The data structure

```
rm(list = ls())
# Import data
table_habitat <- read.csv("Willamette_habitat_features.csv")

# fix some wrong values
table_habitat[42,]$Slope <- 0.000732
table_habitat[42,]$Floodplain_elevation<- 4.214
table_habitat[41,]$Slope <- 0.000586

# Add tiny value for NA
table_habitat[is.na(table_habitat)] <- 1e-8
# Reasign the index
table_habitat <-table_habitat[order(table_habitat$RKM_2008,decreasing = F),]
table_habitat$RKM_2008 <- 1:178

# Add new columns
table_habitat <- table_habitat%>%mutate(ConnectedWet_area=AllWetArea-DisconnectedWater_Area)%>% #Creat ConnectedWet_area
                                 mutate(perc_1_2m=perc_2m-perc_1m)%>% #Creat pure Area_2m
                                 mutate(Habitat_level=as.integer(ntile(table_habitat$Habitat_area, 3))) #Creat Habitat Area leve

# Change to short names
original_name <- names(table_habitat)
names(table_habitat) <- c("No","H_A","D1_A","D2_A","D1_P","D2_P","W_m_A","W_s_A","W_a_A","L_b_A","L_v_A","W_d_A","W_ia_A","W_r_A
table <- rbind(original_name[1:8],names(table_habitat)[1:8],
               original_name[9:16],names(table_habitat)[9:16],
               original_name[17:23],names(table_habitat)[17:23])
pander(table <- cbind(original_name[1:23],names(table_habitat)[1:23]))
```

RKM_2008	No
Habitat_area	H_A
area_1m	D1_A
area_2m	D2_A
perc_1m	D1_P
perc_2m	D2_P
MainChannel_Area	W_m_A
SideChannel_Area	W_s_A
Alcove_Area	W_a_A
BareBar_Area	L_b_A
VegetatedBar_Area	L_v_A
DisconnectedWater_Area	W_d_A
InverseAlcove_Area	W_ia_A
Bedrock_Area	W_r_A
AllWetLength	W_L
AllWetArea	W_A
MainChannelLength	W_m_L
Slope	S
Floodplain_elevation	FE
Polygon_area	A
ConnectedWet_area	W_c_A
perc_1_2m	D12_P
Habitat_level	H_A_L

```
table_habitat_perc<- table_habitat[,-c(3,4)]
table_habitat_perc[,c(2, 5:12,14,18,19)]<- table_habitat[,c(2, 7:14,16,20,21)]/table_habitat[,21]
names(table_habitat_perc) <- c("No","H_P","D1_P","D2_P","W_m_P","W_s_P","W_a_P","L_b_P","L_v_P","W_d_P","W_ia_P","W_r_P","W_L","
glimpse(table_habitat_perc)
```

```
## Observations: 178
## Variables: 21
## $ No     <int> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 1...
## $ H_P    <dbl> 0.219, 0.254, 0.267, 0.236, 0.241, 0.201, 0.200, 0.140, 0.20...
```

```
## $ D1_P   <dbl> 0.727, 0.634, 1.632, 0.418, 0.259, 0.817, 0.911, 0.894, 0.74...
## $ D2_P   <dbl> 1.477, 1.056, 2.198, 0.997, 0.494, 1.486, 1.828, 1.356, 1.22...
## $ W_m_P  <dbl> 0.896, 0.714, 0.618, 0.965, 1.000, 1.000, 0.964, 0.758, 0.63...
## $ W_s_P  <dbl> 4.96e-03, 2.22e-01, 3.68e-01, 3.48e-02, 8.44e-14, 6.63e-14, ...
## $ W_a_P  <dbl> 9.86e-02, 6.37e-02, 1.35e-02, 6.95e-14, 8.44e-14, 6.63e-14, ...
## $ L_b_P  <dbl> 0.1437, 0.4392, 0.1281, 0.1423, 0.0952, 0.5227, 0.2645, 0.07...
## $ L_v_P  <dbl> 0.588, 2.189, 1.234, 0.400, 1.807, 0.801, 0.498, 0.893, 1.08...
## $ W_d_P  <dbl> 1.82e-02, 1.94e-02, 4.31e-14, 6.95e-14, 3.04e-01, 5.35e-03, ...
## $ W_ia_P <dbl> 6.59e-14, 5.45e-14, 4.31e-14, 6.95e-14, 8.44e-14, 6.63e-14, ...
## $ W_r_P  <dbl> 6.59e-14, 5.45e-14, 4.31e-14, 6.95e-14, 8.44e-14, 6.63e-14, ...
## $ W_L    <dbl> 3362, 5969, 7008, 2292, 3593, 3615, 3218, 5039, 5717, 2143, ...
## $ W_P    <dbl> 1.02, 1.02, 1.00, 1.00, 1.30, 1.01, 1.00, 1.02, 1.03, 1.00, ...
## $ W_m_L  <dbl> 1945, 2265, 2036, 2038, 2050, 3520, 2678, 2376, 2112, 2067, ...
## $ S      <dbl> 0.0001454, 0.0003766, 0.0007078, 0.0000534, 0.0000229, 0.000...
## $ FE     <dbl> 7.19, 8.32, 7.46, 10.26, 6.16, 7.58, 8.58, 9.45, 7.06, 9.22,...
## $ P      <dbl> 1.75, 3.65, 2.36, 1.54, 3.21, 2.33, 1.76, 1.99, 2.52, 1.39, ...
## $ W_c_P  <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, ...
## $ D12_P  <dbl> 0.750, 0.422, 0.566, 0.579, 0.234, 0.669, 0.917, 0.461, 0.48...
## $ H_A_L  <int> 2, 3, 3, 2, 1, 2, 2, 1, 2, 1, 3, 3, 3, 2, 2, 2, 2, 1, 1, 1, ...
```

```
# Normalize the variables
table_habitat[,3:22] <- scale(table_habitat[,3:22], center = T, scale = F)
table_habitat_perc[,3:20] <- scale(table_habitat_perc[,3:20], center = T, scale = F)


# Remove some variables
# table_habitat_16 <- table_habitat[,c(1,2,3,4,7,8,9,10,11,13,14,15,17,18,19,20,21,22)]
# glimpse(table_habitat_16)

# table_habitat_category <- table_habitat[,c(23,1,4,7,8,9,10,11,13,14,15,17,18,19,20)]
```

fitting full model

- The full model

```
# ols_regress(model_full)
# summary(model_full)
```

- multicollinearity Diagnostics

According to the result of VIF test (variance inflation factor), the model does have **serious problems of multicollinearity**. The VIF of variables **Polygon__area, BareBar__Area, VegetatedBar__Area, Bedrock__Area, AllWetArea** are huge.

Only **MainChannelLength, InverseAlcove__Area, Slope, Floodplain__elevation** are smaller than 10.

Elimination regression

- Stepwise Variable selection

Removing any predictor can draw down the VIF. We can take more diagnostics and comparisons, gather sufficient evidents to decide the final elimination plan.

Use Stepwise AIC Regression

```
# Stepwise AIC Regression
k <- ols_step_both_aic(model_full)
```

```
## Stepwise Selection Method
## -------------------------
##
## Candidate Terms:
##
## 1 . No
## 2 . D1_A
```

```
## 3 . D2_A
## 4 . D1_P
## 5 . D2_P
## 6 . W_m_A
## 7 . W_s_A
## 8 . W_a_A
## 9 . L_b_A
## 10 . L_v_A
## 11 . W_d_A
## 12 . W_ia_A
## 13 . W_r_A
## 14 . W_L
## 15 . W_A
## 16 . W_m_L
## 17 . S
## 18 . FE
## 19 . A
## 20 . W_c_A
## 21 . D12_P
##
##
## Variables Entered/Removed:
##
## - W_c_A added
## - W_s_A added
## - No added
## - S added
## - W_L added
## - W_m_A added
## - A added
##
## No more variables to be added or removed.
```

```
# plot(k,cex=0.2)
```

Use Stepwise Regression based on p values (use alpha=0.05)

```
# Stepwise Regression based on p values
k <- ols_step_both_p(model_full)
```

```
## Stepwise Selection Method
## ---------------------------
##
## Candidate Terms:
##
## 1. No
## 2. D1_A
## 3. D2_A
## 4. D1_P
## 5. D2_P
## 6. W_m_A
## 7. W_s_A
## 8. W_a_A
## 9. L_b_A
## 10. L_v_A
## 11. W_d_A
## 12. W_ia_A
## 13. W_r_A
## 14. W_L
## 15. W_A
## 16. W_m_L
## 17. S
## 18. FE
## 19. A
## 20. W_c_A
## 21. D12_P
```

```
##
## We are selecting variables based on p value...
##
## Variables Entered/Removed:
##
## - W_c_A added
## - W_s_A added
## - No added
## - S added
##
## No more variables to be added/removed.
##
##
## Final Model Output
## ------------------
##
##                           Model Summary
## --------------------------------------------------------------------
## R                        0.630       RMSE                  9672.584
## R-Squared                0.397       Coef. Var               27.746
## Adj. R-Squared           0.383       MSE               93558882.241
## Pred R-Squared           0.353       MAE                   7042.161
## --------------------------------------------------------------------
##  RMSE: Root Mean Square Error
##  MSE: Mean Square Error
##  MAE: Mean Absolute Error
##
##                                      ANOVA
## -------------------------------------------------------------------------------
##                       Sum of
##                       Squares         DF       Mean Square      F         Sig.
## -------------------------------------------------------------------------------
## Regression    10661385744.598          4    2665346436.149    28.488    0.0000
## Residual      16185686627.609        173      93558882.241
## Total         26847072372.206        177
## -------------------------------------------------------------------------------
##
##                                             Parameter Estimates
## ----------------------------------------------------------------------------------------------------------
##       model            Beta     Std. Error     Std. Beta       t        Sig            lower          upper
## ----------------------------------------------------------------------------------------------------------
## (Intercept)       29236.938       2021.294                  14.464    0.000        25247.366      33226.510
##       W_c_A           0.188          0.023         0.609     8.096    0.000            0.142          0.233
##       W_s_A           0.231          0.053         0.282     4.331    0.000            0.126          0.336
##          No          62.838         21.082         0.263     2.981    0.003           21.228        104.448
##           S    -6611334.155    2725873.064        -0.195    -2.425    0.016    -11991584.276   -1231084.033
## ----------------------------------------------------------------------------------------------------------
```

```r
# plot(k)
```

Bayesian Feature selection

```r
ind.insample <- sample(1:178,120)
X <- data.frame(table_habitat[-c(2,23)])
# y <- table_habitat[23]
y <- table_habitat[2]

y_perc <- y/X$W_c_A
X_perc <- X[c(6:13,15,19,20)]/X$W_c_A
X_perc[12:19] <- X[c(1,4,5,14,16:18,21)]
names(X_perc) <- c("W_m_P","W_s_P","W_a_P","L_b_P","L_v_P","W_d_P","W_ia_P","W_r_P","W_P","P","W_c_P","No","D1_P","D2_P","W_L","
glimpse(X_perc)

source("VarSelectHC.R")
source("summaryout.R")
```

```
# 4 essencial variables (level 1)
# vbase <- c("W_m_A","W_s_A","W_a_A","W_ia_A")
## vbase <- c("W_c_A","W_s_A","No","S")
# 7 Length and other geographical variables (level 2)
# vtest <- c("D1_P","D2_P","W_r_A","W_L","W_m_L","S","F")
## vtest <- c("W_m_A","W_L","A")
# 7 Length no fixed
vtest <- c("W_c_A","W_s_A","No","S","W_m_A","W_L","A")
#----------------------------------------------------------------
#with habitat area as a response
datain <- data.frame(y=y[ind.insample,],X[ind.insample,vtest])  # c(vbase,vtest)
data.holdout <- data.frame(y=y[-ind.insample,],X[-ind.insample,vtest])  # c(vbase,vtest)
modpriorvec=c("HOP","HIP","HUP")

# baseformula <- as.formula(paste(".~ ",paste0(vbase,collapse="+")))
theformula <- as.formula(paste("y ~",paste0(vtest,collapse="+"))) # c(vbase,vtest)

res=VarSelectHC(full.formula=theformula,
                data=datain,
                base.formula=as.formula(. ~ 1),#baseformula,#
                maxdeg=2,
                nodes.to.remove=NULL,
                SH = T,
                model.prior.type=modpriorvec,
                model.prior.pars = "children",
                beta.prior.type = "IP",
                beta.prior.pars = list(alpha=1,nu=1),
                niter=5000)

summary.res <- summaryout(mcmc.out=res,insampledata=datain,modelprior.nams=modpriorvec,
                          shr.adj=T,outsampledata=data.holdout,respnam="y",top.ave=10,betaprtype="IP",
                          parsprbeta=list(alpha=1,nu=1))

#----------------------------------------------------------------
vtest <- c("W_c_P","W_s_P","No","S","W_m_P","W_L","P")
#with proportion of habitat area and other variables
datain.prop <- data.frame(y=y_perc[ind.insample,],X_perc[ind.insample,vtest]) # c(vbase,vtest)
data.holdout.prop <- data.frame(y=y_perc[-ind.insample,],X_perc[-ind.insample,vtest]) # c(vbase,vtest)

theformula <- as.formula(paste("y ~",paste0(vtest,collapse="+"))) # c(vbase,vtest)

res.prop=VarSelectHC(full.formula=theformula,
                 data=datain.prop,
                 base.formula=as.formula(. ~ 1),#baseformula,#
                 maxdeg=2,
                 nodes.to.remove=NULL,
                 model.prior.type=modpriorvec,
                 model.prior.pars = "children",
                 beta.prior.type = "IP",
                 beta.prior.pars = list(alpha=1,nu=1),
                 niter=5000)

summary.res.prop <- summaryout(mcmc.out=res.prop,insampledata=datain.prop,modelprior.nams=modpriorvec,
                               shr.adj=T,outsampledata=data.holdout.prop,respnam="y",top.ave=10,betaprtype="IP",
                               parsprbeta=list(alpha=1,nu=1))

save(file="7plan.RData",
     list=c("res","summary.res","res.prop","summary.res.prop"))
```

formulaHPMs: Vector of characters with variables included in the Highest Probability models (HPMs) identified with each of the model priors considered

TopModels: List of data frames with the formulas for the top "top.ave" models identified and the model posterior probabilities. There is one data frame for each model prior considered.

post.HPM: Vector of model posterior probabilities of the HPM with each model prior considered.

postcumm.Top: Vector of the cummulative model posterior probabilities for the top "top.ave" models for each model prior considered.

MSPE.HPM: Root mean squared prediction error for the HPM's using a holdout data set.

MSPE.ave: Root mean squared prediction error from model averaging using a holdout data set.

Summary

Compare and suggest one best model

```
model_7 <- lm(H_A ~ W_c_A + W_s_A + No + S + W_L + W_m_L + A, data=table_habitat)
model_4<- lm(H_A ~ W_c_A + W_s_A + No + S, data=table_habitat)
model_2<- lm(H_A ~ W_c_A + W_s_A, data=table_habitat)
model_bayes<- lm(H_A ~W_c_A+W_m_A+W_c_A^2+W_c_A*W_m_A+W_m_A^2, data=table_habitat)
model_bayes_perc<- lm(H_P ~ W_s_P+W_m_P+P+W_s_P^2+W_s_P*W_m_P+W_s_P*P+W_m_P^2, data=table_habitat_perc)
model_bayes_perc_cate<- lm(H_A_L ~ W_s_P+W_m_P+P+W_s_P^2+W_s_P*W_m_P+W_s_P*P+W_m_P^2, data=table_habitat_perc)
# model_4<- lm(H_A ~ W_m_A+W_s_A+W_a_A+FE+W_m_A*W_a_A+W_s_A*FE, data = table_habitat)
```

	(1)	(2)	(3)	(4)	(
(Intercept)	30163.105 ***	29236.938 ***	35734.635 ***	0.229 ***	1.96
	(2141.624)	(2021.294)	(913.657)	(0.006)	(0.0
W_c_A	0.169 ***	0.188 ***	0.214 ***		
	(0.028)	(0.023)	(0.028)		
W_s_A	0.209 ***	0.231 ***			
	(0.059)	(0.053)			
No	52.490 *	62.838 **			
	(22.536)	(21.082)			
S	-8840314.083 **	-6611334.155 *			
	(3140762.064)	(2725873.064)			
W_L	1.785 *				
	(0.893)				
W_m_L	1.241				
	(1.309)				
A	-0.012				
	(0.007)				
W_m_A			-0.063		
			(0.036)		
W_c_A:W_m_A			-0.000		
			(0.000)		
W_s_P				0.189	0.7
				(0.135)	(1.5
W_m_P				0.042	-1.2
				(0.047)	(0.5
P				-0.000	-0.1
				(0.006)	(0.0
W_s_P:W_m_P				-0.018	0.2
				(0.550)	(6.4
W_s_P:P				0.059	1.6
				(0.091)	(1.0
N	178	178	178	178	178
R2	0.416	0.397	0.331	0.062	0.1
logLik	-1880.785	-1883.550	-1892.847	233.731	-204.7
AIC	3779.570	3779.101	3795.693	-453.462	423.5

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

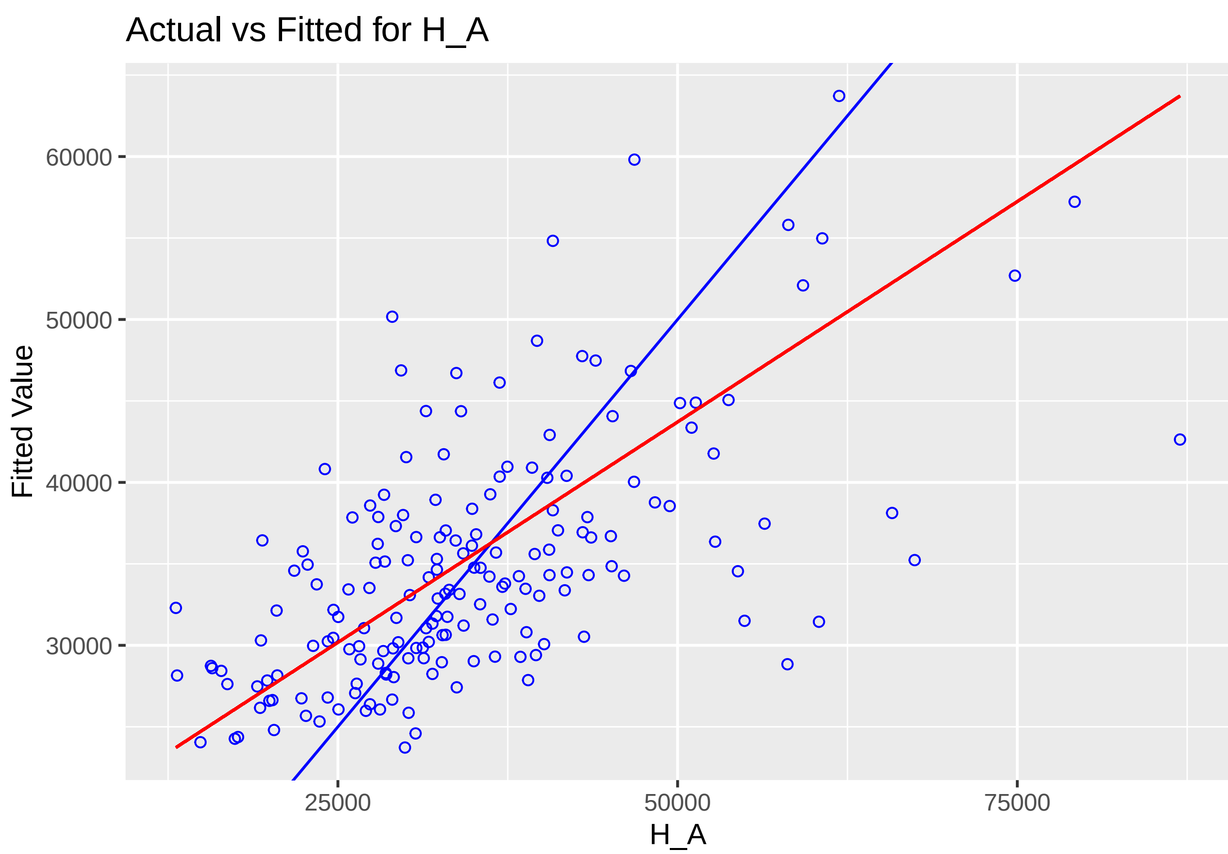
```
vif(model_7)
## W_c_A W_s_A    No     S   W_L W_m_L     A
##  2.44  1.50  2.59  2.51  4.58  1.24  2.82
vif(model_4)
## W_c_A W_s_A    No     S
##  1.62  1.22  2.23  1.86
vif(model_bayes)
##       W_c_A       W_m_A W_c_A:W_m_A
##        2.07        2.26        1.18
vif(model_bayes_perc)
##       W_s_P       W_m_P           P W_s_P:W_m_P     W_s_P:P
##        4.82        1.68        1.26        4.22        2.14
vif(model_bayes_perc_cate)
##       W_s_P       W_m_P           P W_s_P:W_m_P     W_s_P:P
##        4.82        1.68        1.26        4.22        2.14
# print(xtable((table_habitat[40:44,c(1,2,7:9,19)])),floating=FALSE,latex.environments=NULL,booktabs=TRUE)
```

- The 3 variables model

```
ols_regress(model_2)
```

```
##                            Model Summary
## ---------------------------------------------------------------------
## R                        0.604       RMSE                   9875.740
## R-Squared                0.364       Coef. Var                28.329
## Adj. R-Squared           0.357       MSE                97530231.502
## Pred R-Squared           0.334       MAE                    7142.961
## ---------------------------------------------------------------------
##  RMSE: Root Mean Square Error
##  MSE: Mean Square Error
##  MAE: Mean Absolute Error
##
##                                    ANOVA
## --------------------------------------------------------------------------------
##                      Sum of
##                      Squares          DF       Mean Square      F         Sig.
## --------------------------------------------------------------------------------
## Regression     9779281859.298          2    4889640929.649    50.135    0.0000
## Residual      17067790512.908        175      97530231.502
## Total         26847072372.206        177
## --------------------------------------------------------------------------------
##
##                                      Parameter Estimates
## ----------------------------------------------------------------------------------------------
##       model          Beta     Std. Error     Std. Beta      t        Sig         lower        upper
## ----------------------------------------------------------------------------------------------
## (Intercept)     34860.919        740.218                  47.095    0.000    33400.016    36321.823
##       W_c_A         0.148          0.019         0.481     7.731    0.000        0.110        0.186
##       W_s_A         0.218          0.051         0.266     4.281    0.000        0.117        0.318
## ----------------------------------------------------------------------------------------------
```

```
ols_plot_obs_fit(model_2)
```

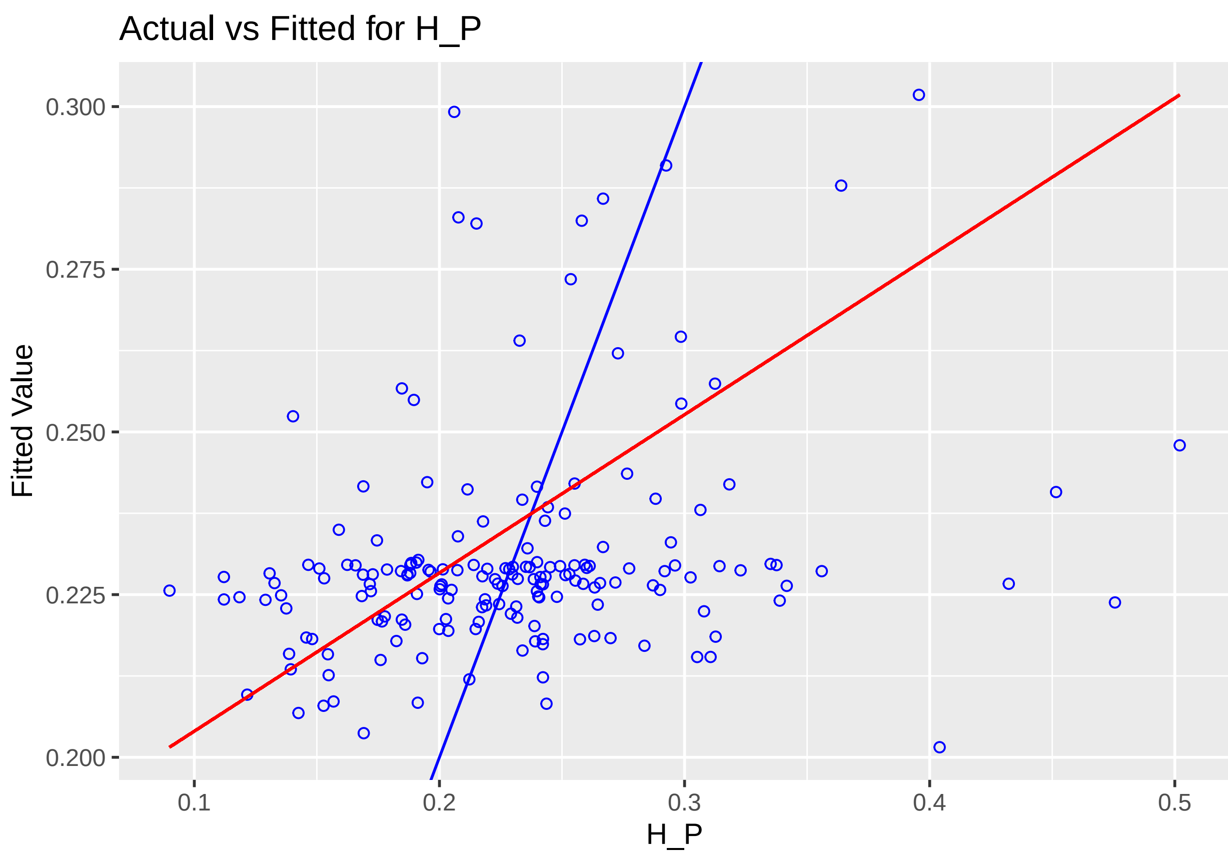


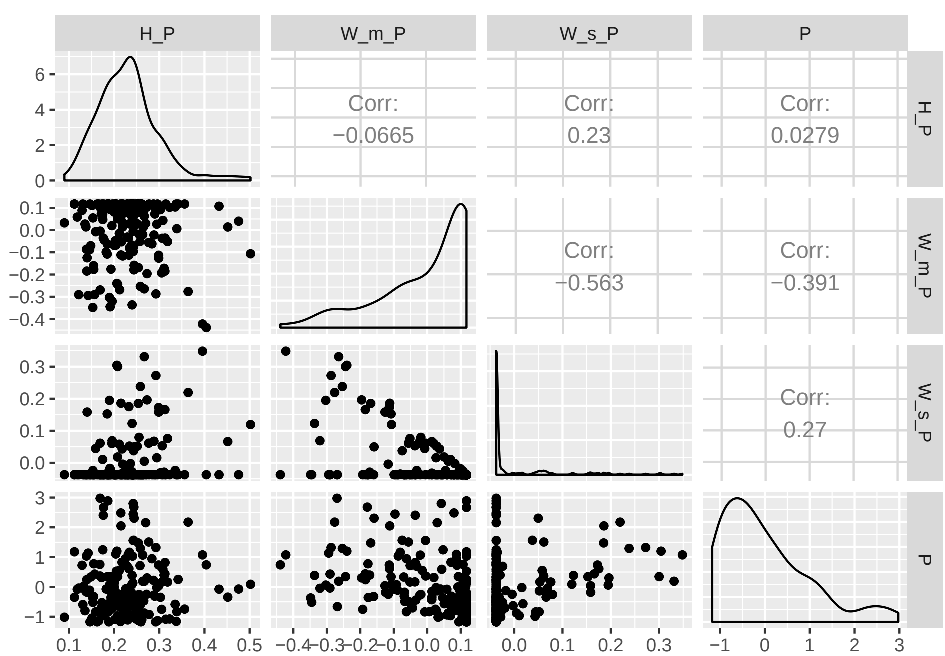

```r
ols_regress(model_bayes_perc)
```

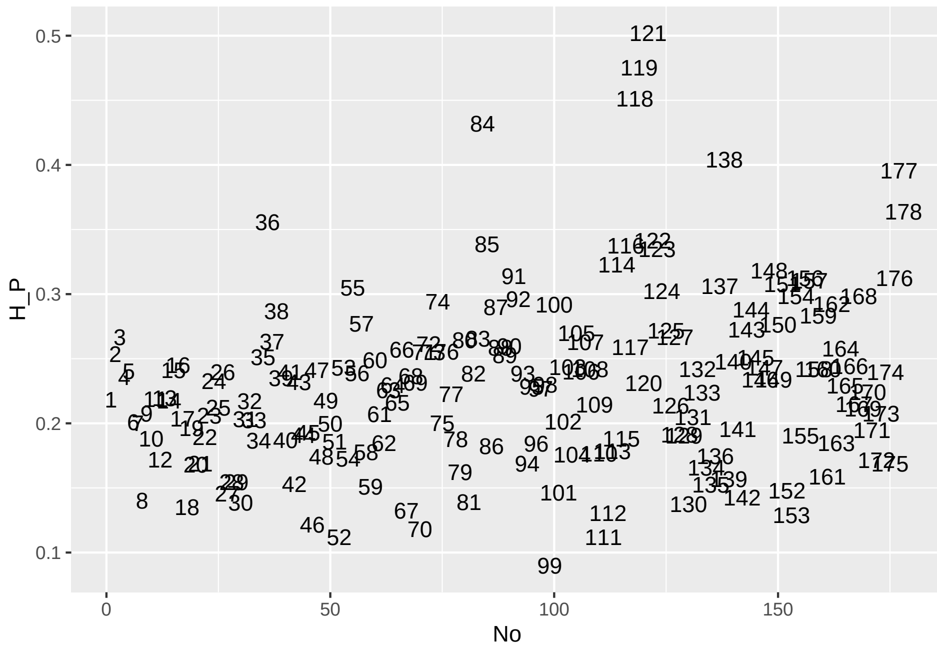
```
##                          Model Summary
## ---------------------------------------------------------------
## R                         0.249       RMSE                0.066
## R-Squared                 0.062       Coef. Var          28.747
## Adj. R-Squared            0.034       MSE                 0.004
## Pred R-Squared           -0.038       MAE                 0.048
## ---------------------------------------------------------------
##  RMSE: Root Mean Square Error
##  MSE: Mean Square Error
##  MAE: Mean Absolute Error
##
##                                ANOVA
## -------------------------------------------------------------------
##                 Sum of
##                Squares          DF    Mean Square      F         Sig.
## -------------------------------------------------------------------
## Regression       0.050           5          0.010    2.265     0.0502
## Residual         0.754         172          0.004
## Total            0.804         177
## -------------------------------------------------------------------
##
##                                  Parameter Estimates
## ---------------------------------------------------------------------------------------
##       model       Beta    Std. Error    Std. Beta        t        Sig      lower    upper
## ---------------------------------------------------------------------------------------
## (Intercept)      0.229         0.006                 38.430     0.000      0.217    0.241
##       W_s_P      0.189         0.135        0.228     1.407     0.161     -0.076    0.455
##       W_m_P      0.042         0.047        0.085     0.889     0.375     -0.051    0.135
##           P      0.000         0.006       -0.001    -0.014     0.989     -0.011    0.011
## W_s_P:W_m_P     -0.018         0.550       -0.005    -0.033     0.974     -1.104    1.068
##     W_s_P:P      0.059         0.091        0.070     0.646     0.519     -0.121    0.238
```

```
## --------------------------------------------------------------------------------
```

```
ols_plot_obs_fit(model_bayes_perc)
```

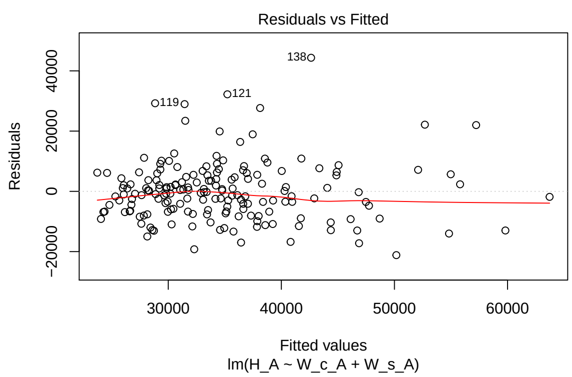


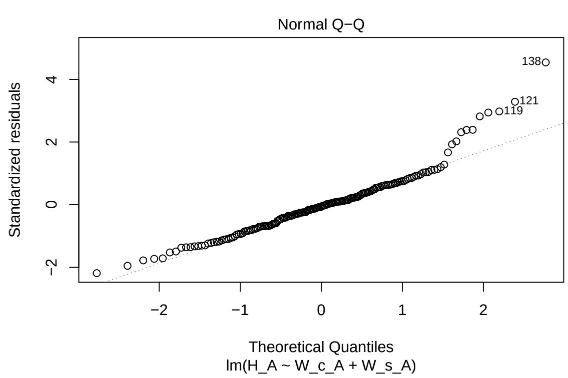


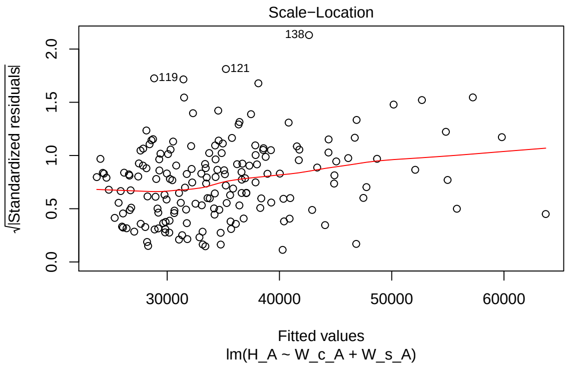


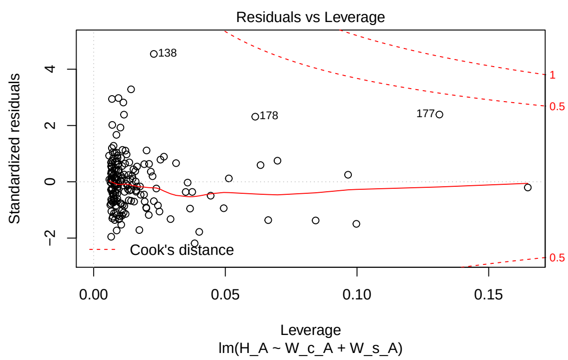
Residual diagnostics

```r
#Model Fit Assessment
plot(model_2)
```



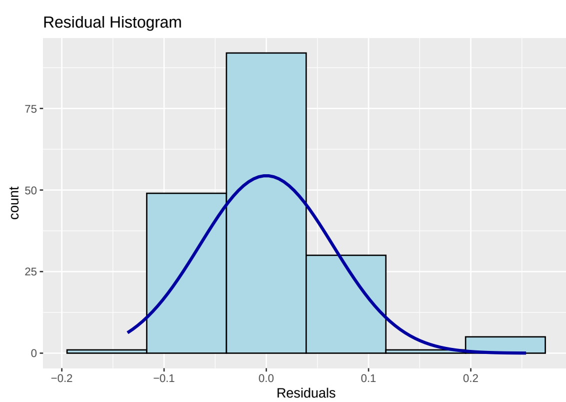







```
# ols_plot_diagnostics(model_2)
# print(xtable(summary(model_4)),floating=FALSE,latex.environments=NULL,booktabs=TRUE)
```

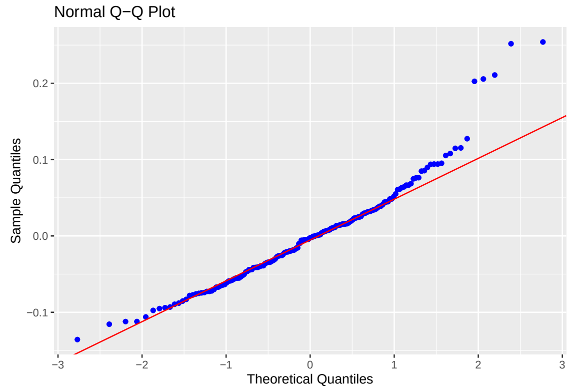
```
ols_plot_resid_hist(model_bayes_perc)
```



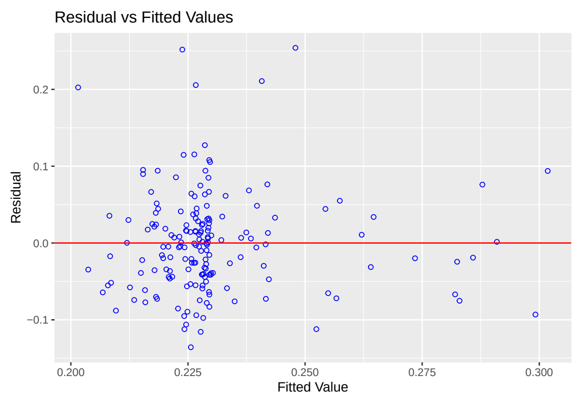

```
ols_test_normality(model_bayes_perc)
```

```
## -----------------------------------------------
##        Test             Statistic       pvalue
## -----------------------------------------------
## Shapiro-Wilk              0.9385         0.0000
## Kolmogorov-Smirnov        0.0801         0.2034
## Cramer-von Mises         52.2031         0.0000
## Anderson-Darling          1.7843         0.0001
## -----------------------------------------------
```

```
ols_plot_resid_qq (model_bayes_perc)
```



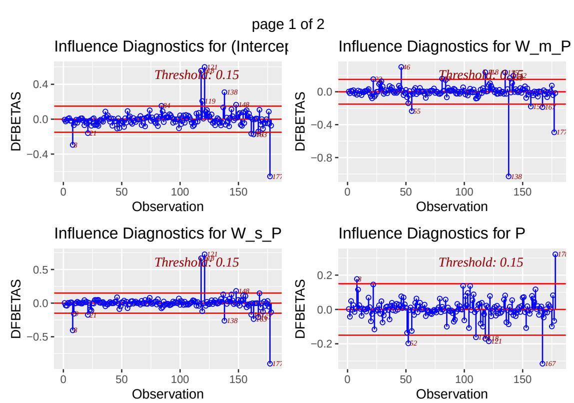

```
ols_plot_resid_fit(model_bayes_perc)
```

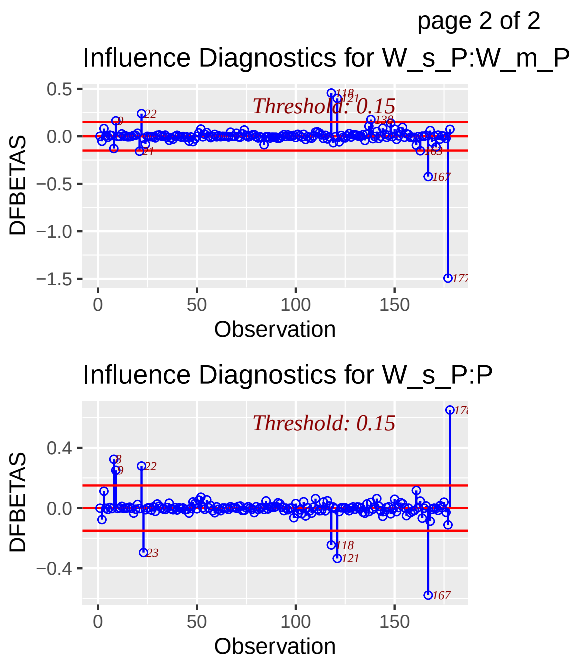


- DFBETAs Panel

DFBETAs measure the difference in each parameter estimate with and without the influential observation.

```
ols_plot_dfbetas(model_bayes_perc)
```

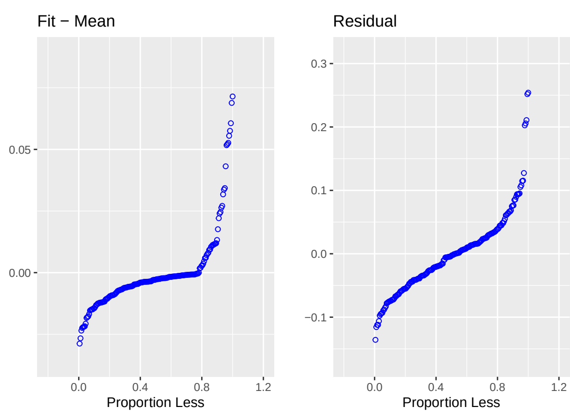




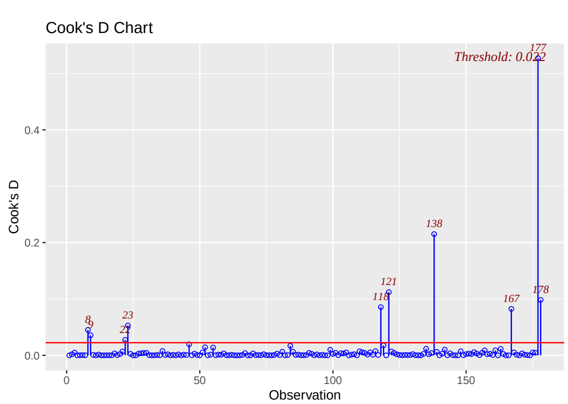
- Collinearity Diagnostics

Plot to detect non-linearity, influential observations and outliers.

```
ols_plot_resid_fit_spread(model_bayes_perc)
```




```
ols_plot_cooksd_chart(model_bayes_perc)
```



Part & Partial Correlations

Correlations Relative importance of independent variables in determining Y. How much each variable uniquely contributes to R2 over and above that which can be accounted for by the other predictors.

Zero Order Pearson correlation coefficient between the dependent variable and the independent variables.

Part Unique contribution of independent variables. How much R2 will decrease if that variable is removed from the model?

Partial How much of the variance in Y, which is not estimated by the other independent variables in the model, is estimated by the specific variable?

- The partial regression and nonlinear diagnostics

Transformation

```
model_3_log<- lm(log(H_P) ~ W_s_P+W_m_P+P+W_s_P^2+W_s_P*W_m_P+W_s_P*P+W_m_P^2, data = table_habitat_perc)
summary(model_bayes_perc)
```

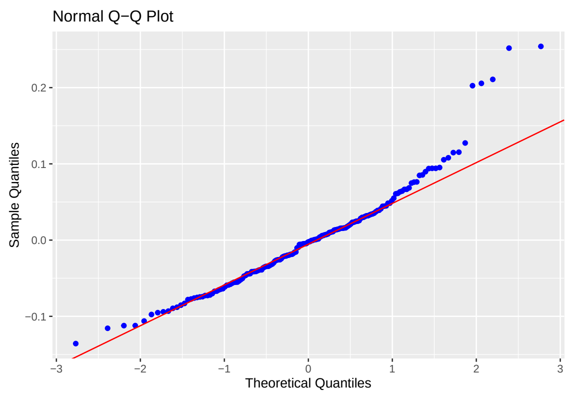
```
##
## Call:
## lm(formula = H_P ~ W_s_P + W_m_P + P + W_s_P^2 + W_s_P * W_m_P +
##     W_s_P * P + W_m_P^2, data = table_habitat_perc)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.13574 -0.04142 -0.00249  0.03072  0.25406
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.2289653  0.0059580   38.43   <2e-16 ***
## W_s_P        0.1894032  0.1345811    1.41     0.16
## W_m_P        0.0419442  0.0472040    0.89     0.38
## P           -0.0000804  0.0057134   -0.01     0.99
## W_s_P:W_m_P -0.0181130  0.5500610   -0.03     0.97
## W_s_P:P      0.0587404  0.0908787    0.65     0.52
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0662 on 172 degrees of freedom
## Multiple R-squared:  0.0618, Adjusted R-squared:  0.0345
## F-statistic: 2.26 on 5 and 172 DF,  p-value: 0.0502
```

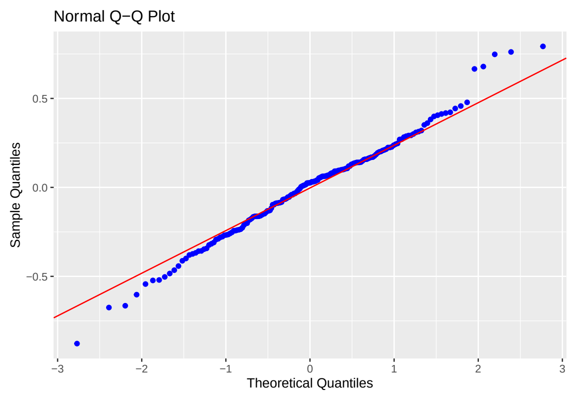
```
summary(model_3_log)
```

```
##
## Call:
## lm(formula = log(H_P) ~ W_s_P + W_m_P + P + W_s_P^2 + W_s_P *
##     W_m_P + W_s_P * P + W_m_P^2, data = table_habitat_perc)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.8778 -0.1648  0.0261  0.1584  0.7925
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.51438    0.02528  -59.89   <2e-16 ***
## W_s_P        0.86982    0.57113    1.52     0.13
## W_m_P        0.26771    0.20032    1.34     0.18
## P            0.00412    0.02425    0.17     0.87
## W_s_P:W_m_P  0.09746    2.33434    0.04     0.97
## W_s_P:P      0.28600    0.38567    0.74     0.46
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 0.281 on 172 degrees of freedom
## Multiple R-squared:  0.0656, Adjusted R-squared:  0.0385
## F-statistic: 2.42 on 5 and 172 DF,  p-value: 0.0379
```

```
# ols_regress(model_4_log)
ols_plot_resid_qq (model_bayes_perc)
ols_plot_resid_qq (model_3_log)
```





GauPro

Random Forest

SVM