ECO 602
Analysis of Environmental Data

FALL 2019 – UNIVERSITY OF MASSACHUSETTS

DR. MICHAEL NELSON

Today’s Agenda

1. Things I forgot to mention on Tuesday
2. Models, null models, predictors, responses
3. 4 Scenarios from the McGarigal readings
4. Model thinking exercise
5. Revisiting the role of data and analyses

General course info

1. On Moodle:
   1. Lecture schedule, readings
   2. Calendar of assignments and due dates.
   3. McGarigal readings
   4. My contact info
   5. Office hours: survey

Reference texts

1. Ben Bolker: Ecological Models and Data in R.
2. Alain Zuur et al.: Analysing Ecological Data
3. Zuur et al.: Mixed Effects Models and Extensions in Ecology with R.

Models: predictors and responses

1. Predictors and responses: hypothesized association between components.
2. We try to explain observed patterns in the response by patterns in the predictors.
3. Relationship may be causal, or correlational.
4. Positive relationship is not necessarily evidence of cause and effect.
5. Beware lurking variables (mor on this later)

Null Models (Null Hypotheses)

1. McGarigal: null model = status quo
2. Model expectations if relationships among components are nonexistent.
3. Stochastic-only model
4. Appropriate null models depend on context

Null Models

1. Null model is needed to quantify support for our alternative hypotheses.
2. Key aspect of inferential statistics: expected and observed measurements.
    1. Contrast between what we **expect** to observe vs what we **actually** observe.

Null Models

Hypothetical system: water pollution and algae blooms

Hypothetical system: no prior knowledge

You think that Phosphorus and Nitrogen water pollution concentration is positively associated with algae growth.

Nobody has ever studied this quantitatively.

Hypothetical system: no prior knowledge

What are predictors and responses?

What is a good null hypothesis?

What is a possible alternative hypothesis?

What pattern(s) do you expect if null hypothesis is true.

Possible outcomes

Outcome 1: no relationship found in our samples.

Outcome 2: strong positive relationship found in our samples

Other possibilities?

Comparing a local pattern to accepted global behavior

Suppose lots of previous research has shown strong support for a positive relationship between water pollution and algae.

You observe anecdotally that pollution in lake water in a region with unique soil characteristics seems to suppress algal growth.

Comparing a local pattern to accepted global behavior

What is a good null hypothesis now?

What evidence would support the null?

What patterns would give good evidence that the null is wrong?

Possible outcomes

Outcome 1: no no relationship in samples.

Outcome 2: strong positive relationship in our samples.

Outcome 3: relationship positive for small amounts of pollution, but negative for high concentrations. Possibly a nonlinear relationship

Example 2: Comparing a local pattern to accepted global behavior

What do we expect to observe if the null hypothesis is true?

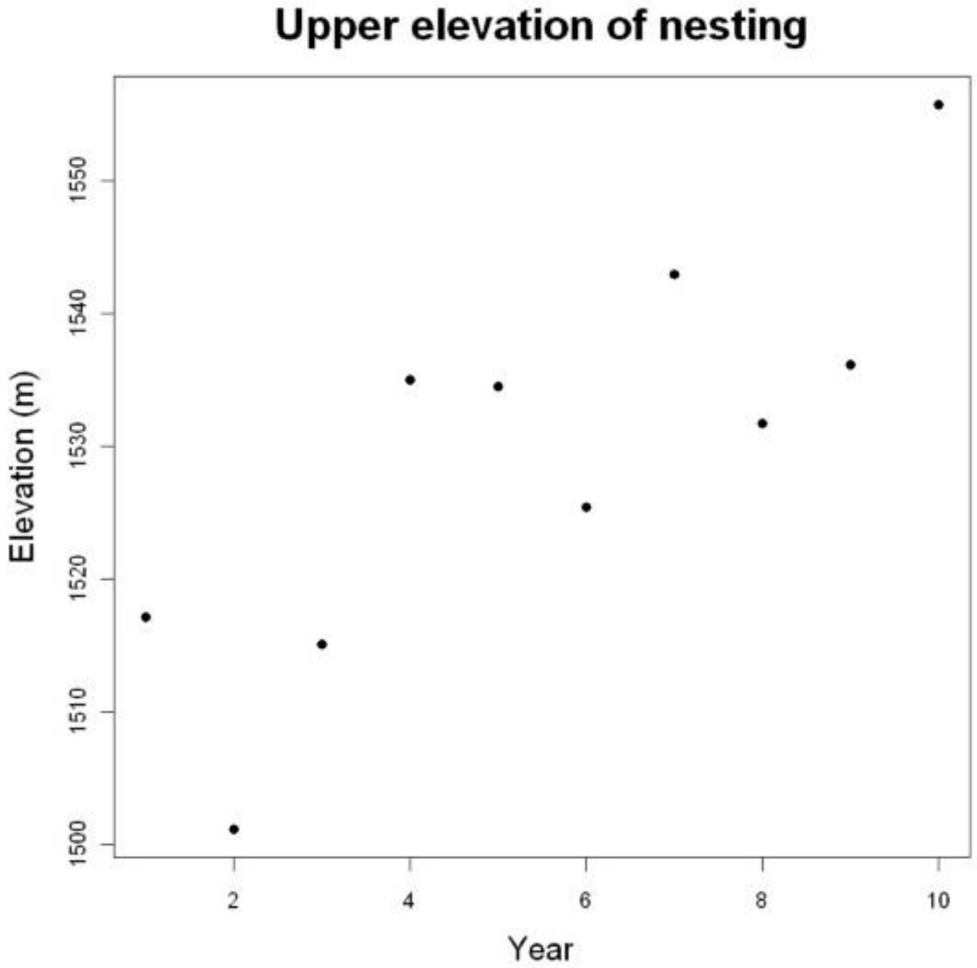
What pattern(s) would show support for rejecting the null?

Testimonies from McGarigal reading

1. For each scenario:
    1. What are the predictors and response?
    2. What is the null model?
    3. Level of support for alternative hypotheses?

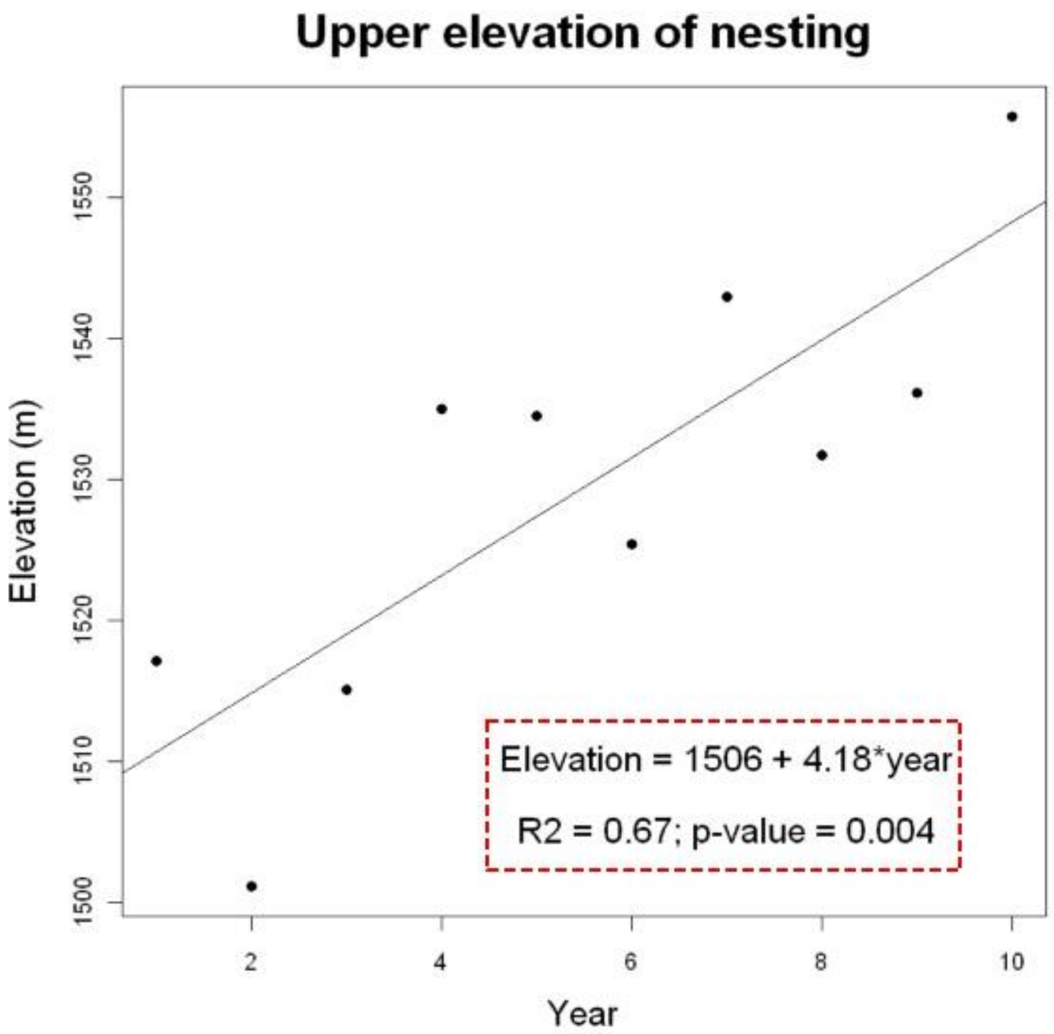
Testimony #1

- Observed increase in upper elevational limit of nesting along an elevational gradient



Testimony #2

- Statistically significant increase in upper elevational limit of nesting along an elevational gradient

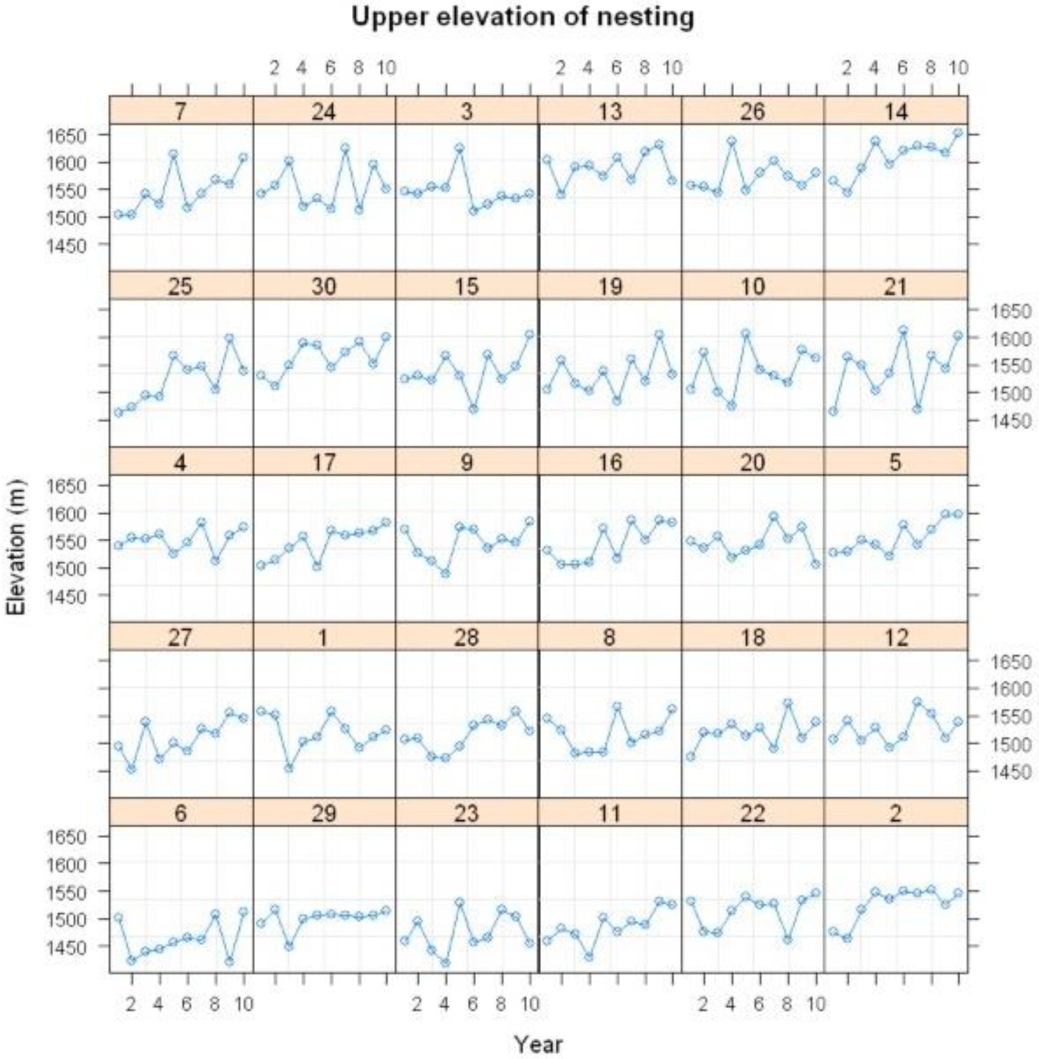


Testimony #3

- Statistically significant increase in upper elevational limit based on 30 survey routes across the Forest along elevational gradients

```
Linear mixed model fit by maximum likelihood
Formula: elev ~ tsteps + (1 + tsteps | plot.id)
   Data: y.sim
  AIC   BIC logLik deviance REMLdev
 2970 2992  -1479     2958    2952
Random effects:
 Groups   Name        Variance   Std.Dev. Corr
 plot.id  (Intercept) 748.643730 27.36135
          tsteps        0.067454  0.25972 1.000
 Residual             887.300857 29.78760
Number of obs: 300, groups: plot.id, 30

Fixed effects:
             Estimate Std. Error t value
(Intercept) 1508.6996     6.2255  242.34
tsteps         4.4981     0.6006    7.49
```



Testimony #4

- Weight of evidence strongly suggests climate as the major culprit, but habitat may play a role as well

	AIC	df	dAIC	weight
climate	2578.5	6	0.0	0.799
habitat	2581.3	6	2.8	0.201
recreation	2603.9	6	25.4	<0.001
deposition	2631.0	6	52.5	<0.001

Model Thinking: questions to ask yourself about a system

1. What question(s) are useful?
2. How would I make a model of this system?
3. What components would I include (or exclude)?
4. How would they be represented?
5. What types of data could I observe and analyze?
6. What is a candidate null hypothesis?

Modeling Exercise

The role of data and analyses: descriptive and inferential stats

1. Descriptive statistics
    1. Suggest patterns, relationships
    2. Help you build a hypothetical model

The role of data and analyses

1. Models are great.
2. Models are not very useful if they can't make meaningful predictions. They must be fit to data.
3. Modelling, data collection, analyses are iterative.S

The role of data and analyses

1. Data needed for your model to:
    1. Quantify strength of relationships: deterministic model
    2. Quantify uncertainty: stochastic model
    3. Quantify strength of evidence for your hypotheses.
    4. Compare alternative models

Why am I harping on model thinking?

1. Model model (not a type) of data and analyses is sometimes not emphasized enough.
2. When you get into details of a statistical model, it is easy to forget our overall questions.
3. Referring often to your conceptual model of a system helps you remember why you should care about your statistical model.

For next class:

Read McGarigal chapter 2.