

Introduction

The carbon fibre dataset consists of 30 observations of carbon fibre experiments in which two temperatures were used High and Low along with three different additives coded with A, B and C. Each additive has 5 observations at each temperature level. There is no quantifiable way to compare temperatures, they're coded as 'Low' and 'High'. The response variable for this analysis is the strength of the carbon fibre after each experiment was conducted. It is not known how the strength metric was measured.

Table 1: Carbon Fibre Data sample for additive A

additive	temperature	strength
a	Low	76
a	Low	90
a	Low	89
a	Low	105
a	Low	89
a	High	74
a	High	77
a	High	97
a	High	114
a	High	88

The question that this analysis is going to answer is whether there is a statistically significant effect between the temperature used or additive on carbon fibre strength and if there is an interaction between temperature and additive used. Table 2 presents the mean difference between high and low temperatures with high temperature showing a higher mean strength but also a lower upper and lower bound of 9-8 carbon fibre strength resulting in less variability, illustrated in Figure 1.

Table 2: Carbon Fibre mean strength, temperature used, upper & lower confidence intervals

Temperature	Mean	Lower CI 95%	Upper CI 95%
High	90	81	98
Low	63	49	76

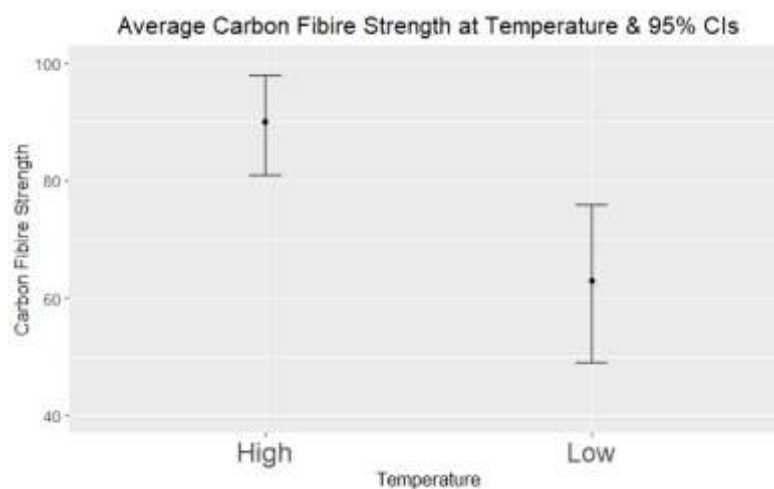


Figure 1: Carbon Fibre strength at temperature used

Carbon Fibre Data

The low temperature setting has a significantly lower mean strength but also a much wider range for upper and lower bounds of 14-13 carbon fibre strength resulting in much higher variability but also a much weaker material.

In Table 3 additive A has the highest mean strength with a confidence interval (CI) range of 9-10, additive B the median mean strength with a CI range of 11 in both directions and additive C with the weakest mean strength and a CI range of 17 also in both directions.

Table 3: Carbon Fibre mean strength, additive used, upper & lower confidence intervals

Additive	Mean	Lower CI 95%	Upper CI 95%
a	89	80	99
b	72	51	93
c	67	50	84

This leaves additive A as the strongest supplement to the material with the lowest range in variability, further illustrated in Figure 2. However, these results do not explain whether a potential interaction between temperature and additive used exist.

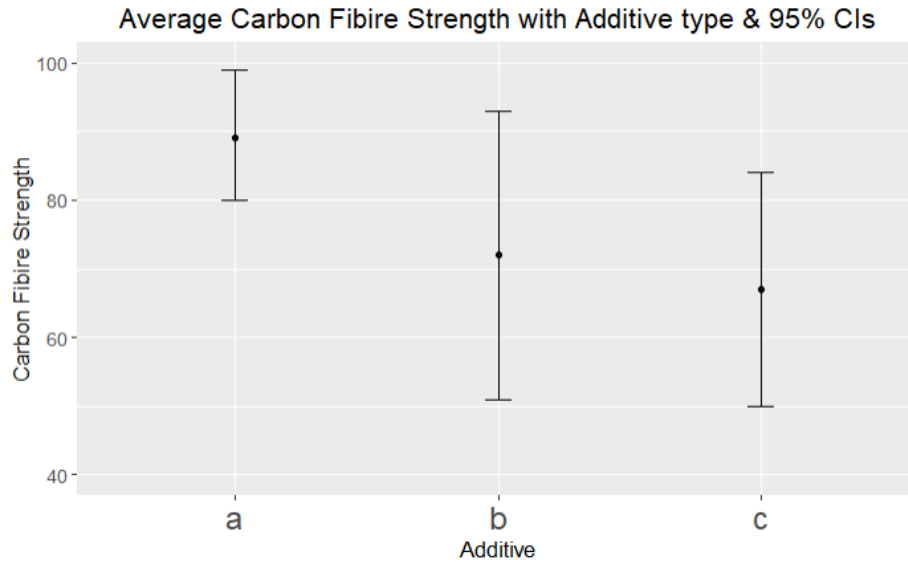


Figure 2: Carbon Fibre strength with additive used

Model 1

A full interaction linear model was fitted to these data with day of temperature as a categorical predictor and additive as a categorical predictor. To assess interaction between temperature and additive both were used as factors.

The model was:

$$y_i = \beta_0 + \beta_1(\delta_{i,low}) + \beta_2(\delta_{i,high}) + \beta_3(\gamma_{i,a}) + \beta_4(\gamma_{i,b}) + \beta_5(\gamma_{i,c}) + \beta_6(\delta_{i,high} * \gamma_{i,a}) + \beta_7(\delta_{i,high} * \gamma_{i,b}) + \beta_8(\delta_{i,high} * \gamma_{i,c}) + \dots + \beta_{11}(\delta_{i,low} * \gamma_{i,c}) + \epsilon_i$$

Where y_i is the live carbon fibre strength for the sample i , $\beta_1(\delta_{i,low}) = 1$ if the sample was synthesised at low temperature, 0 otherwise, repeat for all temperatures. Where $\beta_3(\gamma_{i,a}) = 1$ if

additive A was used, 0 otherwise. Then to check for an interaction between temperature and additive used $\beta_6(\delta_{i,high} * \gamma_{i,a}) = 1$ if both temperature = High and additive = A, 0 otherwise, this is repeated for each combination of temperature and additive.

This hypothesis was tested using the standard formulation for general linear hypothesis, results in Table 5. The ANOVA table for model 1 yields a p-value of 2.14e-05, indicating that there is a difference between means leading to rejecting the null hypothesis and accepting that either temperature, additive or both have an effect on carbon fibre strength.

Then using the drop1() function in R using the F-statistic the best model yielded a p-value of 0.0028 and f-statistic of 7.559 after iteratively dropping a single term. Both the ANOVA table and drop1 p-values indicate the model and interaction between temperature and additive used are statistically significant.

Upon examining the summary of the model which yielded an F-statistic of 10.39 on 5 and 24 degrees of freedom and a p-value of 2.142e-05. Low temperature and additive B yielded a p-value of 0.0007, low temperature and additive C had a p-value of 0.0419, the intercept for this model is 1.17e-12. There is a statistically significant interaction between temperature and additive used, meaning the full interaction model has been effective at determining this. Additive A is just as effective for both temperatures meanwhile being the best for low temperature and additive B showing the best yield with high temperature. Quantitative differences are recorded in Table 5 below.

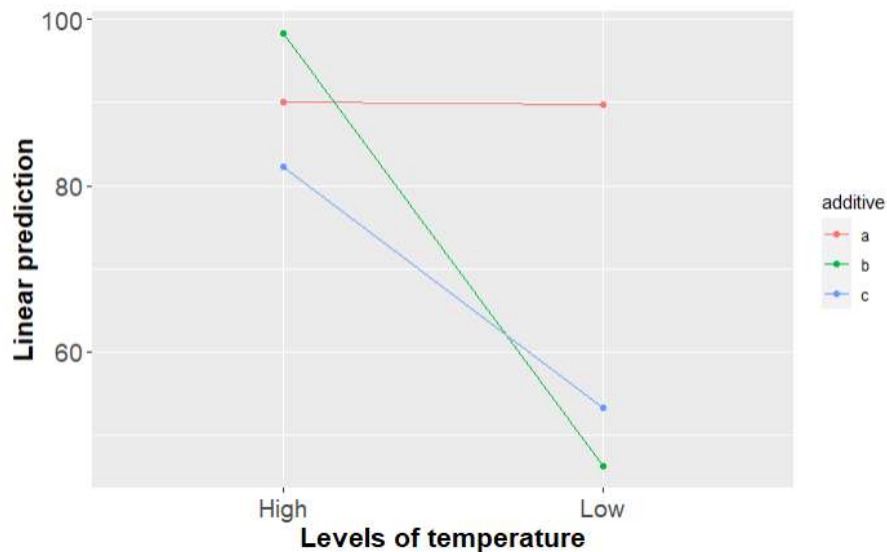


Figure 3: Interaction plot for model 1

Comparison between temperature & additive used on carbon fibre strength

The goal of this analysis was to see how effective temperature and additive were on the strength of carbon fibre and if there was an interaction between the two.

It has been found that both temperature and additive used have a strong statistically significant interaction which has been record in Table 5, for additive A both averages and ranges of carbon fibre strength seem to be almost identical, while making additive A the most potent for low temperature. Additive B has shown to be the best for high temperature.

Carbon Fibre Data

The interaction visualised in Figure 3 indicates that the interaction depends on temperature used especially low temperature which is further reinforced in Table 5 showing a significant difference between temperatures with a p-value of 4.6e-05. The difference in carbon fibre strength isn't as sharp when using high temperature compared to low temperature.

Table 4: model 1 emmeans – temperature & additive with effect on carbon fibre strength

Additive	Temperature	Least squared mean - carbon fibre strength	Lower Confidence Interval - carbon fibre strength	Upper Confidence Interval - carbon fibre strength
A	High	90	70.7	109.3
B	High	98.4	79.1	117.7
C	High	82.2	68.4	96
A	Low	89.8	70.5	109.1
B	Low	46.2	32.4	60
C	Low	53.2	33.9	72.5

Table 5: model 1 GLHT - difference between temperature, additive and carbon fibre strength

Comparing temperature and additives	difference in carbon fibre strength	Standard Error	P-Value
Low - High	-27.13	5.47	4.6e-05
b - a	-17.6	6.7	0.038
c - a	-22.2	6.7	0.008
c - b	-4.6	6.7	0.773

Code sample

```

library(pastecs) #For creating descriptive statistic summaries
library(ggplot2) #For creating histograms with more detail than plot
library(psych) # Some useful descriptive functions
library(semTools) #For skewness and kurtosis
library(FSA) # For percentage
library(car) # For Levene's test for homogeneity of variance
library(effectsize) # To calculate effect size for t-test
library(kableExtra) # Used to generate report ready tables
library(tidyverse) # data wrangling
library(gtsummary) # generate table for model results
library(multcomp) # needed for glht
library(emmeans)
require(ggiraph) # to use ggPredict
require(ggiraphExtra) # to use ggPredict
require(plyr)

data = read.csv("carbon_fibre.csv")
data = subset(data, select = -c(X))
anovatab <-
  function(mod){
    tab=as.matrix(anova(mod))
    rows=dim(tab)[1]
    moddf=sum(tab[,1])-tab[rows,1]
    ssmodel=sum(tab[,2])-tab[rows,2]
    msmodel=ssmodel/moddf
    f=msmodel/tab[rows,3]
    p=1-pf(f,moddf,tab[rows,1])
    tab2=tab[(rows-1):rows,]
    tab2[1,1:5]=c(moddf,ssmodel,msmodel,f,p)
    tab2=rbind(tab2,c(moddf+tab2[2,1],ssmodel+tab2[2,2],rep(NA,3)))
    rownames(tab2)=c('Model','Error','Total')
    colnames(tab2)[1]='df'
    return(print(tab2,na.print = "" , quote = FALSE,digits=3))
  }

tmp_df = data # store reference
data$additive = as.factor(data$additive)
data$temperature = as.factor(data$temperature)

scatter_plot = ggplot(data, aes(x=temperature,
                                y=strength,
                                show.legend = T, color=additive)) +
  geom_point(size=2) +
  ggtitle("Effect of temperature/additive on carbon fibre strength") +
  stat_smooth(method = "lm", formula = y ~ x, size = 1) +
  theme(plot.title = element_text(hjust = 0.5))

# reset reference
data = tmp_df

kbl(data) %>%
  kable_classic(full_width = F)

group_means=by(data$strength, data$temperature, t.test)

group_means=matrix(c(unlist(group_means[['High']][5:4]),
                     unlist(group_means[['Low']][5:4])),
                   nrow=2, ncol=3, byrow=T)

group_means = data.frame(cbind(group_means, c("High", "Low")))
# convert to int
group_means$X1 = round(as.integer(group_means$X1), digit=2)
group_means$X2 = round(as.integer(group_means$X2), digit=2)
group_means$X3 = round(as.integer(group_means$X3), digit=2)

colnames(group_means)=c('Mean','Lower CI 95%','Upper CI 95%','Temperature')

```

Carbon Fibre Data

```
ggplot(group_means, aes(x=Temperature, y=group_means[,1])) +
  geom_errorbar(aes(ymin=`Lower CI 95%`, ymax=`Upper CI 95%`), width=.1) +
  geom_line() +
  geom_point() +
  expand_limits(y=c(5, 20)) +
  ylab("Carbon Fibire Strength") +
  xlab('Temperature') +
  labs(title="Average Carbon Fibire Strength at Temperature & 95% CIs") +
  theme(text = element_text(size=13), axis.text.x=element_text(size=18), plot.title =
element_text(hjust = 0.5)) +
  coord_cartesian(ylim = c(40, 100))

group_means = group_means %>% relocate('Temperature', .before = `Mean`)

kbl(group_means) %>%
  kable_classic(full_width = F)

group_means=by(data$strength, data$additive, t.test)

group_means=matrix(c(unlist(group_means[['a']][5:4]),
                        unlist(group_means[['b']][5:4]),
                        unlist(group_means[['c']][5:4])),
                    nrow=3, ncol=3, byrow=T)

group_means = data.frame(cbind(group_means, c("a", "b", 'c'))))

# convert to int

group_means$X1 = round(as.integer(group_means$X1), digit=2)
group_means$X2 = round(as.integer(group_means$X2), digit=2)
group_means$X3 = round(as.integer(group_means$X3), digit=2)

colnames(group_means)=c('Mean','Lower CI 95%','Upper CI 95%','Additive')

group_means

ggplot(group_means, aes(x=Additive, y=group_means[,1])) +
  geom_errorbar(aes(ymin=`Lower CI 95%`, ymax=`Upper CI 95%`), width=.1) +
  geom_line() +
  geom_point() +
  expand_limits(y=c(5, 20)) +
  ylab("Carbon Fibire Strength") +
  xlab('Additive') +
  labs(title="Average Carbon Fibire Strength with Additive type & 95% CIs") +
  theme(text = element_text(size=13), axis.text.x=element_text(size=18), plot.title =
element_text(hjust = 0.5)) +
  coord_cartesian(ylim = c(40, 100))

group_means = group_means %>% relocate('Additive', .before = `Mean`)

kbl(group_means) %>%
  kable_classic(full_width = F)

data_v2 = subset(data, temperature=='High')

m1 = lm(strength ~ factor(temperature) + factor(additive) +
factor(additive):factor(temperature), data = data)

anovatab(m1)
plot(m1)
# ggPredict(m1, interactive=TRUE)
drop1(m1,test='F')

# assess model
summary(m1)
vcov(m1)
confint(m1)

emmeans(m1, ~additive*temperature)
```

Carbon Fibre Data

```
emmeans(m1, pairwise~additive)
emmeans(m1, pairwise~temperature)

lsmeans(m1, ~additive:temperature, adjust="fdr")

emmip(m1, additive ~ temperature) + theme(axis.text=element_text(size=14),
      axis.title=element_text(size=16, face="bold"))

glht_temperature = glht(m1, mcp("factor(temperature)" = "Tukey", interaction_average=TRUE))
summary(glht_temperature)

glht_additive = glht(m1, mcp("factor(additive)" = "Tukey", interaction_average=TRUE))
summary(glht_additive)

m1 %>% tbl_regression()
```