Algorithms for Climate Informatics
Learning from spatiotemporal data with both spatial and temporal non-stationarity

Claire Monteleoni
CNRS / Université Paris-Saclay /
George Washington University → **U. Colorado Boulder**

August 2005: Hurricane Katrina – Reuters

October 2012: Hurricane Sandy – Reuters


August 2013: Rim Fire, California – Reuters

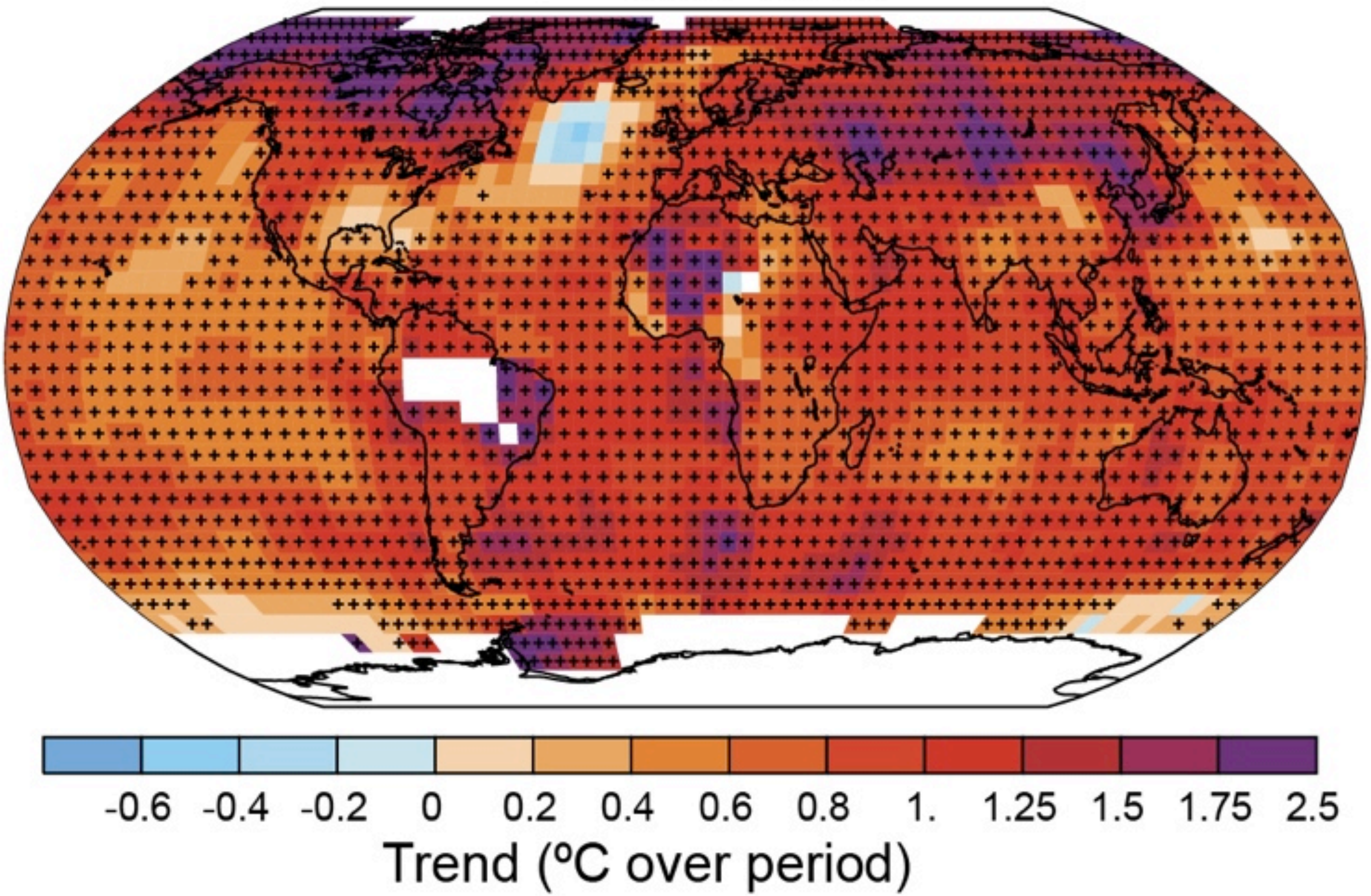
December 2017: Ventura County, California – Associated Press

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

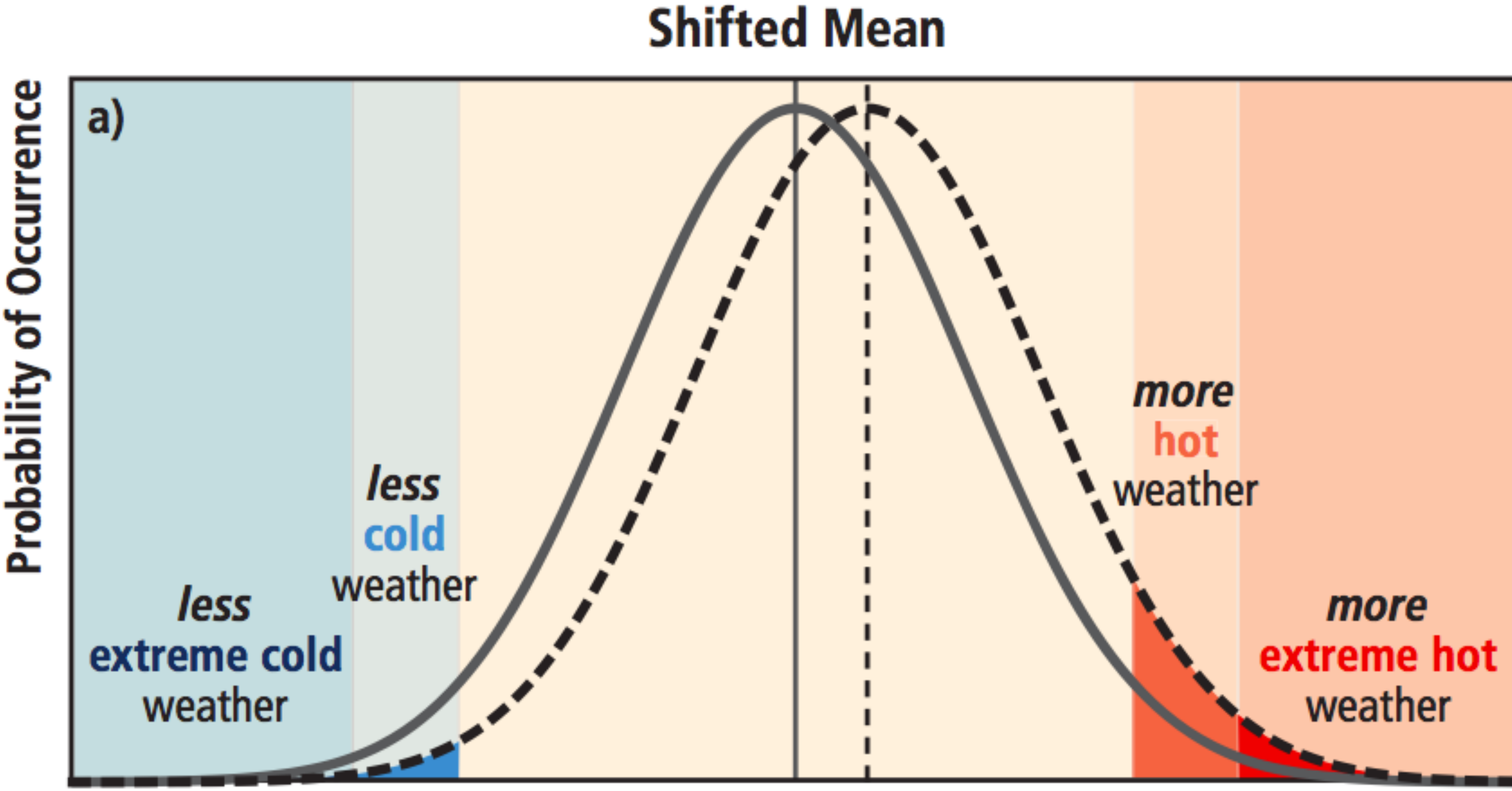
Machine learning can shed light on climate change.

How does climate change affect extreme events?

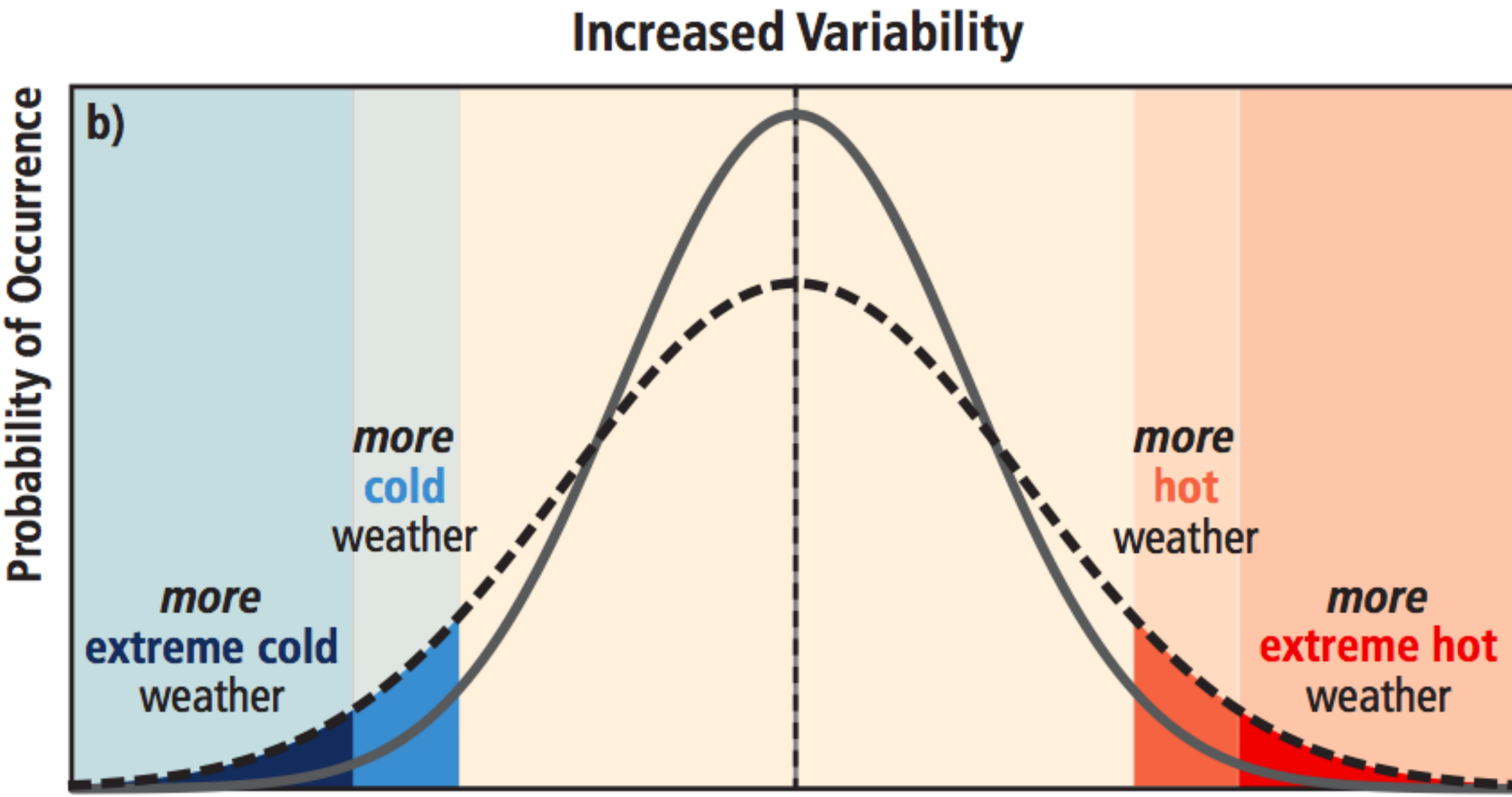
Surface Temperature 1901-2012



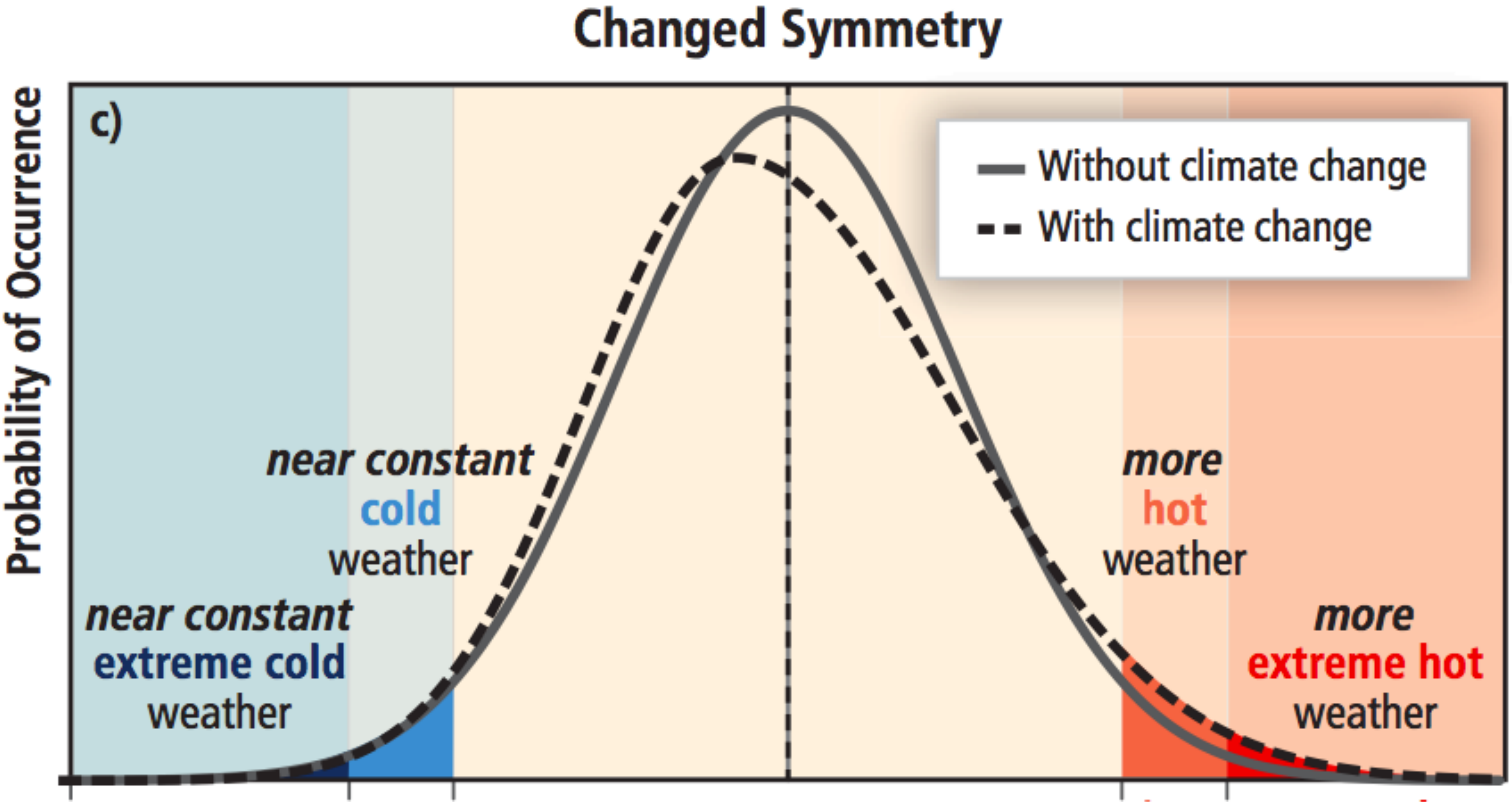
Intergovernmental Panel on Climate Change (IPCC), 2013



Intergovernmental Panel on Climate Change, 2012



Intergovernmental Panel on Climate Change, 2012



Intergovernmental Panel on Climate Change, 2012

Extreme events are rare by definition.

Climate change may affect their distribution.

➔ Past statistics are not sufficient for future prediction.

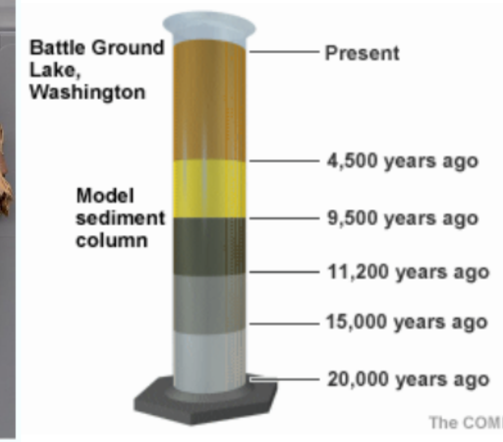
Augment historical data with climate model simulations.

Massive, high-dimensional, big data.

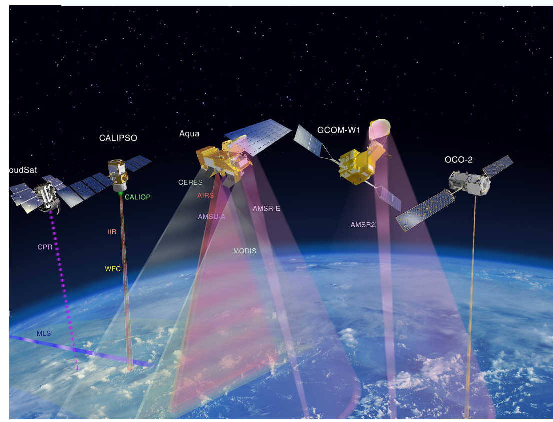
That's where machine learning comes in!

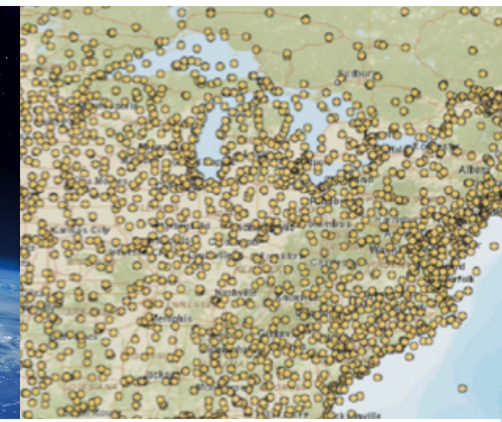
Climate data types

- Past: Historical data
  - Limited amounts
  - Very heterogeneous

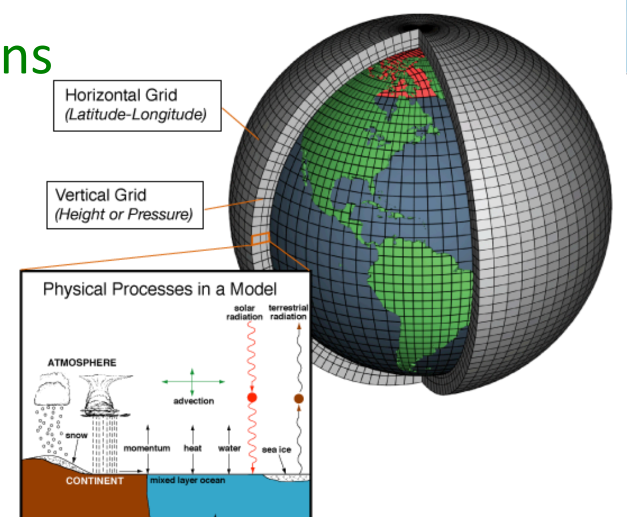


- Present: Observation data
  - Large quantities recently
  - High-dimensional
  - Can be unlabeled, sparse





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



Climate Informatics

2011 First International Workshop on Climate Informatics

New York Academy of Sciences

Climate Informatics Wiki launched

2013 "Climate Informatics" book chapter [Monteleoni et al. 2013]

→ In the first 5 years: participants from over 19 countries and 30 states

2018 Climate Informatics turns 8! With a Hackathon since 2015!

National Center for Atmospheric Research, Boulder, CO, USA

Challenge problems in climate informatics

[Banerjee & M, NIPS Tutorial, 2014]

1. Past: Paleo-climate reconstruction

   What was the climate before we had thermometers?

2. Local: Climate downscaling

   What climate can I expect in my own backyard?

3. Future: Climate model ensembles

   How to reduce uncertainty on future predictions?

4. Spatiotemporal: Space and time

   How to capture dependencies over space and time?

5. Tails/impacts: Extreme events

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

6. Other problems

   Data-rich playground with many opportunities for ML to have an impact!

On the menu

Climate Informatics: a compelling application area for ML

For further info see our NIPS 2014 tutorial

Algorithms for learning when the concept can vary over multiple dimensions

E.g. time, space

Examples of applications posing new questions for ML

On the menu

Applications can pose interesting new questions for ML

- Online + spatial
- Prediction at multiple timescales
- Tracking highly-deformable patterns

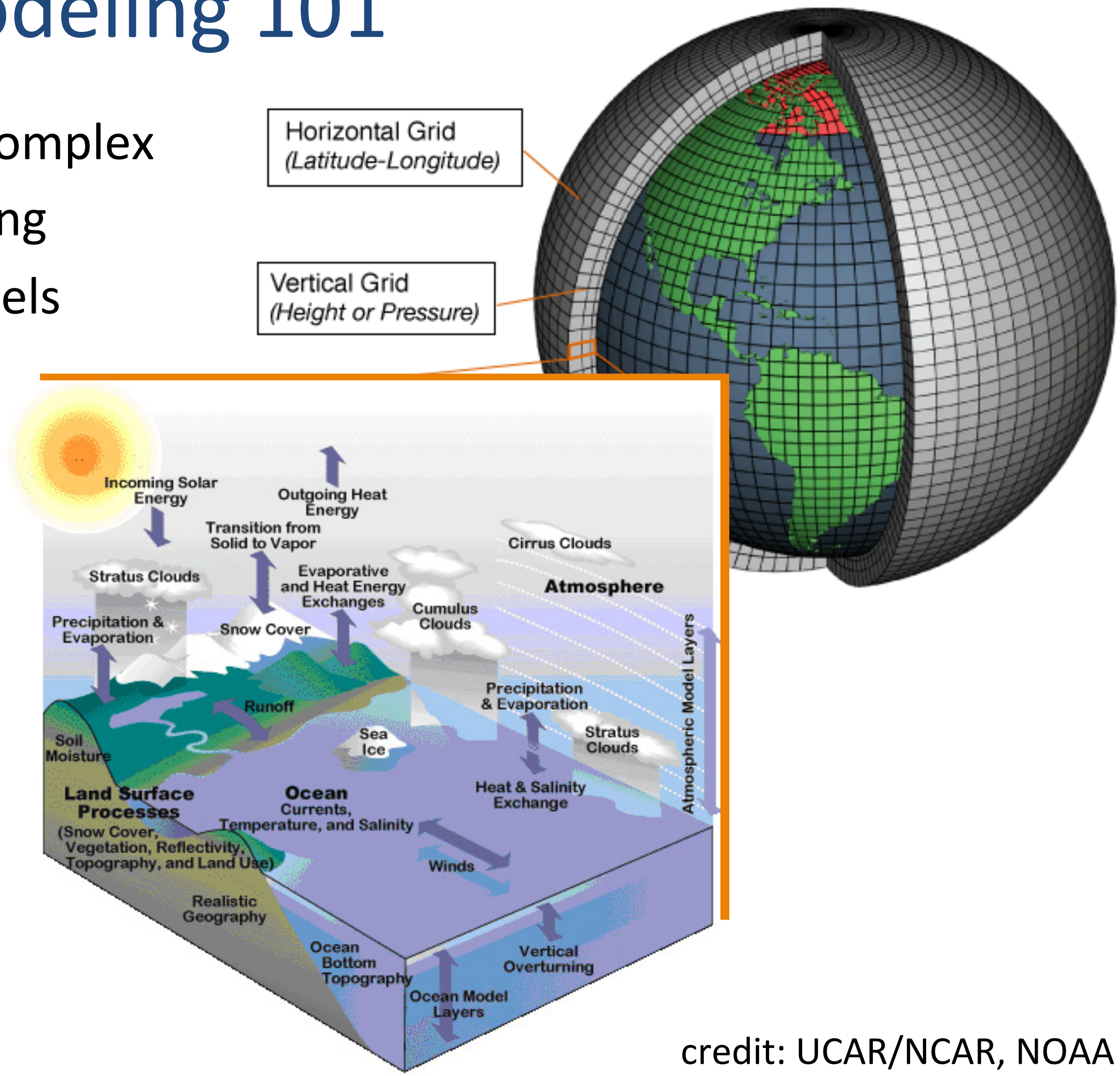
Learning from spatiotemporal data

- Learning from non-stationary time series
  - Simultaneously learn the level of non-stationarity
  - Exploit local temporal structure via multi-task learning
- Learning from non-stationary spatiotemporal data
  - Exploit local spatial structure
    - Distributed online learning
    - Hidden Markov random field

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
- Scale resolution differences



credit: UCAR/NCAR, NOAA

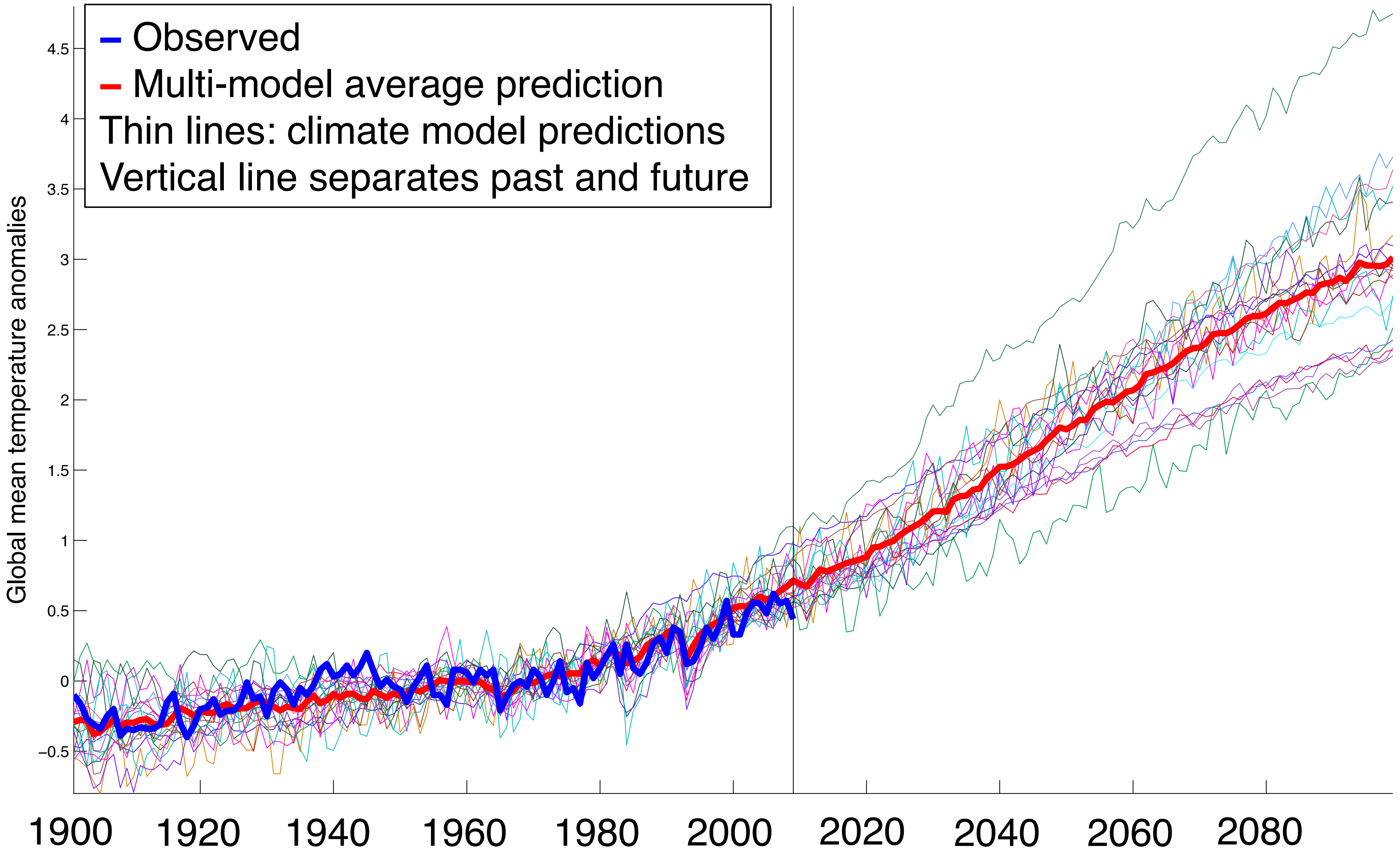
Intergovernmental Panel on Climate Change

- 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 (Canada), 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.



Future fan-out

How to predict future climates?

- No one model predicts best all the time, for al…
- Average prediction is better pre… …del
  [Reichler & Kim, B… …GRL '09]
- Coupl… …roject (CMIP)
  …ns [Meehl et al., Bull. AMS '00]
- …n approaches in climate science, e.g. [Smith et al. JASA '08]
- IPCC Expert Meeting, 2010, on how to combine model predictions

Can we do better using Machine Learning?

Challenge: Improve predictions of the IPCC ensemble

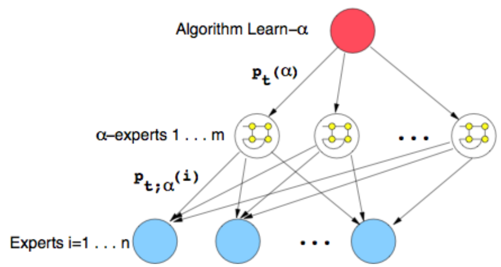
Predict future climates using past observations
and the multi-model ensemble predictions

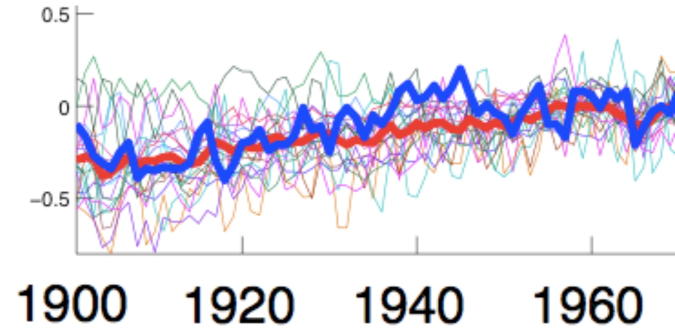
Contributions

- Tracking Climate Models (TCM) [M, Schmidt, Saroha, & Asplund, NASA CIDU 2010; SAM 2011]: Online learning with expert advice
- Neighborhood-Augmented TCM (NTCM) [McQuade & M, AAAI 2012]: Extend TCM to model geospatial neighborhood influence
- MRF-based approach [McQuade & M, book chapter, 2017]
- Multi-resolution temporal structure [McQuade & M, Climate Informatics 2015; DSMM 2016]: online multi-task learning
- Climate Prediction via Matrix Completion [Ghafarianzadeh & M, Late-Breaking Paper, AAAI 2013]: use sparse matrix completion

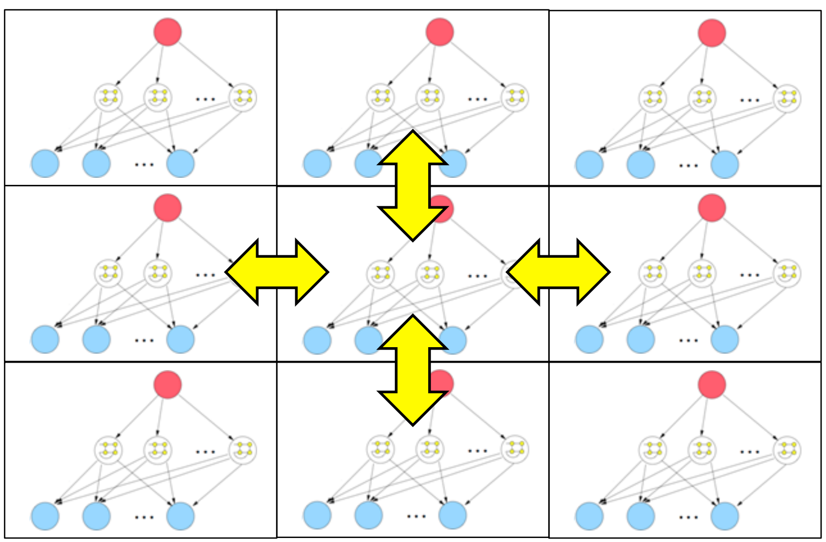
Roadmap

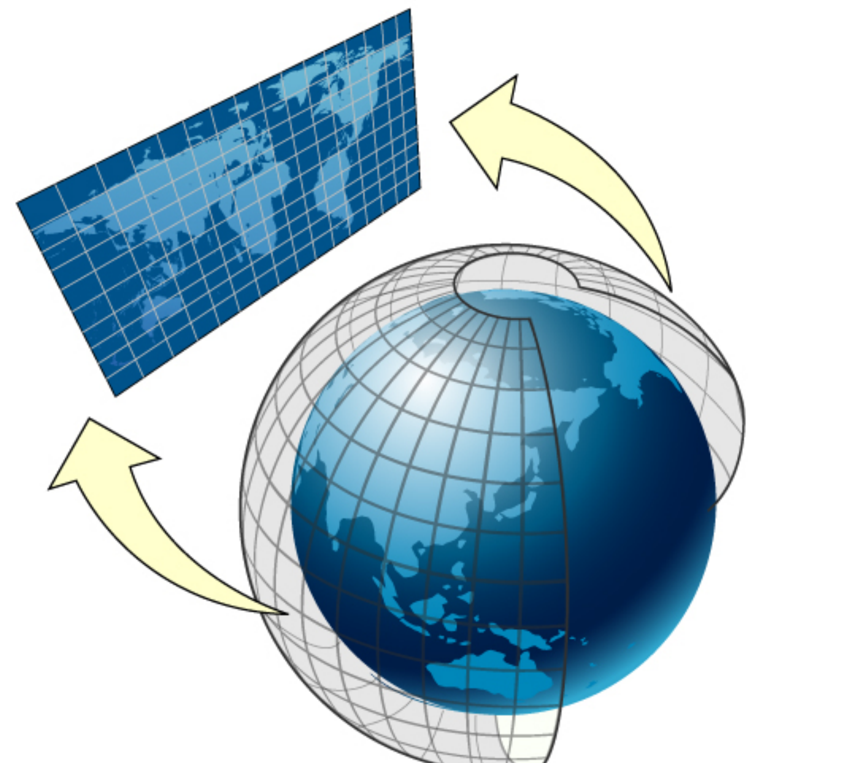
- Learning from data that varies over time





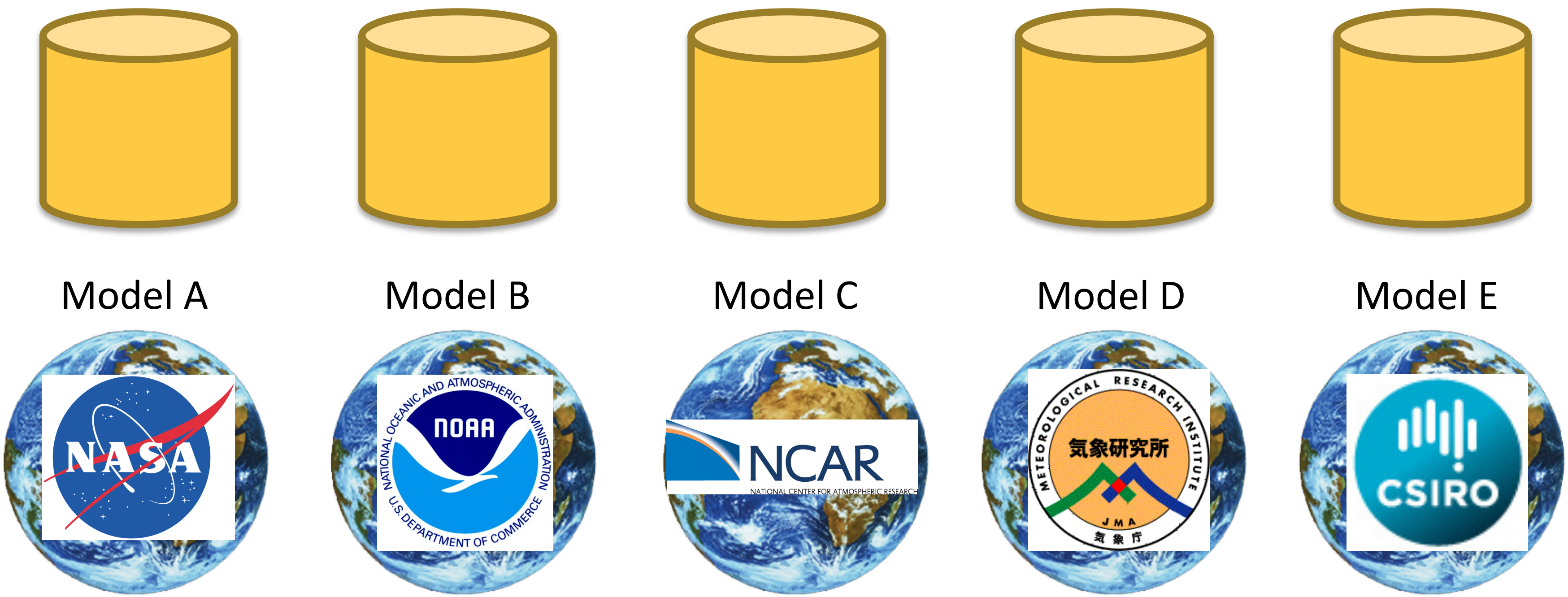
- Learning from spatiotemporal data that varies over time & space



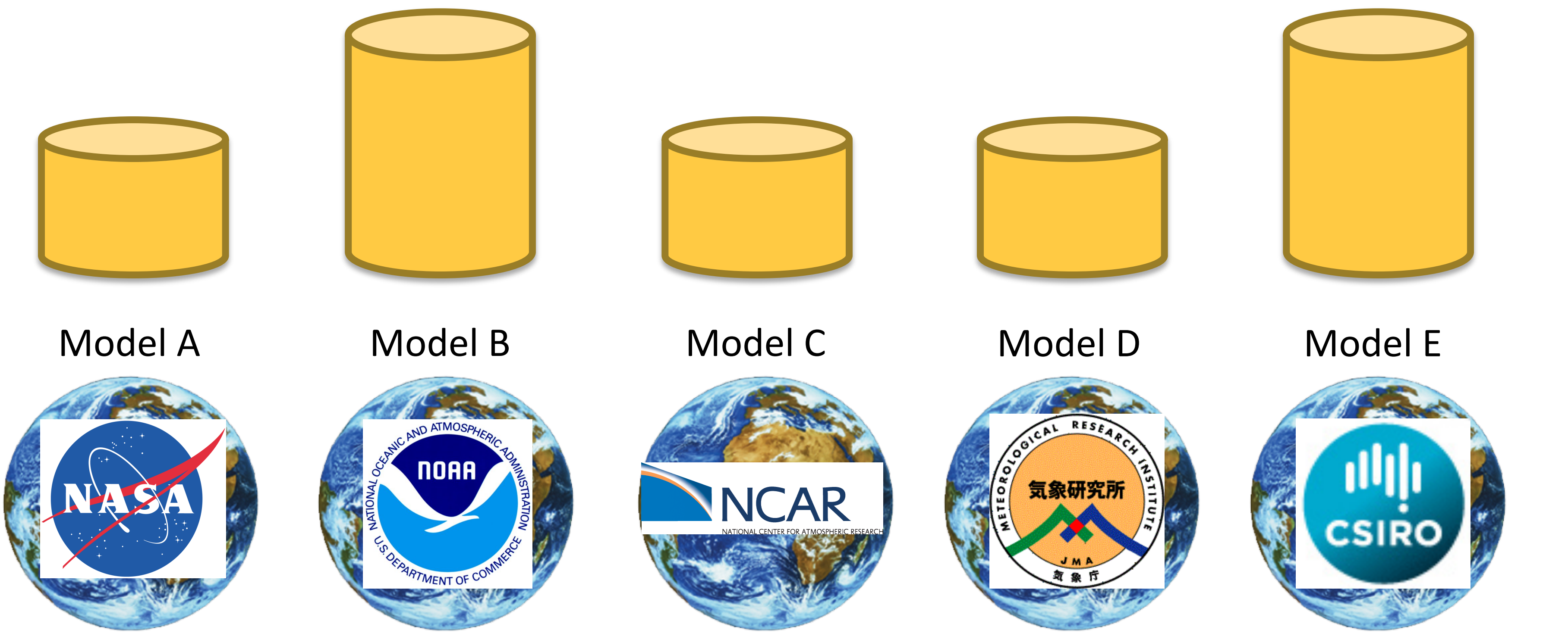


developer.apple.com

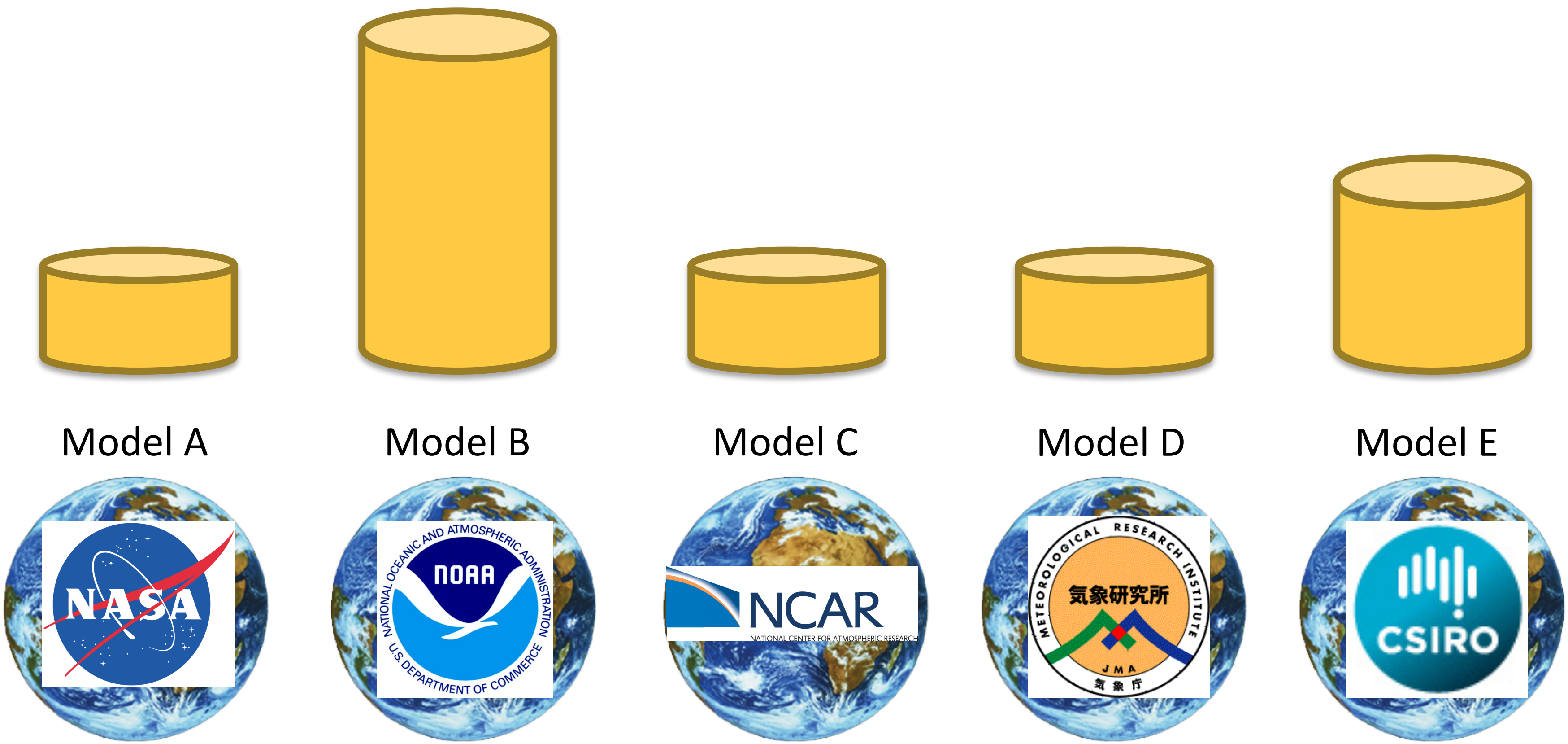
Average prediction



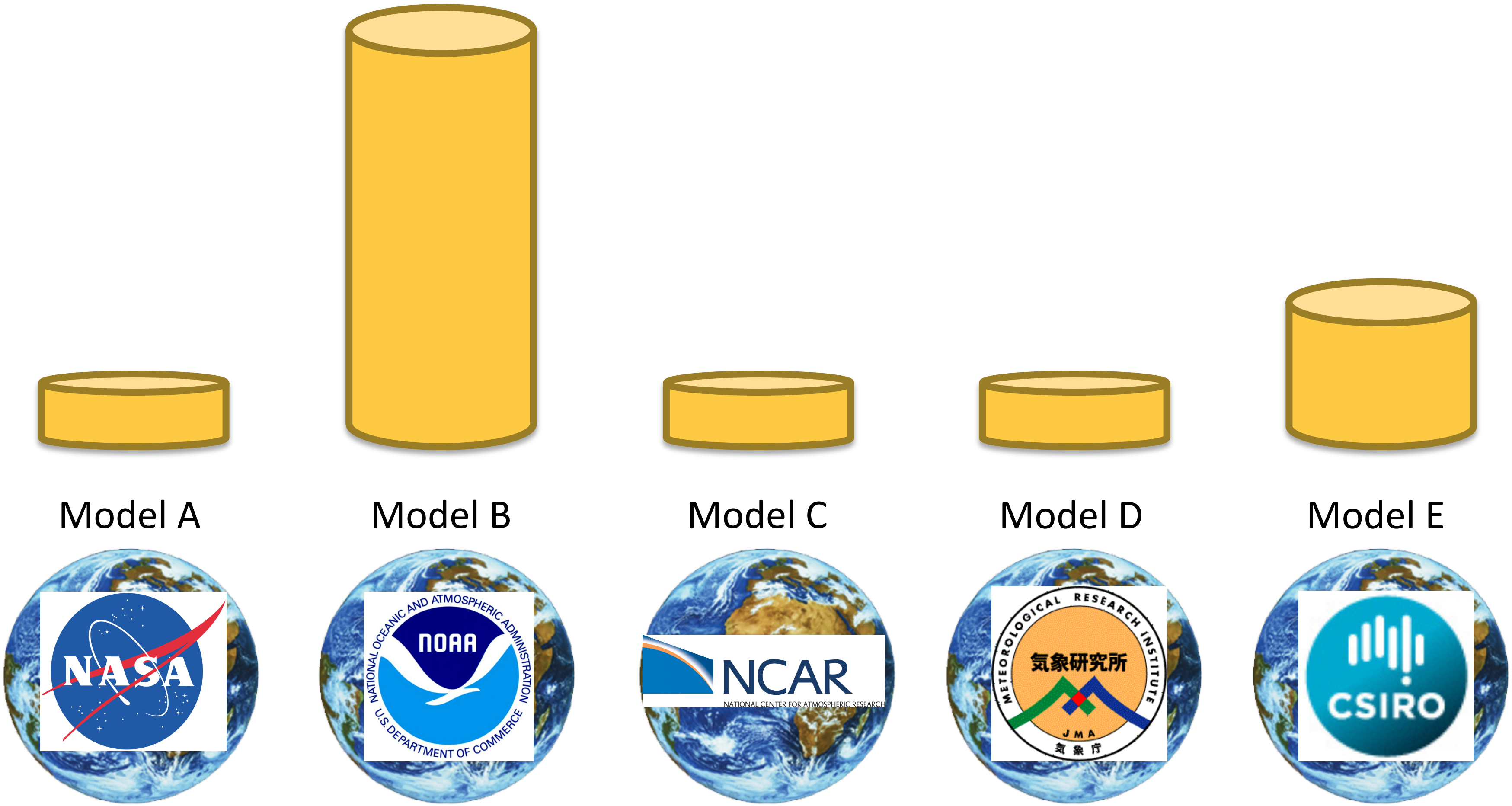
Adaptive, weighted average prediction



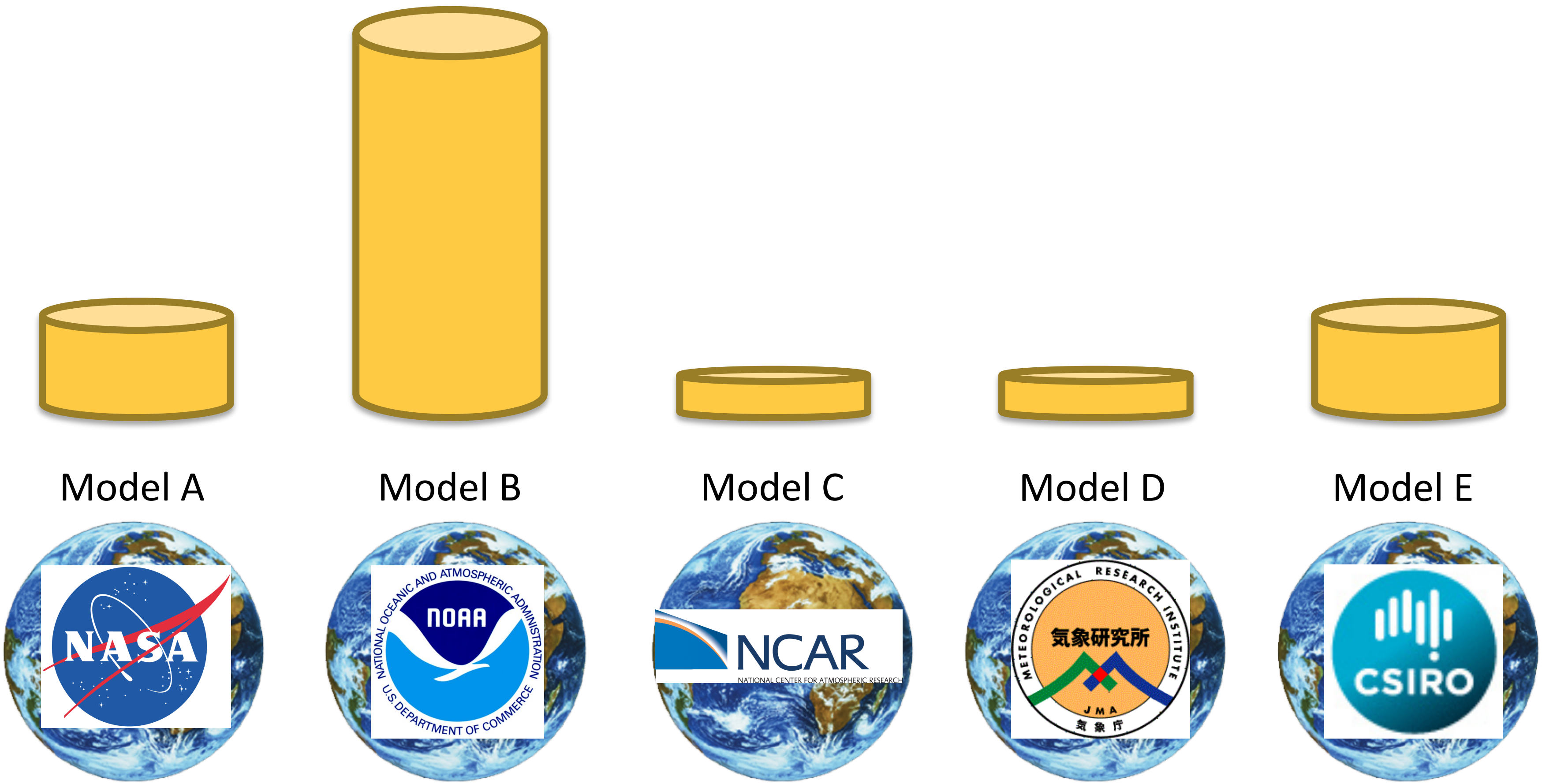
Adaptive, weighted average prediction



Adaptive, weighted average prediction



Adaptive, weighted average prediction



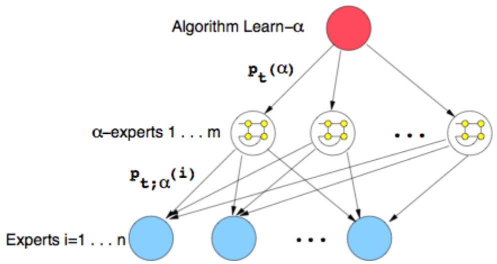
Tradeoff: explore vs. exploit

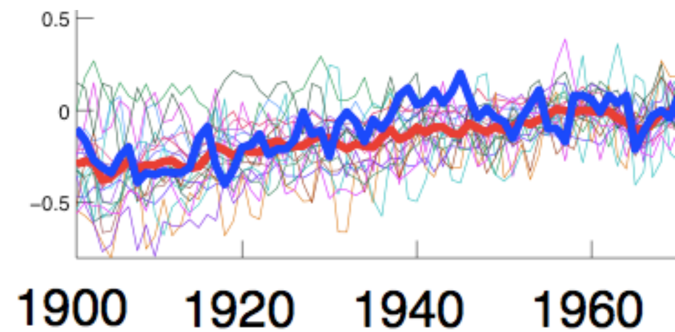
Tradeoff: Quickly finding current best predicting model vs. being ready to quickly switch to other models.

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

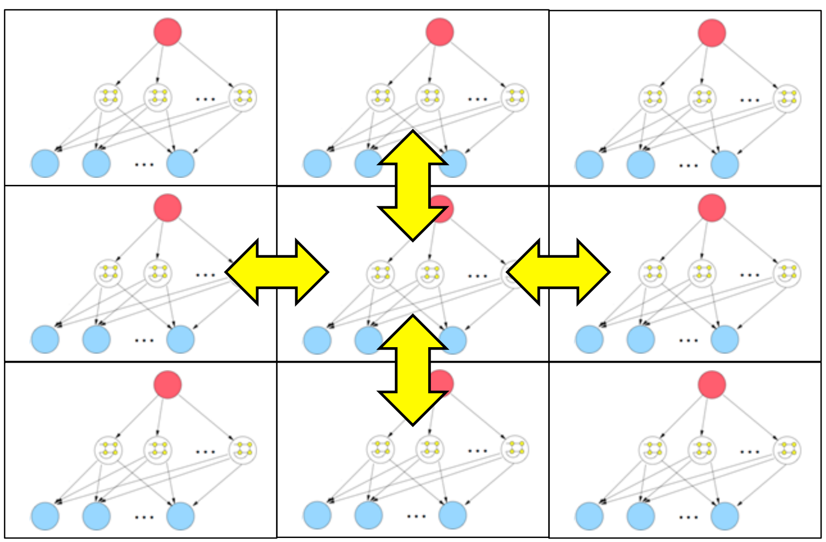
Roadmap

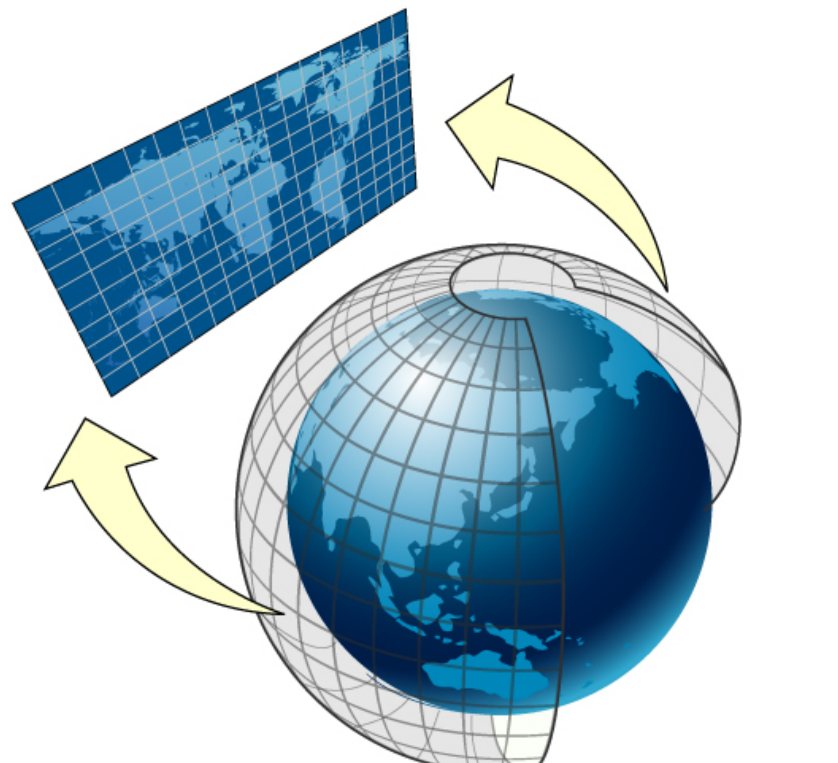
- Learning from data that varies over time





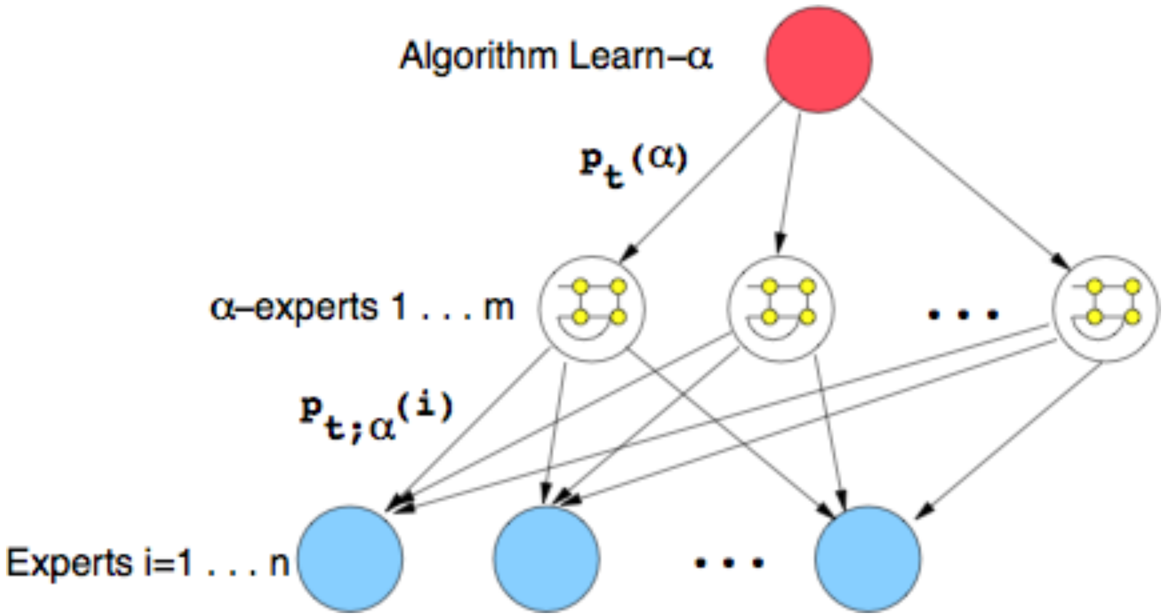
- Learning from spatiotemporal data that varies over time & space





developer.apple.com

Online learning: time-varying data



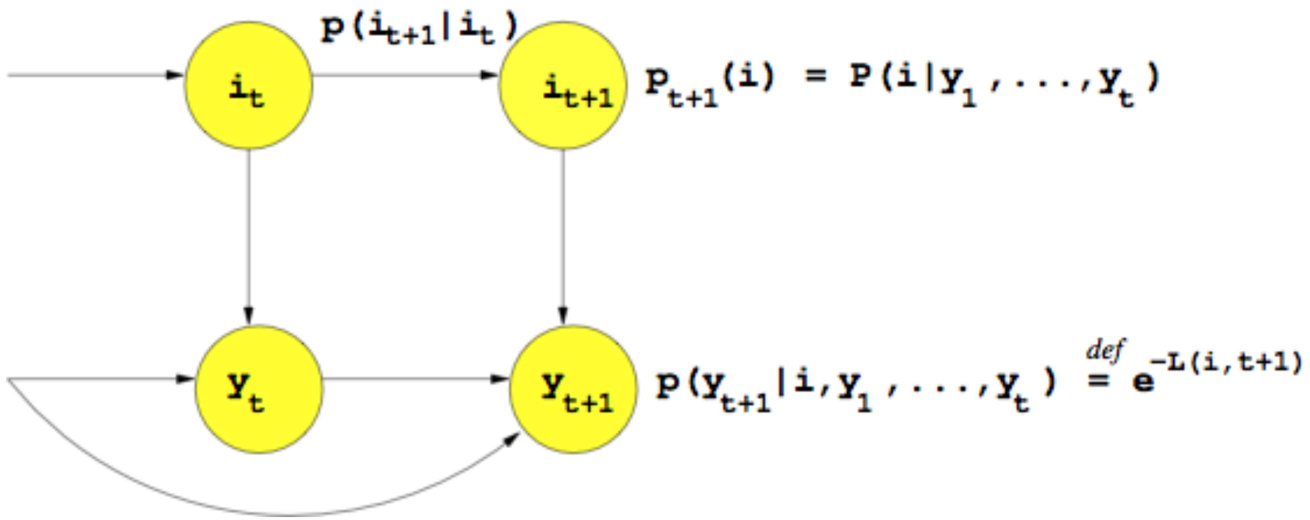
Learn-α Algorithm [M & Jaakkola, NIPS 2003]:

- Learns the switching-rate: level of non-stationarity: α
- Tracks a set of online learning algorithms, each with a different α value
- Each algorithm maintains weights over experts (e.g. climate models)

Online learning with expert advice

Model changing observations via a (generalized) Hidden Markov Model

- where hidden state is identity of the "best expert" (e.g. climate model)



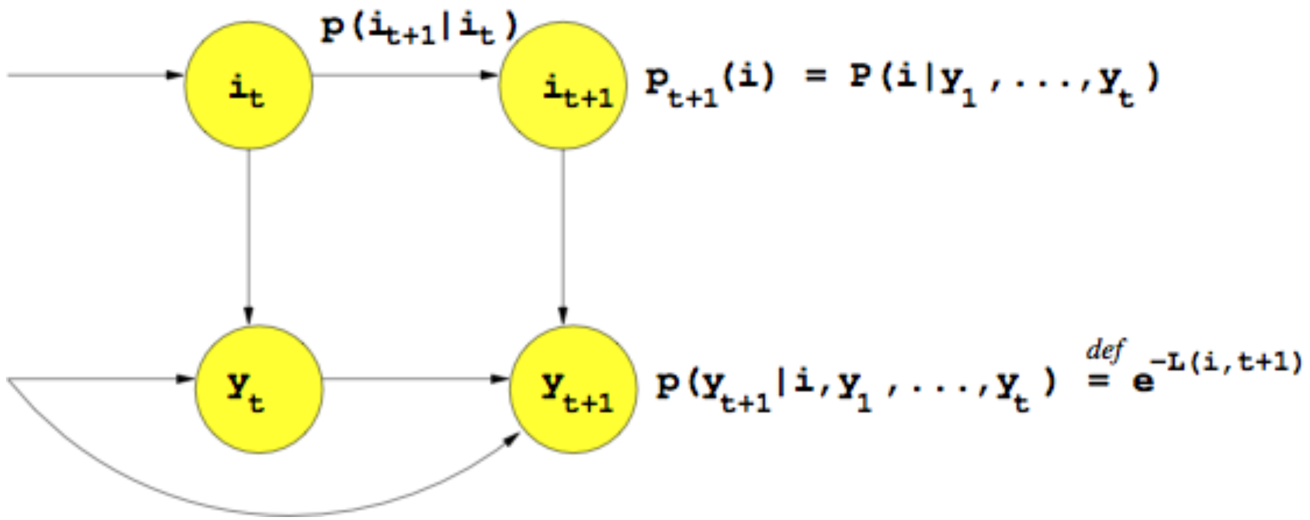
Performing Bayesian updates yields a family of online learning algorithms with transition dynamics P(i | j).

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

Online learning with expert advice

Model changing observations via a (generalized) Hidden Markov Model

- where hidden state is identity of the "best expert" (e.g. climate model)



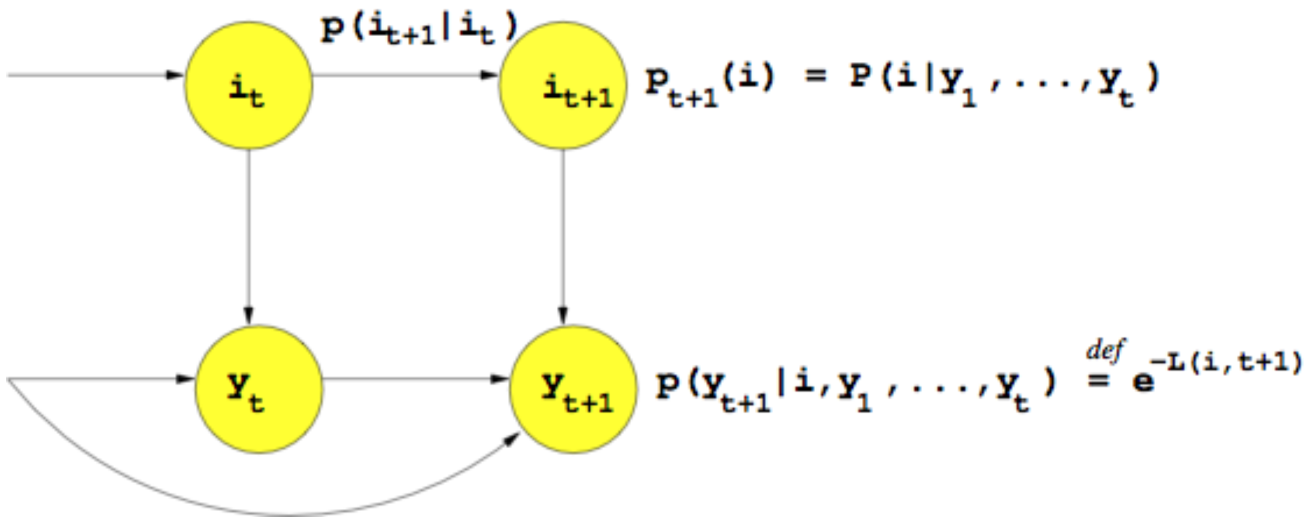
Static-Expert [Littlestone & Warmuth'89], Hedge [Freund & Schapire '97] algorithm: P(i | j) = δ(i,j).

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

Online learning: time-varying data

Model changing observations via a (generalized) Hidden Markov Model

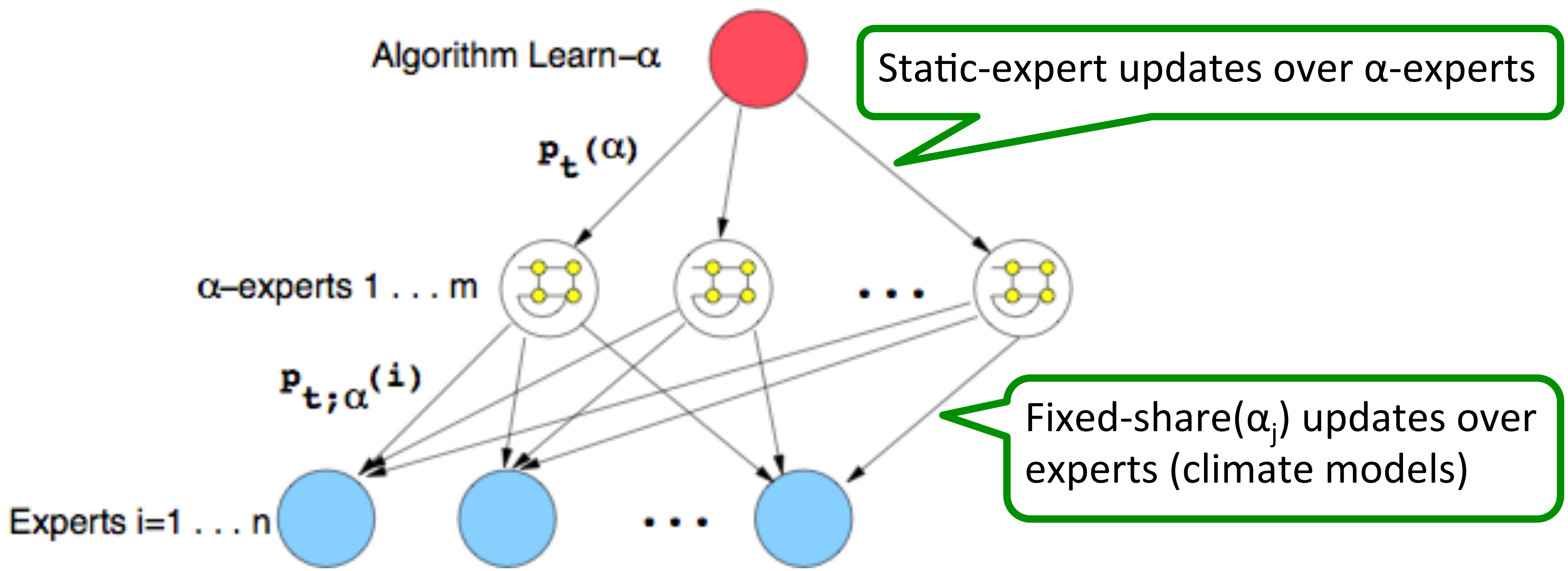
- where hidden state is identity of the "best expert" (e.g. climate model)



[Herbster & Warmuth, '98]: Fixed-Share algorithm models switching w.p. α

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

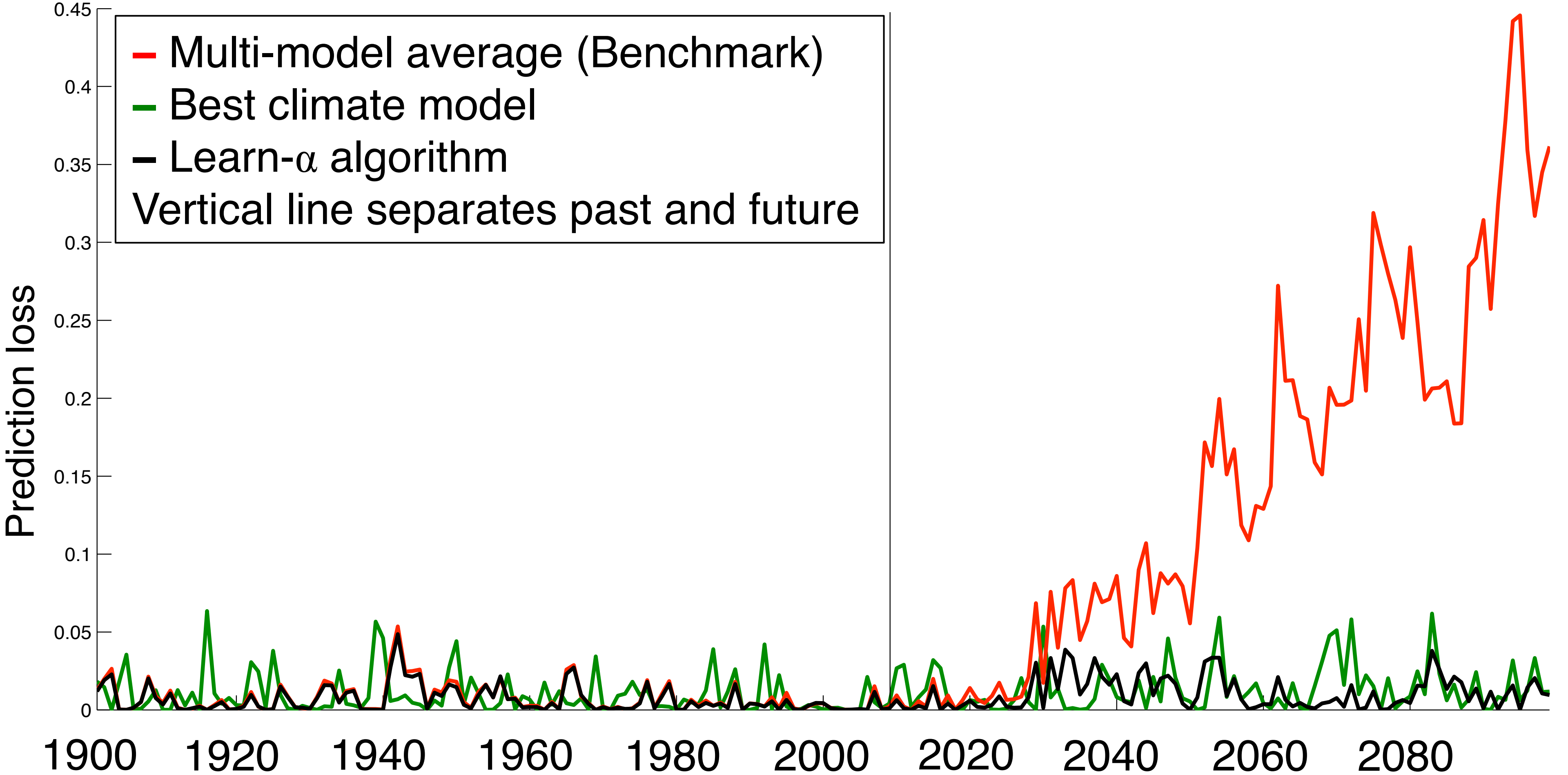
Online learning: time-varying data



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

- Learns the switching-rate: level of non-stationarity: α
- Tracks a set of online learning algorithms, each with a different α value
- Each algorithm maintains weights over experts (e.g. climate models)

Survey of related work: [Koolen & de Rooij, Trans. Info Theory 2013]

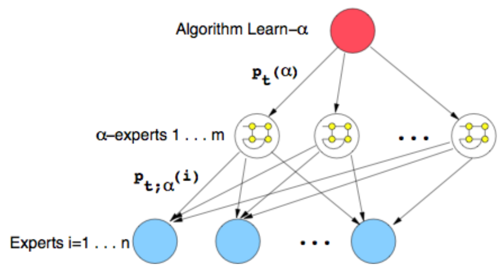


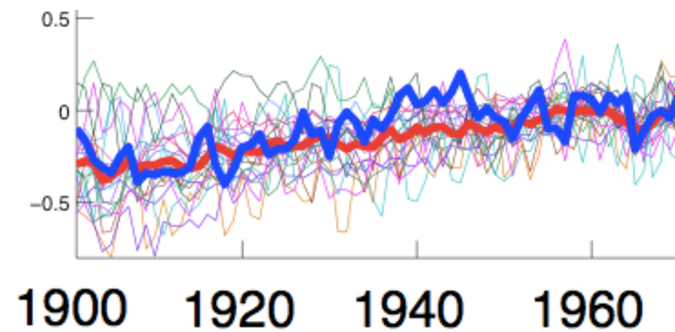
Learning curves

[M, Schmidt, Saroha, & Asplund, NASA CIDU 2010; SAM 2011]

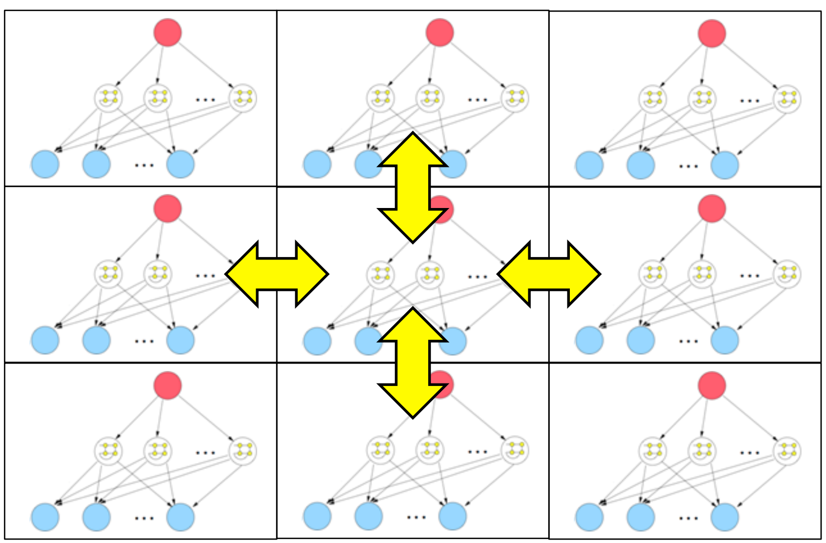
Roadmap

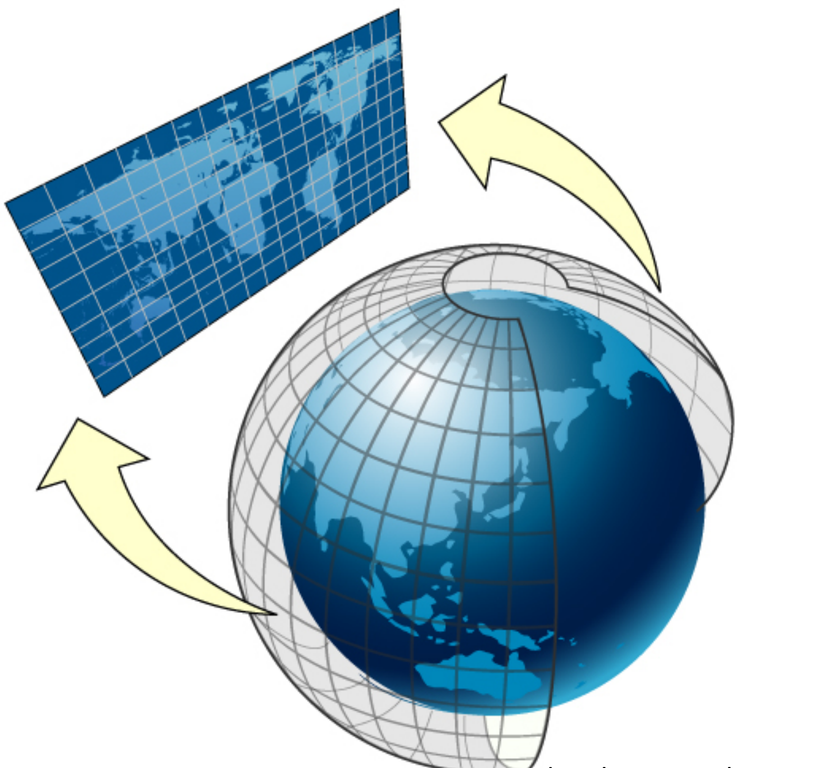
- Learning from data that varies over time





- Learning from spatiotemporal data that varies over time & space
  - Model spatial influence





developer.apple.com

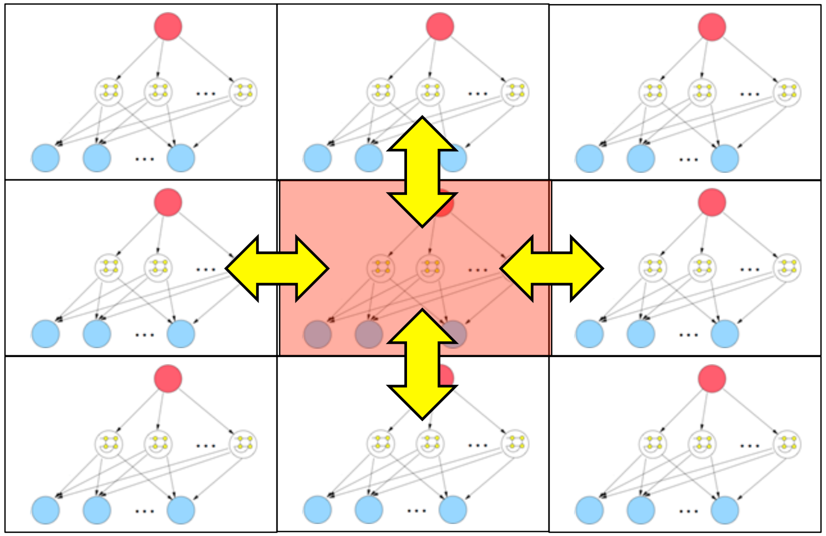
Of general interest for ML:

Online + Spatial

Online learning: spatiotemporal data

[McQuade & M, AAAI 2012]

- Climate predictions are made at higher geospatial resolutions
- Run Learn-α (variant) on multiple sub-regions partitioning globe
- Distribution over climate models varies over both time and space
- Model neighborhood influences among geospatial regions



Incorporating neighborhood influence

Neighborhood-augmented Learn-α

– Non-homogenous HMM transition dynamics:

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

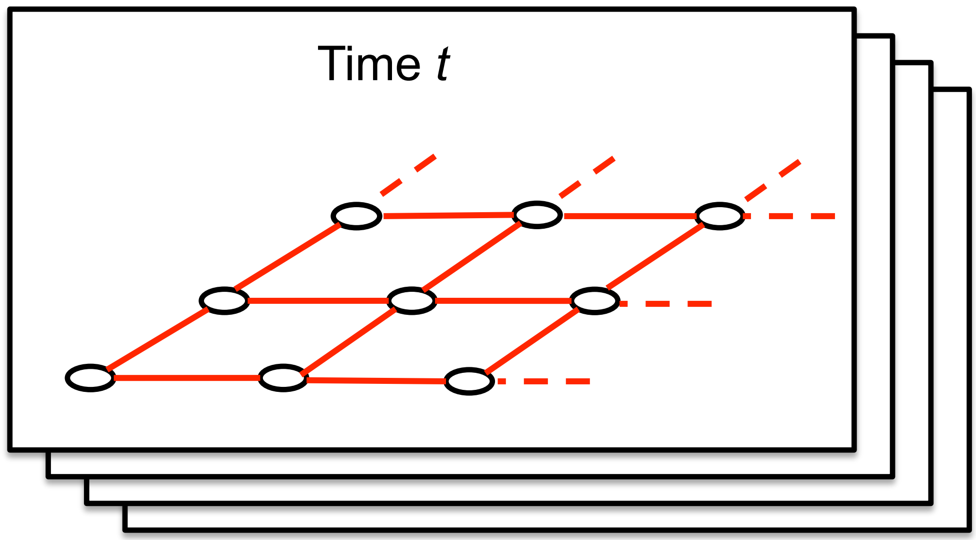
Increase probability of transitioning to a model performing well in neighborhood

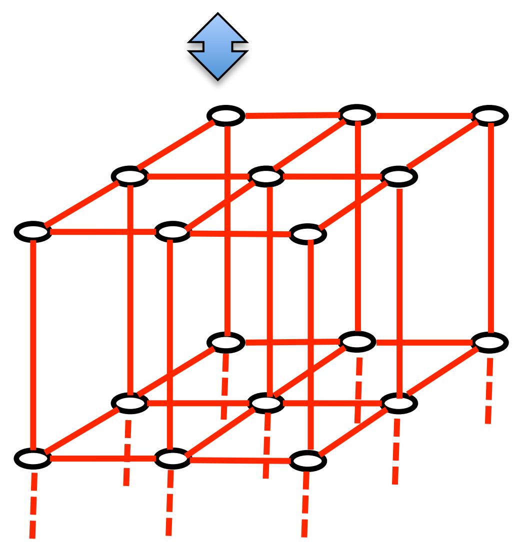
- *S(r)* - neighborhood scheme: set of "neighbors" of region *r*
- $P_{t,s}(i)$ - probability of expert (climate model) *i* in region *s*
- β - regulates geospatial influence
- *Z* - normalization factor

Markov Random Field-based approach

[McQuade & M, book chapter, 2017]

Geospatial lattice

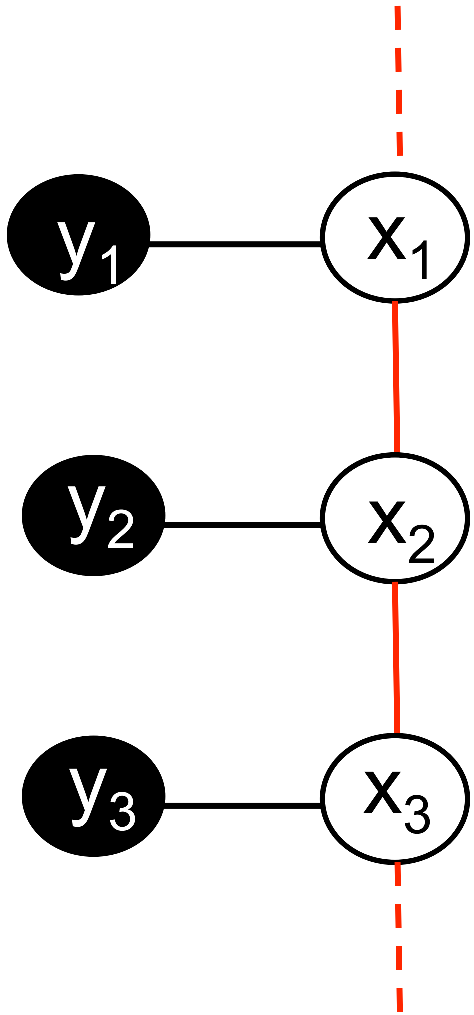




Time *t*

Time *t-1*

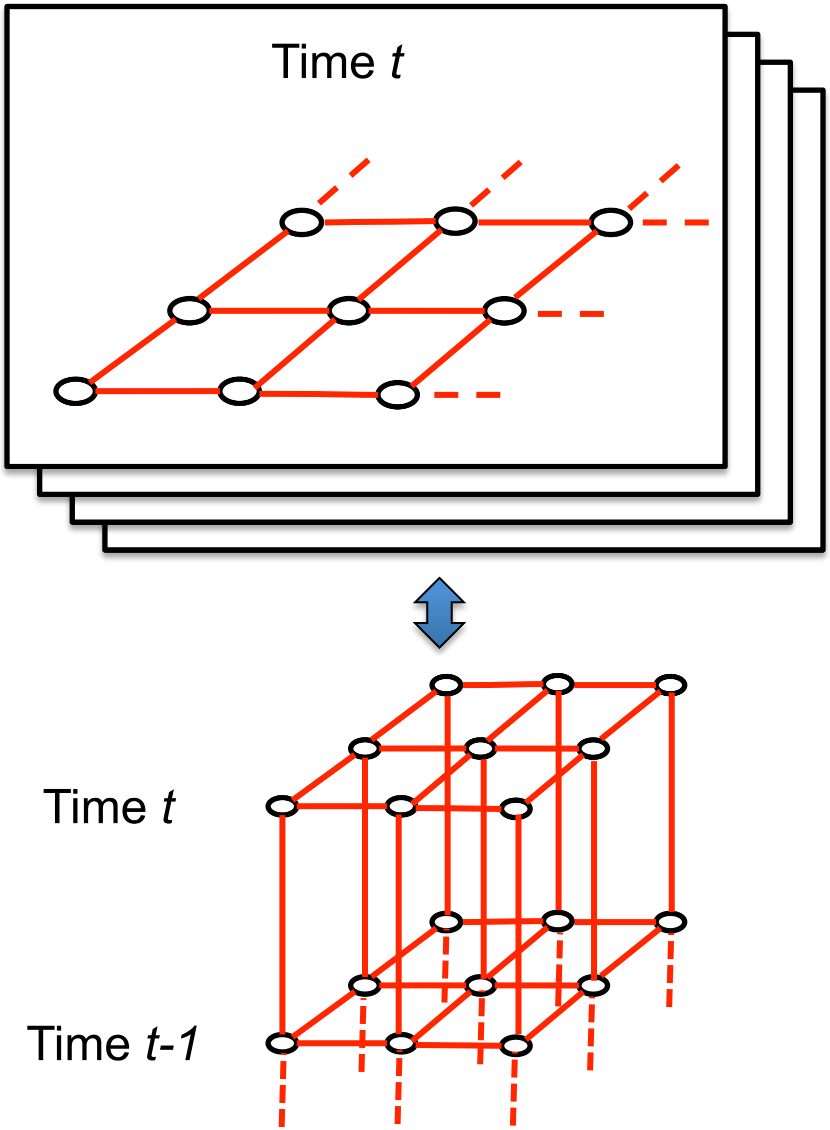
Markov Random Fields



Extend from HMM for time series to hidden Markov random field (MRF) for spatiotemporal field

- Model local spatial dependencies

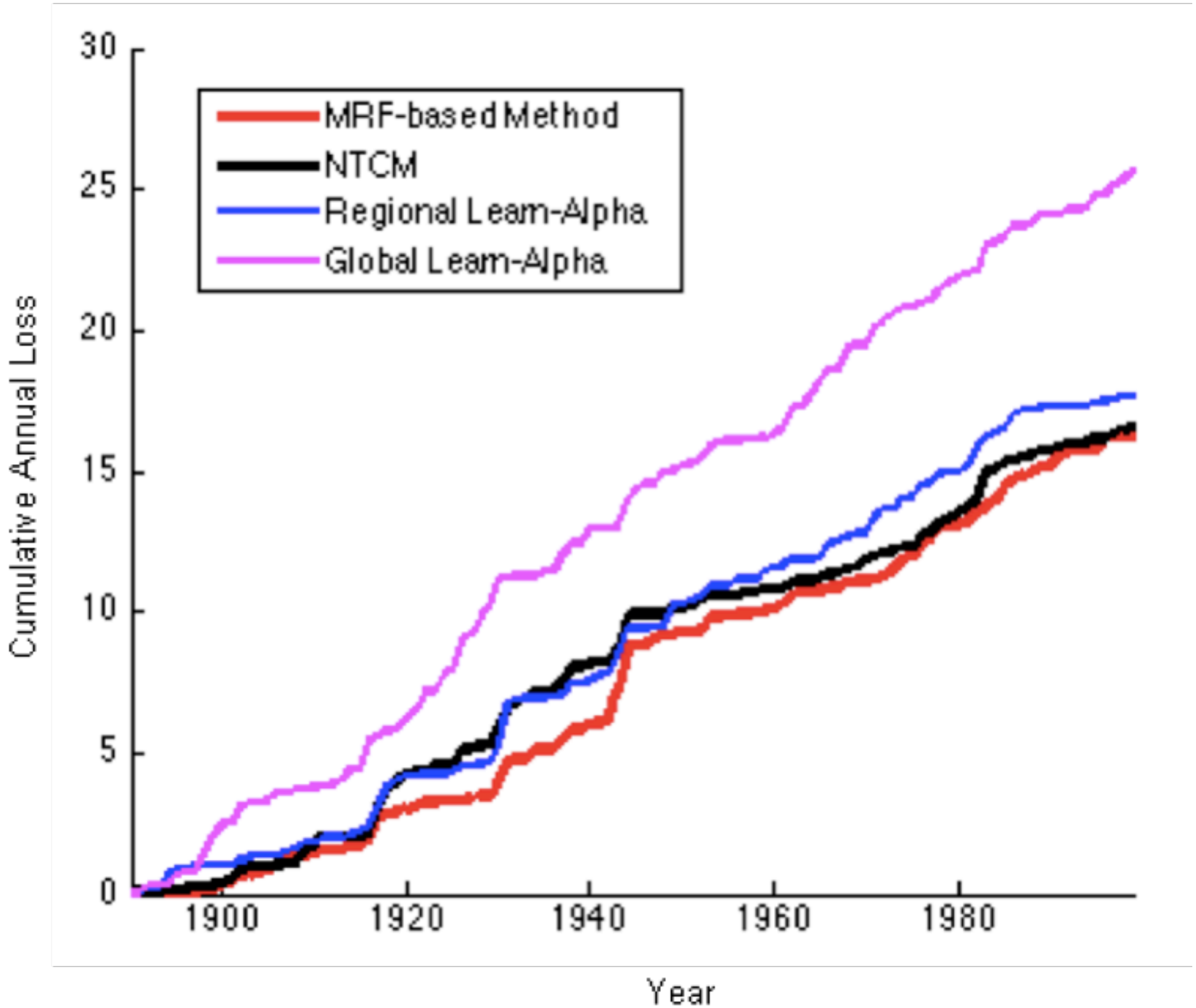
Extending Fixed-Share to a Spatial Lattice



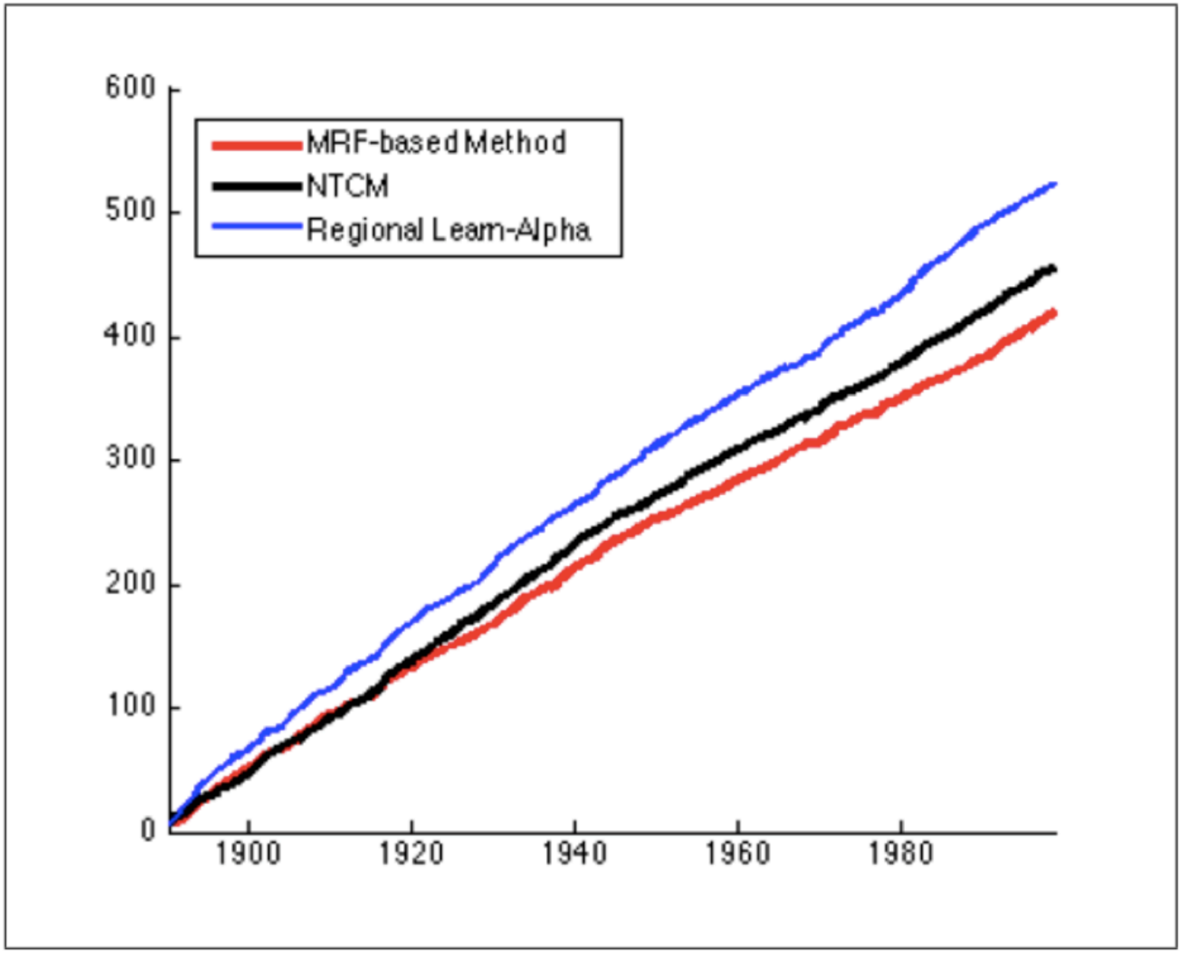
Construct spatial lattice

- Spatial dependencies of same form as temporal dependencies
- Different α_{time} and α_{space} parameters
- Latent variables: best expert at each time and location
- Need to compute marginals

Global prediction loss



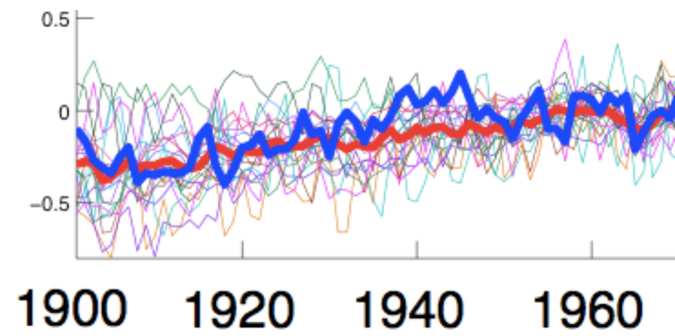
Regional prediction loss

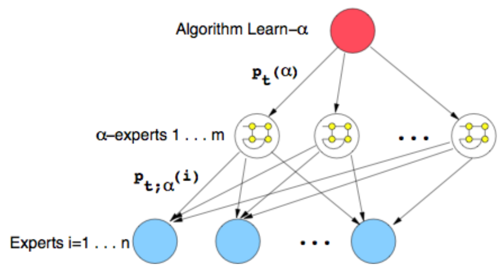


Cumulative mean regional loss of the hindcast.

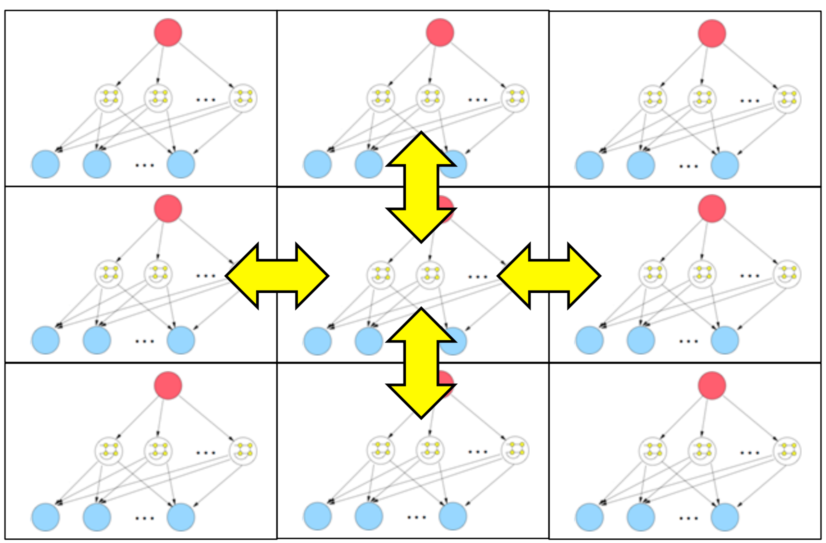
Roadmap

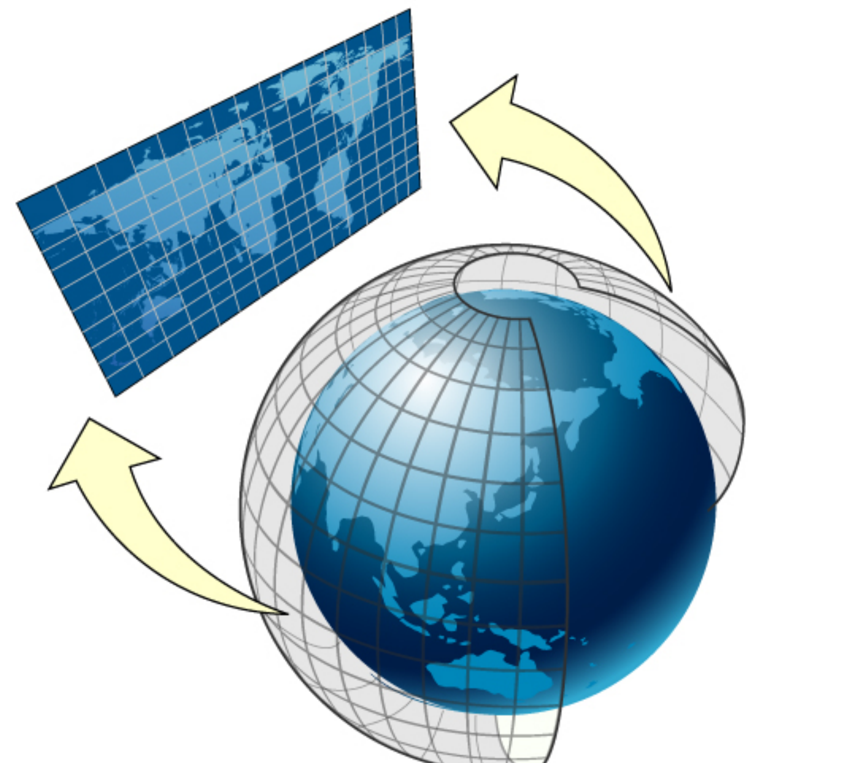
- Learning from data that varies over time





- Learning from spatiotemporal data that varies over time & space
  - Model spatial influence
  - Model temporal influence





developer.apple.com

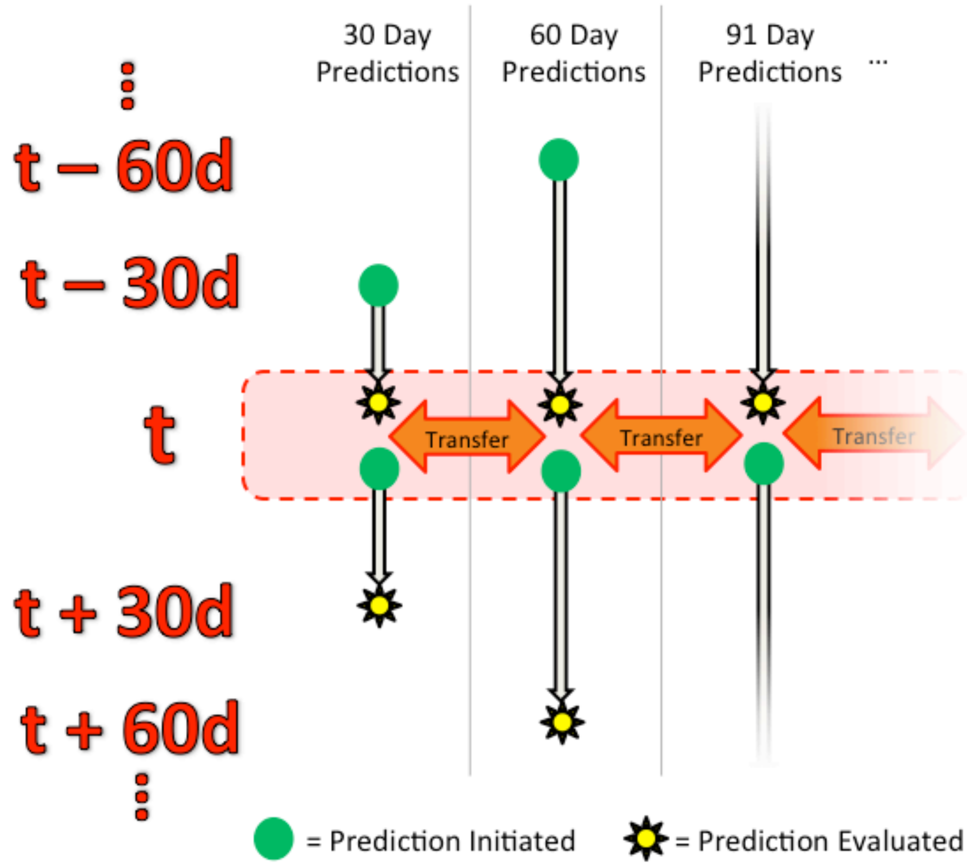
Of general interest for ML:

Prediction at multiple timescales

Seasonal prediction: Online multi-task learning

[McQuade & M, Climate Informatics 2015; SIGMOD DSMM 2016]

- Given forecasts of multiple time periods
- Each forecast period treated as a different task
- Allow influence between tasks



Online multi-task learning

Task-similarity matrix [*cf.* Saha et al., AISTATS 2011]

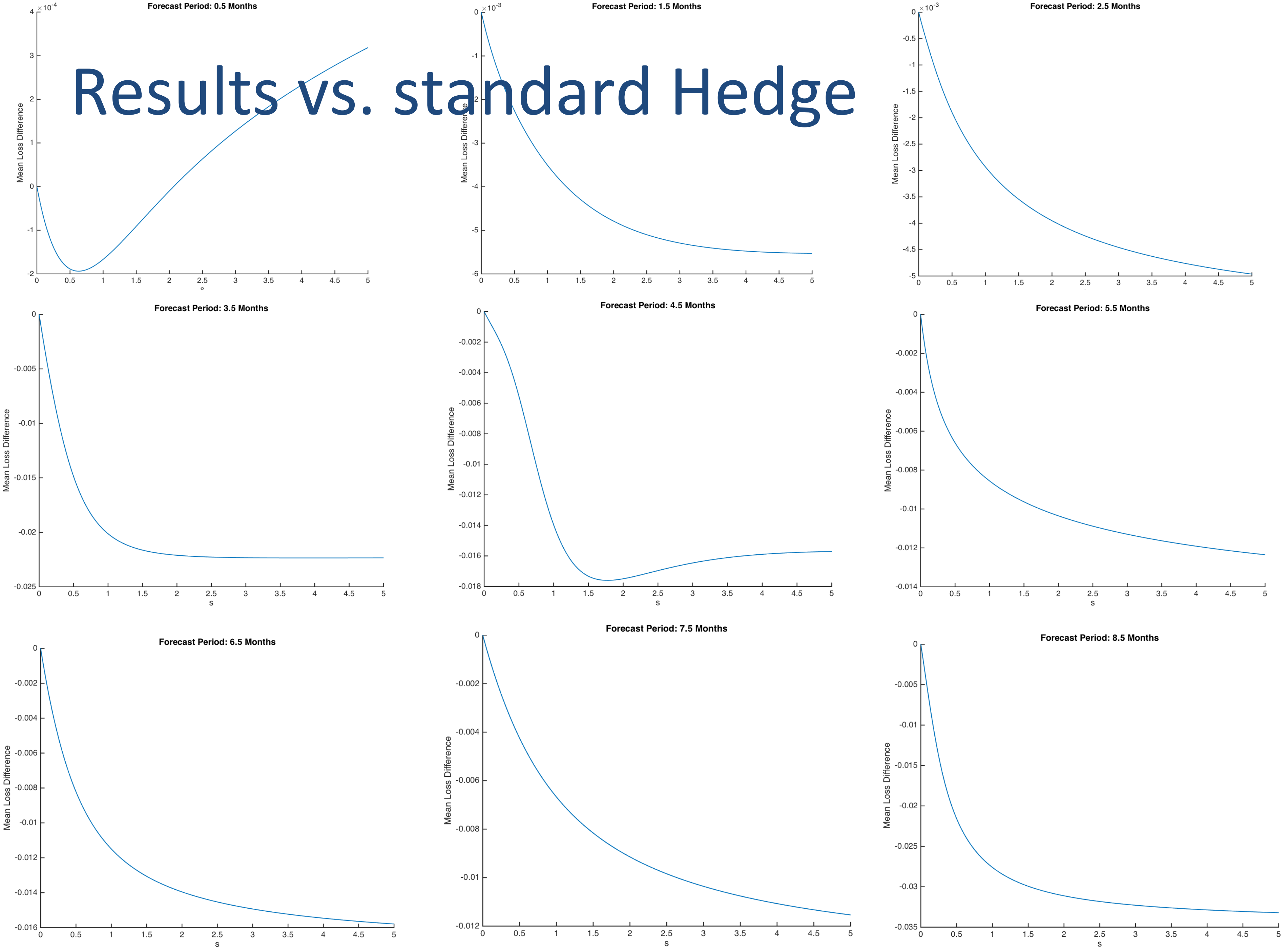
- Allow influence between "neighboring" forecast lengths, parameterized by s

$$S = \begin{array}{cc} \text{Months:} & \begin{array}{cccccc} 0.5 & 1.5 & 2.5 & 3.5 & \cdots & 11.5 \end{array} \\ \begin{array}{c} 0.5 \\ 1.5 \\ \cdots \\ 11.5 \end{array} & \begin{bmatrix} \frac{1}{1+s} & \frac{s}{1+s} & 0 & 0 & \cdots & \\ 0 & \frac{s}{1+2s} & \frac{1}{1+2s} & \frac{s}{1+2s} & \ddots & \\ \vdots & \ddots & \ddots & \ddots & \ddots & \\ & & & & & \frac{1}{1+s} \end{bmatrix} \end{array}$$

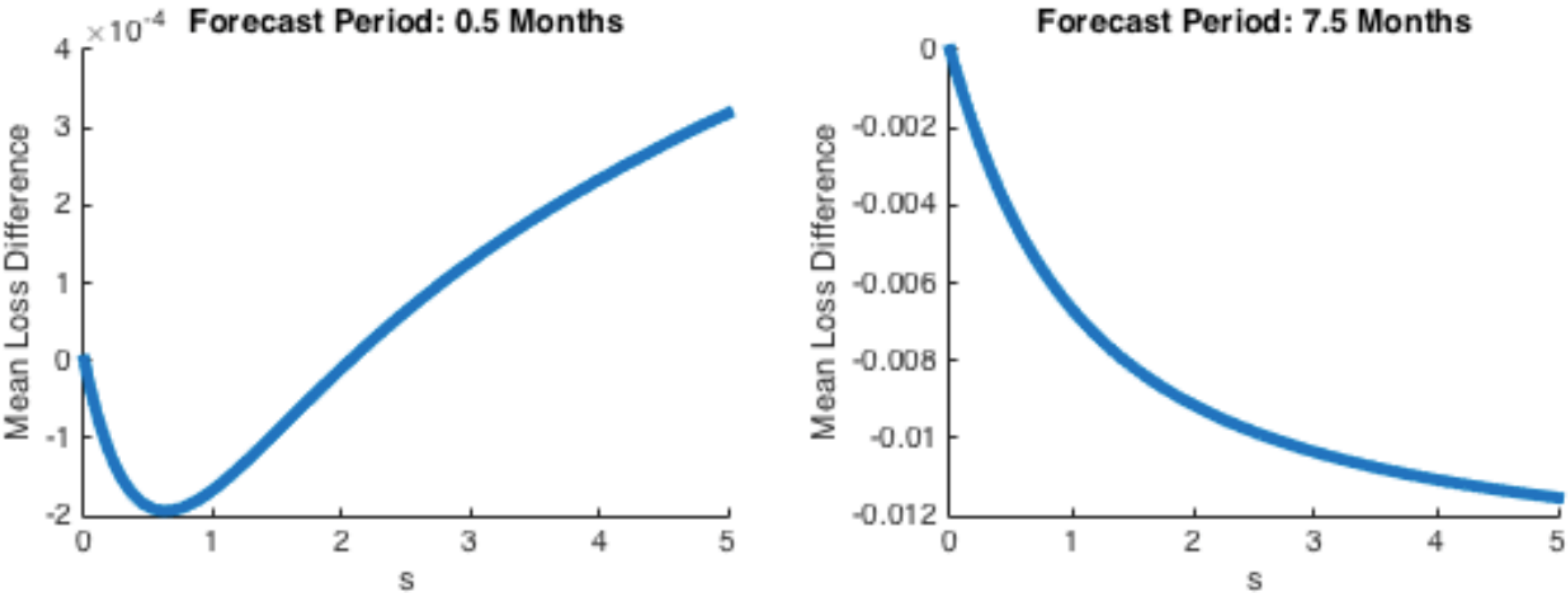
Multi-task update rule (extended from Hedge / Static-Expert algorithm)

$$p_{t+1,j}(i) \propto p_{t,j}(i) e^{-\sum_k S_{j,k} L_{k,t}(i)}$$

Results vs. standard Hedge



Results vs. standard Hedge



For each of the 12 forecast periods, sharing influence from other forecast periods improved predictions.

Only for the 2 forecast periods was loss increased for some s values.

Application to financial volatility prediction [McQuade & M, DSMM 2016]

Future work

About those hurricanes this Fall…

Question:

Can ML improve understanding, prediction of the following?

- Geographical movement (tracks) of storms
- Intensification into severe storms
  - hurricanes, cyclones, typhoons

Sub-question:

Is deep learning an effective approach to such problems?

ML for severe storms: related work

- Meteorology community has used "method of analogs"
  - essentially Nearest Neighbor

- Object tracking in ML/CV seems to rely on some level of rigidity
  - But if you know of algorithms for tracking non-rigid patterns (e.g. in fluids), please let us know!

- ConvLSTM used for precipitation nowcasting [Shi et al., NIPS '15]

- CNN for classification of cyclones from static images [Liu et al., '16]
  - Static: no time dimension
  - Highly curated "easy" data set

Of general interest for ML / CV:

Tracking highly-deformable patterns

Take-home messages

Climate informatics is a compelling application area for ML

For further info see our NIPS 2014 tutorial

Algorithms for learning when the concept can vary over multiple dimensions

We can learn the level of non-stationarity in time, space

We can exploit local structure in space and time

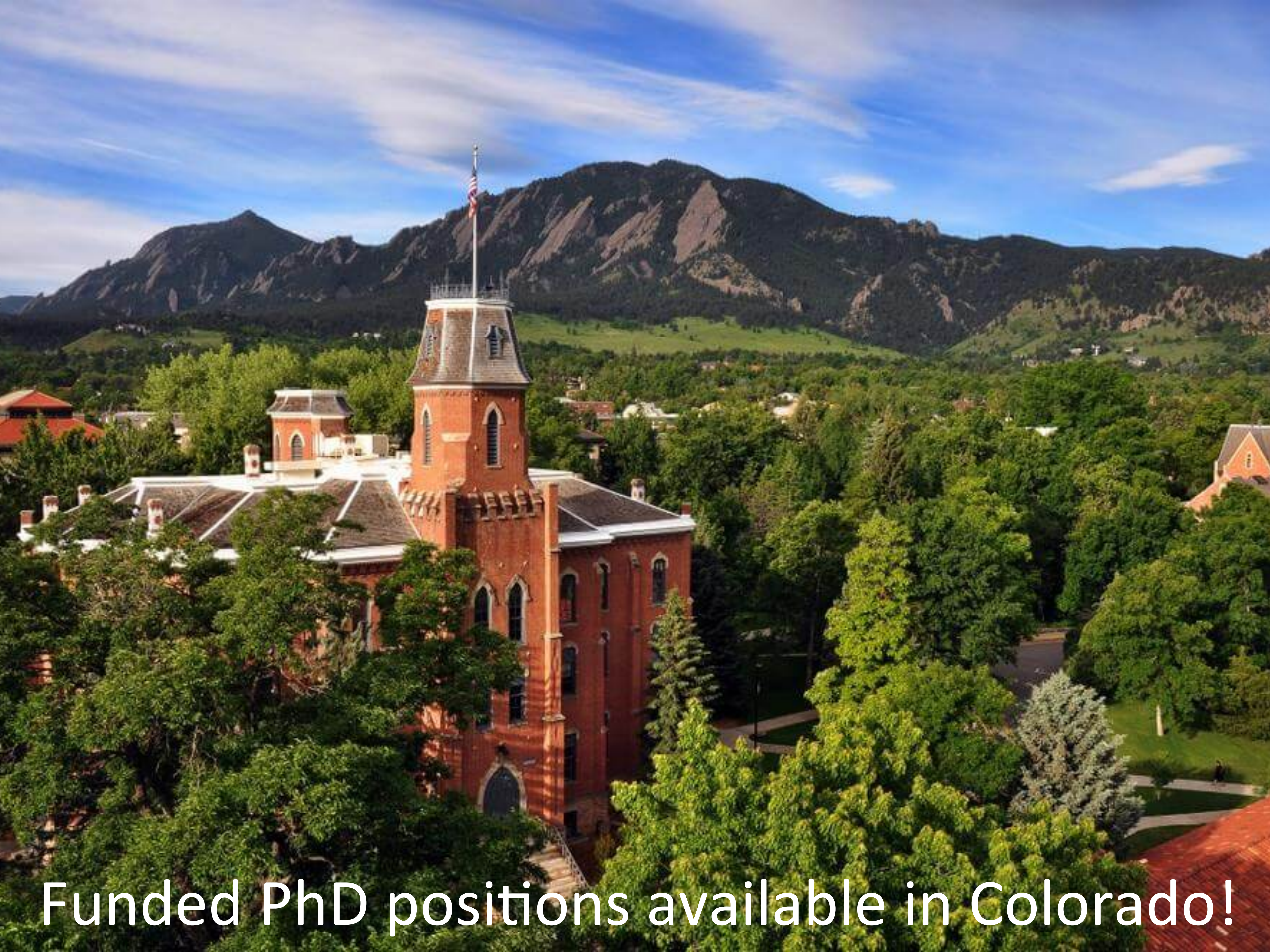
This is a rich area with remaining open questions

Applications can pose interesting new questions for ML

Online + spatial

Prediction at multiple timescales

Tracking highly-deformable patterns


Funded PhD positions available in Colorado!

Thank you! *And thanks to my collaborators:*

Frank Alexander, *Los Alamos National Laboratory*
Eva Asplund, *Barnard College, Columbia University*
Arindam Banerjee, *University of Minnesota*
M. Benno Blumenthal, *International Research Institute for Climate and Society, Columbia U.*
Tim DelSole, *George Mason University & Center for Ocean-Land-Atmosphere Studies*
Auroop R. Ganguly, *Civil and Environmental Engineering, Northeastern University*
Mahsa Ghafarianzadeh, *George Washington University*
Balázs Kégl, *Université Paris-Saclay*
Scott McQuade, *George Washington University*
Doug Nychka, *National Center for Atmospheric Research*
Alex Niculescu-Mizil, *NEC Laboratories America*
Kathleen Pegion, *George Mason University & Center for Ocean-Land-Atmosphere Studies*
Shailesh Saroha, *Amazon.com*
Gavin A. Schmidt, *NASA GISS & Columbia University*
Jason E. Smerdon, *Lamont-Doherty Earth Observatory, Columbia University*
Karsten Steinhaeuser, *University of Minnesota*
Cheng Tang, *George Washington University*
Marco Tedesco, *NSF & CUNY City College and Graduate Center*
Michael Tippett, *Columbia University*

Resources

- Climate Informatics: www.climateinformatics.org
  - Links to resources, Climate Informatics workshops, online community
- 8th International Workshop on Climate Informatics, 2018
  www2.cisl.ucar.edu/events/workshops/climate-informatics/2018/home