



Preddiction lines

Mean height (z-scores)

Shade level

- 1
- 0
- 1

2  
1  
0  
-1

-1.0

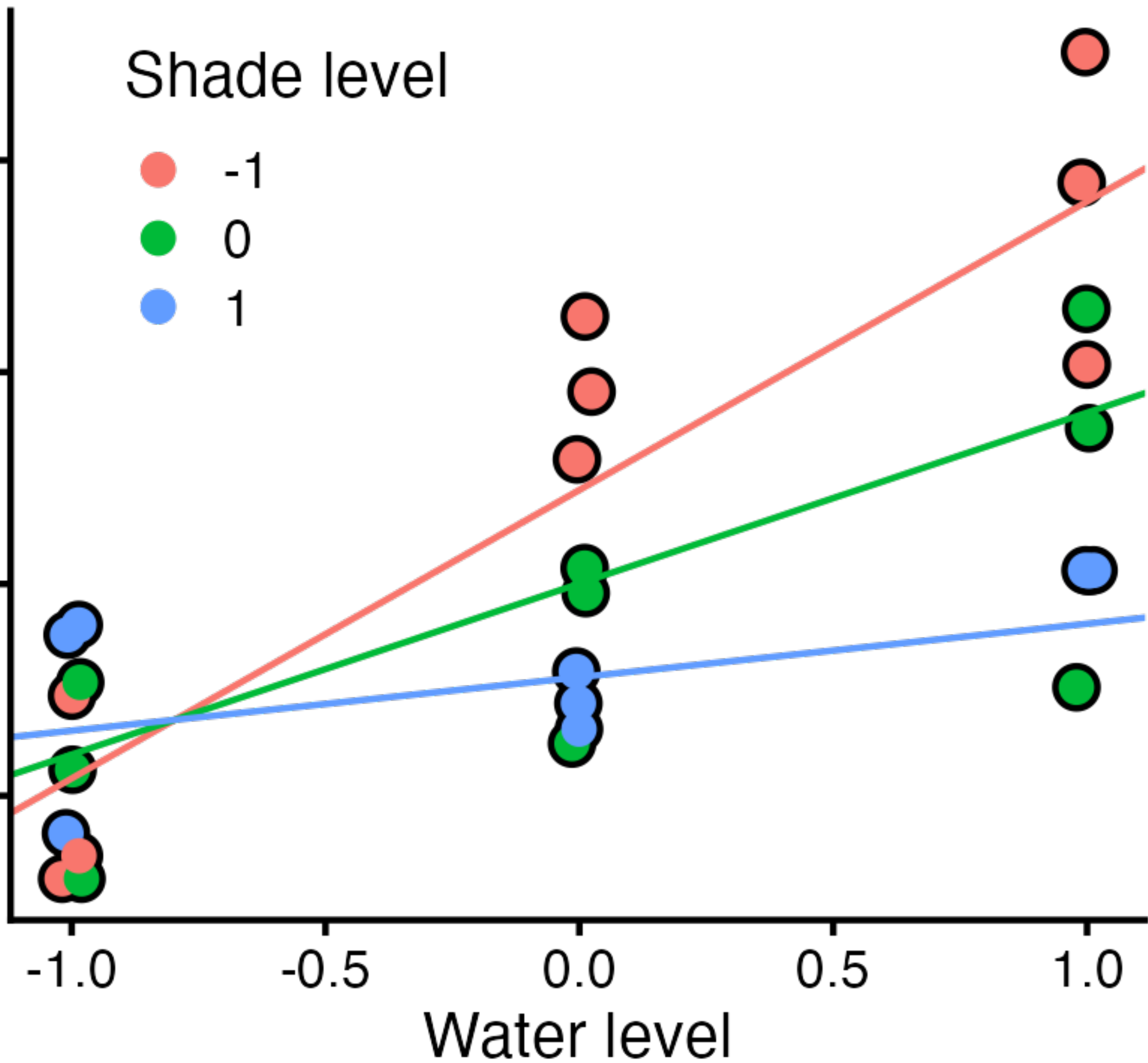
-0.5

0.0

0.5

1.0

Water level

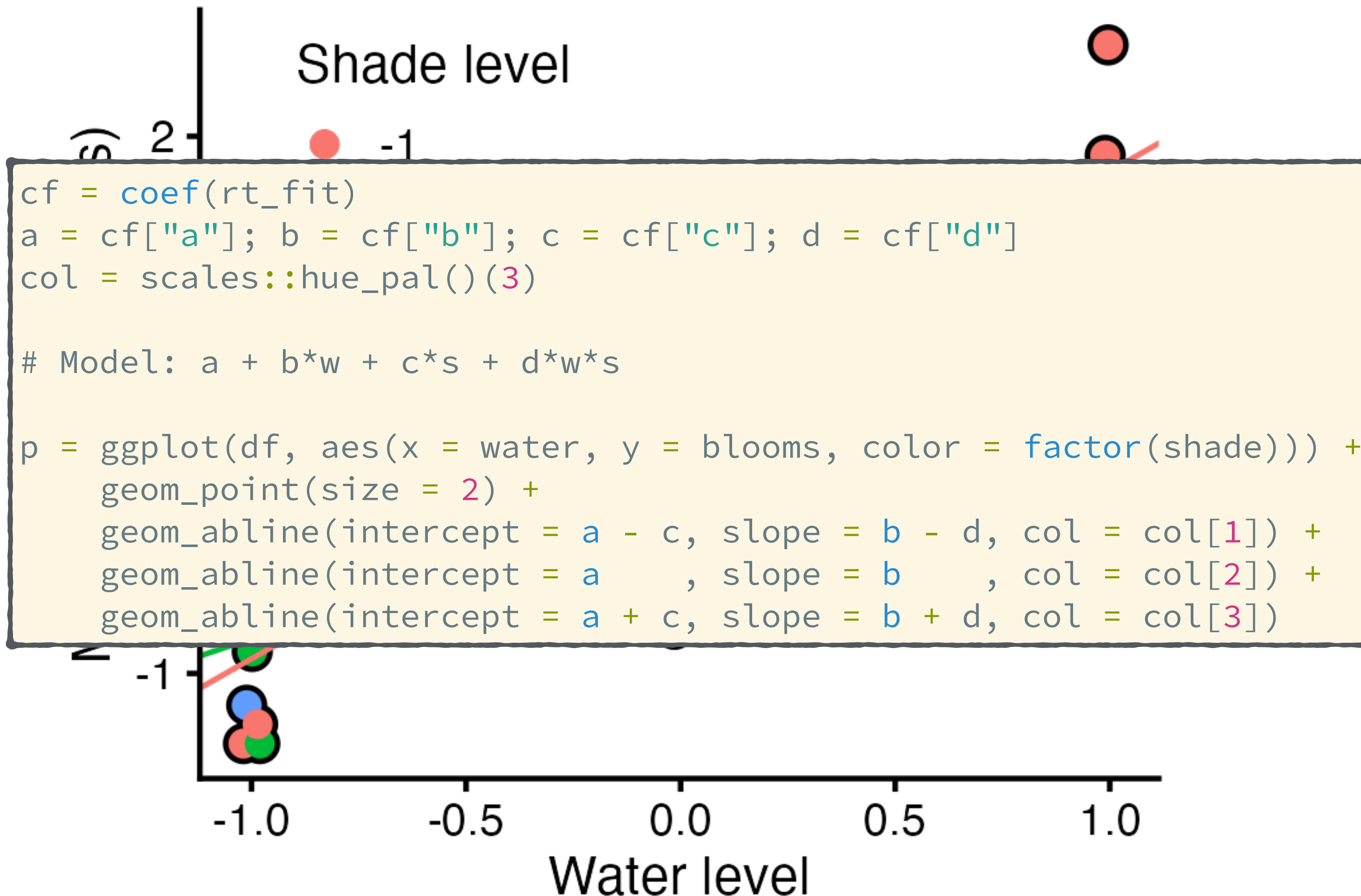


```
cf = coef(rt_fit)
a = cf["a"]; b = cf["b"]; c = cf["c"]; d = cf["d"]
col = scales::hue_pal()(3)

# Model: a + b*w + c*s + d*w*s

p = ggplot(df, aes(x = water, y = blooms, color = factor(shade))) +
  geom_point(size = 2) +
  geom_abline(intercept = a - c, slope = b - d, col = col[1]) +
  geom_abline(intercept = a, slope = b, col = col[2]) +
  geom_abline(intercept = a + c, slope = b + d, col = col[3])
```

# Prediction lines



# Summary

- Multiple linear models allow us to use more than one predictor in a linear model
  - These models do a form of automatic **stratification**
    - **Ex:** difference in size for individuals of the same age, effect of treatment for individuals of the same size
  - The objective is to compare **like-to-like**
- Coefficients can and do change with the inclusion of more predictors
  - Coefficient interpretation is hard, use plots, predictions, scaling and transformations to make models easier to interpret
  - Next week, we talk about principled ways of choosing if a variable should be added to a model, stay tuned!