Historical Development and Practical Application of Correlation and Path Coefficient Analysis in Agriculture

Abenezer Abebe, Endashaw Girma
Ethiopian institute of Agricultural research, Holeta agricultural research center, P.O.Box-31

Abstract
Correlation coefficient measures the mutual relationship between various plant characters and determines the component characters on which selection can be based for the improvement in associated complex character. Several correlation coefficients based on different statistical hypothesis are used today: Pearson correlation coefficient, Spearman ranks correlation coefficient and Spearman semi-quantitative correlation coefficient, Kendall tau-a, -b and -c correlation coefficients, Gamma correlation coefficient. In quantitative genetics there are three types of correlations and these are phenotypic, genotypic and environmental correlations. Path coefficient analysis splits the correlation coefficients into the measures of direct and indirect effects of a set of independent variables on the dependent variable and estimated using a standardized partial regression coefficient known as path coefficient analysis, as suggested by Dewey and Lu in 1959.Based on this, data from national variety trial conducted at Holetta agricultural research center was analyzed both for correlation coefficient(phenotypic and genotypic) and path coefficient. Phenotypic correlation coefficient between yield and plant height, ear height and ear per plant positive and significant and reveals that the relation is due to environmental influence. However, there is no genotypic correlation between yield and other studied trait which indicting that selecting one trait will not improve the other which refers most of the traits were influenced by environment more relatively. Phenotypic path coefficient analysis between, EH (0.406139) and PH (0.402022) showed positive and high direct effects, whereas genotypic path coefficient analysis the direct effect of EH (0.511255(and PH (0.511028) on GY was considerably higher.
Keyword: correlation coefficient, path coefficient, phenotype, genotype, variance, models

Introduction

The correlation coefficient measures the mutual relationship between various plant characters and determines the component characters on which selection can be based for the improvement in associated complex character – yield (Sokoto et al., 2012; Mohammadi et al., 2012). Such correlations can be either negative or positive and correlated characters are of prime importance because of genetic causes of correlations through pleiotropic action or developmental interactions of genes and changes brought about by a natural or artificial selection (Singh, 1993; Falconer and Mackay 1996; Sharma 1998). Correlation coefficients measure the absolute value of correlation between variables in a given body of data. Correlation does not say anything about the cause and effect of relationship (Wright, S., 1921). There are three types correlations in quantitative genetics and these are phenotypic, genotypic and environmental correlations. The association between two characters that can be directly observed is the correlation of phenotypic values or phenotypic correlations (rp). Genetic correlation (rg) is the associations of breeding values (i.e additive genetic variance) of the two characters. Genetic correlation measures the extent to which degree the same genes or closely linked genes cause co-variation (simultaneous variations) in two different characters. The correlation of environmental deviations together with non-additive genetic deviations (i.e. dominance and epistatic genetic deviations) is referred to as environmental correlations (re) (Falconer and Mackay 1996; Sharma 1998). Correlation alone does not give the exact picture of direct and indirect effect of characters upon each other; thus, path coefficient analysis is preferable, since it can identify the direct and indirect causes of associations and can measure the relative importance of each (Sharma, 1998).

Path analysis is a method for partitioning the correlations among variables (Freedman, 1987). Path coefficient analysis can be defined as "the ratio of standard deviation of the total effect" (Falconer and Mackay, 1996). Path analysis is simply standardized partial regression coefficient, which splits the correlation coefficients into the measures of direct and indirect effects of a set of independent variables on the dependent variable .It requires a cause and effect situation among variables. Path coefficient analysis is a very important statistical tool that indicate which variables (causes) exert influence on other variables (responses), while recognizing the impacts of multi co linearity The path coefficient analysis furnishing the cause and effect of different yield component would provide better index for selection rather than mere correlation coefficients (Freedman, 1987).

Objective:

- To review the historical development ,different models and importance of correlation and path analysis
- To analyze correlation and path analysis using preferable model

Literature Review

Historical development correlation

Sir Francis Galton, the man responsible for the correlation coefficient and cousin to Charles Darwin, first established his scientific credentials in a survey of Africa conducted from 1850-52. Galton's definition reveals the properties of the correlation coefficient. It is a measure of strength of a linear relationship; the closer it is to 1, the closer two variables can be predicted from one another by a linear equation. It is a measure of direction: a positive correlation indicates X, Y increase together; a negative correlation indicates one decrease as the other increases (Hauke and Kossowski, 2011).

Karl Pearson provided the mathematical framework we are familiar with today. His interest in biometrics and statistics stemmed from a reading of Galton's Natural Inheritance and the influence of another professor at University College, W. F. R. Weldon. He was derived the "best value" of the correlation coefficient through a method similar to the modern approach of maximum likelihood estimators, although the first paper on likelihood estimation as a general approach did not appear until 1912.(Hauke and Kossowski ,2011).

Kendall (1938) introduced Kendall's tau coefficient (τ) that can be used as an alternative to Spearman's rho for data in the form of ranks. It is a simple function of the minimum number of neighbor swaps needed to produce one ordering from another. He was stated the coefficient they have introduced provides a kind of average measure of the agreement between pairs of members and thus has evident recommendation as a measure of the concordance between two rankings.

The history and properties of Pearson's correlation coefficient were also described by Pearson (1920), Weida (1927), Walker (1928), Stigler (1988), and Piovani (2008). It is worth noting that some authors use the term "Fisher's correlation coefficient", e.g. Plata (2006), as R.A. Fisher also worked in the area of correlation (Fisher 1915, 1921). His contribution was described by Anderson (1996), who mentioned another statistician interested in properties of the coefficient of correlation, namely W.S. Gosset, known as 'Student' (1908).

Pearson's coefficient of correlation was discovered by Bravais in 1846, but Karl Pearson was the first to describe, in 1896, the standard method of its calculation and show it to be the best one possible. Pearson also offered some comments about an extension of the idea made by Galton .He called this method the "product-moments'' method (or the Galton function for the coefficient of correlation r). An important assumption in Pearson's 1896 contribution was the normality of the variables analyzed, which could be true only for quantitative variables.

In 1904 Spearman adopted Pearson's correlation coefficient as a measure of the strength of the relationship between two variables that cannot be measured quantitatively. He noted: "The most fundamental requisite is to be able to measure our observed correspondence by a plain numerical symbol. There is no reason whatever to be satisfied either with vague generalities such as "large", "medium", "small," or, on the other hand, with complicated tables and compilations (Spearman 1904).

Generally, several correlation coefficients based on different statistical hypothesis are known and most frequently used today: Pearson correlation coefficient, Spearman rank correlation coefficient and Spearman semi-quantitative correlation coefficient, Kendall tau-a, -b and -c correlation coefficients, Gamma correlation coefficient (Rosner B,1995).

Historical development path analysis

The concept of path coefficient analysis was originally developed by (Wright in 1921), but the technique was first used for plant selection by Dewey and Lu (1959).The method of Path coefficient was published first by Professor Sewall Wright about thirty five years ago (Li, 1956). Path analysis originally developed by genetist Sewall Wright in 1920s to examine the effect of hypothesized models in phylogenetic studies as an aid to quantitative development of genetics (Lleras, 2005; Freedman, 1987). Wright's involved writing system of the equation based up on the correlations among variables influencing the outcome and the solving for unknown parameter in the model. According to Wright path analytic method was intended to measure the direct effect along each separate path in such system thus finding degree to which variation of a given effect is determined by each particular cause (Freedman, 1987). Wright's also acknowledged the fact that causal relation was uncertain and cautioned and this method was not intended to deduce causal relations simply from correlation coefficients (Wright, 1934). Rather the method utilized information provided by artificial correlations in conjunction with qualitative information regarding the causal relationships to find the consequences of hypothesized structures.

After several decades, path analysis was introduced to scientific research by Blalock, Duncan and others. A Sociologists Peter Blau and Otis Dudley Duncan were among the first utilized path analysis into research process (Lleras, 2005). In 1967, Blau and Duncan proposed a path model for education and stratification. This is one of the most influential applications of statistical modeling technique to social data. During the 1970's path analysis became more popular and numerous paper were published in path analytical method in sociology, psychology, political science, ecology and other fields. Since the early 1980's, path analysis has involved into a variety of causal or structural equation modeling (SEM) programs and computer packages. Unlike earlier path models, which were based on least squares regression, this new modeling of causal modeling utilize the general linear model approach. The path model for education and social stratification proposed by Blau and Duncan (1967) was one of the first applications of that method in the social sciences, and certainly the most influential (Freedman, 1987; Sobel, 1982). (Wright (1934) stressed the assumptions and the limitations of the technique. Many investigators seem to focus only on the statistical calculations. This essay can be viewed as an attempt to put the spotlight back on the assumptions (Freedman, 1987).

Models or methods for correlation analysis

Correlation can be measured in different indices (coefficient) based on different statistical hypothesis and these are Pearson correlation coefficient, Spearman rank correlation coefficient and Spearman semi-quantitative correlation coefficient, Kendall tau-a, -b and -c correlation coefficients, Gamma correlation coefficient (Rosner B,1995).

Karl Pearson (1857 - 1936) coined the Pearson product-moment correlation coefficient ($\mathbf{r_{prs}}$= Pearson correlation coefficient)and a major contributor to the early development of statistics. Assumes both variable (variables x and Y) are interval or ratio variables and are well approximated by a normal distribution, and their joint distribution is bivariate normal. The formula developed $\mathrm{r} = \frac{\sum(\mathrm{xi}-\bar{\mathrm{x}})(\mathrm{yi}-\bar{\mathrm{y}})}{\sqrt{\sum(\mathrm{xi}-\bar{\mathrm{x}})^2\,\sum(\mathrm{yi}-\bar{\mathrm{y}})^2}}$ the Pearson correlation coefficient can take values from -1 to +1 and considering strong correlation if the correlation coefficient is greater than 0.8 and a weak correlation if the correlation coefficient is less than 0.5. Student t-test was used to determine if the value of Pearson correlation coefficient is statistically significant, at a significance level of 5%. For a significance level equal with 5%, a p-value associated to $\mathbf{t_{Prs}}$, df less than 0.05 means that there is evidence to reject the null hypothesis in favor of the alternative hypothesis.

Spearman's rank correlation coefficient ($\mathbf{r_{Spm}}$= Spearman rank correlation coefficient ρ (rho)) named after Charles Spearman (1863 - 1945). A non-parametric measure of correlation between variable which assess how well an arbitrary monotonic function could describe the relationship between two variables, without making any assumptions about the frequency distribution of the variables. It is satisfactory for testing the null hypothesis of no relationship, but is difficult to interpret as a measure of the strength of the relationship (Bland, 1995). It does not require assumptions about the frequency distribution of the variables, linear relationship between variable and interval or ration scale measurement of variable. However, it needs the value of variables to be converted to ranks (rank X for measurement of variable X, rank Y. $\mathbf{r_{Spm}} = 1 - \frac{6\sum \mathrm{D}^2}{\mathrm{n}(\mathrm{n}^2-1)}$ Where D is the differences between each pair of ranks (e.g. D = Rank X – Rank Y) and n is the volume of the sample. Tested by applying the Student t-test (for sample sizes > 20) using different levels of significance. It is a simple solution when the researchers want to know the significance within a certain range or less than a certain value.

Kendalls tau correlation coefficients ($\boldsymbol{\tau}_{\mathbf{Ken,a,b,c}}$ = the Kendall tau-a,b,c correlation coefficient) named after Maurice George Kendall (1907 - 1983). Kendall-tau is a non-parametric correlation coefficient that can be used to assess and test correlations between non-interval scaled ordinal variables. It is considered to be equivalent to the Spearman rank correlation coefficient. While Spearman rank correlation coefficient is like the Pearson correlation coefficient but computed from ranks, the Kendall tau correlation rather represents a probability. The formulas of Kendalls tau correlation coefficients are as follows: Kendall tau-a correlation coefficient ($\tau_{\mathrm{Ken,a}}$) $\tau Ken, a = \frac{C-D}{\frac{n(n-1)}{2}}$, Kendall tau-b correlation coefficient (τKen,b): $\tau Ken, a = \frac{C-D}{\sqrt{\frac{n(n-1)}{2-t}(\frac{n(n-1)}{2-u}}}$, Where t is the number of tied X values and u is the number of tied Y values. Kendall tau-c correlation coefficient ($\tau_{\mathrm{Ken,c}}$): $\tau Ken, c = \frac{2(C-D)}{n^2}$ where C is the number of concordant pairs (C = (<, <) or (>, >)), D is the number of discordant pairs (D = (<, >) or (>, <)). The value of $\boldsymbol{\tau}_{\mathbf{Ken,a,b,c}} = 1$ agreement between the two rankings is perfect and the two rankings are the same, The value of $\boldsymbol{\tau}_{\mathbf{Ken,a,b,c}} = -1$ disagreement between the two rankings is perfect and one ranking is the reverse of the other .Statistical significance of the Kendalls tau correlation coefficient is testes by the Z test, at a significance level of 5%.

Gamma correlation coefficient (Γ = the Gamma correlation coefficient) also known as Goodman and Kruskal's gamma. The Gamma correlation coefficient (Γ, gamma) is a measure of association between variables that comparing with Kendalls tau correlation coefficients is more resistant to tied data (Goodman and Kruskal ,1963) being preferable to Spearman rank or Kendall tau when data contain many tied observations (Siegel and Castellan,1988).The formula for Gamma correlation coefficient: $\Gamma = \frac{(\mathrm{C}-\mathrm{D})}{(\mathrm{C}+\mathrm{D})}$. Statistical significance of Gamma correlation coefficient was tested by the Z-test, at a significance level of 5%.

Finally, among all stated models or methods to measure degree of the relationship between linearly related variables Pearson correlation coefficient is the one most widely used.

Models or methods for path analysis

Path coefficient analysis is measure of direct and indirect effects of each character on bean yield will be estimated using a standardized partial regression coefficient known as path coefficient analysis, as suggested by Dewey and Lu (1959).

rij =Pij+ Σrikpkj Where:- rij = Mutual association between the independent character (i) and dependent Character (j) as measured by the correlation coefficient ,Pij = Component of direct effects of the independent character (i) on dependent character, (j)as measured by the path coefficient and,

∑ Rikpkj = Summation of components of indirect effect of a given independent character (i) on the given dependent character (j) via all other independent character (k).

Residual effect will be estimated by the formula: √1 – R2 ,Where: - R2 =Σpijrij

pij = Component of direct effects of the independent character (i) and dependent character (j) as measured by the path coefficient.rij = Mutual association between the independent character (i) and dependent character (j) as Measured by the correlation coefficient.

Result and discussion

Correlation and Path analysis

Maize preliminary variety trial for ten genotype and two standard checks namely jibat and wenchi was conducted at Holeta Agricultural research station using randomized complete block design with three replications. The data's collected were yield, 50%anthesise date, anthesise silking interval, plant height, ear height, ear per plant, ear position, husk cover, stand count at harvest (number of plant) and ear rot. The phenotypic and genotypic correlation coefficient and path coefficient analysis is calculated or extracted from the data's collected. SAS software (9.0) is used to analysis phenotypic and genotypic correlation coefficient, while for path analysis excel is used.

Partitioning the correlation coefficient into direct and indirect effects was done through path analysis technique (Dewey and Lu, 1959). Many investigators have used this technique on soybeans sugar beet (Naser and Leilah, 1993) and wheat (Mohamed, 1999; Leilah and Al-Khateeb, 2005). The path coefficient analysis was conducted following the procedure developed by Wright (1921) and applied by Dewey & Lu (1959).

Phenotypic and Genotypic Correlation and path analysis

Phenotypic and genotypic correlations is calculated for the characters by working out the variance components of each character and the covariance components for each pair of characters using the formulae (Robinson and Comstock, 1952):Genotypic coefficient variation (GCV) = $\frac{\sqrt{\delta_g^2}}{grand\ mean} * 100$, Genotypic variance $\delta_g^2 = \frac{Msg-Mse}{r}$, Msg and Mse mean square of genotype and error, r =Number of replication,Phenotypic variance $\delta_p^2 = \delta_p^2 + \delta_e^2$, Environmental variance(δ_e^2) =MSe,Phenotypic Coefficient of Variation (PCV) = $\frac{\sqrt{\delta_p^2}}{grand\ mean} * \mathbf{100}$

Table 7.Mean of traits of maze genotypes

ENT	GYD t/ha	ASI	PH	EH	EPO	EPP
1	6.423	4.000	244.000	125.000	0.510	1.367
2	8.837	3.000	271.333	145.333	0.537	1.620
3	8.237	3.667	264.667	148.333	0.560	1.663
4	8.330	3.667	295.667	176.333	0.593	1.517
5	10.027	8.000	264.333	145.667	0.550	1.527
6	9.177	3.667	255.333	138.000	0.543	1.180
7	7.077	3.000	260.667	130.000	0.497	1.060
8	9.210	3.667	259.667	144.000	0.553	1.493
9	8.977	3.333	275.000	151.333	0.550	1.763
10	7.640	3.333	252.000	136.667	0.543	1.363
11	9.267	4.000	267.000	138.333	0.520	1.533
12	8.003	2.667	239.333	120.333	0.500	1.313
G.mean	8.434	3.833	262.417	141.611	0.538	1.450
Msg	3.16	5.67	658.49	625.38	0.002	0.12
Mse	0.76	4.53	58.14	110.89	0.001	0.03
F-test	**	Ns	**	**	*	**

Note; Msg= mean square of genotype, Mse= mean square of error G.mean =grand mean

Table 8.variance components of 12 maize genotypes

Traits	δ_g^2	δ_e^2	δ_p^2	G. mean	GCV	PCV
GYD t/ha	0.8	0.76	1.56	8.434	10.61	14.81
ASI	0.38	4.53	4.91	3.833	16.08	57.81
PH	200.12	58.14	258.26	262.417	5.39	6.12
EH	171.5	110.89	282.39	141.611	9.25	11.87
EPO	0.0003	0.001	0.0013	0.538	3.22	6.70
EPP	0.03	0.03	0.06	1.45	11.95	16.89

*****GYD=grain yield, ASI=anthesis slicking interval, PH=plant height, EH=ear height, EPO=ear position, EPP=eat per plant***

Genotypic correlation coefficient(r) $=\frac{cov\ g(XY)}{\sqrt{var\ g\ X * var\ g\ Y}}$,cov_g=genotypic covariance variance, var_g=genotypic variance and phenotypic correlation coefficient=$\frac{cov\ p(XY)}{\sqrt{var\ p\ X * var\ p\ Y}}$, cov_p= phenotypic covariance, var_p=phenotypic variance where, X and Y different traits under studied.

Table 9. Phenotypic (below diagonal) and genotypic (above diagonal) correlation coefficient

	GY	ASI	PH	EH	EPO	EPP
GY	1.00000	0.49363	0.38023	0.41089	0.45026	0.45484
ASI	0.20282	1.00000	0.09500	0.16442	0.23288	0.17663
PH	0.28340*	0.01522	1.00000	0.93746**	0.73298**	0.50830
EH	0.28540*	0.02323	0.88372**	1.00000	0.91962**	0.55400
EPO	0.27503	0.01709	0.60446**	0.90344**	1.00000	0.54933
EPP	0.42600**	-0.07590	0.33960*	0.38960*	0.38194*	1.00000

*significant at 5% levels of significance, ** significant at 1% level of significance

Table 10.Direct effect (bold diagonal) and indirect effect of traits on yield (phenotype)

	ASI	PH	EH	EPO	EPP
ASI	**0.316935**	0.00612	0.009435	0.006563	-0.01705
PH	0.004824	**0.402022**	0.358914	0.232127	0.076265
EH	0.007362	0.355275	**0.406139**	0.90344	0.087494
EPO	0.005416	0.243006	0.366923	**0.384023**	0.085774
EPP	-0.02406	0.136527	0.158232	0.146674	**0.224574**

Table 11. Direct effect (bold diagonal) and indirect effect of traits on yield (genotype)

	ASI	PH	EH	EPO	EPP
ASI	**0.407476**	0.048548	0.084061	0.115787	0.080485
PH	0.03871	**0.511028**	0.479281	0.364433	0.231617
EH	0.066997	0.479068	**0.511255**	0.457229	0.252442
EPO	0.094893	0.374573	0.47016	**0.497194**	0.250314
EPP	0.071973	0.259756	0.283235	0.273124	**0.455671**

The magnitude of difference between PCV and GCV is relatively high for most of the traits revealing more influence of the environment in the expression of these traits. The higher differences between GCV% and PCV% indicated the low possibility to genetic improvement of the traits. Phenotype coefficient of variation is greater than that of genotype which reveals variation is due to environment and not reproducible. Hossain and Joarder, (2006) reported that higher differences among both GCV and PCV for most of the traits attributed to higher modifying effect of environment on the association of characters.

The analysis indict that phenotypic correlation coefficient between yield and plant height, ear height and ear per plant positive and significant and reveals that the relation is due to environmental influence. However, there is no genotypic correlation between yield and other studied trait which indicting that selecting one trait will not improve the other. Plant height shows positive and significant phenotypic and genotypic correlation with ear height and ear position and selecting or improving one of the traits will improve the other. High significant and strong positive correlation is obtained between ear height and ear position. Phenotypic correlation includes both genotypic and environmental effect. Hence significant phenotypic correlation without significant genotypic correlation has no value.

Estimation of phenotype and genotypic direct effects by conventional path analysis in (Table 4) and (Table.5) respectively, where the five (ASI, PH, EH, EPO, and EPP) yield-related characters were considered as first-order variables with GY as the response variable, and analysis of path were made to identify the patterns of relationships among the variables. For example, phenotypic path coefficient analysis between, EH (0.406139) and PH (0.402022) showed positive and high direct effects, whereas genotypic path coefficient analysis the direct effect of EH (0.511255(and PH (0.511028) on GY was considerably higher. Out the five studied trait no one had a negative direct effect on GY, but a phenotypic path coefficient analysis reveal that EPP had a negative indirect effect on GY (table 4). Besides such inconsistent patterns of indirect effects, high positive phenotypic and genotypic path coefficient analysis between EH and EPO was (0.90344) and relatively high (0.47016) respectively. This exhibit that the indirect positive effect of EH on GY was very high phenotypically, as indicated above in (table 4), whereas relatively low genotypically (table 5). This ensures that this trait is influenced by environment more relatively, because the genotypic path coefficient value predicts the effect of genotype from phenotype in quantitative genetics (Dewey and Lu, 1959).

Conclusion

Correlation coefficients measure the absolute value of correlation between variables in a given body of data. Correlation alone does not give the exact picture of direct and indirect effect of characters upon each other; it does not say anything about the cause and effect of relationship. However, path coefficient analysis is simply standardized partial regression coefficient, which splits the correlation coefficients into the measures of direct and indirect effects.

The object of this study is to review the historical development of different models and importance of correlation and path analysis and to analyze data using preferable model. Several correlation coefficients models or methods based on different statistical hypothesis are known and most frequently used today but, Pearson correlation coefficient is the one most widely used and for path coefficient analysis the model suggested by Dewey and Lu in 1959.

Accordingly, the phenotypic and genotypic correlation is analyzed for preliminary variety trial with twelve genotypes for six traits and obtained that the phenotypic correlation value is greater than the genotypic one .this reveals the variation is non-heritable or due to environmental which is not improved through selection.

Reference

Anderson, T.W., 1996. RA Fisher and multivariate analysis. *Statistical Science*, pp.20-34.

Bland, 1995 An Introduction to Medical Statistics, Oxford University Press; Oxford, New York, Tokyo, p. 205-225.

Blau, P.M. and Duncan, O.D., 1967. The American occupational structure.

Br a v a i s. A.,1846. Analyse mathématique sur les probabilités des erreurs de situation d'un point. Mémoires présentés par divers savants à l'Académie Royale des Sciences .

Comstock, R.E. and Robinson, H.F., 1952. Estimation of average dominance of genes. Estimation of average dominance of genes.de France 9: 255–332.

Dewey JR, Lu KH 1959. A correlation and path coefficient analysis of components of crested wheat grass seed production. Agron. J. 51:515-520.

Falconer, D.S. and Trudy F.C. Mackay. 1996. Introduction to Quantitative Genetics 4th ed. Longman Group Limited Malaysia.

Fisher, R.A., 1915. Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. Biometrika, *10*(4), pp.507-521.

Fisher, R.A., 1921. On the probable error of a coefficient of correlation deduced from a small sample. Metron, 1, pp.3-32.

Freedman, D.A., 1987. As others see us: A case study in path analysis. Journal of educational statistics, 12(2), pp.101-128.

Goodman L. A., Kruskal W.H.,1963 Measures of association for cross-classifications III: Approximate sampling theory, J. Amer. Statistical Assoc. 58, , p. 310-364.

Hauke j., kossowski t., 2011.Comparison of values of Pearson's and Spearman's correlation coefficient on the same sets of data. Quaestiones Geographicae 30(2), pp. 87–93.

Hossain ARMM and N. Joarder. 1987. Studies on some mechanical cells in the basal internodes of some rice cultivars in relation to lodging. Pakistan J Agric. Res., 8: 24-28.

Kendall M.G.,1938. A new measure of rank correlation. Biometrika 30: 81–89. kendall M.G.,1948. Rank correlation methods. 4th ed. Griffin, London.

Leilah, A.A. and Al-Khateeb, S.A., 2005. Statistical analysis of wheat yield under drought conditions. Journal of Arid environments, 61(3), pp.483-496.

Li, C.C., 1956. The concept of path coefficient and its impact on population genetics. Biometrics, 12(2), pp.190-210.

Lleras, C., 2005. Path analysis. Encyclopedia of social measurement, 3(1), pp.25-30.

Mohamed, N.A., 1999. Some statistical procedures for evaluation of the relative contribution for yield components in wheat. Zagazig Journal of Agricultural Research, 26(2), pp.281-290.

Mohammadi M, Sharifi P, Karimizadeh R, Kazem M, Shefazadeh MK 2012. Relationships between grain yield and yield components in bread wheat under different water availability (dryland and supplemental irrigation conditions). Not. Bot. Hortic. Agrobo. 40(1):195-200.

Mohammadi SA 2013. Relationship between grain yield and related agronomic traits in bread wheat recombinant inbred lines under water deficit condition. Ann. Biol. Res. 4(4):7-11.

Nasr, S.M. and Leilah, A.A., 1993. Integrated analysis of the relative contribution for some variables [leaves-roots] in sugar beet using some statistical techniques. Bulletin of Faculty of Agriculture, Cairo Univ.(Egypt).

Pearl, R., 1936. Karl Pearson, 1857–1936. Journal of the American Statistical Association, *31*(196), pp.653-664.

Pearson, K., 1896. Mathematical contributions to the theory of evolution. III. Regression, heredity, and panmixia. *Philosophical Transactions of the Royal Society of London. Series A, containing papers of a mathematical or physical character*, *187*, pp.253-318.

Pearson, K., 1920. Notes on the history of correlation. *Biometrika*, *13*(1), pp.25-45.

Piovani, J.I., 2008. The historical construction of correlation as a conceptual and operative instrument for empirical research. *Quality & Quantity*, *42*(6), pp.757-777.

Plata, S., 2006. A note on Fisher's correlation coefficient. *Applied mathematics letters*, *19*(6), pp.499-502.

Rosner B., Fundamentals of Biostatistics, 4th Edition, Duxbury Press, Belmont, California, USA, 1995.

Sharma J.R. 1998. Statistical and biometrical techniques in plant breeding. New Age International (P) limited, publishers. New Delhi. 432p.

Siegel S., Castellan N. J., 1988 Nonparametric Statistics for the Behavioural Sciences, 2nd Edition, McGraw-Hill,.

Singh, B.D. 1993. Plant Breeding Principles and Methods. Kalyani Publishers. Ludhiana. New Delhi.

Sobel, 1982. Asymptotic confidence intervals for indirect effects in structural equation models. Sociological methodology, 13, pp.290-312.

Sokoto MB, Abubakar, IU, Dikko AU (2012). Correlation analysis of some growth, yield, yield components and grain quality of wheat (Triticum aestivum L.). Niger. J. Basic Appl. Sci. 20(4):349-356.

Spearman C.E, 1904. The proof and measurement of association between two things. American Journal of Psychology 15: 72–101.

Stigler, S.M., 1989. Francis Galton's account of the invention of correlation. *Statistical Science*, pp.73-79.

Student, 1908. Probable error of a correlation coefficient. *Biometrika*, pp.302-310.

Walker, H.M., 1928. The relation of Plana and Bravais to theory of correlation. *Isis*, *10*(2), pp.466-484.

Weida, F.M., 1927. On various conceptions of correlation. *Annals of Mathematics*, pp.276-312.

Wright, S 1921. Systems of mating. Genetics 6(1):11-78. Zafarnaderi N, Aharizad S,

Wright, S.,1934. The method of path coefficients. The annals of mathematical statistics, 5(3), pp.161-215.