Introduction to Bayesian Methods

Dr. Michel d. S. Mesquita[1,2]

[1]Uni Research, Bergen, Norway

[2]Bjerknes Centre for Climate Research, Norway

NCAR Regional Climate Downscaling Tutorial - July 26, 2013

Objectives

- To grasp the need for Bayesian statistics in climate research
- To become familiar with Bayesian thinking
- To understand Bayesian data analysis methods applied to climate studies

Lecture and lab sessions

- **July 26, 2013**
  - Lecture: Introduction to Bayesian Methods
- **Online lab sessions**
  - Module name: “Bayesian Data Analysis for Climate Model Evaluation”
  - www.m2lab.org

What you need to get started:

- Software: R (www.r-project.org)
- Books:
  - Bolstad, W.M. (2007) Introduction to Bayesian Statistics, 2nd ed., Hoboken, Wiley.
  - Kruschke, J.M. (2010) Doing Bayesian Data Analysis: A Tutorial with R and BUGS, Waltham, Elsevier.
- Papers:
  - Knutti et al. (2010) Challenges in combining projections from multiple climate models. Journal of Climate, 23, pp. 2739-2758.
  - Lopez et al. (2006) Two Approaches to Quantifying Uncertainty in Global Temperature Changes. Journal of Climate, 19, pp. 4785-4796.
  - Smith et al. (2009) Bayesian modeling of uncertainty in ensembles of climate models. Journal of the American Statistical Association, 104, pp. 97-116.
  - Stephenson et al. (2011) Statistical problems in the probabilistic prediction of climate change. Environmetrics, 23, pp. 364-372.
  - Tebaldi et al. (2005) Quantifying uncertainty in projections of regional climate change: a Bayesian approach to analysis of multimodel ensembles. Journal of Climate, 18, pp. 1524-1540.

Why Bayesian statistics?

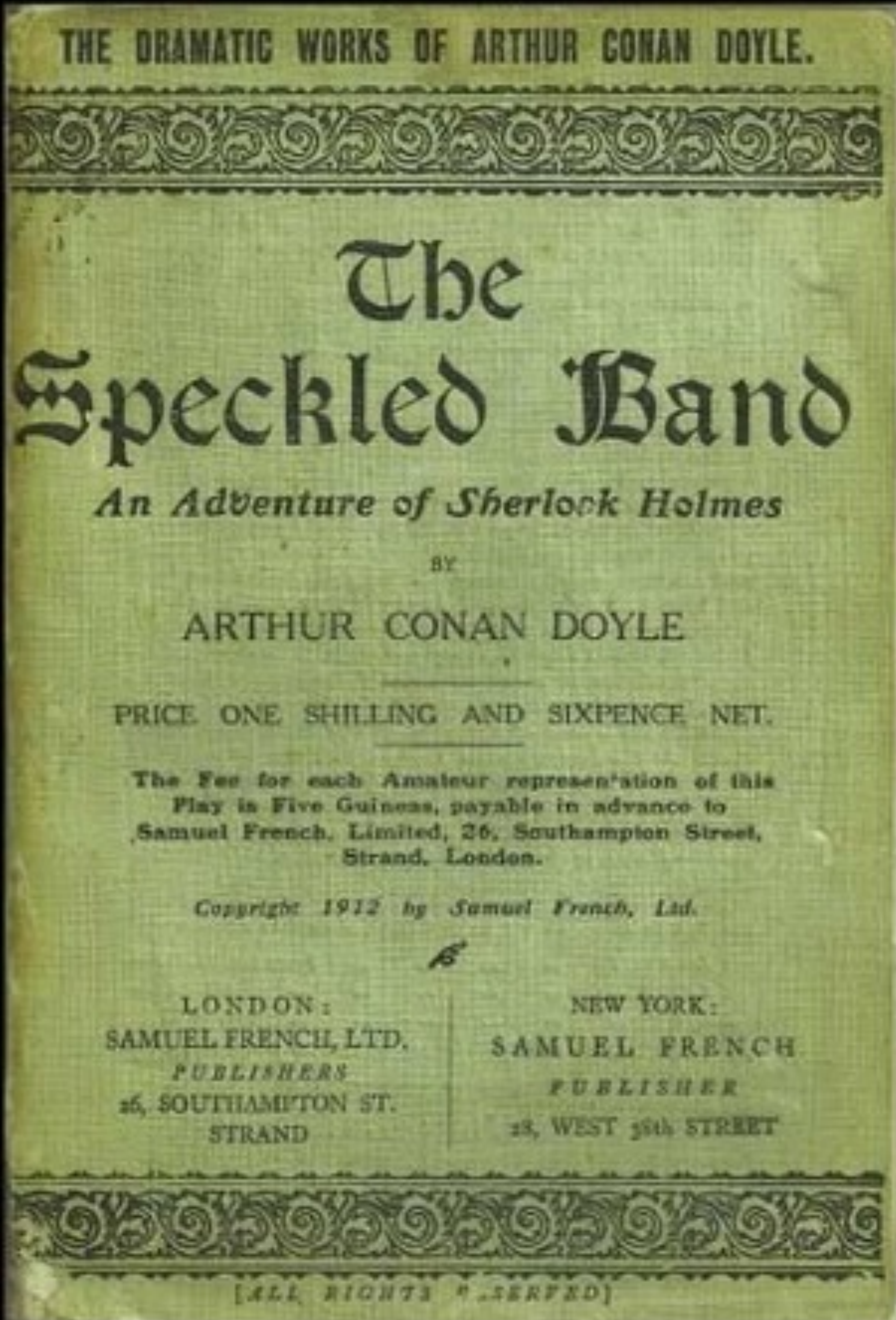
Lopez et al. (2006)

- *"The Bayesian approach is motivated by observing that the earth's climate is a nonrepeatable experiment, and probabilities cannot be determined through a frequentist approach, as that would require a sample from a large number of planets earth and their corresponding climates"*

Stephenson et al. (2011)

- *"Because we only have one realisation of real climate, the relationship of real climate to the MME requires a subjective Bayesian interpretation of probability"*

Background


THE DRAMATIC WORKS OF ARTHUR CONAN DOYLE.

The Speckled Band

An Adventure of Sherlock Holmes

BY

ARTHUR CONAN DOYLE

PRICE ONE SHILLING AND SIXPENCE NET.

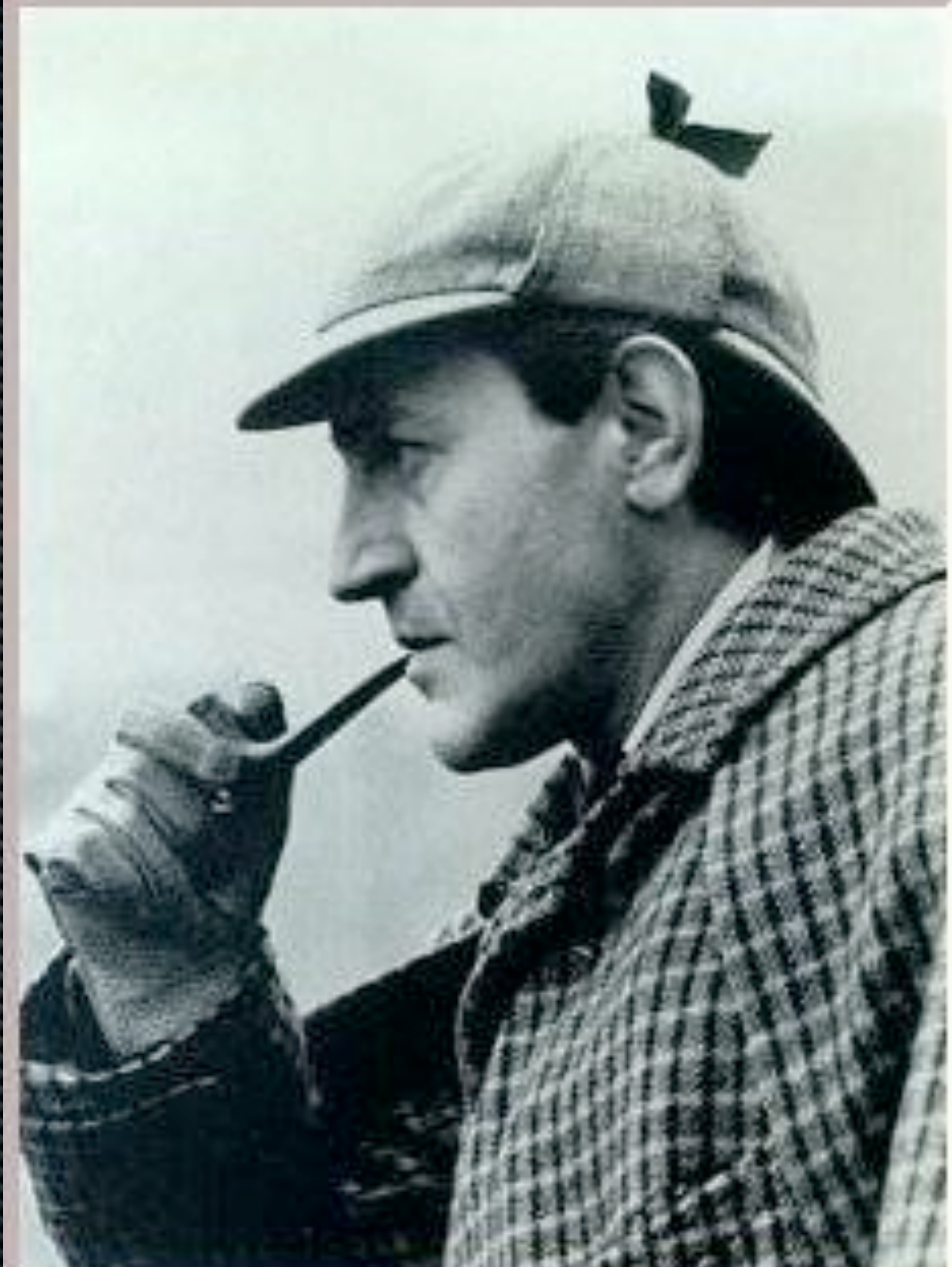
The Fee for each Amateur representation of this Play is Five Guineas, payable in advance to Samuel French, Limited, 26, Southampton Street, Strand, London.

Copyright 1912 by Samuel French, Ltd.

LONDON: SAMUEL FRENCH, LTD. PUBLISHERS 26, SOUTHAMPTON ST. STRAND

NEW YORK: SAMUEL FRENCH PUBLISHER 28, WEST 38th STREET

[ALL RIGHTS RESERVED]



Douglas Wilmer who played Sherlock Homes in the 1964 BBC1 series

Wikipedia

www.sherlockholmesonline.org/

"Good-morning, madam," said Holmes cheerily. "My name is Sherlock Holmes. This is my intimate friend and associate, Dr. Watson, before whom you can speak as freely as before myself.

...

Her features and figure were those of a woman of thirty, but her hair was shot with premature grey, and her expression was weary and haggard. Sherlock Holmes ran her over with one of his quick, all-comprehensive glances.

"You must not fear," said he soothingly, bending forward and patting her forearm. "We shall soon set matters right, I have no doubt. You have come in by train this morning, I see."

"You know me, then?"

http://www.image.ucar.edu/~nychka/Temp/TenLectures/Paleoclimate/speckled.txt

"No, but I observe the second half of a return ticket in the palm of your left glove. You must have started early, and yet you had a good drive in a dog-cart, along heavy roads, before you reached the station."

The lady gave a violent start and stared in bewilderment at my companion.

"There is **no mystery**, my dear madam," said he, smiling. "**The left arm of your jacket is spattered with mud in no less than seven places. The marks are perfectly fresh. There is no vehicle save a dog-cart which throws up mud in that way, and then only when you sit on the left-hand side of the driver**."

"Whatever your reasons may be, **you are perfectly correct**,"

http://www.image.ucar.edu/~nychka/Temp/TenLectures/Paleoclimate/speckled.txt

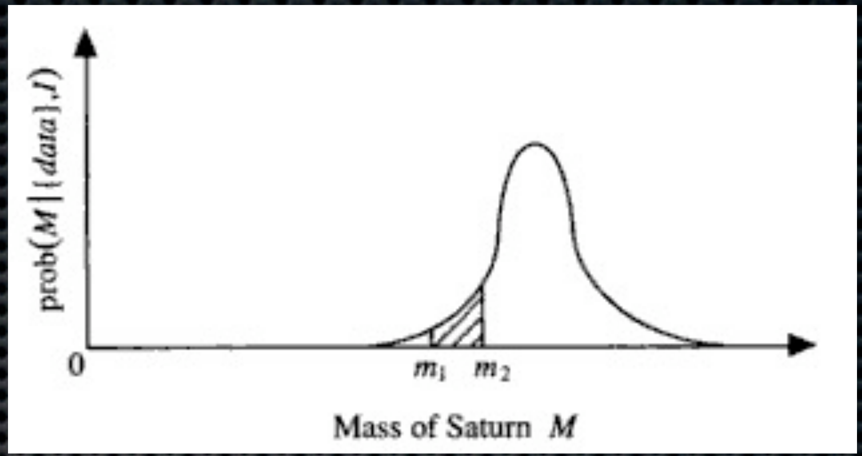
http://www.youtube.com/watch?v=kmlwZ49ibQA

History

- James Bernoulli (1713)
- Reverend Thomas Bayes - posthumous paper (1763)
- Laplace (1812)
- Laplace/Bayes: probability represented by a degree-of-belief or plausability - **how much they thought something was true, based on the evidence at hand**
- 19th century: probability redefined - relative frequency, given infinitely many repeated (experimental) trials

Sivia and Skilling (2006)

Laplace: mass of saturn



- Estimate the mass of Saturn given orbital data that were available from various astronomical observatories.
- Computed the posterior probability (PDF) for the mass M, given the data and all relevant background information I (such as knowledge of the laws of classical mechanics): prob(M|{data},I)
- Shaded area: how much Laplace believed that the mass of Saturn lay in the range between m1 and m2: '...it is a bet of 11,000 to 1 that the error of this result is not 1/100th of its value.'

Mass of Saturn

- Laplace solved this problem successfully since 150 years of data has changed his estimate by only 0.63%
- Frequentist approach - not allowed to use probability theory to solve the problem. The mass of Saturn is a constant and not a random variable; therefore, it has no frequency distribution and so probability theory cannot be used!
- **Why should we have to think about many repetitions of an experiment that never happened? What we really want to do is to make the best inference of the mass given the (few) data that we actually have; this is precisely the Bayes and Laplace view of probability!**

Bayes' theorem

Posterior probability

Likelihood function

Prior probability

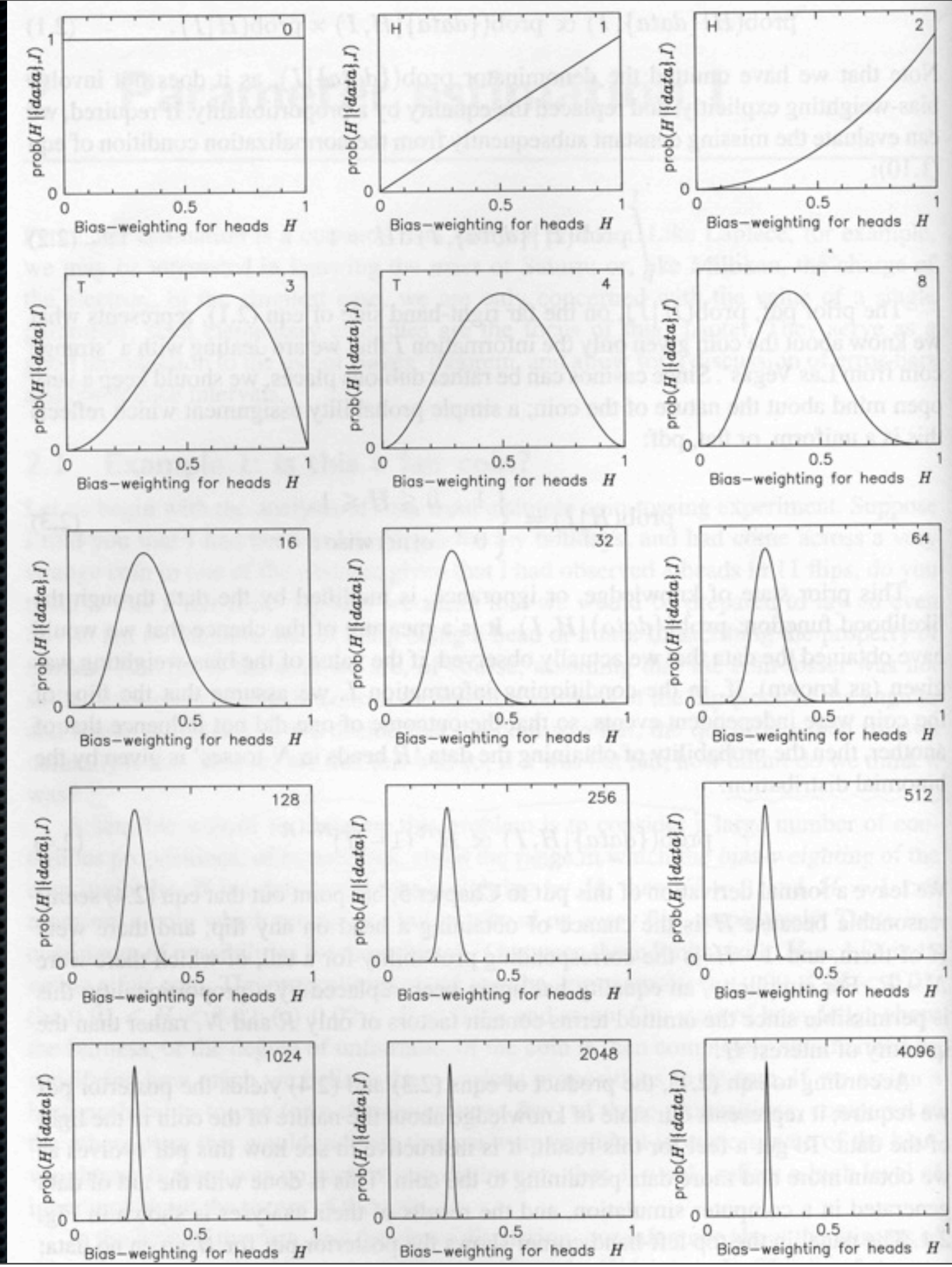
$$prob(X|Y,I) = \frac{prob(Y|X,I) \times prob(X|I)}{prob(Y|I)}$$

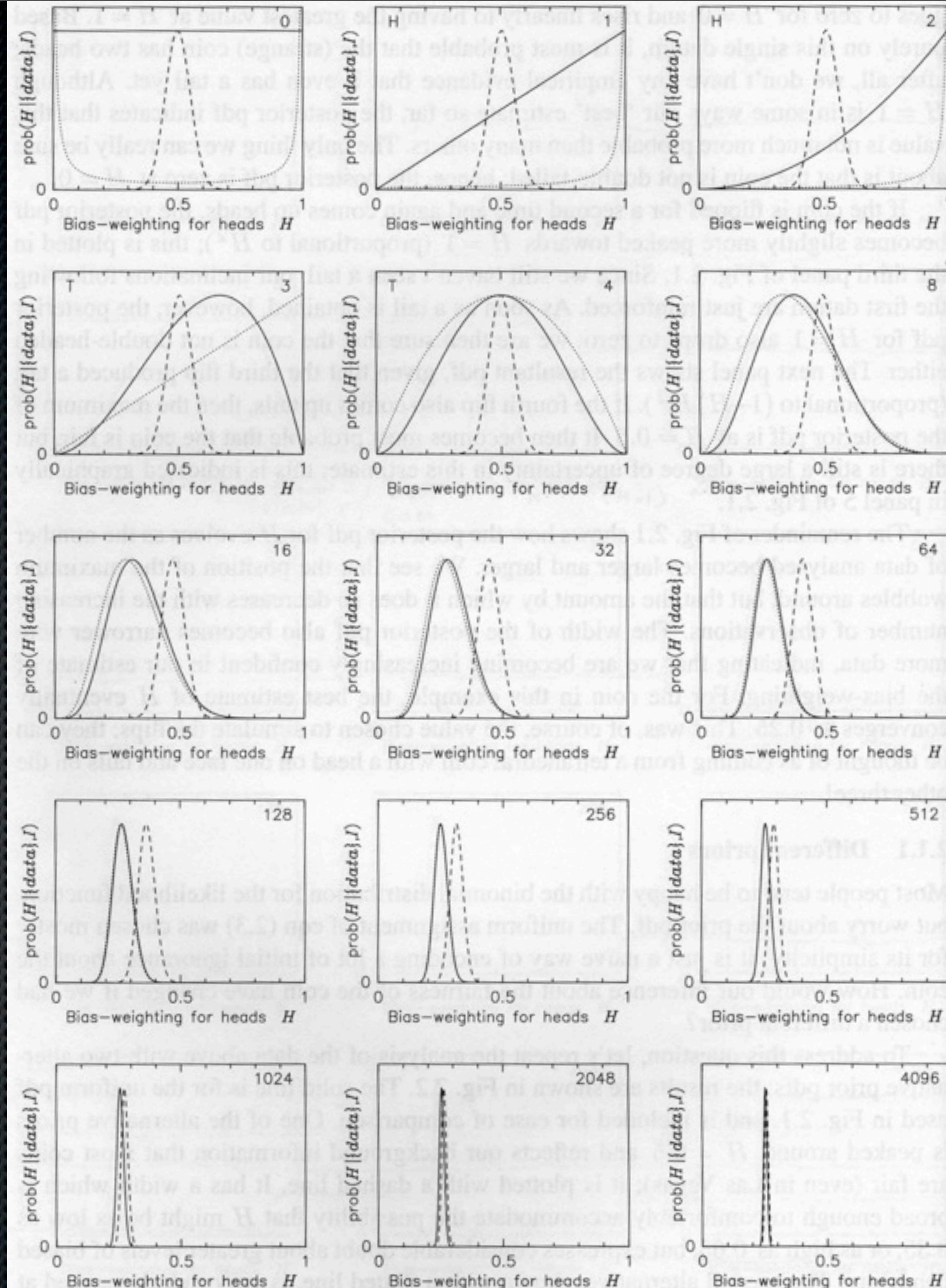
Evidence (marginal likelihood)

$$prob(hypothesis|data,I) \propto prob(data|hypothesis,I) \times prob(hypothesis|I)$$

Is this a fair coin?

- How can we measure that?
- Binomial distribution:
    - prob({data}|H,I) $\propto H^{R} (1-H)^{N-R}$
- Let's see some examples from Sivia & Skilling (2006)

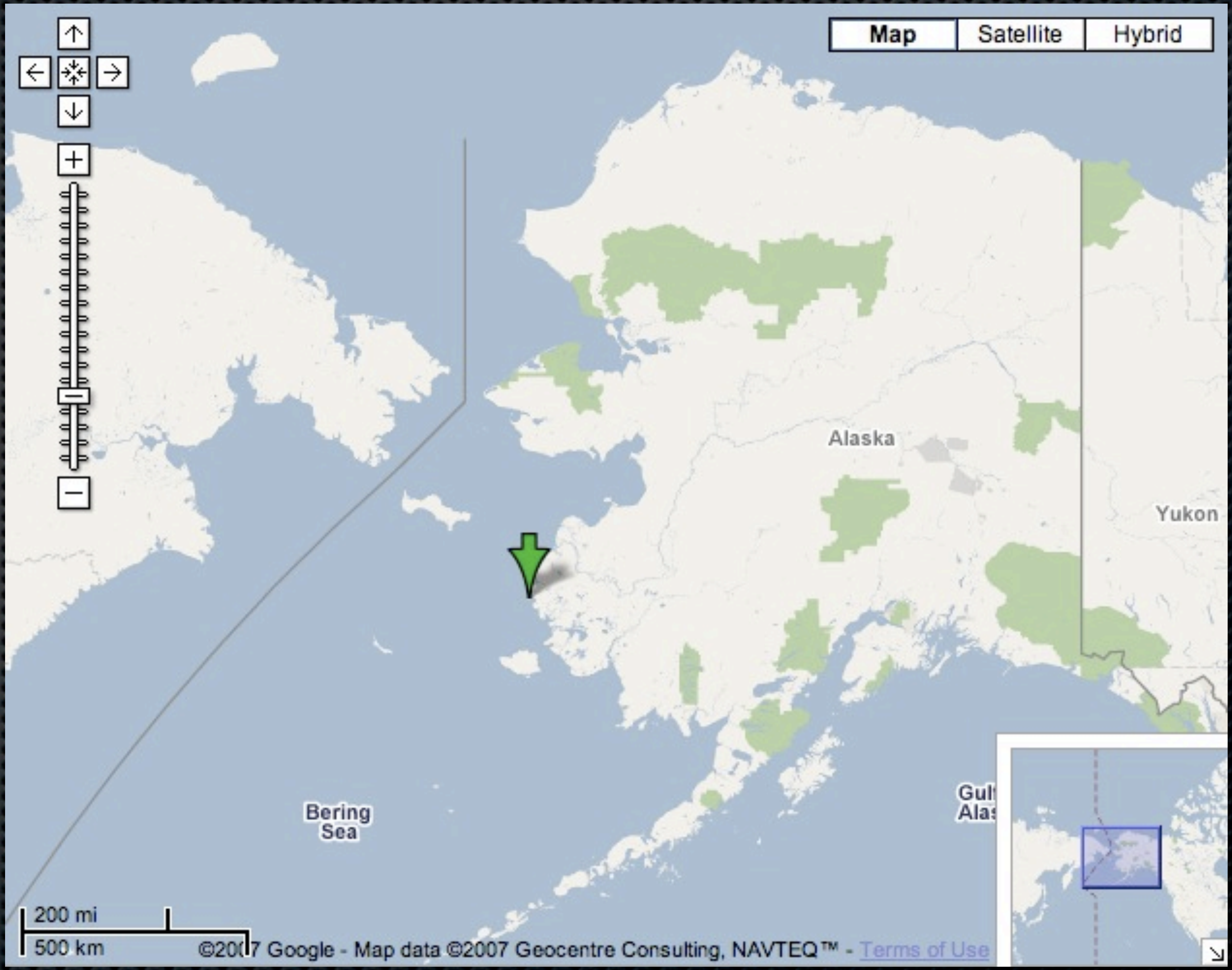

0
H
1
H
2
T
3
T
4
8
16
32
64
128
256
512
1024
2048
4096
prob(H|{data},I)
Bias−weighting for heads H
0
0.5
1


0
1
2
3
4
8
16
32
64
128
256
512
1024
2048
4096
prob(H|{data},I)
Bias-weighting for heads H
0
0.5
1

How can we use the Bayesian Approach in Climate Research?

Extreme wind analysis

- Hooper Bay, AK



Prior: spatial

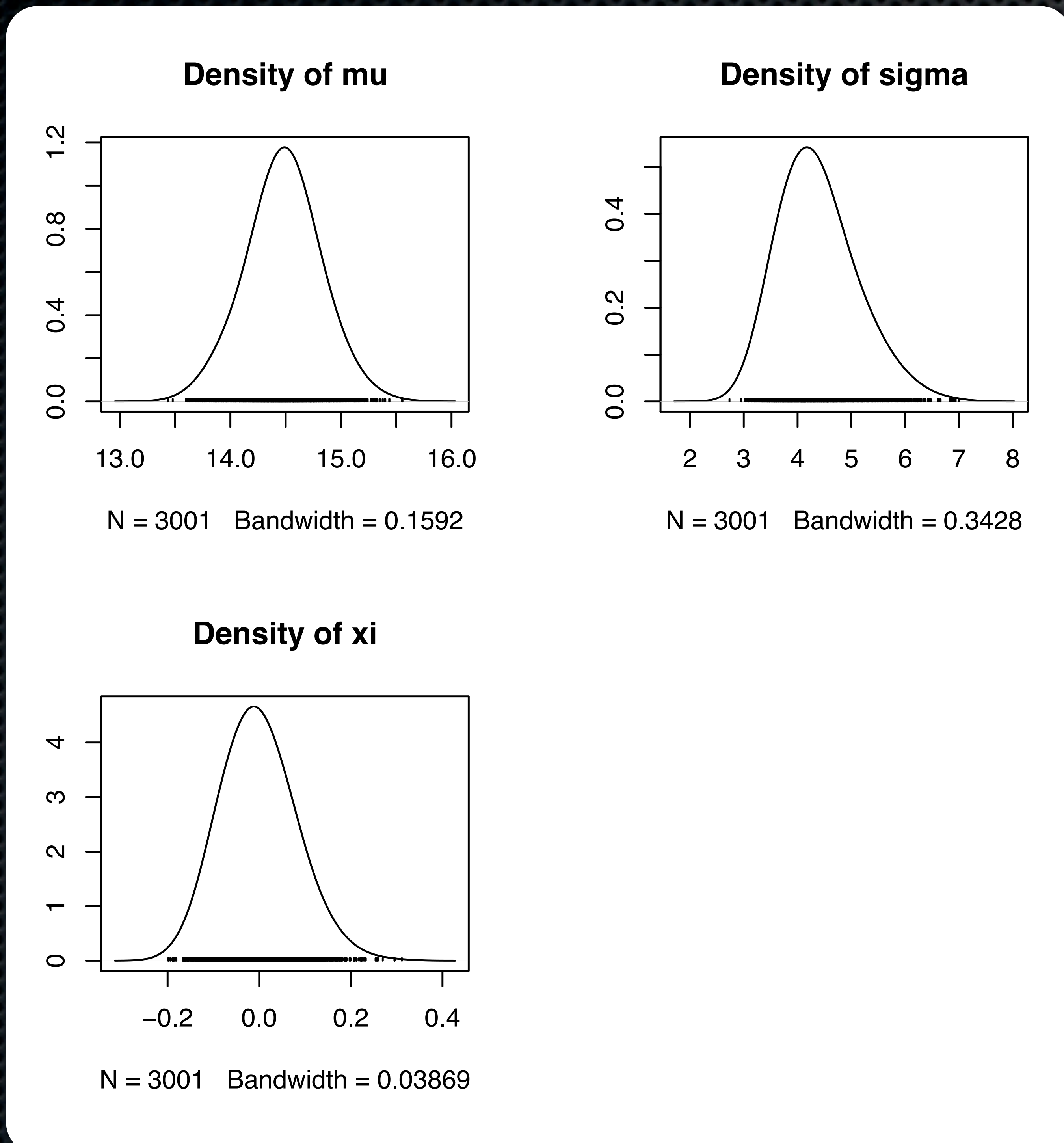
- Generalized Extreme Value Distribution (GEV)
- Trivariate Normal Distribution $\theta' = (\mu, \log \sigma, \xi)$
- Japanese Reanalysis

GEV

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

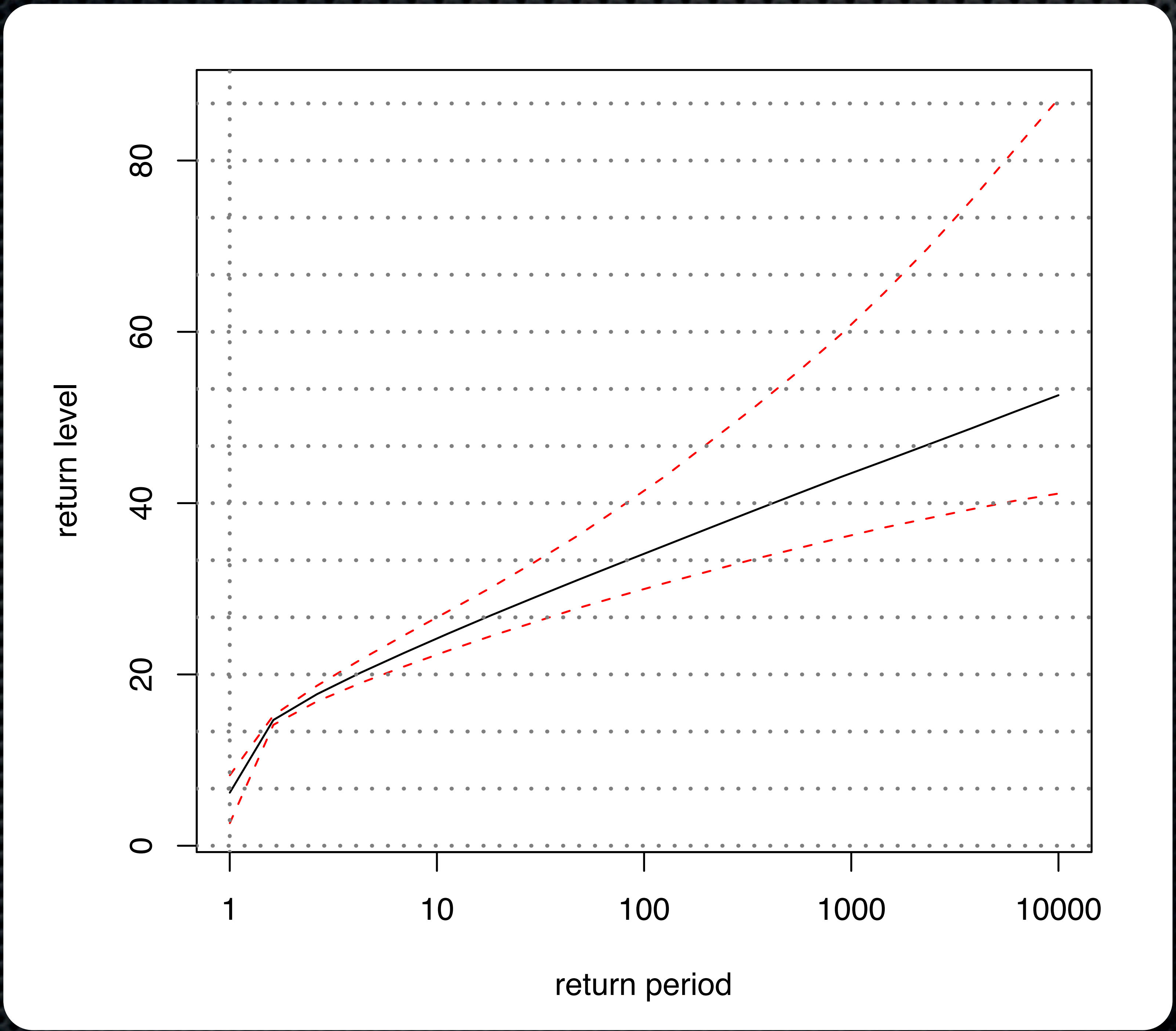
(μ, σ, ξ) location, scale and shape parameters

GEV Parameters for September




```
Iterations = 200:3200
Thinning interval = 1
Number of chains = 1
Sample size per chain = 3001

1. Empirical mean and standard deviation for each variable,
   plus standard error of the mean:

           Mean      SD Naive SE Time-series SE
mu    14.469784 0.33826 0.006175       0.030784
sigma  4.488668 0.76589 0.013981       0.071234
xi    -0.006331 0.07845 0.001432       0.006008

2. Quantiles for each variable:

         2.5%     25%       50%      75%   97.5%
mu    13.7960 14.2417 14.458002 14.69047 15.1314
sigma  3.3101  3.9668  4.380486  4.88127  6.3583
xi    -0.1611 -0.0606 -0.006913  0.04487  0.1480
```

Return Frequency



GEV in R

- **“evdbayes” package**
  - http://cran.r-project.org/web/packages/evdbayes/evdbayes.pdf
  - https://edit.ethz.ch/cces/projects/hazri/EXTREMES/talks/colesDavisonDavosJan08.pdf
- **Practical example (Dr. Fawcett, Newcastle University):**
  - http://www.mas.ncl.ac.uk/~nlf8/part6.pdf
  - “evd” package: http://cran.cnr.berkeley.edu/web/packages/evd/evd.pdf

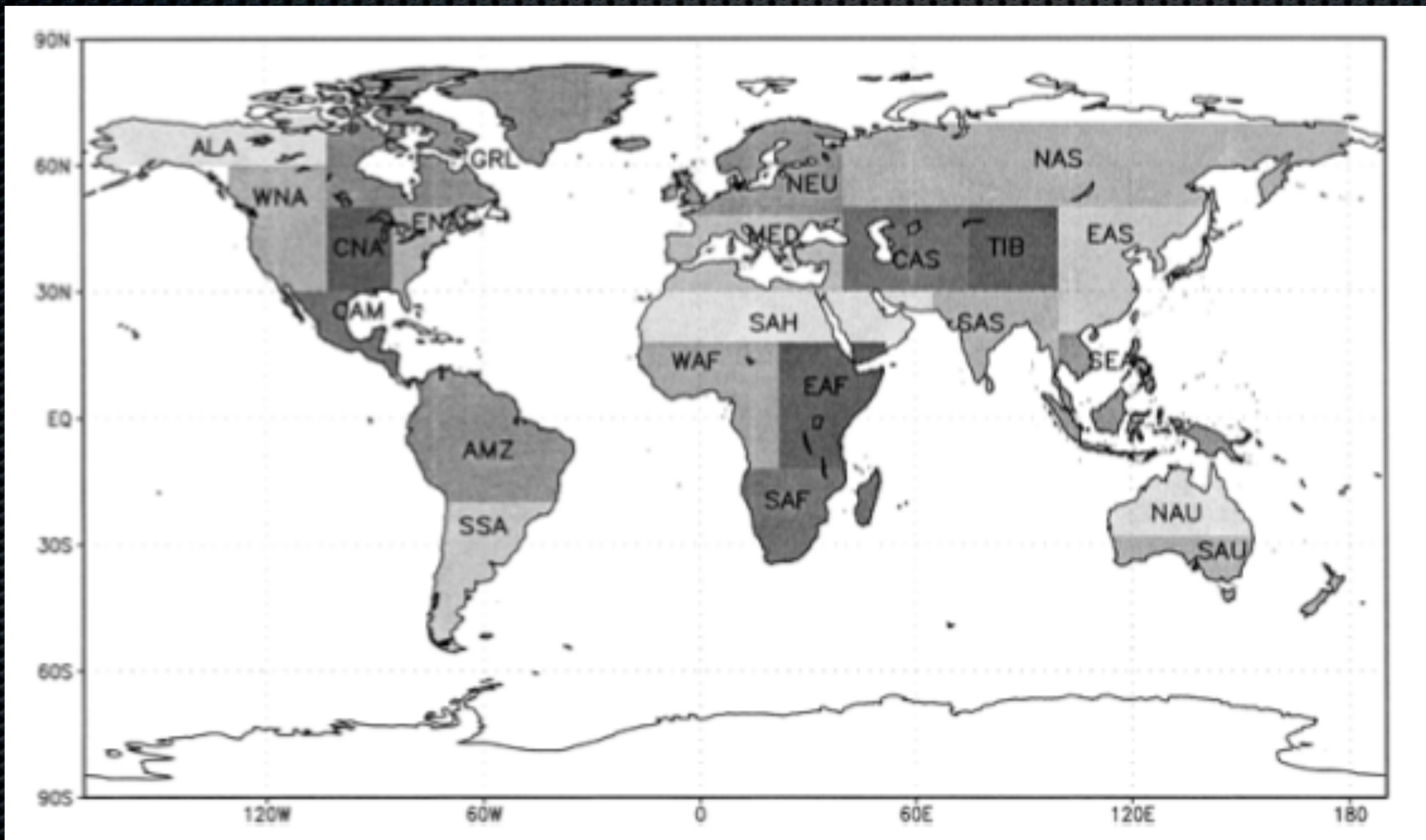
Climate model uncertainty

- The work by Claudia Tebaldi
  - http://www.image.ucar.edu/~tebaldi/papers.html
  - Tebaldi et al. (2005) Quantifying Uncertainty in Projections of Regional Climate Change: A Bayesian Approach to the Analysis of Multimodel Ensembles. *Journal of Climate*, 18, pp.1524-1540.
  - R package: “REA”

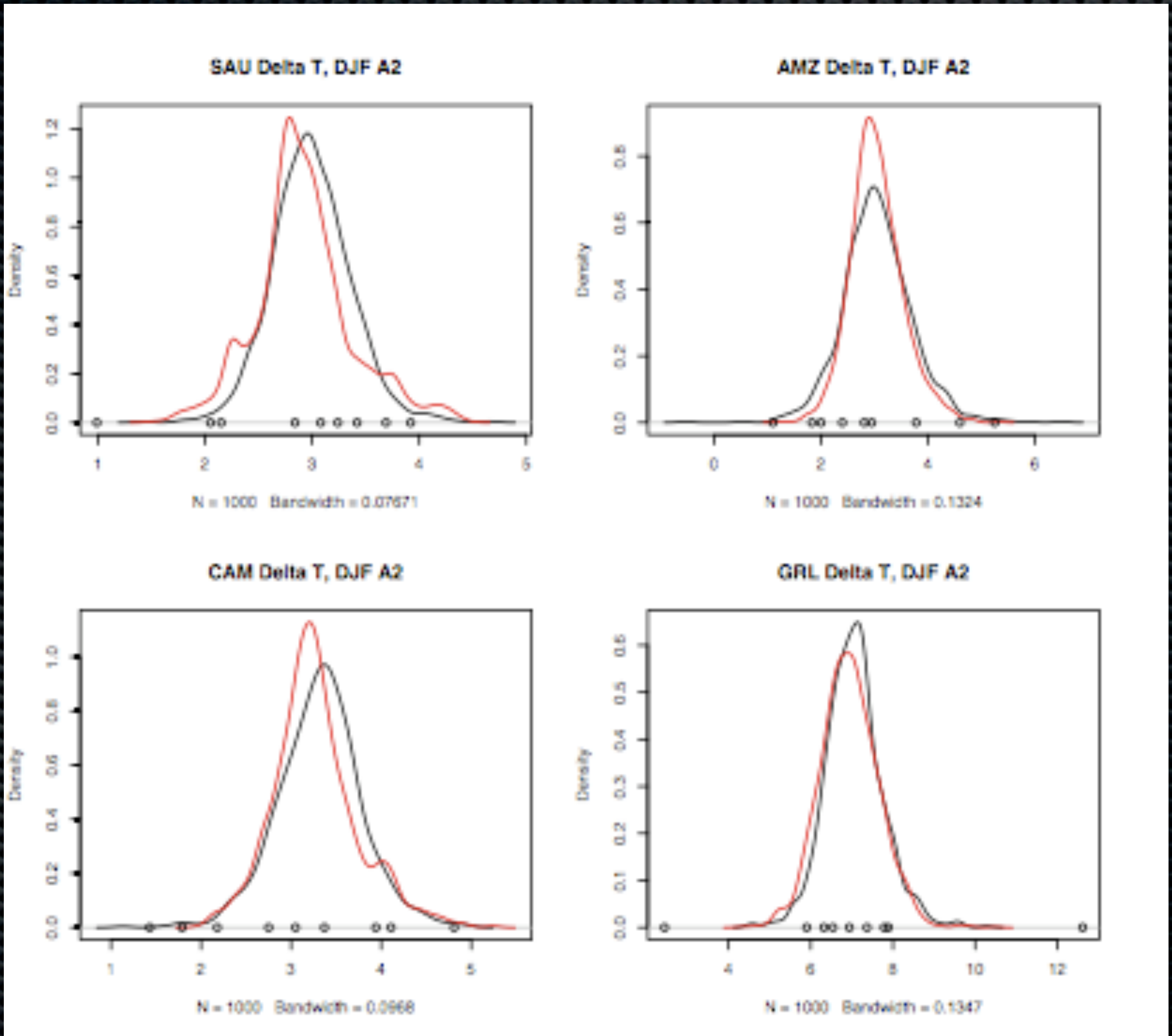
Framework

- GCMs - different future climate projections
  - How can we combine them into a probability distribution of future climate change?
  - Univariate and Multivariate Models - Smith et al (2007)

Smith et al (2007)



Temp. change (DJF - A2)



Sensitivity studies

- Mesquita et al. (2012)
  - Mesquita, M.d.S., Ådlandsvik, B., Bruyère, C. and Sandvik, A. (2012) Bayesian assessment of horizontal resolution in a nested-domain WRF simulation. 13th Annual WRF Users' Workshop, 25-29 June 2012, Boulder, CO, USA.
    - https://www.regonline.com/AttendeeDocuments/1077122/43357632/43357632_1045166.pdf
  - Mesquita, M.d.S. (2012) A Bayesian approach for evaluating regional climate models. Ten Lectures on Statistical Climatology, 6-10 August 2012, University of Washington, Seattle, USA.
  - Mesquita, M.d.S. (2012) An alternative approach for evaluating regional climate models using Bayesian Probability. ECRA: High Impact Events and Climate Change, Cambridge University, UK, on June 15.

Motivation

- Large number of parameterization scheme choices in sophisticated limited-area models
- The Weather Research and Forecasting (WRF) version 3.0:
  - mp_physics: 13 options
  - cu_physics: 6 options
  - (mp_physics, cu_physics) = C(13,1)*C(9,1) = **117 combinations!**
- There many other combinations to consider: radiation schemes, turbulence schemes, land schemes, boundary layer schemes, SST update, slab ocean, resolution and domain size
- The choice of domain size and position can also affect the results

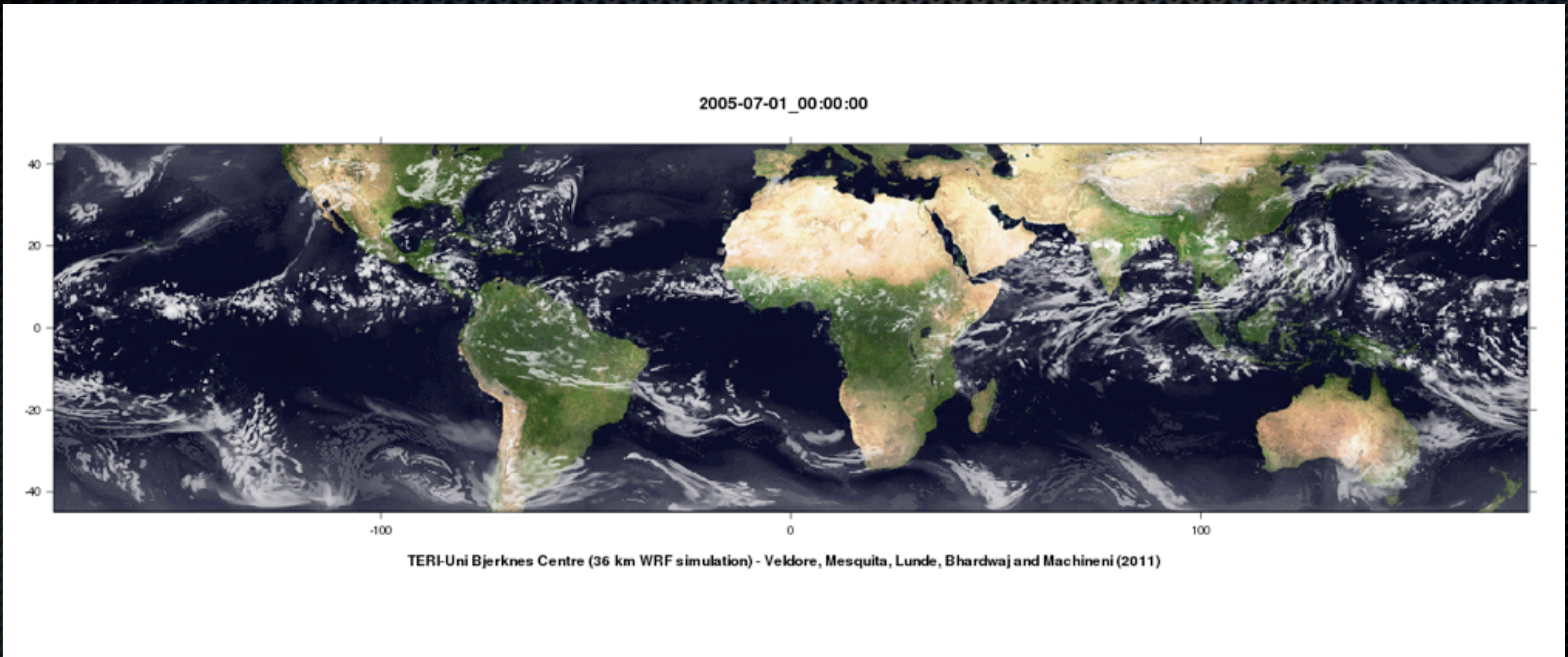
Challenges

- It is computationally expensive to run several combinations of schemes, resolutions and domain characteristics
- Even with a few combinations, one would need a large sample to make inferences and to obtain robust statistics
- It would be more appealing to use a method that could:
  - Make use of small samples for statistical inference
  - Provide a “richer” picture of the estimated

Data and setup

- Model: WRF tropical-channel domain
- Resolution: 36 and 50 km
- Combinations of 2 Cumulus and 2 Microphysical parameterization schemes
  - mp_physics = 3 (Hong-Dudhia-Chan)
  - mp_physics = 6 (Hong-Lin)
  - cu_physics = 1 (Kain-Fritsch)
  - cu_physics = 2 (Betts-Miller-Janjic)
- LBC data: ERA Interim
- Years: 2001-2002 (2001 spin-up)
- Focus: Brazilian northeast region (3-13S; 35-45W)

What is the tropical-channel domain?



Bayesian Inference

- Bayesian inference is the process of fitting a probability to a set of data and summarizing the result by a probability distribution on the parameters of the model and on unobserved quantities such as predictions for new observations
- Three-step approach to Bayesian data analysis:
  - 1. Setting up a full probability model
  - 2. Conditioning on data
  - 3. Evaluating the fit of the model

Gelman et al. (2004)

Joint inference for the mean and variance

- Recall the Bayes' rule:

$$posterior = \frac{Likelihood\_fuction \times prior}{evidence}$$

- For any joint distribution p(θ , σ^2), posterior inference proceeds using Bayes' rule:

$$p(\theta,\sigma^2 \mid y_1,\ldots,y_n) = \frac{p(y_1,\ldots,y_n \mid \theta,\sigma^2)p(\theta,\sigma^2)}{p(y_1,\ldots,y_n)}$$

Hoff (2009)

Posterior distribution

- The posterior distribution can be decomposed:

$$p(\theta,\sigma^2 \mid y_1,\ldots,y_n) = p(\theta \mid \sigma^2, y_1,\ldots,y_n)\,p(\sigma^2 \mid y_1,\ldots,y_n)$$

Hoff (2009)

$$p(\theta,\sigma^2 \mid y_1,\ldots,y_n) = p(\theta \mid \sigma^2, y_1,\ldots,y_n)\,p(\sigma^2 \mid y_1,\ldots,y_n)$$

$$\left\{\theta \mid y_1,\ldots,y_n,\sigma^2\right\} \sim normal(\mu_n, \sigma^2/\kappa_n)$$

$$\left\{1/\sigma^2 \mid y_1,\ldots,y_n\right\} \sim gamma(\upsilon_n/2, \upsilon_n\sigma_n^2/2)$$

Hoff (2009)

$$\left\{\theta \mid y_1,\ldots,y_n,\sigma^2\right\} \sim normal(\mu_n,\sigma^2/\kappa_n)$$

$$\mu_n = \frac{\kappa_o\mu_o + n\bar{y}}{\kappa_n}$$

$$\left\{1/\sigma^2 \mid y_1,\ldots,y_n\right\} \sim gamma(\upsilon_n/2,\upsilon_n\sigma_n^2/2)$$

$$\sigma_n^2 = \frac{1}{\upsilon_n}\left[\nu_o\sigma_o^2 + (n-1)s^2 + \frac{\kappa_o n}{\kappa_n}\left(\bar{y}-\mu_o\right)^2\right]$$

Hoff (2009)

Prior: μ_0 and σ_0^2

$$\mu_n = \frac{\kappa_o \mu_o + n\bar{y}}{\kappa_n}$$

$$\sigma_n^2 = \frac{1}{\upsilon_n}\left[\nu_o \sigma_o^2 + (n-1)s^2 + \frac{\kappa_o n}{\kappa_n}(\bar{y} - \mu_o)^2\right]$$

$$\kappa_n = \kappa_o + n$$

$$\upsilon_n = \upsilon_o + n$$

Hoff (2009)

- JFM (cptec.inpe.br)
  - Tmax = 28 C to 34 C
  - Tmin = 22 C to 24 C
  - Range → 22 C to 34 C

Prior parameters:

- μ_0 = 27 C
- σ_0 = 4.58 C (σ_0^2 = 21)
- $\kappa_0 = \nu_0 = 1$

Monte Carlo sampling

$$\sigma^{2(1)} \sim inverse\_gamma(\upsilon_n/2, \upsilon_n\sigma_n^2/2), \qquad \theta^{(1)} \sim normal(\mu_n, \sigma^{2(1)}/\kappa_n),$$

$$\vdots \qquad\qquad \vdots$$

$$\sigma^{2(S)} \sim inverse\_gamma(\upsilon_n/2, \upsilon_n\sigma_n^2/2), \qquad \theta^{(S)} \sim normal(\mu_n, \sigma^{2(1)}/\kappa_n),$$

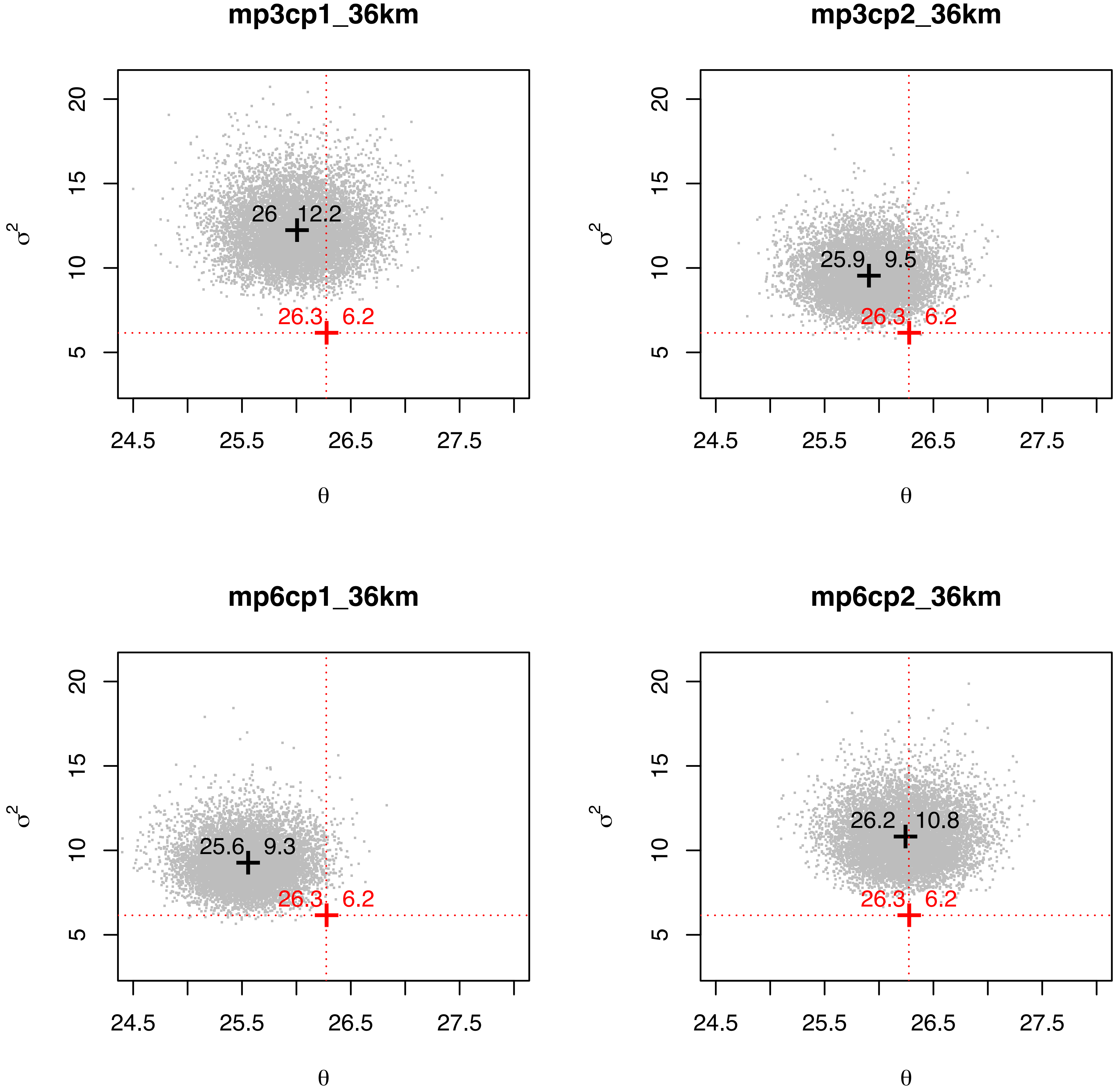
- The sequence of pairs $\{(\sigma^{2(1)}, \theta^{(1)}), \ldots, (\sigma^{2(S)}, \theta^{(S)})\}$ are independent samples from the joint posterior distribution of $p(\theta, \sigma^2 | y_1, \ldots, y_n)$

Hoff (2009)

	CP1	CP2
MP3	30, 50 km	30, 50 km
MP6	30, 50 km	30, 50 km

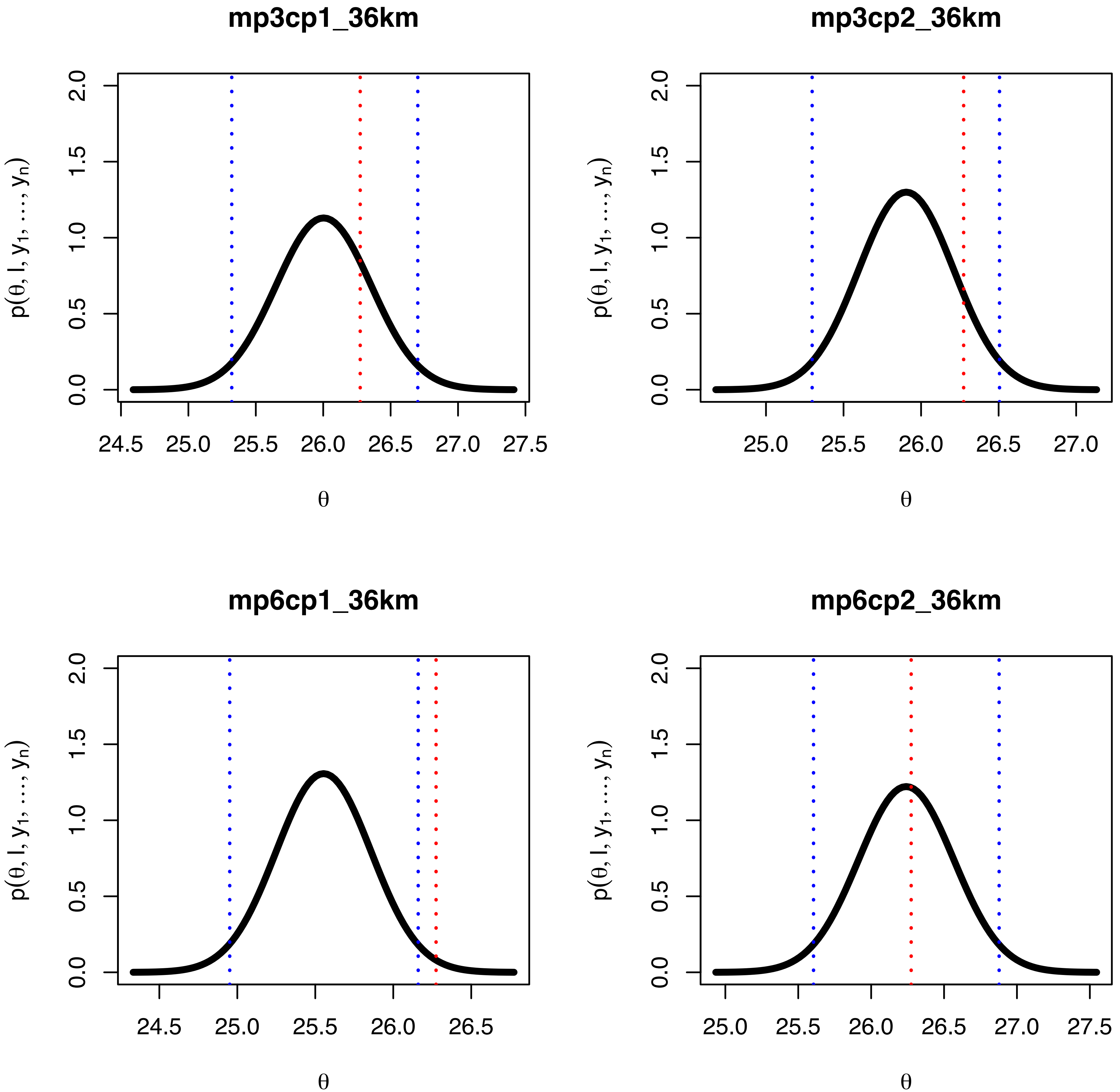
- 8 combinations (4 x 2 res.)
- Years: 2001-2002 (2001 spin-up)
- Season: January – February – March (JFM)
- Parameters to be estimated:
  - 2 m Temperature **Mean** and **Variance**
- Results will be shown with respect to samples drawn from the posterior distribution of the mean and variance using a Monte Carlo method
  - Number of samples: 10 000 per combination

Posterior distribution (36 km)



Mesquita (2012)

Mean (36 km)



Mesquita (2012)

Conclusion

- Bayesian probability has been used in climate research more and more, but:
  - Not many are familiar with the approach
  - The frequentist framework is still the rule
- It can be very useful, e.g.: extreme data analysis, climate uncertainty, sensitivity studies
- The Bayesian framework allows us to use smaller sample sizes, thus saving computational resources. It also provides a “full” view of the parameters – by looking at their probabilistic distributions

Thank you!

Email: michel.mesquita@uni.no

Contact: https://sites.google.com/site/mmeclimate

Educational site: www.m2lab.org