

# Phenols for manuscript

Zhang Zhenglin

## Contents

Necessary libraries	1
RES	2
Regional Survey	5
Import graphing data	9
Graphing RES	9
Graphing growers' fields	10

## Necessary libraries

```
library(knitr)
library(rlang)
library(ggplot2)
theme_set(theme_bw())
library(emmeans)
library(multcomp)
library(PLS205)
library(lme4)
library(lmerTest)
library(multcompView)
library(car)
library(Rmisc)
library(dplyr) #https://r4ds.had.co.nz/ (Chapter 3, Chapter 5, look at filter and select)
# https://bookdown.org/ansellbr/WEHI\_tidyR\_course\_book/
library(stringr)
library(data.table)
library(GGally)
library(formatR)
library(readxl)
library(mgcv)
library(openxlsx)
```

## RES

```
RES <- read_excel("RES_25Jan2024.xlsx", sheet = 1)
RES <- mutate_if(RES, is.character, as.factor)
RES$Year <- as.factor(RES$Year)
```

```
str(RES)
```

```
## tibble [12 x 9] (S3: tbl_df/tbl/data.frame)
## $ Field      : Factor w/ 2 levels "CR","RF": 2 2 2 1 1 1 2 2 2 1 ...
## $ Study      : Factor w/ 1 level "RES": 1 1 1 1 1 1 1 1 1 1 ...
## $ Blk        : Factor w/ 6 levels "RES 1","RES 2",...: 1 2 3 1 2 3 4 5 6 4 ...
## $ Year       : Factor w/ 2 levels "2021","2022": 1 1 1 1 1 1 2 2 2 2 ...
## $ TotalP     : num [1:12] 0.212 0.2 0.21 0.233 0.22 ...
## $ TotalV     : num [1:12] 0.5 0.472 0.521 0.64 0.593 ...
## $ TotalC     : num [1:12] 0.464 0.413 0.579 0.739 0.659 ...
## $ TotalS     : num [1:12] 0.667 0.587 0.668 0.774 0.727 ...
## $ Total_phenols: num [1:12] 1.84 1.67 1.98 2.39 2.2 ...
```

```
Total_P_RES <- lmer(TotalP ~ Field+(1|Blk), 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")
anova(Total_P_RES)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##          Sum Sq   Mean Sq NumDF DenDF F value Pr(>F)
## Field 8.521e-05 8.521e-05      1      5  0.4993 0.5114
```

```
cld(Total_P_means_RES)
```

```
## Field emmean    SE    df lower.CL upper.CL .group
## RF      0.189 0.0121 6.06    0.159    0.218 1
## CR      0.194 0.0121 6.06    0.165    0.224 1
##
## Degrees-of-freedom method: kenward-roger
## Confidence level used: 0.95
## significance level used: alpha = 0.05
## NOTE: If two or more means share the same grouping symbol,
##       then we cannot show them to be different.
##       But we also did not show them to be the same.
```

```
summary(Total_P_effects_RES)
```

```
## contrast estimate      SE df t.ratio p.value
## CR - RF      0.00533 0.00754  5    0.707  0.5114
##
## Degrees-of-freedom method: kenward-roger
```

```
Total_V_RES <- lmer(TotalV ~ Field+(1|Blk), 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")
anova(Total_V_RES)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##      Sum Sq Mean Sq NumDF DenDF F value Pr(>F)
## Field 0.014414 0.014414      1      5  7.0509 0.04514 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
cld(Total_V_means_RES)
```

```
## Field emmean      SE    df lower.CL upper.CL .group
## RF      0.458 0.0311 7.04    0.385    0.532 1
## CR      0.528 0.0311 7.04    0.454    0.601 2
##
## Degrees-of-freedom method: kenward-roger
## Confidence level used: 0.95
## significance level used: alpha = 0.05
## NOTE: If two or more means share the same grouping symbol,
##       then we cannot show them to be different.
##       But we also did not show them to be the same.
```

```
summary(Total_V_effects_RES)
```

```
## contrast estimate      SE df t.ratio p.value
## CR - RF      0.0693 0.0261 5    2.655 0.0451
##
## Degrees-of-freedom method: kenward-roger
```

```
Total_C_RES <- lmer(TotalC ~ Field+(1|Blk), 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")
anova(Total_C_RES)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##      Sum Sq Mean Sq NumDF DenDF F value Pr(>F)
## Field 0.051635 0.051635      1      5  8.557 0.03281 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
cld(Total_C_means_RES)
```

```
## Field emmean      SE    df lower.CL upper.CL .group
## RF      0.459 0.0354 9.62    0.379    0.538 1
## CR      0.590 0.0354 9.62    0.511    0.669 2
##
## Degrees-of-freedom method: kenward-roger
## Confidence level used: 0.95
```

```
## significance level used: alpha = 0.05
## NOTE: If two or more means share the same grouping symbol,
##       then we cannot show them to be different.
##       But we also did not show them to be the same.
```

```
summary(Total_C_effects_RES)
```

```
## contrast estimate      SE df t.ratio p.value
## CR - RF      0.131 0.0448  5   2.925  0.0328
##
## Degrees-of-freedom method: kenward-roger
```

```
Total_S_RES <- lmer(TotalS ~ Field+(1|Blk), 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")
anova(Total_S_RES)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##           Sum Sq Mean Sq NumDF DenDF F value Pr(>F)
## Field 0.010606 0.010606      1      5  3.9083  0.105
```

```
cld(Total_S_means_RES)
```

```
## Field emmean      SE df lower.CL upper.CL .group
## RF      0.610 0.0315 7.72    0.537    0.683  1
## CR      0.669 0.0315 7.72    0.596    0.742  1
##
```

```
## Degrees-of-freedom method: kenward-roger
## Confidence level used: 0.95
## significance level used: alpha = 0.05
## NOTE: If two or more means share the same grouping symbol,
##       then we cannot show them to be different.
##       But we also did not show them to be the same.
```

```
summary(Total_S_effects_RES)
```

```
## contrast estimate      SE df t.ratio p.value
## CR - RF      0.0595 0.0301  5   1.977  0.1050
##
## Degrees-of-freedom method: kenward-roger
```

```
Total_phenols_RES <- lmer(Total_phenols ~ Field+(1|Blk), 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")
anova(Total_phenols_RES)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##           Sum Sq Mean Sq NumDF DenDF F value Pr(>F)
## Field 0.21114 0.21114      1      5  6.4376 0.05207 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
cld(Total_phenols_means_RES)
```

```
## Field emmean SE df lower.CL upper.CL .group
## RF 1.72 0.107 7.84 1.47 1.96 1
## CR 1.98 0.107 7.84 1.73 2.23 1
##
## Degrees-of-freedom method: kenward-roger
## Confidence level used: 0.95
## significance level used: alpha = 0.05
## NOTE: If two or more means share the same grouping symbol,
## then we cannot show them to be different.
## But we also did not show them to be the same.
```

```
summary(Total_phenols_effects_RES)
```

```
## contrast estimate SE df t.ratio p.value
## CR - RF 0.265 0.105 5 2.537 0.0521
##
## Degrees-of-freedom method: kenward-roger
```

## Regional Survey

```
Growers <- read_excel("Regional_survey_25Jan2024.xlsx", sheet = 1)
Growers <- mutate_if(Growers, is.character, as.factor)
Growers$Year <- as.factor(Growers$Year)
```

```
str(Growers)
```

```
## tibble [18 x 12] (S3: tbl_df/tbl/data.frame)
## $ Field      : Factor w/ 2 levels "CR","RF": 2 2 2 2 1 1 1 1 1 1 1 ...
## $ Plot       : Factor w/ 18 levels "Gallagher Continuous",...: 14 2 18 6 9 16 3 1 7 15 ...
## $ Study      : Factor w/ 1 level "Grower": 1 1 1 1 1 1 1 1 1 1 1 ...
## $ Site       : num [1:18] 3 1 4 2 3 4 2 1 7 5 ...
## $ 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 ...
## $ Total_phenols: num [1:18] 1.29 1.02 1.48 1.72 1.88 ...
## $ 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_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")
anova(Total_P_Growers)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
```

```
##          Sum Sq Mean Sq NumDF DenDF F value Pr(>F)
## Field 0.004827 0.004827      1      8 13.968 0.005727 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
cld(Total_P_means_Growers)
```

```
## Field emmean      SE    df lower.CL upper.CL .group
## RF      0.181 0.0132 9.94    0.151    0.210 1
## CR      0.214 0.0132 9.94    0.184    0.243 2
##
## Degrees-of-freedom method: kenward-roger
## Confidence level used: 0.95
## significance level used: alpha = 0.05
## NOTE: If two or more means share the same grouping symbol,
##       then we cannot show them to be different.
##       But we also did not show them to be the same.
```

```
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")
anova(Total_V_Growers)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##          Sum Sq Mean Sq NumDF DenDF F value Pr(>F)
## Field 0.059995 0.059995      1      8 14.559 0.005122 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
cld(Total_V_means_Growers)
```

```
## Field emmean      SE    df lower.CL upper.CL .group
## RF      0.468 0.0437 10.1    0.370    0.565 1
## CR      0.583 0.0437 10.1    0.486    0.680 2
##
## Degrees-of-freedom method: kenward-roger
## Confidence level used: 0.95
## significance level used: alpha = 0.05
## NOTE: If two or more means share the same grouping symbol,
##       then we cannot show them to be different.
##       But we also did not show them to be the same.
```

```
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")
anova(Total_C_Growers)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##      Sum Sq Mean Sq NumDF DenDF F value  Pr(>F)
## Field 0.22144 0.22144      1      8  9.7897 0.01404 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
cld(Total_C_means_Growers)
```

```
## Field emmean      SE   df lower.CL upper.CL .group
## RF      0.443 0.0665 13.5    0.300    0.587    1
## CR      0.665 0.0665 13.5    0.522    0.808    2
##
## Degrees-of-freedom method: kenward-roger
## Confidence level used: 0.95
## significance level used: alpha = 0.05
## NOTE: If two or more means share the same grouping symbol,
##       then we cannot show them to be different.
##       But we also did not show them to be the same.
```

```
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")
anova(Total_S_Growers)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##      Sum Sq Mean Sq NumDF DenDF F value  Pr(>F)
## Field 0.12129 0.12129      1      8 16.381 0.003697 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
cld(Total_S_means_Growers)
```

```
## Field emmean SE df lower.CL upper.CL .group
## RF 0.669 0.075 9.26 0.500 0.838 1
## CR 0.833 0.075 9.26 0.664 1.002 2
##
## Degrees-of-freedom method: kenward-roger
## Confidence level used: 0.95
## significance level used: alpha = 0.05
## NOTE: If two or more means share the same grouping symbol,
## then we cannot show them to be different.
## But we also did not show them to be the same.
```

```
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 = "tukey")
anova(Total_phenols_Growers)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
## Sum Sq Mean Sq NumDF DenDF F value Pr(>F)
## Field 1.2843 1.2843 1 8 16.096 0.003884 **
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
cld(Total_phenols_means_Growers)
```

```
## Field emmean SE df lower.CL upper.CL .group
## RF 1.76 0.19 10.2 1.34 2.18 1
## CR 2.29 0.19 10.2 1.87 2.72 2
##
## Degrees-of-freedom method: kenward-roger
## Confidence level used: 0.95
## significance level used: alpha = 0.05
## NOTE: If two or more means share the same grouping symbol,
## then we cannot show them to be different.
## But we also did not show them to be the same.
```

```
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
```



## Import graphing data

```
manuscript_graphing <- read_excel("Phenols_graphing_manuscript_25Jan2024.xlsx", sheet = 1)
```

```
RES_graphing <- manuscript_graphing %>% 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 %>% 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 ...
```

## Graphing RES

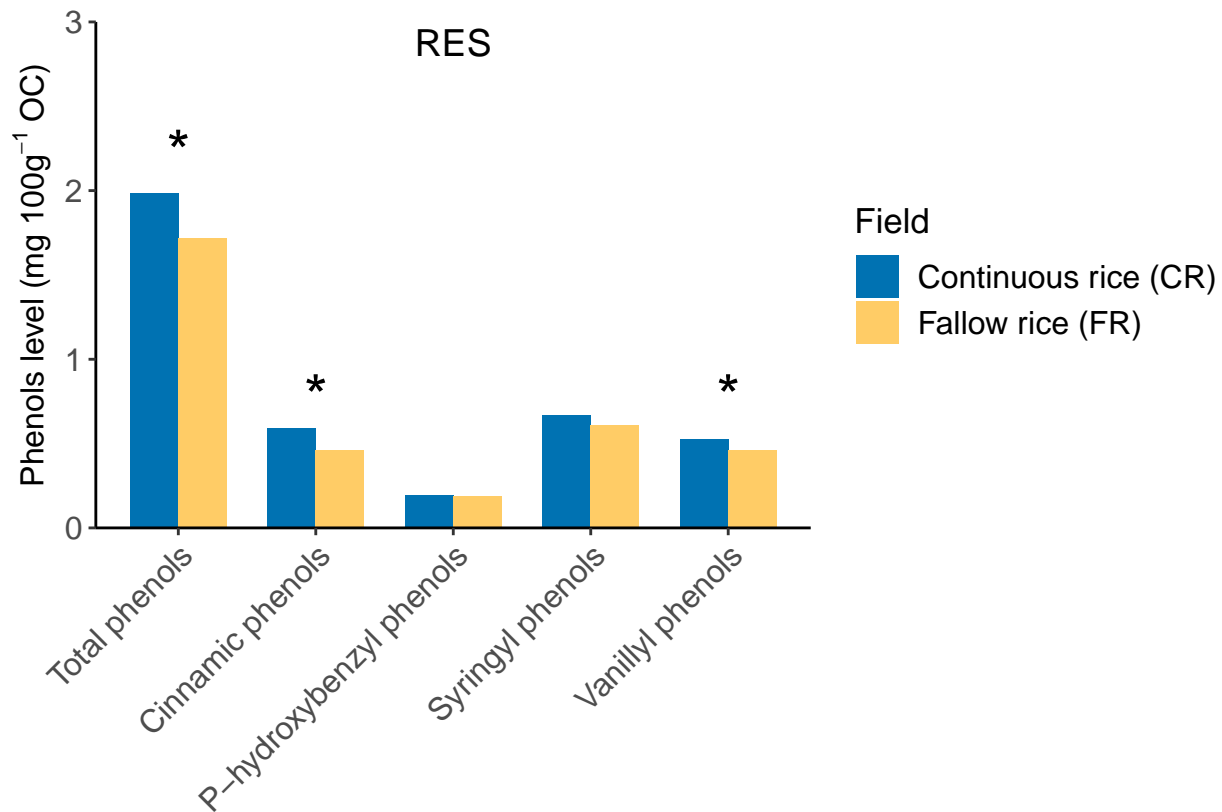
```
RES_phenols_graph <-
ggplot(RES_graphing, aes(x = Type, y = Phenols_level, fill = Field)) +
  geom_bar(stat = "identity", position = "dodge", width = 0.7) +
  labs(x = "Type", y = "Phenols_level", fill = "Field") +
  scale_fill_manual(values = c("#0072B2", "#FFCC66"), labels = c("Continuous rice (CR)", "Fallow rice (FR)")) +
  scale_y_continuous(name=expression("Phenols level (mg 100g"^-1)*" OC"), limits = c(0,3), expand = c(0,0.1)) +
  scale_x_discrete(name="", labels = c("Total phenols", "Cinnamic phenols", "P-hydroxybenzyl phenols", "P-coumaroyl 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 = "*"
  )
```

```

label = "RES",
size = 5,
vjust = 0
)

```

RES\_phenols\_graph



```

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

```

## 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)")) +
  scale_y_continuous(name=expression("Phenols level (mg 100g"^-1)*" OC)"), limits = c(0,3), expand = c(0,0)) +
  scale_x_discrete(name="", labels = c("Total phenols", "Cinnamic phenols", "P-hydroxybenzyl phenols", "Syringyl phenols", "Vanillyl 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))+

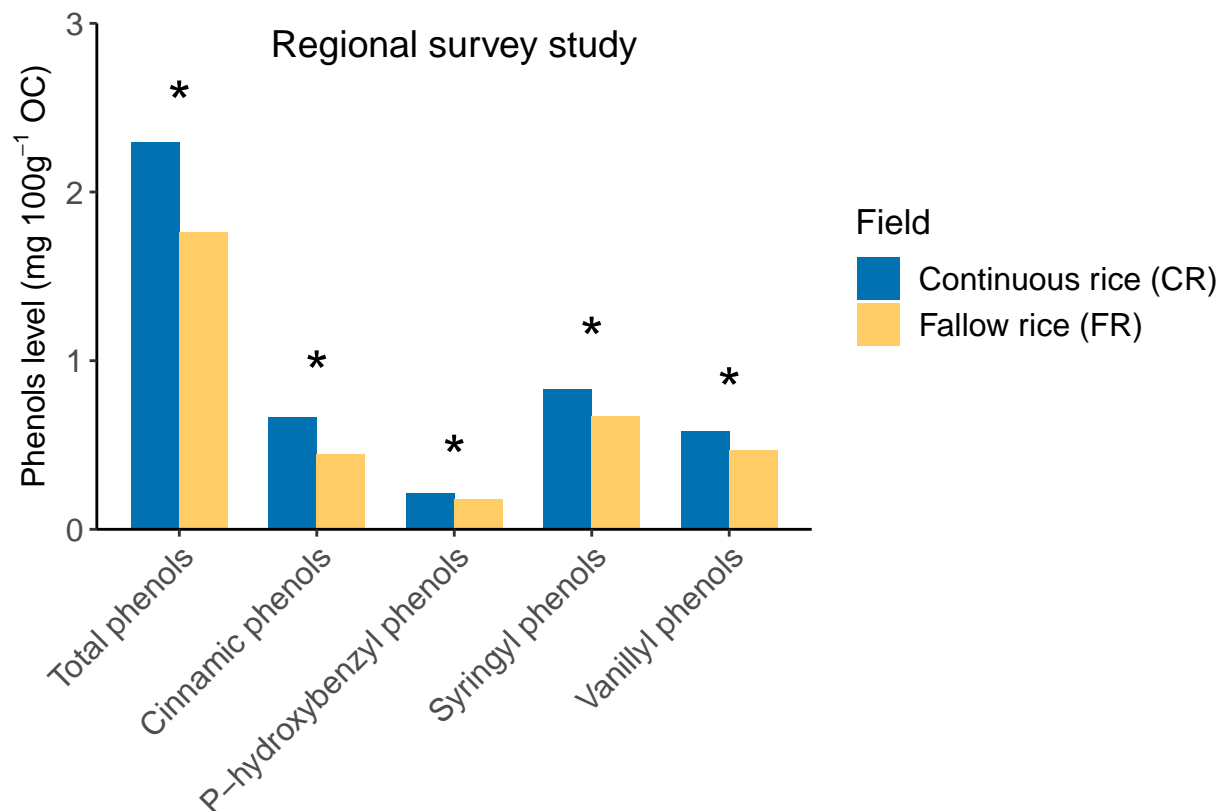
```

```

  annotate(
    "text",
    x = c(1,2,3,4,5), # 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
  )
)

```

Grower\_phenols\_graph



```

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

```