Response variables:

In `RW28 data for prelims.xlsx`

1. Height (ht)
2. Volume (vol)
3. Diameter at breast height (DBH)
   1. Root collar diameter (RCD) is only used for trees too short (young) to measure DBH. Can use RCD and ht to estimate volume (as cone) where DBH is absent.

Predictor variables:

In `RW28 data for prelims.xlsx`

1. Total P (totp) = Cumulative P added in previous rotation (cumP) plus establishment P fertilizer added at recent planting.

Hypothesis 1: Tree growth measurements (ht, vol, dbh) will increase with increasing `totP` but the growth response to `totP` will be different for plots with establishment P fertilization (cumP ≠ totP) compared to plots with only carryover P (cumP =totP).

Different levels of cumP and establishment P (totP-cumP) may be used to define categorical levels for fertilization treatments\* (as in `RW 28 data\_foliar\_and\_soil\_nutrients.xlsx`, sheet = `treatments`, field = `TRT\_DESCRIP`)

(\*see response to question 2 in Prelim response examples below).

In `RW 28 data\_foliar\_and\_soil\_nutrients.xlsx`, sheet = `nut\_soil`

1. SOIL\_P:

Mehlich III extractable P (use only values where YST = 0: collected before addition of establishment P at planting)

* 1. Hypothesis 2a: tree growth will increase with increasing Mehlich III soil P values in upper 15 cm of soil (and within beds on bedded sites) and this response will be different for plots with establishment P fertilization compared to plots with only carryover P.

Filter P\_SOIL: only values where

DEPTH = 1 (0-15 cm)

and

SAMPLING\_LOCATION = 1 (within bedded row), or 18 (composite of bed and interbed), or 19 (no bedding).

* 1. Hypothesis 2b: tree growth will increase with increasing Mehlich III soil P values in upper 15 cm of soil (and across site, including bed and interbed) and this response will be different for plots with establishment P fertilization compared to plots with only carryover P.

Filter P\_SOIL: only values where

DEPTH = 1 (0-15 cm)

If bedded and not composited (SAMPLING\_LOCATION = 1 or 2), take average of SOIL\_P for SAMPLING\_LOCATION = 1 (within bedded row), and SAMPLING\_LOCATION = 2 (between bedded rows).

* 1. Hypothesis 2c: tree growth will increase with increasing Mehlich III soil P values in upper 30 cm of soil (and within beds on bedded sites) and this response will be different for plots with establishment P fertilization compared to plots with only carryover P.

Take average P\_SOIL in upper 30 cm: for DEPTH = 1 (0-15 cm) and DEPTH = 2 (15-30 cm)

and

filter SAMPLING\_LOCATION = 1 (within bedded row), or 18 (composite of bed and interbed), or 19 (flat, no bedding).

* 1. Hypothesis 2d: tree growth will increase with increasing Mehlich III soil P values in upper 30 cm of soil (and across site) and this response will be different for plots with establishment P fertilization compared to plots with only carryover P.

Take average P\_SOIL in upper 30 cm: for DEPTH = 1 (0-15 cm) and DEPTH = 2 (15-30 cm)

If bedded and not composited (SAMPLING\_LOCATION = 1 or 2), take average of mean SOIL\_P for SAMPLING\_LOCATION = 1 (within bedded row), and SAMPLING\_LOCATION = 2 (between bedded rows).

**Example analysis and comments from Ben’s preliminary exam-**

1. The RW28 is a complicated design. Using the RW28 data provided, explore the height and volume (where possible) differences due to treatments. Look at cumulative response and response compared to the control. Are there any differences due to site/soil so far?

Height and volume from most recent measurements at each site were modeled against the total P quantities using a quadratic model\*. Treatments are grouped as “Carryover” (P added in previous rotation) = C, “Establishment” (E) = P added in previous rotation plus P added at establishment of RW28 trials, an “control” (K), no P or other fertilizer added in previous or current rotation. The model was defined with a difference parameterization, as , where ‘E’ serves as a dummy variable to indicate the “Establishment” group. Hypothesis tests were conducted to determine whether the slope and quadratic term of the model associated with the E group are significantly different from the control and carryover groups. Trendlines are plotted with standard error indicated in gray.

For responses compared to the control, the height or volume values for each replicate block were subtracted from height or volume of each treatment in the corresponding replicate block, and pairwise t-tests were used to determine differences between each treatment and control.

Site **280601**

Cumulative height at YST 1,

A graph of different heights

AI-generated content may be incorrect.

Overall model fit R2 = 0.2758

|  |  |
| --- | --- |
| **coefficient** | **p-value** |
| Intercept | 1.83e-15 |
| Totp | 0.915 |
| Totp^2 | 0.845 |

Hypothesis test that E group slope and quadratic term are significantly different.

|  |  |
| --- | --- |
| **coefficient** | **p-value** |
| E\*Totp | 0.6809 |
| E\*Totp^2 | 0.5644 |

A graph of a patient's disease

AI-generated content may be incorrect.

Only the 361 treatment is different from the control, with weak significance p = 0.047.

Site **284501**

Cumulative height at YST 1,

A graph of different types of height and weight

AI-generated content may be incorrect.

Overall model fit R2 = 0.5255

|  |  |
| --- | --- |
| **coefficient** | **p-value** |
| Intercept | 1.8e-08 |
| Totp | 0.290 |
| Totp^2 | 0.889 |

Hypothesis test that E group slope and quadratic term are significantly different.

|  |  |
| --- | --- |
| **coefficient** | **p-value** |
| E\*Totp | 0.5146 |
| E\*Totp^2 | 0.4245 |

No significant difference between parameters for E and C groups

A graph of a number of blue rectangular objects

AI-generated content may be incorrect.

No treatments are are significantly different from controls

Site **281303**

Cumulative height at YST 3, removed outlier plot from treatment group 150 (landing)

A graph of height vs total p by group

AI-generated content may be incorrect.

Overall model fit R2 = 0.6044

|  |  |
| --- | --- |
| **coefficient** | **p-value** |
| Intercept | 6.05e-13 |
| Totp | 0.0178 |
| Totp^2 | 0.0486 |

Hypothesis test that E group slope and quadratic term are significantly different.

|  |  |
| --- | --- |
| **coefficient** | **p-value** |
| E\*Totp | NA |
| E\*Totp^2 | NA |

Here, the use of total P as a continuous variable is problematic, with only three values. ANOVA would be better approach here perhaps.

A graph of a number of blue rectangular objects

AI-generated content may be incorrect.

The 150 (carryover only) and 151 (carryover plus establishment) treatments are different from the control, with moderate to weak significance p= 0.011 and 0.045, respectively.

Cumulative vol at YST 3

A graph of a graph of a graph

AI-generated content may be incorrect.

Overall model fit R2 = 0.1007

|  |  |
| --- | --- |
| **coefficient** | **p-value** |
| Intercept | 4.848346e-08 |
| Totp | 3.074987e-01 |
| Totp^2 | 3.445039e-01 |

Hypothesis test that E group slope and quadratic term are significantly different.

|  |  |
| --- | --- |
| **coefficient** | **p-value** |
| E\*Totp | NA |
| E\*Totp^2 | NA |

A graph of a number of blue squares

AI-generated content may be incorrect.

No significant differences are observed between control and other treatments, lowest p value for 150 at p = 0.115.

Site **282401**

Cumulative height at YST 4, with plot 2960 removed (assuming measurement error)

A graph of different heights

AI-generated content may be incorrect.

Overall model fit R2 = 0.3717

|  |  |
| --- | --- |
| **coefficient** | **p-value** |
| Intercept | 3.13e-08 |
| Totp | 0.0385 |
| Totp^2 | 0.0603 |

Hypothesis test that E group slope and quadratic term are significantly different.

|  |  |
| --- | --- |
| **coefficient** | **p-value** |
| E\*Totp | 0.8494 |
| E\*Totp^2 | 0.7174 |

Cumulative volume at YST 4,

For comparison with control, treatments with only 1 replicate were removed (720, 960)

A graph of a patient's response

AI-generated content may be incorrect.

No significant differences were found according to p-value thresholds. However, standard error bars indicate that the establishment P treatments may be different from the control, albeit with no or very weak significance.

A graph of a graph showing different trends

AI-generated content may be incorrect.

Overall model fit R2 = 0.446

|  |  |
| --- | --- |
| **coefficient** | **p-value** |
| Intercept | 0.02706295 |
| Totp | 0.03250390 |
| Totp^2 | 0.08119848 |

Hypothesis test that E group slope and quadratic term are significantly different.

|  |  |
| --- | --- |
| **coefficient** | **p-value** |
| E\*Totp | 0.7571 |
| E\*Totp^2 | 0.9449 |

A graph of a patient's response

AI-generated content may be incorrect.

Similar to heights, no significant differences are observed, but error bars indicate some differences among all but 360 treatments compared to control.

Site **284201**

Cumulative height at YST 4,

A graph of different types of height and weight

AI-generated content may be incorrect.

Overall model fit R2 = 0.6718

|  |  |
| --- | --- |
| **coefficient** | **p-value** |
| Intercept | 9.47e-10 |
| Totp | 0.0167 |
| Totp^2 | 0.0161 |

Hypothesis test that E group slope and quadratic term are significantly different.

|  |  |
| --- | --- |
| **coefficient** | **p-value** |
| E\*Totp | 0.7678 |
| E\*Totp^2 | 0.8607 |

For comparisons of treatment to control, trt group 480 was removed (only one replicate)

A graph of a patient's treatment

AI-generated content may be incorrect.

As above, no significant differences are observed between treatment and controls

Cumulative volume at YST 3,

A graph of a graph of a number of individuals

AI-generated content may be incorrect.

Overall model fit R2 = 0.5099

|  |  |
| --- | --- |
| **coefficient** | **p-value** |
| Intercept | 0.000195 |
| Totp | 0.095056 |
| Totp^2 | 0.102070 |

Hypothesis test that E group slope and quadratic term are significantly different.

|  |  |
| --- | --- |
| **coefficient** | **p-value** |
| E\*Totp | 0.6809 |
| E\*Totp^2 | 0.5644 |

A graph of a number of blue rectangular objects

AI-generated content may be incorrect.

Only the 961 treatment is significantly different from the control, with weak significance p = 0.04813002.

No significant differences were observed for the model coefficients corresponding to the E (establishment P) group compared to the carryover groups in any site, suggesting that there is not a difference between the carryover and establishment P groups. In other words, the timing of P application at establishment does not appear to be different than the same total P applied prior to establishment, suggesting a carryover effect of P applied in previous rotations. For response differences compared to the controls, we see differences between some treatments and the control, but with no obvious trend. Treatment 361 (low carryover plus establishment P) in site 2806101, treatments 150 (high carryover) and 151(high carryover plus establishment P) in site 281303 showed significant height differences compared to controls, and only treatment 961 (medium carryover with additional P prior to rotation, plus establishment P) in site 284201 showed significant volume response differences compared to the controls. These differences suggest that the responses to treatment may vary by site and soil, but given data is insufficient to clarify this with much confidence.

1. Now give one example of, what you would consider, an inappropriate way to analyze the data and why.

Given the high number of distinct treatment levels (8 per site, except in 281303) and the low number of replicates at each site, regarding the treatments as categorical variables and performing an anova test to determine differences between all treatments could be problematic. With 7 degrees of freedom occupied by the treatments, this leaves only 8 degrees of freedom for the residuals. For the ANOVA to accurately explain the remaining variation of the residuals after accounting for treatment effects, it is important to have at least as many residual degrees of freedom as treatment levels. Therefore, taking the total P concentration as a continuous variable, fitting the data to a linear or (or nonlinear) model, and using contrasts to look for trend differences between the two main groups (carryover and carryover+establishment P) provides more statistical power to examine differences between 2 groups, rather than 8. Alternatively, pairwise t=tests between each treatment group and the control may be used to determine responses compared to the control, as above, but again these results should be interpreted with caution given the small sample size.

For the 2810303 group, however, the use of total P as a continuous variable is problematic, because the total P values only occur at 3 distinct points, therefore the model is not provided sufficient data to predict a trendline between the lowest and highest total P values. Here, an ANOVA test, followed by a Tukey multiple comparisons test, may be appropriate to determine differences among treatments.

\*The quadratic model used above may not be the best fit for this data, as suggested by the significance of coefficients for the overall model for each site. It may be that other non-linear models provide a better fit for the data. The Mitscherlich model was tested for these data, assuming a possible asymptote, but these were very poor fits for the data, and the visual scatterplots from most sites suggest a trend that does not asymptote, but follows a polynomial trend that appears to increase to a maximum in the mid-range of total P values, then decrease with increasingly higher total P values.