Correlation Analysis of Phenols and N availability

Zhang Zhenglin

Contents

Necessary libraries

```
library(knitr)
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)
```

```r
library(data.table)
library(GGally)
library(formatR)
library(readxl)
library(mgcv)
library(writexl)
```

Data Organisation

Read from excel

```r
preplant <- read_excel("Cor_Base.xlsx", sheet = 1)
str(preplant)
```

```
## tibble [12 x 13] (S3: tbl_df/tbl/data.frame)
##  $ Field            : chr [1:12] "RF" "RF" "RF" "CR" ...
##  $ Blk_Site         : chr [1:12] "RES 1" "RES 2" "RES 3" "RES 1" ...
##  $ Blk              : chr [1:12] "1" "2" "3" "1" ...
##  $ Year             : chr [1:12] "2021" "2021" "2021" "2021" ...
##  $ 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 ...
##  $ crop_fertiliser_N: num [1:12] 58.2 52.5 49.1 60.2 44.1 ...
##  $ crop_soil_N      : num [1:12] 120.7 107.2 109.3 98.8 90 ...
##  $ soil_tracer_N    : num [1:12] 32.8 37.6 38 30.1 22.6 ...
##  $ yield            : num [1:12] 15.3 15.1 14.2 13.3 12.7 ...
```

```r
topdress <- read_excel("Cor_Base.xlsx", sheet = 2)
str(topdress)
```

```
## tibble [12 x 13] (S3: tbl_df/tbl/data.frame)
##  $ Field            : chr [1:12] "RF" "RF" "RF" "CR" ...
##  $ Blk_Site         : chr [1:12] "RES 1" "RES 2" "RES 3" "RES 1" ...
##  $ Blk              : chr [1:12] "1" "2" "3" "1" ...
##  $ Year             : chr [1:12] "2021" "2021" "2021" "2021" ...
##  $ 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 ...
##  $ crop_fertiliser_N: num [1:12] 11.96 11.77 7.9 10.61 9.19 ...
##  $ crop_soil_N      : num [1:12] 102.1 96.6 94.2 80.8 93.4 ...
##  $ soil_tracer_N    : num [1:12] 6.02 6.3 2.89 4.79 4.76 ...
##  $ yield            : num [1:12] 11 10.74 10.57 8.96 10.18 ...
```

```
#cor_data_preplant <- preplant%>%select(-Field,-Blk_Site,-Blk,-Year)
#cor_data_topdress <- topdress%>%select(-Field,-Blk_Site,-Blk,-Year)

cor_data_preplant <- preplant%>%select(-Blk_Site,-Blk,-Year)
cor_data_topdress <- topdress%>%select(-Blk_Site,-Blk,-Year)
```

Correlation matrix (with raw and absolute values)

Delinated by field

```
preplant_cor_matrix <-ggpairs(cor_data_preplant, aes(col=Field, alpha=0.5))+
    ggtitle("(a) Absolute_Preplant") +
    theme(plot.title = element_text(size = 30, face = "bold"))

topdress_cor_matrix <-ggpairs(cor_data_topdress, aes(col=Field, alpha=0.5))+
    ggtitle("(c) Absolute_Topdress") +
    theme(plot.title = element_text(size = 30, face = "bold"))
```

```
ggsave(preplant_cor_matrix, filename = "huehue_preplant.pdf", height = 30, width = 30, units = "cm")
```

```
## Warning in geom_point(): All aesthetics have length 1, but the data has 100 rows.
## i Please consider using `annotate()` or provide this layer with data containing
##   a single row.

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
ggsave(topdress_cor_matrix, filename = "huehue_topdress.pdf", height = 30, width = 30, units = "cm")
```

```
## Warning in geom_point(): All aesthetics have length 1, but the data has 100 rows.
## i Please consider using `annotate()` or provide this layer with data containing
##   a single row.

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

all together

```
all_preplant_cor_matrix <-ggpairs(cor_data_preplant %>% select(-Field))
all_topdress_cor_matrix <-ggpairs(cor_data_topdress %>% select(-Field))

ggsave(all_preplant_cor_matrix, filename = "all_preplant.pdf", height = 30, width = 30, units = "cm")
```

```
## Warning in geom_point(): All aesthetics have length 1, but the data has 81 rows.
## i Please consider using `annotate()` or provide this layer with data containing
##   a single row.
```

```
ggsave(all_topdress_cor_matrix, filename = "all_topdress.pdf", height = 30, width = 30, units = "cm")
```

```
## Warning in geom_point(): All aesthetics have length 1, but the data has 81 rows.
## i Please consider using `annotate()` or provide this layer with data containing
##   a single row.
```

Correlation matrix (with differences, obtained betweent treatments by blk)

Obtain differences by blk and year

For preplant

```
difference_preplant <- preplant %>%
  group_by(Year, Blk) %>%                 # Group by Year and Blk
  filter(all(c("CR", "RF") %in% Field)) %>%  # Ensure both "CR" and "RF" exist
  summarise(
    TotalP = TotalP[Field == "CR"] - TotalP[Field == "RF"],
    TotalIV = TotalV[Field == "CR"] - TotalV[Field == "RF"],
    TotalC = TotalC[Field == "CR"] - TotalC[Field == "RF"],
    TotalS = TotalS[Field == "CR"] - TotalS[Field == "RF"],
    Total_phenols = Total_phenols[Field == "CR"] - Total_phenols[Field == "RF"],
    crop_fertiliser_N = crop_fertiliser_N[Field == "CR"] - crop_fertiliser_N[Field == "RF"],
    crop_soil_N = crop_soil_N[Field == "CR"] - crop_soil_N[Field == "RF"],
    soil_tracer_N = soil_tracer_N[Field == "CR"] - soil_tracer_N[Field == "RF"],
    yield = yield[Field == "CR"] - yield[Field == "RF"]
  ) %>%
  ungroup()
```

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

For topdress

```
difference_topdress <- topdress %>%
  group_by(Year, Blk) %>%                # Group by Year and Blk
  filter(all(c("CR", "RF") %in% Field)) %>%  # Ensure both "CR" and "RF" exist
  summarise(
    TotalP = TotalP[Field == "CR"] - TotalP[Field == "RF"],
    TotalIV = TotalV[Field == "CR"] - TotalV[Field == "RF"],
    TotalC = TotalC[Field == "CR"] - TotalC[Field == "RF"],
    TotalS = TotalS[Field == "CR"] - TotalS[Field == "RF"],
    Total_phenols = Total_phenols[Field == "CR"] - Total_phenols[Field == "RF"],
    crop_fertiliser_N = crop_fertiliser_N[Field == "CR"] - crop_fertiliser_N[Field == "RF"],
    crop_soil_N = crop_soil_N[Field == "CR"] - crop_soil_N[Field == "RF"],
    soil_tracer_N = soil_tracer_N[Field == "CR"] - soil_tracer_N[Field == "RF"],
    yield = yield[Field == "CR"] - yield[Field == "RF"]
  ) %>%
  ungroup()
```

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

Remove unneeded columns

```
cor_data_difference_preplant <- difference_preplant%>%select(-Blk,-Year)
cor_data_difference_topdress <- difference_topdress%>%select(-Blk,-Year)
```

all together

```
all_difference_preplant_cor_matrix <-ggpairs(cor_data_difference_preplant) +
  ggtitle("(b) Difference_Preplant") +
  theme(plot.title = element_text(size = 30, face = "bold"))
```

```
all_difference_topdress_cor_matrix <-ggpairs(cor_data_difference_topdress) +
  ggtitle("(d) Difference_Topdress")+
  theme(plot.title = element_text(size = 30, face = "bold"))
```

```
ggsave(all_difference_preplant_cor_matrix, filename = "all_difference_preplant.pdf", height = 30, width
```

```
## Warning in geom_point(): All aesthetics have length 1, but the data has 81 rows.
## i Please consider using `annotate()` or provide this layer with data containing
##   a single row.
```

```
ggsave(all_difference_topdress_cor_matrix, filename = "all_difference_topdress.pdf", height = 30, width
```

```
## Warning in geom_point(): All aesthetics have length 1, but the data has 81 rows.
## i Please consider using `annotate()` or provide this layer with data containing
##   a single row.
```