The Journal of Animal & Plant Sciences, 20(4): 2010, Page: 220-224
ISSN: 1018-7081

PATH ANALYSIS OF SEED YIELD COMPONENTS USING DIFFERENT CORRELATION COEFFICIENTS IN SAFFLOWER (*Carthamus tinctorius* L.)

M. Topal, E. Ozturk* and T. Polat*

Department of Biometry and Genetics, and *Department of Field Crops, Faculty of Agriculture, Ataturk University, 25240, Erzurum, TURKEY
Corresponding author e-mail address: mtopal@atauni.edu.tr

ABSTRACT

Coefficients of Pearson correlation (r) and Spearman's rho (r_s) among plant height, head diameter, number of seeds/head, oil yield, 1000-seed weight on seed yield in safflower (*Carthamus tinctorius* L.) was investigated by breaking up direct and indirect effects in this study during 2001 and 2002. Direct and indirect effects estimated with parametric and nonparametric path analysis exhibited similarity but parametric path analysis was preferred to nonparametric path analysis. Because, residual effects of parametric path analysis are lower than that of nonparametric path analysis. Furthermore, determination coefficients of parametric path analysis were higher than that of nonparametric path analysis in both years. The direct effect of oil yield and indirect effect of number of seed/head on seed yield via oil yield were found large by parametric and nonparametric path analysis in 2001/02. It was concluded that oil content, number of seed/head and plant height were important selection characters for seed yield of safflower under drought conditions.

Key Words: parametric and non parametric path analysis, safflower; spearman's rho

INTRODUCTION

In agricultural studies, to maximize the output and to minimize the input, it is very important to know that which factors have an effect on the agricultural yield and whether they directly affect the yield or not? Direct and indirect effects of yield factors are determined with path analysis. The total correlations between predictor variables and response variable are partitioned into direct and indirect effects by path analysis.

Path model is a diagram relating to independent and dependent variables. Path coefficient is a standardized regression coefficient showing the direct effect of an independent variable on a dependent variable in the path model. The theory and application of path analysis was described (Wright, 1960 a,b; Tabachnick and Fidell, 1996; Dofing and Knight, 1992). Path coefficients determined using phenotypic and genetic correlation coefficients among several agronomic traits (Kang, *et al*. 1983). Williams *et al*. (1990) investigated a concise format for tables of path analysis coefficients. Path analysis to reciprocally interacting variables was discussed by Wright, (1960b) Mokhtassi *et al*. (2006); Mozaffari and Asadi (2006); Pahlavani (2005); Mahasi *et al* (2006); Omidi (2000) investigated direct and indirect effects of seed yield components on seed yield in safflower.

The purpose of this investigation was to present direct and indirect effects of seed yield components on safflower seed yield. Pearson correlation coefficient is used for path analysis generally but direct and indirect effects were estimated with Spearman's rho in the present investigation. Nonparametric path coefficients and indirect effects were calculated with partitioning of Spearman's rho coefficient between seed yield and seed yield components. Direct and indirect effects determined with respect to two relation measures were compared in point of how and which level affected the yield.

MATERIALS AND METHODS

A field study was conducted in non-irrigation conditions on the farm of the Agricultural Research and Extension Center of Atatürk University at the Erzurum ecological conditions (29°55' N and 41°16' E, 1850 m elevation) in Turkey, during the 2001 and 2002 growing seasons. The experiments were performed in loamy-clay soil with 0.73% organic matter, pH of 7.6, and available P and K levels of 68.6 and 1898 kg/ha, respectively. Mean temperature of experimental area was 14.9°C and average rainfall was 192.3 mm. The experiment was laid out as a randomized complete block design with three replicates. Fourteen oilseed safflower genotypes were used. The genotypes were hybrids and open-pollinated. The experimental plots were 1.6 m wide and 5 m long and consisted of 4 rows spaced 0.4 m apart. The crop was planted on 1 May and 2 May during 2001 and 2002, respectively. Harvesting was done at the stage of physiological maturation (in the third week of Sep. in both years). Data were recorded on ten randomly selected plants from the mid-rows. Seed yield (kg/da), plant height (cm), head diameter (cm), 1000-seed weight (g), oil yield (kg/da), number of heads/plant and number of seeds/head were recorded.

Direct and indirect effects of plant height (X_1), head diameter (X_2), number of seed/head (X_3), number of heads/plant (X_4), oil yield (X_5) and 1000-seed weight (X_6) on seed yield (Y) were investigated with parametric and nonparametric path analysis. Path diagram between independent variable (X_1, X_2, X_3, X_4, X_5 and X_6) and dependent variable (Y) is given as Figure 1.

In the path diagram path coefficients (direct effects) are shown by one-headed arrows, and correlations coefficients between independent variables are denoted by double-headed arrows. Each independent variable has one direct effect, and one indirect effect for each of the other independent variables connected with the dependent variable. Y variable was direct affected by error (P_{YU}) and indirect effects of error on egg yield via X_1, X_2, X_3, X_4, X_5, X_6 no showed because error was assumed to be independent and identically from X_1, X_2, X_3, X_4, X_5, X_6.

The total correlations between seed yield and yield components are partitioned into direct and indirect effects by following equations;

$$
\begin{aligned}
r_{Y1} &= P_{Y1} + r_{12}P_{Y2} + r_{13}P_{Y3} + r_{14}P_{Y4} + r_{15}P_{Y5} + r_{16}P_{Y6} \\
r_{Y2} &= r_{12}P_{Y1} + P_{Y2} + r_{23}P_{Y3} + r_{24}P_{Y4} + r_{25}P_{Y5} + r_{26}P_{Y6} \\
r_{Y3} &= r_{13}P_{Y1} + r_{23}P_{Y2} + P_{Y3} + r_{34}P_{Y4} + r_{35}P_{Y5} + r_{36}P_{Y6} \qquad [1] \\
r_{Y4} &= r_{14}P_{Y1} + r_{24}P_{Y2} + r_{34}P_{Y3} + P_{Y4} + r_{45}P_{Y5} + r_{46}P_{Y6} \\
r_{Y5} &= r_{15}P_{Y1} + r_{25}P_{Y2} + r_{35}P_{Y3} + r_{45}P_{Y4} + P_{Y5} + r_{56}P_{Y6} \\
r_{Y6} &= r_{16}P_{Y1} + r_{26}P_{Y2} + r_{36}P_{Y3} + r_{46}P_{Y4} + r_{56}P_{Y5} + P_{Y6}
\end{aligned}
$$

In the equation [1], Coefficients given by P_{Yi} are path coefficients (direct effects) between independent variable (*i*th) and dependent variable (Y). $r_{ij}P_{Yi}$ represent indirect effects of independent variable on dependent variable via *j*th independent variable. r_{ij} represent correlation coefficients between *i*th and *j*th traits. The sum of direct and indirect effects give coefficient of correlation between Y and X_i (Topal and Esenbuga, 2000).

The residual effect, P_{YU}, is calculated with the following equation.

$$
P_{YU} = \begin{bmatrix} 1-(P_{Y1}^2+P_{Y2}^2+P_{Y3}^2+P_{Y4}^2+P_{Y5}^2+P_{Y6}^2+2r_{12}P_{Y1}P_{Y2}+2r_{13}P_{Y1}P_{Y3}+2r_{14}P_{Y1}P_{Y4}+2r_{15}P_{Y1}P_{Y5} \\ +2r_{16}P_{Y1}P_{Y6}+2r_{23}P_{Y2}P_{Y3}+2r_{24}P_{Y2}P_{Y4}+2r_{25}P_{Y2}P_{Y5}+2r_{26}P_{Y2}P_{Y6}+2r_{34}P_{Y3}P_{Y4}+2r_{35}P_{Y3}P_{Y5} \\ +2r_{36}P_{Y3}P_{Y6}+2r_{45}P_{Y4}P_{Y5}+2r_{46}P_{Y4}P_{Y6}+2r_{56}P_{Y5}P_{Y6}) \end{bmatrix}^{1/2}
$$

[2]

$$P_{YU} = \sqrt{1-R^2}$$

Where, R^2 is coefficient of determination.

Equation 1 is written, in matrix form, as follows:

$$
\begin{bmatrix} r_{Y1} \\ r_{Y2} \\ r_{Y3} \\ r_{Y4} \\ r_{Y5} \\ r_{Y6} \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} & r_{15} & r_{16} \\ r_{21} & r_{22} & r_{23} & r_{24} & r_{25} & r_{26} \\ r_{31} & r_{32} & r_{33} & r_{34} & r_{35} & r_{36} \\ r_{41} & r_{42} & r_{43} & r_{44} & r_{45} & r_{46} \\ r_{51} & r_{52} & r_{53} & r_{54} & r_{55} & r_{56} \\ r_{61} & r_{62} & r_{63} & r_{64} & r_{65} & r_{66} \end{bmatrix} \begin{bmatrix} P_{Y1} \\ P_{Y2} \\ P_{Y3} \\ P_{Y4} \\ P_{Y5} \\ P_{Y6} \end{bmatrix}
$$

Standardizations of dependent variable Y and the independent variables X_1, X_2, X_3, X_4, X_5, X_6 are as follows:

$$Z_Y = \frac{Y_i - \overline{Y}}{S_Y}, \quad Z_{ik} = \frac{X_{ik} - \overline{X}_k}{S_k} \qquad k=1,.,6$$

Where, $\overline{Y}$ and $\overline{X}$ are the respective means of *Y* and X_k, and S_Y and S_k are the respective standard deviations of *Y* and X_k. Intercept parameter is no in the standardized regression model. Standardized variables are centered with mean 0 ($\mu_i=0$) and variance 1 ($\sigma_i^2 = 1$) (Neter at al. 1995). The standardized path coefficient and concrete path regression were discussed (Wright, 1996a,b).

Standardized regression model is as follows:

$$Y = b_{Y1}X_1 + b_{Y2}X_2 + b_{Y3}X_3 + b_{Y4}X_4 + b_{Y5}X_5 + b_{Y6}X_6 \qquad [3]$$

Standardized regression coefficients are calculated as follows

$$
\begin{aligned}
r_{Y1} &= r_{11}b_{Y1} + r_{12}b_{Y2} + r_{13}b_{Y3} + r_{14}b_{Y4} + r_{15}b_{Y5} + r_{16}b_{Y6} \\
r_{Y2} &= r_{12}b_{Y1} + r_{22}b_{Y2} + r_{23}b_{Y3} + r_{24}b_{Y4} + r_{25}b_{Y5} + r_{26}b_{Y6} \\
r_{Y3} &= r_{13}b_{Y1} + r_{23}b_{Y2} + r_{33}b_{Y3} + r_{34}b_{Y4} + r_{35}b_{Y5} + r_{36}b_{Y6} \qquad [4] \\
r_{Y4} &= r_{14}b_{Y1} + r_{24}b_{Y2} + r_{34}b_{Y3} + r_{44}b_{Y4} + r_{45}b_{Y5} + r_{46}b_{Y6} \\
r_{Y5} &= r_{15}b_{Y1} + r_{25}b_{Y2} + r_{35}b_{Y3} + r_{45}b_{Y4} + r_{55}b_{Y5} + r_{56}b_{Y6} \\
r_{Y1} &= r_{16}b_{Y1} + r_{26}b_{Y2} + r_{36}b_{Y3} + r_{46}b_{Y4} + r_{56}b_{Y5} + r_{66}b_{Y6}
\end{aligned}
$$

Equation [1] and [4] showed relationship between standardized regression analysis and path analysis.

RESULTS AND DISCUSSION

Descriptive statistics of seed yield (Y), plant height (X_1), head diameter (X_2), number of seed/head (X_3), number of heads/plant (X_4), oil yield (X_5) and 1000-seed weight (X_6) are summarized table 1.

Correlation coefficients among seed yield (Y), plant height (X_1), head diameter (X_2), number of seed/head (X_3), number of heads/plant (X_4), oil yield (X_5) and 1000-seed weight (X_6) given table 2.

Correlation matrix in each one of two coefficient measure (r, r_s) clearly shows (Table 2) that correlation between seed yield (Y) and oil content (X5) was found higher than the other correlations in both years showing linear relationship between seed yield and oil content. Tunçtürk and Çiftçi (2004) have observed similar result. Seed yield (Y) has significant negative correlation with number of heads/plant ($P<0.01$). It has been determined that as number of heads/plant decreased, seed yield increased. Seed yield has significant positive correlation with number of seed/head but negative correlation with 1000-seed weight. It was observed that as 1000-seed weight increased the seed yield decreased, and when number of seed/head increased the seed yield increased too.

Parametric and nonparametric path coefficients and indirect effects of plant height (X_1), head diameter (X_2), number of seed/head (X_3), number of heads/plant (X_4), oil yield (X_5) and 1000-seed weight (X_6) on seed yield (Y) computed according to Pearson correlation (r) and spearman's rho (r_s) given table 3.

Path analysis structured in respect of parametric and nonparametric coefficient measure (r, r_s) in both years (Table 3) shows that the larger path coefficients on seed yield found oil content (X_5) and plant height (X_1) respectively. Number of heads/plant (X_4) has the large negative direct effect on seed yield. Mahasi *et al.* (2006) reported that number of heads/plant (X_4) had negative direct effect on seed yield. The indirect effect of Number of heads/plant (X_4) on seed yield via oil content (X_5) has the large negative indirect effects, and via plant height (X_1) and number of seed/head (X_3) has little negative. Mozaffari and Asadi (2006) found that indirect effect of number of heads/plant via plant height and number of seed/head on seed yield was negative in safflower. Number of seed/head (X_3) has positive little direct effect on seed yield and lesser negative indirect effect via 1000-seed weight, but has positive indirect effect via oil content on seed yield. Tunçtürk and Çiftçi (2004) have observed that Number of seed/head (X_3) had positive little direct effect on seed yield, but had positive indirect effect via oil content on seed yield. Head diameter (X_2)

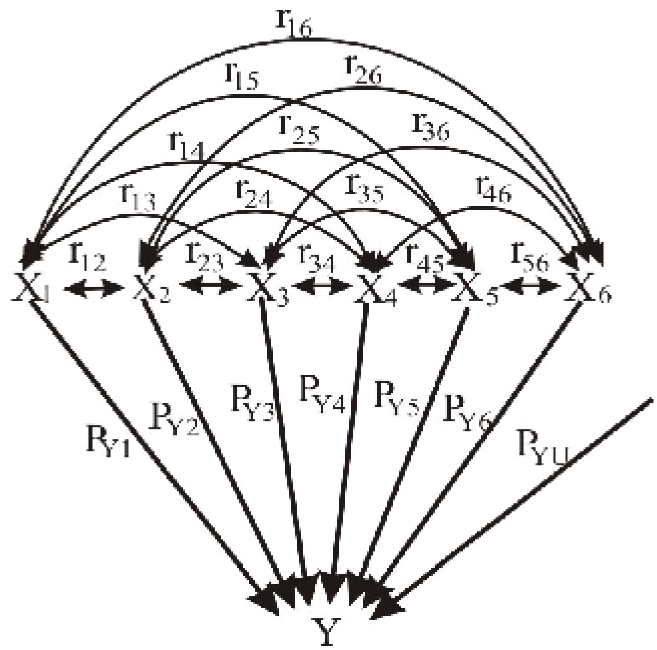

Figure 1. Path diagram for predictor variables, X_1 to X_6, and response variable Y. P_{Yi}: path coefficient (direct effect), r_{ij}: correlation coefficient, P_{YU}: residual effect

Table 1: Descriptive statistics of seed yield and yield components

Year	2001				2002			
	$\overline{X} \pm S_{\overline{X}}$	C.V.(%)	Min.	Max.	$\overline{X} \pm S_{\overline{X}}$	C.V.(%)	Min.	Max.
Seed Yield (kg da^{-1}) (Y)	97.06 ± 4.65	31.08	39.45	143.55	117.62 ± 3.42	18.92	73.48	160.78
Plant Height (cm) (X_1)	55.66 ± 0.68	7.96	46.50	66.40	77.53 ± 1.21	10.13	68.00	105.40
Head Diameter (cm) (X_2)	1.81 ± 0.02	6.30	1.60	2.30	2.29 ± 0.02	6.81	1.98	2.66
Number of seed/head (X_3)	27.94 ± 0.83	19.22	19.90	40.60	36.65 ± 1.12	19.78	25.50	54.70
Number of heads/plant (X_4)	14.22 ± 0.58	26.44	5.60	20.30	18.75 ± 0.49	16.80	11.80	24.60
Oil content (%) (X_5)	25.33 ± 1.14	29.25	10.88	40.13	33.53 ± 1.13	21.77	20.27	48.68
1000-seed weight (g) (X_6)	38.69 ± 0.40	6.67	32.35	43.90	41.21 ± 0.60	9.49	32.45	49.50

$\overline{X}$: mean $S_{\overline{X}}$: Standard Error of mean C.V.: coefficient of variation

Table 2: Pearson correlation (r) and Spearman's rho (r_s) among seed yield and seed yield components in 2001 and 2002

Year		2001						2002					
		Y	X_1	X_2	X_3	X_4	X_5	Y	X_1	X_2	X_3	X_4	X_5
r	X_1	0.248						0.245					
	X_2	0.117	-0.196					0.307*	0.611**				
	X_3	0.526**	0.042	0.290				0.508**	0.366*	0.630**			
	X_4	-0.512**	-0.004	-0.165	-0.258			-0.534**	-0.434**	-0.471**	-0.297		
	X_5	0.944**	0.123	0.272	0.468**	-0.470**		0.827**	-0.184	-0.052	0.342*	-0.255	
	X_6	-0.301	-0.384*	0.171	0.094	0.318*	-0.215	-0.027	-0.480**	-0.096	-0.191	0.227	0.080
r_s	X_1	0.118						0.423**					
	X_2	0.069	-0.242					0.369*	0.499**				
	X_3	0.494**	0.049	0.071				0.514**	0.316*	0.581**			
	X_4	-0.489**	0.038	-0.140	-0.302			-0.428**	-0.356*	-0.537**	-0.250		
	X_5	0.933**	0.008	0.211	0.473**	-0.410**		0.800**	0.103	0.028	0.348*	-0.190	
	X_6	-0.250	-0.256	0.308*	0.067	0.315*	-0.161	0.128	-0.189	0.044	-0.020	0.224	0.092

*Significant at the 0.05 probability level **Significant at the 0.01 probability level

Table 3: Direct and indirect effects of seed yield components on seed yield using Parson correlation (r) and Spearman's rho (r_s) in 2001-2002

Year		2001				2002			
Direct Effect	Indirect effe	r	P_r	r_s	Pr_s	r	P_r	r_s	Pr_s
X_1		0.248	0.090	0.118	0.083	0.245	0.333	0.423	0.255
	X_2		0.029		0.028		0.029		0.018
	X_3		0.006		0.002		0.017		0.042
	X_4		0.001		-0.005		0.074		0.068
	X_5		0.105		0.007		-0.151		0.070
	X_6		0.018		0.003		-0.057		-0.030
X_2		0.117	-0.146	0.069	-0.114	0.307	0.047	0.369	0.036
	X_1		-0.017		-0.020		0.203		0.127
	X_3		0.043		0.003		0.030		0.077
	X_4		0.013		0.018		0.081		0.103
	X_5		0.233		0.186		-0.043		0.019
	X_6		-0.008		-0.004		-0.011		0.007
X_3		0.526	0.148	0.494	0.044	0.508	0.047	0.514	0.133
	X_1		0.004		0.004		0.122		0.080
	X_2		-0.042		-0.008		0.030		0.021
	X_4		0.021		0.039		0.051		0.048
	X_5		0.400		0.416		0.281		0.235
	X_6		-0.004		-0.001		-0.023		-0.003
X_4		-0.512	-0.080	-0.489	-0.130	-0.534	-0.171	-0.428	-0.192
	X_1		-0.001		0.003		-0.144		-0.091
	X_2		0.024		0.016		-0.022		-0.019
	X_3		-0.038		-0.013		-0.014		-0.033
	X_5		-0.402		-0.361		-0.210		-0.128
	X_6		-0.015		-0.004		0.027		0.035
X_5		0.944	0.855	0.933	0.880	0.827	0.822	0.800	0.676
	X_1		0.011		0.001		-0.061		0.026
	X_2		-0.040		-0.024		-0.002		0.001
	X_3		0.069		0.021		0.016		0.046
	X_4		0.038		0.053		0.043		0.036
	X_6		0.010		0.002		0.009		0.015
X_6		-0.301	-0.046	-0.250	-0.014	-0.027	0.120	0.128	0.158
	X_1		-0.035		-0.021		-0.160		-0.048
	X_2		-0.025		-0.035		-0.005		0.002
	X_3		0.014		0.003		-0.009		-0.003
	X_4		-0.025		-0.041		-0.039		-0.043
	X_5		-0.184		-0.142		0.066		0.062
P_{YU}			0.235		0.298		0.336		0.409
R^2			0.945		0.912		0.887		0.833

^{Pr}: Parametric path coefficients ^{Prs}: Nonparametric path coefficients

and 1000-seed weight (X_6) has lesser direct and indirect effect on seed yield. Mozaffari and Asadi (2006) found that head diameter had little indirect effect via plant height, and number of heads/plant, and oil content on seed yield. Plant height (X_1) has positive little direct and lesser indirect effect on seed yield.

Number of seeds in capitilum, 100 seeds weight, capitilum diameter, capitilum weight and days to 50% flowering were important selection traits for yield in drought stress condition and capitilum weight has highest negative direct effect, the others have highest positive direct effects on yield in safflower (Mozaffari and Asadi 2006). Pahlavani (2005) explained that seed yield may affect oil and protein content in safflower. Mokhtassi *et al*. (2006) reported that seed yield was significantly correlated with total biomass and number of days to the beginning of branching which variation in these two traits could explain 94% of the total variation in seed yield according to stepwise multiple regression analysis and were most important selection criteria for seed yield in safflower. Omidi (2000) revealed that number of heads/plant and biomass affected seed yield and to be linear relationship between oil yield and seed yield in spring safflower.

Consequently, oil content, number of seed/head and plant height are important selection characters for seed yield and have positive correlation with seed yield and these traits have positive direct effect on seed yield in safflower yield in drought condition. It is determined that the parametric and nonparametric path coefficients showed the similar results in both years, but parametric path analysis is preferred to nonparametric path analysis.

Because, calculated residual effects for parametric path analysis are lower than that for nonparametric path analysis. Also, determination coefficients for parametric path analysis are higher than that for nonparametric path analysis in this study. However, it seems that the relation measures (correlations) between seed yield and yield components are misleading, as the path analysis does not taken into consideration. Because, when non-significant and negative correlations are take into consideration together with the effects of other factors may be significant in path analysis. Owing to this trait, the path analysis is very important to investigate the factors which have an effect on yield. From the result, when the relation between seed yield and yield components were investigated, the path analysis revealed the more potent and real results than the coefficient measures (r, r_s).

REFERENCES

Dofing, S. M., and C. W. Knight (1992). Alternative model for path analysis of small-grain yield. *Crop Science,* 32, 487-489.

Kang, M. S., J. D. Miller and P. Y. P. Tai (1983) Genetic and phenotypic path analysis and heritability in sugarcane. *Crop Science*, 23(4), 643-647.

Kang, M. S. (1992) Letters to the editor. *Agronomy J*., 84, 917-918.

Mahasi, M. J., R. S. Pathak, F. N. Wachira, T. C. Riungu, M. G. Kinyua and J. W. Kamundia (2006). Correlations and path coefficient analysis in exotic safflower (*Carthamus tinctorius L.*) genotypes tested in the arid and semi arid lands (Asals) of Kenya. *Asian Journal of Plant Sciences*, 5 (6), 1035-1038.

Mokhtassi B., A., G. Akbari, M. Mirhadi, E. Zand and S. Soufizadeh (2006). Path analysis of the relationships between seed yield and some morphological and phonological traits in safflower (*Carthamus tinctorius L*.). *Euphytica*, 148 (3), 261-268.

Mozaffari, K. and A. Asadi (2006). Relationships among traits using correlation, principal components and path analysis in safflower mutants sown in irrigated and drought stress condition. *Asian Journal of Plant Sciences*, 5 (6), 977-983.

Neter, J., W. Wasserman and M. H. Kutner (1989). Applied Linear Regression Models. (2nd ed.) Boston, 284-294.

Omidi T. A. H. (2000). Correlation between traits and path analysis for grain and oil yield in spring safflower. *Sesame and Safflower Newsletter*, 15, 78-83.

Pahlavani, M. H. (2005). Some technological and morphological characteristics of safflower (*Carthamus tinctorius L*.) from Iran. *Asian Journal of Plant Sciences*, 4 (3), 234-237.

Tabachnick, B. G. and L. S. Fidell (1996). *Using Multivariate Statistics*. (4th ed), California State University, Northridge, 709-712.

Topal, M. and N. Esenbuga (2001). A study on direct and indirect effects of some factors on weaning weight of awassi lambs. Turkish J. Vet. and Anim. Sci., 25 (3), 377-382.

Tunçturk, M. and V. Ciftci (2004). Relationships among traits using correlation and path coefficients analysis in safflower (*Carthamus tinctorius L.*) sown different fertilization levels and row spacing. *Asian Journal of Plant Sciences*, 3 (6), 683-686.

Williams, W. A., M. B. Jones and M. W. Demment (1990). A concise table for path analysis statistics. *Agron. J.* 82, 1022-1024.

Wright, S. (1960a). Path coefficients and path regressions: alternative or complementary concepts?. *Biometrics*, 16 (2), 189-202.

Wright, S. (1960b). The treatment of reciprocal interaction, with or without lag, in path analysis. *Biometrics*, 16 (2), 423-444.