Statistical Simulation of Climate Time Series for Hypothesis Testing

Amy Braverman
Jet Propulsion Laboratory, California Institute of Technology

July 26, 2019

Based on joint work with Ansu Chatterjee (U of MN),
Megan Heyman (Rose-Hulman), and Noel Cressie (U of Wollongong)

Introduction

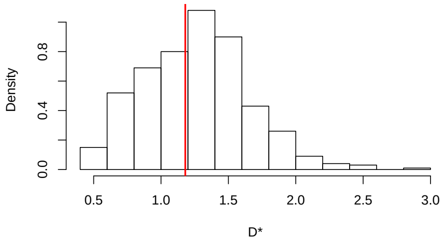
Main idea: decompose a time series into a signal series and a noise series, then resample the noise.

For instance: comparing two time series, a “benchmark", **Z**, and a “target", **X**,

- decompose each into a signal series and a noise series
- compute D, the distance between the benchmark and target signals’ series
- transform D into a probability...

Introduction

- consider D to be the value of a test statistic for testing H_0 : the benchmark and the target share the same signal
- the p-value of this test is a measure of the agreement of the data with the hypothesis, H_0 (Wasserstein and Lazar, 2016)

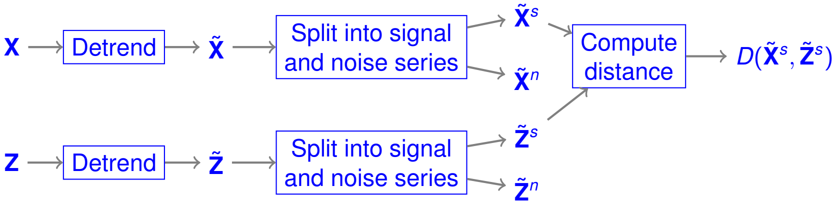


Introduction

- *This p-value is the probability of getting a value of D at least as extreme as the one obtained from the benchmark and the target, if they really did share the same signal.*

- The main technical challenge is obtaining the samplling distribution of D under H_0.

Statistical model and test statistic

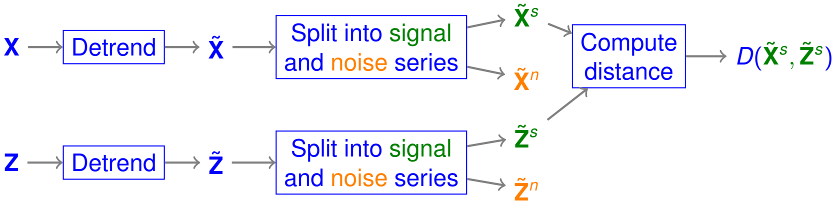


$$\mathbf{X} = (X(1), X(2), \ldots, X(T))', \quad X(t) = \hat{\alpha}_X + \hat{\beta}_X t + \tilde{X}(t), \quad t = 1, \ldots, T.$$

$$\mathbf{Z} = (Z(1), Z(2), \ldots, Z(T))', \quad Z(t) = \hat{\alpha}_Z + \hat{\beta}_Z t + \tilde{Z}(t), \quad t = 1, \ldots, T.$$

Convenient assumption: T is an exact power of 2; $T = 2^J$.

Statistical model and test statistic

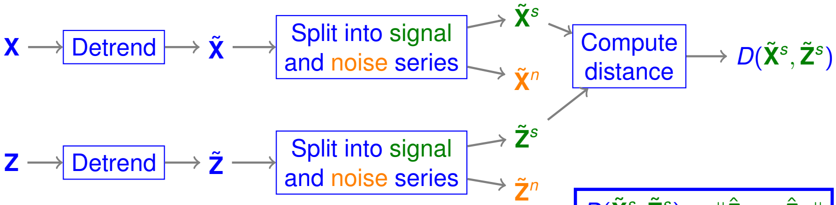


Wavelet decomposition:

$$\tilde{X}(t) = \tilde{X}^s(t) + \tilde{X}^n(t) = \sum_{j=0}^{J_0-1} \sum_{k=1}^{2^j-1} \hat{\gamma}_{Xjk} W_{jk}(t/T) + \sum_{j=J_0}^{J-1} \sum_{k=1}^{2^j-1} \hat{\gamma}_{Xjk} W_{jk}(t/T),$$

$$\tilde{Z}(t) = \tilde{Z}^s(t) + \tilde{Z}^n(t) = \sum_{j=0}^{J_0-1} \sum_{k=1}^{2^j-1} \hat{\gamma}_{Zjk} W_{jk}(t/T) + \sum_{j=J_0}^{J-1} \sum_{k=1}^{2^j-1} \hat{\gamma}_{Zjk} W_{jk}(t/T),$$

Statistical model and test statistic



$$D(\tilde{\mathbf{X}}^s, \tilde{\mathbf{Z}}^s) = \|\hat{\boldsymbol{\Gamma}}_X - \hat{\boldsymbol{\Gamma}}_Z\|$$

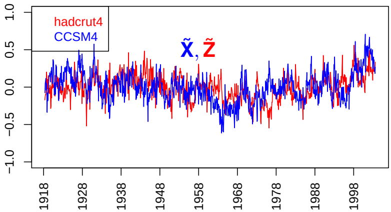
Wavelet decomposition:

$$\tilde{X}^s(t) = \sum_{j=0}^{J_0-1} \sum_{k=1}^{2^j-1} \hat{\gamma}_{Xjk} W_{jk}(t/T) \longrightarrow \hat{\boldsymbol{\Gamma}}_X = \left(\hat{\gamma}_{X00}, \ldots, \hat{\gamma}_{X(J_0-1)(2^{(J_0-1)}-1)}\right)',$$

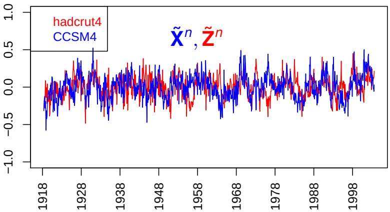
$$\tilde{Z}^s(t) = \sum_{j=0}^{J_0-1} \sum_{k=1}^{2^j-1} \hat{\gamma}_{Zjk} W_{jk}(t/T) \longrightarrow \hat{\boldsymbol{\Gamma}}_Z = \left(\hat{\gamma}_{Z00}, \ldots, \hat{\gamma}_{Z(J_0-1)(2^{(J_0-1)}-1)}\right)'.$$

Sampling distribution under H_0

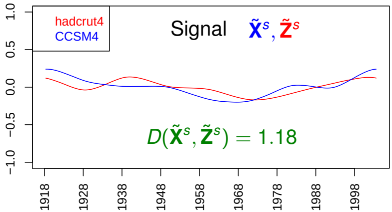
Detrended global surface temperature anomaly



Noise



Is 1.18 large or small?



Sampling distribution under H_0

$\{\mathbf{Z}_b^*, \mathbf{X}_b^*\}_{b=1}^B$

$\mathbf{Z}_1^*$	$\mathbf{X}_1^*$
$\mathbf{Z}_2^*$	$\mathbf{X}_2^*$
$\vdots$	$\vdots$
$\mathbf{Z}_B^*$	$\mathbf{X}_B^*$

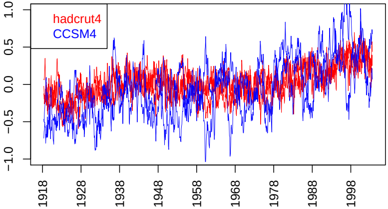
→ Detrend, split, compute distance →

$\{D_b^*\}_{b=1}^B$

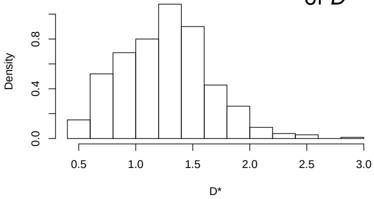
D_1^*
D_2^*
$\vdots$
D_B^*

Alternative plausible realizations of **Z** *and* **X** *give rise to sampling distribution of* D^*

Linked ensemble members



Sampling Distribution of D^*



Sampling distribution under H_0

$\mathbf{X} \rightarrow$ Detrend $\rightarrow \tilde{\mathbf{X}} \rightarrow$ Split into signal and noise series $\rightarrow \tilde{\mathbf{X}}^s$, $\tilde{\mathbf{X}}^n$

Detrend $\rightarrow (\hat{\alpha}_X, \hat{\beta}_X)$

$\tilde{\mathbf{X}}^n \rightarrow$ Fit time series model $\rightarrow$ ARIMA(p_X, d_X, q_X) $\rightarrow$ (bth draw) $\rightarrow \tilde{\mathbf{X}}_b^{n*}$

$\mathbf{Z} \rightarrow$ Detrend $\rightarrow \tilde{\mathbf{Z}} \rightarrow$ Split into signal and noise series $\rightarrow \tilde{\mathbf{Z}}^s$, $\tilde{\mathbf{Z}}^n$

Detrend $\rightarrow (\hat{\alpha}_Z, \hat{\beta}_Z)$

$\tilde{\mathbf{Z}}^n \rightarrow$ Fit time series model $\rightarrow$ ARIMA(p_Z, d_Z, q_Z) $\rightarrow$ (bth draw) $\rightarrow \tilde{\mathbf{Z}}_b^{n*}$

bth linked ensemble member:

$$\mathbf{Z}_b^* = \hat{\alpha}_Z \mathbf{1} + \hat{\beta}_Z \mathbf{t} + \tilde{\mathbf{Z}}^s + \tilde{\mathbf{Z}}_b^{n*},$$

$$\mathbf{X}_b^* = \hat{\alpha}_X \mathbf{1} + \hat{\beta}_X \mathbf{t} + \tilde{\mathbf{Z}}^s + \tilde{\mathbf{X}}_b^{n*}.$$

$\mathbf{1}$ = length T vector of one's; $\mathbf{t} = (1, 2, \ldots, T)'$.

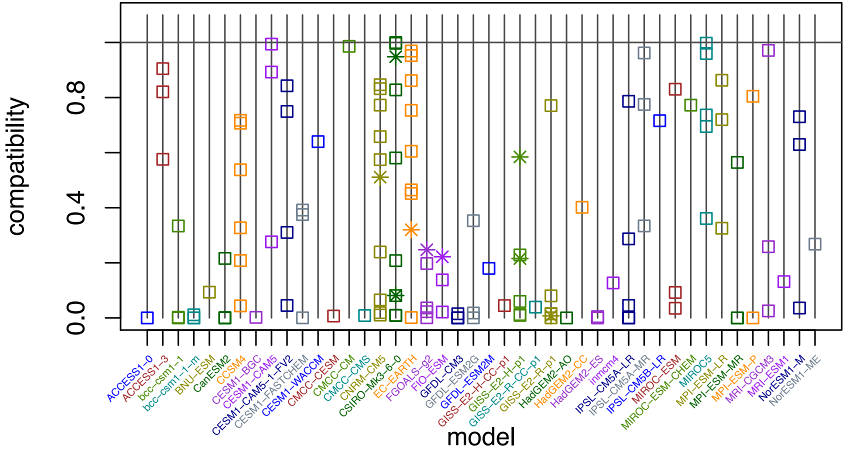
Application to CMIP5 historical runs

- Observational benchmark: HadCRUT4 (Morice et al., 2012) monthly global average surface temperature anomaly, May 1918 through August 2003 (1024 months).
- Model targets: 139 model runs using 44 different models from the CMIP5 historical experiment, May 1918 through August 2003.
- All time series computed relative to their 1961–1990 averages.
- “Signal" defined as the coarsest three wavelet decomposition levels ($J_0 = 3$), which corresponds to approximately ten-year cycles.
- Resampling performed with $B = 5000$ trials.

Application to CMIP5 historical runs

Model	*Center*	*Members*	*Model*	*Center*	*Members*
ACCESS1-0	CSIRO-BOM (Australia)	1	GFDL-ESM2M	GFDL (USA)	1
ACCESS1-3	CSIRO-BOM (Australia)	3	GISS-E2-H p1	NASA GISS (USA)	1
BCC-CSM-1	Beijing Climate Center (PRC)	3	GISS-E2-H-CC p1	NASA GISS (USA)	6
BCC-CSM-1-M	Beijing Climate Center (PRC)	3	GISS-E2-R p1	NASA GISS (USA)	1
BNU-ESM	Beijing Normal Univ. (PRC)	1	GISS-E2-R-CC p1	NASA GISS (USA)	6
CanSM2	CCCMA (Canada)	5	HadGEM2-AO	NIMR/KMA (UK/Korea)	1
CCSM4	NCAR (USA)	6	HadGEM2-CC	MOHC/INPE (UK/Brazil)	1
CESM1-BGC	NCAR/DOE/NSF (USA)	1	HadGEM2-ES	MOHC/INPE (UK/Brazil)	4
CESM1-CAM5	NCAR/DOE/NSF (USA)	3	INMCM4	INM (Russia)	1
CESM1-CAM5-1-FV2	NCAR/DOE/NSF (USA)	4	IPSL-CM5A-LR	IPSL (France)	6
CESM1-FASTCHEM	NCAR/DOE/NSF (USA)	3	IPSL-CM5A-MR	IPSL (France)	3
CESM1-WACCM	NCAR/DOE/NSF (USA)	1	IPSL-CM5B-LR	IPSL (France)	1
CMCC-CESM	CMCC (Italy)	1	MIROC-ESM	MIROC (Japan)	3
CMCC-CM	CMCC (Italy)	1	MIROC-ESM-CHEM	MIROC (Japan)	1
CMCC-CMS	CMCC (Italy)	1	MIROC5	MIROC (Japan)	5
CNRM-CM5	CNRM (France)	10	MPI-ESM-LR	MPI (Germany)	3
CSIRO-Mk3-6-0	CSIRO (Australia)	10	MPI-ESM-MR	MPI (Germany)	3
EC-EARTH	EC-EARTH Consortium (Europe)	9	MPI-ESM-P	MPI (Germany)	2
FGOALS-g2	LASG (PRC)	5	MRI-CGM3	MRI (Japan)	3
FIO-ESM	FIO (PRC)	3	MRI-ESM1	MRI (Japan)	1
GFDL-CM3	GFDL (USA)	5	NorESM1-M	NCC (Norway)	3
GFDL-ESM2G	GFDL (USA)	3	NorESM1-ME	NCC (Norway)	1

Results



∗ indicates runs for which ARIMA assumptions failed.

Discussion

- Nearly all models' outputs with more than one ensemble member show considerable variation.
- This is a property of their time series, particularly their noise series.
- We exploit their structures to mimic "internal variability" through resampling.
- Is this the right definition of internal variability?
- See Braverman, A., Chatterjee, S., Heyman, M., and Cressie, N.: Probabilistic evaluation of competing climate models, Adv. Stat. Clim. Meteorol. Oceanogr., 3, 93-105, https://doi.org/10.5194/ascmo-3-93-2017, 2017.

Question or comments? Contact Amy.Braverman@jpl.nasa.gov.