Analyses of species distributions in peatlands

Alicia Valdés

14 April, 2023

Contents

Load the R packages that you will use

If you do not have the R packages installed, you need to install them. These are the ones that I often use for reading in data, data manipulation, etc. But you can of course use others!

```
library(tidyverse)
library(readxl)
library(knitr)
library(ggeffects)
library(car)
library(glmmTMB)
```

Data preparation

Read data from Excel file

This reads the data from the sheet "SDM Data" in the Excel file "Modelling_SDM_species_data_AV.xlsx". Note that you need to change the path to the folder where you have the Excel file

```
data_peat<-read_excel("data/edited/Modelling_SDM_species_data.xlsx",
                      sheet="SDM Data")
```

Have a look at the data

This shows the first rows of your data file in "tibble" format. You can also see the variable type for each variable (double or character).

```
data_peat
```

```
## # A tibble: 115 x 29
##    n_samples depth depth_corrected fen   tot_Sphagnum  Erio Carex Erica
##        <dbl> <chr>           <dbl> <chr>        <dbl> <dbl> <dbl> <dbl>
##  1         1 0                   0 N              100     0     0     0
##  2         2 14-15              14 N               87     3     2     8
##  3         3 20-21              20 N               74     6     6    14
##  4         4 30-31              30 N               87     5     4     4
##  5         5 40-41              40 N               90     2     3     5
##  6         6 50-51              50 N               41    46     9     4
##  7         7 55-56              55 N               74    19     6     1
##  8         8 60-61              60 N               94     4     1     1
##  9         9 65-66              65 N               95     3     1     1
## 10        10 70-71              70 N               90     3     4     3
##    other_veg Balticum Medium Cuspidata Austinii Fuscum Rubellum Acutifolia
##        <dbl>    <dbl>  <dbl>     <dbl>    <dbl>  <dbl>    <dbl>      <dbl>
##  1         0       32     68         0        0      0        0          0
##  2         0       63     22         2        0      0        0          0
##  3         0       24     48         0        0      0        0          2
##  4         0       23     24         0        0      0       32          8
##  5         0       17      5         3        0      0        0         65
##  6         0        7     17         9        0      0        8          0
##  7         0       42     16        16        0      0        0          0
##  8         0       46     36        12        0      0        0          0
##  9         0       56     25        14        0      0        0          0
## 10         0       49     10        31        0      0        0          0
##    `Diseased Acutifolia` Angustifolium Tenellum Papillosum Fallax Stems   age
```

```
##                     <dbl>         <dbl>    <dbl>      <dbl>  <dbl> <dbl> <dbl>
##  1                      0             0        0          0      0     0     0
##  2                      0             0        0          0      0     0    75
##  3                      0             0        0          0      0     0   136
##  4                      0             0        0          0      0     0   244
##  5                      0             0        0          0      0     0   352
##  6                      0             0        0          0      0     0   452
##  7                      0             0        0          0      0     0   505
##  8                      0             0        0          0      0     0   555
##  9                      0             0        0          0      0     0   606
## 10                      0             0        0          0      0     0   643
##     temp imp_temp moist nutrient  fire   dry
##    <dbl>    <dbl> <chr>    <dbl> <dbl> <dbl>
##  1  6.35        0 NA           0     0     0
##  2  6.92        1 NA           0     0     0
##  3  7.39        1 NA           0     0     0
##  4  8.21        0 NA           0     0     0
##  5  8.00        1 -0.38        1     0     0
##  6  7.81        1 -0.3         1     0     0
##  7  7.71        1 0.01         1     0     0
##  8  7.62        1 -0.26        1     0     0
##  9  7.52        1 0.22         1     0     0
## 10  7.45        1 -0.25        1     0     0
## # i 105 more rows
```

You can see that I renamed the variables with easier (shorter) names, and also without spaces in the variable names (otherwise you can run into problems sometimes).

Convert some variables to factors

It is better to convert some variables (those that are Y/N or 0/1) to factors.

```
data_peat<-data_peat%>%
  mutate(fen=as.factor(fen),imp_temp=as.factor(imp_temp),
         nutrient=as.factor(nutrient),fire=as.factor(fire),dry=as.factor(dry))
# with mutate you create new variables that are equal to the old variables
# but are coded as factors
```

Convert moist to numeric

For some reason, moist appears as a character variable. It should be numeric, so we convert it.

```
data_peat<-data_peat%>%
  mutate(moist=as.numeric(moist))
```

Create presence variables

So far you have abundance data (% cover), but I think you also wanted to look at presences/absences. So we can create new columns with presence/absence data. So for each column, we create a column that starts with "pres_". This column will be 0 if abundance of the species is 0, and 1 otherwise

```
data_peat<-data_peat%>%
  mutate(pres_tot_Sphagnum=ifelse(tot_Sphagnum>0,1,0),
         # This creates a new column called "pres_tot_Spagnum"
         # if tot_Sphagnum is larger than 0,
         # this column will be equal to 1 (presence)
         # otherwise, this column will be equal to 0 (absence)
         pres_Erio=ifelse(Erio>0,1,0),   # Same operation for Erio
         pres_Carex=ifelse(Carex>0,1,0)) # Same for Carex...
# You can do the same for other species
```

Now you can see that the new columns have been created.

```
data_peat
```

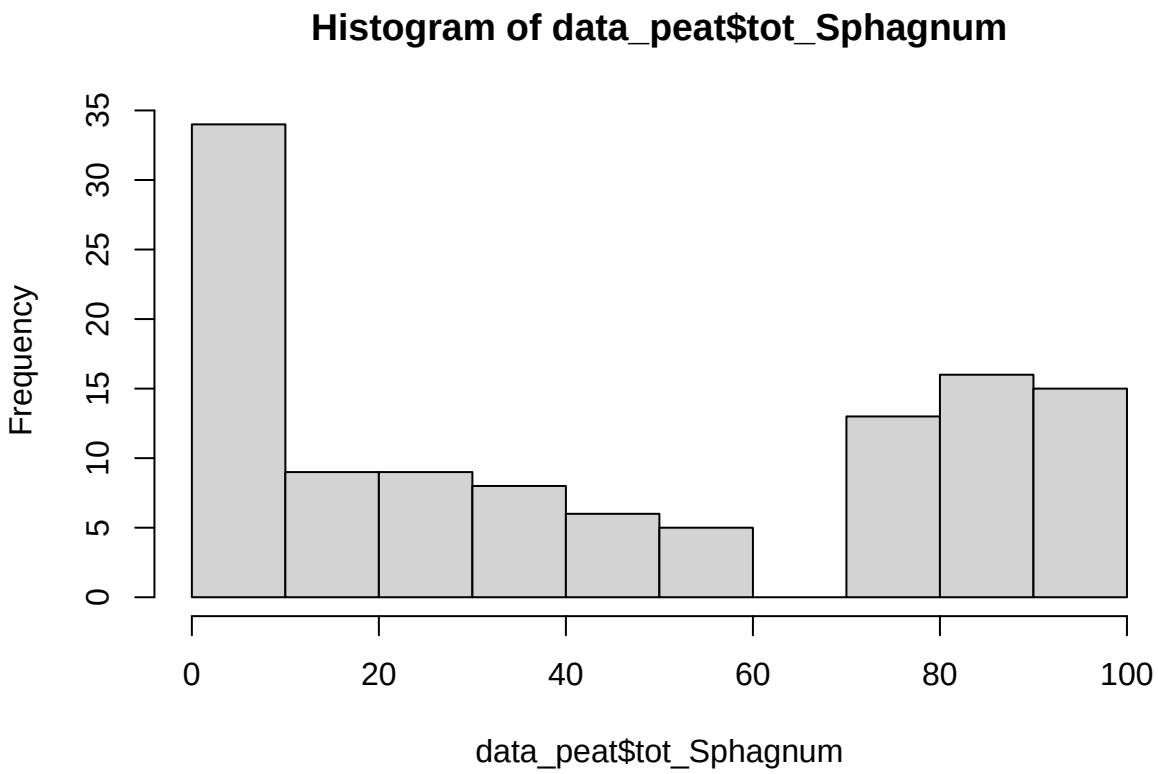
```
## # A tibble: 115 x 32
##    n_samples depth depth_corrected fen   tot_Sphagnum  Erio Carex Erica
##        <dbl> <chr>           <dbl> <fct>        <dbl> <dbl> <dbl> <dbl>
##  1         1 0                   0 N              100     0     0     0
##  2         2 14-15              14 N               87     3     2     8
##  3         3 20-21              20 N               74     6     6    14
##  4         4 30-31              30 N               87     5     4     4
##  5         5 40-41              40 N               90     2     3     5
##  6         6 50-51              50 N               41    46     9     4
##  7         7 55-56              55 N               74    19     6     1
##  8         8 60-61              60 N               94     4     1     1
##  9         9 65-66              65 N               95     3     1     1
## 10        10 70-71              70 N               90     3     4     3
##    other_veg Balticum Medium Cuspidata Austinii Fuscum Rubellum Acutifolia
##        <dbl>    <dbl>  <dbl>     <dbl>    <dbl>  <dbl>    <dbl>      <dbl>
##  1         0       32     68         0        0      0        0          0
##  2         0       63     22         2        0      0        0          0
##  3         0       24     48         0        0      0        0          2
##  4         0       23     24         0        0      0       32          8
##  5         0       17      5         3        0      0        0         65
##  6         0        7     17         9        0      0        8          0
##  7         0       42     16        16        0      0        0          0
##  8         0       46     36        12        0      0        0          0
##  9         0       56     25        14        0      0        0          0
## 10         0       49     10        31        0      0        0          0
##    'Diseased Acutifolia' Angustifolium Tenellum Papillosum Fallax Stems   age
##                    <dbl>         <dbl>    <dbl>      <dbl>  <dbl> <dbl> <dbl>
##  1                     0             0        0          0      0     0     0
##  2                     0             0        0          0      0     0    75
##  3                     0             0        0          0      0     0   136
##  4                     0             0        0          0      0     0   244
##  5                     0             0        0          0      0     0   352
##  6                     0             0        0          0      0     0   452
##  7                     0             0        0          0      0     0   505
##  8                     0             0        0          0      0     0   555
##  9                     0             0        0          0      0     0   606
## 10                     0             0        0          0      0     0   643
##     temp imp_temp moist nutrient fire  dry   pres_tot_Sphagnum pres_Erio
##    <dbl> <fct>    <dbl> <fct>    <fct> <fct>             <dbl>     <dbl>
```

```
##  1  6.35 0        NA    0        0     0                    1         0
##  2  6.92 1        NA    0        0     0                    1         1
##  3  7.39 1        NA    0        0     0                    1         1
##  4  8.21 0        NA    0        0     0                    1         1
##  5  8.00 1        -0.38 1        0     0                    1         1
##  6  7.81 1        -0.3  1        0     0                    1         1
##  7  7.71 1         0.01 1        0     0                    1         1
##  8  7.62 1        -0.26 1        0     0                    1         1
##  9  7.52 1         0.22 1        0     0                    1         1
## 10  7.45 1        -0.25 1        0     0                    1         1
##    pres_Carex
##         <dbl>
##  1          0
##  2          1
##  3          1
##  4          1
##  5          1
##  6          1
##  7          1
##  8          1
##  9          1
## 10          1
## # i 105 more rows
```

Look at the distribution of some variables

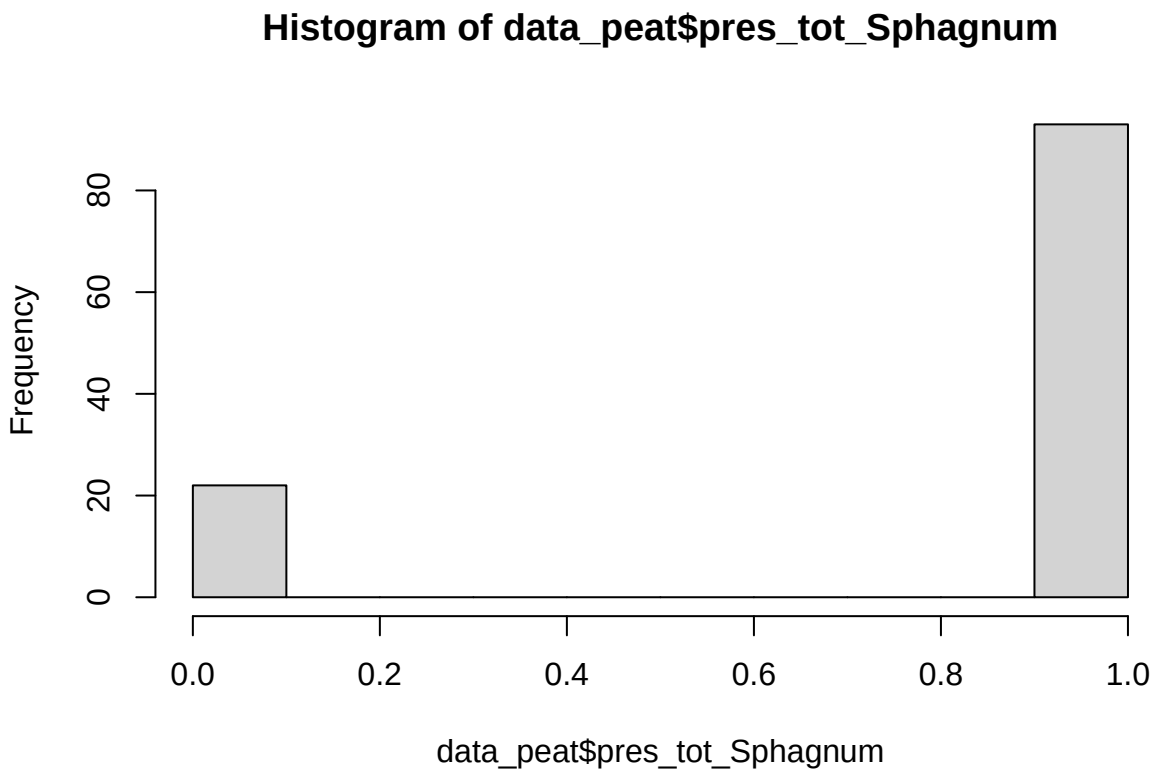
Histogram for total Sphagnum abundance

```
hist(data_peat$tot_Sphagnum)
```



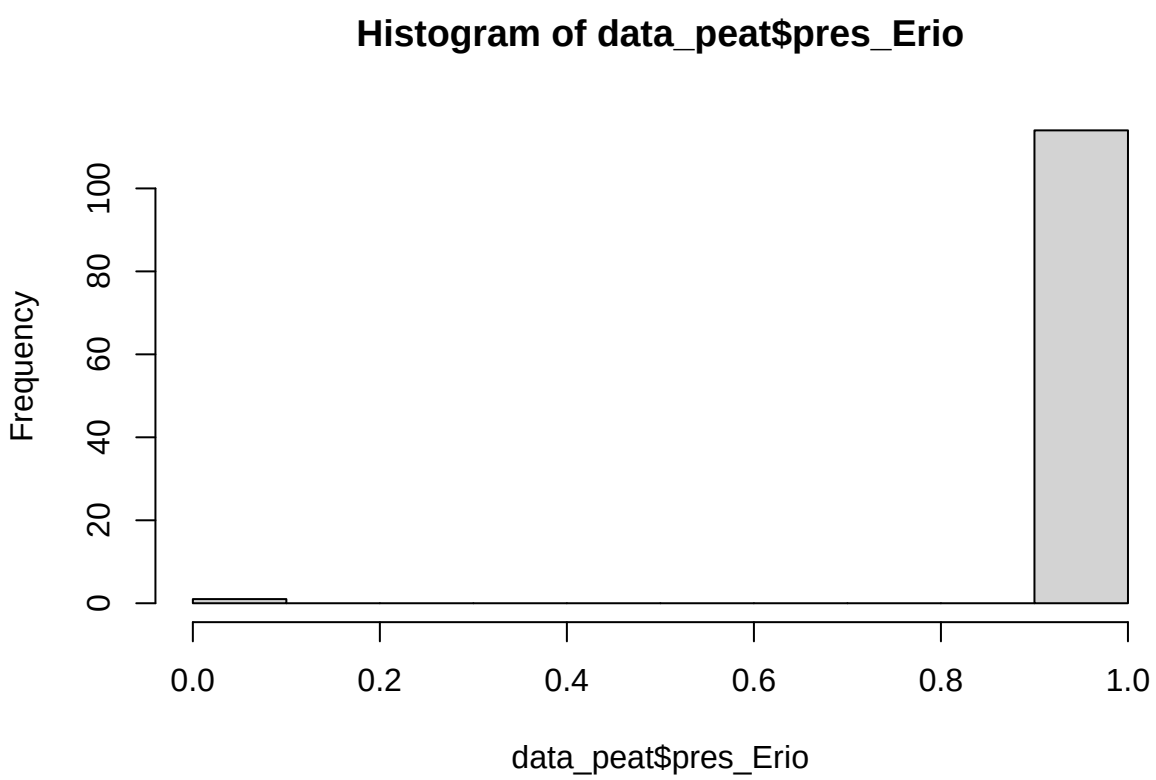
Histogram for total Sphagnum presence (you can see that this is either 0 = absence or 1 = presence)

```
hist(data_peat$pres_tot_Sphagnum)
```



Histogram for total Erio presence

```
hist(data_peat$pres_Erio)
```



There seems to be very few cases where pres_Erio=0. We can see them.

```
data_peat%>%filter(pres_Erio==0) # Show the cases where pres_Erio is 0
```

```
## # A tibble: 1 x 32
##   n_samples depth depth_corrected fen   tot_Sphagnum  Erio Carex Erica other_veg
##       <dbl> <chr>           <dbl> <fct>        <dbl> <dbl> <dbl> <dbl>     <dbl>
## 1         1 0                   0 N              100     0     0     0         0
##   Balticum Medium Cuspidata Austinii Fuscum Rubellum Acutifolia
##      <dbl>  <dbl>     <dbl>    <dbl>  <dbl>    <dbl>      <dbl>
## 1       32     68         0        0      0        0          0
##   'Diseased Acutifolia' Angustifolium Tenellum Papillosum Fallax Stems   age
##                   <dbl>         <dbl>    <dbl>      <dbl>  <dbl> <dbl> <dbl>
## 1                     0             0        0          0      0     0     0
##    temp imp_temp moist nutrient fire  dry   pres_tot_Sphagnum pres_Erio
##   <dbl> <fct>    <dbl> <fct>    <fct> <fct>             <dbl>     <dbl>
## 1  6.35 0           NA 0        0     0                     1         0
##   pres_Carex
##        <dbl>
## 1          0
```

```
# Note the double equal sign!
```

There is only one row where Erio is absent - You won't be able to fit presence/absence models for this species where there is only one absence! I think you might have the same problem with other species. But you might fit models for the abundance of those species.

Check the correlations among your independent variables

Construct a correlation matrix with all your (numeric) independent variables

```
cor(data_peat%>%
      select(age,temp,moist),    # Select the 3 numeric independent variables
    use="pairwise.complete.obs")
```

```
##             age      temp     moist
## age   1.0000000 0.6549057 0.1826218
## temp  0.6549057 1.0000000 0.2159498
## moist 0.1826218 0.2159498 1.0000000
```

```
# This means tha the correlation between each pair of variables
# is computed using all complete pairs of observations on those variables
```

As you mentioned the correlation among age and temp is 0.655 - we can see later if this is a problem in the models.

Fit a couple of models for presence / absence

Total Sphagnum presence

Let's fit a model for total Sphagnum presence

```
mod_pres_tot_Sphagnum<-glm(pres_tot_Sphagnum~   # Your response variable
                             age+temp+moist+
                             nutrient+fire+dry, # Your independent variables
                           family="binomial",   # Use the binomial family
                           data=data_peat)      # Your data
```

And see the results

```
summary(mod_pres_tot_Sphagnum)
```

```
##
## Call:
## glm(formula = pres_tot_Sphagnum ~ age + temp + moist + nutrient +
##     fire + dry, family = "binomial", data = data_peat)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.2978   0.1716   0.4071   0.5123   1.3310
##
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)  5.4314858  2.9934304   1.814   0.0696 .
## age         -0.0003262  0.0001725  -1.891   0.0586 .
## temp        -0.2926687  0.4085821  -0.716   0.4738
## moist        0.1308321  0.6962661   0.188   0.8510
```

```
## nutrient1    1.1397054  0.8525063   1.337   0.1813
## fire1       -0.6592563  0.7034827  -0.937   0.3487
## dry1        -0.1582005  0.7156088  -0.221   0.8250
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 103.500  on 107  degrees of freedom
## Residual deviance:  81.899  on 101  degrees of freedom
##   (7 observations deleted due to missingness)
## AIC: 95.899
##
## Number of Fisher Scoring iterations: 5
```

The effect of age is almost significant (p=0.0586).

To see if there are problems with multicollinearity, you can check the Variance Inflation Factors (VIFs). You can read a bit about VIFs here: https://www.statology.org/variance-inflation-factor-r/ Basically, if they are under some pre-defined value (some say 5, others say 2), all is good.

```
vif(mod_pres_tot_Sphagnum)
```

```
##      age     temp    moist nutrient     fire      dry
## 1.826948 1.811131 1.599711 1.568560 1.104748 1.664673
```

Here they are all under 2 so everything OK! You can check the VIFs for all your models to see if there could be any problem with multicollinearity.

You can try to fit a model only with age, to see if the effect is significant then.

```
mod_pres_tot_Sphagnum_age<-glm(pres_tot_Sphagnum~   # Your response variable
                                 age,               # Your independent variable
                               family="binomial",   # Use the binomial family
                               data=data_peat)      # Your data
```

And see the results

```
summary(mod_pres_tot_Sphagnum_age)
```

```
##
## Call:
## glm(formula = pres_tot_Sphagnum ~ age, family = "binomial", data = data_peat)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.3961   0.2367   0.3360   0.6476   1.3474
##
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)  3.8083521  0.7326311   5.198 2.01e-07 ***
## age         -0.0004663  0.0001170  -3.987 6.68e-05 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 112.265  on 114  degrees of freedom
## Residual deviance:  90.841  on 113  degrees of freedom
## AIC: 94.841
##
## Number of Fisher Scoring iterations: 5
```

The effect of age is significant here, and the probability of presence of total Sphagnum decreases with age. We can see the predicted effect of age.

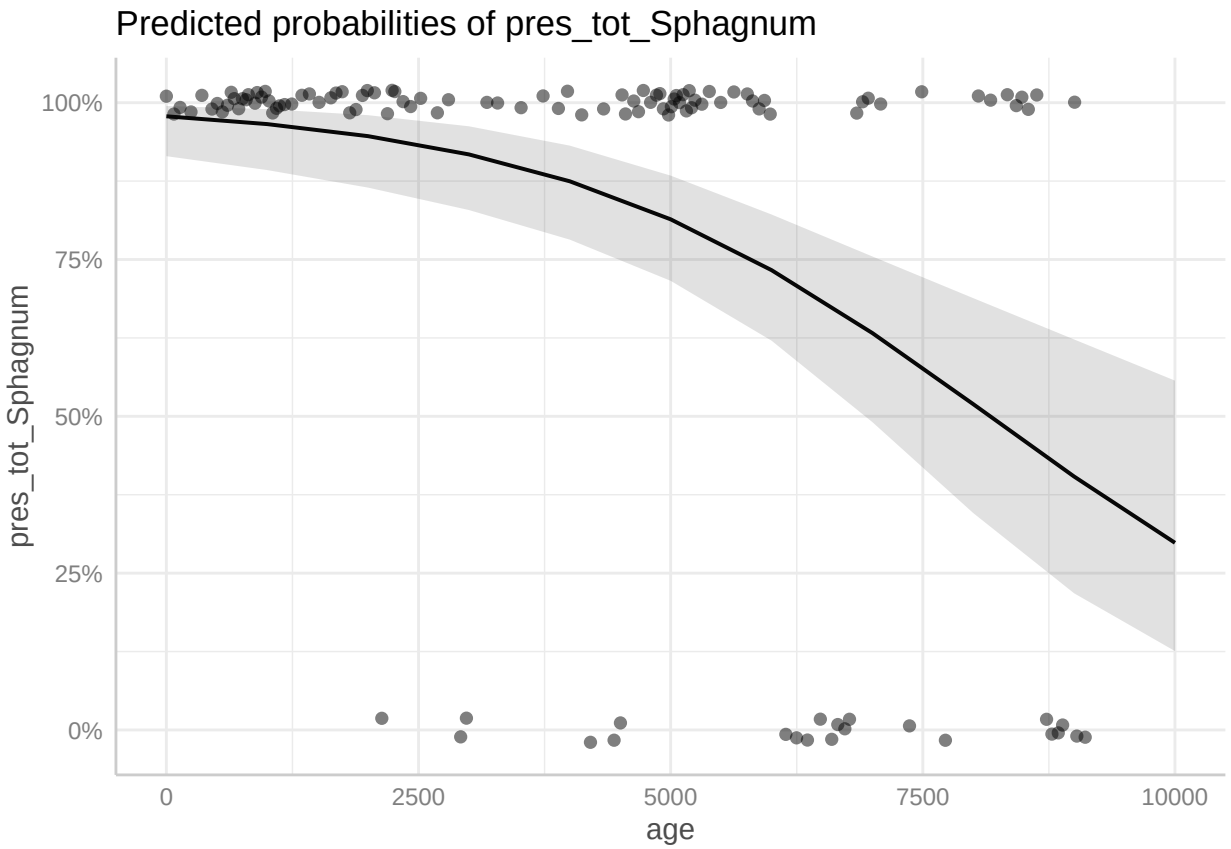
```
ggpredict(mod_pres_tot_Sphagnum_age)
```

```
## $age
## # Predicted probabilities of pres_tot_Sphagnum
##
##   age | Predicted |       95% CI
## --------------------------------
##     0 |      0.98 | [0.91, 0.99]
##  1000 |      0.97 | [0.89, 0.99]
##  3000 |      0.92 | [0.83, 0.96]
##  4000 |      0.87 | [0.78, 0.93]
##  5000 |      0.81 | [0.72, 0.88]
##  6000 |      0.73 | [0.62, 0.82]
##  7000 |      0.63 | [0.49, 0.75]
## 10000 |      0.30 | [0.13, 0.56]
##
## attr(,"class")
## [1] "ggalleffects" "list"
## attr(,"model.name")
## [1] "mod_pres_tot_Sphagnum_age"
```

This shows the predicted probability of presence of total Sphagnum for different ages, and the confidence intervals. You can change the age values if you want to, and do many other things with this ggeffects package (see https://cloud.r-project.org/web/packages/ggeffects/vignettes/content.html). I mainly use this for plotting the predicted effects, see below.

```
plot(ggpredict(mod_pres_tot_Sphagnum_age),
     # This plots the predicted effect from your model
     add.data=TRUE, # And this adds the data
     jitter=0.02)   # This adds a slight jitter to the data points
```

```
## $age
```




```
                    # so you can see them better
```

Erio presence

Let's fit a model for total Erio presence.

```
mod_pres_Erio<-glm(pres_Erio~   # Your response variable
                     age+temp+moist+
                     nutrient+fire+dry, # Your independent variables
                   family="binomial",   # Use the binomial family
                   data=data_peat)      # Your data
```

This works, but you get a warming message: "glm.fit: algorithm did not converge". And if you have a look at the results

```
summary(mod_pres_Erio)
```

```
##
## Call:
## glm(formula = pres_Erio ~ age + temp + moist + nutrient + fire +
##     dry, family = "binomial", data = data_peat)
##
## Deviance Residuals:
##       Min         1Q     Median         3Q        Max
```

```
## 2.409e-06  2.409e-06  2.409e-06  2.409e-06  2.409e-06
##
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)  2.657e+01  3.541e+05       0        1
## age         -7.579e-13  2.197e+01       0        1
## temp         1.758e-09  5.229e+04       0        1
## moist        1.421e-09  8.300e+04       0        1
## nutrient1   -4.819e-10  9.010e+04       0        1
## fire1        1.848e-09  9.660e+04       0        1
## dry1        -4.415e-06  7.788e+04       0        1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 0.0000e+00  on 107  degrees of freedom
## Residual deviance: 6.2657e-10  on 101  degrees of freedom
##   (7 observations deleted due to missingness)
## AIC: 14
##
## Number of Fisher Scoring iterations: 25
```

You get all p=1. This does not work because as we saw before, there is only one case where Erio is absent - you cannot fit presence/absence models in this case (and probably also for some other species where there are few absences or few presences).

Fit a couple of models for abundance

Total Sphagnum abundance

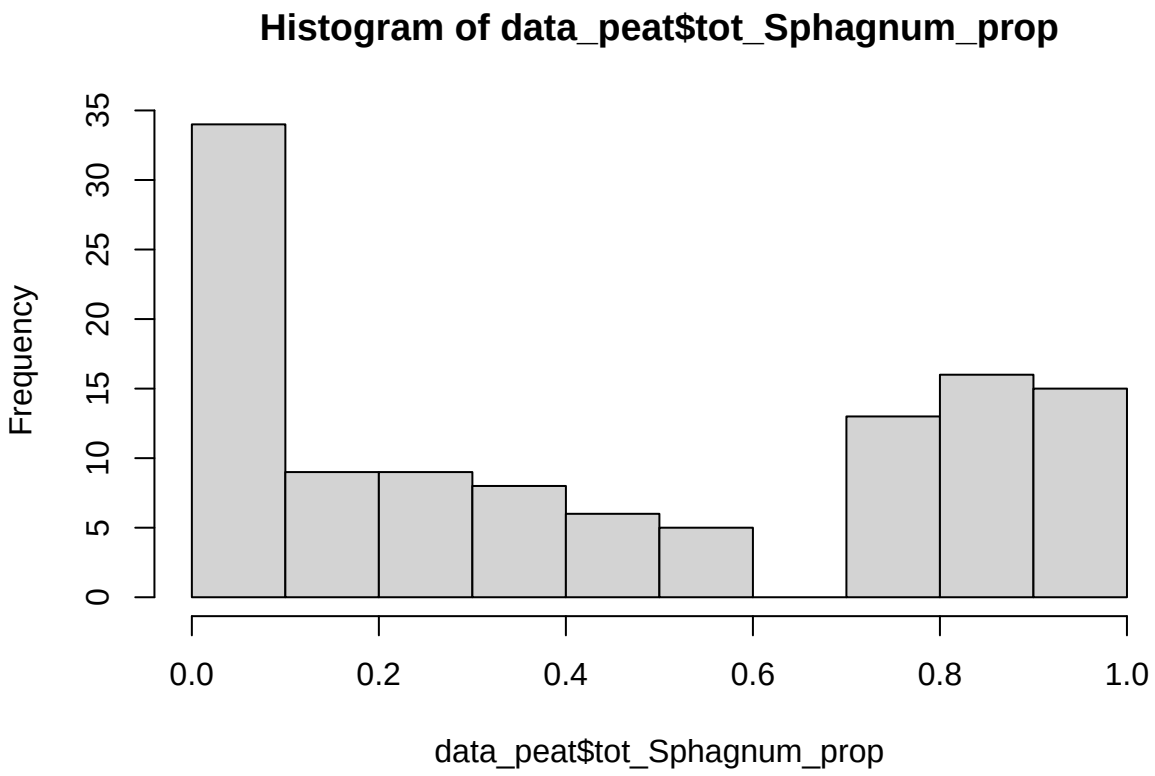
Your species abundances are continuous proportions (see Douma & Weedon 2019), and therefore they would be better analyzed with beta regression, instead of a binomial GLM. However, for beta regression you need values of proportions bounded between 0 and 1, but not including 0 and 1. I think that the best solution for your data is to use zero-augmented / zero-inflated beta regression (see that paper and the others that I sent you). This is like a two-part model, allowing to mix two distribution functions for the same response data: a beta function for the proportional part (response $y > 0$), and an additive function for the binomial part (response $y = 0$; with logit- and log-link respectively).

You need to convert your abundances to proportions between 0 and 1 (they are percentages now). For example, we convert total Sphagnum abundance to a proportion.

```
data_peat<-data_peat%>%
  mutate(tot_Sphagnum_prop=tot_Sphagnum/100)
```

If we make a histogram we see that the values are now between 0 and 1 (including 0 and 1).

```
hist(data_peat$tot_Sphagnum_prop)
```



Let's fit a model for total Sphagnum abundance. I used the function glmmTMB in the glmmTMB package, but probably you can use others (also in a Bayesian setting if you feel like it).

```
mod_abund_tot_Sphagnum<-glmmTMB(tot_Sphagnum_prop~    # Your response variable
                                  age+temp+moist+
                                  nutrient+fire+dry,  # Independent variables
                                family="beta_family", # Use the beta family
                                ziformula=~.,
                                data=data_peat)       # Your data

# ziformula=~. sets the zero-inflation formula identical to the formula
# for the proportion part. This basically means that the same effects will be
# tested on presence/absence and on abundance.
# If you specify ziformula=~0 this means that there is no zero-inflation, but
# this will throw an error with your data, because you do have zeros.
# You can also specify a different formula for the zero-inflation part,
# for example, ziformula=~age will only test the effect of age
# on presence/absence, but all the effects on abundance
```

And see the results

```
summary(mod_abund_tot_Sphagnum)
```

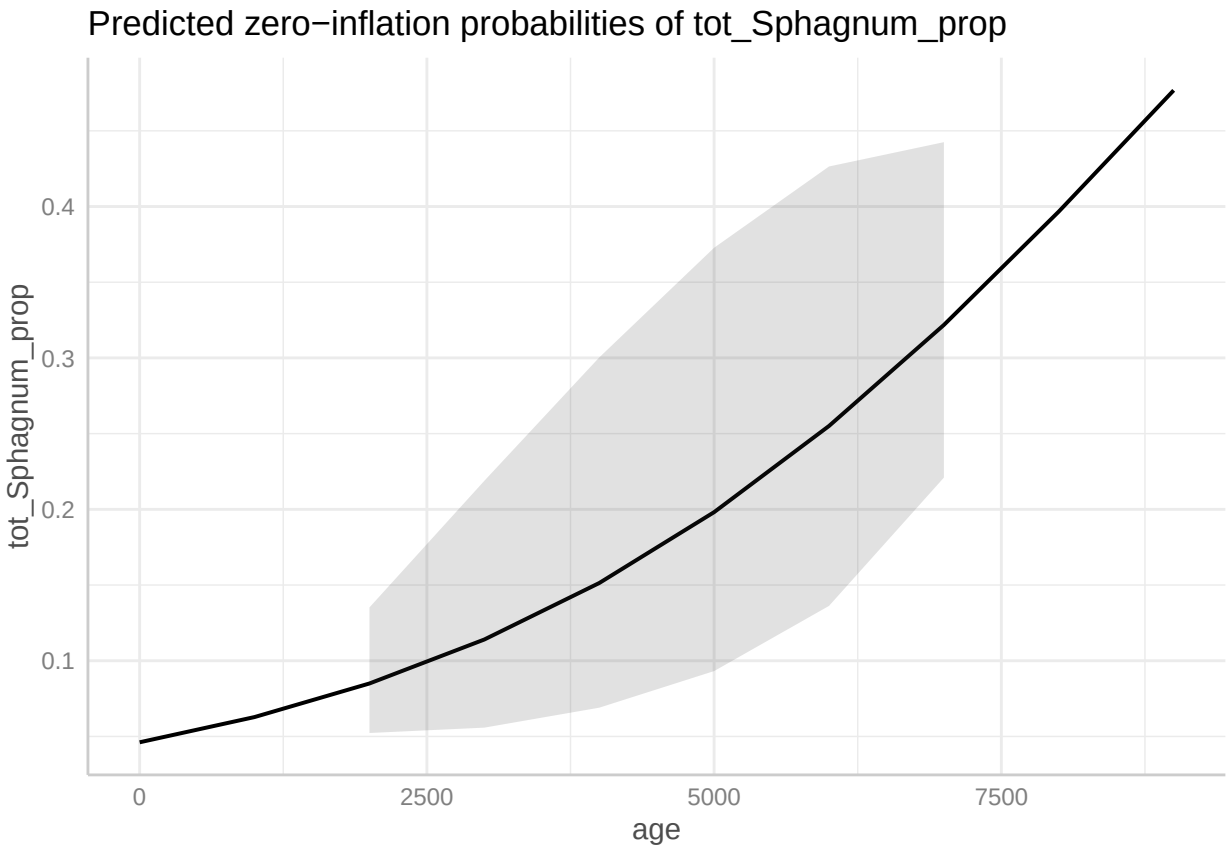
```
##  Family: beta  ( logit )
## Formula:
```

```
## tot_Sphagnum_prop ~ age + temp + moist + nutrient + fire + dry
## Zero inflation:                      ~.
## Data: data_peat
##
##      AIC      BIC   logLik deviance df.resid
##     76.4    116.6    -23.2     46.4       93
##
##
## Dispersion parameter for beta family (): 2.46
##
## Conditional model:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)  9.117e-01  1.303e+00   0.700  0.48422
## age         -4.403e-05  8.234e-05  -0.535  0.59285
## temp        -1.801e-01  1.953e-01  -0.922  0.35661
## moist       -2.880e-01  2.865e-01  -1.005  0.31489
## nutrient1    1.015e+00  3.129e-01   3.244  0.00118 **
## fire1       -1.350e-01  3.547e-01  -0.381  0.70346
## dry1         7.583e-01  2.503e-01   3.030  0.00244 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Zero-inflation model:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept) -5.4314844  2.9936427  -1.814   0.0696 .
## age          0.0003262  0.0001725   1.891   0.0587 .
## temp         0.2926684  0.4086025   0.716   0.4738
## moist       -0.1308289  0.6963095  -0.188   0.8510
## nutrient1   -1.1397037  0.8525940  -1.337   0.1813
## fire1        0.6592571  0.7035135   0.937   0.3487
## dry1         0.1581995  0.7156631   0.221   0.8251
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Here, the part "Conditional model" shows the effects on abundance, and "Zero-inflation model" shows the effects on presence/absence. If you look at the coefficients for the zero-inflation part, you can see that they look pretty similar to those in the model for Total Sphagnum presence (mod_pres_tot_Sphagnum), just with the opposite sign, because here they show the probability of Spaghnum absence (i.e. the probability of having a zero). The conditional model part shows significant effects of nutrient input and dry period on total Sphagnum abundance: abundance is higher when there is a nutrient input (nutrient=1) and when there is a dry period (dry=1).

Let's plot some of the predicted effects. For example, the effect of age on the probability of zero-inflation (i.e. probability of absence):

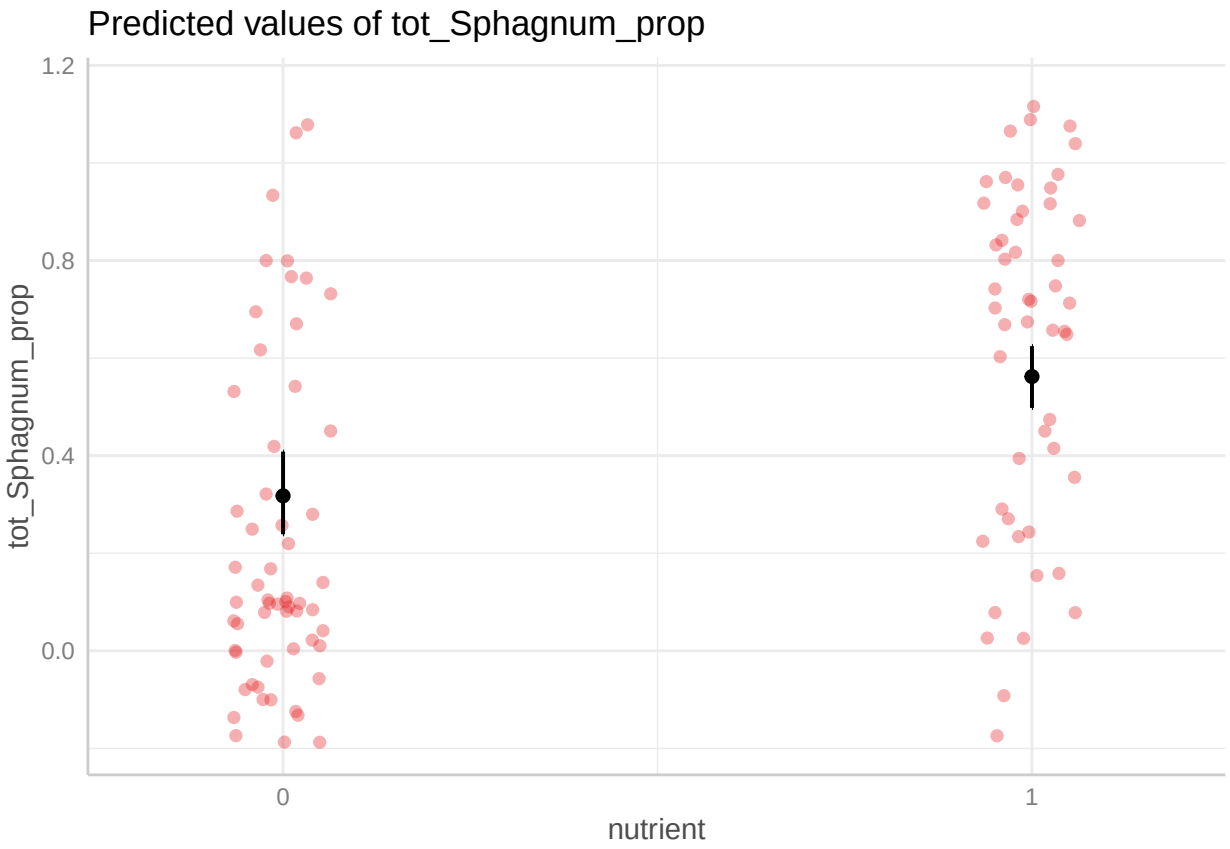
```
plot(ggpredict(mod_abund_tot_Sphagnum,
               # This plots the predicted effect from your model
               type="zi_prob",  # Plots the predicted zero-inflation probability
               terms=c("age")), # For the effect of age
     add.data=FALSE) # Not adding the data here
```



I am not sure why the error band looks strange on this one (I would need to look into this), but you can see that the effect of age is the opposite as before, because this is the probability of getting a zero.

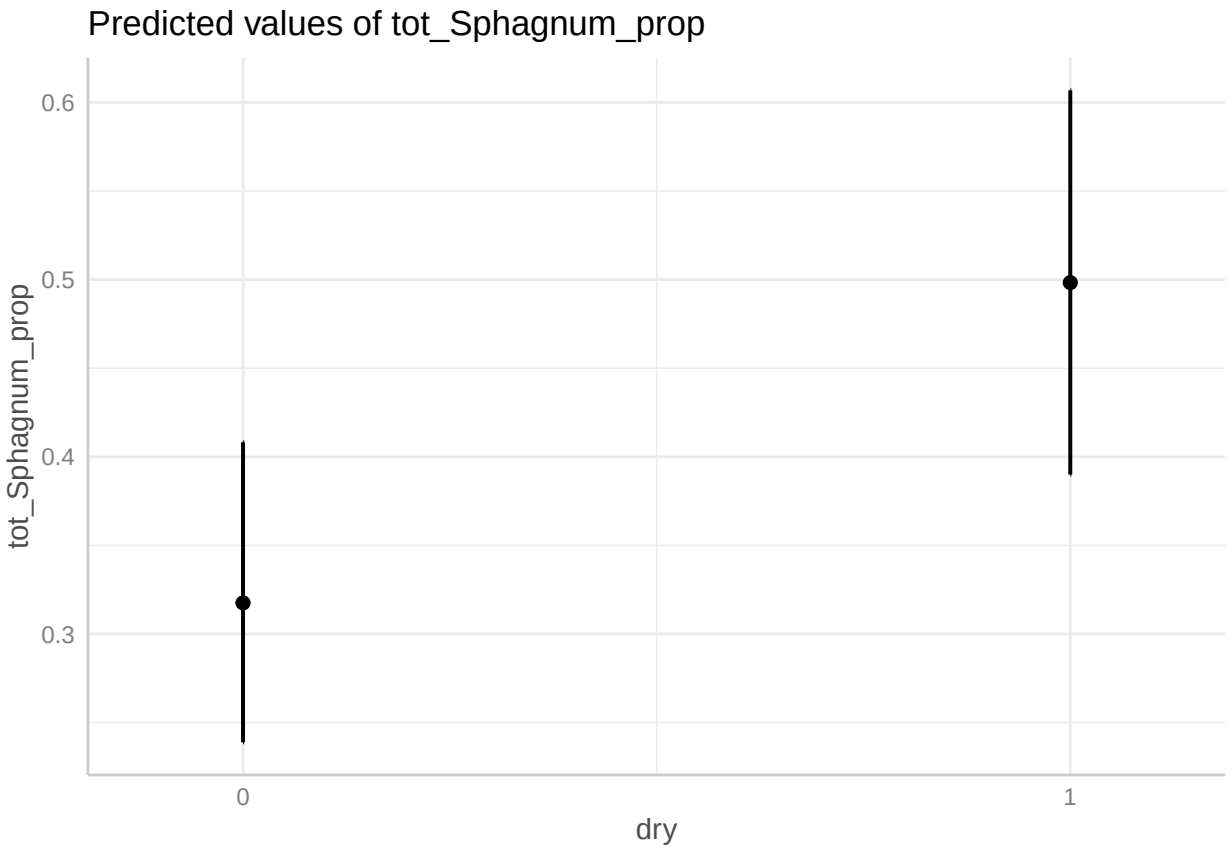
The effect of nutrient input on the abundance:

```
plot(ggpredict(mod_abund_tot_Sphagnum,
               # This plots the predicted effect from your model
               type="fixed",  # Plots the predicted effects for the conditional
               # model only (i.e. effects on abundance)
               terms=c("nutrient")), # For the effect of nutrient
     add.data=TRUE) # You can add the data if you want
```



The effect of dry period on the abundance:

```
plot(ggpredict(mod_abund_tot_Sphagnum,
               # This plots the predicted effect from your model
               type="fixed",  # Plots the predicted effects for the conditional
               # model only (i.e. effects on abundance)
               terms=c("dry")), # For the effect of dry
     add.data=FALSE) # I did not add the data here
```



These plots are just for showing some examples, but for publication-ready plots I prefer to use ggplot and build them from scratch with raw data points + predictions from ggpredict(). But we can get into that later!

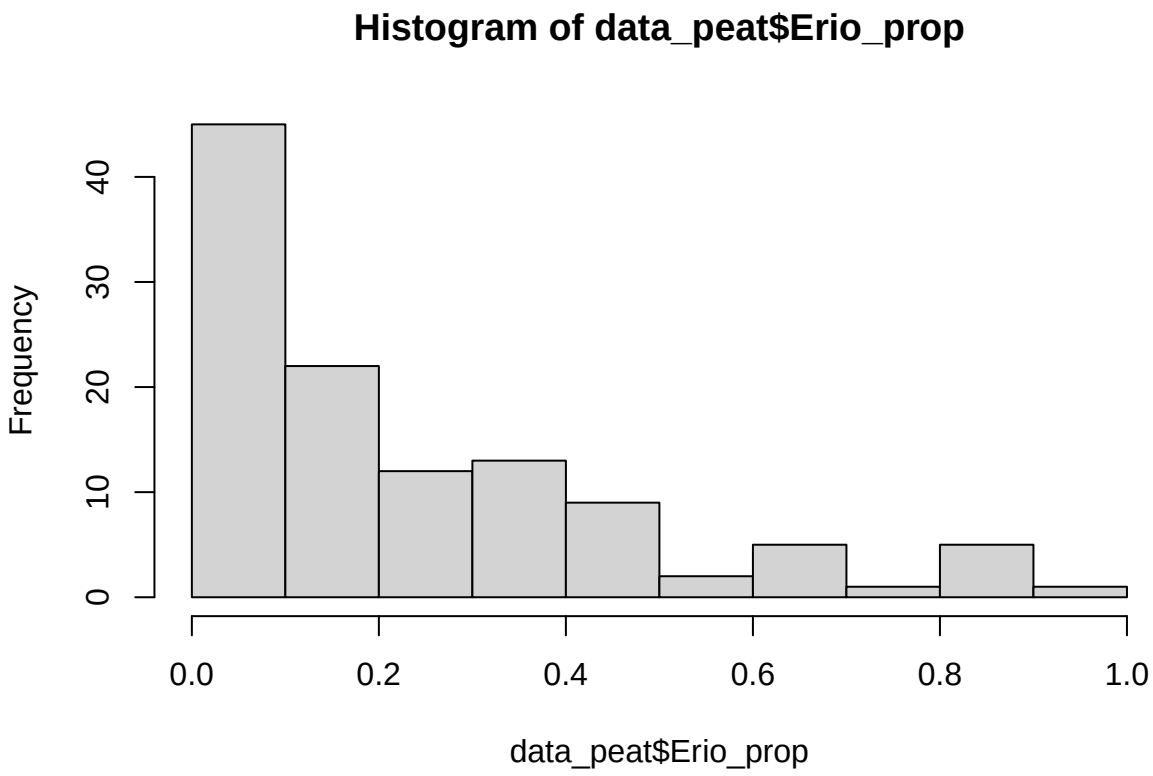
Erio abundance

Convert Erio abundances to a proportion.

```
data_peat<-data_peat%>%
  mutate(Erio_prop=Erio/100)
```

See the histogram:

```
hist(data_peat$Erio_prop)
```



Let's fit a model for Erio abundance:

```
mod_abund_Erio<-glmmTMB(Erio_prop~    # Your response variable
                          age+temp+moist+
                          nutrient+fire+dry,  # Independent variables
                        family="beta_family", # Use the beta family
                        ziformula=~.,
                        data=data_peat)       # Your data
```

And see the results

```
summary(mod_abund_Erio)
```

```
##  Family: beta  ( logit )
## Formula:          Erio_prop ~ age + temp + moist + nutrient + fire + dry
## Zero inflation:             ~.
## Data: data_peat
##
##      AIC      BIC   logLik deviance df.resid
##    -82.5    -42.3     56.3   -112.5       93
##
##
## Dispersion parameter for beta family ():  3.9
##
## Conditional model:
```

```
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept) -2.087e+00        NaN     NaN      NaN
## age          2.145e-05        NaN     NaN      NaN
## temp         1.721e-01        NaN     NaN      NaN
## moist        3.196e-01        NaN     NaN      NaN
## nutrient1   -7.430e-01        NaN     NaN      NaN
## fire1        7.265e-02        NaN     NaN      NaN
## dry1        -4.806e-01        NaN     NaN      NaN
##
## Zero-inflation model:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.270595        NaN     NaN      NaN
## age         -0.012377        NaN     NaN      NaN
## temp        -1.867539        NaN     NaN      NaN
## moist        0.024751        NaN     NaN      NaN
## nutrient1   -0.119101        NaN     NaN      NaN
## fire1       -0.029446        NaN     NaN      NaN
## dry1         0.003262        NaN     NaN      NaN
```

There is also a problem with fitting this model, because as we saw before, there is only one case where Erio is absent (and therefore Erio_prop=0), so we cannot fit the zero-inflation part in this case. We can fit a beta regression for the proportion part:

```
mod_abund_Erio_nozi<-glmmTMB(Erio_prop~    # Your response variable
                               temp+moist+
                               nutrient+fire+dry+fen,  # Independent variables
                             family="beta_family", # Use the beta family
                             ziformula=~0,         # no zero-inflation
                             data=subset(data_peat,Erio_prop>0)) # Your data
# Here, we use only the data where Erio_prop is larger than zero, i.e. we
# remove the only one case where Erio is absent, and look at effects
# on the abundance only (not presence/absence)
```

And see the results

```
summary(mod_abund_Erio_nozi)
```

```
##  Family: beta  ( logit )
## Formula:          Erio_prop ~ temp + moist + nutrient + fire + dry + fen
## Data: subset(data_peat, Erio_prop > 0)
##
##      AIC      BIC   logLik deviance df.resid
##   -110.0    -88.6     63.0   -126.0      100
##
##
## Dispersion parameter for beta family (): 4.45
##
## Conditional model:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -1.26925    0.81503  -1.557 0.119401
## temp         0.01490    0.10575   0.141 0.887965
## moist        0.29878    0.21333   1.401 0.161350
## nutrient1   -0.28875    0.25469  -1.134 0.256919
```

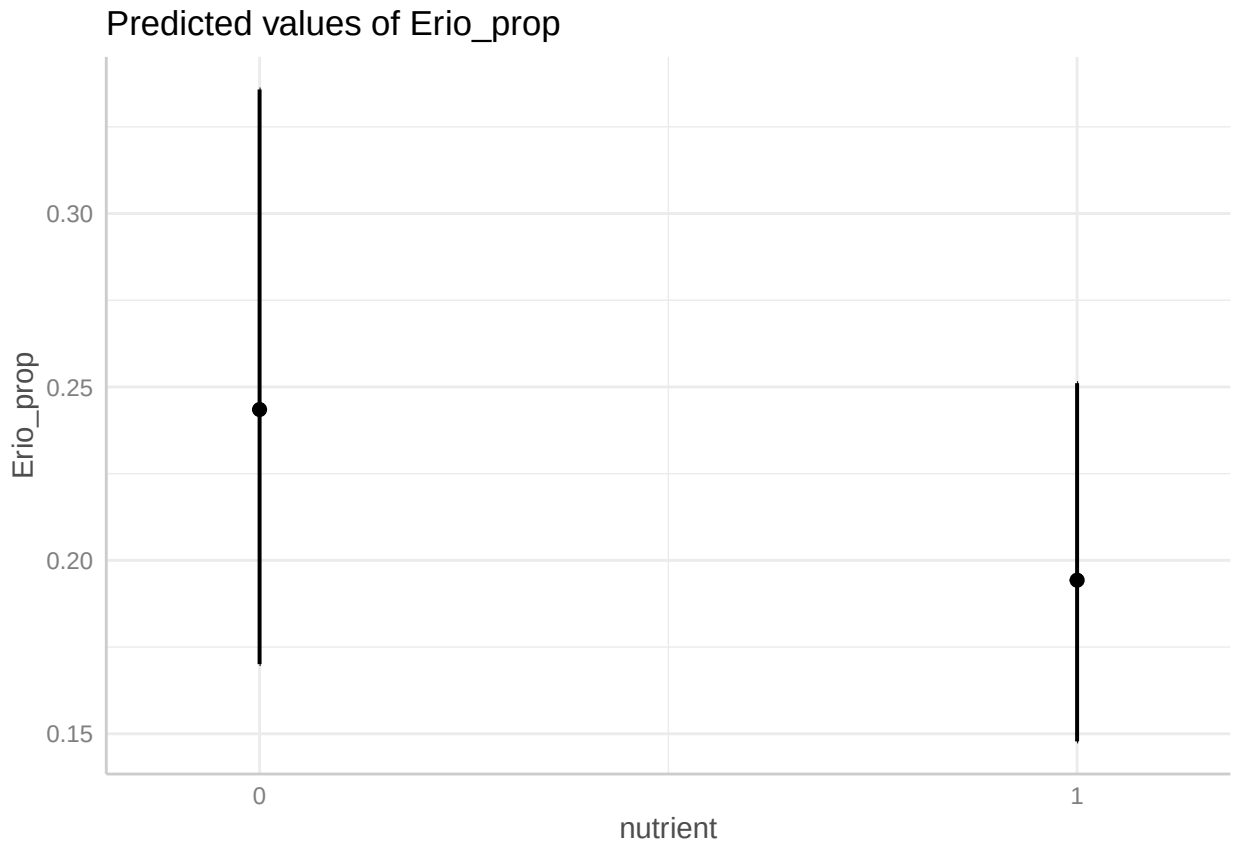
```
## fire1        0.07468    0.24629   0.303 0.761719
## dry1        -0.34993    0.19857  -1.762 0.078033 .
## fenY         1.03600    0.27830   3.723 0.000197 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Here, we only see the part "Conditional model", showing the effects on abundance. There are significant effects of nutrient input and dry period on total Erio abundance: abundance is lower when there is a nutrient input (nutrient=1) and when there is a dry period (dry=1). So the effects are opposite to Sphagnum!

Let's plot some of the predicted effects.

The effect of nutrient input on Erio abundance:

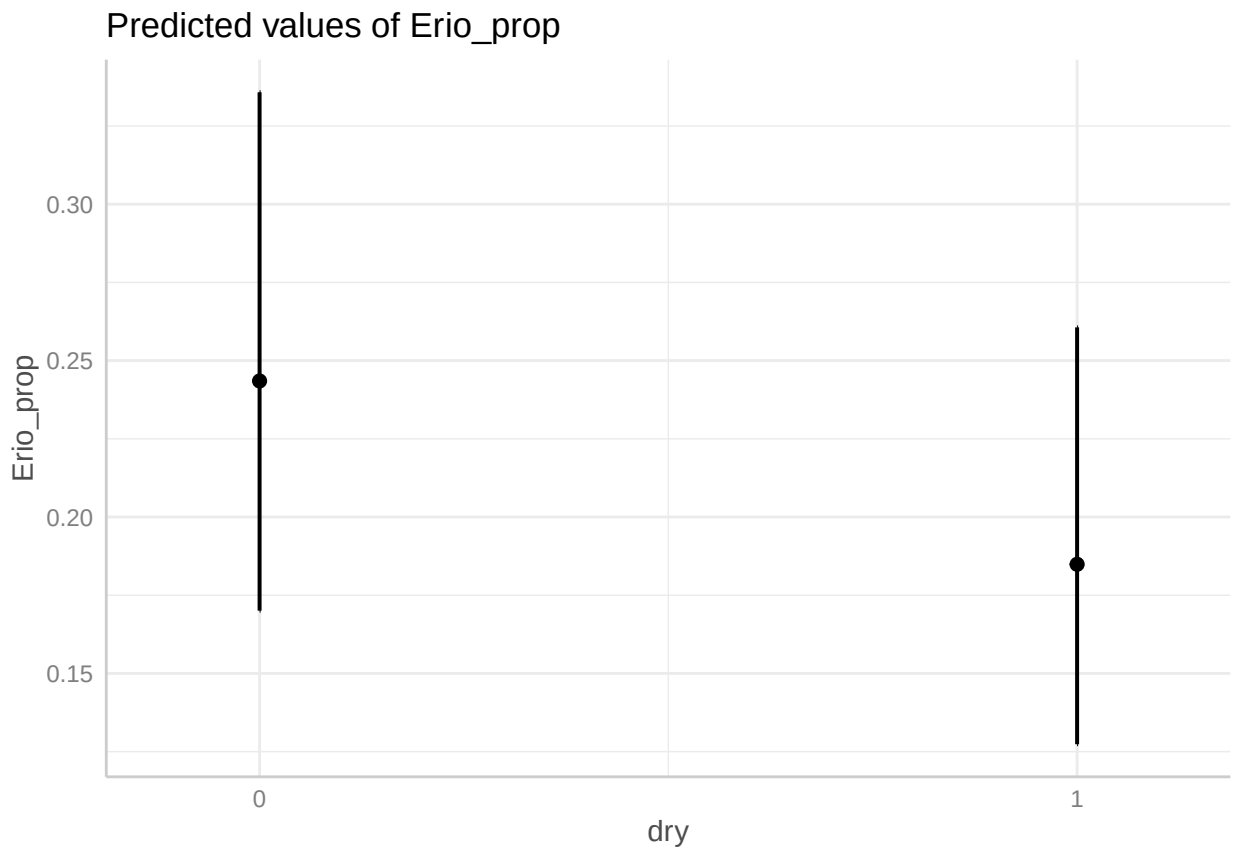
```
plot(ggpredict(mod_abund_Erio_nozi,
               # This plots the predicted effect from your model
               type="fixed",  # Plots the predicted effects for the conditional
               # model only (i.e. effects on abundance)
               terms=c("nutrient")), # For the effect of nutrient
     add.data=FALSE)
```



The effect of dry period on Erio abundance:

```
plot(ggpredict(mod_abund_Erio_nozi,
               # This plots the predicted effect from your model
               type="fixed",  # Plots the predicted effects for the conditional
               # model only (i.e. effects on abundance)
```

```
                terms=c("dry")), # For the effect of dry
    add.data=FALSE)
```



I guess you can try this kind of models (zero-augmented / zero-inflated beta regression) for all the species you are interested in, and then you would be looking both at effects on presence/absence and on abundance. For species where you have only few absences or presences (such as Erio), I think you would only be able to look at abundance.

Session info

This just shows info on your R session, might be useful.

```
sessionInfo()
```

```
## R version 4.2.3 (2023-03-15 ucrt)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 10 x64 (build 19045)
##
## Matrix products: default
##
## locale:
## [1] LC_COLLATE=English_Sweden.utf8  LC_CTYPE=English_Sweden.utf8
## [3] LC_MONETARY=English_Sweden.utf8 LC_NUMERIC=C
## [5] LC_TIME=English_Sweden.utf8
```

```
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
##  [1] glmmTMB_1.1.7   car_3.1-2       carData_3.0-5   ggeffects_1.2.1
##  [5] knitr_1.42      readxl_1.4.2    lubridate_1.9.2 forcats_1.0.0
##  [9] stringr_1.5.0   dplyr_1.1.1     purrr_1.0.1     readr_2.1.4
## [13] tidyr_1.3.0     tibble_3.2.1    ggplot2_3.4.2   tidyverse_2.0.0
##
## loaded via a namespace (and not attached):
##  [1] tidyselect_1.2.0    xfun_0.38           TMB_1.9.3
##  [4] haven_2.5.2         splines_4.2.3       lattice_0.20-45
##  [7] colorspace_2.1-0    vctrs_0.6.1         generics_0.1.3
## [10] htmltools_0.5.5     yaml_2.3.7          utf8_1.2.3
## [13] rlang_1.1.0         nloptr_2.0.3        pillar_1.9.0
## [16] glue_1.6.2          withr_2.5.0         lifecycle_1.0.3
## [19] munsell_0.5.0       gtable_0.3.3        cellranger_1.1.0
## [22] evaluate_0.20       labeling_0.4.2      tzdb_0.3.0
## [25] fastmap_1.1.1       fansi_1.0.4         highr_0.10
## [28] Rcpp_1.0.10         scales_1.2.1        abind_1.4-5
## [31] farver_2.1.1        lme4_1.1-32         hms_1.1.3
## [34] digest_0.6.31       stringi_1.7.12      insight_0.19.1
## [37] numDeriv_2016.8-1.1 grid_4.2.3          cli_3.6.1
## [40] tools_4.2.3         magrittr_2.0.3      pkgconfig_2.0.3
## [43] MASS_7.3-58.2       Matrix_1.5-3        timechange_0.2.0
## [46] minqa_1.2.5         rmarkdown_2.21      rstudioapi_0.14
## [49] R6_2.5.1            boot_1.3-28.1       nlme_3.1-162
## [52] compiler_4.2.3
```