SAND2019-8568C

Evaluating Proxy Influence in Paleoclimate Reconstructions

Trevor Harris[1], Bo Li[1], Nathan Steiger[2], Jason Smerdon[2], Naveen Narisetty[1], Derek Tucker[3]
Department of Statistics, University of Illinois at Urbana-Champaign[1], Lamont-Doherty Earth Observatory[2], Sandia National Laboratories[3]

INTRODUCTION

Climate Field Reconstructions (CFR)

CFRs attempt to estimate spatiotemporal fields of climate variables over a historical period. CFRs are an important tool for studying the mechanisms of climate change

Data Assimilation (DA)

DA methods are a type of CFR that combines climate models with proxies. Proxies can include tree rings, ice cores, and corals. DA methods are advantageous because of their inherent uncertainty quantification with ensembles.

Proxy Influence

It is not known how much the proxies influence reconstructions over the base climate model. If reconstructions are just a product of the climate model then climate change forecasts based on DA are not fully incorporating historical data.

Statistical Challenge

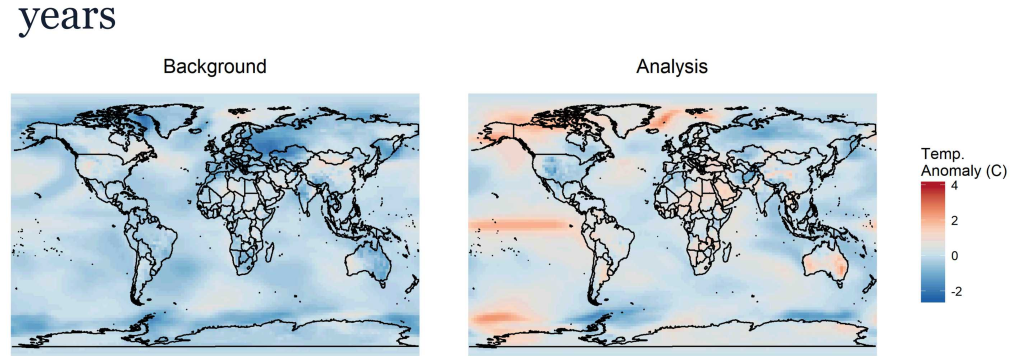
Use a specifically designed DA experiment to isolate proxy influence and recast the climatological problem as comparing two sets of spatiotemporal random fields. Use Functional Data Analysis and Data Depth to compare their distributions.

PROJECT GOALS

1. Develop a new statistical test for comparing the distribution of two sets of random fields.
2. Quantify and assess the influence of proxies in assimilated CFRs .

DATA

- Background: 100 fields on 144 x 96 grid points.
- Analysis: 100 fields on 144 x 96 grid points over 998 years



METHOD

Data and Hypothesis

- Observe $X_1, \ldots, X_n \sim P$ and $Y_1, \ldots, Y_m \sim Q$
- $H_0\colon P = Q;\quad H_A\colon P \neq Q$

Depth Definitions

- Define the Univariate Tukey depth

$$D(x(t), P_t) = 1 - |1 - 2P_t(x(t))|$$

- Define the Integrated Tukey depth

$$D(x, P) = \int_{[0,1]^p} D(x(t), P_t)\,dt$$

- For each X_k define the two empirical measures

$$\hat{F}_X(x_k) = \frac{1}{n}\sum_{i=1}^{n} 1(D(X_i, P_n) \leq D(x_k, P_n))$$

$$\hat{G}_Y(x_k) = \frac{1}{m}\sum_{i=1}^{m} 1(D(Y_i, P_n) \leq D(x_k, P_n))$$

- Measure the outlyingness of P over Q with

$$K_P(X, Y) = \max_{x_k \in X} |\hat{F}_X(x_k) - \hat{G}_Y(x_k)|$$

- For each Y_k define the two empirical measures

$$\tilde{F}_X(x_k) = \frac{1}{n}\sum_{i=1}^{n} 1(D(X_i, P_n) \leq D(y_k, P_n))$$

$$\tilde{G}_Y(x_k) = \frac{1}{m}\sum_{i=1}^{m} 1(D(Y_i, P_n) \leq D(y_k, P_n))$$

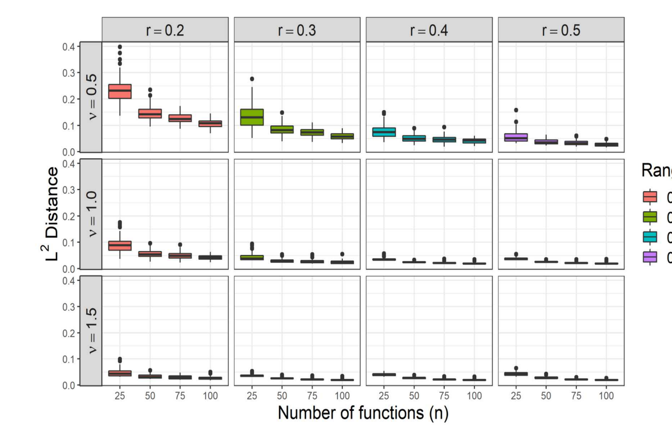
- Measure the outlyingness of Q over P with

$$K_Q(X, Y) = \max_{y_k \in Y} |\tilde{F}_X(y_k) - \tilde{G}_Y(y_k)|$$

Test Statistic

$$K(X, Y) = \max\{K_P(X, Y), K_Q(X, Y)\}$$

RESULTS

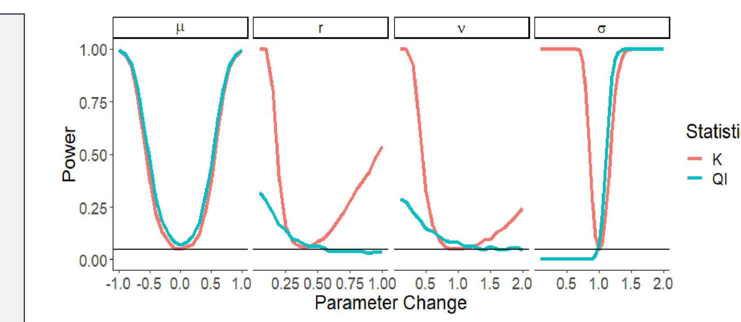


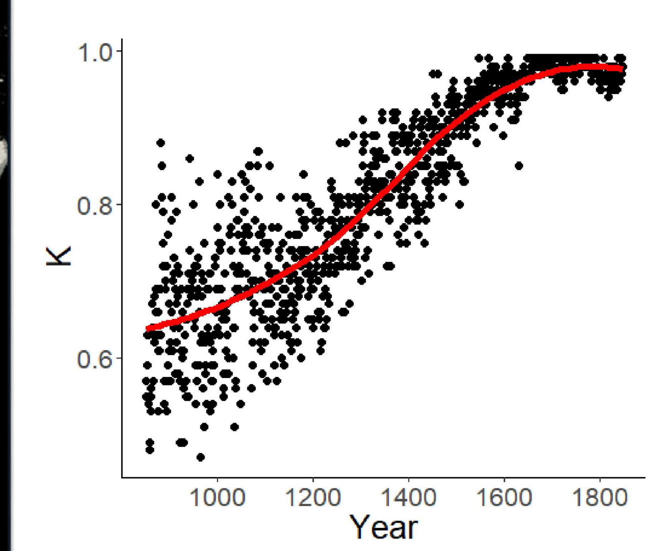
Simulations conducted using Gaussian Processes and Non-Gaussian processes

K Statistic converges to a distribution that is well approximated by the ***Kolmogorov distribution***

K has comparable or ***superior size control and power*** *over the next closest method: the Quality Index*

In particular K excels at correlation changes and variance decreases

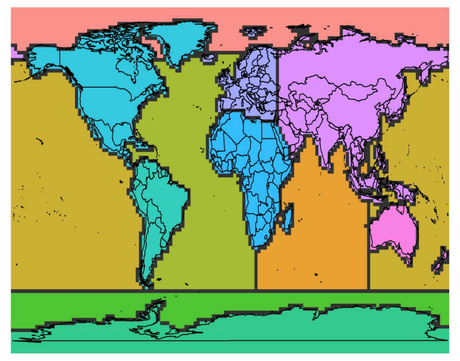




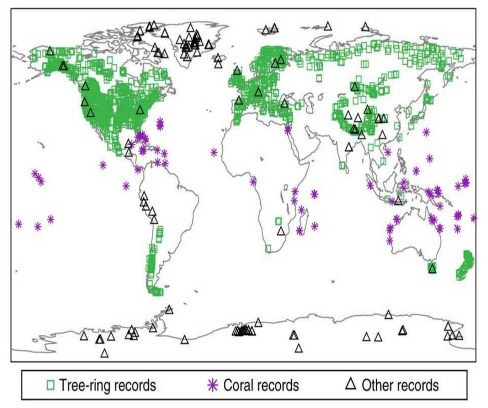
K values when comparing the background and analysis states in PHYDA

Strong upward trend *measures increasing proxy influence as the number of proxies increases*

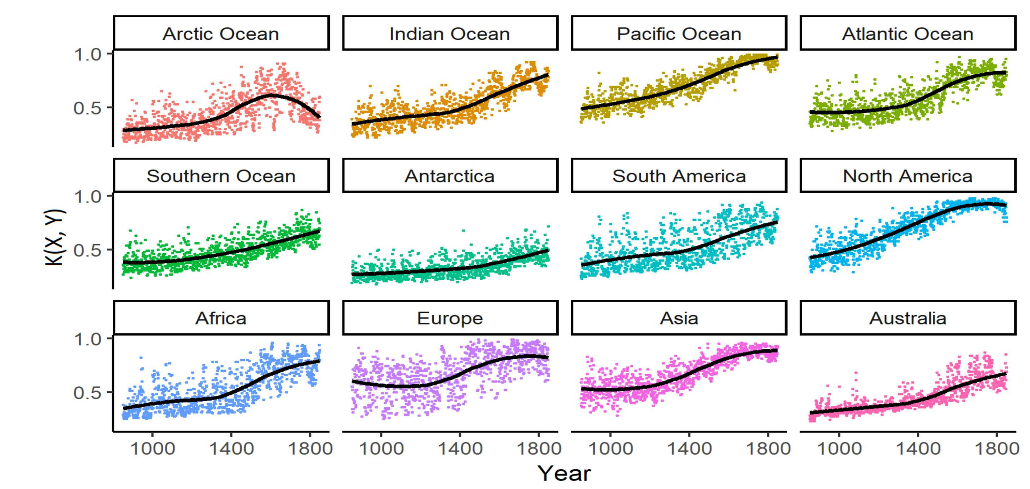
FDR adjusted p-values all significant



Split reconstructions into continental and oceanic regions to study local effects.



Proxies are not collected uniformly *across regions.*

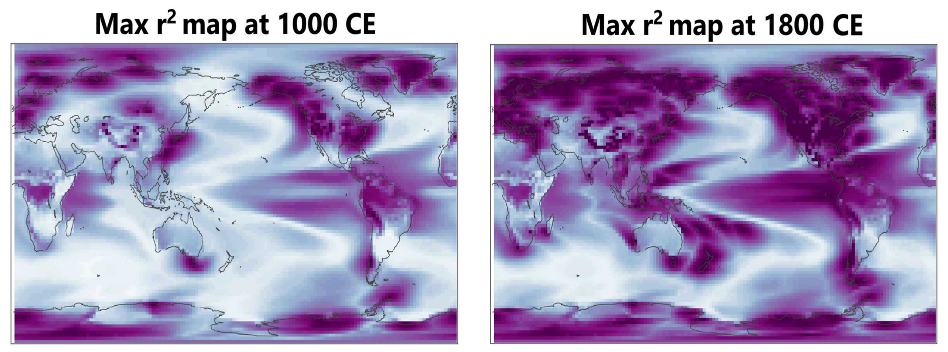


Proxy influence tends to increase more strongly in regions with many proxies than in regions with few proxies.

Arctic experiences a proxy influence reversal due to tree rings dominating ice core information

PROXY CORRELATION

- Some regions experience strong proxy influence despite a lack of local proxies, such as the Pacific.
- Lack of local proxies can be compensated for by strong correlation with regions with numerous proxies.



- Climate in the Pacific is highly connected with climate in North America due to the El-Niño Southern Oscillation.
- Teleconnection and local connection strength increases over time as more proxies become available for reconstruction.

CONCLUSIONS

- Strong evidence of divergence between the background and analysis states associated with PHYDA.
- Degree of separation depends on geographic location and time period.
- Increasing proxy information is associated with a commensurate influence over the reconstructions.
- Influence of the model prior is minimized as proxy networks become considerably more dense.

ACKNOWLEDGEMENTS

This research was supported in part by the National Science Foundation through awards OISE-1743738, AGS-1602581, AGS-1602920, and AGS-1805490 and the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. Background images from NASA's Earth Images.

Sandia National Laboratories

ILLINOIS