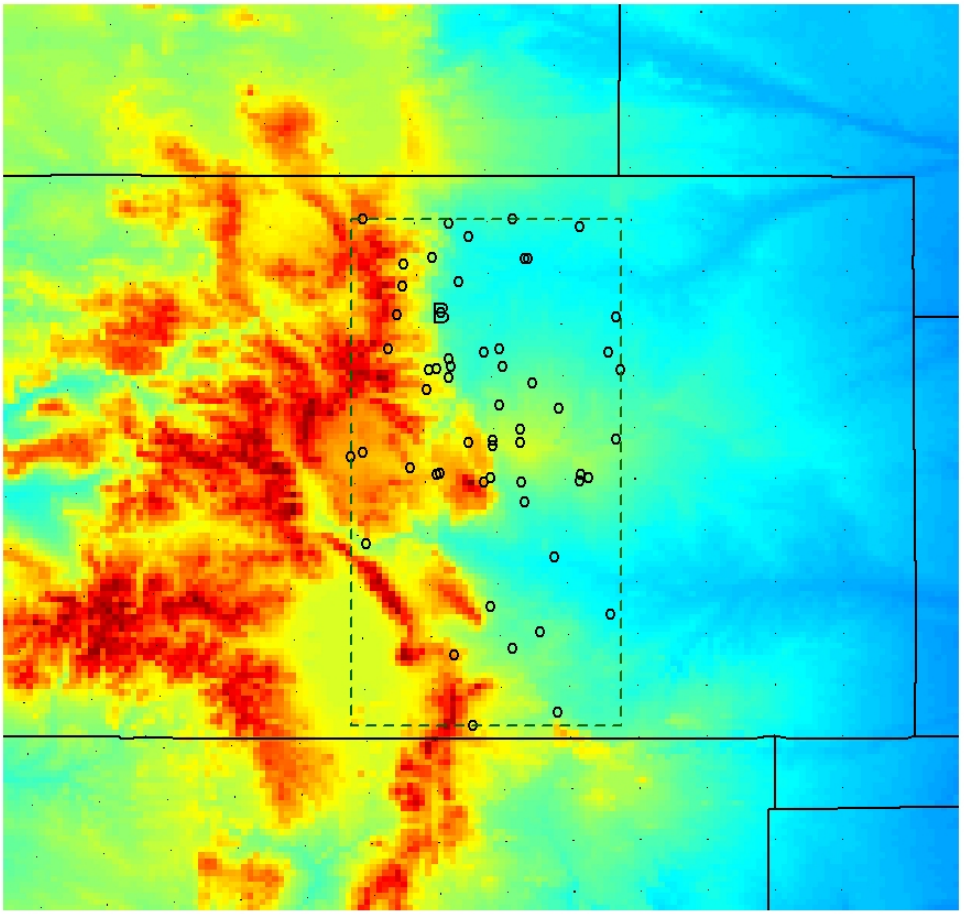
- How does one estimate extremes where no observations are made?
- How does one determine a possible "25 year event" from 20 years worth of data?

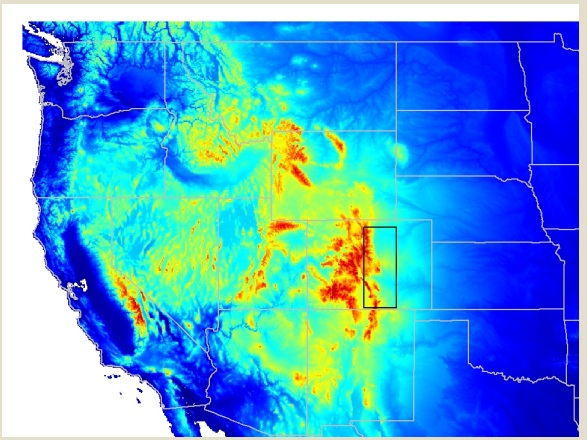
Typically extremes are described by the return period: "A 25 year event" = probability of seeing this value (or higher) in a given year is 1/25 or 4%

# The Western US

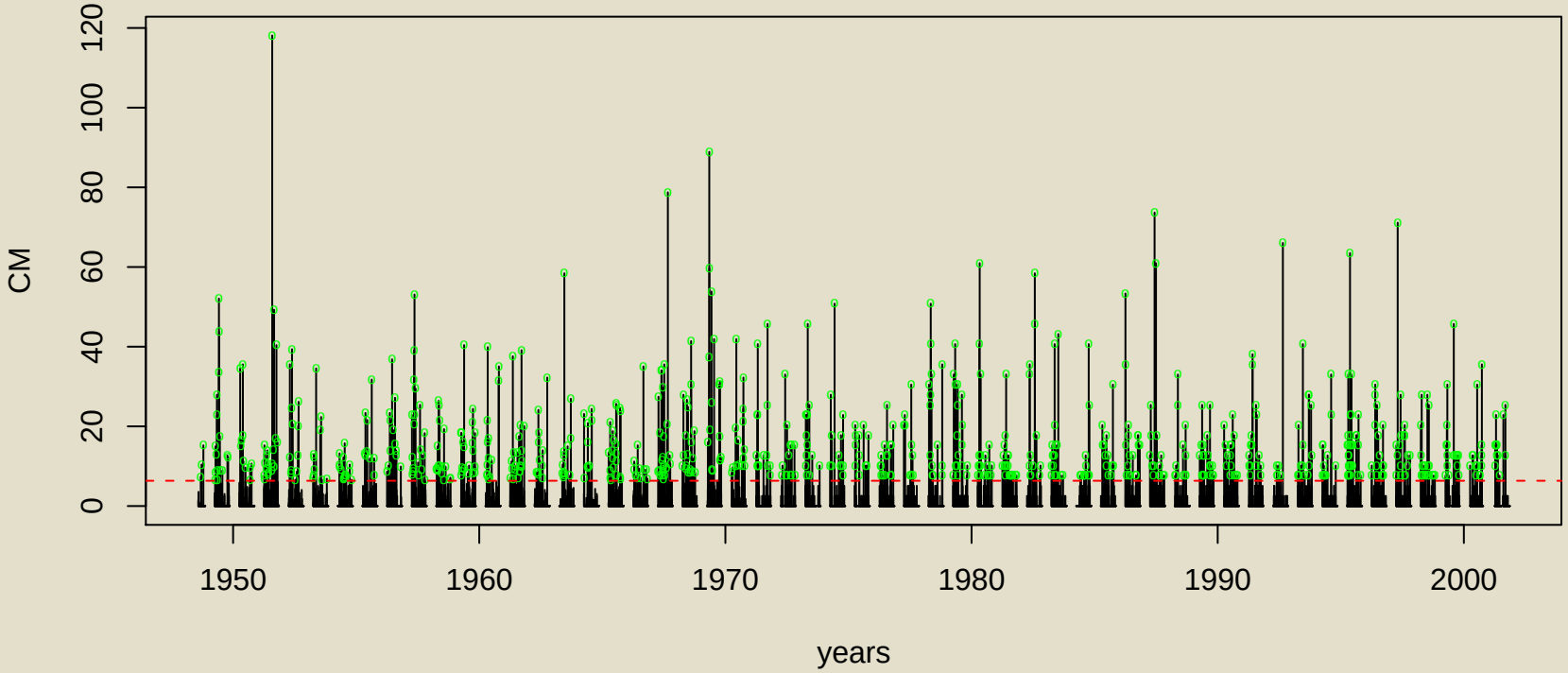


# Colorado Front Range

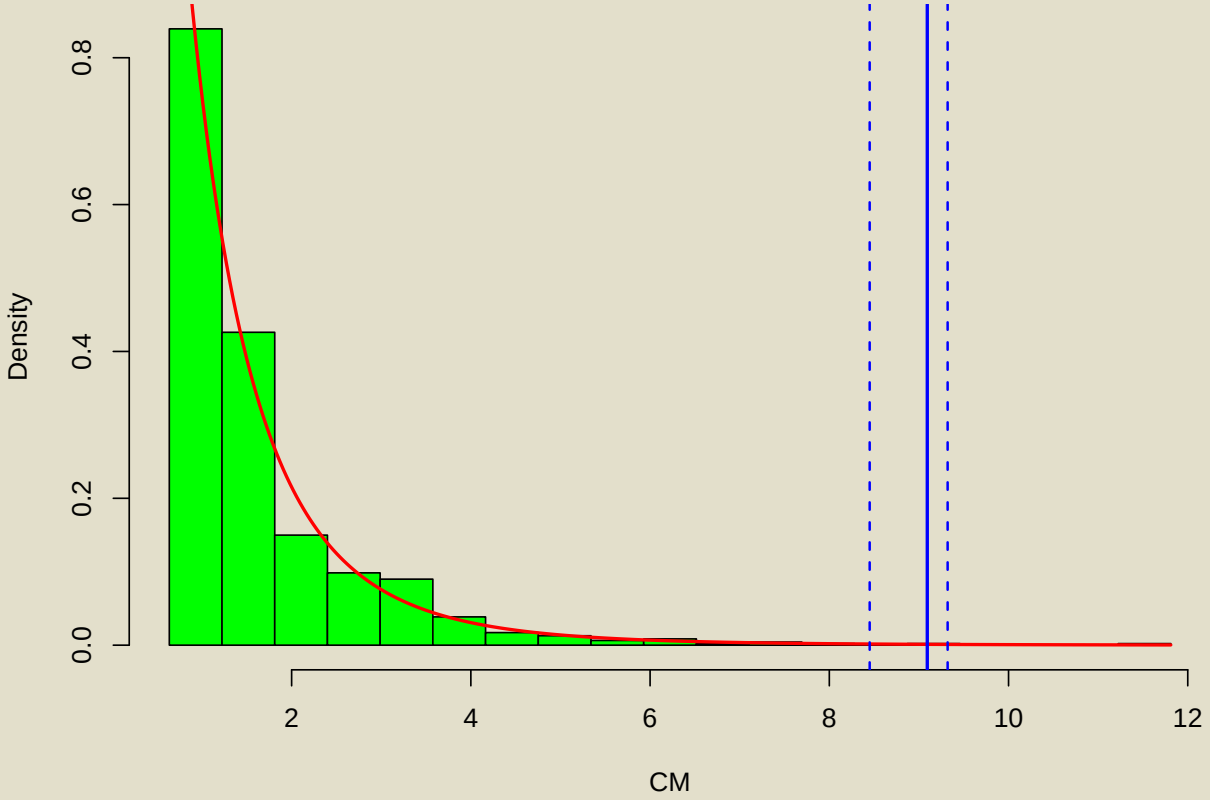




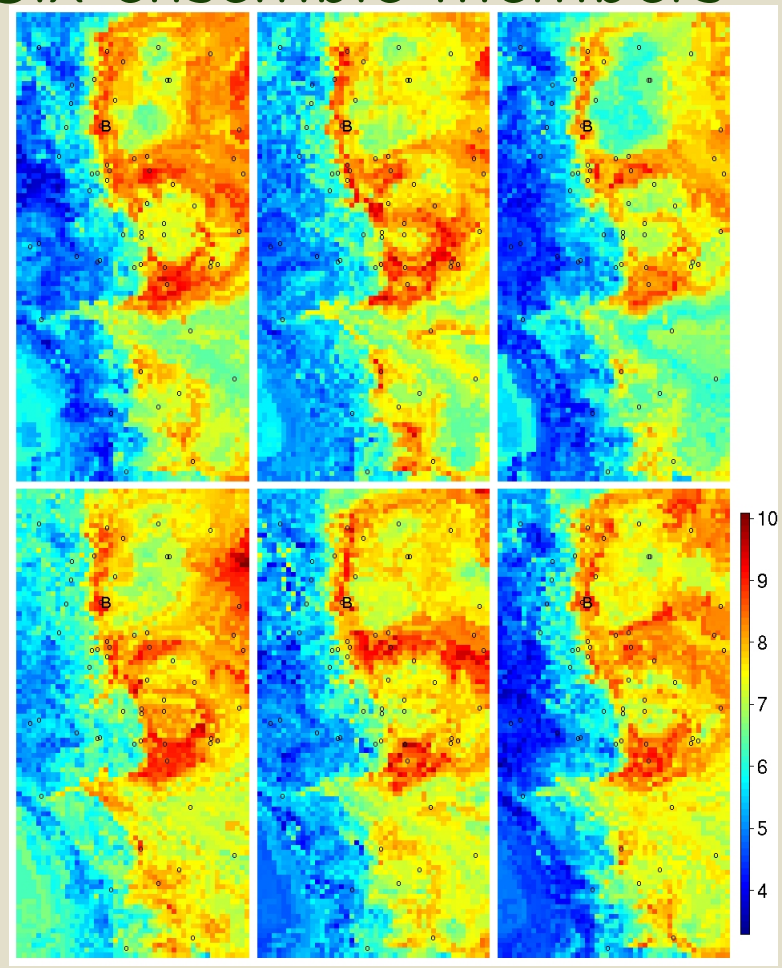
# Observed daily precipitation for Boulder



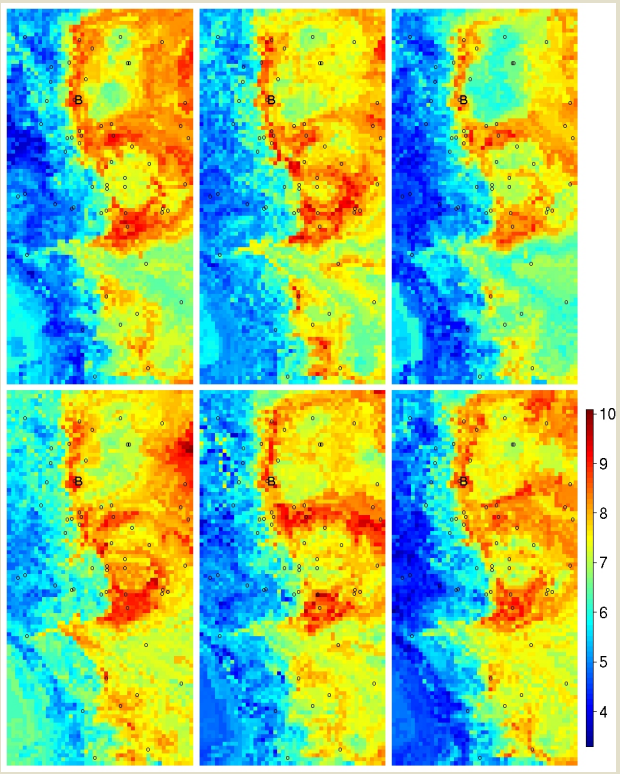
# Distribution fit to exceedances

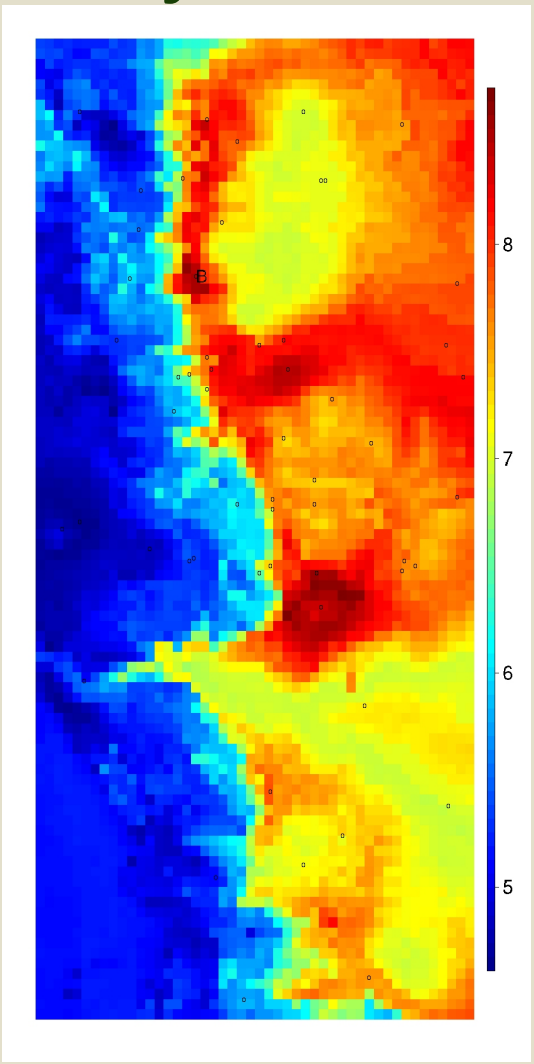


# Six ensemble members



# Ensemble mean of the 25 year return level



# *Part 3 Not sure what you have observed!*

# Data and the IPCC

## What will the climate be like in 2100?

How much data do we have to answer this question.

# Data and the IPCC

## What will the climate be like in 2100?

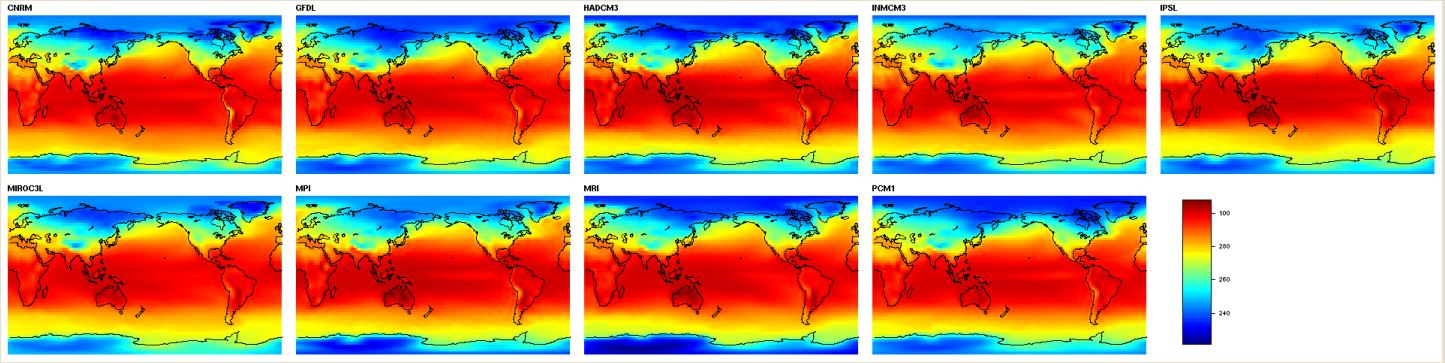
How much data do we have to answer this question.

The most recent experiments to support the fourth report of the International Panel on Climate Change amount to an archive of approximately 100 Tb.

- More than 20 different climate models/modeling centers represented.
- Several different future scenarios.
- Multiply responses e.g. temperature, precipitation,

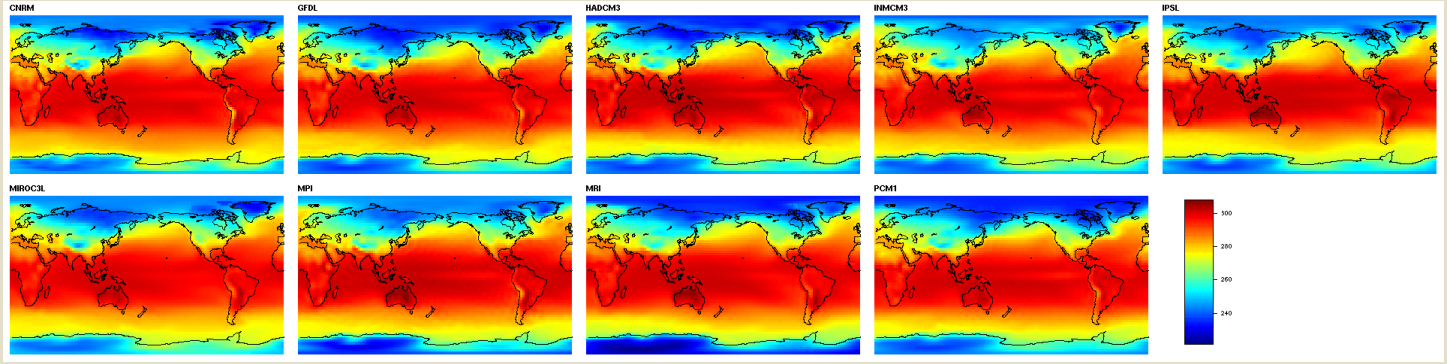
# Some Data

## Present winter temperatures, 9 AOGCMs

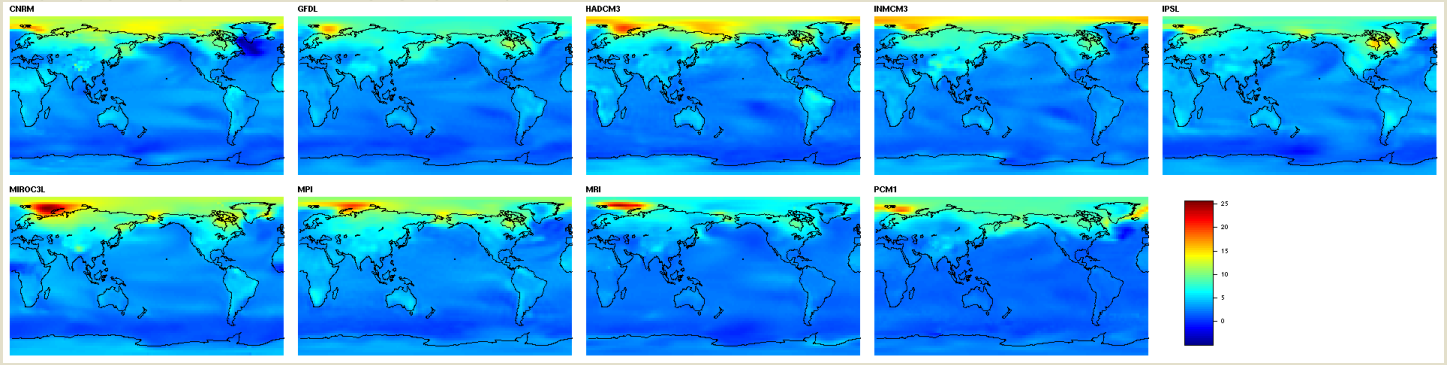


# Some Data

## Present winter temperatures, 9 AOGCMs

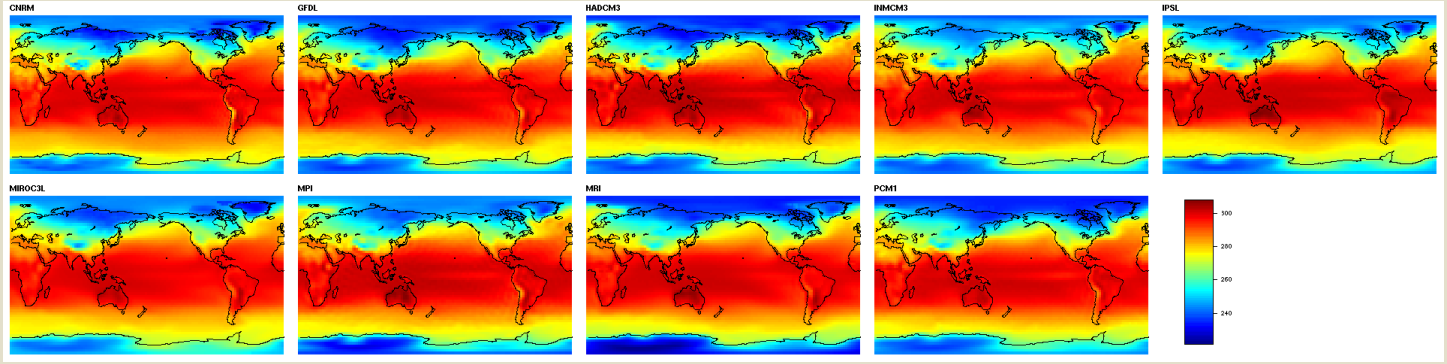



## Future - Present

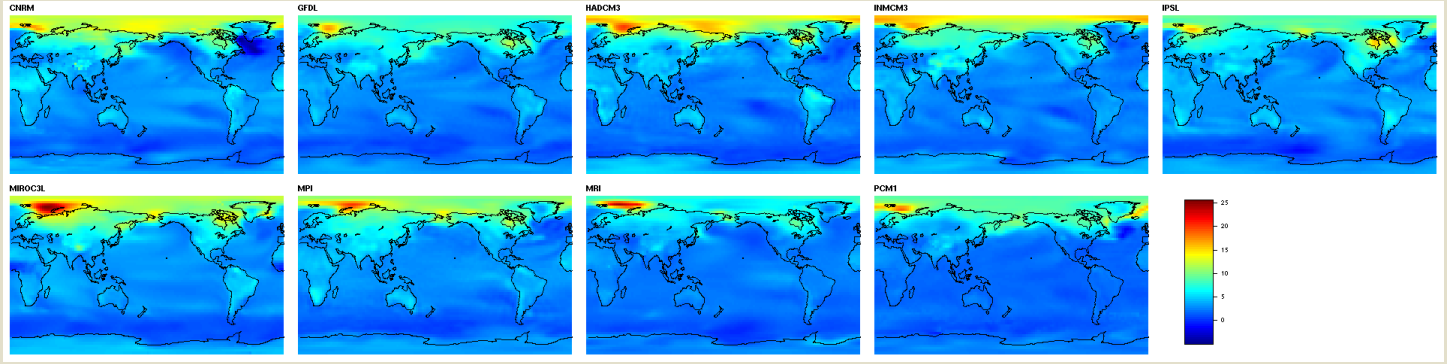


# Some Data

## Present winter temperatures, 9 AOGCMs

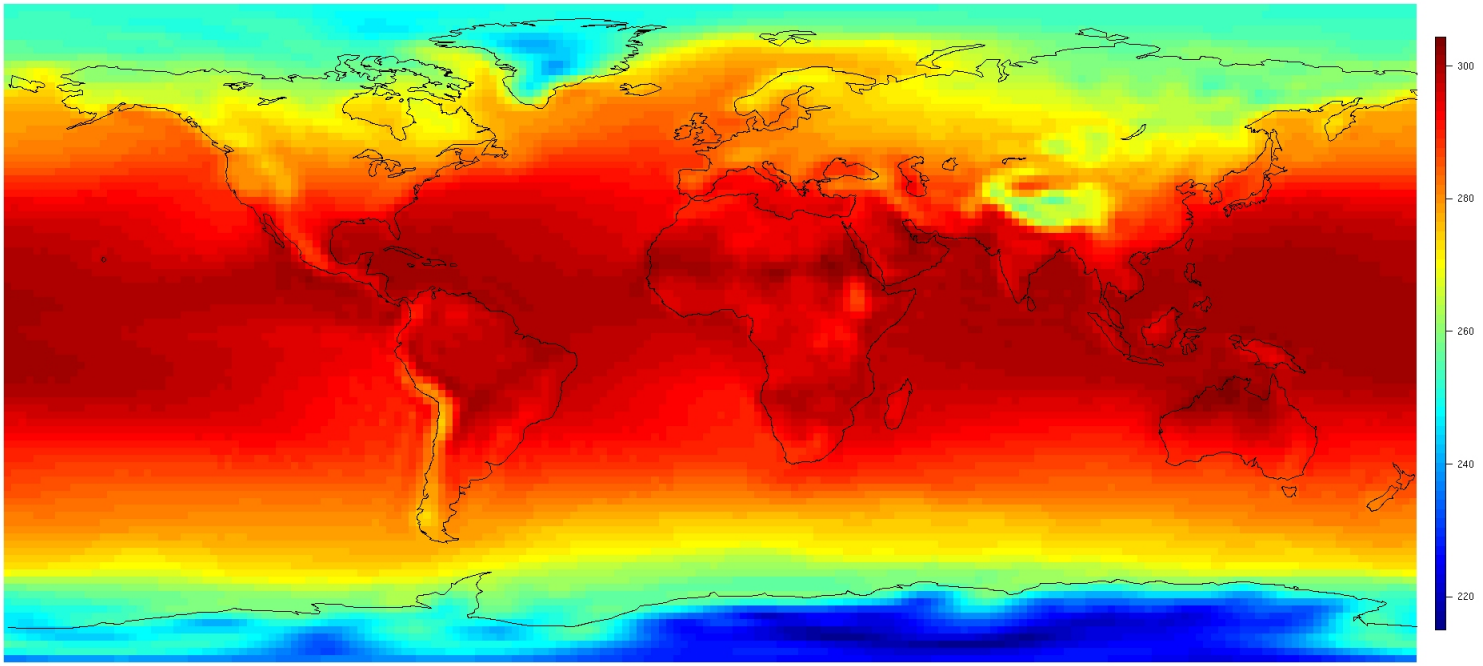



## Future - Present

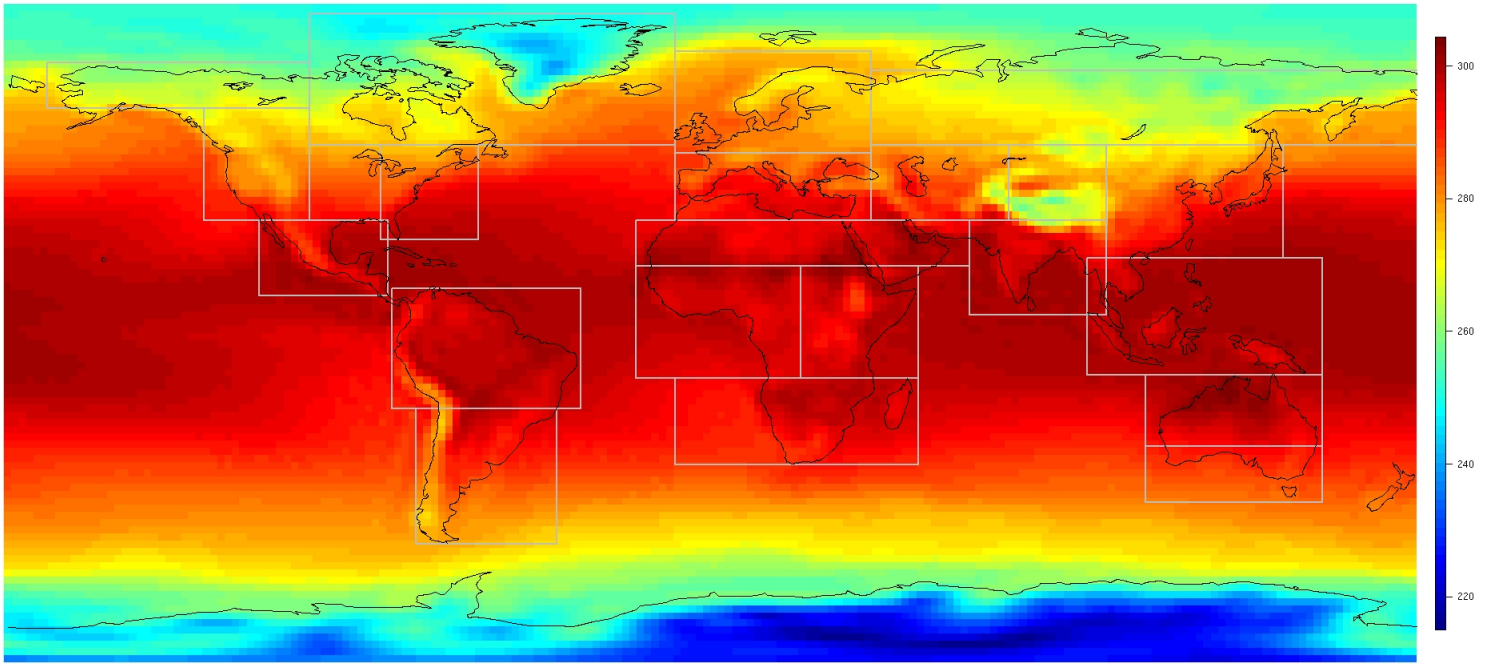



*Is this really a massive data set?*

# NCEP Northern Hemisphere Winter

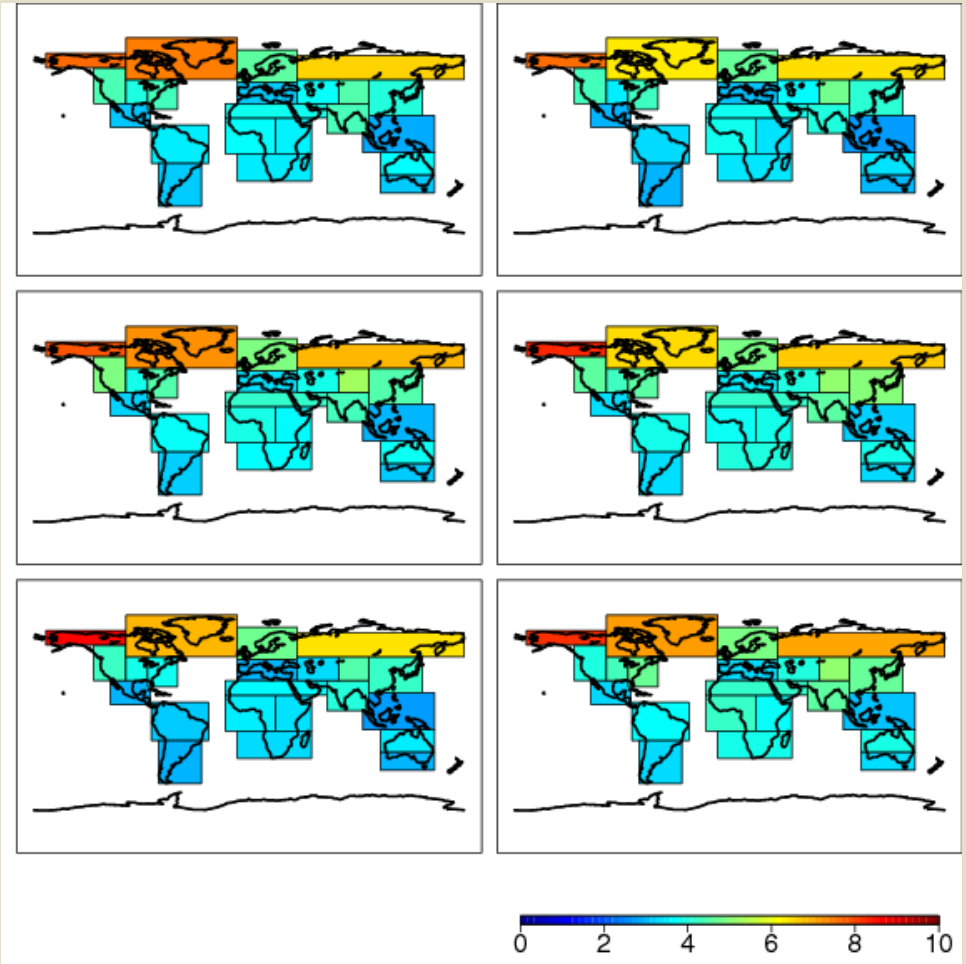


# Standard IPCC regions

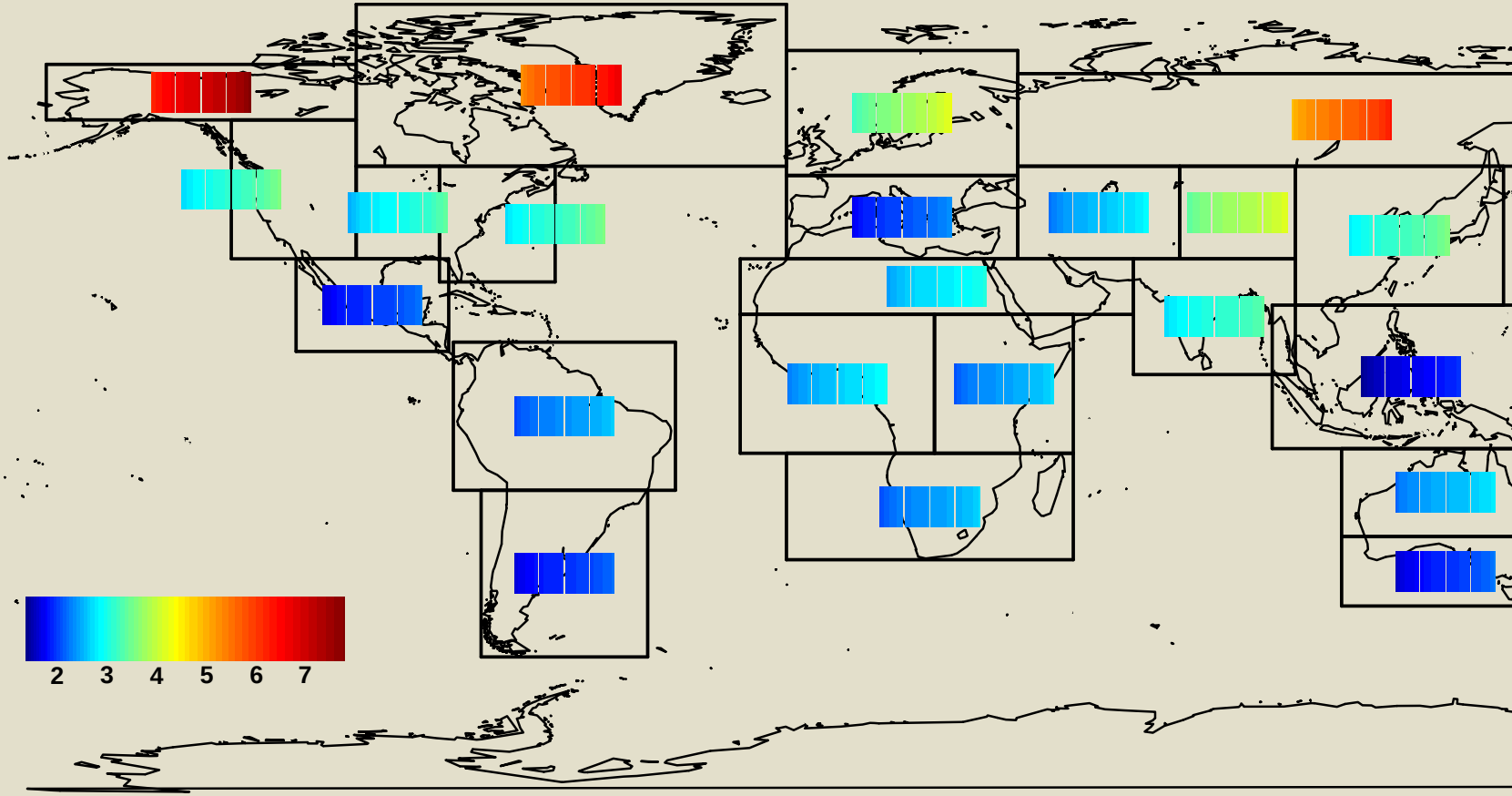


# Some ensemble members for the regional change

*Future – Present DJF temperature (A2)*



# Ensemble distributions: A summary for the AR4



# Summary

- Statistical ensembles are a useful way to estimate spatial fields and characterize uncertainty.
- Statistical methods can be used to estimate complex indirect features.
- The size of "massive" data sets may not be massive. Statistics can be used to gauge the representativeness of a sample.



...