Untitled

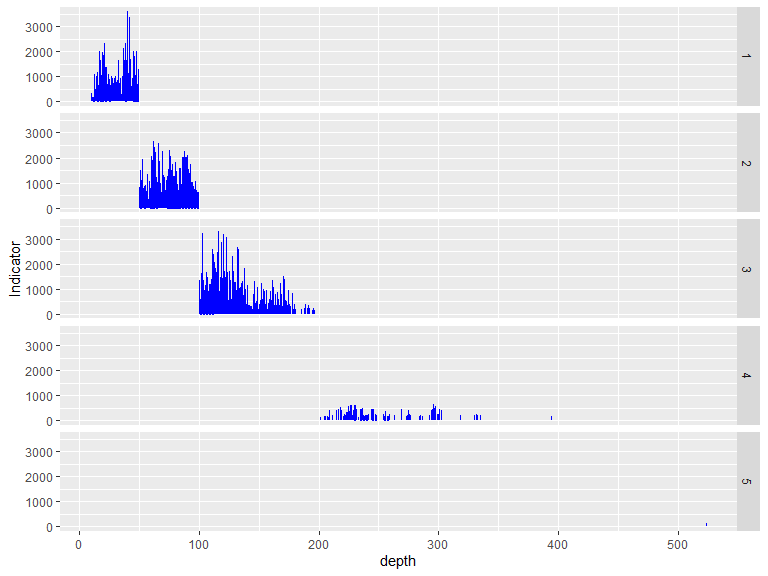
Enrico N Armelloni

1/10/2020

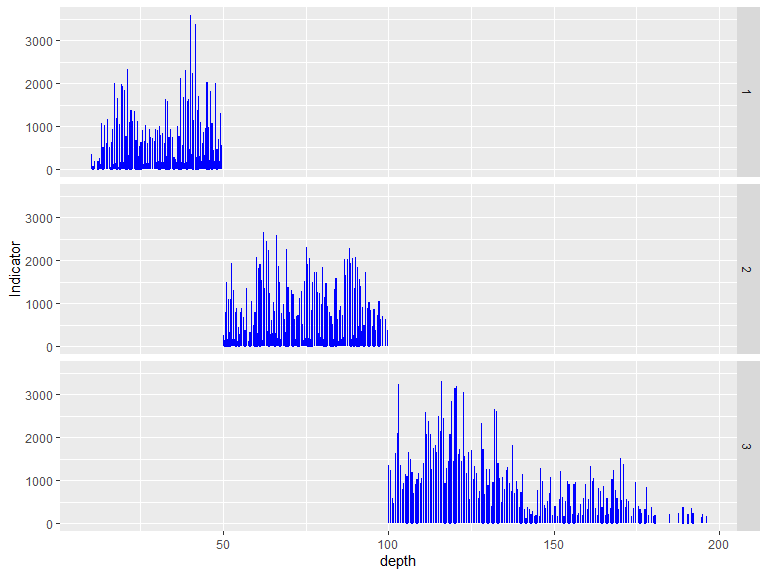
## Red mullet in GSA 17-18

Dati utilizzati: MEDITS DATA 1994-2017 (solo cale positive, con catture triglia di fango)

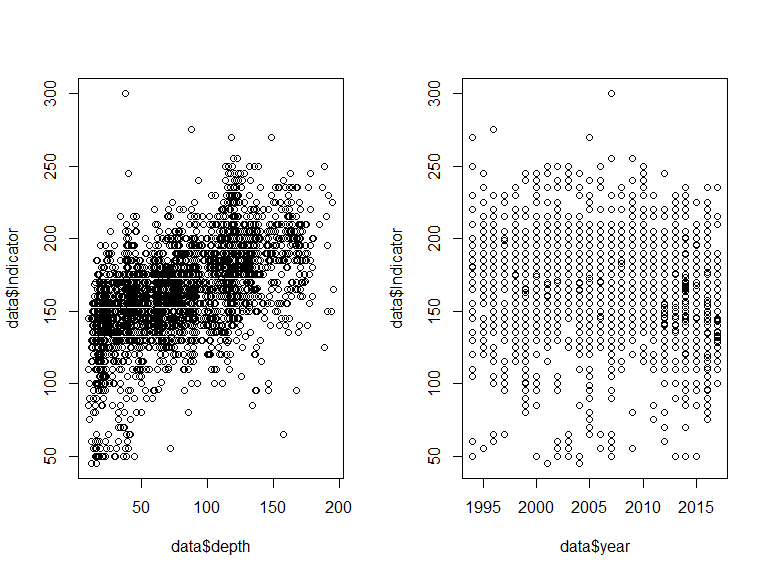
Nelle GSA 17 e 18 la triglia si concentra sulla piattaforma continentale, soprattutto entro i 100 metri di profondità.



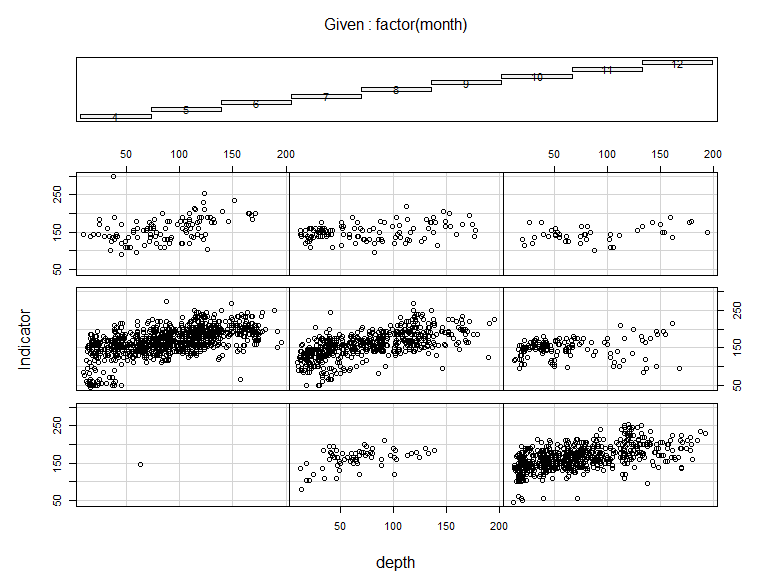
Anche a causa della conformazione del bacino le densità registrate oltre i 100 metri sono esigue ma concentrate spazialmente nella porzione meridionale, quindi si decide di includere nelle analisi le cale comprese nella batimetria 0-200 m.



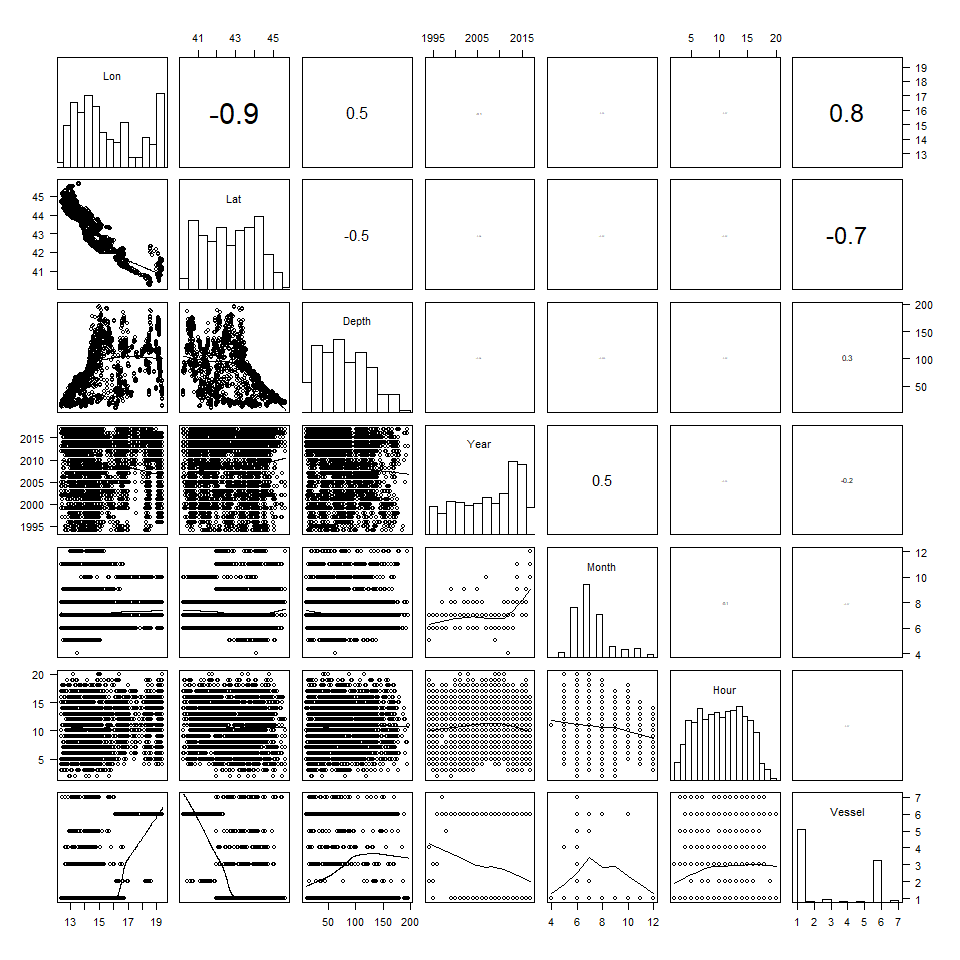
Plottando l’indicatore L95 per cala contro le variabili anno e profondità si osserva la presenza di numerosi outlier, presenti nella maggior parte degli anni ed anche a profondità superiori ai 100 metri. Gli outlier non sono stati rimossi dal dataset.



Per investigare ulteriormente la relazione tra mensilità, indicatore e profondità della cala si presenta anche un coplot, dove emerge la presenza di valori bassi dell’indicatore anche a profondità > 100 metri durante i mesi estivi (soprattutto giugno e luglio).



Sulla base di quanto svolto dai colleghi, si è effettuata una step forward model selection per identificare il modello GAM migliore in termini di fitting, residui e statistiche (AIC, GCV). Prima di questo, è stata fatta una esplorazione dei dati per identificare potenziali collinearità tra variabili indipendenti. Il seguente pairplot mostra come ci sia potenziale collinearità tra le variabili Lat, Lon, vessel (ogni GSA ha la propria imbarcazione.



Analisi VIF

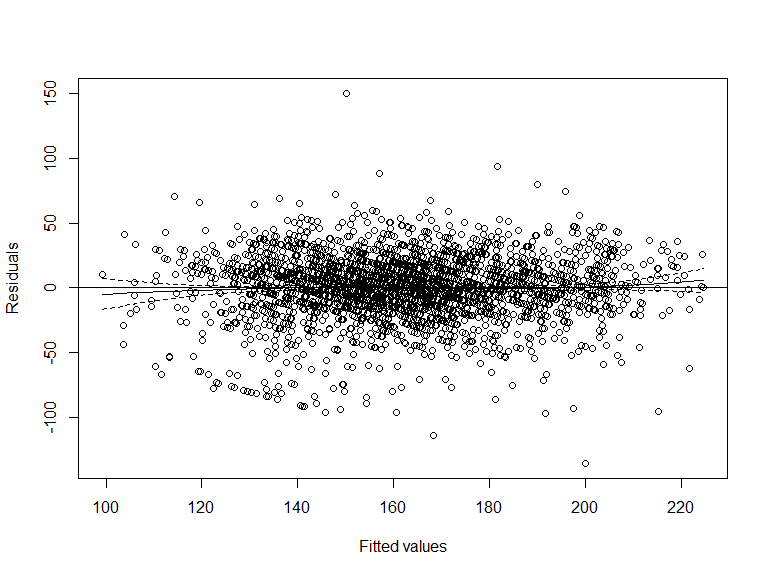
## X Y depth year month hour vessel   
## 8.612273 6.248885 1.462603 1.368444 1.332019 1.024198 3.258361

La variabile vessel viene esclusa

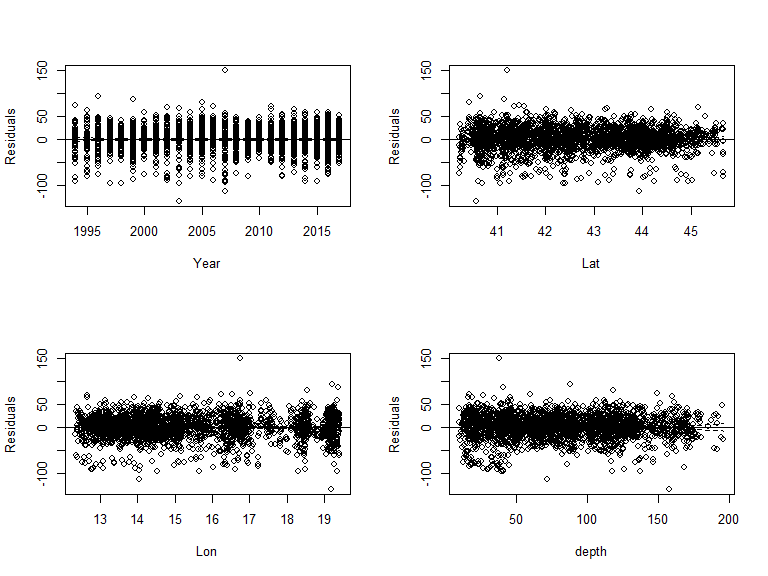
Alla fine del processo di model selection, il modello GAM migliore è risultato il seguente:

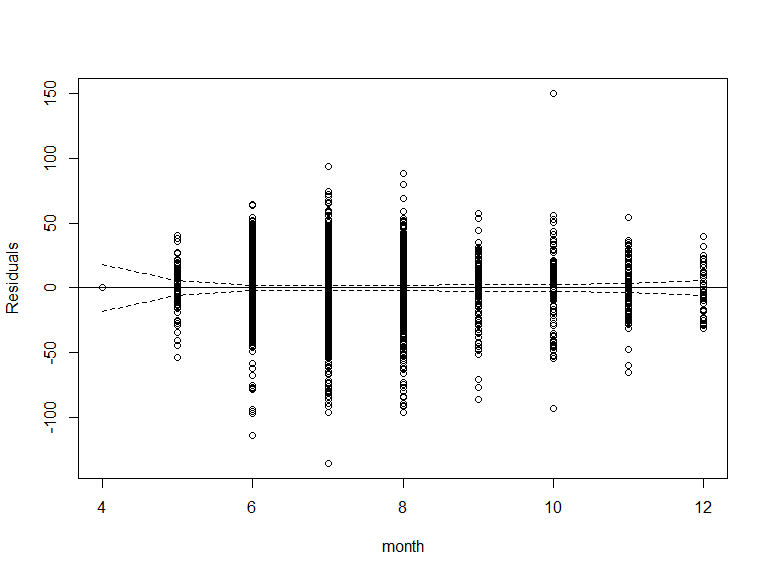
##   
## Family: gaussian   
## Link function: identity   
##   
## Formula:  
## Indicator ~ s(year, k = 16) + factor(month) + s(X, Y) + s(depth) +   
## s(depth, by = month)  
##   
## Parametric coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 142.020 40.152 3.537 0.000412 \*\*\*  
## factor(month)5 7.200 22.041 0.327 0.743956   
## factor(month)6 4.472 18.056 0.248 0.804410   
## factor(month)7 -4.934 14.411 -0.342 0.732121   
## factor(month)8 -13.888 10.827 -1.283 0.199715   
## factor(month)9 -17.229 7.436 -2.317 0.020585 \*   
## factor(month)10 -20.100 4.215 -4.769 1.95e-06 \*\*\*  
## factor(month)11 -16.839 2.383 -7.066 2.04e-12 \*\*\*  
## factor(month)12 -23.564 5.287 -4.457 8.67e-06 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Approximate significance of smooth terms:  
## edf Ref.df F p-value   
## s(year) 12.200 13.847 4.396 7.77e-08 \*\*\*  
## s(X,Y) 25.787 28.324 13.159 < 2e-16 \*\*\*  
## s(depth) 8.120 8.627 12.681 < 2e-16 \*\*\*  
## s(depth):month 3.859 4.512 9.383 5.08e-08 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Rank: 71/72  
## R-sq.(adj) = 0.433 Deviance explained = 44.5%  
## GCV = 650.93 Scale est. = 636.55 n = 2625

Analisi grafica residui – omogeneità I residui sono ben distribuiti anche se piuttosto ampi. Si evidenziano alcuni valori negativi outlier.

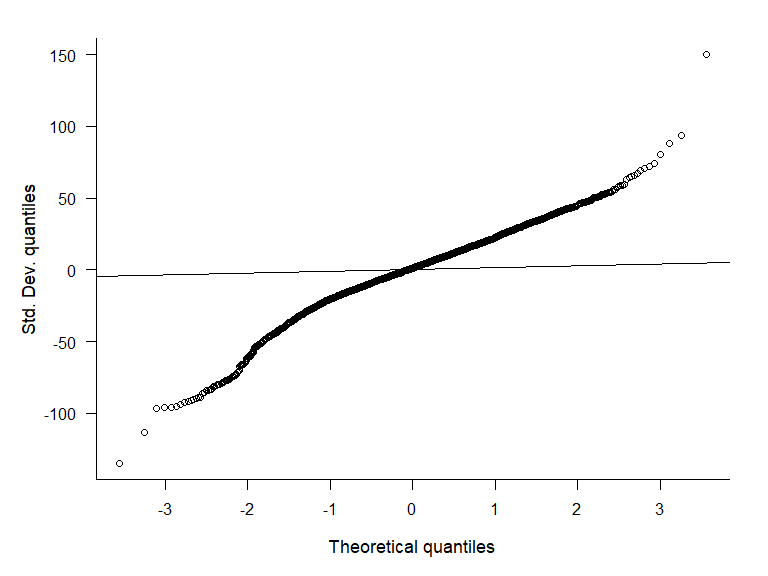
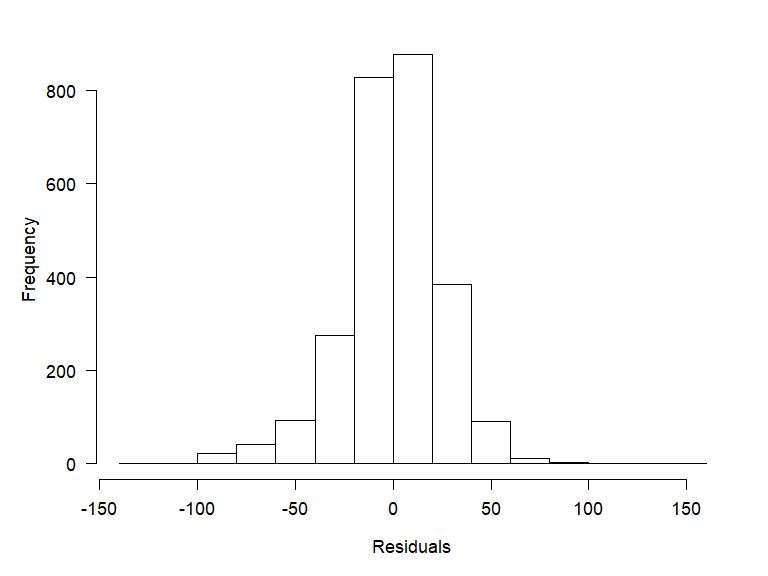


Residui per fattore Nonostante alcuni valori outlier, soprattutto riguardo al fattore mese, non si osservano particolari pattern

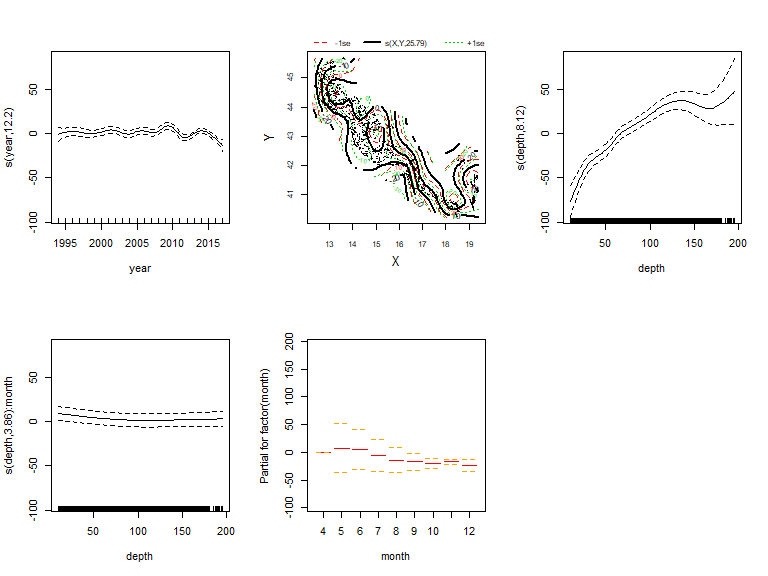




Analisi residui – normalità Si evidenzia una coda nel QQ-plot



Plot variabili indipendenti incluse nel modello finale



Il seguente plot mostra la serie temporale del L95 medio per anno (in blu), i risultati del modello (in rosso) ed i risultati della prova di standardizzazione sul mese di luglio.

## `summarise()` ungrouping output (override with `.groups` argument)  
## `summarise()` ungrouping output (override with `.groups` argument)

