Regression and Path Analysis with Application for Real Data

Dr. Adel Ahmed Haddaw

Mustansiriyah University Iraq - Baghdad

Abstract: *The purpose of this paper is to usepath analysis via multiple regression model withapplication for real earthquake data in Iraq and Iran. A path analysis involves theanalysis and comparison of two models – a full model with all of the possible paths included and a reduced model which has some of the paths deleted. The main result of this paper is that the reduced model fits the data as well as does the full model.*

1. Introduction

Statistically, the most important of a path analysis is which causal variables to include in the model, how to order the chain of those variables and which paths are not important to the model. Whilethe most important of regression analysis iswhich independent variablesare important and more an effect in the model, how the regression model to predict the dependent variable based on the values of the independent variables, [3].

The purpose of this paper is to usepath analysis via multiple regression model withapplication for real earthquake data in Iraq and Iran, [1].

2. Material and methods

In this section we present methodology that includes the hypothesized causal ordering for how Lat., long and Depth cause ML (The Mount local earthquake). Usually a path analysis involves theanalysis and comparison of two models – a full model with all of the possible paths included and a reduced model which has some of the paths deleted, because they arehypothesized to not contribute to the model, [2]. So the graph can be shown in Figure (1).
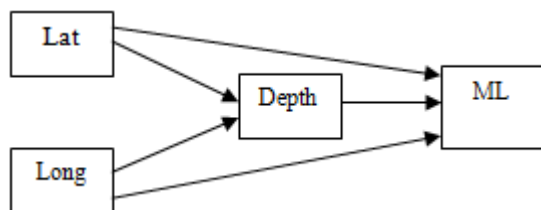

Figure 1: Relationship among Variables

The path coefficients for the full model (with all the arrows) are derived from a series of (layered) multiple regression analysis, [5]. For eachmultipleregression, the criterion is the variable in the box (all boxes after the leftmost layer) and the predictors are all the variables that have arrows leading to the box.

For the full model above, we will need two layers ofmultiple regressions.
1) With Depth as the criterion and Lat. and Long as the predictors.
2) With ML as the criterion and Lat., Long. and Depth as the predictors.

3. Results and Discussion

SPSS and AMOS, [4] have been used to start with a correlation matrix of the all variables, providing an opportunity to reanalyze their variables using model used is shown in Table1. 1.

Table 1: A correlation matrix

		ML	Lat	Long	Depth
Pearson Correlation	ML	1.000	-0.106	0.402	0.216
	Lat	-0.106	1.000	-.879	-0.157
	Long	0.402	-0.879	1.000	0.257
	Depth	0.216	-0.157	0.257	1.000

It can be seen that the largest value of correlation is between ML and Long.(0.402) while the smallest value of correlation is between ML and Lat.(-0.106).

Getting the -First Layer- multiple regression for the full model.

Model	R	R Square	Adjusted R Square
1	0.295	0.087	0.049

	Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-79.865	58.118		-1.374	0.176
	Lat.	0.716	0.685	0.302	1.045	0.301
	Long.	1.423	0.788	0.523	1.806	0.077

Getting the -Second Layer- multiple regression for the full model.

Model	R	R Square	Adjusted R Square
1	0.658	0.432	0.396

	Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-35.222	6.987		-5.041	0.000
	Lat	0.377	0.082	1.075	4.613	0.000
	Long	0.539	0.096	1.337	5.611	0.000
	Depth	0.006	0.017	0.042	0.365	0.717

Portraying the Full Path Model
1) The path coefficients are the β weights from the multiple regression analyses.

2) The e values are computed as $\sqrt{(1- R^2)}$ (for example $e_{Depth} = \sqrt{(1- 0.432)} = 0.753$. So the graph can be shown in Figure (2).
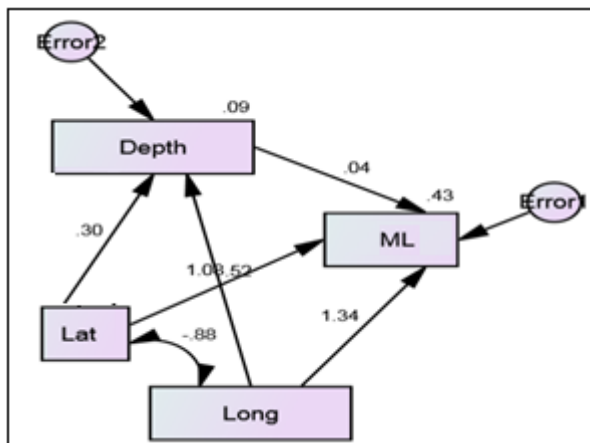

Figure 2

Examining this model
1) Depth not influences ML.
2) Lat. has direct an effect upon ML, but has no an indirect effect through- Depth.
3) Long has direct an effect upon ML, but has no an indirect effect through- Depth.

Other test hypotheses about which model paths do not portray causal links among the variables. Below is such a reduced mode.

The model posits that is Depth notinfluences on ML, once again, two multiple regression model would be used to obtain the path coefficients.

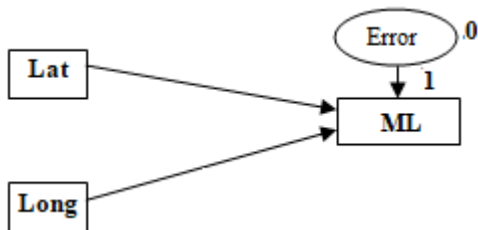For this layer we would use the following analysis: So the graph can be shown in Figure (3).

Figure 3

Model	R	R Square	Adjusted R Square
1	0.656	0.431	0.407

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std.Error	Beta		
1	(Constant)	-35.717	6.792		-5.259	0.000
	Lat	0.381	0.080	1.088	4.763	0.000
	Long	0.548	0.092	1.359	5.948	0.000

Portraying the reduced Path Model
1) The path coefficients are the β weights from the multiple regression analyses.
2) The e values of dependent variable are computed as $\sqrt{(1- R^2)}$ (for example $e_{ML} = \sqrt{(1- 0.431)} = 0.754$

3) Lat and Long have influences ML same as in full model. Portraying the Reduced or Hypothesized Path Model.So the graph can be shown in Figure (4).
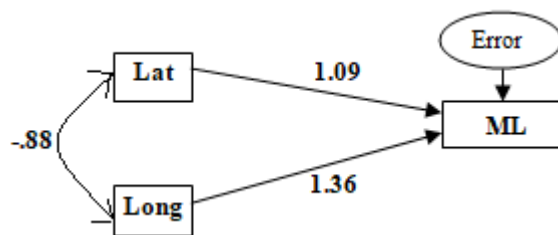

Figure 4

4. Conclusion

We would conclude that the reduced model fits the data as well as does the full model. That is, a causal model deleting the Depth variable as moderator variable did fit the data very well.

References

[1] Diana, Suhr. Step your way through path analysis. Northern Colorado (2000).
[2] Joel, S, Steele. Path. Analysis introduction and example (2017).
[3] Kubra, K.O and Beyza, A.DA review of structure equation modeling applications in Turkish education (2018).
[4] Randall, E.S and Richard, G. LA beginner's guide to structural equation modeling (2011).
[5] Shipley .Exploratory path analysis with application in ecology and evolution. The American Naturalist (1997).