Question 4: Multiple Regression Analysis

Initial approach: Multiple regression models using stepwise method to identify best predictors for:

- Invertebrate abundance (Model 1)
- Species richness (Model 2)

(i) Invertebrate Abundance Model:

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of Estimate
1	.893	.798	.795	78.79344
2	.901	.812	.805	76.71344

Figure 15: Model summary statistics showing R^2 values and goodness of fit for invertebrate abundance regression model. Model 1 includes only stored sediment, while Model 2 includes both stored sediment and bed shear stress.

Coefficients						
	Unstandardised Coefficients		Standardised Coefficients			
Variable	B	SE	Beta	t	Sig.	VIF
(Constant)	557.426	29.068		19.176	<.001	
Stored sediment (g/m^2)	-.356	.025	-.859	-14.364	<.001	1.084
Bed shear stress (N/m^2)	.569	.278	.122	2.046	.045	1.084

Figure 16: Multiple regression coefficients for invertebrate abundance model, showing significant predictors and their relative importance.

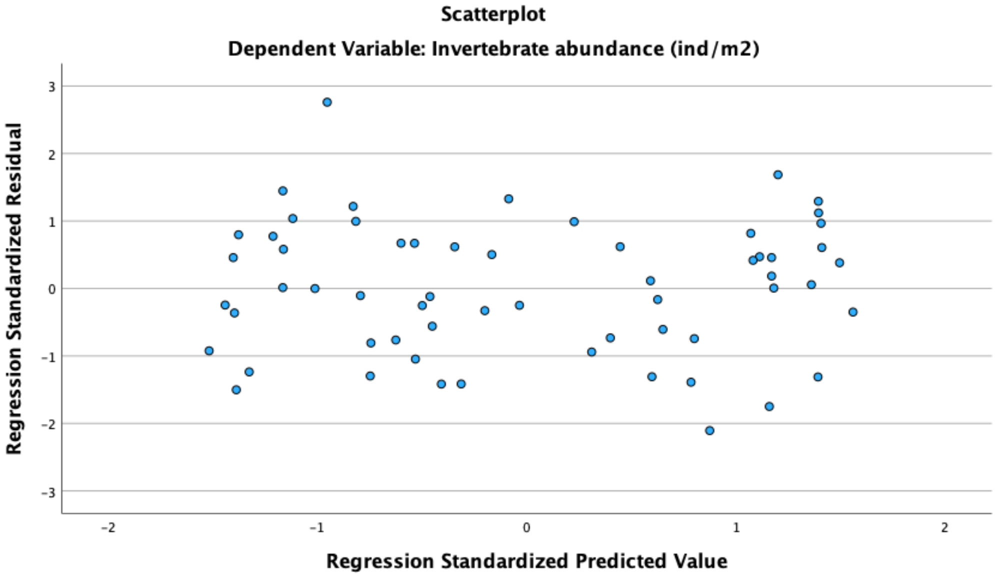

Figure 17: Scatterplot of standardised residuals showing model fit and assumptions.

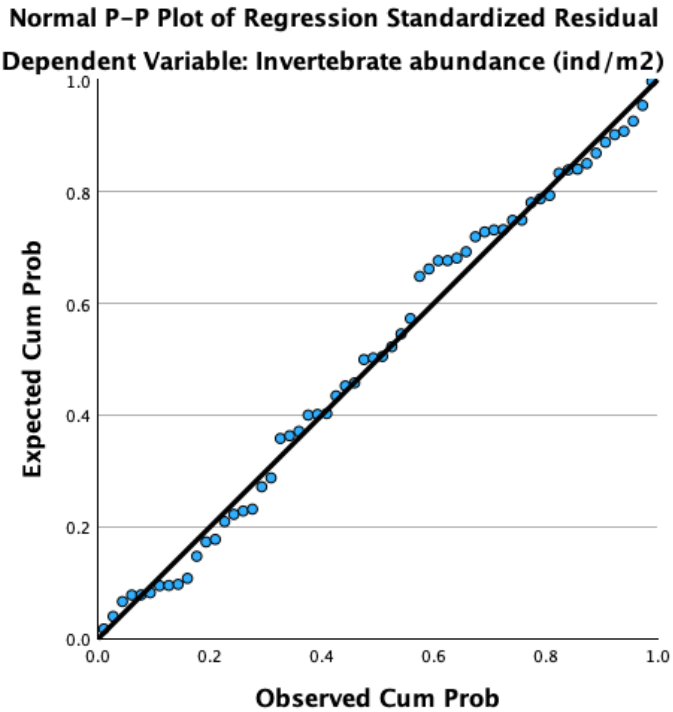

Figure 18: Normal P-P plot demonstrating normality of residuals for abundance model.

The stepwise regression for invertebrate abundance identified a two-predictor model. The final model achieved an R^2 of 0.812 (adjusted R^2 = 0.805), explaining 81.2% of the variance. Stored sediment emerged as the primary predictor ($\beta = -0.859$, $p < 0.001$), showing a strong negative relationship. Bed shear stress contributed a smaller but significant positive effect ($\beta = 0.122$, $p = 0.045$). Both velocity and nitrate were excluded due to non-significance in the stepwise process.

(ii) Species Richness Model:

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of Estimate
1	.844	.712	.707	2.13006

Figure 19: Model summary statistics for species richness regression analysis, showing final model with stored sediment as predictor.

Coefficients						
	Unstandardised Coefficients		Standardised Coefficients			
Variable	B	SE	Beta	t	Sig.	VIF
(Constant)	13.605	.483		28.165	<.001	
Stored sediment (g/m^2)	-.008	.001	-.844	-11.974	<.001	1.000

Figure 20: Multiple regression coefficients for species richness model, showing stored sediment as the only significant predictor.

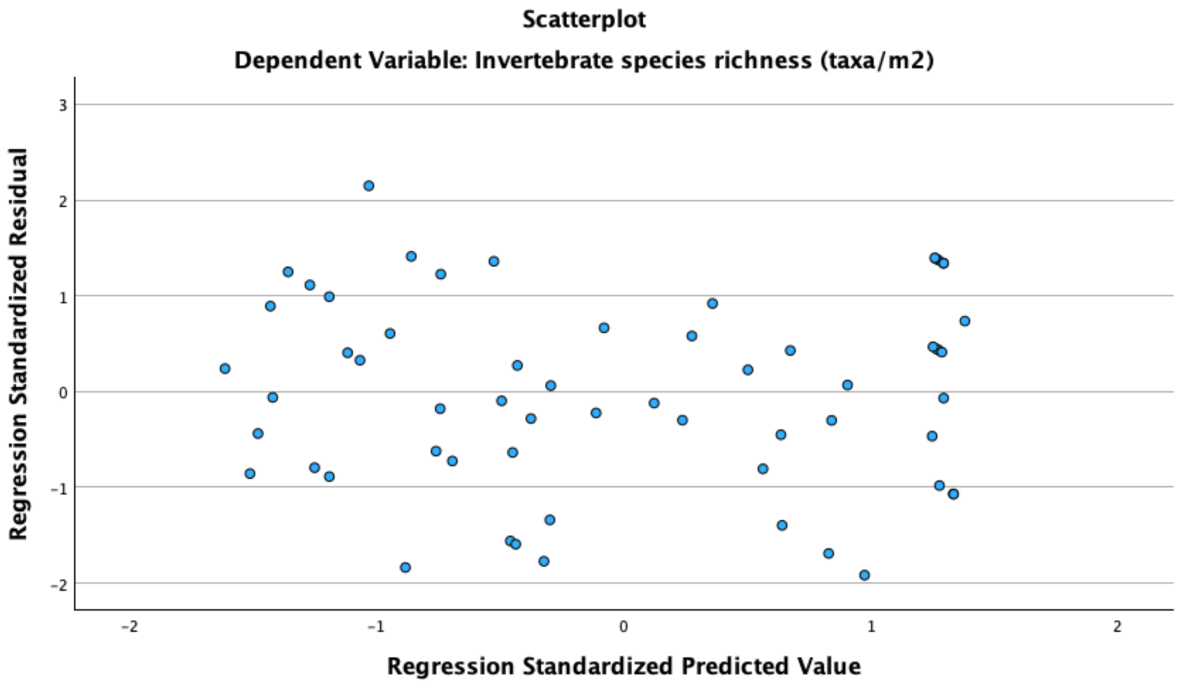

Figure 21: Scatterplot of standardised residuals for species richness model.

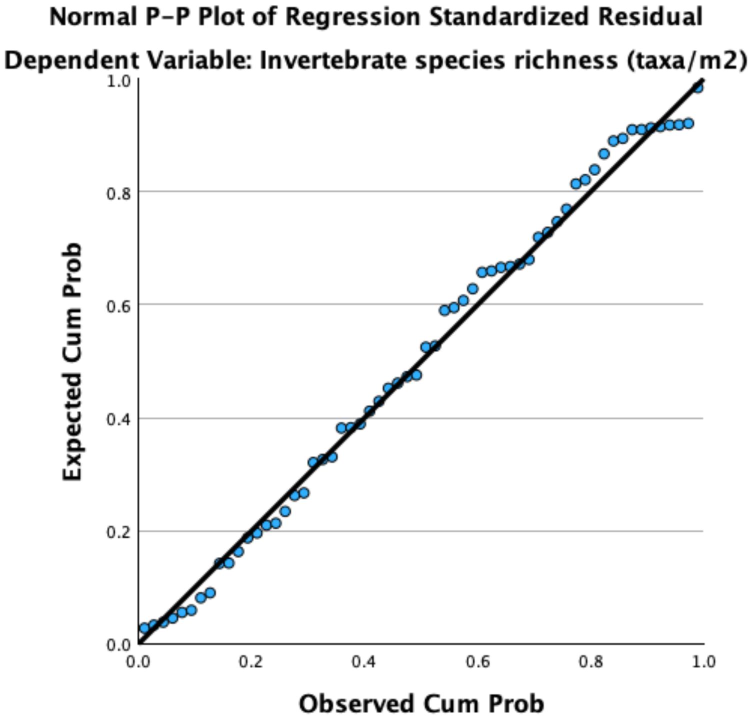

Figure 22: Normal P-P plot showing distribution of residuals for species richness model.

For species richness, the stepwise procedure retained only stored sediment in the final model (R^2 = 0.712, adjusted R^2 = 0.707). This single predictor explained 71.2% of the variance in species richness, demonstrating a strong negative relationship ($\beta = -0.844$, $p < 0.001$). All other variables were excluded as non-significant predictors during the stepwise process.

Model Critique:

Strengths:

- High explanatory power evidenced by strong R^2 values (0.812 and 0.712)
- Normal distribution of residuals confirmed by P-P plots
- Low VIF values (<1.1) indicating absence of multicollinearity
- Statistical significance achieved for all retained predictors ($p < 0.05$)

Limitations:

- Species richness model's reliance on a single predictor suggests possible oversimplification
- Potential interaction effects between variables remain unexplored
- Evidence of mild heteroscedasticity in residual plots
- Limited sample size may affect model robustness
- Possible omission of other influential variables

The analysis reveals stored sediment as the dominant factor in both models, showing consistent negative relationships with both abundance and richness. The additional influence of bed shear stress on abundance, but not on richness, suggests differing mechanisms affecting these two aspects of community structure. These findings align with the patterns observed in the earlier site-specific analyses (Questions 1-3), particularly regarding the inverse relationship between sediment levels and biodiversity metrics.

Question 5: Badlands Location Effect

To test the influence of badlands location, a dummy variable was created (0 = outside badlands [Sites 1 and 3], 1 = inside badlands [Sites 2 and 4]) and added to the previous regression models. For both models, the stepwise regression retained stored sediment as the primary predictor, with badlands location showing negative but relatively modest effects on both invertebrate abundance ($\beta = -0.045$) and species richness ($\beta = -0.127$). Though these results suggest lower biodiversity metrics within badland areas, the effect was not strong enough to be retained in the final models when accounting for other environmental variables, particularly stored sediment.