The Impact of Severity Types and Nitrogen Level on Gudgrass Growth

By Harley Clifton and Mark Braun

I. Introduction

A non-native weed species, Baddgrass, has infested areas across central Montana. A university's research station of size 200m by 200m was used as a study area to conduct an experiment. This experiment aimed to rid the study area of Baddgrass, allowing for the regrowth of Gudgrass, which is a native plant species. First, an herbicide called Bio-B-Gone was used to kill all exposed plant life in the study area. Next, Gudgrass was planted in the hope of revegetating the area. A land reclamation scientist collected data on the amount of vegetation present two years after the Bio-B-Gone application and subsequent planting of Gudgrass. For each sampled quadrat, the number of Gudgrass plants (count) and Nitrogen Level was recorded. The parameter of interest was the mean number of Gudgrass plants per quadrat in the study area.

II. Sampling Design

In our sampling plan, individual units were the 1600 5m by 5m quadrats enclosed in the university's agricultural research station. We stratified the region into four strata corresponding to the different levels of severity type: low (1), moderate (2), high (3), and very high (4). Each stratum had a different number of quadrats. The units (quadrats) in each stratum were numbered from one to that stratum's size. The first unit in each stratum was the most northwest quadrat, and the numbering continued moving east across the rows before moving to the row directly south (Figure 1). This repeated until all quadrats of that severity type were numbered, and this process repeated for each of the strata. Next, a simple random sample was taken with replacement from each stratum using a random number generator.

The sample size from each stratum was calculated based on the different costs of sampling from each stratum and prior information about variability within strata. The calculations resulted in taking a sample of size n = 62 from stratum h_1, n = 37 from stratum h_2, n = 28 from stratum h_3, and n =

7 from stratum h_4. The overall cost of this sampling plan was under our budget by \$10, resulting in a total cost of \$7190.

III. Statistical Procedures Used

Once the data were collected on the number of Gudgrass plants and nitrogen levels for each sampling unit, regression analysis was used to find the overall mean number of Gudgrass plants per quadrant at the mean nitrogen level for each stratum.

First, the counts of Gudgrass plants and nitrogen level for each stratum were visualized in R (R Core Team, 2021) with the ggplot2 package (Wickham, 2016) in Figure 2. From these scatterplots, it was determined that Gudgrass plants and nitrogen level had a positive relationship for each of the severity types. Then, an additional scatterplot was created that plotted Gudgrass plants vs. Nitrogen level, with points colored by severity type (Figure 3, ggplot2 package, Wickham 2016). From Figure 3, it was obvious that the number of Gudgrass plants and nitrogen level had a similar, positive relationship for each of the severity types. Thus, since the data looked plausibly linear, stratified linear regression analysis was deemed appropriate for this data.

The survey package in R was used to estimate the total and mean number of Gudgrass plants per quadrat for each stratum using Linear Regression (Table 1, Lumley 2004). Since there was no prior information about nitrogen level, the mean estimate of Gudgrass plants was calculated using the mean nitrogen level from the sampled units for each stratum. Next, the stratum total estimates and total standard error estimates were weighted based on their respective stratum population sizes to calculate the overall estimated mean number of Gudgrass plants per quadrat. The standard error of this estimate was utilized to get a 95% confidence interval for this estimate, which can be found in Table 2.

The number of Gudgrass plants and nitrogen level was plotted with the overall estimated linear regression line in Figure 4, and the estimated coefficients of the regression line were calculated (Model 1). The Multiple R-Squared values for these model suggested room for improvement, so a Quadratic Regression was fit for each of the strata (Figure 6), and the overall model (Figure 5). The stratum total estimates and total standard error estimates were weighted based on their respective stratum population sizes to calculate the overall estimated mean number of Gudgrass plants per quadrat, the same as before. The Multiple R-Squareds for the Quadratic Model indicated improvement and therefore became our final model. The resulting estimated coefficients for the Quadratic Model for each stratum are displayed in Model 2.

IV. Summary of Statistical Findings

The individual stratum mean and total Gudgrass estimates with Linear Regression can be found in Table 1. The Multiple R-squared values for the Linear Regression ranged from 0.9349 to 0.9874 (Model 1). Therefore, at the least, the linear regression model with nitrogen explained 93.49% of the total variation in the mean number of Gudgrass plants per quadrat. This suggests that nitrogen explains most of the total variation, leaving only little variation in the mean number of Gudgrass plants unexplained. Although this value was high, there was still room for improvement in the model. Thus, the analysis strategy was revised.

Since the data curved slightly upward, a Quadratic Regression Model was fit for each individual stratum (Figure 6). The Multiple R-squared values for the Quadratic Regression ranged from 0.991 to 0.9986 (Model 2). Therefore, at the least, the quadratic regression models with nitrogen explained 99.1% of the total variation in the mean number of Gudgrass plants per quadrat. This suggests that this model leaves a negligible amount of variation in the mean number of Gudgrass plants unexplained. Thus, the R-Squared indicated that the Quadratic Regression model was a better fit than the Linear Regression Model. Further, the Quadratic Model had a smaller overall standard error estimate than the Linear Model, and therefore a narrower confidence interval for the overall estimated mean count of Gudgrass plants per quadrat suggesting further evidence that the Quadratic Model is superior (Table 2, Table 4).

With Quadratic Regression estimation, we estimate the true mean number of Gudgrass plants per quadrat in each stratum to be approximately 50.80464, 45.79905, 50.16564, and 86.71198 plants, with standard errors of 0.2343080, 0.2289214, 0.2379155, and 0.8371624 Gudgrass plants for Severity Types Low, Moderate, High, and Very High respectively (Table 3). After weighting these by the appropriate stratum population sizes, we estimated the overall true mean number of Gudgrass plants to be 50.50348 plants, with a standard error of 0.2545931 Gudgrass plants. We are 95% confident that the true overall mean number of Gudgrass plants per quadrat in this university's agricultural station is between 49.99980 plants and 51.00716 plants (Table 4).

V. Scope of Inference

Since we performed a stratified simple random sample, inferences will apply to the entire university's agricultural region. However, since this was an observational study, no causal inferences can be made about these results. Still, it is worth noting that nitrogen level is positively associated with the number of Gudgrass plants present in this study area; as nitrogen level increases, so does the average number of Gudgrass plants per quadrat.

VI. Conclusion

Overall, the methods employed for this study were very successful in estimating the true population mean number of Gudgrass plants per quadrat. The Quadratic Models fit each stratum's data almost perfectly and produced very precise mean estimates for each stratum, leading to a very precise population mean estimate and associated confidence interval. A higher budget definitely would have been helpful in increasing the precision of the mean estimate even further, as it would have allowed for the sampling of more units, especially from strata with higher variability.

If this study were to be run again, it would be useful to incorporate the prior information about nitrogen levels in each stratum to inform sampling plans and decisions with analysis. We did not have any previous information about nitrogen level in each stratum to inform our mean Gudgrass counts per quadrat, so we had to use the respective means of the sampled nitrogen measurements for each stratum. The mean nitrogen levels found in this study could inform what level of nitrogen to estimate the mean count of Gudgrass plants at in later studies.

Appendix

Figures

Figure 1: Severity Type Map.

Severity Type Map

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
1	3	3	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	3	3	2	2	2	1	1	1	1	1
2	3	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	1	1	1	1	1
3	3	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	1	1	1	1	1
4	3	3	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	1	1	1	1	1	1
5	3	3	3	3	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1
6	3	3	3	3	3	3	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1
7	4	4	4	3	3	3	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1
8	4	4	4	3	3	3	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	3	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1
9	4	4	4	3	3	3	3	2	2	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	3	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1
10	4	4	4	4	3	3	3	2	2	1	1	1	1	1	1	1	1	1	1	1	2	2	3	3	3	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1
11	4	4	4	4	3	3	3	2	2	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	1	1	1	1	2	1	1	1	1	1	1	1	1
12	4	4	4	4	3	3	3	2	2	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1
13	4	4	4	4	3	3	3	2	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
14	4	4	4	4	3	3	3	2	1	1	1	1	1	1	2	2	3	3	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
15	3	3	3	3	3	3	2	2	1	1	1	1	1	2	2	3	3	3	3	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
16	3	3	3	3	2	2	2	1	1	1	1	1	1	2	2	3	3	3	3	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
17	3	2	2	2	2	2	1	1	1	1	1	1	1	2	2	3	3	3	3	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
18	2	2	2	2	2	2	1	1	1	1	1	1	1	2	2	3	3	3	3	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
19	2	2	2	2	2	2	1	1	1	1	1	1	1	1	2	2	3	3	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
20	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	2	2	2	1	1	1	1	1	1
21	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	2	2	2	1	1	1	1	1	1
22	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	2	2	2	2	1	1	1	1	1	1	1	1	1	2	2	2	2	2	1	1	1	1	1	1
23	1	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	2	2	2	2	2	2	2	2	1	1	1	1	1	1
24	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1
25	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1
26	1	1	1	1	1	1	1	1	1	2	2	2	2	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1
27	1	1	1	1	1	1	1	2	2	2	2	2	2	2	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	1	1	1	2	2	2	2	2	2	2
28	1	1	1	1	1	1	1	2	3	3	3	3	3	2	2	1	1	1	1	1	2	2	2	3	3	3	3	2	2	2	1	1	2	2	3	3	3	3	3	2
29	1	1	1	1	1	1	2	3	3	3	3	3	3	3	2	1	1	1	1	1	2	2	2	3	3	3	3	2	2	2	1	2	2	3	3	3	3	3	3	3
30	1	1	1	1	1	1	2	3	3	3	3	3	3	3	2	2	1	1	1	1	2	2	2	3	3	3	3	3	2	2	2	2	2	3	3	4	4	4	4	3
31	1	1	1	1	2	2	3	3	3	3	3	3	3	3	3	2	2	1	1	1	2	2	2	3	3	3	3	3	2	2	2	2	2	3	4	4	4	4	4	3
32	1	1	1	1	2	2	3	3	3	3	3	3	3	3	3	3	3	2	2	2	2	2	3	3	3	3	3	3	2	2	2	2	3	3	4	4	4	4	4	3
33	1	1	1	2	2	2	2	3	3	3	3	3	3	3	3	3	3	2	2	2	2	3	3	3	3	3	3	3	2	2	2	2	3	3	3	4	4	4	4	4
34	1	1	1	2	2	2	2	3	3	3	3	3	3	3	3	3	3	2	2	2	3	3	3	3	3	3	3	3	3	2	2	2	2	3	3	4	4	4	4	4
35	2	2	1	1	1	2	2	2	3	3	3	3	3	3	3	3	3	3	2	2	3	3	3	3	3	3	3	3	3	2	2	2	2	3	3	3	3	3	3	3
36	2	2	1	1	1	1	1	2	2	3	3	3	3	3	3	3	3	3	2	3	3	3	3	3	3	3	3	3	3	3	2	2	2	3	3	3	3	3	3	3
37	2	2	1	1	1	1	1	1	1	2	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	2	2	3	3	3	3	3	3	3
38	1	1	1	1	1	1	1	1	1	1	2	2	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	2	2	2	2	2	2	2	2
39	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	3	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	2	2	2	2	2	2	2	1
40	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	3	3	3	3	2	2	2	2	2	2	3	3	3	4	4	4	3	3	2	2	2	1	1	1	1

Figure 2: Scatterplots of Gudgrass Plants and Nitrogen Levels for Each Severity Type, created with the ggplot2 package.

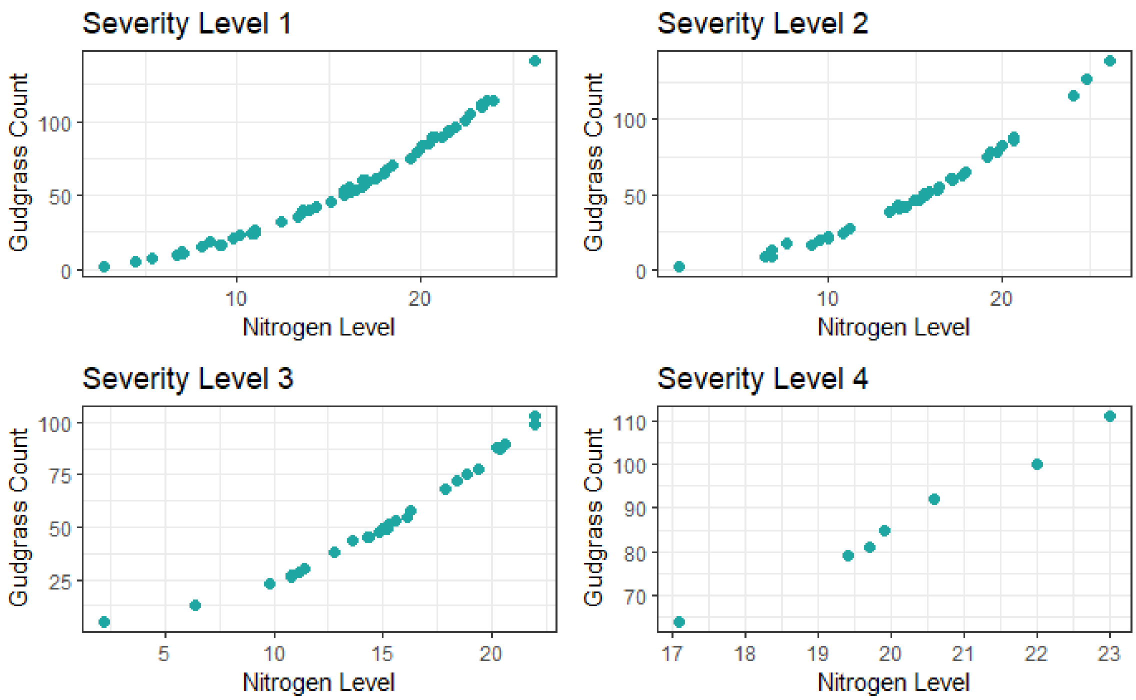


Figure 3: Scatterplot of Gudgrass Plants vs. Nitrogen Level colored by Severity Type, created with the ggplot2 package.

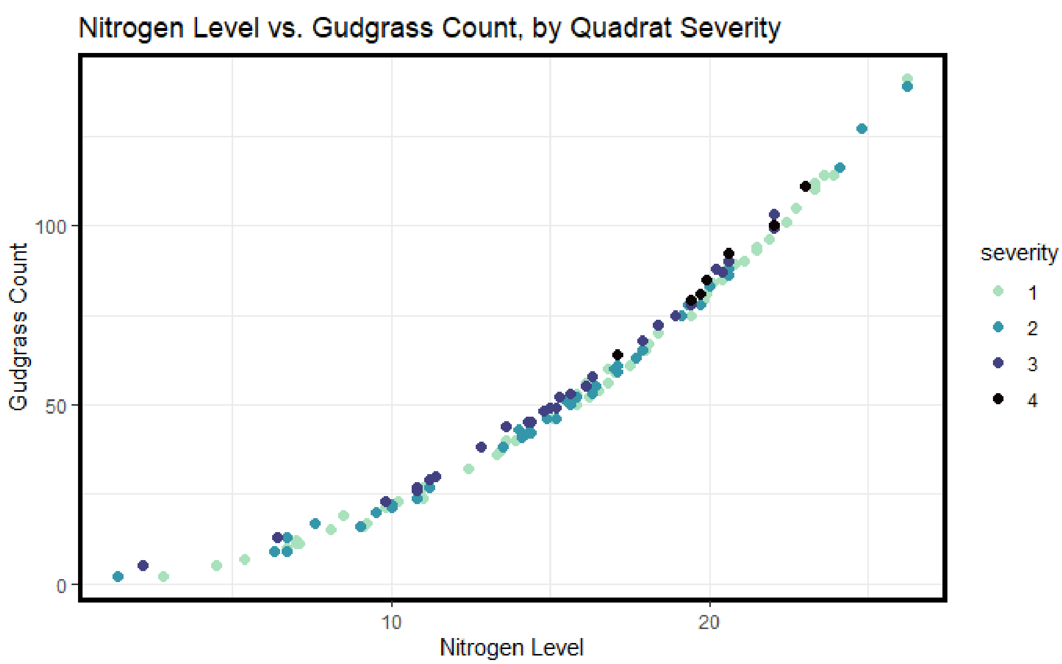


Figure 4: Plot of Gudgrass Plants and Nitrogen Level with the Overall Linear Regression Line, created with the ggplot2 package.

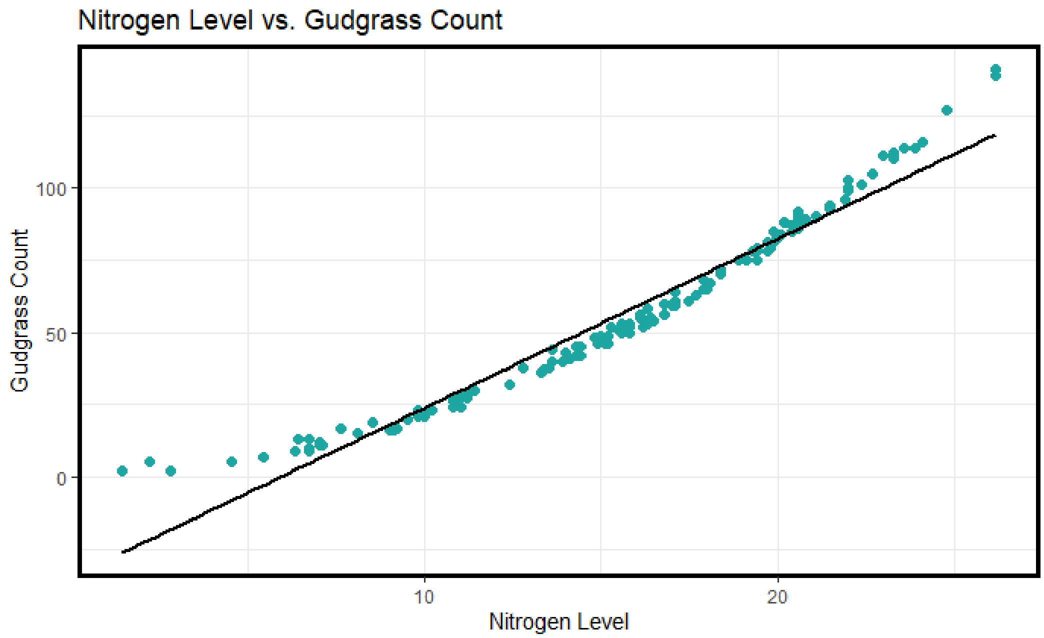


Figure 5: Quadratic Regression Plot of Gudgrass Plants and Nitrogen Level with the Overall Quadratic Regression Line, created with the ggplot2 package.

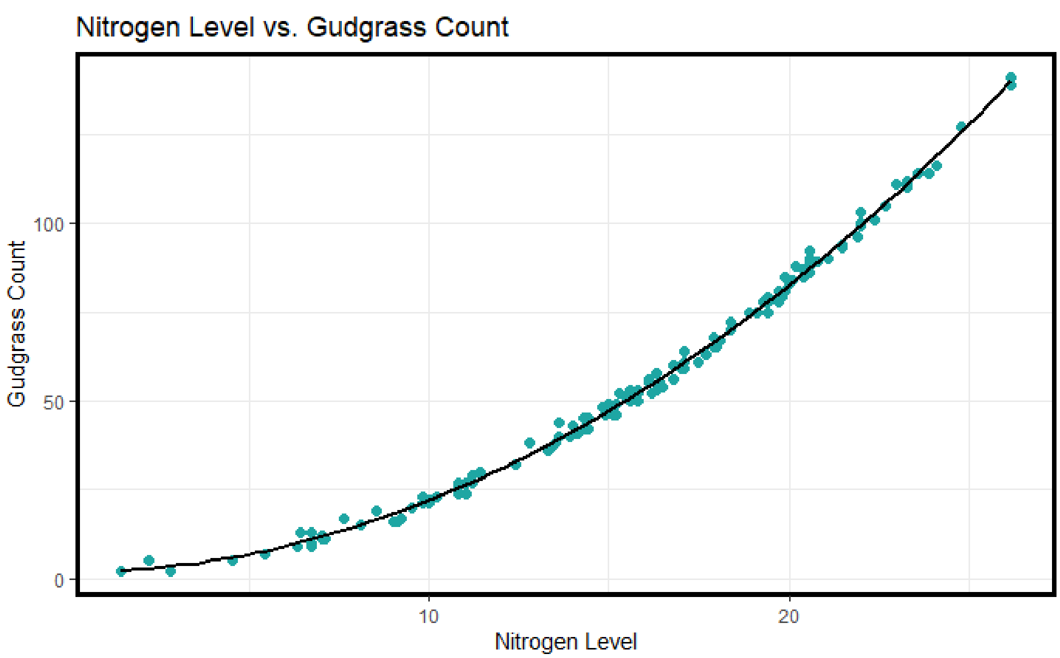
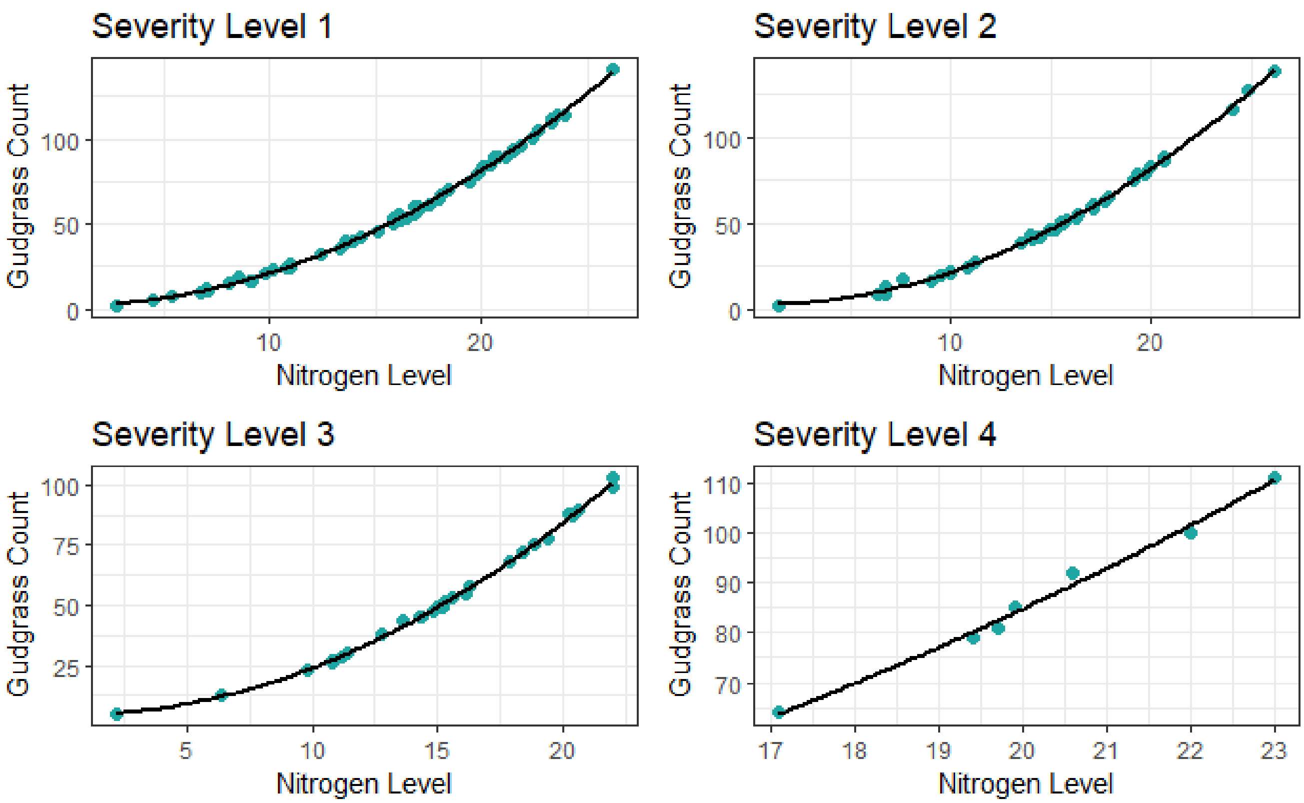


Figure 6: Quadratic Regression of the Number of Gudgrass Plants and Nitrogen Levels for Each Severity Type, created with the ggplot2 package.



Model Summaries

Model 1: Estimated Coefficients for the Linear Model with Severity Types and Nitrogen Level on the Mean Number of Gudgrass Plants per Quadrat.

Stratum	Intercept	Linear Term	Multiple R^2
h_1	-36.1493	5.9391	0.9597
h_2	-33.4844	5.7339	0.9349
h_3	-27.4915	5.4009	0.9442
h_4	-73.8038	7.9649	0.9874

Model 2: Estimated Coefficients for the Quadratic Model with Severity Types and Nitrogen Level on the Mean Number of Gudgrass Plants per Quadrat.

Stratum	Intercept	Linear Term	Quadratic Term	Multiple R^2
h_1	1.44767	-0.05817	0.20345	0.9986
h_2	2.4227	-0.1233	0.2043	0.9983
h_3	4.22446	-0.02359	0.20124	0.9981
h_4	18.5387	-1.2864	0.2299	0.991

Tables

Table 1: Total and Mean Number of Gudgrass Plants per Quadrat for each Stratum using Linear Regression.

Stratum	Severity Type	Total Estimate	Standard Error of Total Estimate	Mean Estimate	Standard Error of Mean Estimate
h_1	Low	43823	669.86411	57.210	0.8744962
h_2	Medium	23678	624.98624	51.811	1.3675848
h_3	High	17472	373.16949	54.429	1.1625218
h_4	Very High	4896	36.29389	87.429	0.6481052

Table 2: Estimated Overall Mean Number of Gudgrass Plants per Quadrat using Linear Regression.

Overall Mean Estimate	SE of Overall Mean Estimate	95% Confidence Interval for Overall Mean Estimate	
56.167328	1.065196	54.05996	58.27469

Table 3: Total and Mean Number of Gudgrass Plants per Quadrat for each Stratum using Quadratic Regression.

Stratum	Severity Type	Total Estimate	Standard Error of Total Estimate	Mean Estimate	Standard Error of Mean Estimate
h_1	Low	38916	179.47991	50.80464	0.2343080
h_2	Medium	20930	104.61708	45.79905	0.2289214
h_3	High	16103	76.37089	50.16564	0.2379155
h_4	Very High	4856	46.88110	86.71198	0.8371624

Table 4: Estimated Overall Mean Number of Gudgrass Plants per Quadrat using Quadratic Regression.

Overall Mean Estimate	**SE of Overall Mean Estimate**	**95% Confidence Interval for Overall Mean Estimate**	
50.50348	0.2545931	49.99980	51.00716

References

H. Wickham. ggplot2: Elegant Graphics for Data Analysis (2016). Springer-Verlag New York.

T. Lumley. survey: Analysis of Complex Survey Samples (2004). Journal of Statistical Software 9(1): 1-19

R Core Team. R: A language and environment for statistical computing (2021). R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.