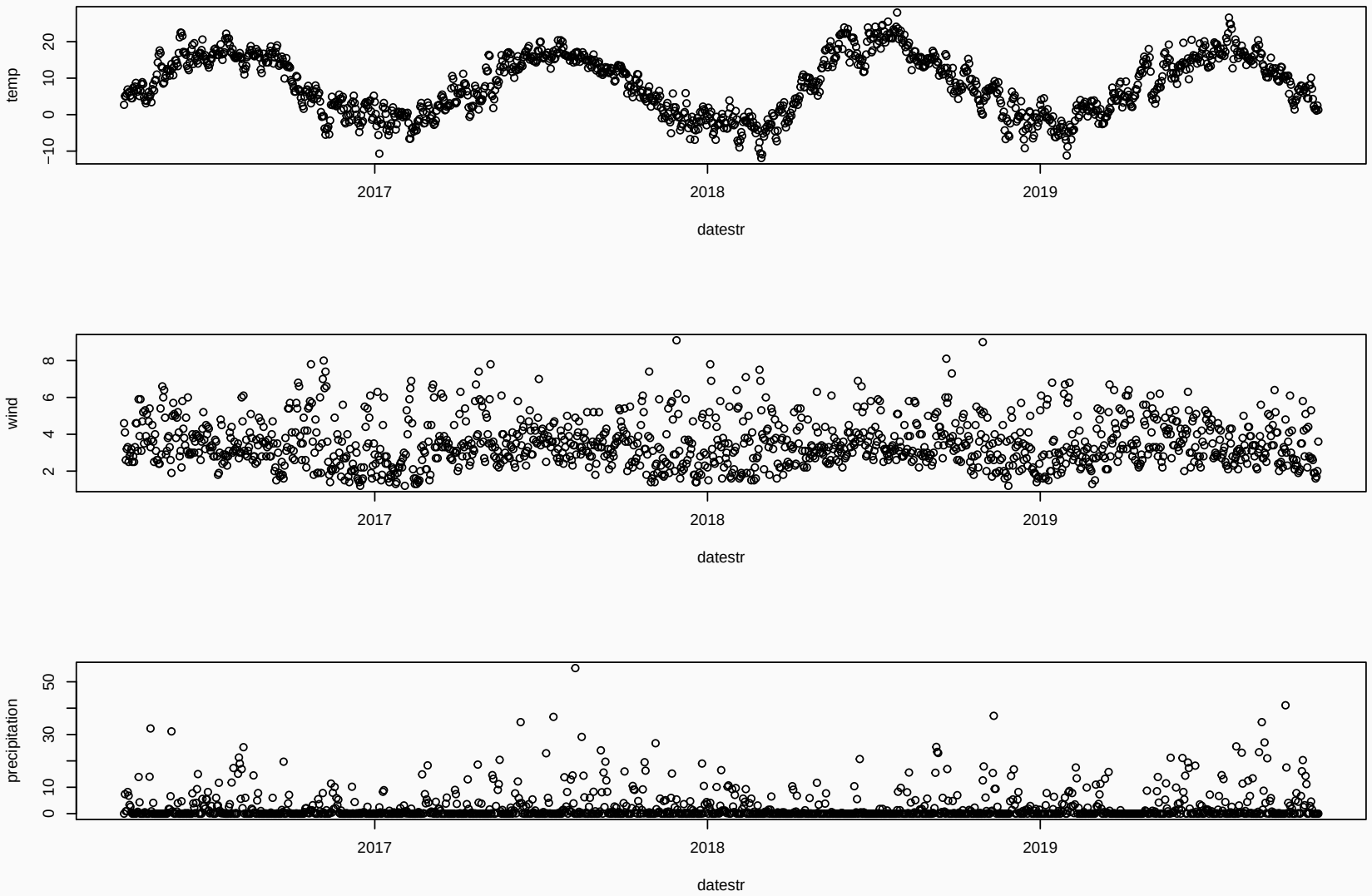
Weather and Daily City Bike Usage

Joshua Krusell

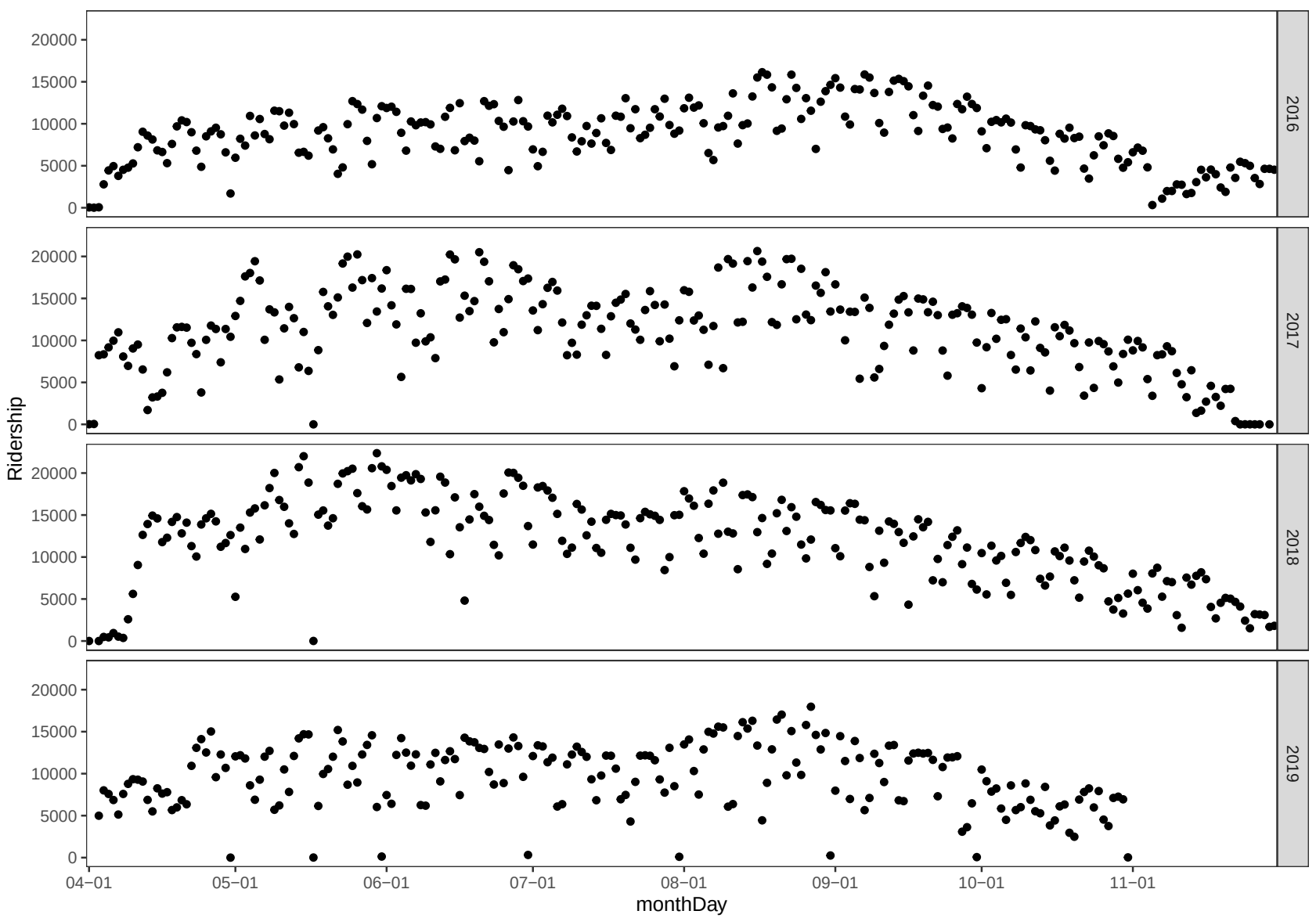
2019-11-05

Replication code available at https://github.com/jsks/RAV

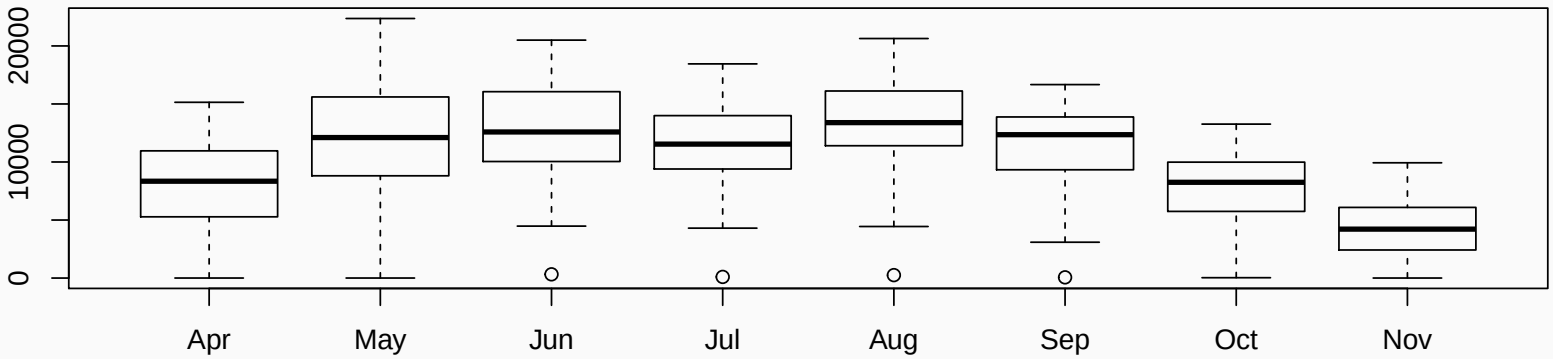
Weather Data

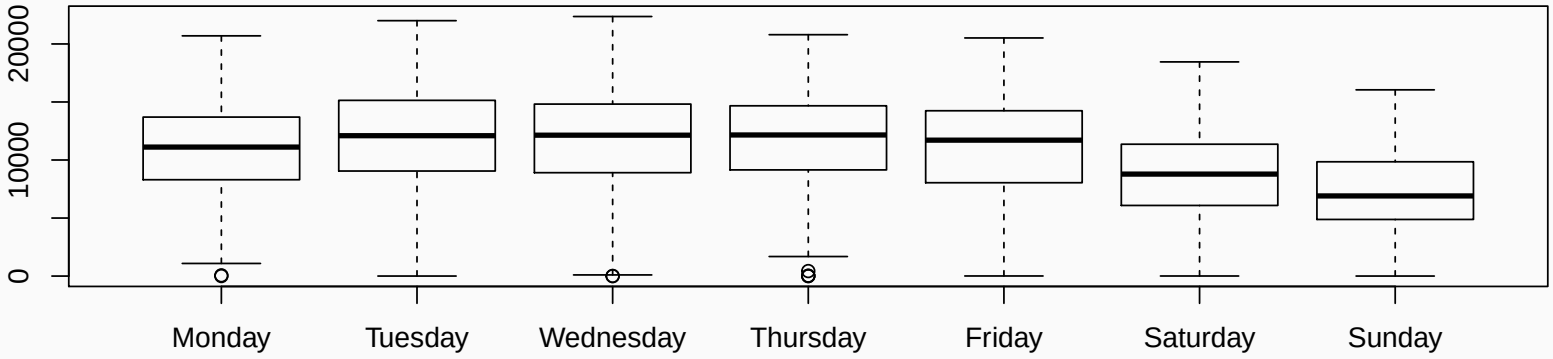


Bicycle Usage

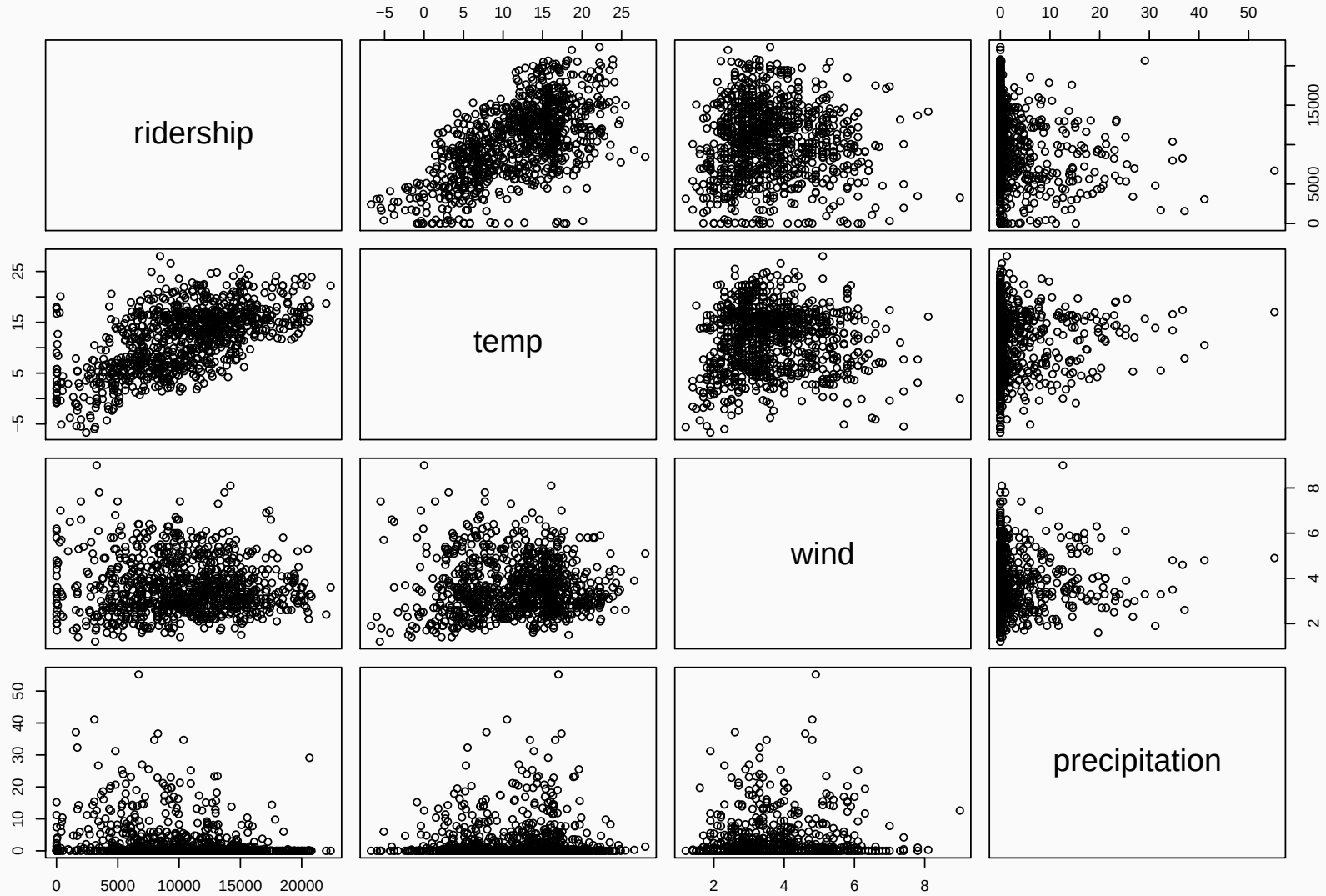


Bicycle Usage





Scatterplot Matrix



Model

Let:

- $y \in \mathbb{N}^n$ be the daily ridership count
- $X \in \mathbb{R}^{n \times m}$ the matrix of predictors
- $\zeta^k \in \mathbb{R}^{J_k}$ the set of partially pooled random intercepts for K grouping factors

$$y_i \sim \text{NegBinomial}(\mu_i, \phi)$$

$$\log(\mu_i) = \alpha + X_i\beta + \sum_{k=1}^{K} \zeta^k_{j_i}$$

$$\mathbb{E}[Y] = \mu \text{ and } \text{Var}[Y] = \mu + \frac{\mu^2}{\phi}$$

Priors

Weakly informative priors:

$$\alpha \sim \text{Normal}(0, 10)$$

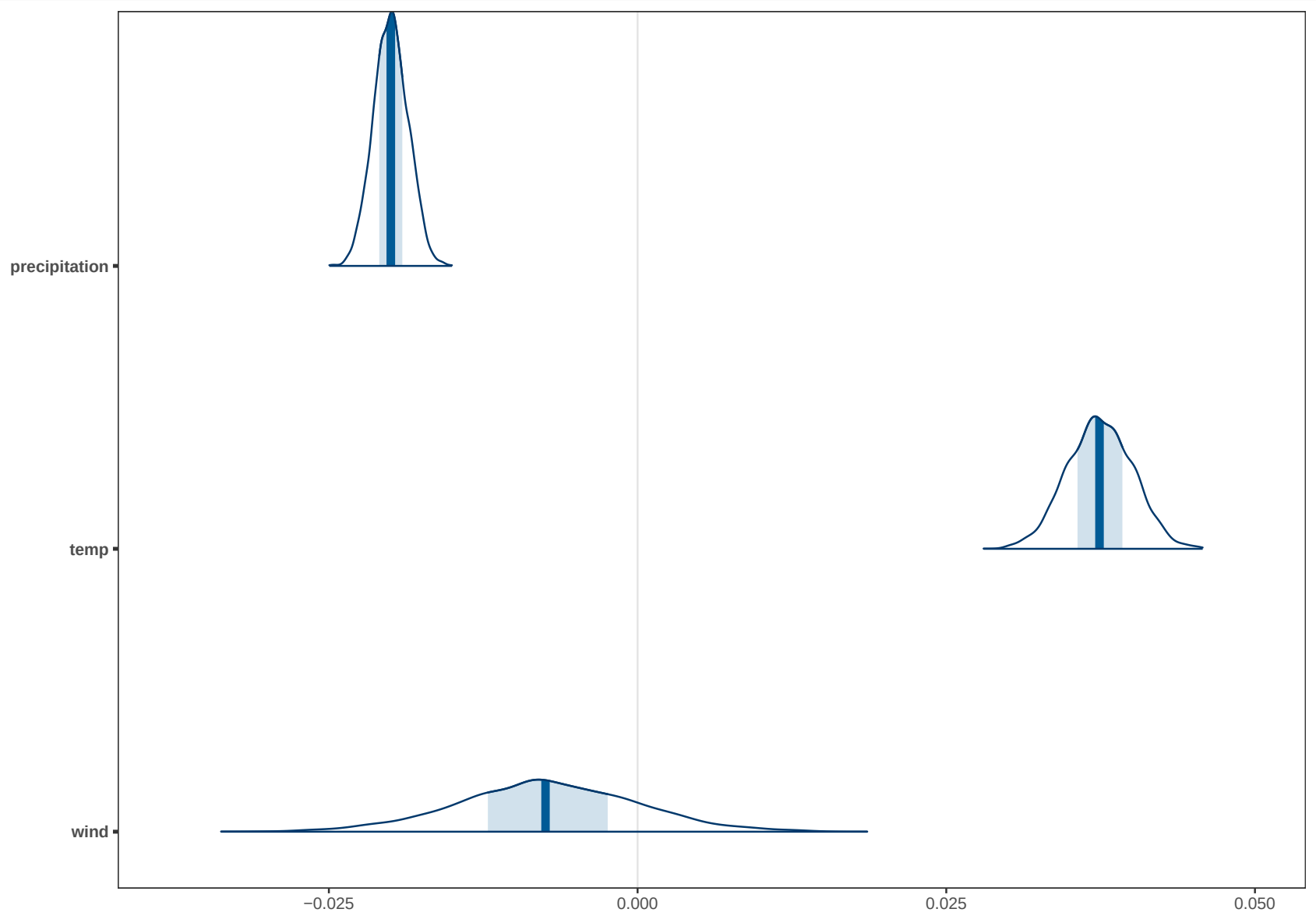
$$\beta \sim \text{Normal}(0, 5)$$

$$\zeta^k \sim \text{Normal}(0, \sigma^k)$$

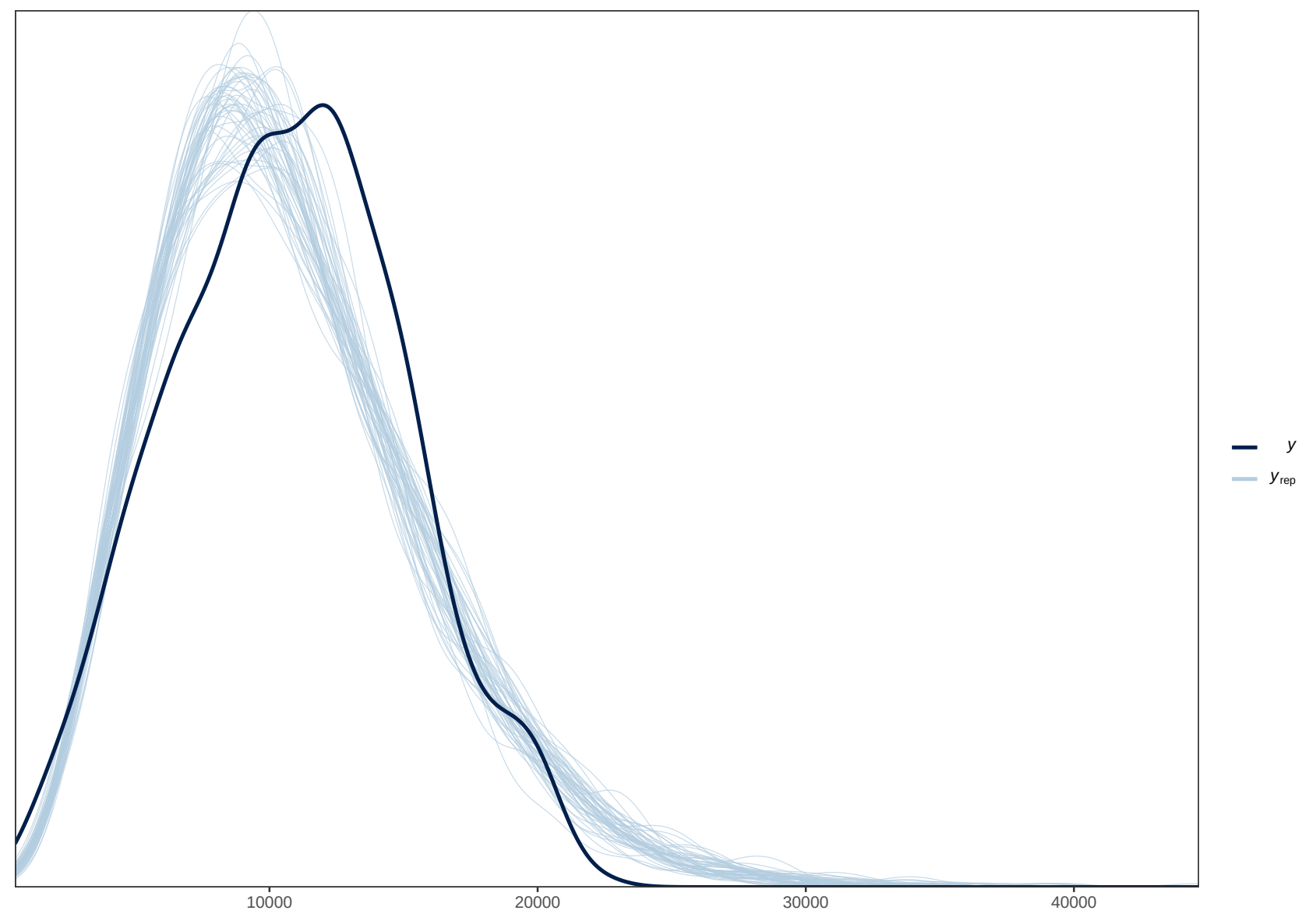
$$\sigma^k \sim \text{HalfCauchy}(0, 1)$$

$$\phi \sim \text{HalfCauchy}(0, 1)$$

Results

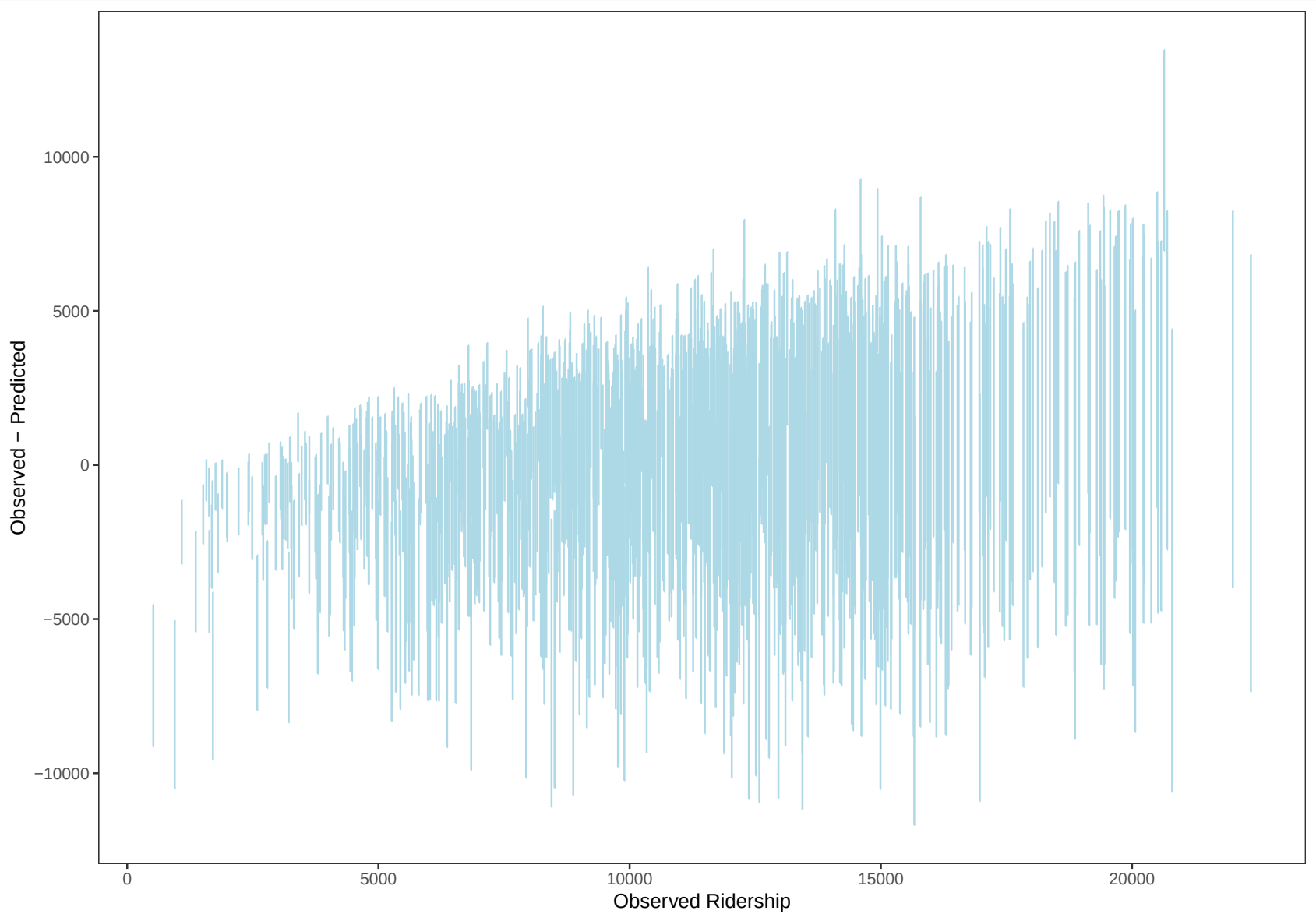


Posterior Predictive Checking

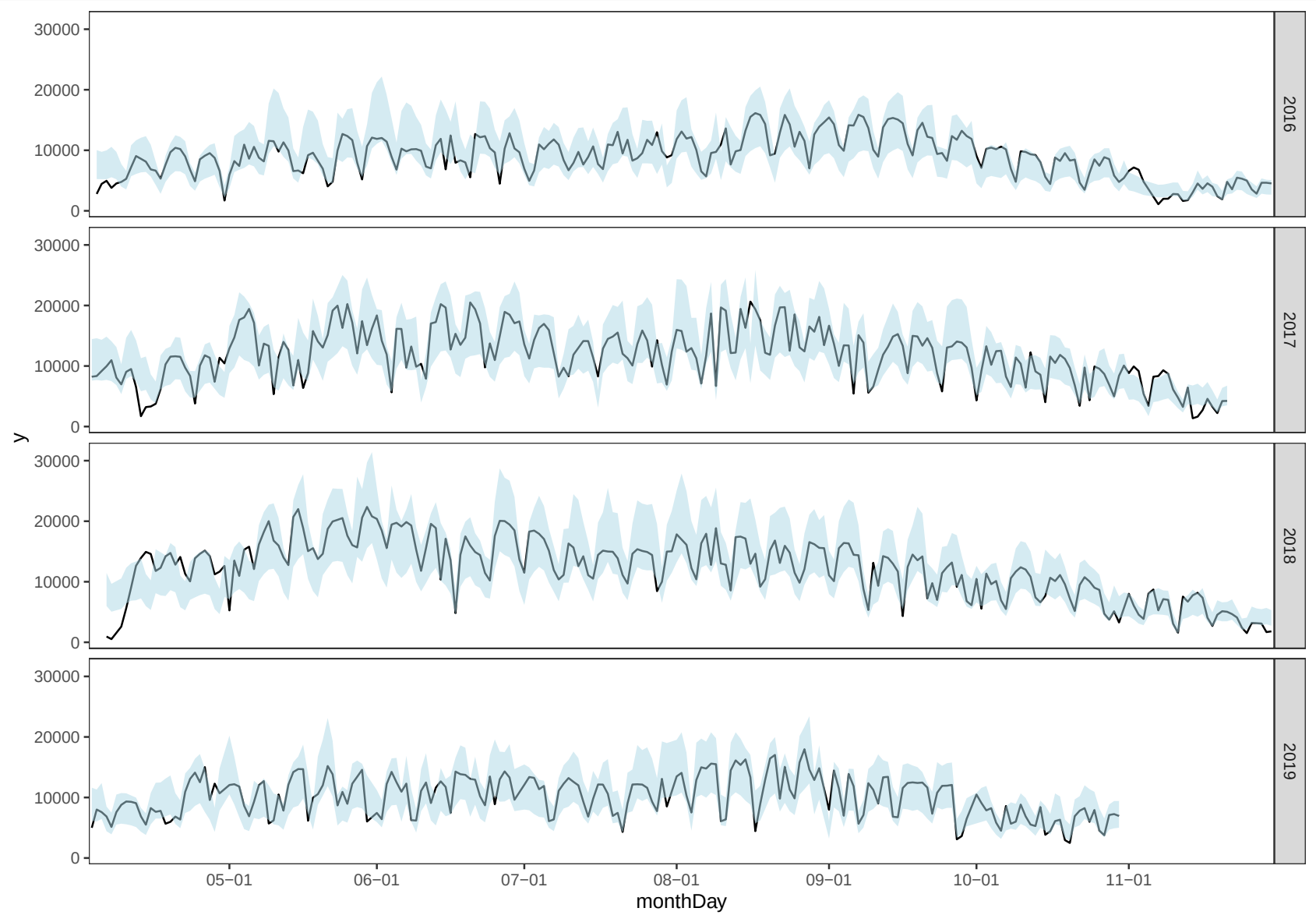


Posterior Predictive Error

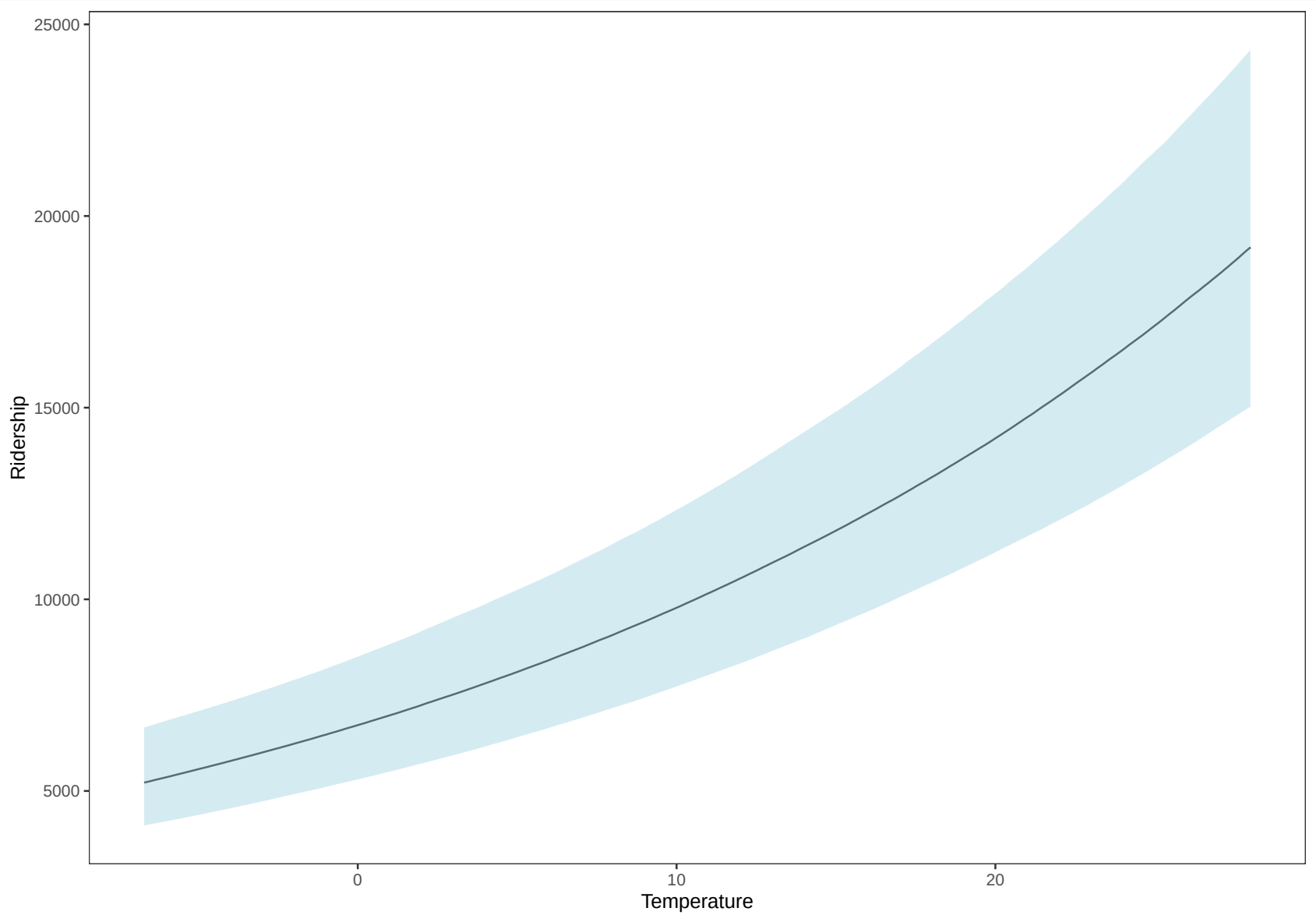
RMSE: 3394.83 - 3679.83



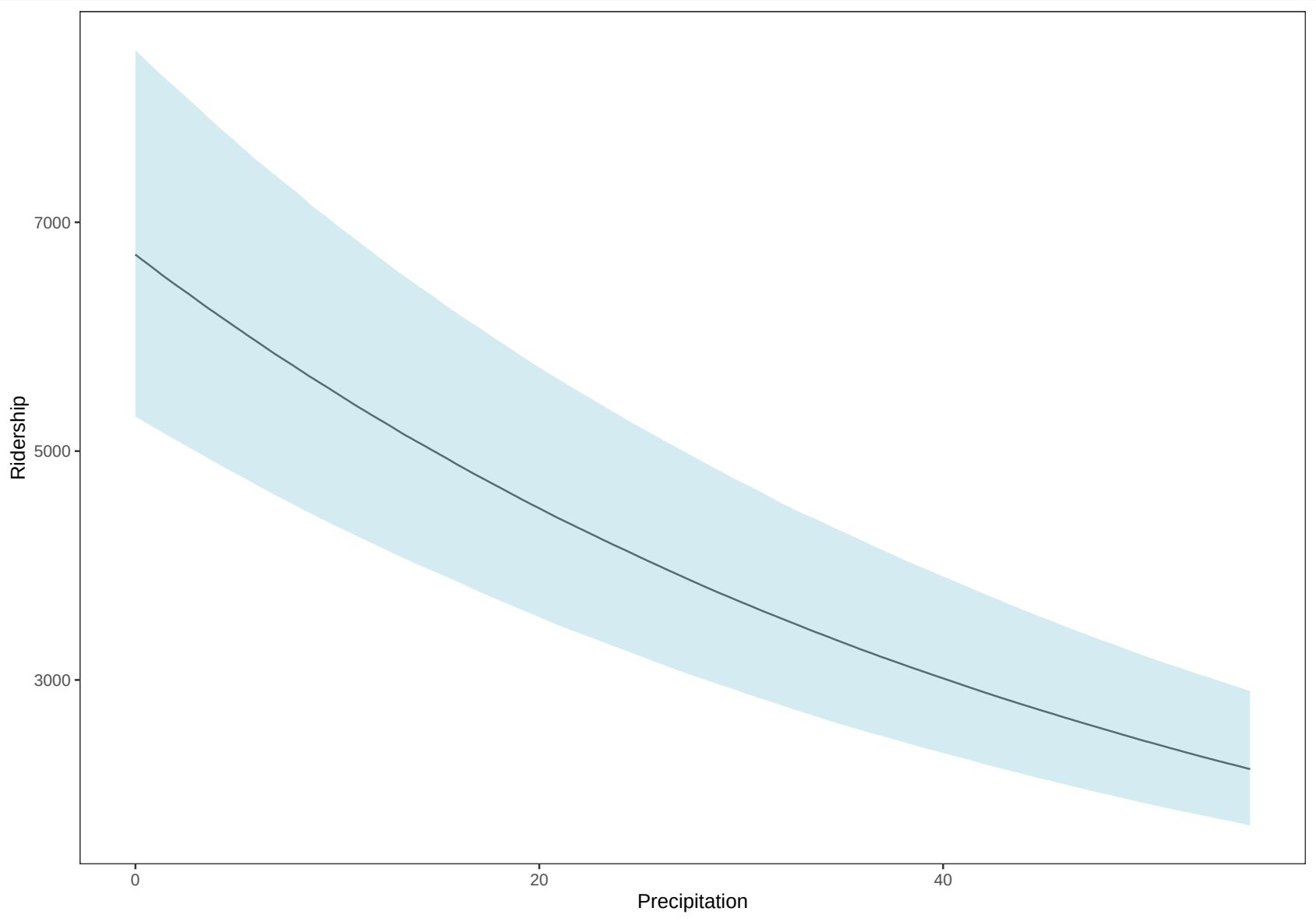
Posterior Predictions



Marginal Effect - Temperature



Marginal Effect - Precipitation



Limitations

- Time endogeneity
- Weather aggregation
- Holidays
- Non-parametric functional terms: gaussian processes!