Phenols for manuscript

Zhang Zhenglin

Contents

Necessary libraries

Data Organisation

Read from excel

```
phenols_master <- read_excel("ForR_Phenols (24 July - OC content added).xlsx", sheet = 1)
phenols_master <- mutate_if(phenols_master, is.character, as.factor)
phenols_master$Year <- as.factor(phenols_master$Year)
phenols_master$Total_phenols <- phenols_master$TotalP+phenols_master$TotalV+phenols_master$TotalC+phenol

str(phenols_master)
```

```
## tibble [42 x 15] (S3: tbl_df/tbl/data.frame)
##  $ All_Logs_number: num [1:42] 7 8 9 10 11 12 13 14 15 16 ...
##  $ Field          : Factor w/ 2 levels "CR","RF": 2 2 2 2 1 1 1 1 2 2 ...
##  $ Plot           : Factor w/ 42 levels "104","106","107",..: 38 26 42 30 33 40 27 25 1 5 ...
##  $ Study          : Factor w/ 2 levels "Grower","RES": 1 1 1 1 1 1 1 1 2 2 ...
##  $ Site           : Factor w/ 15 levels "Gallagher","Jenkins 2021",..: 13 1 14 2 13 14 2 1 5 6 ...
##  $ Nrate          : Factor w/ 3 levels "0","225","Grower": 3 3 3 3 3 3 3 3 1 1 ...
##  $ TotalP         : num [1:42] 0.162 0.14 0.177 0.202 0.182 ...
##  $ TotalV         : num [1:42] 0.387 0.313 0.408 0.495 0.443 ...
##  $ TotalC         : num [1:42] 0.302 0.161 0.328 0.353 0.648 ...
##  $ TotalS         : num [1:42] 0.435 0.402 0.563 0.666 0.609 ...
##  $ Lamda8_Phenols : num [1:42] 1.124 0.877 1.299 1.514 1.7 ...
##  $ Year           : Factor w/ 2 levels "2021","2022": 1 1 1 1 1 1 1 1 1 1 ...
##  $ Lat            : num [1:42] 39 38.9 38.8 39.5 39 ...
##  $ Long           : num [1:42] -122 -122 -122 -122 -122 ...
##  $ Total_phenols  : num [1:42] 1.29 1.02 1.48 1.72 1.88 ...
```

Sub dataset for "RES" and "Growers"

```r
RES <- phenols_master %>% filter (Study == "RES")
str(RES)
```

```
## tibble [24 x 15] (S3: tbl_df/tbl/data.frame)
##  $ All_Logs_number: num [1:24] 15 16 17 18 19 20 21 22 23 24 ...
##  $ Field          : Factor w/ 2 levels "CR","RF": 2 2 2 1 1 1 2 2 2 1 ...
##  $ Plot           : Factor w/ 42 levels "104","106","107",..: 1 5 10 4 8 12 2 6 9 3 ...
##  $ Study          : Factor w/ 2 levels "Grower","RES": 2 2 2 2 2 2 2 2 2 2 ...
##  $ Site           : Factor w/ 15 levels "Gallagher","Jenkins 2021",..: 5 6 7 5 6 7 5 6 7 5 ...
##  $ Nrate          : Factor w/ 3 levels "0","225","Grower": 1 1 1 1 1 1 2 2 2 2 ...
##  $ TotalP         : num [1:24] 0.213 0.181 0.218 0.244 0.227 ...
##  $ TotalV         : num [1:24] 0.525 0.444 0.531 0.65 0.618 ...
##  $ TotalC         : num [1:24] 0.386 0.325 0.611 0.829 0.657 ...
##  $ TotalS         : num [1:24] 0.676 0.57 0.696 0.822 0.745 ...
##  $ Lamda8_Phenols : num [1:24] 1.59 1.34 1.84 2.3 2.02 ...
##  $ Year           : Factor w/ 2 levels "2021","2022": 1 1 1 1 1 1 1 1 1 1 ...
##  $ Lat            : num [1:24] 39.5 39.5 39.5 39.5 39.5 ...
##  $ Long           : num [1:24] -122 -122 -122 -122 -122 ...
##  $ Total_phenols  : num [1:24] 1.8 1.52 2.06 2.55 2.25 ...
```

```r
Growers <- phenols_master %>% filter(Study == "Grower")
str(Growers)
```

```
## tibble [18 x 15] (S3: tbl_df/tbl/data.frame)
##  $ All_Logs_number: num [1:18] 7 8 9 10 11 12 13 14 89 90 ...
##  $ Field          : Factor w/ 2 levels "CR","RF": 2 2 2 2 1 1 1 1 1 1 ...
##  $ Plot           : Factor w/ 42 levels "104","106","107",..: 38 26 42 30 33 40 27 25 31 39 ...
##  $ Study          : Factor w/ 2 levels "Grower","RES": 1 1 1 1 1 1 1 1 1 1 ...
##  $ Site           : Factor w/ 15 levels "Gallagher","Jenkins 2021",..: 13 1 14 2 13 14 2 1 4 15 ...
##  $ Nrate          : Factor w/ 3 levels "0","225","Grower": 3 3 3 3 3 3 3 3 3 3 ...
##  $ TotalP         : num [1:18] 0.162 0.14 0.177 0.202 0.182 ...
##  $ TotalV         : num [1:18] 0.387 0.313 0.408 0.495 0.443 ...
```

```
##  $ TotalC        : num [1:18] 0.302 0.161 0.328 0.353 0.648 ...
##  $ TotalS        : num [1:18] 0.435 0.402 0.563 0.666 0.609 ...
##  $ Lamda8_Phenols : num [1:18] 1.124 0.877 1.299 1.514 1.7 ...
##  $ Year          : Factor w/ 2 levels "2021","2022": 1 1 1 1 1 1 1 1 2 2 ...
##  $ Lat           : num [1:18] 39 38.9 38.8 39.5 39 ...
##  $ Long          : num [1:18] -122 -122 -122 -122 -122 ...
##  $ Total_phenols : num [1:18] 1.29 1.02 1.48 1.72 1.88 ...
```

Initial visualisation for all phenols

```
ggplot(phenols_master, aes(y=Total_phenols, x=Site, color=Field))+
  geom_boxplot()+
  theme(axis.text.x = element_text(angle = 45, hjust = 1, size= 9))
```
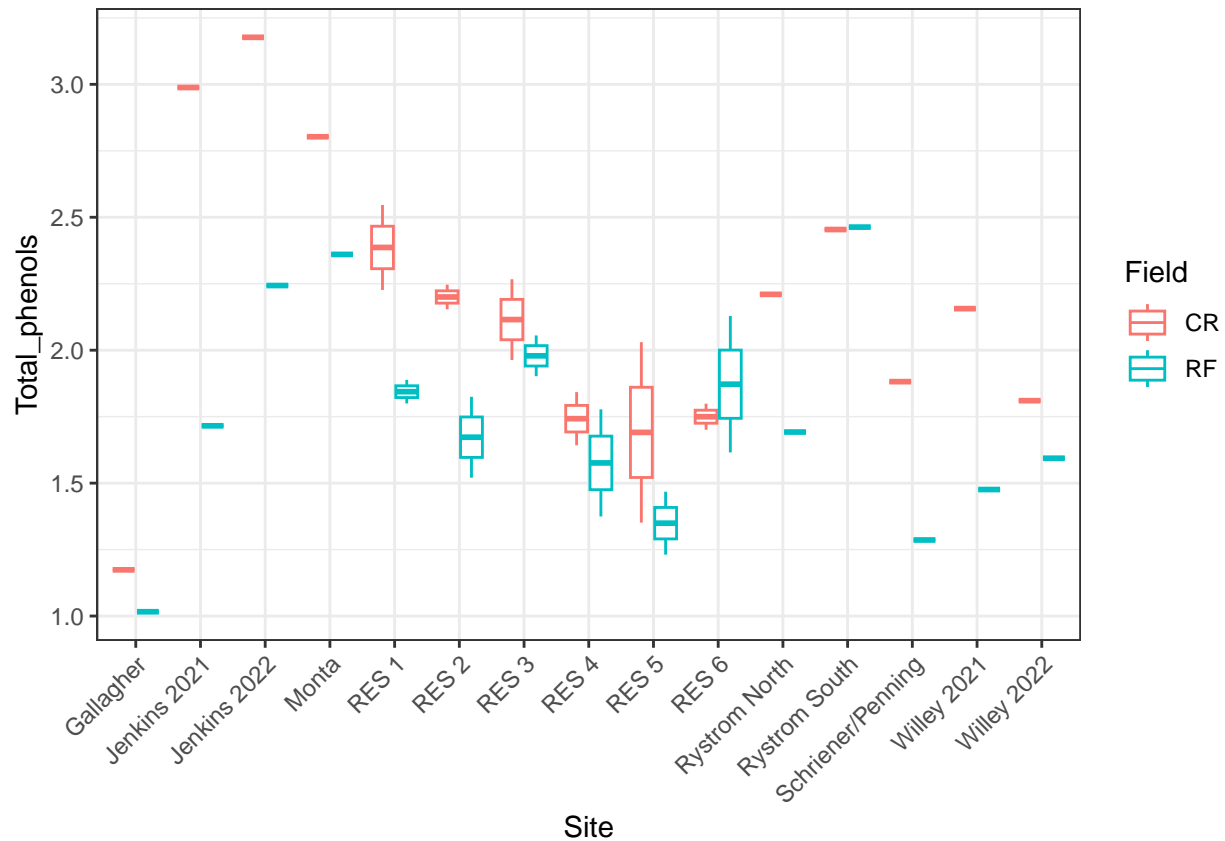

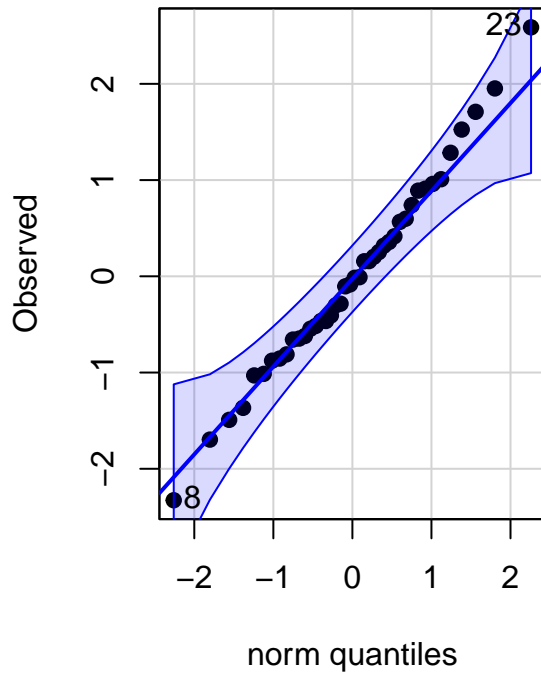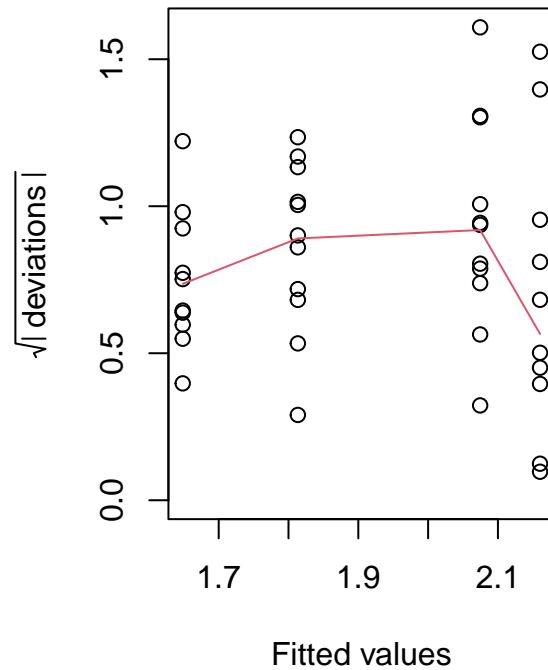## Inital model to get a sensing

Not much year effect

```
phenols_sensing_model <- lm(Total_phenols ~ Field*Year, data=phenols_master)
pls205_diagnostics(phenols_sensing_model)
```

Plot (EU) Normal Q–Q

Scale–Location

```r
anova(phenols_sensing_model)
```

```
## Analysis of Variance Table
##
## Response: Total_phenols
##            Df Sum Sq Mean Sq F value   Pr(>F)
## Field       1 1.5206 1.52058  7.6136 0.008863 **
## Year        1 0.0164 0.01637  0.0820 0.776178
## Field:Year  1 0.1646 0.16462  0.8243 0.369661
## Residuals  38 7.5893 0.19972
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

RES only statistical tests

```r
Total_P_RES <- lmer(TotalP ~ Field+(1|Site), data=RES)
Total_P_means_RES <- emmeans(Total_P_RES, spec ='Field')
Total_P_effects_RES <- contrast(Total_P_means_RES, method = 'pairwise', adjust = "tukey")
summary(Total_P_effects_RES)
```

```
## contrast estimate      SE df t.ratio p.value
## CR - RF   0.00533 0.00782 17   0.682  0.5047
```

```
##
## Degrees-of-freedom method: kenward-roger

Total_V_RES <- lmer(TotalV ~ Field+(1|Site), data=RES)
Total_V_means_RES <- emmeans(Total_V_RES, spec ='Field')
Total_V_effects_RES <- contrast(Total_V_means_RES, method = 'pairwise', adjust = "tukey")
summary(Total_V_effects_RES)


##  contrast estimate     SE df t.ratio p.value
##  CR - RF    0.0693 0.0226 17   3.071  0.0069
##
## Degrees-of-freedom method: kenward-roger

Total_C_RES <- lmer(TotalC ~ Field+(1|Site), data=RES)
Total_C_means_RES <- emmeans(Total_C_RES, spec ='Field')
Total_C_effects_RES <- contrast(Total_C_means_RES, method = 'pairwise', adjust = "tukey")
summary(Total_C_effects_RES)


##  contrast estimate     SE df t.ratio p.value
##  CR - RF     0.131 0.0427 17   3.075  0.0069
##
## Degrees-of-freedom method: kenward-roger

Total_S_RES <- lmer(TotalS ~ Field+(1|Site), data=RES)
Total_S_means_RES <- emmeans(Total_S_RES, spec ='Field')
Total_S_effects_RES <- contrast(Total_S_means_RES, method = 'pairwise', adjust = "tukey")
summary(Total_S_effects_RES)


##  contrast estimate     SE df t.ratio p.value
##  CR - RF    0.0595 0.0306 17   1.943  0.0688
##
## Degrees-of-freedom method: kenward-roger

Total_phenols_RES <- lmer(Total_phenols ~ Field+(1|Site), data=RES)
Total_phenols_means_RES <- emmeans(Total_phenols_RES, spec ='Field')
Total_phenols_effects_RES <- contrast(Total_phenols_means_RES, method = 'pairwise', adjust = "tukey")
summary(Total_phenols_effects_RES)


##  contrast estimate     SE df t.ratio p.value
##  CR - RF     0.265 0.0984 17   2.696  0.0153
##
## Degrees-of-freedom method: kenward-roger
```

Growers' only statistical tests

```
Total_P_Growers <- lmer(TotalP ~ Field+(1|Site), data=Growers)
Total_P_means_Growers <- emmeans(Total_P_Growers, spec ='Field')
Total_P_effects_Growers <- contrast(Total_P_means_Growers, method = 'pairwise', adjust = "tukey")
summary(Total_P_effects_Growers)
```

```
##  contrast estimate      SE df t.ratio p.value
##  CR - RF     0.0328 0.00876  8   3.737  0.0057
##
## Degrees-of-freedom method: kenward-roger
```

```
Total_V_Growers <- lmer(TotalV ~ Field+(1|Site), data=Growers)
Total_V_means_Growers <- emmeans(Total_V_Growers, spec ='Field')
Total_V_effects_Growers <- contrast(Total_V_means_Growers, method = 'pairwise', adjust = "tukey")
summary(Total_V_effects_Growers)
```

```
##  contrast estimate      SE df t.ratio p.value
##  CR - RF     0.115 0.0303  8   3.816  0.0051
##
## Degrees-of-freedom method: kenward-roger
```

```
Total_C_Growers <- lmer(TotalC ~ Field+(1|Site), data=Growers)
Total_C_means_Growers <- emmeans(Total_C_Growers, spec ='Field')
Total_C_effects_Growers <- contrast(Total_C_means_Growers, method = 'pairwise', adjust = "tukey")
summary(Total_C_effects_Growers)
```

```
##  contrast estimate      SE df t.ratio p.value
##  CR - RF     0.222 0.0709  8   3.129  0.0140
##
## Degrees-of-freedom method: kenward-roger
```

```
Total_S_Growers <- lmer(TotalS ~ Field+(1|Site), data=Growers)
Total_S_means_Growers <- emmeans(Total_S_Growers, spec ='Field')
Total_S_effects_Growers <- contrast(Total_S_means_Growers, method = 'pairwise', adjust = "tukey")
summary(Total_S_effects_Growers)
```

```
##  contrast estimate      SE df t.ratio p.value
##  CR - RF     0.164 0.0406  8   4.047  0.0037
##
## Degrees-of-freedom method: kenward-roger
```

```
Total_phenols_Growers <- lmer(Total_phenols ~ Field+(1|Site), data=Growers)
Total_phenols_means_Growers <- emmeans(Total_phenols_Growers, spec ='Field')
Total_phenols_effects_Growers <- contrast(Total_phenols_means_Growers, method = 'pairwise', adjust = "t
summary(Total_phenols_effects_Growers)
```

```
##  contrast estimate      SE df t.ratio p.value
##  CR - RF     0.534 0.133  8   4.012  0.0039
##
## Degrees-of-freedom method: kenward-roger
```

Create graphing data frame

```
manuscript_graphing <-
  phenols_master %>% group_by(Field, Study) %>%
  mutate(Total_phenols_sd = sd(Total_phenols),
         TotalP_sd = sd(TotalP),
         TotalV_sd = sd(TotalV),
         TotalC_sd = sd(TotalC),
         TotalS_sd = sd(TotalS)) %>%
  summarise(Total_phenols = mean(Total_phenols),
            Total_phenols_sd = mean(Total_phenols_sd),
            TotalP = mean(TotalP),
            TotalP_sd = mean(TotalP_sd),
            TotalV = mean(TotalV),
            TotalV_sd = mean(TotalV_sd),
            TotalC = mean(TotalC),
            TotalC_sd = mean(TotalC_sd),
            TotalS = mean(TotalS),
            TotalS_sd = mean(TotalS_sd))
```

```
## 'summarise()' has grouped output by 'Field'. You can override using the
## '.groups' argument.
```

```
write.csv(manuscript_graphing, "manuscript_graphing_11_Jan_2024.csv", row.names = TRUE)
```

With the CSV, I manually reorganised the dataframe so that I can plot phenols by type, with each bar graph showing levels for CR and RF respectively. I was just lazy to manually wrangle this in R. It was too much effort for such a simple dataframe lol.

```
manuscript_graphing_reorg <- read_excel("Phenols_graphing_manuscript_11Jan2024.xlsx", sheet = 1)

manuscript_graphing_reorg$Type <- str_wrap(manuscript_graphing_reorg$Type, width = 10)

RES_graphing <- manuscript_graphing_reorg %>% filter(Study == "RES")
str(RES_graphing)
```

```
## tibble [10 x 5] (S3: tbl_df/tbl/data.frame)
##  $ Type         : chr [1:10] "Total_phenols" "Total_phenols" "TotalP" "TotalP" ...
##  $ Field        : chr [1:10] "CR" "RF" "CR" "RF" ...
##  $ Study        : chr [1:10] "RES" "RES" "RES" "RES" ...
##  $ Phenols_level: num [1:10] 1.981 1.716 0.194 0.189 0.528 ...
##  $ SD           : num [1:10] 0.3312 0.2754 0.0346 0.0285 0.0891 ...
```

```
Growers_Graphing <- manuscript_graphing_reorg %>% filter(Study == "Grower")
str(Growers_Graphing)
```

```
## tibble [10 x 5] (S3: tbl_df/tbl/data.frame)
##  $ Type         : chr [1:10] "Total_phenols" "Total_phenols" "TotalP" "TotalP" ...
##  $ Field        : chr [1:10] "CR" "RF" "CR" "RF" ...
##  $ Study        : chr [1:10] "Grower" "Grower" "Grower" "Grower" ...
##  $ Phenols_level: num [1:10] 2.295 1.761 0.214 0.181 0.583 ...
##  $ SD           : num [1:10] 0.6355 0.4976 0.0472 0.0303 0.1536 ...
```

```
poster<- Growers_Graphing
```

Graphing RES

```
RES_phenols_graph <-
ggplot(RES_graphing, aes(x = Type, y = Phenols_level, fill = Field)) +
  geom_bar(stat = "identity", position = "dodge", width = 0.7) +
  labs(x = "Type", y = "Phenols_level", fill = "Field") +
  scale_fill_manual(values = c("#0072B2","#FFCC66"),labels = c("Continuous rice (CR)", "Fallow rice (FR)
  #geom_errorbar(aes(ymin=Phenols_level-SD, ymax=Phenols_level+SD), width=0.2,position=position_dodge(0
  scale_y_continuous(name=expression("Phenols level (mg 100g"^{-1}*" OC)"), limits = c(0,3), expand = c
  scale_x_discrete(name="", labels = c("Total phenols", "Cinnamic phenols", "P-hydroxybenzyl phenols",
  theme_classic() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1, size= 12))+
  theme(axis.text.y = element_text(size= 12),axis.title = element_text(size=12))+
  theme(legend.text = element_text(size = 12),legend.title = element_text(size = 13))+
    annotate(
  "text",
  x = c(1,2,5),  # X-axis positions for annotations
  y = c(2.1,0.65,0.65),  # Y-axis positions for annotations
  label = "*",
  size = 8,
  vjust = 0  # Adjust vertical position of asterisks
)+
  annotate(
  "text",
  x = c(3),  # X-axis positions for annotations
  y = c(2.8),  # Y-axis positions for annotations
  label = "RES",
  size = 5,
  vjust = 0
)

ggsave(RES_phenols_graph, filename = "RES_phenols_graph_22Jan2024.png", height = 15, width = 20, units
```

Graphing growers' fields

```
Grower_phenols_graph<-
ggplot(Growers_Graphing, aes(x = Type, y = Phenols_level, fill = Field)) +
  geom_bar(stat = "identity", position = "dodge", width = 0.7) +
  labs(x = "Type", y = "Phenols_level", fill = "Field") +
  scale_fill_manual(values = c("#0072B2","#FFCC66"),labels = c("Continuous rice (CR)", "Fallow rice (FR)
  #geom_errorbar(aes(ymin=Phenols_level-SD, ymax=Phenols_level+SD), width=0.2,position=position_dodge(0
  scale_y_continuous(name=expression("Phenols level (mg 100g"^{-1}*" OC)"), limits = c(0,3), expand = c
  scale_x_discrete(name="", labels = c("Total phenols", "Cinnamic phenols", "P-hydroxybenzyl phenols",
  theme_classic() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1, size= 12))+
  theme(axis.text.y = element_text(size= 12),axis.title = element_text(size=12))+
```

```r
  theme(legend.text = element_text(size = 12),legend.title = element_text(size = 13))+
    annotate(
  "text",
  x = c(1,2,3,4,5),  # X-axis positions for annotations
  y = c(2.4,0.8,0.3,1,0.7),  # Y-axis positions for annotations
  label = "*",
  size = 8,
  vjust = 0  # Adjust vertical position of asterisks
)+
  annotate(
  "text",
  x = c(3),  # X-axis positions for annotations
  y = c(2.8),  # Y-axis positions for annotations
  label = "Regional survey study",
  size = 5,
  vjust = 0
)


ggsave(Grower_phenols_graph, filename = "Growers_phenols_graph_22Jan2024.png", height = 15, width = 20,
```