STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **16**, September 2026, pp 2001–2016.
Published online in International Academic Press (www.IAPress.org)

Determinants of Electronic Examination Success: An Application of Robust Structural Equation Modeling

Salwa Salah aldeen Qasim [1,*], Mohammed Ameen Oqbah Mohammed [2], Husam Waleed Yaseen [3]

[1] *College of Physical Education and Sports Sciences, University of Mosul, Iraq*
[2] *College of Education for Pure Sciences, University of AL-Hamdaniya, Iraq*
[3] *College of Computer Science and Mathematics, University of Mosul, Iraq*

Abstract In this paper, robust procedures for estimating path coefficients in the modeling framework of structural equations are proposed based on robust correlation coefficients. These are two estimates: robust correlation coefficients and Pearson correlation coefficients. The traditional approach performs better than the robust approach using real data of the researcher's collected through the questionnaire. To find out factors contributing to success in electronic tests, students were sampled using the questionnaire from postgraduate applicants at the University of Mosul. Following the collection of data, 205 valid responses were obtained that contained six independent variables that represented factors which affect electronic tests. The proposed biweight midcorrelation method produced more stable path estimates and improved model fit in the presence of outliers. The results indicate that self-efficacy, quality of use, and user satisfaction are the most influential factors affecting the success of electronic examinations.

Keywords Total Variation, Image Reconstruction, Zooming, Compressive Imaging, Fast Fourier Transform, Convolution, Alternating Direction Method

DOI: 10.19139/soic-2310-5070-3629

1. Introduction

SEM is an analytical technique that is relatively recently applied across the disciplines of science. It offers significant advantages and great flexibility compared to other statistical techniques. The method of SEM is based on establishing of a graphical model that allows an accurate and systematic evaluation of the complex theoretical structure as well as the hypothesized relationships between variables. It also allows for simultaneous analysis of interrelationships among multiple variables rather than analyzing isolated relationships alone. It is also noteworthy that structural equation modeling has gained traction due to the strong theoretical foundation and the ability to combine measurements and structural relationships in a single analytical framework. This interaction allows for better interpretation of the findings through the use of latent variables and complex causal mechanisms. SEM therefore is of tremendous importance to empirical research and allows for more confident and more robust conclusions.The motivation for this study stems from the increasing use of electronic examination systems, which often generate questionnaire data containing outliers that may affect the accuracy of conventional statistical analyses. The comparison of Pearson correlation to biweight midcorrelation in estimating path coefficients in a path-analysis context is the major contribution of this study, particularly in addressing the issue of robust correlation techniques for enhancing the reliability of parameter estimates when influential observations are present.The fact that in higher education, there are more and more electronic examination systems, where the information obtained

*Correspondence to: Salwa Salah aldeen Qasim (Email: salwa.kasim@uomosul.edu.iq). College of Physical Education and Sports Sciences, University of Mosul, Iraq.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

from questionnaires often contains outliers which could negatively impact on the accuracy of traditional statistical analysis methods, is the motivation for this research. Although the path analysis technique is widely used in social sciences and has been applied to educational studies, little research has focused on assessing the influence of powerful correlation methods on path coefficient estimates for educational research. Accordingly, the determinants of electronic examination success have been investigated by using the conventional Pearson approach and the biweight midcorrelation approach in a path analysis framework. The most significant finding of this study is that strong correlation estimation yields more accurate path coefficients and higher model fit in the face of influential cases, and thus more reliable grounds for decision-making in educational testing.

2. Structural Equation Modeling

Structural Equation Modeling describes the comprehensive statistical estimate, analysis and testing approach of describing and explaining relationships between observed variables and latent variables. It allows for multiple relationships estimation and characterization within a single analysis framework. SEM therefore is a large-scale statistical testing strategy for hypotheses about the relationship between independent variables and observed and latent dependent variables. It provides an explicitly defined means for describing, estimating and verifying linear networks relationships in direct and indirect effects of variables [1]. A structural equation model is a hypothetical or conceptual representation of direct and indirect linear relations between a set of latent variables. It can also be considered as a unified path model representing the relationships of variables in a path diagram illustrating the structure and nature of relationships. In this sense SEM is an extension of the standard linear model that contains several equations and latent constructs in one model [2]. Multiple regression, path analysis, and CFA are also part of the SEM, with many of these being SEM specific. This versatility also leads to more complex research questions, e.g., measurement error, latent constructs, and hierarchical relations, which contribute to robustness and explainability of empirical results, in SEM.

2.1. Steps for Modeling Using Structural Equation Modeling

SEM is a multi-step process that will enable you to better understand the proposed research model and for accurate modeling. Each step is based on the first step and maintains the logical and methodological consistency by analysis. The basic components of SEM that are required to design and test structural models have been described by [3]. These involve data characteristics analysis, reliability and validity analysis, general fit analysis, model respecification, model reanalysis, similar model analysis, and report of results. First of all, the testing takes into account the sample size, normality and outliers that directly affect the use of SEM technique. Next, the model is tested for reliability and validity in order to verify that the observed measures represent accurate latent constructs. The model fit indices are then used to determine the applicability of the hypothesized model to the observed data. If the model is not strong enough, it is possible to resituate the model with theoretical justification. Finally, other comparable models are considered to assure reliability of the results and regularly published findings to aid comprehension and empirical reproducibility

2.2. Variables in Structural Equation Modeling

SEM addresses the causal relationship among variables that can be identified using an equation system of structural equations that are analogous to multiregression equations. These structural relationships are defined and modelled for a better, more complete understanding of the phenomenon being investigated. The hypothesized model can be tested statistically by simultaneously analyzing the entire system of variables to determine the fit with the observed data [4]. In the SEM framework, these variables are typically categorized into observed variables and latent variables, but they are all significant elements of the modeling process. The observed variables are measured directly; latent variables are subsets of constructs that can't be measured directly, but can be inferred from a variety of indicators. This difference allows SEM to capture measurement error and deliver more accurate estimates of the relationship between constructs. So, inclusion and careful formulation of variables will enable the structural model to be more predictable and stable, thereby allowing for accurate causal inferences

2.3. Types of Variables in Structural Equation Modeling

Variables in SEM are categorized into different types based on their role and role within the model. This division helps us understand the construction of this model and explain the relationships among variables.

First: Latent versus Observed Variables

Researchers from other disciplines are often interested in finding ideas that are not immediately apparent or quantifiable. These types of constructs are called latent variables, also known as theoretical variables or structural variables. Latent variables are not directly measured, but are inferred indirectly from a set of observable indicators developed with measurement equipment or other data collection tools. Rather, observed variables are directly quantifiable and act as a measure of latent constructs. In this framework, variables can also function as independent, dependent, or mediating variables depending on the position and role in the structural model.

Second: Exogenous versus Endogenous Latent Variables

In SEM latent variables are also classified into exogenous and endogenous. These exogenous latent variables are conceptually equivalent to independent variables because they are responsible for other variables in the model. Variations in exogenous variables are not at all explained in the model, but are assumed to be due to other factors that are not considered in the analysis. In contrast, endogenous latent variables are those that are explained by relationships within the model as they directly or indirectly affect exogenous variables [4, 17].

It also improves SEM applications by providing a more comprehensive structure of variables with both direct and indirect effects of observed and latent constructs, and allows researchers to model complex causal chains.

Recent studies have emphasized the importance of robust statistical methods in educational and behavioral research, particularly when questionnaire data contain outliers or violate normality assumptions. These studies have shown that robust correlation techniques improve parameter stability and provide more reliable estimates than conventional correlation methods, thereby enhancing the validity of path analysis and structural modeling results. Recent studies have emphasized that robust estimation techniques improve the stability and reliability of structural equation modeling, particularly when the data contain outliers or violate normality assumptions [13, 24, 25].

3. Outliers

Outliers are values in a set of data that are very different from the rest of the data, and are not what you would expect. These values can be out of the ordinary or even conflicting with the rest of the sample and may be caused by measurement error, data entry error or true but extreme variation of the phenomenon being measured. However, most observations do not include such observations and, as noted by [5, 27], can lead to distortions in statistical statements and analyses if not dealt with correctly. Initial examination of the data set is important to any empirical study, since it will provide a basis for identifying values that deviate from the rest of the observations. The goal of this step is to identify possible outliers and avoid performing formal statistical analyses on them. Data screening and cleaning processes are designed to exclude or appropriately deal with abnormal data points, extreme data points or stray values, which can be significant deviations from the data structure. The presence of outliers is known to have a significant effect on parameter estimation and hypothesis testing, so detecting and dealing with outliers is crucial to the validity and reliability of the statistical results, as noted in [6]. Furthermore, in higher-order statistical analysis methods like Structural Equation Modeling, outliers can significantly affect the degrees of the correlations, model-fitting indices, and parameter estimates. Therefore, employing robust statistical methods or appropriate diagnostic tools to manage outliers enhances the accuracy and robustness of analytical outcomes and supports more credible research findings.

3.1. Detection of Outliers

Many researchers have investigated outliers and how they may affect the validity of empirical research. Outlier detection is commonly performed in regressions, experimental designs, and multivariate datasets, primarily by the assessment of observations as partial or complete. There are several approaches to finding outliers that are generally classified into two subclasses: univariate and multivariate. Univariate methods assess each variable individually for extreme values and multivariate methods investigate the correlations among variables in a dataset for information about observations that are out of step with the multivariate pattern [7]. Outlines the various methods used to detect outliers including, standard deviation, z-scores, modified z-scores, Tukey's method, boxplot, adjusted boxplot, binary boxplot, and interquartile range-based methods. Each of these methods has advantages as determined by the nature of the data and the analysis's objectives [8]. On top of that, outliers are useful to ensure the reliability of statistical analysis as well as in identifying and resolving them. Outliers may influence parameter estimates, correlations, and model fit in multivariate analyses, such as SEM. Thus, I am giving in to that fact. combining multiple detection methods and applying robust statistical techniques can improve the reliability and robustness of research findings, particularly in complex datasets where extreme values may obscure underlying relationships, as seen in the following figure.

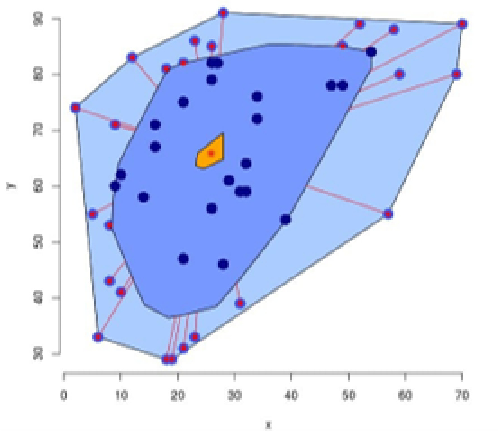

Figure 1. Shows the binary box diagram.

4. Robust Correlation

This is one of the most basic statistics whose coefficient of correlation indicates the degree and direction of the relationship between two random variables, ranging in strength from + to -1. The most widely used test is the Pearson correlation coefficient, which is highly sensitive to anomalous values or outliers that are significantly different from the larger distribution of data. Such outliers can compromise the interpretation of the correlation and lead to false assumptions. There has been a need to enhance robust correlation techniques that are less sensitive to anomalous data as due to the limitations of traditional correlation techniques. This concept of robustness was first discussed by Box [9] who argued that statistical methods are robust if the performance is stable even with a large deviation or unusual observation in the data. Next research further developed this idea, with examples such as [10] demonstrating robust approaches in the presence of observations that can affect correlate strongly estimates through simulation experiments. Moreover, the more powerful correlation techniques are more accurate and stable estimates in the presence of outliers and/or non-normal distributions. Not to mention, these techniques help minimize the effect of outlier observations, and thereby enhance the accuracy of future models, like regressions or SEMs, where having a precise estimation of the correlation is essential to make valid inferences about the parameters and hypothesis. Thus, strong correlation is an integral part of improving the reliability and interpretation of statistics in empirical research.

5. Methods for Estimating Path Coefficients Using the Correlation Matrix

Path coefficients can be estimated in many ways depending on the analytical framework and the nature of the data. Path coefficient estimates are considered in this paper in the following two main ways:

5.1. Estimation of Path Coefficients Using the Simple Correlation Coefficient

In most cases, the simple correlation coefficient indicates that two random variables are linearly related. This coefficient indicates the direction and strength of the relationship between variables and is determined using the standard formula [11]. This measure is used in classical path analysis, which allows you to make a start on estimating the direct effects of variables in a causal model.
Moreover, the easy-correlation method is simple and widely used, but it is vulnerable to data outliers and deviations in data distribution from normality. Therefore, this approach can give inaccurate or unstable estimates of path coefficients in data sets with abnormal or extreme values. In order to overcome such drawbacks, strong correlations can be used to more precisely and robustly estimate path coefficients, which represent the true relationships among variables.
Such powerful techniques are particularly valuable when modelling the direct and indirect effects that are needed to validate and interpret the model in structural equation studies.

$$r_{x,y} = \frac{Cov(x,y)}{\sigma_x \sigma_y} = \frac{E(x-\mu_x)(y-\mu_y)}{\sigma_x \sigma_y} \tag{1}$$

$Cov(x,y)$: The covariance between x, y and x, y is represented by the standard deviation of the y and x variables respectively.
The simple correlation coefficient, also known as the Pearson correlation coefficient, is sensitive to anomalous or extreme values in the dataset. Such outliers can be highly confounding on the intensity and direction of the correlation leading to misleading conclusions about the relationships between variables.
In order to overcome these limitations, multiple more robust approaches have been proposed. These methods aim to provide more precise and reliable estimates of correlation that are not too sensitive to extreme values or deviations from normality. Strong correlation provides stability and guarantees validity in subsequent procedures such as regressions, path analyses, and Structural Equation Modeling for which precise estimation of relationships is necessary in order to make valid judgments.

5.2. Correlation: Biweight Miscorrelation (Robust)

The robust correlation measure biweight midcorrelation is constructed from two vectors of data. This method weights the observations in each vector to minimize the effects of extreme or outlying values. By doing so, it provides a more reliable estimate of the correlation between these two variables, particularly in data with anomalies or non-normal distributions.
follows $(x_i, y_i \quad , i = 1, 2, 3, \ldots, m$ [12]:

$$x = \begin{bmatrix} x_1 & x_2 \ldots & x_m \end{bmatrix} y = \begin{bmatrix} y_1 & y_2 \ldots & y_m \end{bmatrix}$$

The components of these vectors are defined as follows:

$$u_i = \frac{x_i - Med(x)}{9Mad(x)} \tag{2}$$

$$v_i = \frac{y_i - Med(y)}{9Mad(y)} \tag{3}$$

$Med(x)$: represents the magnitude of the mean vector x
$Med(y)$: represents the magnitude of the mean vector y
$Mad(x)$: represents the average absolute deviation of the x
$Mad(y)$vector: represents the average absolute deviation of the y-vector.
The defined weights will be used as follows:

$$w_i^{(x)} = \left(1 - u_i^2\right)^2 I\left(1 - |u_i|\right) \tag{4}$$

$$w_i^{(y)} = \left(1 - v_i^2\right)^2 I\left(1 - |v_i|\right) \tag{5}$$

Since:
I : represents a single function defined as:

$$I\left(x\right) = \begin{cases} 1 & , if \quad x > 0 \\ 0 & , \quad otherwise \end{cases}$$

That is, I1-ui equals 1 if $(1 - u_i) > 0$, which is zero otherwise, and the $wi(x)$ weights are about 1 if the x vector is equal to the x vector's mean value, and it is about zero if the x vector's magnitude of the x vector's media is approaching 9 $Mad(x)$.
Thus, the weighted binomial for both x and y vectors (x, y), can be defined as follows:

$$\overline{x}_i = \frac{\left(x_i - Med(x)\right) w_i^{(x)}}{\sqrt{\sum_{j=1}^{m}\left[\left(x_j - Med(x)\right) w_j^{(x)}\right]^2}} \tag{6}$$

$$\overline{y}_i = \frac{\left(y_i - Med(y)\right) w_i^{(y)}}{\sqrt{\sum_{j=1}^{m}\left[\left(y_j - Med(y)\right) w_j^{(y)}\right]^2}} \tag{7}$$

$$bicor\left(x, y\right) = \sum_{i=1}^{m} \overline{x}_i \overline{y}_i \tag{8}$$

It can be formulated as follows:

$$bicor\left(x, y\right) = \frac{\sum_{i=1}^{m}\left(x_i - Med\left(x\right)\right) w_j^{(x)} \left(y_i - Med\left(y\right)\right) w_j^{(y)}}{\sqrt{\sum_{j=1}^{m}\left[\left(x_j - Med\left(x\right)\right) w_j^{(x)}\right]^2 \sum_{j=1}^{m}\left[\left(y_j - Med\left(y\right)\right) w_j^{(y)}\right]^2}} \tag{9}$$

The biweight midcorrelation was chosen because it is robust and has the ability to minimize the impact of outliers on the correlation structure. This property renders it especially appropriate to use for questionnaire data where the extreme answers might introduce an error into the estimation of conventional Pearson correlation. It should be noted that the proposed approach does not constitute a full likelihood-based robust Structural Equation Modeling estimator. Instead, this study adopts a path-analytic framework in which path coefficients are derived from correlation matrices and compared using both Pearson and biweight miscorrelation coefficients. Therefore, the methodological contribution of this study lies in evaluating the influence of robust correlation estimation on path coefficient estimation rather than proposing a new robust SEM estimation procedure [16].

6. Criteria for Choosing the Best Model

There is a set of statistical criteria that will be covered in this study for choosing the best models:

6.1. Akaike's Information Criteria

This can be expressed mathematically as follows [13]:

$$AIC = 2k - 2\ln L \tag{10}$$

Since:- AIC: Stands for the Akakie criterion of information, L: stands for the maximum possible function.

6.2. Bayesian Information Criteria

The formula can be explained as follows [14]:

$$BIC\,(k) = n\log\sigma_a^2 \; - (n-k)\ln\left(1-\frac{k}{n}\right) \; + k\ln n + k\ln\left[\left(\frac{\sigma_Y^2}{\sigma_a^2}-1\right)/k\right] \tag{11}$$

Since:- σ_Y^2: represents the amount of variance, as the model that corresponds to the lowest value for this criterion is chosen.

6.3. Chi Square Standard

$\left(\chi^2\right)$ The suitability of a model for the observed data is often evaluated using the chi-square (χ) criterion. A model is considered less appropriate when the ratio of the calculated chi-square value to the tabular chi-square value is high, indicating a poor fit between the hypothesized model and the data. Conversely, a smaller ratio suggests that the model adequately represents the underlying data structure [15, 18, 19].
a) Root Mean Square Error of Approximation (RMSEA): If the RMSEA value is less than or equal to 0.05, the model is considered to fit the data very well. Values between 0.05 and 0.08 indicate an acceptable or reasonably good fit, while values exceeding 0.08 suggest a poor fit, leading to the rejection of the model [20, 21, 26].
b) Root Mean Square Residual (RMSR): An RMSR value of between 0 and 1 means that it is a better fit. In particular, the best model corresponds to the lowest RMSR (which should be kept between 0 and 0.1). Furthermore, the use of several model fit measures, instead of only one, contributes to the robustness of the model evaluation. A chi-square, RMSEA, RMSR, and other fit indices (Comparative Fit Index/CFI and Tucker-Lewis Index/TLI) give a thorough idea of model's ability to explain the observed relationships between the variables. This multidimensional assessment is very important to draw valid and reliable conclusions from structural equation modeling [22, 23].

6.4. Incremental Fit Criterion (IFC)

a) Normal Fit Criterion (NFC) The value of this standard is between (0-1) and a high value indicates the model's conformity to the data.
b) Non-Normal Fit Criterion (NNFC) value is between (0-1) and the model is identical to the data when its value is greater than or equal to 0.95.

6.5. Root Mean Square Error (RMSE)

This represents the square root of the average of the error squares, because the root is measured in the original units of the same values for the variables and calculated by the formula:

$$RMSE = \sqrt{MSE} = \sqrt{\frac{\sum_{i=1}^{n} e_i^2}{n}} \tag{12}$$

6.6. Application Aspect

This research was applied to real data collected by the researchers through a questionnaire that was distributed to a sample of students applying for postgraduate studies. The questionnaire deals with the factors affecting the success of electronic tests at the University of Mosul. After that, the data gathered were processed. The number of data reached 205 observations, The sample size of 205 observations is considered adequate for the proposed structural model. This satisfies the commonly recommended rule of at least 10 observations per estimated parameter in Path Analysis within the SEM Framework, ensuring sufficient statistical power and stable parameter estimation, The sample comprised 205 postgraduate applicants, which is considered sufficient for the proposed analytical framework. According to the widely accepted guideline in path analysis and Structural Equation Modeling, a minimum of ten observations per estimated parameter is generally adequate to obtain stable parameter estimates and sufficient statistical power. Therefore, the sample size used in this study is appropriate for the complexity of the proposed model and supports the reliability of the estimated path coefficients, and it includes 6 independent factors of the variables that represent the factors affecting these tests. These run as follows: X_1: is Self-Efficacy, X_2: Represents The System Quality, X_3 : Represents the Information, X_4: Represents How to Use X_5: Represents User Satisfaction, Y : Represents net benefits or what the result is. The proposed structural model was developed based on previous studies on information systems success and electronic learning environments. Self-efficacy, system quality, information quality, quality of use, and user satisfaction have consistently been identified as key determinants influencing users' acceptance and successful utilization of electronic systems. Therefore, these constructs were adopted to examine their direct and indirect effects on the success of electronic examinations. The relationship between these variables can be illustrated by the following figure:

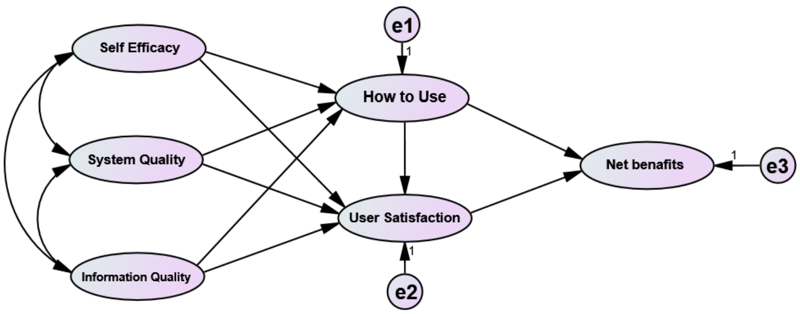

Figure 2. Shows a model for measuring the success of the electronic test system.

Also, these variables are divided by the figure above into separate variables, or what are called external variables. These are $X_1,\ X_2,\ X_3$ and X_4, X_5 are intermediate variables for the external variables and at the same time external variables for the *y* variable, the internal variable is represented by the y variable, after which the abnormal data (outliers) are revealed using the boxplot and as shown in the following figure:

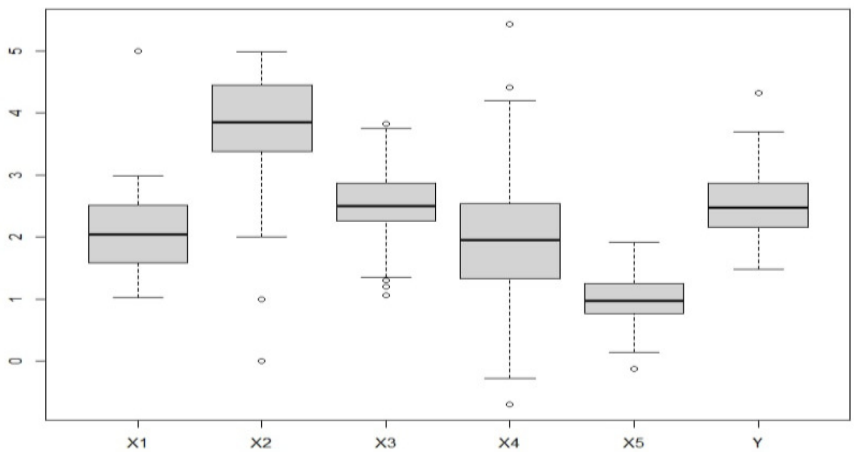

Figure 3. Boxplot to detect outliers in the internal, external, and intermediate variables.

Outlier detection using the bivariate boxplot revealed the presence of a limited number of influential observations. Visual inspection indicated that these observations were genuine extreme responses rather than data entry errors. Therefore, they were retained in the dataset, and robust correlation estimation was employed to minimize their influence instead of excluding them from the analysis. We notice from the above figure that there are outliers between the external variables and the internal and median variables, and because the correlation relationship is a binary relationship between two variables, the outliers will be found through the bivariate box graph. A bivariate boxplot that depends on the ranks as the box moves from observations with rank 3n/4 to observations with rank 3n/4, where the central box of the box is drawn in the middle. In general, the concept of a bi-variate boxplot expresses what the scientist Tukey (1975) described as Biplot and as shown in the following figure:

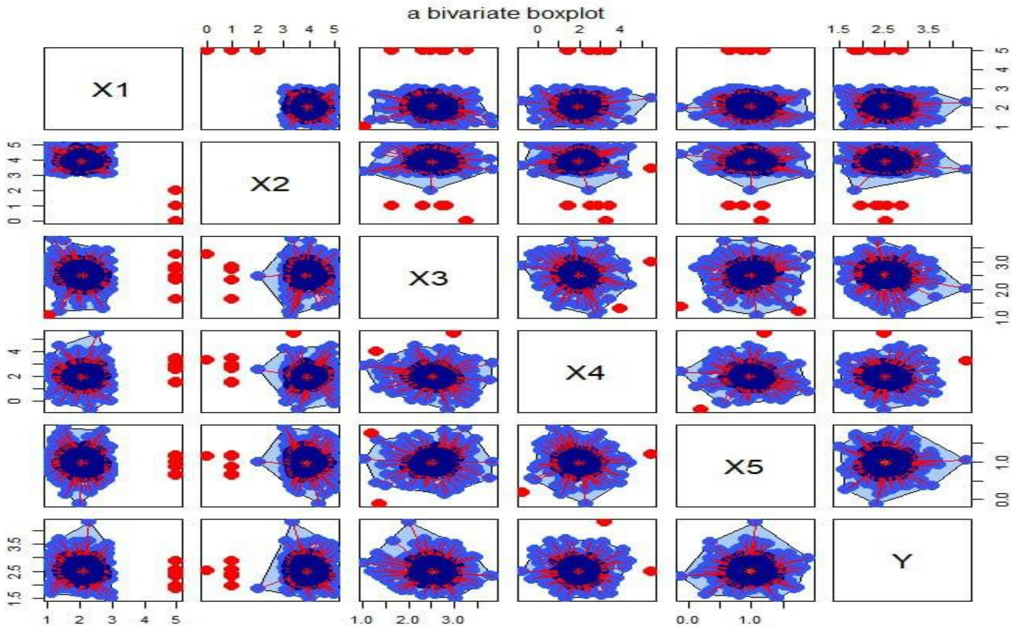

Figure 4. Boxplot for Bivariate data.

The bivariate boxplot revealed the presence of influential observations in the dataset. A visual inspection confirmed that these observations represented genuine extreme responses rather than data entry errors. Consequently, they were retained in the analysis, and the biweight midcorrelation was employed to reduce their influence instead of removing them, thereby preserving the integrity of the original data. We also note from Figure (4) the presence of clear outliers between the study variables, and to find out the relationship between the variables, the correlation matrix between the external, internal and intermediate variables was found using the program R. The results are shown in the following table:

Table 1. Shows the Correlation Coefficients between the study variables

Variables	X_1	X_2	X_3	X_4	X_5	Y
X_1	1	0.36	0.18	0.51	0.38	*0.62*
X_2	0.36	1	0.52	0.42	0.52	0.33
X_3	0.18	0.52	1	0.44	0.36	0.44
X_4	0.52	0.42	0.44	1	0.51	0.42
X_5	0.38	0.52	0.36	0.51	1	0.43
Y	0.62	0.33	0.44	0.42	0.43	1

We note from the above table that there are significant differences at the level 0.05 and at the level of 0.01. This is evidenced by the presence of differences between the values of the correlation coefficients resulting from other side effects. These can be direct and indirect effects between the variables. Therefore, we can also note that the analysis of the correlation between the variables is insufficient to study the effects of direct and indirect variables. It is also insufficient for analyzing the factors affecting the success of electronic tests, as the relationship between

two variables does not mean that the first variable has a direct effect on the second variable. It also means that the existence of the relationship or not between the two variables may be due to other factors and side effects.
Because of the importance of studying the relationship between independent variables and studying their impact on the dependent variable, it is necessary to choose an appropriate statistical method that takes into account these relationships. The independent variables and their effect on the dependent variable, which are the intermediate variables.
In this research, we will rely on the path analysis method that depends on simple correlation analysis and the robust correlation in calculating the path coefficients to study the factors affecting the success of electronic tests at the University of Mosul. We first start using the traditional correlation matrix (Pearson correlation) to find the path coefficients and the results can be presented as follows:

$$\mathrm{P}_{12} = r_{12} = 0.36, \mathrm{P}_{13} = r_{13} = 0.18, \mathrm{P}_{23} = r_{23} = 0.52$$

From the correlation matrix in Table Eq. (1) and by applying Eq. (9), we notice that the correlation coefficients between the external variables are equal to the path coefficients (Direct Effect) between them, which are as follows:
As for the rest of the path coefficients, they are calculated using the simple correlation coefficients calculated in Table (1). If the X_4 variable represents the internal variable, the results are as follows:

$$r_{14} = P_{41}r_{11} + P_{42}r_{12} + P_{43}r_{13}$$

$$r_{24} = P_{41}r_{21} + P_{42}r_{22} + P_{43}r_{23}$$

$$r_{34} = P_{41}r_{31} + P_{42}r_{32} + P_{43}r_{33}$$

According to the internal variable, the path coefficient (Direct Effect) is calculated using matrices as follows:

$$\begin{bmatrix} P_{41} \\ P_{42} \\ P_{43} \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{23} & r_{33} \end{bmatrix}^{-1} \begin{bmatrix} r_{14} \\ r_{24} \\ r_{34} \end{bmatrix} = \begin{bmatrix} 0.414 \\ 0.110 \\ 0.305 \end{bmatrix}$$

But if X_5 represents the internal variable, the correlation coefficients are calculated as follows:

$$r_{15} = P_{51}r_{11} + P_{52}r_{12} + P_{53}r_{13} + P_{54}r_{14}$$

$$r_{25} = P_{51}r_{21} + P_{52}r_{22} + P_{53}r_{23} + P_{54}r_{24}$$

$$r_{35} = P_{51}r_{31} + P_{52}r_{32} + P_{53}r_{33} + P_{54}r_{34}$$

$$r_{45} = P_{51}r_{41} + P_{52}r_{42} + P_{53}r_{43} + P_{54}r_{44}$$

According to the internal variable, the path coefficient (Direct Effect) is calculated using matrices as follows:

$$\begin{bmatrix} P_{51} \\ P_{52} \\ P_{53} \\ P_{54} \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} \\ r_{21} & r_{22} & r_{23} & r_{24} \\ r_{31} & r_{32} & r_{33} & r_{34} \\ r_{41} & r_{42} & r_{43} & r_{44} \end{bmatrix}^{-1} \begin{bmatrix} r_{15} \\ r_{25} \\ r_{35} \\ r_{45} \end{bmatrix} = \begin{bmatrix} 0.102 \\ 0.334 \\ 0.034 \\ 0.307 \end{bmatrix}$$

Whereas if the internal variable is:Y

$$r_{Y4} = P_{Y4}Y_{44} + P_{Y5}Y_{45}, r_{Y5} = P_{Y4}Y_{54} + P_{Y5}Y_{55}$$

Using arrays, the paths can be obtained as follows:

$$\begin{bmatrix} P_{Y4} \\ P_{Y5} \end{bmatrix} = \begin{bmatrix} r_{44} & r_{45} \\ r_{54} & r_{55} \end{bmatrix}^{-1} \begin{bmatrix} r_{Y4} \\ r_{Y5} \end{bmatrix} = \begin{bmatrix} 0.269 \\ 0.293 \end{bmatrix}$$

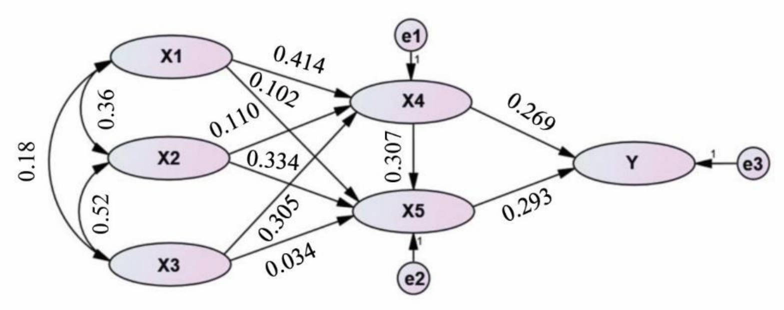

Figure 5. The path coefficients computed by simple correlation analysis.

Table 2. The Robust Correlation Coefficients between the study variables

Variables	X_1	X_2	X_3	X_4	X_5	Y
X_1	1	0.46	0.33	0.53	0.36	*0.76*
X_2	0.46	1	0.46	-0.15	-0.19	0.52
X_3	0.33	0.46	1	0.25	0.16	0.36
X_4	0.53	-0.15	0.25	1	0.76	0.35
X_5	0.36	-0.19	0.16	0.76	1	0.57
Y	0.76	0.52	0.36	0.35	0.57	1

The Simple Correlations and paths between the variables can be illustrated by the following figure:

In the same way, the following robust correlation matrix is used:

This negative coefficient for System Quality in the robust model is a result of the outliers being eliminated from the analysis as well as controlling for the other variables, meaning it is not possible to directly conclude that ease of use will increase if system quality is improved. This discovery suggests that user training and simplicity of the system are also crucial for successful electronic examination systems. Note that some path coefficients from the fortified correlation matrix (Table (2)) are not only different in value but even in direction from those estimated from the Pearson correlation matrix. One of the most interesting cases is that of System Quality and How to Use, which is positively associated in the classical analysis but negatively directly associated in the robust model. This change is not a contradiction, it means that the statistical properties of the robust estimation are different. Unlike Pearson correlation, robust correlation minimizes the influence of outlying observations that may distort conventional estimates. Moreover, path coefficients in Structural Equation Modeling represent conditional direct effects after simultaneously controlling for the remaining predictors. Consequently, when the influence of extreme observations is reduced and shared variance among the exogenous variables is properly accounted for, the net direct contribution of **System Quality** may change direction because of suppression effects and multicollinearity among the predictors. Therefore, the robust model is considered to provide a more realistic representation of the underlying structural relationships, particularly in datasets containing influential observations, and its estimates should be interpreted as the net effects after controlling for the remaining variables rather than as simple bivariate associations, and by using Eq. (9) to get from it the fortified path coefficients as follows:

$$\mathrm{RP}_{12}=\mathrm{r}_{12}=0.454, \mathrm{RP}_{13}=\mathrm{r}_{13}=0.327, R\mathrm{P}_{23}=\mathrm{r}_{23}=0.454$$

The remaining path coefficients are calculated using the robust correlation coefficients in Table (2), if X_4 represents the internal variable and calculates the path coefficient (Direct Effect) using matrices, the results are as follows*:*

$$r_{14}=P_{41}r_{11}+P_{42}r_{12}+P_{43}r_{13}$$

$$r_{24}=P_{41}r_{21}+P_{42}r_{22}+P_{43}r_{23}$$

$$r_{34}=P_{41}r_{31}+P_{42}r_{32}+P_{43}r_{33}$$

$$\begin{bmatrix} RP_{41} \\ RP_{42} \\ RP_{43} \end{bmatrix} = \begin{bmatrix} 0.716 \\ -0.617 \\ 0.299 \end{bmatrix}$$

But if X_5 represents the internal variable, the path coefficient (Direct Effect) is calculated as follows:

$$\begin{bmatrix} RP_{51} \\ RP_{52} \\ RP_{53} \\ RP_{54} \end{bmatrix} = \begin{bmatrix} -0.014 \\ -0.067 \\ 0.006 \\ 0.753 \end{bmatrix}$$

If Y is the internal variable, the results are:

$$\begin{bmatrix} RP_{Y4} \\ RP_{Y5} \end{bmatrix} = \begin{bmatrix} -0.184 \\ 0.708 \end{bmatrix}$$

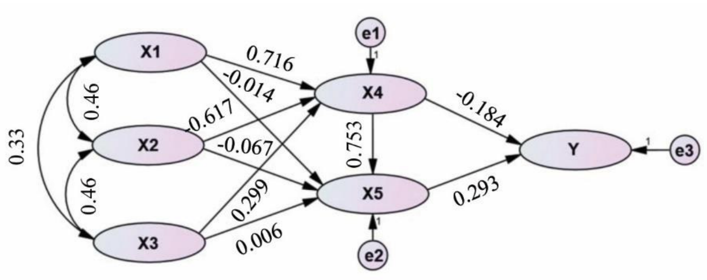

Figure 6. The path Coefficients Computed by Robust Correlation Analysis.

Then, using the simple correlation and the robust correlation, the indirect effect of the external variables on the internal variable was estimated. The paths represented by dynamic paths and the value of the complex path were calculated from the outcome of all the direct path coefficients that make up the complex path, which is divided into Three types of indirect path (I), unanalyzed path U and imaginary path S.

First, the indirect effects of the external variables on the internal variable that represents the user's satisfaction were calculated. In this regard, we can note that there are no (Indirect Effects) of the independent variables on the internal variable using the simple correlation matrix and the robust correlation as follows: X_1, X_2, X_3, X_4 to the input variable X_5:

Table 3. Indirect Effect on the variable X_5

Path	I Using the Simple Correlation Matrix	I Using the Robust Correlation Matrix
Effect through X_1X_4	$P_{14} * P_{45}= 0.127$	$RP_{14} * RP_{45} = 0.539$
Effect through X_2X_4	$P_{24} * P_{45} = 0.033$	$RP_{24} * RP_{45}= 0.464$
Effect through X_3X_4	$P_{34} * P_{45} = 0.093$	$RP_{34} * RP_{45} = 0.225$
Effect through X_4X_4	$I = 0$	$I = 0$

We can note from the table that there are no indirect effects of the independent variables on the internal variable. Therefore, through the simple correlation matrix and the Robust matrix $.X_1, X_2, X_3, X_4X_5$ and $I = 0$.

As for the indirect effects of the external variables on the internal variable that represents the result of the applicant in the test, as there are Indirect Effects for each of the independent variables Y,X_1, X_2, X_3, X_4 in the variable Ythrough the variable X_5. There is also an Indirect Effect to X_4 that represents the advanced use method on the test by X_5 representing the user's satisfaction as shown in the following table:

The Direct and Indirect Effects on X_4 are as follows:

Table 4. Indirect effects in the variable Y using the Simple Correlation Matrix and the Robust Matrix

	Using the Simple Correlation Matrix			Using the Robust Correlation Matrix		
Variables	through X_4	through X_5	through X_4, X_5	through X_4	through X_5	through X_4, X_5
X_1 Self-Efficacy	$RP_{14} * RP_{4Y} = 0.111$	$RP_{15} * RP_{5Y} = 0.029$	$RP_{14} * RP_{45} * RP_{5Y} = 0.037$	$P_{14} * P_{4Y} = 0.131$	$P_{15} * P_{5Y} = -0.009$	$P_{14} * P_{45} * P_{5Y} = 0.381$
X_2 System Quality	$RP_{24} * RP_{4Y} = 0.029$	$RP_{25} * RP_{5Y} = 0.097$	$RP_{24} * RP_{45} * RP_{5Y} = 0.009$	$P_{24} * P_{4Y} = 0.113$	$P_{25} * P_{5Y} = -0.047$	$P_{24} * P_{45} * P_{5Y} = 0.328$
X_3 Information Quality	$RP_{34} * RP_{4Y} = 0.082$	$RP_{35} * RP_{5Y} = 0.009$	$RP_{34} * RP_{45} * RP_{5Y} = 0.027$	$P_{34} * P_{4Y} = 0.055$	$P_{35} * P_{5Y} = 0.004$	$P_{34} * P_{45} * P_{5Y} = 0.159$
X_4 Quality of Use	0	$RP_{45} * RP_{5Y} = 0.089$	0	0	$P_{45} * P_{5Y} = 0.533$	0
X_5 User Satisfaction	0	0	0	0	0	0

Table 5. The Direct and Indirect Effects on X_4 using the Simple Correlation Matrix and the Robust Correlation Matrix

	Using the Simple Correlation Matrix			Using the Robust Correlation Matrix		
Variables	DE	IE	TE	DE	IE	TE
X_1 Self-efficacy	$P_{14} = 0.414$	0	0.414	$RP_{14} = 0.716$	0	0.716
X_2 System Quality	$P_{24} = 0.110$	0	0.11	$RP_{24} = (-0.617)$	0	-0.617
X_3 Information Quality	$P_{34} = 0.305$	0	0.305	$RP_{34} = 0.299$	0	0.299

Table 6. The Direct, Indirect and Total Effects on X_5 using the Simple Correlation Matrix and the Robust Correlation Matrix

	Using the Simple Correlation Matrix			Using the Robust Correlation Matrix		
Variables	DE	IE	TE	DE	IE	TE
X_1 Self-Efficacy	0.102	0.127	0.229	-0.014	0.539	0.525
X_2 System Quality	0.334	0.033	0.367	-0.067	0.464	0.397
X_3 Information Quality	0.034	0.093	0.127	0.006	0.225	0.231
X_4 Quality of Use	0.307	0	0.307	0.753	0	0.753

Table 7. The Direct and Indirect Effects on *Y* using the Simple Correlation Matrix and the Robust Correlation

	Using the Simple Correlation Matrix						Using the Robust Correlation Matrix					
Variables	DE		IE		TE		DE		IE		TE	
X_1 Self-Efficacy	0	0	0.177	0.031	0.177	0.031	0	0	0.503	0.253	0.503	0.253
X_2 System Quality	0	0	0.135	0.018	0.135	0.018	0	0	0.394	0.155	0.394	0.155
X_3 Inf. Quality	0	0	0.118	0.013	0.118	0.013	0	0	0.218	0.047	0.218	0.047
X_4 Quality of Use	0.269	0.072	0.089	0.007	0.358	0.128	-0.184	0.033	0.533	0.284	0.349	0.121
X_5 User Satisfaction	0.293	0.085	0	0	0.293	0.085	0.708	0.501	0	0	0.708	0.501

The Direct and Indirect Effects on X_5 are as follows:
The Direct and Indirect Effects on Y are as follows:
We can note from the table when using the Simple Correlation Matrix that there are no Direct Effects of the applicant's self-efficacy and the quality of the system used as well as the quality of the information provided on the applicant's result for electronic tests. On the hand, there are direct effects on the applicant's result for electronic tests with a path of (-0.184) which is equivalent to 3.38%. The higher the Quality of Use with a single unit, the lower the Result by 0.184 unit, and the more satisfied the User is by 1 unit, the more the applicant's electronic testing result improved. As for using the robust matrix, we notice from the table also that there are no direct effects of the applicant's self-efficacy and the quality of the system used as well as the quality of the information provided

on the applicant's result for electronic tests. There are direct effects on the result of the applicant for electronic tests with a path of 0.269, which is equivalent to 7.23 %. The higher the quality of use by one unit, the better the results, and also the greater the user's satisfaction, the better result we will get. In other words, the applicant's success for electronic tests, and the external variables can be arranged according to their direct impact on the test taker's result through the following table:

Table 8. Ranking of the external variables according to the direct effect on the test-taker's result

Variable	Ratio using Simple Correlation	The ratio using the Robust Correlation
Self-Efficacy	0%	0%
System Quality	0%	0%
Information Quality	0%	0%
Quality of Use	7.23%	3.38%
User Satisfaction	8.58%	50.12%

As for the Indirect Effects when using the Robust Simple Correlation Matrix, we can notice from the table that there is an Indirect Effect of the applicant's result for electronic tests with a path factor of 0.503, which is equivalent to 50.3%. The higher the applicant's self-efficacy, the better his/her result. There is also an indirect effect of the quality of the system used on the applicant's result with a path factor of 0.394, which is equivalent to 15.52%, as well as the quality of information for the applicant, the method of using the system and his/her satisfaction have Indirect Effects on the result he will reach. Also, when using the Robust matrix, the Indirect Effects on the score of the applicant for electronic test was 0.177 (equivalent to 3.13%). This is lower than it is when using the Simple Correlation Matrix. As for the Total Effects, we notice when using the Simple Correlation Matrix that it is much less for all the independent variables that affect the result of the test taker than it is when using the Robust Correlation Matrix. The external variables can be arranged according to their Direct impact on the test taker's result through the following table:

Table 9. Ranking of the exogenous variables according to the Indirect Effect on the test-taker's score

Variable	Ratio using Simple Correlation	The ratio using the Robust Correlation
Quality of Use	0.7%	28.4%
Self-Efficacy	3.13%	25.3%
System Quality	1.82%	15.52%
Information Quality	1.39%	4.75%
User Satisfaction	0%	0%

Thus, the order of the effect of external variables on the result of the applicant for postgraduate studies can be clarified through the following table:

Table 10. Ranking of the exogenous variables according to the Total Effect on the test-taker's score

Variable	Ratio using Simple Correlation	The ratio using the Robust Correlation
Quality of use	12.81%	12.18%
Self-efficacy	3.13%	25.30%
System quality	1.82%	15.52%
Information quality	1.39%	4.75%
User satisfaction	8.58%	50.12%

The criteria values can be clarified by relying on the Simple Correlation Matrix and the Robust Correlation Matrix through the following table:

The classical model exhibited relatively poor fit because Pearson correlation is sensitive to outliers and deviations from normality. Although the robust model does not indicate a perfect fit, it provides a substantially better representation of the observed data according to the adopted fit criteria. We can note from the values of the information criteria and the statistical criteria used that the robust method is superior in showing the effects of

Table 11. The statistical Criteria for the methods used in the analysis

Standards	Simple Correlation	Robust Correlation
AIC	3171.627	2994.037
BIC	3231.439	3053,651
$RMSE$R	0.796	0.382
$RMSR$	0.200	0.094
$Chi-Square$	392.599	92,540
CFI	0.569	0.790

the independent variables on the result of the applicant for electronic tests. The observed differences between the classical and robust estimates further demonstrate the practical importance of adopting robust estimation techniques in the presence of outliers. The Pearson based model is more sensitive to extreme observations while the robust method gave more stable estimates of the model parameters and better statistical fit of the model. The strong model provides therefore a more reliable indication of the factors associated with the success of electronic examinations. The power of the robust model goes beyond the improvement of statistical fit. The robust method minimizes the impact of outliers on the parameter estimates, resulting in more representative parameter estimates for the majority of users and more reliable parameters to use to improve electronic examination systems in practice.

7. Conclusions and Recommendations

The study concluded that robust methods outperform classical approaches in estimating path coefficients when using correlation matrices. The results demonstrated that robust techniques provide more accurate and reliable estimates, particularly in the presence of outliers or non-normal data distributions, thereby enhancing the validity of the structural model. Based on these findings, the researcher recommends exploring additional robust methods for estimating path coefficients and conducting comparative analyses. Such investigations can further validate the effectiveness of robust approaches, identify their strengths and limitations, and provide guidance for selecting the most appropriate method depending on the characteristics of the dataset and research objectives. Employing multiple estimation techniques can also improve the credibility of results and support more informed decisions in both theoretical and applied research contexts. University administrators should improve not only the technical quality of electronic examination systems but also user training, interface simplicity, and technical support. These measures are expected to enhance user satisfaction and improve examination outcomes. The proposed methodology should be interpreted as a robust path analysis approach rather than a full robust SEM estimator. Future research may extend this work by incorporating likelihood-based robust SEM estimation methods, such as robust maximum likelihood estimators, to further validate the proposed findings.

REFERENCES

1. D. Suhr, *The Basics of Structural Equation Modeling*, Presented at the SAS User Group of the Western Region of the United States (WUSS), Irvine, CA, pp. 1–9, 2006.
2. L. J. Williams, R. J. Vandenberg, and J. R. Edwards, *Structural Equation Modeling in Management Research: A Guide for Improved Analysis*, Academy of Management Annals, vol. 3, no. 1, pp. 543–604, 2009.
3. C. L. Shook, D. J. Ketchen Jr., G. T. M. Hult, K. M. Kacmar, et al., *An Assessment of the Use of Structural Equation Modeling in Strategic Management Research*, Strategic Management Journal, vol. 25, no. 4, pp. 397–404, 2004.
4. B. M. Byrne, *Structural Equation Modeling with Mplus: Basic Concepts, Applications, and Programming*, Routledge, New York, 2013.
5. T. A. A. R. S. Al, H. A. A. Hasso, and T. H. Ali, *Using Linear Wavelets in Analyzing the GARCH Model with the Simulation*, Pakistan Journal of Statistics, vol. 42, no. 1, pp. 27–44, 2026.
6. G. Keller, *Statistics for Management and Economics*, Cengage Learning, 2005.
7. P. R. Freeman, *On the Number of Outliers in Data from a Linear Model*, Trabajos de Estadística y de Investigación Operativa, vol. 31, no. 1, pp. 349–365, 1980.
8. S. R. Ahmed, H. H. Taha, and H. Hayawi, *Employing A Wavelet To Predict Gold Prices Using Generalized Self-Regression Models Conditioned On Heterogeneity Of Variance*, Statistics, Optimization & Information Computing, vol. 16, no. 1, pp. 147–159, 2026.

9. E. Dan and O. A. Ijeoma, *Statistical Analysis/Methods of Detecting Outliers in a Univariate Data in a Regression Analysis Model*, International Journal of Education and Research, vol. 1, no. 5, pp. 1–24, 2013.
10. H. W. Yaseen, T. H. Ali, S. R. Ahmed, H. Hayawi, et al., *The Impact of Wavelet-Based Denoising on Beta Regression Model Fit*, Statistics, Optimization & Information Computing, vol. 15, no. 4, pp. 3042–3058, 2026.
11. S. M. Shaker, *Improving the Robust M Method in Estimating the Parameters of the Linear Regression Model*, Iraqi Journal of Statistical Sciences, no. 16, pp. 219–242, 2009.
12. G. E. P. Box and S. L. Andersen, *Preliminary Results on a Robust Test for Variances*, Department of Statistics, North Carolina State University, 1953.
13. K. Lai and A. Simoes, *Reflecting on the "Robust" Standard Errors for Two-Stage SEM Estimation With Categorical Data: Mistakes and Correction*, Structural Equation Modeling: A Multidisciplinary Journal, vol. 30, no. 5, pp. 691–707, 2023.
14. M. Al-Hashimi, H. Hayawi, and M. Q. Y. Alawjar, *Bayesian Optimization for Enhanced Prediction of Earthquakes with Machine Learning Techniques*, Statistics, Optimization & Information Computing, vol. 15, no. 5, pp. 4305–4322, 2026.
15. S. J. Devlin, R. Gnanadesikan, and J. R. Kettenring, *Robust Estimation and Outlier Detection with Correlation Coefficients*, Biometrika, vol. 62, no. 3, pp. 531–545, 1975.
16. J. L. Rodgers and W. A. Nicewander, *Thirteen Ways to Look at the Correlation Coefficient*, The American Statistician, vol. 42, no. 1, pp. 59–66, 1988.
17. A. N. Alkhateeb, et al., *Enhancing Surface Water Potability Assessment through a v-SVM Hybrid Statistical Learning Model*, Statistics, Optimization & Information Computing, vol. 15, no. 6, pp. 5481–5494, 2026.
18. M. Al-Hashimi, H. Hayawi, and M. Alawjar, *Predicting Sunspot Numbers Based on Neutrosophic Time Series*, Statistics, Optimization & Information Computing, vol. 15, no. 6, pp. 5431–5444, 2026.
19. P. Langfelder and S. Horvath, *Fast R Functions for Robust Correlations and Hierarchical Clustering*, Journal of Statistical Software, vol. 46, pp. 1–17, 2012.
20. A. M. Algboory, A. N. Alkhateeb, and Z. Y. Algamal, *A Modified Jackknife Liu-Type Estimator for the Gamma Regression Models Data*, Statistics, Optimization & Information Computing, vol. 14, no. 5, pp. 2142–2154, 2025.
21. M. M. Al-Hashimi and A. N. Alkhateeb, *Spatial Analysis of Brain and Other CNS Cancers Incidence in Iraq During 2000–2015*, Malaysian Journal of Public Health Medicine, vol. 20, no. 3, pp. 27–34, 2020.
22. W. W. S. Wei, *Time Series Analysis*, Wiley, 2013.
23. H. Akaike, *Likelihood of a Model and Information Criteria*, Journal of Econometrics, vol. 16, no. 1, pp. 3–14, 1981.
24. N. R. Bassiouni and A. R. Bassiouni, *Using the Path Analysis Method to Determine Factors Affecting Egypt's Inflation Rate*, Journal of Financial and Commercial Research, vol. 22, p. 39, 2021.
25. H. Du and P. M. Bentler, *40-Year Old Unbiased Distribution Free Estimator Reliably Improves SEM Statistics for Nonnormal Data*, Structural Equation Modeling: A Multidisciplinary Journal, vol. 29, no. 6, pp. 872–887, 2022.
26. Yaseen, Husam, Qasim, Salwa, and Algamal, Zakariya, *Bootstrap Liu Estimator for Almon Distributed Lag Model*, Statistics, Optimization & Information Computing, vol. 15, no. 4, pp. 3129–3137, 2026.
27. Yaseen, Husam W., Ali, T.H., Ahmed, S.R., and Hayawi, H., *The Impact of Wavelet-Based Denoising on Beta Regression Model Fit*, Statistics, Optimization & Information Computing, vol. 15, no. 4, pp. 3042–3058, 2026.