

# Association Analysis and Structured Equation Modelling in Groundnut (*Arachis hypogaea* L.)

## Abstract

Groundnut belongs to “bean” or legume family. It has a rich nutty flavour, crunchy texture, sweet taste and has an oil with relatively extended shelf life. India is one among the top countries producing groundnut in the world ranking. The breeding programmes of groundnut aim at developing early maturing and high yielding varieties. In crop improvement, knowledge on influence of various plant characters on yield is very important. Correlation and Path analysis is a tool which provides information on magnitude and intensity of association among yield components and yield. The present study analyses effects of the yield components on yield in the segregating population of groundnut. The traits days to maturity, number of matured pods, kernel yield, hundred pod and kernel weight are positively correlated with pod yield while the vegetative traits height of main axis and number of branches are negatively correlated with pod yield. Path analysis revealed that days to accumulation of 25 flowers, days to maturity, number of matured pods, kernel yield, hundred pod and kernel weight exerted low to moderate direct effects on pod yield. Structural Equation Model (SEM) is a powerful statistical technique for Path analysis that allows for more complex models than multiple regressions with its single dependent variable. In groundnut segregating population, structural equation modelling splits the yield component into traits before harvest and after harvest. Though both the variables are correlated, the traits before harvest have higher influence on pod yield.

Keywords: Legume, Correlation, Path analysis, Structural equation modelling

## Introduction

Groundnut (*Arachis hypogaea* L.) is an important legume crop mainly cultivated for edible oil widely across the world. The crop is a native of South America, which is cultivated in tropical, sub-tropical, and warm temperate regions of the world. The seeds are used for human consumption as confectionaries and as rich source of oil and protein (Weiss, 1983). Additionally, the plant residues are also used as fodder for cattle in many regions of the world. In Africa and Asia, groundnut is intercropped between sorghum, maize, and soybean and between mature coconut trees in a few areas (Stalker *et al.* 1987). The demand of short duration varieties of groundnut with high yielding capacity suitable for the expanding crop

rotations in our country is increasing day by day. To bridge the requirement of groundnut cultivating farmers, improvement of such traits is essential which could encourage a greater number of groundnut cultivars.

Yield is a complex character that is governed by polygenes and is affected by many genetic and non-genetic factors. It is highly influenced by the environment and has a very low heritability. Direct selection for yield shows less efficiency in improving productivity of the crop. The estimate of degree of association between yield and its attributing traits and also among yield components is an important criterion for selection. In groundnut the correlation study is much more helpful compared to other crops as the pods are formed underground. In such a condition the external plant characters with strong association with yield may help to effect the proper selection. (Kiranmai *et al.* 2016). Correlation coefficient is a statistical measure which is used to find out the degree (strength) and direction of relationship between two or more variables (Falconer, 1981). The correlation studies taken alone are often misleading and the actual dependence of grain yield on the correlated yield component characters needs confirmation, which can easily be untangled and unravelled by path coefficient analysis (Wright 1921). The path coefficient analysis allows partitioning of correlation coefficient into direct and indirect contributions (effects) of various traits towards dependent variable and thus helps in assessing the cause-effect relationship as well as effective selection. Structural equation modelling (SEM) is a combination of regression or path analysis and factor analysis. SEM is a theoretical construct that is represented by latent factors built by regression or path coefficients between the factors. SEM is visualized by a graphical path diagram (Wright, 1921).

The current investigation employs correlation and path analysis to evaluate the association and effect of yield contributing components on pod yield in the segregating population of groundnut. Further, the estimates of direct and indirect effects from path analysis were subjected to structural equation modelling to visualise them graphically. The results will be used for practising selection in groundnut for the enhancement of various traits.

## **Materials and Methods**

The experimental material consisted of the varieties ICGV07222 and Chico that were obtained from Department of oilseeds, Coimbatore and ICRISAT, Hyderabad. The line ICGV07222 is a Spanish bunch variety developed from ICRISAT, Hyderabad with a maturity duration of 95 to 122 days. The Chico variety is also a Spanish bunch a registered in USA

that matures in 75 days. The varieties ICGV07222 was taken as female parent and Chico as male parent were hybridized by hand emasculation and dusting in the fields at the Department of Oilseeds, Centre of Plant Breeding and Genetic, Tamil Nadu Agricultural University, Coimbatore during 2018. Further, true F<sub>1</sub>s were identified and left undisturbed for selfing to yield the F<sub>2</sub> seeds. The F<sub>2</sub> plants of the cross ICGV07222 × Chico were raised during *rabi* 2019, at the Department of Oilseeds, Centre of Plant Breeding and Genetic, Tamil Nadu Agricultural University, Coimbatore, each in a row of 3.0 m length with a spacing of 30 × 10 cm. The morphological traits recorded on the F<sub>2</sub> plants include days to accumulation of 25 flowers, height of main axis, number of branches, days to maturity, number of matured pods, shelling percentage, kernel yield per plant, hundred pod weight, hundred kernel weight and pod yield per plant. The correlation analysis, path analysis and structured equation modelling were performed using the Statistical package RStudio (RStudio Team, 2023). The significance of genotypic correlation coefficient was tested by referring to the standard table given by Snedecor and Cochran (1967). The direct and indirect effects were classified based on the scale given by Lenka and Misra (1973).

## Results and Discussion

Yield is a complex trait that depends on many independent contributing traits. In such cases the information on type of association between yield and its components themselves greatly help in assessing the contribution of different components towards yield (Ravi Kumar *et al.* 2012). The path coefficient allows partitioning of path coefficient analysis into direct and indirect contributions (effects) of numerous characters towards dependent variable and thus helps in assessing the cause-effect relationship as well as effective selection.

Simple correlation is a measure of intensity of association and the interdependency between variables. In selection for pod and kernel yield, correlation studies provide an insight about its component traits and improve the precision of selection.

The correlation estimates are represented in Table 1. and represented as correlogram in Fig. 1. The trait pod yield per plant exhibited positively significant correlation with days to maturity, number of matured pods, kernel yield per plant, hundred pods and kernel weight indicating the influence of these traits on pod yield. On par results for days to maturity has been reported by Alam *et al.* (1985), Dhaliwal *et al.* (2010), Vange and Maga (2014) and Sawargaonkar *et al.* (2010). The present findings for hundred pod and kernel weight are contradicted by Karikari and Tadore (2004) who reported negative correlation between seed weight and yield in groundnut. However, similar positive correlation was reported by Maunde *et al.* (2015) in groundnut.

Thus, increase in the maturity duration, number of pods and pod and kernel weight would bring about an increase of pod yield and are useful traits as selection criteria for the improvement of the trait.

Pod yield per plant is also negatively correlated with height of main axis and number of branches revealing excessive vegetative growth leads to decrease in the yield. This result is supported by Wigglesworth (1996) who reported that the negative association among the vegetative components could result from competition for ambient resources such as nutrients, light, moisture, etc.

In regarding to the inter associations, the trait days to accumulation of 25 flowers is found to be negatively associated to height of main axis and number of branches and positively associated with days to maturity. In a similar way, height of main axis with days to maturity, number of matured pods and hundred pod weight.

Daysto maturity are significantly related to kernel yield and hundred pods weight in a positive way implying longer the crop duration higher is the yield .Positively significant inter-correlation is also recorded among number of branches with days to maturity and number of matured pods and number of matured pods with shelling percentage. Similar inter association of yield contributing traits in groundnut is earlier reported by Kiranmai *et al.* (2016), Ramakrishnan *et al.*(2017) and Rajarathinamet *al.* (2017).

The correlation between any two characters would not give a complete picture of a complex situation like yield of plant which is jointly determined by a number of traits either directly or indirectly. In such situations, path coefficient analysis would be useful, as it permits the separation of direct effect from indirect effects through other related traits. A path coefficient is simply a standardized partial regression coefficient and it measures the direct influence of one trait upon another. Ultimately, we can reduce the time looking for a greater number of component traits by restricting selection to one or few important traits (Dewey and Lu, 1959). In the present study, direct and indirect effects of yield contributing components on pod yield per plant were worked out and are represented in Table 2.

The component of residual effect of path analysis was estimated as 0.657 (Fig. 3). The lower values of residual effect indicated that the characters chosen for path analysis were adequate and appropriate.

All the traits under study show direct effects on the plant yield and they exhibit only negligible indirect effects through other traits on the plant yield. A similar finding of Khan *et al.* (2022) in groundnut revealing negligible effects of yield contributing traits on yield.

The trait days to accumulation of 25 flowers, hundred pods and kernel weight recorded low direct effects on pod yield. Makanda *et al.* (2009), also reported direct effect of seed weight on pod yield in groundnut. The direct effect was negative for height of main axis, number of branches and shelling percentage with an intensity of high, moderate and low respectively. Moderate direct effects were observed for days to maturity and kernel yield. These findings are in line with that reported by Korat *et al.* (2010), Khanpara *et al.* (2010) and Dhaliwal *et al.* (2010).

In the present investigation, it is inferred that, the trait days to maturity, kernel yield, hundred pod and kernel weight are the major yield contributing traits which have the high positive direct effects indicating the scope of improving pod yield per plant by selection of these traits in the segregating materials.

Further, to investigate the precise association and effects of the yield contributing traits on total yield Structural Equation modelling is taken up. SEM is a multivariate statistical analysis for generalizing the theoretical models and path diagrams from causal relationships and effect values among the components (Cole and Preacher, 2014).

In the present investigation, the total independent variables are structured into traits before harvest and traits after harvest (Fig. 4). The set model recorded non-significant chi-square value indicating that it fits into the structured equation truly (Table 3).

The first latent variable, traits before harvest (BH) includes days to accumulation of 25 flowers, height of main axis, number of branches and days to maturity. Among these traits, height of main axis and number of branches recorded negative significance and days to maturity recorded positive significance on yield through days to accumulation of 25 flowers. The traits number of matured pods, shelling percentage, kernel yield, hundred pods and kernel weight are included in the second latent variable of the structure. The components kernel yield and hundred pod weight showed positively significant effect on yield through number of matured pods.

Both the latent variables recorded significantly positive regression conferring their linear relationship on pod yield. The two latent variables effect the yield trait together as indicated by the positively significant covariance between them.

When considering the estimates of variances, all the traits recorded significant variations while the traits before harvest have significant variances. This is interpreted as the traits before harvest play an important role in deciding the total yield per plant (Table 4).

Thus, the Structure equation model, splits the components of yield into two latent variables that fit truly with the model set. The significance of the trait variables before harvest indicated that improvement in these trait plays a vital role in the enhancement of yield.

The present investigation is an attempt to predict the complexity of maturity in groundnut that is strongly correlated to yield in a positive way (Fig. 2) while breeding objectives demand reducing the maturity duration without any decline in yield. Bailey and Bear (1973) have reported that early onset of flowering and the accumulation of a given number of flowers (about 30) are important constituent of early maturity in groundnut. In the same study they have also reported that high proportion of first 25 flowers develop into mature pods. In the present investigation we also observe that the trait days to accumulation of 25 flowers is not significantly correlated to pod yield. Thus, shortening the duration of flowering and accumulation of 25 flowers will eventually lead to reduction in maturity duration of groundnut. Hence, the trait - days to accumulation of 25 flowers is the major selection factor for early maturity in groundnut.

### **Conclusion**

Groundnut is an unpredictable legume. The improvement of the crop requires detailed understanding of the yield components. Association studies reveals the estimates of connotation of various traits on pod yield. The traits days to maturity, number of matured pods, hundred kernel and hundred pod weight show a strong correlation with yield and could be taken as selection criteria for improving the yield. When it comes to early maturity in groundnut, it is negatively correlated with yield. In such case, days to accumulation of 25 flowers that have least effect on yield could be used as selection criteria.

### **References**

- AlamM., RahmanA. &KhairA. (1985). Genetic variability and charachter association in groundnut [*Arachis hypogaea* L. of Bangladesh]*Bangladesh Journal of Agriculture (Bangladesh)*.
- BaileyW.K. &BearJ.E.(1973). Components of earliness of maturity in peanuts, *Arachis hypogaea* L. J. Amer. Peanut Res. & Ed. Assoc. 5: 32-39.
- ColeD.A. & Preacher K.J.(2014). Manifest variable path analysis: Potentially serious and misleading consequences due to uncorrected measurement error. *Psychological Methods*, 19(2), 300– 315.<https://doi.org/10.1037/a003380>

- Dewey D.R. & Lu K. (1959). A Correlation and Path-Coefficient Analysis of Components of Crested Wheatgrass Seed Production Agronomy journal, 51(9): 515-518. <https://doi.org/10.2134/agronj1959.00021962005100090002x>
- Dhaliwal G., Nagda A. & Mittal V. (2010). Intertrait associations and path analysis studies in groundnut (*Arachis hypogaea* L.) Crop Improvement, 37(1), 57-60.
- Falconer D.S. (1981). Introduction to quantitative genetics New York. British Library, 2nd Edition, Oliver and Boyd, Edinburg, London, 316.
- Khan M.M., Rafii M.Y., Ramlee S.I., Jusoh M. & Al Mamun M. (2022). Path-coefficient and correlation analysis in Bambara groundnut (*Vigna subterranea* [L.] Verdc.) accessions over environments. Scientific reports, 12(1):245. <https://doi.org/10.1038/s41598-021-03692-z>
- Karikari S.K. & Tabona T.T. (2004). Constitutive traits and selective indices of Bambara groundnut (*Vigna subterranea* (L) Verdc.) landraces for drought tolerance under Botswana conditions. Phys. Chem. Earth A/B/C 29(15-18), 1029-1034. <https://doi.org/10.1016/J.PCE.2004.08.002>
- Khanpara M.D., Shinde P.P., Jivani L.L., Vachhani J.H. & Kachhadia V.H. (2010). Character association and path coefficient analysis in groundnut (*Arachis hypogaea* L.) Pl. Archives, 10(2): 695-698.
- Kiranmai, M.S., Venkataravana P. & Push H. (2016). Correlation and path analysis studies in groundnut under different environment Legume Research-An International Journal, 39(6): 1048-1050.
- Korat V., Pithia M., Savaliya J., Pansuriya A. & Sodavadiya P. (2010). Studies on characters association and path analysis for seed yield and its components in groundnut (*Arachis hypogaea* L.) Legume Research, 33(3): 211-216.
- Lenka D. & Misra B. (1973). Path-coefficient analysis of yield in rice varieties. Indian journal of agricultural sciences, 43(4): 376-379.
- Makanda I., Tongoona P., Madamba R., Icishahayo D. & Derera J. (2009). Path coefficient analysis of bambara groundnut pod yield components at four planting dates. Res. J. Agric. Biol. Sci. 5(3), 287-292.
- Maunde S.M., Tanimu B. & Mahmud M. (2015). Correlation and path coefficient analysis of yield characters of bambara (*Vigna subterranea* L. Verdc.). Afr. J. Environ. Sci. Technol. 9(1), 12-15. <https://doi.org/10.5897/AJEST12.102>
- Rajarathinam P., Narayana M., Alagirisamy M. & Moshin I.S. (2017). Variability, correlation and path coefficient analysis in groundnut (*Arachis hypogaea* L.)

- S. Vignesh and A. Philip Arokiadoss (ed.) Statistical Approaches on Multidisciplinary Research, Volume I, Surragh Publishers, India. <https://doi.org/10.5281/zenodo.262967>
- Ramakrishnan P., Manivannan N., Mothilal A. & Mahalingam L. (2017). Correlation studies in back cross derived population for foliar disease resistance in groundnut (*Arachis hypogaea* L.). Int. J. Curr. Microbiol. App. Sci, 6(5): 266-272. <https://doi.org/10.20546/ijcmas.2017.605.031>
- Ravi Kumar D., Sekhar M. R., Reddy K. & Ismail S. (2012). Character association and path analysis in groundnut (*Arachis hypogaea* L.). International Journal of Applied Biology and Pharmaceutical Technology, 3(1): 385-387.
- RStudio Team. (2023). RStudio Desktop IDE (Version 2023.06.0-421) PBC. <https://posit.co/products/open-source/rstudio/>
- Sawargaonkar S., Giri R. & Hudge B. (2010). Character association and path analysis of yield component traits and late Leaf spot disease traits in groundnut (*Arachis hypogaea* L.). Agricultural Science Digest, 30(2): 115-119.
- Snedecor G. W. & Cochran W. (1967). Statistical methods 6th edition. The Iowa State University.
- Stalker H. & Moss J. (1987). Speciation, cytogenetics, and utilization of *Arachis* species. Advances in Agronomy, 41: 1-40. [https://doi.org/10.1016/S0065-2113\(08\)60801-9](https://doi.org/10.1016/S0065-2113(08)60801-9)
- Vange T. & Maga T. J. (2014). Genetic characteristics and path coefficient analysis in ten groundnut varieties (*Arachis hypogaea* L.) evaluated in the Guinea Savannah agro-ecological zone. African Journal of Agricultural Research, 9(25): 1932-1937. <https://doi.org/10.5897/AJAR2013.7340>
- Weiss, E. A. (1983). Oilseed crops. Longman Group.
- Wigglesworth D. J. (1996). The potential for genetic improvement of bambara groundnut (*Vigna subterranea* L. Verdc.) in Botswana. In Proceedings of the International Symposium on Bambara groundnut (pp. 23-25).
- Wright S. (1921). Correlation and causation. Journal of Agricultural Research, 20, 557-585.

Diagonally up t value and (p - value)		DTF	HMA	NB	DM	NMP	SP	KY	HPW	HKW	PY
Diagonally down Genotypic correlation											
DTF	1.00	Inf (0.00)	-3.28 (0.00)	-2.76 (0.01)	0.43 (0.07)	1.09 (0.28)	0.82 (0.41)	3.24 (0.07)	5.77 (0.12)	-0.38 (0.70)	6.23 (0.14)
HMA	-0.23*	1.00	Inf (0.00)	-0.66 (0.51)	-2.45 (0.02)	-1.95 (0.05)	0.06 (0.95)	-0.41 (0.68)	-2.83 (0.01)	0.80 (0.43)	-7.02 (0.00)
NB	-0.19*	-0.05	1.00	Inf (0.00)	2.46 (0.01)	2.05 (0.04)	0.46 (0.64)	-1.23 (0.22)	-0.95 (0.34)	-0.07 (0.94)	-3.00 (0.00)
DM	0.33*	-0.17*	0.17*	1.00	Inf	2.34	-0.14	3.36	3.81	1.34	6.70

Table1. Correlation table with genotypic correlation, calculated t value and p – value for yield contributing traits in Groundnut

					(0.00)	(0.02)	(0.89)	(0.00)	(0.00)	(0.18)	(0.00)
NMP	0.08	-0.14*	0.14*	0.16*	1.00	Inf	6.33	0.90	-0.19	1.65	2.18
						(0.00)	(0.00)	(0.37)	(0.85)	(0.1)	(0.03)
Model Test	User Model										
SP	0.06	0.00	0.03	-0.01	0.41*	1.00	Inf	-0.94	-1.09	0.94	-1.32
							(0.00)	120.365 <sup>NS</sup>	(0.28)	(0.35)	(0.19)
		Degrees of freedom									
KY	0.22	0.03	0.09	0.23*	0.06	-0.07	1.00	Inf	0.59	-0.69	5.68
		P-value (Chi-square)						(0.56)	(0.56)	(0.49)	(0.00)
		Regression of Latent Variables									
HPW	0.23	-0.20*	-0.07	0.26*	0.01	0.08	0.04	1.00	Inf	0.90	5.67
		Estimate		Std. Er.			z-Value		(0.00)	(0.37)	(0.00)
		Before Harvest (BH)									
HKW	0.03	0.06	-0.01	0.09	0.12	0.07	-0.05	0.06	1.00	Inf	2.11
									(0.00)	(0.00)	0.97
PY	0.21	-0.45*	-0.21*	0.43*	0.15*	-0.09	0.37*	0.37*	0.15*	1.00	Inf
											(0.00)

\*Significant at 5%

DTF – Days to accumulation of 25 flowers, HMA – Height of the main axis (cm), NB – No. of branches, DM – Days to maturity, NMP – No. of matured pods, PY – Pod yield per plant (g), KY – Kernel yield per plant (g), HPW – Hundred pods weight (g), HYW – Hundred kernels weight (g), SP – Shelling percentage (%)

Table 2. Path analysis of yield contributing traits showing Direct (the bold numbers on the diagonal) and indirect effects on yield in Groundnut

	DTF	HMA	NB	DM	NMP	SP	KY	HPW	HKW	PY
DTF	<b>0.186</b>	-0.009	0.042	0.009	0.006	-0.097	0.051	0.046	-0.024	0.21
HMA	-0.042	<b>-0.329</b>	0.010	-0.050	-0.012	-0.001	-0.007	-0.024	0.008	-0.446*
NB	-0.036	0.015	<b>-0.219</b>	0.050	0.014	-0.004	-0.020	-0.008	-0.001	-0.209*
DM	0.006	0.057	-0.038	<b>0.291</b>	0.015	0.001	0.053	0.032	0.014	0.430*
NMP	0.013	0.043	-0.032	0.048	<b>0.092</b>	-0.048	0.012	-0.002	0.017	0.141*
SP	0.011	-0.001	-0.007	-0.003	0.038	<b>-0.116</b>	-0.015	-0.009	0.010	-0.094
KY	0.042	0.010	0.019	0.068	0.005	0.008	<b>0.226</b>	0.005	-0.007	0.375*
HPW	0.071	0.065	0.015	0.076	-0.002	0.009	0.010	<b>0.121</b>	0.009	0.374*
HKW	-0.005	-0.019	0.001	0.028	0.011	-0.008	-0.011	0.008	<b>0.144</b>	0.148*

Residual effect = 0.657, \*Significant at 5%

DTF – Days to accumulation of 25 flowers, HMA – Height of the main axis (cm), NB – No. of branches, DM – Days to maturity, NMP – No. of matured pods, PY – Pod yield per plant (g), KY – Kernel yield per plant (g), HPW – Hundred pods weight (g), HYW – Hundred kernels weight (g), SP – Shelling percentage (%)

Table 3. Test statistic and estimate of set model in Structured Equation Modelling for Groundnut

\*Significant at 5%

DTF – Days to accumulation of 25 flowers, HMA – Height of the main axis (cm), NB – No. of branches, DM – Days to maturity, NMP – No. of matured pods, PY – Pod yield per plant (g), KY – Kernel yield per plant (g), HPW – Hundred pods weight (g), HYW – Hundred kernels weight (g), SP – Shelling percentage (%)

HMA	-0.972*	0.225	-4.322	0
NB	-0.804*	0.251	-3.205	0.001
DM	3.485*	0.716	4.867	0
	Estimate	Std.Er.	z-value	P(> z )
After Harvest (AH)				
NMP	1			
SP	-0.302	0.264	-1.142	0.253
KY	2.054*	0.824	2.493	0.013
HPW	10.232*	4.001	2.557	0.011
HKW	0.797	0.605	1.318	0.187

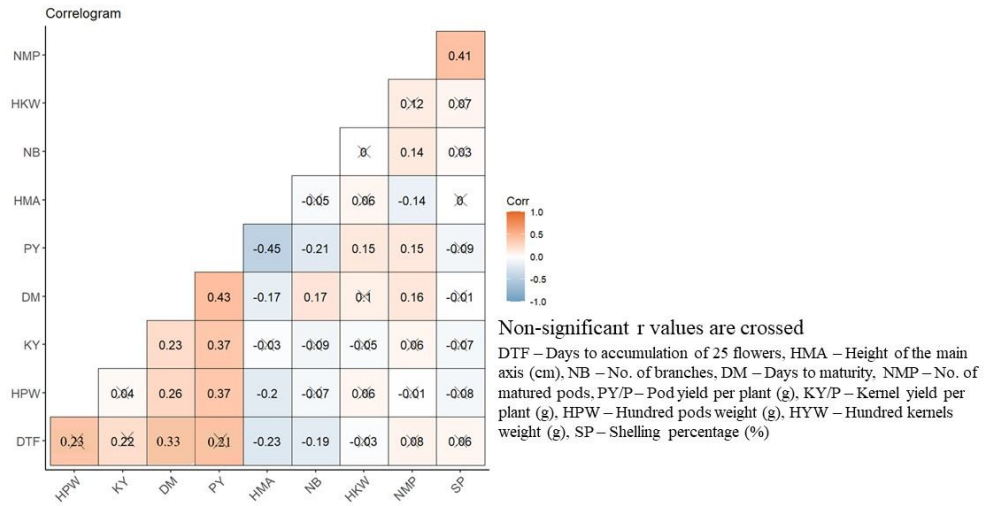
Table 4. Regression, Covariance and Variance of yield component traits in SEM of Groundnut

Regression					Covariance				
	Estimate	Std.Er.	z-value	P(> z )		Estimate	Std.Er.	z-value	P(> z )
PY ~(Pod Yield)					BH ~				
BH	8.917*	2.519	3.541	0	AH	1.055*	0.42	2.515	0.012
AH	9.326*	3.693	2.525	0.012					
Variances									
DTF	5.92*	0.605	9.791	0	KY	19.94*	2.166	9.207	0
HMA	6.95*	0.7	9.932	0	HPW	343.82*	40.477	8.494	0
NB	12.62*	1.251	10.092	0	HKW	45.53*	4.547	10.013	0
DM	51.47*	5.459	9.428	0	PY	26.39	29.468	0.896	0.37
NMP	19.17*	1.918	9.993	0	BH	0.59*	0.298	1.993	0.046
SP	9.34*	0.933	10.011	0	AH	0.17	0.269	0.647	0.518

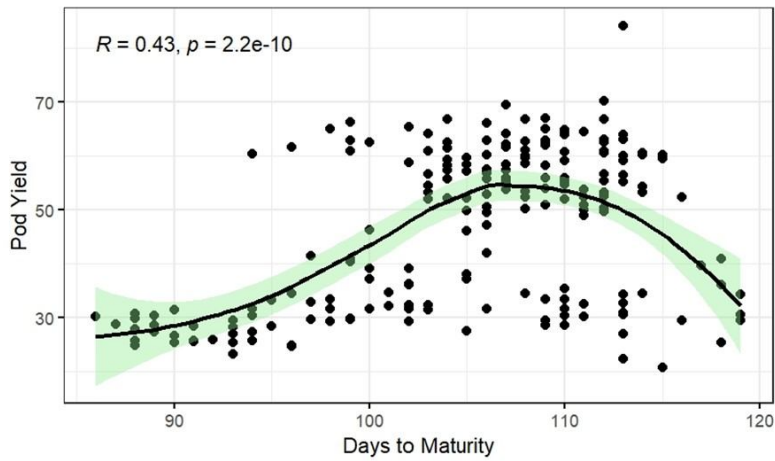
\*Significant at 5%

DTF – Days to accumulation of 25 flowers, HMA – Height of the main axis (cm), NB – No. of branches, DM – Days to maturity, NMP – No. of matured pods, PY – Pod yield per plant (g), KY – Kernel yield per plant (g), HPW – Hundred pods weight (g), HYW – Hundred kernels weight (g), SP – Shelling percentage (%)

**Fig.1 Correlogram between Yield component traits and Pod Yield in Groundnut**

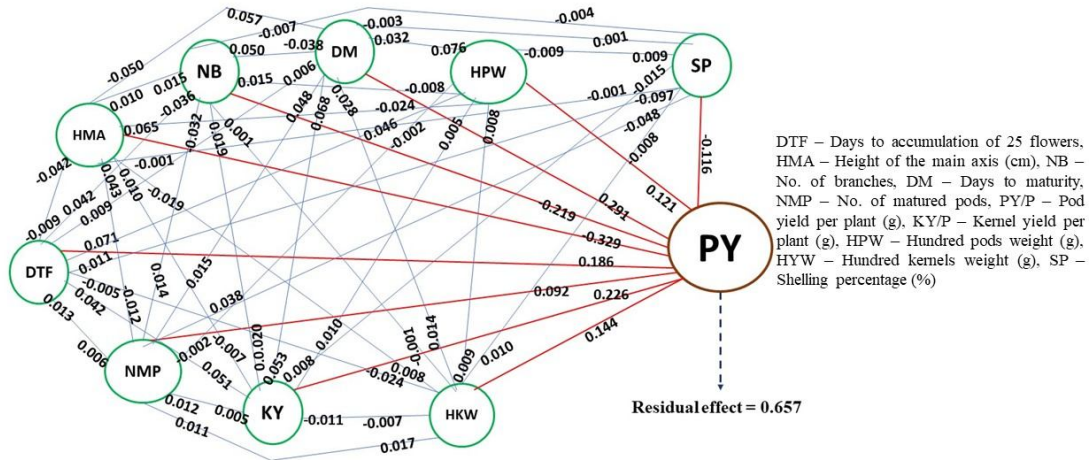


**Fig.2 Scatter plot depicting association between Days to maturity and Pod yield**



UN

**Fig.3 Path diagram for Yield component traits and Pod Yield in Groundnut**



**Fig.4 Structural Equation Modelling of Yield component traits and Pod Yield in Groundnut**

