TMA4267 Linear Statistical Models
3-factor 2-level experiment of the rheological properties of noodles
NTNU

Written by
Alexander Edward Hatle
Ole Gunnar Røsholt Hovland

2023

Abstract

This study aimed to investigate the effects of cook time, washing soda concentration, and water volume on the softness of homemade noodles using a Design of Experiments (DoE) approach. A full factorial experimental design was employed to examine the main effects and interactions of the factors. A linear regression model was fitted to the experimental data, and its goodness of fit was assessed. However, the model was found to be a poor fit for the data, with insignificant predictor variables and interactions, low R-squared and adjusted R-squared values, and a large residual standard error. Several sources of error were identified in the experiment, potentially impacting the accuracy and reliability of the results. The study concluded that the current model does not provide meaningful insights into the effects of cook time, washing soda concentration, and water volume on noodle softness. Further investigation, with careful consideration of experimental error sources and conducting multiple trials, is necessary to optimize the noodle-making process and better understand the relationships between these factors and noodle softness.

Contents

1 Introduction **1**
1.1 Selection of factors . . . 1
1.2 Selection of response variable . . . 2
1.3 Selection of design . . . 2
1.4 Experimental design . . . 3

2 Analysis of data **4**
2.1 Fitting a linear model . . . 4
2.2 Analysis of plots . . . 5

3 Error sources **6**

4 Conclusion **7**

1 Introduction

Noodle dishes are enjoyed worldwide, and their texture plays a crucial role in their overall culinary experience. One of the key attributes of noodle texture is their rheological properties, i.e., the "softness" or "chewyness". , which can vary depending on the cooking process, ingredients, and recipe. In this study, we aim to systematically investigate the effects of three cooking parameters – cook time, washing soda concentration, and water volume – on the softness of homemade noodles using a Design of Experiments (DoE) approach.

DoE is a powerful statistical technique that allows researchers to efficiently plan and analyze experiments by considering the impact of multiple factors on an output or response variable. In this case, the response variable is the force required to cut through a noodle, which serves as a proxy for noodle softness. By selecting appropriate factors and levels, we can study these factors' individual and interactive effects on the response variable and optimize the noodle-making process to achieve the desired softness.

The measure of chewyness will be done experimentally by putting a noodle on a digital kitchen weight and slowly applying pressure with a dull knife while filming the weight to pinpoint how much force was needed to cut through the noodle.

1.1 Selection of factors

We have chosen to perform a 3-factor 2-level experiment. The three factors we will look at are how much water to add to the dough, how much washing soda to add, and how long we cook the noodles.

Table 1.1: Factors used in the experiment

	A	B	C
1	4 minute cooktime	1.5 tsps washing soda	100 mL water
-1	2 minute cooktime	0.5 tsps washing soda	70 mL water

The reasoning for selecting these factors and their levels can be summarized as follows:

1. Cook time: Cooking time can significantly impact the softness of the noodles, as it affects the degree of hydration and gelatinization of the starch in the dough. Shorter cook times may result in firmer, less cooked noodles, while longer ones may lead to softer, more cooked ones. The two levels of 2 and 4 minutes represent a reasonable range of cooking times for homemade noodles.
2. Washing soda: We produce washing soda (sodium carbonate) by heating baking soda (sodium bicarbonate). Washing soda is a more potent alkaline substance than baking soda and can influence the texture of noodles by affecting the gluten structure and the degree of starch gelatinization. Higher concentrations of washing soda can lead to softer, more elastic noodles, while lower concentrations may result in firmer, less tender noodles. The chosen levels of 0.5 and 1.5 tsps of washing soda provide a range of alkaline concentrations that may significantly affect noodle softness.
3. Water: The amount of water in the dough can impact the hydration of the starch and gluten proteins, affecting the noodles' texture. Less water may result in firmer, drier noodles, while more water can produce softer, more hydrated ones. The levels of 70 mL and 100 mL of water represent a range that could produce meaningful differences in noodle softness.

Prior to the experiment, we expected an interaction between A and C, that longer cooktimes with higher water content could lead to even softer noodles, as there's more time for water to be absorbed and hydrate the noodle structure. Another interaction that may occur is between A and B, that a longer cooktime might enhance the alkaline properties of washing soda, resulting in softer noodles.

1.2 Selection of response variable

To quantify noodle softness, we require a response variable that accurately reflects the textural properties of the noodles. After careful consideration, we have chosen the force required to cut through a noodle as the response variable for this experiment.

The force needed to cut through a noodle serves as a reliable proxy for noodle softness, with higher force values indicating firmer noodles and lower force values representing softer noodles. This response variable provides a quantifiable measure of the textural properties of noodles, enabling us to systematically study the effects of our chosen factors on noodle softness.

We will use a method involving a knife and a kitchen weight to measure the force required to cut through a noodle. A noodle will be placed on the weight, a knife will be positioned over the noodle, and we will slowly apply more and more weight to the knife until the noodle is cut through. We will also film the kitchen weight and record the maximum weight in grams, which we will use as our response variable.

1.3 Selection of design

For this experiment, we have chosen a 2^3 full factorial design to systematically investigate the effects of the three factors and their interactions on noodle softness. A full factorial design allows us to examine all possible combinations of the factors and their levels, providing comprehensive insights into their individual and combined influences on the response variable.

This choice of design enables us to estimate the main effects of each factor, as well as the two-way and three-way interactions among the factors, providing a comprehensive understanding of how the cook time, washing soda concentration, and water volume influence noodle softness.

To further improve the reliability of the results, we choose to perform the experiment twice. This helps account for potential variability and measurement errors, allowing us to obtain more accurate and robust estimates of the factor effects.

Finally, we apply randomization in the order of cooking. This minimizes the impact of any lurking variables or biases that could affect the response variable. By randomizing the order in which the experimental runs are performed, we can ensure that the results are not influenced by any uncontrolled factors or systematic errors.

1.4 Experimental design

The following, Listing 1, is a listing of the experimental design.

Listing 1: The experimental design. The response variable y is in grams. Note: columns run.no and run.no.std.rp are annotations. They are not part of the data frame

```
   run.no run.no.std.rp  A  B  C Blocks    y
1       1           1.1 -1 -1 -1     .1 2266
2       2           2.1  1 -1 -1     .1 3043
3       3           3.1 -1  1 -1     .1 3574
4       4           4.1  1  1 -1     .1 3369
5       5           5.1 -1 -1  1     .1 2449
6       6           6.1  1 -1  1     .1 2640
7       7           7.1 -1  1  1     .1 2795
8       8           8.1  1  1  1     .1 2474
9       9           1.2 -1 -1 -1     .2 4425
10     10           2.2  1 -1 -1     .2 4011
11     11           3.2 -1  1 -1     .2 2829
12     12           4.2  1  1 -1     .2 3149
13     13           5.2 -1 -1  1     .2 3034
14     14           6.2  1 -1  1     .2 3097
15     15           7.2 -1  1  1     .2 3067
16     16           8.2  1  1  1     .2 2860
class=design, type= full factorial
```

In Listing 1 the first replicate is assigned to Block .1, and the second replicate is assigned to Block .2. The treatments are randomly assigned within each block, so there is no systematic bias in the assignment of treatments.

The advantage of using a randomized block design is that it allows you to control for the effects of variables that are expected to influence the response variable but are not of interest in the study. By grouping the experimental units into blocks based on these variables, we reduce the amount of variability in the response variable due to these factors and improve the precision of the estimates of treatment effects.

2 Analysis of data

2.1 Fitting a linear model

The following, Listing 2, is the summary of fitting a linear model to the experimental design.

Listing 2: console output to `summary(fitted)`

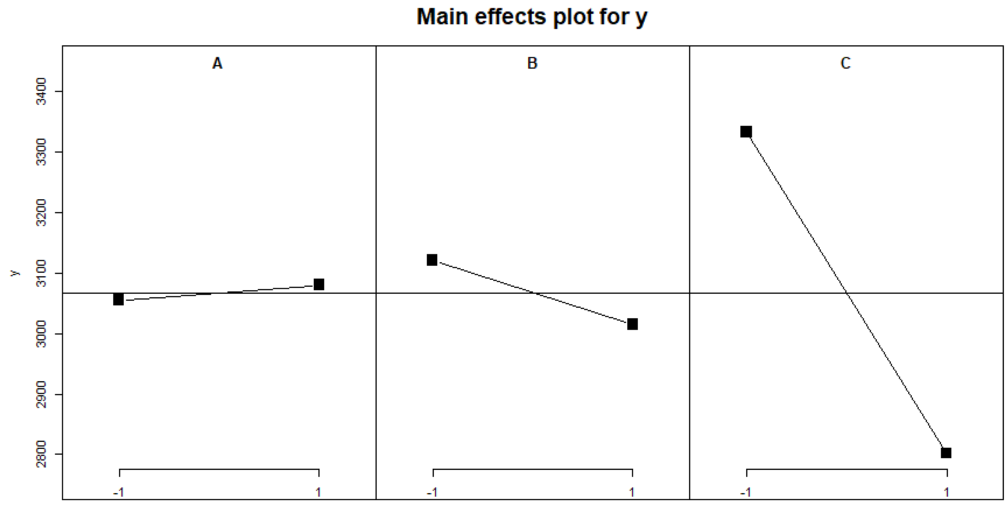
```
1  Call:
2  lm.default(formula = y ~ (A + B + C)^3, data = f2_design)
3
4  Residuals:
5      Min      1Q  Median      3Q     Max
6  -1079.5  -244.5     0.0   244.5  1079.5
7
8  Coefficients:
9              Estimate Std. Error t value Pr(>|t|)
10 (Intercept)  3067.62     165.07  18.583 7.25e-08 ***
11 A1             12.75     165.07   0.077    0.940
12 B1            -53.00     165.07  -0.321    0.756
13 C1           -265.62     165.07  -1.609    0.146
14 A1:B1         -64.37     165.07  -0.390    0.707
15 A1:C1         -47.00     165.07  -0.285    0.783
16 B1:C1          50.00     165.07   0.303    0.770
17 A1:B1:C1      -33.38     165.07  -0.202    0.845
18 ---
19 Signif. codes:  0 "***" 0.001 "**" 0.01 "*" 0.05 "." 0.1 " " 1
20
21 Residual standard error: 660.3 on 8 degrees of freedom
22 Multiple R-squared:  0.2769,    Adjusted R-squared:  -0.3557
23 F-statistic: 0.4377 on 7 and 8 DF,  p-value: 0.8535
```

In Listing 2, we use a linear regression model with the formula `y ~ (A + B + C)^3` where `y` is the response variable and `A`, `B`, and `C` are the predictor variables. The coefficients table shows each predictor variable's estimated coefficients and interactions. The intercept (when all predictor variables are at zero) (line 10) is 3067.62 with a standard error of 165.07 and a significant t-value of 18.583 (`p-value`< 0.001). Despite this, our data has a plethora of issues:

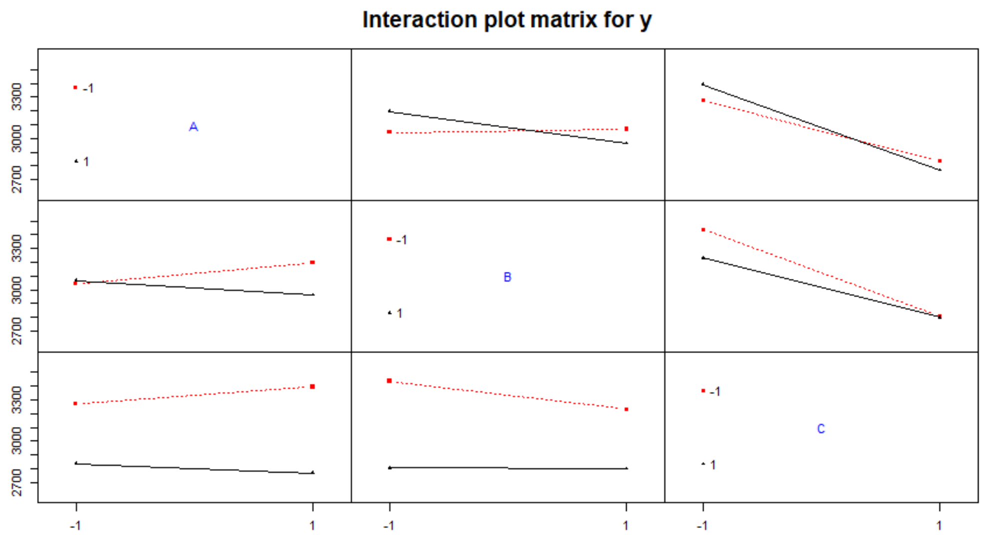
1. None of the predictor variables or their interactions are significant (p-values > 0.05), except for a marginal significance for the main effect of `C1` (`p-value` $= 0.146$).
2. The R-squared value of 0.2769 indicates that the model explains only 27.69% of the variance in the response variable, which is low. The adjusted R-squared value of -0.3557 (line 22) suggests that adding predictors to the model did not improve its fit and might be overfitting the data.
3. The F-statistic of 0.4377 (line 23) and the p-value of 0.8535 (line 23) indicate that the model is not significant as a whole and cannot reject the null hypothesis that all coefficients are zero.
4. The residual standard error of 660.3 (line 21) is relatively large compared to the range of the response variable, indicating that the model has high variability and might not fit the data well.

Overall, this model is not a good fit for the data. This is a strong indication that our data-gathering hosted too many errors.

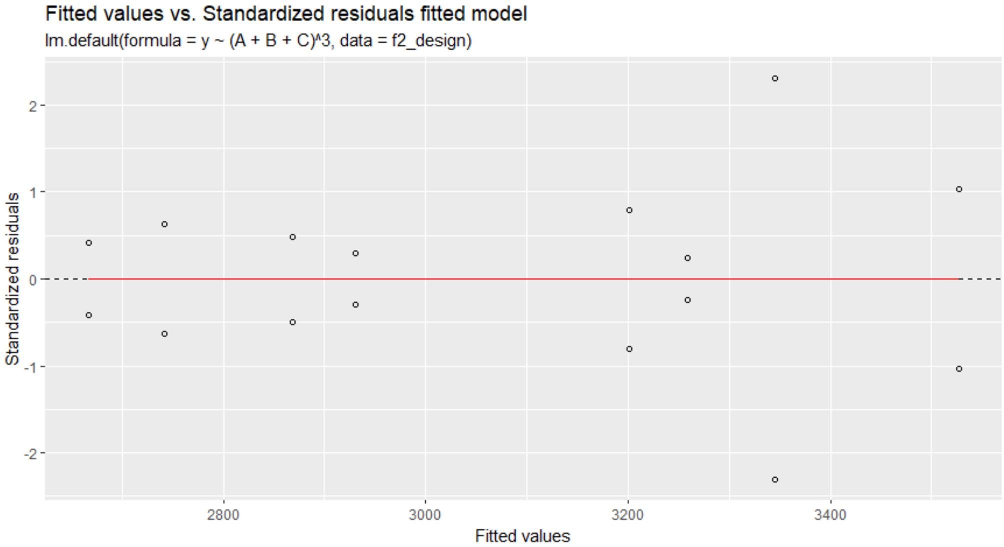
2.2 Analysis of plots



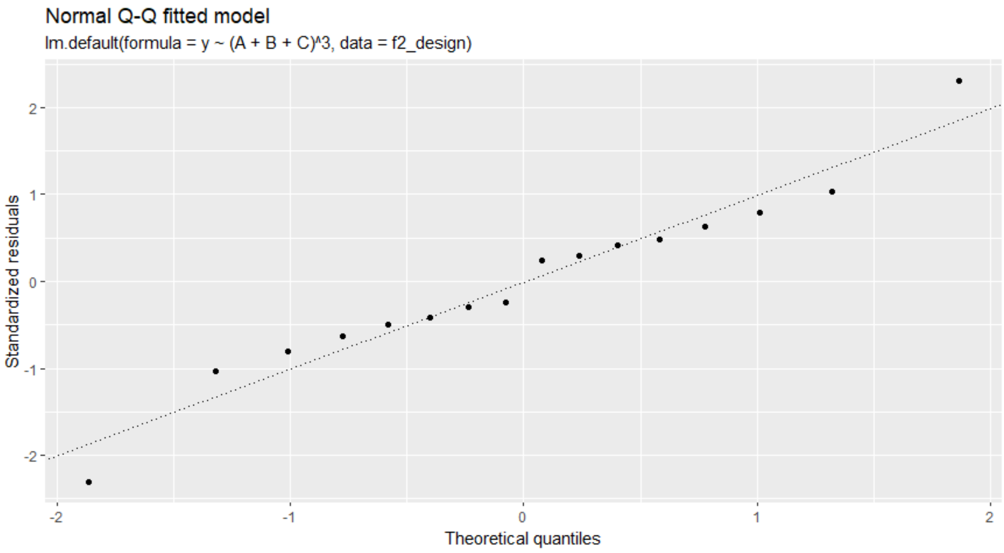
The main effects plots reveal that water volume (C) has the most substantial impact on noodle softness, followed by washing soda concentration (B), while cook time (A) appears to have the least influence. However, since the model is not a good fit, it's essential to consider potential sources of error and variability in the experiment, as well as interactions between the factors, before drawing definitive conclusions.



Most of the lines in the interaction plot matrix appear to be close to parallel, suggesting that the interaction effects between the factors are relatively small. However, the interaction between cook time (A) and washing soda (B) appears to be the least parallel, with an angle of about 10 degrees. This indicates that there may be a moderate interaction effect between these two factors on noodle softness.



We observe some outliers in this plot, which may suggest that the model does not fit the data as well for these observations and may not fully capture the underlying relationships between the factors and the response variable.



The middle points seem to follow the line relatively well, indicating that the residuals for these observations are consistent with a normal distribution. However, the first and last points are further away from the line, suggesting that there may be some deviations from normality in the residuals for these observations. This could potentially indicate the presence of outliers or influential observations in the data.

3 Error sources

In our noodle softness experiment, we identified several sources of error that may have impacted the accuracy and reliability of our results. A discussion of these error sources and their potential effects on the experiment is provided below:

1. Kitchen weight update lag: The kitchen weight used to apply force on the knife updated around 3 times per second, which may have resulted in inaccurate weight measurements. This could

have led to inconsistencies in the force applied across different experimental runs, potentially affecting the relationship between the factors and the response variable.

2. Time between cooking and cutting: There was a variable time interval between cooking and cutting the noodles, which could have affected the temperature of the noodles. Colder noodles are generally harder to cut, and the inconsistency in noodle temperature might have introduced variability in the force required to cut the noodles, regardless of the experimental factors.
3. Overestimation of applied force: We used the maximum displayed weight value during the cutting process, which might have resulted in applying more force than necessary to cut through the noodles. This overestimation of force could have affected the response variable measurements and potentially masked the actual effects of the factors on noodle softness.
4. Reusing cooking water: We did not change the water between each experiment, leading to an accumulation of free starches in the water. This may have affected the starch gelatinization and hydration of the noodles during cooking, introducing variability in the softness of the noodles that was not directly related to the experimental factors.
5. Inconsistent noodle thickness: The thickness of the noodles was estimated visually, which could have resulted in variability in the actual thickness of the noodles across different experimental runs. Inconsistent noodle thickness can impact the force required to cut through the noodles and potentially confound the effects of the experimental factors on noodle softness.
6. Pre-cooking noodle drying: The noodles may have dried to different extents before cooking, affecting their water content. Water content variation can influence the hydration of starch and gluten proteins, leading to differences in noodle softness that are not directly related to the experimental factors.

To improve the reliability and accuracy of the results in future experiments, it is essential to address these sources of error by using real-time updating kitchen weights, controlling the time between cooking and cutting, ensuring consistent noodle thickness, refreshing the cooking water between experiments, and preventing pre-cooking noodle drying. Additionally, conducting multiple trials and averaging the results can help to account for any remaining variability and further enhance the validity of the findings.

4 Conclusion

The goal of this study was to investigate the effects of cook time, washing soda concentration, and water volume on the softness of homemade noodles. We employed a Design of Experiments (DoE) approach to systematically investigate the impact of these factors and their interactions on noodle softness. Our analysis involved fitting a linear regression model to the experimental design and evaluating the model's fit to the data.

Our results indicated that the model was not a good fit for the data, and several issues were identified, including insignificant predictor variables and interactions, low R-squared and adjusted R-squared values, insignificant F-statistic and p-value, and a relatively large residual standard error. These issues suggest that the model does not provide meaningful insights into the effects of the chosen factors on noodle softness.

We also discussed potential sources of error in the experiment, which may have impacted the accuracy and reliability of our results. To improve the validity of the findings in future experiments, it is crucial to address these error sources and conduct multiple trials to account for any remaining variability.

In conclusion, this study's results do not provide definitive insights into the effects of cook time, washing soda concentration, and water volume on noodle softness. Further investigation, with careful consideration of experimental error sources, is necessary to optimize the noodle-making process and better understand the relationships between these factors and noodle softness.