

1 Define model candidates

Location is added as a variable in every model to control for location with a fixed intercept.

- Model 0: erosion rate ~ 1 (no effect)
- Model 1: erosion rate $\sim$ slope + length + location (only topography)
- Model 2: erosion rate $\sim$ slope * length + location (only topography, with interaction)
- Model 3: erosion rate $\sim$ cover + location (only vegetation cover)
- Model 4: erosion rate $\sim$ slope * length + cover + location (topography and vegetation cover)
- Model 5: erosion rate $\sim$ slope * length + cover + area + location (topography, vegetation cover and contributing area)

2 Compare AIC to select best model

Model with the lowest AIC is chosen. Δ shows difference with the lowest AIC. Weight shows the probability of a model being the best model. Select model 2.

Model	AIC	Δ	Weight
Model 2	-25.6	0.00	0.831
Model 4	-21.9	2.73	0.128
Model 5	-18.7	6.87	0.027
Model 1	-16.6	8.99	0.009
Model 3	-15.2	10.39	0.005
Model 0	18.6	44.21	0.000

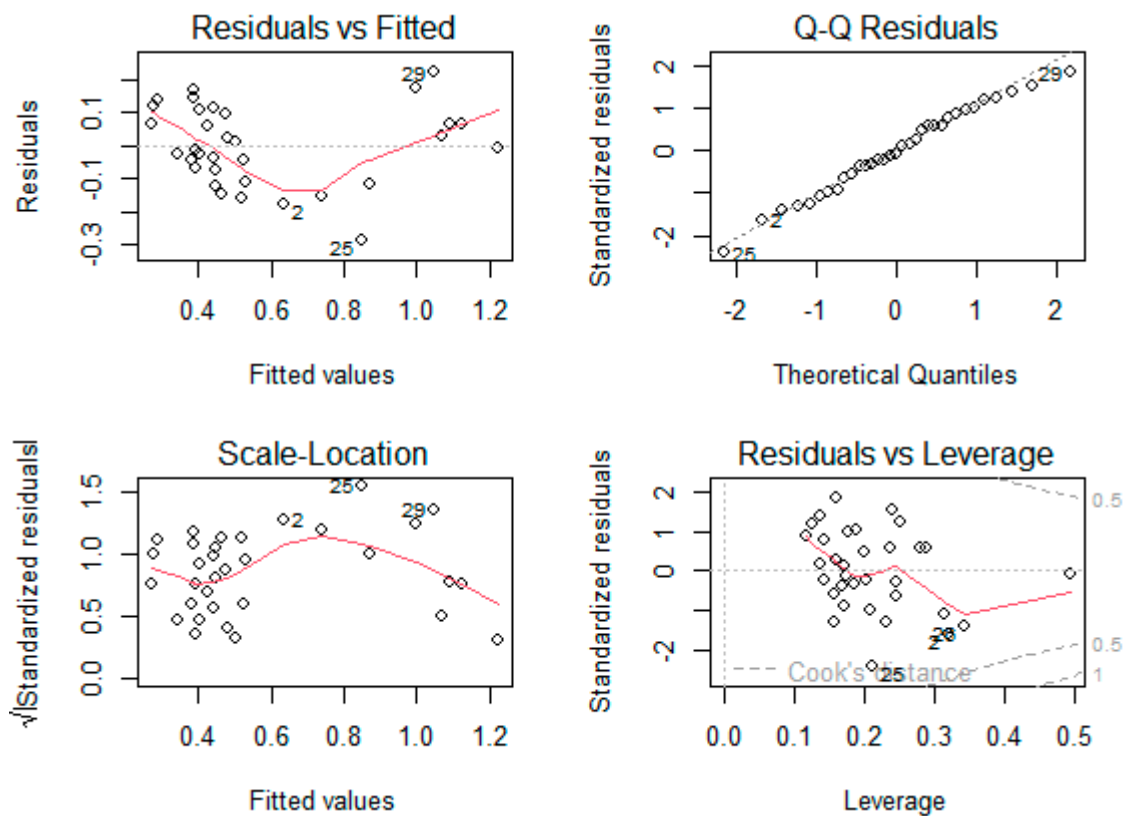
3 Scale data and report VIF

Data used in the interaction term should be scaled (slope and length) to remove collinearity in the model. The Variance Inflation Factor is used to investigate collinearity, with $VIF < 5$ being no collinearity.

Variable	VIF
Slope	1.467988
Length	1.385536
Location	1.282899
Slope:Length	1.225301

4 Diagnostic plots

We use diagnostic plots to investigate assumptions. QQ plot shows that the data is mostly normally distributed. Scale-location and Residuals plots suggest that there is some degree of heteroscedasticity (meaning variance is not constant across all levels of the independent variables). Some degree of heteroscedasticity and non-linearity is normal in ecological data. To account for the heteroscedasticity, we can use robust standard errors to prevent biased standard errors in the model summary.



Step 5 model summary with robust standard errors

	Estimate β	Std. Error	t-value	Pr(> t)
(Intercept)	-1.915687	0.885172	-2.164	0.03982 *
Slope	0.203703	0.070432	2.892	0.00763 **
Length	0.067817	0.020888	3.247	0.00321 **
Location_MOO	-0.038590	0.114577	-0.337	0.73897
Location_MOV	-0.005529	0.072295	-0.076	0.93963
Location_VFO	0.577612	0.066813	8.645	4.01e-09 ***
Slope:length	-0.005748	0.001660	-3.462	0.00187 **

- Residual standard error: 0.1325 on 26 degrees of freedom.
- Multiple R-squared: 0.8463, Adjusted R-squared: 0.8109.
- F-statistic: 23.87 on 6 and 26 DF, p-value: 2.054e-09.

Step 6 Leave one out cross-validation (LOOCV)

LOOCV R-squared = 0.747888