Design of Experiments on Grass Growth in Controlled Environment

Zaman, Muhammad Adib Uz and Parmer, Anup Bharatkumar and Naghdi, Saman

Northern Illinois University, Northern Illinois University, University of the Persian Gulf International Branch

4 January 2018

Design of Experiments on Grass Growth in Controlled Environment

Abstract

Most design of experiments in agricultural applications are complex operations in nature because of numerous process variables, feed material attributes, and raw material attributes that can have significant impact on the performance of the process. Design of experiments (DOE)-based approach offers a solution to this conundrum and allows for an efficient estimation of the main effects and the interactions with minimal number of experiments. This study investigates on the most effective factors contributing in grass growth. All the factors are set in two levels to create a full-factorial 2^k design. A systematic methodology is proposed for construction of the model and for precise prediction of the responses which is lawn growth. The results indicate that water is the most significant factor that the cultivator can directly control and cheap seeds found to be suitable for the grass growth applications under consideration.

Introduction and Literature Review

Sowing grass seed is an American tradition. The general home-owning population has been conditioned to value having thick green lawns as part of their landscape. There is a multi-billion dollar industry (Lawn Starter, 2017) designed to provide the supplies required to succeed in this endeavor. Various seed and fertilizer options are available at a variety of price points, all touting the benefits they want the customer to believe and pay for. Water is an obvious requirement to grow grass. Nature can provide all of the rain required, but it can also fail. In arid climates, and in the absence of rain in moist climates, watering can become necessary to sprout and nurture new grass as shown by Rabiei Hosseinabad et al. (2015).

Irrigation can strain local municipal water supplies (Polycarpou, 2017), and costs the home owner. A variety of advice can be found suggesting the proper amount to water to sustain a lawn. Clearly, the actual requirements may vary by region, and by species of grass being grown. Additionally, there is a market for products that intend to maximize the retention of water so that it is available to the plant for a longer period of time.

This study looks to study the effects of several factors that can be incorporated to growing grass from seed. Type of seed, use of fertilizer, use of water-retaining soil enhancement, frequency of watering, and quantity of water was studied to determine if any of these factors have a significant effect on the growth of grass seed.

ANOVA test which is used on the univariate analysis of the results essentially handles the factors used in the experiment or the total of the square of the result variables in order to determine the contribution of their interactions on the experiment and determines the total variances. And then makes possible the election of the most suitable factor/parameter by calculating the contribution percentage of the change (Gencel, 2007). The theory of single replicate incomplete factorial designs has been implemented and tested to check what information it could provide regarding the interplay of optimization parameters. In literature only tables of low order incomplete factorial experiments are to be found (2^{k-p} and 3^{k-p}) and were used

(Connor and Zelen, 1959). The most important process of the DOE is determining the independent variable values at which a limited number of experiments will be conducted. For this purpose, Taguchi proposed an improved DOE. This approach adopts the fundamental idea of DOE, but simplifies and standardizes the factorial and fractional factorial designs so that the conducted experiments can produce more consistent results (Roy, 2001). The effect of the agriculture on environment is very important. Agricultural lands are mostly treated with chemical fertilizers. This causes heavy metal contamination in the soil. Numerous consumers are started to prefer to use organically produced food because of pesticide residues (Foley et al., 2005; Feili et al., 2014).

The Problem

Persons who want to successfully grow grass are bombarded by confusing marketing and packaging. It is important to know which factors the grass grower can control that will actually impact the success of the grass sowing. Marketing and packaging preys on the consumers' uncertainty in order to maximize the sale of grass growing products. It is necessary to quantify the impact of available products in order to determine if they contribute significantly and positively to the successful growing of grass in order to enable the consumer to minimize cost and maximize success of his or her growing endeavors.

Apparatus:

The items listed below were used to create environments to grow each sample under a variety of factor combinations.

- 10oz plastic cup drilled for drainage
- Gravel for drainage
- Top soil
- Factor treatments (Seed, Fertilizer, Peat Moss)
- Flags to label each treatment
- Corrugated trays for drainage
- Syringe for measured watering

Methodology

DOE Modeling

Factorial designs are frequently used to identify the main effects as well as interactions amongst the various factors. For quantitative factors, the data can be represented through the commonly used "linear regression model."1 For two factors, it can be represented as:

$y = \beta_0 + \beta_1 x_1 + \beta_2 x_1 x_2 + \varepsilon$

where, β's are the regression coefficients. This first-order model can be generalized to a higher order model by addition of terms containing higher powers of x. In general, method of least square is used to estimate $\hat{\beta}$ with the assumption that expected value and the variance of the error (ε) are E(ε) 50 and V(ε) = s^2, respectively. In matrix notation, the model can be represented as

$$y = X\beta + \varepsilon$$

where y, β, and ε are the column matrices of (n × 1), (p × 1), and (n × 1) vectors, respectively, X is a (n × p) matrix, and n is the number of observations. Further, p is the number of parameters in the model. The method chooses $\hat{\beta}$ so that the sum of squares of the error e is minimized. The least squares estimate of β is then given by

$$\hat{\beta} = (X'X)^{-1}X'y$$

And, the fitted regression model is

$$y = X\hat{\beta} \quad (4)$$

To evaluate the design and model statistically, it is necessary to estimate the variance (s^2). This point needs to be mentioned that the validity of the model can be measured by R-squared value. If the value of R-squared is close to one, the model is accurate and reliable [5].

Each cup was drilled through the bottom to produce a ¼ inch hole and filled with a ½ inch of gravel to allow any excess water to drain from the system. The standard growing medium was top soil commercially available in forty pound bags. All samples received top soil from the same bag. Peat moss was chosen as a water-retaining soil enhancement. A 2:1 topsoil to peat moss ratio was mixed to fill the cups receiving the "peat moss" treatment. Each cup was filled to approximately 3" of depth and tapped against a work surface to settle the soil into the cups.

Two different varieties of seed were chosen. Both were "Sun and Shade" mixes from the manufacturer, Scotts. One was plain seed, while the other was an enhanced seed that contained a blue coating referred to in this study as "fancy seed" compared to the plain "cheap" seed. The mix of species varied between the two types of seed. The cheap seed was chosen because it is marketed as servicing the same application (sun and shade) and was by the same manufacturer. The only discernable difference to the casual consumer is that the fancy seed retails for about double the cost of the cheap seed and claims to absorb "2X more water than uncoated seed, feeds to jumpstart growth, and protects seedlings against disease." (Scotts, 2017) All samples received the same volume of 1/8 teaspoon of one seed or the other. Once applied, the soil was gently tamped to ensure good soil contact with the seed. The cheap seed was tamped first to avoid any coating cross-contamination.

A Scotts brand fertilizer was chosen. The variety was chosen because it had no special properties such as "weed-n-feed". Scotts brand was chosen to coordinate with the seed choice assuming that the fertilizer should be compatible with the seed since they are from the same manufacturer. The sample that received fertilizer each received 1/16 teaspoon of fertilizer. The fertilizer was sprinkled evenly over each sample designated to receive this treatment. Once all samples were planted, all samples were then watered according to their designated watering level. Two levels were chosen. The low level was 10mL, while the high level was 20mL. The samples were randomly assigned positions in the corrugated trays, and placed in a bright window to germinate.

A watering schedule was adhered to. On Thursdays, samples designated to receive water twice per week were watered to their designated level (Low – 10mL or High- 20mL). On Mondays, all samples were watered to their designated level. A demarcated syringe was used to meter and apply the water. Water used was filtered municipal water from Crystal Lake, Illinois.

On Thursdays, the trays were rotated 180 degrees. On Mondays, the trays traded positions in order to mitigate any effects from variation in either light exposure or temperature that may have affected the samples based on the tray locations.

Germination date for each sample was determined when a predominance of seeds sprouted. Height of each sample was measured 10 days after the recorded germination date. Height was measured to the "predominant height" ignoring outliers of excessive height or shortness. Quarter inch was the smallest unit of measure utilized. All measurements were performed by the same person to maximize consistency in interpreting the height measurements.

Data Collection

The data collection table 1 is given below:

Table 1. Data collection table for height after 10 days

Seed	Peat moss	Fertilizer	Water level	Water frequency	Height after 10 days
Fancy	No peat	Fertilizer	High	2X	1.25
Cheap	No peat	Fertilizer	Low	1X	1
Cheap	No peat	Fertilizer	Low	2X	1.75
Cheap	With peat	Fertilizer	Low	2X	1
Fancy	With peat	Fertilizer	High	1X	1
Cheap	With peat	Fertilizer	High	2X	1
Cheap	With peat	No fertilizer	Low	1X	2
Cheap	With peat	No fertilizer	Low	2X	2.5
Fancy	With peat	Fertilizer	High	2X	1.75
Cheap	No peat	No fertilizer	High	1X	2.5
Fancy	With peat	No fertilizer	Low	1X	2.5
Fancy	No peat	Fertilizer	Low	1X	0.5
Fancy	With peat	No fertilizer	High	2X	2.5
Fancy	No peat	No fertilizer	Low	1X	2.5
Cheap	No peat	Fertilizer	High	1X	0.75
Fancy	No peat	No fertilizer	Low	2X	2.75
Fancy	No peat	Fertilizer	High	1X	1.25
Fancy	No peat	No fertilizer	High	2X	3.25
Fancy	With peat	No fertilizer	Low	2X	2.25
Cheap	With peat	Fertilizer	High	1X	1.25
Cheap	No peat	No fertilizer	Low	2X	2
Cheap	With peat	No fertilizer	High	1X	3
Fancy	With peat	Fertilizer	Low	1X	1
Cheap	No peat	No fertilizer	Low	1X	1.75
Fancy	With peat	No fertilizer	High	1X	2.5
Cheap	With peat	No fertilizer	High	2X	3
Cheap	No peat	Fertilizer	High	2X	1.5
Fancy	With peat	Fertilizer	Low	2X	1.5

Cheap	No peat	No fertilizer	High	2X	2.75
Cheap	With peat	Fertilizer	Low	1X	0.75
Fancy	No peat	No fertilizer	High	1X	3
Fancy	No peat	Fertilizer	Low	2X	0.75

Results

The following plot shows the normality plot of the response variable. The normality plot indicates the data follows normality since the p-value is less than 0.05. knowing that the data follows normality, it enables us to utilize ANOVA analysis to determine the significant factors in growing lawn.

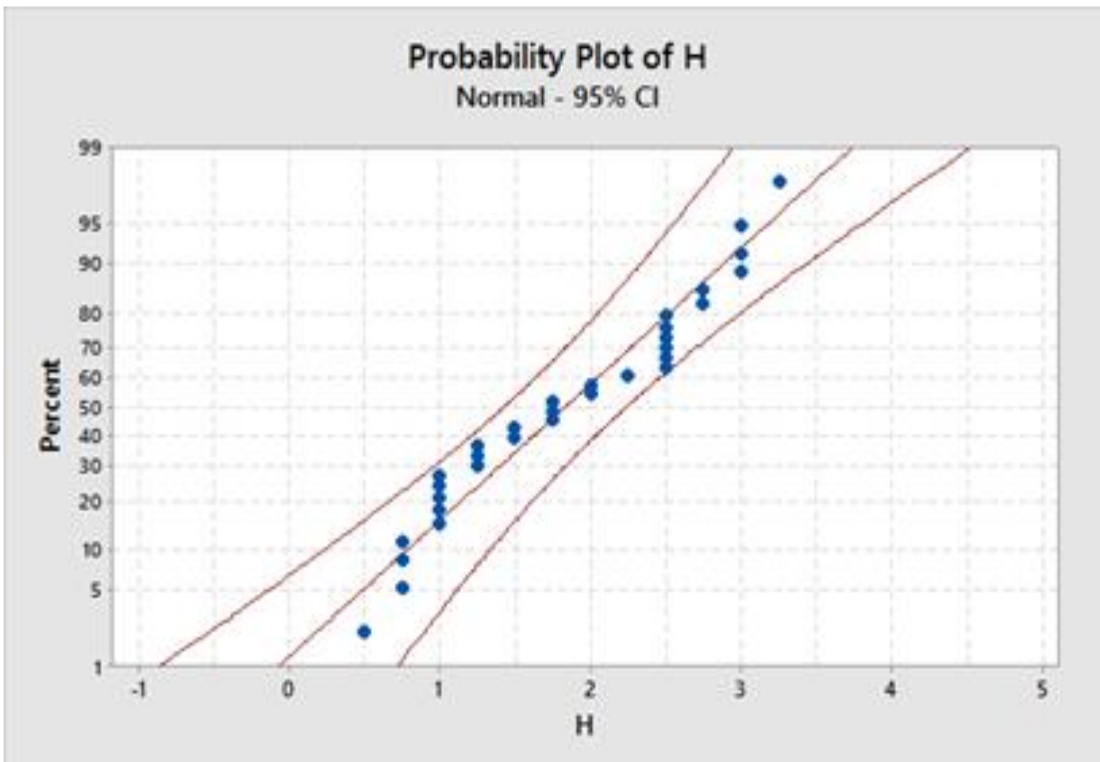

Figure 1. Normality plot of the data collected as response variable

The 2^5 full factorial ANOVA table for the full factorial model is shown below:

Response 2 Height after 10 days
ANOVA for selected factorial model
Analysis of variance table [Partial sum of squares - Type III]

Source	**Sum of Squares**	**df**	**Mean Square**	**F Value**	**p-value Prob > F**
A-Seed	*0.096*	*1*	*0.096*		
B-Peat moss	*1.953E-003*	*1*	*1.953E-003*		

C-Fertilizer	*16.17*	*1*	*16.17*
D-Water level	*1.03*	*1*	*1.03*
E-Water frequency	*0.56*	*1*	*0.56*
AB	*0.018*	*1*	*0.018*
AC	*0.096*	*1*	*0.096*
AD	*1.953E-003*	*1*	*1.953E-003*
AE	*0.018*	*1*	*0.018*
BC	*0.018*	*1*	*0.018*
BD	*0.018*	*1*	*0.018*
BE	*0.049*	*1*	*0.049*
CD	*0.24*	*1*	*0.24*
CE	*0.096*	*1*	*0.096*
DE	*0.018*	*1*	*0.018*
ABC	*1.03*	*1*	*1.03*
ABD	*0.24*	*1*	*0.24*
ABE	*0.096*	*1*	*0.096*
ACD	*0.33*	*1*	*0.33*
ACE	*0.018*	*1*	*0.018*
ADE	*0.049*	*1*	*0.049*
BCD	*0.018*	*1*	*0.018*
BCE	*1.953E-003*	*1*	*1.953E-003*
BDE	*1.953E-003*	*1*	*1.953E-003*
CDE	*1.953E-003*	*1*	*1.953E-003*
ABCD	*0.049*	*1*	*0.049*
ABCE	*0.33*	*1*	*0.33*
ABDE	*0.096*	*1*	*0.096*
ACDE	*1.953E-003*	*1*	*1.953E-003*
BCDE	*1.953E-003*	*1*	*1.953E-003*
ABCDE	*1.953E-003*	*1*	*1.953E-003*
Pure Error	0.000	0	
Cor Total	20.70	31	

The following table 2 contains the factor effects, sum of squares and percentage contribution

Table 2. Table of Terms, their effects, sum of squares and % contribution

Term	Effects	Sum of squares	% Contribution
A-Seed	0.109	0.096	0.462
B-Peat moss	0.016	0.002	0.009
C-Fertilizer	-1.422	16.174	**78.130**
D-Water level	0.359	1.033	**4.991**
E-Water frequency	0.266	0.564	**2.727**

AB	-0.047	0.018	0.085
AC	-0.109	0.096	0.462
AD	-0.016	0.002	0.009
AE	-0.047	0.018	0.085
BC	0.047	0.018	0.085
BD	-0.047	0.018	0.085
BE	-0.078	0.049	0.236
CD	-0.172	0.236	**1.142**
CE	0.109	0.096	0.462
DE	-0.047	0.018	0.085
ABC	0.359	1.033	**4.991**
ABD	-0.172	0.236	**1.142**
ABE	0.109	0.096	0.462
ACD	0.203	0.330	**1.594**
ACE	0.047	0.018	0.085
ADE	0.078	0.049	0.236
BCD	0.047	0.018	0.085
BCE	0.016	0.002	0.009
BDE	-0.016	0.002	0.009
CDE	-0.016	0.002	0.009
ABCD	-0.078	0.049	0.236
ABCE	0.203	0.330	**1.594**
ABDE	0.109	0.096	0.462
ACDE	-0.016	0.002	0.009
BCDE	0.016	0.002	0.009
ABCDE	0.016	0.002	0.009

The following figure 2 shows the normal plot of effects with the significant factors.

The normal plot of effects suggests that only factors C, D, E, CD, ABC, ACD, ABD and ABCE appear to be significant. The rest of the factor effects might be treated as errors.

For the ANOVA, our initial model (full factorial) does not contain any error terms, we use the normal probability plot of effects as a good indicator for errors.

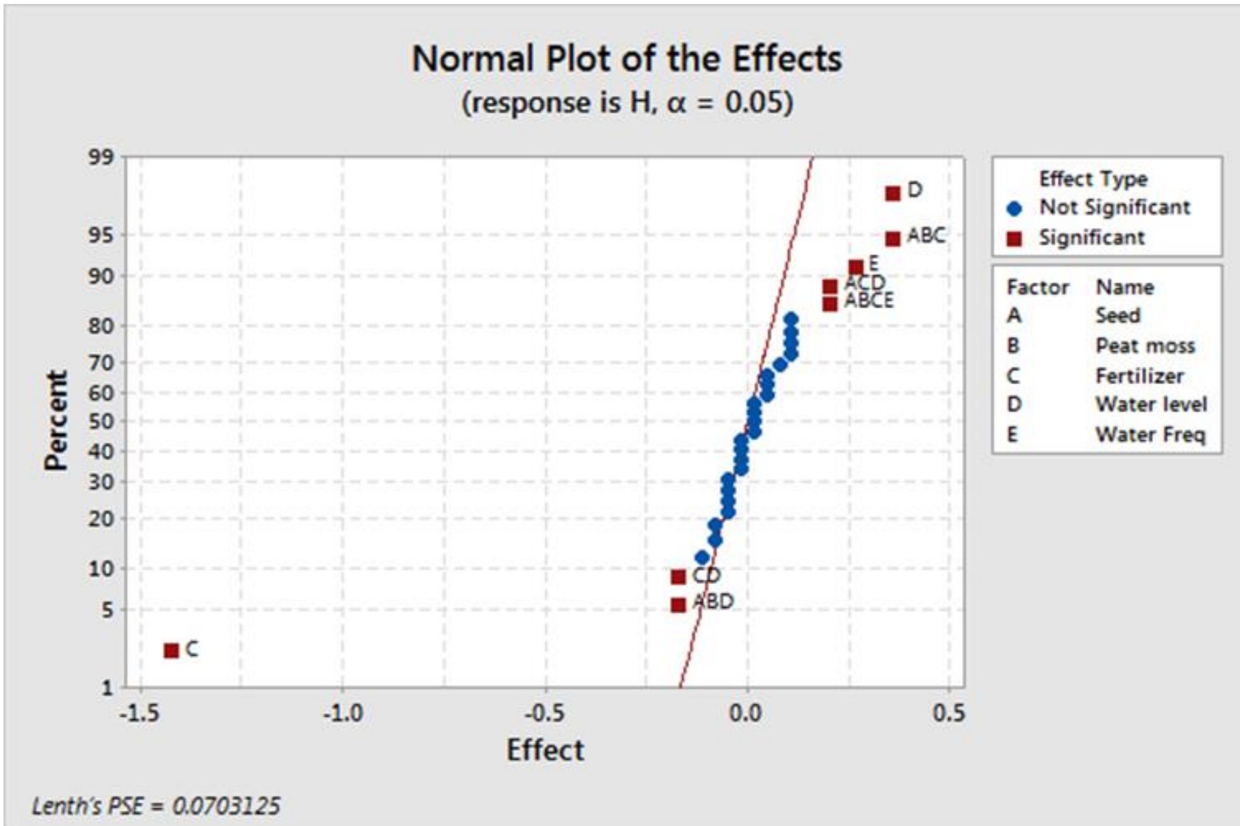

Figure 2. Normal probability plot of effects

The 1st reduced model is shown below considering all the main effects and the suggested interactions between them.

1st reduced model

Response 2 Height after 10 days
ANOVA for selected factorial model
Analysis of variance table [Partial sum of squares - Type III]

Source	**Sum of Squares**	**df**	**Mean Square**	**F Value**	**p-value Prob > F**
Model	20.04	10	2.00	63.17	< 0.0001
A-Seed	*0.096*	*1*	*0.096*	*3.02*	*0.0970*
B-Peat moss	*1.953E-003*	*1*	*1.953E-003*	*0.062*	*0.8064*

C-Fertilizer	*16.17*	*1*	*16.17*	*509.97*	*< 0.0001*
D-Water level	*1.03*	*1*	*1.03*	*32.58*	*< 0.0001*
E-Water frequency	*0.56*	*1*	*0.56*	*17.80*	*0.0004*
CD	*0.24*	*1*	*0.24*	*7.45*	*0.0126*
ABC	*1.03*	*1*	*1.03*	*32.58*	*< 0.0001*
ABD	*0.24*	*1*	*0.24*	*7.45*	*0.0126*
ACD	*0.33*	*1*	*0.33*	*10.41*	*0.0041*
ABCE	*0.33*	*1*	*0.33*	*10.41*	*0.0041*
Residual	0.67	21	0.032		
Cor Total	20.70	31			

Factors A and B are not significant and therfore will not be considered in the final model shown below:

Final reduced model

Response 2 Height after 10 days

ANOVA for selected factorial model

Analysis of variance table [Partial sum of squares - Type III]

Source	**Sum of Square**	**df**	**Mean Square**	**F Value**	**p-value Prob > F**
C-Fertilizer	*16.17*	*1*	*16.17*	*487.12*	*< 0.0001*
D-Water level	*1.03*	*1*	*1.03*	*31.12*	*< 0.0001*
E-Water frequency	*0.56*	*1*	*0.56*	*17.00*	*0.0004*
CD	*0.24*	*1*	*0.24*	*7.12*	*0.0137*
ABC	*1.03*	*1*	*1.03*	*31.12*	*< 0.0001*
ABD	*0.24*	*1*	*0.24*	*7.12*	*0.0137*
ACD	*0.33*	*1*	*0.33*	*9.94*	*0.0045*
ABCE	*0.33*	*1*	*0.33*	*9.94*	*0.0045*
Residual	0.76	23	0.033		
Cor Total	20.70	31			

In the above ANOVA table, only the significant factors are considered and other factors are discarded from the model and put as errors.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.

In this case C, D, E, CD, ABC, ABD, ACD, ABCE are significant model terms.

The "Pred R-Squared" of 0.9286 is in reasonable agreement with the "Adj R-Squared" of 0.9503.

Final Equation in Terms of Coded Factors:

Height after 10 days=

+1.84
-0.71 * C
+0.18 * D
+0.13 * E
-0.086 * C * D
+0.18 * A * B * C
-0.086 * A * B * D
+0.10 * A * C * D
+0.10 * A * B * C * E

The following table shows the normal plot of residuals.

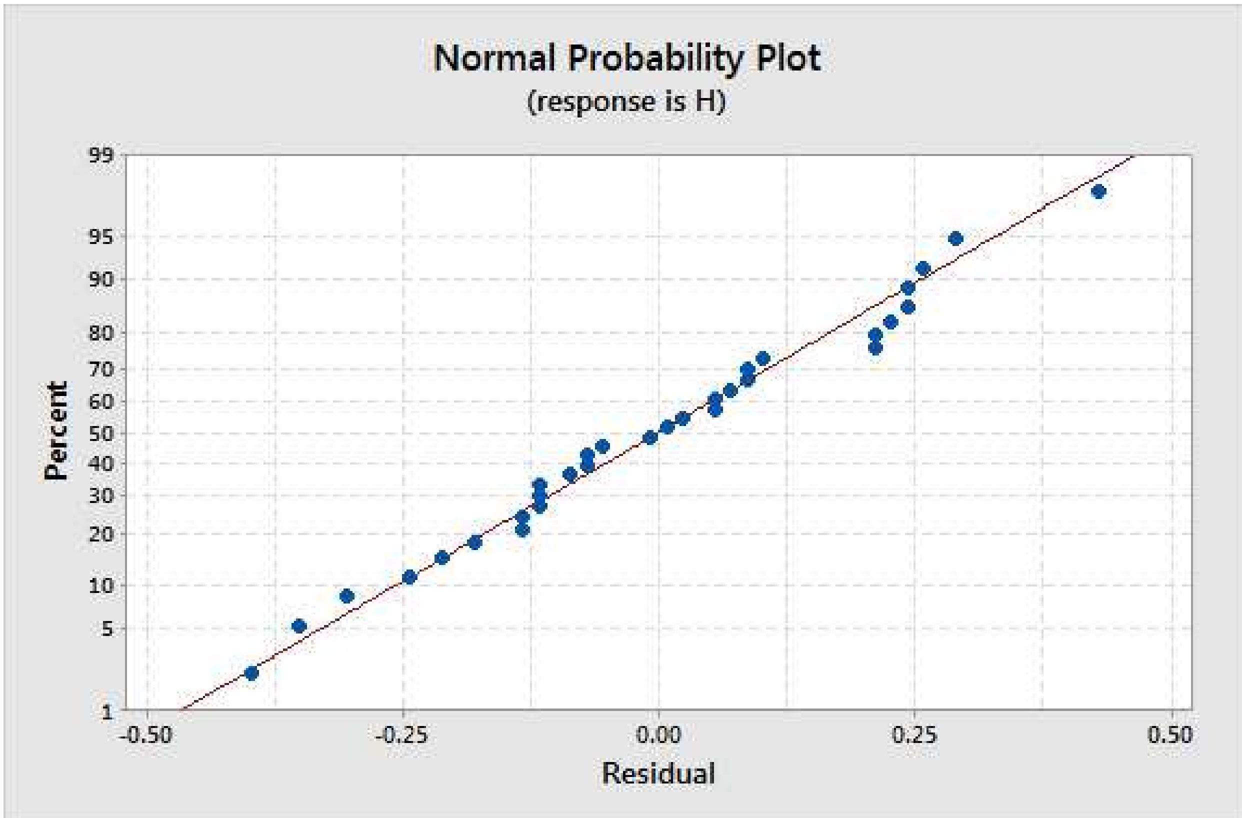

Figure 3. Normal probability plot of residuals

As the above plot shows that, there is no abnormality in the residuals and the normality assumption of residuals with a mean of 0 and variance of σ^2 are satisfied.

Discussion and interpretation of results:

The normal plot of effects shows not all main factors are significant. The only factors that happen to be significant are C (fertilizer factor), D (water level factor), E (water frequency factor), CD (the interaction of fertilizer and water level), ABC (the interaction of seed and peat moss and fertilizer), ACD (the interaction of seed and fertilizer and water level), ABD (the interaction of seed and peat moss and water level) and ABCE (the interaction of seed and peat moss and fertilizer and water freq). Factor A (seed) and factor B (peat moss) do not appear to be significant. Although seed and peat moss as main factors are not significant, their interactions with other significant factors happen to be significant which shows that other factors affect the non-significant factors high enough.

Further, the interaction between CD (fertilizer and water level) and the interaction between ABD (seed and peat moss and water level), are near the line in normal plot of effects. Therefore, it was unclear that if they should go to the error section or they will remain significant. In order to figure it out, we have run the first round of reduced model to observe if factor CD and ABD remain significant and as it is showed in the ANOVA table they remain to be significant. Hence, we kept them in our final reduced model.

In order to investigate reliability of our model, we have used the residual plot to see if they follows normality. As it is showed in the residual graph, almost all of them are plotted near the line proving this fact that residuals follows normality. Therefore, we can conclude that we were consistent in our analysis and our model is reliable. Also it proves that our dependent variable, which is set as "Height of Grass within 10 days", is closely correlated with our independent variables.

The team's final recommendation is to use cheap seed, with no enhancements. Fertilizer, at least at the level used in this study, appears to stunt the growth of the tender seedlings. Water is the most significant factor that the cultivator can directly control. In the absence of natural precipitation, the cultivator should monitor the soil wetness and irrigate as necessary to maintain adequate moisture levels for the grass to develop. This study does not recommend any specific watering schedule as the needs will vary by local climate, season, weather, sun exposure and possibly other uncontrollable factors that may need to be responded to.

References:

Connor, W., & Zelen, M. (1959). Fractional Factorial Experiment Designs for Factors at three levels. *Applied Mathematics Series, 54*.

Feili, H. R., Ahmadian, P., & Rabiei Hosseinabad, E. (2014). Life Cycle Assessment Of Municipal Solid Waste Systems To Prioritize And Compare Their Methods With Multi-Criteria Decision Making. *The Open Access Journal of Resistive Economics (OAJRE), 39*.

Foley, J., Defries, R., Asner, G., Barford, C., Bonan, G., Carpenter, S., . . . N., R. (2005). Global Consequence of landuse. *Science, 309*, 570-574.

Gencel, I. (2007). Very Additive Using the Taguchi Method and Alcohol Problems in the Optimization of an Application in the Industry. Kocaeli: Kocaeli University.

Lawn Starter. (2017, May 2). *U.S. Lawn Care Industry Statistics*. Retrieved from Lawnstarter.com: https://www.lawnstarter.com/lawn-care/lawn-care-industry-statistics

Moraga, R. J., & Rabiei Hosseinabad, E. (2017). A System Dynamics Approach in Air Pollution Mitigation of Metropolitan Areas with Sustainable Development Perspective: A Case

Study of Mexico City. *Journal of Applied Environmental and Biological Sciences, 7*(12), 164-174.

Polycarpou, L. (2017, May 3). *Water The Problem of Lawns*. Retrieved from State of the Planet: http://blogs.ei.columbia.edu/2010/06/04/the-problem-of-lawns/

Rabiei Hosseinabad, E., Feili, H., Ahmadian, P., Majidi, B., & Karimi, J. (2015). The Economic Analysis between Wind Energy and Biogas Energy to Determine Economic Policy in the Renewable Energy Systems in Iran. *1st International Congress On the Development of Agricultural Science and Natural Resources.* Warsaw, Poland.

Roy, R. (2001). *Design of Experiments using the Taguchi Approach.* Wiley-Interscience Publication.

Scotts. (2017, May 2). *Scotts Turf Builder Grass Seed Sun & Shade Mix*. Retrieved from Scotts.com: http://www.scotts.com/smg/goprod/turf-builder-sun-shade-grass-seed/prod10300003