Statistical Downscaling of Sea Surface Temperature Projections with a Multivariate Gaussian Process Model

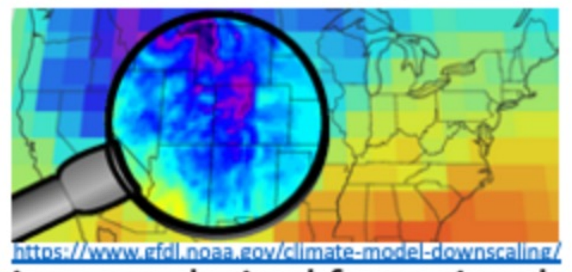
Ayesha Ekanayaka[1], Emily L. Kang[1], Peter Kalmus[2], Amy Braverman[2]
[1]Division of Statistics and Data Science, Department of Mathematical Sciences, University of Cincinnati
[2]Jet Propulsion Laboratory, California Institute of Technology

ekanaykk@mail.uc.edu

University of Cincinnati

Introduction

- Global Climate Models (GCMs) produce projections for future climate.
- GCM projections are usually made at coarse spatial resolutions ($\sim 1°$).



https://www.gfdl.noaa.gov/climate-model-downscaling/

- Downscaled GCM projections are desired for regional studies.
- There are two types of downscaling for climate model outputs.
  - Dynamical downscaling (DD)
    - Use Regional Climate Models (RCMs) and boundary conditions are generated from GCMs.
    - Computationally expensive.
  - Statistical Downscaling (SD)
    - Establish a statistical relationship between coarse-resolution model projections and fine-resolution observations.
    - Computationally less expensive.
- We propose an improved, computationally efficient SD method with a multivariate Gaussian process model.
- We present a demonstration using Sea Surface Temperature (SST).

Data

- We choose Great Barrier Reef (GBR) region as the study region.
- We use,
  - Monthly averaged NASA/JPL Multiscale Ultrahigh Resolution (MUR) satellite SST data at 1km resolution from June 2002 to December 2020.
  - Monthly SST outputs from 19 Coupled Model Intercomparison Project Phase 6 (CMIP6) GCM models at 100km resolution under ssp126 climate scenario.
  - ssp 126 is an extreme mitigation scenario which assumes less warming towards next century.
  - Our choice of the scenario here is only for the demonstration purpose.

http://www.mappery.com/map-of/Great-Barrier-Reef-Marine-Park-map

Model

- Let $Y_{t,1}(s)$ be the MUR observational SST at location s on month t and $Y_{t,2}(s)$ be the deterministically downscaled SST from the standard SD method.
- We assume the bi-variate process, $\boldsymbol{Y_t(s)} = (Y_{t,1}(s), Y_{t,2}(s))'$ follow,

$$\boldsymbol{Y_t(s)} = \boldsymbol{\mu_t(s)} + \boldsymbol{\nu_t(s)}$$

- $\boldsymbol{\mu_t(s)}$ is the mean component to describe variation with spatio-temporal interaction and $\boldsymbol{\nu_t(s)}$ is a multivariate Gaussian process to capture small-scale spatial variations. We assume,

$$\nu_{t,i}(s) = \sum_{k=1}^{K} \xi_i^k U^k(s) + \sum_{l=1}^{L} \eta_{t,i}^l S^l(s) + \varepsilon_{t,i}(s) \text{ for } i = 1,2$$

- ξ_i^k for $k = 1, \dots, K$ are a set of fixed effects and $\eta_{t,i}^l$ for $l = 1, \dots, L$ are a set of random effects to describe additional spatial variability but assumed to be independent across time.
- Motivated by Basis Graphical Lasso (BGL) we set $U^k(s)$ and $S^l(s)$ to be known orthogonal basis functions (Krock et al.,2021).
- Let $\boldsymbol{\eta_t^l} = (\eta_{t,1}^l, \eta_{t,2}^l)'$ for $l = 1, \dots, L$ and assume, $\boldsymbol{\eta_t^l} \sim \boldsymbol{N(0, Q_l^{-1})}$.
- We assume the precision matrix $\boldsymbol{Q}$ of the full vector $\boldsymbol{\eta_t} = (\boldsymbol{\eta_t^1}, \dots, \boldsymbol{\eta_t^L})'$ is, $\boldsymbol{Q} = diag(\boldsymbol{Q_1}, \dots, \boldsymbol{Q_L})$.
- Last term $\varepsilon_{t,i}(s)$ is white noise with zero mean and τ_i^2 variance which is independent from random effects.

Implementation

- Let $\boldsymbol{U(s)} = (U^1(s), \dots, U^K(s))'$ and $\boldsymbol{S(s)} = \left(S^1(s), \dots, S^L(s)\right)'$.
- The model for the full data vector can be written as,

$$\begin{pmatrix} \boldsymbol{Y_t(s_1)} \\ \vdots \\ \boldsymbol{Y_t(s_N)} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu_t(s_1)} \\ \vdots \\ \boldsymbol{\mu_t(s_N)} \end{pmatrix} + \begin{pmatrix} \boldsymbol{U'(s_1) \otimes I_2} \\ \vdots \\ \boldsymbol{U'(s_N) \otimes I_2} \end{pmatrix} \begin{pmatrix} \boldsymbol{\xi^1} \\ \vdots \\ \boldsymbol{\xi^K} \end{pmatrix} + \begin{pmatrix} \boldsymbol{S'(s_1) \otimes I_2} \\ \vdots \\ \boldsymbol{S'(s_N) \otimes I_2} \end{pmatrix} \begin{pmatrix} \boldsymbol{\eta_t}^1 \\ \vdots \\ \boldsymbol{\eta_t}^L \end{pmatrix} + \begin{pmatrix} \boldsymbol{\varepsilon_t(s_1)} \\ \vdots \\ \boldsymbol{\varepsilon_t(s_N)} \end{pmatrix}$$

where, $\boldsymbol{\xi^k} = (\xi_1^k, \xi_2^k)'$ and $\boldsymbol{\eta_t}^{\boldsymbol{l}} = (\eta_{t,1}^l, \eta_{t,2}^l)'$.

- Estimation is done in two-steps.
- Step1: Detrend and obtain the detail residuals vector $\boldsymbol{Z_t} = \boldsymbol{Y_t} - \boldsymbol{\hat{\mu}_t} - \boldsymbol{U\hat{\xi}}$. Then the log-likelihood for the residuals is written as,

$$\log(\det(\boldsymbol{\Sigma})) + \frac{\sum_{t=1}^{T} \boldsymbol{Z_t'}\Sigma^{-1}\boldsymbol{Z_t}}{T} = \log(\det(\boldsymbol{\Sigma})) + tr(\boldsymbol{\hat{\Sigma}\Sigma^{-1}}) \text{ where,}$$

$$var(\boldsymbol{Z_t}) = \boldsymbol{\Sigma}, \boldsymbol{\hat{\Sigma}} = \frac{\sum_{t=1}^{T} \boldsymbol{Z_t'Z_t}}{T}, \boldsymbol{\Sigma} = \boldsymbol{SQ^{-1}S'} + \boldsymbol{D} \text{ and } \boldsymbol{D} = diag(\tau_1^2, \tau_2^2) \otimes \boldsymbol{I_N}$$

- Step2: Solve for,

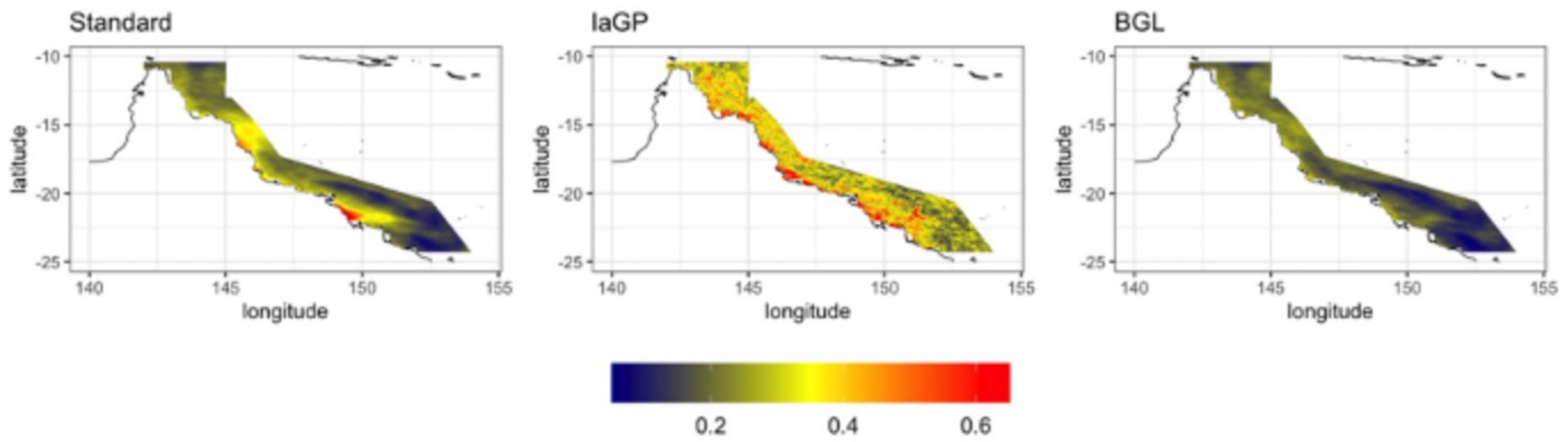
$$\boldsymbol{\hat{Q}} \in \arg\min \log(\det(\boldsymbol{SQ^{-1}S'} + \boldsymbol{D})) + tr\left(\boldsymbol{\hat{\Sigma}}\left(\boldsymbol{SQ^{-1}S'} + \boldsymbol{D}\right)^{-1}\right) + P(\boldsymbol{Q}) \text{ where,}$$

$P(\boldsymbol{Q}) = P(\boldsymbol{Q_1}, \dots, \boldsymbol{Q_L}) = \rho \sum_{l=1}^{L-1} \sum_{i \neq j} |(\boldsymbol{Q_l})_{ij} - (\boldsymbol{Q_{l+1}})_{ij}|$ and ρ is a penalty which penalize $\boldsymbol{Q_l}$ matrices at adjacent levels if the off-diagonals are not similar (Krock et al.,2021).

Results

- We compare Mean Square Error (MSE) of our proposed method with,
  - Interpolated raw GCM projections
  - State-of-art (Standard) SD method for SSTs (Van Hooidonk et al.,2015)
  - local approximate Gaussian Process (laGP) (Gramacy et al., 2015)

ssp126				
Season	GCM	Standard	laGP	BGL
Summer	0.396	0.305	0.419	**0.297**
Autumn	0.472	0.105	0.221	**0.088**
Winter	0.878	0.250	0.335	**0.163**
Spring	0.380	0.212	0.356	**0.153**
Overall	0.531	0.218	0.353	**0.175**



Conclusions and Future Work

- Our downscaling method is computationally feasible for large data sets, accounts for spatio-temporal dependencies, provides meaningful uncertainty estimates, and produce improved downscaling results compared with the state-of-the-art methods.
  - Downscaling for the GBR region (a total of 309,700 high-resolution pixels) was performed in Matlab on a Macbook Air with an 8-core 3.2GHz processor and 8 GB RAM.
  - Given the second process $Y_{t,2}(s)$ for $t > T_o$ for all $s \in \mathcal{D}$, computation time to fit the model and estimate SSTs for a single future month is 6.8 seconds.
- A possible extension is to generalize the current model to the framework of autoregressive co-kriging for multi-fidelity model output and then consider the observations, regional climate model output, and global climate model output as the high-, medium-, and low-fidelity data, respectively.

Acknowledgements

- This work is partially funded by, NASA/JPL , NSF (DMS-2053668), Simons Foundation's Collaboration Award (#317298 and #712755), University Research Council (URC) and TAFT research center at University of Cincinnati.