Detecting the Mine Site Impact on Rarerus Plantus Chlorophyll Level

Author: Franco Meng, 2023-04-30

Abstract/Introduction

Mining, has always been in the centre of the environmental concern. The aim for this short paper is to determine, whether the mining activity has had an impact on a particular plant species known as Rarerus Plantus.

The data provided by DAA, contains 139 Rarerus Plantus's sample records, each sample has 37 bimonthly observations of chlorophyll measurement over 6 years (Nov 16 – Nov 22), the data also indicates whether each plant was located near (<50m) or far (>50m) from the mine site activity. The mine activity commenced in June 2017.

For the analysis, multiple inspections, plots, transformations were conducted.
The dependent variable: chlorophyll measurements, do not follow normal distribution.

The non parametric test: Friedman test was used, followed by post hoc test: Pairwise comparisons using Wilcoxon rank sum test. From the test I conclude:

If the mean yearly level of Chlorophyll deemed to be the same over the year in nature, without the influence of the mine site. I have 95% confidence that mining activity has had an impact on the plant Rarerus Plantus. The mining activities may also have had a long-lasting impact.

It is important to note that the amount of chlorophyll in plants can be also impacted by nutrient availability, light, and environmental stress. P.Palta (1990)

Methodology / Results

R statistical tool and R markdown, with multiple packages (ggplot2, ImerTest, reshape2, etc) were used. Statistical significance was determined at 5% level.

After the inspections, transformations of the data, no missing value was found, chlorophyll measurements were treated as continuous variable, where location, time, were treated as categorical variable. Two plots were attached in Appendix:

Figure 1 : Data was split into two groups based on the condition of <50m of >50m, each bar represents the mean value of all the chlorophyll values collected within a certain group, at a certain time.

The mean Chlorophyll measurement tend to vary seasonally each year. January has reported the constant lowest mean chlorophyll levels every year. One particular chlorophyll measurement of 48.9, was recorded for a sample on Jan 22. This was identified as outlier for the analysis, and all the records for that plant has been removed. After the removal of the outlier, the average of all the chlorophyll measurements is 0.7268.

Figure 2 : a histogram of all the chlorophyll measurements . (outlier excluded)

The dependent variable: chlorophyll measurements, do not follow normal distribution.

Step 1

Based on many studies on chlorophyll concentration in plants, it is expected see some seasonal variations in chlorophyll levels over the time in nature.

To confirm that, a subset of the raw data, including Nov.16, Jan.17, Mar.17, and May.17 measurements taken before the commencement of mining activity (Jun.17), were firstly extracted for analysis.

Due to the non-normality of the data, the non-parametric equivalent of repeated measure anova: Friedman test was used, to compare the mean ranks between the related groups. The following three assumptions have been made.

1. The observations of all bimonthly measures for each plant, were independent of each other.
2. The plant sampling was randomly selected from the population.
3. The homogeneity of variance: the variance of the differences between bimonthly measurement was equal.

```
Friedman chi-squared = 168.56, df = 3, p-value < 2.2e-16
```

From the p value result above, I have 95% confidence to conclude, statistically based on the given data, before the start of mining activity, there already has had significant mean differences of chlorophyll level in Rarerus Plantus bimonthly.

Step 2

Due to the above tested results, it may not be the best practice to use all the bimonthly measure to examine the mining impact, as the chlorophyll level has shown statistically significant variation over time bimonthly already.

Because only 4 bimonthly measurements (Nov-16, Jan-17, Mar-17, May-17), were recorded before the mining activities begin. An average of these 4 bimonthly measurements was calculated, for each plant, and the same was done for the each year, after the mining activity has begun, in order to compare with the mean chlorophyll level before the mining activity.

The July and September bimonthly records were not included for the analysis, as there was no July and September data recorded before the start of mining activity.

A total of 6 mean measurements for each plant has been calculated, to represent each year's mean chlorophyll levels between November to May. I have further assumed there was no significant change of these "Nov-May" means in nature, over the 6 recorded years.

A paired box plot was produced as below (Appendix: Figure 3), to visualize the 6 mean measurements over the year, the mean chlorophyll level has shown some variations over the years in both <50m and >50m groups.

The data was again transformed into long format, grouped into two groups: near (<50m), and far (<50m). Friedman tests were performed on both groups. Results are presented below :

```
near (<50m) the mine:      Friedman chi-squared = 34.702, df = 5, p-value = 1.726e-06
------------------------------------------------------------------------------
far (>50m) from the mine:  Friedman chi-squared = 94.828, df = 5, p-value < 2.2e-16
```

If I had assumed the yearly mean chlorophyll levels stays the same over the year without the impact of mine activities, and the pre mine collected data between Nov16 - May17, reflected the true natural chlorophyll level in Rarerus Plantus.

I then have 95% confidence to conclude that statistically, the mine site has had an impact on the Rarerus Plantus Chlorophyll Levels, regardless whether the plants located within 50 meters, or over 50 meters.

Discussion

1. The Pairwise Wilcoxon Rank Sum Test, with Bonferroni correction was used as Post-hoc test. There were ties occur in the <50m group and this group has fewer than 50 plant observations, the test returned a warning message "cannot compute exact p-values with ties"

Therefore, one Pairwise Wilcoxon Rank Sum Test has been carried out on the whole data, both near and far group inclusive. From the p value matrix result below we have 95% confidence to conclude, The mean chlorophyll levels for each year since the start of the mine activity, are all significantly different from the pre mine site mean chlorophyll levels, except for the fourth year Nov.20 - May.21

P value result of Pairwise Wilcoxon Rank Sum Test: (Mine activity started in June.17)

	Nov.17-May.18	Nov.18-May.19	Nov.19-May.20	Nov.20-May.21	Nov.21-May.22
Nov.16-May.17	0.00416	8.7e-13	5.4e-07	0.27049	6.5e-06

If I had assumed the yearly mean chlorophyll levels stays the same over the year in nature, and the pre mine collected data between Nov16 - May17, reflect the true natural chlorophyll level in Rarerus Plantus population.**

I am 95% confident that statistically, the mine activity has had a long lasting impact.

2. The mixed model regression was conducted, to have better insight on the mean chlorophyll variation trend over the years, the time was treated as continuous variable in the mixed model. This can be used if any prediction needs to be made, the output was attached in Appendix.

3. There was no details on mine location, weather condition, and how the chlorophyll measuring was carried out, or if the date/time of each bimonthly observation was constant, the measurement unit was also unknown.

4. There was no details on why 50m has been decided as near or far classifications, and how far was the furthest sample located.

5. Many factors may contribute to the variation of mean chlorophyll detected in this paper, such as rainfall, extreme weather, global warming. A better way to record data for " controlled group", should be discussed and designed.

References

P.Palta, Jiwan. 1990. "Leaf Chlorophyll Content." *Remote Sensing Reviews* 5 (1): 207–13. https://doi.org/10.1080/02757259009532129.

Appendix

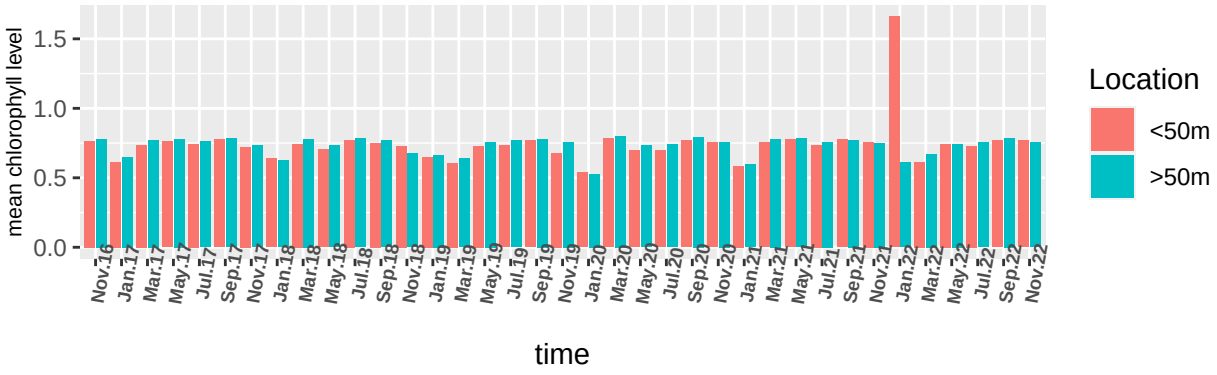

Figure 1 : Mean Rarerus Plantus Chlorophyll leavel of each bimonthly measurement

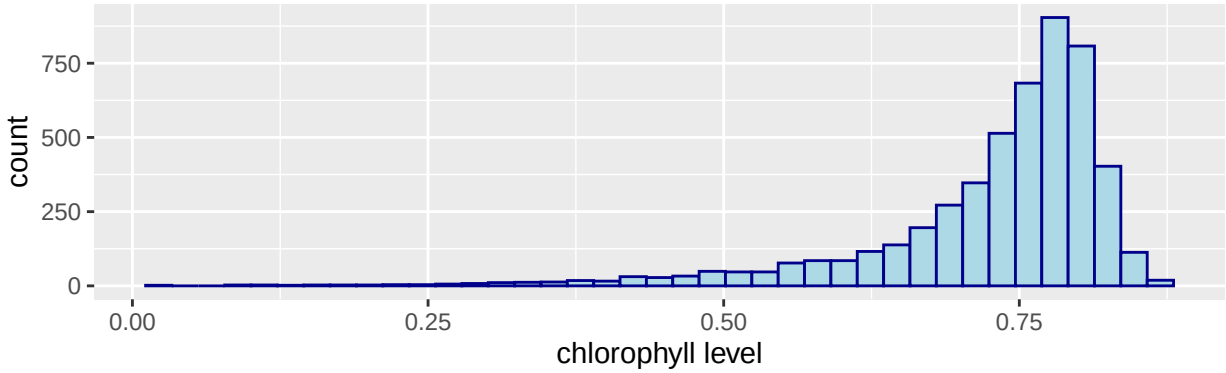

Figure 2 : Histogram of observed Rarerus Plantus Chlorophyll leavel

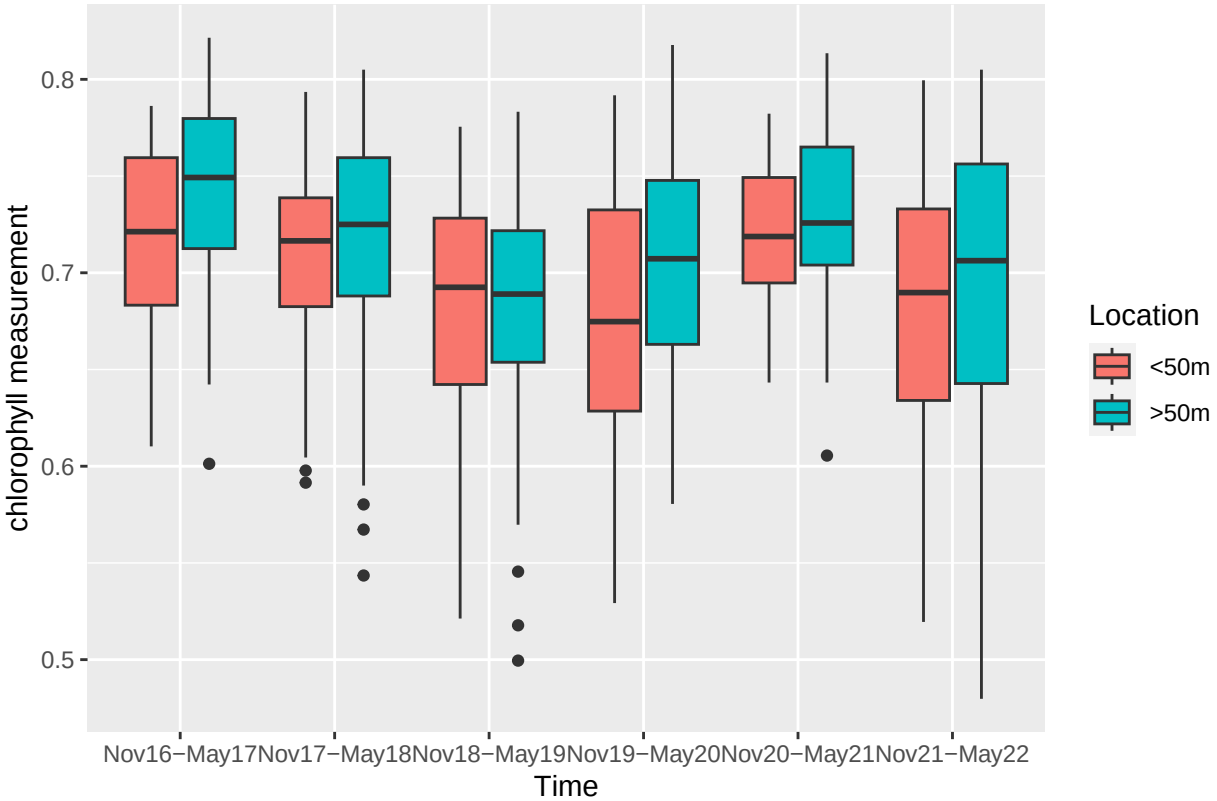

Figure 3 : Paired box plots of Rarerus Plantus Chlorophyll leavels

```
 Linear mixed model fit by maximum likelihood . t-tests use Satterthwaite's method ['lmerModLmerTest']
 Formula: value ~ time + distance + (1 | ID)
 Data: mixed_long3

 AIC      BIC      logLik  deviance   df.resid
-------------------------------------------------
-2490.1  -2466.5   1250.0  -2500.1      823
-------------------------------------------------
 Scaled residuals:
-------------------------------------------------
 Min      1Q     Median      3Q     Max
-3.5792 -0.5172  0.1148  0.6592  2.1422

Random effects:
Groups   Name        Variance Std.Dev.
ID       (Intercept) 0.001342 0.03663
Residual             0.002214 0.04705
Number of obs: 828, groups:  ID, 138

Fixed effects:
                Estimate Std. Error     df  t value    Pr(>|t|)
                ------------------------------------------------
(Intercept)   7.148e-01  7.018e-03  2.276e+02 101.855  < 2e-16 ***
                ------------------------------------------------
time         -5.210e-03  9.575e-04  6.900e+02  -5.442 7.34e-08 ***
                ------------------------------------------------
distance>50m  1.566e-02  7.511e-03  1.380e+02   2.085   0.0389 *

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
              (Intr)  time
 time        -0.477
 distanc>50m -0.721  0.000
```