Research on Influencing Factors of Land Rental Prices for Alfalfa Planting in Minnesota

Huyi Xiong

Faculty of Arts and Science, University of Toronto, Toronto, Canada
huyi.xiong@mail.utoronto.ca

Abstract

Agriculture is essential for human beings to survive. It not only provides food to eat and feed but also brings profits through exportation. Not all people own their lands, so they have to rent for planting. This study aims to analyze the factors contributing to the overall rental prices for alfalfa planting. It investigated the average rental prices of lands planting alfalfa in Minnesota under R package alr4 with 67 observations in the 1970s. Based on the pairwise correlation and scatterplot matrix, this paper suggested a simple linear regression model as a startup. After analyzing four diagnosis plots, the initial model failed the constant variance assumption. Then this paper built a new linear model containing all variables and their interactions. This new model produced the exact model under backward elimination AIC and BIC methods. A comparison of the initial model to the final model under ANOVA also had evidence supporting the final model. The average specialization rent is positively associated with the average rent for all tillable lands, density of cattle and pasture percentage; negatively associated with the interactions between the tillable and pastures as well as between the cattle and the fields. This study demonstrates a model available projecting the future rents as the changes in its predictors. It brings out an overview to farmers for budget preparation and land allocations.

Keywords: *linear regression; model selection; agricultural land rental price; alfalfa planting.*

1.INTRODUCTION

1.1. Research Background and Motivation

From the outset of civilization, food, fiber, and other agricultural commodities provided the basis for civilization. In the absence of sufficient food, people would be forced to perform food production in order to survive and would not have the bandwidth to develop their thinking [9]. Also, it brings significant incomes through exportation. Our data was collected in Minnesota. Minnesota's early economic foundation was land development. Minnesota farmers have started to focus on diversified crops and animal husbandry since the 1870s. Dairy farms have played a crucial role in this diversification process. These efforts were designed to maintain profitability. However, growth has been slowly occurring due to an enormous amount of investment requirements and the updates to the farming techniques. In addition to increasing the population, the formation of milk cooperatives after two decades also contributed to the spread and dominance of dairy industries. Alfalfa's ability to withstand cold and drought also stimulated the dairy industry in Minnesota as well [3].

The plant, alfalfa (also called Medicago sativa), is selected as the main subject of this research. It has been a staple crop of dairy cattle for centuries. Its nutritional advantages made it a preferred source of vitamins, minerals, and protein, as compared to alternative sources of feed [4]. Alfalfa production ranks fourth among all crops in the United States, with 18 million acres each year. For the year 2018, it surpassed wheat as the third most valuable crop in the nation. America has relied on alfalfa extensively for livestock and transportation [2]. Land rent is not only an economic problem, but also a social problem. Objective aspects like weather conditions and soil fertility, and subjective aspects like capitalization and government policies are both devoted to the changes in land rental prices. And these prices would effectively affect the incoming crop production amount by farmers' budgeting and city property structures. Thus, the projection of land rent is necessary for farms planting alfalfa, and this study would focus on the observations in Minnesota to project the local land rent.

G. Vilas Bhau et al. (Eds.): MSEA 2022, ACSR 101, pp. 741–746, 2023.
https://doi.org/10.2991/978-94-6463-042-8_106

1.2. Literature Review

A farm's potential for production can be derived from the price and value of agricultural land. In the year 2006-2007, the value of land has shown rapid development, having risen in almost all European countries. Increases in the market prices are beneficiary to the landowner but trouble the tenants [7]. Therefore, it is crucial to identify which aspects influence the agricultural land. This study would analyze and conclude potential factors to rental prices under common situations from previous papers. From the countries' perspective, Wang explored the alfalfa planting area in China using a naïve empirical model. In the article, crop prices and percentage of arable land occupied by people influence cultivation acreage positively, whereas transportation costs and the price of the crop alternatives would decrease production. As a result of the 2012 alfalfa subsidy policy, the alfalfa cultivation area has significantly declined [11]. The study, however, included only a few observations from primary sources. Most of the data were inducted from other industry profiles. To generate a more consistent result, we need a more informative dataset. Andrew and his colleagues conducted a stochastic model involving three thousand regions around the U.S. They discovered that both agricultural and non-agricultural benefits from renting would contribute to the land price valuation [6]. This research contains a tremendous amount of research objects, but they are too widely spread. Since land conditions vary among regions, it might not be appropriate for a smaller county.

On top of that, to maximize a single net acre of profitability, David and Angelos built a linear model on the statistical data from a specific farm. It demonstrates the importance of landscape managing skills over alfalfa planting [1]. Also, lime is a kind of fertilizer to combine the acidity and alkalinity of the soil. Alfalfa is best adapted to soils with a pH range between 6.5 and 7.5. Before planting, the pH of the soil needs to be maintained. Farmers would add lime to a low pH land for land sustainability [5]. It is necessary to contain lime in the research. For improvements in our study, data from primary sources would be used for credibility. It would focus on the land rent in a specific region – Minnesota. Then apply a linear model to establish a relationship between the land rent and the other potential predictors, including lime. Since alfalfa production is positively associated with the area of tillable land, we might expect a positive relationship on the average rent between tillable land and alfalfa planted land.

1.3. Research Contents and Framework

This article analyzes the land rental prices for land rent in Minnesota for alfalfa planting in the last century. It uses a simple linear regression model as an initial model. Then check four model diagnoses to figure out possible modifications to the model. A full model with all variables and interaction terms is proposed. With backward AIC and BIC elimination, the final model is presented. Then apply the ANOVA method to the initial and the final model to see whether the final model is supported.

The structure would be as follows: Part 1 Introduction: the importance of agriculture; Part 2: methodology - data characteristics and potential models; Part 3: discussion and results; Part 4: conclusion.

2.METHODOLOGY

2.1. Data Description

The data to be analyzed would be land rent, coming from primary sources, which is under R package alr4. Land rent is a data frame containing a total of 67 observations. It examines the variation in the rent paid (in U.S. dollars) in the state of Minnesota by counties for alfalfa-planted agricultural land [10][12]. Douglas Tiffany collected the data in 1977 (2014). First of all, to prevent and avoid technical issues that might damage our original data, we made a copy of the dataset and saved it as landrent1. Then, to make the data more meaningful, we renamed the columns according to their representation, as shown in Table 1.

Table 1: Data description.

Variable	Representation	New variable name
X_1	the rents among all tillable land on average	tillable
X_2	the number of dairy cattle per square mile	cattle
X_3	the pasture area as a percentage of total farmland	pasture
X_4	whether lime is a necessity for alfalfa planting, where X_4 equals 1 when lime is needed and X_4 equals zero when lime is not needed	lime
Y	the rents per acre for alfalfa planting land on average	avg_rent

Next, this paper describes the eigenvalues of each variable, and the results are shown in Table 2.

Table 2: First four rows of land rent dataset.

	tillable	cattle	pasture	lime	avg_rent
1	15.50	17.25	0.24	0	18.38
2	22.29	18.51	0.20	1	20.00
3	12.36	11.13	0.12	0	11.50
4	31.84	5.54	0.12	1	25.00

Table 2 presents the first four rows of land rent dataset. We denoted that tillable, cattle, pasture, and avg_rent are numerical and continuous variables since they are infinitely uncountable. lime is a categorical and nominal variable since it is unordered and describes two scenarios where 1 stands for lime requirement and 0 stands for no lime requirement. Meanwhile, these variables can also be classified based on dependency. We treated the response variable being avg_rent and the independent variables being the rest.

2.2. Data Analysis

As seen from the histogram of average rent per acre planted to alfalfa in Figure 1, there is a long-tailed distribution in the right positive direction. And the mean is also to the right of the peak. This means we have a right-skewed distribution of Y. The minimum value of the dependent variable of this paper is 5, the maximum value is 99.17, the mean is 42.17, and the median is 39.17.

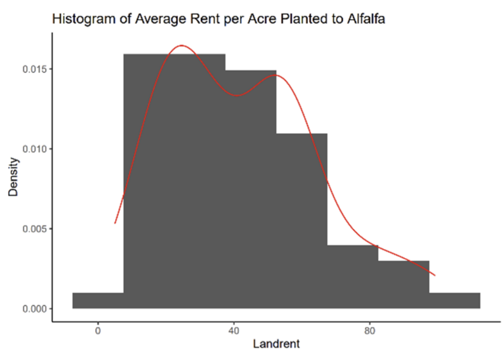

Figure 1: Histogram of average rent per acre planted to alfalfa.

To facilitate further data analysis, this paper examines whether data cleaning is required. It is worth noting that the current dataset is already a data frame and has no missing values.

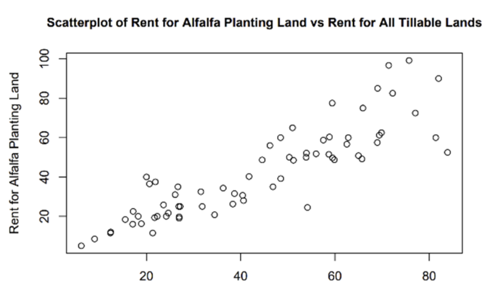

Figure 2: Scatterplot of rent for alfalfa planting land vs rent for all tillable lands.

Meanwhile, by looking at the scatterplot of rent for alfalfa planting land by the rent for all tillable lands in Figure 2, there are no obvious outliers or influential points. Thus, no cases would be removed. This scatterplot shows a relatively strong, linear, positive association between rents for alfalfa planting and rents for all tillable lands.

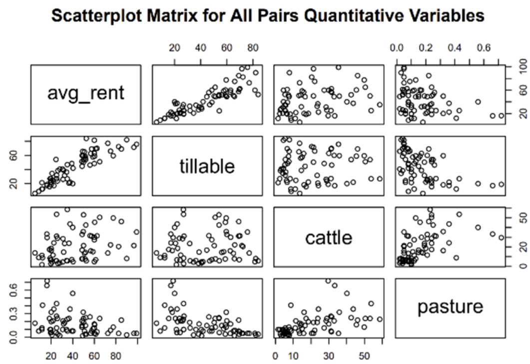

Figure 3: Scatterplot matrix for all pairs quantitative variables.

Further, the scatterplot matrix of all variables is described in this paper, as shown in Figure 3.

Table 3: Model evaluation metrics with its corresponding selection criterion.

	avg_rent	tillable	cattle	pasture
avg_rent	1.00	0.88	0.31	-0.32
tillable	0.88	1.00	0.05	-0.50
cattle	0.31	0.05	1.00	0.52
pasture	-0.32	-0.50	0.52	1.00

Moreover, Table 3 figured out the pairwise correlation to see how each quantitative predictor ranks. From the highest to the lowest for absolute correlation to the sale price, the results are avg_rent, pasture and cattle. Also, their relationship looks linear. We noticed there might be covariances between explanatory variables due to a moderate negative relationship between pasture and tillable, and a moderate positive association between pasture and cattle. In this case, we may suggest their covariates in the following modelling process.

2.3. Model Fitting and Analysis

This study fits statistical models to address the problem. It proposes an additive simple linear regression model as a startup. This starting model includes every single variable we have in the original dataset, and we do not include any of their covariates, noted as model 1.

$$\text{avg\_rent} = \beta_0 + \beta_1 \text{tillable} + \beta_2 \text{cattle} + \beta_3 \text{pasture} + \beta_4 \text{lime} + \epsilon \qquad (1)$$

Table 4: Summary table.

	Estimate	P-value
(Intercept)	-2.83	0.55
tillable	0.88	< 0.01
cattle	0.43	< 0.01

pasture	-11.38	0.34
lime	-1.01	0.72
Residual standard error: 3.91 Mean square error: 86.69 Multiple R-squared: 0.84		

Table 4 uses a 0.05 benchmark for the significance test. Variables tillable and cattle possess P-values less than 0.05, which indicates that these predictors are significant. Both of them have a strong, positive linear relationship with the average rent for alfalfa planting. For the density of cattle, when it increases by 1 unit, the land rent for alfalfa planting will increase by 0.43 dollars on average. The increasing speed is doubled for a 1 unit increase in average rent for all tillable land. The mean square error, equals the square of residual standard error, which by calculation, is approximately 86.69. R^2 gives the percentage of variation in average rent for alfalfa planting explained by the regression line. The summary table shows that the model explains around 84 percent of variations.

Table 5: 95% Confidence intervals.

	2.5% Lower bound	97.5% Upper bound
(Intercept)	-12.17	6.52
tillable	0.75	1.02
cattle	0.22	0.65
pasture	-35.16	12.39
lime	-6.71	4.68

Table 5 presents the 95% confidence intervals, showing the probability that a true parameter will fall between the 2.5% lower and 97.5% upper bound. Only tillable and cattle do not include zero in their confidence intervals. Although this paper presents a simple linear regression with all variables (Model 1), it still needs to determine whether this model is trustworthy.

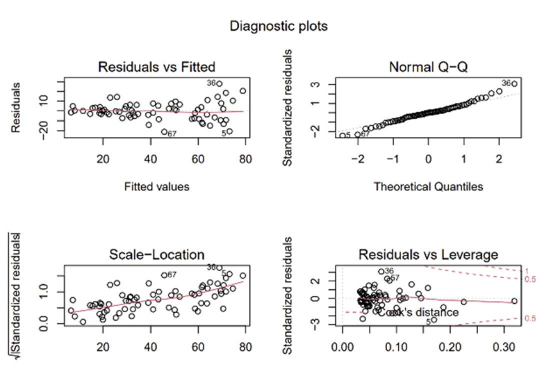

Figure 4: Model diagnosis.

Figure 4 checks the model diagnosis using four plots. Firstly, the Residuals vs Fitted plot checks the assumption that a simple linear model can be used. There is no trend in this plot, and all points are randomly scattered around a horizontal line. And the red line is nearly horizontal. This tells that the model has a linear relationship. Secondly, the Normal Q-Q plot. We're checking the assumption that the errors do not correlate. We see most points are on the dashed 45-degree line, which means the residuals of the regression are normally distributed. In other words, the normal error MLR assumptions are satisfied. This means we might propose a multi-linear regression as well. Thirdly, Scale-Location plot. We're checking the assumption that the error variances are constant. This plot has a noticeable trend: the residuals spread widely along the x-axis, and the red line is not horizontal. This means the homoscedasticity assumption does not hold.

Last but not least, the Residuals vs Leverage plot checks the assumption of the normal distribution of errors. There are no points within the upper-right corner nor the lower-right corner, indicating there are no influential points against our regression. Due to the violation of the linear regression assumption, we then added three covariates for more explanations according to the moderate correlations found in Table 3, noted as model 2.

$$\text{avg\_rent} = \beta_0 + \beta_1 \text{ tillable} + \beta_2 \text{ cattle} + \beta_3 \text{ pasture} + \beta_4 \text{ lime} + \beta_5 \text{ pasture:tillable} + \beta_6 \text{ pasture:cattle} + \beta_7 \text{ pasture:cattle:tillable} + \epsilon \quad (2)$$

Thus, this paper used the backward elimination method toward model 2 to select the best model.

Table 6: Backward elimination method.

Start	AIC = 317.56 Model 2 (Equation 2)
Step	AIC = 309.94 $\text{avg\_rent} = \beta_0 + \beta_1 \text{ tillable} + \beta_2 \text{ cattle} + \beta_3 \text{ pasture} + \beta_5 \text{ tillable:pasture} + \beta_6 \text{ cattle:pasture} + \epsilon \quad (3)$
(Intercept)	-12.84
tillable	0.99
cattle	0.81
pasture	54.41
tillable:pasture	-1.12
cattle:pasture	-1.58

3.RESULTS AND DISCUSSION

3.1. Model Determination

The final model (model 3) for backward AIC is as follows:

estimated avg_rent = -12.84 + 0.99 tillable

+ 0.81 cattle + 54.41 pasture

-1.12 tillable:pasture

-1.58 cattle:pasture (4)

The result is consistent with the model summary in Table 4, since both significant predictors are the same in the final model; it is also consistent with the pairwise correlation analysis for Table 3 and Figure 3, where the two potential pairs of covariates were kept. The final regression model for backward BIC is precisely the same as backward AIC. Then we compared the initial model used in selecting with its final model using ANOVA.

Table 7: Analysis of variance table.

	Res.Df	RSS	Df	Sum of Sq	F	P-value
Model 2	61	4694.1				
Model 3	59	4639.7	2	54.43	0.35	0.71

As can be seen from Table 7, the P-value is greater than 0.05, which indicates that a simpler model was chosen, supporting our analysis above. Therefore, the model after back-elimination is the best model for our study. Based on our research, model 3 is the best since it passes several verification methods: AIC, BIC and ANOVA.

3.2. Discussion

Similar results could be retrieved on a broader market in the whole American and even global markets. From the research by Daniel and John, with an increasing demand for milk, a potential increase would happen in the alfalfa prices [8]. Since alfalfa is a leading source of feed, if more milk is required by the market, then the density of dairy cows would increase. Also, more alfalfa would be necessary to feed the dairy cows. In this case, farmers would be eager to plant alfalfa to gain profits, and even non-agricultural business people would like to fulfil the alfalfa planting market. However, the total land area available for planting is limited. People then would compete to offer a higher rental price due to the market crazes. Therefore, it indirectly supports a positively associated relationship between cattle density and the land rent for alfalfa planting as well.

4.CONCLUSION

4.1. Main Findings

Under backward elimination methods, AIC and BIC select the same final model. As we only had four variables, BIC would not have a heavier penalty than AIC, which results in the same model. The final model suggests the basic average rent is -12.84 if all predictors do not involve. Holding other variables constant, when the average rent for all arable land increases by 1 unit, the average rent for alfalfa cultivation increases by an average of 0.99. Similarly, the density of dairy cows brings an increasing ratio of 0.81, and the proportions of pasture get that of 54.41 on average. However, although three variables have positive relations with the average rent, this does not apply to their interaction terms. The two interactions: between rents for all tillable lands and pasture proportion, and, between cattle density and pasture proportions, are negative, at a ratio of -1.12 and -1.58, respectively.

4.2. Future Studies

Although the model of this paper is efficient and stable, there are some limitations. Firstly, the data points are averaged, so they may not provide a precise projection. As we have mentioned in the data details part: the data of average rent per acre planted to alfalfa spreads widely, ranging from 5 to 99 dollars and 17 cents. The averaged data might not reflect a direct relation between factors to the exact price. Instead, we can only get an average cost, which might have the potential to be high above our rental budget. Also, there are only a few variables or factors collected in the dataset. Four elements are not enough to project the rental price among the whole market. And more factors like weather conditions and economic conditions. Therefore, for further research, this paper suggested displaying individual rent for each household on top of the average rent. Also, consider more factors that affect the rental price.

More importantly, according to the literature review, we suggest future investigators explore whether the average rents for all tillable land and land planted for alfalfa for lime needed and limed not needed have overlapped. Last but not least, for business purposes, this article suggested that the market provide land rental price insurance and land rent price promotions that target securing the land price. This can benefit both the company and the customers at the same time.

REFERENCES

[1] Debertin, D. L., & Pagoulatos, A. (1985). Optimal Management Strategies for Alfalfa Production Within a Total Farm Plan. *Journal of Agricultural and Applied Economics*, *17*(2), 127-138.

[2] Fernandez, A. L., Sheaffer, C. C., Tautges, N. E., Putnam, D. H., & Hunter, M. C. (2019). Alfalfa, wildlife, and the environment. *National Alfalfa and Forage Alliance*.

[3] Granger, S., & Kelly, S. (2005). *Historic context study of Minnesota farms, 1820-1960*. Minnesota Department of Transportation.

[4] Jones, T. (2021, September 29). *Alfalfa - Benefits, Nutrition Facts & Side Effects*. Healthline. https://www.healthline.com/nutrition/alfalfa

[5] Lollato, R., & Min, D. (n.d.). *Agronomy eUpdate August 21st, 2020: Issue 815*. Eupdate.agronomy.ksu.edu. Retrieved May 6, 2022, from https://eupdate.agronomy.ksu.edu/article_new/tips-for-fall-planting-of-alfalfa-403-6#:~:text=Alfalfa%20grows%20best%20in%20well

[6] Plantinga, A. J., Lubowski, R. N., & Stavins, R. N. (2002). The effects of potential land development on agricultural land prices. *Journal of urban economics*, *52*(3), 561-581.

[7] Střeleček, F., Lososova, J., & Zdeněk, R. (2010). The relations between the rent and price of agricultural land in the E.U. countries. *Agricultural Economics*, *56*(12), 558-568.

[8] Sumner, D. A., & Rosen-Molina, J. T. (2011). Alfalfa in the context of global crop price prospects. In *Western Alfalfa and Forage Conference* (pp. 11-13).

[9] Tauger, M. B. (2010). *Agriculture in world history*. Routledge.

[10] Tiffany, D. G., Thomas, K. H., & Hasbargen, P. R. (1980). *Land Rental Arrangements for Minnesota Farmland* (No. 1690-2016-137313).

[11] Wang, W. (2020). Improving China's alfalfa industry development: an economic analysis. *China Agricultural Economic Review*.

[12] Weisberg, S. (2014). *Applied Linear Regression*, 4th edition. Hoboken NJ: Wiley.