Spatial smoothing of zero-inflated abundance data

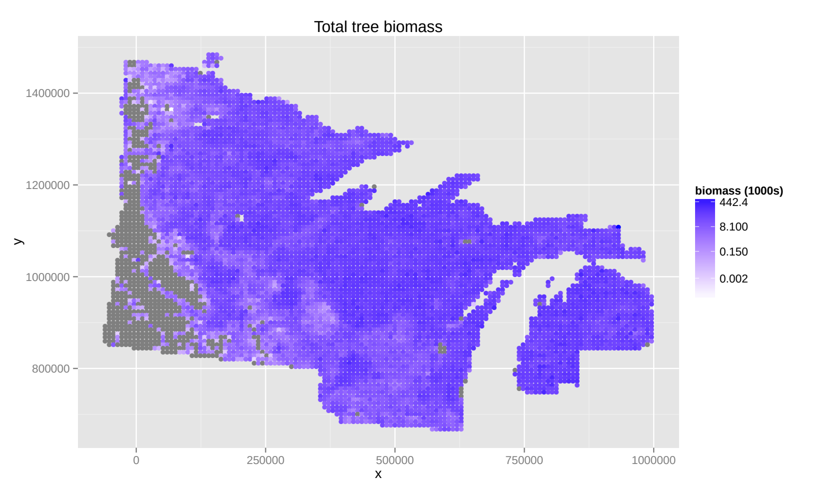
Wesley Brooks

Department of Statistics
University of Wisconsin–Madison

.

Motivation

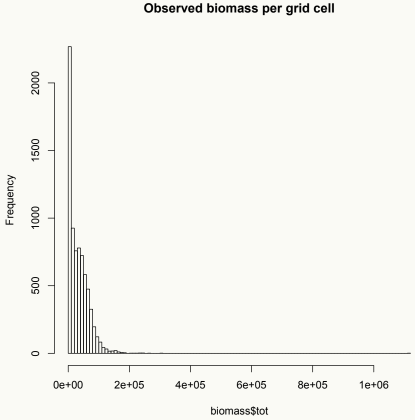
Take a look at some data



Motivation

Take a look at some data

The data is non-negative with a long right tail, suggesting a log transformation or a gamma model.



Motivation

Take a look at some data

The data is the total tree biomass per grid cell from the Public Land Survey of the upper midwest.

- Observations are at corner points (90 per grid cell)
- Want to know the actual biomass on the grid cell from these local samples
- Goal is to calculate the total biomass of each taxon across the upper midwest
- Many grid cells have zero observed biomass

Motivation

Take a look at some data

Want a model for which power-law variation (long right tail) and exact zeroes are handled naturally, not as an exception.

- We use the Tweedie family of distribution
- Tweedie distribution has a tuning parameter that slides smoothly from Poisson to Gamma distribution
- Estimating that tuning parameter is difficult, and current state of the art is a priori specification
- A better solution may be to estimate the parameter by matching the theoretical and observed mean-variance relationship.

Title

Tweedie distribution

- Exponential dispersion model:

$$f(y;\theta;\phi) = a(y,\phi)\,\mathsf{exp}\left[\phi^{-1}\left\{y\theta - \kappa(\theta)\right\}\right]$$

Title

Tweedie distribution

- Tweedie model's parameters:

$$
\begin{aligned}
\lambda &= \frac{\mu^(2-p)}{\phi(2-p)} \\
\alpha &= \frac{2-p}{1-p} \\
\gamma &= \phi(p-1)\mu^{p-1}
\end{aligned}
$$

Title

Tweedie distribution

- Probability of exact zero:

$$P(Y = 0) = \text{exp}\left\{-\frac{\mu^{2-p}}{\phi(2-p)}\right\}$$