preliminary\_analysis\_Odonata

Diana Bowler

24 januar 2018

# Analysis of the Odonata data from Saarland

First step: read in the raw data file from the sMon data portal. Sheet 5 of the workbook contains the raw data of occurrences. Each row is the data frame is a species observation.

library(gdata)  
df <- read.xls("raw-data/Odonata\_Saarland\_Trockur.xls",sheet=5)

# Data frame formatting

Step one: format the date and extract month, dat and year information.

library(lubridate)  
df$Date <- as.Date(df$Date,file="%Y-%m-%d")  
df$month <- month(df$Date)  
df$day <- yday(df$Date)  
df$year <- year(df$Date)

Step two: remove those with missing date entries

df <- subset(df,!is.na(month))

# Data filtering

Filter one: focus on the months that were most sampled. The results suggest this is between April and September.

table(df$month)

##   
## 1 2 3 4 5 6 7 8 9 10 11 12   
## 219 1 15 536 5461 4852 3544 2698 1008 256 45 3

df <- subset(df, month %in% c(4:9))

Filter two: begin the time series from the first year with reasonable amount of data (i.e., number of records). 1981 seems like a good place to start. Also, we cut off 2017 incase the database is not up-to-date yet.

table(df$Year)

##   
## 1970 1971 1976 1977 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990   
## 12 37 2 1 1 174 187 362 141 452 147 302 397 219 380   
## 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005   
## 511 580 587 728 230 496 862 221 412 189 201 113 183 277 305   
## 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017   
## 557 183 362 994 1728 1957 1240 526 1016 310 229 288

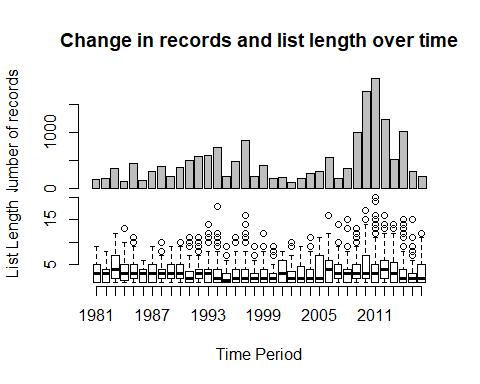
df <- subset(df, Year>1980)  
df <- subset(df, Year<2017)

# Data availability/Sampling effort checking

We can look at effort by: total number of records and average number of species seen per sampling day (“list length”)

library(sparta)  
  
dataDiagnostics(taxa= as.character(df$Species),   
 site= as.character(df$MTB\_Q),   
 time\_period = df$Date)

## Calculating diagnostics  
##   
 |   
 | | 0%  
 |   
 |============= | 20%  
 |   
 |==================== | 30%  
 |   
 |================================ | 50%  
 |   
 |========================================================== | 90%  
 |   
 |=================================================================| 100%

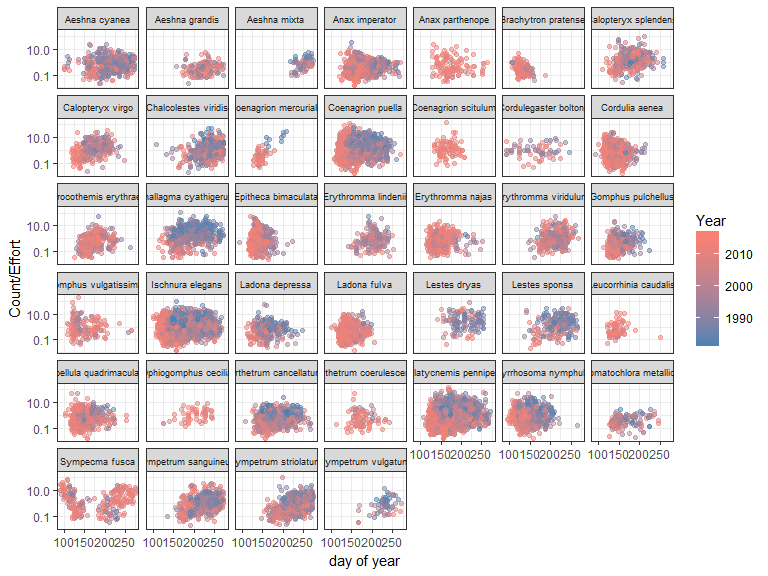


## ## Linear model outputs ##  
##   
## There is a significant change in the number of records over time:  
##   
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) -33800.15916 12633.289840 -2.675484 0.01139379  
## time\_period 17.15959 6.321301 2.714566 0.01034811  
##   
##   
## There is a significant change in list lengths over time:  
##   
## Estimate Std. Error z value Pr(>|z|)  
## (Intercept) 1.036485e+00 2.944291e-02 35.20320 1.786515e-271  
## time\_period 1.883019e-05 2.368396e-06 7.95061 1.855953e-15

# Study Phenology

We need to check the pattern of activity over the year to see how we should model phenology.

#using abundance data corrected by effort  
  
#first get number of species seen per sampling day as a measure of sampling effort:  
library(plyr)  
samplingEffort <- ddply(df, .(MTB\_Q,Date), summarise, LL=length(Anzahl\_min))  
#add it to the data frame  
df$samplingEffort <- samplingEffort$LL[match(interaction(df$Date,df$MTB\_Q),  
 interaction(samplingEffort$Date,samplingEffort$MTB\_Q))]  
#plots counts/sampling effort over time  
df$yday <- yday(df$Date)  
library(ggplot2)  
ggplot(subset(df,Species %in% names(table(df$Species))[table(df$Species)>50]))+  
 geom\_point(aes(x=yday, y=Anzahl\_min/samplingEffort,colour=Year),alpha=0.5)+  
 facet\_wrap(~Species)+  
 scale\_y\_log10()+  
 theme\_bw()+  
 scale\_colour\_gradient(low="steelblue",high="salmon")+  
 ylab("Count/Effort")+  
 xlab("day of year")+  
 theme(strip.text = element\_text(size=rel(0.6)))



# Model fitting

Use the sparta package (<https://github.com/BiologicalRecordsCentre/sparta>) to fit an occupancy detection model. (Note: the data file does also count information on counts, but for the moment we only use information on species presence)

library(sparta)  
  
# settings  
sparta\_options <- c('ranwalk', # prior on occupancy is set by last year's posterior  
 'jul\_date', # use the Julian date as a covariate on the detection probability  
 'catlistlength', # categorises the visits into three sets of 'qualities'  
 'halfcauchy') # prior on the precisions  
  
#odonata\_occmod <- occDetModel(taxa= as.character(df$Species),   
# site= as.character(df$MTB\_Q),   
# time\_period = df$Date,  
# modeltype = sparta\_options,  
# n\_iterations = 10000)

# Plotting the predictions

Plot the predicting occupancy rates (% of sites occupied per year) for each species.

Step one: Get the list of model files that sparta spat out from the previous code chunk. These are contained within the model-outputs folder.

source("R/sparta\_wrapper\_functions.R")  
  
models<-getSpartaModels("model-outputs")

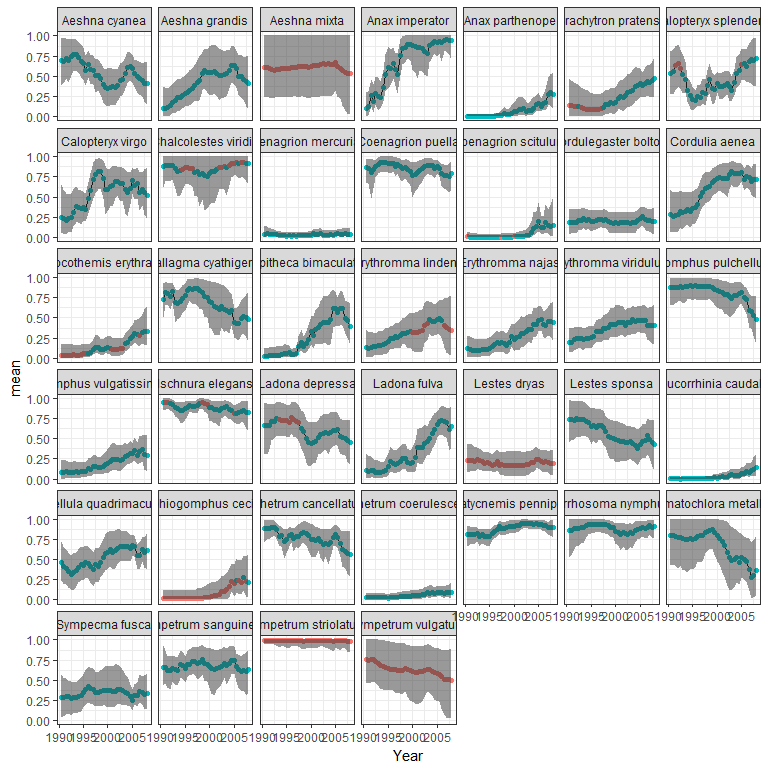
Step two: Pull out the annual predictions of model occupancy per species

#in sparta wrapper functions  
  
myAnnualPredictions <- annualPredictions(models)

# Plotting the time series

Plot the predicted time series for species with more than 50 records. Pink means that this model did not converge (i.e., Rhat was not less than 1.1). We can try running the model for more iterations to see if it will converge.

#in sparta wrapper functions  
plotPredictions(myAnnualPredictions,df)



# Quantify the linear long-term trends of each species

#in sparta wrapper functions  
library(knitr)  
kable(calculateTrends(models),digits=3)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | mean.trend | CI.lower | CI.upper | nuRecords |
| Aeshna affinis | 2.652924e+04 | -0.239 | 2358.382 | 5 |
| Aeshna cyanea | -4.130000e-01 | -0.678 | -0.071 | 411 |
| Aeshna grandis | 2.473000e+00 | 0.364 | 5.662 | 112 |
| Aeshna isoceles | 3.827217e+03 | -0.640 | 7925.718 | 18 |
| Aeshna juncea | 4.332400e+01 | 2.695 | 267.980 | 39 |
| Aeshna mixta | 1.020000e-01 | -0.648 | 1.107 | 59 |
| Anax imperator | 4.614000e+00 | 1.959 | 10.678 | 623 |
| Anax parthenope | 2.696580e+02 | 15.750 | 1356.920 | 86 |
| Brachytron pratense | 1.406000e+01 | 0.187 | 63.026 | 101 |
| Calopteryx splendens | 5.700000e-01 | -0.047 | 1.477 | 270 |
| Calopteryx virgo | 1.206000e+00 | 0.236 | 2.492 | 238 |
| Chalcolestes viridis | 7.900000e-02 | -0.186 | 0.434 | 320 |
| Coenagrion mercuriale | 9.430000e-01 | -0.722 | 6.380 | 33 |
| Coenagrion puella | -9.500000e-02 | -0.265 | 0.077 | 1069 |
| Coenagrion pulchellum | 1.560000e-01 | -0.902 | 3.266 | 33 |
| Coenagrion scitulum | 1.894692e+04 | 7.783 | 21688.735 | 70 |
| Cordulegaster bidentata | -1.380000e-01 | -0.936 | 2.179 | 13 |
| Cordulegaster boltonii | 9.100000e-02 | -0.511 | 0.984 | 50 |
| Cordulia aenea | 2.336000e+00 | 0.742 | 5.069 | 670 |
| Crocothemis erythraea | 1.561100e+01 | 1.151 | 42.771 | 236 |
| Enallagma cyathigerum | -4.250000e-01 | -0.631 | -0.182 | 514 |
| Epitheca bimaculata | 2.585200e+01 | 8.018 | 60.102 | 467 |
| Erythromma lindenii | 2.859000e+00 | 0.384 | 9.217 | 105 |
| Erythromma najas | 6.227000e+00 | 1.573 | 16.770 | 284 |
| Erythromma viridulum | 1.290000e+00 | 0.263 | 3.052 | 253 |
| Gomphus pulchellus | -3.160000e-01 | -0.534 | -0.097 | 348 |
| Gomphus vulgatissimus | 5.511000e+00 | 0.842 | 15.801 | 122 |
| Ischnura elegans | -1.060000e-01 | -0.237 | 0.015 | 1238 |
| Ischnura pumilio | 4.800000e+00 | 0.138 | 20.363 | 44 |
| Ladona depressa | -3.210000e-01 | -0.649 | 0.044 | 237 |
| Ladona fulva | 1.421900e+01 | 3.927 | 38.148 | 387 |
| Lestes barbarus | 1.840500e+01 | -0.396 | 131.037 | 9 |
| Lestes dryas | -1.600000e-02 | -0.643 | 0.838 | 92 |
| Lestes sponsa | -4.690000e-01 | -0.777 | -0.025 | 196 |
| Lestes virens | 2.958000e+00 | -0.703 | 17.463 | 14 |
| Leucorrhinia caudalis | 9.118640e+02 | 13.030 | 4942.324 | 60 |
| Leucorrhinia pectoralis | 8.698649e+03 | -0.884 | 43.290 | 2 |
| Libellula quadrimaculata | 9.950000e-01 | 0.078 | 2.291 | 464 |
| Onychogomphus forcipatus | 2.785900e+01 | -0.321 | 108.179 | 33 |
| Ophiogomphus cecilia | 1.476344e+03 | 2.475 | 8621.153 | 39 |
| Orthetrum brunneum | 2.773000e+00 | -0.272 | 13.336 | 43 |
| Orthetrum cancellatum | -2.680000e-01 | -0.457 | -0.027 | 663 |
| Orthetrum coerulescens | 5.950000e+00 | 0.174 | 26.551 | 65 |
| Platycnemis pennipes | 2.080000e-01 | 0.086 | 0.378 | 1401 |
| Pyrrhosoma nymphula | -3.100000e-02 | -0.243 | 0.250 | 591 |
| Somatochlora flavomaculata | 5.288396e+18 | -0.511 | 71259169.035 | 4 |
| Somatochlora metallica | -5.260000e-01 | -0.848 | -0.084 | 119 |
| Sympecma fusca | 2.030000e-01 | -0.367 | 1.794 | 191 |
| Sympetrum danae | -2.710000e-01 | -0.799 | 0.401 | 38 |
| Sympetrum flaveolum | -5.760000e-01 | -0.975 | 0.523 | 13 |
| Sympetrum fonscolombii | 4.038330e+02 | -0.090 | 1340.405 | 37 |
| Sympetrum sanguineum | 1.000000e-02 | -0.304 | 0.370 | 378 |
| Sympetrum striolatum | -9.000000e-03 | -0.116 | 0.032 | 422 |
| Sympetrum vulgatum | -2.410000e-01 | -0.844 | 0.175 | 57 |