1. Introduction

The focus of the original study was to develop a predictive model utilizing neural networks to accurately forecast both the air permeability of woven fabrics and the moisture content of these fabrics post vacuum drying. This research addresses a crucial aspect in textile engineering, contributing to advancements in fabric quality control and manufacturing processes. In the textile industry, a fabric's air permeability is extremely important as their characteristics, such as thermal properties, are determined by the air flow rate. These characteristics are directly related to the comfort of the textile, which then affects the acceptance and performance of the final product, as well as the cost of production. Therefore, the study equates the prediction of a fabric's behavior during vacuum drying with the prediction of its air permeability. This forms the foundation of their statistical model.

Initially, a linear regression model was constructed from the data set obtained from the experiment, with variables of weft densities, warp densities, and mass per unit area, and a regression equation was derived from the model. The study also conducts further analysis through sequential sum-of-squares, which revealed insightful findings. Then, in a secondary experiment, the original linear regression model was refined by excluding a variable, mass per unit area. Analyses were also performed using this refined model. In this paper, I will describe the study's methods of data collection and aim to replicate the analytical results.

Possible extensions could be made to the original study's model and analysis. This study conducted a secondary experiment, where mass per unit area data was excluded from the model for analysis. Another extension could be to conduct similar tertiary experiments to exclude weft density and warp density, in turn, and observe how the model changes. In this paper, I will complete these extensions to the model and analysis, adding my findings to the original study's results.

Finally, I will describe any discrepancies between the study's analyses and my results. I will also discuss whether my assumptions were reasonable. My conclusions demonstrate that this study effectively uses neural network-based predictive models in the textile industry. By accurately forecasting air permeability and moisture content of woven fabrics, manufacturers can optimize production processes and ensure product quality. Future research may explore additional variables or refine the model even further to enhance predictive accuracy and applicability across various fabric types.

2. Description of Data

The study's goal was to establish a general relation between air permeability and warp yarn density, weft yarn density, and mass per unit area. Thus, the warp yarn density, the weft yarn density, and the mass per unit area were the chosen variables for the study. The researchers chose to use these parameters as the warp and weft densities are predefined in the weaving process, and the mass per unit area can be easily and accurately measured. The aim was to use these parameters to predict air permeability values for the given fabric types. With this experimental setup, we would then have 3 independent variables with one dependent variable.

The researchers produced a set of 30 total samples of woven fabric for the experiment, and all fabric types used were made of 100% cotton yarns with plain weave structure. The warp and weft yarn linear densities vary in incremental steps with 6 different weft densities and 5 different warp densities. The air permeability of each fabric was estimated under constant pressure drop conditions using the measurement of air velocity. Each fabric underwent 5 trials, and the results from the trials were averaged to obtain an average air permeability value. The researchers reported the fabric parameters, averaged air permeability values, and the standard deviation of the measurements in a table:

```
dt <- read.csv("airpermeabilitywoven.csv")

dt2 <- read.csv("/Users/chaney/Desktop/Project /airpermeabilitywoven.csv")

library(flextable)
flextable(read.csv("/Users/chaney/Desktop/Project /airpermeabilitywoven.csv"))
```

```
## Warning: fonts used in `flextable` are ignored because the `pdflatex` engine is
## used and not `xelatex` or `lualatex`. You can avoid this warning by using the
## `set_flextable_defaults(fonts_ignore=TRUE)` command or use a compatible engine
## by defining `latex_engine: xelatex` in the YAML header of the R Markdown
## document.
```

sampNum	warp	weft	mass	airPermMn	airPermSD	waterCont
1	54	20	101.0	45.74	1.80	53.65
2	54	25	110.3	27.02	2.44	45.65
3	54	30	119.5	15.68	0.68	40.34
4	54	35	127.3	8.76	0.39	38.21
5	54	40	136.7	4.37	0.15	35.66
6	54	45	144.1	2.90	0.40	32.35
7	57	20	101.2	47.94	1.83	50.53
8	57	25	111.1	27.52	0.78	44.91
9	57	30	120.5	14.84	0.97	40.93
10	57	35	130.7	8.99	0.48	37.49
11	57	40	142.0	3.98	0.24	33.62
12	57	45	146.3	2.68	0.12	32.33
13	60	20	107.9	33.98	1.16	50.19
14	60	25	118.7	17.01	0.79	45.68
15	60	30	129.5	9.75	0.41	39.31
16	60	35	141.4	4.56	0.41	36.00
17	60	40	151.2	2.12	0.07	33.80
18	60	45	153.7	1.70	0.05	32.04
19	63	20	109.0	27.68	0.55	47.89
20	63	25	119.1	15.24	0.64	44.52
21	63	30	129.1	7.23	0.13	40.20

sampNum	warp	weft	mass	airPermMn	airPermSD	waterCont
22	63	35	148.5	3.22	0.42	35.72
23	63	40	150.3	1.89	0.08	33.40
24	63	42	154.4	1.61	0.05	32.31
25	66	20	111.2	27.28	1.68	47.77
26	66	25	123.0	14.42	1.39	43.59
27	66	30	134.0	6.90	0.21	37.66
28	66	35	144.4	3.19	0.17	37.09
29	66	40	155.0	1.65	0.06	35.69
30	66	42	160.2	1.38	0.05	29.49

From the table, air permeability values noticeably decrease when warp and weft yarn densities increase. This behavior can be expected because the pores decrease in size for tighter fabric types, which is also were the resistance to airflow would be higher.

3. Methods (or Models) and Analysis

Though the researchers suspected that the relationships between the variables were nonlinear, they still attempted to fit a multiple linear regression to the data. This linear regression would be a good starting place for further analysis and would give them an idea of how important each of the three variables were.

The data table below shows the swapped warp and weft columns, as well as the new table using the author's naming conventions of C1, C2, C3, and C4 for the variables.

```r
# Switch warp and weft columns in data table because they denoted warp as C2 and weft as C1
dt3 <- dt2[,c(1,3,2,4,5,6,7)]

# Change column names to C1, C2, C3, C4
colnames(dt3) <- c("sampNum", "C1","C2", "C3", "C4", "C4_SD", "waterCont")

flextable(dt3)
```

```
## Warning: fonts used in `flextable` are ignored because the `pdflatex` engine is
## used and not `xelatex` or `lualatex`. You can avoid this warning by using the
## `set_flextable_defaults(fonts_ignore=TRUE)` command or use a compatible engine
## by defining `latex_engine: xelatex` in the YAML header of the R Markdown
## document.
```

sampNum	C1	C2	C3	C4	C4_SD	waterCont
1	20	54	101.0	45.74	1.80	53.65
2	25	54	110.3	27.02	2.44	45.65

sampNum	C1	C2	C3	C4	C4_SD	waterCont
3	30	54	119.5	15.68	0.68	40.34
4	35	54	127.3	8.76	0.39	38.21
5	40	54	136.7	4.37	0.15	35.66
6	45	54	144.1	2.90	0.40	32.35
7	20	57	101.2	47.94	1.83	50.53
8	25	57	111.1	27.52	0.78	44.91
9	30	57	120.5	14.84	0.97	40.93
10	35	57	130.7	8.99	0.48	37.49
11	40	57	142.0	3.98	0.24	33.62
12	45	57	146.3	2.68	0.12	32.33
13	20	60	107.9	33.98	1.16	50.19
14	25	60	118.7	17.01	0.79	45.68
15	30	60	129.5	9.75	0.41	39.31
16	35	60	141.4	4.56	0.41	36.00
17	40	60	151.2	2.12	0.07	33.80
18	45	60	153.7	1.70	0.05	32.04
19	20	63	109.0	27.68	0.55	47.89
20	25	63	119.1	15.24	0.64	44.52
21	30	63	129.1	7.23	0.13	40.20
22	35	63	148.5	3.22	0.42	35.72
23	40	63	150.3	1.89	0.08	33.40
24	42	63	154.4	1.61	0.05	32.31
25	20	66	111.2	27.28	1.68	47.77
26	25	66	123.0	14.42	1.39	43.59
27	30	66	134.0	6.90	0.21	37.66
28	35	66	144.4	3.19	0.17	37.09
29	40	66	155.0	1.65	0.06	35.69
30	42	66	160.2	1.38	0.05	29.49

Thus, an initial linear regression model was constructed using data from columns C1, C2, and C3, which represent weft densities, warp densities, and mass per unit area, respectively. Column C4 contains average air permeability values. The regression equation derived from this model is C4 =

105 - (0.679) C1 - (0.392) C2 - (0.355) C3, with an impressive R-squared value of 0.84. The figure below gives the results of the multiple linear regression analysis.

```r
# Fit the multiple linear regression model
dt3_model1 <- lm(formula = C4 ~ C1 + C2 + C3, data = dt3)

summary(dt3_model1)
```

```
##
## Call:
## lm(formula = C4 ~ C1 + C2 + C3, data = dt3)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.8875 -4.2474 -0.6555  2.4051 14.7703
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 105.0349    16.3545   6.422 8.36e-07 ***
## C1           -0.6796     0.7280  -0.933    0.359
## C2           -0.3921     0.5693  -0.689    0.497
## C3           -0.3550     0.3655  -0.971    0.340
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.594 on 26 degrees of freedom
## Multiple R-squared:  0.8404, Adjusted R-squared:  0.822
## F-statistic: 45.65 on 3 and 26 DF,  p-value: 1.678e-10
```

In this section, I want to check whether there is evidence of a linear relationship between the dependent variable and the independent variable.

At $\alpha = 0.05$:

$H_0 : \beta_1 = 0$ $H_a : \beta_1 \neq 0$

From the summary table: the P-value is 0.359, which is larger, and larger than $\alpha = 0.05$. Thus, H_0 is not rejected. There is not sufficient evidence of a linear relationship between weft densities and air permeability values.

At $\alpha = 0.05$:

$H_0 : \beta_2 = 0$ $H_a : \beta_2 \neq 0$

From the summary table, the P-value is 0.497, which is larger, and larger than $\alpha = 0.05$. H_0 is not rejected. There is not sufficient evidence of a linear relationship between warp densities and air permeability values.

At $\alpha = 0.05$:

$H_0 : \beta_3 = 0$ $H_a : \beta_3 \neq 0$

From the summary table, the P-value is 0.340, which is larger, and larger than $\alpha = 0.05$. H_0 is not rejected. There is not sufficient evidence of a linear relationship between mass per unit area and air permeability values.

From the model summary table, the sample coefficient of determine (R-Square) is 84.04% . This means that 84.04% of the variation in air permeability values is explained by the variation in the weft densities, warp densities, and mass per unit area. Therefore, according to this R-Squared value, this linear model seems useful.

Since the P-value is very high compared to α, and the R-Squared value is 84.04%, we can assume that there might be collinearity in this linear model.

```r
# ANOVA
anova(dt3_model1)
```

```
## Analysis of Variance Table
##
## Response: C4
##           Df Sum Sq Mean Sq  F value    Pr(>F)
## C1         1 3825.4  3825.4 122.2469 2.516e-11 ***
## C2         1  430.5   430.5  13.7569 0.0009938 ***
## C3         1   29.5    29.5   0.9432 0.3404115
## Residuals 26  813.6    31.3
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

In the ANOVA, the inclusion of mass per unit area (C3) did not yield a significant enhancement to the model, and the P-value is still higher than α.

```r
library(car)
```

```
## Loading required package: carData
```

```r
vif(dt3_model1)
```

```
##        C1        C2        C3
## 34.792272  5.593738 38.745959
```

While we check the variance inflation factor (VIF), we can see that the value of the VIF is extremely high in all of these three independent variables. It shows that my assumption about the collinearity is correct. Therefore, this linear model is not the best model to explain the relationship between the independent variable and dependent variable.

```r
# Fit the new multiple linear regression model
dt3_model2 <- lm(formula = C4 ~ C1 + C2, data = dt3)

summary(dt3_model2)
```

```
##
## Call:
## lm(formula = C4 ~ C1 + C2, data = dt3)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
```

```
## -6.4566 -4.0214 -0.9593  3.6748 15.2906
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 111.0858    15.1053   7.354 6.54e-08 ***
## C1           -1.3763     0.1233 -11.160 1.28e-11 ***
## C2           -0.8932     0.2406  -3.713 0.000941 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.588 on 27 degrees of freedom
## Multiple R-squared:  0.8347, Adjusted R-squared:  0.8224
## F-statistic: 68.15 on 2 and 27 DF,  p-value: 2.808e-11
```

In this section, I want to check whether there is evidence of a linear relationship between the dependent variable and the independent variable.

At $\alpha = 0.05$:

$H_0 : \beta_1 = 0$ $H_a : \beta_1 \neq 0$

From the summary table, the P-value is 0.000, which is extremely small, and small than $\alpha = 0.05$. Thus, H_0 is rejected. There is sufficient evidence of a linear relationship between weft densities and air permeability values.

At $\alpha = 0.05$:

$H_0 : \beta_2 = 0$ $H_a : \beta_2 \neq 0$

From the summary table, the P-value is 0.000, which is extremely small, and small than $\alpha = 0.05$. H_0 is rejected. There is sufficient evidence of a linear relationship between warp densities and air permeability values.

From the model summary table, the sample coefficient of determination (R-Square) is 83.47%. This means that 84.04% of the variation in air permeability values is explained by the variation in the weft densities and warp densities. Therefore, according to this R-Squared value, this linear model seems useful. Since this P-value is also very close to the original linear model, we can assume that C3 does not really make a impact in this linear regression.

```r
# ANOVA
anova(dt3_model2)
```

```
## Analysis of Variance Table
##
## Response: C4
##           Df Sum Sq Mean Sq F value    Pr(>F)
## C1         1 3825.4  3825.4 122.505 1.541e-11 ***
## C2         1  430.5   430.5  13.786  0.000941 ***
## Residuals 27  843.1    31.2
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

In the ANOVA, weft densities and warp densities yield a significant enhancement to the model, and the P-value is close to 0.000 and smaller than α.

```r
vif(dt3_model2)
```

```
##       C1       C2
## 1.000658 1.000658
```

When we check the variance inflation factor (VIF), we can see that the value of the VIF is very close to 1. It shows that there is no collinearity in this model. Therefore, this linear model seems to be the best model to explain the relationship between the independent variable and dependent variable.

Then, the paper conducted a few nonlinear approaches for data analysis using artificial neural networks.

4. Extension/Enhancement

The authors of this study conducted a secondary experiment, where mass per unit area data (C3) was excluded from the model for analysis. Another extension could be to conduct similar tertiary experiments to exclude weft density (C1) and warp density (C2), in turn, and observe how the model changes.

First, I excluded the weft density (C1) and conducted an analysis.

```r
# Fit the multiple linear regression model
dt3_model3 <- lm(formula = C4 ~ C2 + C3, data = dt3)

summary(dt3_model3)
```

```
##
## Call:
## lm(formula = C4 ~ C2 + C3, data = dt3)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.4047 -4.5118 -0.7317  2.7577 14.5208
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 98.61319   14.80193   6.662 3.77e-07 ***
## C2           0.08348    0.25351   0.329    0.744
## C3          -0.69123    0.06184 -11.178 1.24e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.581 on 27 degrees of freedom
## Multiple R-squared:  0.8351, Adjusted R-squared:  0.8229
## F-statistic: 68.36 on 2 and 27 DF,  p-value: 2.709e-11
```

In this section, I want to check whether there is evidence of a linear relationship between the dependent variable and the independent variable.

At $\alpha = 0.05$:

$H_0 : \beta_2 = 0$ $H_a : \beta_2 \neq 0$

From the summary table, the P-value is 0.744, which is larger, and larger than $\alpha = 0.05$. H_0 is not rejected. There is not sufficient evidence of a linear relationship between warp densities and air permeability values.

At $\alpha = 0.05$:

$H_0 : \beta_3 = 0$ $H_a : \beta_3 \neq 0$

From the summary table, the P-value is 0.000, which is extremely small and smaller than $\alpha = 0.05$. H_0 is rejected. There is sufficient evidence of a linear relationship between mass per unit area and air permeability values.

From the model summary table, the sample coefficient of determination (R-Square) is 83.51%. This means that 83.51% of the variation in air permeability values is explained by the variation in the warp densities and mass per unit area. Therefore, according to this R-Squared value, this linear model seems useful.

Since the P-value is very high compared to α, and the R-Squared value is 83.51%, we can assume that there might be collinearity in this linear model.

```r
# ANOVA
anova(dt3_model3)
```

```
## Analysis of Variance Table
##
## Response: C4
##           Df Sum Sq Mean Sq F value    Pr(>F)
## C2         1  366.9   366.9  11.783  0.001942 **
## C3         1 3891.2  3891.2 124.944 1.237e-11 ***
## Residuals 27  840.9    31.1
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

In the ANOVA, warp densities and mass per unit area yield a significant enhancement to the model, and the P-value is close to 0.000 and smaller than α.

```r
vif(dt3_model3)
```

```
##       C2       C3
## 1.114369 1.114369
```

When we check the variance inflation factor (VIF), we can see that the value of the VIF is very close to 1. This shows that there is no collinearity in this model. Therefore, this linear model seems a good model to explain the relationship between the independent variable and dependent variable. However, when compared to the previous model, the previous model is still better than this model because the H_0 has been rejected in all of the variables.

Then, I excluded the warp density (C2) and conducted an analysis.

```r
# Fit the multiple linear regression model
dt3_model4 <- lm(formula = C4 ~ C1 + C3, data = dt3)
```

```r
summary(dt3_model4)
```

```
##
## Call:
## lm(formula = C4 ~ C1 + C3, data = dt3)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.4752 -4.6972 -0.5329  2.4213 14.6578
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  96.9082    11.2132   8.642 2.95e-09 ***
## C1           -0.2309     0.3218  -0.718 0.479137
## C3           -0.5831     0.1531  -3.809 0.000732 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.539 on 27 degrees of freedom
## Multiple R-squared:  0.8375, Adjusted R-squared:  0.8255
## F-statistic: 69.59 on 2 and 27 DF,  p-value: 2.216e-11
```

In this section, I want to check whether there is evidence of a linear relationship between the dependent variable and the independent variable.

At $\alpha = 0.05$:

$H_0 : \beta_1 = 0$ $H_a : \beta_1 \neq 0$

From the summary table, the P-value is 0.479, which is larger, and larger than $\alpha = 0.05$. H_0 is not rejected. There is not sufficient evidence of a linear relationship between represent weft densities and air permeability values.

At $\alpha = 0.05$:

$H_0 : \beta_3 = 0$ $H_a : \beta_3 \neq 0$

From the summary table, the P-value is 0.000, which is extremely small, and smaller than $\alpha = 0.05$. H_0 is rejected. There is sufficient evidence of a linear relationship between mass per unit area and air permeability values.

From the model summary table, the sample coefficient of determination (R-Square) is 83.75%. This means that 83.75% of the variation in air permeability values is explained by the variation in the weft densities and mass per unit area. Therefore, according to this R-Squared value, this linear model seems useful.

Since the P-value is very high compared to α, and the R-Squared value is 83.75%, we can assume that there might be collinearity in this linear model.

```r
# ANOVA
anova(dt3_model4)
```

```
## Analysis of Variance Table
##
## Response: C4
##           Df Sum Sq Mean Sq F value    Pr(>F)
## C1         1 3825.4  3825.4 124.674 1.267e-11 ***
## C3         1  445.2   445.2  14.508  0.000732 ***
## Residuals 27  828.4    30.7
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

In the ANOVA, weft densities and mass per unit area yield a significant enhancement to the model, and the P-value is close to 0.000 and smaller than α.

```r
vif(dt3_model4)
```

```
##       C1       C3
## 6.931221 6.931221
```

When we check the variance inflation factor (VIF), we can see that the value of the VIF is extremely high in both of the two independent variables. This shows that my assumption about the collinearity makes sense. Therefore, this linear model is not the best model to explain the relationship between the independent variable and dependent variable.

5. Discussion

In my extension and enhancement section, I conducted two additional analyses by removing two independent variables in turn. I then assessed whether these adjustments produced a better linear model compared to the author's. However, none of the models I tested were found to be better than the author's. Additionally, I extended each of the original models, as well as my own, and conducted an analysis on the null hypothesis. In the author's regression model, the coefficient for C3 was incisive, extremely small, and smaller than $\alpha = 0.05$, leading to rejection of the null hypothesis (H_0). This provides sufficient evidence of a linear relationship between mass per unit area and air permeability values. To confirm this evidence, I also conducted an additional analysis using ANOVA. Moreover, I further extended the analyses by addressing the common issue of collinearity in multiple linear regression, which can inflate the p-values of t-tests. After examining the Variance Inflation Factor (VIF), I found that all VIF values were around 1, indicating the absence of collinearity. However, both of the regression models that I constructed encountered their own issues. When I excluded the weft density (C1) and conducted an analysis, the p-value for C2 was extremely high. Similarly, when I excluded the warp density (C2) and conducted an analysis, the p-value for C2 remained notably high, suggesting the possibility of collinearity in these linear models. This indicates that these two models are not better than the modified one proposed by the author.

6. References

Çay, A., Vassiliadis, S., Rangoussi, M., & Tarakçıoğlu, I. (2007). Prediction of the air permeability of woven fabrics using neural networks. International Journal of Clothing Science and Technology, 19(1), 18–35. https://doi.org/10.1108/09556220710717026