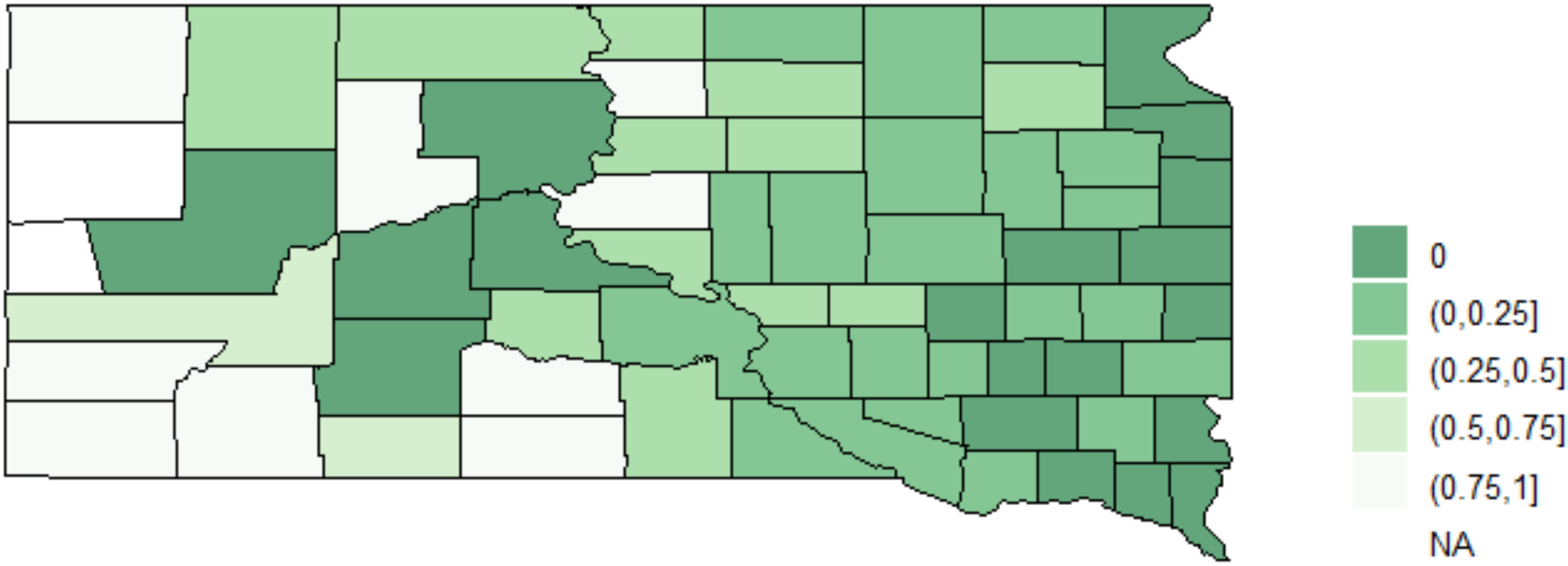
Iowa State University

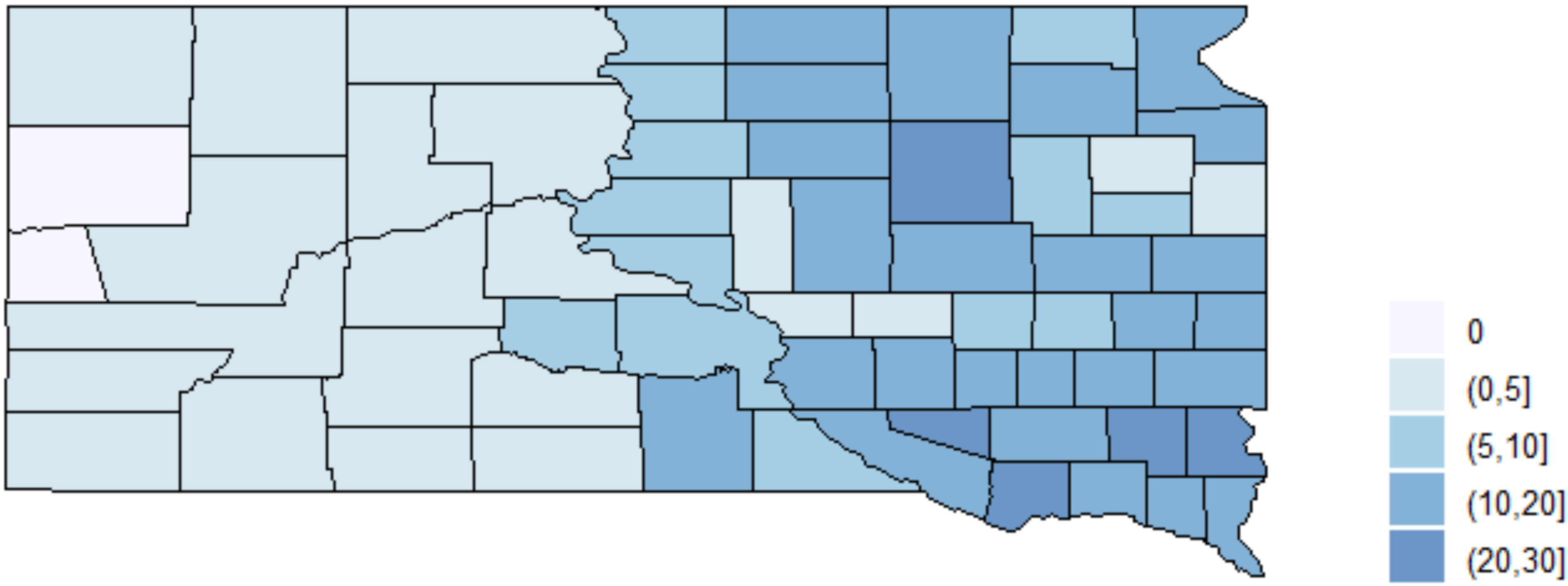
Soil Loss Data

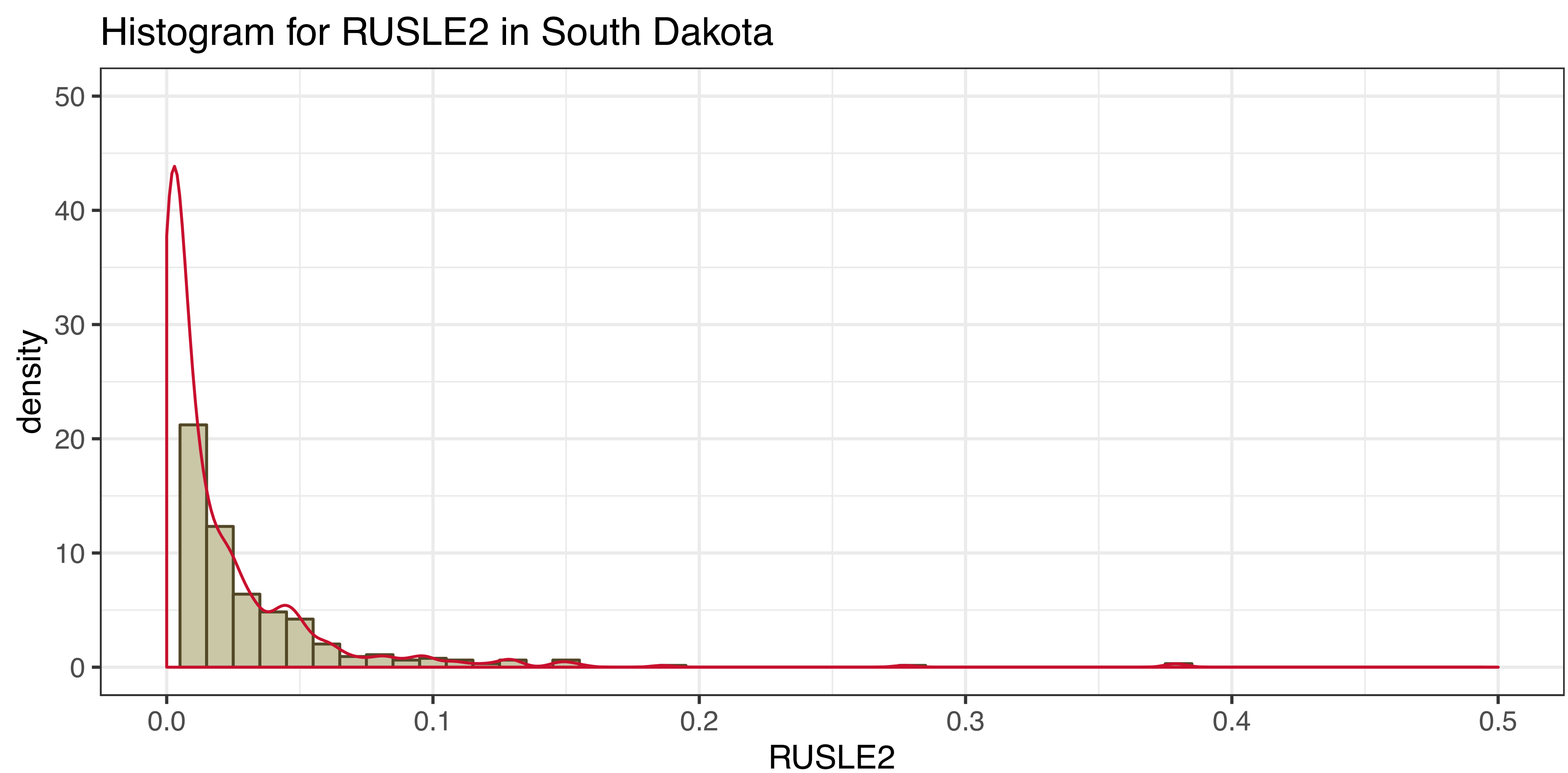
3

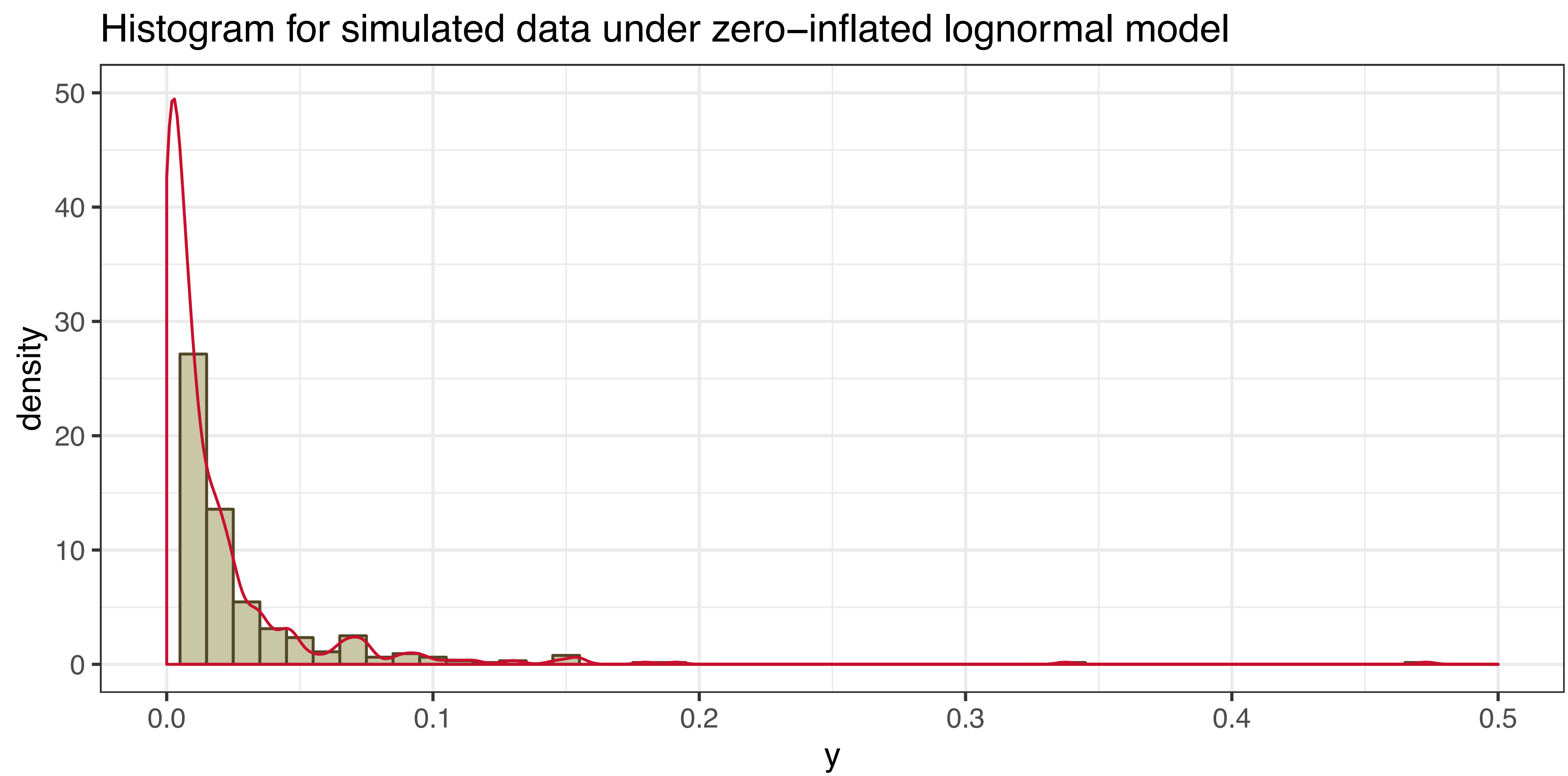
Heatmap of proportion of zeros of RUSLE2 observations in South Dakota



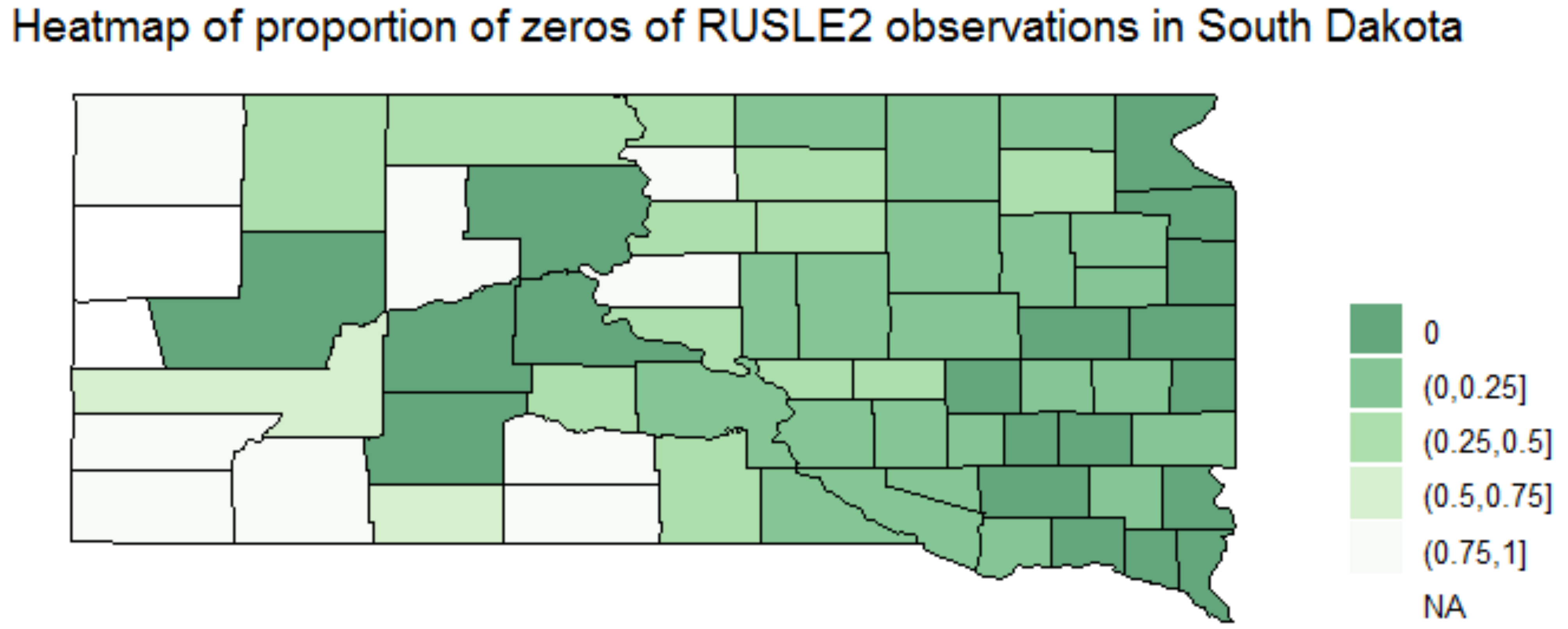
Heatmap of sample sizes of RUSLE2 observations in South Dakota

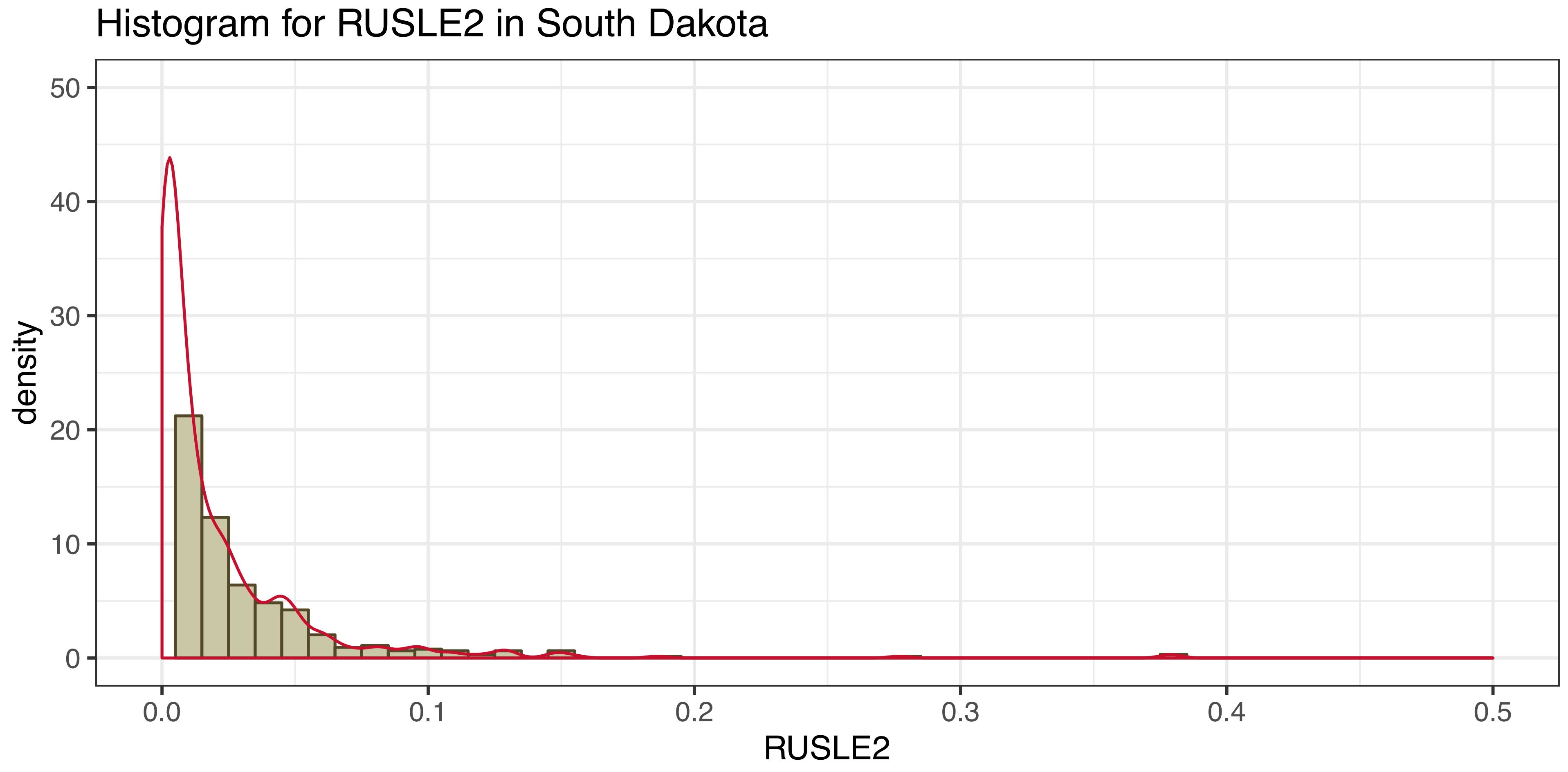


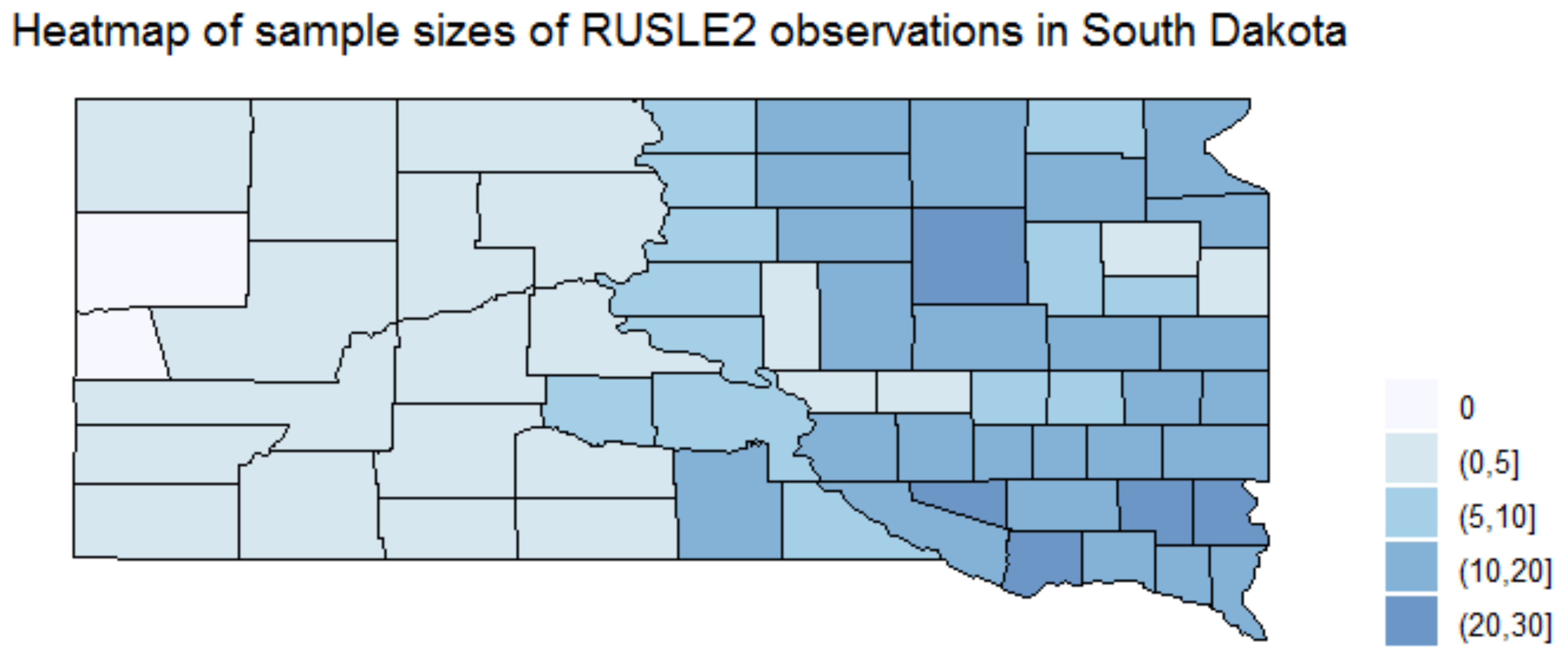

Histogram for RUSLE2 in South Dakota
density
50
40
30
20
10
0
0.0
0.1
0.2
0.3
0.4
0.5
RUSLE2

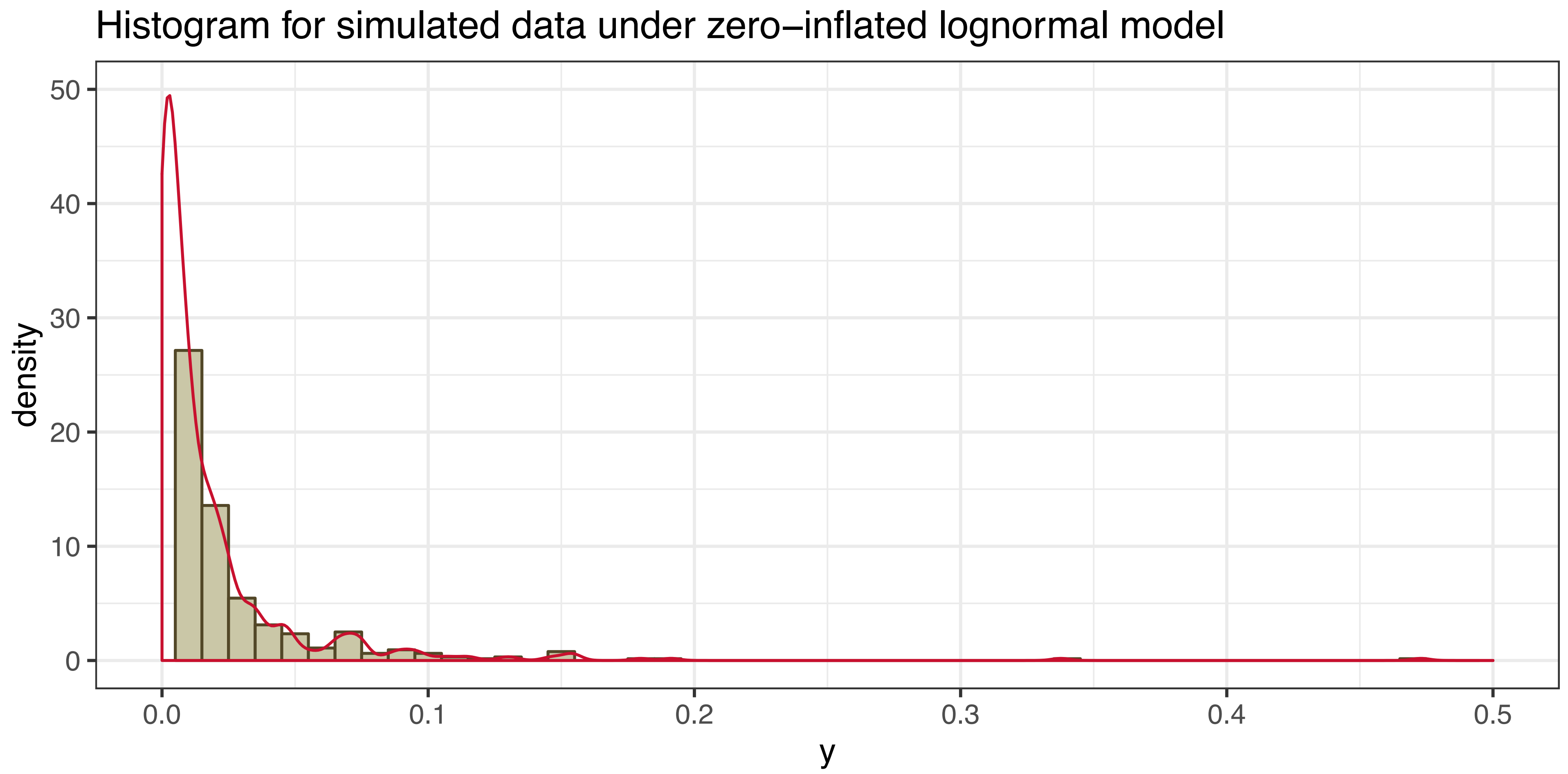

Histogram for simulated data under zero−inflated lognormal model
density
50
40
30
20
10
0
0.0
0.1
0.2
0.3
0.4
0.5
y

Soil Loss Data









Model

- Variable of interest: $y_{ij}^* = y_{ij}\delta_{ij} \geq 0$, population parameter: $\bar{y}_{N_i}^* = \frac{1}{N_i}\sum_{j=1}^{N_i} y_{ij}^*$
  - Positive part: $\log(y_{ij}) = \beta_0 + \boldsymbol{z}_{1ij}'\boldsymbol{\beta}_1 + u_i + e_{ij}$
  - Binary part: $\delta_{ij} \sim \text{Bernoulli}(p_{ij}), g(p_{ij}) = \alpha_0 + \boldsymbol{z}_{2ij}'\boldsymbol{\alpha}_1 + b_i, g(\cdot)$ is a specific parametric link function.
  - $(u_i, b_i, e_{ij}) \sim N(\boldsymbol{0}, \text{diag}(\sigma_u^2, \sigma_b^2, \sigma_e^2))$
- Observed data: $(\boldsymbol{y}^*, \boldsymbol{z}) = \{y_{ij}^*, i = 1, \ldots, D, j \in s_i\} \cup \{\boldsymbol{z}_{ij} : i = 1, \ldots, D, j = 1, \ldots, N_i\}$
- Empirical Bayes predictor: $\hat{y}_{ij}^{*\text{EB}} = \hat{y}_{ij}^{*\text{MMSE}}|_{\hat{\boldsymbol{\theta}}=\boldsymbol{\theta}}, \hat{y}_{ij}^{*\text{MMSE}} = E_{\boldsymbol{\theta}}\{y_{ij}^* \,|\, (\boldsymbol{y}^*, \boldsymbol{z})\}$
- MSE estimator: analytic “one-step” and parametric bootstrap.