

Hinweise zum Plotten von GAMs

Im folgenden sollen ein paar Hinweise zum Plotten des GAM-Modells gegeben werden. Als Beispiel dient der Datensatz `ghuepfer001.txt` aus der 2. Hausübung. Das zugehörige GAM-Modell heie `g1` (Hinweis: Hier sind noch viele Variablen enthalten. In der Hausübung blieben Ihnen ggf. weniger Variablen übrig nach der Variablenselektion):

```
summary(g1)
```

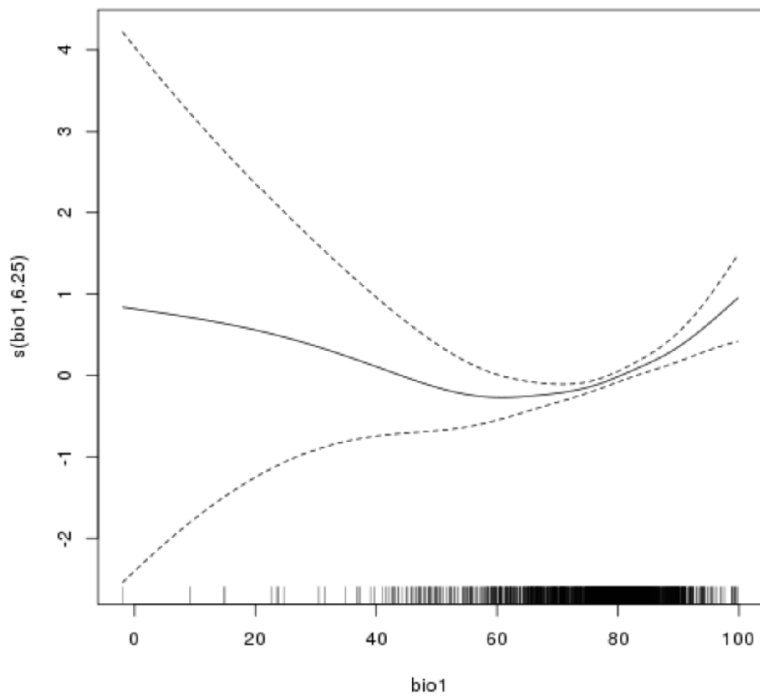
```
##
## Family: gaussian
## Link function: identity
##
## Formula:
## OMOCESTUS.HAEMORRHOIDALIS ~ s(bio1) + s(bio2) + s(bio5) + s(bio6) +
##   s(bio7) + s(Acker) + s(Felsen) + s(Laubwald) + s(Nadelwald) +
##   s(SAWald) + s(Sumpf) + s(WaldrandGeb) + s(WasserFl) + s(Wiesen)
##
## Parametric coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.11365    0.00686   16.6   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##             edf Ref.df   F p-value
## s(bio1)      6.25  7.32 3.29 0.00153 **
## s(bio2)      2.46  3.13 1.88 0.12860
## s(bio5)      4.53  5.76 1.91 0.07838 .
## s(bio6)      7.59  8.15 4.15 5.6e-05 ***
## s(bio7)      2.38  3.11 0.74 0.53203
## s(Acker)     1.91  2.41 4.20 0.01084 *
## s(Felsen)    1.00  1.00 0.10 0.74800
## s(Laubwald)  1.00  1.00 4.79 0.02868 *
## s(Nadelwald) 2.81  3.54 6.65 8.1e-05 ***
## s(SAWald)    2.86  3.60 4.43 0.00244 **
## s(Sumpf)     1.00  1.00 0.54 0.46110
## s(WaldrandGeb) 2.76  3.40 0.67 0.58585
## s(WasserFl)  1.77  2.20 3.58 0.02472 *
## s(Wiesen)    4.05  5.01 4.40 0.00054 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Rank: 126/127
## R-sq.(adj) = 0.0997   Deviance explained = 11.9%
## GCV = 0.092828   Scale est. = 0.090739   n = 1927
```

Nachfolgend sollen die einzelnen nichtlinearen Effekte dargestellt werden. Dabei können Sie zum einen die Option `scale=0` als auch `select=ZAHL` verwenden. Ersteres hat zur Folge, dass nicht alle Kurven auf eine gemeinsame y-Skala gezeichnet werden. Letzteres wählt aus, welche Kurve Sie gerade zeichnen lassen wollen: `select=1` wählt die erste Kurve aus, `select=4` die 4., usw...

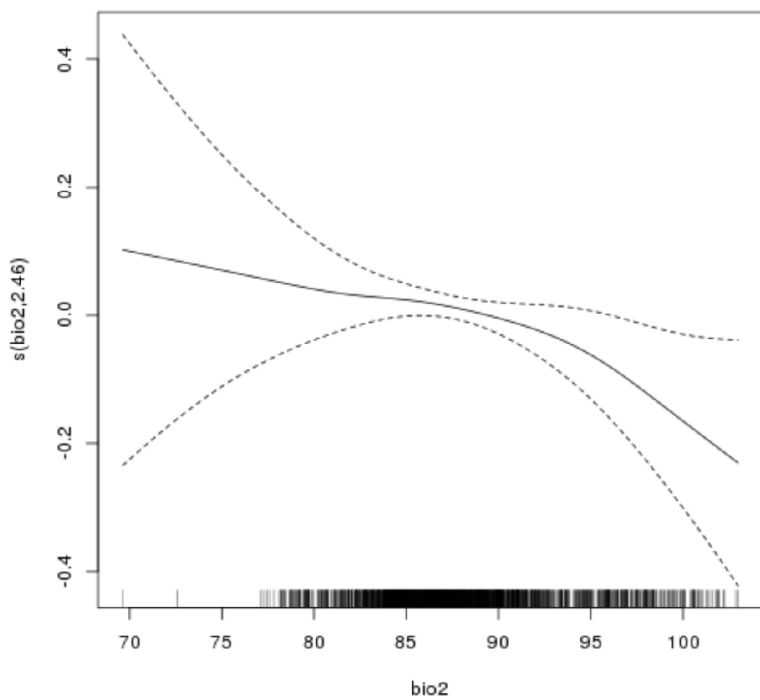
bio1, bio2, bio5, bio6

So werden die ersten 4 Kurven gezeichnet. Dabei passiert noch nichts ungewöhnliches.

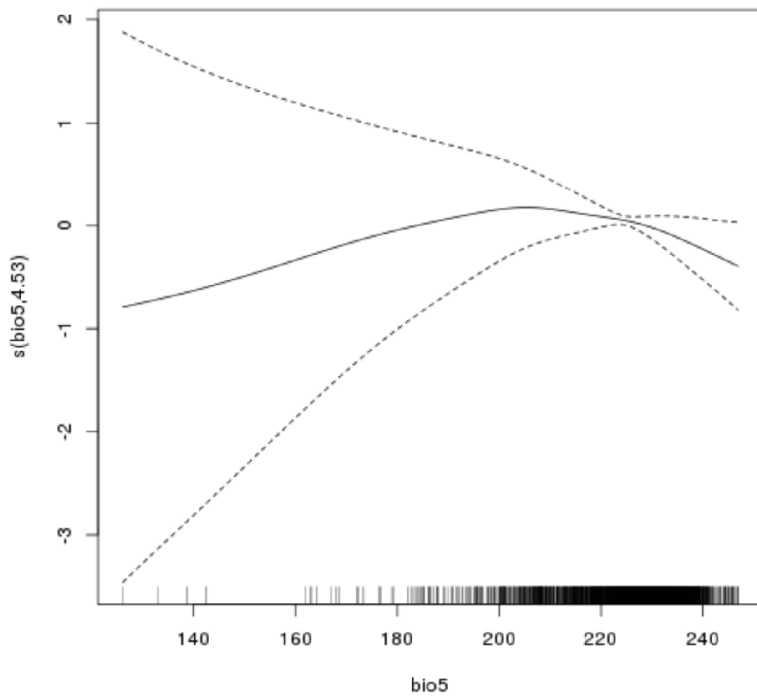
```
plot(g1, select=1, scale=0)
```



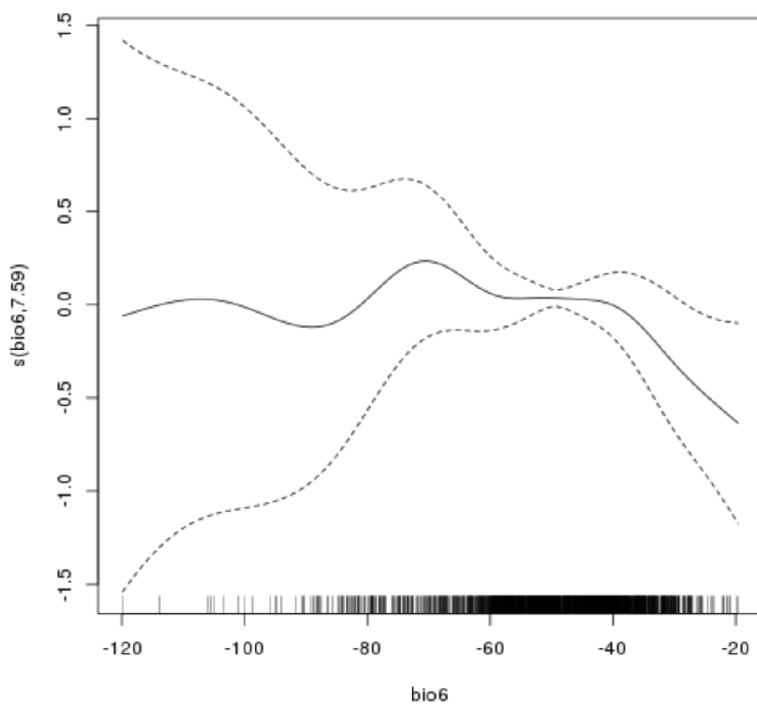
```
plot(g1, select=2, scale=0)
```



```
plot(g1, select=3, scale=0)
```



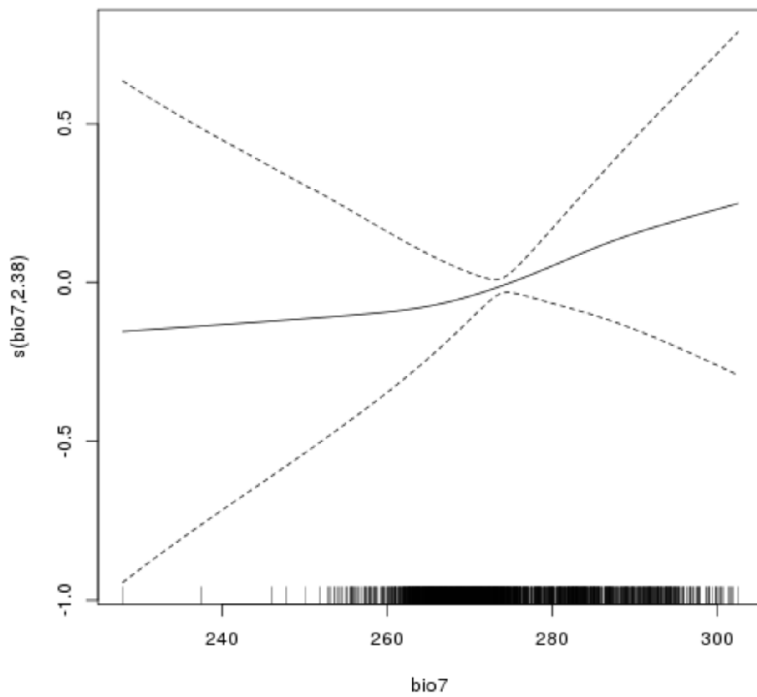
```
plot(g1, select=4, scale=0)
```



bio7

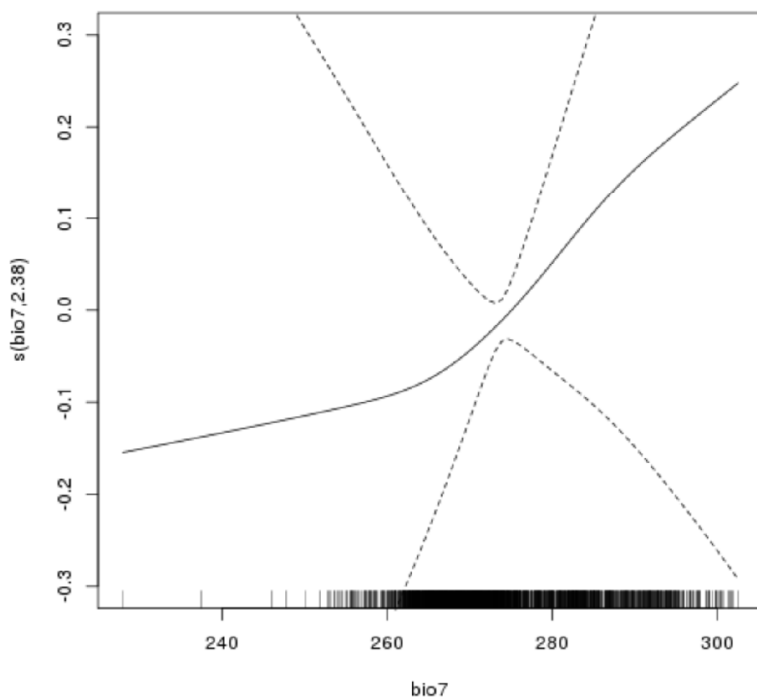
Hier sind die Konfidenzbänder nun so groß, so dass die funktionale Form nicht mehr erkennbar ist.

```
plot(g1, select=5, scale=0)
```



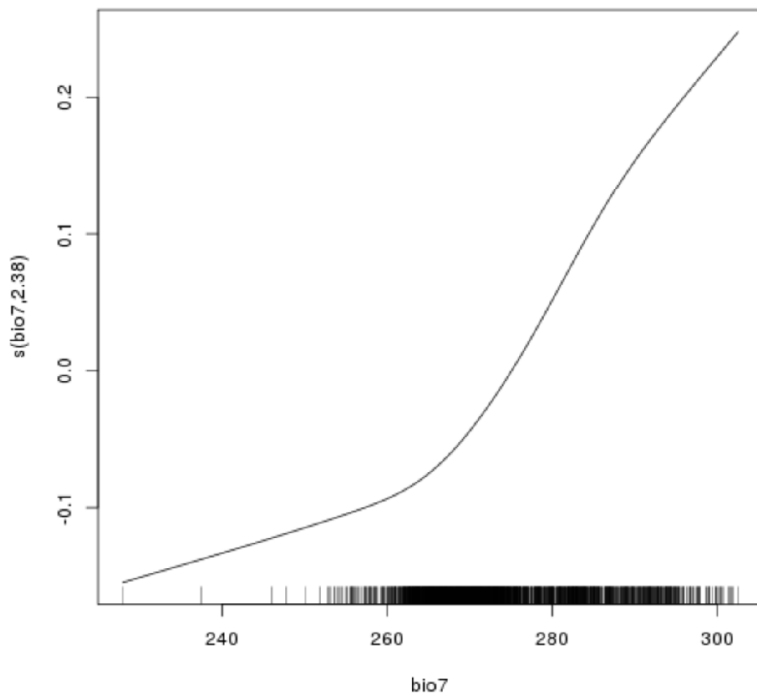
Als Abhilfe könnte man die y-Skala verkleinern, z.B. von -0.3 bis +0.3. Die Option *ylim* bekommt jeweils die untere und obere Grenze des zu zeichnenden Wertebereichs (ggf. einfach verschiedene Wertepaare ausprobieren).

```
plot(g1, select=5, scale=0, ylim=c(-0.3, 0.3))
```



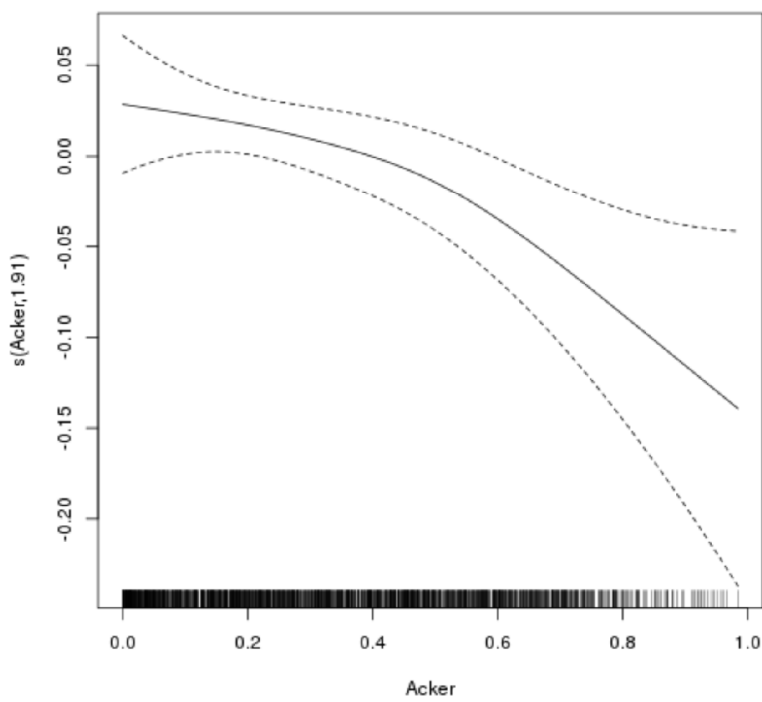
Oder die Konfidenzbänder ausschalten (sollte man nur im Ausnahmefall machen).

```
plot(g1, select=5, scale=0, se=FALSE)
```

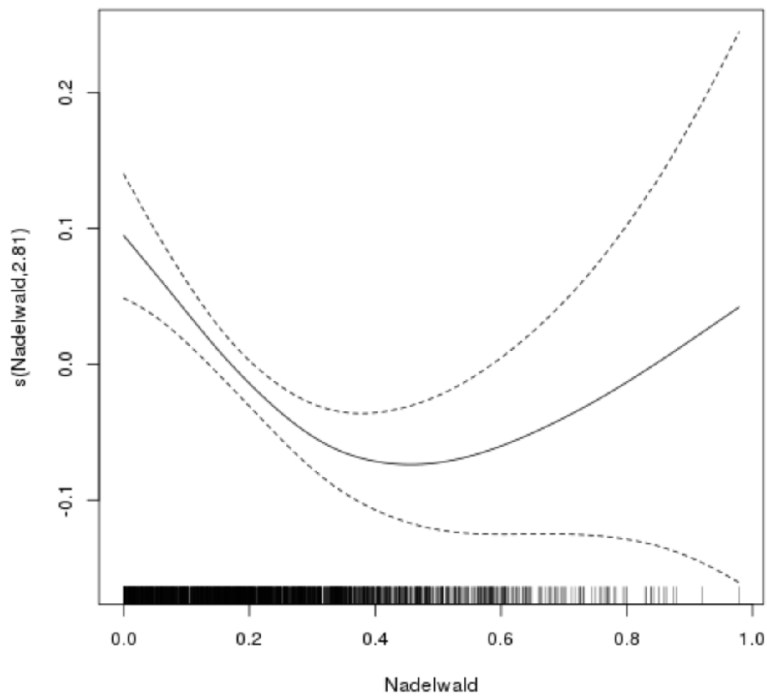


Acker, Nadelwald, SAWald

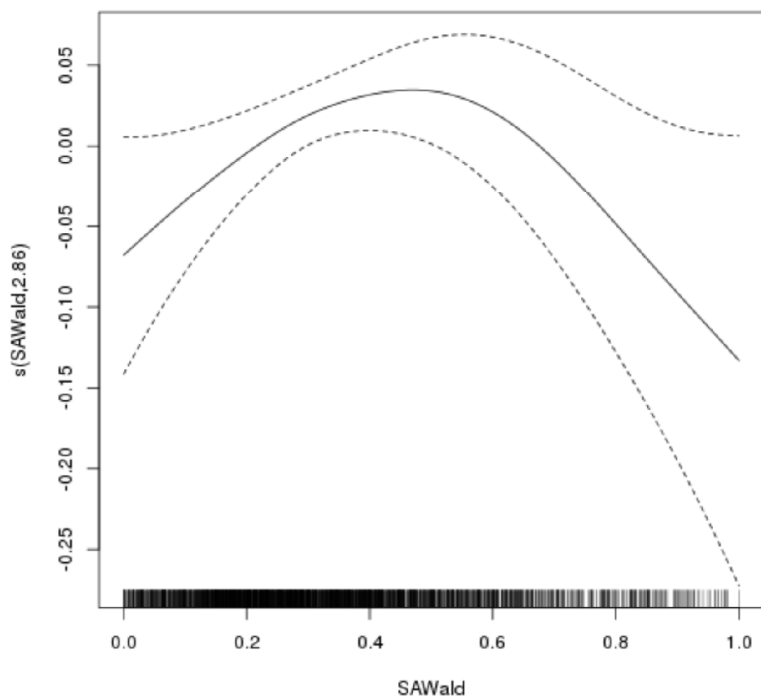
```
plot(g1, select=6, scale=0)
```



```
plot(g1, select=9, scale=0)
```



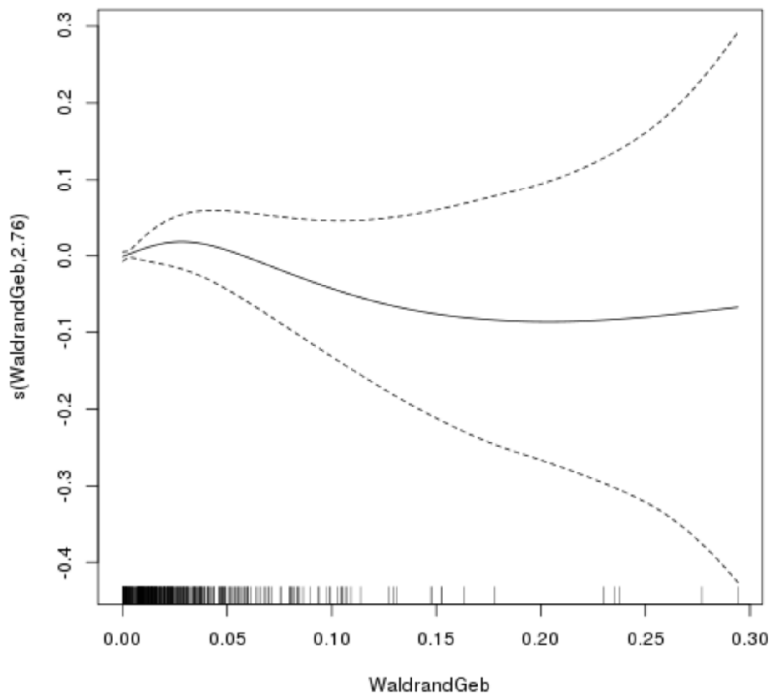
```
plot(g1, select=10, scale=0)
```



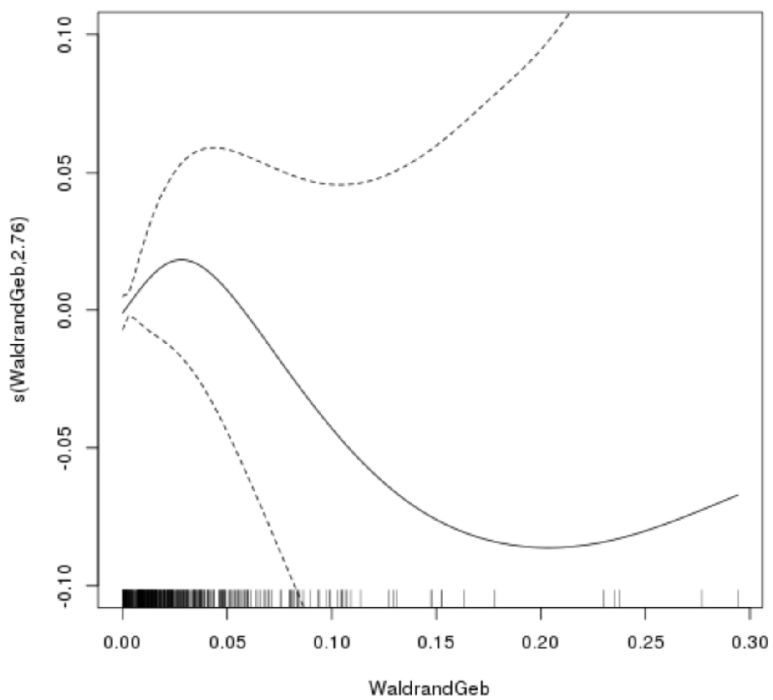
WaldrandGeb, WasserFI

Hier ließe sich evtl wieder die Skala verkleinern.

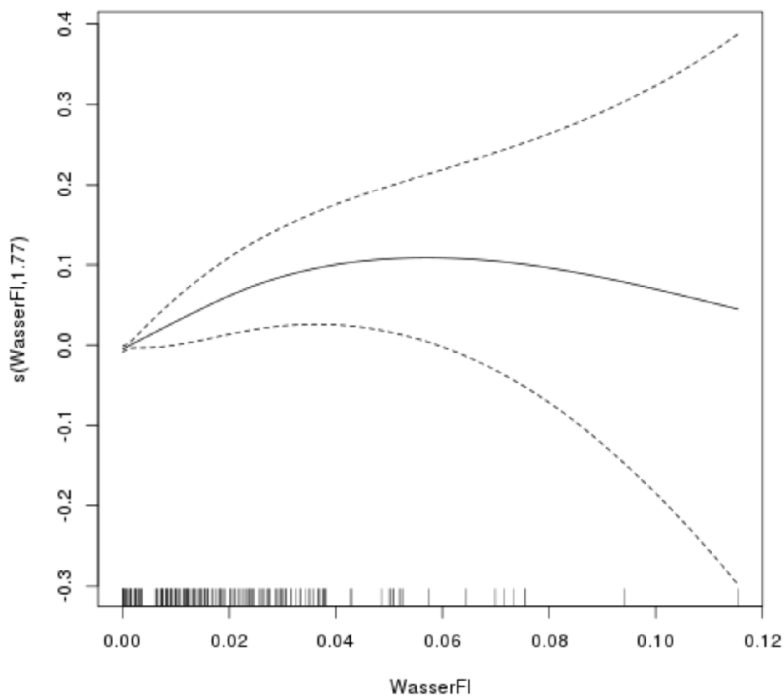
```
plot(g1, select=12, scale=0)
```



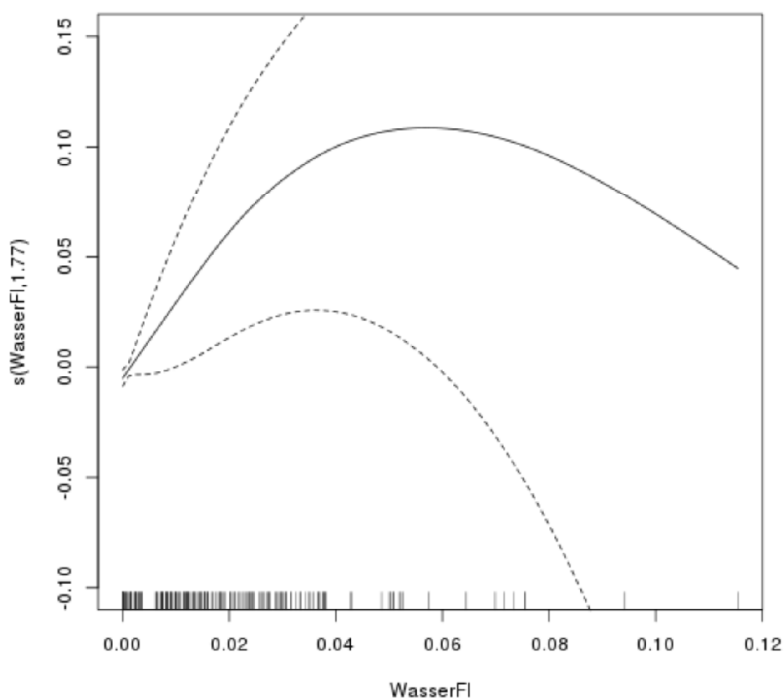
```
plot(g1, select=12, scale=0, ylim=c(-0.1, 0.1))
```



```
plot(g1, select=13, scale=0)
```



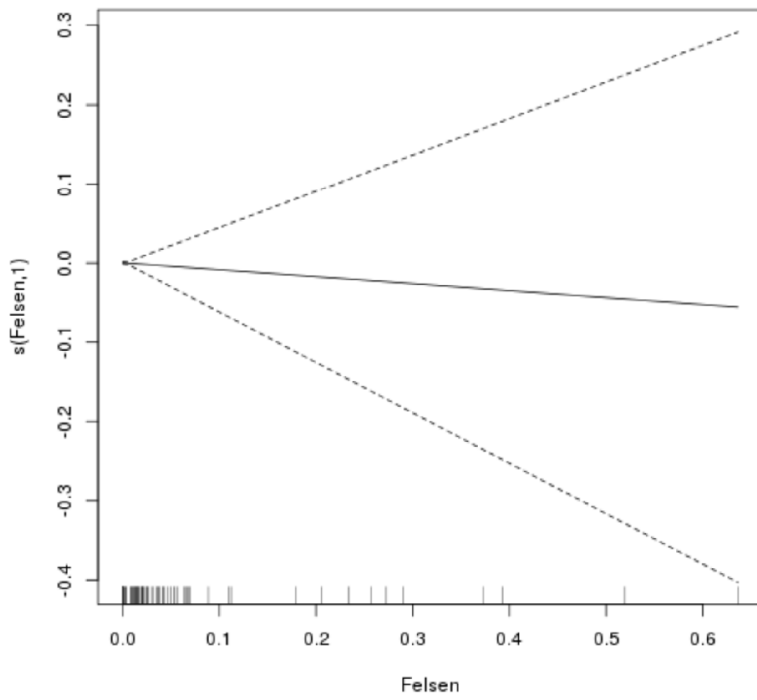
```
plot(g1, select=13, scale=0, ylim=c(-0.1,0.15))
```



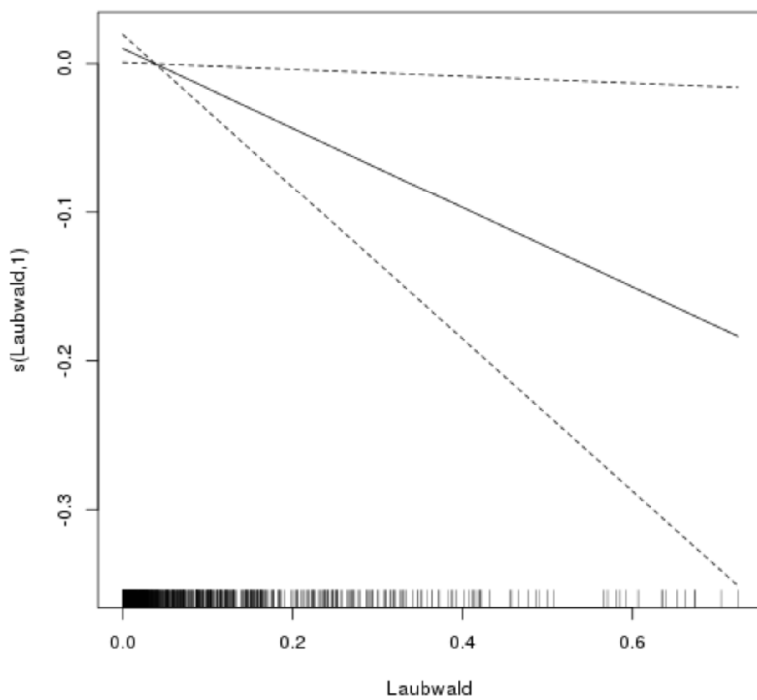
Felsen, Laubwald, Sumpf

Diese Variablen werden als Geraden geschätzt ($edf=1$).

```
plot(g1, select=7, scale=0)
```

```
plot(g1, select=8, scale=0)
```



```
plot(g1, select=11, scale=0)
```

