

Spatial smoothing of zero-inflated abundance data

Wesley Brooks

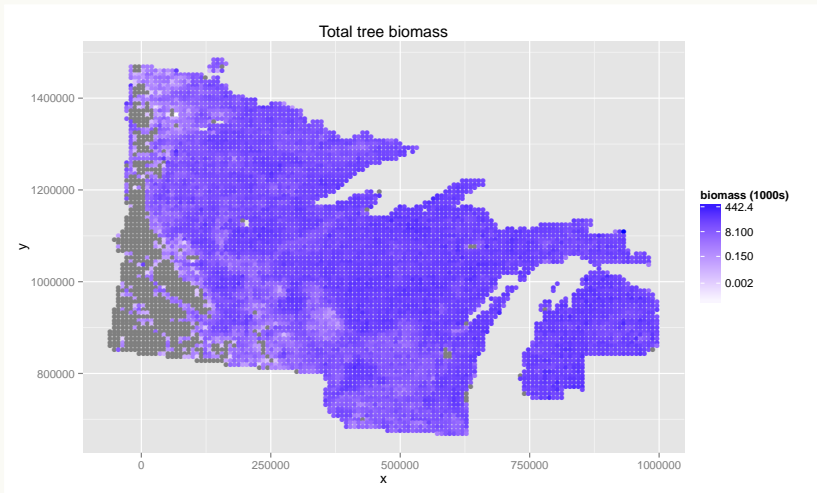
Department of Statistics
University of Wisconsin–Madison

.

Motivation

Motivation

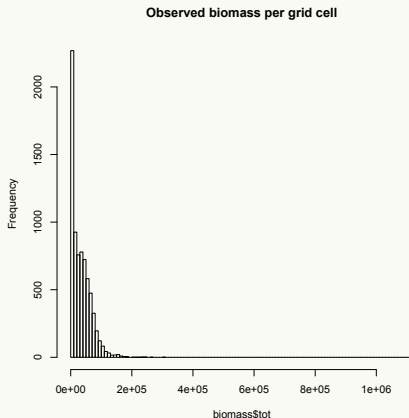
A look at the data



Motivation

A look at the data

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



Motivation

A look at the 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

Our approach

Motivation

Recall our approach

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
- ▶ Estimate that tuning parameter via MLE

Biomass data

Biomass data

Drawing biomass samples

- ▶ Create a model for each taxon
- ▶ Use MLE matching to estimate $\hat{\theta}$
- ▶ Use parametric bootstrap to get draws from biomass
- ▶ distribute draws from the "total" taxon based on the relative biomass of the other taxa

Biomass data

Complications

- ▶ Smoothing biomass via a GAM for Wisconsin with 500 knots:
 - Too few knots
 - Takes too long to complete (some taxa never finished running)
- ▶ Possible solution: adapt R-INLA for Tweedie likelihood

Biomass data

Complications

► Incomplete taxa:

- Cherries
- Willow
- Walnuts
- Hickory
- Fir
- Spruce
- Cedar
- Hemlock
- Pine
- Birches

Biomass data

Complications

► Complete taxa:

- tot
- Beech
- Ironwoods
- Basswood
- Ashes
- Elms
- Poplar
- Tamarack
- Maple
- Oaks