Analysis of Fish Movement Patterns

Heidi Golden

02 Feb 2016

Contents

1 Prior to Loading Data

The raw data from our PIT antenna readers was "cleaned" using a script written by my colleage Cameron MacKenzie and compiled into a singel data file, which was further cleaned by removing repeat individuals on the same day at each tagging station and removing marker and test tags used in the field to check our equipment.

The file was then loaded into Excel, where I sorted the data by "Year" and "Date" and added the columns "M.lake0.river1" and "R.spr0.fall1" using the following criterion: M.lake0.river1 - M stands for migration. Migration equals 1 if a fish leaves the lake and is detected in the river and equals 0 if a fish is detected in the lake but not in the river.

R.spr0.fall1 - R stands for return migrant, and is conditional upon M. Right now this value does not reflect the conditionality on its own, but equals 0 if the date observed occured in the spring (< July 28th) or fall (>= July 28th) for each year.

I'm sure there's a way to do this in R, but trying to create conditional statements with POSIX* data in R making me crazy, and taking waaay too long.

2 Set working directory and read data into R

```
> setwd("/Users/heidigolden/Dropbox/UConn/Classes/Stats_Lunch_Spring2016/")
> data <- read.csv("TagTracks2010to2013_StatsLunch.csv")
> head(data)
```

```
  Year antenna       tag    date     time                  dt M.lake0.river1
1 2010    Kup7 168455220 6/20/10 21:22:44 (06/20/10 21:22:44)              1
2 2010    Kup7 166251223 6/22/10  0:27:11 (06/22/10 00:27:11)              1
3 2010    Kup7 174476505 6/22/10 22:52:23 (06/22/10 22:52:23)              1
4 2010    Kup7 174476795 6/22/10  7:56:57 (06/22/10 07:56:57)              1
5 2010    Kup6 166251481 6/23/10 22:43:22 (06/23/10 22:43:22)              1
6 2010    Kup7 166251482 6/23/10 22:13:04 (06/23/10 22:13:04)              1
  R.spr0.fall1
1            0
2            0
3            0
4            0
5            0
6            0
```

3 Load Indexed Tag Numbers

Load the indexed PIT tag numbers for fish that returned "early" to the lake and fish that became "trapped" in 2011.

```
> early.trap <- read.csv("Kup_early_trapped2011.csv")
> head(early.trap)

    ref_tag early0.trapped1
1 157362761               0
2 157362762               0
3 174439540               0
4 174440600               0
5 174440601               0
6 174440620               0

> str(early.trap)

'data.frame':        880 obs. of  2 variables:
 $ ref_tag        : Factor w/ 879 levels "157361321","157361327",..: 26 27 104 363 364 365 3
 $ early0.trapped1: int  0 0 0 0 0 0 0 0 0 0 ...
```

4 Create New Column Indexing "early/trapped" Fish

Use vector indexing by subsetting combined with "match" to add the index values "early" or "trapped" to the corresponding tag numbers:

```
> index <- as.character(early.trap$ref_tag)
> values <- early.trap$early0.trapped1
> data$tag <- as.character(data$tag)
> data$early_trap <- values[match(data$tag, index)]
> head(data)

  Year antenna       tag    date     time                   dt M.lake0.river1
1 2010    Kup7 168455220 6/20/10 21:22:44 (06/20/10 21:22:44)               1
2 2010    Kup7 166251223 6/22/10  0:27:11 (06/22/10 00:27:11)               1
3 2010    Kup7 174476505 6/22/10 22:52:23 (06/22/10 22:52:23)               1
4 2010    Kup7 174476795 6/22/10  7:56:57 (06/22/10 07:56:57)               1
5 2010    Kup6 166251481 6/23/10 22:43:22 (06/23/10 22:43:22)               1
6 2010    Kup7 166251482 6/23/10 22:13:04 (06/23/10 22:13:04)               1
  R.spr0.fall1 early_trap
1            0         NA
2            0         NA
3            0         NA
4            0         NA
5            0         NA
6            0         NA

> str(data)
```

```
'data.frame':        16632 obs. of  9 variables:
 $ Year          : int  2010 2010 2010 2010 2010 2010 2010 2010 2010 2010 ...
 $ antenna       : Factor w/ 15 levels "GCL","GCL_inlet",..: 13 13 13 13 12 13 12 13 12 12 .
 $ tag           : chr  "168455220" "166251223" "174476505" "174476795" ...
 $ date          : Factor w/ 462 levels "10/1/13","10/10/13",..: 97 104 104 104 108 108 108
 $ time          : Factor w/ 14588 levels "0:00:00","0:00:02",..: 8099 223 9258 14088 9131 8
 $ dt            : Factor w/ 15709 levels "(05/10/13 23:59:00)",..: 3301 3492 3494 3493 3599
 $ M.lake0.river1: int  1 1 1 1 1 1 1 1 1 1 ...
 $ R.spr0.fall1  : int  0 0 0 0 0 0 0 0 0 0 ...
 $ early_trap    : int  NA NA NA NA NA NA NA NA NA NA ...
```

```
> data$early_trap[1:200]
```

```
  [1] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
 [26] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
 [51] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
 [76] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
[101] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA  0
[126] NA NA NA NA  0 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
[151] NA NA NA NA NA NA  0 NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
[176] NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA
```

```
> write.csv(data, "TagTracks2010to2013_StatsLunch2.csv")
```

5 Create New Data Frame with ”Experiment” Fish only

```
> expmt <- data[!(is.na(data$early_trap)), ]
> head(expmt)
```

```
    Year antenna       tag    date     time                   dt M.lake0.river1
125 2010    Kup6 174476635  7/2/10 12:10:29 (07/02/10 12:10:29)              1
130 2010    Kup6 174475855  7/3/10  4:50:15 (07/03/10 04:50:15)              1
157 2010    Kup6 174476540  7/7/10 10:17:14 (07/07/10 10:17:14)              1
224 2010    Kup4 174476685 7/10/10 13:36:19 (07/10/10 13:36:19)              1
325 2010    Kup3 174476800 7/18/10  6:02:53 (07/18/10 06:02:53)              1
333 2010    Kup4 174476323 7/19/10  5:14:17 (07/19/10 05:14:17)              1
    R.spr0.fall1 early_trap
125            0          0
130            0          0
157            0          0
224            0          0
325            0          0
333            0          0
```

```
> write.csv(expmt, "ExpmtFish2010to2013_StatsLunch2.csv")
```

6 Change Date to POSIX*

```
> # Just in case we'll need the dates later:
> data$date <- strptime(data$date, "%Y-%m-%d")
> data$time <- strptime(data$time, "%H:%M:%S")
> str(data)

'data.frame':        16632 obs. of  9 variables:
 $ Year          : int  2010 2010 2010 2010 2010 2010 2010 2010 2010 2010 ...
 $ antenna       : Factor w/ 15 levels "GCL","GCL_inlet",..: 13 13 13 13 12 13 12 13 12 12 .
 $ tag           : chr  "168455220" "166251223" "174476505" "174476795" ...
 $ date          : POSIXlt, format: NA NA ...
 $ time          : POSIXlt, format: "2016-02-04 21:22:44" "2016-02-04 00:27:11" ...
 $ dt            : Factor w/ 15709 levels "(05/10/13 23:59:00)",..: 3301 3492 3494 3493 3599
 $ M.lake0.river1: int  1 1 1 1 1 1 1 1 1 1 ...
 $ R.spr0.fall1  : int  0 0 0 0 0 0 0 0 0 0 ...
 $ early_trap    : int  NA NA NA NA NA NA NA NA NA NA ...
```

7 Some Simple Queries to Check the Data

How many trapped and non-trapped fish were seen in the river: a. Before the drought (2010) b. During the drought year (2011) c. One year after the drought (2012) d. Two years after the drought (2013)

```
> # a. Trapped fish before the drought
> tr2010 <- subset(expmt, expmt$Year == 2010 & expmt$early_trap == 1)
> str(tr2010) # 0 fish

'data.frame':        0 obs. of  9 variables:
 $ Year          : int
 $ antenna       : Factor w/ 15 levels "GCL","GCL_inlet",..:
 $ tag           : chr
 $ date          : Factor w/ 462 levels "10/1/13","10/10/13",..:
 $ time          : Factor w/ 14588 levels "0:00:00","0:00:02",..:
 $ dt            : Factor w/ 15709 levels "(05/10/13 23:59:00)",..:
 $ M.lake0.river1: int
 $ R.spr0.fall1  : int
 $ early_trap    : int

> # b. Trapped fish during the drought
> tr2011 <- subset(expmt, expmt$Year == 2011 & expmt$early_trap == 1)
> str(tr2011) # 3 fish - this makes no sense!!

'data.frame':        3 obs. of  9 variables:
 $ Year          : int  2011 2011 2011
 $ antenna       : Factor w/ 15 levels "GCL","GCL_inlet",..: 6 6 6
```

```
 $ tag           : chr  "177350987" "174440084" "174440084"
 $ date          : Factor w/ 462 levels "10/1/13","10/10/13",..: 397 400 403
 $ time          : Factor w/ 14588 levels "0:00:00","0:00:02",..: 5515 10027 14208
 $ dt            : Factor w/ 15709 levels "(05/10/13 23:59:00)",..: 13896 14076 14116
 $ M.lake0.river1: int  0 0 0
 $ R.spr0.fall1  : int  1 1 1
 $ early_trap    : int  1 1 1

> # c. Trapped fish year following the drought
> tr2012 <- subset(expmt, expmt$Year == 2012 & expmt$early_trap == 1)
> str(tr2012) # 92 fish

'data.frame':        92 obs. of  9 variables:
 $ Year          : int  2012 2012 2012 2012 2012 2012 2012 2012 2012 2012 ...
 $ antenna       : Factor w/ 15 levels "GCL","GCL_inlet",..: 7 7 7 7 7 7 7 7 7 6 ...
 $ tag           : chr  "174440002" "174440022" "174440028" "174440071" ...
 $ date          : Factor w/ 462 levels "10/1/13","10/10/13",..: 48 48 48 48 48 48 51 51 51
 $ time          : Factor w/ 14588 levels "0:00:00","0:00:02",..: 10048 3739 10881 5284 9658
 $ dt            : Factor w/ 15709 levels "(05/10/13 23:59:00)",..: 119 76 25 92 112 24 162
 $ M.lake0.river1: int  0 0 0 0 0 0 0 0 0 0 ...
 $ R.spr0.fall1  : int  0 0 0 0 0 0 0 0 0 0 ...
 $ early_trap    : int  1 1 1 1 1 1 1 1 1 1 ...

> # d. Trapped fish two years after the drought
> tr2013 <- subset(expmt, expmt$Year == 2013 & expmt$early_trap == 1)
> str(tr2013) # 0 fish

'data.frame':        0 obs. of  9 variables:
 $ Year          : int
 $ antenna       : Factor w/ 15 levels "GCL","GCL_inlet",..:
 $ tag           : chr
 $ date          : Factor w/ 462 levels "10/1/13","10/10/13",..:
 $ time          : Factor w/ 14588 levels "0:00:00","0:00:02",..:
 $ dt            : Factor w/ 15709 levels "(05/10/13 23:59:00)",..:
 $ M.lake0.river1: int
 $ R.spr0.fall1  : int
 $ early_trap    : int

> # a. Non-trapped fish before the drought
> ntr2010 <- subset(expmt, expmt$Year == 2010 & expmt$early_trap == 0)
> str(ntr2010) # 74 fish

'data.frame':        74 obs. of  9 variables:
 $ Year          : int  2010 2010 2010 2010 2010 2010 2010 2010 2010 2010 ...
 $ antenna       : Factor w/ 15 levels "GCL","GCL_inlet",..: 12 12 12 9 8 9 8 9 9 9 ...
 $ tag           : chr  "174476635" "174475855" "174476540" "174476685" ...
 $ date          : Factor w/ 462 levels "10/1/13","10/10/13",..: 198 242 263 162 190 194 194
```

```
 $ time          : Factor w/ 14588 levels "0:00:00","0:00:02",..: 2368 11823 1518 3124 13613
 $ dt            : Factor w/ 15709 levels "(05/10/13 23:59:00)",..: 4325 4414 4797 5053 5560
 $ M.lake0.river1: int  1 1 1 1 1 1 1 1 1 1 ...
 $ R.spr0.fall1  : int  0 0 0 0 0 0 0 0 0 0 ...
 $ early_trap    : int  0 0 0 0 0 0 0 0 0 0 ...

> # b. Non-trapped fish during the drought
> ntr2011 <- subset(expmt, expmt$Year == 2011 & expmt$early_trap == 0)
> str(ntr2011) # 512 fish

'data.frame':        512 obs. of  9 variables:
 $ Year          : int  2011 2011 2011 2011 2011 2011 2011 2011 2011 2011 ...
 $ antenna       : Factor w/ 15 levels "GCL","GCL_inlet",..: 4 4 4 4 4 4 4 4 9 4 ...
 $ tag           : chr  "174475948" "174476068" "166251202" "174439426" ...
 $ date          : Factor w/ 462 levels "10/1/13","10/10/13",..: 50 50 53 53 56 59 59 59 68
 $ time          : Factor w/ 14588 levels "0:00:00","0:00:02",..: 10241 10241 10241 10241 10
 $ dt            : Factor w/ 15709 levels "(05/10/13 23:59:00)",..: 123 123 169 169 194 208
 $ M.lake0.river1: int  0 0 0 0 0 0 0 0 1 0 ...
 $ R.spr0.fall1  : int  0 0 0 0 0 0 0 0 0 0 ...
 $ early_trap    : int  0 0 0 0 0 0 0 0 0 0 ...

> # c. Non-trapped fish year following the drought
> ntr2012 <- subset(expmt, expmt$Year == 2012 & expmt$early_trap == 0)
> str(ntr2012) # 203 fish

'data.frame':        203 obs. of  9 variables:
 $ Year          : int  2012 2012 2012 2012 2012 2012 2012 2012 2012 2012 ...
 $ antenna       : Factor w/ 15 levels "GCL","GCL_inlet",..: 7 7 7 7 7 7 7 7 9 ...
 $ tag           : chr  "157362871" "174441113" "174441137" "174441190" ...
 $ date          : Factor w/ 462 levels "10/1/13","10/10/13",..: 48 48 48 48 48 48 48 48 54
 $ time          : Factor w/ 14588 levels "0:00:00","0:00:02",..: 4048 14097 2106 6043 13944
 $ dt            : Factor w/ 15709 levels "(05/10/13 23:59:00)",..: 85 46 72 106 34 41 73 84
 $ M.lake0.river1: int  0 0 0 0 0 0 0 0 0 1 ...
 $ R.spr0.fall1  : int  0 0 0 0 0 0 0 0 0 0 ...
 $ early_trap    : int  0 0 0 0 0 0 0 0 0 0 ...

> # d. Non-trapped fish two years after the drought
> ntr2013 <- subset(expmt, expmt$Year == 2013 & expmt$early_trap == 0)
> str(ntr2013) # 109 fish

'data.frame':        109 obs. of  9 variables:
 $ Year          : int  2013 2013 2013 2013 2013 2013 2013 2013 2013 2013 ...
 $ antenna       : Factor w/ 15 levels "GCL","GCL_inlet",..: 7 7 7 7 7 7 7 7 7 7 ...
 $ tag           : chr  "157361611" "157362853" "166251202" "174440838" ...
 $ date          : Factor w/ 462 levels "10/1/13","10/10/13",..: 137 137 137 137 137 137 137
 $ time          : Factor w/ 14588 levels "0:00:00","0:00:02",..: 10514 12329 992 999 1542 1
 $ dt            : Factor w/ 15709 levels "(05/10/13 23:59:00)",..: 1542 1386 1538 1438 1400
```

```
$ M.lake0.river1: int  0 0 0 0 0 0 0 0 0 0 ...
$ R.spr0.fall1  : int  0 0 0 0 0 0 0 0 0 0 ...
$ early_trap    : int  0 0 0 0 0 0 0 0 0 0 ...
```

8 Conclusions Regarding Data

There should be many more tagged fish in 2011 that were "trapped" than show up here in the dataset. Upon further examination of the files in Excel, I see the KUS antenna is missing from 2011. This error needs reconsiling through consultation with my collaborator. I'll work on this tomorrow. HG