Investigating Carp Movement Pattern with Linear Mixed Effects Model

A Summer Collaboration with the Fisheries Department

Yannan Pan, Xiaowan Liu
School of Statistics, University of Minnesota

Introduction

The big idea(s)

What factors may influence carp's 24- hour activity? And **How**?

Our involvement

- Serve as support for theoretical analysis.
- Draw exploratory plots, build a LME models to test the hypothesis, and gave an analysis report.

Using statistics to answer these questions

Plots: average distance per period by phase, sex, day&night, lake, individual. (See the general trend)
Regression: Anova, lsmeans, pairwise comparisons, interaction plots
(Delve into the details)

The Collaborative Process

The first meeting

Went well! We briefly talked about the background and goals, data, deliverables and communication as we expected. After our first meeting, a collaboration plan was developed with additional aspects such as proposed analysis methodology, responsibilities and timelines.

Communication

Regular meeting needed! Every week, we scheduled a meeting with our clients to keep all of us updated.

Email exchange! Before every meeting, we would send a summary of our previous work to the clients; During the meeting, we would discuss in more details and address next goals; After the meeting, our nice clients would always send us a follow-up email to keep everyone on the same page. We really appreciated our mutual cooperation.

We also did a final report integrating all the work we've done, which had been arranged in a **proper order** and with **clear explanation**.

Learning about collaboration

Ask Questions! Problems arise when "statisticians" meet "biologists". Asking about what we didn't understand, rephrasing or repeating what we seemed to understand greatly helped us collaborating with non-statisticians.

Organizing management! Change happens all the time. We strongly felt that we should organize the work such as the data, the model, the summary chronologically.

Data Exploration and Methods

About the data

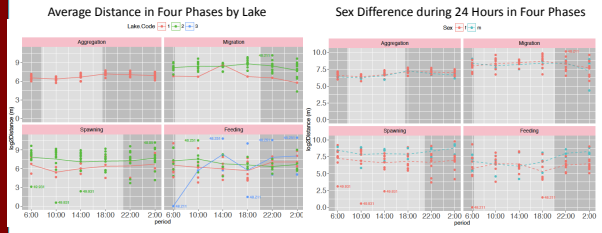
Fish were caught six times a day for 10 months. The time and the distance were recorded.

Data preparation

1. Weighted distance assigned to the 4-hour periods.
2. Time to period and day&night. Date to phase.
3. Distance to logDistance. (Data was skewed)

```
##       ID Sex Lake.Code       Dis       phase daynight period   logDis
## 1 48.211   f         1  66.01224 Aggregation       Day      1 6.066353
## 2 48.211   f         1 159.84158 Aggregation       Day      2 7.329497
## 3 48.211   f         1 168.71287 Aggregation       Day      3 7.406952
## 4 48.211   f         1 188.47780 Aggregation     Night      4 7.565885
## 5 48.211   f         1 164.54894 Aggregation     Night      5 7.371114
## 6 48.211   f         1  42.15350 Aggregation     Night      6 5.431406
```

Exploratory plots



The statistical methods

Why Linear Mixed-Effects Model:

1. The data was measured repeatedly over time (every four hours a day)
2. We want to treat each individual carp as random effects.

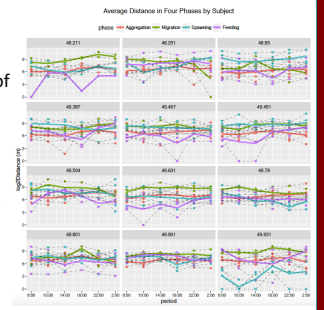
Model: logDistance ~ Phaselake*Sex*Daynight + (ID:day)

1. Significant? Type 2 Anova
2. Specific influence? Least squares means & Pairwise comparisons
3. Interaction? Interaction plots

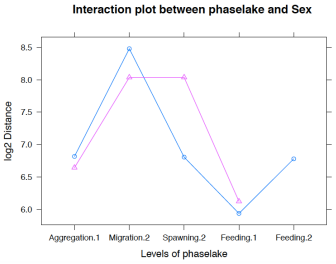
Statistical Analyses

Individual variation

- To assess random effects coming from individual variation, we used "ggplot" to show the trajectories of individuals.
- The dashed lines are the the average distance of each phase, and the dashed lines connect points measured on the same day.
- We found the measurement day might be a source of random effects.



Interaction effects



- Considering our design was unbalanced in terms of phase and lake, we created a combination variable.
  - We used interaction plots to interpret the interaction effects intuitively.
- Males are more active during Spawning phase (May - June) in shallow area!

Results

What did it mean?

- Phase and lake have significant effects on the moving pattern of carp.
- In long lake, carp are more active at night during aggregation phase (Jan - March).

```
## Analysis of Variance Table of type II  with  Satterthwaite
## approximation for degrees of freedom
##                    Sum Sq Mean Sq NumDF  DenDF F.value    Pr(>F)
## Sex                 1.205  1.2054     1  20.60  0.7942 0.3831211
## daynight            9.158  9.1580     1 981.52  6.0339 0.0142057 *
## phaselake          45.131 11.2827     4  22.94  7.4339 0.0005405 ***
## Sex:daynight        0.685  0.6852     1 982.07  0.4515 0.5018003
## Sex:phaselake      16.597  5.5324     3 174.42  3.6451 0.0138635 *
## daynight:phaselake 38.061  9.5154     4 982.66  6.2694 5.549e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

What's next?

Provide advice or suggestions when writing up the statistical part of the client's thesis.

University of Minnesota
Driven to Discover℠