Partial regression and partial residual plots

FW8051 Statistics for Ecologists

Department of Fisheries, Wildlife and Conservation Biology

Learning Objective

Understand approaches for visualizing fitted multiple regression models

Visualizing Multiple Regression

$$Y \sim \beta_0 + X_1\beta_1 + X_2\beta_2 + \epsilon$$

β_1 reflects the “effect” of X_1 after accounting for X_2.

How can we visualize this “effect”?

Visualizing Multiple Regression

$$Y \sim \beta_0 + X_1\beta_1 + X_2\beta_2 + \epsilon$$

β_1 reflects the "effect" of X_1 after accounting for X_2.

How can we visualize this "effect"?

- Added variable or partial *regression* plots
- Component + residual or partial *residual* plots

See the paper by Larano and Corcobado (2008) and Section 3.14 in the Book.

Added Variable Plots (for X_i)

1. Regress Y against X_{-i} (i.e., all predictors except X_i), and obtain the residuals

Added Variable Plots (for X_i)

1. Regress Y against X_{-i} (i.e., all predictors <u>except</u> X_i), and obtain the residuals
2. Regressing X_i against all other predictors (X_{-i}) and obtain the residuals

Added Variable Plots (for X_i)

1. Regress Y against X_{-i} (i.e., all predictors except X_i), and obtain the residuals
2. Regressing X_i against all other predictors (X_{-i}) and obtain the residuals
3. Plot the residuals from [1] against the residuals from [2].

Added Variable Plots (for X_i)

1. Regress Y against X_{-i} (i.e., all predictors except X_i), and obtain the residuals
2. Regressing X_i against all other predictors (X_{-i}) and obtain the residuals
3. Plot the residuals from [1] against the residuals from [2].

Plots the part of Y not explained by other predictors (i.e., X_{-i}) against the part of X_i not explained by the other predictors (X_{-i}).

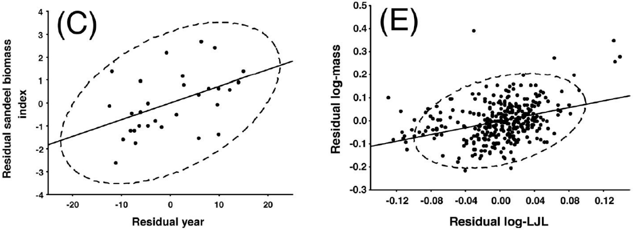
Added Variable Plots (for X_i)

1. Regress Y against X_{-i} (i.e., all predictors <u>except</u> X_i), and obtain the residuals
2. Regressing X_i against all other predictors (X_{-i}) and obtain the residuals
3. Plot the residuals from [1] against the residuals from [2].

Plots the part of Y not explained by other predictors (i.e., X_{-i}) against the part of X_i not explained by the other predictors (X_{-i}).

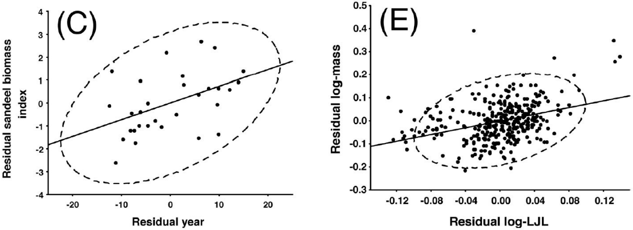
Lets us visualize the effect of X_i after accounting for all other predictors.

Partial regression plots



Shows the slope and the true scatter of points around the partial line in an analogous way to bi-variate plots in simple linear regression

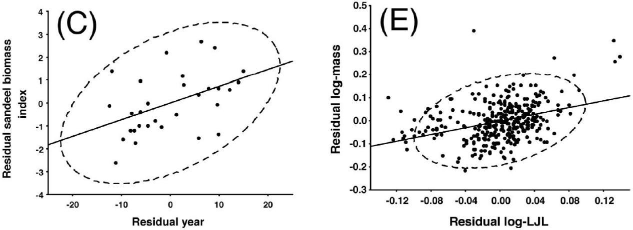
Partial regression plots



Shows the slope and the true scatter of points around the partial line in an analogous way to bi-variate plots in simple linear regression

- Tells us about the importance of X_2 (given everything else already in the model)

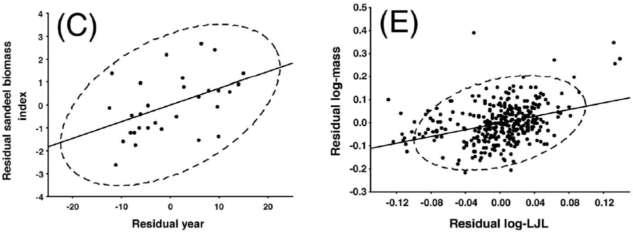
Partial regression plots



Shows the slope and the true scatter of points around the partial line in an analogous way to bi-variate plots in simple linear regression

- Tells us about the importance of X_2 (given everything else already in the model)
- Can help with diagnosing non-linearities

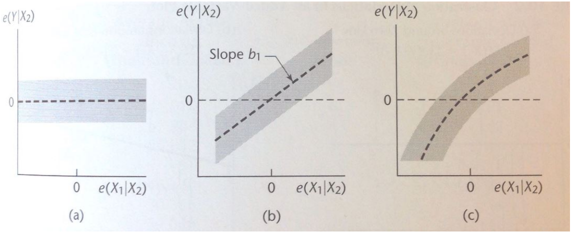
Partial regression plots



Shows the slope and the true scatter of points around the partial line in an analogous way to bi-variate plots in simple linear regression

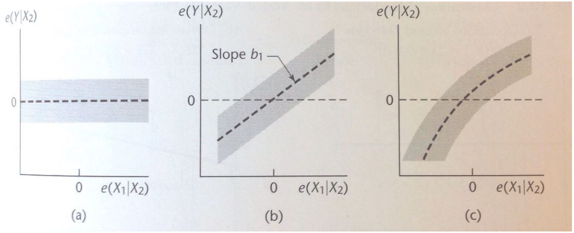
- Tells us about the importance of X_2 (given everything else already in the model)
- Can help with diagnosing non-linearities
- Helps visualize influential points and outliers

Added variable plot for X_1 (with one other predictor, X_2)



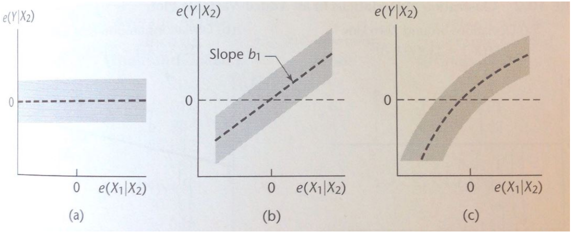
- Panel (a) suggests X_1 provides no additional information useful for predicting Y beyond that contained in X_2

Added variable plot for X_1 (with one other predictor, X_2)



- Panel (a) suggests X_1 provides no additional information useful for predicting Y beyond that contained in X_2
- Panel (b) suggest a linear relationship is appropriate (after accounting for X_2); the slope here is the same as that in the multiple regression model containing both X_1 and X_2

Added variable plot for X_1 (with one other predictor, X_2)

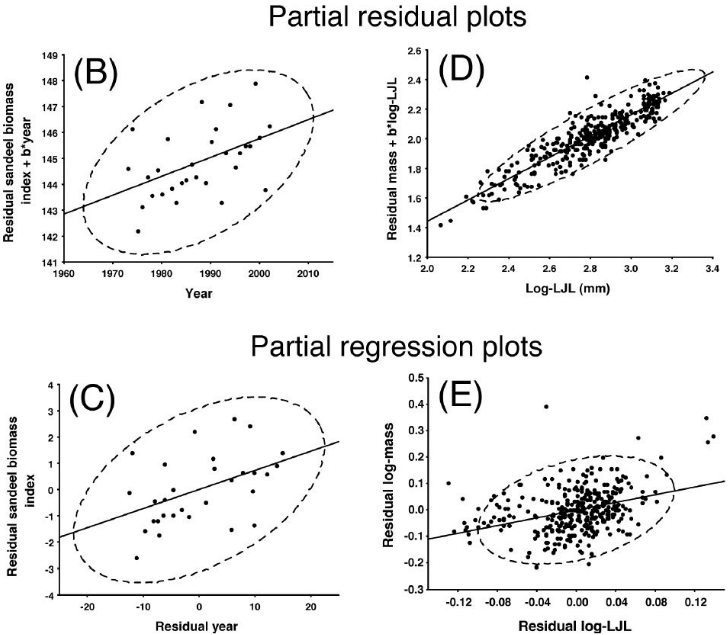


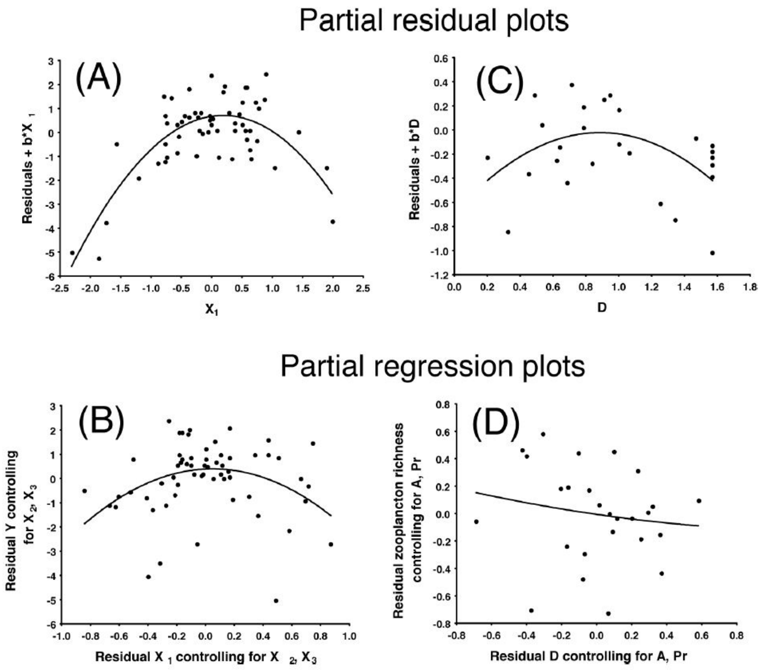
- Panel (a) suggests X_1 provides no additional information useful for predicting Y beyond that contained in X_2
- Panel (b) suggest a linear relationship is appropriate (after accounting for X_2); the slope here is the same as that in the multiple regression model containing both X_1 and X_2
- Panel (c) suggests we may need to allow for a non-linear relationship between X_1 and Y

Component + residual plots or partial residual plot

Plots $X_i\beta_i + \hat{\epsilon}_i$ versus X_i.

- Better for diagnosing non-linearities
- X-axis depicts the scale of the focal variable (rather than the scale residuals)
- Not as good at depicting the amount of variability explained by the predictor (given everything else in the model).
- Easy to generalize to other regression models (see visreg package on Canvas)


Partial residual plots
(B)
Residual sandeel biomass index + b*year
Year
(D)
Residual mass + b*log-LJL
Log-LJL (mm)
Partial regression plots
(C)
Residual sandeel biomass index
Residual year
(E)
Residual log-mass
Residual log-LJL


Partial residual plots
(A)
Residuals + b*X 1
X1
(C)
Residuals + b*D
D
Partial regression plots
(B)
Residual Y controlling for X2, X3
Residual X 1 controlling for X 2, X3
(D)
Residual zooplancton richness controlling for A, Pr
Residual D controlling for A, Pr

Effect plots

See Section 3.14.3 in the Book. Consider a focal predictor X_i and the set of all other predictors X_{-i}.

Effect plots

See Section 3.14.3 in the Book. Consider a focal predictor X_i and the set of all other predictors X_{-i}.

We can plot adjusted means by varying a focal variable over its range of observed values, while holding all non-focal variables at constant values (e.g., at their means or modal values).

Effect plots

See Section 3.14.3 in the Book. Consider a focal predictor X_i and the set of all other predictors X_{-i}.

We can plot adjusted means by varying a focal variable over its range of observed values, while holding all non-focal variables at constant values (e.g., at their means or modal values).

Depict $E[Y_i|X_{-i} = x_{-i}]$ versus X_i.

Effect plots

See Section 3.14.3 in the Book. Consider a focal predictor X_i and the set of all other predictors X_{-i}.

We can plot adjusted means by varying a focal variable over its range of observed values, while holding all non-focal variables at constant values (e.g., at their means or modal values).

Depict $E[Y_i|X_{-i} = x_{-i}]$ versus X_i.

Alternatively, we can plot marginal means. These are formed in much the same way, except that predictions are averaged across different levels of each categorical variable.

Effect plots

See Section 3.14.3 in the Book. Consider a focal predictor X_i and the set of all other predictors X_{-i}.

We can plot adjusted means by varying a focal variable over its range of observed values, while holding all non-focal variables at constant values (e.g., at their means or modal values).

Depict $E[Y_i|X_{-i} = x_{-i}]$ versus X_i.

Alternatively, we can plot marginal means. These are formed in much the same way, except that predictions are averaged across different levels of each categorical variable.

These two types of means are equivalent if there are no categorical predictors in the model.