An Overview of Regression Analysis in Polymers Chain Reaction (PCR)

[1]Saif H. Jalil; [2]Irtefaa A. Neamah

[1,2] Department of Mathematics/ Faculty of Computer Science and Mathematics

University of Kufa

Abstract

This paper examines the various uses of regression analysis in the polymer chain reaction (PCR) test to detect various infections and diseases. We found that regression is of great importance in developing this test method, especially after using Kaplan-Meier analyses, where the accuracy of the results was more than 95%, as well as the two Cox proportional hazards methods, Ct method and ANCOVA. As a result, regression has a major role in developing the work of this interaction, according to the published research papers that were addressed in this study.

Keywords: Regression, Simple Linear Regression, Multiple Linear Regression, Polymerase Chain Reaction (PCR).

1. Introduction

Regression analysis is the analysis of relationship between dependent and independent variable as it depicts how dependent variable will change when one or more independent variable changes due to factors, most regression models propose that Y_i is a function of X_i and β with ε_i representing an additive error term that may stand in for un-modeled determinants of Y_i or random statistical noise [1]:

$$Y_i = f(X_i, \beta) + \varepsilon_i \qquad \qquad \text{…(1)}$$

The first form of regression was the method of least squares, published by Legendre in 1805 and Gauss in 1809. Both Legendre and Gauss applied the method to the problem of determining the orbits of bodies around the Sun, through astronomical observations. Gauss published a further development of the theory of least squares in 1821. In the fifties and sixties of the last century, economists used electromechanical desktop calculators to calculate regressions. before 1970. Regression methods are still an area of active research and new methods have been developed for robust regression, regression that includes correlated responses such as time series and growth curves, and regression in which predictor (independent variable) or response variables are curves, images, graphs, or data objects other complex. Below is a large group of research that used multiple regression analysis in various fields, and which was able through this analysis to reach important results and have a significant impact on human life at various educational, environmental, health, scientific and other levels:

Dempster emphasized some concepts in data analysis structure such as logical, epistemic an hypothesis structure as well as examined them in a discussion of exponential families models in missing values data. [2] . Here, the researcher studied the relationship between heterogeneity bias and strict homogeneity in the lag regression, and the results indicated that leads and delays can occur, which were generated by the longitudinal intercept [3] . Kung-Yee Liant et al. discuss a combination of predictive models with nonlinear models in terms of efficiency and robustness in health data analysis [4] . In 1997, researchers Leo Breiman and Jerome H. Friedman were able to introduce a new procedure called the curd and whey method that could be used to significantly reduce prediction errors when there is a correlation between responses while maintaining accuracy [5]. The researcher used the (MRT) method to explore the relationships between multiple species and environmental characteristics in order to describe and predict them, as this method worked on analyzing complex data such as imbalance, missing values and prediction as well [6]. The quantification of mRNAs through the use of real-time polymerase chain reaction to monitor product formation is used in various fields, including medical diagnosis and neuroscience, and the procedures assume that the efficiency of the polymerase reaction under study is either constant or equal and the efficiency of PCR is calculated through the curves This study proposes a linear regression of all cycle number data as an assumption-free method [7].The researcher was able to determine the impact of the multiple linear relationship on multiple environmental regression through the use of simulation as well as real environmental data, and he explained the usefulness of various statistics to explain multiple environmental regression [8]. In this research, a linear regression analysis was performed on project data in North East Scotland in order to solve the problem of inaccurate estimation of concrete productivity. The results showed a strong linear relationship between the selected operating conditions and between productivity [9]. And due to the lack of processing procedures PCR for real-time analysis in the field of statistical processing, the researchers worked on presenting and comparing four statistical models for data analysis, and the results showed that the procedures presented here have provided statistical elements to estimate the relative expression of genes [10]. The aim of this study is to determine whether tumor-related markers have a role in detecting the spread of malignant cells and to identify patients with colorectal cancer in its early stages, by conducting Cox proportional hazards regression analysis, and the results showed that 41 out of 125 of Patients were positive for disease spread in the early stages, and thus the markers used in this study are an important diagnostic tool for

improving the management and regulation of colorectal cancer [11].Through this study, the researchers tried to find a relationship between the home environment and reading achievement in Zimbabwe. The results showed that the proxy was the strongest predictor of reading achievement through structural equation modeling [12]. The possibility of accurate determination of target DNA from environmental samples is still possible although current methodologies are very sensitive, and in order to facilitate the measurement of high-throughput DNA, a new quantification of DNA has been developed which is a method based on a different approach which is curve fitting Non-linearity in order to extract additional information through quantitative PCR amplification curves. The results clearly showed that the DNA content that was determined was more accurate and the analysis of variance revealed fewer statistical artifacts compared to the traditional Ct-based method and this new approach is accurate and inexpensive [13].Noroviruses are considered one of the most important causes that lead to gastro-intestinal infections for different age groups, but due to the diversity of the sequences of these genetic groups, the design of diagnostic PCR tests for the whole of this virus is complex, and through this research a new reverse transcription protocol was developed One-step detection of this virus in humans by using Kaplan-Meier time-dependent analyzes and Cox multivariate regression analyzes for immunosuppression as a single risk factor [14]. Several linear regression equations have been reviewed in order to link water and moisture content in flower honey in order to arrive at the weighted average regression equation, which is a general linear relationship for calculating honey water on the basis of honey moisture content, and it was obtained through comparison and addition Among the results of ten independent studies that included 638 observations from different countries, so this proposed equation would confirm that glucose and fructose solutions are the main determinants of water activity in honey [15]. In this study, the latest meta-analysis methods were implemented in order to measure the effectiveness of a comprehensive and long-term applied analytical intervention for children with autism, and the results indicated significant positive effects from an intellectual perspective, which were superior to nonverbal IQ [16]. And because the multiple linear relationship may increase the difficulty of interpreting multiple regression, so the researchers tried in this research paper to find multiple indicators to understand the contributions that predict not only the regression model but also to each other [17]. The researchers used a new approach to estimate the energy performance of the building in the early design stages through the Monte Carlo simulation program, and the results showed that the linear regression model can be an effective tool for decision support during the early design stages [18]. This study aimed to enhance the performance of industrial enterprises with mathematical models and in order to increase the ability in terms of competition and adaptation. The researchers found through regression analysis that there is a relationship between the dependent variable and the independent variables, which are very important predictions for increasing the volume of profit [19]. This technique has been used by the researcher to elicit honest answers when the questions are sensitive. Least squares and maximum probability estimators have been used for multivariate regression analysis. The investigators used this technique to measure the degree of racial hatred in America in 1991 [20]. Real-time quantitative polymerase chain reaction is a sensitive quantitative determination of genes and is a widely used method in various fields, but the linear and non-linear thresholding method Linearity and model fitting methods in PCR data analysis require great effort and background defluorination may not be accurate, so the results will appear distorted. Therefore, this paper proposes an alternative method, which is the difference-taking method of linear regression, meaning that for every two cycles of the polymerase chain reaction, the fluorescence in the previous cycle is subtracted from the one in the last cycle, and this linear regression is then applied to the natural logarithm of the transformed data, and with this The method avoids the error in subtracting the unknown background and is thus more accurate and reliable [21]. This research paper proposes a model to predict the demand for heating energy, depending on many factors that affect the heat consumption of the building through conducting multiple dynamic simulations, and the calculated results using the building energy dependence methodology showed that the model provides a very good accuracy [22]. Modeling energy consumption is important in order to reduce energy consumption and emissions, so this research paper was able to find a new model to predict and quantify energy consumption in the early stages of building design, using Monte Carlo simulation techniques, where ten thousand simulations of seven shapes of buildings were conducted. The results showed a good agreement between the data predictions according to the developed regression model and the simulation of the Ministry of Energy, where the error was less than %5 [23]. This research is concerned with analyzing the effects of regression in the individual study where people are exposed to different failures, and regression models are taken into account in order to link the risk or severity function to the covariates and the date and time of the previous failure [24] .

In this paper, new methods have been used to compare seven PCR analysis methods in their performance, quantification cycle estimation, and amplification efficiency for a large set of test data the linearity and accuracy were evaluated through the use of chi-square, which is equivalent to the variance and as a miniature quantity in the installation of least squares, and this method is used to determine the statistical efficiency [25].The relationship between time and event in terms of results is considered common, especially in medical research, and in order to deal with these results in addition to monitoring and recording observations, survival analysis methods are used, and the study aimed in this research paper to review the general and basic concepts Survival analysis methods as well as their use and interpretation through the use of simulated data [26]. Depending on the definition of the chemical to predict the quality of wine, multiple linear regression has been applied, through three different models, which are the ripening coefficients of grapes, the parameters of wine and the mixture between them, and the mixture between them has been analyzed, and the multivariate statistical results were promising for prediction Quality wine in an inexpensive and easy way too [27]. PCR is a laboratory procedure that determines the amount of target DNA molecules, and the reaction product has discovered a more modern technique known as quantitative qPCR. This study compares the accuracy of the estimate achieved by the eight different models when applied to the same group, and this study also compares the approach The results showed that taking the difference is the best method for pre-treatment and that the estimation accuracy achieved by the mixed models is slightly better than what was achieved by the linear regression models [28]. Due to the importance of designing the geothermal heat exchanger in the geothermal heat source, it was developed using the finite size method in order to evaluate the effects of thermal conductivity and ground porosity, etc., and through the use of the study of Darcy's velocity and well depth, nine correlations were obtained that were able to pass Significance test The results indicated the possibility of estimating the heat flow with a root mean square error of 15.23 W/m [29].

The shortage of fresh water in different regions of the world has become a big problem, and one of the reasons for this shortage is largely due to urbanization and climate change, and therefore the purpose of this study is to try to improve our behavior by using water reliably, and the focus has been on social factors and economic instead of time factor as in common studies, and results have been obtained that prove the quality of the new approach quantitatively, and the analytical procedures can be used in this study in different cities to analyze water use [30]. Finding a relationship between the different variables is very important for prediction and analysis, and regression analysis is one of the most important methods used for this. The research paper here focuses on the different models for regression analysis and how to use them in different data , Ridge and Lasso regression are used when the data becomes complex and there are problems such as improper scaling and overfitting, so the use of these models can lead to more accurate results [31]. In order to use operating rooms effectively, accurate timetables must be adhered to for each operation and taking into account the sequence of patients' cases. Data from six Dutch academic hospitals between 2012 and 2016 have been used, and results have been obtained indicating great accuracy in predictions, and the performance of this model is significantly better than the fixed ratio model, which returns material benefits as well as productivity [32]. This study has given a multivariate statistical analysis with regard to the various categories of students, through the creation of a multivariate regression model, in order to show the important factors affecting the quality of teaching in universities, as well as in order to develop solutions for them [33]. The use of woody plants has led to the need to obtain a quick and effective estimation of carbon stocks in forests in terms of cost, and therefore this study aimed to find the factors associated with The carbon stock in the Shore forest in the state of Nepal and the evaluation of these factors, and the correlation between the variables was observed through the use of variations, and a positive correlation with the carbon stock was obtained through the graph, while the height, ownership and geographical location had no indications Statistic [34]. Research related to sound explores many factors that certainly affect human perception in terms of emotional and behavioral terms, and this exploration needs to study the relationship between personal, social, cultural and other factors and between the impact of sound on the other hand.Therefore, this paper focuses on regression analysis in order to investigate multiple variables in order to study relationships and predict the effects of vocal range and evaluate it in terms of satisfaction and comfort [35].Estimating risks and predicting future opportunities is of great importance in the success of business, and the technique of linear regression analysis is one of the most important statistical methods that examine the relationship between variables through prediction based on data and finding the relationship between variables, and this research examines the ability of linear regression to explain Suitable for research and business data [36]. In this research paper, the researchers studied how to manage energy through the use of regression analysis in addition to various algorithms based on a set of input variables and only two output variables, which are heating load and cooling load, and knowing the effect of each input variable on the output variables. The results showed that the model is accurate, and it can be used to know the cooling and heating loads for any building in the early stages [37]. Multivariate regression is used in many applications, including genomics, as well as time-series prediction, and the use of the results reached helps improve prediction error, and the best method is chosen by establishing a connection between linear regression and complex decision problems, and from During Bayes' experimental procedure, it was found that it is a procedure free of parameter tuning, and its results are good in simulations, as well as in predicting the prices of multiple stocks [38]. The research paper analyzes the relationship between the development in occupational health indicators and between each of the risk rate and the performance of factors affecting the labor market and lost days using the regression contraction, and the study focused on occupational accidents in Spain in the time period 1995-2017, which is considered This period is one of the most important stages of economic growth, and a conclusion has been reached confirming that the regression method AdaLasso is one of the strong and most consistent methods and is able to explain the different relationships with a small number of variables [39]. This study worked on developing a data table based on Microsoft Excel, which is considered a model that researchers can rely on , as the data table contains the degree of confidence and standard error, as well as the appropriate quality measures in order to know the effects of isothermal experimental design, and it has been proven that the test F is the standard The appropriateness of the appropriateness of quality was evaluated and the possibility of variance of the model and experimental data was evaluated the same [40]. The high level of dynamic loading, building shape, repair effects, low qualifications of drivers, among other factors, are among the main reasons for the failure of mining excavators [41]. The objective of this study is to build a kernel-type estimator file for the expected regression function, and to reach the natural state that is close to the proposed estimator and under moderate conditions, and the limited sample is used to clarify the performance of the model and the estimation method through the use of real and simulated data as well [42]. Linear regression is one of the most important and most popular statistical and machine learning algorithms, where linear regression is used to find a linear relationship between one or more variables. In this paper, the different work done by the researchers on polynomial regression is discussed and their performance is compared to improve prediction and accuracy [43] .

This study deals with multiple linear regression analysis, and focuses specifically and mainly on the methods of multiple regression analysis in the polymerase chain reaction (PCR). It is one of the most important techniques used in various medical fields and biological laboratories and has been used more and more widely as a method Successful and rapid in detecting infection with the emerging corona virus.

2. Regression Analysis

Multivariate regression is a machine learning algorithm that includes multiple data variables for analysis, where regression attempts to discover a formula by which to explain how the factors in the variables respond to other changes at the same time and below is the general equation of a multivariate regression model:

$$y = \beta 0 + \beta 1.x1 + \beta 2.x2 + \dots + \beta n.xn \qquad (\dots 2)$$

Where n represents the number of independent variables, $\beta 0 \sim \beta n$ represents the coefficients and x1~xn, is the independent variable [44].

Linear regression is a modeling technique in which an association is expected between a dependent variable and one or more independent variables. This analysis is the most widely used among various statistical methods [45]. Nonlinear regression is a type of regression analysis in which the data fits a specific model and is expressed mathematically, and the relationship between the two variables is nonlinear [46].Univariate analysis is the simplest type of data analysis and deals with data of only one variable and its main purpose is description and does not deal with causes or relationships [47]. As for Cox regression analysis, this analysis is a very popular method and was developed in 1972, this analysis is used in the simultaneous evaluation of the impact of many factors on survival, and it is of great importance in many applications and is also known as the proportional hazards model [48] .

3.   Polymerase Chain Reaction (PCR)

Polymerase chain reaction (PCR) is a common molecular biology technique that has been used to clone enzymatic DNA without using an organism such as yeast or Escherichia coli. This technology was discovered by Kari Mullis in the year 1983, and this technology has taken place in the fields of medicine and biological research laboratories for a wide range of applications, including the discovery of DNA-based evolution, DNA cloning and functional analysis of genes. The polymerase chain reaction has many uses for diagnosing various diseases, including: diagnosing cancer, especially leukemia and lymphoma, as well as pneumonia. It is also used to diagnose genetic fingerprinting, paternity testing, early diagnosis of malignant tumors, and direct PCR tests on genomic DNA samples for detection of malignant cells.

To perform the polymerase chain reaction, several components are required:

1- DNA template or c-DNA containing the region profile of the DNA to be amplified.

2- Determine the beginning and end of the area to be amplified.

3- Taq polymerase.

4- The nucleotides that make up the DNA - the new DNA polymerase.

5- Finally, the DNA-polymerase environment that provides a suitable chemical [49].

The principle of the PCR is based on the fact that the higher denaturing temperatures when they are near 95 ° C, the strands of the DNA molecule, which is the target, will be separated due to the breakage of the A-T and G-C bonds. Thus, reverse batteries stick to the end of three sides of the target DNA that has been stranded because it has been separated, and then the new polymerase extends two strands of DNA by adding dNTPs and the double strand that is separated to restructure the molecule itself upon extension under temperature 72 ° C , this process is repeated several times, thus generating many copies of the target DNA [50].

4.   Discussion

All papers included in this review are observational studies that used regression models from 1970 to 2020. Table 1 provides a summary of each study reviewed in this paper whose research scope was PCR. Which includes the main points in each paper in terms of the type of variable, the method, used to reach the results, the types of multivariate regression used in each research paper, and what the data is in terms of being real data or simulated data, and whether it is a mixture of the two types, and also includes the summary In this table, the quality of the disease and finally the accuracy of the results that were reached. As shown in the table, continuous variables were used in three research papers and the other research papers were separate variables. As for the nature of the data, about 80% of the data were real and 20% were simulated to generate the data to be modeled, and the best accuracy has been reviewed in this The research papers were based on discrete variables, where the average accuracy was 95%, and nearly 5% of the predictions were inaccurate. It appears from the summary that the real data had a much greater accuracy than the data that was modeled through simulation, and the CT method used with continuous variables has less accuracy than the same method, but with discrete variables. In the first research paper of Table No. 1 the computer simulation shows the effect of the toxin on DNA and that the error is four times if the efficiency of the polymerase chain reaction varies over the range 0.04. As for the second research, there was an attempt to make the polymerase chain reaction analysis present in the world of statistical processing in terms of confidence interval and statistical significance. The results resulted in the development of practical statistical solutions using SAS programs. This result is close to what the researcher reached in the third and fourth paper in terms of epidemiological diseases or early detection of colon and rectal cancer before it spreads, through certain indicators.

Table 1: A Comparison of Some PCR Directions in Regression Analysis

Reference	Variable Type	Method	Regression Type	Dataset	disease	Results and Accuracy
[7]	Continues	Ct method and standard curve	linear regression	Simulated data vary between 1.81 and 2.02	poison	PCR efficiencies that vary over a 0.2 range second approach may be off by 0.26.

[11]	discrete	Cox proportional hazards	multivariate analysis	Real data for a sample of 125 patients	Colon and rectal cancer	(32.8%) early-stage patients were positive for disseminated tumor Cells. 95% confidence interval
[10]	discrete	Ct method and ANCOVA	multiple regression analysis	Real data	Arabidopsis thaliana	95% confidence levels
[13]	Continues	corrected and uncorrected Ct-based (threshold cycle) methods	Multiple regression and Non-linear regression	Real data	three different soils	80% confidence levels
[14]	discrete	Kaplan-Meier analyses	multivariate Cox regression analyses	Real data for 77 patient	Non-bacterial gastroenteritis	95% confidence interval
[28]	Continues	the linear mixed model	linear regression	Real data	clinical treatments	between 75% and 95% confidence interval

5. Conclusion

Regression analysis has been studied in several ways when using polymer chain reaction (PCR). Some of them were continuous data, while others were intermittent. The accuracy of the results varied from one method to another and from one research to another. It concludes from the above that the effectiveness of using regression in its various forms and types gave effective and encouraging results in the polymer chain reaction, and this encourages researchers to delve deeper in this direction, which is considered very important in the last three years and after the Corona pandemic, which was caused by Covid 19. Therefore, speech relied entirely on this test in detecting infection with this virus, which confirms the importance of this study, which linked this test to the method of statistical regression analysis.

References

[1] M. Thakur, "Regression Analysis Formula," 2021. [Online]. Available: https://www.wallstreetmojo.com/regression-analysis-formula/.

[2] A. Dempster, "An Overview of Multivariate Data Analysis," *JOURNAL OF MULTIVARIATE ANALYSIS 1,* pp. 316-346, 1971.

[3] G. Chamberlain, "MULTIVARIATE REGRESSION MODELS," *Journal of Econometrics,* pp. 5-46, 1982.

[4] K. Y. Liangt, B. Qaqish and S. L. Zeger, "Multivariate Regression Analyses for Categorical Data," *J. R. Statist. Soc. B,* pp. 3-40, 1992.

[5] L. Breiman and J. H. Friedman, "Predicting Multivariate Responses in Multiple Linear Regression," *J. R. Statist. Soc. B,* pp. 3-54, 1997.

[6] G. Death, "MULTIVARIATE REGRESSION TREES: A NEW TECHNIQUE FOR," *Ecology,,* pp. 1105-1117, 2002.

[7] C. Ramakers, J. . M. Ruijter, R. H. Lekanne Deprez and A. F. M. Moorman, "Assumption-free analysis of quantitative real-time polymerase chain," *Elsevier,* p. 62–66, 2003.

[8] M. H. Graham, "CONFRONTING MULTICOLLINEARITY IN ECOLOGICAL," *Ecology,* p. 2809–2815, 2003.

[9]   P. Dunlop and S. Smith, "ESTIMATING KEY CHARACTERISTICS OF THE," *Civil Engineering and Environmental Systems,* p. 273–290, 2003.

[10] J. . S. Yuan, A. Reed, F. Chen and C. . N. S. Jr, "Statistical analysis of real-time PCR data," *BMC Bioinformatics,* pp. 1-12, 2006.

[11] J. M. Lloyd, C. M. McIver, S. A. Stephenson, P. J. Hewett, N. Rieger and J. E. Hardingham, "Identification of Early-Stage Colorectal Cancer Patients at Risk of," *clincancerres.aacrjournals,* 2006.

[12] G. Y. Kanyongo, J. Certo and B. . I. Launcelot, "Using regression analysis to establish the," *International Education Journal,* pp. 632-641, 2006.

[13] R. H. Gulden, D. L. Booth, R. Campbell, J. R. Powell, M. M. Hart, K. E. Dunfield, J. T. Trevors, K. P. Pauls, J. . N. Klironomos and C. J. Swanton, "An empirical approach to target DNA quantification in environmental," *Elsevier,* p. 1956–1967, 2007.

[14] C. H. Gendo, G. Harste, B. J. Saathoff, F. Mattner, H. Deppe and A. Heim, "New Real-Time PCR Detects Prolonged Norovirus Excretion in Highly," *American Society for Microbiology,* p. 2855–2862, 2009.

[15] A. A. Pérez, . V. Sánchez, R. Baeza, M. C. Zamora and J. Chirife, "Literature Review on Linear Regression Equations for Relating Water Activity," *Food and Bioprocess Technology,* 2009.

[16] J. . V. Ortega, "Applied behavior analytic intervention for autism in early childhood: Meta-analysis,," *Elsevier,* p. 387–399, 2010.

[17] A. Kraha, . H. Turner, . K. Nimon, . L. . R. Zientek and R. . K. Henson, "Tools to support interpreting multiple regression in the face of multicollinearity," *Frontiers,* 2012.

[18] J. . S. Hygh, J. F. DeCarolis, D. . B. Hill and S. . R. Ranjithan, "Multivariate regression as an energy assessment tool in early building design," *Elsevier,* pp. 165-175, 2012.

[19] T. Zsuzsanna and . L. Marian, "Multiple regression analysis of performance indicators in the," *Elsevier,* p. 509 – 514, 2012.

[20] K. Imai, "Multivariate Regression Analysis for the Item," *Journal of the American Statistical Association,* 2012.

[21] X. Rao, D. Lai and X. Huang, "A New Method for Quantitative Real-Time," *JOURNAL OF COMPUTATIONAL BIOLOGY,* p. 703–711, 2013.

[22] T. Catalina, V. Iordache and . B. Caracaleanu, "Multiple regression model for fast prediction of the heating energy demand," *Elsevier,* pp. 302-312, 2013.

[23] S. Asadi, S. . S. Amiri and M. Mottahedi, "On the Development of Multi-Linear Regression Analysis to Assess Energy Consumption," *Energy and Buildings,* pp. 1-29, 2014.

[24] R. L. Prentice, B. J. Williams and A. V. Peterson, "On the regression analysis of multivariate failure time data," *Biometrik,* pp. 373-379, 2014.

[25] J. Tellinghuisen and A. N. Spiess, "Comparing real-time quantitative polymerase chain reaction analysis," *Elsevier,* p. 76–82, 2014.

[26] B. George, S. Seals and I. Aban, "Survival analysis and regression models," *American Society of Nuclear Cardiology,* 2014.

[27] J. L. ALeixandre-Tudo, I. ALvarez, M. J. Garcia, V. Lizama and J. L. ALeixandre, "Application of Multivariate Regression Methods to Predict," *Food Analysis, Food Quality and Nutrition,* p. 217–227, 2015.

[28] P. CHen and X. Huang, "Comparison of Analytic Methods for Quantitative," *JOURNAL OF COMPUTATIONAL BIOLOGY,* p. 988–996, 2015.

[29] S. Chen, J. Mao and X. Han, "Heat transfer analysis of a vertical ground heat exchanger using," *Energy and Buildings,* 2016.

[30] S. . A. Eslamian, S. S. Li and F. Haghighat, "A new multiple regression model for predictions of urban water use," *Elsevier,* p. 11, 2016.

[31] A. Gupta, A. Sharma and A. Goel, "Review of Regression Analysis Models," *International Journal of Engineering Research & Technology,* 2017.

[32] E. R. Edelman, S. M. . J. van Kuijk, A. E. . W. Hamaekers, M. J. . M. de Korte, G. . G. van Merode and W. F. . F. A. Buhre, "Improving the Prediction of Total Surgical Procedure Time Using Linear Regression Modeling," *Frontiers,* 2017.

[33] X. Li, P. Zhao and Y. Yang, "Application of Multivariate Regression Analysis in Teaching Management," *Atlantis Press,* 2018.

[34] I. Sharma and S. Kakchapati, "Linear Regression Model to Identify the Factors Associated with," *Hindawi,* p. 8, 2018.

[35] M. Yang, "A review of regression analysis methods: Establishing the," in *PROCEEDINGS of the 23rd International Congress on Acoustics*, Aachen, Germany, 2019.

[36] A. Pant and R. S. Rajput, "Linear Regression Analysis Using R for Research and Development," *Researchgate,* pp. 180-195, 2019.

[37] A. . V. R. Reddy and M. S. Kumar, "A Comparative Analysis of Regression," in *International Conference on Intelligent Computing and Communication Technologies*, 2020.

[38] Y. Wang and S. . D. Zhao, "A nonparametric empirical Bayes approach to large-scale," *Elsevier,* 2020.

[39] V. Gallego, A. S´anchez, I. Mart´on and S. Martorell, "Analysis of occupational accidents in Spain using shrinkage," *Elsevier,* 2020.

[40] J. . G. Outram, S. J. Couperthwaite, W. Martens and G. J. Millar, "Application of non-linear regression analysis and statistical testing to," *Elsevier,* 2020.

[41] V. S. Velikanov, N. . V. Dyorina, E. I. Rabina and T. Y. Zalavina, "Practical application of regression analysis to study load," *Journal of Physics: Conference Series,* 2020.

[42] M. Mohammedi, S. Bouzebda and A. Laksaci, "The consistency and asymptotic normality of the kernel type," *Elsevier,* 2020.

[43] D. . H. Maulud and A. M. Abdulazeez, "A Review on Linear Regression Comprehensive in Machine Learning," *Journal of Applied Science and Technology Trends,* p. 140 –147, 2020.

[44] P. Korwar, "Introduction to Multivariate Regression Analysis," Aug 2020. [Online].

[45] K. Kumari and S. Yadav, "Linear Regression Analysis Study," *Journal of the Practice of Cardiovascular Sciences,* 2018.

[46] . W. Kenton, "investopedia.com/terms/n/nonlinear-regression.asp," 30 April 2021. [Online]. Available: https://www.investopedia.com/terms/n/nonlinear-regression.asp.

[47] S. Glen, 10 December 2014. [Online]. Available: https://www.statisticshowto.com/univariate/.

[48] A. Brembilla, . A. Olland, . M. Puyraveau, . G. Massard, . F. Mauny and P. E. Falcoz, "Use of the Cox regression analysis in thoracic surgical research," *Journal of Thoracic Disease,* p. 3891–3896, 2018.

[49] M. Rahman, M. Uddin, R. Sultana, A. Moue and M. Setu, "Polymerase Chain Reaction (PCR): A Short Review," *researchgate,* pp. 30-36, 2013.

[50] J. Singh, N. Birbian, S. Sinha and A. Goswami, "A critical review on PCR, its types and applications," *International Journal of Advanced Research in Biological Sciences,* p. 65–80, 2014.

[51] A. V. Reddy and M. S. Kumar, "Please include any non-standard abbreviations referred to wi," *System REALIability,* pp. 300-310, 2020.