

# Statistical Modeling and Inference for Climate Applications

Jon Hobbs

Jet Propulsion Laboratory, California Institute of Technology

Jonathan.M.Hobbs@jpl.nasa.gov

## 1. Bayesian Inference

- Hierarchical Model

$$\text{Data} = Y(\mathbf{s}_i, t)$$

$$\text{Process} = \alpha(\mathbf{s}_i, t)$$

$$\text{Parameters} = \boldsymbol{\theta}$$

$$\begin{aligned} [\text{Data}, \text{Process}, \text{Parameters}] &= [\text{Data}|\text{Process}, \text{Parameters}] \\ &\times [\text{Process}|\text{Parameters}] \\ &\times [\text{Parameters}] \end{aligned}$$

- Posterior Distribution

$$[\boldsymbol{\theta}|\mathbf{Y}] \propto \int [\mathbf{Y}|\boldsymbol{\alpha}, \boldsymbol{\theta}] [\boldsymbol{\alpha}|\boldsymbol{\theta}] [\boldsymbol{\theta}] d\boldsymbol{\alpha}$$

## 2. Spatial Models

Model the joint distribution for the spatial field  $\mathbf{Y}_t$

- Covariance model  
Model  $Cov(Y(\mathbf{s}_i, t), Y(\mathbf{s}_j, t))$  as a function of spatial distance
- Spatial random effects  
Model  $Y(\mathbf{s}_i, t)$  as a linear combination of a low-dimensional random vector and spatial basis functions.
- Conditionally-specified model  
Specify conditional distributions given neighboring values

$$Y(\mathbf{s}_i, t)|\mathbf{y}_t(N_i)$$

## References

- Berliner, L. M. (2003). Physical-statistical modeling in geophysics. *Journal of Geophysical Research*, 108. doi: 10.1029/2002JD002865
- Cressie, N. (2018). Mission CO<sub>2</sub>ntrol: a statistical scientist's role in remote sensing of atmospheric carbon dioxide. *Journal of the American Statistical Association*, 113(521), 152–168. doi: 10.1080/01621459.2017.1419136
- Cressie, N., & Johannesson, G. (2008). Fixed rank kriging for very large spatial data sets. *Journal of the Royal Statistical Society, Series B*, 70, 208–226. doi: 10.1111/j.1467-9868.2007.00633.x
- Cressie, N., & Wikle, C. K. (2011). *Statistics for spatio-temporal data*. Hoboken, NJ: John Wiley & Sons, Inc.
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis* (Third ed.). Boca Raton, FL: Chapman & Hall/CRC.
- Heaton, M. J., Datta, A., Finley, A. O., Furrer, R., Guinness, J., Guhaniyogi, R., . . . Zammit-Mangion, A. (2018). A case study competition among methods for analyzing large spatial data. *Journal of Agricultural, Biological, and Environmental Statistics*. doi: 10.1007/s13253-018-00348-w
- Katzfuss, M., & Guinness, J. (2021). A general framework for Vecchia approximations of Gaussian processes. *Statistical Science*, 36, 124–141. doi: 10.1214/19-STS755
- Wikle, C. K., & Berliner, L. M. (2005). Combining information across spatial scales. *Technometrics*, 47, 80–91. doi: 10.1198/004017004000000572
- Wikle, C. K., & Berliner, L. M. (2007). A Bayesian tutorial for data assimilation. *Physica D: Nonlinear Phenomena*, 230, 1–16. doi: 10.1016/j.physd.2006.09.017
- Wikle, C. K., & Zammit-Mangion, A. (2023). Statistical deep learning for spatial and spatiotemporal data. *Annual Review of Statistics and Its Application*, 10, 247–270. doi: 10.1146/annurev-statistics-033021-112628
- Wilks, D. S. (2020). *Statistical methods in the atmospheric sciences* (Fourth ed.). Amsterdam: Elsevier.