# Climate Change: Challenges for Machine Learning

Arindam Banerjee
University of Minnesota, Twin Cities

Claire Monteleoni
George Washington University

NIPS 2014

August 2005: Hurricane Katrina – Reuters

October 2012: Hurricane Sandy – Reuters

August 2013: Rim Fire, California – Reuters

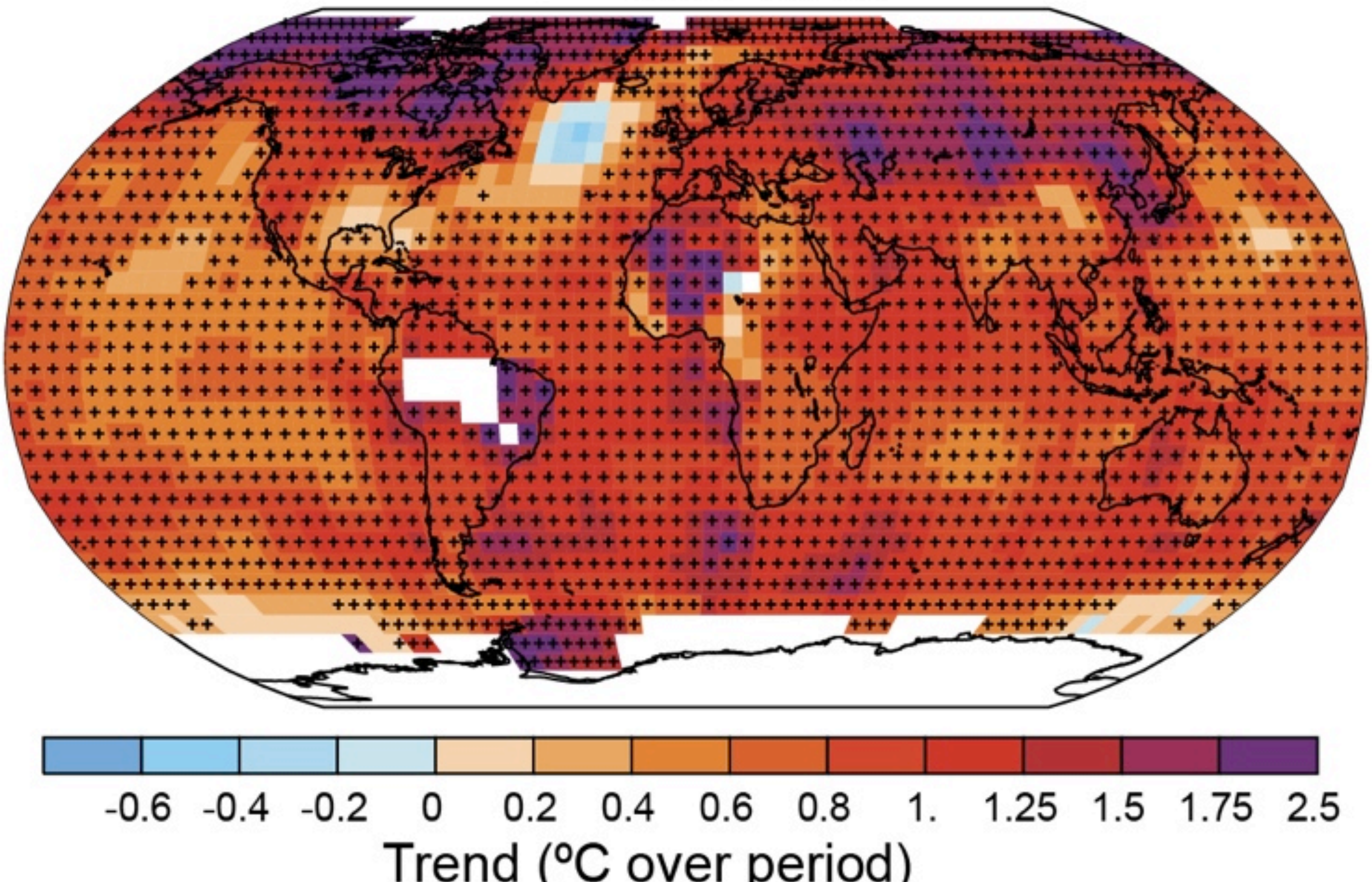
January 2014: Drought, Folsom Lake – California Department of Water Resources

Machine learning can shed light on climate change.

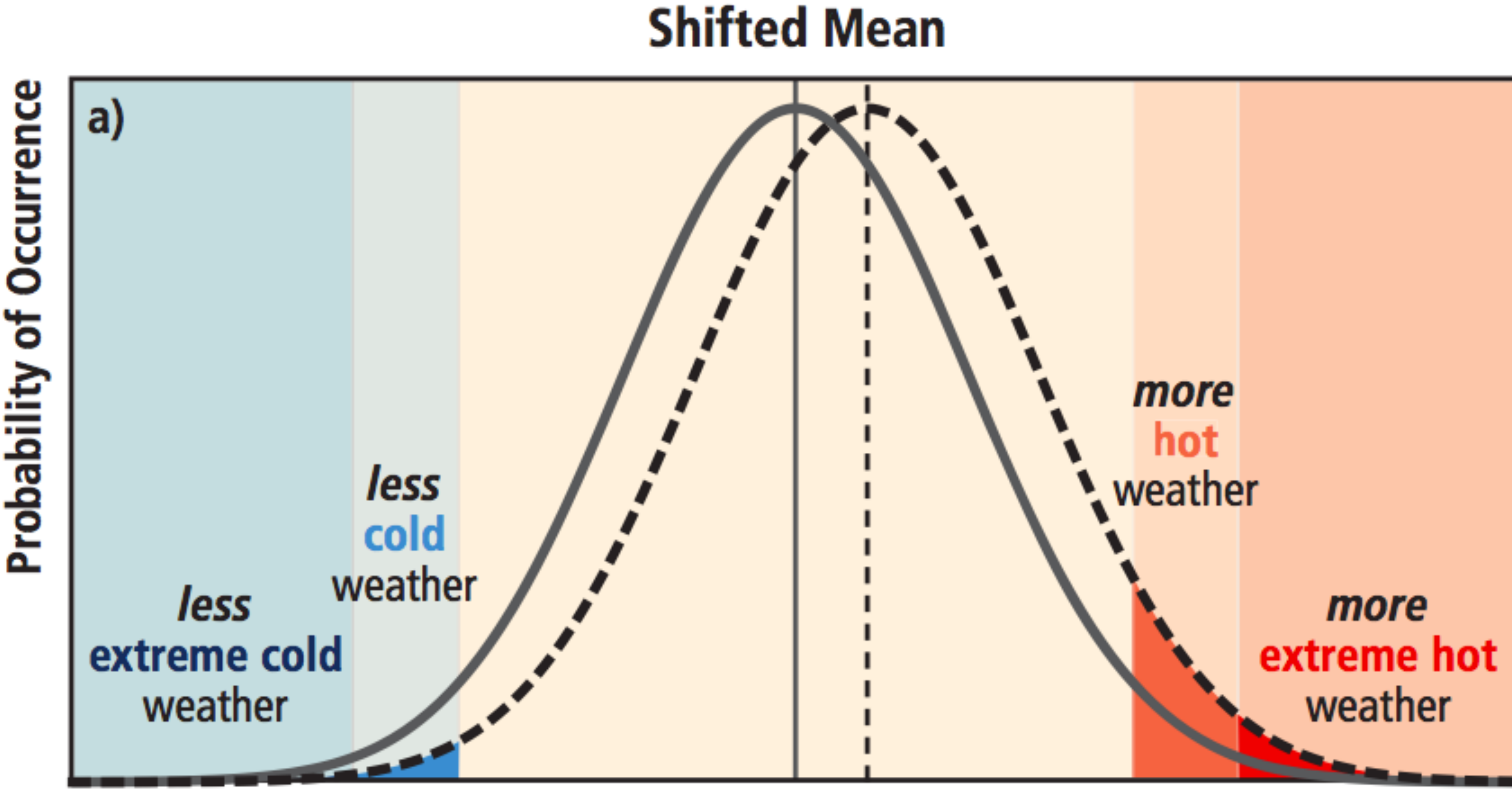
Despite the scientific consensus on climate change, drastic uncertainties remain. For instance:

How does climate change affect extreme events?

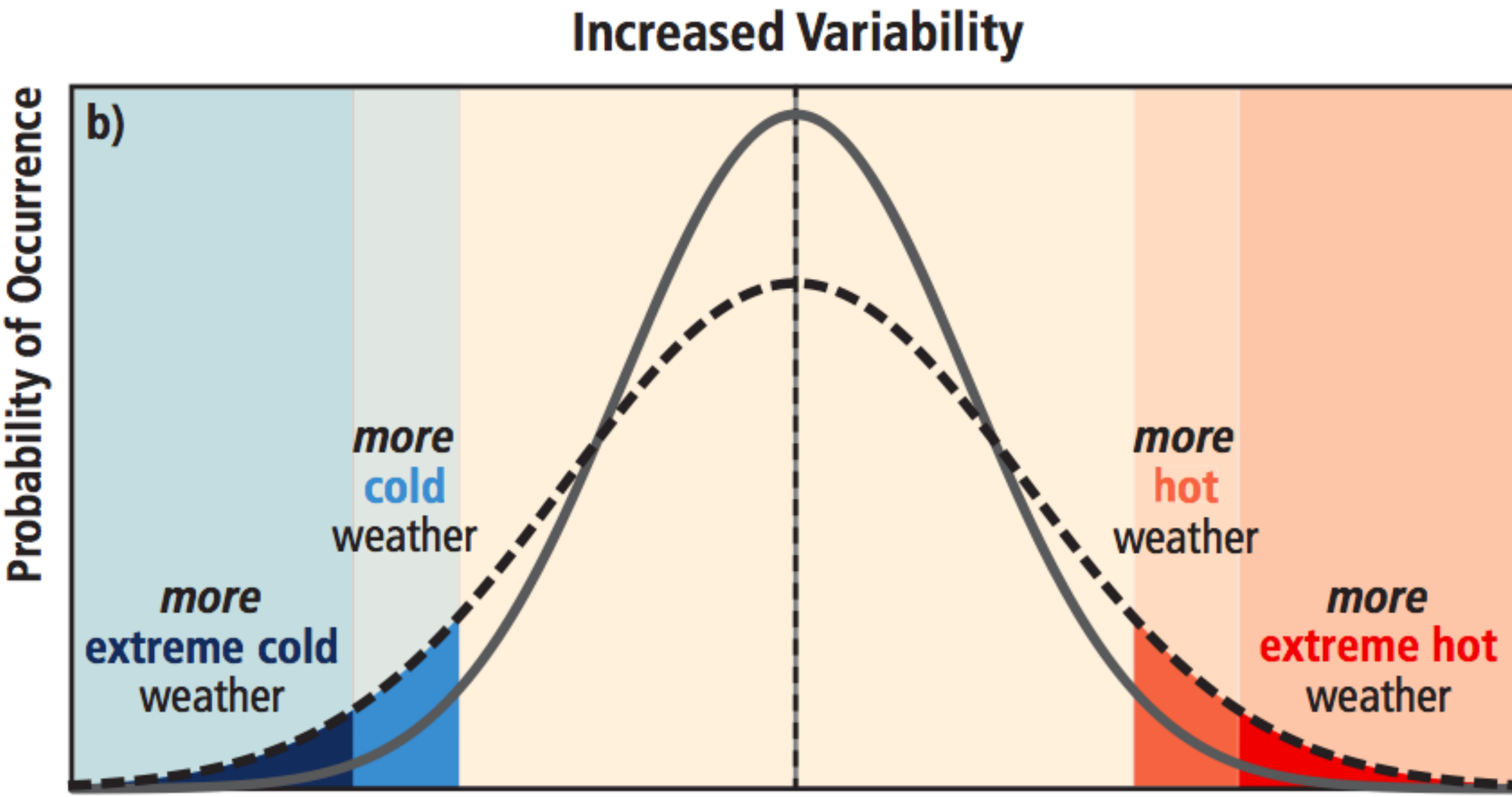
Surface Temperature 1901-2012



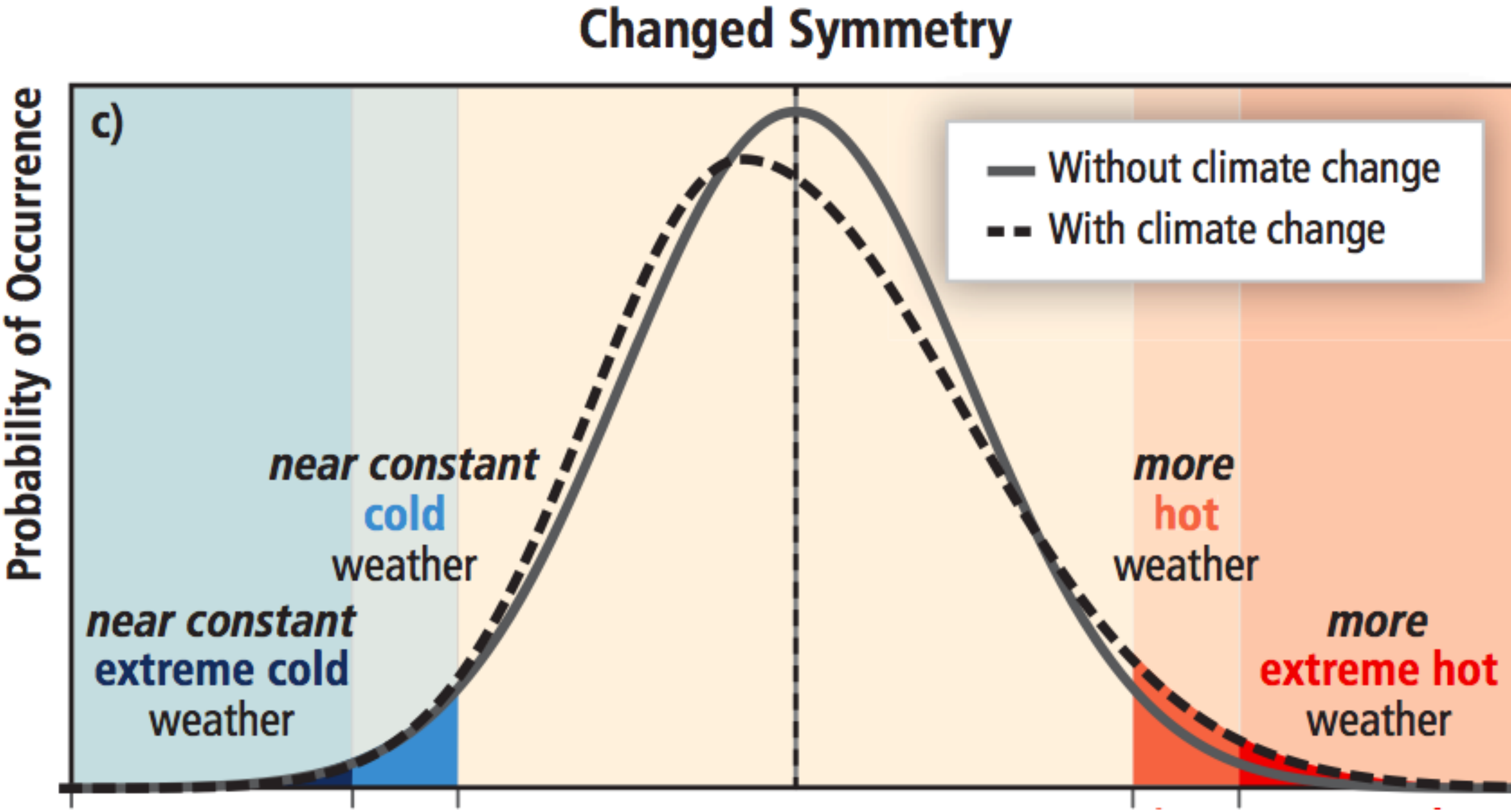

Intergovernmental Panel on Climate Change (IPCC), 2013




Intergovernmental Panel on Climate Change, 2012




Intergovernmental Panel on Climate Change, 2012




Intergovernmental Panel on Climate Change, 2012

# Uncertainty in extremes, especially regional

Warmer atmosphere can hold more water vapor
→ heavier precipitation, storms, flooding

Global warming may increase surface evaporation
→ heat waves, droughts

Possible changes in El Niño-Southern Oscillation
→ changes in floods in some regions, droughts in others

World Climate Research Programme 2013, grand challenge:
understanding and improving predictions of extreme events

Extreme events are rare by definition.

Climate change may affect their distribution.

→ Past statistics are not sufficient for future prediction.

Augment historical data with climate model simulations.

Massive, high-dimensional, big data.

That's where machine learning comes in!

# Main types of climate data

- Past: Historical data
  - Limited amounts
  - Very heterogeneous

- Present: Observation data
  - Increasingly measured. Large quantities for recent times.
  - Can be unlabeled, sparse, measured at higher resolution than relevant information

- Past, Present, Future: Climate model simulations
  - Vast, high-dimensional
  - Encodes scientific domain knowledge
  - Some information is lost in discretizations
  - Future predictions cannot be validated

# What is climate?

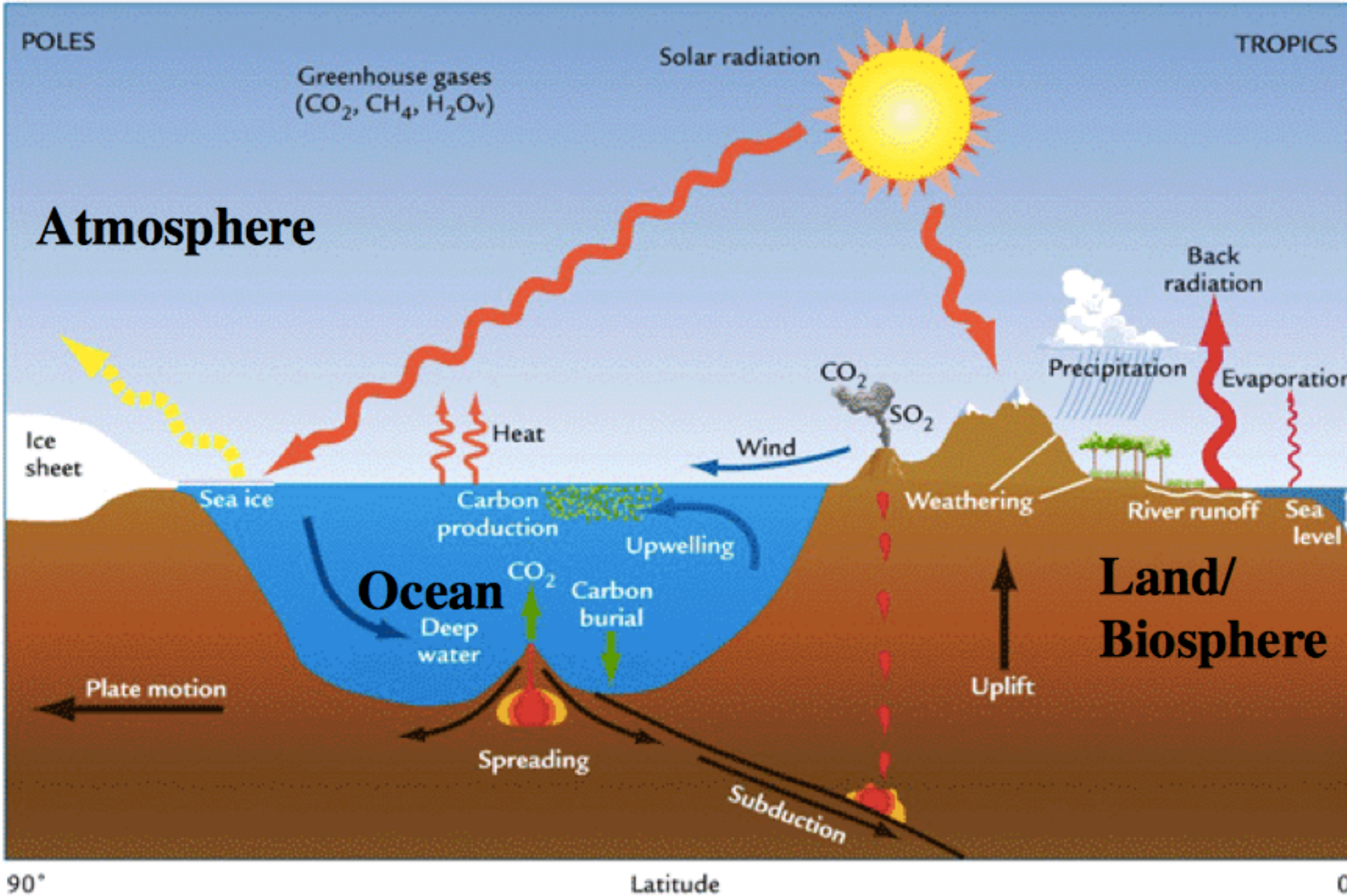
- Climate is what you expect, weather is what you get.
- Weather: A thunderstorm, or an unusually high rainfall.

- Climate: E.g. the 30 year average rainfall in a region.



Credit: D. Nychka

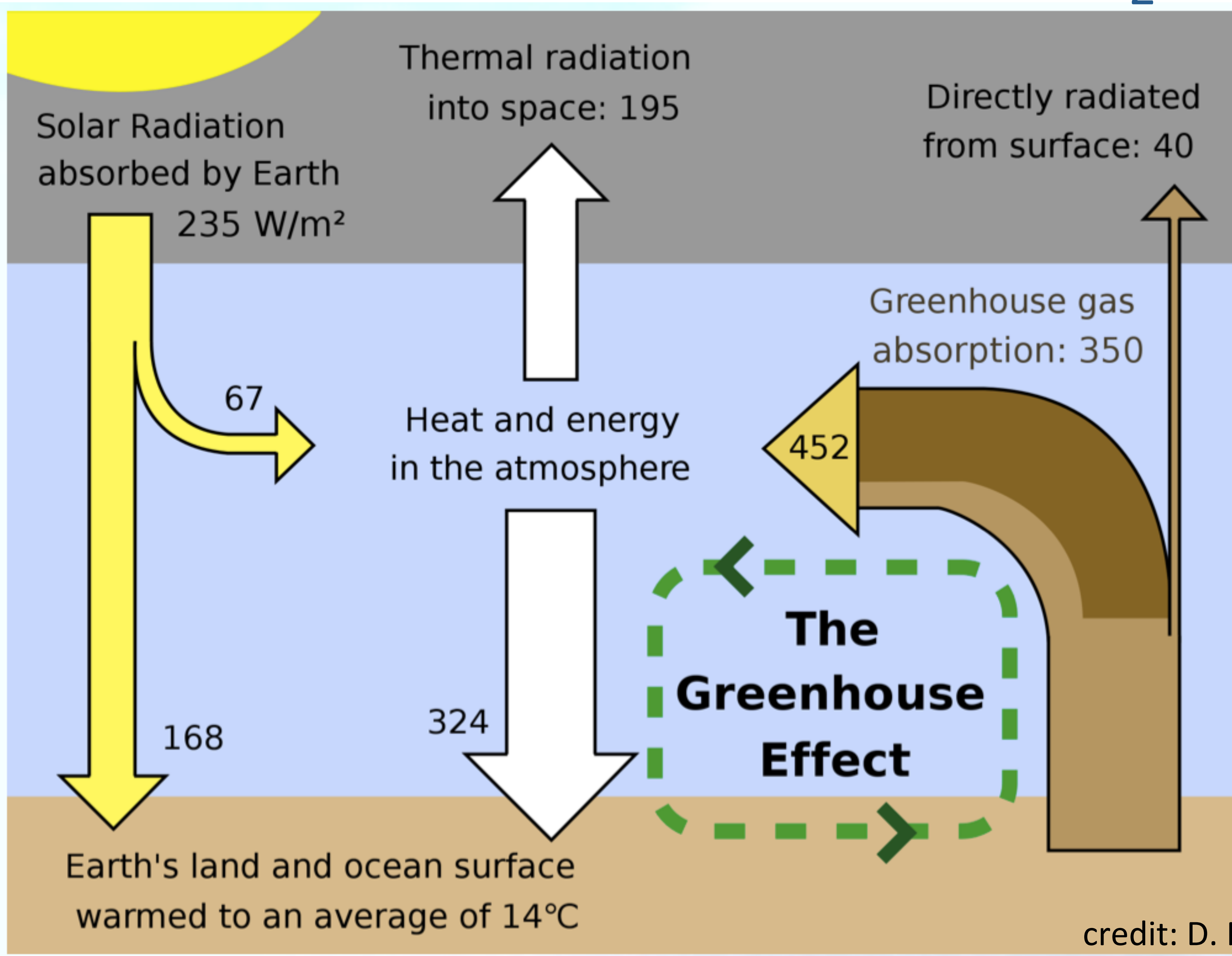
# The Climate System






credit: J. J. Hack/A. Gettelman

# Role of Greenhouse gasses, e.g. $CO_2$, $CH_4$




credit: D. Nychka

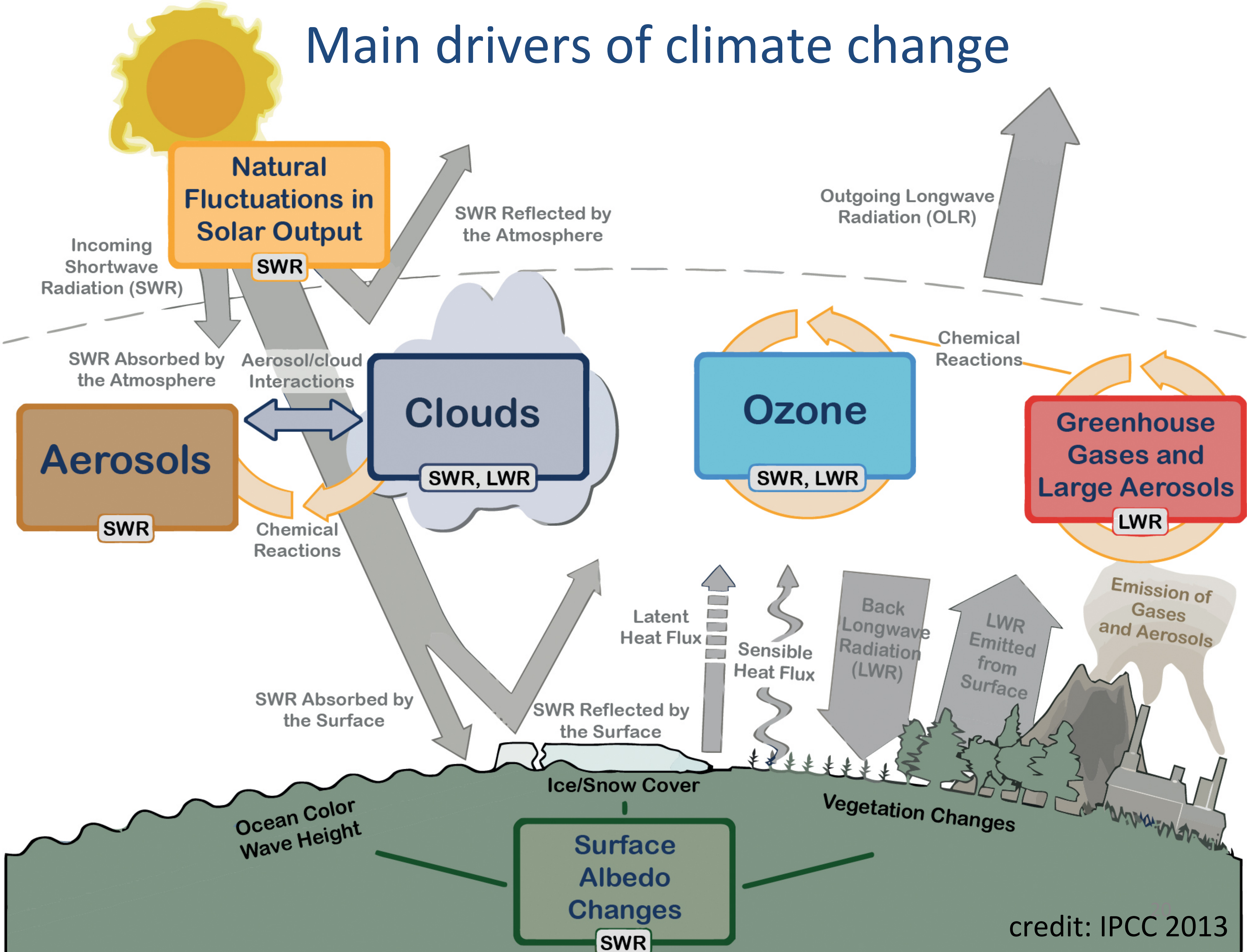
# Climate forcings

- Increasing greenhouse gasses changes the climate: a forcing.
  - Human activity can cause this by burning fossil fuel, etc.

- Changes in land use are also a forcing.

- Other (natural) forcings:
  - Changes in the sun's intensity
  - Volcanic eruptions

# Main drivers of climate change




credit: IPCC 2013

# History of climate modeling

Scientific basis for atmospheric simulation

- Rooted in laws of classical mechanics/thermodynamics
  - developed during 18th and 19th centuries (see Thompson, 1978)
- Early mathematical model described by Arrhenius (1896)
  - surface energy balance model

[Lorenz 1996] Description using nonlinear dynamical systems:

- $\mathbf{x}_t$ - the current state of the system is thought of as weather
  - e.g. a vector of relevant variables
- Climate is $E[\mathbf{x}_t]$ if the system is stationary.
  - Compute windowed averages over ~30 years.

- Current view: climate is a distribution, and can change over time.



Credit: D. Nychka, J. J. Hack/A. Gettelman

# History of climate modeling

[Lorenz 1996] Description using nonlinear dynamical systems:

- $\mathbf{x}_t$ - the current state of the system is thought of as weather
  - e.g. a vector of relevant variables

$$\mathbf{x}_{t+1} = G(\mathbf{x}_t, F_t)$$

- $F_t$ are external "forcings" (human activities, solar radiation, etc.)

- G is based on physics, usually deterministic.
  - Climate models!

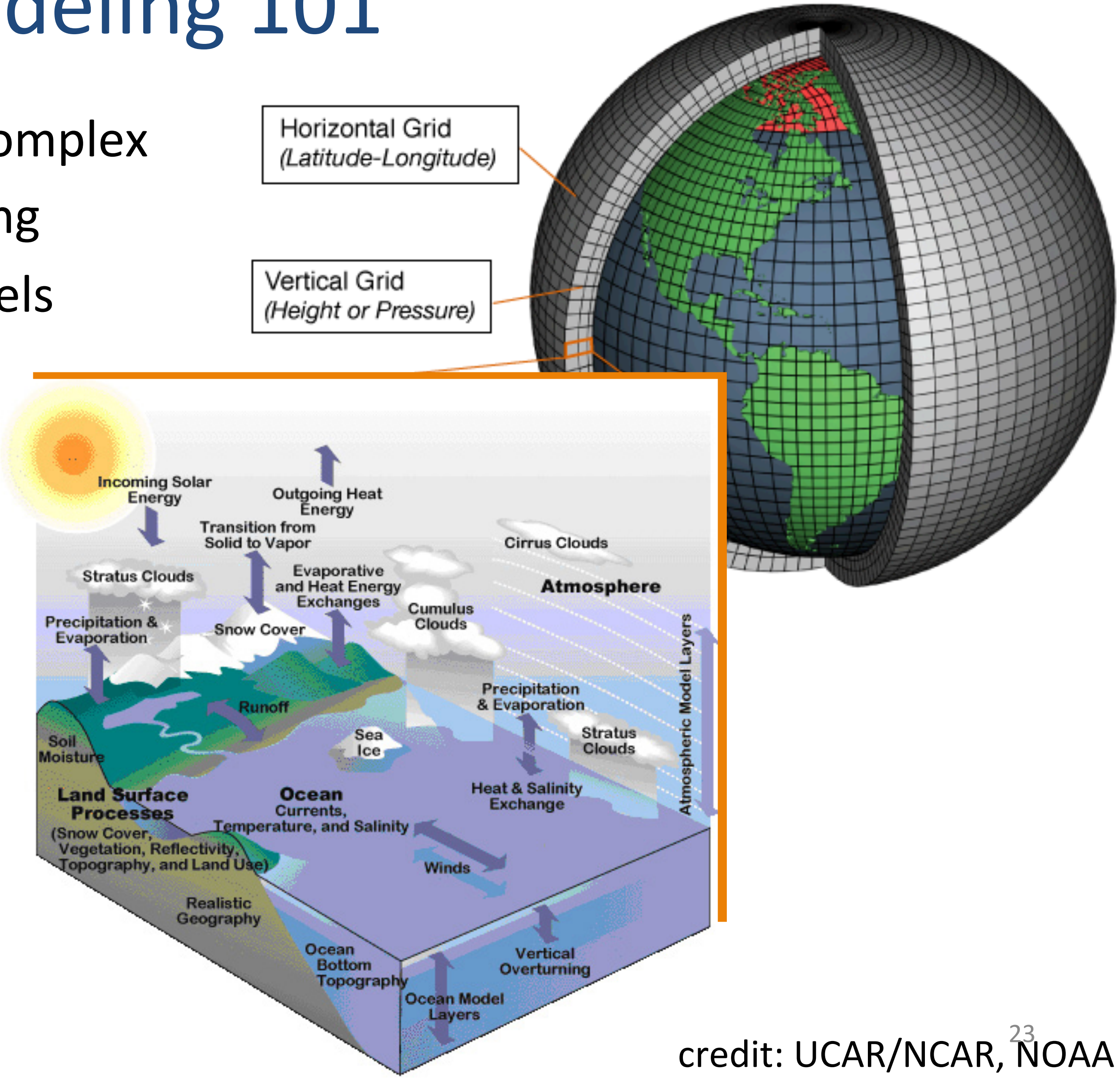


Credit: D. Nychka

# Climate Modeling 101

Climate model: a complex system of interacting mathematical models

- Not data-driven
- Based on scientific first principles
  - Meteorology
  - Oceanography
  - Geophysics
  - …
- Discretization into grid boxes




credit: UCAR/NCAR, NOAA

# Intergovernmental Panel on Climate Change

- IPCC: Intergovernmental Panel on Climate Change
  - Nobel Peace Prize 2007 (shared with Al Gore).
  - Interdisciplinary scientific body, formed by UN in 1988.
  - Fourth Assessment Report, 2007, on global climate change

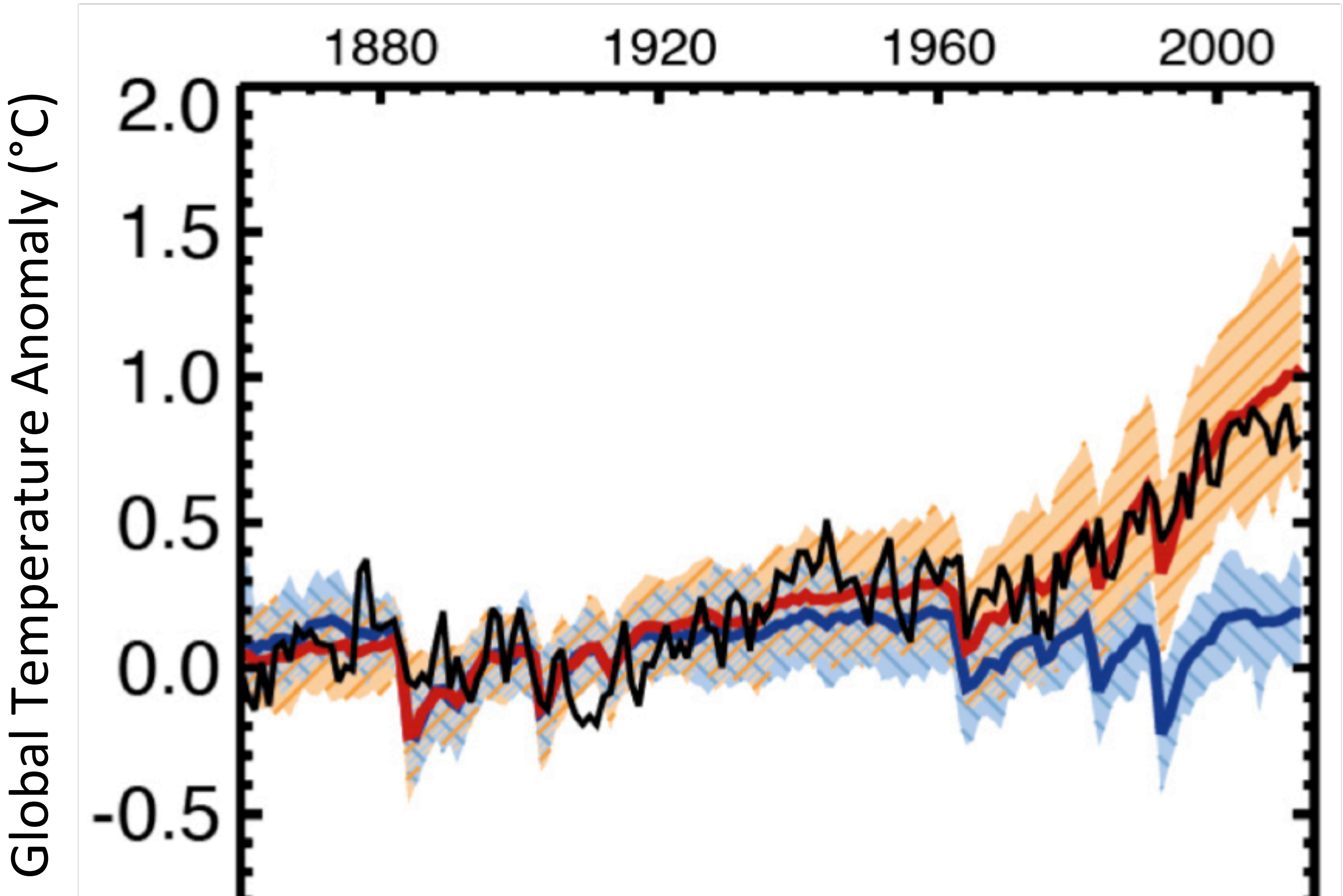
    450 lead authors from 130 countries, 800 contributing authors, over 2,500 reviewers.
  - Fifth Assessment Report, September 2013. Over 830 authors.

- Climate models contributing to IPCC reports include:

  Bjerknes Center for Climate Research (Norway), Canadian Centre for Climate Modelling and Analysis, Centre National de Recherches Météorologiques (France), Commonwealth Scientific and Industrial Research Organisation (Australia), Geophysical Fluid Dynamics Laboratory (Princeton University), Goddard Institute for Space Studies (NASA), Hadley Centre for Climate Change (United Kingdom Meteorology Office), Institute of Atmospheric Physics (Chinese Academy of Sciences), Institute of Numerical Mathematics Climate Model (Russian Academy of Sciences), Istituto Nazionale di Geofisica e Vulcanologia (Italy), Max Planck Institute (Germany), Meteorological Institute at the University of Bonn (Germany), Meteorological Research Institute (Japan), Model for Interdisciplinary Research on Climate (Japan), National Center for Atmospheric Research (Colorado), among others.

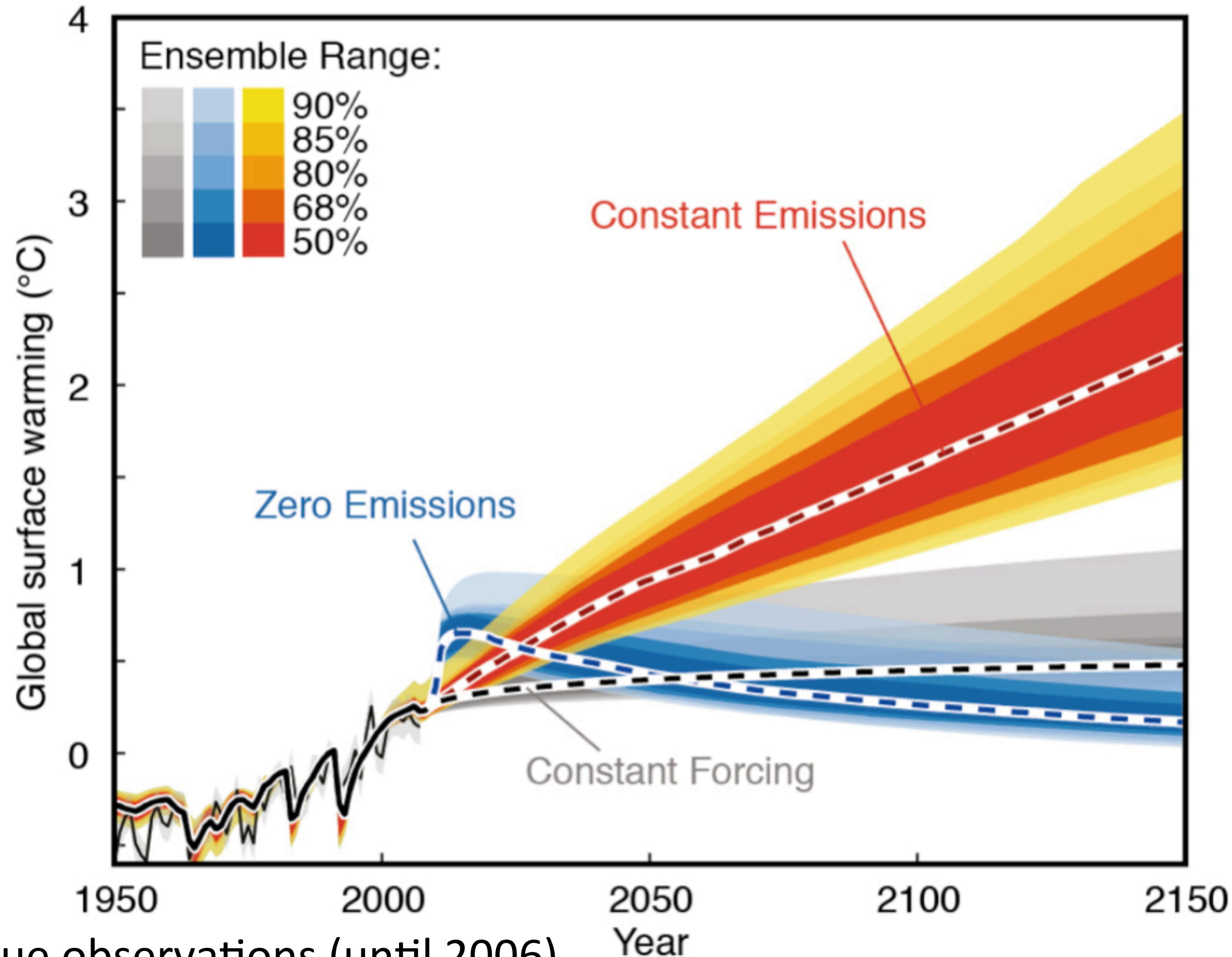
# IPCC findings: human influence on climate

**Black**: true observations.
Orange/red: Climate model simulations with human-induced greenhouse gasses.
Blue: Climate model simulations *without* human-induced greenhouse gasses.




IPCC 2013

# Modeling future scenarios




Black: True observations (until 2006).

Orange/red: Constant emissions.

Grey: Constant atmospheric composition (constant forcing).

Blue: Zero emissions starting 2010 (impossible).

credit: IPCC 2013

# Climate Data is Big Data

GCMs/ESMs (CMIP3/5) (Tb/day)

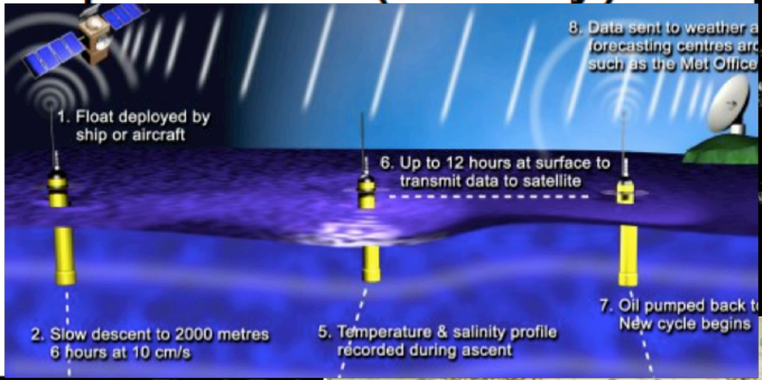
Satellite retrievals (Tb/day)

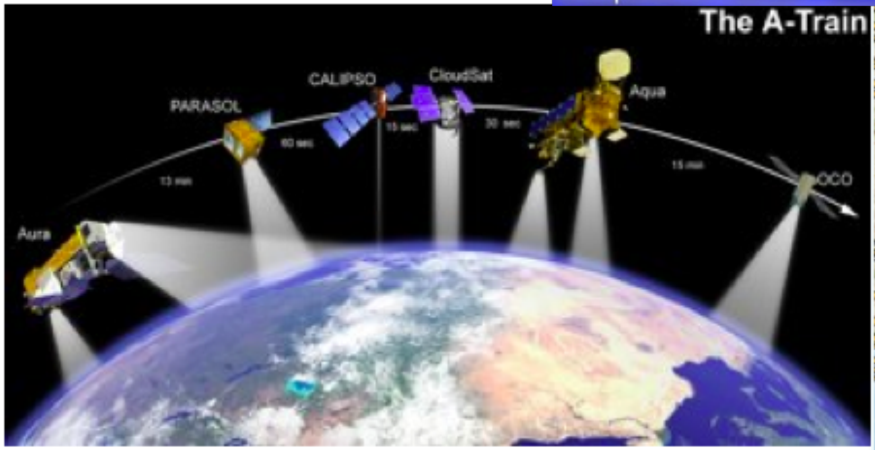
Next-gen reanalysis products (Tb/day)

In-situ data

Paleo-data

Regional models






credit: G. Schmidt/D. Nychka

# Challenges of climate data

Massive, high-dimensional, spatiotemporal

Data streams, non-stationary (over time and space)

Unlabeled, sparsity, missing data, heterogeneity

Low intrinsic dimension, latent structure

# Outline of Tutorial

For each climate problem, we'll discuss existing approaches in climate science, and ML, and open problems.

1. Paleo-climate reconstruction

   What was the climate before we had thermometers?

2. Climate downscaling

   What climate can I expect in my own backyard?

3. Climate model ensembles

   How best to harness the predictions of the IPCC ensemble?

4. Extreme events

   What are extreme events and how will climate change affect them?

5. Space and time

   How to capture dependencies over space and time?

6. Conclusion and further challenges for ML in climate science

# We'll touch on most of these ML topics

- Graphical models
  - MRF/CRF, topic models, inference, structure learning
- Hierarchical Bayesian models
- Matrix completion
- Sparse representations
- Causality
- Multitask learning
- Unsupervised learning
- Online learning
- Analysis of quantiles and extremes
- Spatial statistics
- Deep learning

# Why should I (NIPS attendee) care?

- Very impactful problems for society; climate change mitigation and adaptation. Chance to affect IPCC.
- Data-rich "big data" playground, public data sets
- Largely open field for ML, with many low-hanging fruit
- Climate scientists are already extremely computationally sophisticated, writing massive software, running HPC.
  - Allows for fruitful collaborations focused on the ML value-add.
  - Climate model simulations provide a vast wealth of data/knowledge.
- Physics provides some inertia, predictability!
- Funding opportunities


# Climate Informatics

2011 First International Workshop on Climate Informatics

New York Academy of Sciences

Climate Informatics Wiki launched

2013 “Climate Informatics” book chapter [M et al. 2013]

2015 Please join us in September as Climate Informatics turns 5!

National Center for Atmospheric Research, Boulder CO

In the first 4 years: participants from over 16 countries, 28 states

# *Paleo*-climate Reconstruction

# Paleo-climate reconstruction

Problem:

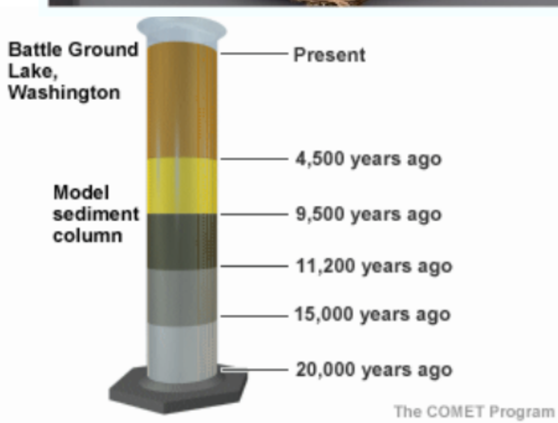
- To understand climate change we need to understand past climates.
- NOTE: climate has fluctuated at much greater scales in the past than in the 20th Century.
- However the variance on measurements is higher in the past.
  - We did not have a global grid of measurements
  - Measurements corrupted or lost

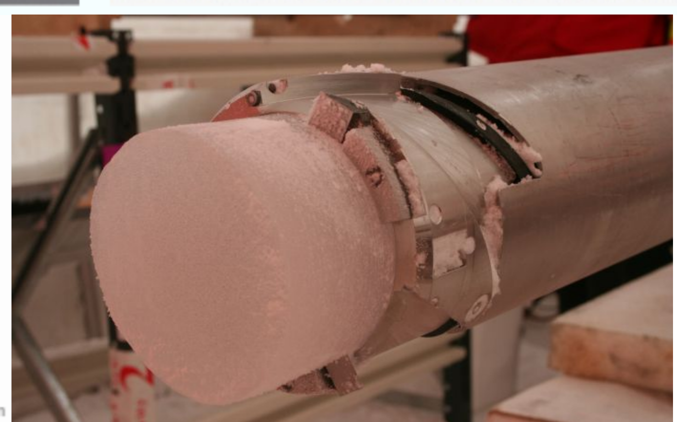
Challenge: use paleo-proxies to reconstruct temperatures, $CO_2$

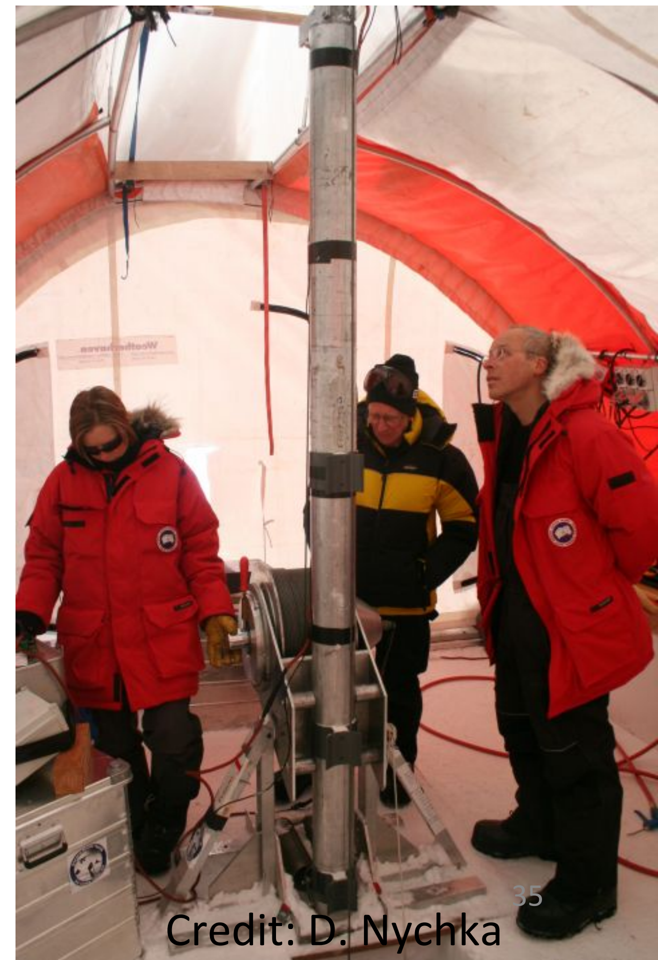
E.g. tree rings, coral, ice cores, lake sediment cores, provide estimates.

# Paleo-climate reconstruction

Challenge: use paleo-proxies to reconstruct temperature, $CO_2$ concentrations. E.g. tree rings, coral, ice cores, lake sediment cores.

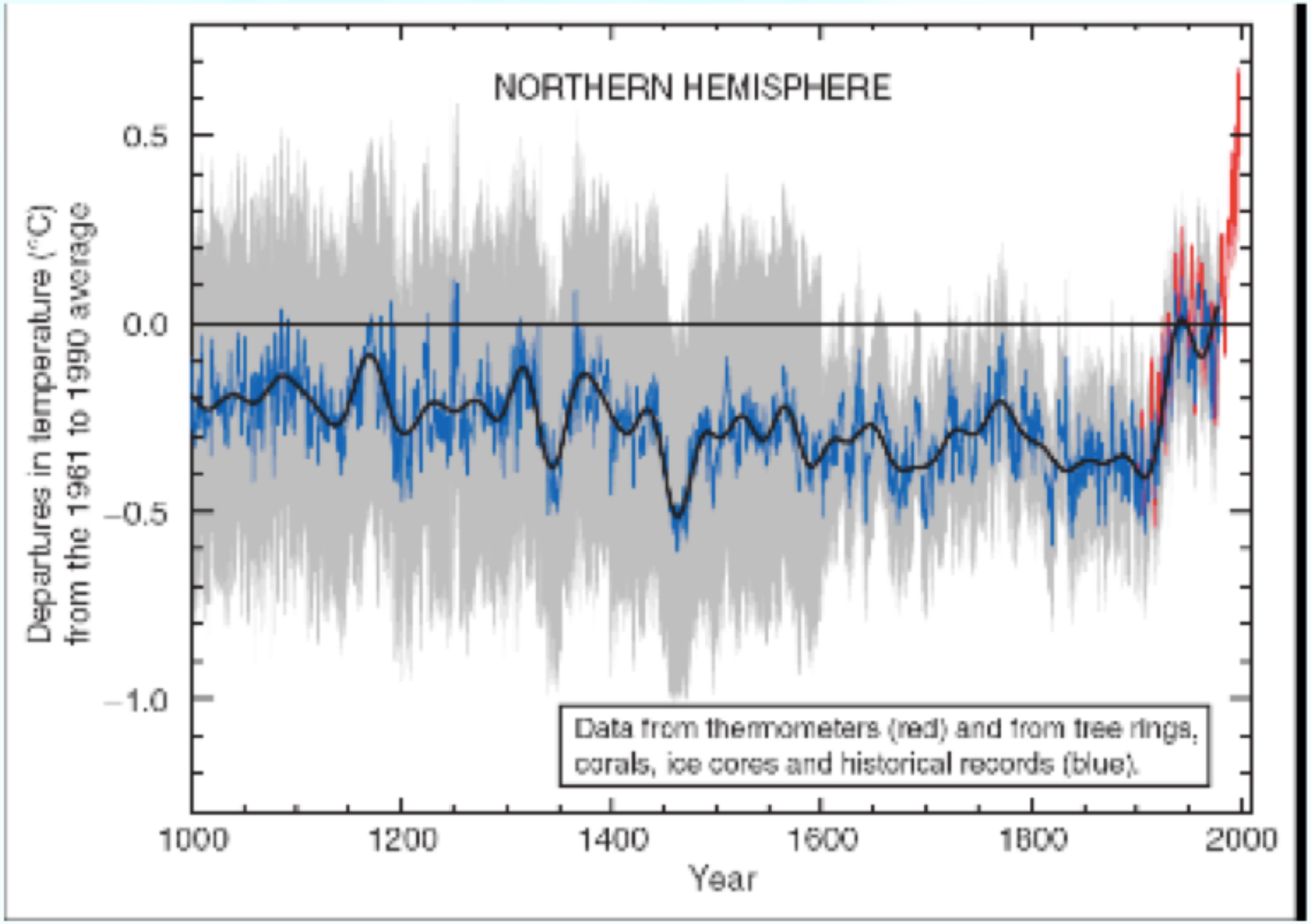





Credit: D. Nychka

# “The Hockey Stick Curve”

Northern Hemisphere temperatures 1000AD - 2000AD




Public attention focused on the mean curve, not the variance.

Mann, Bradley and Hughes 1999, JGL.

Credit: D. Nychka

# Bayesian approach

[Li, Nychka & Ammann, JASA, 2010]

Bayesian hierarchical model used to generate ensembles

P(D, T, θ) = P(D | T, θ) P(T | θ) P(θ)

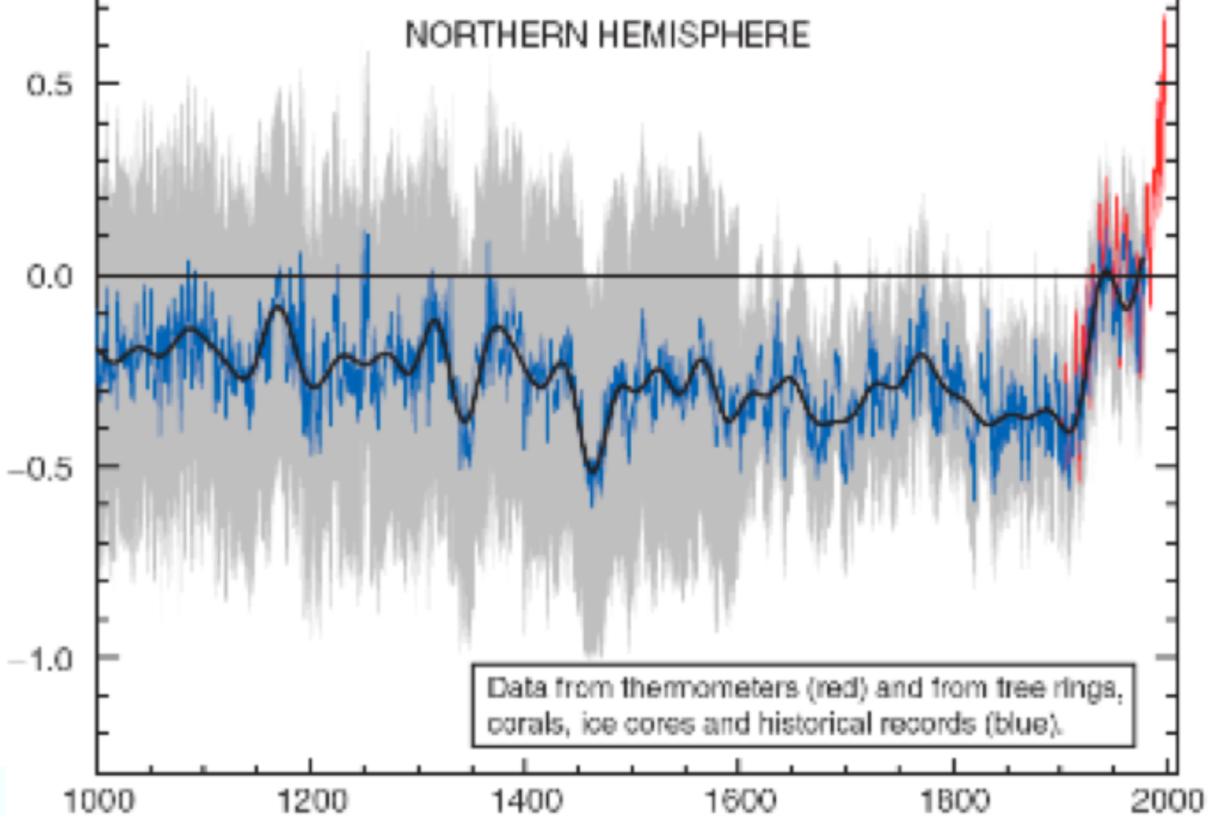
- P(D | T, θ): Data model: The relationship of the proxies to surface temperatures
  - e.g. linear plus noise. $D_{j,t} = f_j(T_t) + \epsilon_{j,t}$
  - The linear filter is deterministic, based on the type of proxy, *j*. Scalar parameters.
- P(T | θ): Process model: How do surface temperatures evolve in time (and space)
  - Physical model incorporating forcings, plus noise. Scalar parameters.
- P(θ): Priors over statistical parameters (above).
  - Chosen based on physical information

Estimate temperature by sampling from the conditional distribution of temperature, given the observed proxy values.

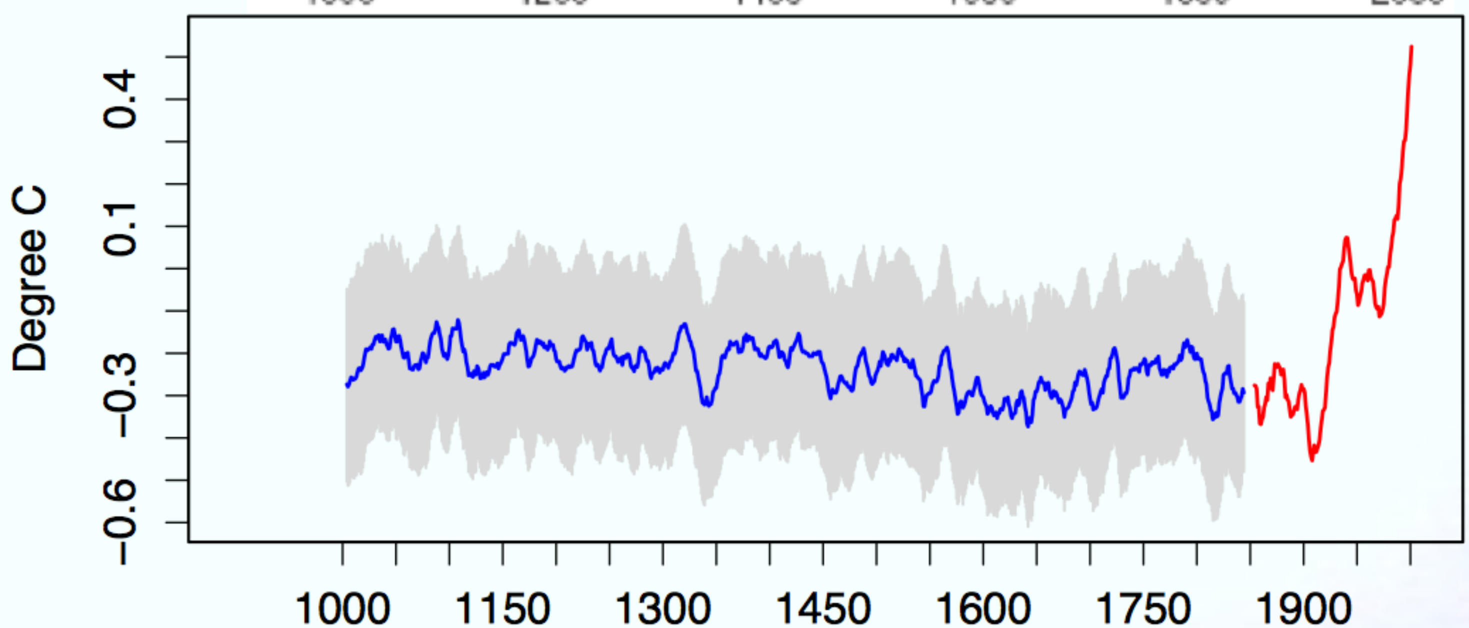
- Using MCMC



Credit: D. Nychka

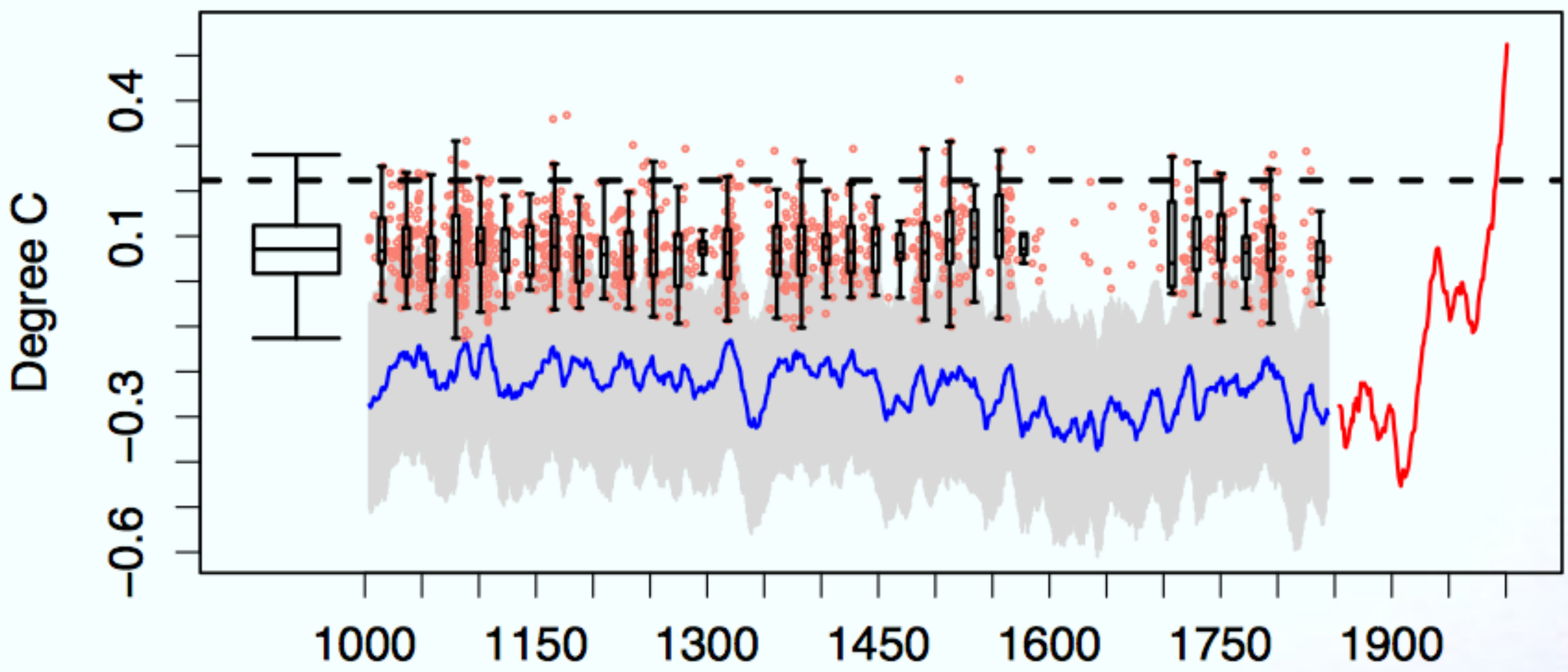



"Hockey stick curve"




Bayesian hierarchical approach

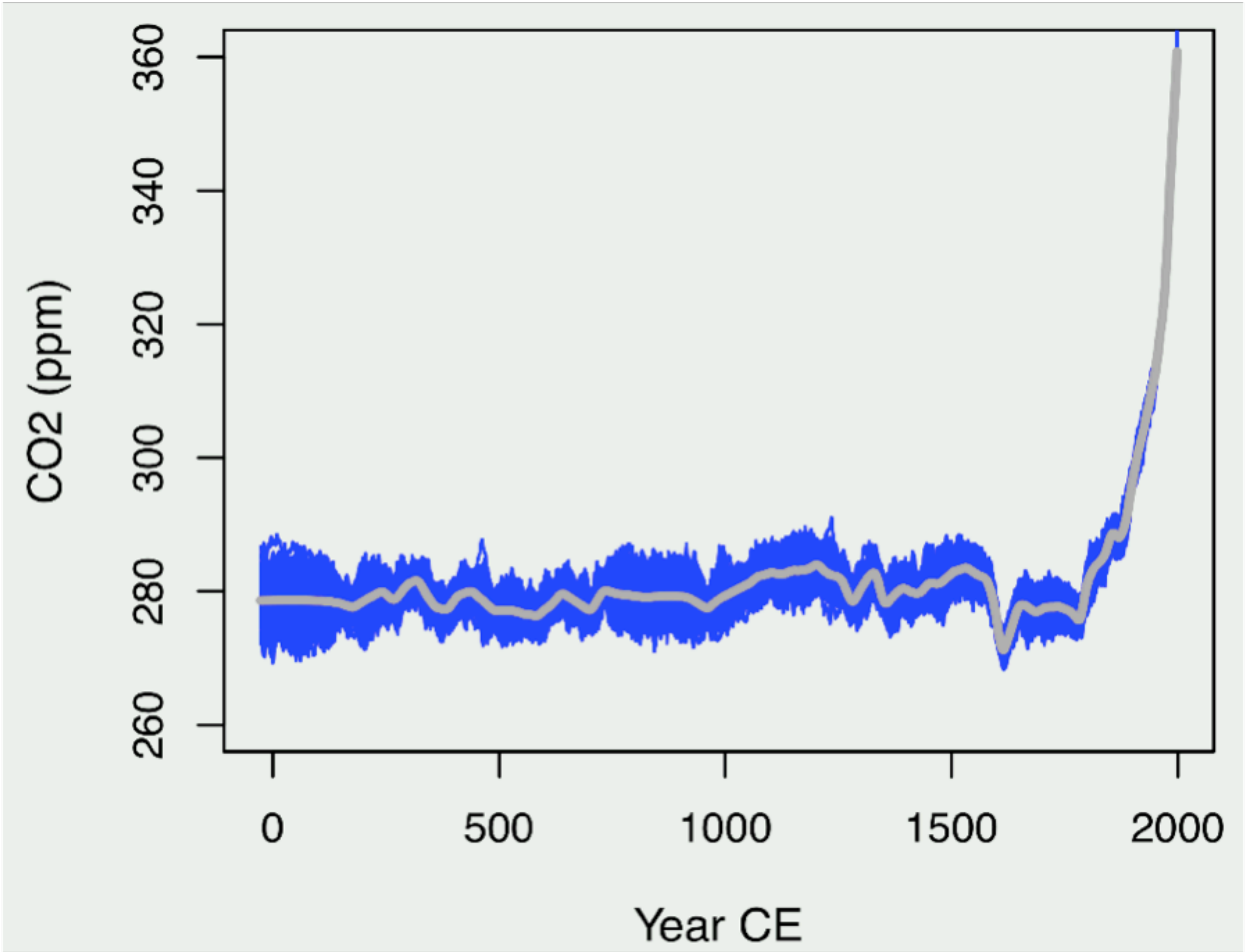
Credit: D. Nychka




o max value (pre-1850) of each ensemble draw
\- - - 95% upper bound on these 1000 maxima


Credit: D. Nychka

# Bayesian reconstructions of $CO_2$




[Nychka et al., JSM, 2014]

Challenge: How to best harness paleo-proxies to reconstruct past climates?

Possible ML approaches:

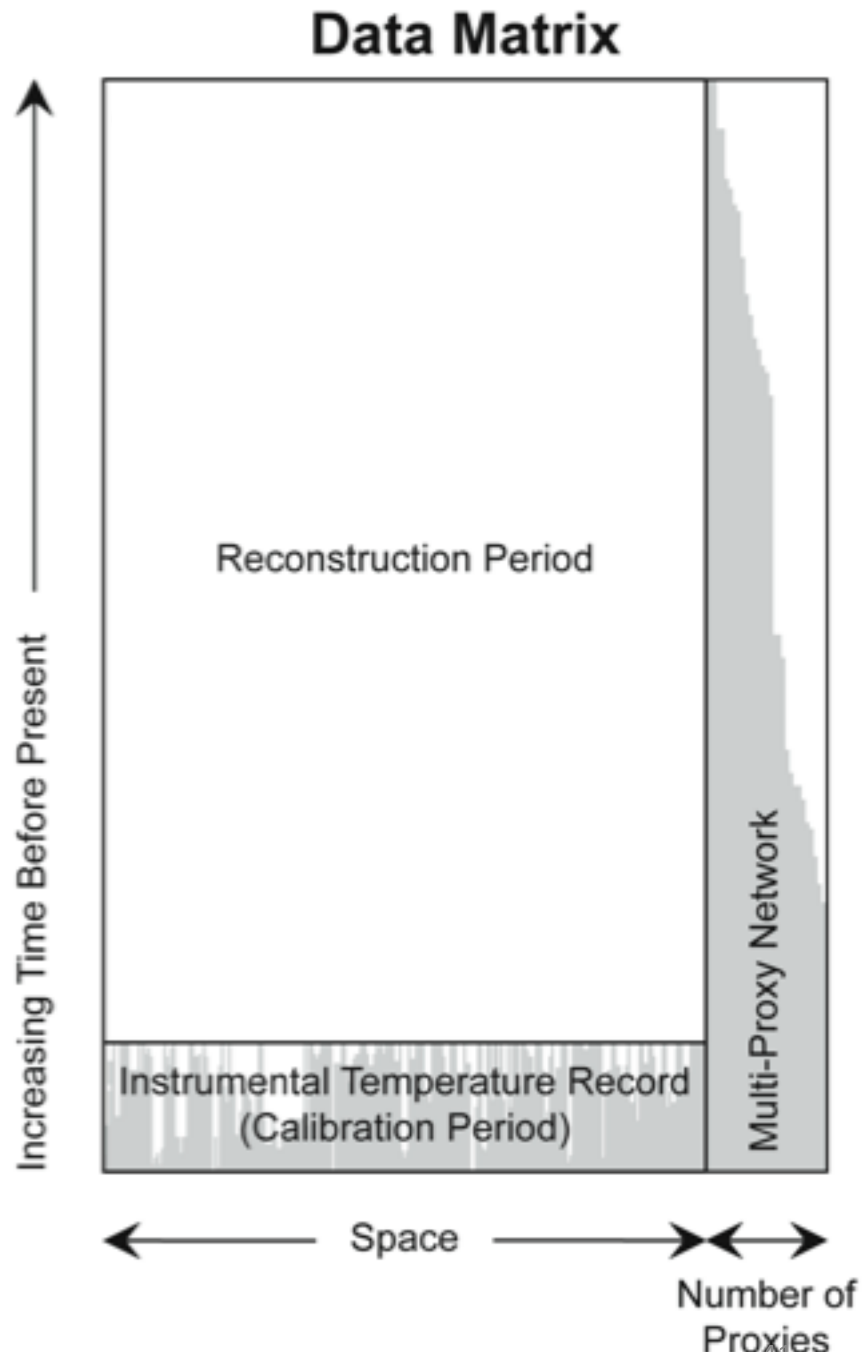
Can sparse matrix completion techniques play a role?

Discover latent structure?

Related ML issues:

Data fusion (many small data sets!)

Multi-view learning

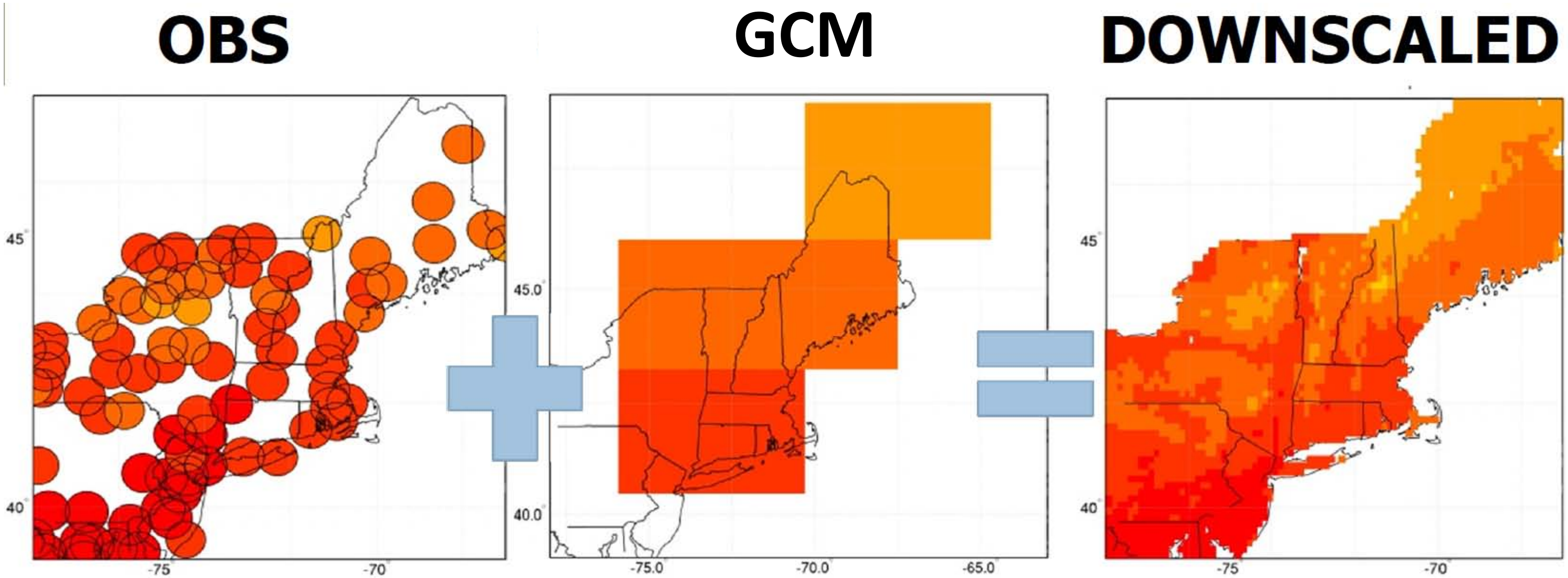



[Smerdon & Kaplan, Journal of Climate, 2007]

# Climate Downscaling

# Climate Downscaling

# Climate Downscaling



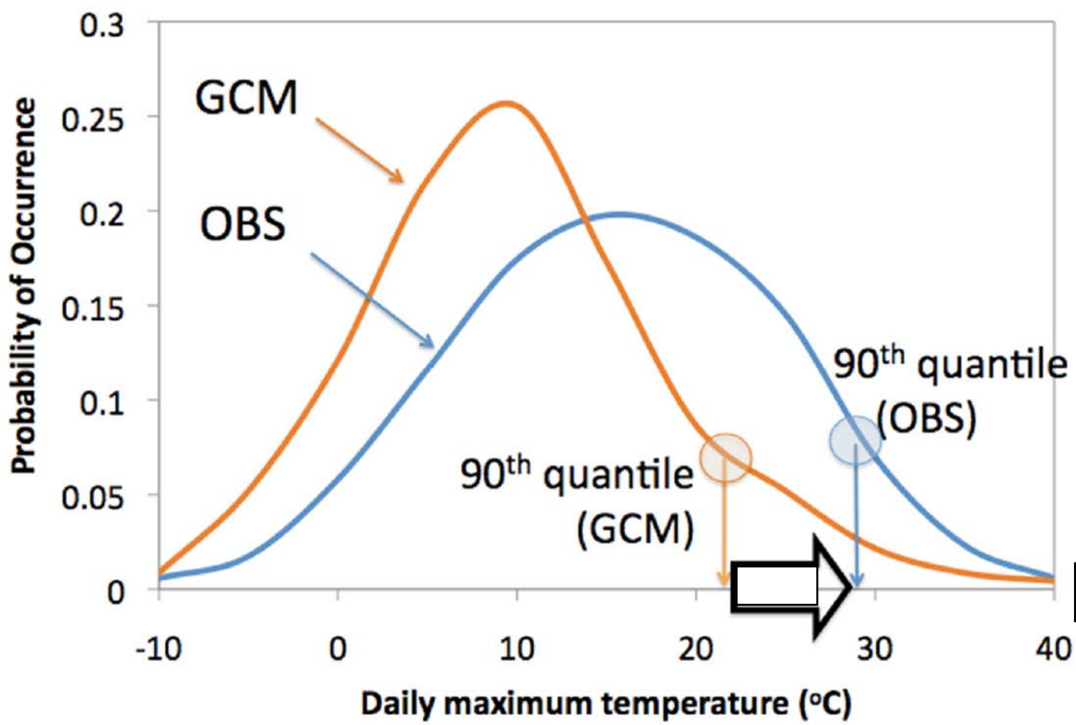

Credit: K. Hayhoe

- Sub-grid scale variables: temperature, precipitation, …
  - Local observations (OBS): Weather stations, remote sensing, …
  - Global models (GCM): Coarse resolution fields

Goal: Understanding Local Climate Change

# Matching Statistics

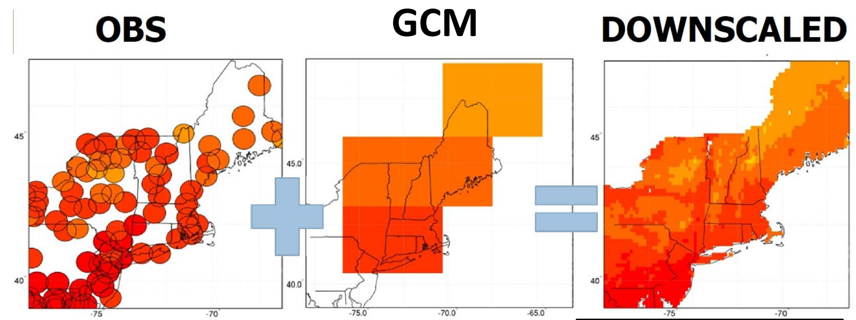
- Climate variables: statistics vs (daily) values
- Nonlinear dynamics, hard to project exact value
  – Example: Value of (temp, precip) after one year, *not* predictable
- Statistics are more stable, (hopefully) changes smoothly
  – Example: Matching quantiles between GCM and OBS [Li et al., JGR 2010]




Credit: K. Hayhoe

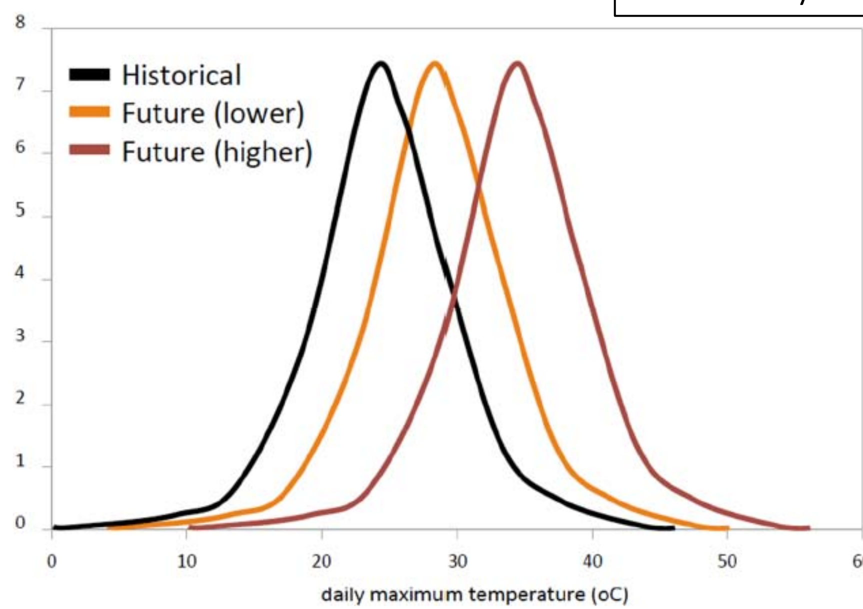
# Downscaling: Local Climate

- OBS is for past and current
  - What will it be in future?



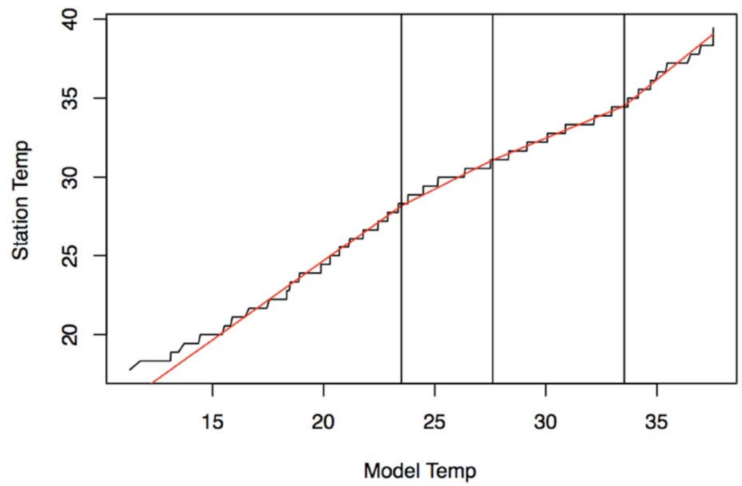

Credit: K. Hayhoe

- Delta method [Ramirez et al., 2010]
  - Distribution shifts
  - No change in shape

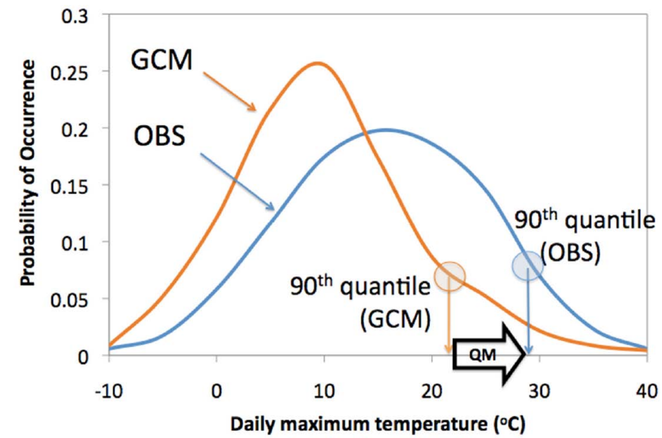


# Downscaling: Local Climate

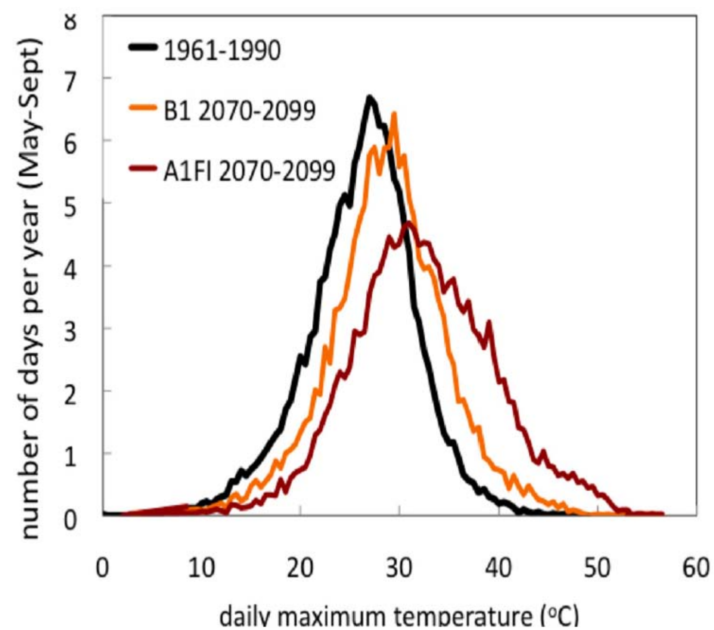
- Quantile mapping [Li et al., JGR 2010]
  - Bias-correction
- Statistical asynchronous regression
  - Piecewise linear, by month/season
    [O'Brien et al., JGR 2001, Dettinger et al., 2004]




- Quantile Regression
  [Koenker & Bassett, 1978, Friederichs et al., MWR 2007]

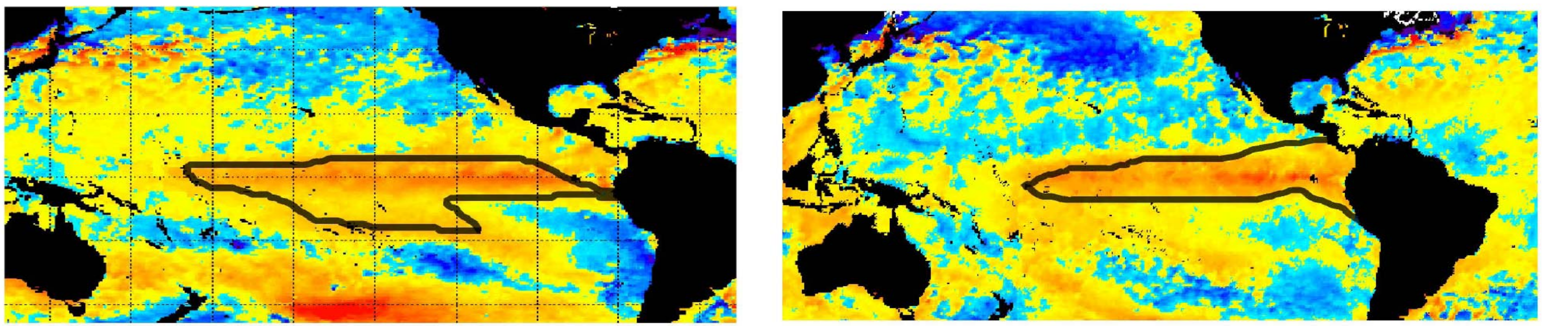



Credit: K. Hayhoe



# Downscaling Approaches

- Weather typing: Method of analogues [Zorita et al., J Clim 1999]
  - Matching with previous analogous situations



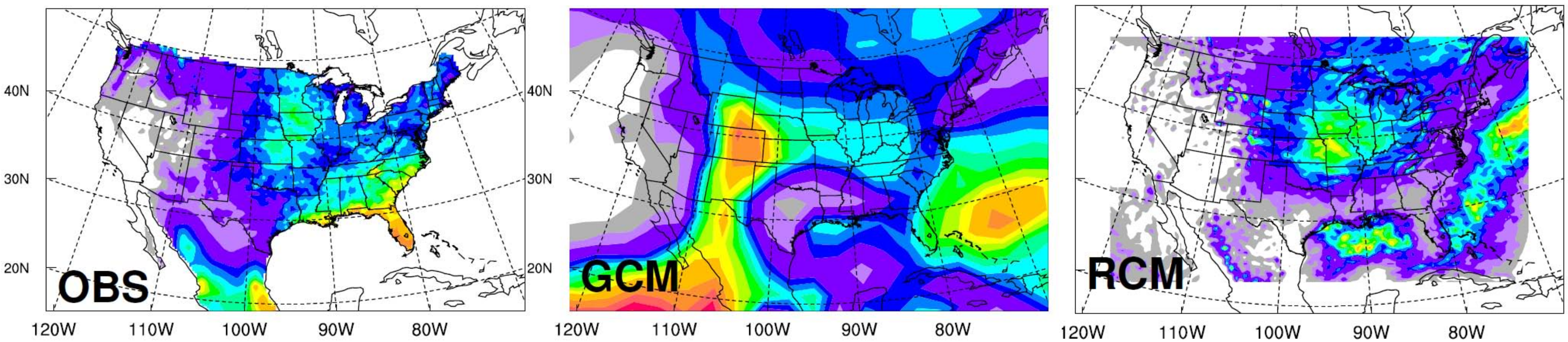
Sea Surface Temperature (SST) anomalies: El Nino in 1997-98 and 2002-03

- Weather generators
  - Match statistical attributes, not values [Semenov et al., CC 1997]
- Multiple linear regression [Jeong et al., CC 2012]
  - Use statistical influences, relationships
  - Spatial smoothing, structural constraints

# Dynamical Downscaling

- Regional Climate Models (RCMs) [Christensen et al., CC 2007]
  - Run RCMs with suitable boundary conditions
  - Usually better than GCMs for local climate

Average summer (JJA) rainfall, 1990-1995



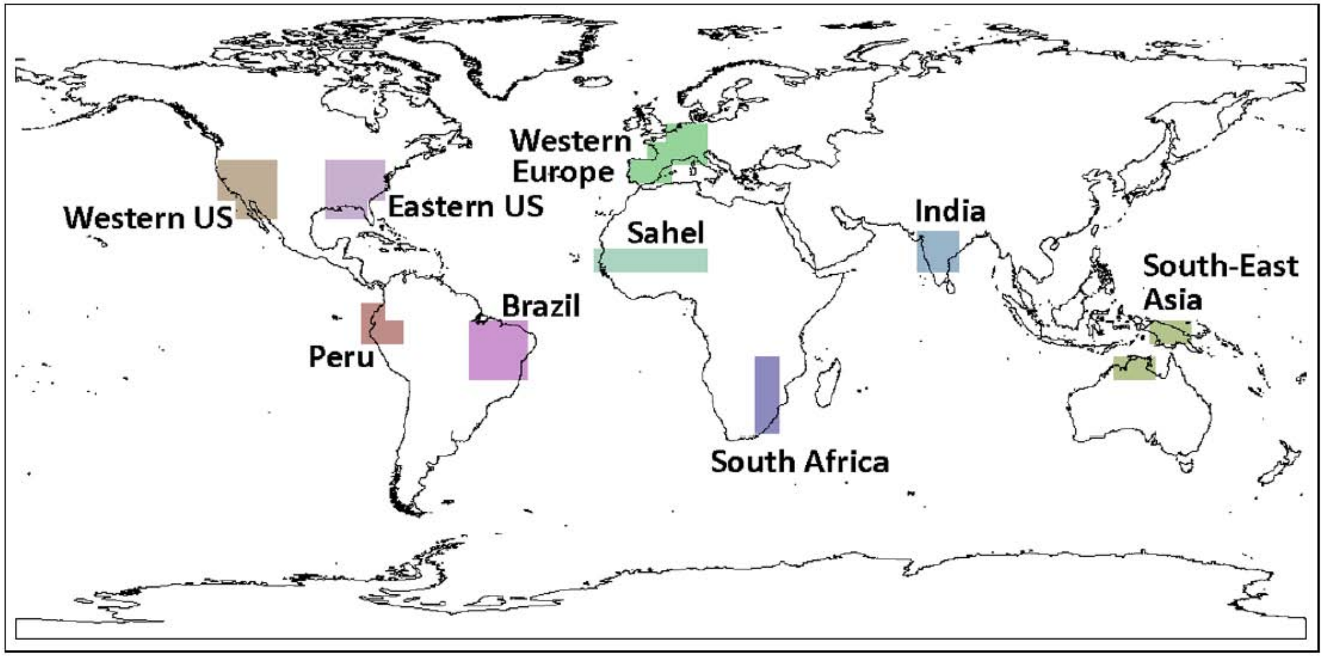

Credit: X. Liang

  - Boundary conditions for future climate? GCMs
  - Bias of GCMs affect boundary conditions

# Sparse, Structured Regression

- Influence of oceans on land temperature, precipitation
  - Y, over land: temp/precip in 9 regions
  - X, over oceans: temperature, sea level pressure, relative humidity, wind speeds, etc.

[Chatterjee et al., SDM 2012, Steinhaeusar et al., CD 2012]



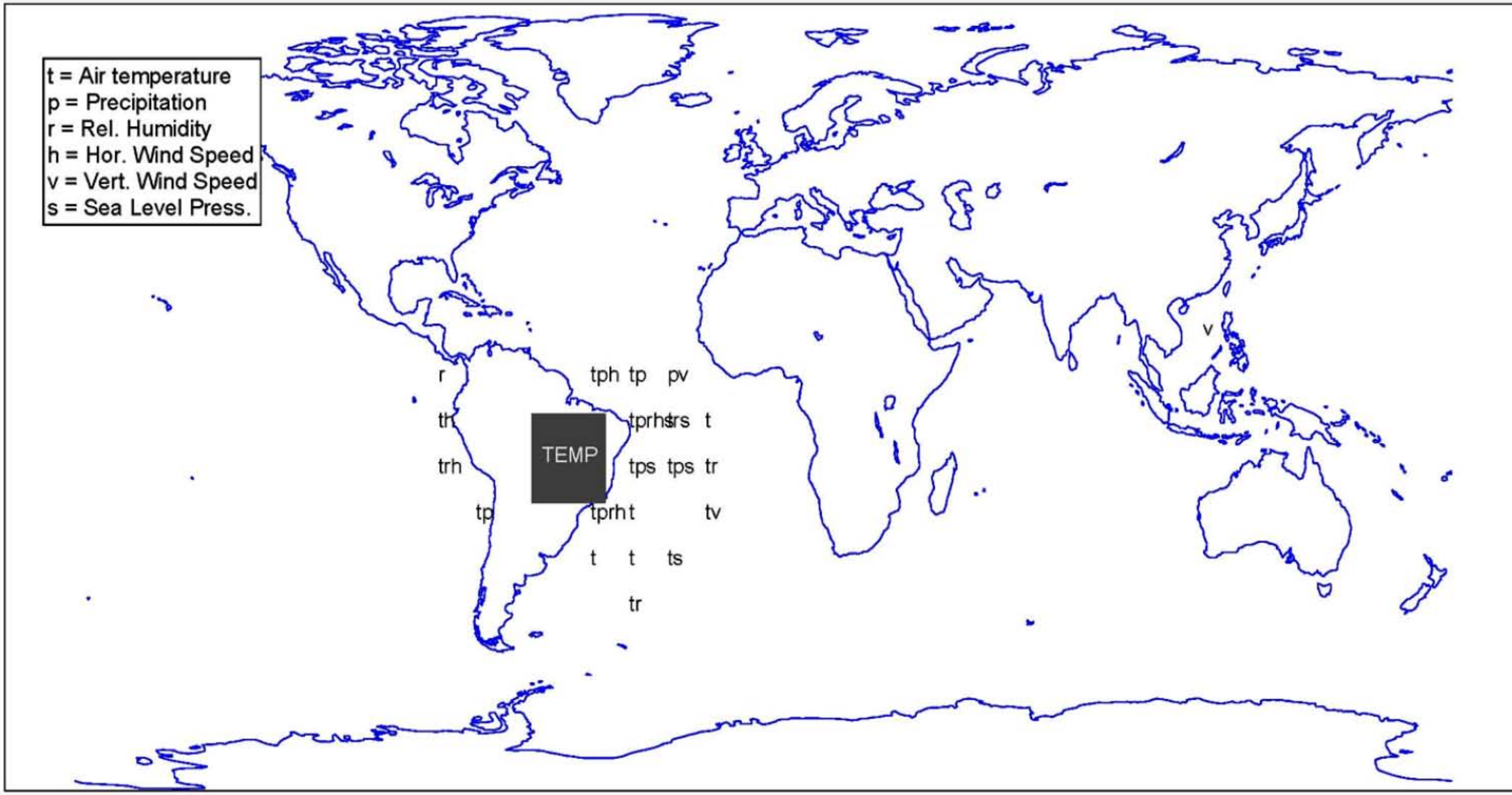
# Sparse, Structured Regression

- High-dimensional regression: Sparse Group Lasso (SGL)

$$\hat{\theta}_{sgl} = \underset{\theta \in \mathbb{R}^p}{\operatorname{argmin}} \left\{ \frac{1}{n} \|y - X\theta\|_2^2 + \lambda \left( \alpha \|\theta\|_1 + (1-\alpha) \|\theta\|_{1,\mathcal{G}} \right) \right\}$$

- Multiple ocean locations, multiple variables in each location
  - Few ocean locations are relevant
    - Group level sparsity for SGL
  - For relevant ocean locations, few covariates are relevant
    - Covariate level sparsity for SGL

# Results: SGL vs Baselines



# LatticeKrig: Spatial Downscaling

- Spatial statistical model [Benestad et al., NCC 2012, Nychka et al., JCGS 2014]

$$y_i = z_i^T \beta + g(x_i) + \varepsilon_i$$

Features

Spatial function

Noise

- Data at $n$ spatial locations, $(x_i, y_i, z_i)$, over time

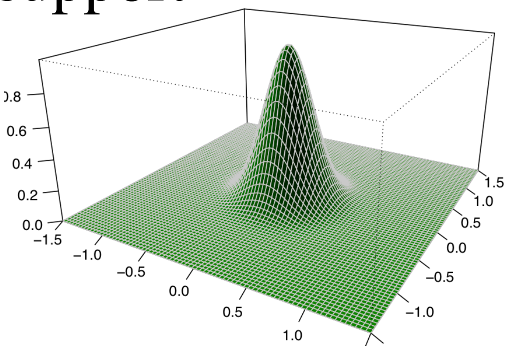
  $x_i$: spatial locations, $y_{i:}$: observations, $z_i$ : features

- Estimate $g(x)$ based on observations
  - Uncertainty quantification in estimate

# Model: Random effects, Multi-resolution

- Linear model with $m$ basis functions, compact support

$$g(x) = \sum_j c_j \varphi_j(x)$$



Credit: D. Nychka

- Coefficients $c_j$ are random, jointly Gaussian

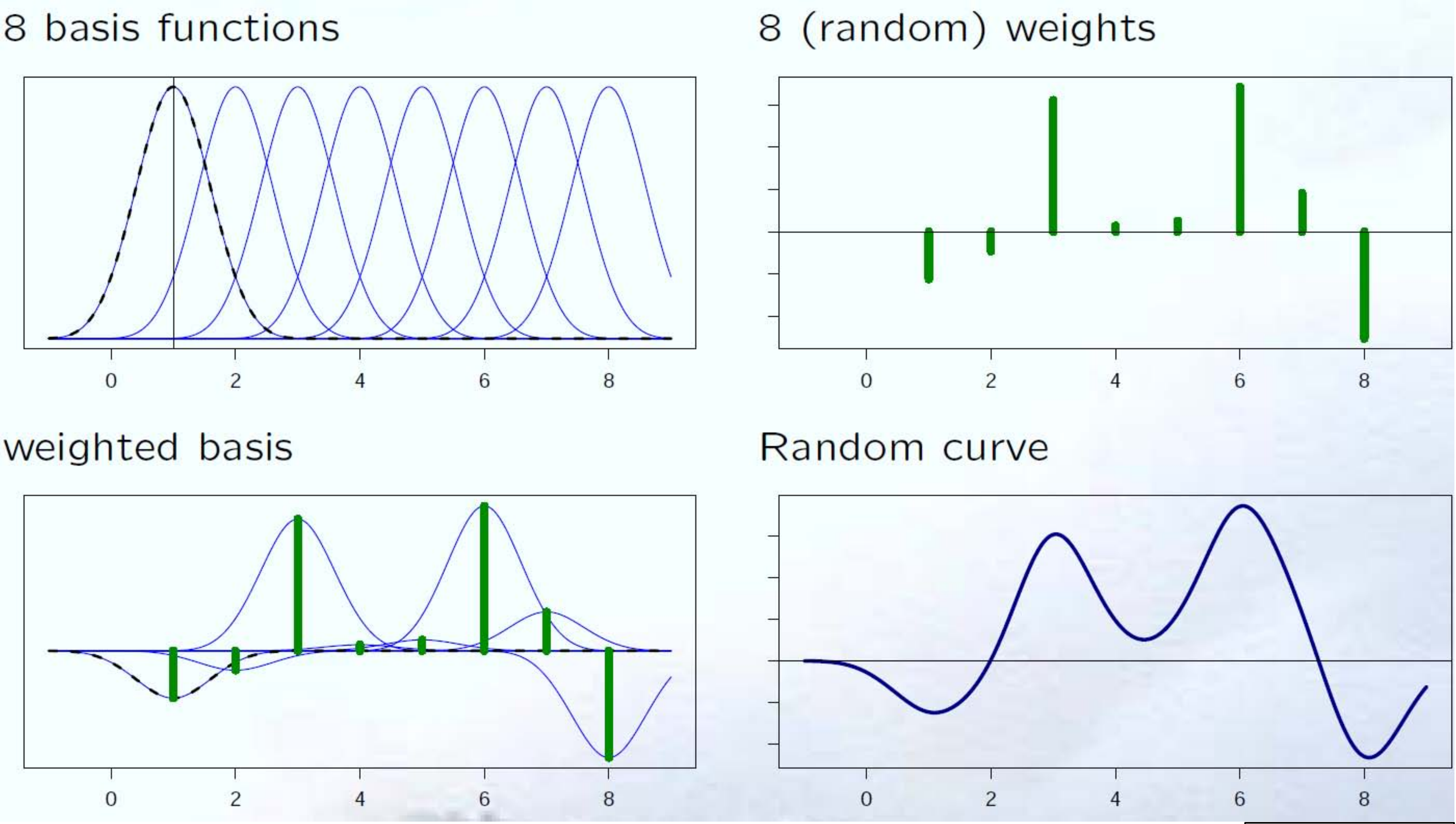
$$\mathbf{c} \sim N(0, \Sigma)$$

  – Precision matrix: $\Theta = \Sigma^{-1}$ is assumed to be sparse
  – Gaussian Markov random field

- Multi-resolution: Sum of $L$ independent random effects

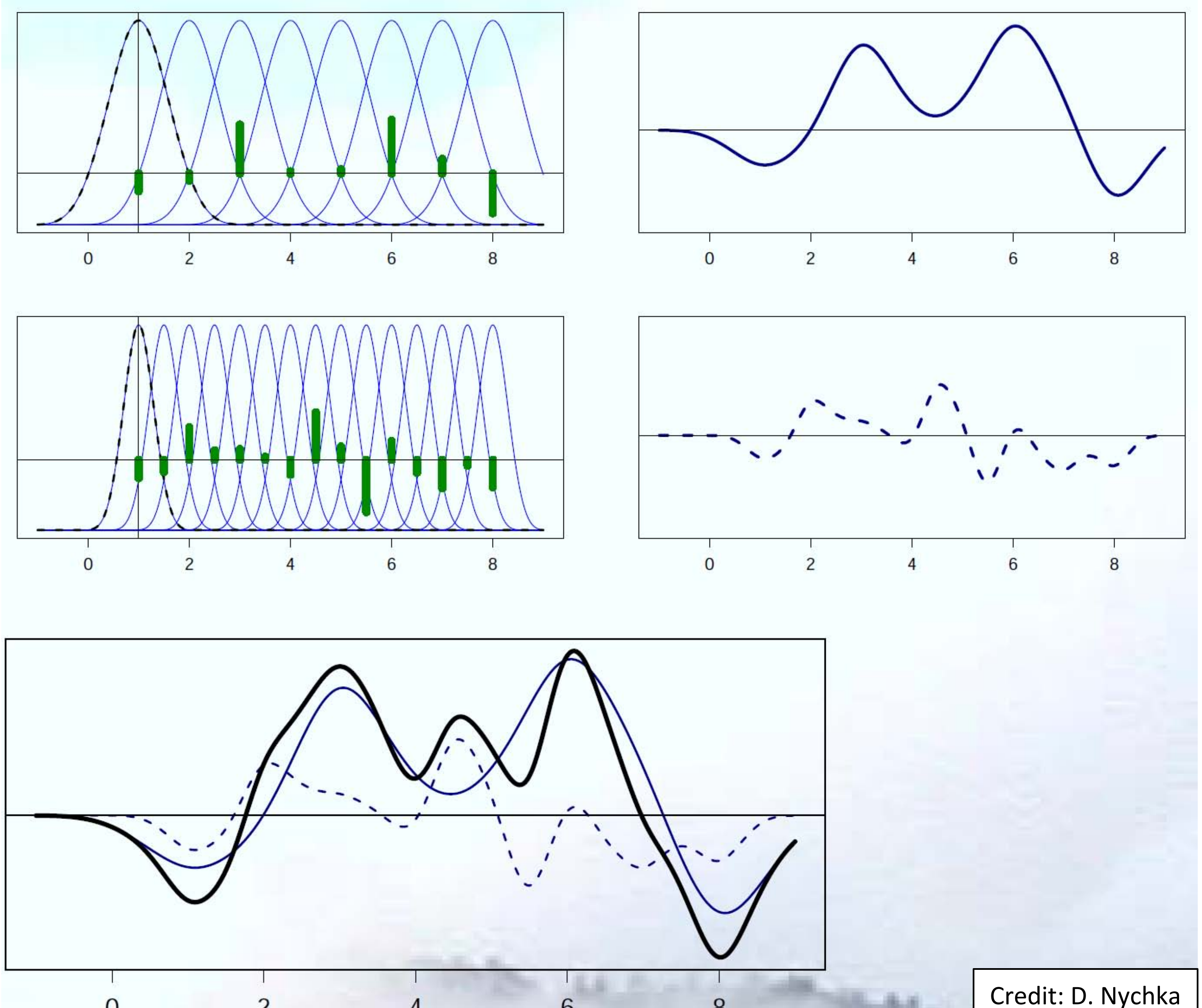
$$g(x) = \sum_{l=1}^{L} g_l(x) \qquad g_l(x) = \sum_j c_j^l \phi_{j,l}(x)$$

# Random Coefficients: 1-D example




Credit: D. Nychka

# Multi-resolution 1-D example

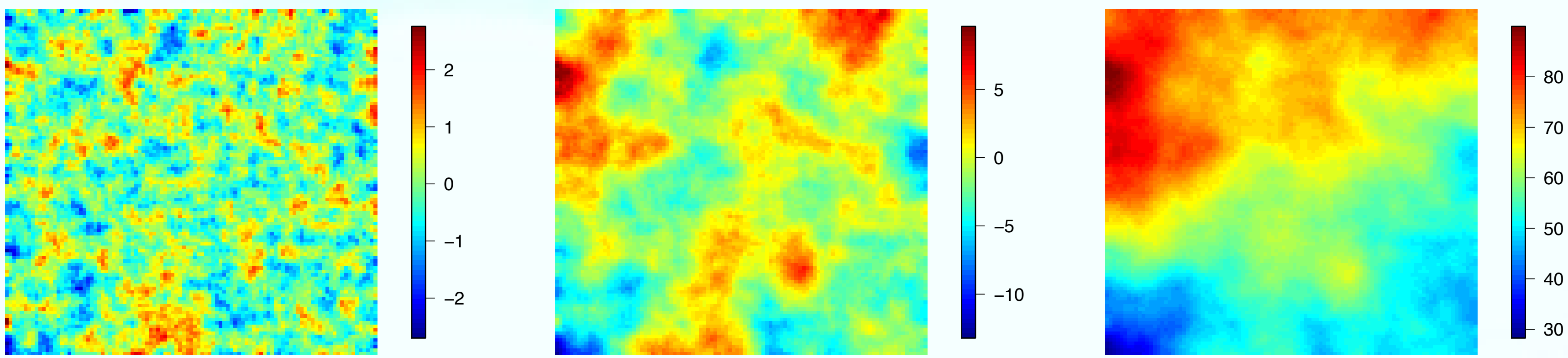



Credit: D. Nychka

# Spatial, Gaussian MRF

- Gaussian MRF covariance $\Sigma = (H^T H)^{-1}$
  - H is sparse, precision is $H^T H$ (sparse)
  - Spatial Auto-Regressive process, $\kappa > 0$
  - Conditional Auto-Regressive, $\kappa = 0$: Laplacian

$$
\begin{array}{ccccc}
\cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & -1 & \cdot & \cdot \\
\cdot & -1 & (4 + \kappa^2) & -1 & \cdot \\
\cdot & \cdot & -1 & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot
\end{array}
$$




$\kappa = .5$ $\kappa = .1$ $\kappa = .01$

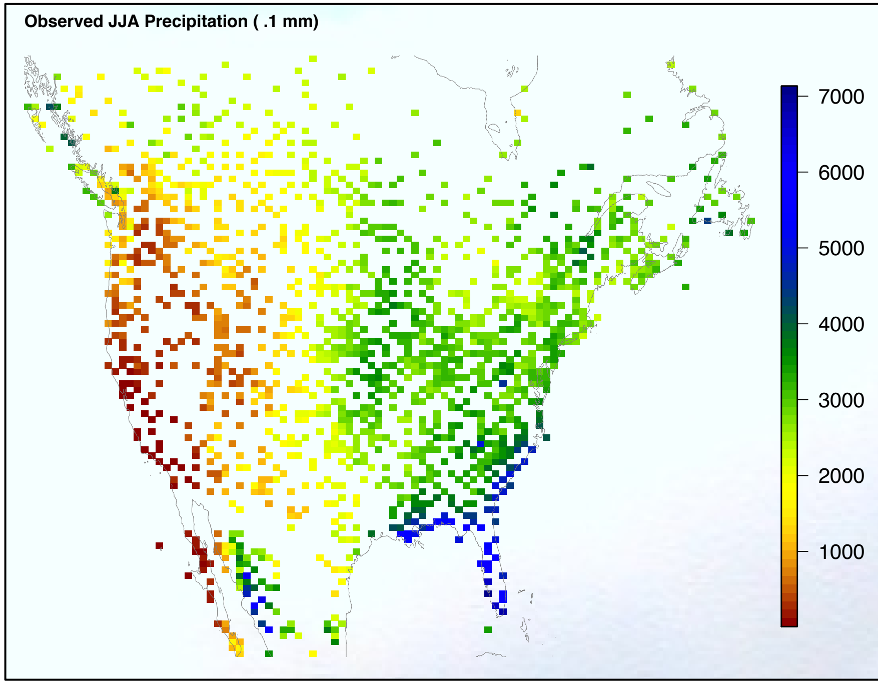
Credit: D. Nychka

Simulated fields of random coefficients **c**

# LatticeKrig: Statistical Downscaling

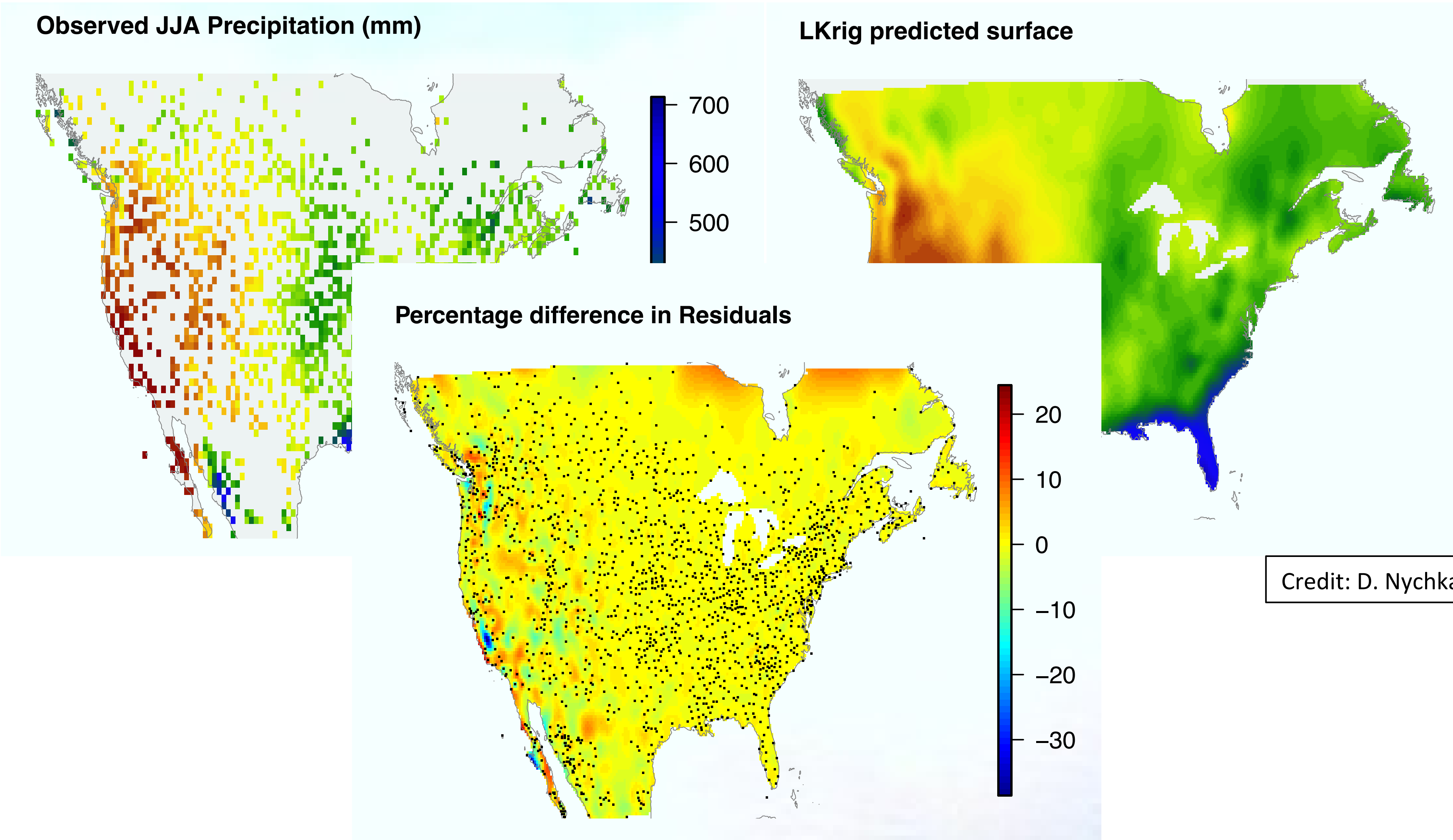
- North America (NA) summer (JJA) precipitation
- 1720 stations, JJA mean for 1950-2010

- Too coarse for precipitation
- Differences in spatial coverage
- Dependency on
  - Local topography
  - Multi-resolution spatial variations




Credit: D. Nychka

# Downscaling with LatticeKrig



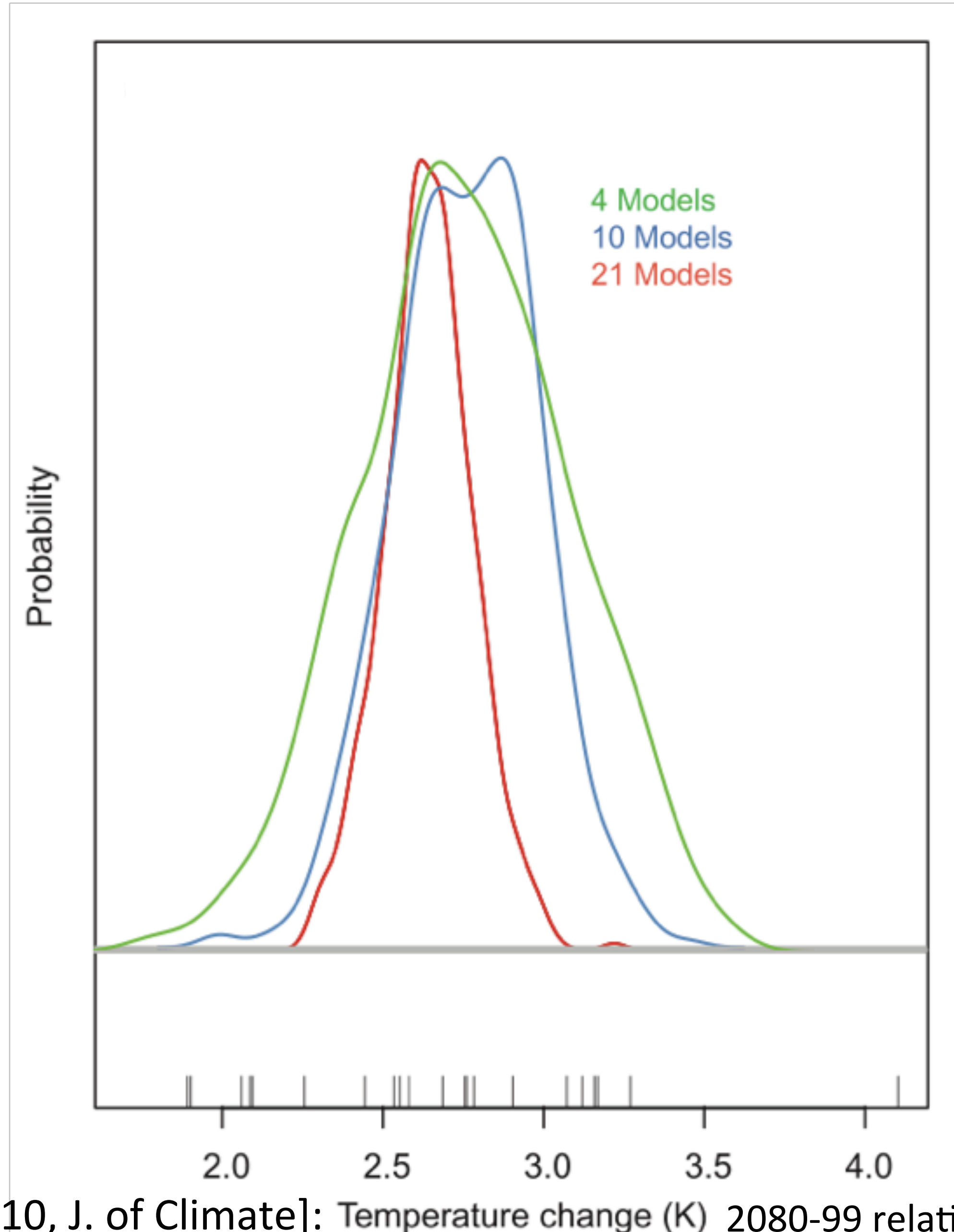

Credit: D. Nychka

# Challenges for Statistical Downscaling

- Simple vs complex processes, e.g., temp vs precip
  - Heterogeneous in space-time, nonsmooth structure, multi-resolution
  - Feedback mechanisms and interactions, e.g., land, ocean, atmosphere
- Choice of predictors, feature selection, interpretability
  - Set of possible predictors can be large, e.g., in tropics
  - Dependence on climate processes vs simple covariates
  - Nonlinear dependencies, nonlinear dynamics
- Oscillators, teleconnections
  - Phases of oscillators may jointly determine `state'
  - Low frequency variability
- Quantiles, Extremes
  - Understanding tail behavior
- Predictability

# Climate Model Ensembles

# Why ensembles?




[Knutti et al. 2010, J. of Climate]: 2080-99 relative to 1980-99

# Ensembles used in climate science

- Ensembles of opportunity
  - Different models from different modeling groups, e.g. the IPCC ensemble

- Initial condition ensembles
  - Perturb initial conditions of a single model
  - Significant changes possible (cf. Butterfly Effect)
  - "Pure ensemble" – perturb only last few significant digits of an initial condition. Changes the weather but should not change the climate. Used to robustify estimates of climate.

- Perturbed physics ensembles (PPE)
  - Change parameter values of a single model
  - Can create drastic changes in predictions

NOTE: weather forecasting also makes use of ensembles (e.g. Bayesian model averaging).

# Multi-model ensemble used by IPCC

- IPCC: Intergovernmental Panel on Climate Change
  - Nobel Peace Prize 2007 (shared with Al Gore).
  - Interdisciplinary scientific body, formed by UN in 1988.
  - Fourth Assessment Report, 2007, on global climate change

    450 lead authors from 130 countries, 800 contributing authors, over 2,500 reviewers.
  - Fifth Assessment Report, September 2013.

- Climate models contributing to IPCC reports include:

  Bjerknes Center for Climate Research (Norway), Canadian Centre for Climate Modelling and Analysis, Centre National de Recherches Météorologiques (France), Commonwealth Scientific and Industrial Research Organisation (Australia), Geophysical Fluid Dynamics Laboratory (Princeton University), Goddard Institute for Space Studies (NASA), Hadley Centre for Climate Change (United Kingdom Meteorology Office), Institute of Atmospheric Physics (Chinese Academy of Sciences), Institute of Numerical Mathematics Climate Model (Russian Academy of Sciences), Istituto Nazionale di Geofisica e Vulcanologia (Italy), Max Planck Institute (Germany), Meteorological Institute at the University of Bonn (Germany), Meteorological Research Institute (Japan), Model for Interdisciplinary Research on Climate (Japan), National Center for Atmospheric Research (Colorado), among others.

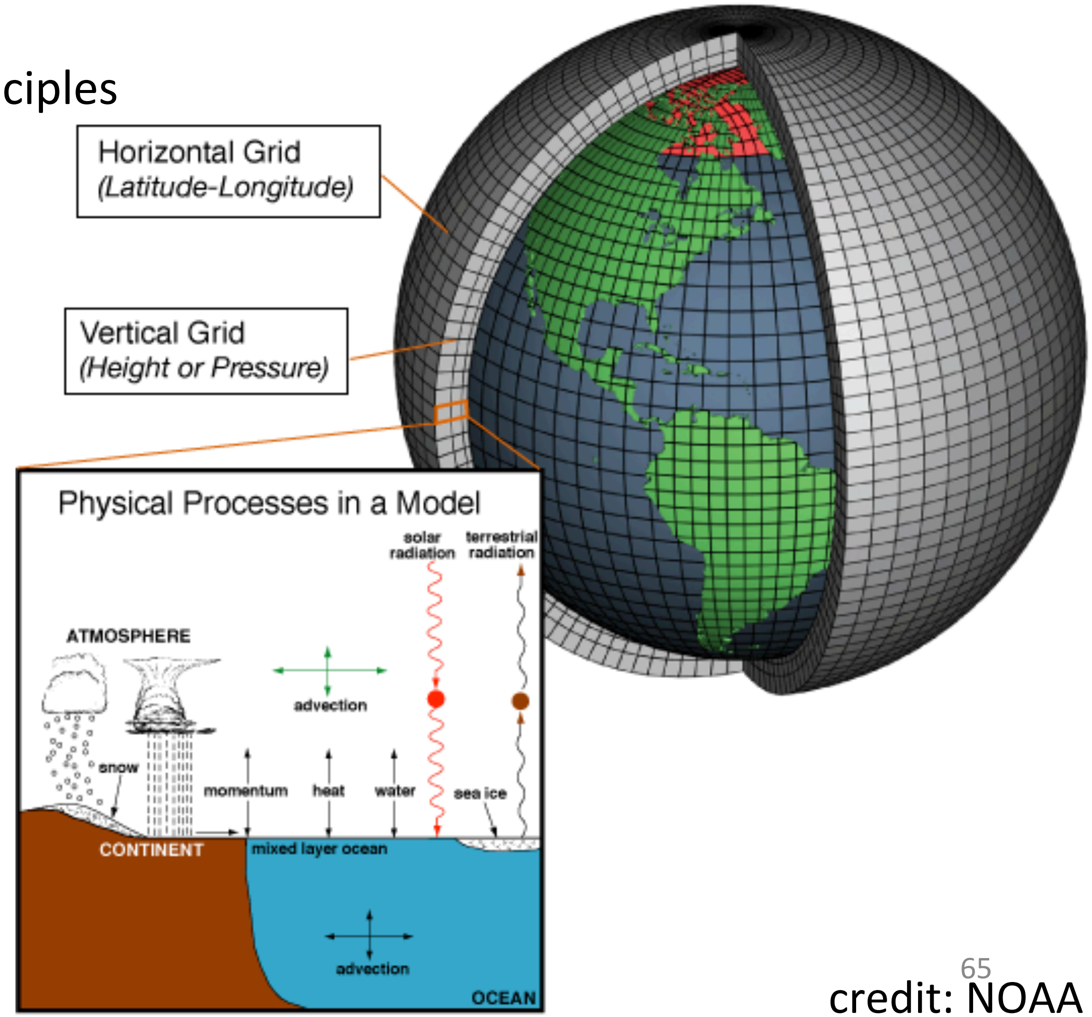
# Climate models (GCMs)

Climate model: a complex system of interacting mathematical models

- Not data-driven
- Based on scientific first principles
  - Meteorology
  - Oceanography
  - Geophysics
  - …

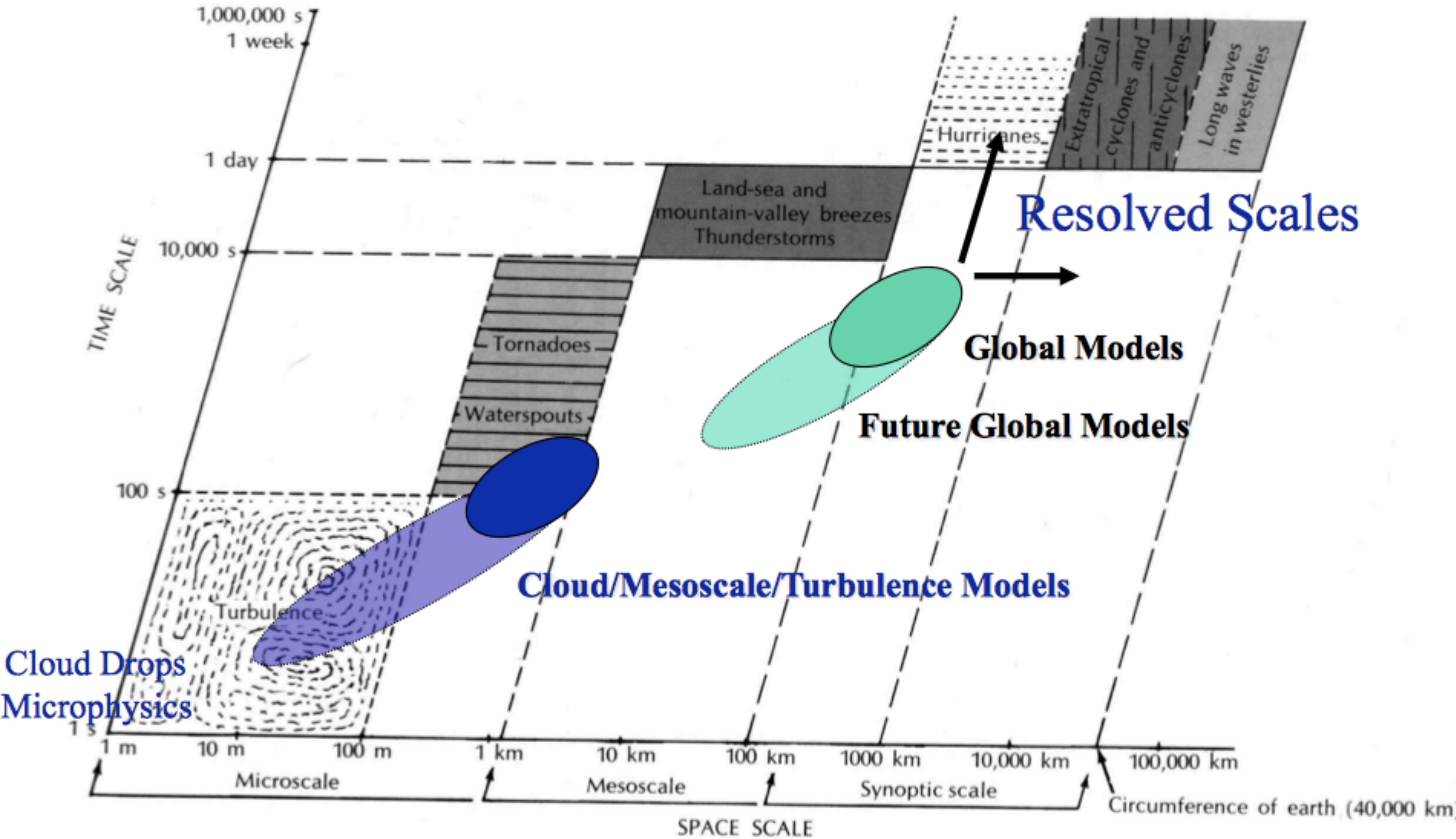
Climate model differences

- Assumptions
- Discretizations
- Scale interactions
  - Micro: rain drop
  - Macro: ocean




credit: NOAA

# Scale resolution problem




credit: Anthes et al. 1975, presented by J. J. Hack/A. Gettelman

# Scale resolution problem

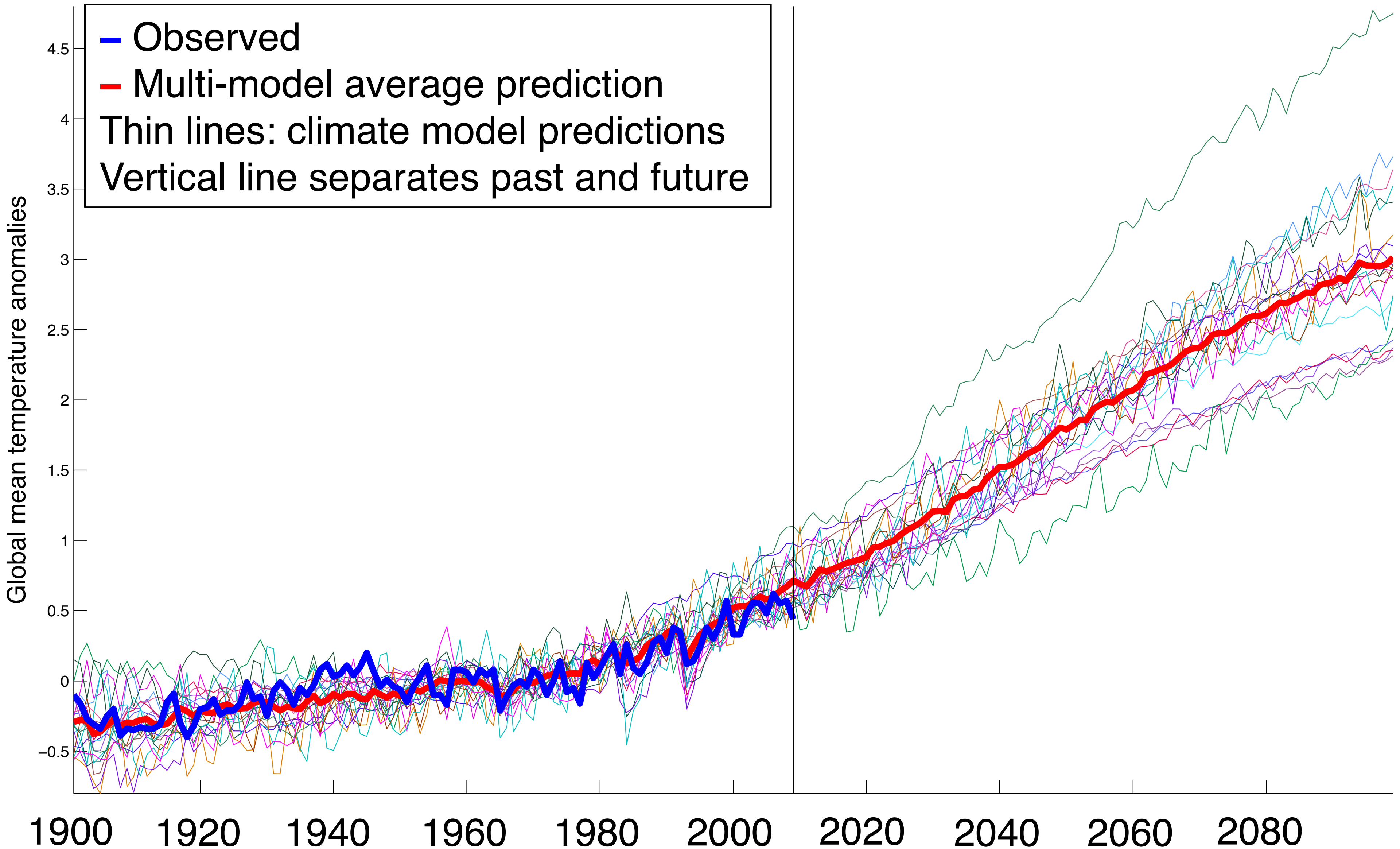
Some important physical processes cannot be resolved at correct scale, and are therefore approximated ("parameterizations").

E.g.

- Moist Processes: Moist convection, shallow convection, large scale condensation
- Radiation and Clouds: Cloud parameterization, radiation
- Surface Fluxes: Fluxes from land, ocean and sea ice (from data or models)
- Turbulent mixing: Planetary boundary layer parameterization, vertical diffusion, gravity wave drag



credit: J. J. Hack/A. Gettelman




Future fan-out.

# Improving predictions of the IPCC ensemble

- Coupled Model Intercomparison Project (CMIP)
  [Meehl et al., Bull. AMS, '00]
- No one model predicts best all the time, for all variables.
- Average prediction over all models is better predictor than any single model. [Reichler & Kim, Bull. AMS '08], [Reifen & Toumi, GRL '09]
- Bayesian approaches in climate science e.g. [Smith et al. JASA '08]
- IPCC held 2010 Expert Meeting on how to better combine model predictions.

Can we do better, using Machine Learning?

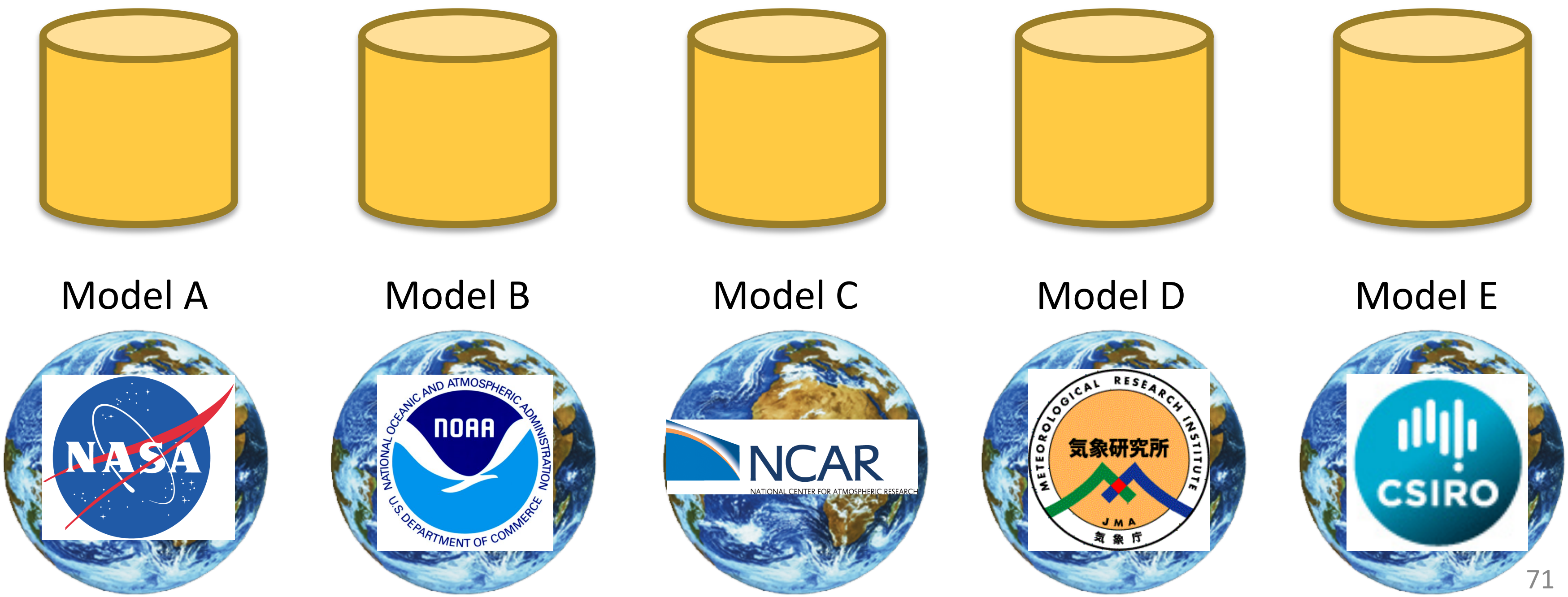
Challenge: How should we predict future climates?

– While taking into account the multi-model ensemble predictions

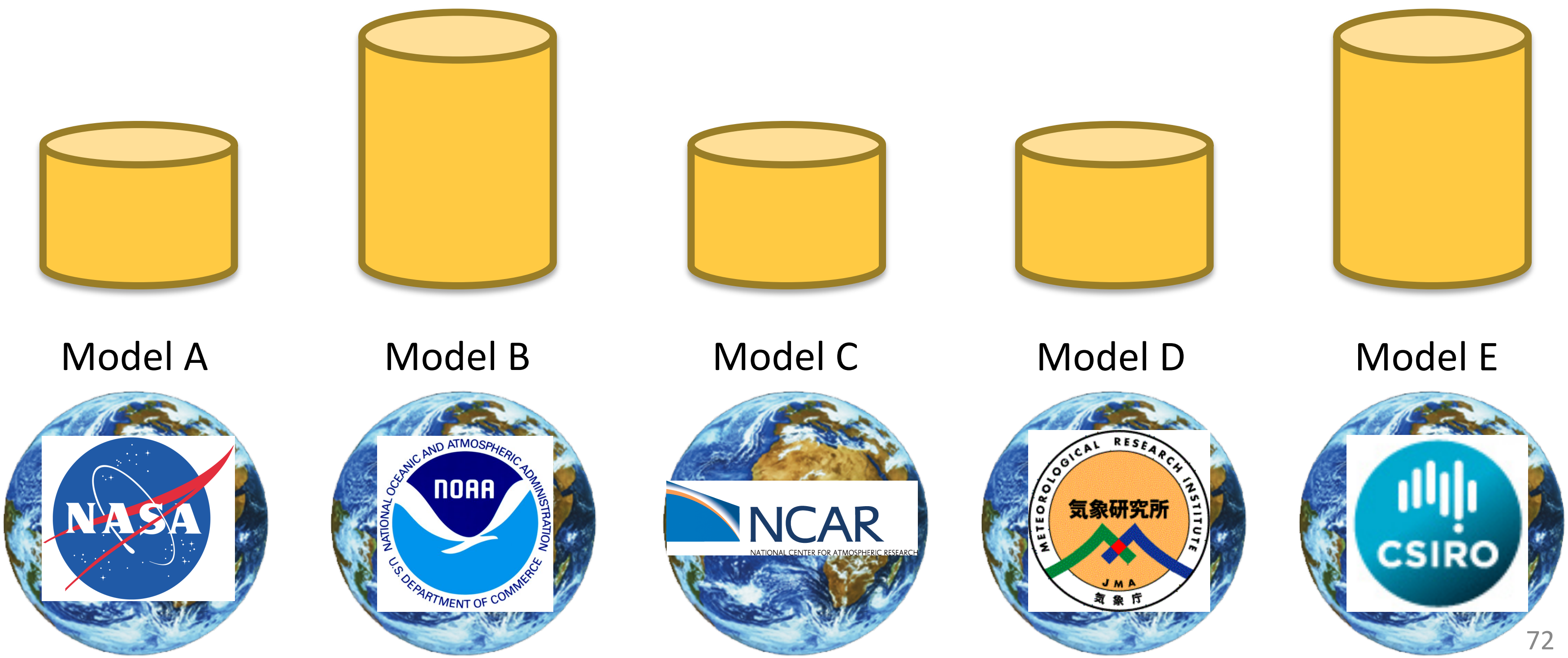
# Machine learning approaches

- Tracking Climate Models (TCM) [M, Schmidt, Saroha, & Asplund, SAM 2011; NASA CIDU 2010]: Online learning with expert advice.
- Neighborhood-Augmented TCM (NTCM) [McQuade & M, AAAI 2012]: Extend TCM to model geospatial neighborhood influence.
- Multi-model regression with spatial smoothing [Subbian & Banerjee, SDM 2012].
- Climate Prediction via Matrix Completion [Ghafarianzadeh & M, Late-Breaking Paper, AAAI 2013]: use sparse matrix completion.
- Multi-task Sparse Structure Learning [Goncalves et al. CIKM 2014].
- MRF-based approach [McQuade & M, submitted 2014].

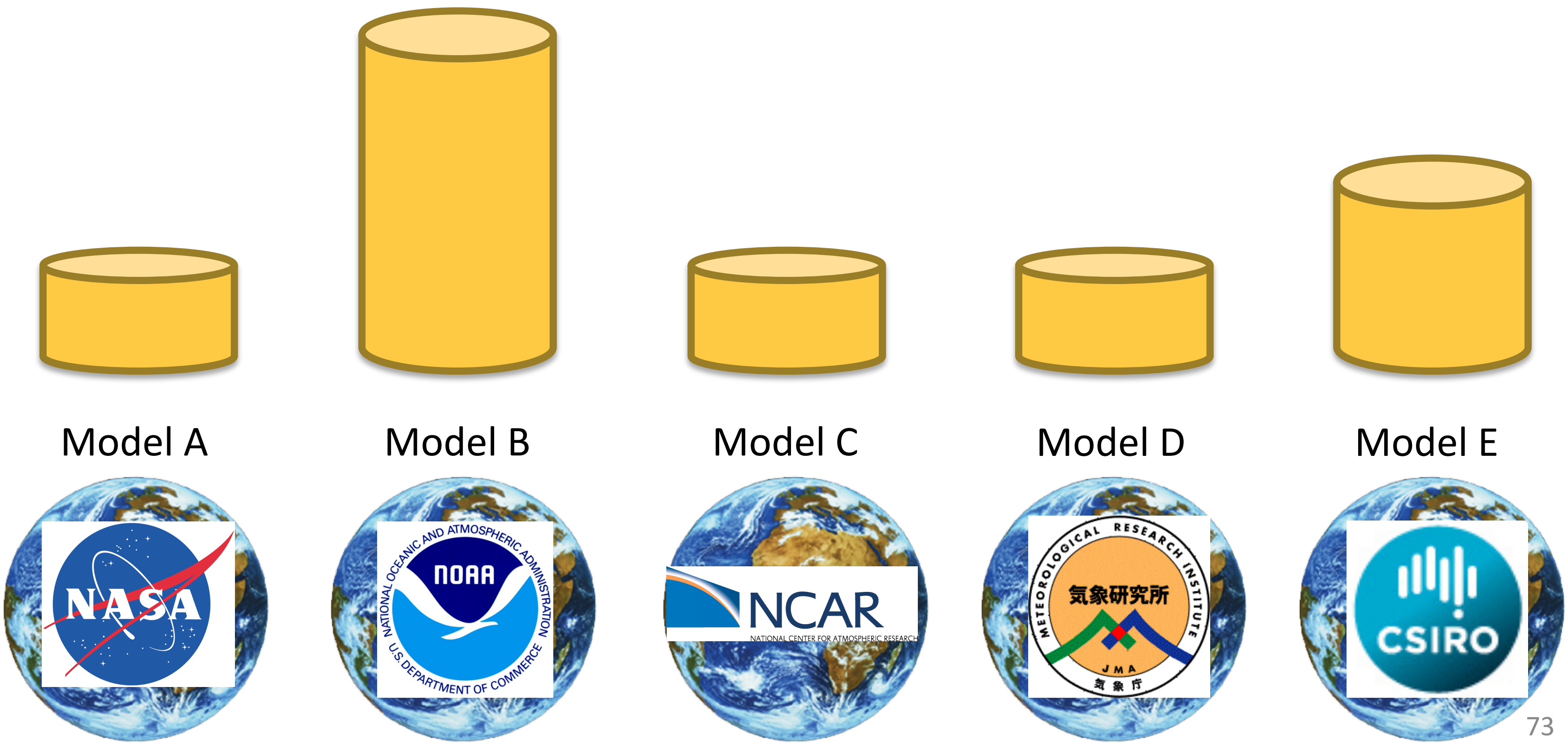
# Average prediction



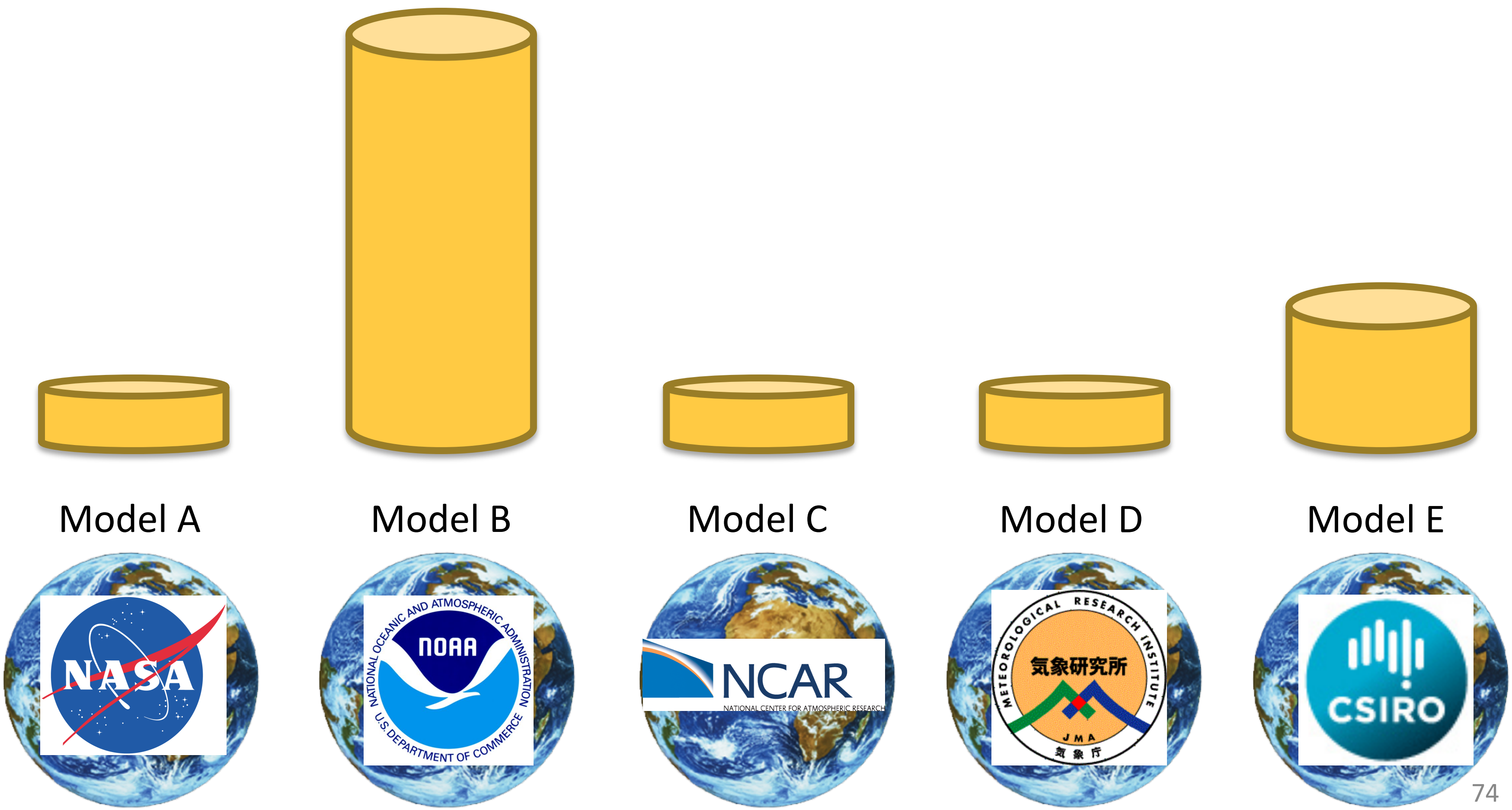
# Adaptive, weighted average prediction



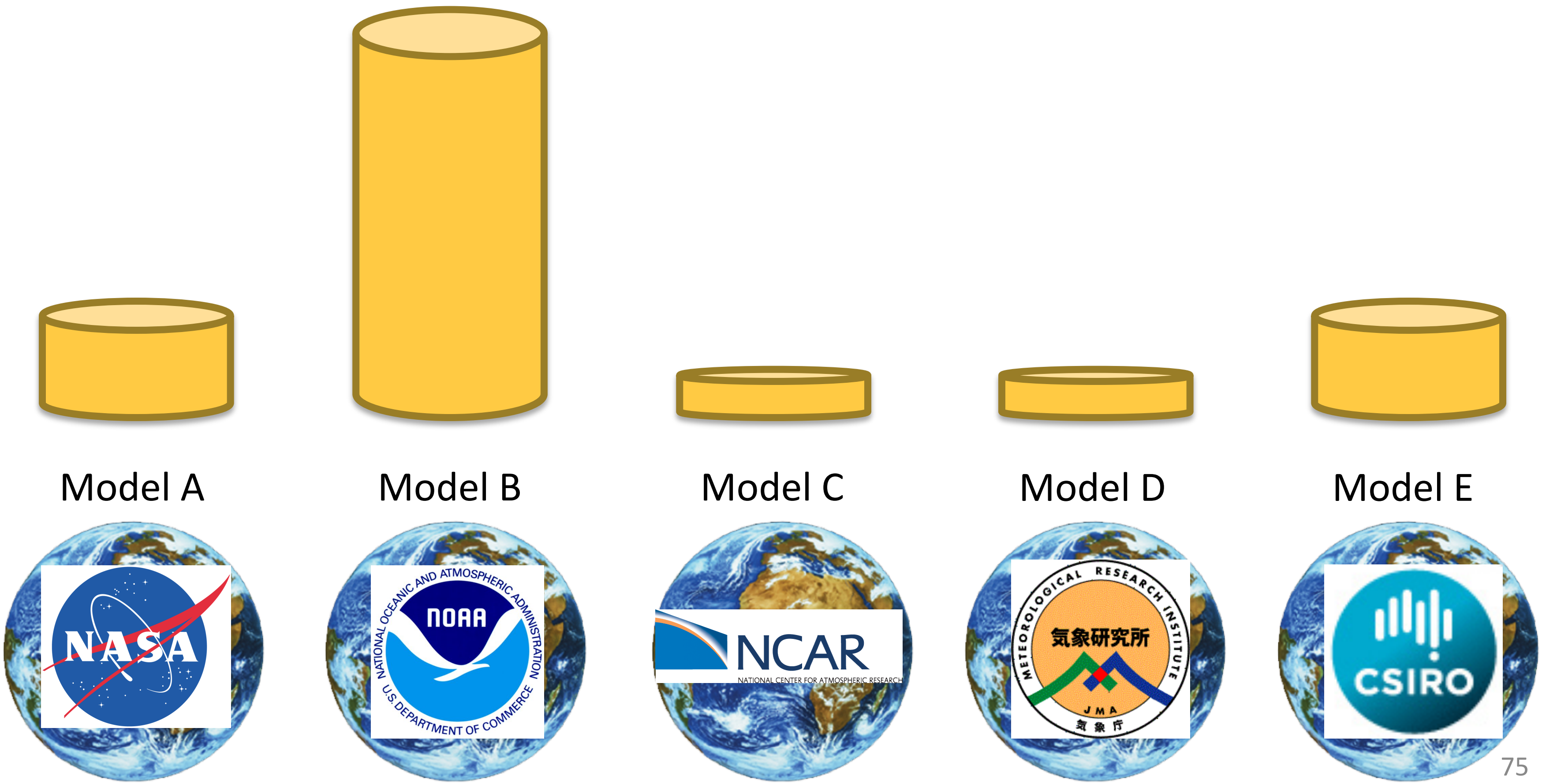
# Adaptive, weighted average prediction



# Adaptive, weighted average prediction



# Adaptive, weighted average prediction

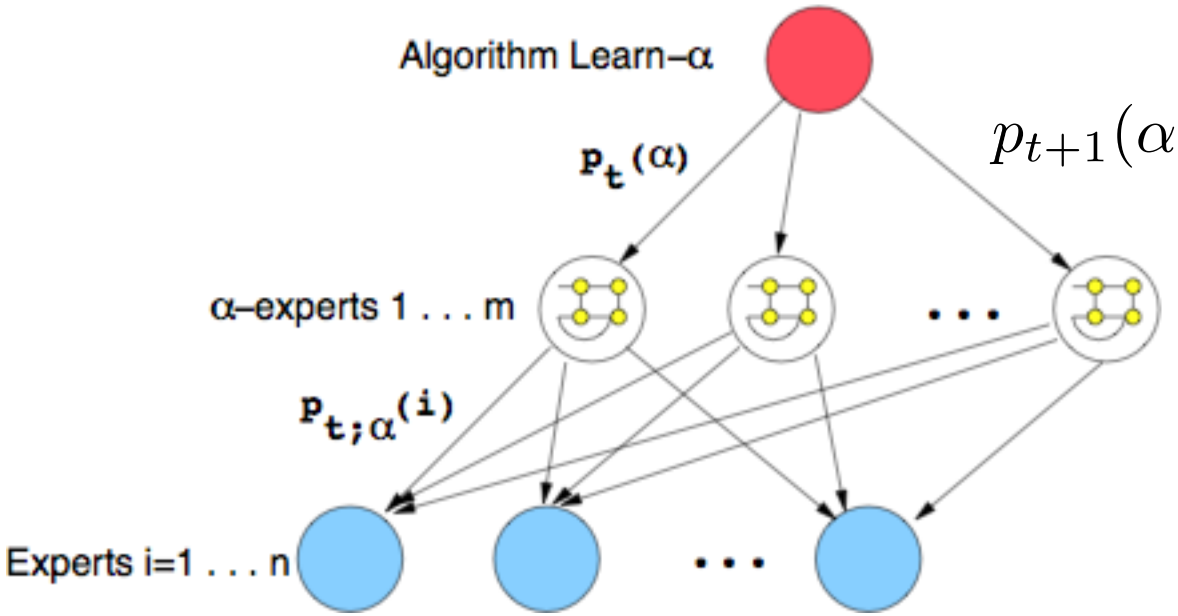


# Tradeoff: explore vs. exploit

Tradeoff: Quickly finding current best predicting model vs. being ready to quickly switch to other models.

Tradeoff hinges on how often the identity of the best model switches.

# Online learning: non-stationary data

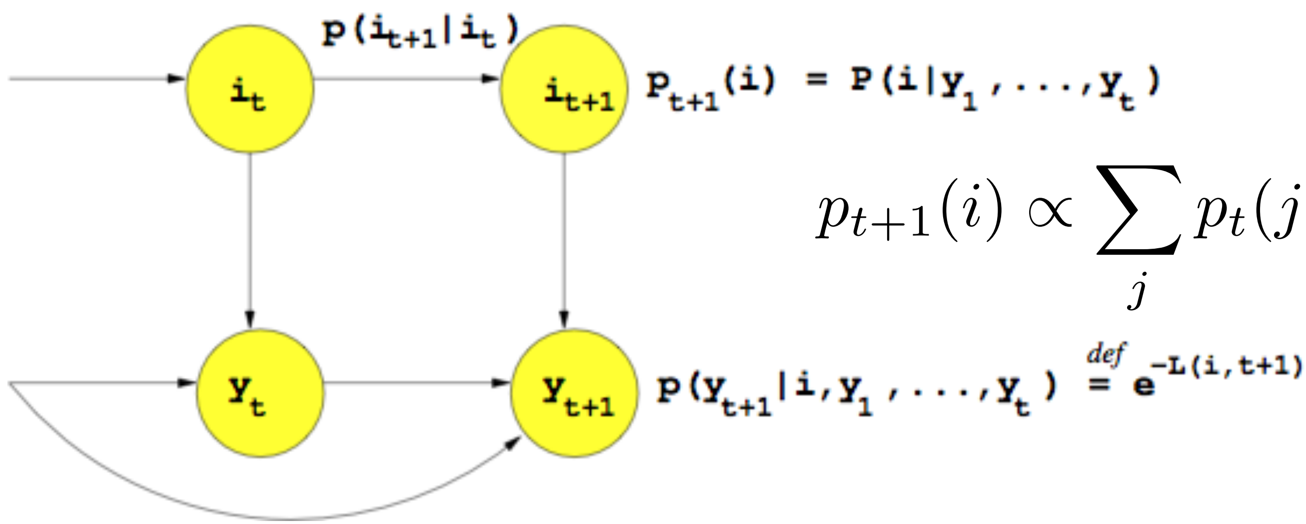



$$p_{t+1}(\alpha) \propto p_t(\alpha)e^{-L(\alpha,t)}$$

Learn-α Algorithm [M & Jaakkola, NIPS 2003]:

- Learns the switching rate: level of non-stationarity: α.
- Tracks a set of meta-experts, online learning algorithms, each with a different value of the α parameter.

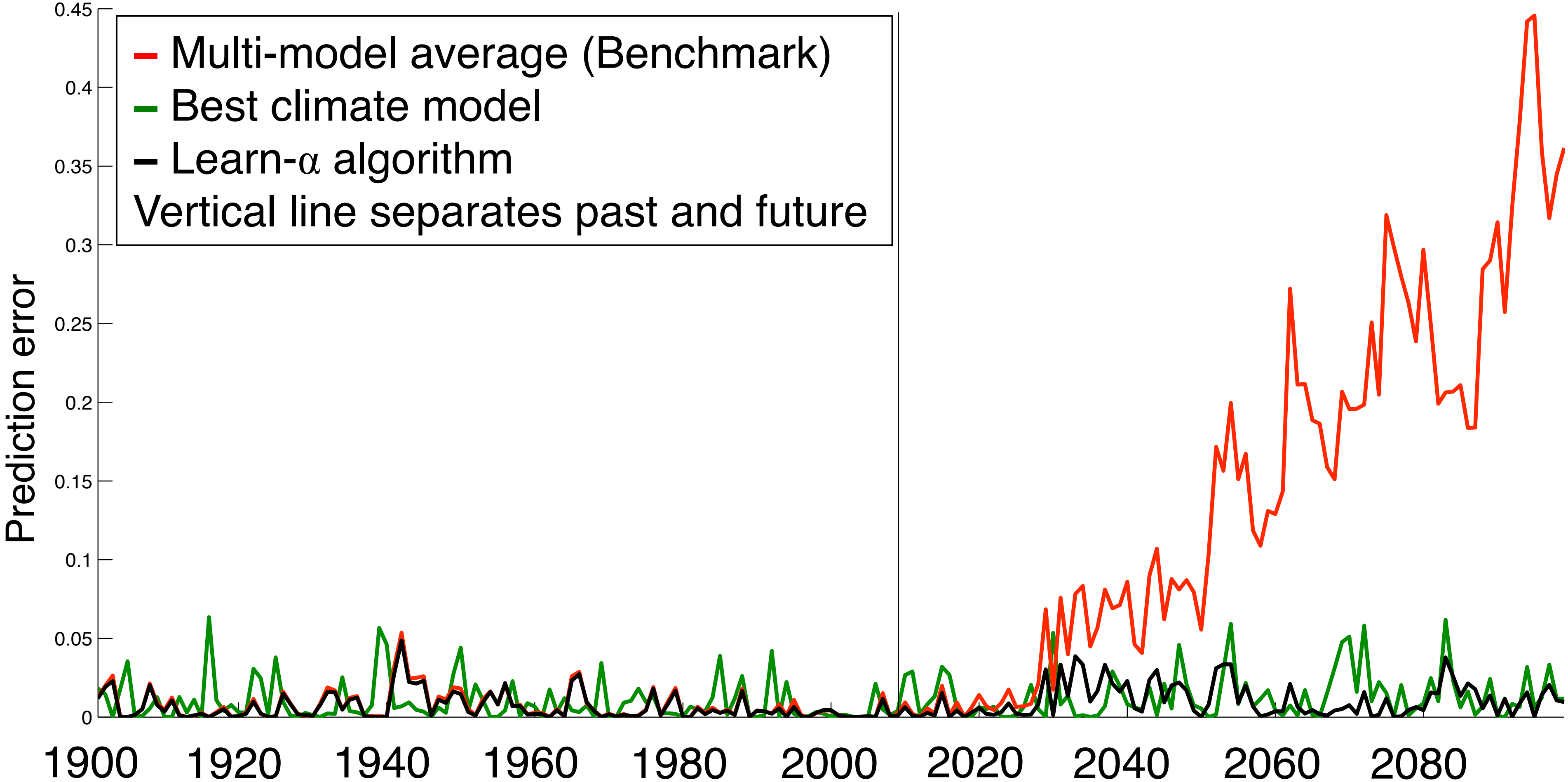
# Online learning: non-stationary data




$$p_{t+1}(i) \propto \sum_j p_t(j) e^{-L(j,t)} p(i|j)$$

- [M & Jaakkola, 2003]: In a family of online learning algorithms, weight updates, $p_t(i)$, equivalent to Bayesian updates of a generalized Hidden Markov Model.
  - Hidden variable: identity of "best expert."
  - Transition dynamics, $p(i \mid j)$, model non-stationarity.
- [Herbster & Warmuth, 1998]: Fixed-Share algorithm models switching w.p. $\alpha$.

$$P(i|j;\alpha) = \begin{cases} (1-\alpha) & i = j \\ \frac{\alpha}{n-1} & i \neq j \end{cases}$$

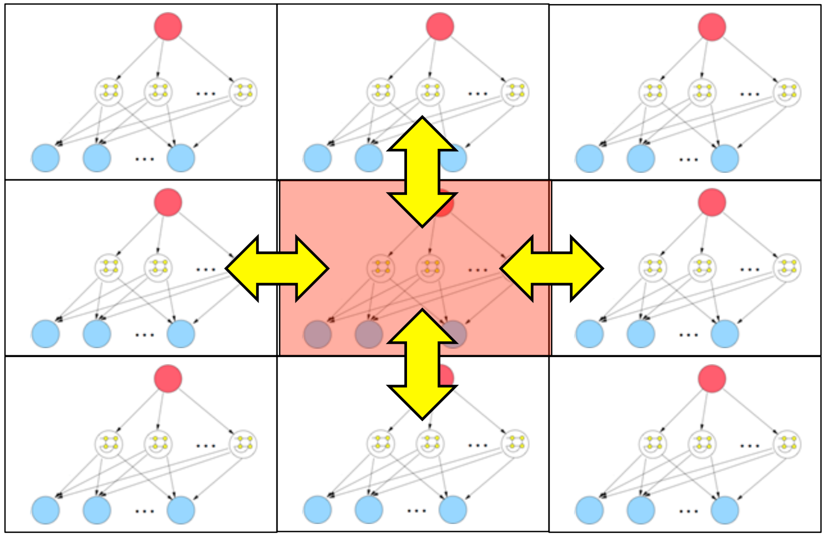



# Learning curves

# Incorporating neighborhood influence

[McQuade & M, AAAI 2012]

- Climate predictions are made at higher geospatial resolutions.
- Run instances of Learn-α (variant) on multiple sub-regions that partition the globe.
- Model neighborhood influences among geospatial regions.



# Incorporating neighborhood influence

Neighborhood-augmented Learn-α.

Non-homogenous HMM transition dynamics:

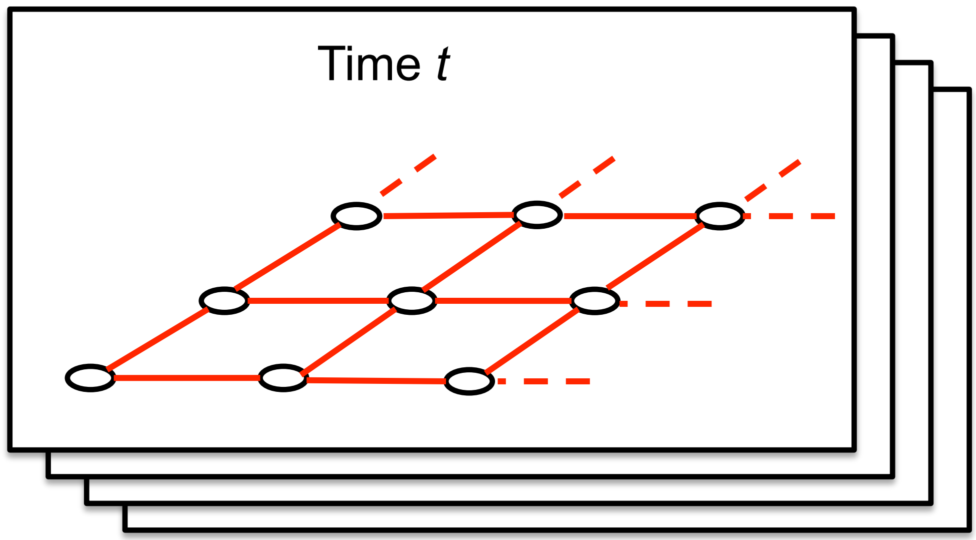
$$P(i \mid k; \alpha) = \begin{cases} (1-\alpha) & \text{if i=k} \\ \frac{\alpha}{Z}\left[(1-\beta) + \beta \frac{1}{|S(r)|} \sum_{s \in S(r)} P_{t,s}(i)\right] & \text{if i}\neq\text{k} \end{cases}$$

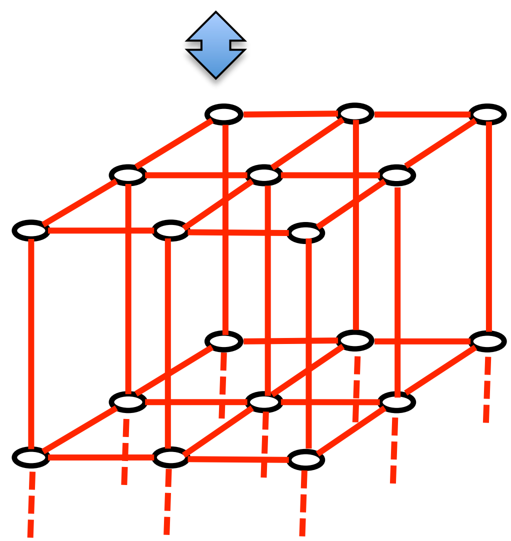
- *S(r)* - neighborhood scheme: set of "neighbors" of region *r*
- $P_{t,s}(i)$ - probability of expert (climate model) *i* in region *s*
- $\beta$ - regulates geospatial influence
- *Z* - normalization factor

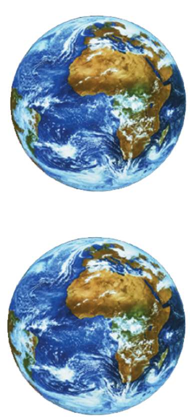
# MRF-based approach

[McQuade & M, submitted]

Geospatial lattice



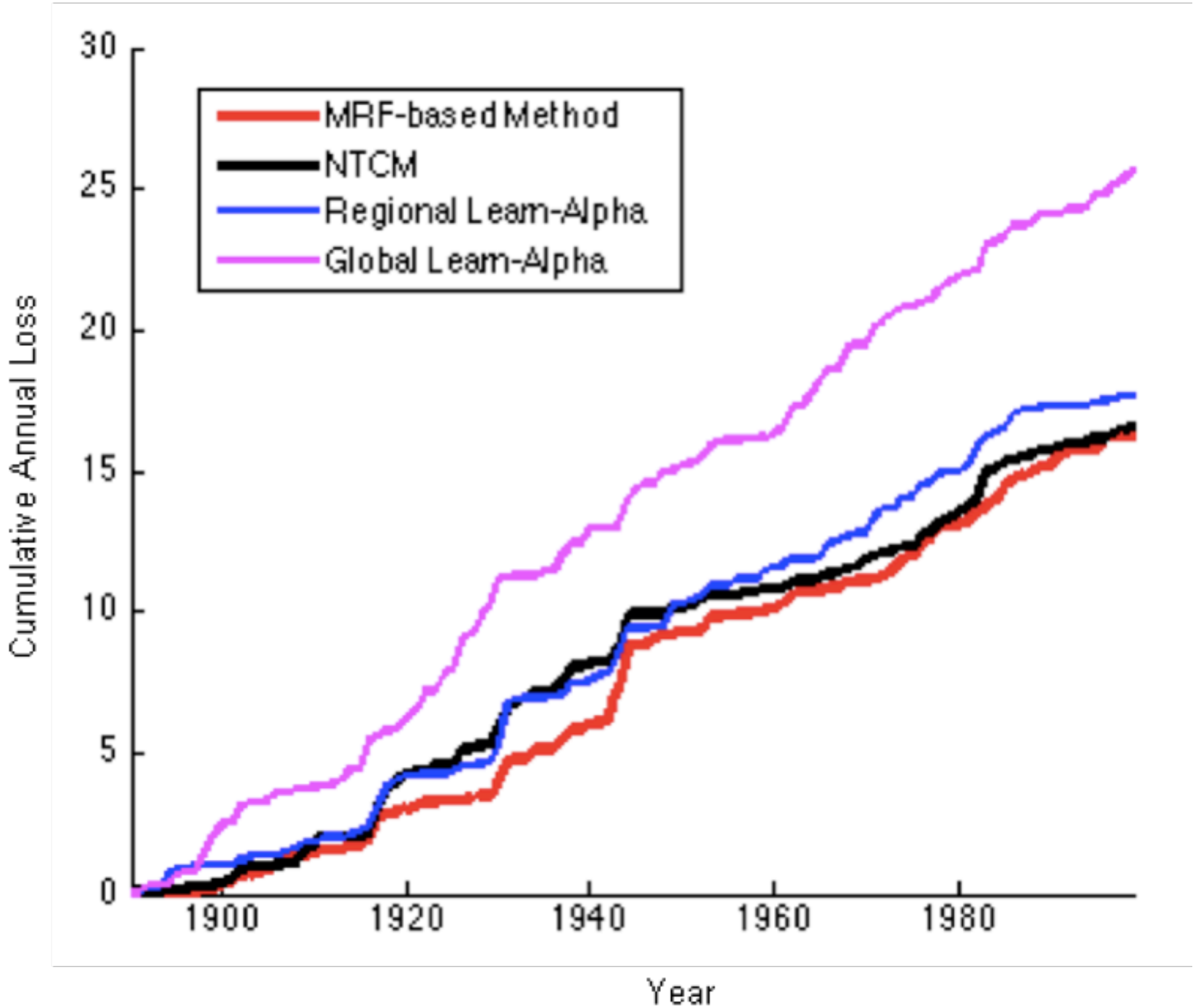





Time $t$

Time $t$-1

# MRF-based approach



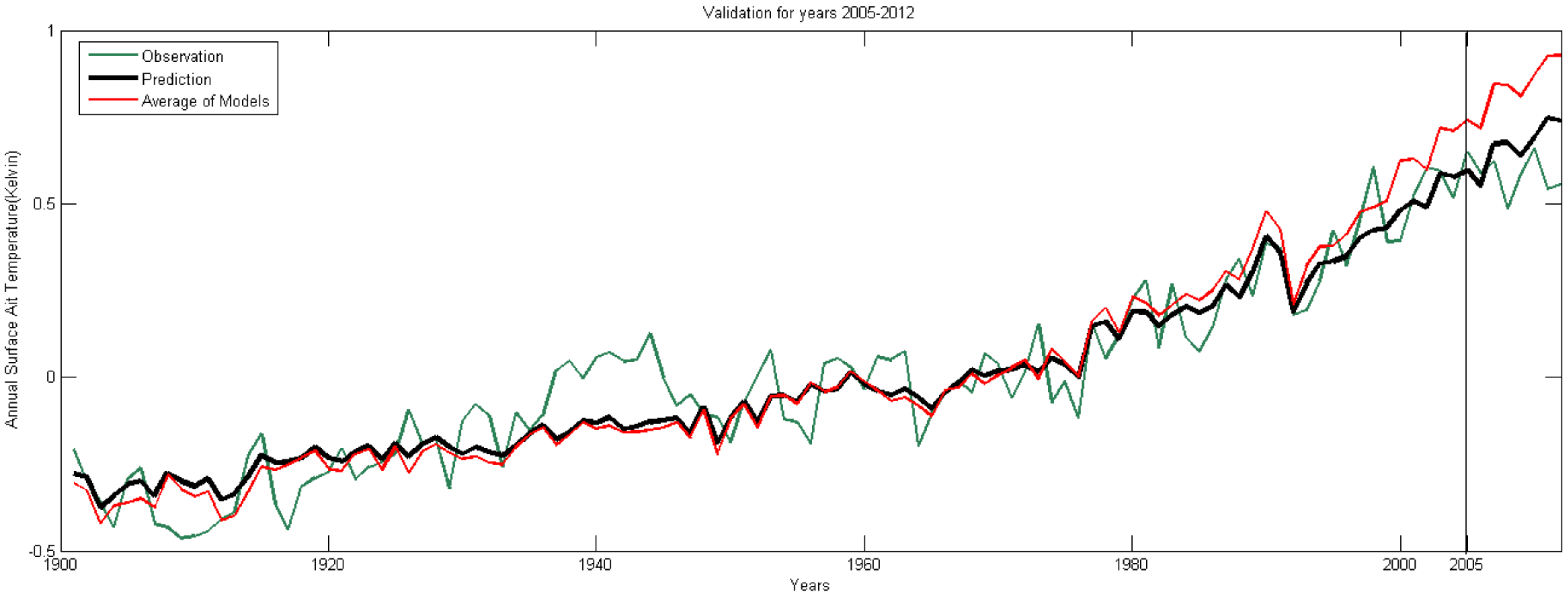
# Climate Prediction via Matrix Completion

[Ghafarianzadeh & M, Late-Breaking Paper, AAAI 2013]

- Goal: combine/improve the predictions of the multi-model ensemble of GCMs, using sparse matrix completion.
- Exploits past observations, and the predictions of the multi-model ensemble of GCMs.
- Learning approach is batch, unsupervised.
- Create a sparse (incomplete) matrix from climate model predictions and observed temperature data.
- Apply a matrix completion algorithm to recover it.

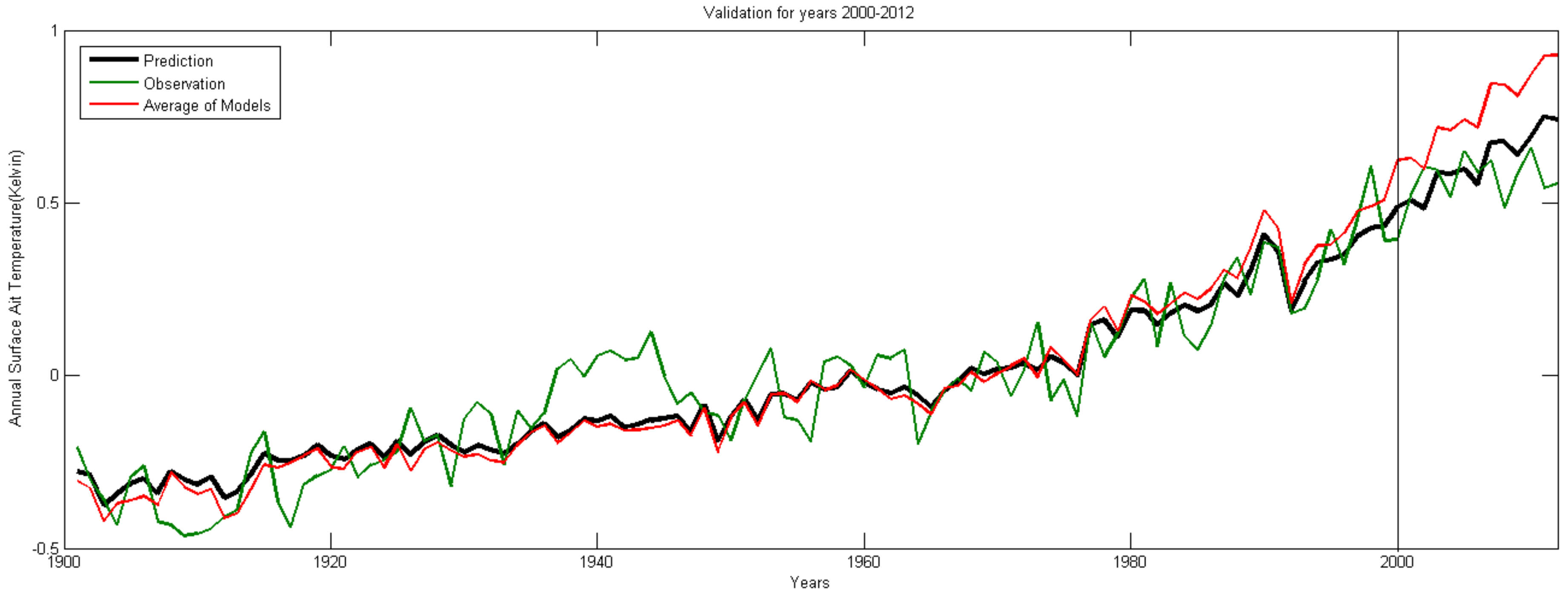
  [Keshavan, Montanari & Oh, JMLR '10] OptSpace algorithm: minimization of nuclear norm; uses spectral techniques and manifold optimization

- Yields predictions of unobserved temperatures.



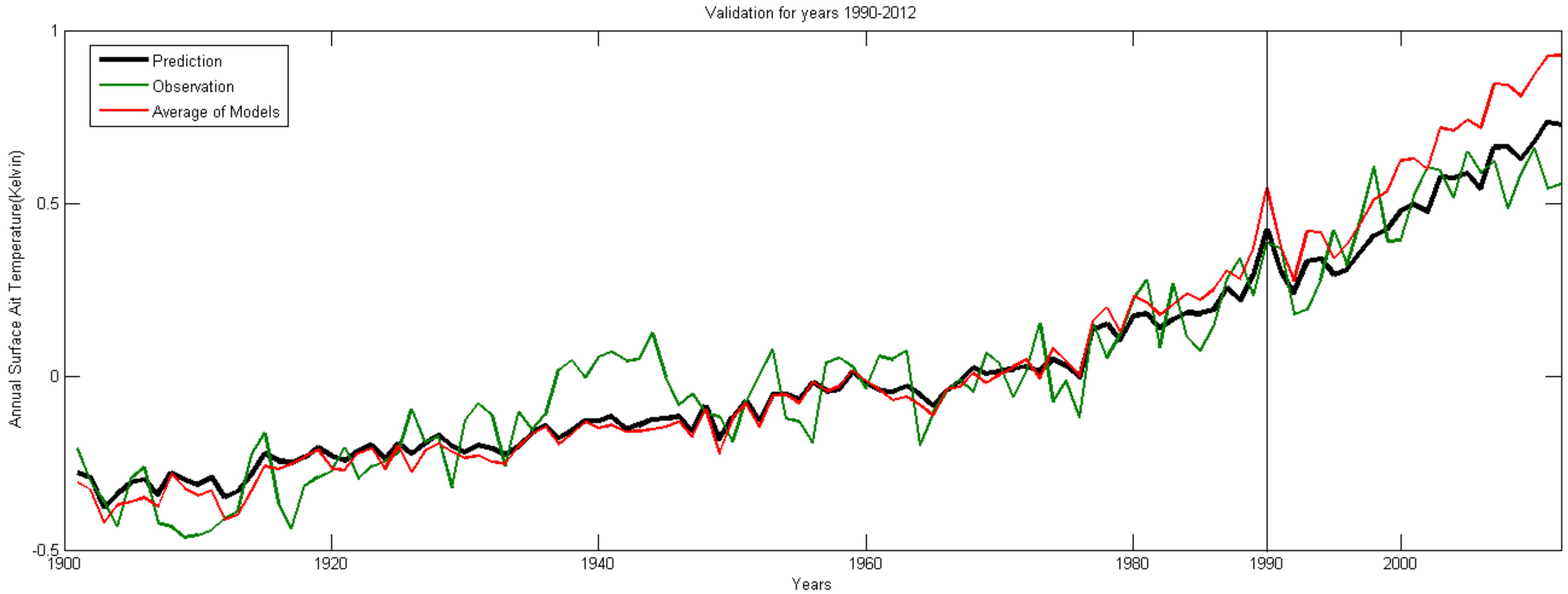

Green: observation, Red: mean prediction of climate models, Black: matrix completion

# Validation period: 2005-2012



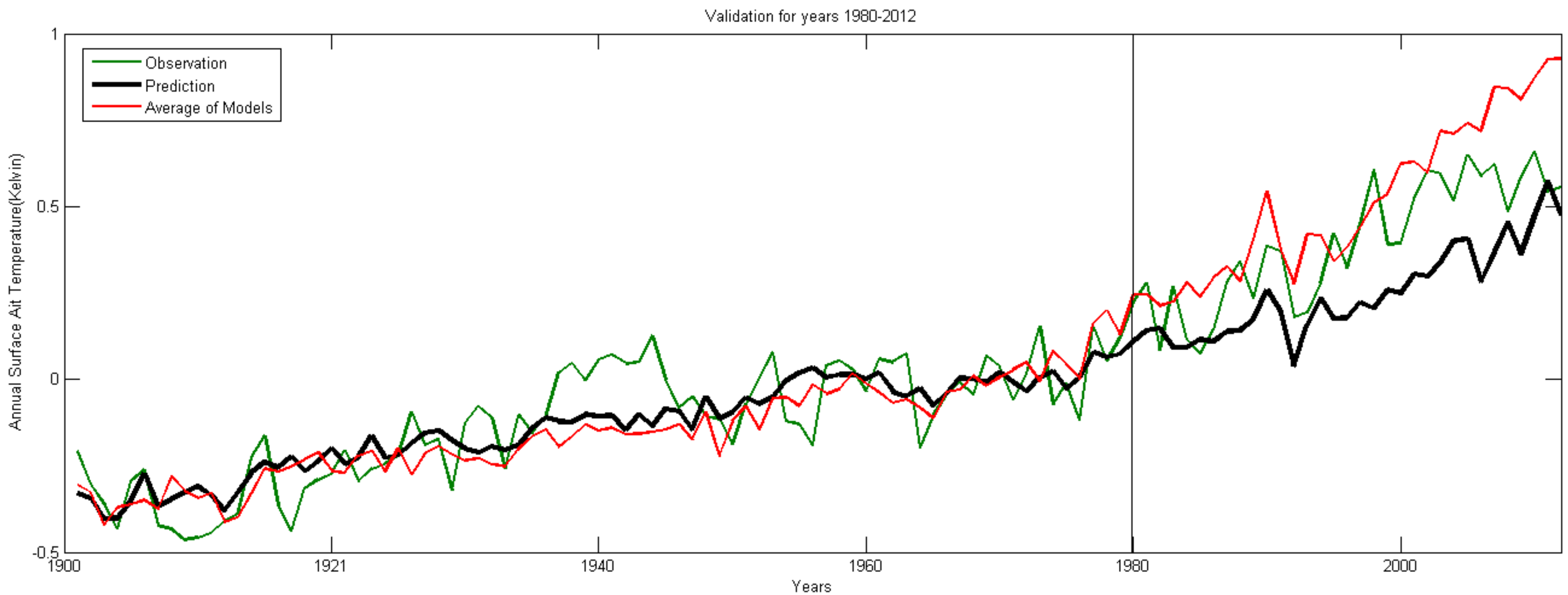

Green: observation, Red: mean prediction of climate models, Black: matrix completion

# Validation period: 2000-2012



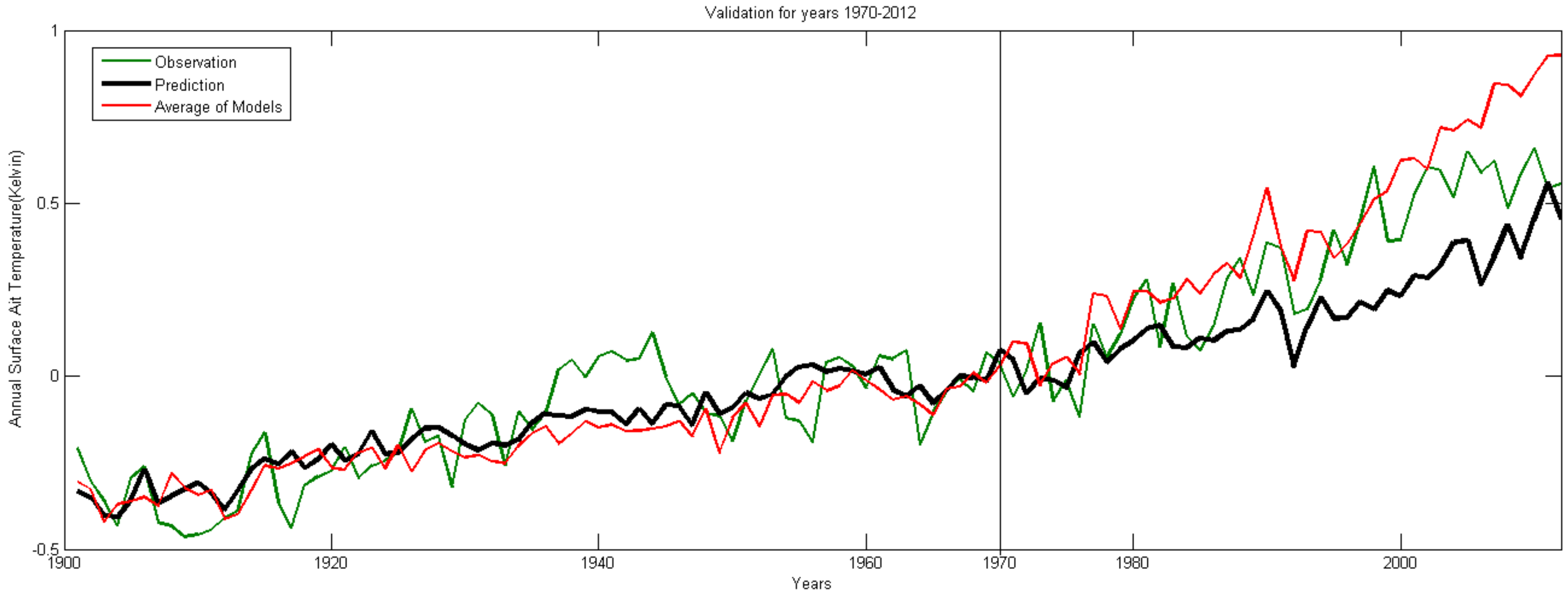

Green: observation, Red: mean prediction of climate models, Black: matrix completion

# Validation period: 1990-2012




Green: observation, Red: mean prediction of climate models, Black: matrix completion

# Validation period: 1980-2012




Green: observation, Red: mean prediction of climate models, Black: matrix completion

# Validation period: 1970-2012

# Outlook

- These results suggest some low intrinsic dimensionality.
- We induced some sparsity in the input matrix
  - Need not ensure low intrinsic dimensionality
- [Jia, DelSole & Tippett, J. Climate '13] also suggest low intrinsic dimensionality:
  - Only a small number (~2) climatological "predictive components" [DelSole & Tippett, Rev. Geophys. '07] determine the predictive "skill" of climate models (measured w.r.t. observations).
    - General warming trend, and El Niño-Southern Oscillation
- GCM ensemble (or subsets) as lower dimensional subspace
  - Can serve as a proxy for the high dimensional, complicated (dependencies, redundancies) space of climatological components in each GCM.
- Suggests future work on tracking a small subset of the ensemble.
  - Subset can change over time and space

# Challenges in climate modeling

Challenge: Improve the predictions of the multi-model ensemble

- Online learning approaches
  - Tracking a small subset of the ensemble, changing over time and space
  - {Semi,un}-supervised online learning with experts
- Challenge: try other ML approaches!
- Challenge: predict multiple variables simultaneously
- Challenge: Calibrate and compare climate models in a principled manner

Challenge: Improve the predictions of an individual climate model

- Challenge: resolve scale interactions ("climate model parameterization")
  - Multiscale algorithms for multiscale climate interactions
- Challenge: harness both physics and data!
  - Hybrid methods between physics-based models and data driven models
  - Data assimilation

# Climate Extremes

# How to define extremes?

① Threshold in single variable [IPCC special report 2012, p.4]

② Multiple degrees of severity

③ Related to multiple variables (complex extreme events)

④ Accumulation of non-extremes [IPCC 2012, p.6]

⑤ Subject to local climate characteristics [IPCC 2012, p.7]

# Machine learning approaches

- Causal attribution of climate extremes [Lozano, Li, Niculescu-Mizil, Liu, Perlich, Hosking, Abe, KDD 2009].
- Copula-Granger causality [Chen, Liu, Liu & Carbonell, AAAI ’10] for non-Gaussian time series.
- Sparse-GEV [Liu, Bahadori, Li, ICML 2012]. Latent state model for generalized extreme value time-series.
- Drought detection using MRFs [Fu, Banerjee, Liess, Snyder, SDM 2012].
- Unsupervised detection of extreme events via topic modeling [Tang & M, Climate Informatics 2014].

# Causal attribution of temp. extremes

Challenge: which factors have contributed to temperature extremes and to what extent?

[Lozano et al. KDD 2009]:
Which factors Granger-caused extreme temperatures?

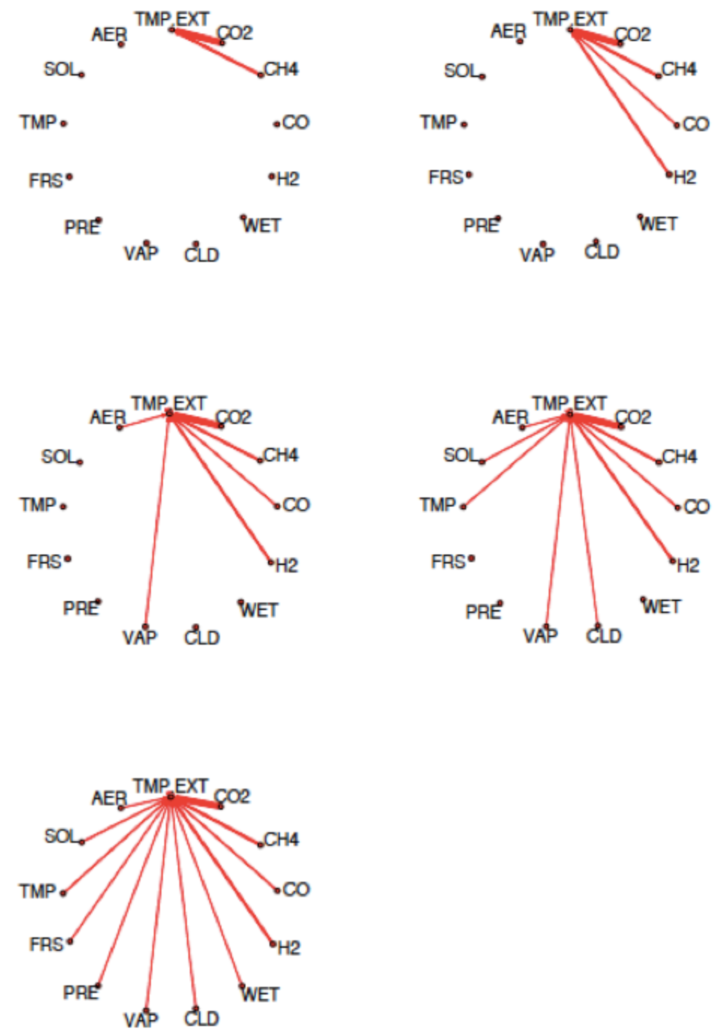
- Granger-causality: X Granger-causes Y if past values of $X_{1:t}$ influence $Y_{t+1}$
- Extension to spatiotemporal variables via group Lasso [Arnold et al. KDD '07]
- Imposes sparsity, and spatial penalties enforce spatial smoothness, neighborhood similarity
- Takes estimates from existing GEV models as input

[Liu, Bahadori & Li, ICML, 2012]: Sparse-GEV models to directly infer the (sparse temporal) dependence structure in multivariate extreme value time series.

# Causal attribution of temp. extremes

[Lozano et al. KDD 2009]:

| Variables (Variable group) | Type | Source |
|---|---|---|
| Methane ($CH_4$)<br>Carbon-Dioxide ($CO_2$)<br>Hydrogen ($H_2$)<br>Carbon-Monoxide (CO) | Greenhouse Gases | NOAA |
| UV (AER) | Aerosol Index | NASA |
| Temperature (TMP)<br>Temp Range (TMP)<br>Temp Min (TMP)<br>Temp Max (TMP)<br>Precipitation (PRE)<br>Vapor (VAP)<br>Cloud Cover (CLD)<br>Wet Days (WET)<br>Frost Days (FRS) | Climate | CRU |
| Global Horizontal (SOL)<br>Direct Normal (SOL)<br>Global Extraterrestrial (SOL)<br>Direct Extraterrestrial (SOL) | Solar Radiation | NCDC |
| 1-year return level for temperature extreme (TMP.EXT) | Climate | Estimated using temp from CDIAC |

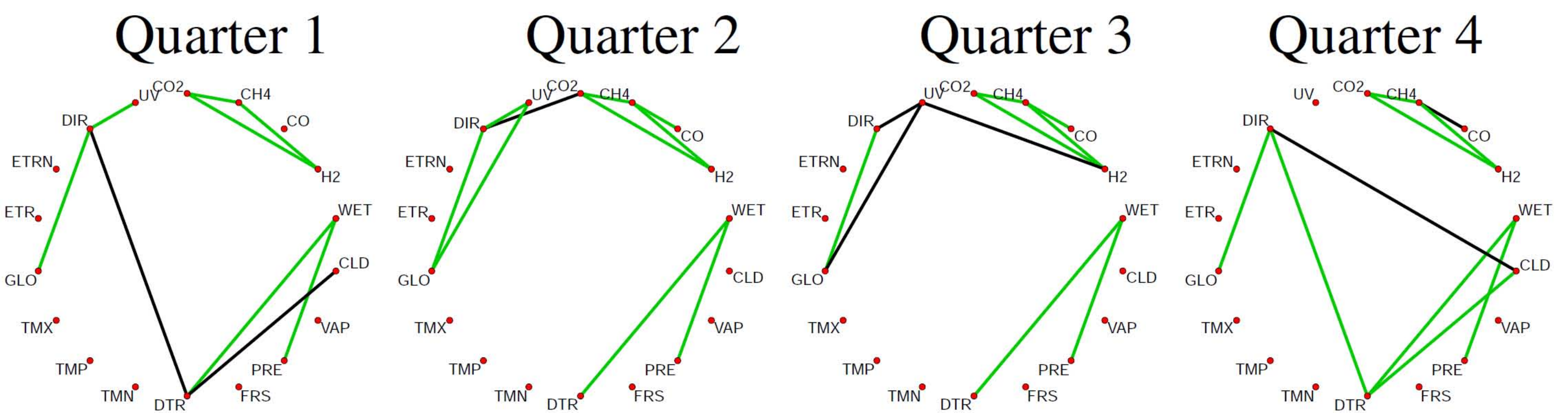



[Liu at al., ICML 2010]: Granger-causal attribution using MRF-based sparse regression.

# Copula-Granger Causality

[Chen, Liu, Liu & Carbonell, AAAI '10] [Bahadori & Liu, SDM '13]

- How to handle non-Gaussian spatiotemporal climate data?
- Map observations to (non-paranormal) Gaussian Copula space:
  - Assume $(f_1(X_1),\ldots,f_p(X_p)) \sim N(\mu,\Sigma)$, where $\{f_j\}$ univariate, monotone
- Determine Granger causality over these copula data
  - Using lasso variants applied to $(f_1(X_1(t)),\ldots,f_p(X_p(t))$

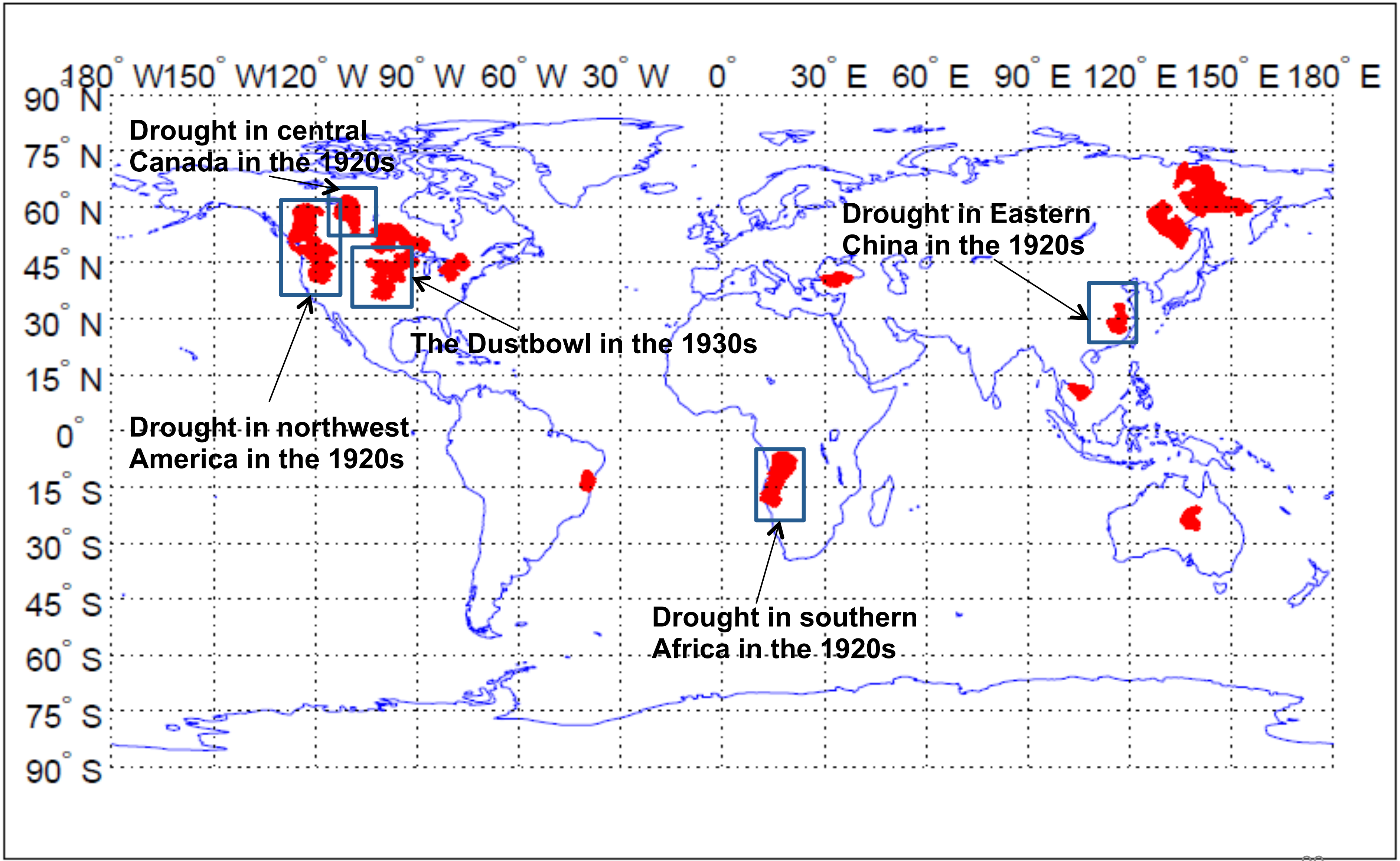


# Mega-Droughts

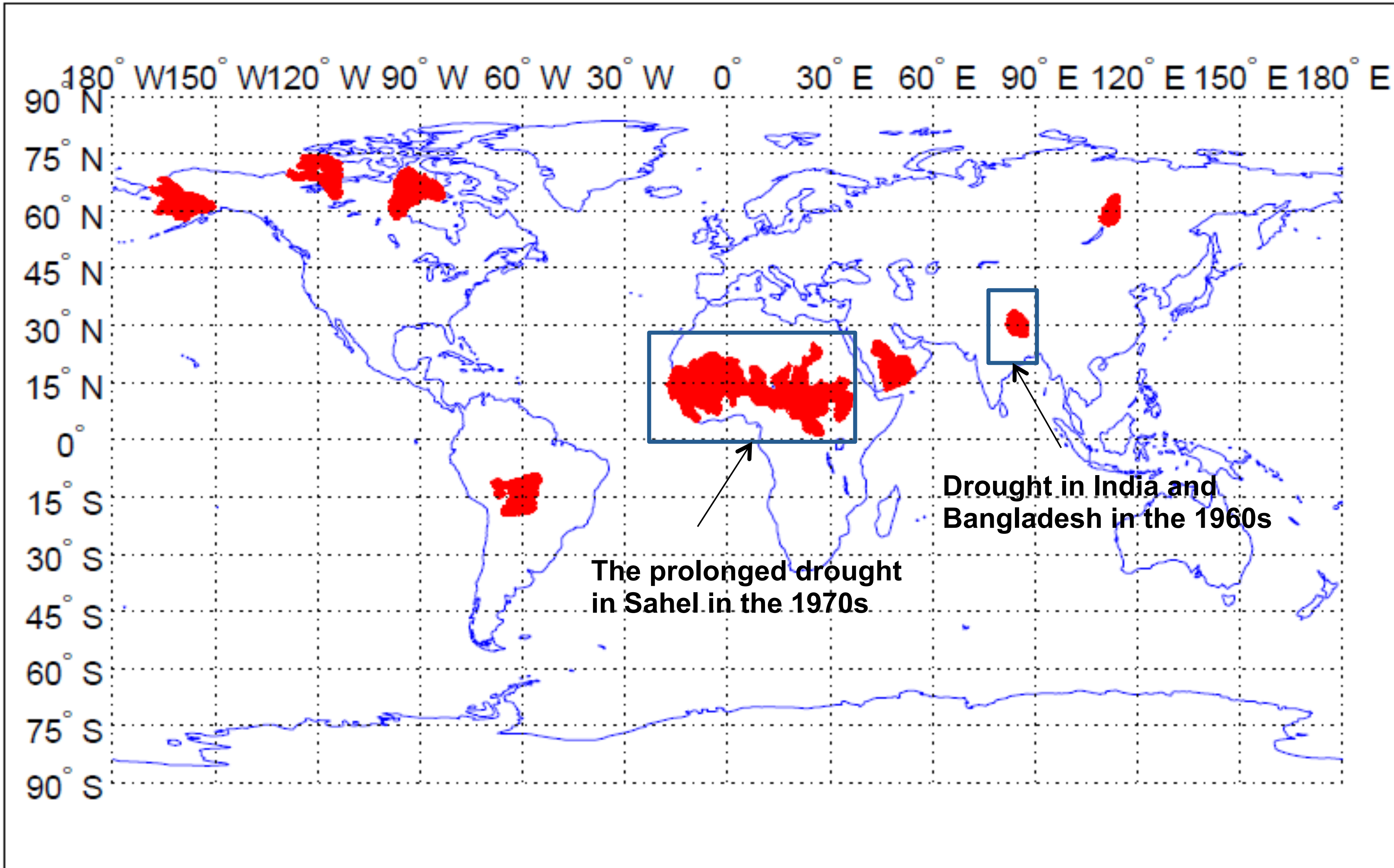
- Mega-Droughts
  - Persistent over space and time
  - Catastrophic consequences
- Examples
  - Late 1906s Sahel drought
  - 1930s North American Dust Bowl
- Related work in climate science, e.g. Palmer Drought Severity Index
  - [Palmer '65]: Geophysical index, primarily based on soil moisture
- Discrete hidden Markov random field (HMRF)

  [Fu et al., SDM '12, Wang et al., UAI'13, Wang et al., NIPS'14]
  - Each latent node $z_i$ is "wet" or "dry," observed $x_i$ is precipitation
  - MRF gives smoothness in space and time
    - MAP inference with two states: "wet" or "dry"
    - Post process to find significant space-time "dry" regions
    - Significantly outperforms naïve thresholding algorithm

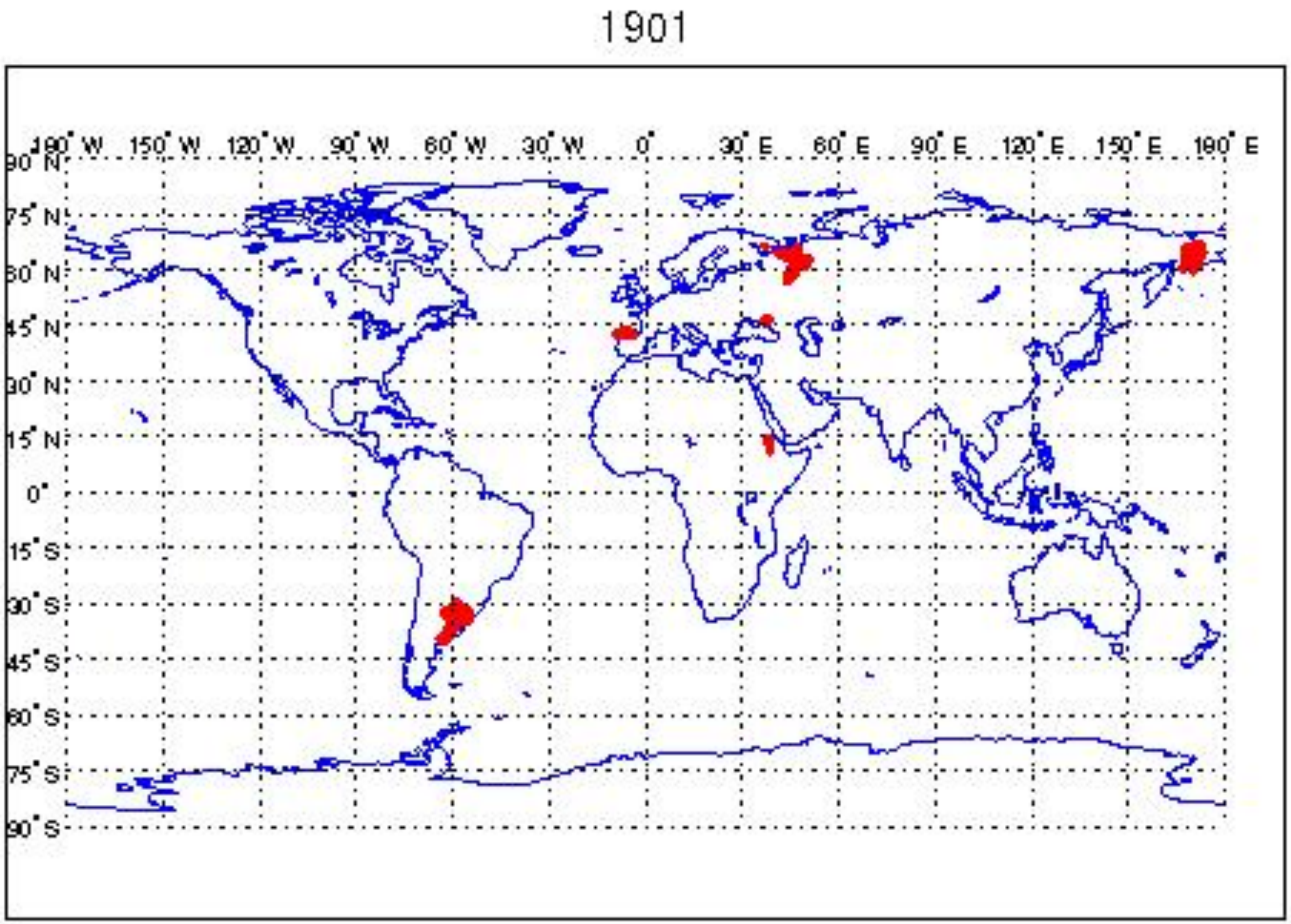
# Results: Droughts starting in 1920-30s



# Results: Droughts starting in 1960-70s



# Major Droughts: 1901-2006

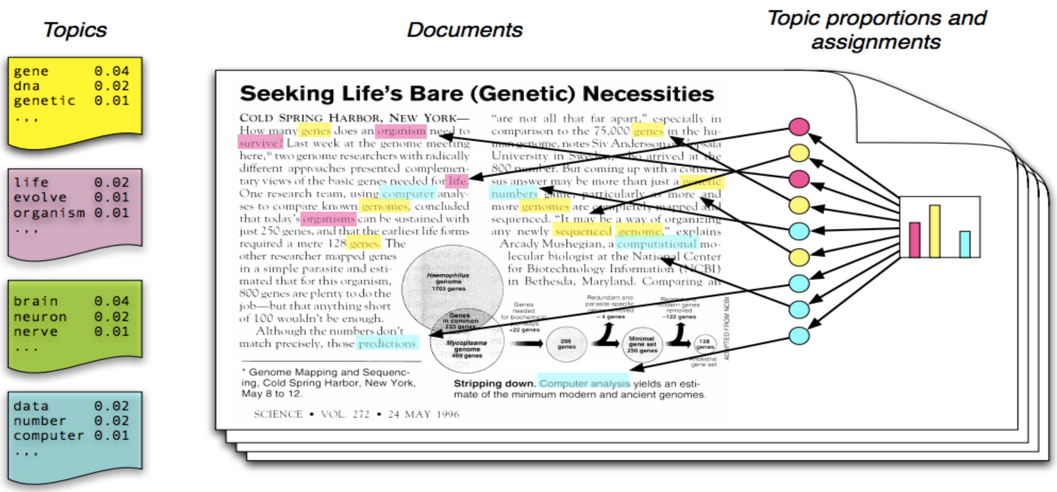


# Topic modeling approach

[Tang & M, Climate Informatics 2014]

| | | |
|---|---|---|
| Geophysical Models | Statistical Models | Model |
| Extreme and Non-extreme values | Extreme values | Data type |
| Single variable | Multiple variables | Variables |
| Single event type | Multiple event types | Events |

# Climate topic modeling




Document
(bag of words)

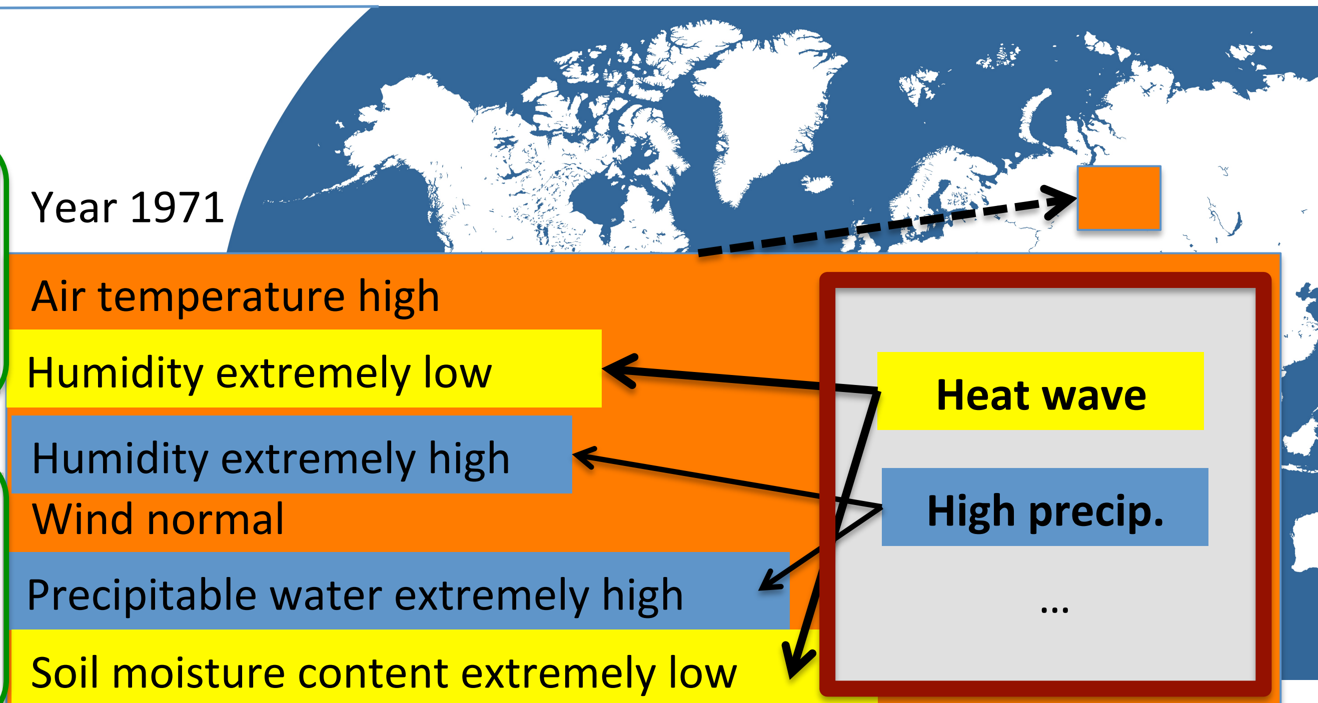
Topics

Words

credit: D. Blei

| TOPIC_4 | 0.25196 |
|---|---|
| shum1 | 0.18842 |
| pr_wtr1 | 0.16720 |
| slp5 | 0.15101 |
| rhum1 | 0.13596 |
| pres5 | 0.13455 |

| TOPIC_2 | 0.24983 |
|---|---|
| rhum5 | 0.19172 |
| pr_wtr5 | 0.18384 |
| shum5 | 0.15476 |
| slp1 | 0.12487 |
| pres1 | 0.11816 |

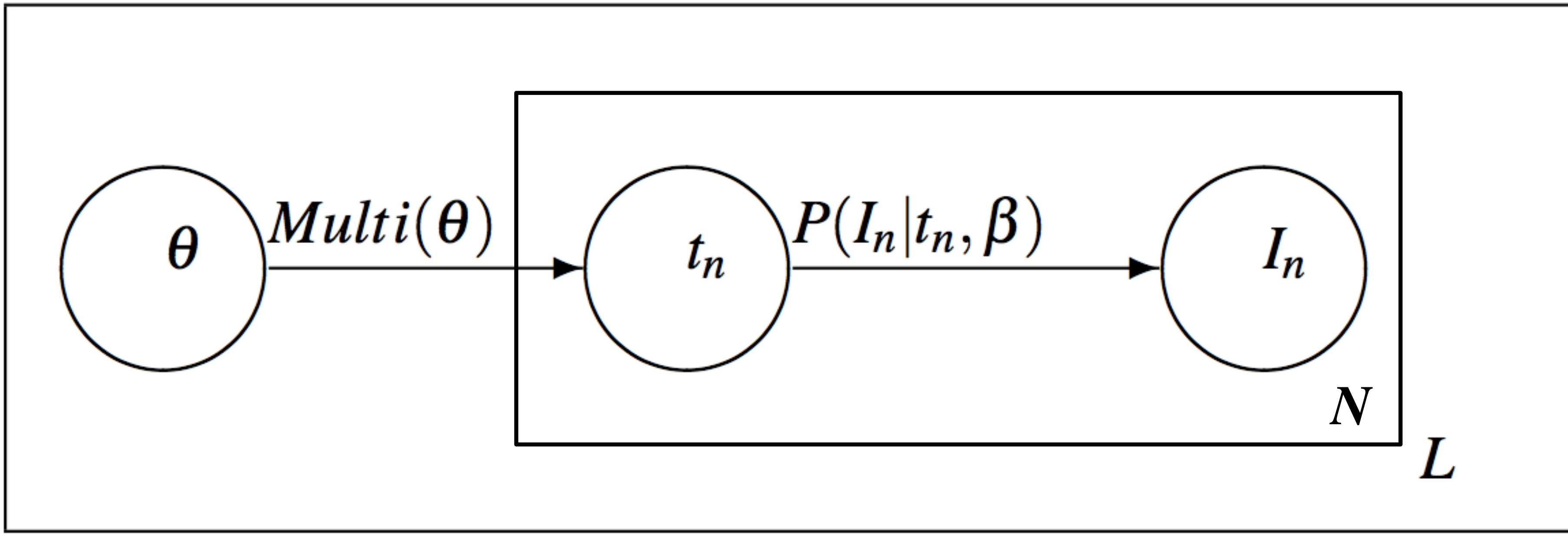



Geo-locations

***Climate topics***

Climate Descriptors

# Climate topic modeling using LDA




- *L*: number of spatial regions
- *N:* number of observations in region
- $t_n$: climate topic
- $I_n$: climate descriptor: discretized observed climate variable
- Dirichlet prior on θ

# How to use these results?

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| TOPIC_1 | 0.24775 | TOPIC_2 | 0.24983 | TOPIC_3 | 0.25046 | TOPIC_4 | 0.25196 |
| uwnd1 | 0.20495 | rhum5 | 0.19172 | vwnd5 | 0.17427 | shum1 | 0.18842 |
| vwnd1 | 0.16806 | pr_wtr5 | 0.18384 | pres1 | 0.16477 | pr_wtr1 | 0.16720 |
| slp3 | 0.10586 | shum5 | 0.15476 | slp1 | 0.12670 | slp5 | 0.15101 |
| pres3 | 0.09527 | slp1 | 0.12487 | rhum1 | 0.10940 | rhum1 | 0.13596 |
| rhum3 | 0.07771 | pres1 | 0.11816 | shum1 | 0.09419 | pres5 | 0.13455 |
| pr_wtr3 | 0.06630 | uwnd5 | 0.09307 | pr_wtr1 | 0.06436 | uwnd5 | 0.09769 |
| shum3 | 0.05672 | vwnd5 | 0.04259 | uwnd5 | 0.05586 | vwnd1 | 0.06195 |
| slp5 | 0.04081 | vwnd3 | 0.03911 | uwnd3 | 0.04704 | vwnd3 | 0.03125 |
| pres5 | 0.04037 | uwnd3 | 0.01767 | pr_wtr3 | 0.04614 | slp4 | 0.00924 |
| uwnd3 | 0.03619 | shum3 | 0.01572 | shum3 | 0.03379 | vwnd2 | 0.00600 |

- Defining climate extreme events automatically
- Modeling and detecting complex extreme climate events
- Feature selection for complex extreme events
- Use the results to find spatial covariability of extreme events

# Remaining Challenges for ML

- What are the effects of climate change on extreme events, especially regional?
- How will distributions of relevant variables change with climate change?
- Detecting / predicting climate extremes, anomaly detection
- Real-time learning from data streams, tracking extreme events

# Space and Time

# Climate Networks

- Network(s) of interactions between dynamical processes

- Correlation Measures
  - Pearson correlation
  - Mutual Information

$$\boxed{\rho_{ij} = \frac{Cov(X_i, X_j)}{\sigma(X_i)\sigma(X_j)}}$$

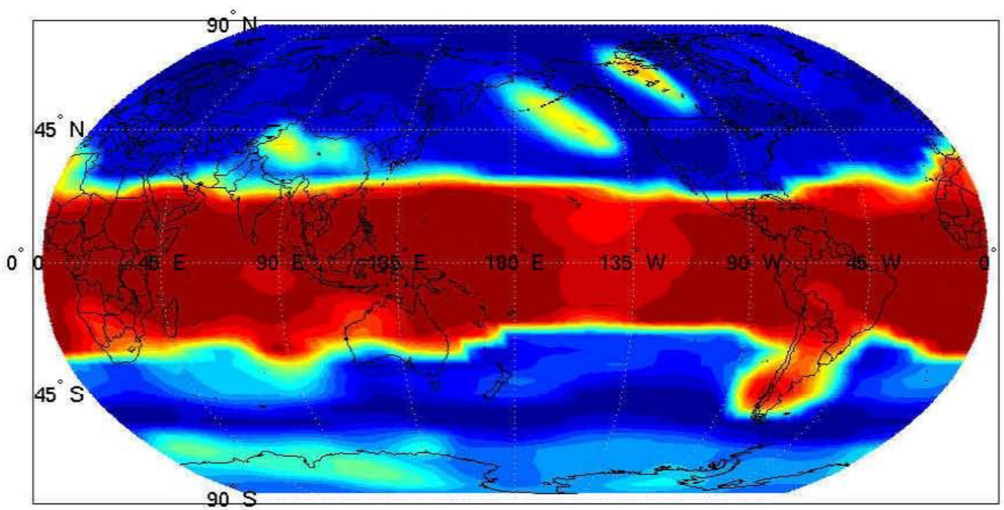
$$\boxed{MI_{ij} = \sum_{uv} p_{ij}(u,v) \log \frac{p_{ij}(u,v)}{p_i(u)\,p_j(v)}}$$

- Graph over locations, thresholded
  - Data: One variable, typically
  - Aggregated or segregated
  - E.g., El-Nino/La-Nina, Summer/Winter, …

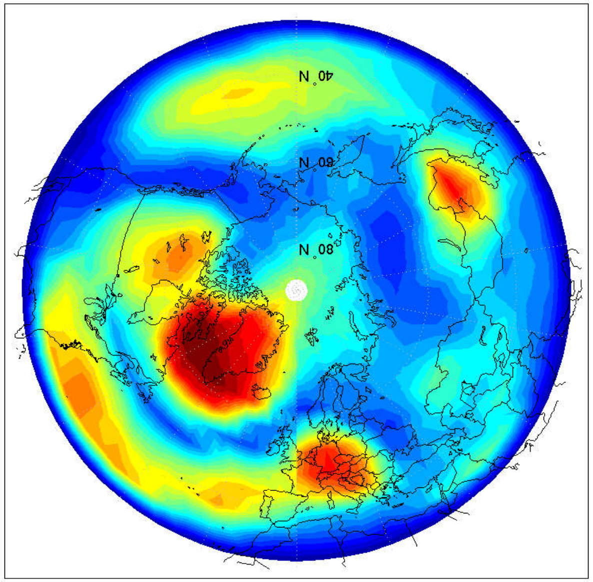
# Correlation Networks

- Goal: Construct graph $G = (V, E)$ over nodes $V$
  - Single climate variable, e.g., 500-hPa pressure level height
  - Correlation Thresholding

$$E_{ij} = \mathbf{1}\left\{\left|\rho_{ij}\right| \geq 0.5\right\}$$




Total degree at each location
[Tsonis et al., Physica '04; Tsonis et al., BAMS '06]



Extratropics Only (>30°N)
[Tsonis et al., NPG '12]

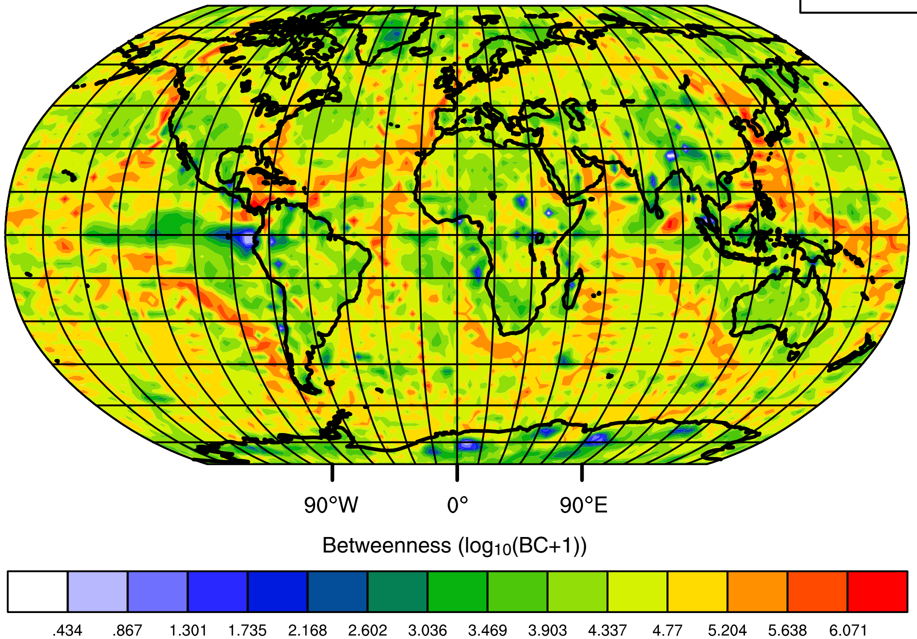
# "Backbone" of Climate Network

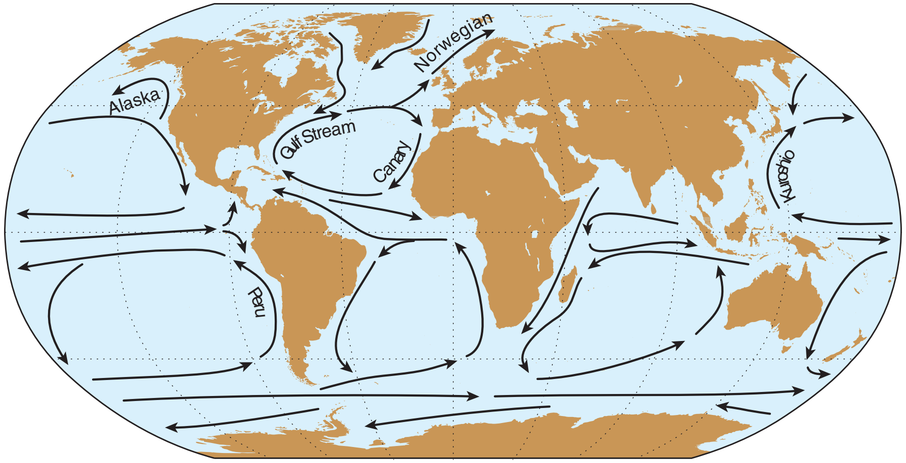
- Threshold Mutual Information $E_{ij} = \mathbf{1}\{|MI_{ij}| \geq \tau\}$
  [Donges et al., EPL '09]
  - Surface Air Temperature (SAT)
- Betweenness centrality

$$BC_v = \sum_{i,j \neq v} \frac{\eta_{ij}(v)}{\eta_{ij}}$$

# shortest paths between $i$ & $j$ through $v$

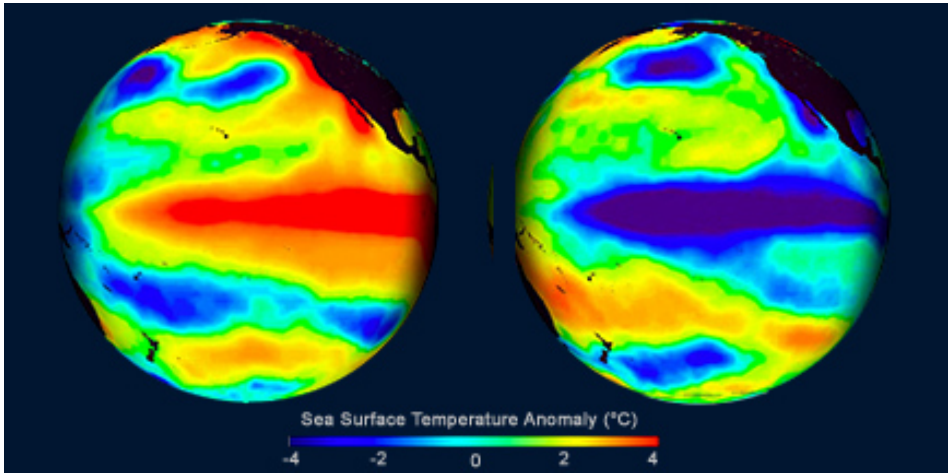
Total # shortest paths between $i$ & $j$




Ocean Surface Currents

# Oscillations, Teleconnections, Dipoles




Credit: S. Albers

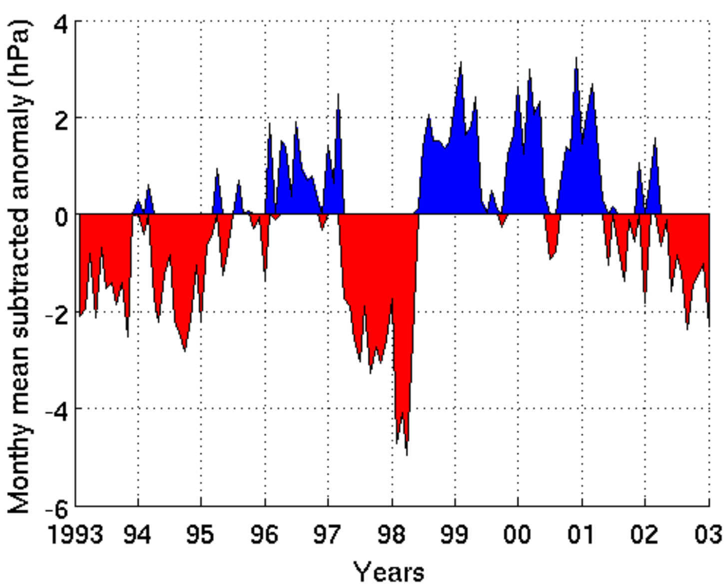
- Crucial for understanding the climate system
- Causes temperature and precipitation anomalies worldwide

Several oscillations: ENSO (El Nino Southern Oscillation), NAO (North Atlantic Oscillation), AO (Arctic Oscillation), AAO (Antarctic Oscillation), AMO (Atlantic Multidecadal Oscillation), PDO (Pacific Decadal Oscillation), MJO (Madden-Julian Oscillation), IOD (Indian Ocean Dipole), PNA (Pacific-North American Pattern), …

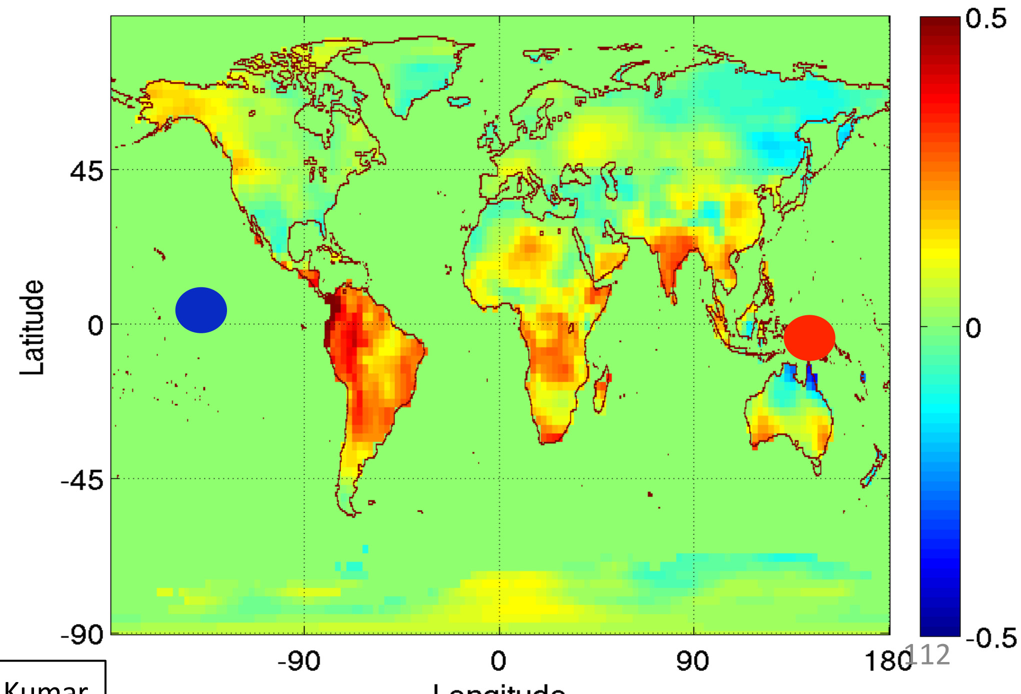
[van Loon et al., MWR'78; Wallace et al., MWR'81; von Starch et al.,'02; Walker, MIMD '23]

# Southern Oscillation Index: SOI

- Sea Level Pressure (SLP) difference: Tahiti and Darwin
  - Air pressure fluctuations between east and west tropical Pacific
- Smoothed time series of SOI and ENSO
  - Negative SOI: warm ocean waters, El Nino episodes
  - Positive SOI: cold ocean waters, La Nina episodes



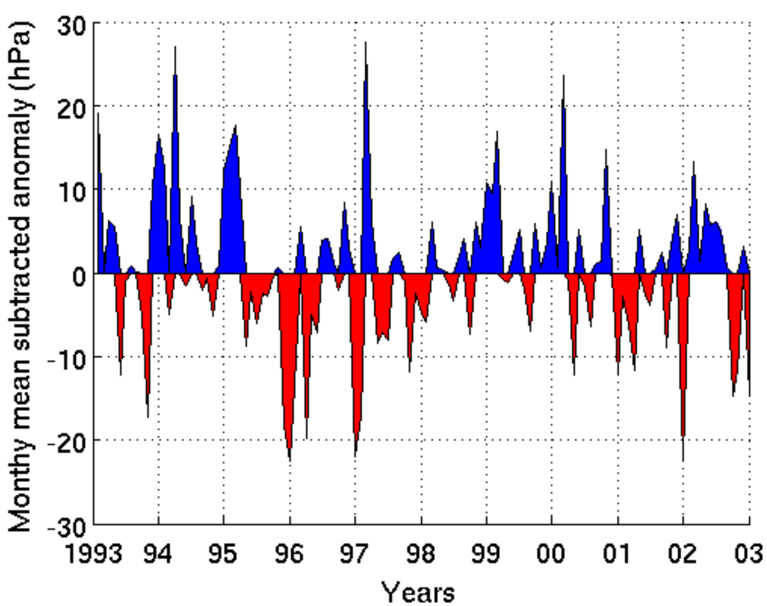

Correlation of land temperature anomalies with SOI



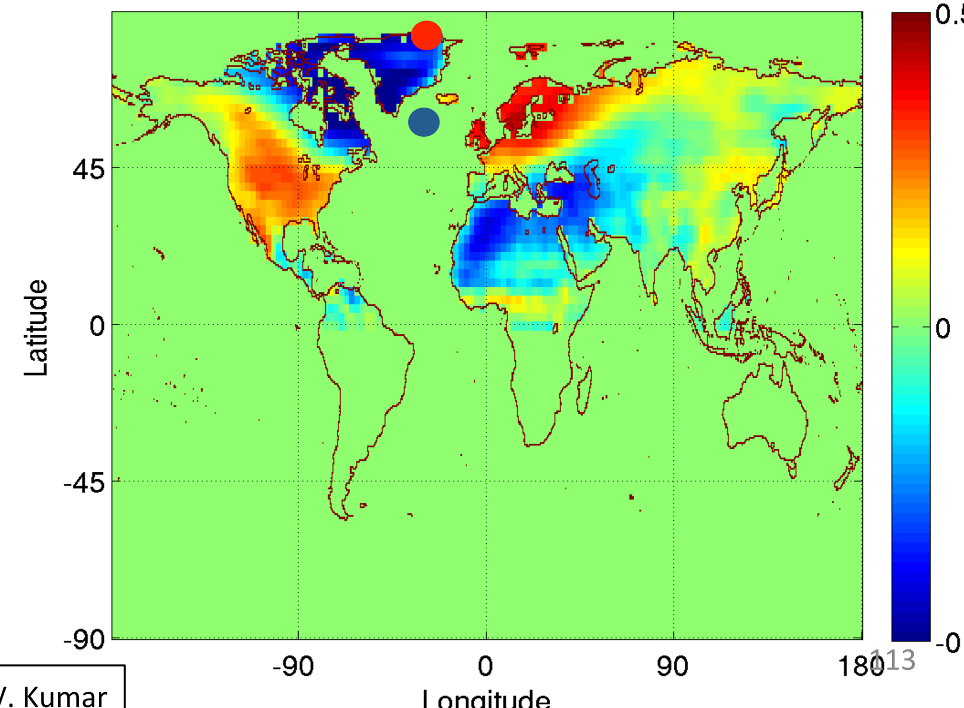

Credit: V. Kumar

# North Atlantic Oscillation: NAO

- Pressure difference in north Atlantic
- NAO affects the northern hemisphere
  - Positive NAO: High pressure over east US, west Europe
    - High temp over east US, north Europe, low temp in Greenland
    - Often, low temp in south Europe, Middle East

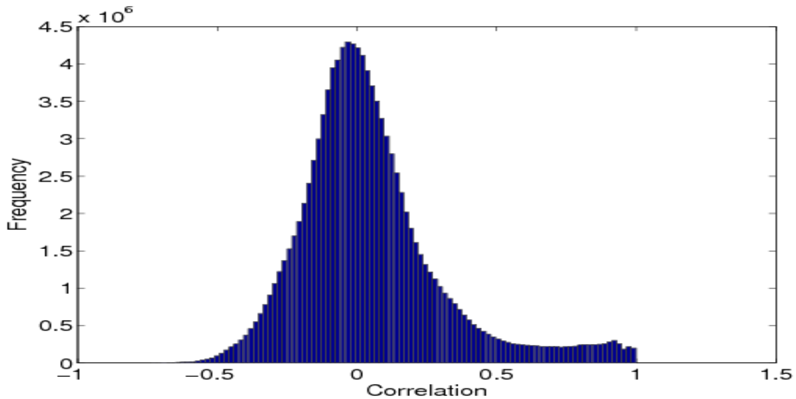



Correlation of land temperature anomalies with NAO

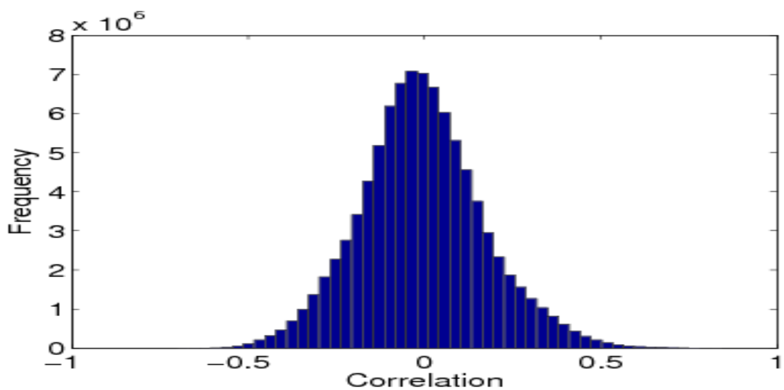



Credit: V. Kumar

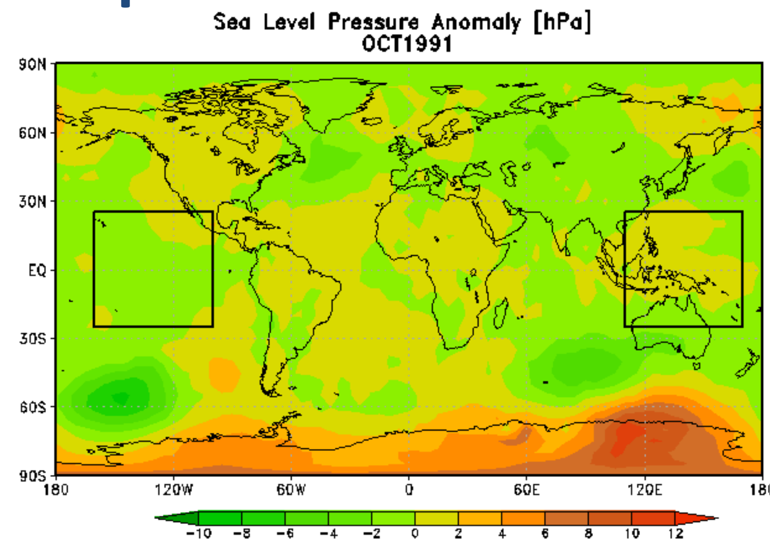
# Automated Discovery of Dipoles

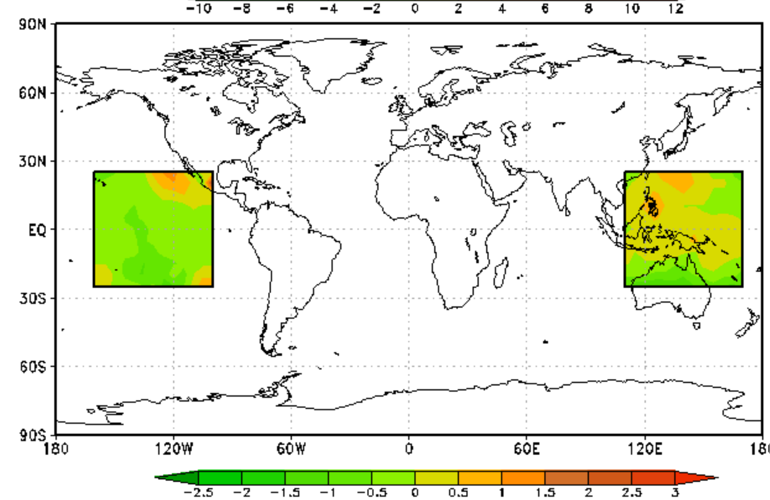



Pair-wise correlations (~10000 locations)




Pair-wise correlations (locations > 5000 km apart)



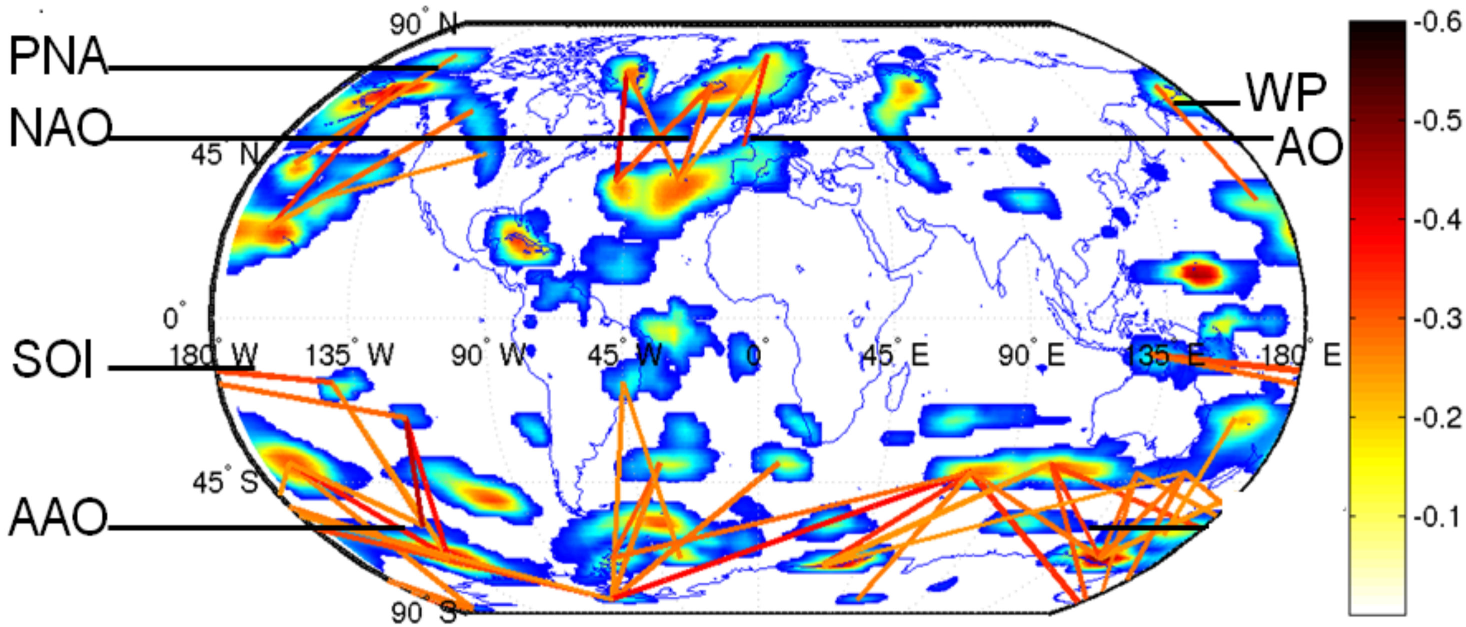



**Consistency in space and time is key to reduce the search space**

Credit: V. Kumar

[Kawale et al., SDM '11; Steinbach et al., KDD '03]

# Automatic Discovery of Dipoles



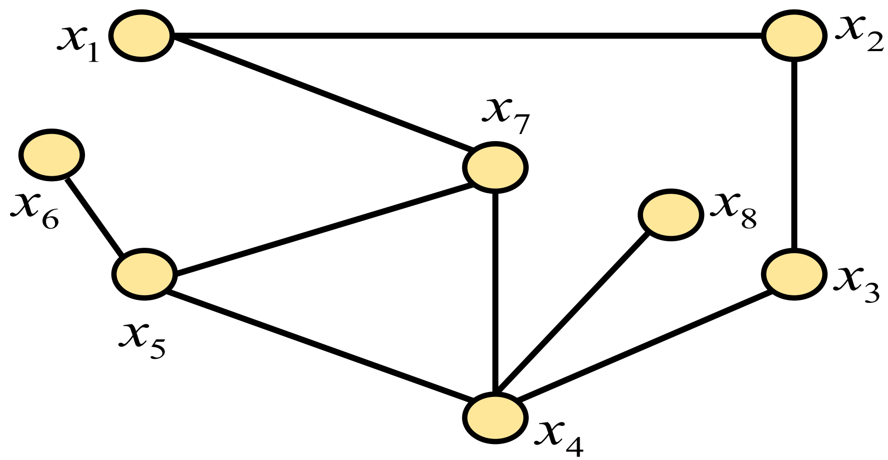

Credit: V. Kumar

- Detection of Global Dipole Structures
  - Most known dipoles discovered
  - Some `new' dipoles: Previously unknown phenomenon?
  - A new dipole near Australia [Liess et al., J Clim'14]

# "Sparse" Statistical Dependencies

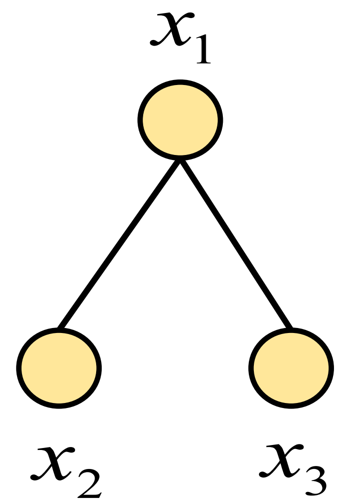
Graph of statistical dependencies: $a_{ij} = 0 \Leftrightarrow X_i \perp X_j \mid X_{-\mathrm{i},-\mathrm{j}}$

- Example: $X_1\ X_6 \ \mid\ X_2, X_3, X_4, X_5, X_7, X_8$
- Conditional independence: $X_6$ has no explicit influence on $X_1$
- Knowing $(X_2, X_7)$ is sufficient to (statistically) characterize $X_1$



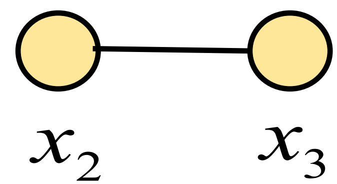
# Approaches: Correlation, Mutual Information

- Simple dependency structure




- Why correlation will not work
  - $X_2$, $X_3$ can be (strongly) correlated: $\mathrm{Corr}(X_2,X_3) \neq 0$
  - Correlation does not capture conditional independence

- Why Mutual Information (MI) will not work
  - Marginalizing $X_1$ makes $(X_2,X_3)$ dependent
  - In general, $p(X_2,X_3)$ will not factorize: $\mathrm{MI}(X_2,X_3) \neq 0$

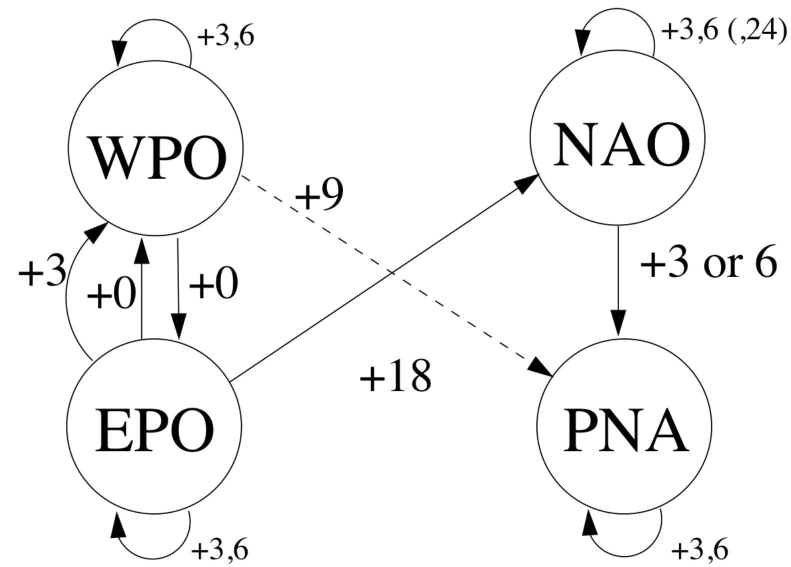



- Conditional MI (CMI) will work, but difficult
  - $X_2$, $X_3$ are conditionally independent: $\mathrm{MI}(X_2,X_3|X_1) = 0$
  - Infeasible in high-dimensions: Exponentially many CMI computations

# Relationship Between Oscillations

**Question:** [Ebert-Uphoff et al., J Clim '12]
Causal relationships between the four modes?

WPO = Western Pacific Oscillation
EPO = Eastern Pacific Oscillation
PNA = Pacific North America Pattern
NAO = North Atlantic Oscillation




Credit: I. Ebert-Uphoff

**Data:**

- NCEP-NCAR reanalysis data, 1950-2011
- Daily 500 mb geopotential height for all four modes (Oct-Mar)

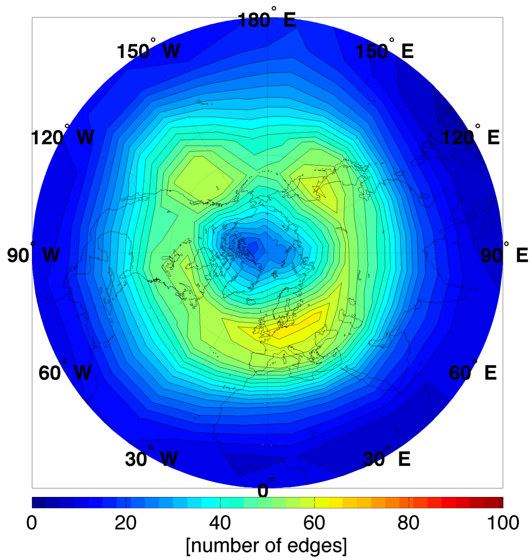
**Results:**

- Most links consistent with mechanisms in literature.
- Some time scales are new.
- One new link: NAO → PNA (with 3-6 days lag)

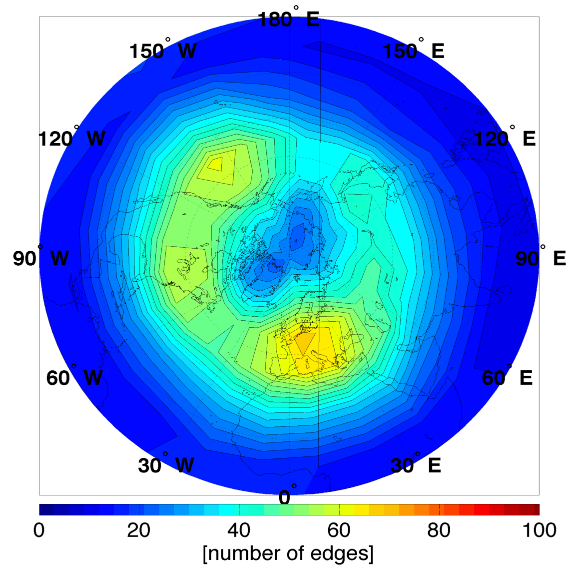
# Trends: Now vs. Future

Credit: I. Ebert-Uphoff

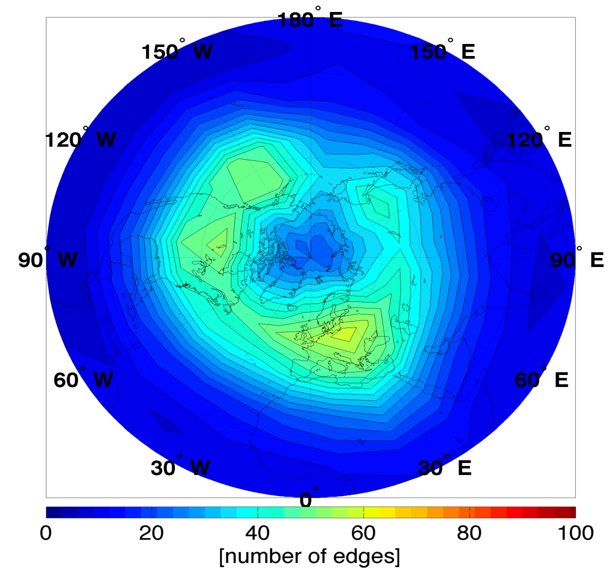
**a) 1950-2000 observed (NCEP-NCAR reanalysis)**




**b) CCSM4 model data Years: 1950-2000**




**c) CCSM4 model data Years: 2050-2100**

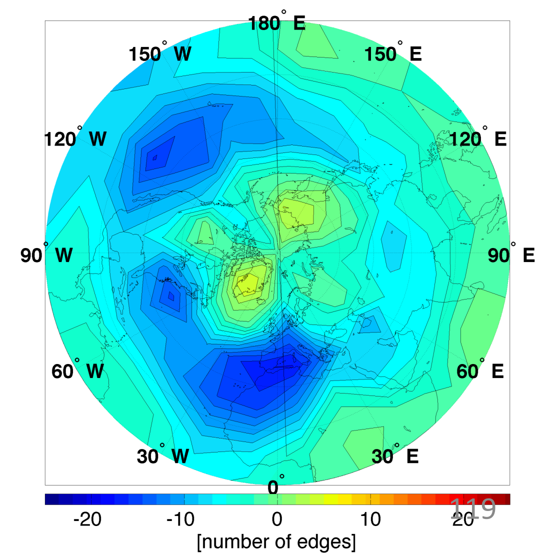



**Observations:** In warmer climate [Deng et al., GRL'14]

- Information flow diminishes (hubs disappear)
- Remaining hubs move poleward

**Literature:** Mid-latitude storm tracks

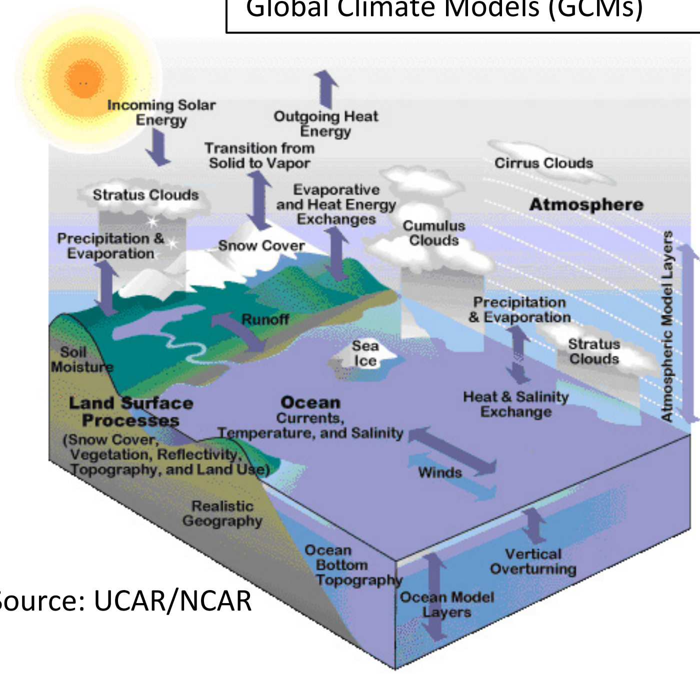
- Move poleward in warming climate




**Difference : c) – b)**

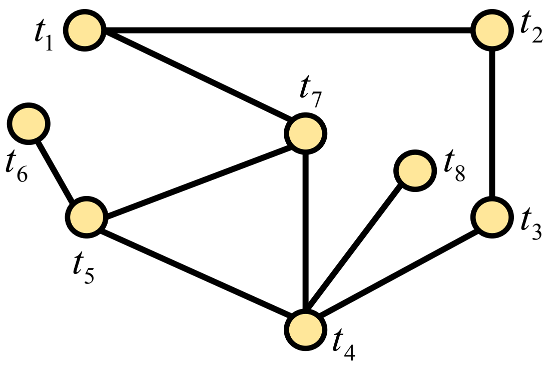
# Combining GCMs: Spatial Multi-task Learning

Global Climate Models (GCMs)




Source: UCAR/NCAR

Tasks: South America regional temperature




Dependency between different regions (tasks)

Combining GCM outputs as Multi-task learning

- Tasks: Climate model weights for a region
- Task based regularization
  - Model weights on related tasks should be similar
- Task is a location: Which locations are related?

# Multi-Task Sparse Structure Learning

- Combining GCMs in each location is a task t
  - Linear model for each location (task) t, $y^t = X^t w^t + e^t$
  - Joint error $[e^t]$ over all locations (tasks) is Gaussian

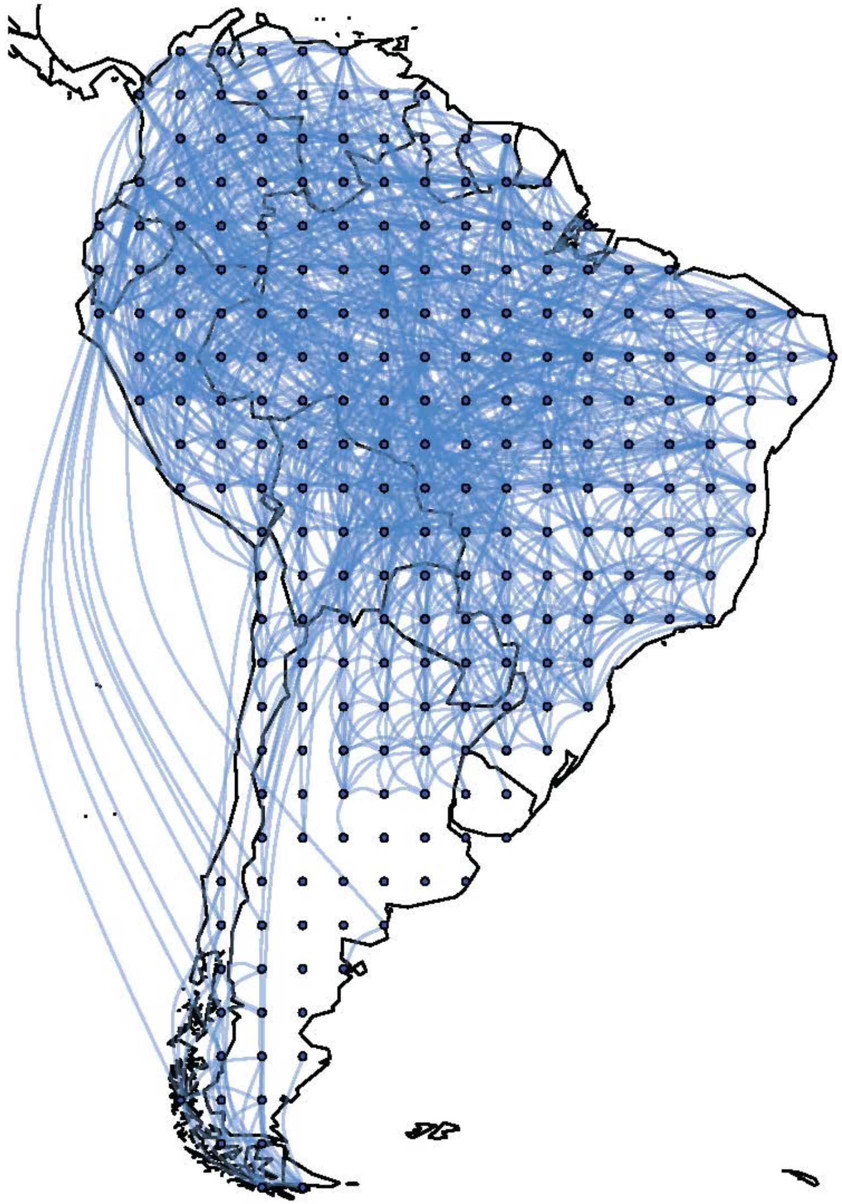
$$[y^t] = [X^t w^t] + [e^t] \ , \quad [e^t] \sim N(0, \Theta_0^{-1})$$

- Sparse precision matrix for spatial Gaussian
  - Non-zero entries of precision reveal task relationships

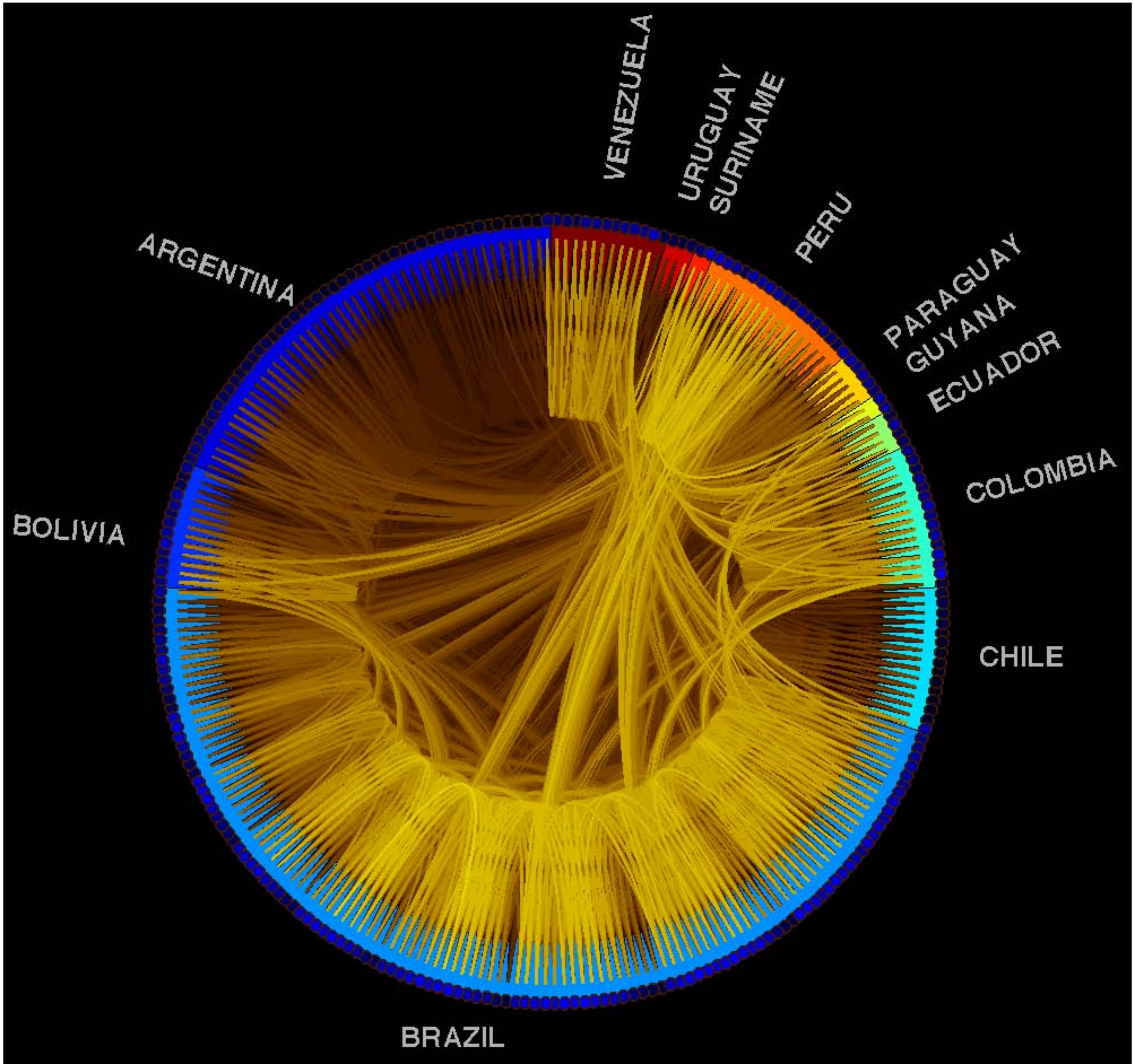
- Residual Multi-task Sparse Structure Learning (r-MSSL)
  - Muli-task Graphical Lasso: Estimate both $W = \{w^t\}$ and $\Theta$

  [Goncalves et al., CIKM '14]

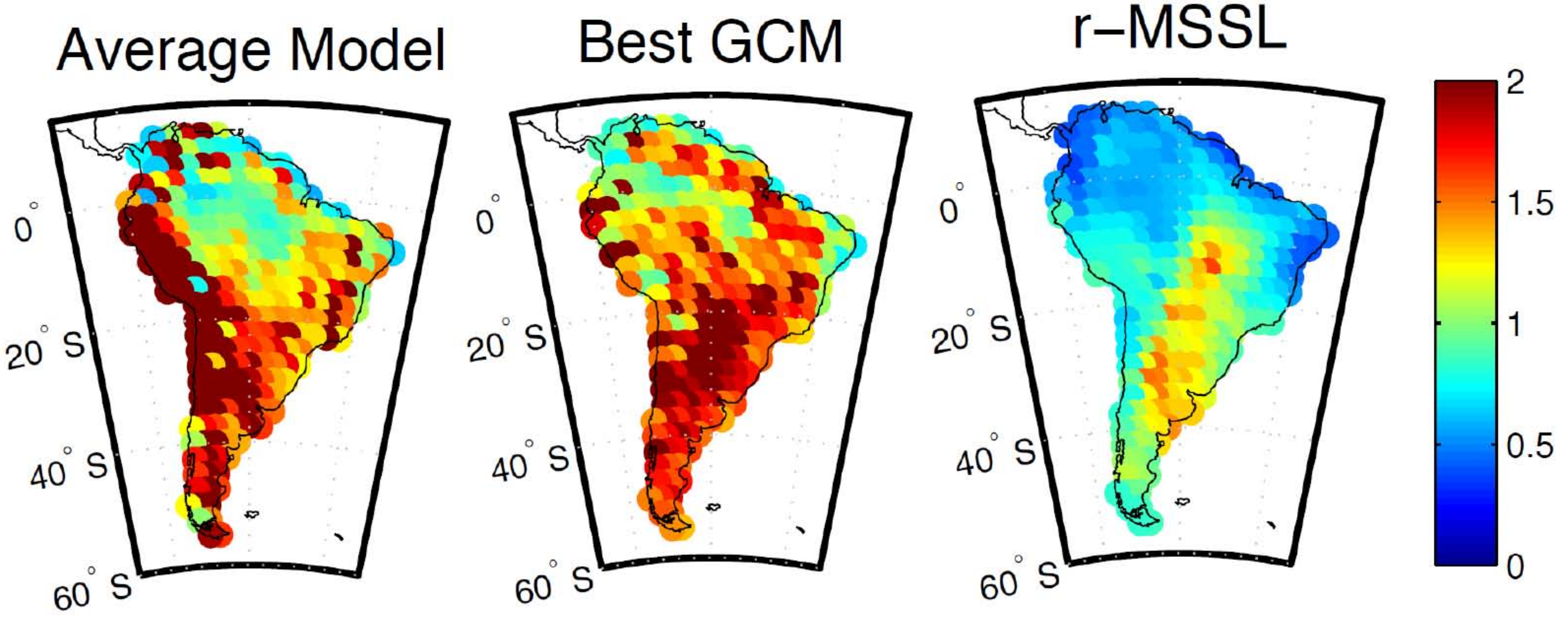
# Dependency: Temperature in South America



# Dependency: Temperature in South America



# RMSE Comparison: r-MSSL vs Baselines




| Average Model | Best GCM | r-MSSL |
|---|---|---|
| **1.621 (±0.020)** | **1.410 (±0.037)** | **0.780 (±0.039)** |

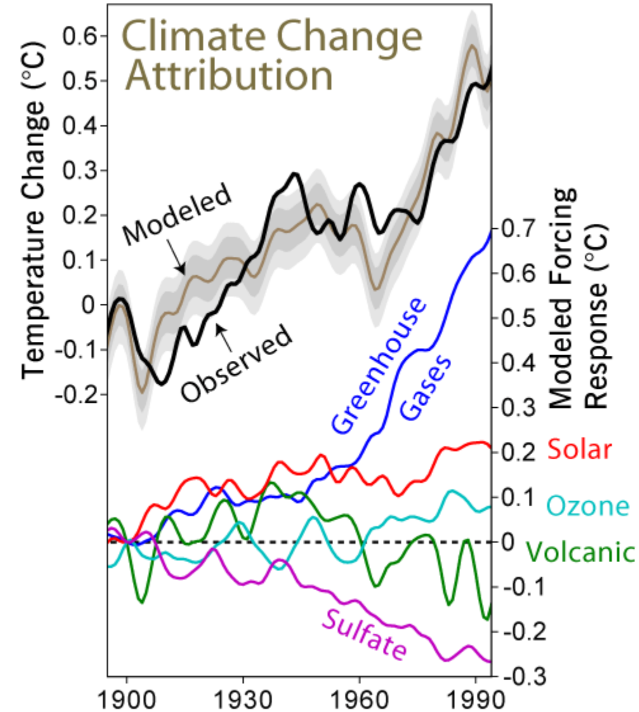
# Challenges for Spatiotemporal Analysis

- Predictability
  - Nonlinear dynamics, chaos
  - Estimate predictability before building models
- Influences over space, time
  - Non-stationary, non-linear dependencies, superposition effects
  - Many possible predictors, different scales
- Long-range dependencies, feedbacks, memory
  - Spatial teleconnections, e.g., pressure dipoles
  - Temporal dependencies: local lag, long memory
- Spatial diversity, temporal trends
  - Avoid fixed, blackbox approach
  - Interpretable, in terms of climate processes

# Additional challenges and conclusion

# Challenges

- Climate change attribution
  - Dependencies on external forcings
  - Beyond linear models, correlated noise




Credit: Wikipedia

- Causal discovery
  - Mechanistic understanding
  - Chain of events at climate process level
  - E.g., Warm ocean ⟺ forest fires

Credit: Wikimedia

# Challenges

- Extreme events
  - Heat waves, heavy precipitation
  - Droughts, floods
  - Hurricanes, tornadoes
- Behavior and Trends
  - Over space, time
  - Vulnerabilities, impacts
- Tracking changes
  - Polar ice
  - Forest cover, biodiversity
  - Water resources
  - …

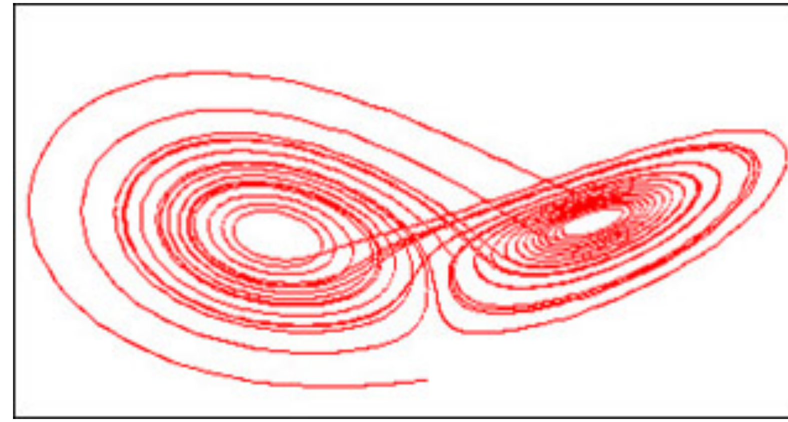
Credit: NBC News

Credit: S. Alur

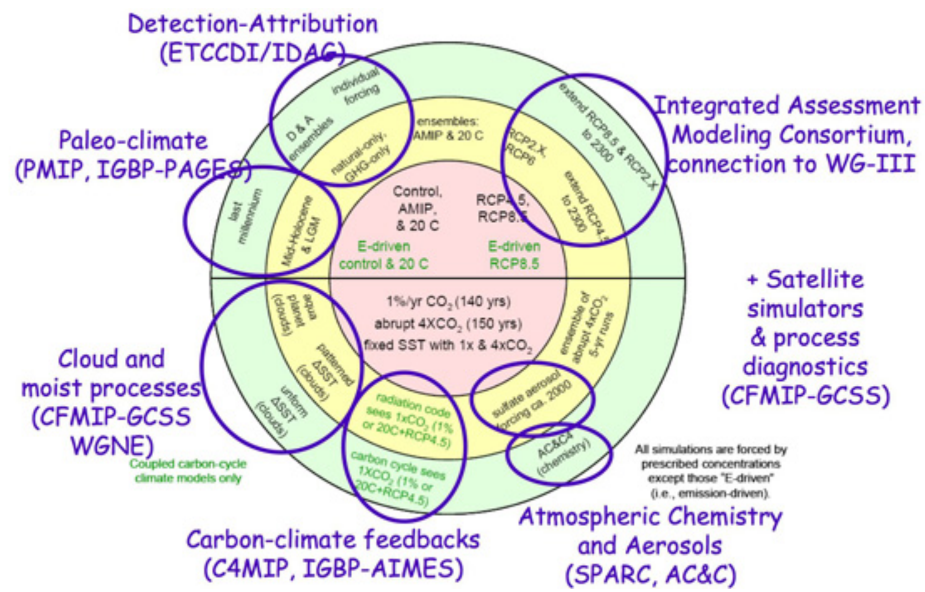
# Challenges

- Nonlinear dynamics, Predictability
  - Weather vs Climate
    - Few days vs statistics/trends
  - Predictability at decadal scales
    - Mutual information $I(X_t, X_{t+\tau})$, large $\tau$

- Climate model evaluation
  - Goal: Improve ESMs/GCMs
  - Skills on climate processes
  - Use multiple metrics



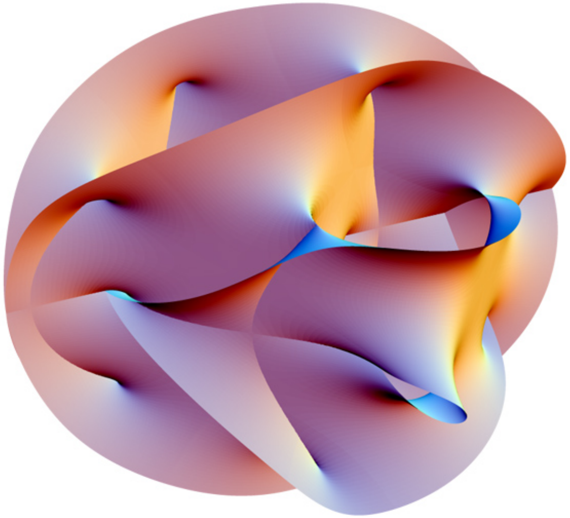
Credit: D. Rothman



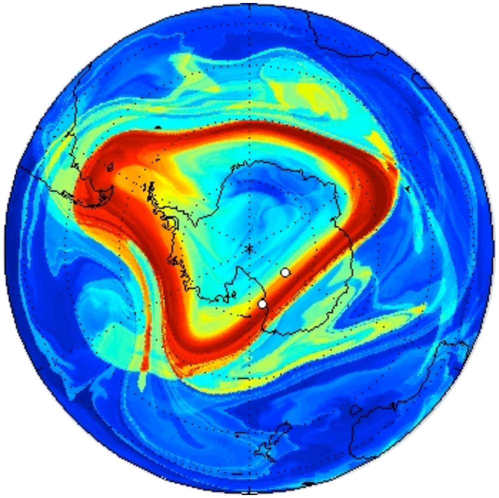

Credit: S. Easterbrook

# Challenges

- Machine Learning for Physical Processes
  - Variables obey physics
  - Physics guided data analysis
  - Representation, geometry/manifolds, …
  - Dynamic models, differential equations, …
  - …



Credit: M. Brooks



Credit: A. de La Camara, et al.

# Conclusion

- Climate Change: Challenges where ML can help
  - Paleoclimate: The past
  - Downscaling: Local climate
  - Climate model ensembles: The future, quantify uncertainty
  - Extremes: Tail behavior, impacts
  - Space and Time: Complex influences and dependencies
- Many more challenges: Only scratched the surface
  - Datasets (and basic tools) are all available
  - Work with a domain scientist
- Climate Informatics
  - Small but growing community
  - Like bioinformatics in the early days

*Many thanks to our coauthors (listed in bibliography). Special thanks to colleagues and students who helped with this tutorial:*

John Carlis, Soumyadeep Chatterjee, Hal Daumé, Tim DelSole,
Imme Ebert-Uphoff, Auroop Ganguly, Andre Goncalves, Evan Kodra,
Vipin Kumar, Yan Liu, Scott McQuade, Mahesh Mohan, Doug Nychka,
Vidyashankar Sivakumar, Cheng Tang, Nan Zhang

# Thank you!

# Resources

- Climate Informatics: www.climateinformatics.org
  - Links to resources, Climate Informatics workshops, online community
- Climate Informatics Wiki
  - Data sets here: sites.google.com/site/1stclimateinformatics/materials
- 4th International Workshop on Climate Informatics, 2014
  www2.image.ucar.edu/event/ci2014
- 4th Workshop on Understanding Climate Change from Data, 2014
  www2.image.ucar.edu/event/fourth-climatechange
- IPCC AR5 Report: www.ipcc.ch/report/ar5/
- WCRP Grand Challenges:
  www.wcrp-climate.org/grand-challenges

# References: Introduction


- IPCC (2012). Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change [Field, C.B., V. Barros, T.F. Stocker, D. Qin, D.J. Dokken, K.L. Ebi, M.D. Mastrandrea, K.J. Mach, G.-K. Plattner, S.K. Allen, M. Tignor, and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, UK, and New York, NY, USA, 582 pp.
- IPCC (2013). Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp.
- E. Lorenz (1996). Predictability – A problem partly solved. Seminar on Predictability, Vol. I, ECMWF.
- C. Monteleoni, G.A. Schmidt, F. Alexander, A. Niculescu-Mizil, K. Steinhaeuser, M. Tippett, A. Banerjee, M.B. Blumenthal, A.R. Ganguly, J.E. Smerdon, and M. Tedesco (2013). Climate Informatics. Computational Intelligent Data Analysis for Sustainable Development; Data Mining and Knowledge Discovery Series. Yu, T., Chawla, N., and Simoff, S. (Eds.), CRC Press, Taylor & Francis Group. Chapter 4, pp. 81–126, 2013.
- IPCC Fifth Assessment Report: www.ipcc.ch/report/ar5/
- World Climate Research Program Grand Challenges:

  www.wcrp-climate.org/grand-challenges

# References: Paleo-climate reconstruction


- M. E. Mann, R. S. Bradley, M. K. Hughes, (1999), Northern hemisphere temperatures during the past millennium: Inferences, uncertainties, and limitations. Geophysical Research Letters 26 (6): 759-762.
- J.E. Smerdon, and A. Kaplan, (2007). Comment on Testing the fidelity of methods used in proxy-based reconstructions of past climate: The role of the standardization interval, Journal of Climate, 20(22), 5666-5670.
- B. Li, D. W. Nychka, C. M. Ammann. The value of multiproxy reconstruction of past climate. Journal of the American Statistical Association. 105.491 (2010): 883-895.
- M. P. Tingley, P. F. Craigmile, M. Haran, B. Li, E. Mannshardt, B. Rajaratnam, (2012) Piecing together the past: statistical insights into paleoclimatic reconstructions. Quaternary Science Reviews, Volume 35, 5 March, Pages 1-22.
- D. Nychka, E. Wahl, W. Gross, and D. Anderson. Reconstructing CO2 for the last 2000 years. (2014) Joint Statistical Meetings.

# References: Climate Model Ensembles


- G. A. Meehl, G. J. Boer, C. Covey, M. Latif, and R. J. Stouffer, (2000). The Coupled Model Intercomparison Project (CMIP). Bull. Amer. Meteor. Soc., 81, 313-318.
- T. Reichler, and J. Kim, (2008). How well do coupled models simulate today's climate? Bull. Amer. Meteor. Soc. 89:303– 311.
- C. Reifen, and R. Toumi, (2009). Climate projections: Past performance no guarantee of future skill? Geophys. Res. Lett. 36.
- R. Smith, C. Tebaldi, D. Nychka, and L. O. Mearns . Bayesian Modeling of Uncertainty in Ensembles of Climate Models. Journal of the American Statistical Association. 01(2009); 104(485):97-116.
- R. Knutti, R. Furrer, C. Tebaldi, J. Cermak, and G. A. Meehl, (2010). Challenges in Combining Projections from Multiple Climate Models. J. Climate, 23, 2739–2758.
- C. Monteleoni, G.A. Schmidt, S. Saroha, and E. Asplund, (2011). Tracking climate models. Statistical Analysis and Data Mining: Special Issue on Best of CIDU 2010. 4(4):72–392.
- S. McQuade and C. Monteleoni (2012). Global Climate Model Tracking using Geospatial Neighborhoods. Proceedings of the Twenty-Sixth Conference on Artificial Intelligence (AAAI), Computational Sustainability and AI Special Track, (2012):335–341.

# References: Climate Model Ensembles


- K. Subbian and A. Banerjee (2013). Climate multi-model regression using spatial smoothing. In SDM, 2013.
- M. Ghafarianzadeh and C. Monteleoni, (2013). Climate Prediction via Matrix Completion. Proceedings of the Twenty-Seventh Conference on Artificial Intelligence (AAAI), Late-Breaking Papers Track, 2013.
- R. Keshavan, A. Montanari, and S. Oh (2009). Matrix Completion from Noisy Entries. Advances in Neural Information Processing Systems 22. pages 952-960.
- C. Monteleoni, and T. Jaakkola, (2003). Online learning of non-stationary sequences. Advances in Neural Information Processing Systems 16. pages 1093–1100.
- Jia, L, T. DelSole, and M. K. Tippett (2014). Can Optimal Projection Improve Dynamical Model Forecasts? J. Climate, 27, 2643-2655.
- DelSole, T., and M. K. Tippett (2007). Predictability: Recent Insights from Information Theory. Rev. Geophys., 45, RG4002.

# References: Extremes


- W.C. Palmer, Meteorological Droughts. (1965). Paper number 45: 1-58, Weather Bureau, US Department of Commerce.
- A. Arnold, Y. Liu, N. Abe. Temporal Causal Modeling with Graphical Granger Methods. Proceedings of the Thirteenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD-07), 2007.
- A. Lozano, H. Li, A. Niculescu-Mizil, Y. Liu, C. Perlich, J. Hosking, N. Abe. Spatial-temporal causal modeling for climate change attribution. International Conference on Knowledge Discovery and Data Mining (KDD' 09), 2009.
- X. Chen, Y. Liu, H. Liu, J. G. Carbonell. Learning spatial-temporal varying graphs with applications to climate data analysis. AAAI, 2010.
- Y. Liu, A. Niculescu-Mizil, A. Lozano, and Y. Lu. Learning temporal graphs for relational time-series analysis. ICML, 2010.
- Q. Fu, A. Banerjee, S. Liess, and P. K. Snyder. Drought Detection of the Last Century: An MRF-based Approach. SDM, 2012.

# References: Extremes


- Y. Liu, M. T. Bahadori, and H. Li. Sparse-GEV: Sparse latent space model for multivariate extreme value time series modeling. ICML, 2012.
- Q. Fu, H. Wang, and A. Banerjee. Bethe-ADMM for Tree Decomposition based Parallel MAP inference. Conference on Uncertainty in Artificial Intelligence (UAI), 2013
- C. Tang and C. Monteleoni, Detecting Extreme Events from Climate Time-Series via Topic Modeling. Machine Learning and Data Mining Approaches to Climate Science: Proceedings of the The Fourth International Workshop on Climate Informatics, 2014 (to appear).
- H. Wang and A. Banerjee. Bregman Alternating Direction Method of Multipliers. Advances in Neural Information Processing Systems (NIPS), 2014.

# References: Downscaling


- H. Li, J. Sheffield, and E. F. Wood (2010). Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching. *Journal of Geophysical Research: Atmospheres*, *115*(D10).
- J. Ramirez and A. Jarvis (2010). Downscaling global circulation model outputs: the delta method. *Decision and policy analysis working paper*, (1), 1-18.
- T. P. O'Brien, D. Sornette and R. L. McPherron (2001). Statistical asynchronous regression: Determining the relationship between two quantities that are not measured simultaneously. *Journal of Geophysical Research: Space Physics*, *106*(A7), 13247-13259.
- P. Friederichs and A. Hense (2007). Statistical downscaling of extreme precipitation events using censored quantile regression. *Monthly weather review*, *135*(6), 2365-2378.
- E. Zorita and H. Von Storch (1999). The analog method as a simple statistical downscaling technique: comparison with more complicated methods. *Journal of climate*, *12*(8), 2474-2489.

# References: Downscaling


- M. A. Semenov and E. M. Barrow (1997). Use of a stochastic weather generator in the development of climate change scenarios. *Climatic change*, *35*(4), 397-414.
- D. I. Jeong, A. St-Hilaire, T. B. M. J. Ouarda, and P. Gachon (2012). Multisite statistical downscaling model for daily precipitation combined by multivariate multiple linear regression and stochastic weather generator. *Climatic change*, *114*(3-4), 567-591.
- J. H. Christensen, T. R. Carter, M. Rummukainen and G. Amanatidis (2007). Evaluating the performance and utility of regional climate models: the PRUDENCE project. *Climatic Change*, *81*(1), 1-6.
- R. E. Benestad, D. Nychka, L. O. Mearns, (2012). Spatially and temporally consistent prediction of heavy precipitation from mean values. Nature Climate Change 2, 544547
- D. Nychka, S. Bandyopadhyay, D. Hammerling, F. Lindgren, and S. Sain (2014). A multi-resolution Gaussian process model for the analysis of large spatial data sets. *Journal of Computational and Graphical Statistics*, (just-accepted)

# References: Climate Networks


- A. A. Tsonis, and P. J. Roebber. "The architecture of the climate network." *Physica A: Statistical Mechanics and its Applications* 333 (2004): 497-504.
- A. A. Tsonis, K. L. Swanson, and P. J. Roebber. "What do networks have to do with climate?." *Bulletin of the American Meteorological Society* 87.5 (2006): 585-595.
- A. A. Tsonis, K. L. Swanson, and G. Wang. "On the role of atmospheric teleconnections in climate." *Journal of Climate* 21.12 (2008): 2990-3001.
- A. A. Tsonis et al. "Community structure and dynamics in climate networks." *Climate dynamics* 37.5-6 (2011): 933-940.
- A. A. Tsonis, and K. L. Swanson. "Review article" On the origins of decadal climate variability: a network perspective"." *Nonlinear Processes in Geophysics* 19.5 (2012): 559-568.
- J. F. Donges, Y. Zou, N. Marwan, & J. Kurths. The backbone of the climate network. *EPL (Europhysics Letters)*, *87*(4), 48007, 2009

# References: Dipole discovery


- H. van Loon and J. C. Rogers. The seesaw in winter temperatures between greenland and northern europe. Part I: General description. Monthly Weather Review, 106(3):296{310, 1978}
- G. Walker. Correlation in seasonal variations of weather, VIII. a preliminary study of world weather. Memoirs of the India Meteorological Department, 24(4):75{131, 1923}
- H. Von Storch and F. Zwiers. Statistical analysis in climate research. Cambridge Univ Pr, 2002
- D. H. Portis, J. E. Walsh, M. El Hamly and P. J . Lamb, Seasonality of the North Atlantic Oscillation, Journal of Climate, vol. 14, pg. 2069- 2078, 2001
- J. Wallace, and D. Gutzler: Teleconnections in the Geopotential Height Field during the Northern Hemisphere Winter. Monthly Weather Review, 109, 784-811, 1981
- J. Kawale, M. Steinbach, V. Kumar. Discovering dynamic dipoles in climate data. In SDM, 2011.

# References: Dipole Discovery


- M. Steinbach, P. Tan, V. Kumar, S. Klooster, C. Potter. Discovery of climate indices using clustering. In KDD, pages 446-455, 2003
- S. Liess, A. Kumar, P. K. Snyder, J. Kawale, K. Steinhaeuser, F. H. M. Semazzi, A. R. Ganguly, N. F. Samatova, and V Kumar. Different Modes of Variability over the Tasman Sea: Implications for Regional Climate. J.Climate, 27, 8466–8486, 2014.

# References: Spatiotemporal analysis


- Q. Fu, A. Banerjee, S. Liess, and P. K. Snyder. Drought Detection of the Last Century: An MRF-based Approach. In SDM, 2012.
- J. Friedman, T. Hastie, and R. Tibshirani. Sparse inverse covariance estimation with the graphical lasso. Biostatistics, 9(3), 432-441, 2007
- T. T. Cai, W. Liu, X. Luo. A constrained $l_1$ minimization approach to sparse precision matrix estimation. Journal of the American Statistical Association, 106, 594-607, 2011.
- S. Chatterjee, K. Steinhaeuser, A. Banerjee, S. Chatterjee, and A. R. Ganguly. Sparse group lasso: Consistency and climate applications. In SDM, pages 47-58, 2012.
- X. Chen, Y. Liu, and J. G. Carbonell. Learning spatial-temporal varying graphs with applications to climate data analysis. In AAAI, 2010.
- Y. Liu, A. Niculescu-Mizil, A. Lozano, and Y. Lu. Learning temporal graphs for relational time-series analysis. In ICML, 2010

# References: Spatiotemporal analysis


- Y. Liu, M. T. Bahadori, and H. Li. Sparse-GEV: Sparse latent space model for multivariate extreme value time series modeling. In ICML, 2012.
- A. R. Goncalves, P. Das, S. Chatterjee, V. Sivakumar, F. J. Von Zuben, and A. Banerjee. Multi-task sparse structure learning. In CIKM, 2014
- I. Ebert-Uphoff and Y. Deng. Causal discovery for climate research using graphical models, 25, 5648-5665, 2012.
- IPCC fifth assessment report. Intergovernmental Panel on Climate Change, 2013.
- T. N. Krishnamurti, C. M. Kishtawal, Z. Zhang, T. Larow, D. Bachiochi, E. Williford, S. Gadgil, and S. Surendaran. Multimodel ensemble forecasts for weather and seasonal climate. Journal of Climate, 12, 4196-4216, 2000.
- K. Subbian and A. Banerjee. Climate multi-model regression using spatial smoothing. In SDM, 2013.