Likelihood-based Climate Model Evaluation

Amy Braverman[1] Noel Cressie[2,1]

[1] Jet Propulsion Laboratory,
California Institute of Technology

[2] University of Wollongong

August 15, 2013

Outline

- Introduction
- Motivating application/previous efforts
- Methodology
- Implementation
- Results
- Should we believe this? (Results of a simulation study)
- Conclusions

Introduction

- Climate models are deterministic, mathematical descriptions of the physics of climate.
- Confidence in predictions of future climate is increased if the physics are verifiably correct.
- A necessary (but not sufficient) condition is that past and present climate be simulated well.
- Quantify the likelihood that a (summary statistic computed from a) set of observations arises from a physical system with the characteristics expressed by a model-generated time series.

Introduction

If the atmosphere behaves as the model specifies, then we would expect the observations to look like the model output to within the inherent variability of the model output.

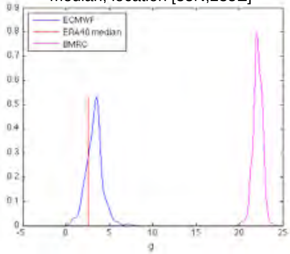
Observations: $\mathbf{Y}_0 = (Y_{01}, \ldots, Y_{0N_0})'$.

Output of model j: $\mathbf{Y}_j = (Y_{j1}, \ldots, Y_{jN_j})'$.

Statistic: $g(\cdot)$: $g(\mathbf{Y}_0) = g_0$, $g(\mathbf{Y}_j) = g_j$.

Estimate the sampling distribution of g_j by resampling.

Sampling distribution of the median, location [35N,235E]

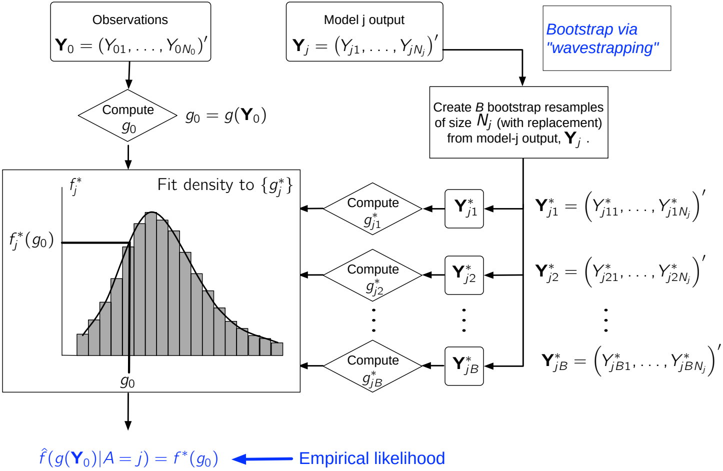


Likelihood of observing g_0 given model j sampling distribution is a figure of merit.

Introduction

- Let $A = j$ be the event that model j best represents the physical system.
- Let $g_0 = g(\mathbf{Y}_0)$ be a statistic computed from the time series of observations.
- Let $f(x|A = j)$ be the sampling distribution (density) of that statistic given $A = j$.
- $f(g_0|A = j)$ is the likelihood of g_0 given $A = j$.
- $P(A = j|g_0) \propto f(g_0|A = j)P(A = j)$.

Methodology



Observations

$\mathbf{Y}_0 = (Y_{01}, \ldots, Y_{0N_0})'$

Model j output

$\mathbf{Y}_j = (Y_{j1}, \ldots, Y_{jN_j})'$

Bootstrap via "wavestrapping"

Compute g_0

$g_0 = g(\mathbf{Y}_0)$

Create B bootstrap resamples of size N_j (with replacement) from model-j output, $\mathbf{Y}_j$.

Fit density to $\{g_j^*\}$

Compute g_{j1}^*

Compute g_{j2}^*

Compute g_{jB}^*

$\mathbf{Y}_{j1}^*$

$\mathbf{Y}_{j2}^*$

$\mathbf{Y}_{jB}^*$

$\mathbf{Y}_{j1}^* = \left(Y_{j11}^*, \ldots, Y_{j1N_j}^*\right)'$

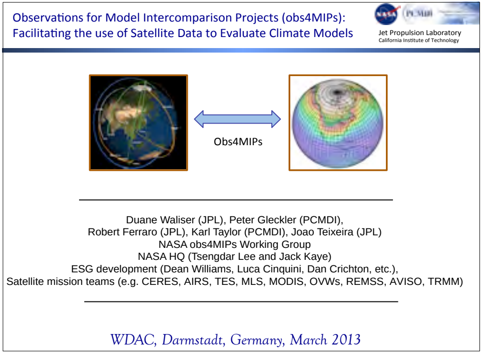
$\mathbf{Y}_{j2}^* = \left(Y_{j21}^*, \ldots, Y_{j2N_j}^*\right)'$

$\vdots$

$\mathbf{Y}_{jB}^* = \left(Y_{jB1}^*, \ldots, Y_{jBN_j}^*\right)'$

$\hat{f}(g(\mathbf{Y}_0)|A = j) = f^*(g_0)$ ← Empirical likelihood

Motivating application/previous efforts



Observations for Model Intercomparison Projects (obs4MIPs):
Facilitating the use of Satellite Data to Evaluate Climate Models

Jet Propulsion Laboratory
California Institute of Technology

Duane Waliser (JPL), Peter Gleckler (PCMDI),
Robert Ferraro (JPL), Karl Taylor (PCMDI), Joao Teixeira (JPL)
NASA obs4MIPs Working Group
NASA HQ (Tsengdar Lee and Jack Kaye)
ESG development (Dean Williams, Luca Cinquini, Dan Crichton, etc.),
Satellite mission teams (e.g. CERES, AIRS, TES, MLS, MODIS, OVWs, REMSS, AVISO, TRMM)

WDAC, Darmstadt, Germany, March 2013

- "Water vapor changes represent the largest feedback affecting climate sensitivity..." (IPCC 2007).
- Coupled Model Intercomparison Project (CMIP5) produces simulations from multiple models covering the decade of the 2000's ("decadal experiments").

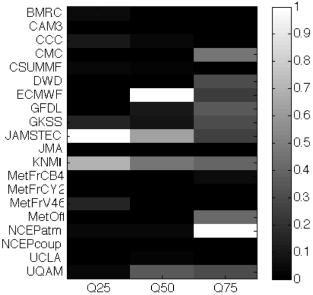
- We seek to evaluate the performance of various models' daily simulations of water vapor against satellite observations using an empirical likelihood approach.

Motivating application/previous efforts

- Earlier version of this work evaluated selected CMIP3 water-vapor time series against AIRS observations for two quartiles and the median of their marginal distributions.

- There, we used a moving-block bootstrap to simulate sampling distributions, but that was computationally very slow.

- For a more comprehensive application to CMIP5, we need a) statistics that are more sensitive to time series structure, and b) a faster way to generate sampling distributions.



Relative likelihood scores for 13 CMIP3 models and three statistics (Braverman, Cressie, and Teixeira, 2011).

Motivating application

For this exercise:

- Four model runs from Institute Pierre Simon Laplace (IPSL) daily simulations of water vapor at 500 hPa for January 1, 2001 through December 31, 2010. Time series of length 3650 days (10 years, 365 days/year) in each of 22 × 96 lat-lon grid cells (20S to 20N latitude).

- Six model runs from the Model for Interdisciplinary Research on Climate (MIROC5) daily simulations of water vapor at 500 hPa for January 1, 2001 through December 31, 2010. Time series of length 3652 (10 years, 365 days/year plus leap days) days in each of 30 × 256 lat-lon grid cells (20S to 20N latitude).

- Observations' daily averages of water vapor at 500 hPa from NASA's Atmospheric Infrared Sounder (AIRS) from October 1, 2002 through September 30, 2012 regridded twice to match the spatial resolution of the two classes of models. Time series of length ≈ 3652.

Methodology

- What g should we use?
- How to simulate the sampling distribution of g?
- How do we know our method works?

Methodology

What g should we use?

- We need a statistic (or set of statistics) that capture(s) important structure in a time series. The universal descriptor of time series structure is the spectrum.
- Let $\mathbf{Y}_j$ be a time series of length N_j, $\mathbf{Y}_j = \left(Y_{j1}, \ldots, Y_{jN_j}\right)'$. Assume N_j is even and let $\boldsymbol{\Phi}_j = \left(\phi_1, \ldots, \phi_{(N_j/2)-1}\right)$ be a matrix for which the columns are Fourier basis vectors, ϕ_k, $k = 1, \ldots, (N_j/2) - 1$.
- Let $\boldsymbol{\alpha}_j$ be the projection of $\mathbf{Y}_j$ onto the space spanned by $\boldsymbol{\Phi}_j$, $\boldsymbol{\alpha}_j = \boldsymbol{\Phi}_j{}'\mathbf{Y}_j$. The set of real and imaginary coefficients in $\boldsymbol{\alpha}_j = (\alpha_1, \ldots, \alpha_{(N_j/2)-1})$, $\alpha_k = (a_k + b_k i)$, provide a complete description of the structure of $\mathbf{Y}_j$ in terms of Fourier basis functions.

Methodology

- Drop the subscript *j* for the moment. There are too many α's to use them all, so we reduce $\boldsymbol{\alpha}$ to five summary statistics: $\bar{\omega}^*(p)$, $p = .05, .25, .50, .75, .95$, whose definition follows.

- Let $\hat{I}(\omega_k) = |\hat{\alpha}(\omega_k)|^2$. Then $\hat{I}(\omega_1), \hat{I}(\omega_2), \ldots, \hat{I}(\omega_{(N_j/2)-1})$, is the periodogram of $\mathbf{Y}_j$. Now let $\hat{I}^{(k)}(\omega_k^*)$ be the *k*th largest value of the periodogram elements, and let ω_k^* be the associated frequency. Find the largest value of *k* such that

$$\frac{\sum_{l=1}^{k} \hat{I}^{(k)}(\omega_k^*)}{\sum_{m=1}^{(N_j/2)-1} \hat{I}^{(m)}(\omega_m^*)} \leq p; \quad \text{call this value } k_p^*.$$

- Finally, let

$$\bar{\omega}^*(p) = \frac{1}{I_p^*} \sum_{m=1}^{k_p^*} \omega_m^* \left[\hat{I}^{(m)}(\omega_m^*)\right], \quad I_p^* = \sum_{m=1}^{k_p^*} \hat{I}^{(m)}(\omega_m^*).$$

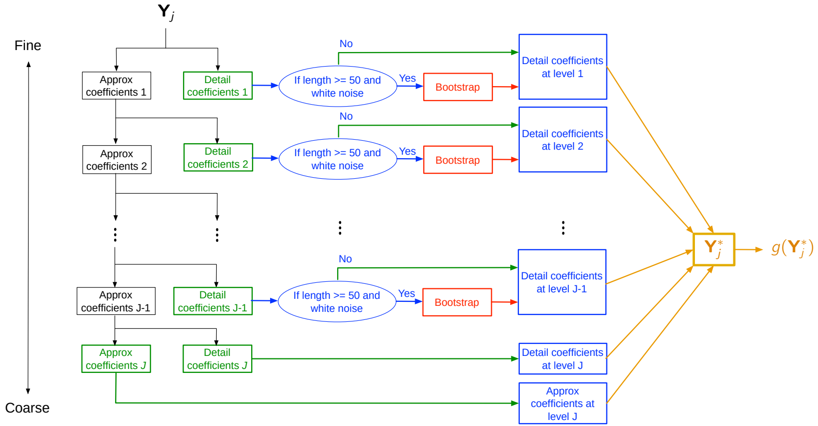
Methodology

How to simulate the sampling distribution of g?

- “Wavestrapping": originally introduced by Percival, Sardy, and Davison (2000). Our version is slightly different, but inspired by theirs.
- Idea:
  - Perform a J-level wavelet decomposition of the time series. This results in a set of approximation coefficients at level J, and detail coefficients at levels $1, \ldots, J$ (see next slide).
  - Test approximation coefficients and all levels’ detail coefficients for white noise. If white noise, bootstrap the coefficients. If not, don’t.
  - Reconstruct pseudo-series from coefficients.
  - Compute g from each pseudo-series and fit a kernel density estimate.

Methodology

Wavestrapping:



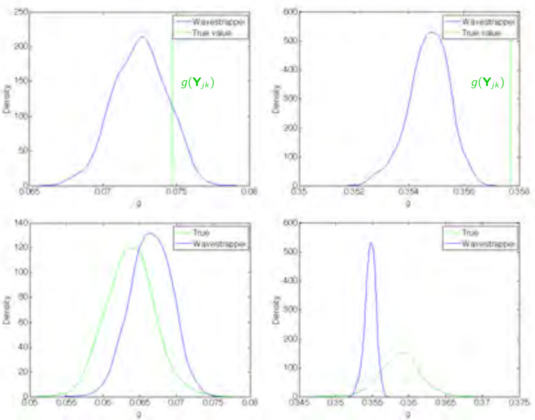
Methodology

Wavestrapping issues:

- Experiments suggest that wavestrapping doesn't work well if the data are skewed. First-differencing tends to correct this and improve stationarity.
- Even so, wavestrapping can yield pseudo-series that don't look like the original. This may lead to sampling distributions that aren't close to the true sampling distribution of g. (See next slide.)
- Diagnostic: trust results only when the time series' own value of g falls in the central 95 percent of the wavestrapped distribution. This is the "95 percent rule".

Methodology

The 95 percent rule:



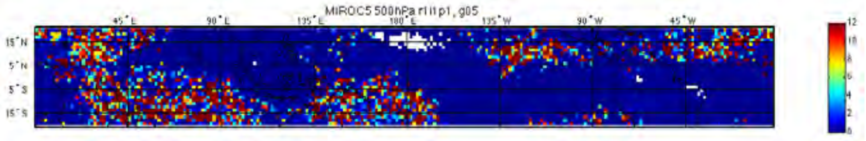
$\mathbf{Y}_{j1}, \ldots, \mathbf{Y}_{jK} = K$ independent realizations.

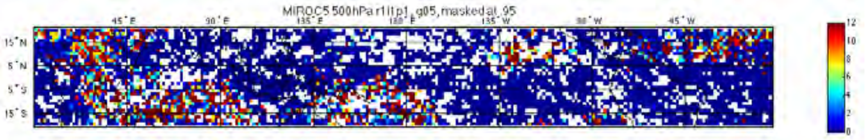
$\mathbf{Y}^*_{jk1}, \ldots, \mathbf{Y}^*_{jkB} = B$ wavestrapped realizations from $\mathbf{Y}_{jk}$.

$g(\mathbf{Y}^*_{jk1}), \ldots, g(\mathbf{Y}^*_{jkB}) \longrightarrow$ wavestrapped sampling distribution obtained from$\mathbf{Y}_{jk}$.

$g(\mathbf{Y}_{j1}), \ldots, g(\mathbf{Y}_{jK}) \longrightarrow$ sampling distribution.

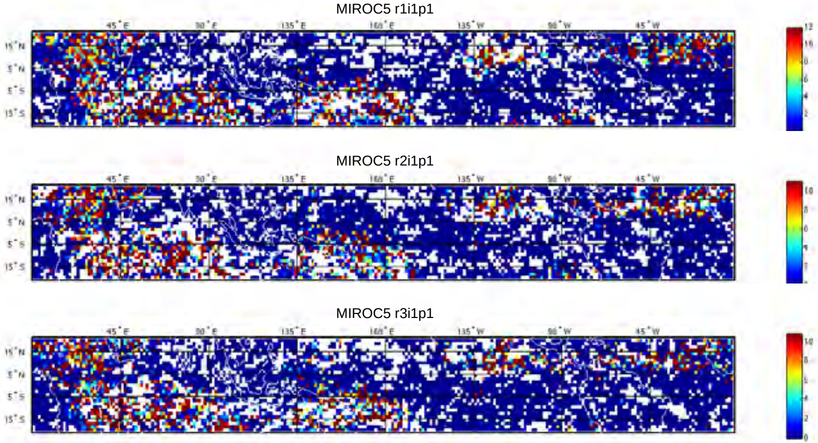
Results





- MIROC5 model, specific humidity at 500 hPa, model run 1 (r1i1p1) scored against AIRS using $g = \bar{\omega}(.05)$.
- White cells are “disqualified" for failing the 95 percent rule (bottom), or having no frequencies that account for 5 percent of the power (top).

Results


MIROC5 r1i1p1

MIROC5 r2i1p1

MIROC5 r3i1p1

Results

Remarks:

- We have only two models here, but for purposes of illustration we could treat different ensemble members' (e.g., r1i1p1, r2i1p1) output as if they were output of different models.
- These are maps of raw empirical likelihood scores. They are difficult to interpret as-is, and need to be converted to *relative* scores: for each grid cell, divide the raw score by the maximum score across all “models".
- However, MIROC5 and IPSL are at different resolutions, and there are many “missing" values. This calls for kriging to supply complete maps with common grid points.
- Stay tuned...

Conclusions (1)

- This method works very well for simple, synthetic time series (e.g., MA and low-order AR– those experiments not shown here) models.

- Initial overall indications are that this method works to varying degrees on complex (high-order AR) time series typical of climate model output, both synthetic and real.

- This may be due to our choice of statistics ($\bar{\omega}(\cdot)$), and issues with wavestrapping methodology.

- The raw likelihood scores show reassuring geographic consistency, but are difficult to interpret. Relative scores are needed. With multiple models at different spatial resolutions and multiple realizations of each, many questions remain about how to combine and compute relative scores.

- It is too early to conclude anything about the CMIP5 climate models themselves, based on this work.

Conclusions (2)

- There are two main ideas here: assessing consistency between output of a deterministic model via empirical likelihood, and wavestrapping as a mechanism for bootstrapping time series.

- The likelihood idea relies on a hypothesis testing framework, but randomness is induced by simulation from climate model output time series rather than sampling from the real world. Does this simulated uncertainty really correspond to the errors climate models make representing the real world?

- Is wavestrapping a good way to introduce this uncertainty? Is it better than alternatives (e.g, moving-block bootstrap, TFT-bootstrap (Kirch and Politis, 2011))?

- How is this going to help the IPCC/climate modeling community?

References

Braverman, A., Cressie, N., and Teixeira, J. (2011). A Likelihood-based comparison of temporal models for physical processes, *Statistical Analysis and Data Mining*, Volume 4, Number 3, pp. 247-258, doi: 10.1002/sam.10113.

Kirch, C., and Politis, D. (2011). TFT-bootstrap: Resampling time series in the frequency domain to obtain replicates in the time domain, *Annals of Statistics*, Volume 39, Number 3, pp. 1427-1470, doi:10.1214/10-AOS868SUPP.

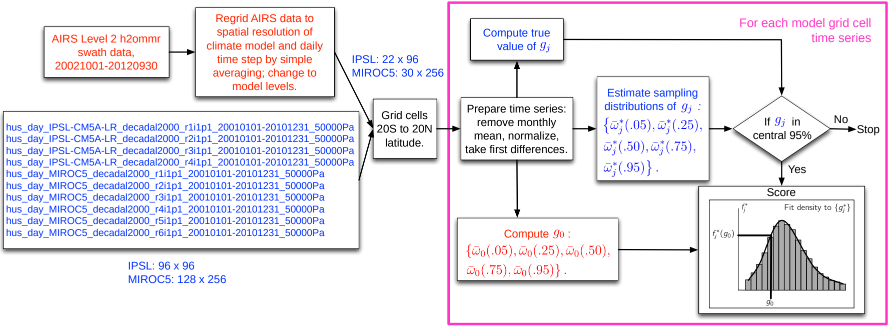
Percival, D. B., Sardy, S., and Davison, A. C. (2000). Wavestrapping time series: Adaptive wavelet-based bootstrapping in *Nonlinear and Nonstationary Signal Processing*, Fitzgerald, W., Smith, R., Walden, A., and Young, P., editors. Cambridge University Press, pp. 442-471.

Questions/comments? Reach me at `Amy.Braverman@jpl.nasa.gov`.

This research is supported by NASA's Earth Science Data Records Uncertainty Analysis (ESDRERR) program.

Backup slides

Implementation



AIRS Level 2 h2ommr swath data, 20021001-20120930

Regrid AIRS data to spatial resolution of climate model and daily time step by simple averaging; change to model levels.

IPSL: 22 x 96
MIROC5: 30 x 256

hus_day_IPSL-CM5A-LR_decadal2000_r1i1p1_20010101-20101231_50000Pa
hus_day_IPSL-CM5A-LR_decadal2000_r2i1p1_20010101-20101231_50000Pa
hus_day_IPSL-CM5A-LR_decadal2000_r3i1p1_20010101-20101231_50000Pa
hus_day_IPSL-CM5A-LR_decadal2000_r4i1p1_20010101-20101231_50000Pa
hus_day_MIROC5_decadal2000_r1i1p1_20010101-20101231_50000Pa
hus_day_MIROC5_decadal2000_r2i1p1_20010101-20101231_50000Pa
hus_day_MIROC5_decadal2000_r3i1p1_20010101-20101231_50000Pa
hus_day_MIROC5_decadal2000_r4i1p1_20010101-20101231_50000Pa
hus_day_MIROC5_decadal2000_r5i1p1_20010101-20101231_50000Pa
hus_day_MIROC5_decadal2000_r6i1p1_20010101-20101231_50000Pa

IPSL: 96 x 96
MIROC5: 128 x 256

Grid cells 20S to 20N latitude.

For each model grid cell time series

Compute true value of g_j

Prepare time series: remove monthly mean, normalize, take first differences.

Estimate sampling distributions of g_j : $\{\bar{\omega}_j^*(.05), \bar{\omega}_j^*(.25), \bar{\omega}_j^*(.50), \bar{\omega}_j^*(.75), \bar{\omega}_j^*(.95)\}$.

If g_j in central 95%

No

Stop

Yes

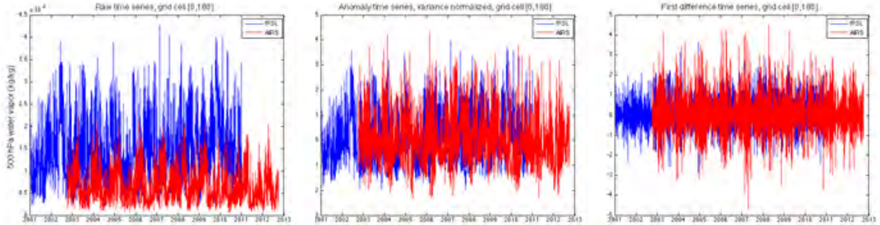
Score

Compute g_0 : $\{\bar{\omega}_0(.05), \bar{\omega}_0(.25), \bar{\omega}_0(.50), \bar{\omega}_0(.75), \bar{\omega}_0(.95)\}$.

- Yields five, 22 × 96 spatial maps for IPSL (one for each *g*) for each of four IPSL input data sets.
- Yields five, 30 × 256 spatial maps for MIROC5 (one for each *g*) for each of six MIROC5 input data sets.

Implementation

Details 1: data preparation



X_{jd} = source j water vapor on day d,
$d = 1, 2, \dots, N_j$, $j = 0, 1, \dots, J$,
where J is the number of sources.

$j = 0 \leftrightarrow$ AIRS, $j = 1, \dots, 10 \leftrightarrow$ ensemble members from two models: IPSL (4), and MIROC5 (6).

$$Y_{jd} = \frac{X_{jd} - \bar{X}_{jm}}{\sqrt{\text{var}(\mathbf{X}_j)}}, \quad \bar{X}_{jm} = \frac{1}{N_{jm}} \sum_{d \in \mathcal{D}_{jm}} X_{jd},$$

$\mathcal{D}_{jm}$ = the set of days in source j month m,
over all years, $m = 1, 2, \dots, 12$,

$N_{jm} = |\mathcal{D}_{jm}|$.

$$W_{jd} = Y_{jd+1} - Y_{jd}.$$

Time series shown here are IPSL, realization 1 for grid cell [0,180].

- We do not wish to test whether climate model output mean and variance are consistent with AIRS, so standardize all time series.

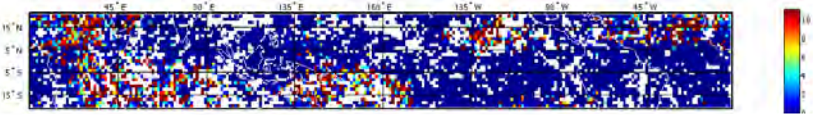
Implementation

Details 2: wavestrapping

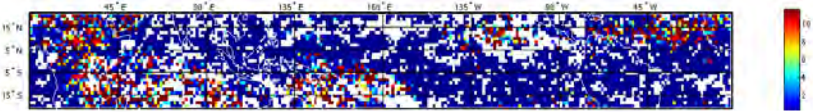
- Wavelet decomposition performed on each first-differenced time series, $\mathbf{W}_j$, using $J = 8$ levels and the Haar wavelet basis in Matlab.
- Lejung-Box white-noise test performed on coefficients (whenever there are at least 50 coefficients) at 0.05 level.
- Wavestrapped sampling distribution simulated with $B = 500$ trials. Kernel density estimate fit to 500 g_j^* values, and evaluated at g_j and g_0 using Matlab's `interp1` function.
- Diagnostic test for whether to trust result requires that g_j fall within the central 95 percent of the wavestrapped sampling distribution of g derived from $\mathbf{W}_j$.
- Timing: wavestrapping to generate a sampling distribution with $J = 8$ and $B = 500$ takes about 8 seconds on my MacBook Air.

Results

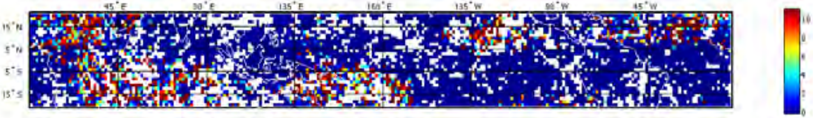
MIROC5 r4i1p1



MIROC5 r5i1p1

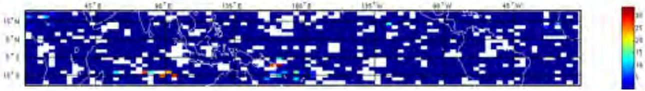


MIROC5 r6i1p1

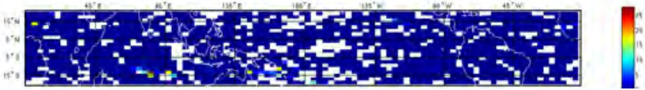


Results

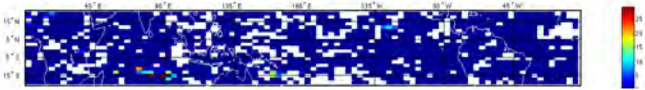
IPSL r1i1p1



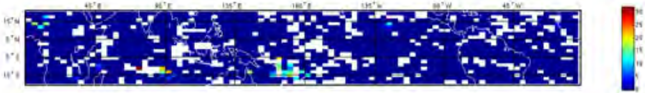
IPSL r2i1p1



IPSL r3i1p1



IPSL r4i1p1

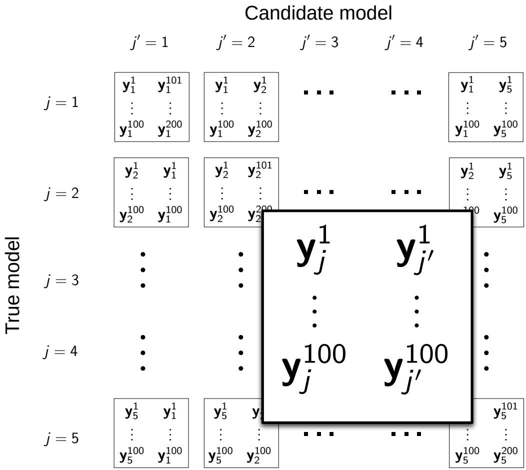


A simulation study

How do we know our method works? *Perform a simulation study.*

- Choose five AR(p) models that: 1) are representative of climate model output time series, and 2) are distinguishable from one another.
  - Fit AR(p) models with $p = 1, 11, 21, 31, \ldots, 201$ to all 30×256 grid cell time series for MIROC5 500hPa r1i1p1 output. Choose the best fitting AR(p) using AIC.
  - Choose five grid cells that are spatially diverse and for which the best fitting AR models should be distinguishable.
  - How do we know that five AR models associated with the five grid cells should be distinguishable? Simulate $K = 100$ time series from each of the five AR models, and compute $\bar{\omega}_j^k(q)$, $q = \{.05, .25, .50, .75, .95\}$, $j = 1, 2, 3, 4, 5$ and $k = 1, \ldots, K$.
  - Examine boxplots of the $\bar{\omega}_j^k(q)$ to determine which values of q lead to $\bar{\omega}_j^k(q)$ values that ought to distinguish among the distributions.
- Now, let each AR model, $j = 1, 2, 3, 4, 5$, successively play the role of the "true" model, and score all the models against the true model.

A simulation study



Candidate model

$j' = 1$ $j' = 2$ $j' = 3$ $j' = 4$ $j' = 5$

True model

$j = 1$ $j = 2$ $j = 3$ $j = 4$ $j = 5$

$j = 1$, $j' = 1$: $\mathbf{y}_1^1 \; \mathbf{y}_1^{101} \; \vdots \; \vdots \; \mathbf{y}_1^{100} \; \mathbf{y}_1^{200}$

$j = 1$, $j' = 2$: $\mathbf{y}_1^1 \; \mathbf{y}_2^1 \; \vdots \; \vdots \; \mathbf{y}_1^{100} \; \mathbf{y}_2^{100}$

$j = 1$, $j' = 5$: $\mathbf{y}_1^1 \; \mathbf{y}_5^1 \; \vdots \; \vdots \; \mathbf{y}_1^{100} \; \mathbf{y}_5^{100}$

$j = 2$, $j' = 1$: $\mathbf{y}_2^1 \; \mathbf{y}_1^1 \; \vdots \; \vdots \; \mathbf{y}_2^{100} \; \mathbf{y}_1^{100}$

$j = 2$, $j' = 2$: $\mathbf{y}_2^1 \; \mathbf{y}_2^{101} \; \vdots \; \vdots \; \mathbf{y}_2^{100} \; \mathbf{y}_2^{200}$

$j = 2$, $j' = 5$: $\mathbf{y}_2^1 \; \mathbf{y}_5^1 \; \vdots \; \vdots \; \mathbf{y}_2^{100} \; \mathbf{y}_5^{100}$

$j = 5$, $j' = 1$: $\mathbf{y}_5^1 \; \mathbf{y}_1^1 \; \vdots \; \vdots \; \mathbf{y}_5^{100} \; \mathbf{y}_1^{100}$

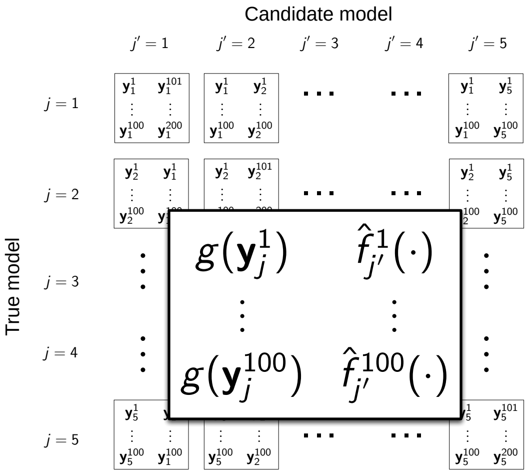
$j = 5$, $j' = 2$: $\mathbf{y}_5^1 \; \mathbf{y}_2^1 \; \vdots \; \vdots \; \mathbf{y}_5^{100} \; \mathbf{y}_2^{100}$

$j = 5$, $j' = 5$: $\mathbf{y}_5^1 \; \mathbf{y}_5^{101} \; \vdots \; \vdots \; \mathbf{y}_5^{100} \; \mathbf{y}_5^{200}$

$$\begin{array}{cc} \mathbf{y}_j^1 & \mathbf{y}_{j'}^1 \\ \vdots & \vdots \\ \mathbf{y}_j^{100} & \mathbf{y}_{j'}^{100} \end{array}$$

y_j^v is the vth time series generated by model j.

A simulation study



Candidate model

$j' = 1$ $j' = 2$ $j' = 3$ $j' = 4$ $j' = 5$

True model

$j = 1$: $\mathbf{y}_1^1 \; \mathbf{y}_1^{101} \; \vdots \; \mathbf{y}_1^{100} \; \mathbf{y}_1^{200}$ | $\mathbf{y}_1^1 \; \mathbf{y}_2^1 \; \vdots \; \mathbf{y}_1^{100} \; \mathbf{y}_2^{100}$ | $\cdots$ | $\cdots$ | $\mathbf{y}_1^1 \; \mathbf{y}_5^1 \; \vdots \; \mathbf{y}_1^{100} \; \mathbf{y}_5^{100}$

$j = 2$: $\mathbf{y}_2^1 \; \mathbf{y}_1^1 \; \vdots \; \mathbf{y}_2^{100}$ | $\mathbf{y}_2^1 \; \mathbf{y}_2^{101} \; \vdots$ | $\cdots$ | $\cdots$ | $\mathbf{y}_2^1 \; \mathbf{y}_5^1 \; \vdots \; \mathbf{y}_5^{100}$

$j = 3$: $\vdots$ | $\vdots$

$j = 4$: $\vdots$ | $\vdots$

$j = 5$: $\mathbf{y}_5^1 \; \vdots \; \mathbf{y}_5^{100} \; \mathbf{y}_1^{100}$ | $\vdots \; \vdots \; \mathbf{y}_5^{100} \; \mathbf{y}_2^{100}$ | $\cdots$ | $\cdots$ | $\mathbf{y}_5^{101} \; \vdots \; \vdots \; \mathbf{y}_5^{100} \; \mathbf{y}_5^{200}$

$$g(\mathbf{y}_j^1) \qquad \hat{f}_{j'}^1(\cdot)$$

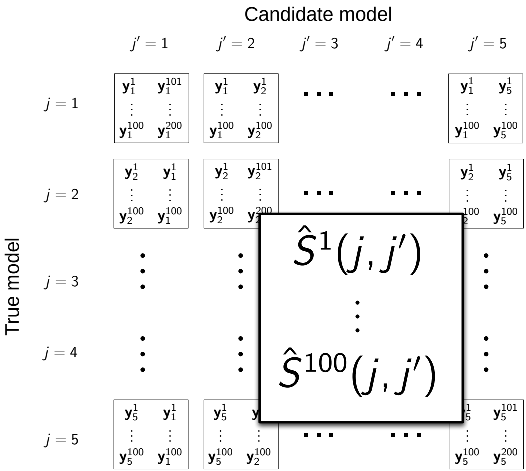
$$\vdots \qquad \vdots$$

$$g(\mathbf{y}_j^{100}) \qquad \hat{f}_{j'}^{100}(\cdot)$$

Evaluate:

$$\hat{S}^v(j,j') \equiv \hat{f}_{j'}^v(g(y_j^v)),$$
$$v = 1,\ldots,500,$$
$$j,j' = 1,\ldots,5.$$

A simulation study



Candidate model

True model

	$j' = 1$	$j' = 2$	$j' = 3$	$j' = 4$	$j' = 5$
$j = 1$	$\mathbf{y}_1^1$ $\mathbf{y}_1^{101}$ $\vdots$ $\vdots$ $\mathbf{y}_1^{100}$ $\mathbf{y}_1^{200}$	$\mathbf{y}_1^1$ $\mathbf{y}_2^1$ $\vdots$ $\vdots$ $\mathbf{y}_1^{100}$ $\mathbf{y}_2^{100}$	$\cdots$	$\cdots$	$\mathbf{y}_1^1$ $\mathbf{y}_5^1$ $\vdots$ $\vdots$ $\mathbf{y}_1^{100}$ $\mathbf{y}_5^{100}$
$j = 2$	$\mathbf{y}_2^1$ $\mathbf{y}_1^1$ $\vdots$ $\vdots$ $\mathbf{y}_2^{100}$ $\mathbf{y}_1^{100}$	$\mathbf{y}_2^1$ $\mathbf{y}_2^{101}$ $\vdots$ $\vdots$ $\mathbf{y}_2^{100}$ $\mathbf{y}_2^{200}$	$\cdots$	$\cdots$	$\mathbf{y}_2^1$ $\mathbf{y}_5^1$ $\vdots$ $\vdots$ $\mathbf{y}_2^{100}$ $\mathbf{y}_5^{100}$
$j = 3$	$\vdots$	$\vdots$			$\vdots$
$j = 4$	$\vdots$	$\vdots$			$\vdots$
$j = 5$	$\mathbf{y}_5^1$ $\mathbf{y}_1^1$ $\vdots$ $\vdots$ $\mathbf{y}_5^{100}$ $\mathbf{y}_1^{100}$	$\mathbf{y}_5^1$ $\mathbf{y}$ $\vdots$ $\vdots$ $\mathbf{y}_5^{100}$ $\mathbf{y}_2^{100}$	$\cdots$	$\cdots$	$\mathbf{y}_5^1$ $\mathbf{y}_5^{101}$ $\vdots$ $\vdots$ $\mathbf{y}_5^{100}$ $\mathbf{y}_5^{200}$

$$\hat{S}^1(j,j')$$

$$\vdots$$

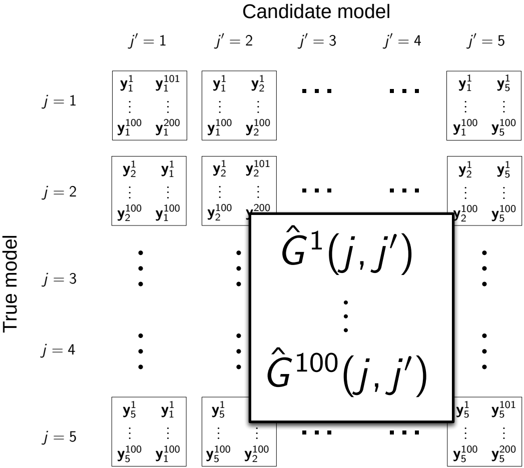
$$\hat{S}^{100}(j,j')$$

Evaluate:

$$\hat{S}^v(j,j') \equiv \hat{f}^v_{j'}(g(y^v_j)),$$

$$v = 1, \ldots, 500,$$
$$j, j' = 1, \ldots, 5.$$

A simulation study



Candidate model

True model

	$j' = 1$	$j' = 2$	$j' = 3$	$j' = 4$	$j' = 5$
$j = 1$	$\mathbf{y}_1^1$ $\mathbf{y}_1^{101}$ $\vdots$ $\vdots$ $\mathbf{y}_1^{100}$ $\mathbf{y}_1^{200}$	$\mathbf{y}_1^1$ $\mathbf{y}_2^1$ $\vdots$ $\vdots$ $\mathbf{y}_1^{100}$ $\mathbf{y}_2^{100}$	$\cdots$	$\cdots$	$\mathbf{y}_1^1$ $\mathbf{y}_5^1$ $\vdots$ $\vdots$ $\mathbf{y}_1^{100}$ $\mathbf{y}_5^{100}$
$j = 2$	$\mathbf{y}_2^1$ $\mathbf{y}_1^1$ $\vdots$ $\vdots$ $\mathbf{y}_2^{100}$ $\mathbf{y}_1^{100}$	$\mathbf{y}_2^1$ $\mathbf{y}_2^{101}$ $\vdots$ $\vdots$ $\mathbf{y}_2^{100}$ $\mathbf{y}_2^{200}$	$\cdots$	$\cdots$	$\mathbf{y}_2^1$ $\mathbf{y}_5^1$ $\vdots$ $\vdots$ $\mathbf{y}_2^{100}$ $\mathbf{y}_5^{100}$
$j = 3$	$\vdots$	$\vdots$			$\vdots$
$j = 4$	$\vdots$	$\vdots$			$\vdots$
$j = 5$	$\mathbf{y}_5^1$ $\mathbf{y}_1^1$ $\vdots$ $\vdots$ $\mathbf{y}_5^{100}$ $\mathbf{y}_1^{100}$	$\mathbf{y}_5^1$ $\vdots$ $\vdots$ $\mathbf{y}_5^{100}$ $\mathbf{y}_2^{100}$	$\cdots$	$\cdots$	$\mathbf{y}_5^1$ $\mathbf{y}_5^{101}$ $\vdots$ $\vdots$ $\mathbf{y}_5^{100}$ $\mathbf{y}_5^{200}$

$$\hat{G}^1(j, j')$$

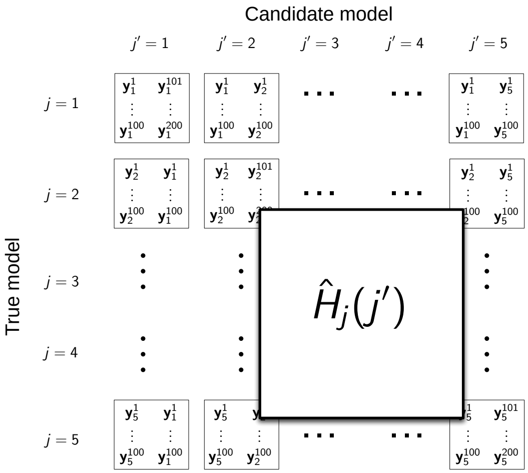
$$\vdots$$

$$\hat{G}^{100}(j, j')$$

$$\hat{G}_j^v(j') = \frac{\hat{S}^v(j, j')}{\max_{\{m\}}\{\hat{S}^v(j, m)\}},$$

$$v = 1, \ldots, 500,$$
$$m = 1, \ldots, 5,$$
$$j, j' = 1, \ldots, 5.$$

A simulation study



Candidate model

True model

	$j' = 1$	$j' = 2$	$j' = 3$	$j' = 4$	$j' = 5$
$j = 1$	$\mathbf{y}_1^1$ $\mathbf{y}_1^{101}$ $\vdots$ $\vdots$ $\mathbf{y}_1^{100}$ $\mathbf{y}_1^{200}$	$\mathbf{y}_1^1$ $\mathbf{y}_2^1$ $\vdots$ $\vdots$ $\mathbf{y}_1^{100}$ $\mathbf{y}_2^{100}$	$\cdots$	$\cdots$	$\mathbf{y}_1^1$ $\mathbf{y}_5^1$ $\vdots$ $\vdots$ $\mathbf{y}_1^{100}$ $\mathbf{y}_5^{100}$
$j = 2$	$\mathbf{y}_2^1$ $\mathbf{y}_1^1$ $\vdots$ $\vdots$ $\mathbf{y}_2^{100}$ $\mathbf{y}_1^{100}$	$\mathbf{y}_2^1$ $\mathbf{y}_2^{101}$ $\vdots$ $\vdots$ $\mathbf{y}_2^{100}$ $\mathbf{y}_2^{200}$	$\cdots$	$\cdots$	$\mathbf{y}_2^1$ $\mathbf{y}_5^1$ $\vdots$ $\vdots$ $\mathbf{y}_2^{100}$ $\mathbf{y}_5^{100}$
$j = 3$	$\vdots$	$\vdots$			$\vdots$
$j = 4$	$\vdots$	$\vdots$			$\vdots$
$j = 5$	$\mathbf{y}_5^1$ $\mathbf{y}_1^1$ $\vdots$ $\vdots$ $\mathbf{y}_5^{100}$ $\mathbf{y}_1^{100}$	$\mathbf{y}_5^1$ $\mathbf{y}_2$ $\vdots$ $\vdots$ $\mathbf{y}_5^{100}$ $\mathbf{y}_2^{100}$	$\cdots$	$\cdots$	$\mathbf{y}_5^1$ $\mathbf{y}_5^{101}$ $\vdots$ $\vdots$ $\mathbf{y}_5^{100}$ $\mathbf{y}_5^{200}$

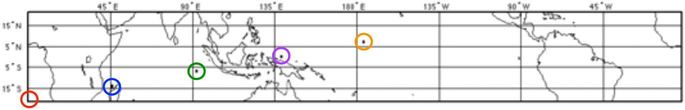
$$\hat{H}_j(j')$$

$$\hat{H}_j(j') = \text{mean}_{\{v\}}\{\hat{G}_j^v(j')\}.$$

$$v = 1, \ldots, 500,$$
$$m = 1, \ldots, 5,$$
$$j, j' = 1, \ldots, 5.$$

A simulation study

Synthetic models:



[-20,0]

Discrete-time AR model: A(z)y(t) = e(t)

A(z) = 1 - 0.8697 z^-1 + 0.2103 z^-2 - 0.09482 z^-3
+ 0.06311 z^-4 - 0.03068 z^-5 - 0.001724 z^-6
+ 0.01479 z^-7 - 0.0166 z^-8 - 0.002901 z^-9
+ 0.02878 z^-10 - 0.07958 z^-11 + 0.06159 z^-12
- 0.03055 z^-13 + 0.0629 z^-14 - 0.04819 z^-15
+ 0.00919 z^-16 + 0.01181 z^-17 + 0.006613 z^-18
- 0.01012 z^-19 - 0.007866 z^-20 + 0.02257 z^-21
- 0.007116 z^-22 + 0.01483 z^-23 - 0.01836 z^-24
- 0.007909 z^-25 + 0.01504 z^-26 - 0.01122 z^-27
- 0.01601 z^-28 - 0.0008788 z^-29 + 0.01156 z^-30
+ 0.03504 z^-31 - 0.05372 z^-32 - 0.01407 z^-33
+ 0.01914 z^-34 + 0.01553 z^-35 - 0.03193 z^-36
- 0.0421 z^-37 + 0.04482 z^-38 + 0.02292 z^-39
- 0.02216 z^-40 - 0.00727 z^-41 - 0.01076 z^-42
+ 0.03132 z^-43 - 0.03072 z^-44 - 0.002662 z^-45
+ 0.0217 z^-46 + 0.01244 z^-47 + 0.002948 z^-48
+ 0.007417 z^-49 - 0.04761 z^-50 - 0.001807 z^-51
- 0.01111 z^-52 + 0.05184 z^-53 + 0.002367 z^-54
- 0.04631 z^-55 + 0.0411 z^-56 - 0.03142 z^-57
+ 0.03263 z^-58 - 0.004381 z^-59 - 0.01737 z^-60
- 0.0003969 z^-61

[-13,45]

Discrete-time AR model: A(z)y(t) = e(t)

A(z) = 1 - 0.9039 z^-1 + 0.2817 z^-2 - 0.08821 z^-3
+ 0.001304 z^-4 + 0.0178 z^-5 - 0.02326 z^-6
- 0.01898 z^-7 + 0.02436 z^-8 - 0.04645 z^-9
+ 0.04858 z^-10 - 0.02563 z^-11

[1,138]

Discrete-time AR model: A(z)y(t) = e(t)

A(z) = 1 - 0.9139 z^-1 + 0.3564 z^-2 - 0.1338 z^-3
+ 0.03928 z^-4 + 0.01054 z^-5 - 0.01363 z^-6
+ 0.003495 z^-7 + 0.02252 z^-8 - 0.04506 z^-9
+ 0.01921 z^-10 - 0.001184 z^-11

[-6,91]

Discrete-time AR model: A(z)y(t) = e(t)

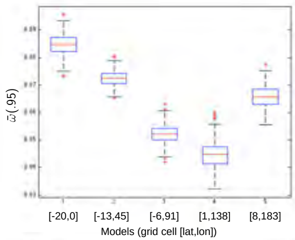
A(z) = 1 - 0.939 z^-1 + 0.3254 z^-2 - 0.1156 z^-3
- 0.00398 z^-4 - 0.004228 z^-5 - 0.009052 z^-6
- 0.0128 z^-7 - 0.006564 z^-8 + 0.004539 z^-9
- 0.01218 z^-10 - 0.02097 z^-11

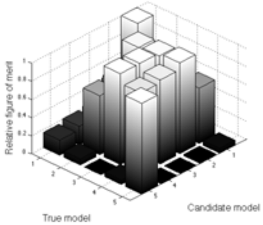
[8,183]

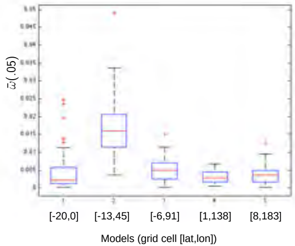
Discrete-time AR model: A(z)y(t) = e(t)

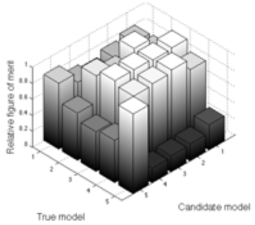
A(z) = 1 - 0.7658 z^-1 + 0.2778 z^-2 - 0.2191 z^-3
+ 0.02208 z^-4 + 0.05671 z^-5 - 0.05538 z^-6
- 0.01158 z^-7 + 0.005982 z^-8 - 0.01589 z^-9
+ 0.008379 z^-10 - 0.02543 z^-11

A simulation study









- Boxplots don’t entirely explain which models are distinguishable.