# Multiple regression and general linear models

# ANOVA and regression

- Closely related to each other
  - same least square principle

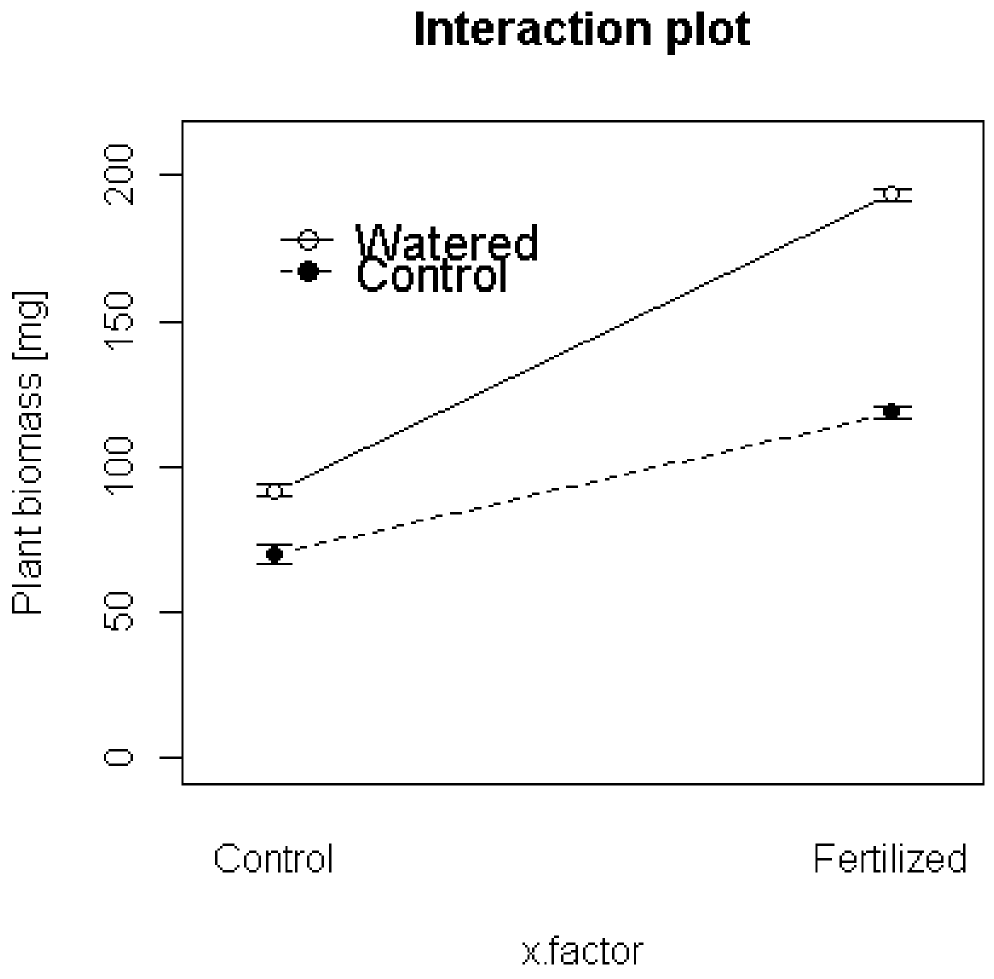
- ANOVA with a predictor of n levels is analogous to a multiple linear regression with n-1 predictors
  - *a priori* defined contrasts

# Models with multiple predictors

- Two-way (Multiple-way) ANOVA
  - response ~ factor.1 + factor.2 + ...
- Multiple regression
  - response ~ predictor.1 + predictor.2 + ...
- Additive effects vs. interaction
  - additivity – response on factor.1 (predictor.1) does not depend on the value of factor.2 (predictor.2)
  - additivity can be statistically tested and rejected in favor of **interaction**

# Interaction

- Significant interaction indicates a relationship between the effects of the predictors
  - $y = a + bx_1 + cx_2 + dx_1x_2 + \varepsilon$
- Test of interaction H0: d = 0
  - d > 0: positive interaction, higher values of response compared to additivity
  - d < 0 : negative interaction, lower values of response compared to additivity
- $df_{int} = df_{x1} * df_{x2}$
- Interaction plot
  - Plotting of interaction
  - Under $H_0$, the lines connecting factor levels would be parallel
- **Interaction does not mean interdependence of predictors!**



# General linear models

- Allow an analysis of the dependence of **a single response variable** on multiple predictors of whatever nature (continuous or categorical)
- This is possible because of the equivalence of ANOVA and regression
- Include e.g. Analysis of covariance (linear model with a single continuous and multiple categorical predictors)
- In R: function lm

# Model selection in LMs

- Not all candidate predictors are significant but only the significant ones should be included in the model
- Statistical theory provides little help for predictor selection but we can compare models differing in their predictor structure
- Stepwise selection
  - Forward stepwise: most significant predictors are added; suitable for observatory data
  - Backward stepwise: non-significant predictors are removed; suitable for experimental data (e.g. with interactions)
  - Both directions: the iterative approach used in modern software

# Akaike Information Criterion (AIC)

- Quantifies the information accounted for by a predictor
  - allows comparisons between predictors with different numbers of df model
  - lower AIC suggests a better fit, absolute values of AIC are not informative
- AIC = 2k – 2log(L) , (log = natural logarithm), k is number of model parameters (i.e. df model in lm)
  - in linear models AIC = 2k – 2 log (n/RSS) + C, where RSS is residual sum of squares, C is constant (can be ignored)
- Combination with an F-test of significance
  - Order importance of predictors based on AIC
  - Exclude those that are not significant in F-test
  - Pragmatic approach not supported by statistical theory