A Nonparametric Copula Based Bias Correction Method for Statistical Downscaling

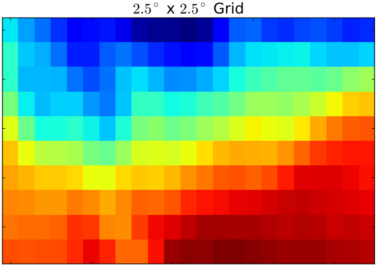
Yi Li, Adam Ding, and Jennifer Dy.

Northeastern University

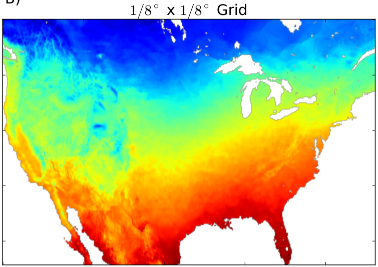
September 2016

Statistical Downscaling

A)



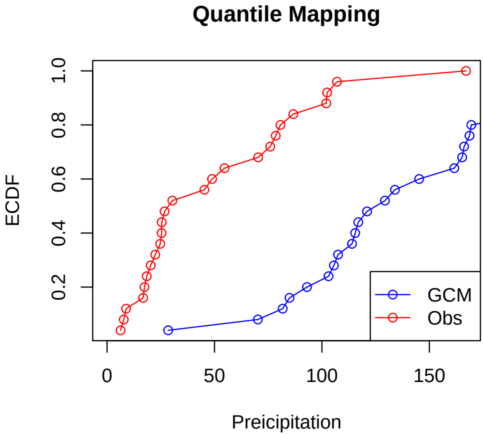
B)



- Coarsely resolved climate models vs Local-scale climate information

BCSD

- Bias correction with Quantile Mapping (our main focus)



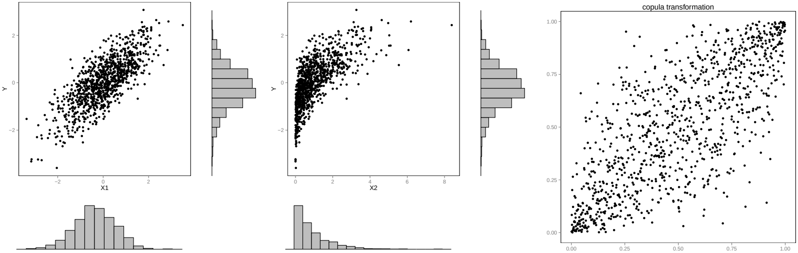
- Spatial disaggregation with historical local scaling factor

Copula

For d random variables $X = (X_1, \cdots, X_d)$, by Sklar's Theorem,

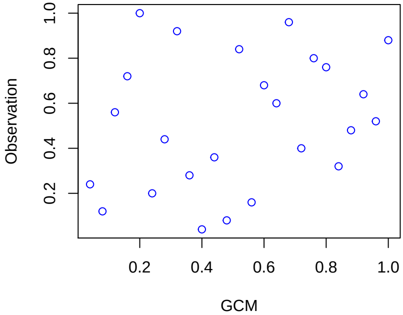
$$F(x_1, \cdots, x_d) = C[F_1(x_1), \cdots, F_d(x_d)], \tag{1}$$

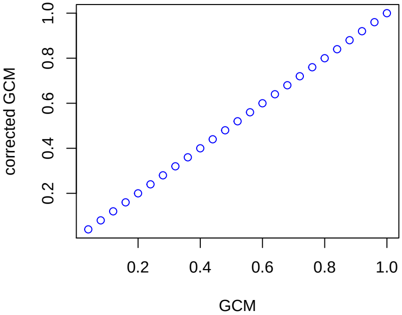
- $F_j(x) = \mathbb{P}(X_j \leq x)$. C is a copula, where $C(u_1, ..., u_d)$ is the CDF of $U_j = F_j(X_j)$.
- This decomposition separates the dependence structure in the data from the marginals.



NCBCSD

- Quantile mapping is using $C(u, v) = \min(u, v)$ (Fréchet-Hoeffding upper bound), which assumes the strongest dependancy.

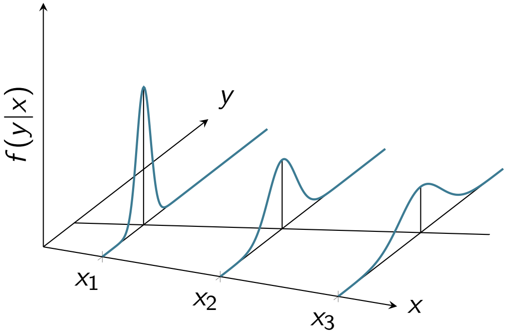




- Nonparametric copula based BCSD method (NCBCSD)

NCBCSD as Modal Regression

$$\mathtt{mode}(Y|X = x) = \mathtt{argmax}_y f_{Y|X}(y|x), \tag{2}$$



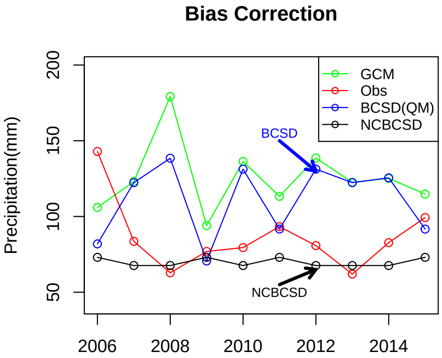
NCBCSD as Modal Regression

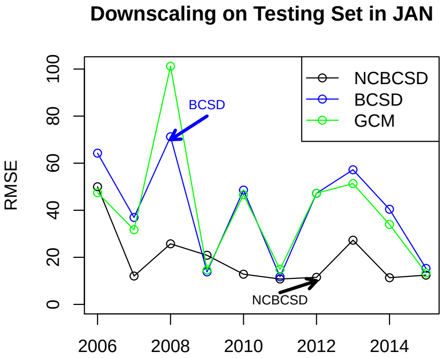
$$\texttt{mode}(Y|X = x) = \texttt{argmax}_y \hat{c}(u, v)\hat{f}_Y(y), \tag{3}$$

where $U = F_X(X)$, $V = F_Y(Y)$ and with the (nonparametric) kernel density estimation (KDE):

$$\begin{aligned} \text{Copula:} \quad & \hat{c}_{kde}(u, v) = \frac{1}{nh^2}\sum_{i=1}^{n} K(\frac{u - u_i}{h})K(\frac{v - v_i}{h}), \\ \text{Marginal:} \quad & \hat{f}_{Y_{kde}}(y) = \frac{1}{nh}\sum_{i=1}^{n} K(\frac{y - y_i}{h}). \end{aligned} \tag{4}$$

Results





- Climate Variables: Monthly precip. of south New England.
- GCM Data: GFDL's Coupled Physical Model (CM3).
- Local Data: Univ. of Idaho Gridded Surface Meteorological Data.

Thank you!