# Washington Lake, Weed Control (Dave Kyle)

Hodge / Dauwalter

11/15/2021

### Weed Removal - The Question

Non-native game fishes predate upon Sockeye Salmon *O. nerka* in Lake Sammamish(?), WA. Removal of aquatic vegetation is an action that would presumably reduce habitat of non-native gamefish predators, thus reducing their populations and predation impact on Sockeye Salmon. This study was designed to answer the question **"Did vegetation removal have an effect on abundance of nonnative game fishes?"**. Vegetation removal treatments were performed at 11 and 22 foot depths at each of 3 sites (Samm Landing, Idylwood Park, and Weowna Park). Fishes were enumerated during scuba surveys, at each treatment site and a paired control site, twice over one summer (survey 1 and 2). A generalized linear mixed model (Poisson link) was used to evaluate the effect of the vegetation treatment on sport fish abundance.

## Resources

Ben Bolker draft chapter on GLMMs: link

The data (from D. Kyle, Trout Unlimited) look like this:

Table 1: Weed removal data.

| | Site | Depth | SurvNum | Type | SurvDate | TimeElapsed | Visibility | VegRemoved | Temp |
|---|---|---|---|---|---|---|---|---|---|
| 8 | Idylwood Park | 11 | 1 | Control | 7/22/2021 | 13 | 12 | 0 | 24 |
| 6 | Idylwood Park | 11 | 1 | Treat | 7/22/2021 | 13 | 12 | 75 | 24 |
| 16 | Idylwood Park | 11 | 2 | Control | 8/5/2021 | 27 | 10 | 0 | 26 |
| 14 | Idylwood Park | 11 | 2 | Treat | 8/5/2021 | 27 | 10 | 75 | 26 |
| 7 | Idylwood Park | 22 | 1 | Control | 7/22/2021 | 13 | 8 | 0 | 24 |
| 5 | Idylwood Park | 22 | 1 | Treat | 7/22/2021 | 13 | 8 | 100 | 24 |
| 15 | Idylwood Park | 22 | 2 | Control | 8/5/2021 | 27 | 6 | 0 | 26 |
| 13 | Idylwood Park | 22 | 2 | Treat | 8/5/2021 | 27 | 6 | 100 | 26 |
| 4 | Samm Landing | 11 | 1 | Control | 7/21/2021 | 16 | 10 | 0 | 24 |
| 2 | Samm Landing | 11 | 1 | Treat | 7/21/2021 | 16 | 10 | 75 | 24 |
| 12 | Samm Landing | 11 | 2 | Control | 8/4/2021 | 30 | 8 | 0 | 25 |
| 10 | Samm Landing | 11 | 2 | Treat | 8/4/2021 | 30 | 8 | 75 | 25 |
| 3 | Samm Landing | 22 | 1 | Control | 7/21/2021 | 16 | 8 | 0 | 24 |
| 1 | Samm Landing | 22 | 1 | Treat | 7/21/2021 | 16 | 8 | 100 | 24 |
| 11 | Samm Landing | 22 | 2 | Control | 8/4/2021 | 30 | 6 | 0 | 25 |
| 9 | Samm Landing | 22 | 2 | Treat | 8/4/2021 | 30 | 6 | 100 | 25 |
| 20 | Weowna Park | 11 | 1 | Control | 9/14/2021 | 15 | 6 | 0 | 23 |
| 18 | Weowna Park | 11 | 1 | Treat | 9/14/2021 | 15 | 6 | 85 | 23 |
| 24 | Weowna Park | 11 | 2 | Control | 9/28/2021 | 29 | 9 | 0 | 18 |

| | Site | Depth | SurvNum | Type | SurvDate | TimeElapsed | Visibility | VegRemoved | Temp |
|---|---|---|---|---|---|---|---|---|---|
| 22 | Weowna Park | 11 | 2 | Treat | 9/28/2021 | 29 | 9 | 85 | 18 |
| 19 | Weowna Park | 22 | 1 | Control | 9/14/2021 | 15 | 6 | 0 | 23 |
| 17 | Weowna Park | 22 | 1 | Treat | 9/14/2021 | 15 | 6 | 100 | 23 |
| 23 | Weowna Park | 22 | 2 | Control | 9/28/2021 | 29 | 9 | 0 | 18 |
| 21 | Weowna Park | 22 | 2 | Treat | 9/28/2021 | 29 | 9 | 100 | 18 |

Table 2: Fish data. Link to Table 1 by rowID (column 1).

| | SMBs | SMBl | LMBs | LMBl | YLPs | YLPl | RKB | BKC | SUNFISH | CCF | CRP | ALL |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 8 | 2 | 0 | 34 | 7 | 160 | 45 | 0 | 0 | 550 | 1 | 0 | 799 |
| 6 | 2 | 1 | 28 | 7 | 55 | 3 | 0 | 0 | 175 | 0 | 0 | 271 |
| 16 | 0 | 0 | 3 | 2 | 18 | 2 | 0 | 0 | 425 | 2 | 0 | 452 |
| 14 | 30 | 0 | 3 | 1 | 180 | 0 | 0 | 0 | 260 | 0 | 1 | 475 |
| 7 | 8 | 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 20 |
| 5 | 0 | 6 | 0 | 0 | 250 | 0 | 0 | 0 | 0 | 0 | 2 | 258 |
| 15 | 0 | 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 7 |
| 13 | 0 | 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 9 |
| 4 | 3 | 2 | 18 | 10 | 60 | 6 | 12 | 5 | 225 | 0 | 0 | 341 |
| 2 | 15 | 1 | 0 | 0 | 8 | 3 | 0 | 0 | 25 | 0 | 0 | 52 |
| 12 | 15 | 0 | 18 | 0 | 1150 | 30 | 2 | 0 | 950 | 0 | 0 | 2165 |
| 10 | 7 | 0 | 0 | 1 | 300 | 5 | 0 | 0 | 130 | 0 | 0 | 443 |
| 3 | 1 | 7 | 0 | 1 | 0 | 8 | 0 | 0 | 12 | 0 | 0 | 29 |
| 1 | 17 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 11 | 0 | 0 | 30 |
| 11 | 18 | 5 | 1 | 0 | 850 | 80 | 0 | 0 | 200 | 0 | 0 | 1154 |
| 9 | 29 | 5 | 0 | 0 | 3 | 3 | 0 | 0 | 6 | 0 | 0 | 46 |
| 20 | 15 | 0 | 18 | 0 | 75 | 4 | 0 | 0 | 1800 | 0 | 0 | 1912 |
| 18 | 6 | 0 | 0 | 0 | 11 | 0 | 0 | 0 | 45 | 0 | 0 | 62 |
| 24 | 7 | 0 | 18 | 9 | 300 | 5 | 0 | 2 | 700 | 0 | 0 | 1041 |
| 22 | 9 | 0 | 1 | 1 | 22 | 0 | 0 | 0 | 75 | 0 | 0 | 108 |
| 19 | 25 | 2 | 4 | 0 | 0 | 0 | 0 | 0 | 5 | 0 | 0 | 36 |
| 17 | 0 | 5 | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 8 |
| 23 | 9 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 11 |
| 21 | 30 | 6 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 38 |

## Scale and center the data

Create scaled and centered data for continuous variables to alleviate model-fit problems (link)

```
DK.s<-DK
DK.s[,names(DK) %in% c("JulDaySurv","VegRemoved","Visibility")]<-
  scale(DK[,names(DK) %in% c("JulDaySurv","VegRemoved","Visibility")])
```

### Correlations Among Variables

Evaluate correlations among continuous variables that may cause multicollinearity problems in the model(s).

```
## The following objects are masked from DK (pos = 4):
```

```
##
##     ALL, BKC, CCF, CRP, Depth, DepthCat, JulDaySurv, JulDayTrt, LMBl,
##     LMBs, RKB, Site, SMBl, SMBs, SUNFISH, SurvCat, SurvDate, SurvNum,
##     Temp, TimeElapsed, TreatDate, Type, VegRemoved, Visibility,
##     Weather, YLPl, YLPs
```

Table 3: Pearson correlations (|r|>0.6) among continuous variables associated with transects.

| | JulDaySurv | SurvNum | JulDayTrt | TimeElapsed | Depth | Visibility | VegRemoved | Temp |
|---|---|---|---|---|---|---|---|---|
| JulDaySurv | 1 | * | 0.964 | * | * | * | * | -0.788 |
| SurvNum | * | 1 | * | 0.984 | * | * | * | * |
| JulDayTrt | 0.964 | * | 1 | * | * | * | * | -0.776 |
| TimeElapsed | * | 0.984 | * | 1 | * | * | * | * |
| Depth | * | * | * | * | 1 | * | * | * |
| Visibility | * | * | * | * | * | 1 | * | * |
| VegRemoved | * | * | * | * | * | * | 1 | * |
| Temp | -0.788 | * | -0.776 | * | * | * | * | 1 |

Screen from global model JulDayTrt and Temp because of high correlation with JulDaySurv. TimeElapsed is correlated with SurveyNum (in binomial sense) and so it is excluded also.

## Explore Random Effect Structure

Explore model fit with different variations of random effects based on design while also considering model complexity given sample size (n=24). Models with no random effect (gm1), site only random effect (gm1.site.reff), site and survey (1 & 2) non-nested random effects (gm2.site.surv.reff), and survey nested within site random effects (gm3.site.surv_reff). It also would fit these models with REML = FALSE, so AIC comparisons may be suspect.

```
library(lme4)
```

```
## Warning: package 'lme4' was built under R version 4.0.5
```

```
## Loading required package: Matrix
```

```
gm1<-glm(ALL~JulDaySurv+Type*DepthCat+Visibility, data=DK.s, family=poisson, na.action="na.fail")
gm1.site.reff<-glmer(ALL~JulDaySurv+Type*DepthCat+Visibility+(1|Site),
                     data=DK.s, family=poisson, na.action="na.fail")
gm2.site.surv.reff<-glmer(ALL~JulDaySurv+Type*DepthCat+Visibility+(1|Site)+(1|SurvCat),
                     data=DK.s, family=poisson, na.action="na.fail")
gm3.site.surv_reff<-glmer(ALL~JulDaySurv+Type*DepthCat+Visibility+(1|Site/SurvCat),
                     data=DK.s, family=poisson, na.action="na.fail")  #Is singular
```

```
## boundary (singular) fit: see ?isSingular
```

Note that (1|Site/SurvCat) in gm3 results in singularity, i.e., each site and survey combination needs its own intercept (n=6).

Compare models using AIC

```
AIC(gm1, gm1.site.reff, gm2.site.surv.reff,gm3.site.surv_reff)#AIC,  suggests random effect needed
```

```
##                     df      AIC
## gm1                  6 5884.469
## gm1.site.reff        7 5275.094
## gm2.site.surv.reff   8 5271.776
## gm3.site.surv_reff   8 3566.273
```

gm3 has lowest AIC but is singular. Fall back to non-nested random effects of site and survey, as that model has the next lowest AIC.

## Global Model

Fit a global model with Julian Day, Visibility, and the two variables of interest Treatment Type (Control, Impact) and Depth Category (11 v. 22 ft). Site and Survey (1 or 2) are maintained as random effects (non-nested).

```
Global.mod<-glmer(ALL~JulDaySurv+Type*DepthCat+Visibility+(1|Site)+(1|SurvCat),
                  data=DK.s,family=poisson,nAGQ=0, na.action="na.fail")
summary(Global.mod)
```

```
## Generalized linear mixed model fit by maximum likelihood (Adaptive
##   Gauss-Hermite Quadrature, nAGQ = 0) [glmerMod]
##  Family: poisson  ( log )
## Formula: ALL ~ JulDaySurv + Type * DepthCat + Visibility + (1 | Site) +
##     (1 | SurvCat)
##    Data: DK.s
##
##      AIC      BIC   logLik deviance df.resid
##   5271.8   5281.2  -2627.9   5255.8       16
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -17.894  -9.571  -5.886   6.123  38.508
##
## Random effects:
##  Groups  Name        Variance Std.Dev.
##  Site    (Intercept) 0.00586  0.07655
##  SurvCat (Intercept) 0.06307  0.25113
## Number of obs: 24, groups:  Site, 3; SurvCat, 2
##
## Fixed effects:
##                             Estimate Std. Error z value Pr(>|z|)
## (Intercept)                  7.17974    0.18349  39.129  < 2e-16 ***
## JulDaySurv                  -0.23324    0.04878  -4.781 1.74e-06 ***
## TypeTreat                   -1.55930    0.02929 -53.242  < 2e-16 ***
## DepthCatTwentyTwo           -2.25628    0.03591 -62.834  < 2e-16 ***
## Visibility                  -0.54697    0.01548 -35.331  < 2e-16 ***
## TypeTreat:DepthCatTwentyTwo  0.38640    0.06499   5.945 2.76e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Correlation of Fixed Effects:
##             (Intr) JlDySr TypTrt DptCTT Vsblty
## JulDaySurv  -0.006
## TypeTreat   -0.028  0.000
## DpthCtTwntT -0.036  0.107  0.142
## Visibility  -0.023  0.166  0.000  0.503
## TypTrt:DCTT  0.012  0.000 -0.451 -0.405  0.000
```

## Model Selection

Run model selection using all combinations of covariates in global model with some constraints. Constraints include Type is in every model (random effects are included in all models by default).

```
library(MuMIn)
mod.sel<-dredge(Global.mod,rank="AICc", REML = FALSE, fixed=c("Type"))
```

Print top models (AICc <4) and model-averaged parameter estimates.

```
(top.mods<-subset(mod.sel,subset = delta<4))
```

```
## Global model call: glmer(formula = ALL ~ JulDaySurv + Type * DepthCat + Visibility +
##     (1 | Site) + (1 | SurvCat), data = DK.s, family = poisson,
##     nAGQ = 0, na.action = "na.fail")
## ---
## Model selection table
##    (Int) DpC     JDS Typ     Vsb DpC:Typ df    logLik   AICc delta weight
## 16 7.180   + -0.2332   + -0.5470       +  8 -2627.888 5281.4  0.00  0.508
## 14 7.178   +           + -0.5437       +  7 -2630.221 5281.4  0.07  0.492
## Models ranked by AICc(x, REML = FALSE)
## Random terms (all models):
## '1 | Site', '1 | SurvCat'
```

Two models are plausible: global model, and a global model but without Julian Day of Survey term.

Compute model-averaged parameter estimates (with and withouth shrinkage):

```
avgm<-model.avg(top.mods)
summary(avgm)
```

```
##
## Call:
## model.avg(object = top.mods)
##
## Component model call:
## glmer(formula = ALL ~ <2 unique rhs>, data = DK.s, family = poisson,
##      nAGQ = 0, na.action = na.fail)
##
## Component models:
##       df   logLik    AICc delta weight
## 12345  8 -2627.89 5281.38  0.00   0.51
## 1345   7 -2630.22 5281.44  0.07   0.49
```

```
##
## Term codes:
##      DepthCat    JulDaySurv          Type    Visibility DepthCat:Type
##             1             2             3             4             5
##
## Model-averaged coefficients:
## (full average)
##                              Estimate Std. Error Adjusted SE z value Pr(>|z|)
## (Intercept)                   7.17876    0.19811     0.21370  33.593   <2e-16
## DepthCatTwentyTwo            -2.25384    0.03601     0.03884  58.022   <2e-16
## JulDaySurv                   -0.11856    0.12168     0.12252   0.968    0.333
## Visibility                   -0.54538    0.01559     0.01681  32.436   <2e-16
## TypeTreat                    -1.55930    0.02929     0.03160  49.340   <2e-16
## DepthCatTwentyTwo:TypeTreat   0.38640    0.06499     0.07013   5.510   <2e-16
##
## (Intercept)                  ***
## DepthCatTwentyTwo            ***
## JulDaySurv
## Visibility                   ***
## TypeTreat                    ***
## DepthCatTwentyTwo:TypeTreat ***
##
## (conditional average)
##                              Estimate Std. Error Adjusted SE z value Pr(>|z|)
## (Intercept)                   7.17876    0.19811     0.21370   33.59  < 2e-16
## DepthCatTwentyTwo            -2.25384    0.03601     0.03884   58.02  < 2e-16
## JulDaySurv                   -0.23324    0.04878     0.05276    4.42 9.85e-06
## Visibility                   -0.54538    0.01559     0.01681   32.44  < 2e-16
## TypeTreat                    -1.55930    0.02929     0.03160   49.34  < 2e-16
## DepthCatTwentyTwo:TypeTreat   0.38640    0.06499     0.07013    5.51 3.60e-08
##
## (Intercept)                  ***
## DepthCatTwentyTwo            ***
## JulDaySurv                   ***
## Visibility                   ***
## TypeTreat                    ***
## DepthCatTwentyTwo:TypeTreat ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Drop Julian Day of Survey because Survey Category random effect captures Julian Day effect, and the parameter estimate is imprecise. Drop Visibility because the parameter estimate is non-sensical (less fish as visibility increases).

## Reduced Model

Fit reduced model based on results of model selection.

```
r.mod<-glmer(ALL~Type*DepthCat+(1|Site)+(1|SurvCat),
             data=DK.s, family=poisson,na.action="na.fail")
summary(r.mod)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
```

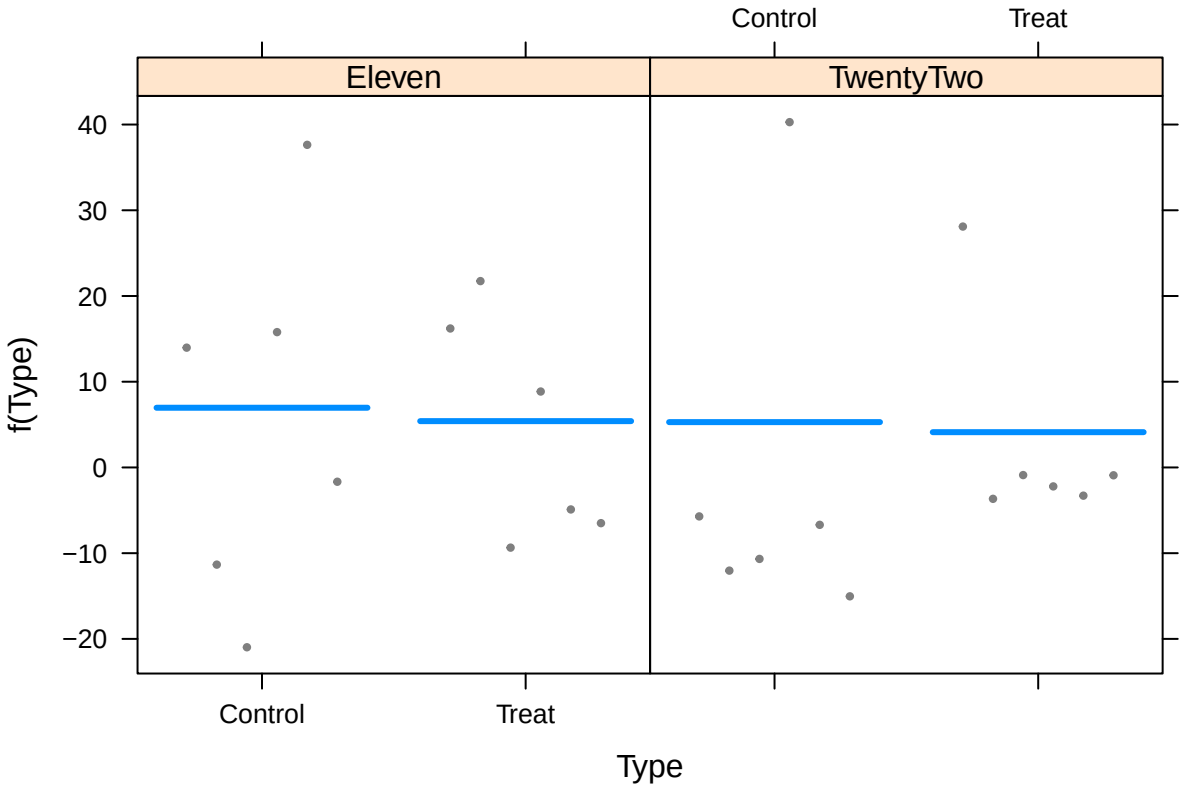
```
##   Approximation) [glmerMod]
##  Family: poisson  ( log )
## Formula: ALL ~ Type * DepthCat + (1 | Site) + (1 | SurvCat)
##    Data: DK.s
##
##      AIC      BIC   logLik deviance df.resid
##   6552.8   6559.9  -3270.4   6540.8       18
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -23.743 -10.835  -6.101   7.780  44.936
##
## Random effects:
##  Groups  Name        Variance Std.Dev.
##  Site    (Intercept) 0.07832  0.2799
##  SurvCat (Intercept) 0.06329  0.2516
## Number of obs: 24, groups:  Site, 3; SurvCat, 2
##
## Fixed effects:
##                             Estimate Std. Error z value Pr(>|z|)
## (Intercept)                  6.96384    0.24065  28.937  < 2e-16 ***
## TypeTreat                   -1.55930    0.02928 -53.247  < 2e-16 ***
## DepthCatTwentyTwo           -1.67487    0.03073 -54.501  < 2e-16 ***
## TypeTreat:DepthCatTwentyTwo  0.38641    0.06499   5.946 2.75e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) TypTrt DptCTT
## TypeTreat   -0.021
## DpthCtTwntT -0.020  0.166
## TypTrt:DCTT  0.010 -0.451 -0.473
```

**Results**

The significant interaction between Treatment Type (Control, Treatment) and Depth (11 v. 22 ft) suggests that the effect of Treatment Type differs depending on whether vegetation was removed at 11 v. 22 ft. At 11 ft., the Depth term and Depth:Treatment Type interaction terms get set to zero. Thus, the parameter estimate of the Treatment Type main effect at 11 ft. is estimated to be -1.56, suggesting that application of the vegetation removal treatment results in a $e^{-1.56} = 0.21$, or a 79% (1 - 0.21 = .79) decrease in abundance of game fish. At 22 ft., the Depth, Treatment Type, and Treatment Type:Depth interaction estimates suggest that vegetation removal results in a $e^{(-1.56+-1.67+0.39)} = 0.06$, or a 94% (1 - 0.06 = 0.94) decrease in abundance.

## Effects Plot with Partial Residuals

```
library(visreg)
visreg(r.mod, "Type", by="DepthCat",type='conditional')
```




This shows the common effect of vegetation removal on abundance of game fish at the two depths.

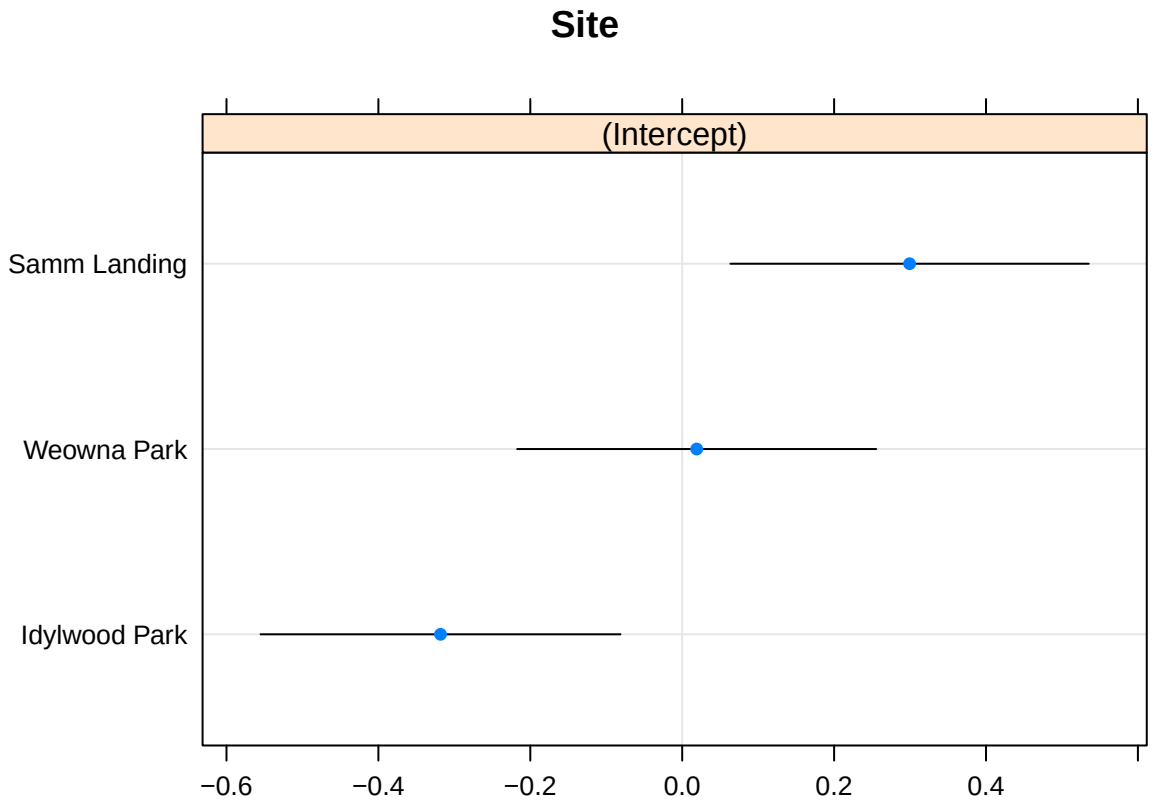
## A Look at the Random Effects

First step, pull random effects from r.mod

```
randoms <- ranef(r.mod)
```

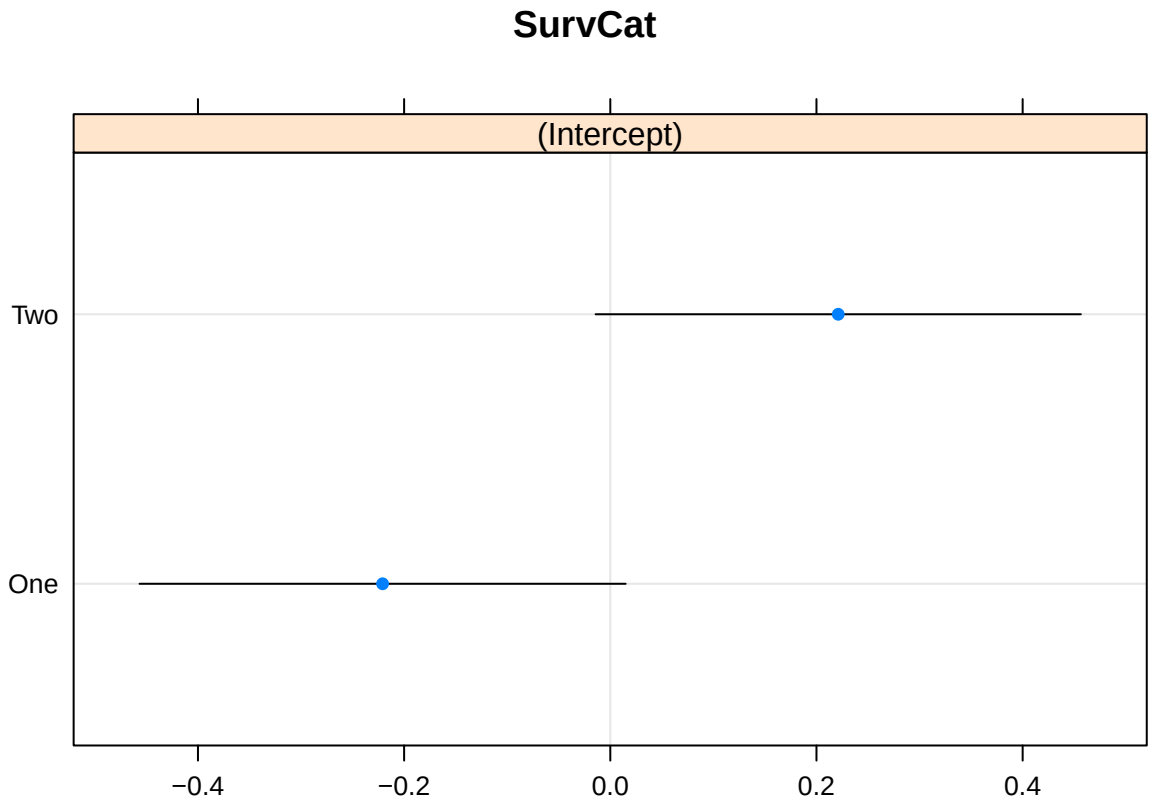
### Catepiller plots of random effects

```
library(lattice)
randoms <- ranef(r.mod)
dotplot(randoms)$Site
```




The random effects show that abundance was generally highest at Samm Landing and lowest at Idylwood Park.

```
dotplot(randoms)$SurvCat
```




The random effect estimates for survey 1 versus 2 show abundance of game fish was generally higher during the 2nd survey.