A Bayesian Hierarchial Model for Maize Yield

Alice Huang

04/04/2023

Abstract

We investigate the effects of water and nitrogen levels on maize yields across different soil types in China. We fit a Bayesian hierarchical model with a Normal likelihood on maize yield, grouping by soil type and investigating the effects of water and nitrogen within each soil type. We show that increasing water and nitrogen results in greater increases in maize yields for coarser soil types that are worse at retaining water and nutrients.

Introduction

Maize, also known as corn in North America, is one of the most commonly consumed crops in the world. It is used to manufacture many different products including animal feed, cornstarch, and corn syrup. Maize yield influences the supply chain and production of many different food products. Maize yield will have significant implications for food security, and business profitability. In this paper, we build a Bayesian hierarchical model on maize yield, with soil type being a level in the hierarchy, and water level and nitrogen level being the main covariates of interest. We are interested in determining how changes in the water level and nitrogen level affect maize yield in different soil types.

Data

The data were downloaded from Data Mendeley at https://data.mendeley.com/datasets/7m3pjrz52x . It is a collection of data from many different studies on maize yield, and was originally featured in a paper by Li et al (2019). In this project, we focus on maize yield data from China, in order to control for correlation due to country's agriculture laws & policies, farming techniques, and technology. Since China is a large country spanning many different types of geographic regions, there was more data available for comparing different soil types. China is one of the biggest producers of maize in the world, so the findings will be significant for global food production.

The data contained measurements of maize yield (tonnes/hectare), water level from both irrigation and rainfall (mm), soil texture, mean annual temperature (degrees Celcius), nitrogen level (kg/hectare), density (plants per hectare), and soil organic content.

There are different types of soil made of varying percentages of sand, silt, and clay. Sand has the largest particle size, followed by silt, and then clay. Sand is the worst at retaining water, followed by silt, then clay. Maize has shallow roots so it is unable to dig deep into the ground to absorb nutrients. Thus the soil's ability to hold onto water and nutrients is especially important for maize growth. See the Appendix for a soil diagram on the different types of soil as classified by their compositions of sand, silt, clay.

Missing Data

First we removed observations with missing values for maize yield. There were 80 observations missing Soil Texture type. We intend to build a Bayesian hierarchical model with soil being a level in the hierarchy. We removed rows with missing Soil Texture from the dataset, as there was no way to categorize their soil type and determine how soil interacts with water and other nutrients.

Some soil types were missing lots of data on Plants Per Hectare (PPH), and did not have a lot of observations to begin with. PPH seems to have been kept fairly constant in other soil groups. Perhaps there were industry conventions for how close plants should be planted in the regions corresponding to those soil groups. Or the researchers kept PPH constant since they were interested in studying other factors. Thus PPH may not be particularly useful as a covariate in a model.

Soil Organic Content (SOC) was missing 168 observations, but there didn't seem to be much correlation between SOC and maize yield, so we ended up not including SOC in the model anyway.

After removing all missing data, there were 709 observations remaining from 52 studies done on 39 locations in China.

Summary Statistics

The median yield was 8.6 t/ha and mean yield was 8.857 t/ha, thus indicating that the distribution of the yield skewed slightly to the right.

Yield	MAT	PPH	SOC	Water	Nitrogen
Min. : 1.000	Min. : 0.500	Min. : 40000	Min. : 4.270	Min. : 173.0	Min. : 0.0
1st Qu.: 6.610	1st Qu.: 7.000	1st Qu.: 56500	1st Qu.: 7.080	1st Qu.: 380.0	1st Qu.: 80.0
Median : 8.600	Median :11.600	Median : 67500	Median : 7.660	Median : 458.2	Median :180.0
Mean : 8.857	Mean : 9.998	Mean : 68073	Mean : 8.691	Mean : 487.4	Mean :188.5
3rd Qu.:10.510	3rd Qu.:12.900	3rd Qu.: 72000	3rd Qu.: 8.990	3rd Qu.: 552.0	3rd Qu.:275.0
Max. :19.200	Max. :21.700	Max. :120000	Max. :21.810	Max. :1112.5	Max. :560.0
NA	NA	NA's :19	NA's :150	NA	NA

We see that there is less data available for the clay and sandy soil texture types. This is to be expected since clay soil represents the finest possible soil and sandy soil is the coarsest possible soil. It is rare to observe such extreme soil types in nature that are suitable for agriculture. Growing maize in clay loam, loam, sandy loam, and silt loam appears to be more common.
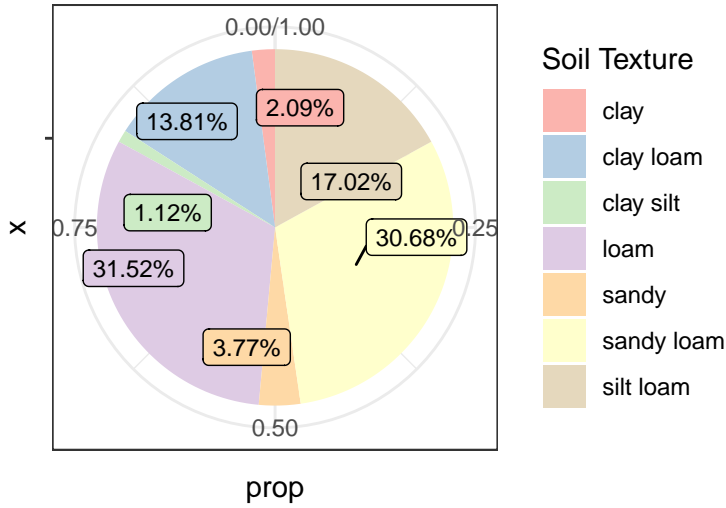
## Relationships Between Covariates and Maize Yield

From the following boxplot, we see that clay silt soil seems to have the greatest median yield. However, this finding should be taken with caution since it is only based on 8 data points. There seems to be a lot of variation in yields between different soil types. For example, we note that the variance of maize yield in the sandy and sandy loam groups seems to be larger than the variance of maize yield in the clay loam soil group. Thus soil type seems to be a significant factor influencing maize yield.
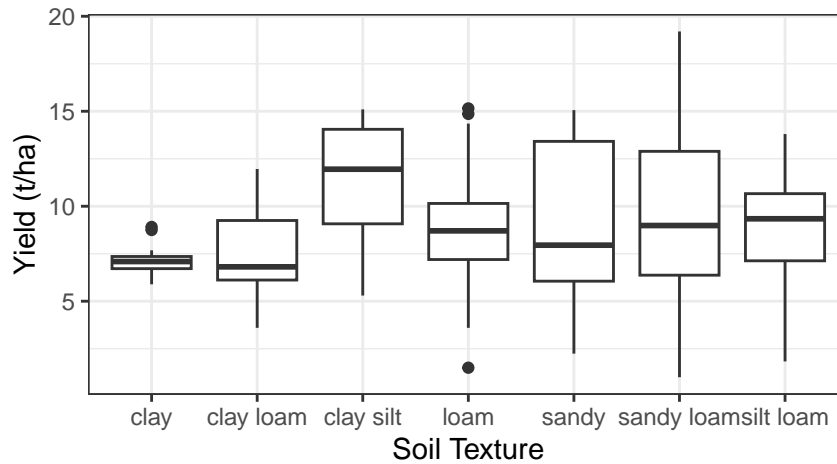
We see that water and soil type jointly influence maize growth. For finer soil types like clay and clay loam, increasing water seemed to either have little effect or slightly negative effect on yield. It is possible that finer soil types retain too much water. Thus when farmers give maize plants in finer soil types more water, after a certain point, they may be overwatering the maize plants. Conversely, for coarser soil types like sandy and sandy loam, increasing water seemed to increase yield. Since coarser soil types are worse at retaining water, increasing water through irrigation may compensate for the soil's decreased ability to retain water, and is thus beneficial for plant growth. However, many areas with sandy soil may also have less rainfall, so it may be more expensive to irrigate plants grown in such areas, in order to reach desired maize yields.

For clay silt soil, there were only 8 observations, and the 8 observations only had 2 unique values of water level: 487 mm and 496 mm. We did not believe there was enough data and variation within data to do any meaningful modeling on yield based on water level for the clay silt soil group. The MCMC chains for

the coefficients corresponding to clay silt soil sometimes had issues with convergence. So we removed the 8 observations on clay silt soil from consideration.
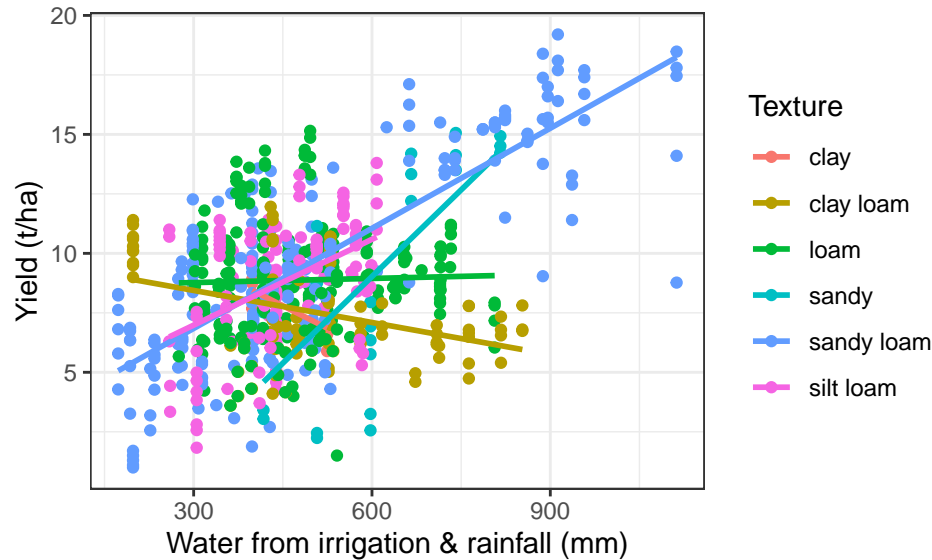

Previous research (Li et al, 2019) suggests that water and Nitrogen are the main factors that influence plant growth. Nitrogen may come from fertilizer added by farmers or it may occur naturally in the soil. It is believed that major improvements in the production of nitrogen-based fertilizers have greatly contributed towards increased crop production over the past century.

For most soil types, as nitrogen increases, yield either doesn't change much, or increases slightly but plateaus after (e.g. loam, clay loam). Perhaps for these soil types, the maize yield was limited by the maize plants' ability to absorb nitrogen. The exception seems to be sandy soil, the coarsest type of soil which does not hold onto water and nutrients well. Increasing Nitrogen levels in sandy soil seemed to increase yields at a higher rate, compared to other types of soil.
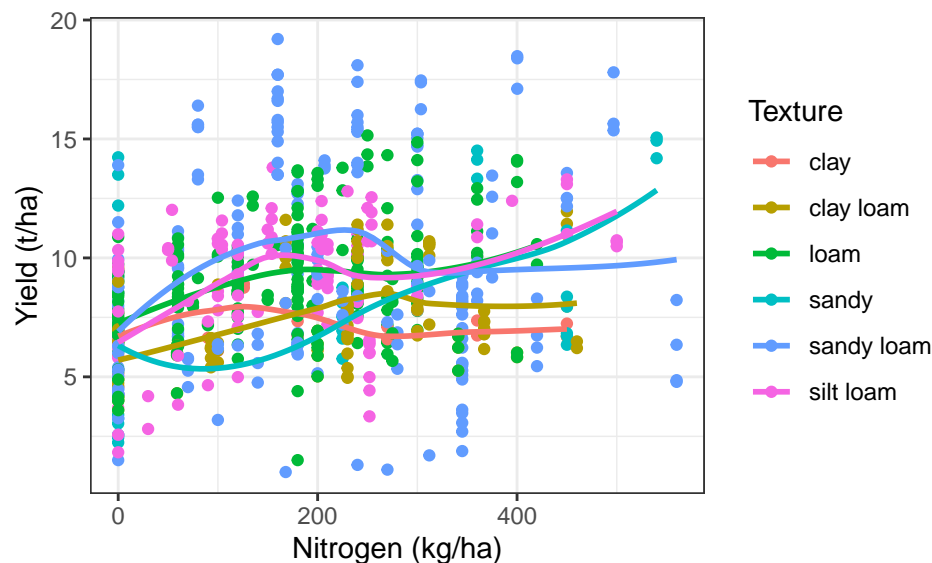

Increasing temperature seemed to have a slight negative effect on yield. Perhaps warmer climates have longer growing seasons, so the soil is more depleted of nutrients during the year. Temperature and soil texture seemed to be correlated due to geographic region. For certain soil types, like sandy loam, there was very little variance in temperature. For some soil types like clay and sandy, there was little variance

in temperature since there were few data points to begin with. Mean annual temperature may not be the best measure of temperature to consider since it includes temperatures throughout the entire year, not just temperatures during growing season. That is why many of the mean annual temperatures recorded in the dataset are lower than one would expect. However, it is difficult to find data on growing season temperatures. Since we intended to stratify based on soil, we decided not to include temperature in the model.
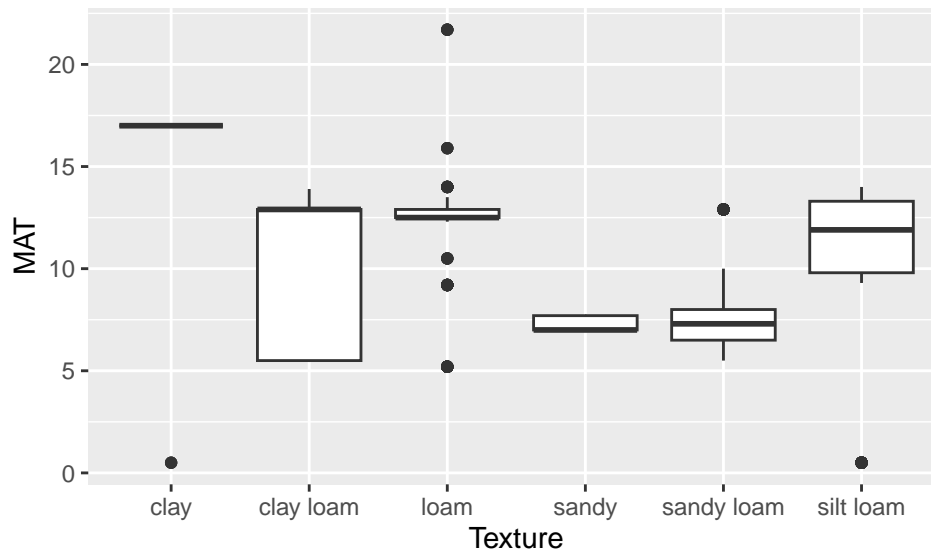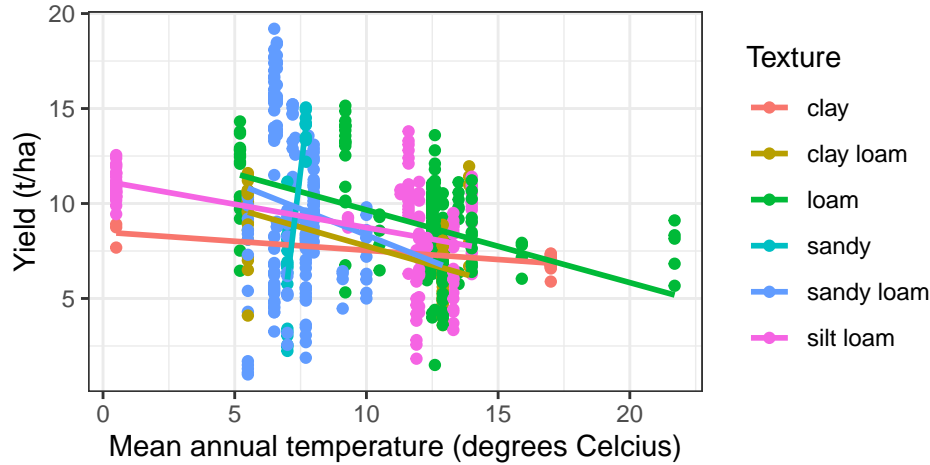

## Determining Appropriate Likelihood for Maize Yield

Maize yield is a continuous variable, and its histogram seems to be fairly symmetric and bell-shaped for the soil types which occupy larger proportions of data. However, the tails seemed to be too large for a normal distribution, so we tried transforming the maize yield so we could use a normal distribution to model it for simplicity. We tried fitting Bayesian hierarchical models on log(maize yield) and maize yield without transformation, however the posterior checks seemed to show worse results than the maize yield under Box Cox transformation. The replicated log maize yields would be far from the observed log maize yields. The other models would give divergent transitions, and sometimes low Bulk Effective Sample Sizes, thus indicating that the posterior means and median estimates could be unreliable. Thus, it seemed that a normal likelihood on maize yield under Box Cox transformation was the best option available.

Histogram of Yield / **Normal Q–Q Plot**
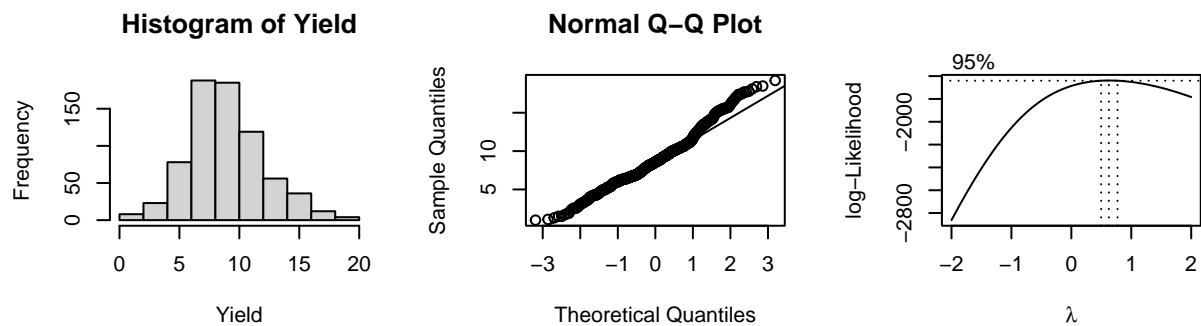
The untransformed yield for loam and sandy loam seemed to have a somewhat symmetric bellshaped distribution. For the clay and sandy soil types, there were a low number of observations, so it couldn't be normally distributed. The yield of maize plants grown in clay loam was skewed right. The yield of maize plants grown in silt loam was skewed left.
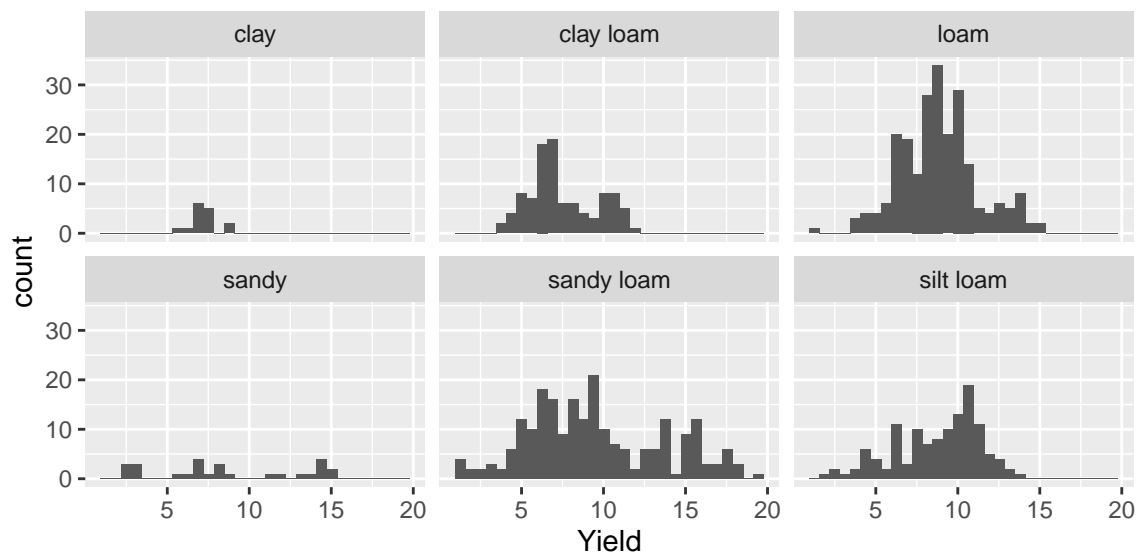
Upon applying the Box Cox transformation to the maize yield (with $\lambda \approx 0.626$), the points on the normal quantile-quantile plot seemed to more closely follow a straight line. However, the tails of the histogram still seemed to be too large, and the Shapiro-Wilk test still indicated evidence against normality.

6

Histogram of Box–Cox Transformed Maize Yield

Normal Q–Q Plot
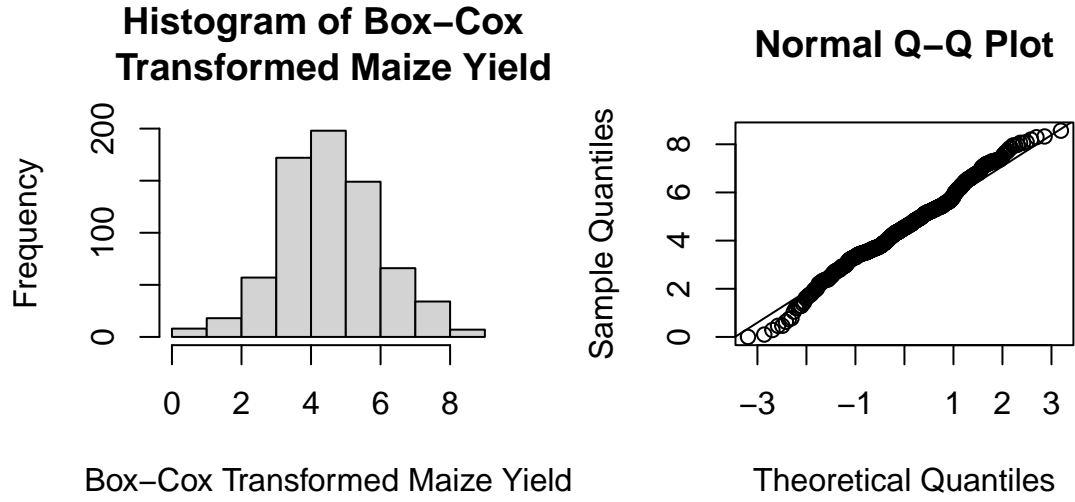
Upon applying the Box Cox transformation to the maize yield, the yield for clay loam and sandy soil types seemed to become more symmetric.
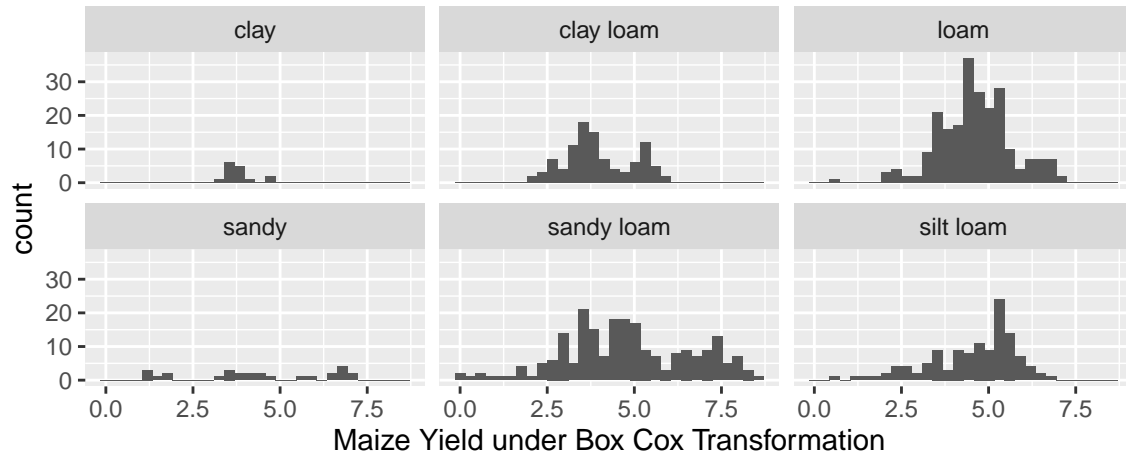
Methods

We fit a Bayesian hierarchical model on the transformed maize yield, grouping by soil texture. Within each soil type, we consider the effects of water and nitrogen levels on maize yield. Previous research and data analysis seems to suggest that water and nitrogen are the main factors that influence maize yield. We stratify on soil type because the maize plants' ability to absorb water and nutrients seems to vary across soil types. Soil also seems to be correlated with temperature, geographic region, and planting density. Due to lack of prior information, we place uninformative standard normal priors on the coefficients, and half normal priors on the variances. We center water and nitrogen for stability. We validate the model using two different posterior predictive checks. The first involves comparing the densities of maize yields simulated under the model with the density of the observed maize yield. The second involves comparing median maize yields simulated under the model with the observed median maize yield.

The model is given as follows:

$$y_j | \alpha_{j[i]}, \beta_{j[i]}, \gamma_{j[i]}, \sigma_y^2 \sim N(\alpha_{j[i]} + \beta_{j[i]} N_i + \gamma_{j[i]} W_i, \sigma_y^2)$$

$$\sigma_y^2 \sim N^+(0, 1)$$

$$\alpha_j \sim N(\mu_\alpha, \sigma_\alpha^2) \qquad \mu_\alpha \sim N(0, 1) \qquad \sigma_\alpha^2 \sim N^+(0, 1)$$

$$\beta_j \sim N(\mu_\beta, \sigma_\beta^2) \qquad \mu_\beta \sim N(0, 1) \qquad \sigma_\beta^2 \sim N^+(0, 1)$$

$$\gamma_j \sim N(\mu_\gamma, \sigma_\gamma^2) \qquad \mu_\gamma \sim N(0, 1) \qquad \sigma_\gamma^2 \sim N^+(0, 1)$$

where y_j is the transformed maize yield (t/ha), W_i is water (mm), N_i is nitrogen (kg/ha), $j[i]$ corresponds to the soil texture j for observation i
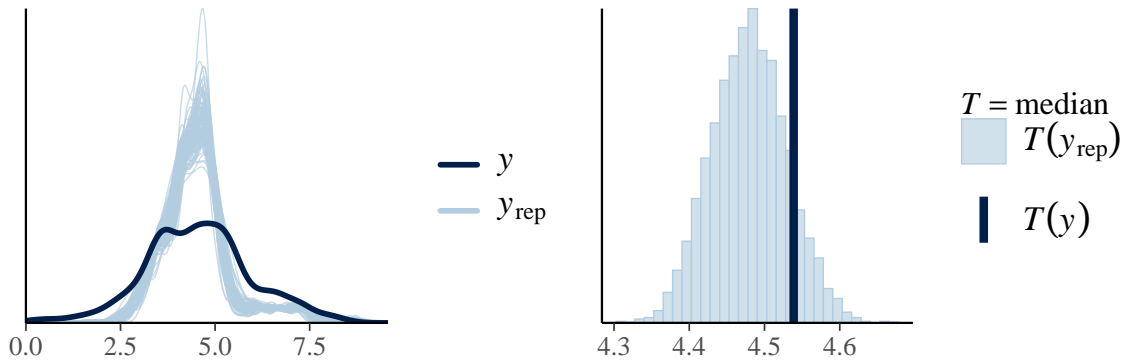
Results

We fit the model in Stan and note that the MCMC chains mix for each of the coefficients. We also see that the distributions for the intercepts and coefficients in each soil type generated by each chain are similar. Thus the model converges. Here, the coefficients 1-6 respectively correspond to clay, clay loam, loam, sandy, sandy loam and silt loam.
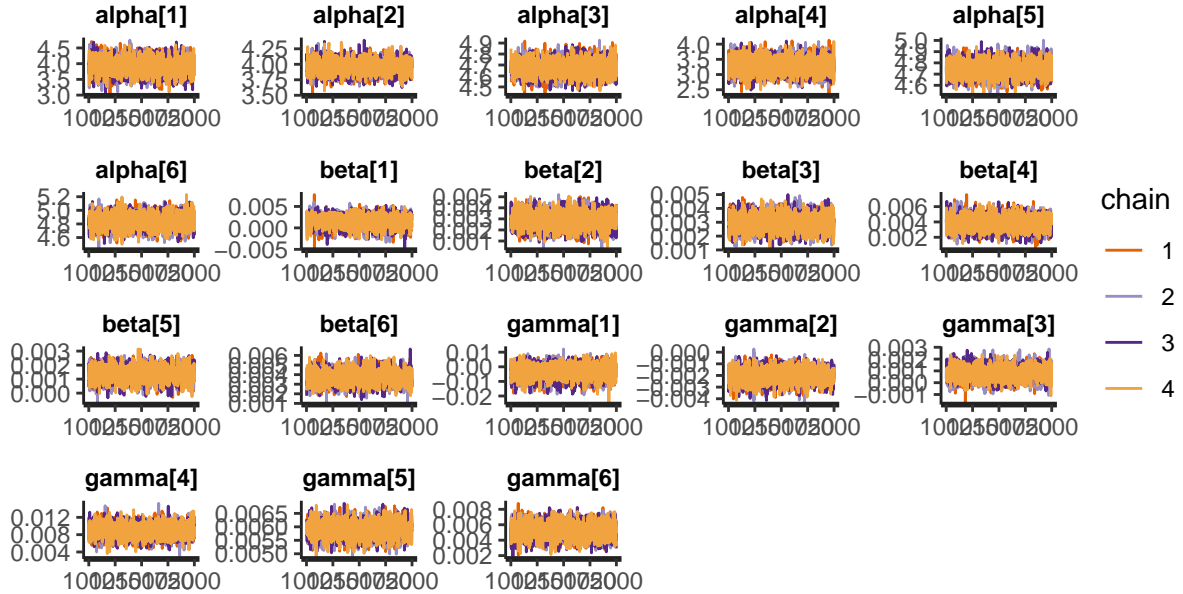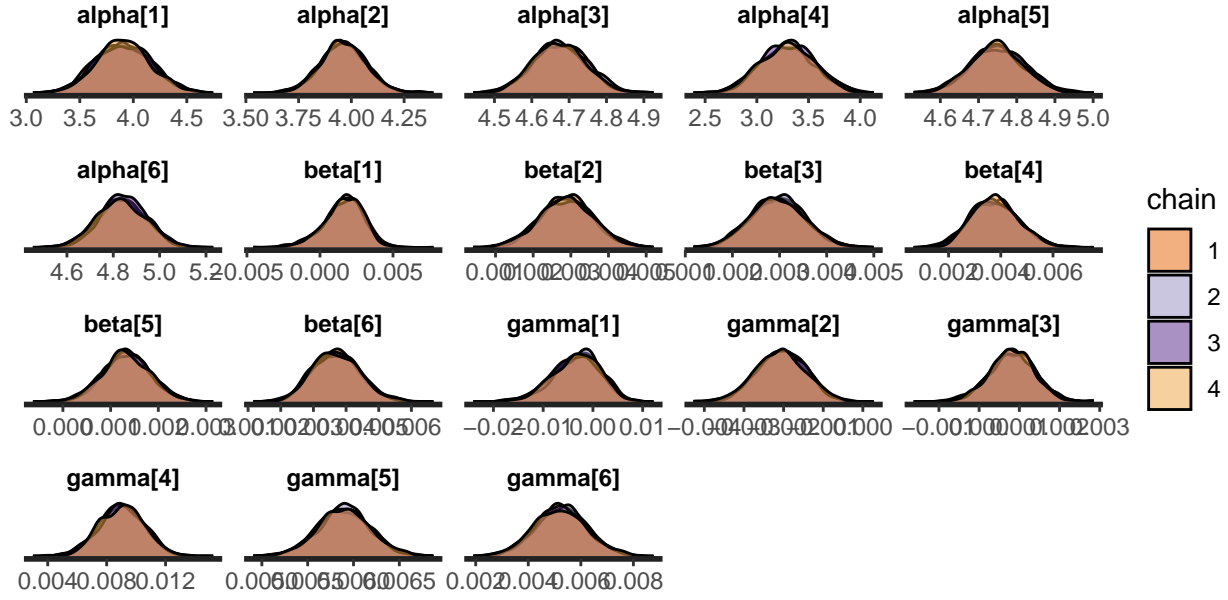
We compare the densities of 100 sampled yields to the actual yield. The densities of the sampled datasets based on the model were far from the true data. Unfortunately the density of the observed maize yield did not fall entirely within the densities of the replicated maize yields. The model puts too much weight on values in the range 4-5, and not enough weight on values in the tails of the original yield's distribution. The observed median yield and replicated median yields were quite close to each other, however. Upon fitting the model on the data for each of the soil types separately, it seemed that the sampled yields and actual yield were closer for the sandy loam and silt loam soil types. However, the model's posterior predictive checks seemed to be worse for loam and clay loam. For each of the soil types, the replicated medians were close to the true median. See Appendix for the posterior checks by soil type.

Here is the summary of the model coefficients, after reversing the transformation on yield estimates and Confidence interval endpoints so that unit increases in coefficients would result in the tabulated changes in yield.

soil	coefficients	mean	se mean	Lower 95% CI	Upper 95% CI
clay	alpha[1]	7.21376	0.004094	6.23143	8.29814
clay loam	alpha[2]	7.32944	0.001679	6.89156	7.79436
loam	alpha[3]	8.88586	0.001099	8.58168	9.19131
sandy	alpha[4]	5.98366	0.004254	5.03503	6.99184

soil	coefficients	mean	se mean	Lower 95% CI	Upper 95% CI
sandy loam	alpha[5]	9.05444	0.001116	8.74378	9.36647
silt loam	alpha[6]	9.25957	0.001589	8.802	9.72501
clay	beta[1]	1.00163	2.1e-05	0.99865	1.00399
clay loam	beta[2]	1.00289	1.2e-05	1.00149	1.00431
loam	beta[3]	1.00296	9e-06	1.00187	1.00409
sandy	beta[4]	1.00378	1.5e-05	1.00207	1.00563
sandy loam	beta[5]	1.00134	8e-06	1.00035	1.00234
silt loam	beta[6]	1.00371	1.3e-05	1.00239	1.0052
clay	gamma[1]	0.99664	8.9e-05	0.98661	1.00496
clay loam	gamma[2]	0.99794	1e-05	0.99672	0.99911
loam	gamma[3]	1.00083	9e-06	0.99968	1.00193
sandy	gamma[4]	1.00903	2.3e-05	1.00606	1.01193
sandy loam	gamma[5]	1.0059	4e-06	1.00534	1.00647
silt loam	gamma[6]	1.00525	1.5e-05	1.00332	1.0072

For clay soil, increasing the water by 1 mm was expected to increase maize yield by 1.00163 tonnes/hectare (CI = (0.99865, 1.00399)). Increasing the nitrogen level by 1 kg/ha was expected to increase the maize yield by 0.99664 tonnes/hectare (CI = (0.98661, 1.00496)).

For clay loam soil, increasing the water by 1 mm was expected to increase maize yield by 1.00289 tonnes/hectare (CI = (1.00149, 1.00431)). Increasing the nitrogen level by 1 kg/ha was expected to increase the maize yield by 0.99794 tonnes/hectare(CI = (0.99672, 0.99911)).

For loam soil, increasing the water by 1 mm was expected to increase maize yield by 1.00296 tonnes/hectare (CI = (1.00187, 1.00409)). Increasing the nitrogen level by 1 kg/ha was expected to increase the maize yield by 1.00083 tonnes/hectare (CI = (0.99968, 1.00193)).

For sandy soil, increasing the water by 1 mm was expected to increase maize yield by 1.00378 tonnes/hectare (CI = (1.00207, 1.00563)). Increasing the nitrogen level by 1 kg/ha was expected to increase the maize yield by 1.00903 tonnes/hectare (CI = (1.00606, 1.01193)).

For sandy loam soil, increasing the water by 1 mm was expected to increase maize yield by 1.00134 tonnes/hectare (CI = (1.00035, 1.00234)). Increasing the nitrogen level by 1 kg/ha was expected to increase the maize yield by 1.0059 tonnes/hectare (CI = (1.00534, 1.00647)).

For silt loam soil, increasing the water by 1 mm was expected to increase maize yield by 1.00371 tonnes/hectare (CI = (1.00239, 1.0052)). Increasing the nitrogen level by 1 kg/ha was expected to increase the maize yield by 1.00525 tonnes/hectare (CI = (1.00332, 1.0072)).

Discussion

Increasing water on maize crops seemed to yield greater increases in maize yields for sandy, loam, and silt loam soil. This result is to be expected, because these types of soil have coarser particles and are worse at retaining water. Thus increasing water compensates for the soil's reduced ability to hold water. Increasing water doesn't increase maize yield as much for soil with finer particles and good water retention, like clay.

For the coarser soil types which are worse at retaining nutrients, such as sandy, sandy loam, and silt loam soils, the model indicates that increasing the nitrogen through fertilization will be associated with greater increases in maize yields. This result is also consistent with previous findings. It should be noted that it is unknown whether the farmers employed mulching during these studies, which can also help keep shallow parts of the soil moist, and may possibly influence the soil's ability to hold onto nutrients.

Even though the model coefficients may appear small and differences between different soil types' coefficients may be very small, the yield is measured in tonnes per hectare, and multiplying the effects over large regions

could result in significant yields overall. However, caution should be taken with applying these results in practice, as local availability of water and nutrients in nature, among other factors, must be considered. Areas with sandy soil may have less rainfall to begin with, so even though adding more water to maize plants in sandy soil likely results in greater increases in yield, it may end up becoming costly for farmers. Growing maize in sandy soil and achieving high yields is possible, however farmers must ensure that they optimize their supply chain and apply water and nitrogen fertilizer efficiently.

Future Work

A more thorough literature review needs to be conducted to see if previous researchers have created Bayesian models for maize yield using more informative priors than the standard normal priors in this model. A model updated with more informative priors may fit better, and the simulated maize yields and observed maize yields may be closer to each other. Due to limited time, the simple equation $E[y_j|\alpha_{j[i]}, \beta_{j[i]}, \gamma_{j[i]}, \sigma_y^2] = \alpha_{j[i]} + \beta_{j[i]}N_i + \gamma_{j[i]}W_i$ was used to model the mean maize yield at each soil level. However, it may be worthwhile to investigate transformations for the nitrogen covariate at each soil level, especially given that some soil types' maize yields did not change linearly in response to nitrogen levels.

In addition to modeling the effects of water and nitrogen, it would be interesting to model other variables that influence maize growth. It would be interesting to aggregate more data on other nutrients from other studies on maize yield. It would also be interesting to investigate the effect of organic vs chemical fertilizers on Nitrogen intake and yield. In addition, it would also be interesting to investigate the effect of different genetic modifications and maize varieties and how that affects maize's ability to efficiently use water, nitrogen, and other nutrients in different soil types. There would also be important practical implications of investigating the effects of different farming techniques like crop rotation, mulching, and soil tilling. There are many more factors influencing maize yield which have not been discussed, and are worth studying.

Overall, there still remains a lot of work that can be done to improve this simple model.

Appendix

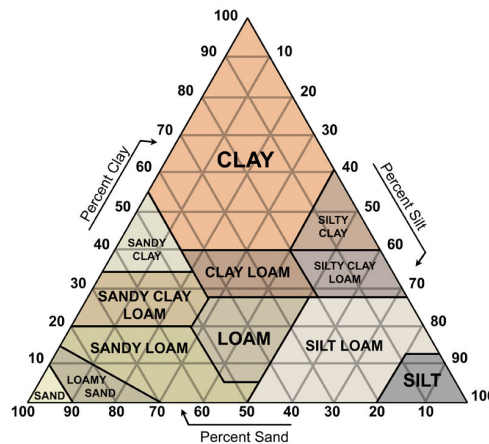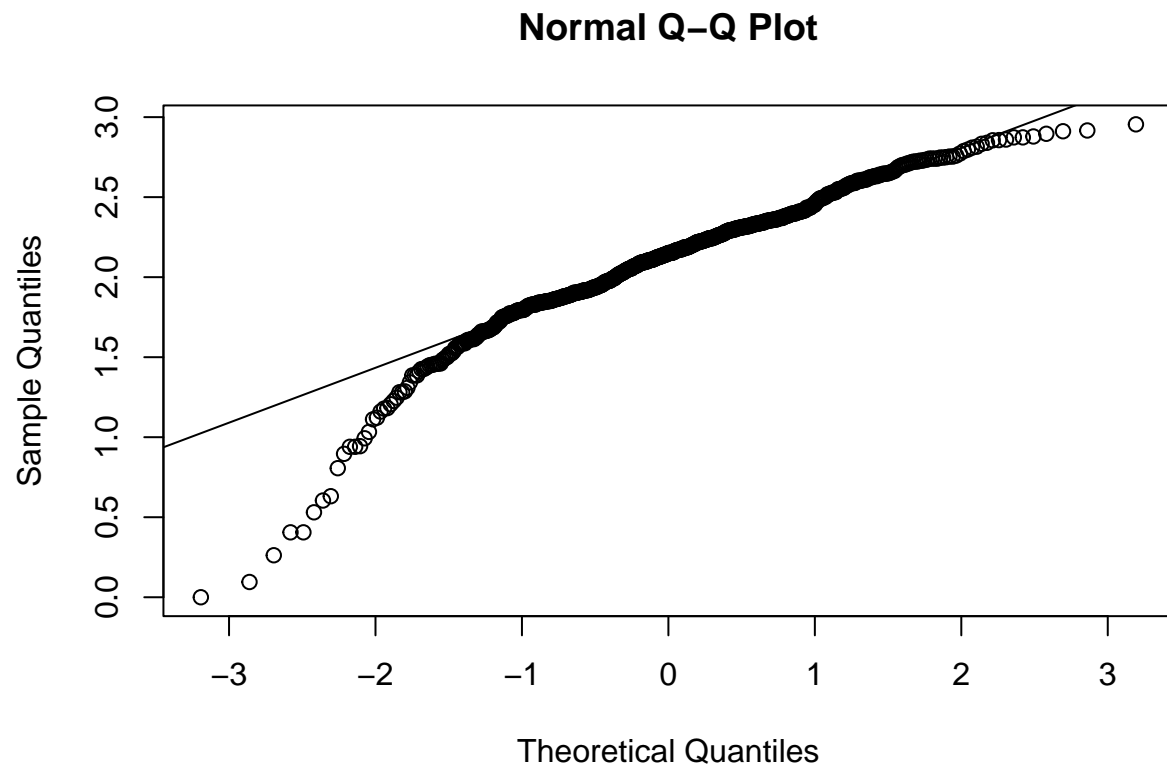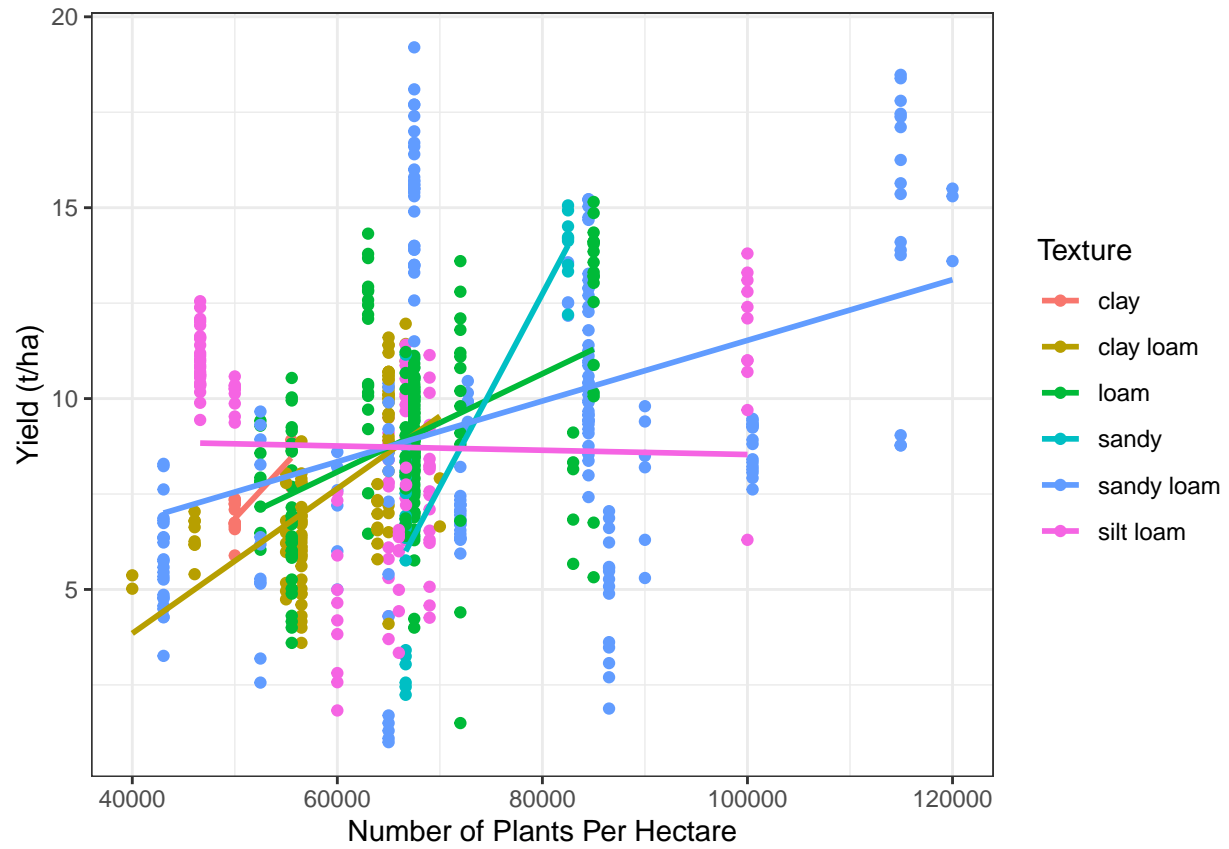Here is a diagram indicating the soil types by percentage of clay, silt and sand.


Figure 1: Soil Types

Log(yield) did not seem to follow a normal distribution.

Normal Q–Q Plot

For most soil types except silt loam, increasing plant density seemed to increase yield. Perhaps farmers already knew how dense to plant their seeds before plants start competing with each other
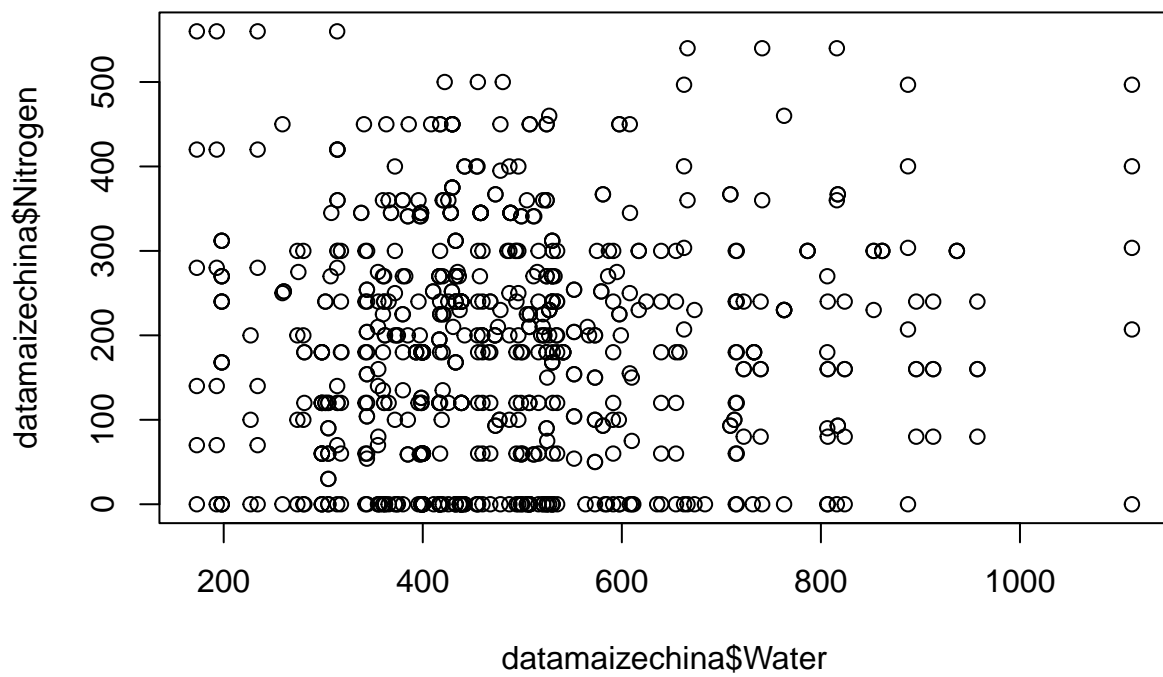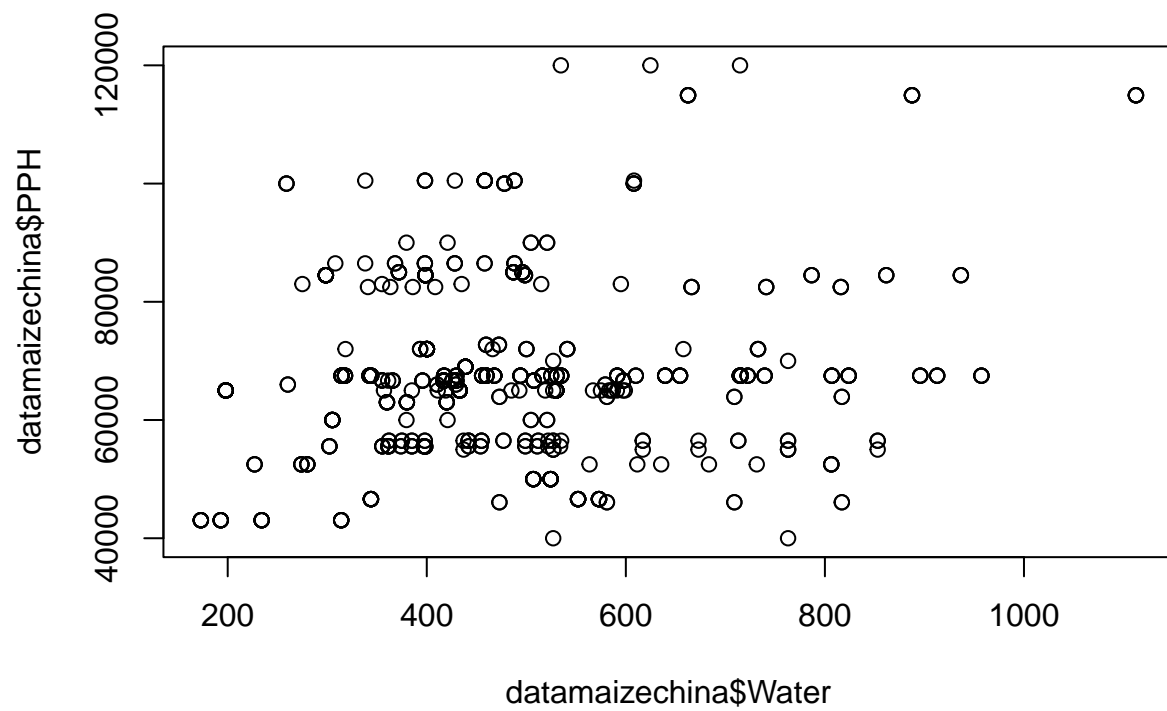
Soil Organic Carbon did not seem particularly significant for most soil types. Within each soil type, there was also little variation in values of Soil Organic Carbon.
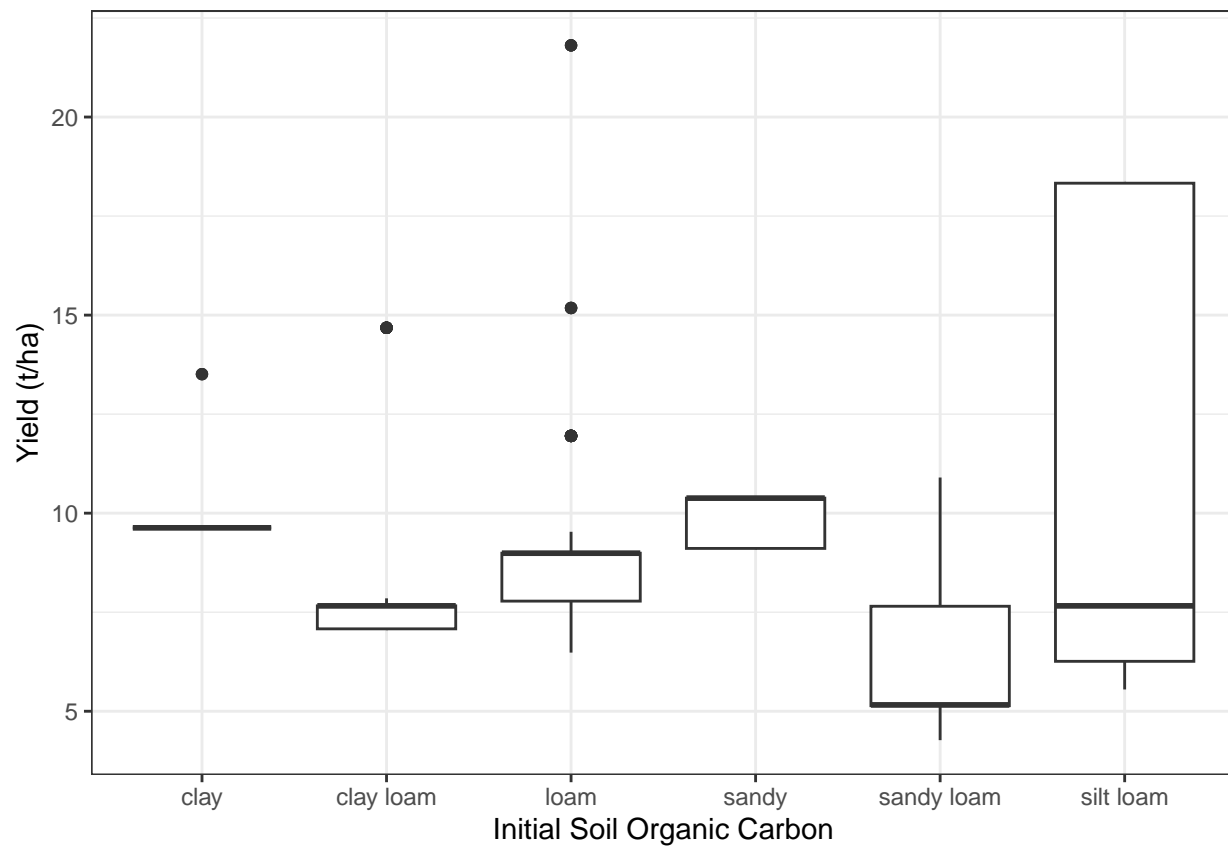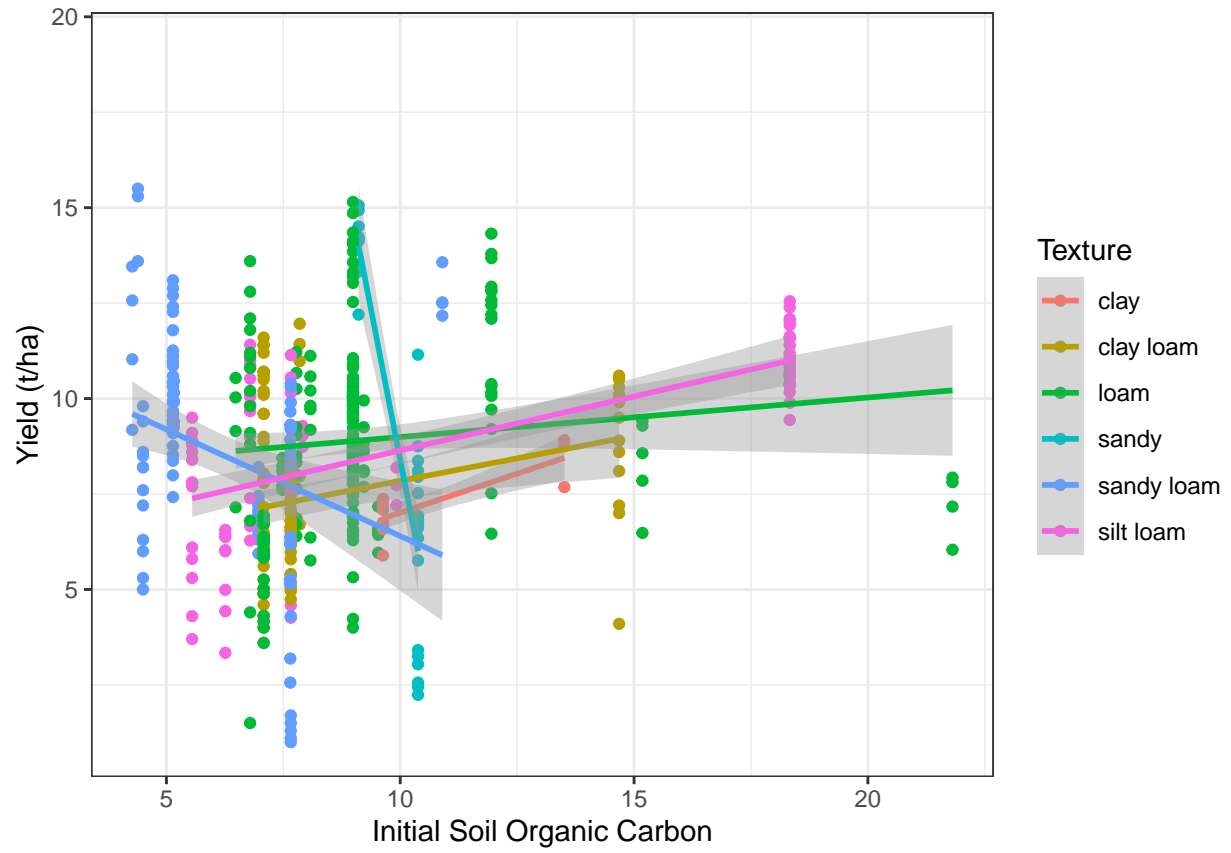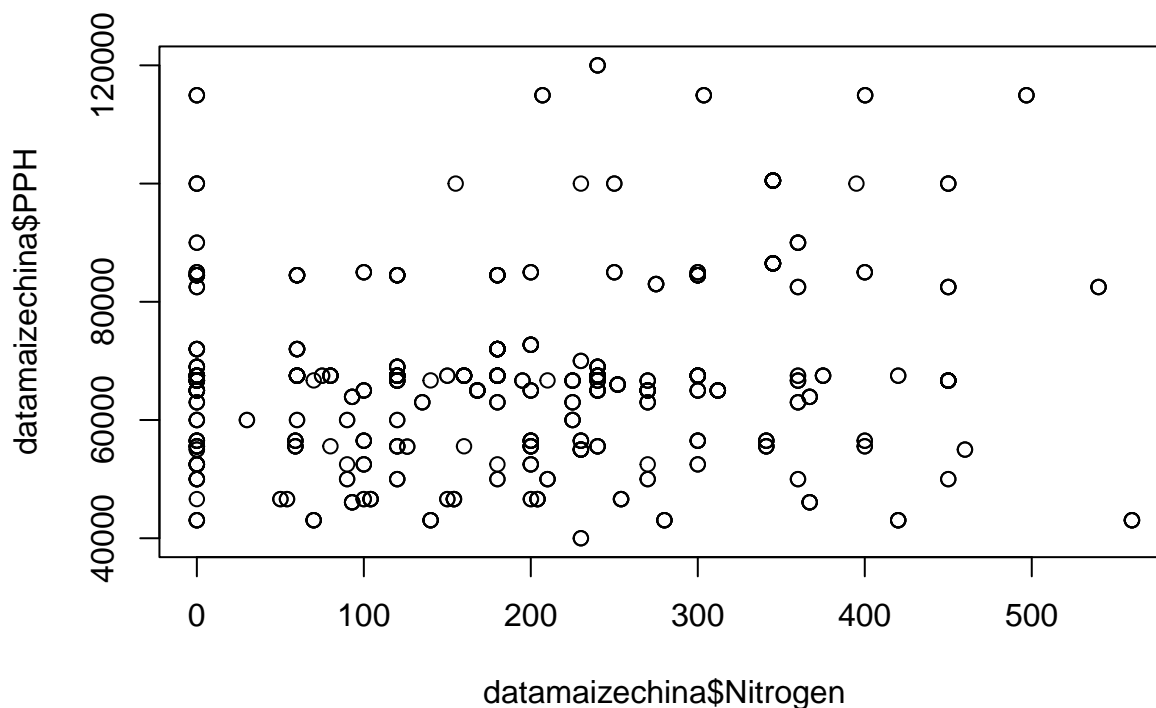
```
## [1] 0.04502987
```

Posterior Checks for Loam Soil

```
##
## SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 1).
## Chain 1:
## Chain 1: Gradient evaluation took 5.6e-05 seconds
## Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 0.56 seconds.
## Chain 1: Adjust your expectations accordingly!
## Chain 1:
## Chain 1:
## Chain 1: Iteration:    1 / 2000 [  0%]  (Warmup)
## Chain 1: Iteration:  200 / 2000 [ 10%]  (Warmup)
## Chain 1: Iteration:  400 / 2000 [ 20%]  (Warmup)
## Chain 1: Iteration:  600 / 2000 [ 30%]  (Warmup)
## Chain 1: Iteration:  800 / 2000 [ 40%]  (Warmup)
## Chain 1: Iteration: 1000 / 2000 [ 50%]  (Warmup)
## Chain 1: Iteration: 1001 / 2000 [ 50%]  (Sampling)
## Chain 1: Iteration: 1200 / 2000 [ 60%]  (Sampling)
## Chain 1: Iteration: 1400 / 2000 [ 70%]  (Sampling)
## Chain 1: Iteration: 1600 / 2000 [ 80%]  (Sampling)
## Chain 1: Iteration: 1800 / 2000 [ 90%]  (Sampling)
## Chain 1: Iteration: 2000 / 2000 [100%]  (Sampling)
## Chain 1:
## Chain 1:  Elapsed Time: 10.635 seconds (Warm-up)
## Chain 1:                 2.959 seconds (Sampling)
```

```
## Chain 1:                    13.594 seconds (Total)
## Chain 1:
##
## SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 2).
## Chain 2:
## Chain 2: Gradient evaluation took 3.8e-05 seconds
## Chain 2: 1000 transitions using 10 leapfrog steps per transition would take 0.38 seconds.
## Chain 2: Adjust your expectations accordingly!
## Chain 2:
## Chain 2:
## Chain 2: Iteration:    1 / 2000 [  0%]  (Warmup)
## Chain 2: Iteration:  200 / 2000 [ 10%]  (Warmup)
## Chain 2: Iteration:  400 / 2000 [ 20%]  (Warmup)
## Chain 2: Iteration:  600 / 2000 [ 30%]  (Warmup)
## Chain 2: Iteration:  800 / 2000 [ 40%]  (Warmup)
## Chain 2: Iteration: 1000 / 2000 [ 50%]  (Warmup)
## Chain 2: Iteration: 1001 / 2000 [ 50%]  (Sampling)
## Chain 2: Iteration: 1200 / 2000 [ 60%]  (Sampling)
## Chain 2: Iteration: 1400 / 2000 [ 70%]  (Sampling)
## Chain 2: Iteration: 1600 / 2000 [ 80%]  (Sampling)
## Chain 2: Iteration: 1800 / 2000 [ 90%]  (Sampling)
## Chain 2: Iteration: 2000 / 2000 [100%]  (Sampling)
## Chain 2:
## Chain 2:  Elapsed Time: 9.209 seconds (Warm-up)
## Chain 2:                2.844 seconds (Sampling)
## Chain 2:                12.053 seconds (Total)
## Chain 2:
##
## SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 3).
## Chain 3:
## Chain 3: Gradient evaluation took 6.8e-05 seconds
## Chain 3: 1000 transitions using 10 leapfrog steps per transition would take 0.68 seconds.
## Chain 3: Adjust your expectations accordingly!
## Chain 3:
## Chain 3:
## Chain 3: Iteration:    1 / 2000 [  0%]  (Warmup)
## Chain 3: Iteration:  200 / 2000 [ 10%]  (Warmup)
## Chain 3: Iteration:  400 / 2000 [ 20%]  (Warmup)
## Chain 3: Iteration:  600 / 2000 [ 30%]  (Warmup)
## Chain 3: Iteration:  800 / 2000 [ 40%]  (Warmup)
## Chain 3: Iteration: 1000 / 2000 [ 50%]  (Warmup)
## Chain 3: Iteration: 1001 / 2000 [ 50%]  (Sampling)
## Chain 3: Iteration: 1200 / 2000 [ 60%]  (Sampling)
## Chain 3: Iteration: 1400 / 2000 [ 70%]  (Sampling)
## Chain 3: Iteration: 1600 / 2000 [ 80%]  (Sampling)
## Chain 3: Iteration: 1800 / 2000 [ 90%]  (Sampling)
## Chain 3: Iteration: 2000 / 2000 [100%]  (Sampling)
## Chain 3:
## Chain 3:  Elapsed Time: 8.454 seconds (Warm-up)
## Chain 3:                2.568 seconds (Sampling)
## Chain 3:                11.022 seconds (Total)
## Chain 3:
##
## SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 4).
```
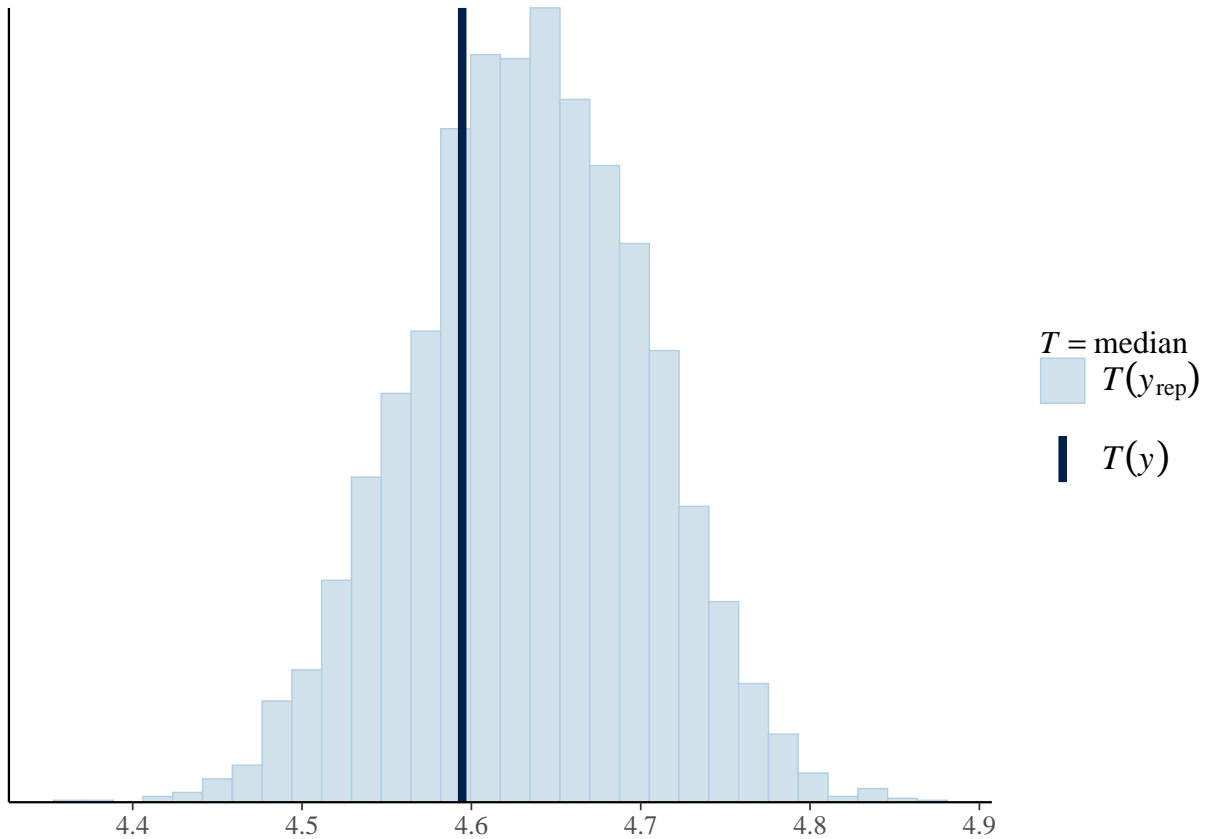
```
## Chain 4:
## Chain 4: Gradient evaluation took 3.8e-05 seconds
## Chain 4: 1000 transitions using 10 leapfrog steps per transition would take 0.38 seconds.
## Chain 4: Adjust your expectations accordingly!
## Chain 4:
## Chain 4:
## Chain 4: Iteration:    1 / 2000 [  0%]  (Warmup)
## Chain 4: Iteration:  200 / 2000 [ 10%]  (Warmup)
## Chain 4: Iteration:  400 / 2000 [ 20%]  (Warmup)
## Chain 4: Iteration:  600 / 2000 [ 30%]  (Warmup)
## Chain 4: Iteration:  800 / 2000 [ 40%]  (Warmup)
## Chain 4: Iteration: 1000 / 2000 [ 50%]  (Warmup)
## Chain 4: Iteration: 1001 / 2000 [ 50%]  (Sampling)
## Chain 4: Iteration: 1200 / 2000 [ 60%]  (Sampling)
## Chain 4: Iteration: 1400 / 2000 [ 70%]  (Sampling)
## Chain 4: Iteration: 1600 / 2000 [ 80%]  (Sampling)
## Chain 4: Iteration: 1800 / 2000 [ 90%]  (Sampling)
## Chain 4: Iteration: 2000 / 2000 [100%]  (Sampling)
## Chain 4:
## Chain 4:  Elapsed Time: 8.691 seconds (Warm-up)
## Chain 4:                2.812 seconds (Sampling)
## Chain 4:                11.503 seconds (Total)
## Chain 4:
```
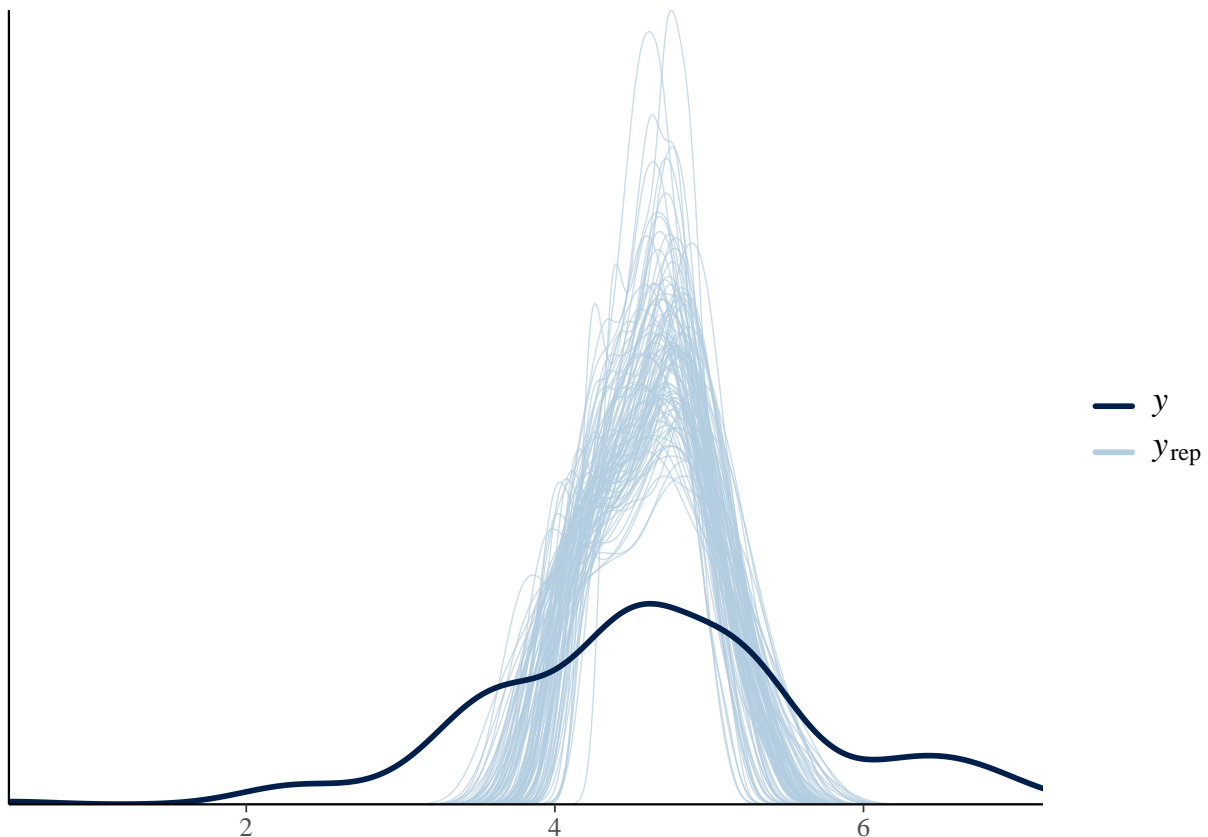
Posterior Checks for Clay Loam Soil

```
##
## SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 1).
## Chain 1:
## Chain 1: Gradient evaluation took 4e-05 seconds
## Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 0.4 seconds.
## Chain 1: Adjust your expectations accordingly!
## Chain 1:
## Chain 1:
## Chain 1: Iteration:    1 / 2000 [  0%]  (Warmup)
## Chain 1: Iteration:  200 / 2000 [ 10%]  (Warmup)
## Chain 1: Iteration:  400 / 2000 [ 20%]  (Warmup)
## Chain 1: Iteration:  600 / 2000 [ 30%]  (Warmup)
## Chain 1: Iteration:  800 / 2000 [ 40%]  (Warmup)
## Chain 1: Iteration: 1000 / 2000 [ 50%]  (Warmup)
## Chain 1: Iteration: 1001 / 2000 [ 50%]  (Sampling)
## Chain 1: Iteration: 1200 / 2000 [ 60%]  (Sampling)
## Chain 1: Iteration: 1400 / 2000 [ 70%]  (Sampling)
## Chain 1: Iteration: 1600 / 2000 [ 80%]  (Sampling)
## Chain 1: Iteration: 1800 / 2000 [ 90%]  (Sampling)
## Chain 1: Iteration: 2000 / 2000 [100%]  (Sampling)
## Chain 1:
## Chain 1:  Elapsed Time: 4.4 seconds (Warm-up)
## Chain 1:                1.625 seconds (Sampling)
```

```
## Chain 1:                      6.025 seconds (Total)
## Chain 1:
##
## SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 2).
## Chain 2:
## Chain 2: Gradient evaluation took 2.9e-05 seconds
## Chain 2: 1000 transitions using 10 leapfrog steps per transition would take 0.29 seconds.
## Chain 2: Adjust your expectations accordingly!
## Chain 2:
## Chain 2:
## Chain 2: Iteration:    1 / 2000 [  0%]  (Warmup)
## Chain 2: Iteration:  200 / 2000 [ 10%]  (Warmup)
## Chain 2: Iteration:  400 / 2000 [ 20%]  (Warmup)
## Chain 2: Iteration:  600 / 2000 [ 30%]  (Warmup)
## Chain 2: Iteration:  800 / 2000 [ 40%]  (Warmup)
## Chain 2: Iteration: 1000 / 2000 [ 50%]  (Warmup)
## Chain 2: Iteration: 1001 / 2000 [ 50%]  (Sampling)
## Chain 2: Iteration: 1200 / 2000 [ 60%]  (Sampling)
## Chain 2: Iteration: 1400 / 2000 [ 70%]  (Sampling)
## Chain 2: Iteration: 1600 / 2000 [ 80%]  (Sampling)
## Chain 2: Iteration: 1800 / 2000 [ 90%]  (Sampling)
## Chain 2: Iteration: 2000 / 2000 [100%]  (Sampling)
## Chain 2:
## Chain 2:  Elapsed Time: 4.389 seconds (Warm-up)
## Chain 2:                1.689 seconds (Sampling)
## Chain 2:                6.078 seconds (Total)
## Chain 2:
##
## SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 3).
## Chain 3:
## Chain 3: Gradient evaluation took 2.6e-05 seconds
## Chain 3: 1000 transitions using 10 leapfrog steps per transition would take 0.26 seconds.
## Chain 3: Adjust your expectations accordingly!
## Chain 3:
## Chain 3:
## Chain 3: Iteration:    1 / 2000 [  0%]  (Warmup)
## Chain 3: Iteration:  200 / 2000 [ 10%]  (Warmup)
## Chain 3: Iteration:  400 / 2000 [ 20%]  (Warmup)
## Chain 3: Iteration:  600 / 2000 [ 30%]  (Warmup)
## Chain 3: Iteration:  800 / 2000 [ 40%]  (Warmup)
## Chain 3: Iteration: 1000 / 2000 [ 50%]  (Warmup)
## Chain 3: Iteration: 1001 / 2000 [ 50%]  (Sampling)
## Chain 3: Iteration: 1200 / 2000 [ 60%]  (Sampling)
## Chain 3: Iteration: 1400 / 2000 [ 70%]  (Sampling)
## Chain 3: Iteration: 1600 / 2000 [ 80%]  (Sampling)
## Chain 3: Iteration: 1800 / 2000 [ 90%]  (Sampling)
## Chain 3: Iteration: 2000 / 2000 [100%]  (Sampling)
## Chain 3:
## Chain 3:  Elapsed Time: 4.276 seconds (Warm-up)
## Chain 3:                1.535 seconds (Sampling)
## Chain 3:                5.811 seconds (Total)
## Chain 3:
##
## SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 4).
```
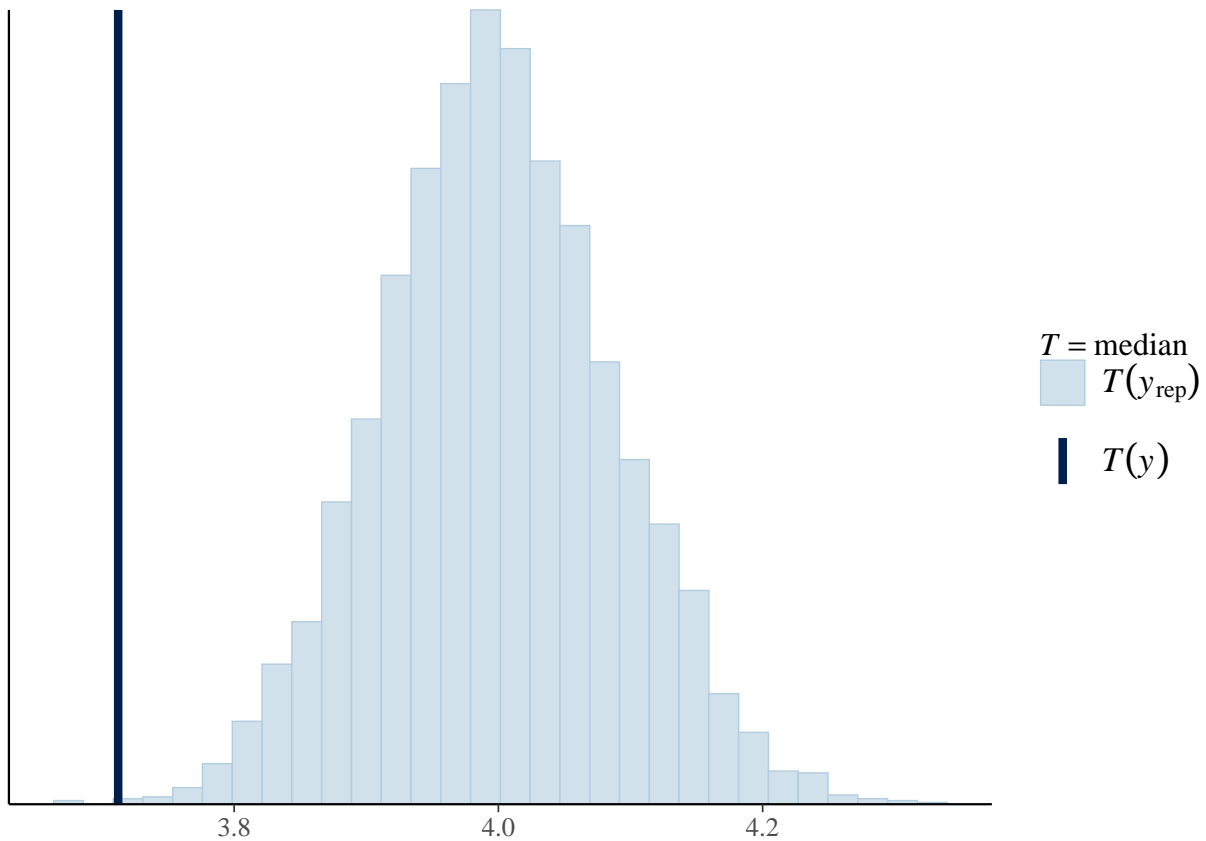
```
## Chain 4:
## Chain 4: Gradient evaluation took 4.3e-05 seconds
## Chain 4: 1000 transitions using 10 leapfrog steps per transition would take 0.43 seconds.
## Chain 4: Adjust your expectations accordingly!
## Chain 4:
## Chain 4:
## Chain 4: Iteration:    1 / 2000 [  0%]  (Warmup)
## Chain 4: Iteration:  200 / 2000 [ 10%]  (Warmup)
## Chain 4: Iteration:  400 / 2000 [ 20%]  (Warmup)
## Chain 4: Iteration:  600 / 2000 [ 30%]  (Warmup)
## Chain 4: Iteration:  800 / 2000 [ 40%]  (Warmup)
## Chain 4: Iteration: 1000 / 2000 [ 50%]  (Warmup)
## Chain 4: Iteration: 1001 / 2000 [ 50%]  (Sampling)
## Chain 4: Iteration: 1200 / 2000 [ 60%]  (Sampling)
## Chain 4: Iteration: 1400 / 2000 [ 70%]  (Sampling)
## Chain 4: Iteration: 1600 / 2000 [ 80%]  (Sampling)
## Chain 4: Iteration: 1800 / 2000 [ 90%]  (Sampling)
## Chain 4: Iteration: 2000 / 2000 [100%]  (Sampling)
## Chain 4:
## Chain 4:  Elapsed Time: 5.224 seconds (Warm-up)
## Chain 4:                2.814 seconds (Sampling)
## Chain 4:                8.038 seconds (Total)
## Chain 4:
```

$T = \text{median}$
$T(y_{\text{rep}})$

$T(y)$
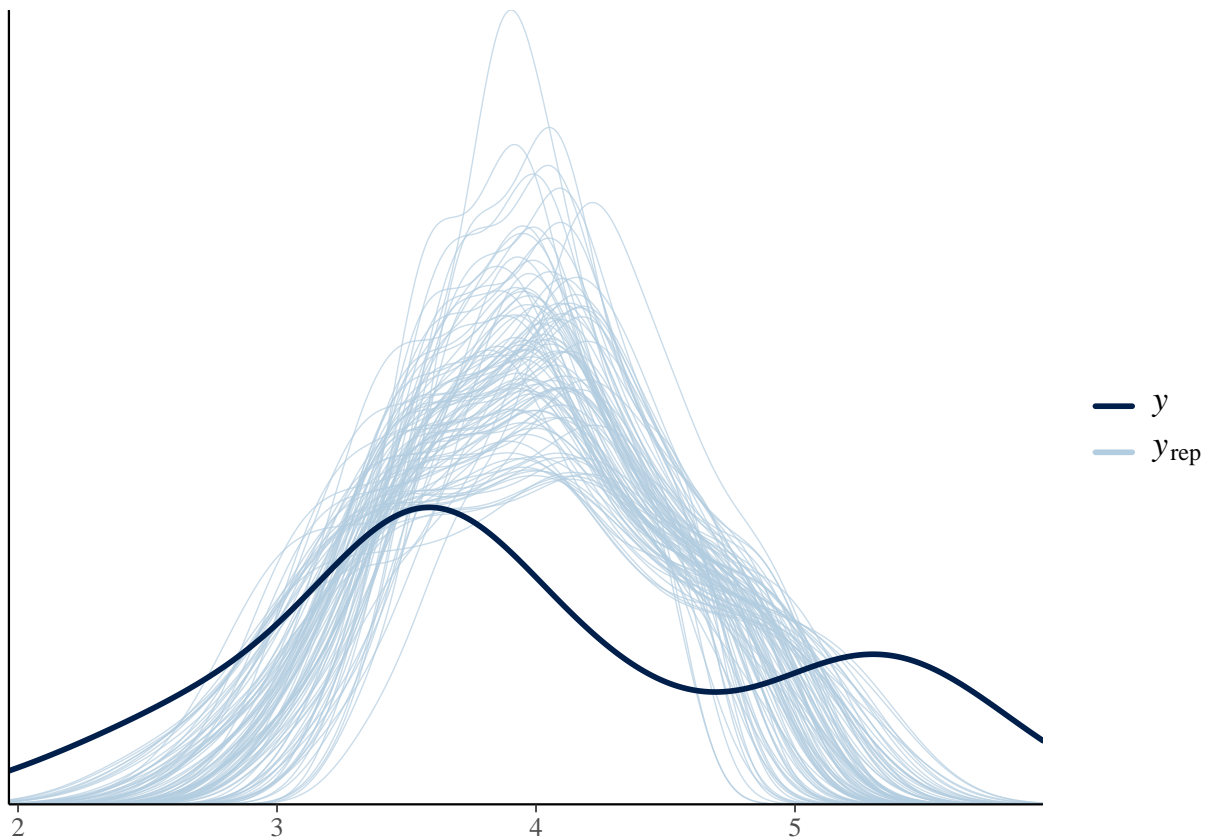
Posterior Checks for Silt Loam Soil

```
##
## SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 1).
## Chain 1:
## Chain 1: Gradient evaluation took 3.2e-05 seconds
## Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 0.32 seconds.
## Chain 1: Adjust your expectations accordingly!
## Chain 1:
## Chain 1:
## Chain 1: Iteration:    1 / 2000 [  0%]  (Warmup)
## Chain 1: Iteration:  200 / 2000 [ 10%]  (Warmup)
## Chain 1: Iteration:  400 / 2000 [ 20%]  (Warmup)
## Chain 1: Iteration:  600 / 2000 [ 30%]  (Warmup)
## Chain 1: Iteration:  800 / 2000 [ 40%]  (Warmup)
## Chain 1: Iteration: 1000 / 2000 [ 50%]  (Warmup)
## Chain 1: Iteration: 1001 / 2000 [ 50%]  (Sampling)
## Chain 1: Iteration: 1200 / 2000 [ 60%]  (Sampling)
## Chain 1: Iteration: 1400 / 2000 [ 70%]  (Sampling)
## Chain 1: Iteration: 1600 / 2000 [ 80%]  (Sampling)
## Chain 1: Iteration: 1800 / 2000 [ 90%]  (Sampling)
## Chain 1: Iteration: 2000 / 2000 [100%]  (Sampling)
## Chain 1:
## Chain 1:  Elapsed Time: 4.944 seconds (Warm-up)
## Chain 1:                1.534 seconds (Sampling)
```

```
## Chain 1:                    6.478 seconds (Total)
## Chain 1:
##
## SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 2).
## Chain 2:
## Chain 2: Gradient evaluation took 3e-05 seconds
## Chain 2: 1000 transitions using 10 leapfrog steps per transition would take 0.3 seconds.
## Chain 2: Adjust your expectations accordingly!
## Chain 2:
## Chain 2:
## Chain 2: Iteration:    1 / 2000 [  0%]  (Warmup)
## Chain 2: Iteration:  200 / 2000 [ 10%]  (Warmup)
## Chain 2: Iteration:  400 / 2000 [ 20%]  (Warmup)
## Chain 2: Iteration:  600 / 2000 [ 30%]  (Warmup)
## Chain 2: Iteration:  800 / 2000 [ 40%]  (Warmup)
## Chain 2: Iteration: 1000 / 2000 [ 50%]  (Warmup)
## Chain 2: Iteration: 1001 / 2000 [ 50%]  (Sampling)
## Chain 2: Iteration: 1200 / 2000 [ 60%]  (Sampling)
## Chain 2: Iteration: 1400 / 2000 [ 70%]  (Sampling)
## Chain 2: Iteration: 1600 / 2000 [ 80%]  (Sampling)
## Chain 2: Iteration: 1800 / 2000 [ 90%]  (Sampling)
## Chain 2: Iteration: 2000 / 2000 [100%]  (Sampling)
## Chain 2:
## Chain 2:  Elapsed Time: 4.417 seconds (Warm-up)
## Chain 2:                1.391 seconds (Sampling)
## Chain 2:                5.808 seconds (Total)
## Chain 2:
##
## SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 3).
## Chain 3:
## Chain 3: Gradient evaluation took 2.6e-05 seconds
## Chain 3: 1000 transitions using 10 leapfrog steps per transition would take 0.26 seconds.
## Chain 3: Adjust your expectations accordingly!
## Chain 3:
## Chain 3:
## Chain 3: Iteration:    1 / 2000 [  0%]  (Warmup)
## Chain 3: Iteration:  200 / 2000 [ 10%]  (Warmup)
## Chain 3: Iteration:  400 / 2000 [ 20%]  (Warmup)
## Chain 3: Iteration:  600 / 2000 [ 30%]  (Warmup)
## Chain 3: Iteration:  800 / 2000 [ 40%]  (Warmup)
## Chain 3: Iteration: 1000 / 2000 [ 50%]  (Warmup)
## Chain 3: Iteration: 1001 / 2000 [ 50%]  (Sampling)
## Chain 3: Iteration: 1200 / 2000 [ 60%]  (Sampling)
## Chain 3: Iteration: 1400 / 2000 [ 70%]  (Sampling)
## Chain 3: Iteration: 1600 / 2000 [ 80%]  (Sampling)
## Chain 3: Iteration: 1800 / 2000 [ 90%]  (Sampling)
## Chain 3: Iteration: 2000 / 2000 [100%]  (Sampling)
## Chain 3:
## Chain 3:  Elapsed Time: 4.535 seconds (Warm-up)
## Chain 3:                1.483 seconds (Sampling)
## Chain 3:                6.018 seconds (Total)
## Chain 3:
##
## SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 4).
```
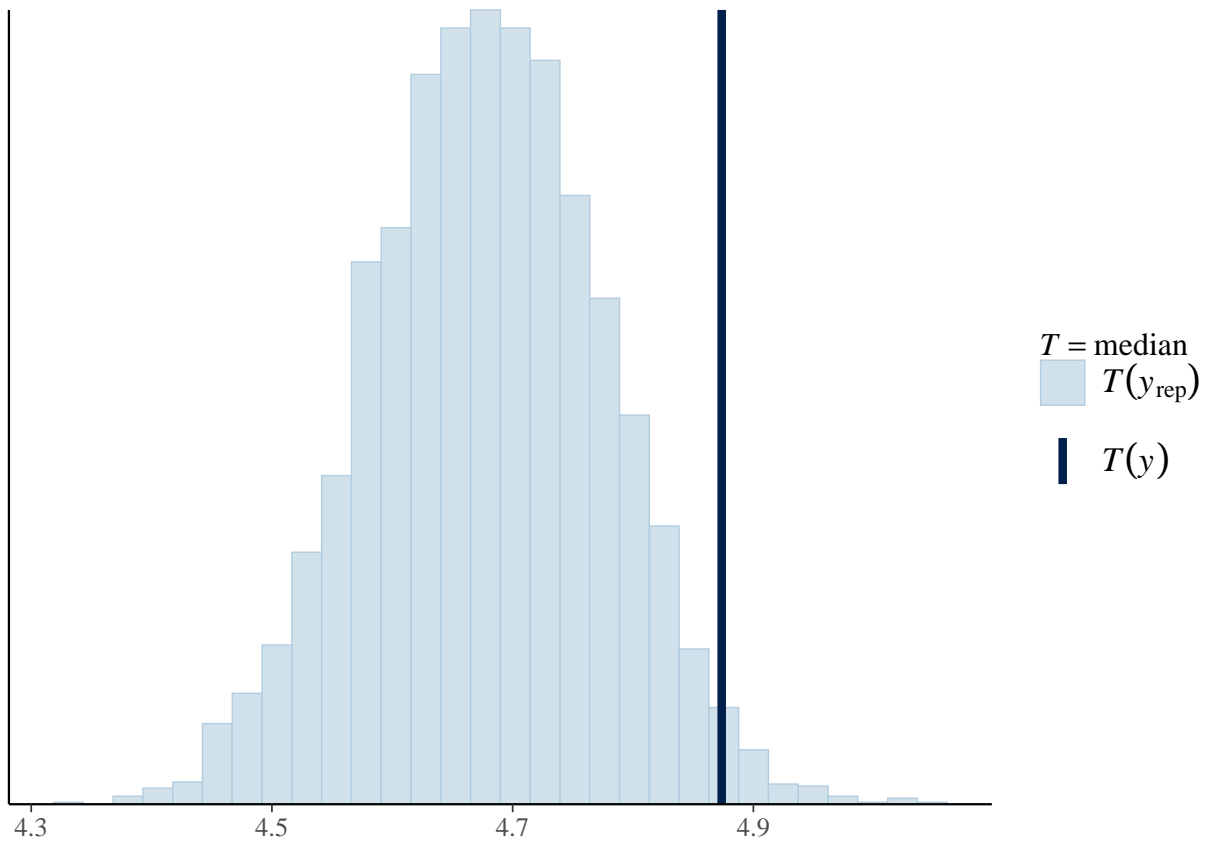
```
## Chain 4:
## Chain 4: Gradient evaluation took 2.5e-05 seconds
## Chain 4: 1000 transitions using 10 leapfrog steps per transition would take 0.25 seconds.
## Chain 4: Adjust your expectations accordingly!
## Chain 4:
## Chain 4:
## Chain 4: Iteration:    1 / 2000 [  0%]  (Warmup)
## Chain 4: Iteration:  200 / 2000 [ 10%]  (Warmup)
## Chain 4: Iteration:  400 / 2000 [ 20%]  (Warmup)
## Chain 4: Iteration:  600 / 2000 [ 30%]  (Warmup)
## Chain 4: Iteration:  800 / 2000 [ 40%]  (Warmup)
## Chain 4: Iteration: 1000 / 2000 [ 50%]  (Warmup)
## Chain 4: Iteration: 1001 / 2000 [ 50%]  (Sampling)
## Chain 4: Iteration: 1200 / 2000 [ 60%]  (Sampling)
## Chain 4: Iteration: 1400 / 2000 [ 70%]  (Sampling)
## Chain 4: Iteration: 1600 / 2000 [ 80%]  (Sampling)
## Chain 4: Iteration: 1800 / 2000 [ 90%]  (Sampling)
## Chain 4: Iteration: 2000 / 2000 [100%]  (Sampling)
## Chain 4:
## Chain 4:  Elapsed Time: 4.455 seconds (Warm-up)
## Chain 4:                1.351 seconds (Sampling)
## Chain 4:                5.806 seconds (Total)
## Chain 4:
```

$T = \text{median}$
$T(y_{\text{rep}})$
$T(y)$
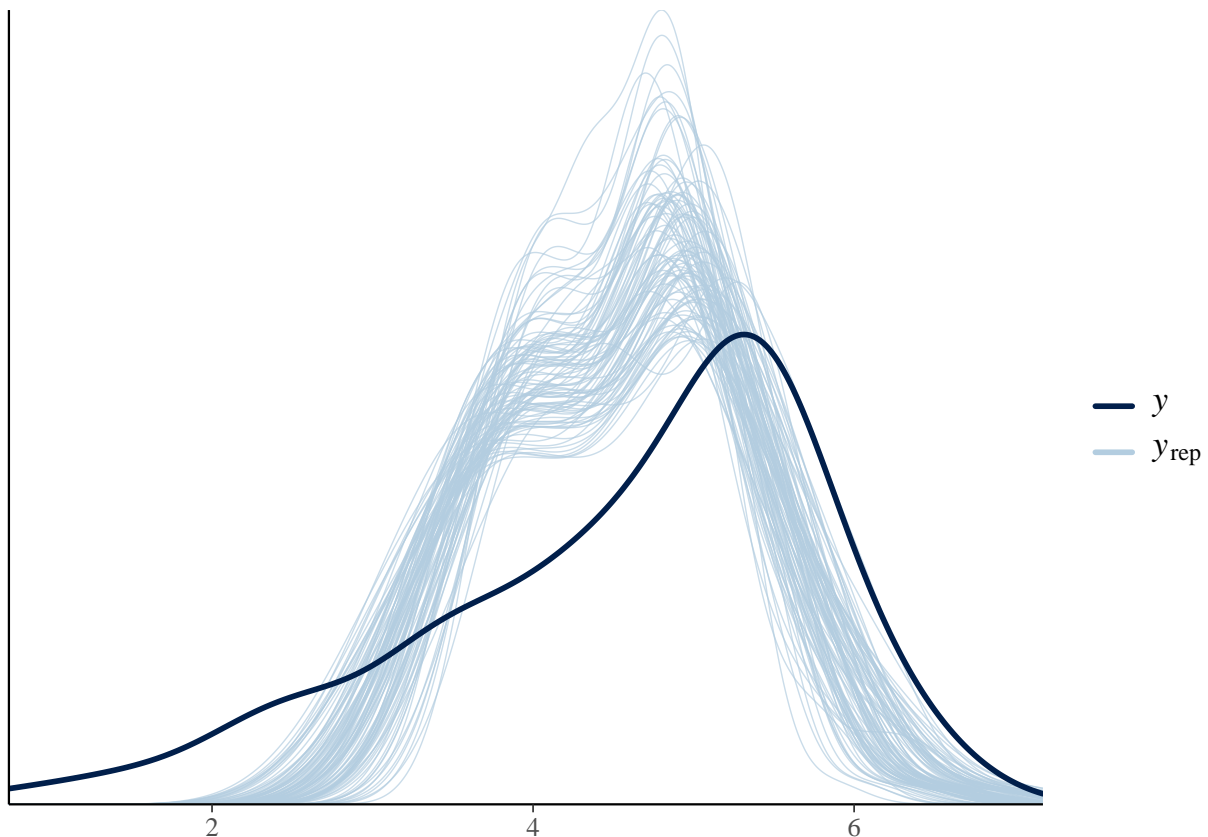
Posterior Checks for Sandy Loam Soil

```
##
## SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 1).
## Chain 1:
## Chain 1: Gradient evaluation took 5.8e-05 seconds
## Chain 1: 1000 transitions using 10 leapfrog steps per transition would take 0.58 seconds.
## Chain 1: Adjust your expectations accordingly!
## Chain 1:
## Chain 1:
## Chain 1: Iteration:    1 / 2000 [  0%]  (Warmup)
## Chain 1: Iteration:  200 / 2000 [ 10%]  (Warmup)
## Chain 1: Iteration:  400 / 2000 [ 20%]  (Warmup)
## Chain 1: Iteration:  600 / 2000 [ 30%]  (Warmup)
## Chain 1: Iteration:  800 / 2000 [ 40%]  (Warmup)
## Chain 1: Iteration: 1000 / 2000 [ 50%]  (Warmup)
## Chain 1: Iteration: 1001 / 2000 [ 50%]  (Sampling)
## Chain 1: Iteration: 1200 / 2000 [ 60%]  (Sampling)
## Chain 1: Iteration: 1400 / 2000 [ 70%]  (Sampling)
## Chain 1: Iteration: 1600 / 2000 [ 80%]  (Sampling)
## Chain 1: Iteration: 1800 / 2000 [ 90%]  (Sampling)
## Chain 1: Iteration: 2000 / 2000 [100%]  (Sampling)
## Chain 1:
## Chain 1:  Elapsed Time: 7.791 seconds (Warm-up)
## Chain 1:                2.734 seconds (Sampling)
```

```
## Chain 1:                    10.525 seconds (Total)
## Chain 1:
##
## SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 2).
## Chain 2:
## Chain 2: Gradient evaluation took 4.2e-05 seconds
## Chain 2: 1000 transitions using 10 leapfrog steps per transition would take 0.42 seconds.
## Chain 2: Adjust your expectations accordingly!
## Chain 2:
## Chain 2:
## Chain 2: Iteration:    1 / 2000 [  0%]  (Warmup)
## Chain 2: Iteration:  200 / 2000 [ 10%]  (Warmup)
## Chain 2: Iteration:  400 / 2000 [ 20%]  (Warmup)
## Chain 2: Iteration:  600 / 2000 [ 30%]  (Warmup)
## Chain 2: Iteration:  800 / 2000 [ 40%]  (Warmup)
## Chain 2: Iteration: 1000 / 2000 [ 50%]  (Warmup)
## Chain 2: Iteration: 1001 / 2000 [ 50%]  (Sampling)
## Chain 2: Iteration: 1200 / 2000 [ 60%]  (Sampling)
## Chain 2: Iteration: 1400 / 2000 [ 70%]  (Sampling)
## Chain 2: Iteration: 1600 / 2000 [ 80%]  (Sampling)
## Chain 2: Iteration: 1800 / 2000 [ 90%]  (Sampling)
## Chain 2: Iteration: 2000 / 2000 [100%]  (Sampling)
## Chain 2:
## Chain 2:  Elapsed Time: 7.998 seconds (Warm-up)
## Chain 2:                3.415 seconds (Sampling)
## Chain 2:                11.413 seconds (Total)
## Chain 2:
##
## SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 3).
## Chain 3:
## Chain 3: Gradient evaluation took 4.3e-05 seconds
## Chain 3: 1000 transitions using 10 leapfrog steps per transition would take 0.43 seconds.
## Chain 3: Adjust your expectations accordingly!
## Chain 3:
## Chain 3:
## Chain 3: Iteration:    1 / 2000 [  0%]  (Warmup)
## Chain 3: Iteration:  200 / 2000 [ 10%]  (Warmup)
## Chain 3: Iteration:  400 / 2000 [ 20%]  (Warmup)
## Chain 3: Iteration:  600 / 2000 [ 30%]  (Warmup)
## Chain 3: Iteration:  800 / 2000 [ 40%]  (Warmup)
## Chain 3: Iteration: 1000 / 2000 [ 50%]  (Warmup)
## Chain 3: Iteration: 1001 / 2000 [ 50%]  (Sampling)
## Chain 3: Iteration: 1200 / 2000 [ 60%]  (Sampling)
## Chain 3: Iteration: 1400 / 2000 [ 70%]  (Sampling)
## Chain 3: Iteration: 1600 / 2000 [ 80%]  (Sampling)
## Chain 3: Iteration: 1800 / 2000 [ 90%]  (Sampling)
## Chain 3: Iteration: 2000 / 2000 [100%]  (Sampling)
## Chain 3:
## Chain 3:  Elapsed Time: 7.673 seconds (Warm-up)
## Chain 3:                2.596 seconds (Sampling)
## Chain 3:                10.269 seconds (Total)
## Chain 3:
##
## SAMPLING FOR MODEL 'anon_model' NOW (CHAIN 4).
```

```
## Chain 4:
## Chain 4: Gradient evaluation took 3.5e-05 seconds
## Chain 4: 1000 transitions using 10 leapfrog steps per transition would take 0.35 seconds.
## Chain 4: Adjust your expectations accordingly!
## Chain 4:
## Chain 4:
## Chain 4: Iteration:    1 / 2000 [  0%]  (Warmup)
## Chain 4: Iteration:  200 / 2000 [ 10%]  (Warmup)
## Chain 4: Iteration:  400 / 2000 [ 20%]  (Warmup)
## Chain 4: Iteration:  600 / 2000 [ 30%]  (Warmup)
## Chain 4: Iteration:  800 / 2000 [ 40%]  (Warmup)
## Chain 4: Iteration: 1000 / 2000 [ 50%]  (Warmup)
## Chain 4: Iteration: 1001 / 2000 [ 50%]  (Sampling)
## Chain 4: Iteration: 1200 / 2000 [ 60%]  (Sampling)
## Chain 4: Iteration: 1400 / 2000 [ 70%]  (Sampling)
## Chain 4: Iteration: 1600 / 2000 [ 80%]  (Sampling)
## Chain 4: Iteration: 1800 / 2000 [ 90%]  (Sampling)
## Chain 4: Iteration: 2000 / 2000 [100%]  (Sampling)
## Chain 4:
## Chain 4:  Elapsed Time: 7.641 seconds (Warm-up)
## Chain 4:                3.008 seconds (Sampling)
## Chain 4:                10.649 seconds (Total)
## Chain 4:
```
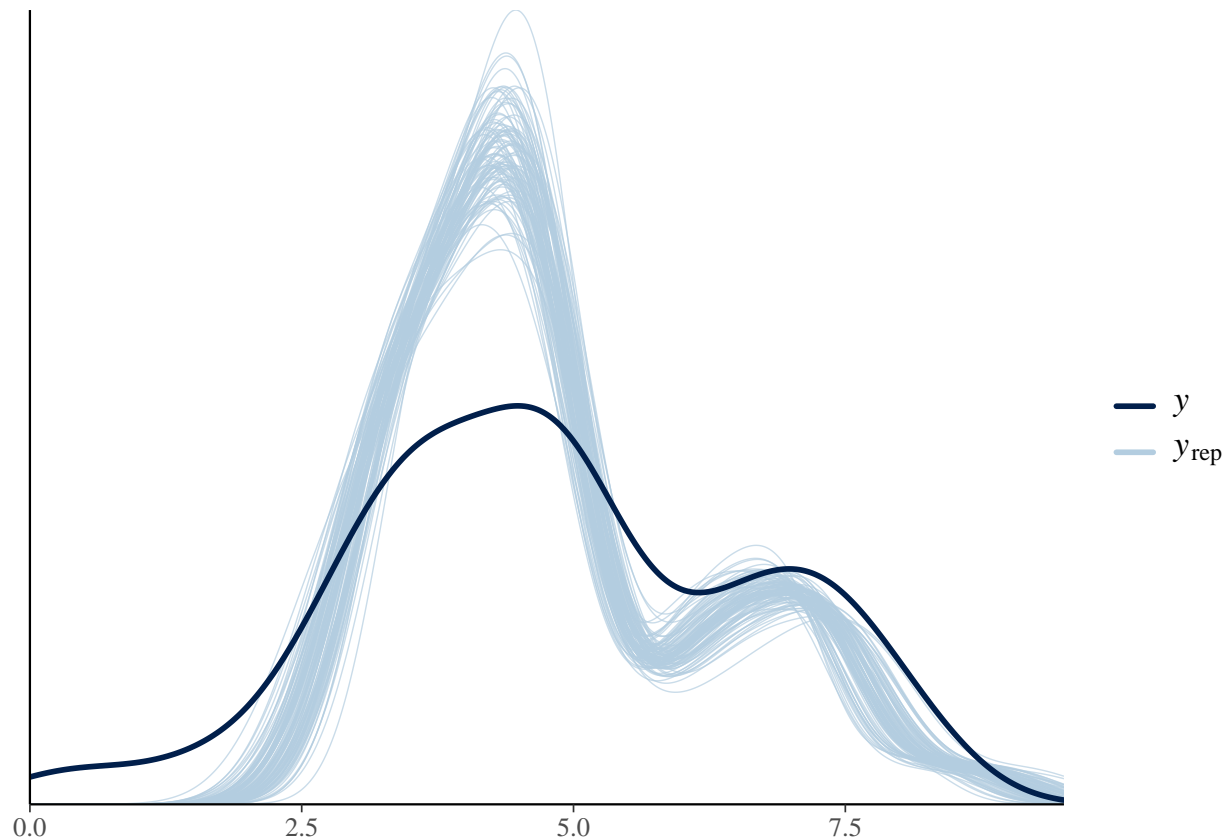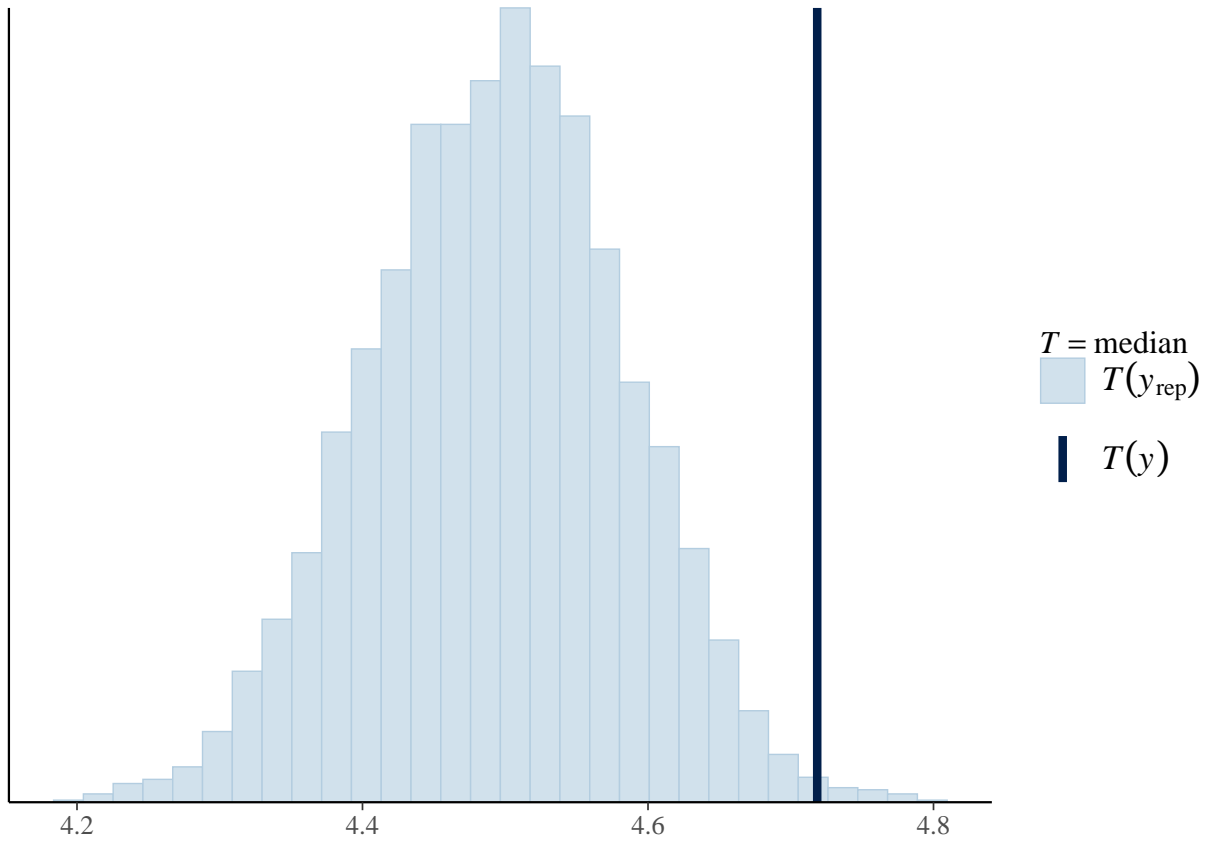
References

Yuan Li, Zhou Li, Song Cui, Scott X. Chang, Chunlin Jia, Qingping Zhang, A global synthesis of the effect of water and nitrogen input on maize (Zea mays) yield, water productivity and nitrogen use efficiency, Agricultural and Forest Meteorology, Volume 268, 2019, Pages 136-145, ISSN 0168-1923, https://doi.org/10.1016/j.agrformet.2019.01.018. (https://www.sciencedirect.com/science/article/pii/S0168192319300188)