Untitled

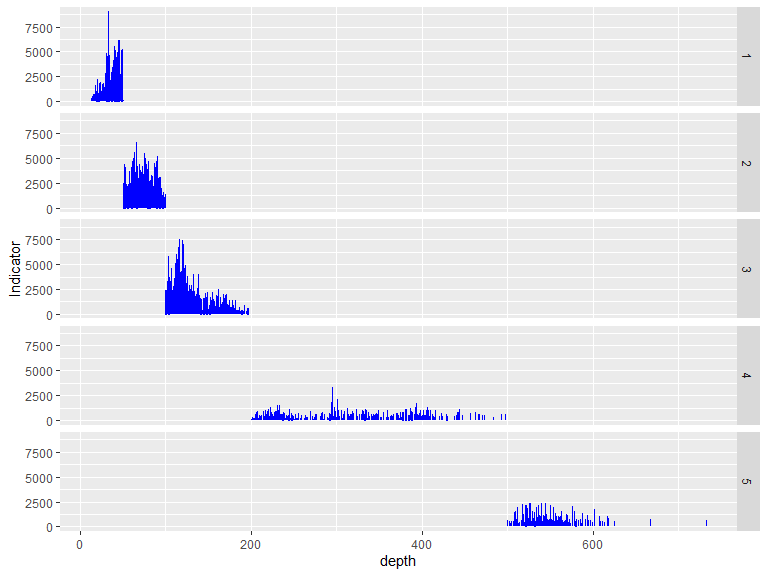
Enrico N Armelloni

1/10/2020

## European hake in GSA 17-18

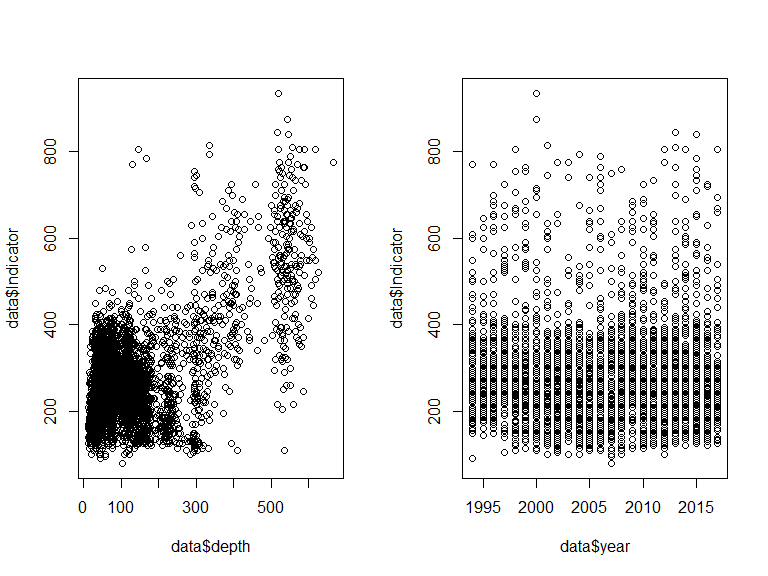
Dati utilizzati: MEDITS DATA 1994-2017 (solo cale positive, con catture nasello)

Nelle GSA 17 e 18 il nasello è distribuito sia sulla piattaforma continentale che sulla scarpata.

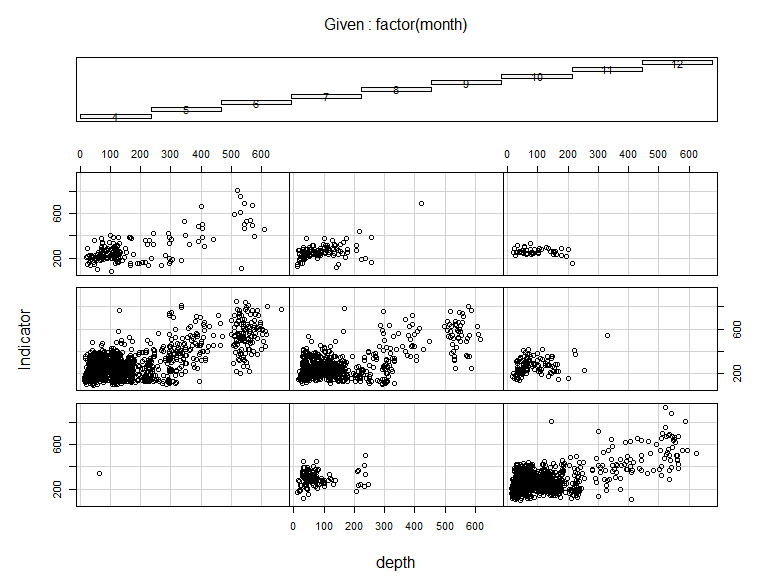


Si decide di includere nelle analisi le cale comprese nella batimetria 0-700 m.

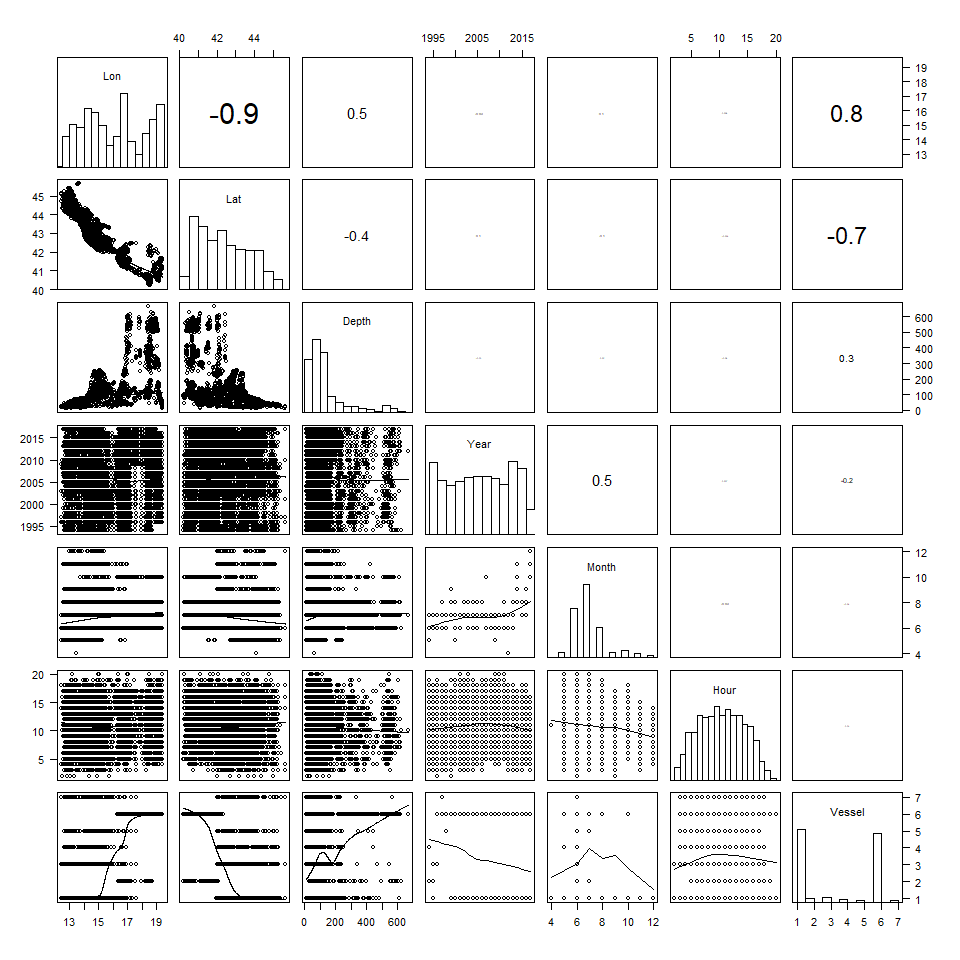
Plottando l’indicatore L95 per cala contro le variabili anno e profondità si osserva che entro i 200 metri si concentrano i valori più bassi, mentre i valori > 500 mm si concentrano nelle zone di scarpata



Per investigare ulteriormente la relazione tra mensilità, indicatore e profondità della cala si presenta anche un coplot, dove emerge la presenza di valori alti dell’indicatore soprattutto durante i mesi estivi (giugno e luglio) e la scarsità di dati relativi alla zona di scarpata nei mesi più tardivi.



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

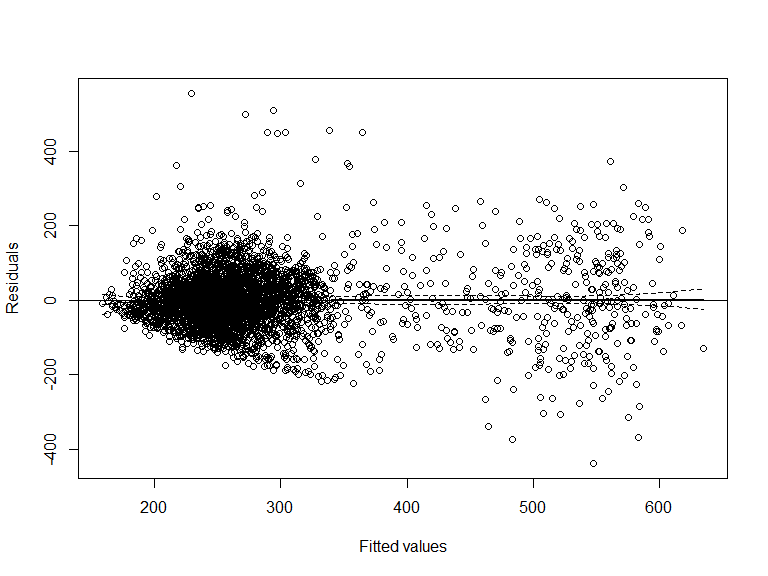
## X Y depth year month hour vessel   
## 8.394991 6.856145 1.324864 1.361516 1.339859 1.019271 2.559122

La variabile vessel viene esclusa

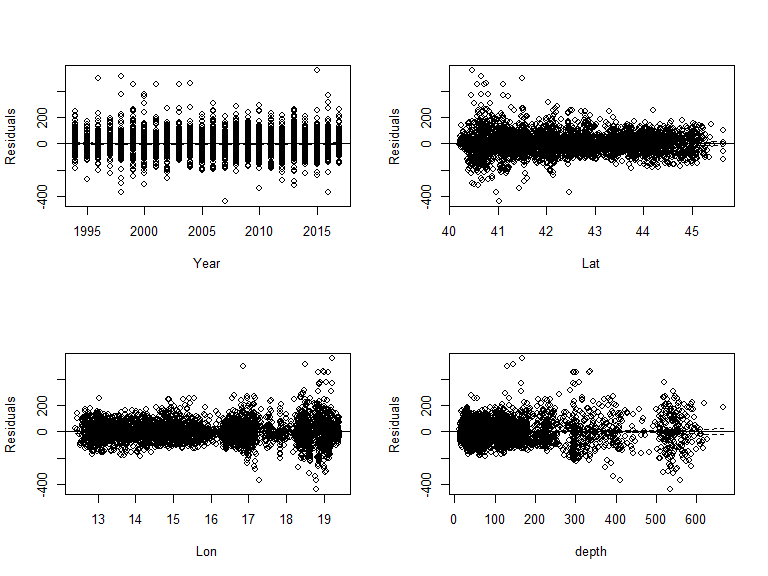
Alla fine del processo di model selection, il modello GAM migliore è risultato il seguente:

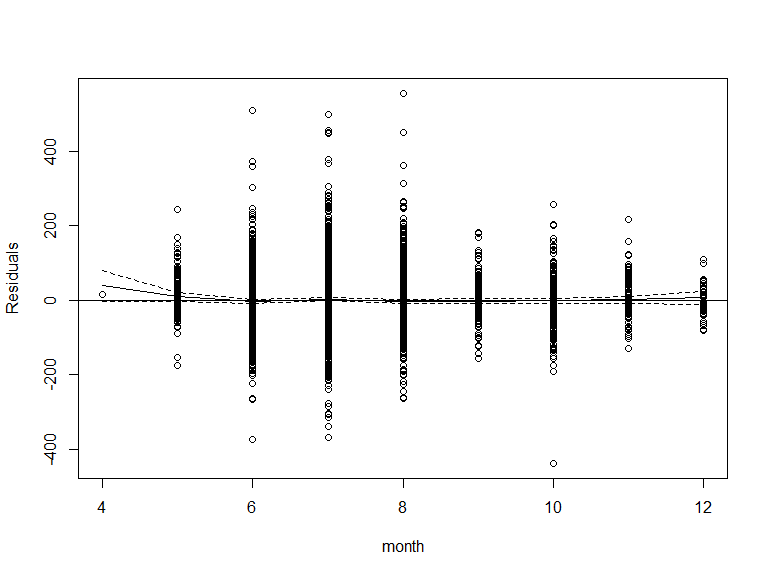
##   
## Family: gaussian   
## Link function: identity   
##   
## Formula:  
## Indicator ~ s(depth) + s(Y, X) + s(month, k = 4) + s(year, k = 14) +   
## s(year, by = as.numeric(depth >= 200), k = 14)  
##   
## Parametric coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 282.863 2.902 97.47 <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(depth) 7.291 8.255 167.588 < 2e-16 \*\*\*  
## s(Y,X) 24.291 27.605 13.019 < 2e-16 \*\*\*  
## s(month) 2.348 2.710 3.281 0.014 \*   
## s(year) 10.170 11.640 5.418 5.48e-09 \*\*\*  
## s(year):as.numeric(depth >= 200) 11.287 12.755 3.262 5.42e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = 0.514 Deviance explained = 52%  
## GCV = 5965.9 Scale est. = 5889 n = 4374

Analisi grafica residui – omogeneità I residui non sono ben distribuiti e rimane evidente la differenza tra piattaforma e scarpata. Si evidenziano alcuni valori negativi outlier.

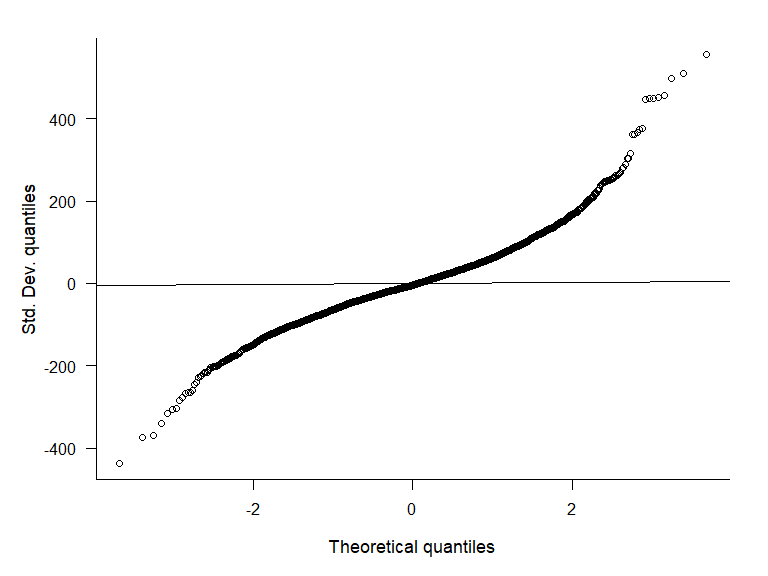
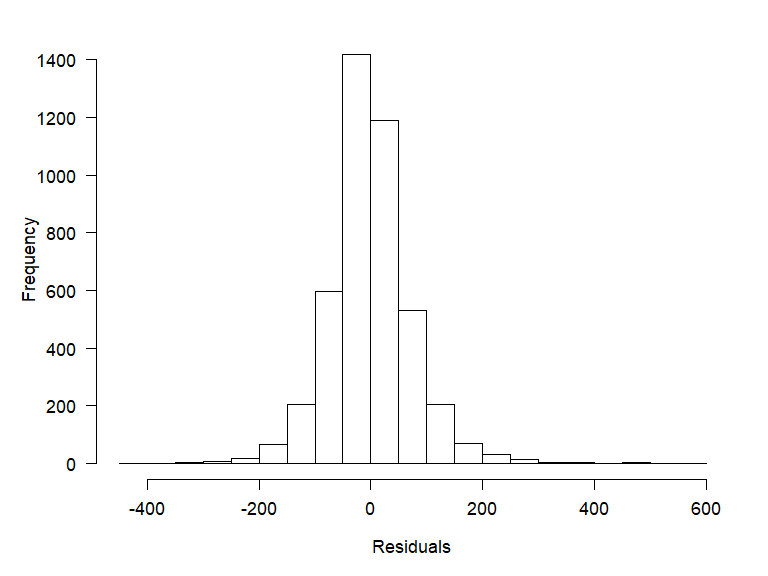


Residui per fattore Anche in questo caso si osservano dei pattern legati alla conformazione geografica del bacino, con le zone più profonde concentrate a sud

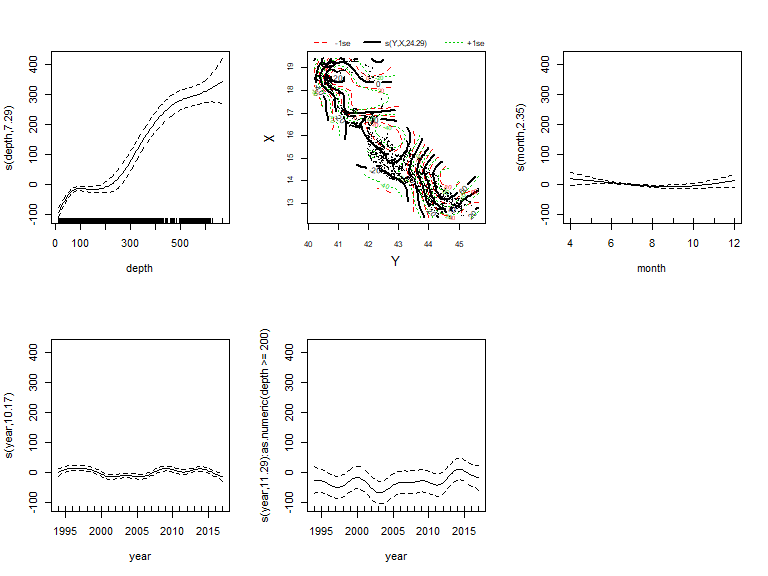




Analisi residui – normalità Si evidenzia una coda nel QQ-plot



Plot variabili indipendenti incluse nel modello finale



Il seguente plot mostra la serie temporale del L95 medio per anno (in blu) ed i risultati del modello (in rosso).In questo caso, data la scarsa influenza del fattore mese non è stata effettuata la standardizzazione.

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

