

Appendix B - R code for reproducing the analysis

1. Introduction

In this knitr document (Xie 2013) we present the code we used generate the results we present and discuss in “Conservation zones promote oak regeneration and shrub diversity in certified Mediterranean oak woodlands”. We assume the reader is familiar with the R packages “Distance” and “dsm”. If you wish to know more about using these packages we recommend the Supplementary information of Miller et al. (2013) and Winiarski et al. (2013).

2. Preamble

We start by loading the workspace and the packages we used.

```
load("ws.RData")
suppressPackageStartupMessages(library(knitr))
suppressPackageStartupMessages(library(Distance))
suppressPackageStartupMessages(library(dsm))
suppressPackageStartupMessages(library(spatial))
suppressPackageStartupMessages(library(mvtnorm))
```

2.1 The data

This dataset includes all observed cork oaks plants, the distance between each plant and the transect and some observation-level and transect-level covariates.

```
names(data)
```

```
## [1] "id"          "estate"      "zone_type"   "zone_code"
## [5] "transect"    "distance"    "height"      "shrubs_dens"
## [9] "shrubs_cover" "tree_dens"   "tree_cover"  "class"
## [13] "effort"      "Sample.Label" "Region.Label" "size"
## [17] "object"
```

Description of each column in the dataset

1. `id` - unique identifier for each observation
2. `estate` - surveyed estate
3. `zone_type` - CZ - conservation zone; NCZ - non-conservation
4. `zone_code` - unique identifier for each surveyed conservation and non-conservation zone
5. `transect` - unique identifier for each surveyed transect
6. `distance` - perpendicular distance between each plant and the transect (m)
7. `height` - height of each plant (cm)
8. `shrubs_dens` - shrub density in the transect (plants/m²)
9. `shrubs_cover` - shrub cover in the transect (% of ground covered)
10. `tree_dens` - tree density in the transect (tree/ha)
11. `tree_cover` - tree cover in the transect (m²/ha)

12. `class` - age class of the plant (seed - seedling, sap - sapling or yt - young tree)
13. `effort` - length of the transect (m)
14. `Sample.Label` - unique identifier for each surveyed transect
15. `Region.Label` - unique identifier for each surveyed conservation and non-conservation zone
16. `size` - cluster size. Since there were no clusters, all observations have value 1.
17. `object` - unique identifier for each observation.

3. Exploratory data analysis

We begin by plotting histograms with the pooled data and with the data split by zone type (conservation zones and non-conservation zones), estate and zone.

3.1 Histogram of observed distances - pooled data

```
par(pin=c(2,2))
hist(data$distance, xlab="distance (m)", main="Histogram of observed distances")
```

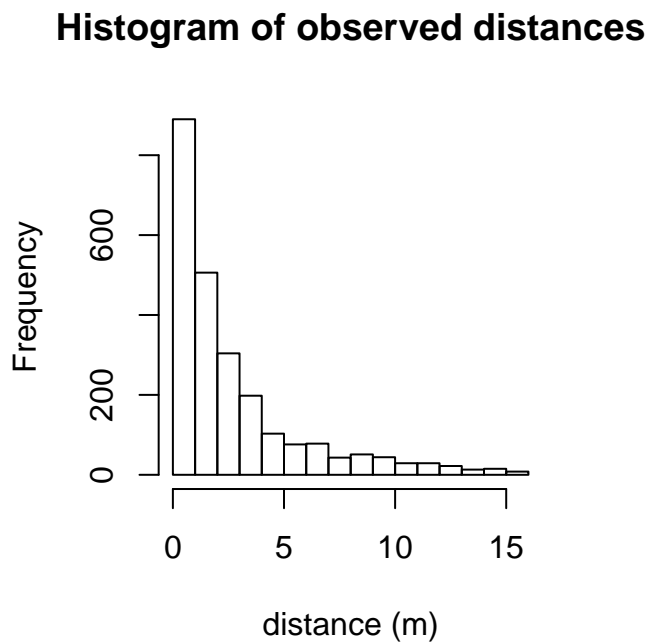


Figure 1 - Histogram of observed distances.

3.2 Histograms of observed distances in conservation and non-conservation zones

```
par(mfrow=c(1,2), mar=c(2,2,2,2), pin=c(1.5,1.5))
hist(subset(data, zone_type=="cz")$distance, xlab="distance (m)", main="Conservation Zones")
hist(subset(data, zone_type=="ncz")$distance, xlab="distance (m)", main="Non-conservation Zones")
```

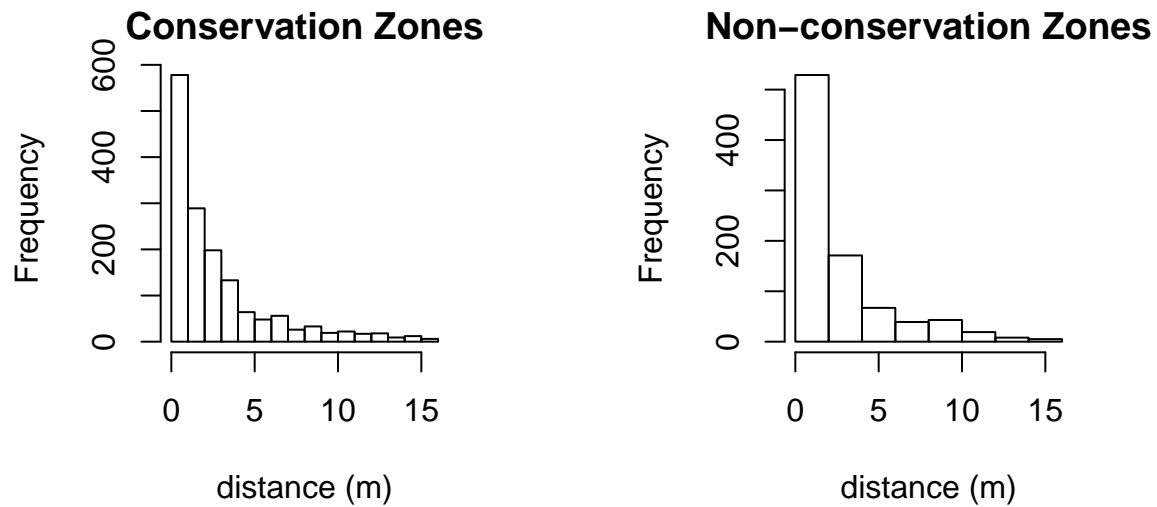


Figure 2 - Histograms of observed distances in conservation and non-conservation zones.

3.3 Histograms of observed distances in each estate

```
par(mfrow=c(3,3),mar=c(2,2,2,2),pin=c(0.8,0.8))
levels(data$estate)->estates
for(i in seq_along(estates)){
  hist(subset(data,estate==estates[i])$distance,main=estates[i],xlab="distance (m)",ylab="Frequency")
}
par(mfrow=c(1,1))
```

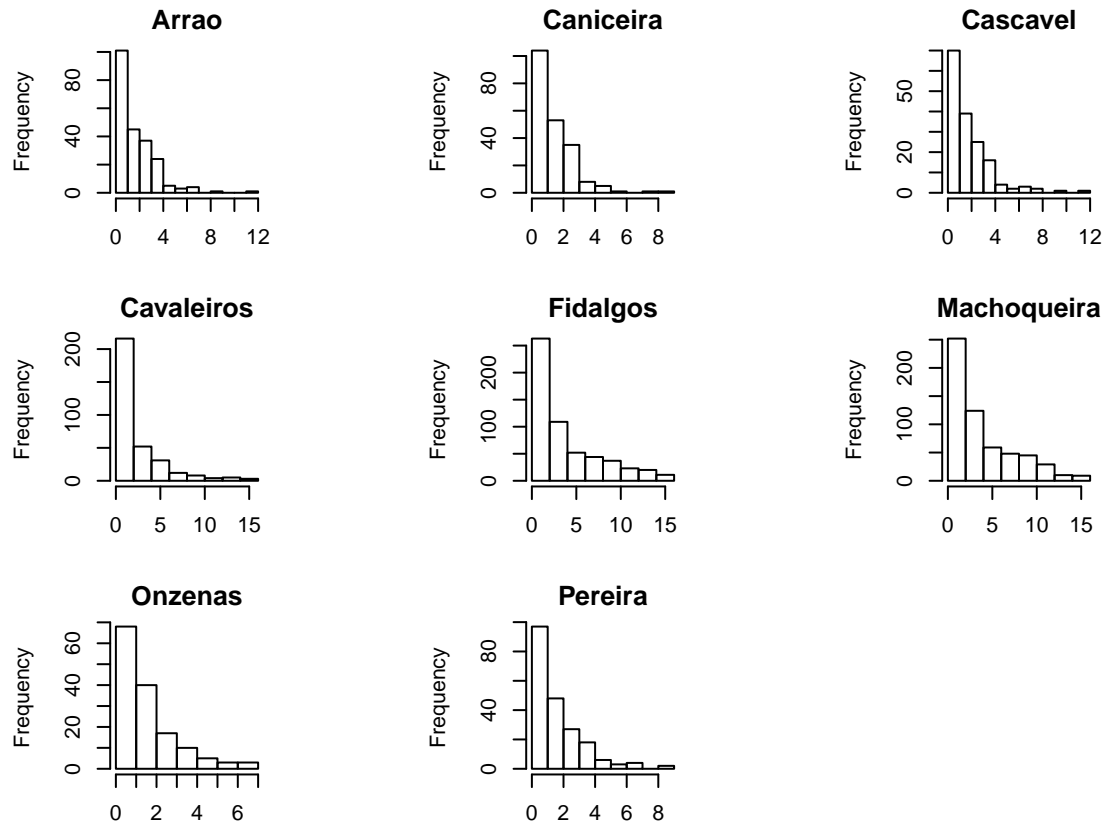


Figure 3 - Histogram of observed distances in each estate.

3.4 Histogram of observed distances split by each surveyed conservation and non-conservation zone

```
par(mfrow=c(6,5),mar=c(2,2,2,2))
#Codes for surveyed conservation and non-conservation zone
zones<-c("ArraoZC1","ArraoZNC1","CanZC1","CanZNC1","CascZC1","CascZNC1","CavZC1","CavZC2",
         "CavZNC1", "FidZC1","FidZNC1","FidZC2","FidZNC2","FidZC3","FidZNC3","FidZC4","FidZNC4",
         "FidZC5","FidZNC5", "MacZC1","MacZNC1","MacZC2","MacZNC2","OnzZC1","OnzZNC1","PerZC1",
         "PerZNC1")
for (i in seq_along(zones)){
  hist(subset(data,data$Region.Label==zones[i])$distance,main=zones[i],xlab="distance (m)")
}
par(mfrow=c(1,1))
```

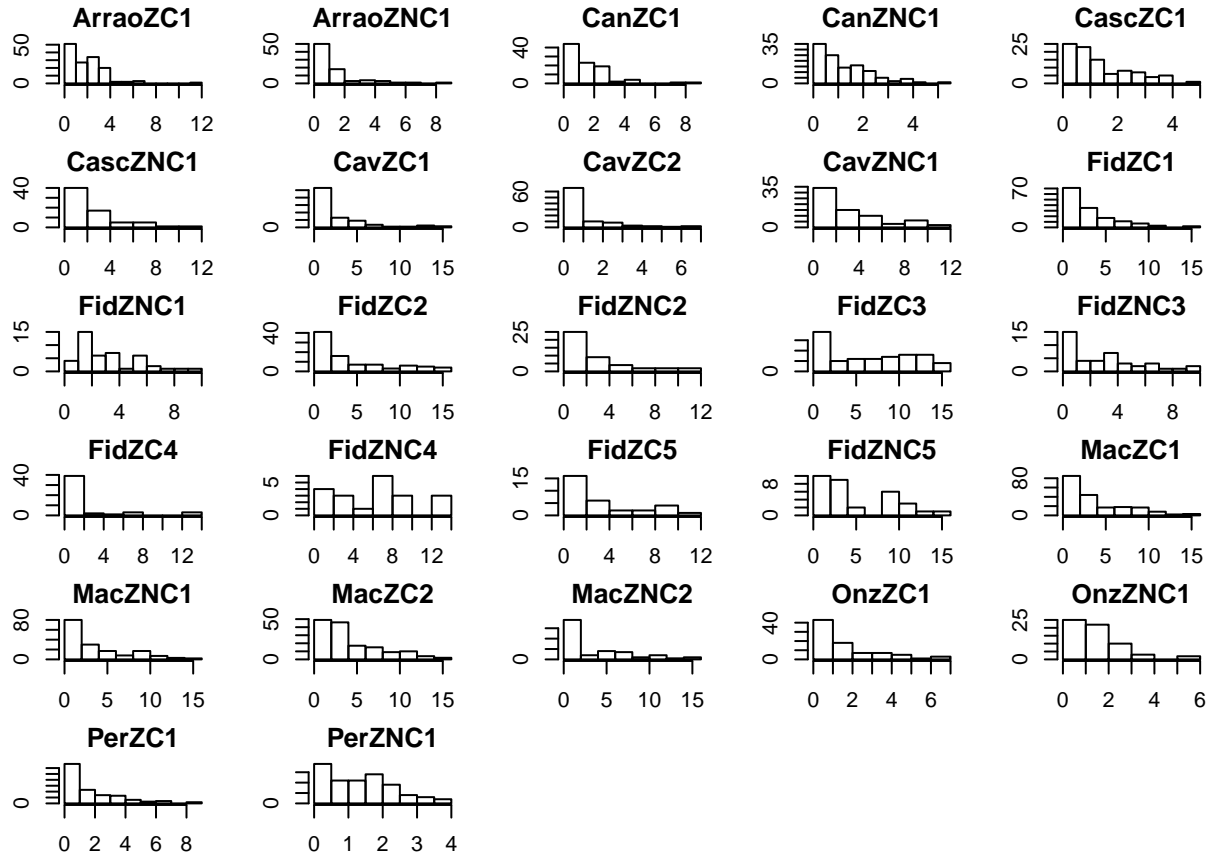


Figure 4 - Histogram of observed distances split by each surveyed conservation and non-conservation zone. Now we investigate whether the distance at which each plant was observed was affected by the height of each plant and by shrub density, shrub cover, tree cover and tree density (measured on each transect).

3.5 Distance of each plant to the center line against its height

```
plot(data$distance,data$height,main="Distance vs Height of each plant",
      ylab="Height (cm)",xlab="Distance (m)")
lines(lowess(data$distance,data$height),col=2)
```

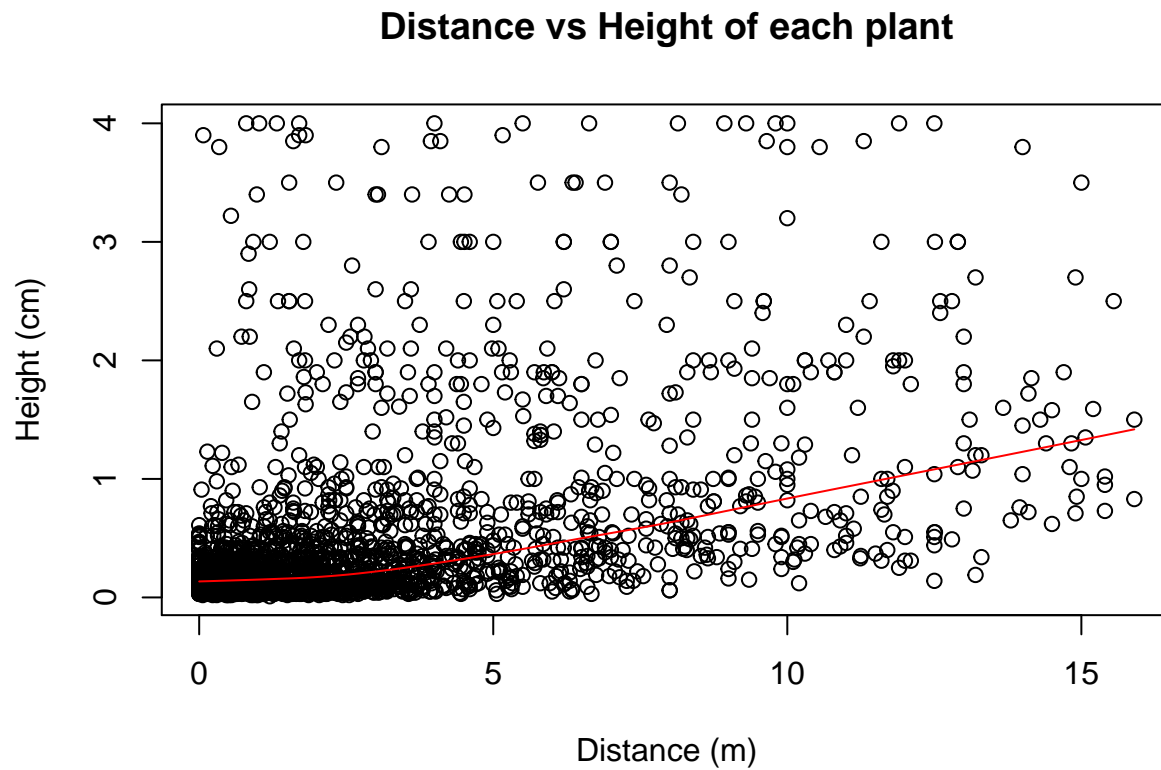


Figure 5 - Distance of each plant to the center line against its height

3.6 Distance of each plant to the center line against shrub cover on the transect

```
plot(data$distance,data$shrub_cover,main="Distance vs Shrub cover",
      ylab="shrub cover (% of ground covered)",xlab="distance (m)")
lines(lowess(data$distance,data$shrub_cover),col=2)
```

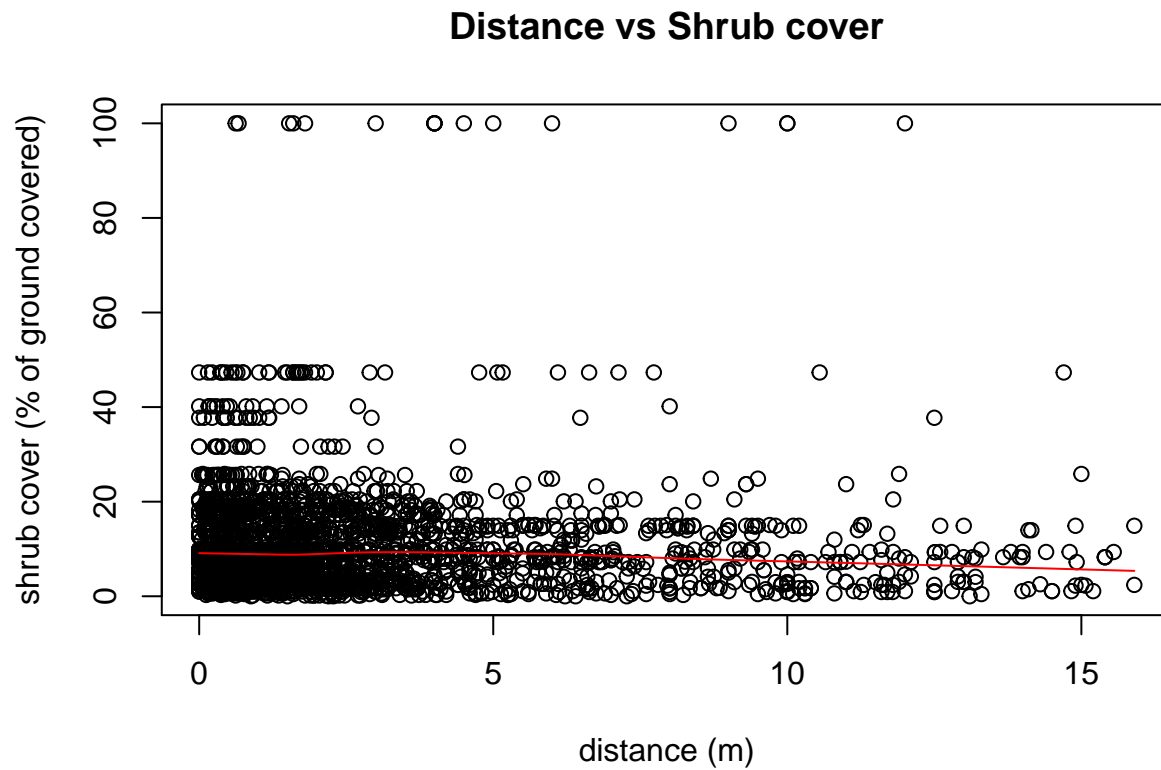


Figure 6 - Distance of each plant to the center line against shrub cover on the transect

3.7 Distance of each plant to the center line against adult tree density on the transect

```
plot(data$distance,data$tree_dens,main="Distance vs Tree density",
      ylab="tree density (trees/ha",xlab="distance (m)")
lines(lowess(data$distance,data$tree_dens),col=2)
```

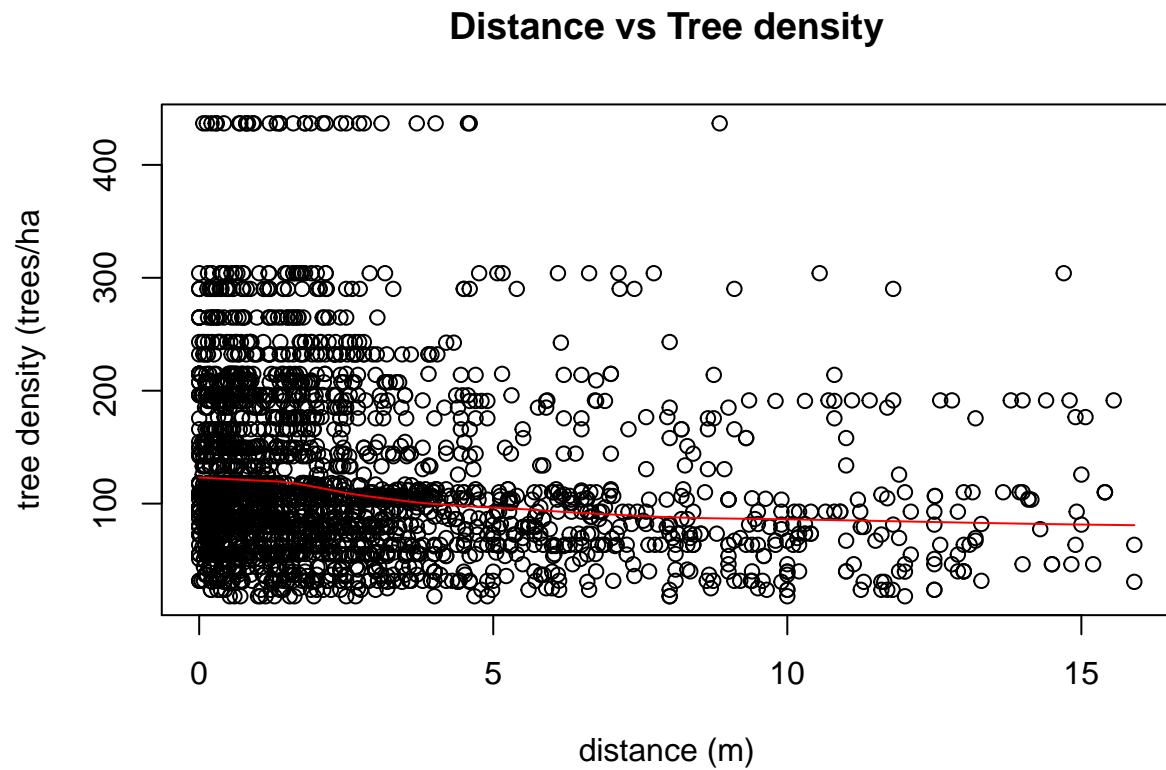


Figure 7 - Distance of each plant to the center line against adult tree density on the transect

3.8 Distance of each plant to the center line against adult tree cover on the transect

```
plot(data$distance,data$tree_cover,main="Distance vs Tree cover",
      ylab="tree cover (m2/ha)",xlab="distance (m)")
lines(lowess(data$distance,data$tree_cover),col=2)
```

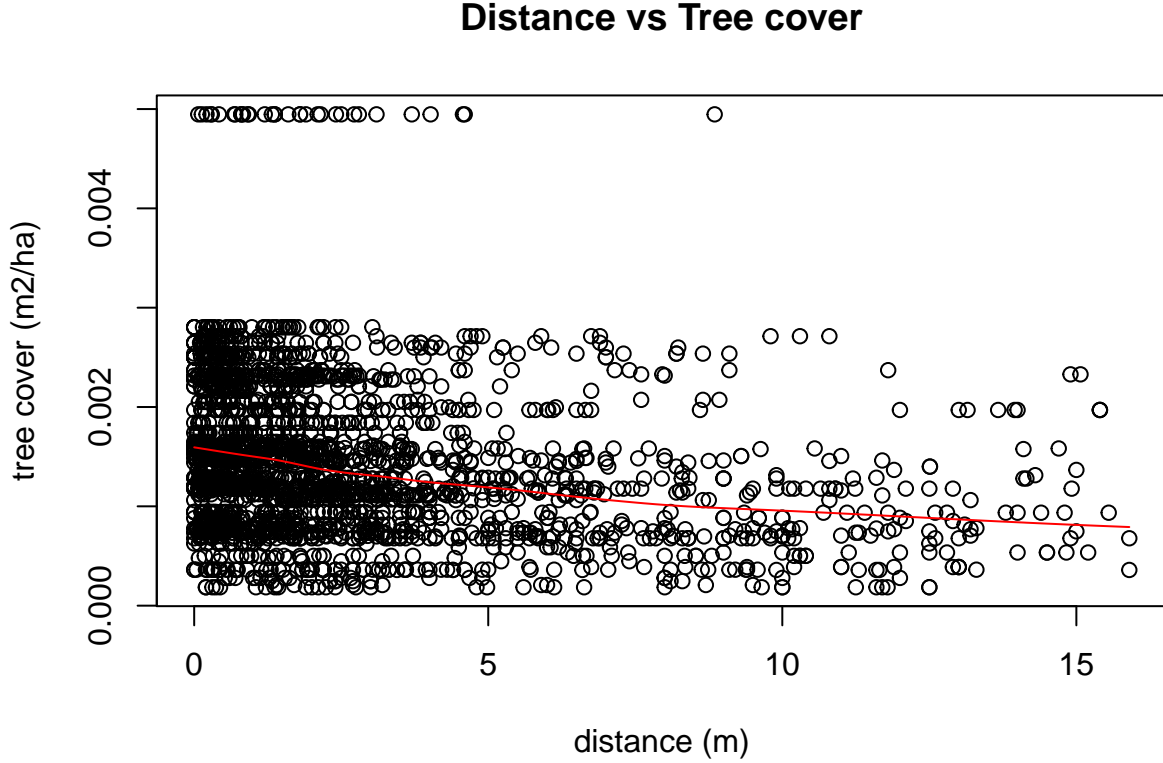



Figure 8 - Distance of each plant to the center line against adult tree cover on the transect

Conclusions

- The distribution of the distances is similar in conservation and non-conservation zones, as well as in seven surveyed estates (Fig. 2). This is not the case when we look at the distribution of the distances split by zone (Fig. 3),
- Plants detected further away from the transect plant tend to be taller, which suggests this variable can affect detectability (Fig. 5).
- The detection of plants far from the transect seems to be higher in areas with lower shrub cover, tree cover and tree density (Fig 6-8).

4. Selecting the detection function

4.1 Candidate detection functions

Uniform, half-normal and hazard-rate detection functions (with cosine, Hermite polynomial and simple polynomial adjustment terms) were fitted to the data. Along with the adjustment terms, tree height, shrub cover, tree cover and zone code (27 level factor) were included as covariates (one or two variables at a time) as they may affect plant detectability.

Below are the top rated candidate functions that are presented in Table 2. Based on a visual inspection of the histograms we decided on a truncation distance of 6 meters.

```

df<-ds(data,truncation=6, key="hr",formula =~height+zone_code,monotonicity="strict")
df1<-ds(data,truncation=6, key="hn",formula =~height+zone_code,monotonicity="strict")
df2<-ds(data,truncation=6, key="hr",formula =~height+shrub_cover,monotonicity="strict")
df3<-ds(data,truncation=6, key="hr",formula =~height,monotonicity="strict")
df4<-ds(data,truncation=6, key="hr",formula =~height+shrub_dens,monotonicity="strict")
df5<-ds(data,truncation=6, key="hr",formula =~height+tree_dens,monotonicity="strict")
df6<-ds(data,truncation=6, key="hn",formula =~height+tree_cover,monotonicity="strict")
df7<-ds(data,truncation=6, key="hn",formula =~height,monotonicity="strict")
df8<-ds(data,truncation=6, key="hn",formula =~height+shrub_dens,monotonicity="strict")
df9<-ds(data,truncation=6, key="hr",formula =~zone_code,monotonicity="strict")
df10<-ds(data,truncation=6, key="hr",formula =~tree_cover,monotonicity="strict")
df11<-ds(data, truncation=6, key = "hr",adjustment="poly",monotonicity="strict")
df12<-ds(data,truncation=6, key="hn",formula =~zone_code,monotonicity="strict")
df13<-ds(data,truncation=6, key="hr",formula =~shrub_cover,monotonicity="strict")
df14<-ds(data,truncation=6, key="hr",formula =~shrub_density,monotonicity="strict")
df15<-ds(data,truncation=6, key="hn",adjustment="cos",order=(2,3),monotonicity="strict")
df16<-ds(data,truncation=6, key="hr",formula =~tree_cover,monotonicity="strict")
df17<-ds(data,truncation=6, key="hr",formula =~tree_cover,monotonicity="strict")
df18<-ds(data,truncation=6, key="hn",adjustment="poly",monotonicity="strict")
df19<-ds(data,truncation=6, key="hn",formula =~shrub_density,monotonicity="strict")

```

Due to convergence problems we were unable to fit candidate detection functions with more than two covariates. For instance, the two following detection functions failed to fit:

```

#Two covariates and adjustment terms

dfx1<-ds(data,truncation=6, key="hr",formula =~height+zone_code,
          monotonicity="strict",adjustment="cos",order=2)

#Three covariates:

dfx2<-ds(data,truncation=6, key="hr",formula =~height+zone_code + shrub_cover,
          monotonicity="strict")

```

The hazard rate detection function with plant height and zone code as covariates had the lowest AIC and

was selected.

```
df<-ds(data,truncation=6, key="hr",formula =~height+zone_code,monotonicity="strict")
```

4.2 Adjusting the truncation distance

To the select the truncation distance we:

1. fit the selected detection function with several truncation distances

```
df5<-ds(data,truncation=5, key="hr",formula =~height+zone_code)
df5.5<-ds(data,truncation=5.5, key="hr",formula =~height+zone_code)
df<-ds(data,truncation=6, key="hr",formula =~height+zone_code)
df6.5<-ds(data,truncation=6.5, key="hr",formula =~height+zone_code)
df7<-ds(data,truncation=7, key="hr",formula =~height+zone_code)
```

2. Plot the detection functions superimposed on histograms with the distances

```
par(mfrow=c(3,2),pin=c(0.9,0.9))
plot(df5,showpoints=F,main="Truncation = 5 m")
plot(df5.5,showpoints=F,main="Truncation = 5.5 m")
plot(df,showpoints=F,main="Truncation = 6 m")
plot(df6.5,showpoints=F,main="Truncation = 6.5 m")
plot(df7,showpoints=F,main="Truncation = 7 m")
par(mfrow=c(1,1))
```

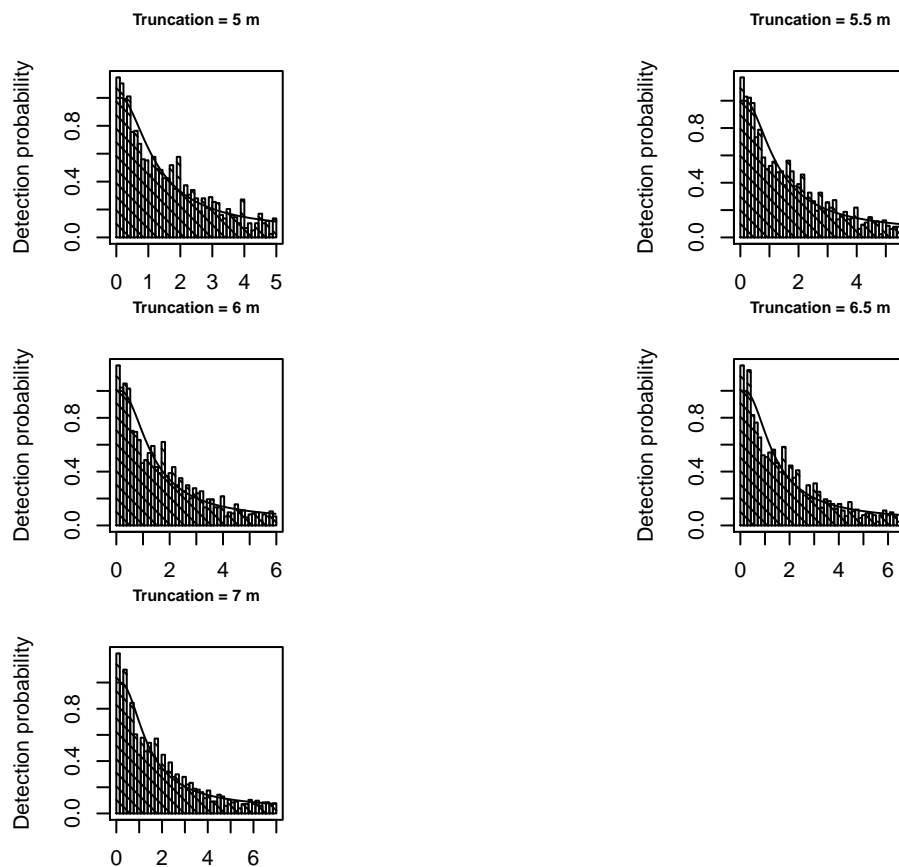


Figure 9 - Detection functions superimposed on histograms.

4. Perform goodness of fit tests

We use `ddf.gof` to perform goodness of fit tests on the detection functions with different truncation distances

```
gof5 <- ddf.gof(df5$ddf)
gof5.5 <- ddf.gof(df5.5$ddf)
gof6 <- ddf.gof(df6$ddf)
gof6.5 <- ddf.gof(df6.5$ddf)
gof7 <- ddf.gof(df7$ddf)

gof_results<-rbind(rbind(gof5$dsgof$ks, gof5$dsgof$CvM), rbind(gof5.5$dsgof$ks, gof5.5$dsgof$CvM),
  rbind(gof6$dsgof$ks, gof6$dsgof$CvM),
  rbind(gof6.5$dsgof$ks, gof6.5$dsgof$CvM),
  rbind(gof7$dsgof$ks, gof7$dsgof$CvM))

gof_results<-cbind(Truncation=rep(c("5m", "5.5m", "6m", "6.5m", "7m"),each=2),
  gof_results)
  row.names(gof_results) <- rep(c("Kolmogorov-Smirnov", "Cramer-von Mises"),5)
kable(gof_results) #Print the results in a table
```

	Truncation	Dn	p
Kolmogorov-Smirnov	5m	0.03548226	0.01296718
Cramer-von Mises	5m	0.2560784	0.1807619
Kolmogorov-Smirnov	5.5m	0.03480392	0.01427836
Cramer-von Mises	5.5m	0.2648459	0.1704052
Kolmogorov-Smirnov	6m	0.03418392	0.01559243
Cramer-von Mises	6m	0.2717049	0.1627775
Kolmogorov-Smirnov	6.5m	0.03349057	0.01720571
Cramer-von Mises	6.5m	0.2814218	0.1526313
Kolmogorov-Smirnov	7m	0.03294664	0.01858732
Cramer-von Mises	7m	0.3053172	0.1305983

The detection function with a truncation distance of 6 meters presented the best fit as judged by the combined results of the Distance sampling Kolmogorov-Smirnov test and Distance sampling Cramer-von Mises test(unweighted).

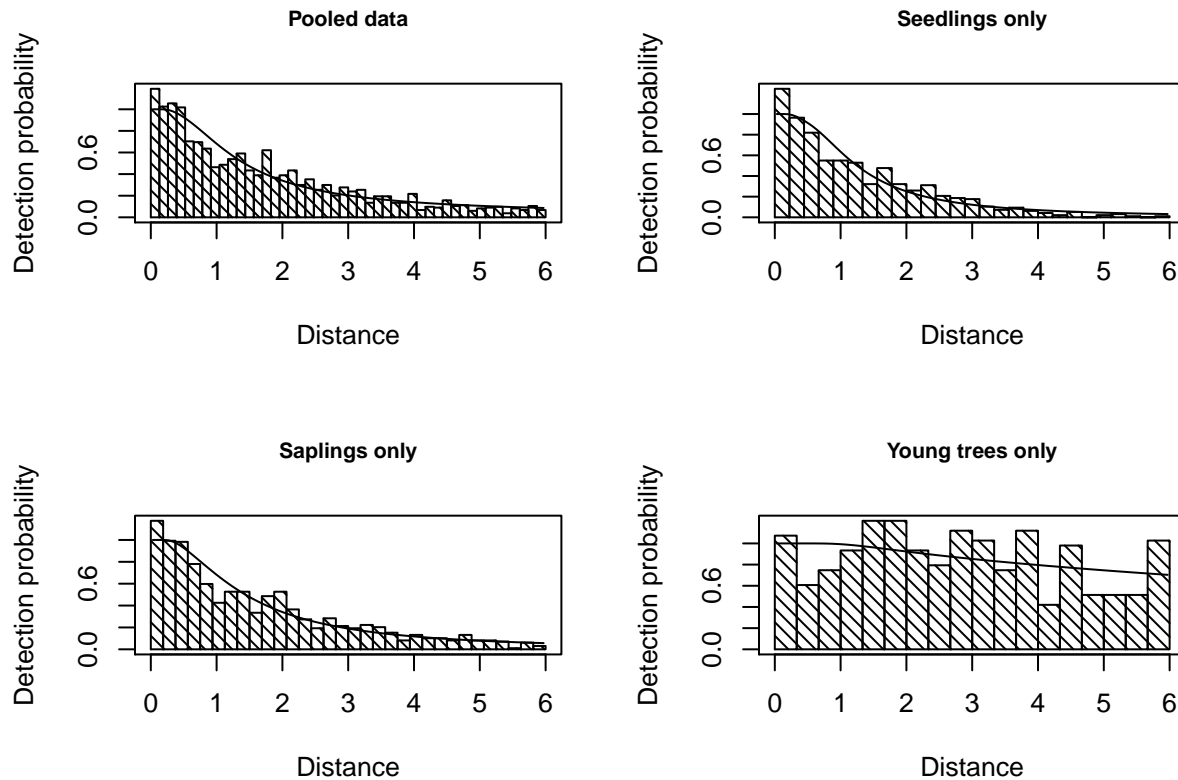
4.3 Summary and plot of the selected detection function

```
summary(df)

##
## Summary for distance analysis
## Number of observations : 2077
## Distance range       : 0 - 6
##
## Model : Hazard-rate key function
## AIC   : 5745.999
##
## Detection function parameters
```

```
## Scale Coefficients:
##               estimate      se
## (Intercept)    0.28829993 0.1305540
## height         3.12576918 0.2827719
## zone_codeArraoZNC1 -0.75144137 0.1986340
## zone_codeCanZC1   -0.28147234 0.1831962
## zone_codeCanZNC1  -0.42798658 0.1723370
## zone_codeCascZC1  -0.30245360 0.1833967
## zone_codeCascZNC1 -0.07794575 0.2212604
## zone_codeCavZC1   -0.96871520 0.1651929
## zone_codeCavZC2   -1.42893080 0.1915214
## zone_codeCavZNC1  -1.68058822 0.2857942
## zone_codeFidZC1   -0.18034662 0.1801708
## zone_codeFidZC2   -0.04464589 0.2516107
## zone_codeFidZC3   -1.26762370 0.2702632
## zone_codeFidZC4   -0.98056589 0.2336511
## zone_codeFidZC5   -0.40298227 0.2933816
## zone_codeFidZNC1    0.39675578 0.2991536
## zone_codeFidZNC2  -0.27631993 0.3048113
## zone_codeFidZNC3  -0.65772342 0.2924730
## zone_codeFidZNC4    0.25280635 0.5615240
## zone_codeFidZNC5  -0.36729274 0.4493595
## zone_codeMacZC1   -0.30368309 0.1817123
## zone_codeMacZC2   -0.05289528 0.1928943
## zone_codeMacZNC1  -0.71011169 0.1778759
## zone_codeMacZNC2  -0.51296516 0.2287680
## zone_codeOnzZC1   -0.54774756 0.1833473
## zone_codeOnzZNC1  -0.20401672 0.2113198
## zone_codePerZC1   -1.42381878 0.1830866
## zone_codePerZNC1  -0.38515161 0.1998909
##
## Shape parameters:
##               estimate      se
## (Intercept) 0.7345699 0.04785475
##
##               Estimate      SE      CV
## Average p      0.3377865    0.01081018 0.03200299
## N in covered region 6148.8548204 229.72578863 0.03736074
```

```
par(mfrow=c(2,2))
plot(df,showpoints=F,main="Pooled data")
plot(df,main="Seedlings only",showpoints=F,subset=height<=0.10)
plot(df,main="Saplings only",showpoints=F,subset=height>0.10&height<=0.5)
plot(df,main="Young trees only",showpoints=F,subset=height>0.5)
```



```
par(mfrow=c(1,1))
```

Figure 10 - Selected detection superimposed on histograms with the distances of the pooled data, seedlings, saplings and young trees.

5. Density surface models

5.1 Dataset “variables”

This `data.frame` includes transect-level covariates that will be used in the density surface models.

1. `transect` - unique identifier for each surveyed transect
2. `zone` - unique identifier for each conservation and non-conservation zones
3. `zone_type` - conservation (`cz`) or non-conservation zone (`ncz`)
4. `estate` - surveyed estate
5. `ado_sb` - cork oak cover (m²/ha)
6. `do_roffici` - cover of *Rosmarinus officinalis* (% of cover)
7. `do_csalvi` - cover of *Cistus salvifolius* (% of cover)
8. `do_ulexsp` - cover of *Ulex* sp (% of cover)
9. `do_ccrispus` - cover of *Cistus crispus* (% of cover)
10. `do_lstoechas` - cover of *Lavandula stoechas* (% of cover)
11. `do_cmonsp` - cover of *Cistus monspelliensis* (% of cover)
12. `do_cvulgaris` - cover of *Calluna vulgaris* (% of cover)
13. `do_hlasi` - cover of *Hallimium lasianthum* (% of cover)
14. `do_ladanifer` - cover of *Cistus ladanifer* (% of cover)

15. `do_dgnidium` - cover of *Daphne gnidium* (% of cover)
16. `do_qlusi` - cover of *Quercus lusitanica* (% of cover)
17. `do_ptrident` - cover of *Pterospartum tridentatum* (% of cover)
18. `do_rper` - cover of *Rubia peregrina* (% of cover)
19. `do_hocym` - cover *Hallimum ocymoides* (% of cover)
20. `do_eaustralis` - cover of *Erica australis* (% of cover)
21. `do_plentistcusc` - cover of *Pistacia lentiscus* (% of cover)
22. `do_mcomm` - cover of *Myrtus communis* (% of cover)
23. `diversity` - shrub diversity
24. `alt` - altitude (m)
25. `slope` - slope (%)
26. `tri` - topographic ruggedness index
27. `tpi` - topographic position index
28. `aspect` - aspect (degrees)
29. `x` - longitude
30. `y` - latitude
31. `Sample.Label` - unique identifier for each surveyed transect
32. `Region.Label` - unique identifier for each surveyed conservation and non-conservation zone
33. `Effort` - length of the transect
34. `soil_lito` - % of cover by litholic soil
35. `soil_podz` - % of cover by podzol soil

5.2 Data preparation

Create a backup detection function

```
backup<-df
```

Create the objects required for running `dsm()`

Set non-conservation zone as the baseline for the variable “zone_type”

```
variables$zone_type<-as.factor(variables$zone_type)
variables$zone_type<-relevel(variables$zone_type,ref="ncz")
```

Extract seedlings, saplings and young trees observations:

```
obs.table.seed<-subset(data,data$class=="seed" & data$distance<=6)
obs.table.sap<-subset(data,data$class=="sap" & data$distance<=6)
obs.table.yt<-subset(data,data$class=="yt" & data$distance<=6)
```

5.3 Local density surface models

First we check if the local variables are correlated by calculating pairwise correlations and variance inflation factors. We use `cor` “`corvif()`” and “`myvif()`” that were written by (Zuur et al 2007). We consider only species occurring in at least 20% of the transects (see the main text for an explanation).

```
corvif(data.frame(variables$ado_sb,
                  variables$do_cmonsp,
                  variables$do_roffici,
```



```

variables$do_csalvi,
variables$do_ulexsp,
variables$do_ccrispus,
variables$shrub_div))

```

Correlations of the variables

```

##
##
##          variables.ado_sb variables.do_cmonsp
## variables.ado_sb          1.00000000        -0.06620807
## variables.do_cmonsp      -0.06620807          1.00000000
## variables.do_roffici       0.02995396        -0.04761247
## variables.do_csalvi       -0.08035617        -0.06940701
## variables.do_ulexsp        0.17328587        -0.09059289
## variables.do_ccrispus     -0.08384775        -0.04592375
## variables.shrub_div       -0.16919408        -0.12007102
##
##          variables.do_roffici variables.do_csalvi
## variables.ado_sb          0.02995396        -0.080356175
## variables.do_cmonsp      -0.04761247        -0.069407007
## variables.do_roffici       1.00000000          0.226025403
## variables.do_csalvi       0.22602540          1.000000000
## variables.do_ulexsp      -0.08579984        -0.006681319
## variables.do_ccrispus      0.02221107          0.074519862
## variables.shrub_div       0.23316969          0.188840169
##
##          variables.do_ulexsp variables.do_ccrispus
## variables.ado_sb          0.173285868        -0.08384775
## variables.do_cmonsp      -0.090592891        -0.04592375
## variables.do_roffici     -0.085799836          0.02221107
## variables.do_csalvi     -0.006681319          0.07451986
## variables.do_ulexsp       1.000000000        -0.04817708
## variables.do_ccrispus    -0.048177076          1.00000000
## variables.shrub_div       0.168523422          0.24781005
##
##          variables.shrub_div
## variables.ado_sb          -0.1691941
## variables.do_cmonsp      -0.1200710
## variables.do_roffici       0.2331697
## variables.do_csalvi       0.1888402
## variables.do_ulexsp       0.1685234
## variables.do_ccrispus     0.2478101
## variables.shrub_div       1.0000000
##
##

```

Variance inflation factors

```

##
##          GVIF
## variables.ado_sb          1.098461
## variables.do_cmonsp       1.029568
## variables.do_roffici      1.132736
## variables.do_csalvi       1.084671
## variables.do_ulexsp       1.113795
## variables.do_ccrispus     1.080129
## variables.shrub_div       1.263793

```

When we fit a density surface model we use the detection function to estimate the density of seedlings, saplings or young trees on each transect. To do this we need to extract the fitted values corresponding to the age class we are analyzing.

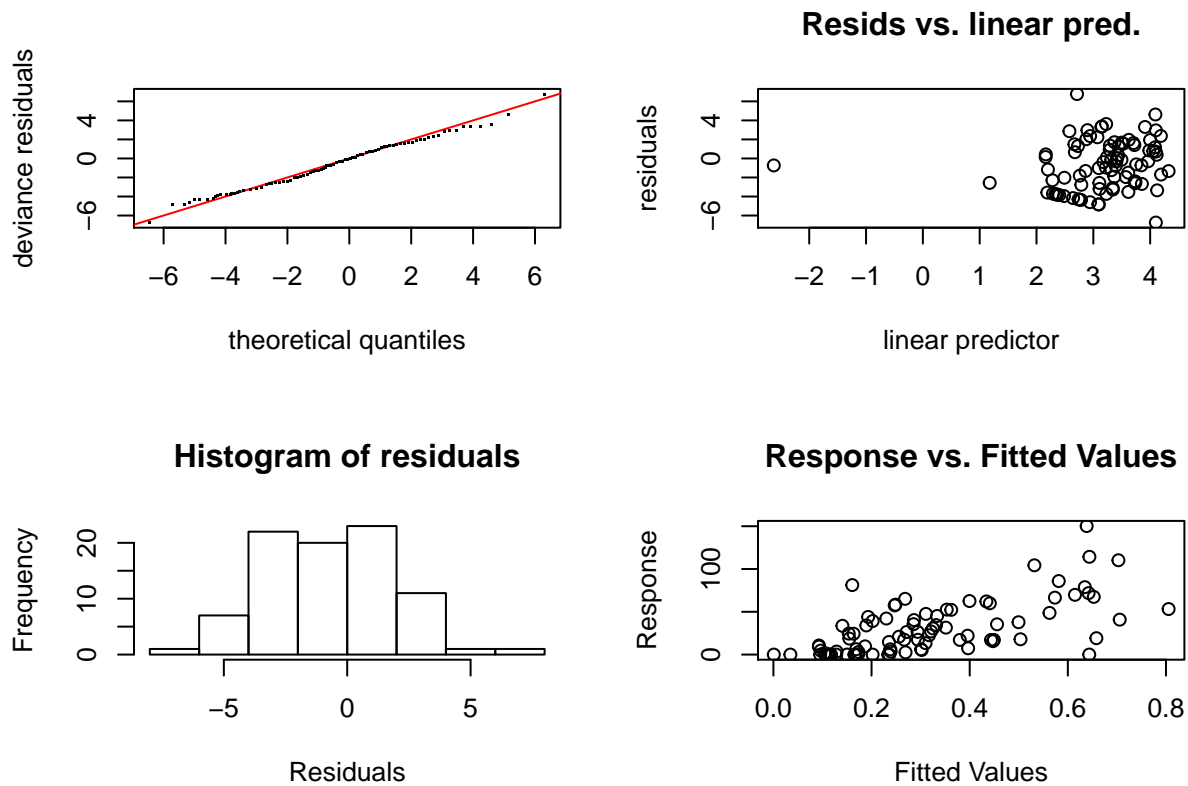
To fit a density surface model with a Tweedie distribution we need to select the value for the parameter “p”. This can be done automatically during fitting by specifying the option `family=tw()`.

Covariates were selected using approximate p-values ($p < 0.05$; Wood, 2006) and an extra penalty for each term was included in the model that allowed smooth terms to be removed during model fitting (option “`select=TRUE`”; (Marra and Wood 2011)).

5.3.1 Seedlings - Local

Tweedie

```
#call the detection function
df<-backup
#Extract the fitted values corresponding to seedlings from the detection function
df$ddf$fitted<- df$ddf$fitted[as.character(subset(data,data$distance<=6 & data$class=="seed")$id)]
#Fit the model
modseed_tw<- dsm(Nhat ~ zone_type+s(ado_sb)+s(estate,bs="re")+
  s(sqrt(do_ulexsp))+s(do_lstoechas)+s(do_cmonsp)+
  s(do_roffici)+s(do_csalvi)+ s(do_ccrispus)+
  s(shrub_div), df, observation.data=obs.table.seed,
  segment.data=variables,engine="gam",family=tw(),
  select=TRUE,method="REML")
gam.check(modseed_tw)
```



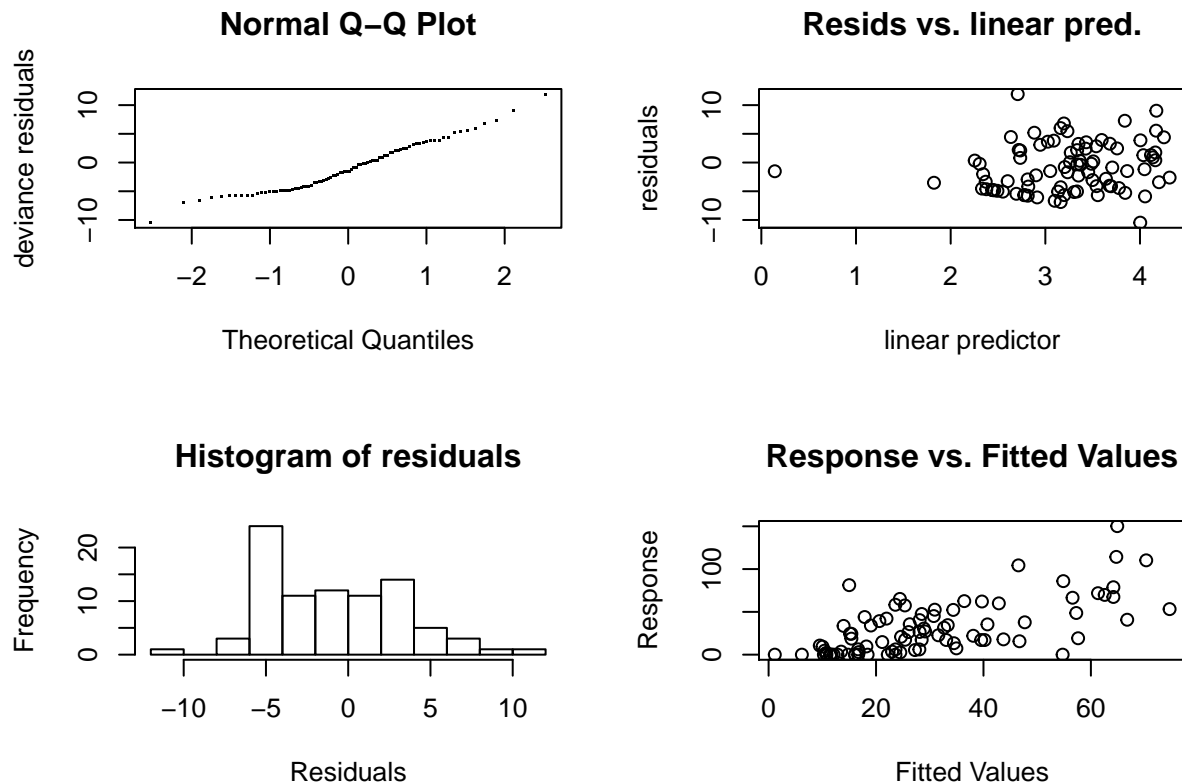
##

```
## Method: REML   Optimizer: outer newton
## full convergence after 14 iterations.
## Gradient range [-4.054614e-05,6.302619e-05]
## (score 349.2044 & scale 6.911073).
## Hessian positive definite, eigenvalue range [6.450639e-07,80.46577].
## Model rank = 82 / 82
##
## Basis dimension (k) checking results. Low p-value (k-index<1) may
## indicate that k is too low, especially if edf is close to k'.
##
##
```

	k'	edf	k-index	p-value
s(ado_sb)	9.00e+00	2.00e+00	7.96e-01	0.07
s(estate)	8.00e+00	4.74e+00	NA	NA
s(sqrt(do_ulexsp))	9.00e+00	3.79e-01	1.02e+00	0.79
s(do_lstoechas)	9.00e+00	4.74e-02	8.42e-01	0.14
s(do_cmonsp)	9.00e+00	7.86e-01	8.86e-01	0.34
s(do_roffici)	9.00e+00	6.98e-05	8.53e-01	0.23
s(do_csalvi)	9.00e+00	5.89e-05	9.37e-01	0.44
s(do_ccrispus)	9.00e+00	1.87e-05	8.59e-01	0.20
s(shrub_div)	9.00e+00	5.70e-05	7.75e-01	0.04

quasi-Poisson

```
df<-backup
df$ddf$fitted<- df$ddf$fitted[as.character(subset(data,data$distance<=6 & data$class=="seed")$id)]
modseed_qpois<- dsm(Nhat ~ zone_type+s(ado_sb)+s(estate,bs="re")+
  s(sqrt(do_ulexsp))+s(do_lstoechas)+s(do_cmonsp)+
  s(do_roffici)+s(do_csalvi)+s(do_ccrispus)+s(shrub_div), df,
  observation.data=obs.table.seed, segment.data=variables,
  engine="gam",family=quasipoisson,select=TRUE,method="REML")
gam.check(modseed_qpois)
```

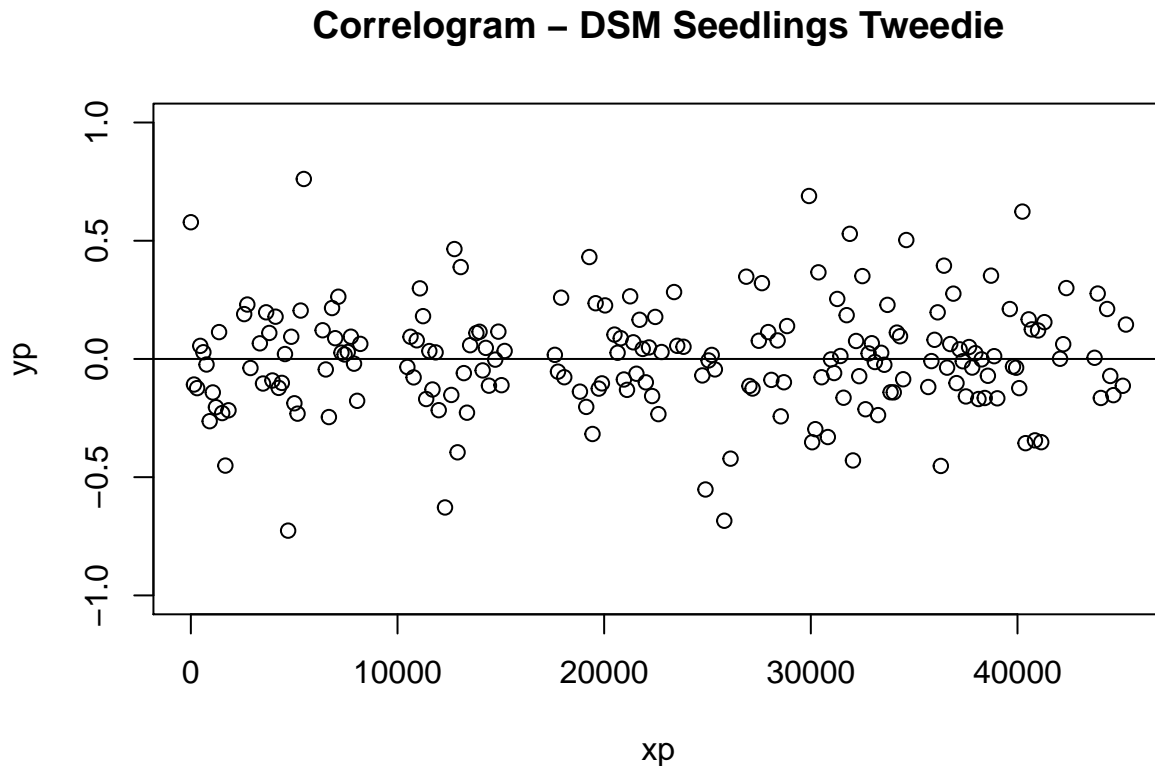


```
##
## Method: REML   Optimizer: outer newton
## full convergence after 12 iterations.
## Gradient range [-2.247222e-05,3.069446e-05]
## (score 182.1582 & scale 19.79854).
## Hessian positive definite, eigenvalue range [3.526147e-07,42.17156].
## Model rank = 82 / 82
##
## Basis dimension (k) checking results. Low p-value (k-index<1) may
## indicate that k is too low, especially if edf is close to k'.
##
##           k'      edf  k-index p-value
## s(ado_sb)    9.00e+00 1.99e+00 8.58e-01 0.10
## s(estate)    8.00e+00 5.03e+00      NA      NA
## s(sqrt(do_ulexsp)) 9.00e+00 5.23e-01 1.07e+00 0.86
## s(do_lstoechas) 9.00e+00 2.89e-05 8.71e-01 0.20
## s(do_cmosp)   9.00e+00 6.66e-01 9.39e-01 0.39
## s(do_roffici) 9.00e+00 5.47e-05 9.27e-01 0.34
## s(do_csalvi)  9.00e+00 3.25e-05 1.01e+00 0.62
## s(do_ccrispus) 9.00e+00 3.47e-05 9.45e-01 0.38
## s(shrub_div)  9.00e+00 4.53e-05 8.40e-01 0.10
```

The Tweedie model presented less patterns in the residuals. Now we check the residuals for spatial autocorrelation by examining a correlogram from a least-squares quadratic trend surface.

```
library(spatial)
dts<-data.frame(x=variables$x,y=variables$y,z=residuals(modseed_tw))
```

```
surface<-surf.ls(2,dts)
correlogram(surface,300,main="Correlogram - DSM Seedlings Tweedie")
```



The points appear to be randomly distributed across $y=0$ which suggests there are no correlation problems in the residuals.

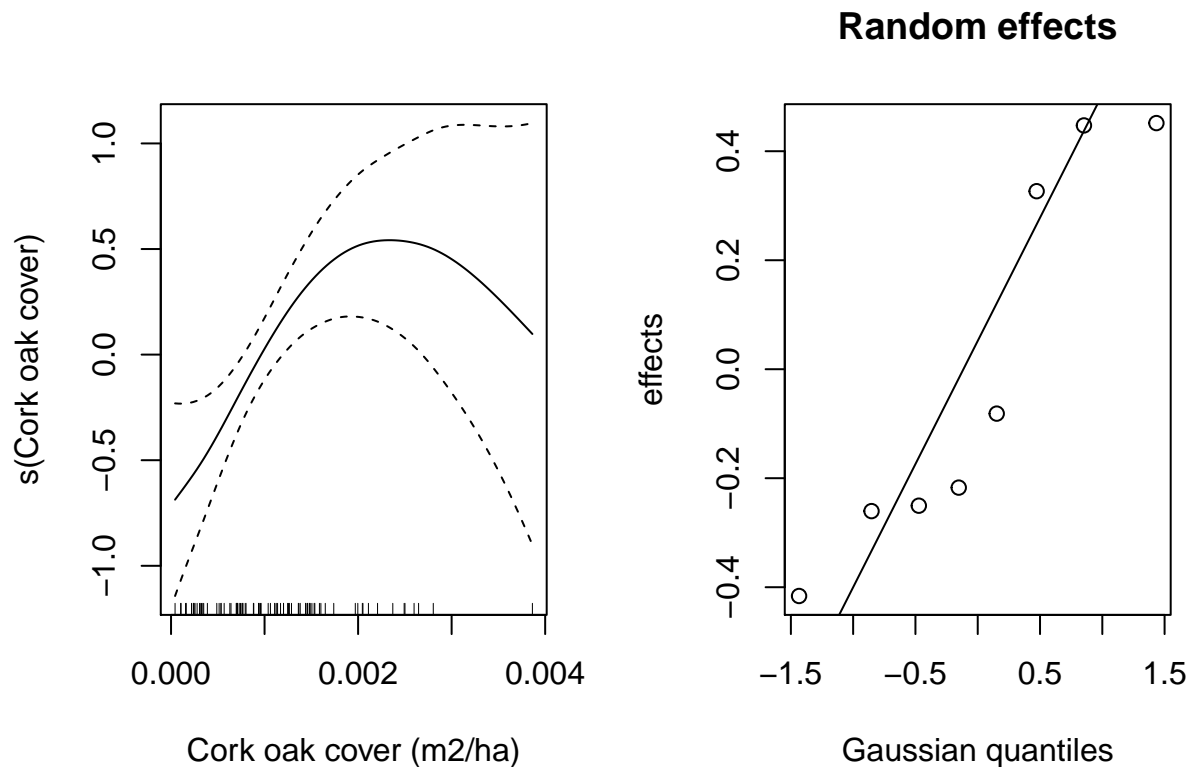
We now take a look at the model and at the selected smoothers.

```
summary(modseed_tw)
```

```
##
## Family: Tweedie(p=1.34)
## Link function: log
##
## Formula:
## Nhat ~ zone_type + s(ado_sb) + s(estate, bs = "re") + s(sqrt(do_ulexsp)) +
##       s(do_lstoechas) + s(do_cmonsp) + s(do_roffici) + s(do_csalvi) +
##       s(do_ccrispus) + s(shrub_div) + offset(off.set)
##
## Parametric coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -3.3503    0.2317  -14.46   <2e-16 ***
## zone_typecz   0.4402    0.1974    2.23   0.0287 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##               edf Ref.df      F p-value
## s(ado_sb)      1.997e+00     9 3.312 0.00143 **
```

```
## s(estate)          4.739e+00      7 2.098 0.00519 **
## s(sqrt(do_ulexsp)) 3.791e-01      9 0.077 0.21657
## s(do_lstoechas)    4.735e-02      9 0.006 0.28535
## s(do_cmonsp)       7.860e-01      8 0.230 0.16179
## s(do_roffici)      6.983e-05      9 0.000 0.91652
## s(do_csalvi)       5.886e-05      9 0.000 1.00000
## s(do_ccrispus)     1.869e-05      9 0.000 0.59482
## s(shrub_div)       5.703e-05      9 0.000 0.81984
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.341   Deviance explained = 37.5%
## -REML = 349.2   Scale est. = 6.9111    n = 86
```

```
#Plot smoothers
par(mfrow=c(1,2))
plot(modseed_tw,select=1,scale=0,ylab="s(Cork oak cover)",xlab="Cork oak cover (m2/ha)")
plot(modseed_tw,select=2,scale=0,main="Random effects")
```



```
par(mfrow=c(1,1))
```

5.3.2 Saplings - Local

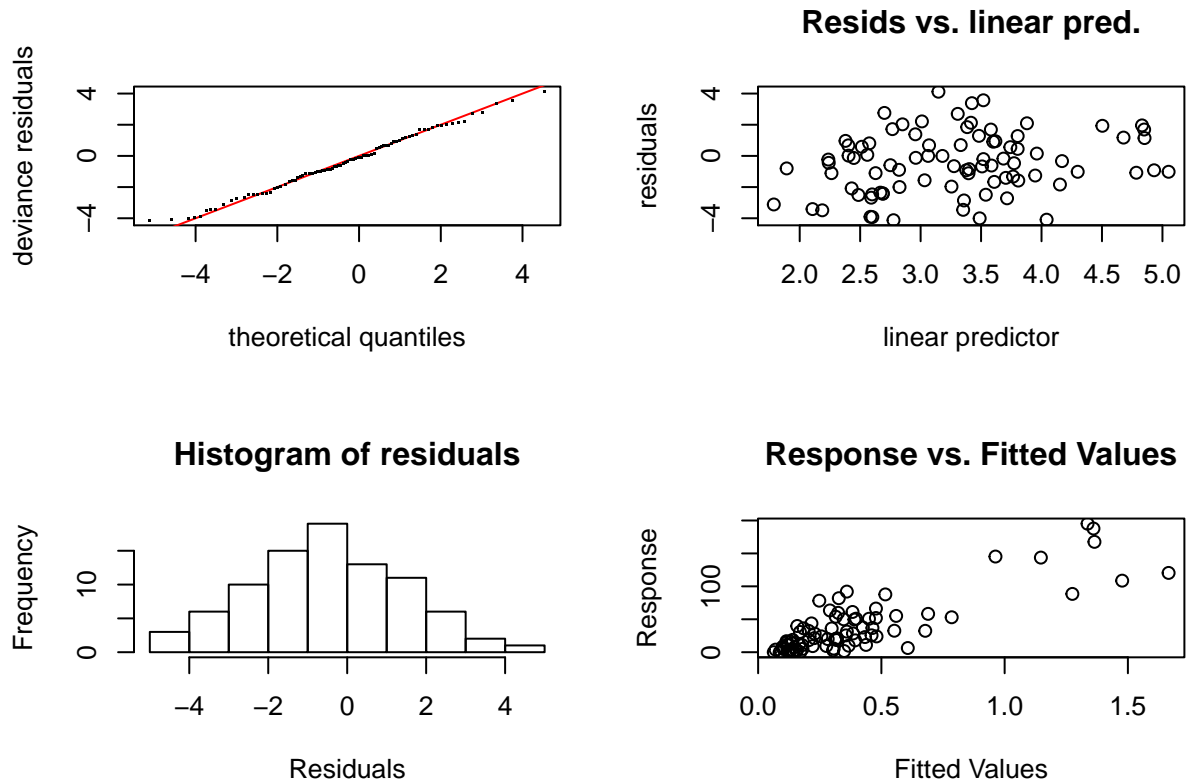
Tweedie

```
df<-backup
df$ddf$fitted<- df$ddf$fitted[as.character(subset(data,data$distance<=6 & data$class=="sap")$id)]
```

```

modsap_tw<- dsm(Nhat ~ zone_type+s(do_ulexsp)+s(do_roffici)+
  s(estate,bs="re")+s(do_cmonsp)+s(do_csalvi)+
  s(ado_sb)+s(do_lstoechas)+s(do_ccrispus)+
  s(shrub_div), df, observation.data=obs.table.sap,
  segment.data=variables,engine="gam",family=tw(),
  select=TRUE,method="REML")
gam.check(modsap_tw)

```



```

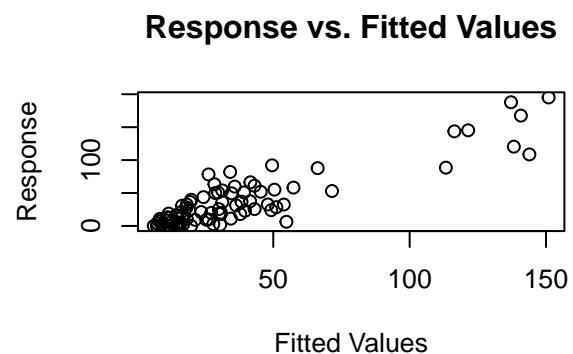
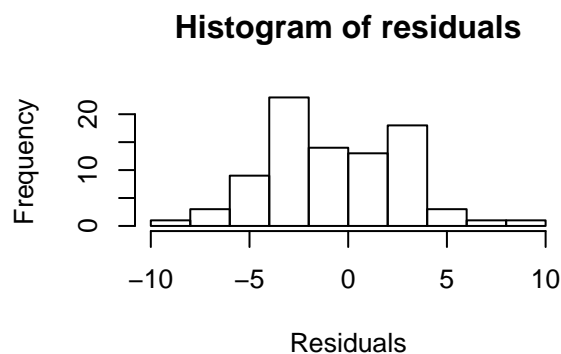
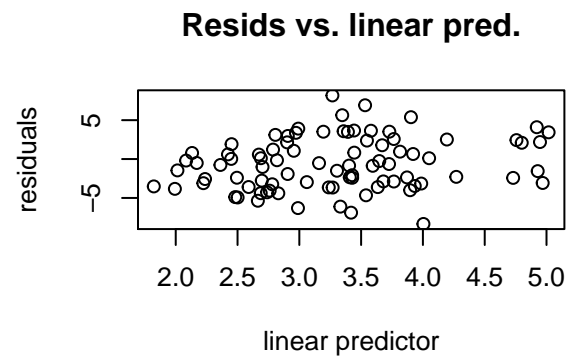
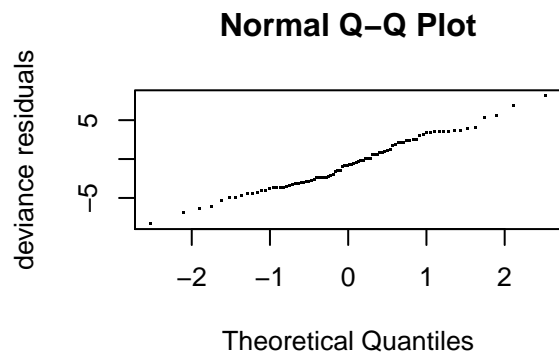
##
## Method: REML   Optimizer: outer newton
## full convergence after 16 iterations.
## Gradient range [-4.453001e-05,8.937978e-05]
## (score 375.8441 & scale 3.766784).
## Hessian positive definite, eigenvalue range [5.057701e-07,80.25762].
## Model rank = 82 / 82
##
## Basis dimension (k) checking results. Low p-value (k-index<1) may
## indicate that k is too low, especially if edf is close to k'.
##
##           k'      edf  k-index p-value
## s(do_ulexsp)  9.00e+00 1.73e+00 9.47e-01  0.47
## s(do_roffici)  9.00e+00 7.68e-01 9.97e-01  0.60
## s(estate)      8.00e+00 5.70e+00      NA    NA
## s(do_cmonsp)   9.00e+00 7.30e-01 1.15e+00  0.99
## s(do_csalvi)   9.00e+00 1.14e+00 1.02e+00  0.76
## s(ado_sb)      9.00e+00 8.59e-01 9.81e-01  0.60
## s(do_lstoechas) 9.00e+00 5.07e-05 9.05e-01  0.34
## s(do_ccrispus) 9.00e+00 2.96e-05 8.38e-01  0.12

```

```
## s(shrub_div)      9.00e+00 1.04e+00 8.90e-01    0.24
```

quasi-Poisson

```
df<-backup
df$ddf$fitted<- df$ddf$fitted[as.character(subset(data,data$distance<=6 & data$class=="sap")$id)]
modsap_qpois<- dsm(Nhat ~ zone_type+s(do_ulexsp)+
  s(do_roffici)+
  s(estate,bs="re")+
  s(do_cmonsp)+s(do_csalvi)+
  s(ado_sb)+
  s(do_lstoechas)+s(do_ccrispus)+s(shrub_div),
  df, observation.data=obs.table.sap, segment.data=variables,engine="gam",
  family=quasipoisson,select=TRUE,method="REML")
gam.check(modsap_qpois)
```



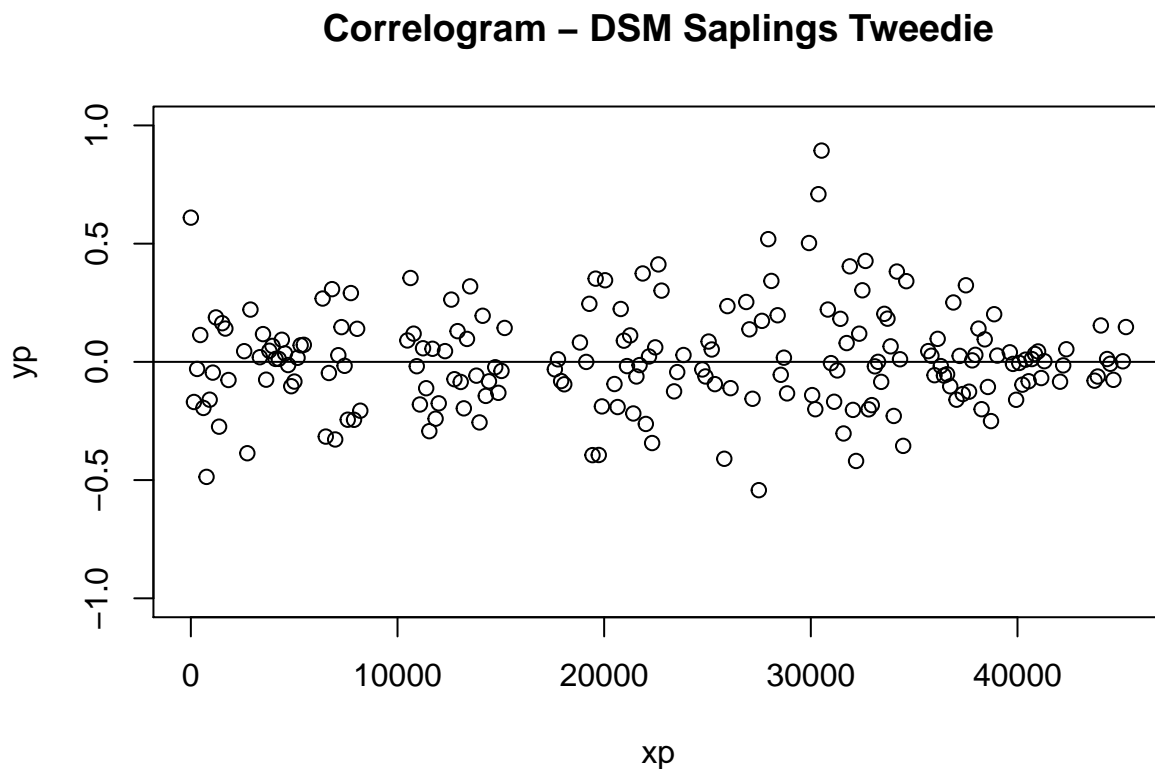
```
##
## Method: REML   Optimizer: outer newton
## full convergence after 13 iterations.
## Gradient range [-2.704343e-05,6.466438e-05]
## (score 178.1558 & scale 12.90889).
## Hessian positive definite, eigenvalue range [1.440588e-06,42.34326].
## Model rank = 82 / 82
##
## Basis dimension (k) checking results. Low p-value (k-index<1) may
## indicate that k is too low, especially if edf is close to k'.
##
##           k'      edf  k-index p-value
```



```
## s(do_ulexsp)      9.00e+00 1.56e+00 9.60e-01    0.44
## s(do_roffici)     9.00e+00 8.03e-01 1.04e+00    0.66
## s(estate)         8.00e+00 5.86e+00      NA      NA
## s(do_cmonsp)      9.00e+00 6.63e-01 1.12e+00    0.90
## s(do_csalvi)      9.00e+00 7.74e-01 9.60e-01    0.48
## s(ado_sb)         9.00e+00 1.47e+00 9.78e-01    0.48
## s(do_lstoechas)   9.00e+00 1.48e-04 9.39e-01    0.32
## s(do_ccrispus)    9.00e+00 2.07e-05 8.71e-01    0.19
## s(shrub_div)      9.00e+00 4.31e+00 1.01e+00    0.62
```

The residuals of the Tweedie model presented less patterns than the ones from the quasi-Poisson model. Now we check for spatial autocorrelation.

```
dts<-data.frame(x=variables$x,y=variables$y,z=residuals(modsap_tw))
surface<-surf.ls(2,dts)
correlogram(surface,300,main="Correlogram - DSM Saplings Tweedie")
```



We see no major problems.

We now take a look at the model and plot the smoothers.

```
summary(modsap_tw)
```

```
##
## Family: Tweedie(p=1.433)
## Link function: log
##
## Formula:
## Nhat ~ zone_type + s(do_ulexsp) + s(do_roffici) + s(estate, bs = "re") +
```

```

##      s(do_cmonsp) + s(do_csalvi) + s(ado_sb) + s(do_lstoechas) +
##      s(do_ccrispus) + s(shrub_div) + offset(off.set)
##
## Parametric coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -3.1665      0.2493 -12.703  <2e-16 ***
## zone_typeecz   0.4104      0.1796   2.285   0.0253 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##              edf Ref.df      F  p-value
## s(do_ulexsp)    1.732e+00     9 2.044  0.00928 **
## s(do_roffici)    7.683e-01     9 0.512  0.04543 *
## s(estate)       5.698e+00     7 5.369 1.43e-06 ***
## s(do_cmonsp)    7.303e-01     9 0.284  0.07935 .
## s(do_csalvi)    1.143e+00     9 0.471  0.06772 .
## s(ado_sb)       8.593e-01     9 0.521  0.08453 .
## s(do_lstoechas) 5.074e-05     9 0.000  0.63504
## s(do_ccrispus)  2.959e-05     9 0.000  0.99698
## s(shrub_div)    1.040e+00     9 0.306  0.16119
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.64   Deviance explained = 58.4%
## -REML = 375.84   Scale est. = 3.7668    n = 86

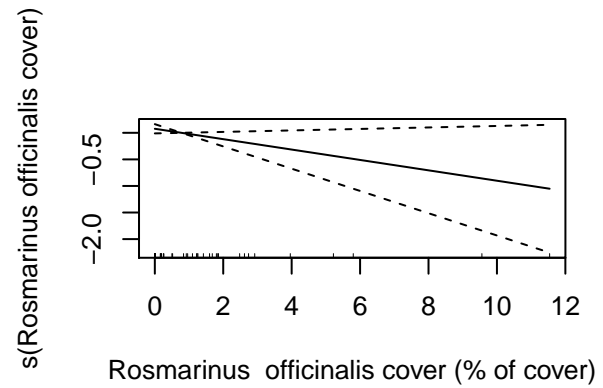
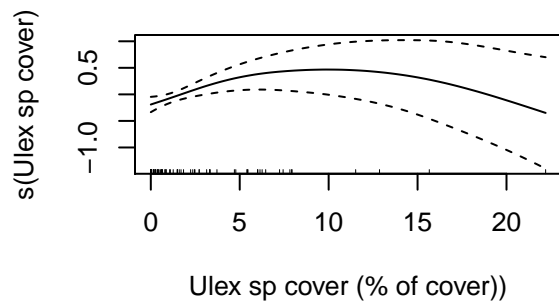
```

#Plot smoothers

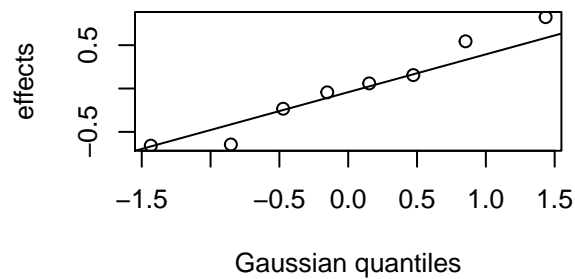
```

par(mfrow=c(2,2))
plot(modsap_tw,select=1,scale=0,ylab="s(Ulex sp cover)",xlab="Ulex sp cover (% of cover)")
plot(modsap_tw,select=2,scale=0,ylab="s(Rosmarinus officinalis cover)",xlab="Rosmarinus officinalis cover (% of cover)")
plot(modsap_tw,select=3,scale=0,main="Random effects")
par(mfrow=c(1,1))

```



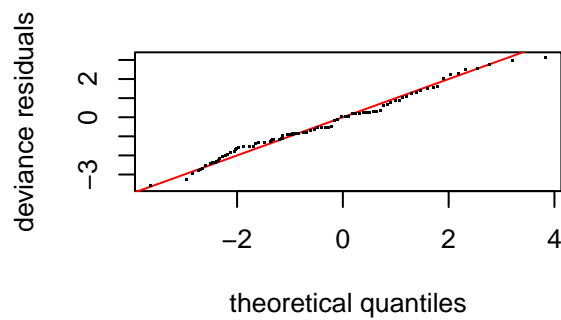
Random effects



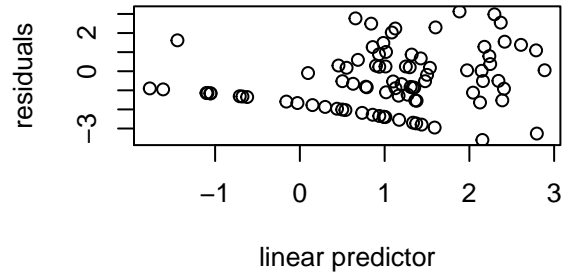
5.3.2 Young trees - Local

Tweedie

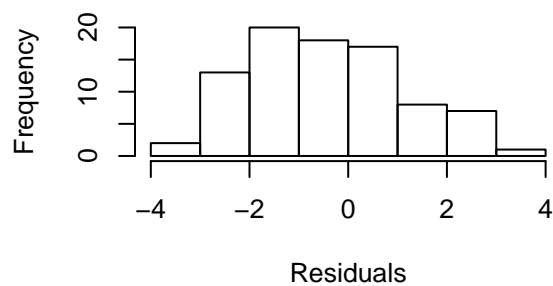
```
#Tweedie
df<-backup
df$ddf$fitted<- df$ddf$fitted[as.character(subset(data,data$distance<=6 & data$class=="yt")$id)]
modyt_tw<- dsm(Nhat ~ zone_type+s(do_ulexsp)+s(do_roffici)+
               s(estate,bs="re")+s(do_cmonsp)+s(do_csalvi)+
               s(ado_sb)+s(do_lstoechas)+s(do_ccrispus)+
               s(shrub_div), df, observation.data=obs.table.yt,
               segment.data=variables,engine="gam",family=tw(),
               select=TRUE,method="REML")
gam.check(modyt_tw)
```



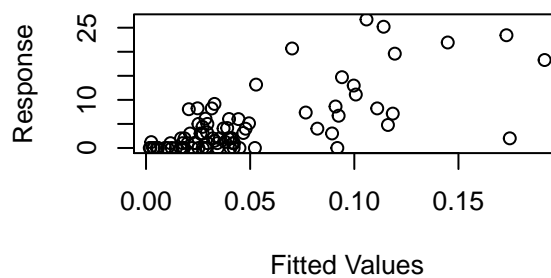
Resids vs. linear pred.



Histogram of residuals



Response vs. Fitted Values



```
##
## Method: REML   Optimizer: outer newton
## full convergence after 25 iterations.
## Gradient range [-5.155678e-05,0.000288191]
## (score 208.6083 & scale 2.621433).
## Hessian positive definite, eigenvalue range [3.464147e-06,82.80138].
## Model rank = 82 / 82
##
## Basis dimension (k) checking results. Low p-value (k-index<1) may
## indicate that k is too low, especially if edf is close to k'.
##
##           k'      edf  k-index p-value
## s(do_ulexsp)  9.00e+00 7.72e-05 9.11e-01 0.44
## s(do_roffici)  9.00e+00 4.49e-04 8.80e-01 0.30
## s(estate)     8.00e+00 5.75e+00      NA    NA
## s(do_cmonsp)  9.00e+00 4.36e-01 9.58e-01 0.61
## s(do_csalvi)  9.00e+00 7.35e-05 7.56e-01 0.04
## s(ado_sb)     9.00e+00 1.91e+00 8.95e-01 0.38
## s(do_lstoechas) 9.00e+00 7.24e-01 9.72e-01 0.72
## s(do_ccrispus) 9.00e+00 1.63e+00 8.04e-01 0.10
## s(shrub_div)  9.00e+00 8.41e-01 9.00e-01 0.42
```

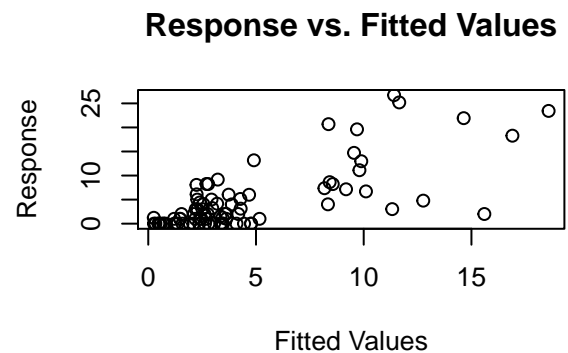
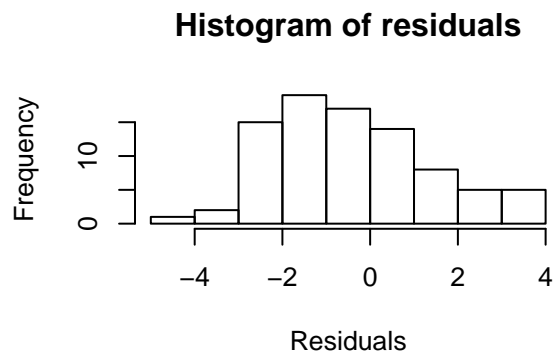
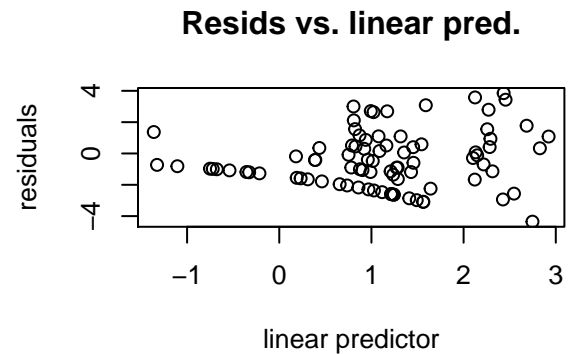
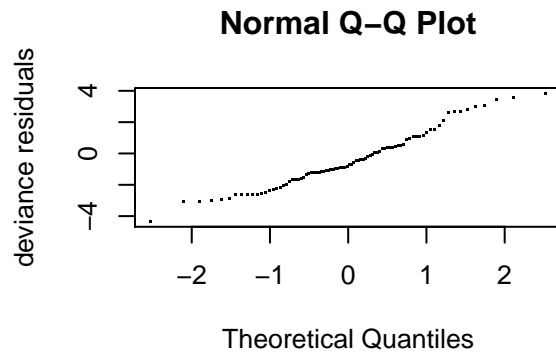
quasi-Poisson

```
df<-backup
df$ddf$fitted<- df$ddf$fitted[as.character(subset(data,data$distance<=6 & data$class=="yt")$id)]
modyt_qpois<- dsm(Nhat ~ zone_type+s(do_ulexsp)+
                  s(do_roffici)+
```

```

s(estate,bs="re")+
s(do_cmonsp)+s(do_csalvi)+
s(ado_sb)+
s(do_lstoechas)+s(do_ccrispus)+s(shrub_div),
df, observation.data=obs.table.yt, segment.data=variables,engine="gam",
family=quasipoisson,select=TRUE,method="REML")
gam.check(modyt_qpois)

```



```

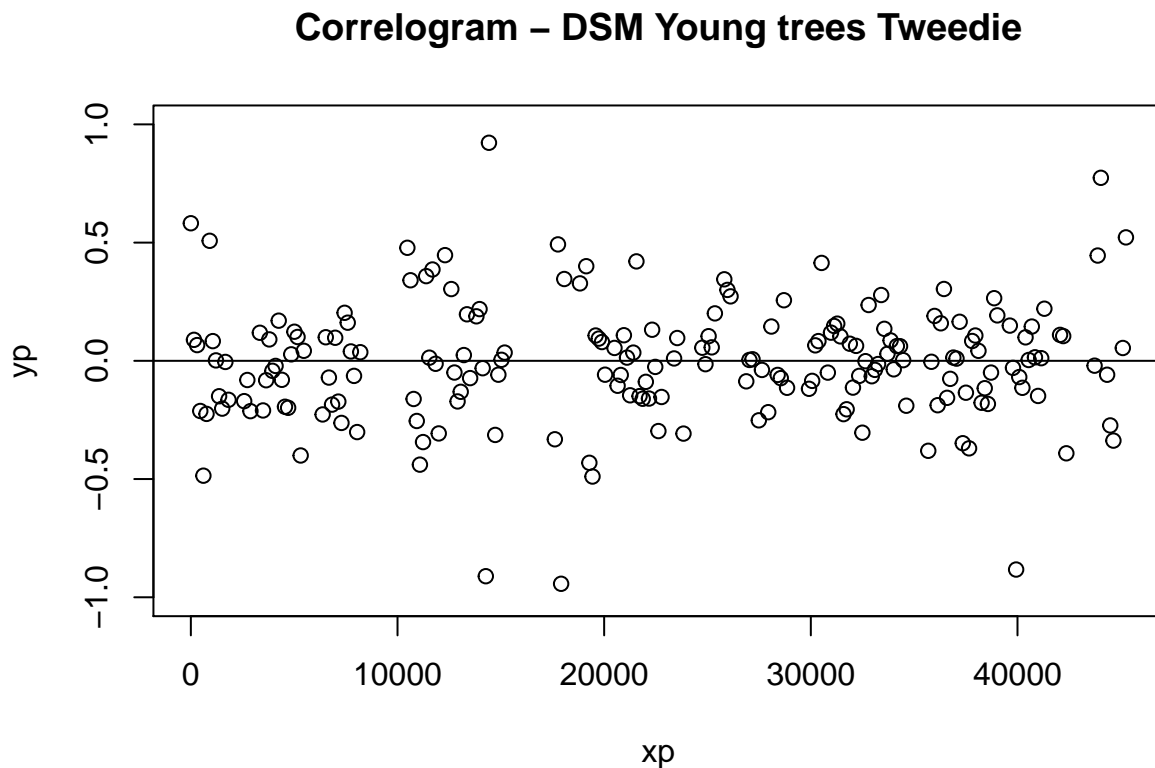
##
## Method: REML   Optimizer: outer newton
## full convergence after 23 iterations.
## Gradient range [-5.969375e-05,5.882648e-05]
## (score 116.0962 & scale 3.547451).
## Hessian positive definite, eigenvalue range [5.069988e-07,42.18589].
## Model rank = 82 / 82
##
## Basis dimension (k) checking results. Low p-value (k-index<1) may
## indicate that k is too low, especially if edf is close to k'.
##
##           k'      edf  k-index p-value
## s(do_ulexsp)  9.00e+00 1.70e-05 9.70e-01 0.52
## s(do_roffici)  9.00e+00 2.04e-01 9.27e-01 0.35
## s(estate)      8.00e+00 5.59e+00      NA    NA
## s(do_cmonsp)   9.00e+00 5.23e-01 1.00e+00 0.65
## s(do_csalvi)   9.00e+00 1.53e+00 8.40e-01 0.08
## s(ado_sb)      9.00e+00 1.92e+00 9.40e-01 0.37
## s(do_lstoechas) 9.00e+00 7.07e-01 1.02e+00 0.68
## s(do_ccrispus) 9.00e+00 1.73e-05 8.28e-01 0.12

```

```
## s(shrub_div)      9.00e+00 8.73e-01 9.04e-01    0.32
```

The residuals of the Tweedie model showed no major issues, unlike the ones from the quasi-Poisson model. Now we check for spatial autocorrelation.

```
dts<-data.frame(x=variables$x,y=variables$y,z=residuals(modyt_tw))
surface<-surf.ls(2,dts)
correlogram(surface,300,main="Correlogram - DSM Young trees Tweedie")
```



We see no major problems.

We now take a look at the model and plot the smoothers.

```
summary(modyt_tw)
```

```
##
## Family: Tweedie(p=1.293)
## Link function: log
##
## Formula:
## Nhat ~ zone_type + s(do_ulexsp) + s(do_roffici) + s(estate, bs = "re") +
##       s(do_cmonsp) + s(do_csalvi) + s(ado_sb) + s(do_lstoechas) +
##       s(do_ccrispus) + s(shrub_div) + offset(off.set)
##
## Parametric coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -5.45093    0.45965 -11.859  <2e-16 ***
## zone_typecz -0.02396    0.24743  -0.097   0.923
## ---
```

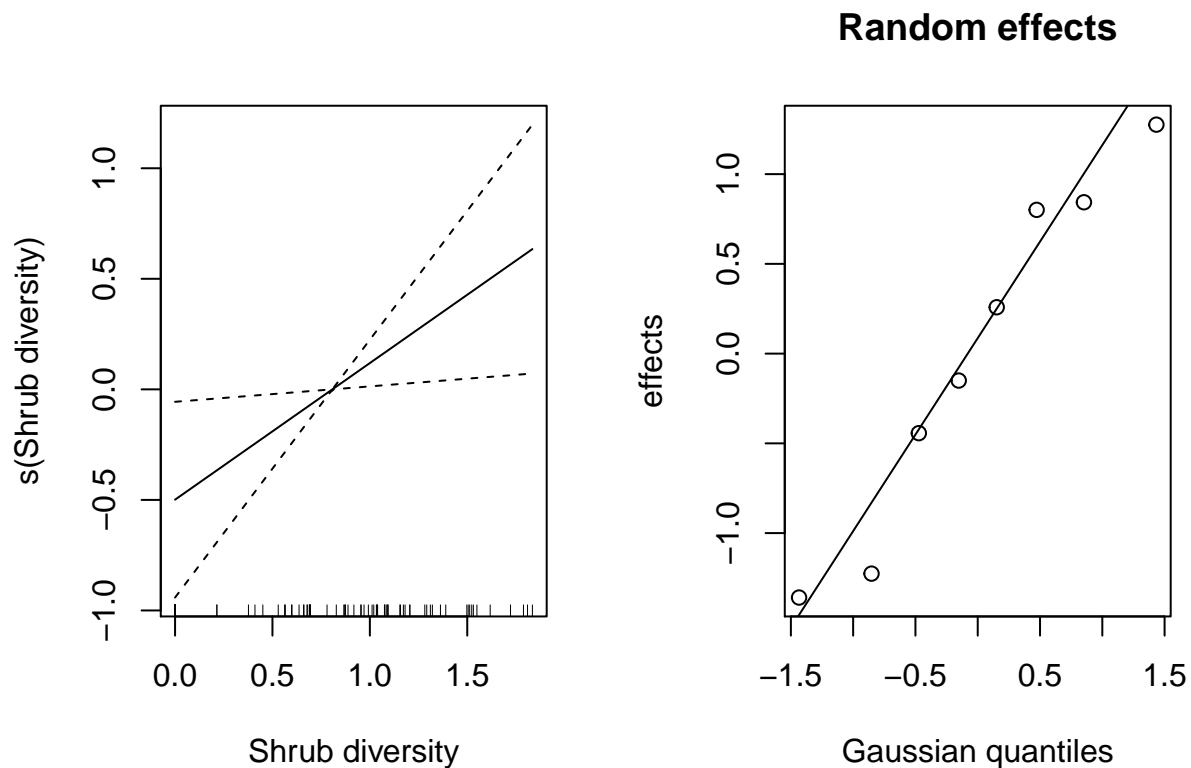
```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##              edf Ref.df    F  p-value
## s(do_ulexsp)   7.718e-05    9 0.000   0.4132
## s(do_roffici)  4.494e-04    9 0.000   0.3702
## s(estate)      5.746e+00    7 5.170 5.07e-06 ***
## s(do_cmonsp)   4.358e-01    9 0.124   0.1680
## s(do_csalvi)   7.350e-05    9 0.000   0.4480
## s(ado_sb)      1.906e+00    9 1.575   0.0701 .
## s(do_lstoechas) 7.242e-01    9 0.455   0.0738 .
## s(do_ccrispus) 1.631e+00    9 0.754   0.0795 .
## s(shrub_div)   8.408e-01    9 2.335   0.0135 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.408   Deviance explained =   50%
## -REML = 208.61   Scale est. = 2.6214    n = 86
```

```
#Plot smoothers
```

```
par(mfrow=c(1,2))
```

```
plot(modyt_tw,select=9,scale=0,ylab="s(Shrub diversity)",xlab="Shrub diversity")
```

```
plot(modyt_tw,select=3,scale=0,main="Random effects")
```



```
par(mfrow=c(1,1))
```

5.4 Topographic density surface models

First it's necessary to check if the topographic variables are correlated by calculating pairwise correlations and variance inflation factors.

```
corvif(data.frame(variables$x,variables$y,
                  variables$aspect,variables$tri,
                  variables$alt,variables$soil_lito,
                  variables$soil_podz))

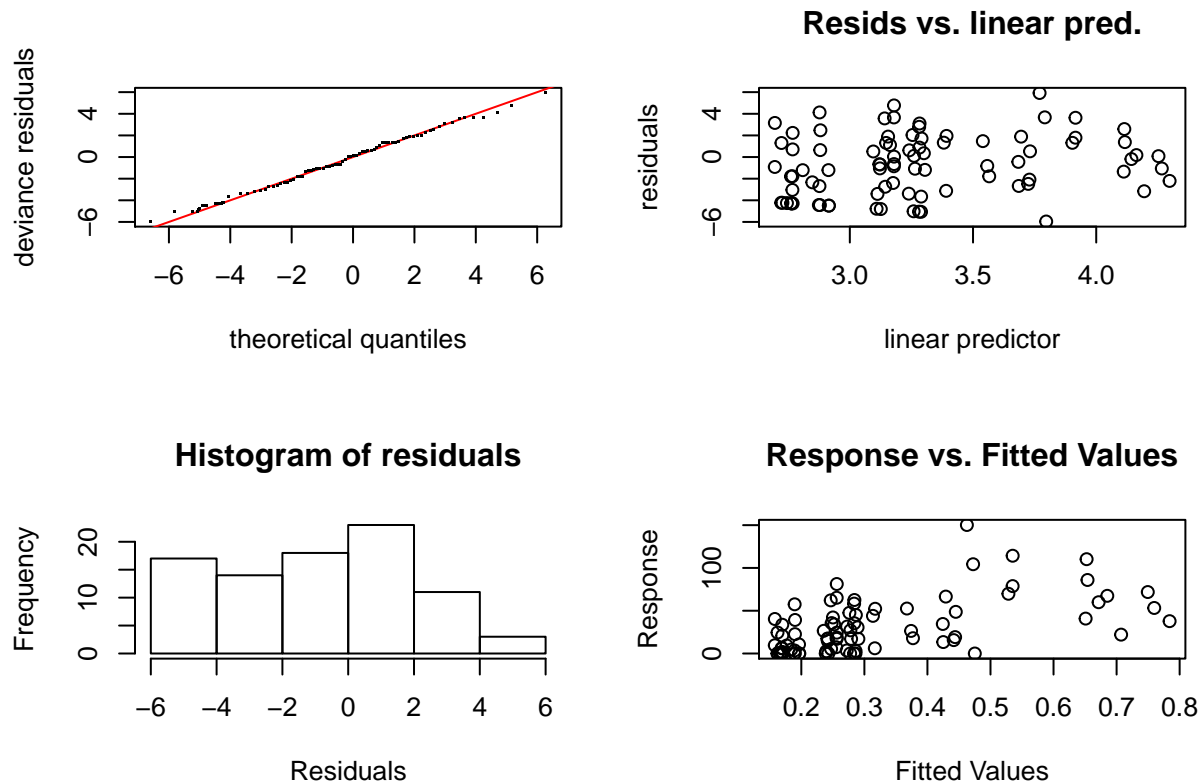
## Correlations of the variables
##
##          variables.x variables.y variables.aspect variables.tri
## variables.x          1.00000000  0.60981655          0.09163273  0.62673295
## variables.y          0.60981655  1.00000000          0.03593386  0.32412086
## variables.aspect      0.09163273  0.03593386          1.00000000  0.18513879
## variables.tri         0.62673295  0.32412086          0.18513879  1.00000000
## variables.alt         0.73298606  0.76080743          0.08744181  0.31793729
## variables.soil_lito    0.28024000  0.25638964          0.06254317  0.06900076
## variables.soil_podz   -0.60983791 -0.49132555         -0.06583941 -0.38954162
##
##          variables.alt variables.soil_lito variables.soil_podz
## variables.x          0.73298606          0.28024000         -0.60983791
## variables.y          0.76080743          0.25638964         -0.49132555
## variables.aspect      0.08744181          0.06254317         -0.06583941
## variables.tri         0.31793729          0.06900076         -0.38954162
## variables.alt         1.00000000          0.13344558         -0.40472682
## variables.soil_lito    0.13344558          1.00000000         -0.66305789
## variables.soil_podz   -0.40472682         -0.66305789          1.00000000
##
##
## Variance inflation factors
##
##          GVIF
## variables.x      4.269248
## variables.y      2.678293
## variables.aspect  1.058190
## variables.tri     1.966998
## variables.alt     3.859409
## variables.soil_lito 2.004590
## variables.soil_podz 2.897606
```

Since all variance inflation factors are below 5 we can proceed.

5.4.1 Seedlings - Topographic

Tweedie

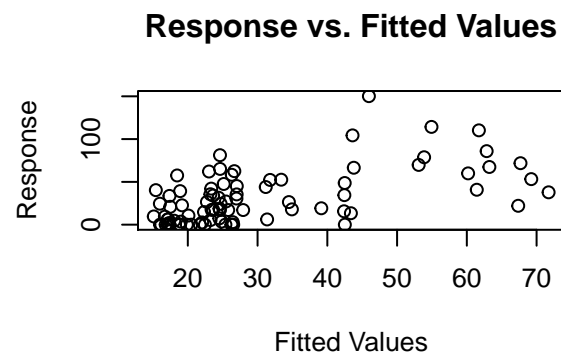
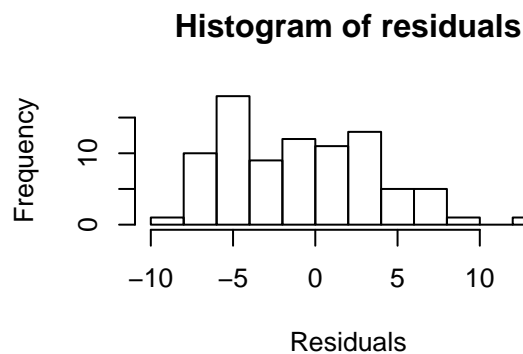
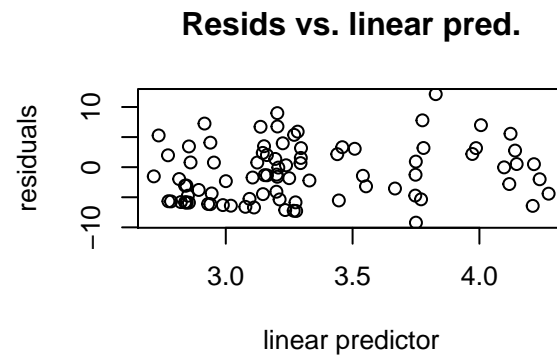
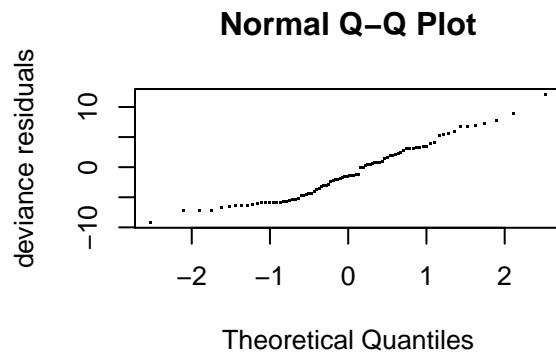

```
df<-backup
df$ddf$fitted<- df$ddf$fitted[as.character(subset(data,data$distance<=6 & data$class=="seed")$id)]
modseedgis_tw<- dsm(Nhat ~ zone_type+s(x,y)+s(aspect)+s(tri)+
  s(alt)+s(soil_lito)+s(soil_podz)+
  s(estate,bs="re"), df, observation.data=obs.table.seed,
  segment.data=variables,engine="gam",family=tw(),select=TRUE,method="REML")
gam.check(modseedgis_tw)
```



```
##
## Method: REML   Optimizer: outer newton
## full convergence after 14 iterations.
## Gradient range [-5.384185e-05,0.0004880651]
## (score 354.5149 & scale 7.398576).
## Hessian positive definite, eigenvalue range [5.152816e-06,85.69667].
## Model rank = 84 / 84
##
## Basis dimension (k) checking results. Low p-value (k-index<1) may
## indicate that k is too low, especially if edf is close to k'.
##
##          k'      edf k-index p-value
## s(x,y)    2.90e+01 3.81e-05 1.04e+00 0.92
## s(aspect)  9.00e+00 2.94e-01 8.02e-01 0.10
## s(tri)     9.00e+00 1.28e-03 1.00e+00 0.71
## s(alt)     9.00e+00 3.01e-05 9.40e-01 0.45
## s(soil_lito) 9.00e+00 6.81e-05 1.02e+00 0.78
## s(soil_podz) 9.00e+00 3.16e-05 8.27e-01 0.12
## s(estate)  8.00e+00 5.20e+00      NA      NA
```

quasi-Poisson

```
df<-backup
df$ddf$fitted<- df$ddf$fitted[as.character(subset(data,data$distance<=6 & data$class=="seed")$id)]
modseedgis_qpois<- dsm(Nhat ~ zone_type+s(x,y)+s(aspect)+s(tri)+
  s(alt)+s(soil_lito)+s(soil_podz)+
  s(estate,bs="re"), df,
  observation.data=obs.table.seed,
  segment.data=variables,engine="gam",
  family=quasipoisson,select=TRUE,method="REML")
gam.check(modseedgis_qpois)
```

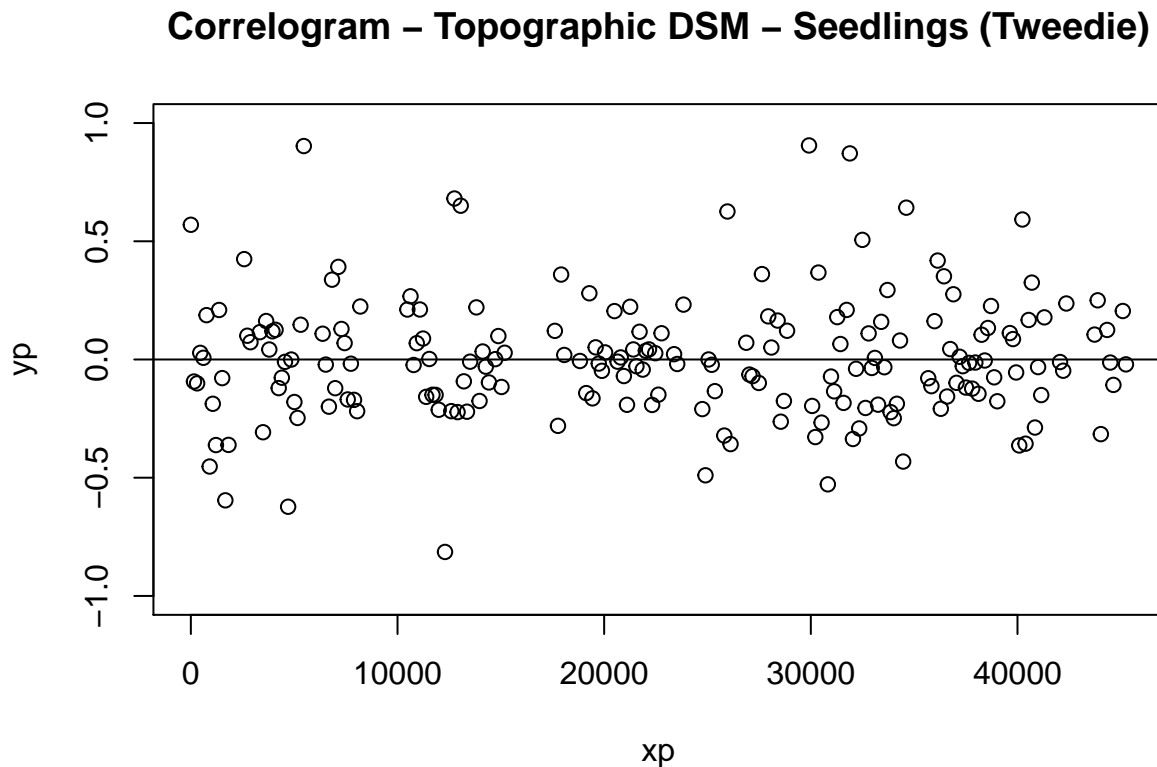


```
##
## Method: REML   Optimizer: outer newton
## full convergence after 13 iterations.
## Gradient range [-1.789357e-05,6.010877e-05]
## (score 186.3852 & scale 22.10125).
## Hessian positive definite, eigenvalue range [1.051911e-06,42.17759].
## Model rank = 84 / 84
##
## Basis dimension (k) checking results. Low p-value (k-index<1) may
## indicate that k is too low, especially if edf is close to k'.
##
##          k'      edf  k-index p-value
## s(x,y)      2.90e+01 9.84e-06 1.14e+00 0.99
## s(aspect)    9.00e+00 4.27e-01 8.63e-01 0.15
## s(tri)       9.00e+00 1.07e-01 1.02e+00 0.67
## s(alt)       9.00e+00 2.05e-05 9.59e-01 0.42
```

```
## s(soil_lito) 9.00e+00 4.51e-05 1.07e+00 0.80
## s(soil_podz) 9.00e+00 6.05e-05 8.76e-01 0.17
## s(estate) 8.00e+00 5.30e+00 NA NA
```

The residuals of the Tweedie model present considerably less patterns. However we still need to determine whether there is spatial autocorrelation in the residuals.

```
dts<-data.frame(x=variables$x,y=variables$y,z=residuals(modseedgis_tw))
surface<-surf.ls(2,dts)
correlogram(surface,300,main="Correlogram - Topographic DSM - Seedlings (Tweedie)")
```



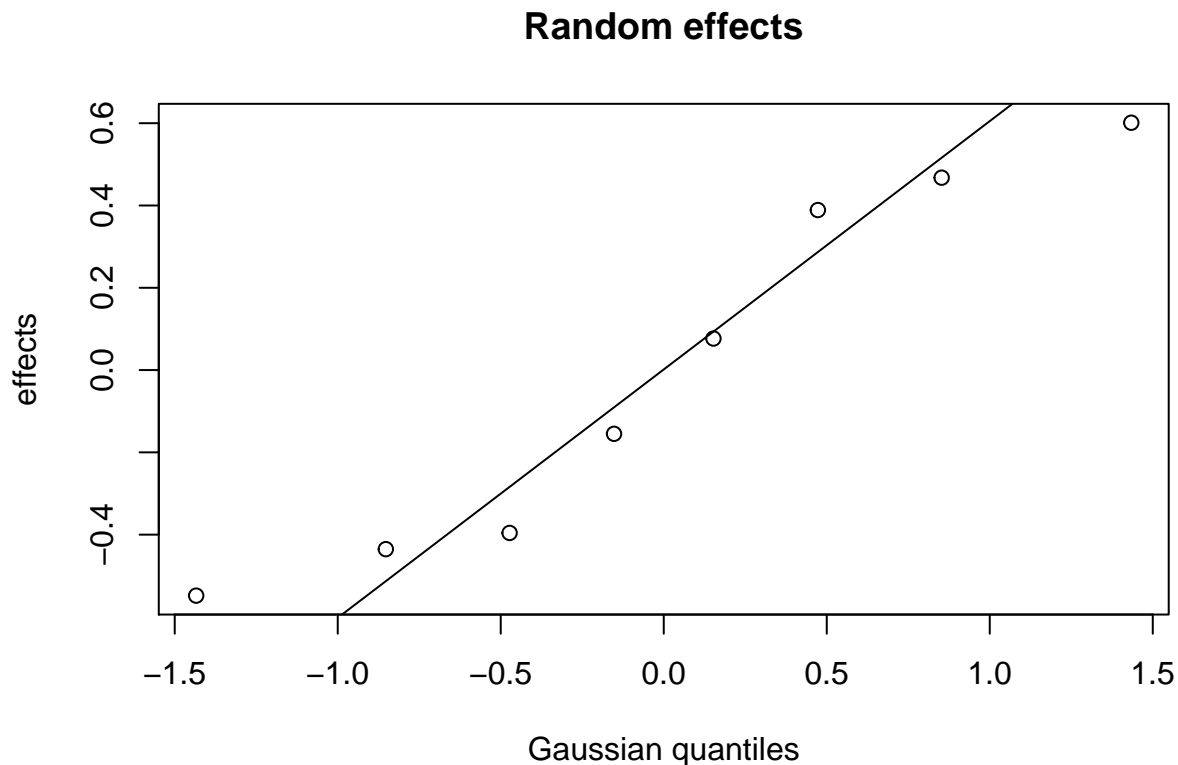
Since we can't find any major problems, we now look at the model.

```
summary(modseedgis_tw)
```

```
##
## Family: Tweedie(p=1.36)
## Link function: log
##
## Formula:
## Nhat ~ zone_type + s(x, y) + s(aspect) + s(tri) + s(alt) + s(soil_lito) +
##       s(soil_podz) + s(estate, bs = "re") + offset(off.set)
##
## Parametric coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -3.1055    0.2445  -12.703  <2e-16 ***
## zone_typecz   0.4114    0.2049   2.008   0.0481 *
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##              edf Ref.df    F  p-value
## s(x,y)        3.810e-05   29 0.000 0.814575
## s(aspect)      2.942e-01    9 0.040 0.294112
## s(tri)         1.279e-03    9 0.000 0.362748
## s(alt)         3.007e-05    9 0.000 0.832724
## s(soil_lito)   6.807e-05    9 0.000 0.606050
## s(soil_podz)   3.161e-05    9 0.000 0.583453
## s(estate)      5.205e+00    7 3.533 0.000156 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.255   Deviance explained = 25.6%
## -REML = 354.51   Scale est. = 7.3986    n = 86
```

```
plot(modseedgis_tw,select=7,scale=0,main="Random effects")
```



5.4.2 Saplings - Topographic

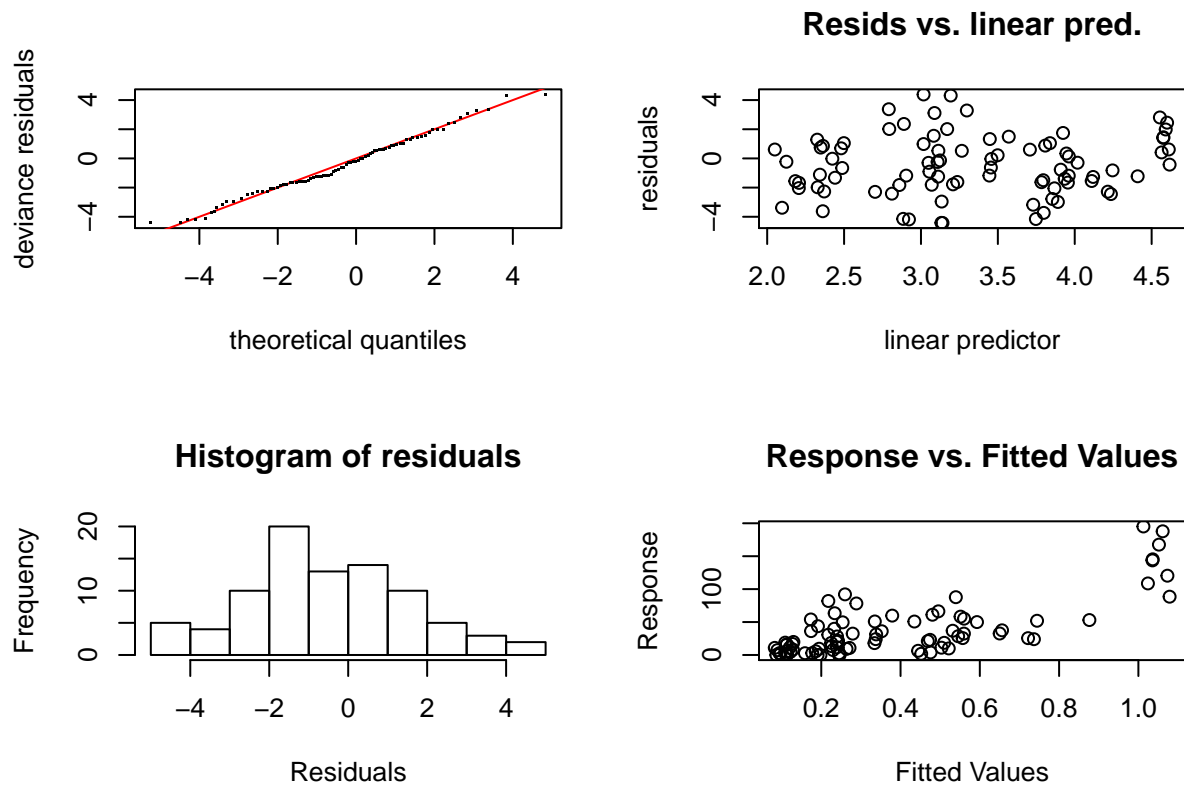
Tweedie

```
df<-backup
df$ddf$fitted<- df$ddf$fitted[as.character(subset(data,data$distance<=6 & data$class=="sap")$id)]
modsapgis_tw<- dsm(Nhat ~ zone_type+s(x,y)+s(aspect)+s(tri)+
  s(alt)+s(soil_lito)+s(soil_podz)+
  s(estate,bs="re"), df, observation.data=obs.table.sap,
```

```

segment.data=variables,engine="gam",
family=tw(),select=TRUE,method="REML")
gam.check(modsapgis_tw)

```



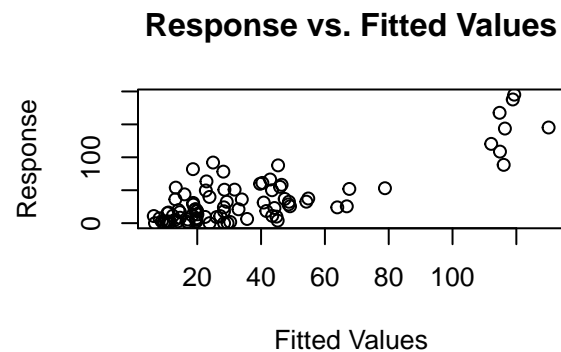
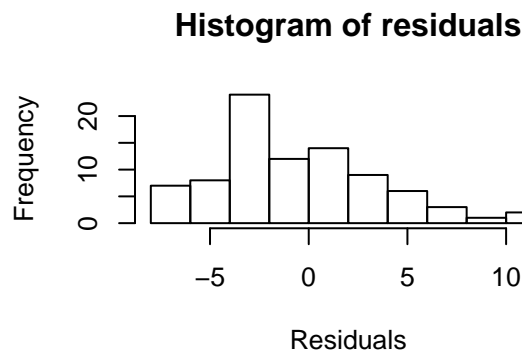
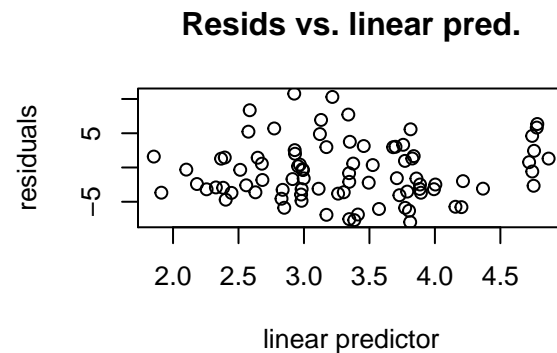
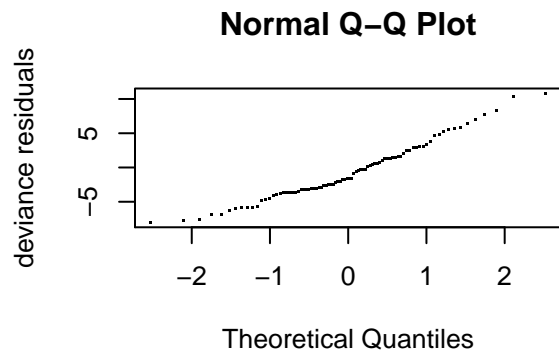
```

##
## Method: REML   Optimizer: outer newton
## full convergence after 12 iterations.
## Gradient range [-0.0001693822,0.0005189584]
## (score 377.3378 & scale 3.842786).
## Hessian positive definite, eigenvalue range [4.57976e-06,86.35144].
## Model rank = 84 / 84
##
## Basis dimension (k) checking results. Low p-value (k-index<1) may
## indicate that k is too low, especially if edf is close to k'.
##
##          k'      edf  k-index p-value
## s(x,y)      2.90e+01 1.82e+00 9.33e-01 0.36
## s(aspect)    9.00e+00 7.42e-01 9.67e-01 0.50
## s(tri)       9.00e+00 2.87e-04 1.02e+00 0.76
## s(alt)       9.00e+00 8.65e-05 9.23e-01 0.40
## s(soil_lito) 9.00e+00 8.80e-05 9.26e-01 0.42
## s(soil_podz) 9.00e+00 1.96e-01 9.35e-01 0.42
## s(estate)   8.00e+00 1.76e+00      NA      NA

```

quasi-Poisson

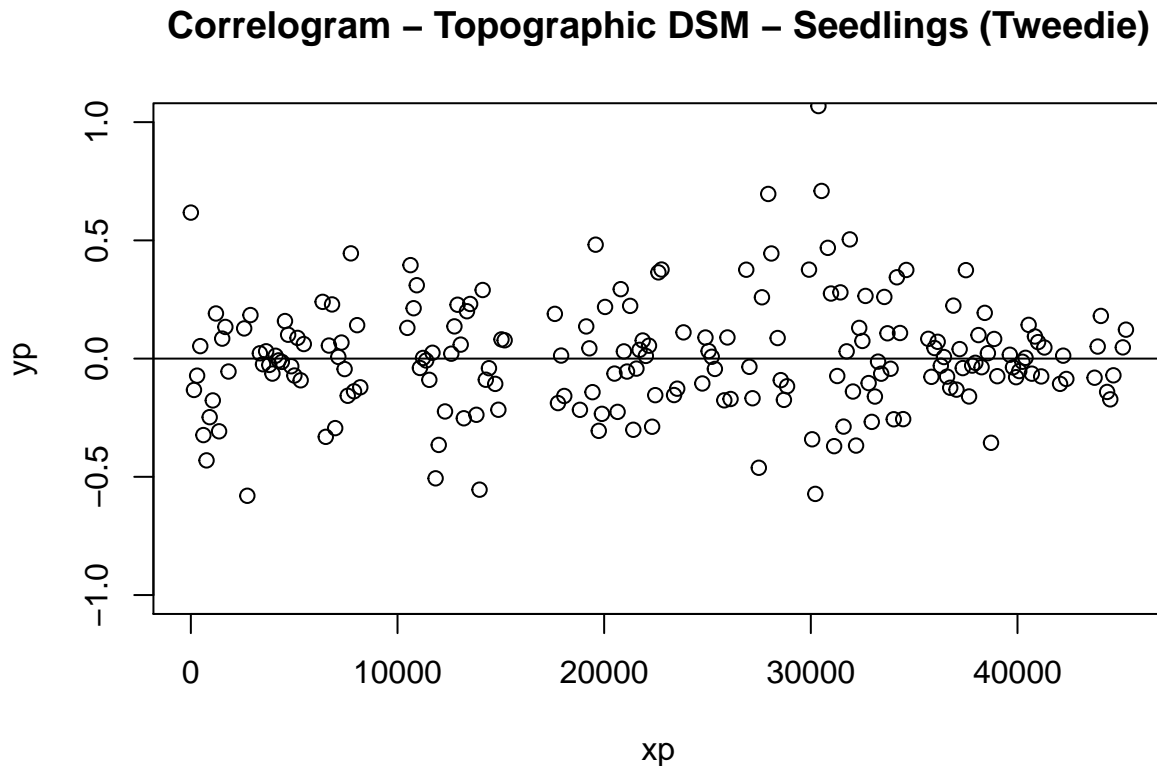
```
df<-backup
df$ddf$fitted<- df$ddf$fitted[as.character(subset(data,data$distance<=6 & data$class=="sap")$id)]
modsapgis_qpois<- dsm(Nhat ~ zone_type+s(x,y)+s(aspect)+s(tri)+
  s(alt)+s(soil_lito)+s(soil_podz)+
  s(estate,bs="re"), df, observation.data=obs.table.sap,
  segment.data=variables,engine="gam",
  family=quasipoisson,select=TRUE,method="REML")
gam.check(modsapgis_qpois)
```



```
##
## Method: REML   Optimizer: outer newton
## full convergence after 19 iterations.
## Gradient range [-0.0001462467,0.0003503313]
## (score 178.494 & scale 20.31981).
## Hessian positive definite, eigenvalue range [6.772085e-06,42.03824].
## Model rank = 84 / 84
##
## Basis dimension (k) checking results. Low p-value (k-index<1) may
## indicate that k is too low, especially if edf is close to k'.
##
##           k'      edf  k-index p-value
## s(x,y)      2.90e+01 1.83e+00 1.01e+00 0.56
## s(aspect)    9.00e+00 1.62e+00 1.02e+00 0.62
## s(tri)       9.00e+00 1.15e-04 1.02e+00 0.56
## s(alt)       9.00e+00 4.42e-05 9.75e-01 0.42
## s(soil_lito) 9.00e+00 9.04e-01 9.20e-01 0.22
## s(soil_podz) 9.00e+00 8.20e-01 9.91e-01 0.46
## s(estate)    8.00e+00 4.84e-04      NA      NA
```

The residuals of the Tweedie model present considerably less patterns. However we still need to determine whether there is spatial autocorrelation in the residuals.

```
dts<-data.frame(x=variables$x,y=variables$y,z=residuals(modsapgis_tw))
surface<-surf.ls(2,dts)
correlogram(surface,300,main="Correlogram - Topographic DSM - Seedlings (Tweedie)")
```



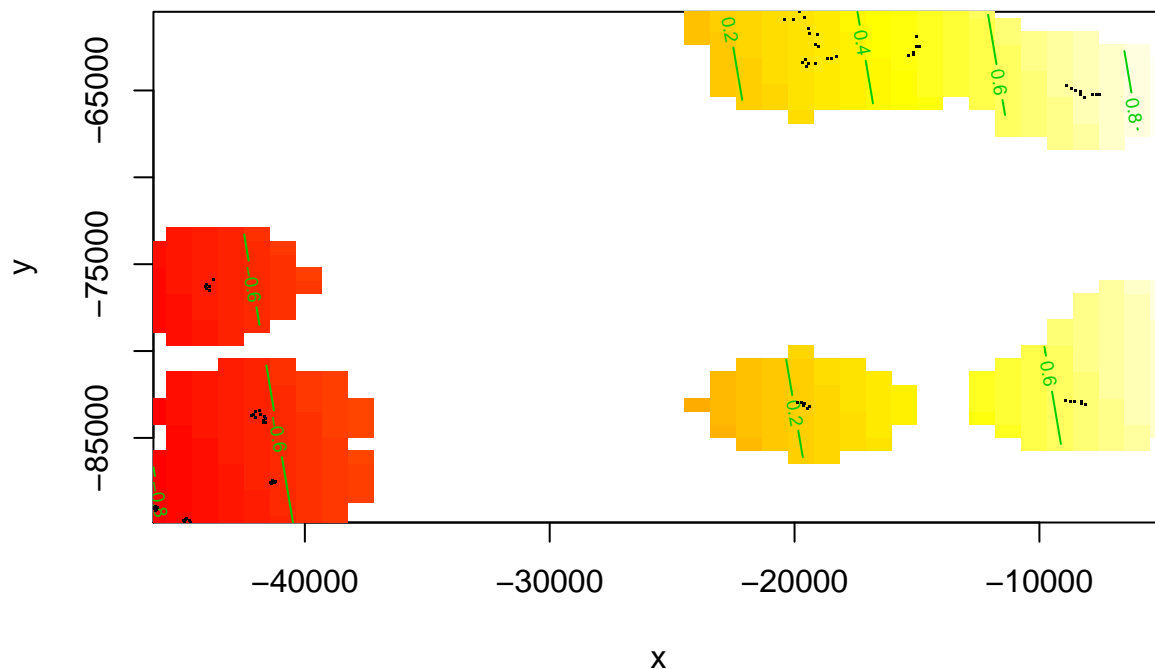
Since we can't find any major problems, we now look at the model.

```
summary(modsapgis_tw)
```

```
##
## Family: Tweedie(p=1.484)
## Link function: log
##
## Formula:
## Nhat ~ zone_type + s(x, y) + s(aspect) + s(tri) + s(alt) + s(soil_lito) +
##       s(soil_podz) + s(estate, bs = "re") + offset(off.set)
##
## Parametric coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -3.4046    0.1646  -20.687  < 2e-16 ***
## zone_typecz   0.6379    0.1818   3.509  0.000743 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##               edf Ref.df    F  p-value
## s(x,y)         1.818e+00   29 2.307 9.76e-06 ***
```

```
## s(aspect)      7.418e-01      9 0.151      0.173
## s(tri)         2.873e-04      9 0.000      0.423
## s(alt)         8.648e-05      9 0.000      0.994
## s(soil_lito)   8.799e-05      9 0.000      0.717
## s(soil_podz)   1.956e-01      9 0.026      0.271
## s(estate)      1.765e+00      7 0.415      0.138
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.522   Deviance explained = 45.3%
## -REML = 377.34   Scale est. = 3.8428      n = 86
```

```
plot(modsapgis_tw,select=1,scheme=2,scale=0,main="")
```

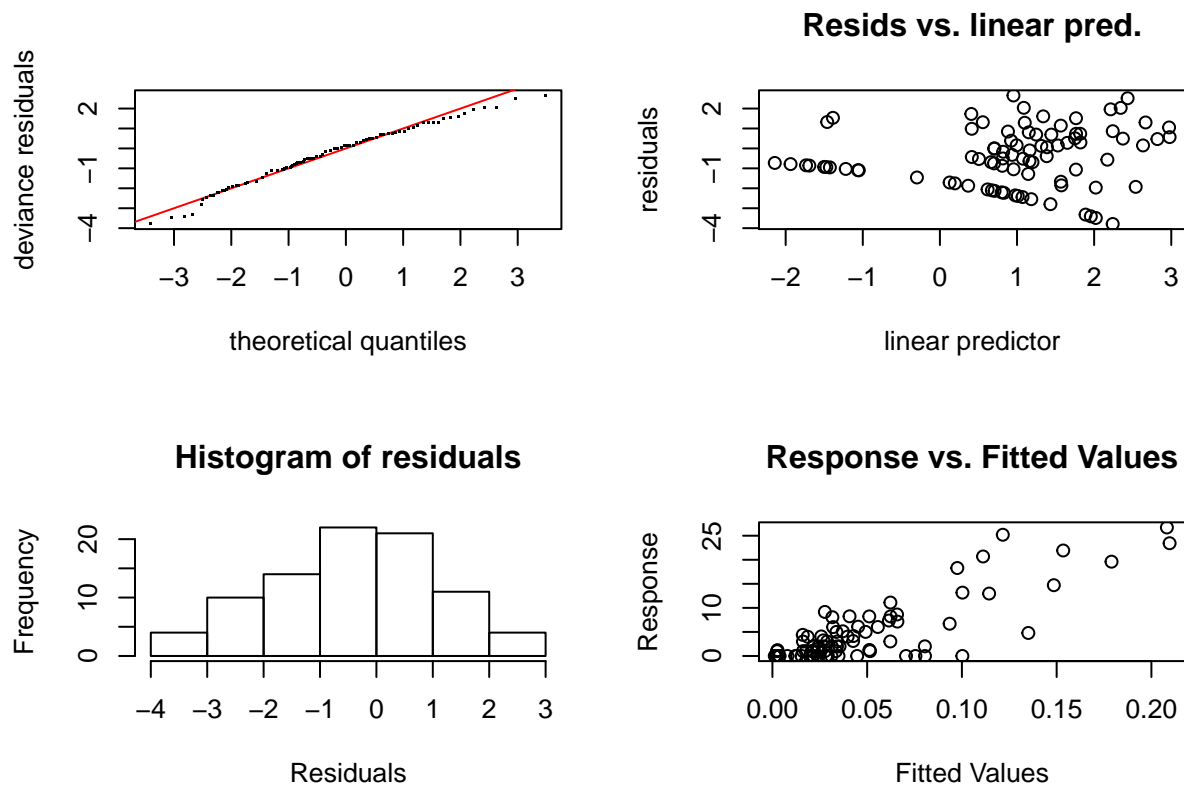


This plot shows how the abundance of saplings changes with the x,y coordinates. Saplings abundances is shown in a colour gradient ranging from yellow (lower) to red (higher). The black dots show the starting points of the transects.

5.4.3 Young trees - topographic

Tweedie

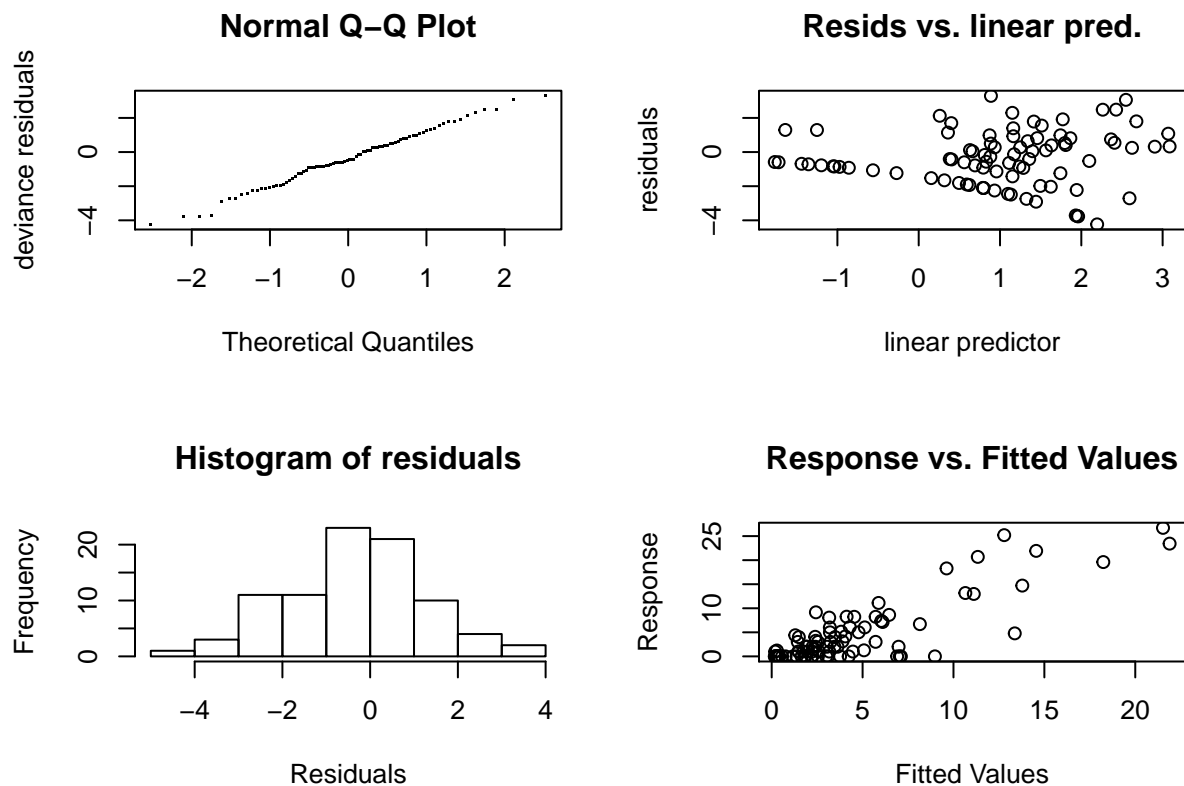
```
df<-backup
df$ddf$fitted<- df$ddf$fitted[as.character(subset(data,data$distance<=6 & data$class=="yt")$id)]
modytgis_tw<- dsm(Nhat ~ zone_type+s(x,y)+s(aspect)+s(tri)+
                  s(alt)+s(soil_lito)+s(soil_podz)+
                  s(estate,bs="re"), df, observation.data=obs.table.yt,
                  segment.data=variables,engine="gam",family=tw(),select=TRUE,method="REML")
gam.check(modytgis_tw)
```

```
##
## Method: REML   Optimizer: outer newton
## full convergence after 13 iterations.
## Gradient range [-1.808416e-05,0.0002037878]
## (score 208.7634 & scale 2.3713).
## Hessian positive definite, eigenvalue range [3.75966e-06,87.39643].
## Model rank = 84 / 84
##
## Basis dimension (k) checking results. Low p-value (k-index<1) may
## indicate that k is too low, especially if edf is close to k'.
##
##          k'    edf k-index p-value
## s(x,y)      29.000 0.167  1.056  0.94
## s(aspect)    9.000 1.752  1.067  0.89
## s(tri)       9.000 0.331  0.894  0.37
## s(alt)       9.000 0.433  0.903  0.34
## s(soil_lito) 9.000 3.460  0.964  0.64
## s(soil_podz) 9.000 1.684  0.770  0.06
## s(estate)   8.000 5.574    NA    NA
```

quasi-Poisson

```
df<-backup
df$ddf$fitted<- df$ddf$fitted[as.character(subset(data,data$distance<=6 & data$class=="yt")$id)]
modytgis_qpois<- dsm(Nhat ~ zone_type+s(x,y)+s(aspect)+s(tri)+
                     s(alt)+s(soil_lito)+s(soil_podz)+
                     s(estate,bs="re"), df, observation.data=obs.table.yt,
                     segment.data=variables,engine="gam",family=quasipoisson,select=TRUE,method="REML")
gam.check(modytgis_qpois)
```

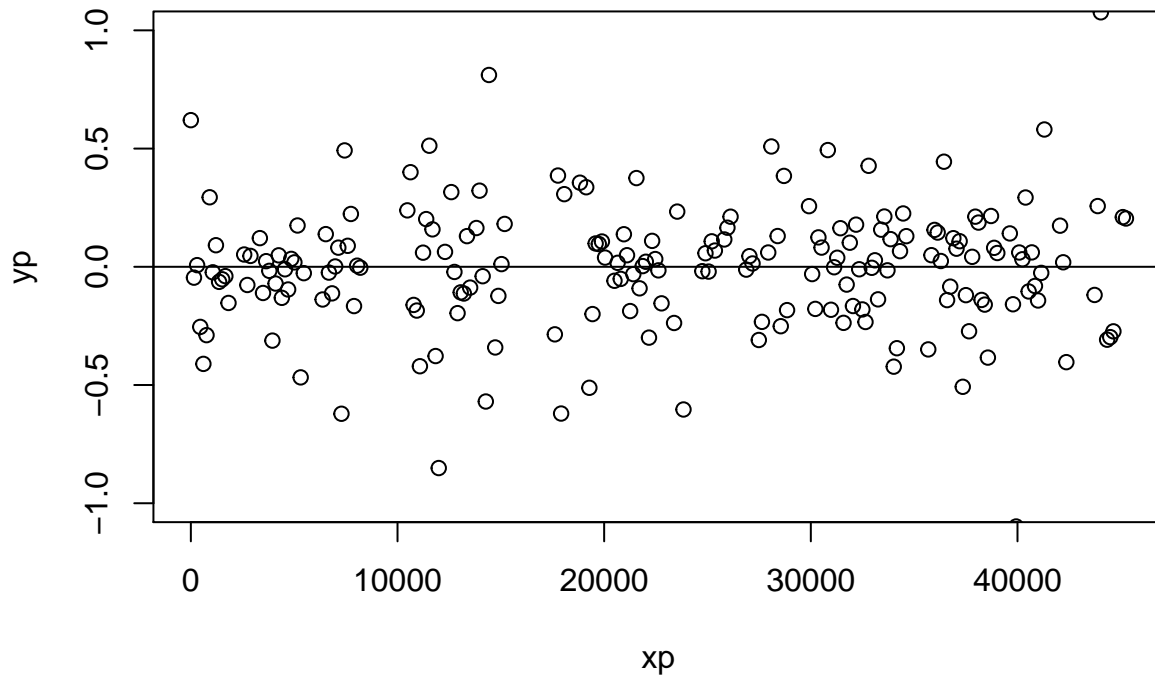


```
##
## Method: REML   Optimizer: outer newton
## full convergence after 15 iterations.
## Gradient range [-4.579769e-05,0.0002702689]
## (score 115.5787 & scale 2.671749).
## Hessian positive definite, eigenvalue range [1.261456e-06,42.23757].
## Model rank = 84 / 84
##
## Basis dimension (k) checking results. Low p-value (k-index<1) may
## indicate that k is too low, especially if edf is close to k'.
##
##          k'      edf  k-index p-value
## s(x,y)    29.00000  0.00043  1.10138  0.96
## s(aspect)  9.00000  1.98403  1.10204  0.92
## s(tri)     9.00000  0.24991  0.90052  0.28
## s(alt)     9.00000  0.50863  0.95865  0.44
## s(soil_lito) 9.00000  3.66885  0.99894  0.64
## s(soil_podz) 9.00000  1.74372  0.80914  0.08
## s(estate)  8.00000  5.50623    NA      NA
```

In this case the residuals of the quasi-Poisson model present less patterns. Now we determine whether there is spatial autocorrelation in the residuals.

```
dts<-data.frame(x=variables$x,y=variables$y,z=residuals(modytgis_qpois))
surface<-surf.ls(2,dts)
correlogram(surface,300,main="Correlogram - Topographic DSM - Young trees (quasi-Poisson)")
```

Correlogram – Topographic DSM – Young trees (quasi-Poisson)



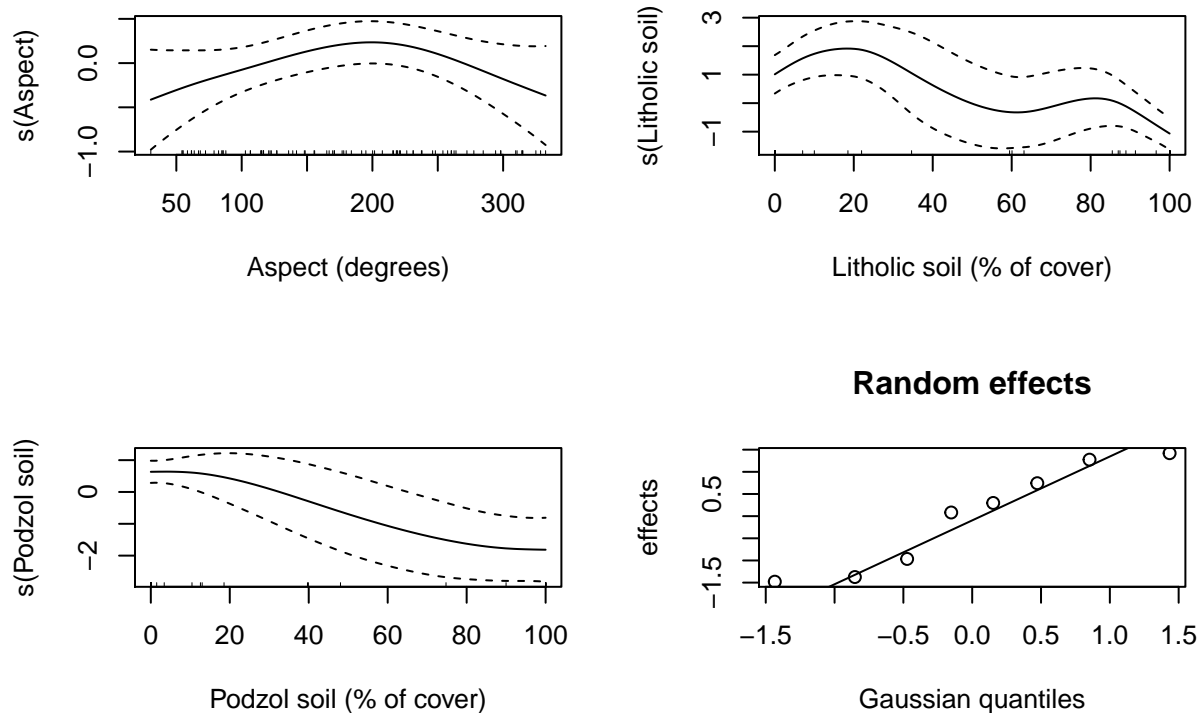
No major problems were found. We now inspect the model.

```
summary(modytgis_qpois)
```

```
##
## Family: quasipoisson
## Link function: log
##
## Formula:
## Nhat ~ zone_type + s(x, y) + s(aspect) + s(tri) + s(alt) + s(soil_lito) +
##       s(soil_podz) + s(estate, bs = "re") + offset(off.set)
##
## Parametric coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -5.91340    0.49020 -12.063  <2e-16 ***
## zone_typecz -0.02114    0.21256  -0.099    0.921
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##               edf Ref.df      F  p-value
## s(x,y)         0.0004304    29  0.000 0.473895
## s(aspect)       1.9840299     9  0.925 0.020315 *
## s(tri)          0.2499073     9  0.055 0.189593
## s(alt)          0.5086311     9  2.047 0.126833
## s(soil_lito)    3.6688506     9 64.603 0.000285 ***
## s(soil_podz)    1.7437177     9 15.703 2.65e-05 ***
## s(estate)       5.5062256     7  3.643 8.41e-05 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.628   Deviance explained = 63.1%
## -REML = 115.58   Scale est. = 2.6717    n = 86
```

```
par(mfrow=c(2,2))
plot(modytgis_tw,select=2,scheme=2,scale=0,xlab="Aspect (degrees)",ylab="s(Aspect)")
plot(modytgis_tw,select=5,scheme=2,scale=0,xlab="Litholic soil (% of cover)",ylab="s(Litholic soil)")
plot(modytgis_tw,select=6,scheme=2,scale=0,xlab="Podzol soil (% of cover)",ylab="s(Podzol soil)")
plot(modytgis_tw,select=7,scheme=2,scale=0,main="Random effects")
```



```
par(mfrow=c(1,1))
```

6. Effects of conservation zones on shrubs

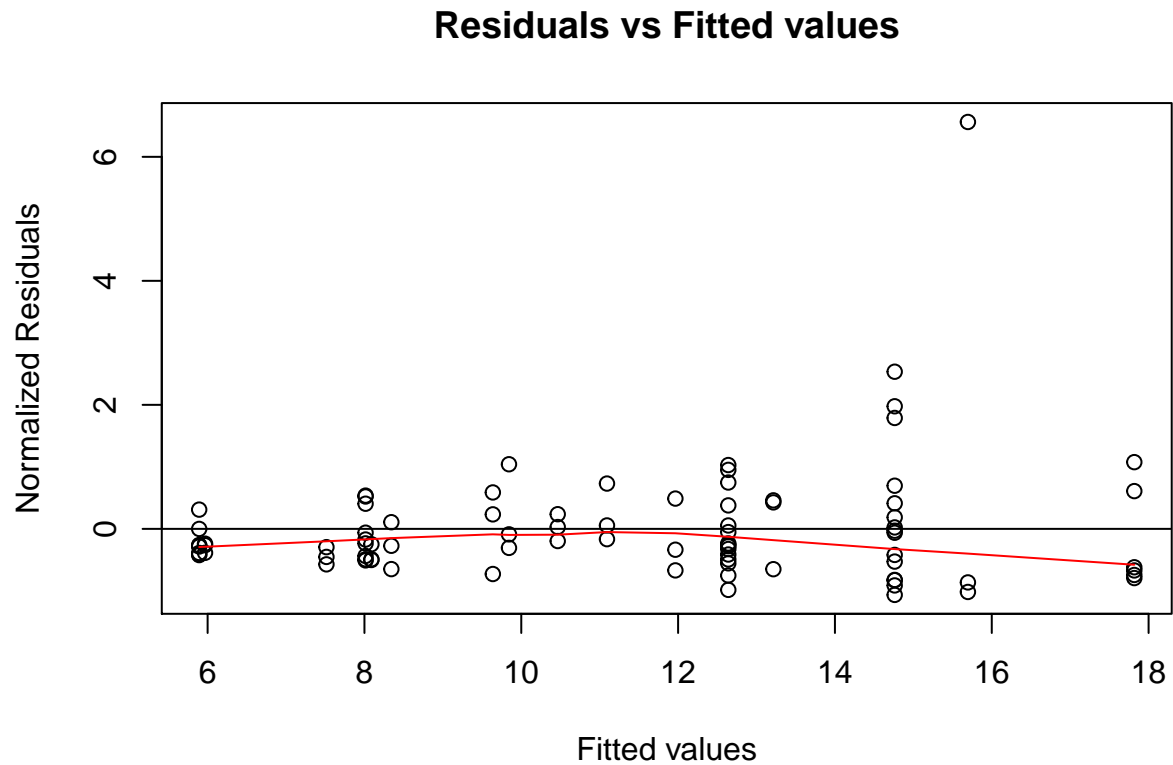
To analyze the effects of conservation on the diversity and cover of shrubs species we fitted linear mixed models.

6.1 Model validation

Each model was validated by (1) plotting the residuals against the values, (2) plotting the residuals against “zone_type” (3) checking the normality of the residuals and (4) checking the residuals for spatial autocorrelation.

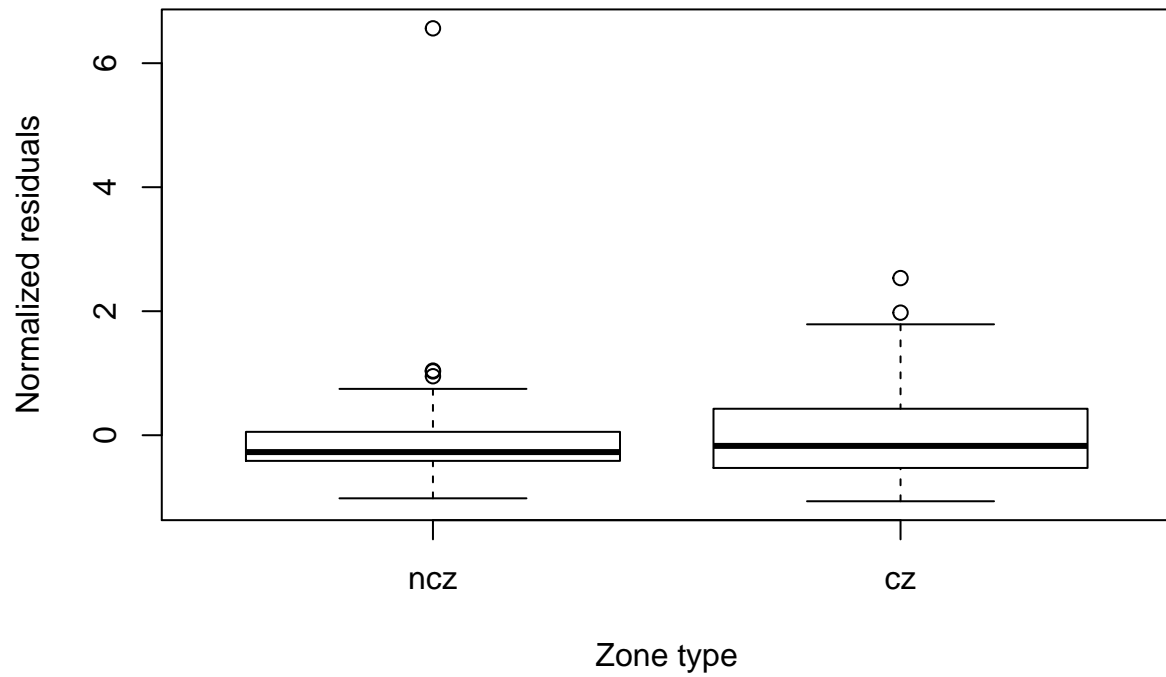
```
#Select model
model<-shrub_cover
```

```
#Plot model residuals against fitted values
EX<-resid(model,type="normalized")
FX<-fitted(model)
plot(y=EX,x=FX,ylab="Normalized Residuals",xlab="Fitted values",main="Residuals vs Fitted values")
abline(h=0)
lines(lowess(y=EX,x=FX),col=2)
```



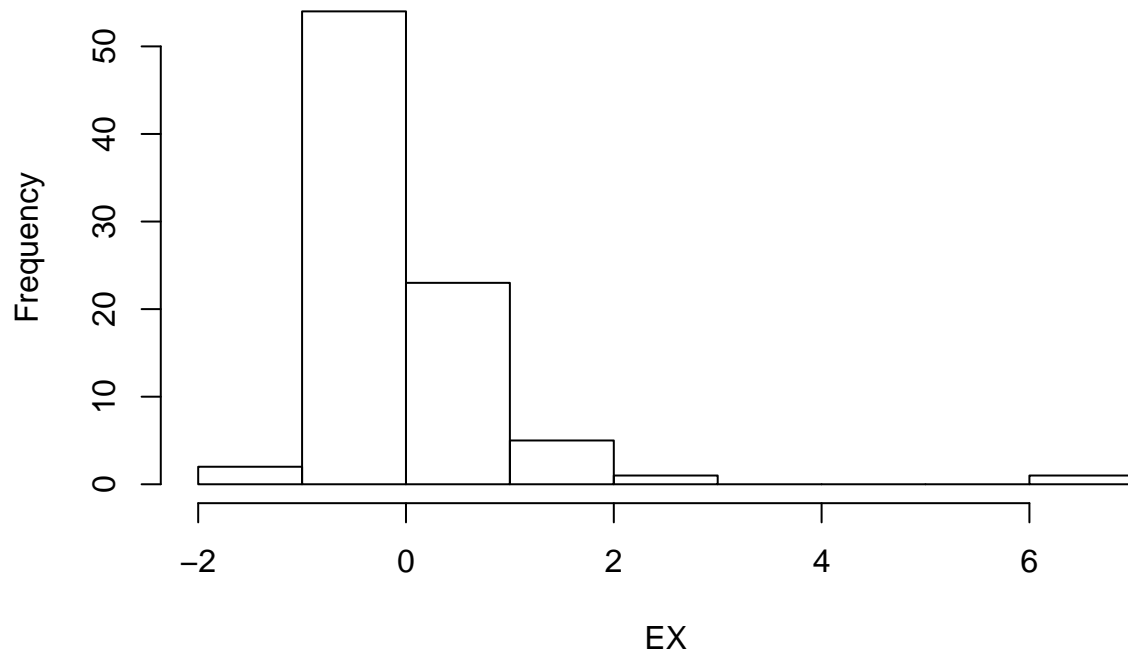
```
#Plot residuals against zone_type
plot(y=EX,x=variables$zone_type,ylab="Normalized residuals",xlab="Zone type",main="Residuals vs Zone type")
```

Residuals vs Zone type



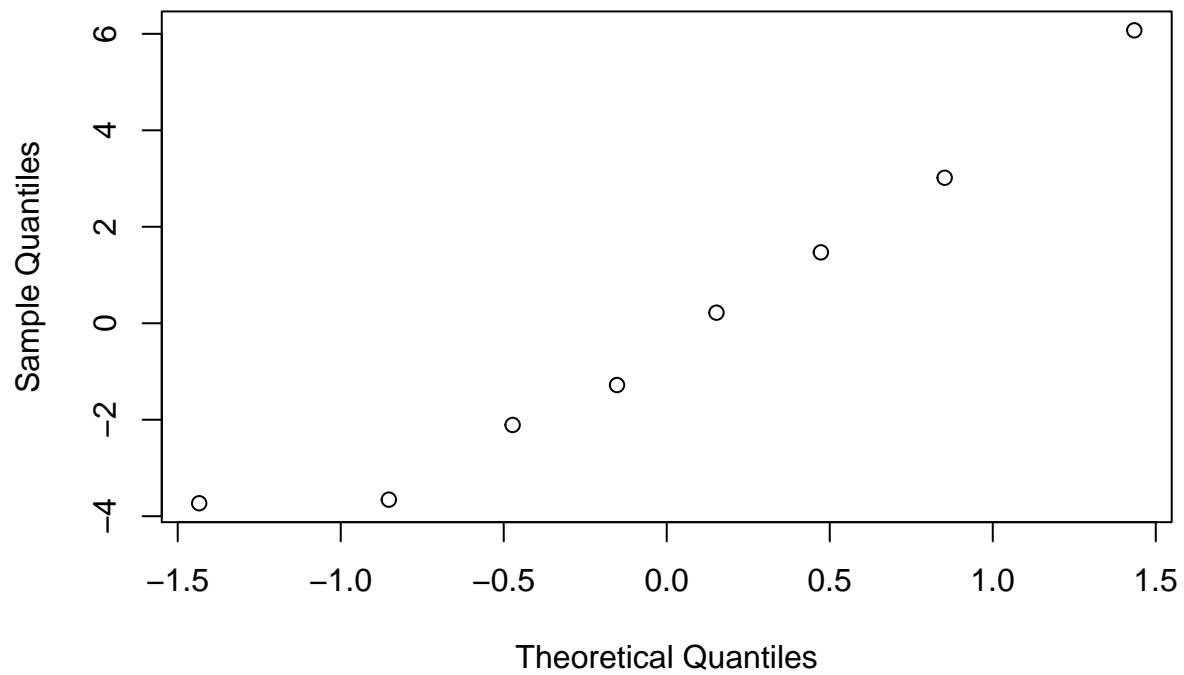
```
#Check the normality of the residuals  
hist(EX)
```

Histogram of EX

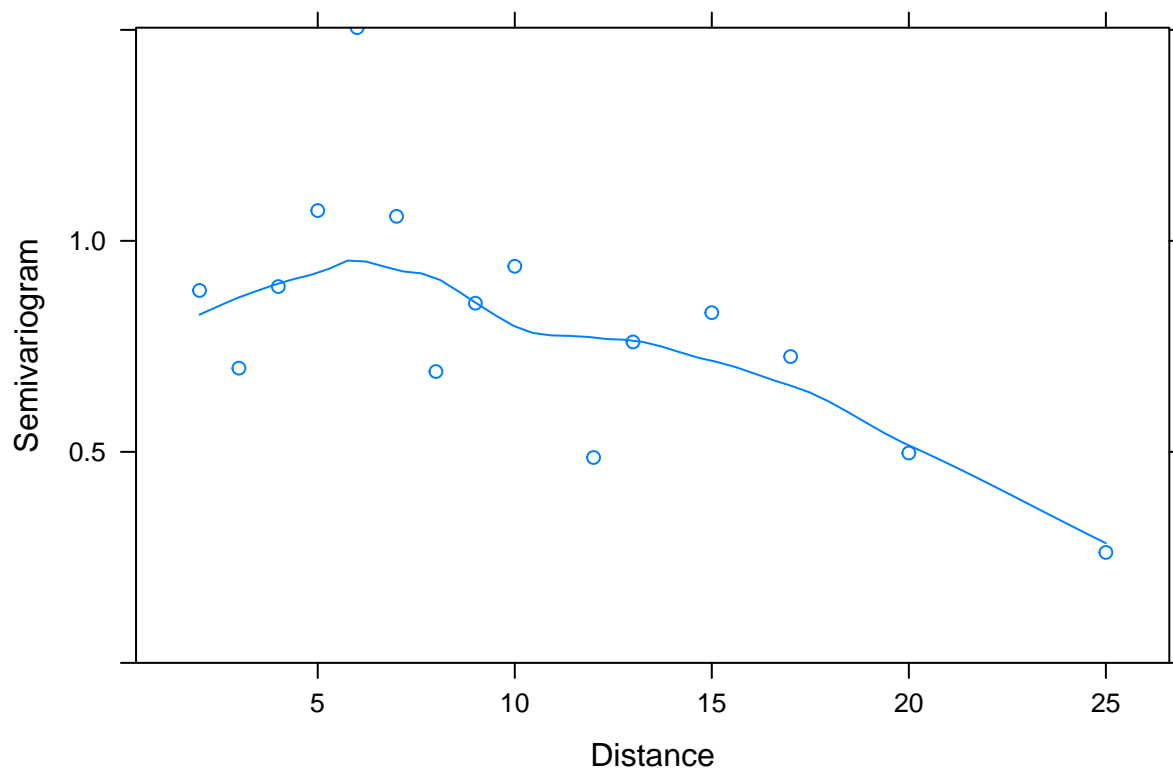


```
#Check the normality of the random effects  
data.frame(ranef(model))->ri  
qqnorm(ri$X.Intercept.,main="qqplot Random Intercepts")
```

qqplot Random Intercepts



```
#Check the residuals for spatial autocorrelation  
plot(Variogram(model))
```



7.2 Dealing with spatial autocorrelation

When spatial autocorrelation was detected, we added a linear, Gaussian, rational quadratic or exponential correlation structures to the model. We selected the best one based on the AIC and by analysing the variogram.

Eg.

```
div1<-lme(shrub_div~zone_type,random = ~ 1 | estate,method="REML",data=variables,correlation = corLin(f
div2<-lme(shrub_div~zone_type,random = ~ 1 | estate,method="REML",data=variables,correlation = corLin(f
div3<-lme(shrub_div~zone_type,random = ~ 1 | estate,method="REML",data=variables,correlation = corRatio
div4<-lme(shrub_div~zone_type,random = ~ 1 | estate,method="REML",data=variables,correlation = corExp(f

AIC(div1,div2,div3,div4)
```

7.3 Final models

```
#Total Shrub cover
shrub_cover<-lme(shrub_cover~zone_type,random = ~ 1 | estate,
                 method="REML",data=variables)
summary(shrub_cover)
```

```
## Linear mixed-effects model fit by REML
## Data: variables
##      AIC      BIC    logLik
## 688.3659 698.0892 -340.1829
```



```
##
## Random effects:
## Formula: ~1 | estate
## (Intercept) Residual
## StdDev: 4.696265 12.84765
##
## Fixed effects: shrub_cover ~ zone_type
## Value Std.Error DF t-value p-value
## (Intercept) 9.623898 2.704154 77 3.558932 0.0006
## zone_typecz 2.121921 2.781717 77 0.762810 0.4479
## Correlation:
## (Intr)
## zone_typecz -0.541
##
## Standardized Within-Group Residuals:
## Min Q1 Med Q3 Max
## -1.0649525 -0.4928849 -0.2533231 0.2918275 6.5618213
##
## Number of Observations: 86
## Number of Groups: 8
```

```
#Shannon-Wiener shrub diversity
shrub_div<-lme(shrub_div~zone_type,random = ~ 1 | estate,
              method="REML",data=variables)
summary(shrub_div)
```

```
## Linear mixed-effects model fit by REML
## Data: variables
## AIC BIC logLik
## 133.4045 143.1278 -62.70225
##
## Random effects:
## Formula: ~1 | estate
## (Intercept) Residual
## StdDev: 0.2096046 0.4676848
##
## Fixed effects: shrub_div ~ zone_type
## Value Std.Error DF t-value p-value
## (Intercept) 0.6149841 0.1077633 77 5.706803 0.0000
## zone_typecz 0.2755090 0.1013101 77 2.719461 0.0081
## Correlation:
## (Intr)
## zone_typecz -0.494
##
## Standardized Within-Group Residuals:
## Min Q1 Med Q3 Max
## -2.4458613 -0.7886007 0.2061238 0.7340563 1.7167312
##
## Number of Observations: 86
## Number of Groups: 8
```

```
#Shrub species richness
shrub_rich<-lme(shrub_rich~zone_type,random = ~ 1 | estate,
```

```

method="REML",data=variables)
summary(shrub_rich)

```

```

## Linear mixed-effects model fit by REML
## Data: variables
##      AIC      BIC    logLik
##  339.527 349.2503 -165.7635
##
## Random effects:
## Formula: ~1 | estate
##      (Intercept) Residual
## StdDev:    0.7195592 1.594568
##
## Fixed effects: shrub_rich ~ zone_type
##              Value Std.Error DF   t-value p-value
## (Intercept)  2.434498  0.3686559 77  6.603714  0.0000
## zone_typecz  1.032552  0.3454215 77  2.989252  0.0038
## Correlation:
##      (Intr)
## zone_typecz -0.493
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -2.166491273 -0.621218646 -0.002073735  0.489555119  2.850541659
##
## Number of Observations: 86
## Number of Groups: 8

```

```

#Rosmarinus offinalis cover
doroffici<-lme(do_roffici~zone_type,random = ~ 1 | estate,
method="REML",data=variables)
summary(doroffici)

```

```

## Linear mixed-effects model fit by REML
## Data: variables
##      AIC      BIC    logLik
##  358.4707 368.194 -175.2354
##
## Random effects:
## Formula: ~1 | estate
##      (Intercept) Residual
## StdDev:    0.5404221 1.818047
##
## Fixed effects: do_roffici ~ zone_type
##              Value Std.Error DF   t-value p-value
## (Intercept)  0.5774086  0.3553163 77  1.6250551  0.1082
## zone_typecz  0.1051462  0.3934339 77  0.2672524  0.7900
## Correlation:
##      (Intr)
## zone_typecz -0.583
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max

```

```
## -0.7654992 -0.2864746 -0.1990175 -0.1411828 5.5819715
##
## Number of Observations: 86
## Number of Groups: 8
```

```
#Cistus salvifolius cover
```

```
docsalvi<-lme(do_csalvi~zone_type,random = ~ 1 | estate,
              method="REML",data=variables)
```

```
#Ulex sp cover
```

```
doulexsp<-lme(do_ulexsp~zone_type,random = ~ 1 | estate,
              method="REML",
              correlation = corGaus(form =~ x + y, nugget = TRUE),data=variables)
summary(doulexsp)
```

```
## Linear mixed-effects model fit by REML
## Data: variables
##      AIC      BIC    logLik
## 475.2773 489.8622 -231.6387
##
## Random effects:
## Formula: ~1 | estate
##      (Intercept) Residual
## StdDev:      1.571784 3.493801
##
## Correlation Structure: Gaussian spatial correlation
## Formula: ~x + y | estate
## Parameter estimate(s):
##      range      nugget
## 17.910691 0.101446
## Fixed effects: do_ulexsp ~ zone_type
##              Value Std.Error DF  t-value p-value
## (Intercept) 1.605991 0.8065348 77 1.991223 0.0500
## zone_typeecz 2.008387 0.7568365 77 2.653661 0.0097
## Correlation:
##      (Intr)
## zone_typeecz -0.494
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -1.4179963 -0.6335396 -0.1832274 0.3569562 5.4349820
##
## Number of Observations: 86
## Number of Groups: 8
```

```
#Cistus crispus cover
```

```
doccrispus<-lme(do_ccrispus~zone_type,random = ~ 1 | estate,
                method="REML",data=variables)
summary(doccrispus)
```

```
## Linear mixed-effects model fit by REML
## Data: variables
##      AIC      BIC    logLik
```

```
##    427.0174 436.7406 -209.5087
##
## Random effects:
## Formula: ~1 | estate
##      (Intercept) Residual
## StdDev:    0.4693493 2.774423
##
## Fixed effects: do_ccrispus ~ zone_type
##              Value Std.Error DF   t-value p-value
## (Intercept) 0.9092586 0.4762534 77 1.9091910 0.0600
## zone_typecz 0.0614145 0.5996704 77 0.1024138 0.9187
## Correlation:
##      (Intr)
## zone_typecz -0.661
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -0.5301371 -0.3394766 -0.2987919 -0.1420276  6.6718714
##
## Number of Observations: 86
## Number of Groups: 8
```

```
#Lavandula stoechas cover
dolstoechas<-lme(do_lstoechas~zone_type,random = ~ 1 | estate,
                 method="REML",data=variables)
summary(dolstoechas)
```

```
## Linear mixed-effects model fit by REML
## Data: variables
##      AIC      BIC    logLik
## 175.367 185.0902 -83.68348
##
## Random effects:
## Formula: ~1 | estate
##      (Intercept) Residual
## StdDev:    0.4470029 0.5809405
##
## Fixed effects: do_lstoechas ~ zone_type
##              Value Std.Error DF   t-value p-value
## (Intercept) 0.28837114 0.1861545 77 1.5490957 0.1255
## zone_typecz 0.05818604 0.1259674 77 0.4619133 0.6454
## Correlation:
##      (Intr)
## zone_typecz -0.355
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -2.12674373 -0.35711452 -0.17012111 -0.06996277  5.06536308
##
## Number of Observations: 86
## Number of Groups: 8
```

```
#Cistus monspeliensis cover
```

```
docmonsp<-lme(do_cmonsp~zone_type,random = ~ 1 | estate,  
             method="REML",data=variables)  
summary(docmonsp)
```

```
## Linear mixed-effects model fit by REML  
## Data: variables  
##      AIC      BIC    logLik  
## 655.9571 665.6804 -323.9785  
##  
## Random effects:  
## Formula: ~1 | estate  
##      (Intercept) Residual  
## StdDev:      1.706794 10.85225  
##  
## Fixed effects: do_cmonsp ~ zone_type  
##              Value Std.Error DF   t-value p-value  
## (Intercept)  3.307433  1.842928 77   1.794662  0.0766  
## zone_typecz -2.768625  2.345352 77  -1.180473  0.2414  
## Correlation:  
##              (Intr)  
## zone_typecz -0.668  
##  
## Standardized Within-Group Residuals:  
##              Min              Q1              Med              Q3              Max  
## -0.4692076346 -0.2784396865 -0.1204079940  0.0005877631  8.7454683093  
##  
## Number of Observations: 86  
## Number of Groups: 8
```

```
#Calluna vulgaris cover
```

```
docvulgaris<-lme(do_cvulgaris~zone_type,random = ~ 1 | estate,  
                method="REML",data=variables)  
summary(docvulgaris)
```

```
## Linear mixed-effects model fit by REML  
## Data: variables  
##      AIC      BIC    logLik  
## 328.3315 338.0547 -160.1657  
##  
## Random effects:  
## Formula: ~1 | estate  
##      (Intercept) Residual  
## StdDev:      0.2610521 1.541896  
##  
## Fixed effects: do_cvulgaris ~ zone_type  
##              Value Std.Error DF   t-value p-value  
## (Intercept) -0.0039729 0.2647126 77 -0.0150083  0.9881  
## zone_typecz  0.4655651 0.3332695 77  1.3969626  0.1664  
## Correlation:  
##              (Intr)  
## zone_typecz -0.661
```

```
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -0.48923265 -0.27758999 -0.18728933  0.02435333  8.22731039
##
## Number of Observations: 86
## Number of Groups: 8
```

```
#Halimium lasianthum cover
```

```
dohlasi<-lme(do_hlasi~zone_type,random = ~ 1 | estate,
            method="REML",data=variables)
summary(dohlasi)
```

```
## Linear mixed-effects model fit by REML
## Data: variables
##      AIC      BIC    logLik
##  336.4287 346.152 -164.2144
##
## Random effects:
## Formula: ~1 | estate
##      (Intercept) Residual
## StdDev:    0.9114725 1.541529
##
## Fixed effects: do_hlasi ~ zone_type
##              Value Std.Error DF   t-value p-value
## (Intercept) 0.2375389 0.4139207 77 0.5738754 0.5677
## zone_typeecz 0.3493892 0.3341171 77 1.0457089 0.2990
## Correlation:
##      (Intr)
## zone_typeecz -0.424
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -1.63027062 -0.19965479 -0.12062964  0.02699633  7.97058797
##
## Number of Observations: 86
## Number of Groups: 8
```

```
#Cistus ladanifer
```

```
doladanifer<-lme(do_ladanifer~zone_type,random = ~ 1 | estate,
                method="REML",data=variables)
summary(doladanifer)
```

```
## Linear mixed-effects model fit by REML
## Data: variables
##      AIC      BIC    logLik
##  456.172 465.8953 -224.086
##
## Random effects:
## Formula: ~1 | estate
##      (Intercept) Residual
```

```
## StdDev:      0.518456 3.304259
##
## Fixed effects: do_ladanifer ~ zone_type
##              Value Std.Error DF   t-value p-value
## (Intercept) -0.1339134 0.5609438 77 -0.2387288  0.8120
## zone_typeecz 1.1720955 0.7141026 77  1.6413545  0.1048
## Correlation:
##              (Intr)
## zone_typeecz -0.669
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -0.4793432 -0.2965840 -0.1246205  0.0581386  6.4268948
##
## Number of Observations: 86
## Number of Groups: 8
```

#Daphne gnidium

```
dodgnidium<-lme(do_dgnidium~zone_type,random = ~ 1 | estate,
               method="REML",data=variables)

summary(dodgnidium)
```

```
## Linear mixed-effects model fit by REML
## Data: variables
##      AIC      BIC logLik
##  88.482 98.20527 -40.241
##
## Random effects:
## Formula: ~1 | estate
##      (Intercept) Residual
## StdDev: 1.162315e-05 0.3735728
##
## Fixed effects: do_dgnidium ~ zone_type
##              Value Std.Error DF   t-value p-value
## (Intercept) 0.04878049 0.05834227 77 0.8361088  0.4057
## zone_typeecz 0.03921951 0.08065408 77 0.4862682  0.6282
## Correlation:
##              (Intr)
## zone_typeecz -0.723
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -0.2355632 -0.2355632 -0.1305783 -0.1305783  7.0990175
##
## Number of Observations: 86
## Number of Groups: 8
```

#Quercus lusitanica

```
doqlusi<-lme(do_qlusi~zone_type,random = ~ 1 | estate,
            method="REML",data=variables)
```

```
summary(doqlusi)
```

```
## Linear mixed-effects model fit by REML
## Data: variables
##      AIC      BIC    logLik
##  363.898 373.6212 -177.949
##
## Random effects:
## Formula: ~1 | estate
##      (Intercept) Residual
## StdDev:    0.4592242 1.889946
##
## Fixed effects: do_qlusi ~ zone_type
##              Value Std.Error DF   t-value p-value
## (Intercept)  0.5038348 0.3488302 77   1.4443555   0.1527
## zone_typeecz -0.3072411 0.4087961 77  -0.7515755   0.4546
## Correlation:
##              (Intr)
## zone_typeecz -0.616
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -0.60795933 -0.21811503 -0.10495591 -0.04912401  8.33409479
##
## Number of Observations: 86
## Number of Groups: 8
```

```
#Pterospartum tridentatum
```

```
doptrident<-lme(do_ptrident~zone_type,random = ~ 1 | estate,
               method="REML",data=variables)
```

```
summary(doptrident)
```

```
## Linear mixed-effects model fit by REML
## Data: variables
##      AIC      BIC    logLik
##   38.70974 48.433 -15.35487
##
## Random effects:
## Formula: ~1 | estate
##      (Intercept) Residual
## StdDev:    0.1582647 0.2616446
##
## Fixed effects: do_ptrident ~ zone_type
##              Value Std.Error DF   t-value p-value
## (Intercept)  0.0232342 0.07125275 77  0.3260814   0.7452
## zone_typeecz 0.0767394 0.05671221 77  1.3531370   0.1800
## Correlation:
##              (Intr)
## zone_typeecz -0.418
##
```



```
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -1.6733076 -0.1718270 -0.1639974  0.1214693  8.1873957
##
## Number of Observations: 86
## Number of Groups: 8
```

#Rubia peregrina

```
dorper<-lme(do_rper~zone_type,random = ~ 1 | estate,
            method="REML",data=variables)
```

```
summary(dorper)
```

```
## Linear mixed-effects model fit by REML
## Data: variables
##      AIC      BIC    logLik
## -247.9089 -238.1857 127.9545
##
## Random effects:
## Formula: ~1 | estate
##      (Intercept)  Residual
## StdDev:  0.03445024 0.04691475
##
## Fixed effects: do_rper ~ zone_type
##              Value Std.Error DF   t-value p-value
## (Intercept) 0.005038731 0.01453862 77  0.3465756  0.7299
## zone_typeecz 0.015792178 0.01017204 77  1.5525079  0.1246
## Correlation:
##      (Intr)
## zone_typeecz -0.368
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -2.0787571 -0.1864993 -0.1688958  0.1501151  7.2999575
##
## Number of Observations: 86
## Number of Groups: 8
```

#Hallimium ocymoides

```
dohocym<-lme(do_hocym~zone_type,random = ~ 1 | estate,
             method="REML",data=variables)
```

```
summary(dohocym)
```

```
## Linear mixed-effects model fit by REML
## Data: variables
##      AIC      BIC    logLik
## -1.358391 8.364877 4.679195
##
## Random effects:
## Formula: ~1 | estate
```

```
##          (Intercept) Residual
## StdDev:  0.07304539 0.2123015
##
## Fixed effects: do_hocym ~ zone_type
##              Value Std.Error DF   t-value p-value
## (Intercept) 0.01056335 0.04364872 77 0.2420083  0.8094
## zone_typeecz 0.04634108 0.04595971 77 1.0082980  0.3165
## Correlation:
##          (Intr)
## zone_typeecz -0.554
##
## Standardized Within-Group Residuals:
##          Min          Q1          Med          Q3          Max
## -0.86714922 -0.16730519 -0.14405121  0.06068243  8.74182658
##
## Number of Observations: 86
## Number of Groups: 8
```

#Erica australis

```
doeaustralis<-lme(do_eaustralis~zone_type,random = ~ 1 | estate,
                  method="REML",data=variables)

summary(doeaustralis)
```

```
## Linear mixed-effects model fit by REML
## Data: variables
##      AIC      BIC   logLik
##  4.378291 14.10156 1.810854
##
## Random effects:
## Formula: ~1 | estate
##          (Intercept) Residual
## StdDev:  0.03191157 0.2247855
##
## Fixed effects: do_eaustralis ~ zone_type
##              Value Std.Error DF   t-value p-value
## (Intercept) 0.00370717 0.03766752 77 0.0984182  0.9219
## zone_typeecz 0.04444448 0.04857252 77 0.9150129  0.3630
## Correlation:
##          (Intr)
## zone_typeecz -0.677
##
## Standardized Within-Group Residuals:
##          Min          Q1          Med          Q3          Max
## -0.35081586 -0.18329140 -0.17074713  0.01442808  8.99142183
##
## Number of Observations: 86
## Number of Groups: 8
```

#Pistacia lentiscus

```
doplentiscus<-lme(do_plentiscus~zone_type,random = ~ 1 | estate,
```

```

method="REML",data=variables)

summary(doplentiscus)

```

```

## Linear mixed-effects model fit by REML
## Data: variables
##      AIC      BIC    logLik
## -266.1013 -256.378 137.0507
##
## Random effects:
## Formula: ~1 | estate
##      (Intercept)  Residual
## StdDev: 0.001158087 0.04525124
##
## Fixed effects: do_plentiscus ~ zone_type
##              Value Std.Error DF   t-value p-value
## (Intercept)  0.01024609 0.007086303 77   1.445901  0.1523
## zone_typeecz -0.01024763 0.009770041 77  -1.048883  0.2975
## Correlation:
##      (Intr)
## zone_typeecz -0.722
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -0.2312657100 -0.2259836012 -0.0048049773  0.0004771315  9.0502479514
##
## Number of Observations: 86
## Number of Groups: 8

```

#Myrtus communis

```

domcomm<-lme(do_mcomm~zone_type,random = ~ 1 | estate,
             method="REML",data=variables)

summary(domcomm)

```

```

## Linear mixed-effects model fit by REML
## Data: variables
##      AIC      BIC    logLik
##  145.1704 154.8937 -68.58522
##
## Random effects:
## Formula: ~1 | estate
##      (Intercept)  Residual
## StdDev:  0.1747358 0.5078585
##
## Fixed effects: do_mcomm ~ zone_type
##              Value Std.Error DF   t-value p-value
## (Intercept)  0.02526916 0.1044145 77  0.2420081  0.8094
## zone_typeecz 0.11085514 0.1099428 77  1.0082979  0.3165
## Correlation:
##      (Intr)

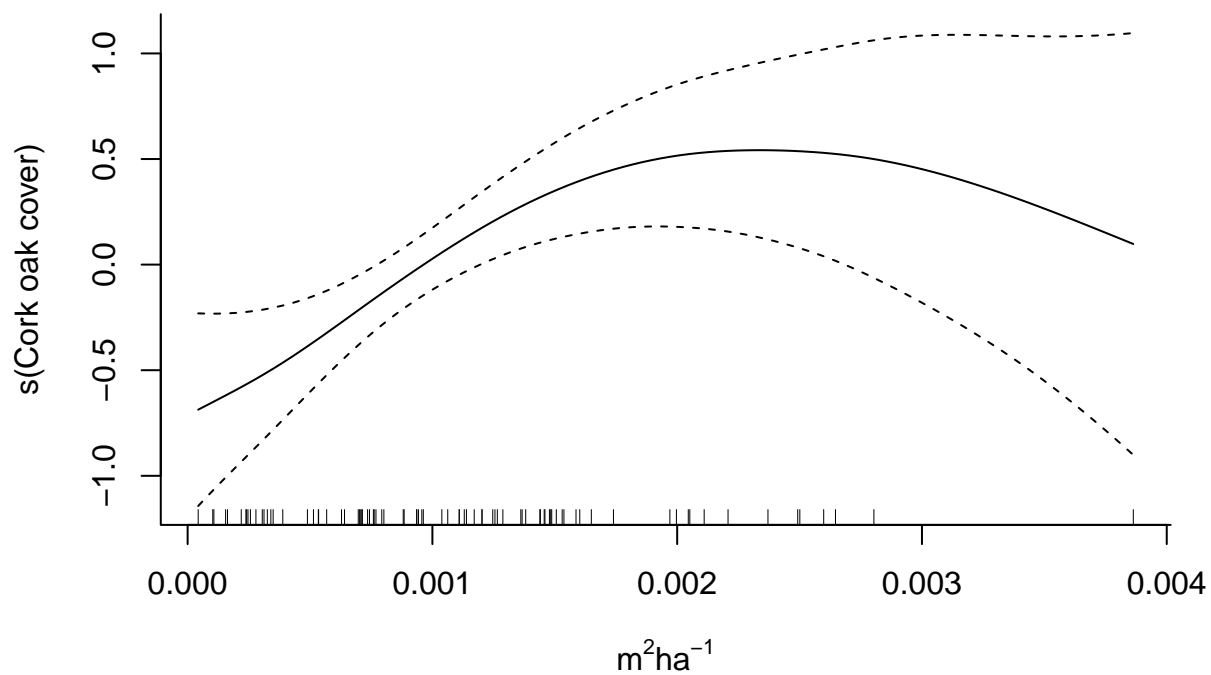
```

```
## zone_type cz -0.554
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -0.86714790 -0.16730531 -0.14405127  0.06068231  8.74182721
##
## Number of Observations: 86
## Number of Groups: 8
```

8. Code for reproducing the figures

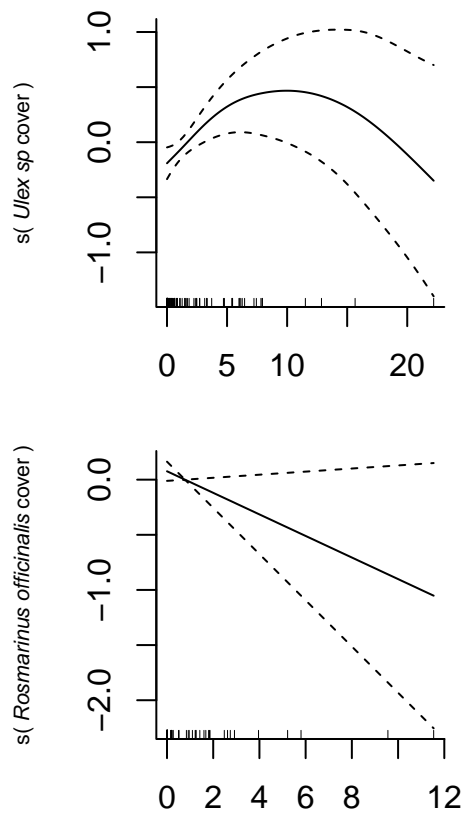
8.1 Figure 2

```
plot(modseed_tw,select=1,scale=0,ylab="s(Cork oak cover)",xlab=expression('m'^2*'ha'^-1),bty="l")
```



8.2 Figure 3

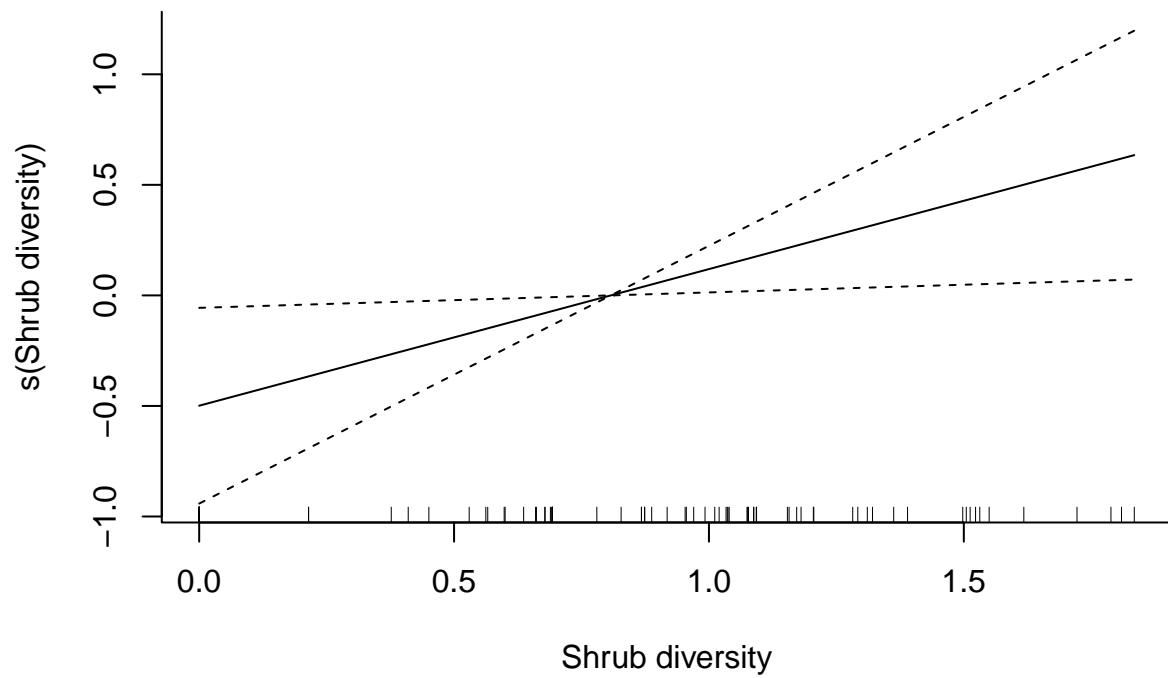
```
par(mfrow=c(2,1),pin=c(1.5,1.5))
plot(modsap_tw,select=1,scale=0,ylab= expression("s(~italic("Ulex sp")~"cover"~")"),
     xlab=expression(italic("Ulex sp")~"cover"),bty="l",cex.lab=0.7)
plot(modsap_tw,select=2,scale=0,ylab= expression("s(~italic("Rosmarinus officinalis")~"cover"~")"),
     xlab=expression(italic("Rosmarinus officinalis")~"cover"),bty="l",cex.lab=0.7)
```



```
par(mfrow=c(1,1))
```

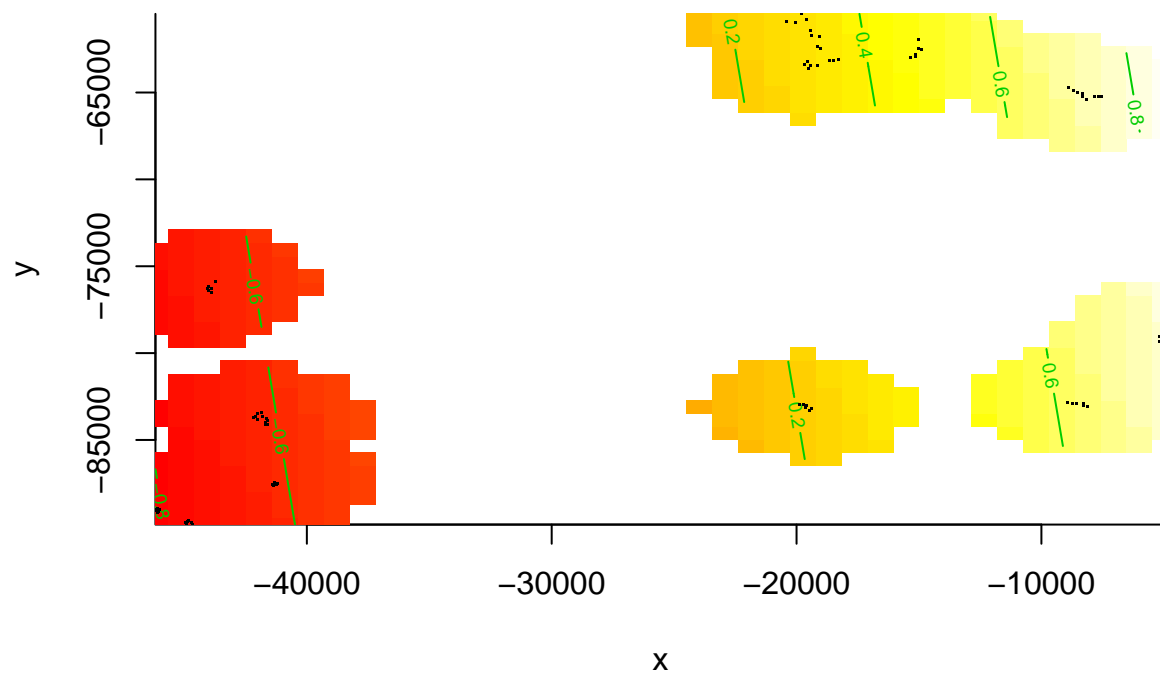
8.3 Figure 4

```
plot(modyt_tw,select=9,scale=0,ylab="s(Shrub diversity)",xlab="Shrub diversity",bty="l",cex.lab=1)
```



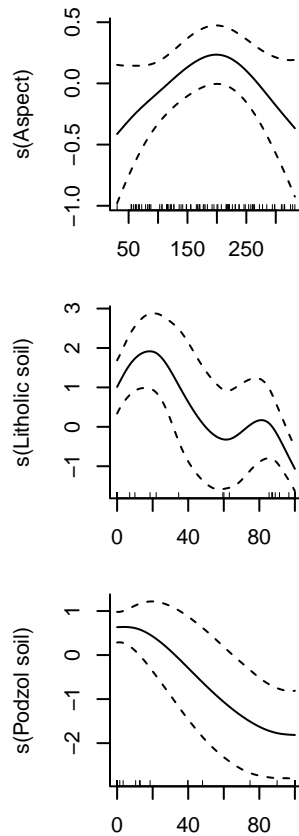
8.4 Figure 5

```
plot(modsapgis_tw,select=1,scheme=2,scale=0,main="",bty="l",cex.lab=1)
```



8.4 Figure 6

```
par(mfrow=c(3,1),pin=c(1,1))
plot(modytgis_tw,select=2,scheme=2,scale=0,xlab="Aspect (degrees)",ylab="s(Aspect)",bty="l",cex.lab=1)
plot(modytgis_tw,select=5,scheme=2,scale=0,xlab="Litholic soil (% of cover)",ylab="s(Litholic soil)",bty="l",cex.lab=1)
plot(modytgis_tw,select=6,scheme=2,scale=0,xlab="Podzol soil (% of cover)",ylab="s(Podzol soil)",bty="l",cex.lab=1)
```



```
par(mfrow=c(1,1))
```

9. References

- Miller, D.L., Burt, M.L., Rexstad, E.A. & Thomas, L. (2013) Spatial models for distance sampling data: recent developments and future directions. *Methods in Ecology and Evolution*, 4, 1001–1010.
- Marra, G., and S. N. Wood. 2011. Practical variable selection for generalized additive models. *Computational Statistics & Data Analysis* 55:2372–2387.
- Winiarski, K., Miller, D., Paton, P. & McWilliams, S. (2013) Spatially explicit model of wintering common loons: conservation implications. *Marine Ecology Progress Series*, 492, 273–283.
- Wood, S. 2006. *Generalized Additive Models: An Introduction with R*. First edition. Chapman and Hall/CRC, Boca Raton, FL.
- Xie, Y. (2013) *Dynamic Documents with R and Knitr*, 1 edition. Chapman and Hall/CRC, Boca Raton.

Zuur, A., E. N. Ieno, N. Walker, A. A. Saveliev, and G. M. Smith. 2009. Mixed Effects Models and Extensions in Ecology with R. Springer, New York, USA. Code available at: <http://www.highstat.com/BGS/GAM/HighstatLibV4.R>