The Importance of Regression Analysis in Heavy Metal Content

Prof Dr Mustafa Gurhan YALCIN *1

Abstract

The acquisition and analysis of numerical data is of great importance in the field of natural sciences. The acquisition of numerical data is a common practice in certain professional clusters and analytical fields. In the descriptions of data on heavy metals in the studied natural environments or the results of chemical analysis, it is very important whether the data are significant or not. In this case, it is advisable to use regression analysis with R^2 value, Durbin-Watson statistic and significance test. It is very important to determine the values of the Y-model.

2020 MSC: 62G08, 62J02, 62-06, 62P12.

KEYWORDS: Regression Analysis, Heavy Metal

1 Introduction

Heavy metal analysis has gained importance in mining, environmental pollution and sustainable environment studies. Chemical analysis results and heavy metal analyses are in demand in multivariate statistical analysis and regression analysis (*cf.*[1]-[2]-[4]-[5]-[6]-[7]-[8]-[9]-[10]-[11]-[12]-[14]-[15]-[16]).

Regression analyses have different areas of use with different methods. In studies where original data are provided, interpretations can be made by performing linear regression analysis of the relationship between the knowers among themselves. opinions can be obtained about whether the data are sufficient or not, and how accurately the data explain the results.

Therefore, the aim of the study is to explain the importance of performing regression analysis in heavy metals. Examples made in this context will be examined.

2 Method

The application of regression analysis, which is one of the statistical methods of chemical analysis data, is a frequently applied method.

It can be easily interpreted by regression analysis with programs such as SPSS software. SPSS software has been frequently used in regression analysis of heavy metal analysis and can be recommended for such analysis.

3 Results

In scientific studies, the significance of relationships between data obtained from analysis results can be readily determined. In such cases, particularly when the R^2 value is approaching 1, it can be inferred that the relationship between the variables under analysis and comparison is significant. A value of less than 0.5 indicates that the relationship between the variables is inconsequential. Conversely, a value approaching 1 indicates a meaningful relationship between the variables. Upon examination of the previously chemically analyzed scientific study, it was evident that the geochemical relationships of PbO-Pt (PbO (%) $= -1 + 3.22E2.Pt$ (%) $= 0.974$) and PbO-Bi ($PbO(\%) = 0.57 + 2.35E2.Bi = 0.972$) were statistically significant. However, the PbO-ZnO relationship was not found to be significant (*cf.*[13]). Similarly, in a different study, the results of the regression analysis of the physical and chemical analysis of the water samples yielded a Durbin-Watson value of 1.881, anR^2 value of 0.997, and a significance value of 0. It can be stated that the established model is statistically significant. The relationship between the residual deleted from regression and the residual standardized from regression was obtained as a simple linear equation, $y = 0.04 + 0.09.x$, with an R^2 value of 0.914 (*cf.*[3]). A comparable outcome was observed in Figure 2. The R^2 value was 0.914, indicating a high degree of significance in the data (*cf.*[3]).

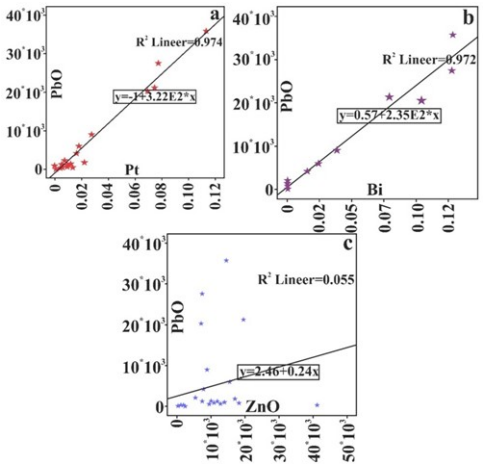


Figure 1: Behaviors of PbO-Pt, PbO-Bi and Pb-ZnO geochemical associations were determined by linear regression analysis (*cf.*[13])

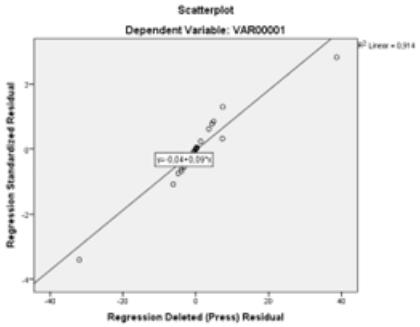


Figure 2: Linear scatter plot $y = 0.04 + 0.09.x$; R^2=0.914 (*cf.*[3])

In a different study, the importance of linear and non-linear model was emphasized (*cf.*[2]). In the tests conducted, it was determined which model was superior.

	Linear Model				Nonlinear Model			
		Std.err.	t	p		Std.err.	t	p
b0	19.51*	0.921	21.17	0.000	19.826	0.052	56.28	0.000
b1	-0.011*	0.002	-4.04	0.007	-0.0007	0.0001	-4.38	0.005
RME (Root Mean Square Error)	0.8325				0.0477			
R^2 (R Square)	0.7316				0.761			
F-statistics	16.35			0.006	19.14			0.005
Breusch-Pagan/Cook-Weisberg heteroscedasticity test	P=0.2478>0.05				P=0.251>0.05			
Shapiro_Wilk W Normality test for residual	P=0.356>0.05				P=0.514>0.05			
Ramsey RESET test 0.05 meaningful at error level	P=0.352>0.05				P=0.359>0.05			

Table 1: Regression result of Linear and Nonlinear model (*cf.*[2])

In addition to these studies, it is also important to examine the results of the model summary and analysis with respect to the variables. It is also crucial to consider the R^2 value, the level of significance, and the Durbin–Watson value.

4 Conclusion

A linear regression analysis of the data obtained from the results of chemical analyses can determine whether the relationship between variables is significant or not. This method is recommended for use in this field.

References

[1] O. Ozer Atakoglu, E. Berberoglu, F. Yalcin, S. Gokaydin, E. Akkopru and M. G. Yalcin, *Source identification of heavy metal contamination*

in beach sediments of the ancient city of Phaselis in Antalya/Turkey, International Journal of Sediment Research, 2024.

[2] B. Aydin, F. Yalcin, O. Ozer Atakoglu and M. G. Yalcin, *Regression analysis and statistical examination of Knoop hardness on abrasion resistance in Lyca beige marbles*, Filomat, **32 (2)**, 2020.

[3] B.Cayhan and M. G. Yalcin, *Regression analysis of chemical compound of Hurma region spring water* , 6th Mediterranean International Conference of Pure and Applied Mathematics and Related Areas (MICOPAM 2023), Proceedings Book of Micopam 2023, 99-102, Paris, France, 2023.

[4] D. G. Nyamsari and M. G. Yalcin, *Statistical analysis and source rock of the Minim-Martap plateau bauxite, Cameroon*, Arabian Journal of Geosciences, 10, 1-16, 2017.

[5] F. Yalcin, O. Ozer, D. G. Nyamsari and M. G. Yalcin, *Statistical evaluation of the geochemical content of beach sand along the Sarisu-Kemer coastline of Antalya, Turkey*, In AIP Conference Proceedings, (Vol. 2116, No. 1), 2019.

[6] M. G. Yalcin, M. Setti, F. Karakaya, E. Sacchi and N. Ilbeyli, *Geochemical and mineralogical characteristics of beach sediments along the coast between Alanya and Silifke (Southern Turkey)*, Clay Minerals, **50 (2)**, 233-248, 2015.

[7] M. G. Yalcin, D. G. Nyamsari, O. Ozer Atakoglu and F. Yalcin, *Chemical and statistical characterization of beach sand sediments: implication for natural and anthropogenic origin and paleo-environment*, International Journal of Environmental Science and Technology, **19 (3)** , 1335-1356, 2022.

[8] M. G. Yalcin, G. Simsek, S. B. Ocak, F. Yalcin, Y. Kalayci and M. Karaman, *Multivariate statistics and heavy metals contamination in beach sediments from the Sakarya Canyon, Turkey*, Asian Journal of Chemistry, **25 (4)**, 2013.

[9] M. G. Yalcin, O. Cevik and M. Karaman, *Use of multivariate statistics methods to determine grain size, heavy metal distribution and origins of heavy metals in Mersin Bay (Eastern Mediterranean) coastal sediments*, Asian Journal of Chemistry,**25 (5)** , 2013.

[10] O. Ozer and M. G. Yalcin, *Correlation of chemical contents of Sutlegen (Antalya) bauxites and regression analysis*, In AIP Conference Proceedings (Vol. 2293, No. 1), AIP Publishing, 2020.

[11] O. Ozer, F. Yalcin, O. K. Tarinc and M. G. Yalcin, *Investigation of suitability of marbles to standards with inequality expressions and statistical approach using some physical and mechanical properties* , Journal of Inequalities and Applications, 1-15, 2020.

[12] O. Ozer Atakoglu, M. G. Yalcin, Y. Leventeli and B. T. San, *Geochemistry of Red Soils in the Kas District of Antalya (TUrkiye) Using Multivariate Statistical Approaches and GIS*, Minerals, **13 (6)**, 772, 2023.

[13] O.Ozer Atakoglu and M. G. Yalcin,*Regression Analysis of Chemical Contents of Erva Pb-Zn Deposit Kayseri (Yahyali)*, The 35th International Conference of Jangjeon Mathematical Society (ICJMS 2024), 20-25, Seul, August 2024.

[14] O. Ozer, F. Yalcin, D. G. Nyamsari and M. G. Yalcin, *Appraisal of metal accumulation in beach sand using contamination indices and multivariate statistical analysis*, Proceedings Book of the 2nd Mediterranean, 32, 2019.

[15] F. Yalcin, S. Unal, M. G. Yalcin, O. Akturk, S. B. Ocak and S. F. Ozmen, *Investigation of the effect of hydrothermal waters on radionuclide activity concentrations in natural marble with multivariate statistical analysis*,**12 (8)**, 1219, Symmetry, 2020.

[16] M. G. Yalcin, E. Mutlu, C. Olguner, O. O. Atakoglu, L. Bat and E. Y. Ozkan,*Spatial geochemical structure of soft sediment on shallow littoral of the Gulf of Antalya, the eastern Mediterranean Sea* Marine Pollution Bulletin, 193, 115155, 2023. FUSUN HOCA

[17] F. E. Harrell, *Regression modeling strategies*, R package version, 6–2, 2012.

[18] A. K. Vasisht, *Logit and probit analysis*, IASRI, Library Avenue, New Delhi, **110 (12)**, 2007.

[19] princeton, *Getting Started in Logit and Ordered Logit Regression*, last access: 8/10/2024, https://www.princeton.edu/ otorres/Logit.pdf

[20] wikipedia, *Probit model*, last access: 8/10/2024, https://en.wikipedia.org/wiki/Probit_model

[21] theanalysisfactor, *The Difference Between Logistic and Probit Regression*, last access: 8/10/2024 https://www.theanalysisfactor.com/the-difference-between-logistic-and-probit-regression/

[22] A. Ozturk and F. Yalcin, *Logit Regression Modelling of The Support For Women in Working LifeSurvey*, The 7th Mediterranean International

Conference of Pure and Applied Mathematics and Related Areas (MICOPAM 2024), which is dedicated to Professor Abdelmejid BAYAD on the Occasion of his 61st Anniversary, will be held at Sherwood Dreams Resort Hotel (Ultra All Inclusive) in Antalya, TURKEY on October 26–29, 2024.

Akdeniz University Faculty of Engineering Department Of Geology 07058 Antalya-Turkiye [*1]

Email(s): gurhanyalcin@akdeniz.edu.tr [*1] (corresponding author)