RESULTS AND DISCUSSION

RESULTS AND DISCUSSION

In the traditional approach to research design we use a big sample enough to detect the smallest worthwhile effect. But hang on; we have wasted resources if the effect turns out to be large, because we need a smaller sample for larger effect. Sample size is a common practice in agronomic research. Important decision in planning experiments is the number of sampling per experimental unit to be used.

Correlation analysis

Sample size and simple correlation coefficients for 10 characters of maize are presented in Tables 1, 2, 3 and 4 for 2004, 2005, 2006 seasons and combined data, respectively.

Results in the first, second and third seasons and over all seasons indicate that increasing sample size increased the value of correlation coefficient. Also, the smallest sample size is about (40-50) for all seasons of experimentation which corresponds to correlation of highly significant at 1% level of significance for all characters. The obtained result clearly showed that the relationship between sample size and correlation coefficient was not significant at sample size less than 20 of all characters. When sample size was equal to 40 the value of correlation coefficient become significant, but this value was less than 0.5 which corresponds to correlation of negligible. On the other hand, when the value of correlation coefficient was more than 0.5 the correlation coefficient is dependable. Hence, it can not be only depend on the level of significance but, it must depend on the

Table 1: Effect of sample size on correlation coefficients and significance levels (p- value) between grain yield and yield factors in maize in 2004 season.

Sample size		Characters									
		No. of rows /ear	No. of Kernels /row	Ear Length (cm)	Ear Diameter (cm)	Weight of 100 Kernel	No. of kernels/ear	Plant height (cm)	No. of Leaves	Stem diameter	Leaf area (cm^2)
10	Correlation	0.252	0.383	0.240	0.374	0.349	0.496	0.316	0.373	0.342	0.259
	P value	0.482	0.275	0.505	0.286	0.323	0.145	0.374	0.288	0.333	0.470
20	Correlation	0.348	0.394	0.358	0.401	0.365	0.395	0.385	0.385	0.402	0.391
	P value	0.132	0.086	0.121	0.080	0.113	0.085	0.094	0.094	0.079	0.088
30	Correlation	0.447	0.412	0.489	0.424	0.433	0.488	0.415	0.417	0.416	0.431
	P value	0.013	0.024	0.006	0.020	0.017	0.006	0.023	0.022	0.022	0.017
40	Correlation	0.452	0.420	0.491	0.432	0.495	0.536	0.499	0.513	0.485	0.444
	P value	0.003	0.007	0.001	0.005	0.001	0.000	0.001	0.001	0.002	0.004
50	Correlation	0.471	0.450	0.505	0.480	0.515	0.594	0.512	0.555	0.561	0.508
	P value	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
100	Correlation	0.553	0.486	0.509	0.520	0.518	0.640	0.523	0.577	0.576	0.571
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
150	Correlation	0.568	0.567	0.517	0.634	0.574	0.644	0.571	0.591	0.599	0.583
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
200	Correlation	0.565	0.571	0.528	0.639	0.586	0.642	0.582	0.594	0.600	0.589
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
250	Correlation	0.570	0.577	0.532	0.642	0.590	0.645	0.584	0.598	0.605	0.583
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
300	Correlation	0.581	0.585	0.542	0.647	0.614	0.652	0.587	0.602	0.618	0.594
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
350	Correlation	0.599	0.592	0.612	0.663	0.649	0.655	0.602	0.626	0.637	0.606
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
400	Correlation	0.603	0.597	0.627	0.682	0.718	0.663	0.628	0.640	0.653	0.636
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
450	Correlation	0.653	0.601	0.635	0.687	0.780	0.695	0.668	0.673	0.666	0.682
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
540	Correlation	0.793	0.719	0.653	0.718	0.889	0.757	0.800	0.769	0.714	0.811
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 2: Effect of sample size on correlation coefficients and significance levels (p- value) between grain yield and yield factors in maize in 2005 season.

Sample size		Characters									
		No. of rows /ear	No. of Kernels /row	Ear Length (cm)	Ear Diameter (cm)	Weight of 100 Kernel	No. of kernels /ear	Plant height (cm)	No. of Leaves	Stem diameter	Leaf area (cm^2)
10	Correlation	0.245	0.334	0.208	0.229	0.245	0.279	0.313	0.288	0.321	0.268
	P value	0.496	0.345	0.564	0.524	0.496	0.435	0.379	0.420	0.373	0.454
20	Correlation	0.386	0.375	0.394	0.335	0.387	0.419	0.387	0.324	0.350	0.302
	P value	0.093	0.103	0.086	0.149	0.092	0.066	0.092	0.164	0.130	0.195
30	Correlation	0.396	0.395	0.419	0.344	0.402	0.469	0.403	0.365	0.377	0.362
	P value	0.030	0.031	0.021	0.063	0.028	0.009	0.027	0.047	0.040	0.049
40	Correlation	0.416	0.414	0.429	0.439	0.422	0.478	0.410	0.442	0.408	0.398
	P value	0.008	0.008	0.006	0.005	0.007	0.002	0.009	0.004	0.009	0.011
50	Correlation	0.423	0.428	0.449	0.460	0.459	0.481	0.425	0.462	0.412	0.444
	P value	0.002	0.002	0.001	0.001	0.001	0.000	0.002	0.001	0.003	0.001
100	Correlation	0.535	0.540	0.542	0.491	0.577	0.551	0.514	0.584	0.596	0.564
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
150	Correlation	0.651	0.642	0.574	0.535	0.643	0.657	0.691	0.691	0.680	0.624
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
200	Correlation	0.709	0.663	0.706	0.642	0.715	0.700	0.706	0.704	0.700	0.715
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
250	Correlation	0.751	0.729	0.720	0.720	0.760	0.791	0.776	0.726	0.773	0.724
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
300	Correlation	0.773	0.786	0.769	0.750	0.796	0.830	0.791	0.727	0.795	0.752
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
350	Correlation	0.789	0.790	0.784	0.769	0.809	0.834	0.799	0.745	0.798	0.772
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
400	Correlation	0.794	0.780	0.789	0.771	0.808	0.843	0.807	0.757	0.805	0.777
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
450	Correlation	0.816	0.812	0.817	0.795	0.834	0.865	0.825	0.766	0.821	0.809
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
540	Correlation	0.832	0.839	0.835	0.815	0.852	0.883	0.838	0.779	0.823	0.834
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 3 : Effect of sample size on correlation coefficients and significance levels (p- value) between grain yield and yield factors in maize in 2006 season.

Sample size		Characters									
		No. of rows /ear	No. of Kernels /row	Ear Length (cm)	Ear Diameter (cm)	Weight of 100 Kernel	No. of kernels/ear	Plant height (cm)	No. of Leaves	Stem diameter	Leaf area (cm^2)
10	Correlation	0.290	0.259	0.270	0.359	0.355	0.217	0.238	0.340	0.321	0.242
	P value	0.416	0.469	0.451	0.309	0.315	0.548	0.509	0.337	0.365	0.501
20	Correlation	0.335	0.297	0.341	0.418	0.456	0.294	0.373	0.471	0.393	0.247
	P value	0.149	0.204	0.141	0.067	0.043	0.208	0.105	0.036	0.086	0.293
30	Correlation	0.379	0.431	0.496	0.491	0.481	0.444	0.484	0.517	0.453	0.392
	P value	0.039	0.017	0.005	0.006	0.007	0.014	0.007	0.003	0.012	0.002
40	Correlation	0.488	0.463	0.545	0.511	0.490	0.475	0.521	0.526	0.508	0.417
	P value	0.001	0.003	0.000	0.001	0.001	0.002	0.001	0.000	0.001	0.007
50	Correlation	0.501	0.486	0.564	0.528	0.511	0.517	0.530	0.540	0.520	0.500
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
100	Correlation	0.535	0.513	0.621	0.602	0.540	0.583	0.567	0.557	0.584	0.546
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
150	Correlation	0.586	0.601	0.626	0.616	0.564	0.654	0.643	0.615	0.641	0.538
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
200	Correlation	0.611	0.616	0.637	0.652	0.583	0.680	0.661	0.636	0.658	0.558
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
250	Correlation	0.618	0.628	0.644	0.631	0.614	0.706	0.691	0.689	0.673	0.574
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
300	Correlation	0.621	0.635	0.646	0.641	0.623	0.706	0.696	0.696	0.681	0.582
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
350	Correlation	0.645	0.639	0.652	0.657	0.633	0.712	0.701	0.704	0.691	0.608
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
400	Correlation	0.652	0.637	0.652	0.670	0.646	0.713	0.703	0.709	0.697	0.647
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
450	Correlation	0.654	0.642	0.656	0.672	0.651	0.715	0.708	0.712	0.700	0.652
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
540	Correlation	0.714	0.695	0.714	0.705	0.682	0.728	0.745	0.751	0.778	0.689
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 4 : Effect of sample size on correlation coefficients and significance levels (p- value) between grain yield and yield factors in maize as a mean of three seasons.

Sample size		Characters									
		No. of rows /ear	No. of Kernels /row	Ear Length (cm)	Ear Diameter (cm)	Weight of 100 Kernels	No. of kernels/ear	Plant height (cm)	No. of Leaves	Stem diameter	Leaf area (cm^2)
10	Correlation	0.262	0.325	0.239	0.321	0.316	0.331	0.289	0.334	0.328	0.256
	P value	0.465	0.363	0.434	0.373	0.378	0.376	0.421	0.348	0.357	0.475
20	Correlation	0.356	0.355	0.364	0.385	0.376	0.369	0.382	0.393	0.382	0.313
	P value	0.125	0.131	0.102	0.099	0.003	0.120	0.097	0.098	0.098	0.192
30	Correlation	0.407	0.413	0.468	0.420	0.439	0.467	0.434	0.433	0.415	0.395
	P value	0.027	0.024	0.015	0.029	0.017	0.010	0.019	0.024	0.025	0.023
40	Correlation	0.452	0.432	0.488	0.461	0.469	0.496	0.477	0.494	0.467	0.420
	P value	0.004	0.006	0.004	0.004	0.003	0.001	0.004	0.002	0.004	0.007
50	Correlation	0.465	0.455	0.506	0.489	0.495	0.531	0.489	0.519	0.498	0.484
	P value	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
100	Correlation	0.541	0.540	0.557	0.538	0.545	0.591	0.535	0.573	0.585	0.560
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
150	Correlation	0.602	0.603	0.572	0.595	0.594	0.452	0.635	0.632	0.640	0.582
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
200	Correlation	0.628	0.617	0.624	0.644	0.628	0.674	0.650	0.645	0.653	0.621
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
250	Correlation	0.646	0.645	0.632	0.664	0.655	0.714	0.684	0.671	0.684	0.627
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
300	Correlation	0.658	0.667	0.652	0.679	0.678	0.729	0.691	0.673	0.698	0.643
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
350	Correlation	0.678	0.674	0.683	0.696	0.697	0.734	0.701	0.692	0.709	0.662
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
400	Correlation	0.683	0.671	0.689	0.708	0.724	0.740	0.713	0.702	0.718	0.687
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
450	Correlation	0.708	0.685	0.703	0.718	0.755	0.753	0.734	0.717	0.729	0.714
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
540	Correlation	0.779	0.751	0.734	0.746	0.808	0.789	0.794	0.766	0.772	0.778
	P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

value of correlation coefficient with the level of significance start with sample size of 100 plants for all characters. On the other hand, a general trend of correlation associated with large sample size. The greater sample size is about 540, which corresponds to real value of correlation.

It is possible to obtain the largest sample size which gave the great value of correlation coefficient for each character on the basis of any increase after this great value is little or no interest. It would seem to classify these characters at (Table 5) into the following:

1- Characters with low variation between plants needed 150 (plants) samples: Number of rows/ear, number of kernels/row, plant height, number of leaves / plant and stem diameter.

2- Characters with medium variation between plants needed 200(plants) samples: Ear length, ear diameter, weight of 100 kernels and leaf area.

3- Character with high variation between plants, needed 250 (plants) samples: Number of kernels/ ear.

In experiments, where several measurements are evaluated, the variability of the measurements had not the same magnitude. Thus, the number of samples will be determined by the most variable or most important measurement.

The change in correlation coefficient of ten characters with different sample size is shown graphically in Figures (2-11) as mean of three seasons.

Table 5: The converint sample size which gave the true value of correlation coefficient for each character overall three seasons.

Character	sample size	correlation coefficient
No. of rows/ear	150	0.602
No. of kernels/ row	150	0.603
Plant height	150	0.635
No. of leaves/plant	150	0.632
Stem diameter	150	0.640
Ear length	200	0.624
Ear diameter	200	0.624
Weight of 100 Kernels	200	0.628
Leaf area (cm^2)	200	0.621
No. of kernel/ear	250	0.714

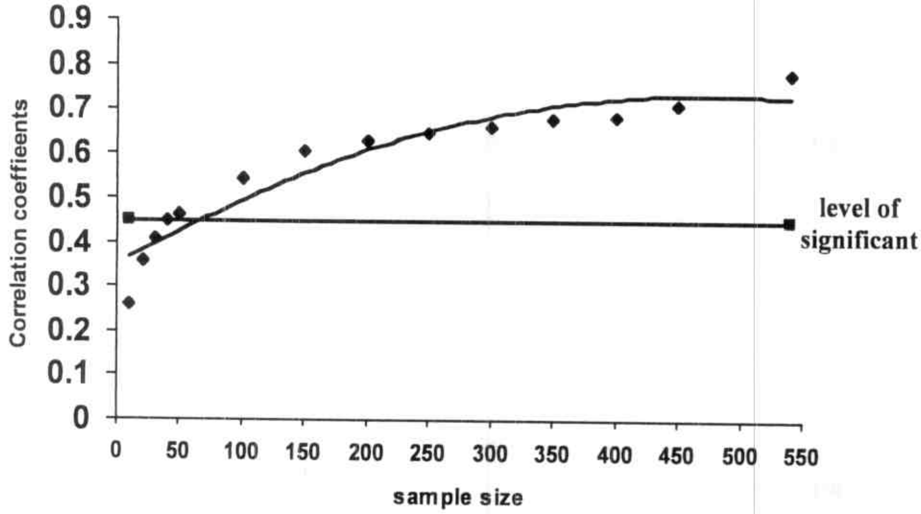


Fig (2): Effect of sample size on correlation coefficients of No. of rows/ ear in maize as mean of three seasons of experimentation

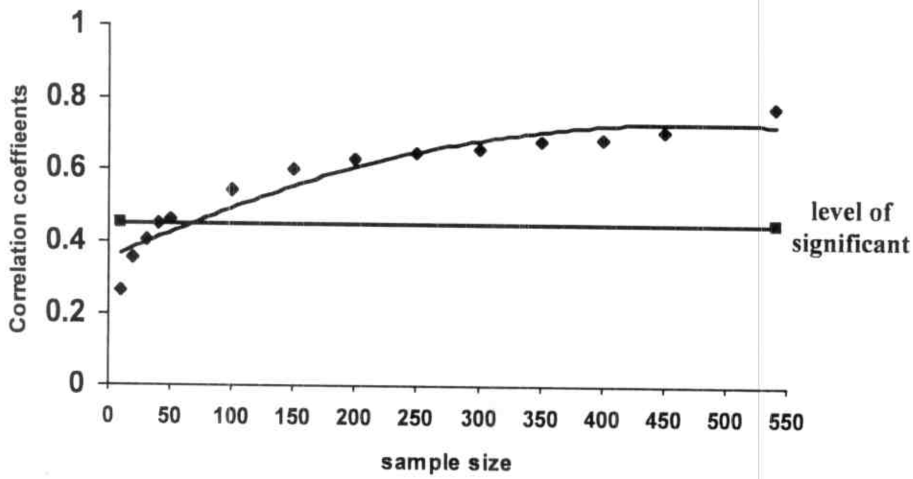


Fig (3): Effect of sample size on correlation coefficients of No. of kernels/row in maize as mean of three seasons of experimentation

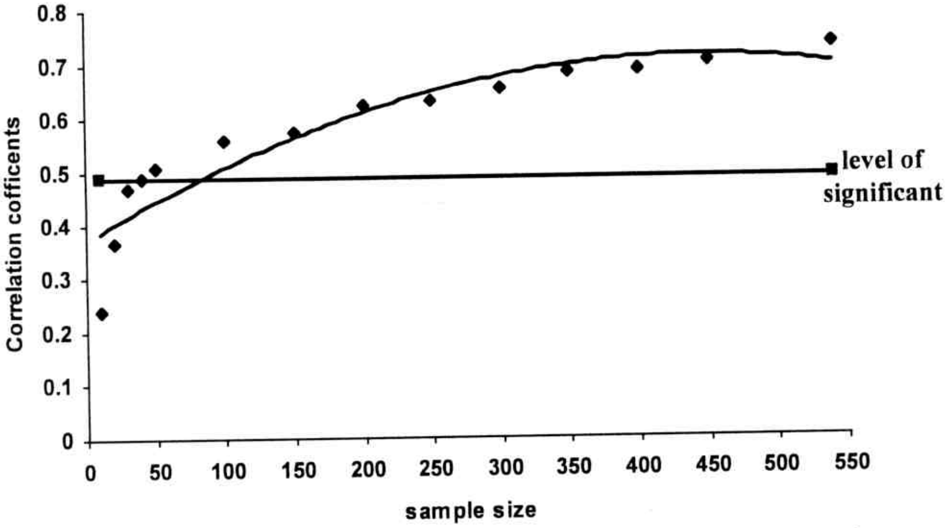


Fig (4): Effect of sample size on correlation coefficients of ear length (cm) in maize as mean of three seasons of experimentation

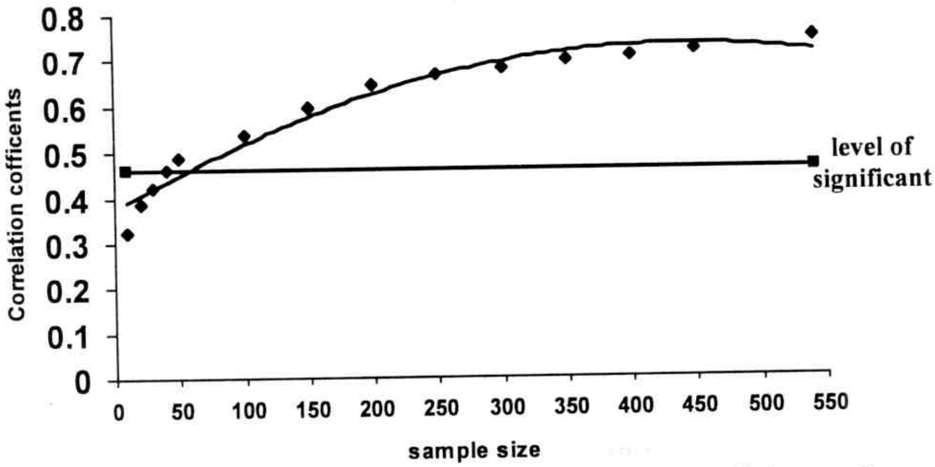


Fig (5): Effect of sample size on correlation coefficients of ear diameter (cm) in maize as mean of three seasons of experimentation

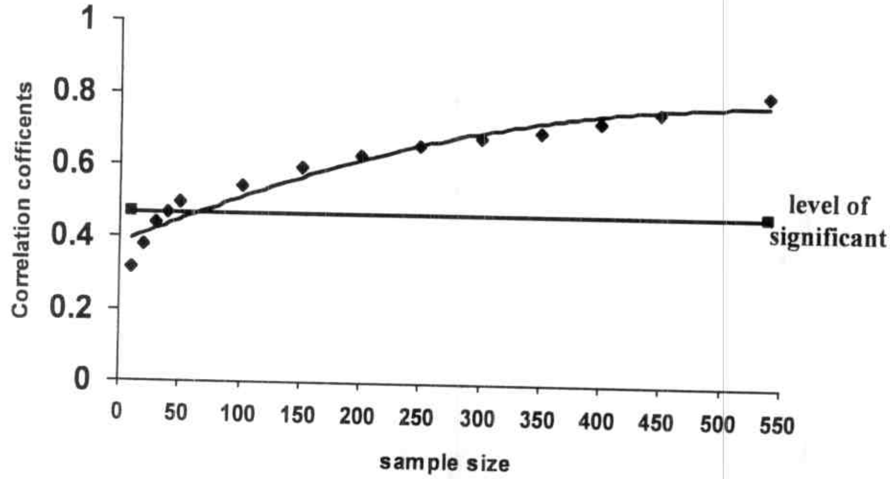


Fig (6): Effect of sample size on correlation coefficients of weight of 100 kernel in maize as mean of three seasons of experimentation

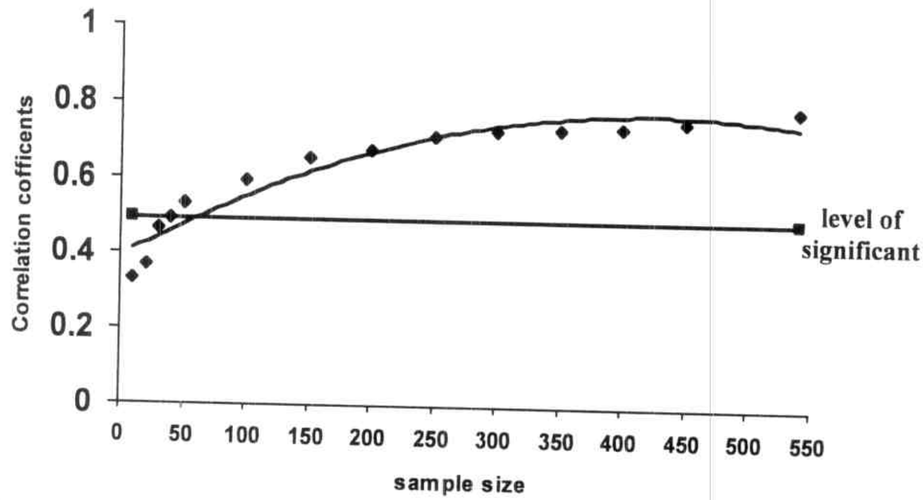


Fig (7): Effect of sample size on correlation coefficients of No .of kernels/ear in maize as mean of three seasons of experimentation

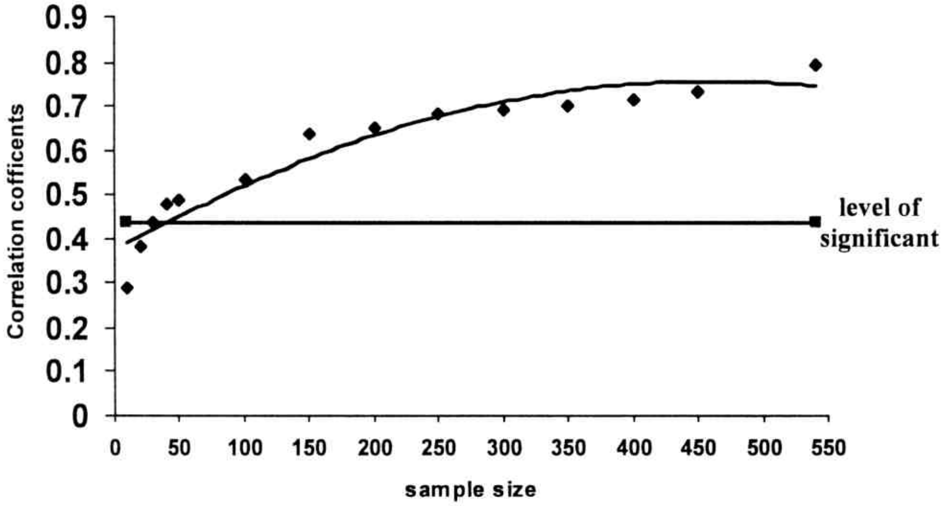


Fig (8): Effect of sample size on correlation coefficients of plant height (cm) in maize as mean of three seasons of experimentation

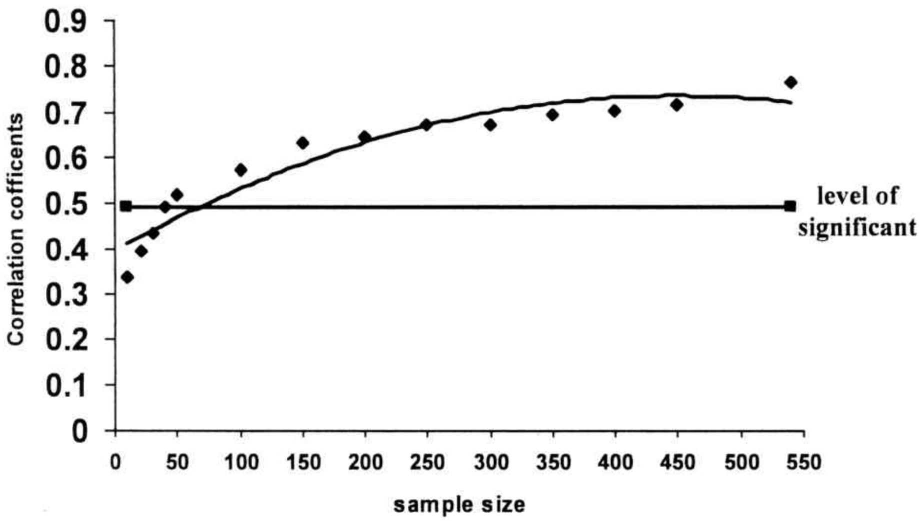


Fig (9): Effect of sample size on correlation coefficients of No. of leaves / plant in maize as mean of three seasons of experimentation

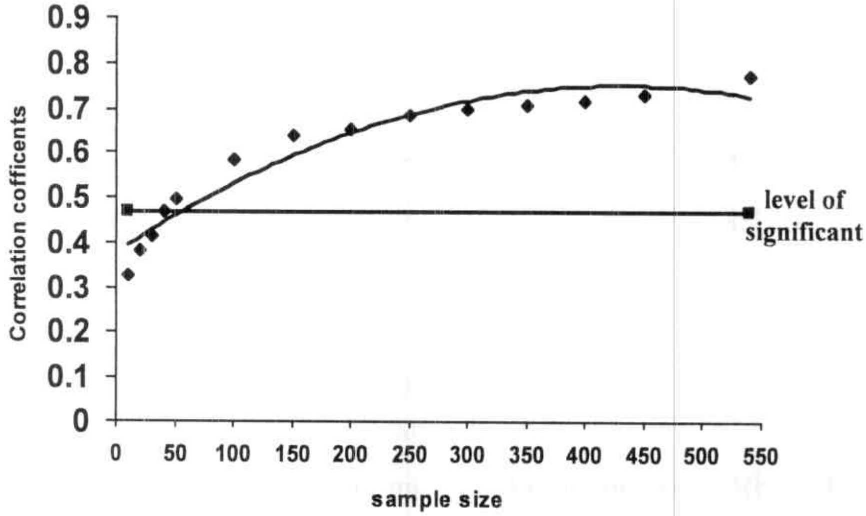


Fig (10): Effect of sample size on correlation coefficients of stem diameter in maize as mean of three seasons of experimentation

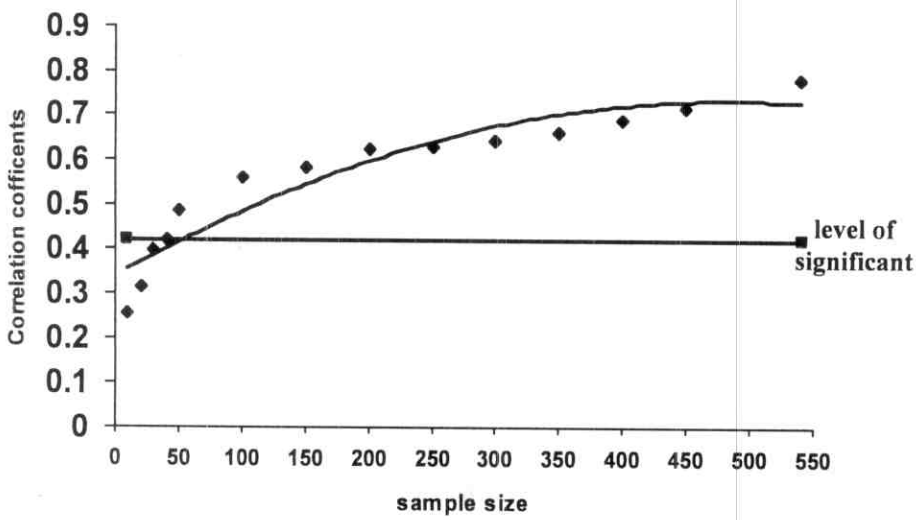


Fig (11): Effect of sample size on correlation coefficients of leaf area in maize as mean of three seasons of experimentation

The graph is just an adaptation of the figure to get the sample size corresponding to any correlation. Results clearly indicate that the relationship between sample size and correlation coefficient followed the quadratic model (second degree) for all yield factors in the three seasons. On the other wise, it could be concluded that quadratic model was the best of the response models tested for describing the relationship between sample size and correlation coefficient of yield factors in maize in the three seasons. Notice that there is positive relationship between sample size and correlation coefficient for all characters.

The present results are in good agreement with those reported by **Wolde – Tsadik (1980), Cohen (1988), Mohamed and Sedhom (1993), Petrovik *et al* (1994) and Kumar and Kumar (2001). Similar results were also reported by Comrey (1973), (1978), Gorsuch (1983), Guilford (1954), Hair *et al* (1979),Lindeman *et al* (1980) and Loo (1983)** who stated that the recommendation for a minimum sample size was 100 to200 observations.

Optimum sample size in correlation was 400. Data reported in Table (6) indicate the simple correlation coefficients matrix for 11 characters. Results indicate that the relationship between all possible pairs of the 11 traits were highly significant at 1% level of significance in all cases. In addition, a number of interesting relationships can be observed. Furthermore, the most important relationships to the maize breeder are between grain yield and number of kernels/ear (0.843), weight of 100 kernels (0.808), plant height (0.807), stem diameter (0.805), number of rows/ ear (0.794), ear length (0.789) and number of kernels/ row

Table 6. A matrix of simple correlation coefficients for 11 characters by using optimum sample size as mean of three seasons.

Characters	1	2	3	4	5	6	7	8	9	10	11
1- No .of Leaves	1.000										
2- Stem diameter	0.637	1.000									
3- Plant height (cm)	0.718	0.736	1.000								
4-No . of rows / ear	0.636	0.674	0.724	1.000							
5- No . kernels / row	0.673	0.635	0.662	0.735	1.000						
6- No . kernels/ ear	0.684	0.693	0.721	0.893	0.913	1.000					
7- Leaf area (cm^2)	0.596	0.699	0.720	0.646	0.636	0.692	1.000				
8 - Ear Length (cm)	0.670	0.673	0.743	0.707	0.601	0.699	0.696	1.000			
9-Weight of 100 kernels	0.712	0.600	0.743	0.664	0.689	0.707	0.600	0.727	1.000		
10- Ear diameter (cm)	0.703	0.636	0.686	0.692	0.620	0.695	0.568	0.627	0.710	1.000	
11- Grain yield	0.757	0.805	0.807	0.794	0.780	0.843	0.777	0.789	0.808	0.771	1.000

(0.780) which were highly significant positive correlation. This indicated that these characters had greatest influence on grain yield, coinciding with the results of **Mohamed and Sedhom (1993), Ashmawy (1994) and Mohamed (2004).**

Multiple linear regression analysis:

To test the validity of this procedure, variance inflation factor (VIF) must be calculated which is a potent aid for detecting the levels of multicollinearity. VIF's values are the diagonal elements of VIF value tell us the degree to which each independent variable is explained by the other independent variables. Thus, large VIF values denote high collinearity. Values exceeding 10.0 are considered likely to cause difficulty in coefficient estimation due to multicollinearity.

The results of variance inflation factor (VIF), coefficient of determination (R^2), adjusted coefficient of determination (adjusted R^2) and standard error (SE) of variables by using multiple linear regression as affected by sample sizes are presented in Tables 7, 8 and 9 for the first, second and third seasons of experimentation , respectively.

According to results in Tables 7, 8 and 9, it is clear that increasing sample size decreased VIF value of all variables .The smallest sample size is about (150- 200) in the three seasons, which corresponds to VIF value of less than 10 of all variables except number of kernels/ row and number of kernels/ ear in the first season, number of rows / ear, number of kernels/ row and number of kernels/ ear in the second season and number of kernels/ row and number of kernels/ ear in the third season which had the highest VIF. Furthermore, sample size of about

Table 7: Variance inflation factor (VIF), R^2 ,adjusted R^2 and SE of multiple regression as affected by sample sizes in relation between grain yield and yield factors in 2004 season .

Sample size	VIF										R^2	Adjust R^2	Std . Error
	No. of rows /ear	No. of Kernels /row	Ear Length (cm)	Ear Diameter (cm)	Weight of 100 kernels	No. of kernels /ear	Plant height (cm)	No. of Leaves	Stem diameter	Leaf area (cm^2)			
20	31.154	27.791	1.928	3.260	2.131	61.702	1.713	4.282	2.648	4.495	0.656	0.275	18.979
30	31.254	24.710	1.707	2.527	1.814	55.623	1.762	2.900	2.272	3.788	0.689	0.525	18.240
40	22.220	14.020	1.729	1.673	1.722	36.366	1.177	2.421	1.686	2.155	0.687	0.579	17.828
50	22.545	13.473	1.481	1.537	1.590	36.268	1.296	2.388	1.767	1.784	0.696	0.618	16.923
100	11.338	14.163	1.496	1.317	1.436	29.462	1.162	2.361	2.076	1.386	0.696	0.662	16.039
150	6.579	14.662	1.613	1.748	1.477	28.656	1.201	1.949	2.123	1.199	0.685	0.667	16.528
200	4.348	11.924	1.732	1.796	1.502	21.413	1.201	1.918	2.195	1.506	0.718	0.705	15.153
250	3.758	8.622	1.744	1.807	1.527	14.788	1.201	1.912	2.252	1.500	0.720	0.707	15.078
300	3.327	7.186	1.767	1.798	1.588	11.780	1.205	1.873	2.238	1.511	0.732	0.732	14.589
350	2.880	6.126	1.948	1.894	1.702	9.460	1.201	1.783	2.326	1.658	0.759	0.752	14.662
400	2.471	5.428	2.087	1.916	1.478	7.558	1.181	1.672	2.303	1.659	0.786	0.780	13.447
450	2.388	5.167	2.028	1.390	1.316	7.075	1.152	1.834	2.303	1.816	0.814	0.809	11.487
540	2.805	5.012	2.246	1.343	1.104	5.718	1.205	1.569	2.331	1.735	0.882	0.879	11.456

350 get a VIF of less than 10 denoting that is no collinearity between predicted variables of all variables in 2004. The other variables had a VIF value less than 10 of all sample sizes in 2005 and 2006 seasons. Multicollinearity is often strong enough to make unclear results **(Hoerl and Kennard, 1970).** On the other hand, the VIF's for ear length, ear diameter, weight of 100 kernels, plant height, number of leaves, stem diameter and leaf area are considerably less than 10 of all sample sizes in 2004 season, indicating that there is no collinearity or multicollinearity between these variables. Thus, evidence from the VIF calculation confirmed a part of the stability of the regression coefficients. Also, VIF of number of rows/ ear exceeded the value of 10 when sample size ranged from 20 to 100. Sample size about 150 gave VIF less than 10. Meanwhile, VIF of some characters were affected by sample size. Hence, the larger samples discarded the effect of collinearity or multicollinearity from the data, VIF was greater than 10 showing evidence confirming a part of the instability of the regression coefficients due to interrelation among the explanatory variables. Also, the characters with a high variation between plants needed a large sample. Also, the characters with a low variation between plants needed a small sample **(Nasr and Leilah, 1993a).** It can be noted that at small sample size, the R^2 value was greater than adjusted R^2 until sample size of =>300 plants the R^2 equal to adjusted R^2 owing to the large sample size discarded the effect of collinearity and multicollinearity. Larger samples are better than smaller samples (all other things being equal) because larger samples tend to minimize the probability of errors, maximizing the accuracy of population estimates and increase the

generalizability of the results **(Guadagnoli and Velicer, 1988)**, and decreased the VIF value (indicator the effect of the other predictor variables have on the variance of regression coefficient, it is directly related to the tolerance value ($VIF_i=1/R^2_i$)) and SE of all variables. This means that these procedures optimize the fit of the model of the given data yet no sample is perfectly reflective of the population. Thus, this over fitting can result in erroneous conclusions if models fit to one data set are applied to others. In multiple regression this manifests itself as inflated R^2 (Shrinkage) and miss- estimated variable regression coefficient **(Cohen and Cohen, 1983).** The relative contribution (R^2)for all yield factors increased by 22.6% of total variation in grain yield of sample size from 20 to 540 with standard error decreased by 7.52 in the first season . Also , the relative contribution (R^2) for all yield factors increased by 18.5%of sample size ranged from 20 to 540 with standard error (SE) decreased by 11.26 in the second season . Also, in the third season the relative contribution (R^2) for all yield factors increased by 21.9 of total variation in grain yield of sample size from 20 to 540 with standard error (SE) decreased by 5.1. Thus, worn that R^2,s cannot be compared between samples due to differences in the variances of the independents and dependent variables **(Achen, 1982)**. Thus, the most valid conclusion of sample size is that more is always better **(Pedhazur, 1997).** Always use adjusted R^2 when comparing models with different numbers of independents. Because of adjusted R^2 discarded the collinearity and multicollinearity of the data.

VIF, R^2, adjusted R^2 and SE of variables by using multiple linear regressions as affected by sample sizes as mean of all three seasons are shown in Table 10. Results of over all the three seasons indicate that increasing sample size increased the value of $R^2$and adjusted $R^2$and decreased the VIF value and SE of all variables. It is clear that increasing sample size decreased VIF value of all variables the smallest sample size is about 200, which corresponds to VIF value of all variables except number of kernels /row, and number of kernels / ear in three seasons, which had a highest VIF denoting there is high collinearity or multicollinearity between predicted variables.

On the other hand, other variables had a VIF value less than 10.0 of all sample sizes means that there is no collinearity. Meanwhile, VIF of some characters was affected by sample size. VIF value was greater than 10 evidence confirmed a part of the instability of the regression coefficients due to interrelation among the explanatory variables means there is collinearity or multicollinearity. Thus, the sample size should be adequate to provide a high probability of detecting as significant an effect size of given magnitude if each an effect actually exists. Also, a sample size that is too small will not allow to properly addressing research questions with proposed statistical analysis .Meanwhile, large sample need to make accurate statistical conclusions. Hence, larger sample is needed in order to obtain a more precise estimate.

The effect of sample size on variance inflation factor (VIF) for these characters is shown graphically in Figures (12-21) over three seasons.

Table 10 : Variance inflation factor (VIF), R^2 ,adjusted R^2 and SE of multiple regression as affected by sample sizes in relation between grain yield and yield factors as mean of three seasons .

Sample size	VIF										R^2	Adjust R^2	Std . Error
	No. of rows /ear	No. of Kernels /row	Ear Length (cm)	Ear Diameter (cm)	Weight of 100 kernels	No. of kernels /ear	Plant height (cm)	No. of Leaves	Stem diameter	Leaf area (cm^2)			
20	19.241	32.982	3.772	3.983	3.293	68.801	2.274	2.662	2.394	2.864	0.593	0.258	22.733
30	18.772	27.753	3.431	2.973	2.493	61.852	2.321	2.223	2.322	2.502	0.603	0.394	21.997
40	14.553	18.653	2.963	2.304	2.394	41.933	1.922	1.994	2.193	2.063	0.612	0.478	21.403
50	14.462	17.722	2.564	2.162	2.472	39.964	2.003	2.002	2.214	2.013	0.632	0.537	20.875
100	9.992	15.053	2.332	1.753	2.012	33.862	1.833	1.983	2.182	1.944	0.664	0.626	21.183
150	8.344	14.473	2.133	1.834	2.194	29.632	1.702	1.873	2.003	1.794	0.675	0.667	19.867
200	7.453	12.034	2.012	1.783	2.333	24.113	1.702	1.864	2.163	1.772	0.705	0.690	17.994
250	5.823	9.433	2.373	1.7523	2.071	18.803	1.614	1.811	2.114	1.762	0.721	0.709	17.957
300	5.313	8.924	2.154	1.702	2.172	17.284	1.592	1.812	2.044	1.593	0.726	0.721	16.783
350	5.152	8.643	2.114	1.792	2.384	16.572	1.613	1.833	2.013	1.574	0.737	0.729	16.706
400	4.144	6.403	2.154	1.793	2.373	11.683	1.743	1.833	1.963	1.563	0.748	0.742	15.877
450	4.234	6.123	1.933	1.643	2.072	11.023	1.454	1.764	2.113	1.523	0.759	0.754	14.783
540	3.323	4.584	1.913	1.693	1.973	6.194	1.504	1.794	2.014	1.483	0.804	0.800	14.763

In Figures (12-21), illustrated the relationship between sample size and VIF values of all characters as mean of three seasons. In Figures, nonlinear pattern is shown. Maximum sample size ranged from (350-550) observations for all characters is clear that the characters had a high variation between plants needed a large sample while the characters had a low variation between plants needed a small sample size. These results point out that number of leaves and stem diameter had a low variation between plants and other characters had high variation between plants. These results help in planning appropriate selection sample size for improving maize crop.

The results of the three seasons agree with each other. The results are in agreement with those of **Comrey and Lee (1992)** who suggest that the adequacy of sample size might be evaluated very roughly on the following scale :50- very poor: 100- poor: 200-fair: 300- good : 500- very good and 1000 or more – excellent. Also, the results are in agreement with those of **Surin (1992).** He found that, systematic sampling has higher efficiency in estimation than simple random sampling when using sample size between 20-24% of the total number of plants. They also agree with the findings of **Nasr and Leilah (1993b), Ashmawy (1994), Nasr (1998b), Nasr and El-Hady (1998), Maxwell (2000), Bezeau and Graves (2001), Kelly and Maxwell (2003), Delucchi, (2004) and Tan *et al* (2005).**

Optimum sample size gave more accuracy, small value of standard error and the most replicable result in multiple linear regression (N=400).

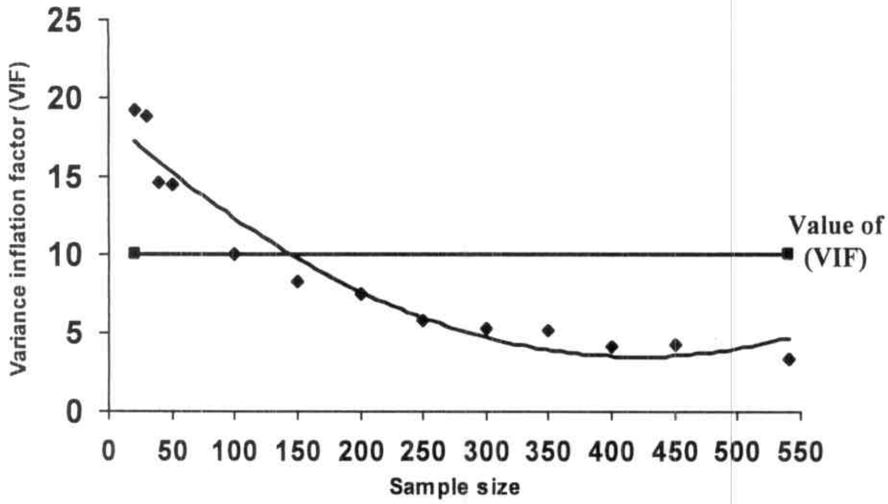


Fig (12): Effect of sample size on variance inflation factor (VIF) of No. of rows / ear in maize as mean of three seasons of experimentation.

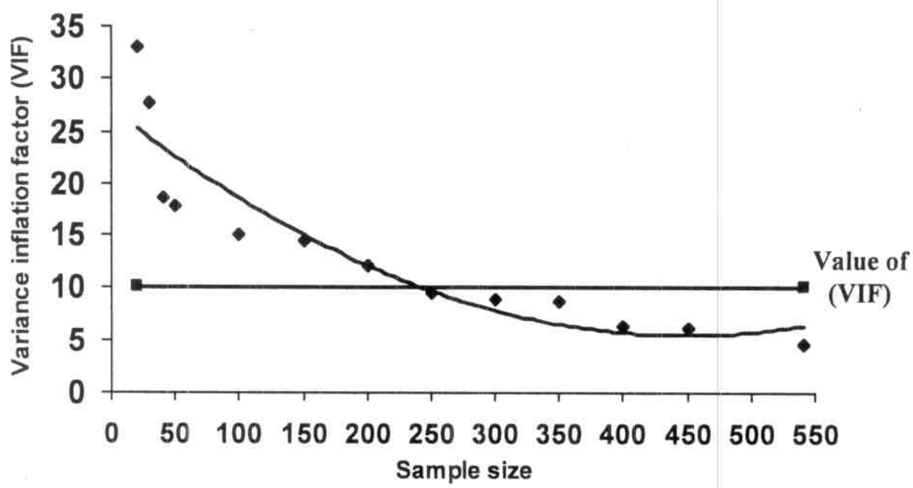


Fig (13): Effect of sample size on variance inflation factor (VIF) of No. of kernels / row in maize as mean of three seasons of experimentation.

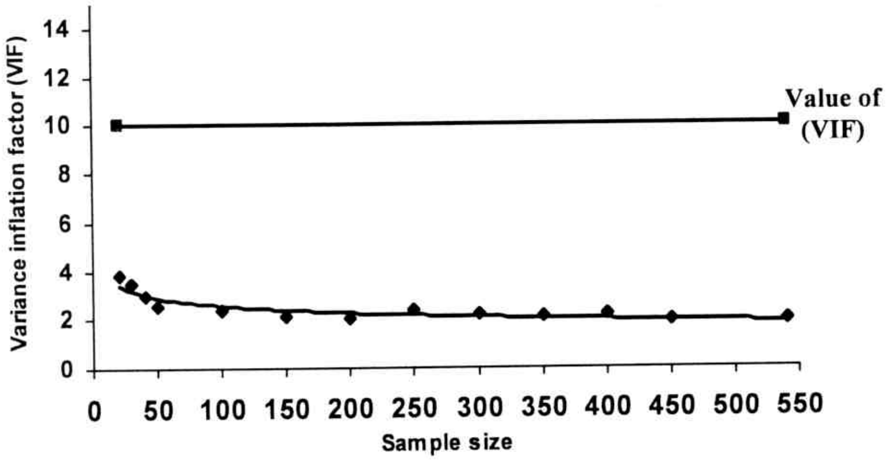


Fig (14): Effect of sample size on variance inflation factor (VIF) of ear length (cm) in maize as mean of three seasons of experimentation

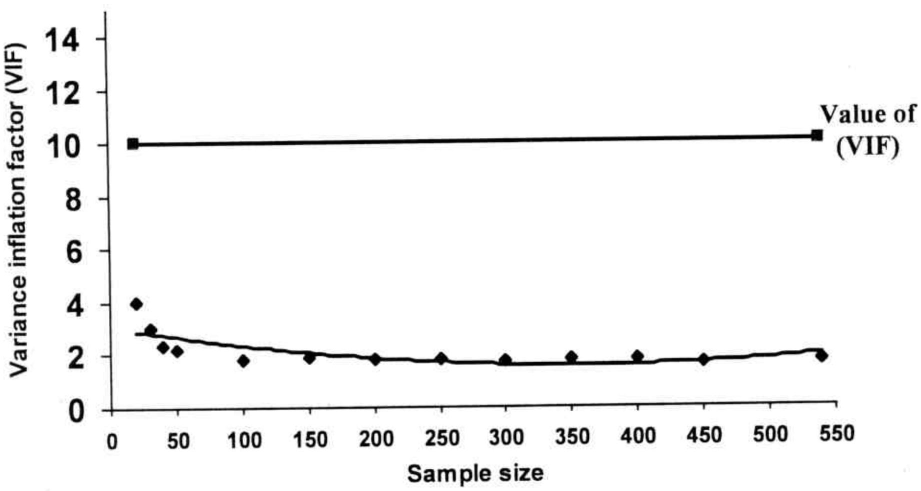


Fig (15): Effect of sample size on variance inflation factor (VIF) of ear diameter (cm) in maize as mean of three seasons of experimentation.

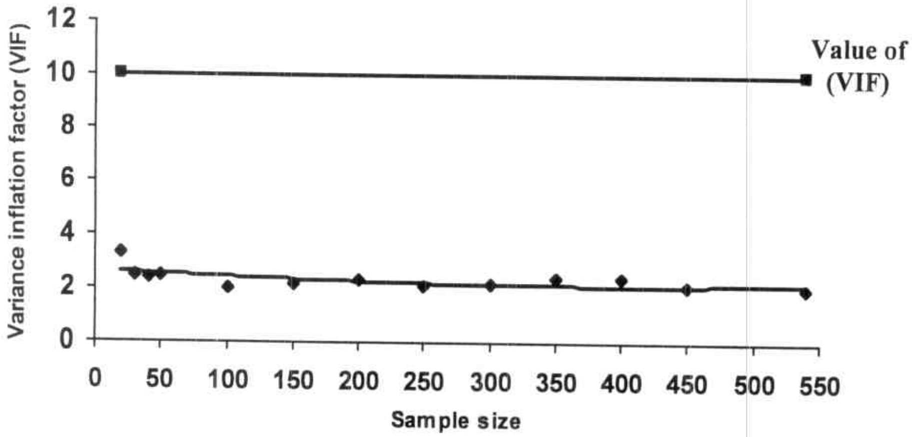


Fig (16): Effect of sample size on variance inflation factor (VIF) of weight of 100 kernels in maize as mean of three seasons of experimentation

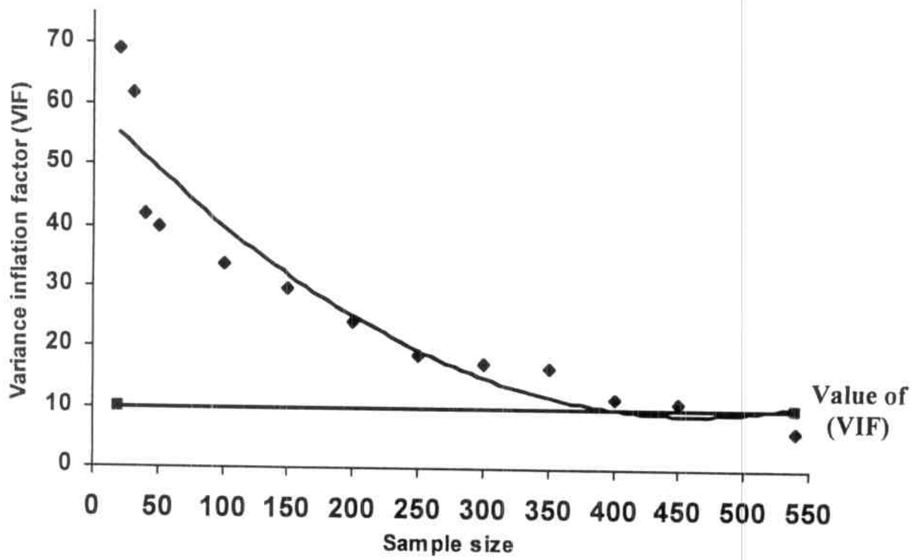


Fig (17): **Effect of sample size on variance inflation factor (VIF) of No. of kernels / ear in maize as mean of three seasons of experimentation**

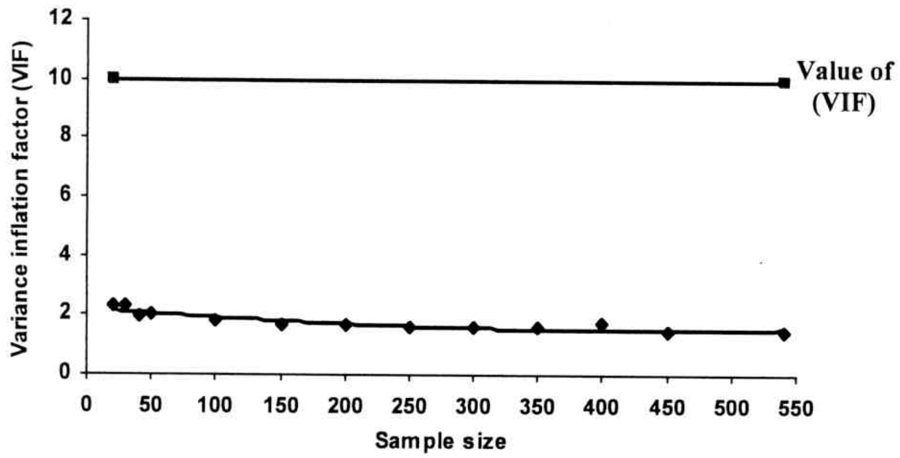


Fig (18): Effect of sample size on variance inflation factor (VIF) of plant height (cm) in maize as mean of three seasons of experimentation

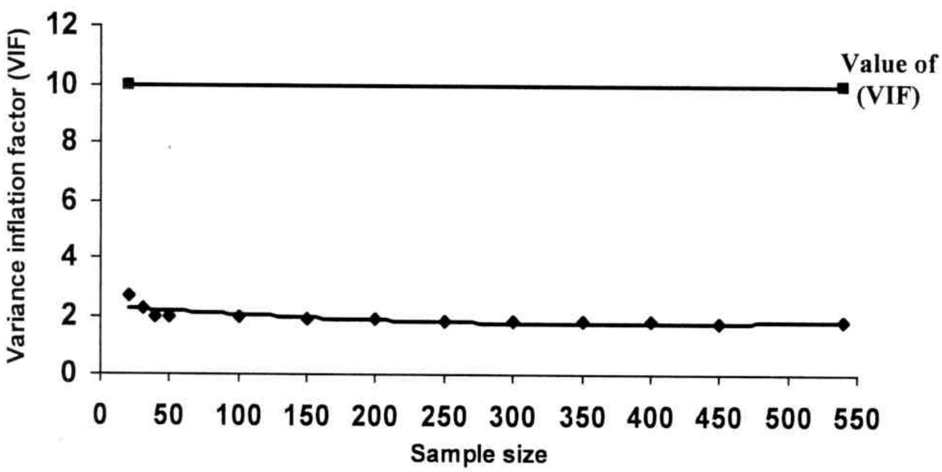


Fig (19): Effect of sample size on variance inflation factor (VIF) of No. of leaves in maize as mean of three seasons of experimentation

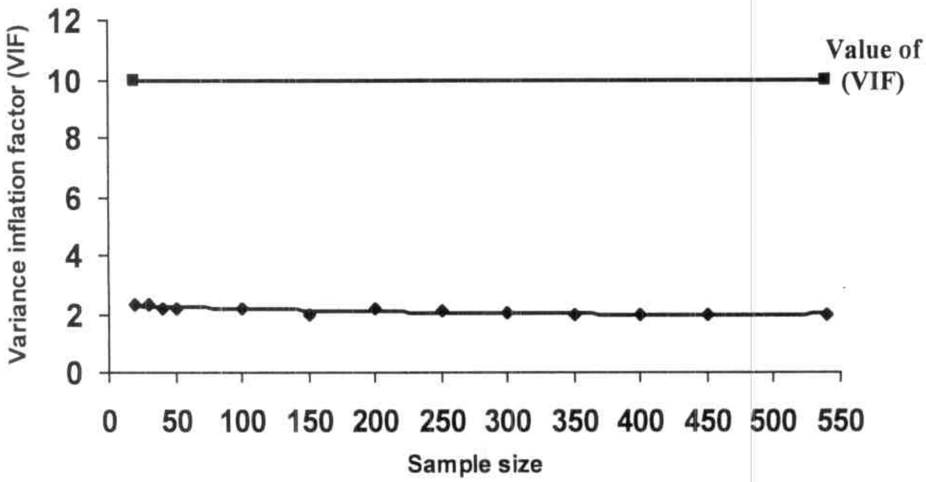


Fig (20): Effect of sample size on variance inflation factor (VIF) of stem diameter in maize as mean of three seasons of experimentation

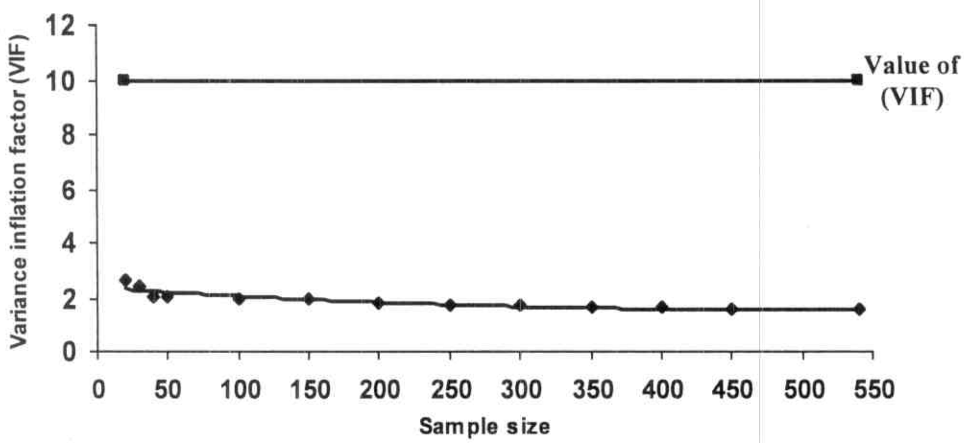


Fig (21): Effect of sample size on variance inflation factor (VIF) of leaf area (cm^2) in maize as mean of three seasons of experimentation

Table (11) indicates the prediction model by using multiple linear regression for grain yield of maize and its attributes. The prediction equation was formulated as follows:

$$Y= 254.869+ 3.175x1+ 12.059x2+ 0.485x3 + 1.062x4 + 0.144x5 + 0.081x6 + 0.209x7 +6.43x8+6.228x9+17.514x10$$

Where: X1= No. of leaves / plant, X2= stem diameter, X3= plant height,

X4= No. of rows / ear, X5= No. of kernels / row, X6= of kernels / ear,

X7= leaf area, X8= ear length, X9= weight of 100 kernels and X10= ear diameter.

The relative contribution for all yield factor explained 78.8% of the total variation in grain yield weight of 100 kernels, ear length, ear diameter, number of rows / ear, leaf area and number of kernels / ear had the highest coefficient of determination , their values were 53.52 , 50.18, 32.41, 27.56, 24.51 and 21.72 , respectively . Number of leaves / plant and stem diameter had a small value which was 4.36 and 4.02, respectively. Also, VIF value of all characters was less than 10.0. Thus, evidence from the VIF calculation confirmed a part of the stability of the regression coefficient due to nonsignificant interrelations among the explanatory variables which means that there is no collinearity or multicollinearity.

These results indicated that using this model with the highly multiple R squared (R^2 = 78.8) for prediction only (i.e. make no try to interpret the partial regression coefficients) and

Table 11. Multiple linear regression of 11 characters in predicting grain yield in maize by using the optimum sample size as mean of three seasons.

Characters	Regression coefficient	Standard Error (SE)	R^2%	Variance Inflation Factor (VIF)
1- No .of Leaves	3.175	0.643	4.36	1.158
2- Stem diameter	12.059	4.038	4.02	1.197
3- Plant height (cm)	0.485	0.027	18.68	1.452
4-No . of rows / ear	1.062	0.881	27.56	3.155
5- No . kernels / row	0.144	0.239	18.31	5.613
6- No . kernels/ ear	0.081	0.015	21.72	8.847
7- Leaf area (cm^2)	0.209	0.828	24.51	1.162
8 - Ear Length (cm)	6.436	0.469	50.18	1.168
9--Weight of 100 kernels	6.228	0.522	53.52	1.968
10- Ear diameter (cm)	17.514	3.023	32.41	1.382
Multiple R	0.887			
R squared	0.788			
Adjusted R squared	0.784			
Standard error of estimated	15.5796			

using a more sophisticated method of analysis such as, ridge regression, principle components or factor analysis to obtain a model that more clearly reflects the simple effects of the predictor **(Joseph *et al*, 1992).** These results are in agreement with those of **Ashmawy (1994), Nasr (1998) and Atia and Mahmoud (2006).**

A graphical representation of total pair observation can be obtained by the treatment of each pair of measurements such a diagram of scattered points is called a scatter diagram (Figures 22-30).

To illustrate, one may consider 540 pairs of observations that relate number of rows/ear, number of kernels/row, ear length, ear diameter, weight of 100 kernels, number of kernels/ear, plant height, number of leaves, stem diameter and leaf area to grain yield/ plant of maize. From the scatter diagram in (the aforementioned) Figures, clear that the points are plotted into four regions or quadrants. It is also evident that most of the points fall into two or these regions, i.e., those which describe the characters of the same type above the average and below the average .Thus, there appears to exist a direct or positive correlation between the characters. The totalities of points that form the scatter very often possess the rough geometrical form of an ellipse. The position of the ellipse indicates the type of association, i.e. whether positive (direct) or negative (inverse). The shape of the ellipse roughly estimates the degree of correlation. The characters are closely related when the ellipse is narrow.

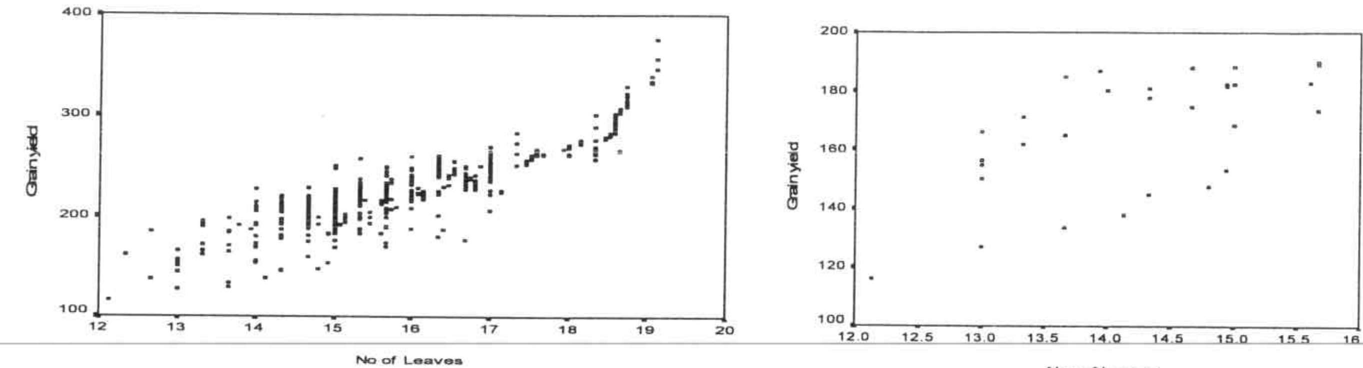


Fig 22 :Scatter diagram for the relationship between No. of leaves (x) and grain yield / plant (y) of 540 plants of maize as a mean of three seasons at experimentation (linear trend).

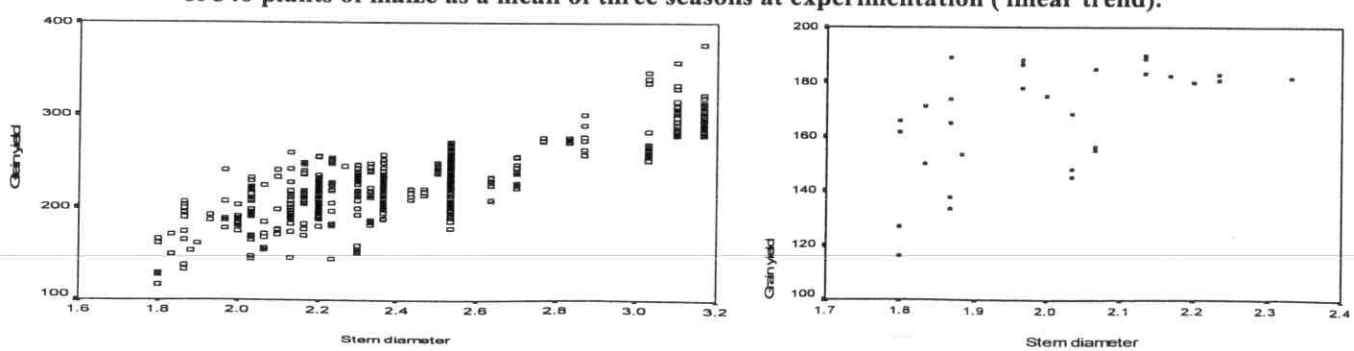


Fig 23 :Scatter diagram for the relationship between stem diameter (x) and grain yield / plant (y) of 540 plants of maize as a mean of three seasons at experimentation (linear trend).

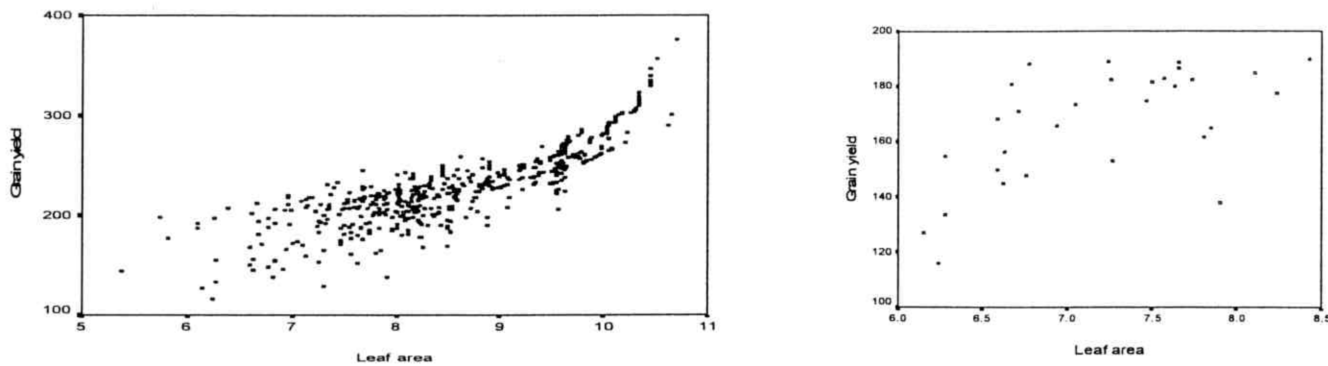


Fig24 :Scatter diagram for the relationship between leaf area (x) and grain yield / plant (y) of 540 plants of maize as a mean of three seasons at experimentation (quadratic trend) .

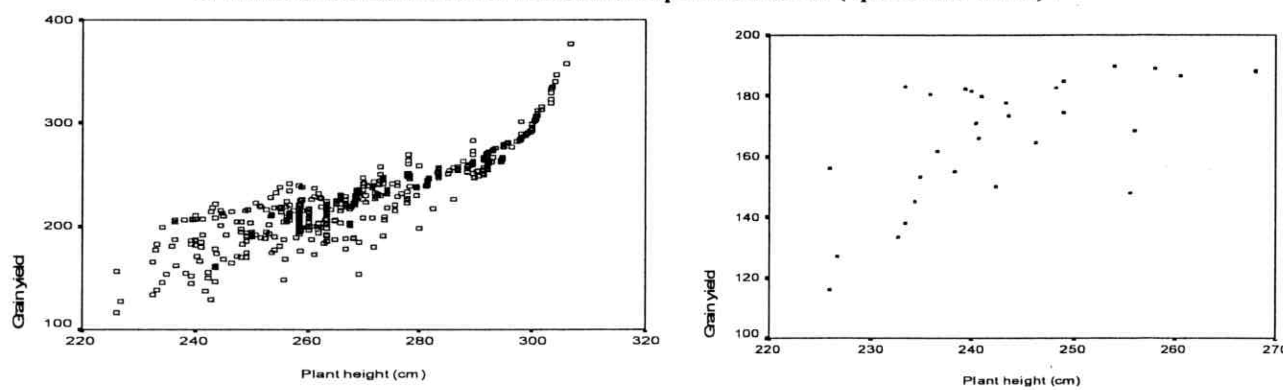


Fig 25 :Scatter diagram for the relationship between plant height (cm) (x) and grain yield / plant (y) of 540 plants of maize as a mean of three seasons at experimentation (quadratic trend) .

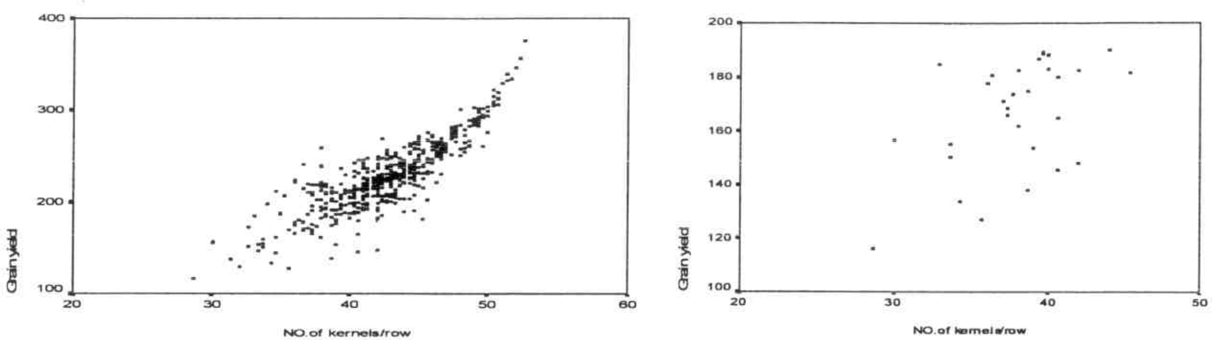


Fig 26 :Scatter diagram for the relationship between No. of kernels/ row (x) and grain yield / plant (y) of 540 plants of maize as a mean of three seasons at experimentation(linear trend).

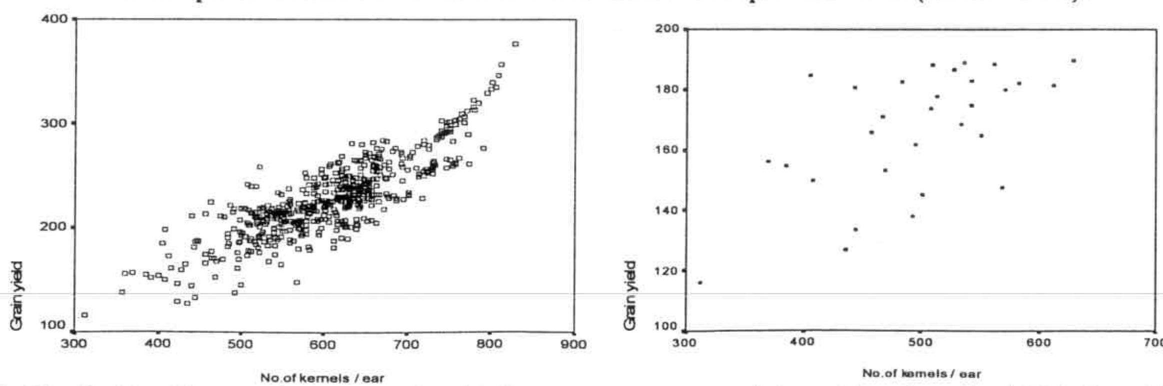


Fig 27 :Scatter diagram for the relationship between No .of kernels/ ear (x) and grain yield / plant (y) of 540 plants of maize as a mean of three seasons at experimentation (quadratic trend) ,

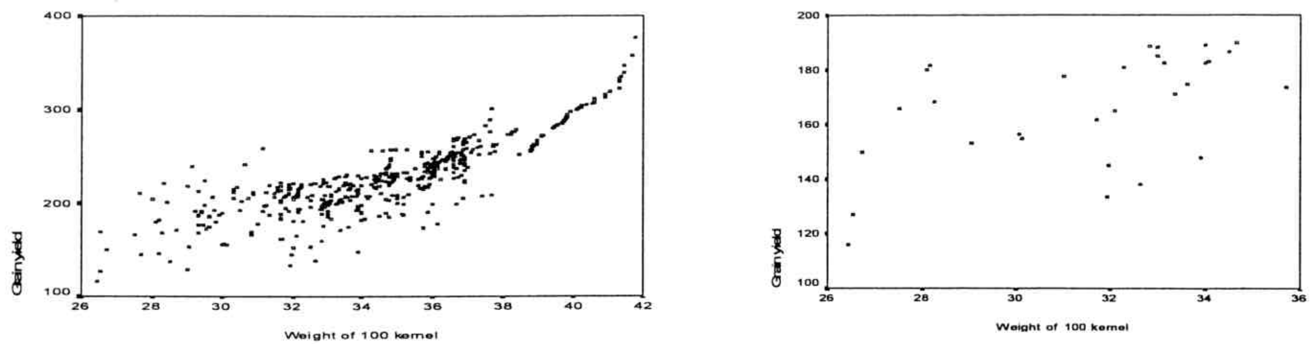


Fig 28 :Scatter diagram for the relationship between weight of 100 kernel (x) and grain yield / plant (y) of 540 plants of maize as a mean of three seasons at experimentation (quadratic trend) .

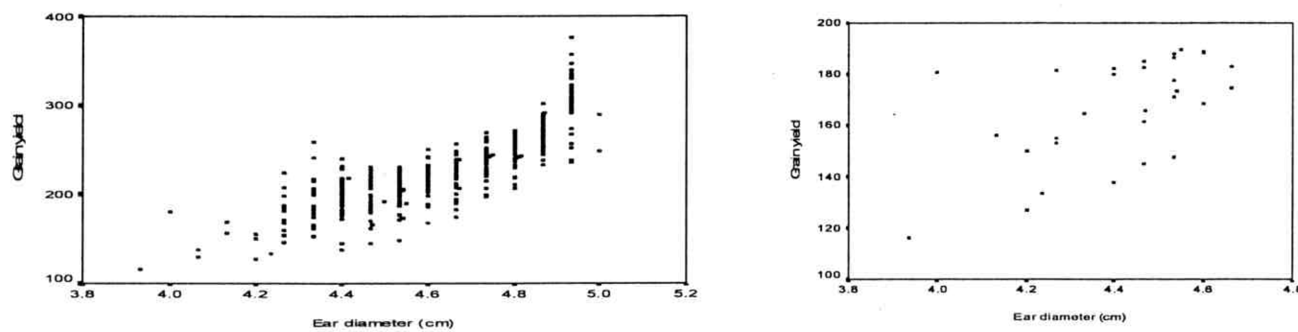
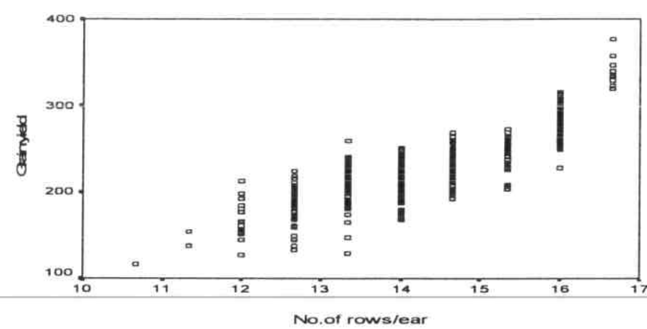


Fig 29 :Scatter diagram for the relationship between ear diameter(cm) (x) and grain yield / plant (y) of 540 plants of maize as a mean of three seasons at experimentation (linear trend).



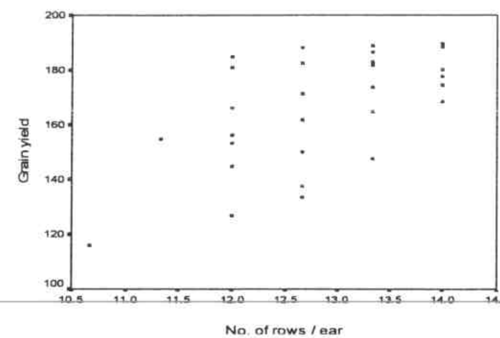


Fig 30 :Scatter diagram for the relationship between No. of rows / ear (x) and grain yield / plant (y) of 540 plants of maize as a mean of three seasons at experimentation (linear trend).

Table12: Variance inflation factor (VIF), R^2, adjusted R^2 and SE of accepted variables by using stepwise multiple regressions affected by sample sizes in relation between grain yield and yield factors in 2004 season.

Sample size	VIF										R^2	Adjust R^2	Std . Error
	No. of rows /ear	No. of Kernels /row	Ear Length (cm)	Ear Diameter (cm)	Weight of 100 Kernel	No. of kernels /ear	Plant height (cm)	No. of Leaves	Stem diameter	Leaf area (cm^2)			
20						1.000					0.438	0.406	21.623
30			1.111			1.111					0.565	0.533	22.116
40			1.235			1.174			1.244		0.623	0.580	21.812
50						1.341	1.241		1.102		0.633	0.609	20.060
100			1.397	1.278		1.283	1.266		1.475	1.541	0.674	0.649	20.116
150				1.753		1.807	1.332	1.850	1.948	1.297	0.715	0.706	22.662
200	1.292			1.473			1.365	1.804	1.998	1.304	0.724	0.713	17.681
250				1.648		1.855	1.369	1.877	2.039	1.339	0.714	0.707	18.068
300	1.391			1.558			1.381	1.774	2.055	1.322	0.728	0.723	16.595
350	1.495		1.901	1.726			1.417	1.894	2.330	1.499	0.756	0.751	15.68
400	1.562		1.958	1.765			1.491	1.938	2.299	1.594	0.785	0.782	15.397
450	1.701		1.993	1.870			1.619	2.086	2.287	1.758	0.813	0.810	15.473
540	2.636		2.266	2.397		2.618	2.451	2.355	2.631	2.789	0.881	0.879	16.984

Table13: Variance inflation factor (VIF), R^{2}, adjusted R^2 and SE of accepted variables by using stepwise multiple regressions affected by sample sizes in relation between grain yield and yield factors in 2005 season.

Sample Size	VIF										R^2	Adjust R^2	Std . Error
	No. of rows /ear	No. of Kernels /row	Ear Length (cm)	Ear Diameter (cm)	Weight of 100 Kernel	No. of kernels /ear	Plant height (cm)	No. of Leaves	Stem diameter	Leaf area (cm^2)			
20		1.000									0.397	0.363	16.524
30			1.123						1.006		0.577	0.528	16.115
40				1.065					1.065		0.568	0.544	15.593
50							1.096		1.096		0.576	0.558	13.466
100	1.527						1.538	1.093	1.227	6.329	0.714	0.702	12.648
150				1.605		1.112	1.845		1.198	1.280	0.719	0.709	10.966
200	1.218			1.697			1.874	1.156	1.247	1.232	0.732	0.724	9.273
250	2.819			1.602		2.794	1.915		1.232	1.213	0.741	0.735	7.822
300		5.630		1.094		6.022	1.529	1.091			0.762	0.758	6.329
350			1.262	1.379		1.372	1.932	1.064			0.762	0.762	5.731
400			1.429	1.327		1.407	1.815	1.164	1.125		0.765	0.768	5.316
450			1.598	1.521		1.328	1.865	1.149	1.109		0.779	0.777	5.825
540			1.495	1.539		1.220	1.818	1.105	1.128		0.786	0.783	4.795

Table 14: Variance inflation factor (VIF), R^2, adjusted R^2 and SE of accepted variables by using stepwise multiple regressions affected by sample sizes in relation between grain yield and yield factors in 2006 season.

Sample Size	VIF										R^2	Adjust R^2	Std . Error
	No. of rows /ear	No. of Kernels /row	Ear Length (cm)	Ear Diameter (cm)	Weight of 100 Kernel	No. of kernels /ear	Plant height (cm)	No. of Leaves	Stem diameter	Leaf area (cm^2)			
20								1.000			0.222	0.179	34.578
30				1.007			1.007				0.438	0.397	32.012
40			1.156						1.156		0.407	0.374	30.691
50			1.382					1.369	1.337		0.471	0.436	30.761
100			2.216	1.910				1.533	1.840	1.694	0.551	0.527	31.029
150			1.156	1.473		2.262	2.864	2.332	2.811	1.589	0.616	0.603	29.789
200				1.554		2.682	2.116	2.322		1.585	0.647	0.637	29.608
250				1.535		2.444	3.125	2.275	2.536		0.686	0.679	29.075
300			2.087	1.677		2.821		2.244	1.975		0.677	0.671	28.721
350				1.756		2.344	3.075	2.427	2.657		0.680	0.676	28.987
400				1.827		2.359	3.040	2.436	2.637		0.686	0.682	28.334
450				1.904		2.458	3.146	2.516	2.707		0.680	0.676	27.251
540				2.089		2.107	3.020	2.813	2.467		0.739	0.737	27.551

error ranged from 34.58 to 27.6 in the third season. Although the sample sizes are the same in the three seasons the relative contribution of accepted variables differed from season to other. Hence, stepwise methods will not necessary produce the best model if there are redundant predictor variables **(Tibshirani, 1996).** It can be classified sample sizes into three sizes. These sizes were small which ranged from 20 to 40, medium which ranged from 50 to 200 and large which ranged from 250 to 540 observations. At small sample size number of accepted variables was two (ear length and stem diameter) in the three seasons. At medium size six variables were accepted in the first and second season, but in third season five variables were accepted. At large sample size number of accepted variables were seven, six and five in the first, second and third seasons, respectively. Accepted variables differed from sample size to other. Also, stepwise methods are even more affected by multicollinearity than regular method this is problem of stepwise regression **(Fox, 2005).**The degree of correlation between the predictor variables affected the frequency with which authentic predictor variables found their way into the final model. The number of candidate predictor variables affected the number of noise variables that gained entry to the model **(Altman and Andersen, 1989).** On the other hand, number of kernels / row and weight of 100 kernels removed for all sample sizes in the first season. In the second season weight of 100 kernels removed for all sample size. Also, number of rows / ear, number of kernels / row and weight of 100 kernels removed for all sample size in the third season. Although these variables are most important variables of yield components, the stepwise analysis removed these variables, since of these

Table 15: Accepted and removed variables according to stepwise multiple linear regression and relative contribution in grain yield in maize by using the optimum sample size as mean of three seasons.

Characters	Regression coefficient	Standard Error (SE)	R^2%	Variance Inflation factor (VIF)
1- No .of Leaves	3.244	0.830	4.361	1.938
2- Stem diameter	8.964	3.136	4.022	2.299
3- Plant height (cm)	0.411	0.053	18.68	1.491
4-No . of rows / ear	3.781	0.924	27.56	1.562
5- Leaf area (cm^2)	0.113	0.012	24.51	1.594
6 - Ear Length (cm)	1.919	0.652	50.18	1.958
7- Ear diameter (cm)	32.058	3.735	32.41	1.765
Removed variables :				
1- No . kernels / row	0.005			12.181
2- No . kernels/ ear	0.015			13.015
3- Weight of 100 kernel	0.011			11.465
Multiple R	0.886			
R squared	0.785			
Adjusted R squared	0.782			
Standard error of estimated	5.397			

To evaluate the results of stepwise analysis, assumptions must be tested. Meeting the assumptions of regression analysis is essential to ensure both that the results obtained were truly representative of the sample and that we have obtained the best possible results. The principal measure used to test the assumption is the residual difference between the actual dependent variable values and its predicted values **(Joseph *et al*, 1992).**

On the other hand, removed variables were number of kernels / row, number of kernels / ear and weight of 100 kernels, these variables are considered characters of yield components. Hence, this equation is not valid. These results were in agreement with **El-Kalla and El-Rayes (1984), Ashmawy (1989) and Shafshak *et al* (1989).**

In stepwise multiple linear regression with more than independent variable, examining the residuals shows the combined effects of all predictor variables, but we can not examine any predictor variable separately. To do so, we use partial regression plots. It shows the relationship of a single predictor variable with the criterion variable. Examination of (Figure 31when sample size less than 30 plants and Figure 32 when sample size of 540 plants) indicates a nonlinear relationship between this specific predictor variable and the criterion variable. This is more useful method when we have several predictor variables, as we can tell which specific variables violate the assumption of linearity and apply the needed remedies only to them .Also, the identification of outliers or influential observations is facilitated on the basis of one predictor variable at a time . Also, we can see that the ear diameter, number of kernels /ear, plant height, number of leaves, stem diameter and leaf area are less well defined , both in slope and scatter of the points , thus explaining there lesser effect in the equation for all six components , nonlinear pattern is shown . Therefore, this equation is not valid to interpret the partial correlation **(Joseph, *et al*, 1992).** Partial regression plots are often used to assess nonlinearity. These are simply plots of each independent on the x axis against the dependent on the y axis. Curvature in the pattern of points in a partial regression plot shows if there is a nonlinear relationship between the dependent and any one of the independents taken individually. Note, however, partial regression plots are preferred for illuminating cases with high leverage. Partial residual plots are preferred for illuminating nonlinearities.

Generally, the distribution of standardized partial regression plot had no shape when sample size was less than 30 plants as shown Figure (31). On the other hand, the distribution of standardized partial regression plot is captured beautifully by the sample size of 540 plants. Hence, the relationship is very clear and will look like the curves in Figure (2).

Figure (32) shows the large sample size is quite well defined the relationships between independent variables and dependent variable. The results could be illustrated in the following graphs:

Figure of number of kernels / ear which will look like the scatter in Figure (1a) called null plot.

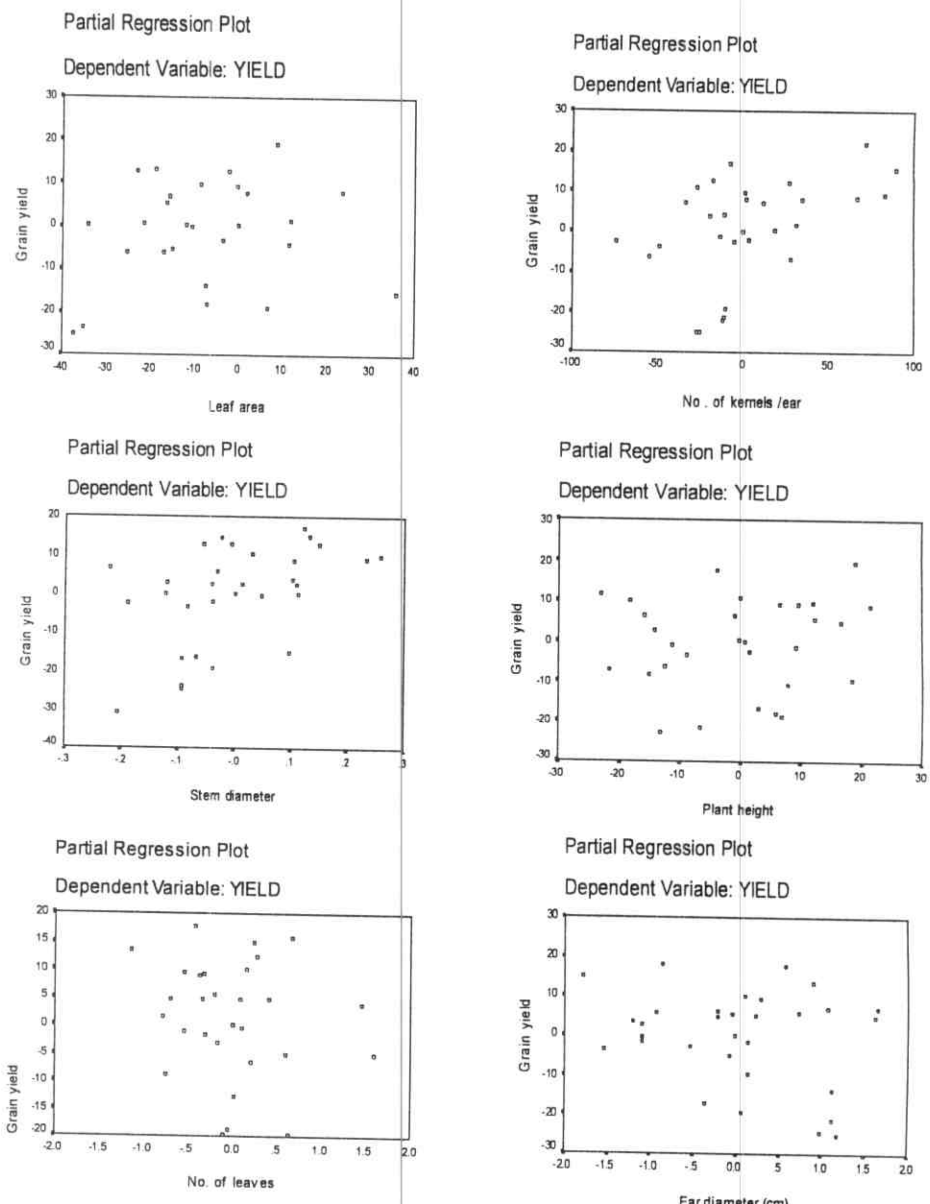


Fig 31: Partial regression plot for leaf area, No. of kernels/ear, stem diameter, plant height, No. of leaves / plant and ear diameter with grain yield at sample size of 30.

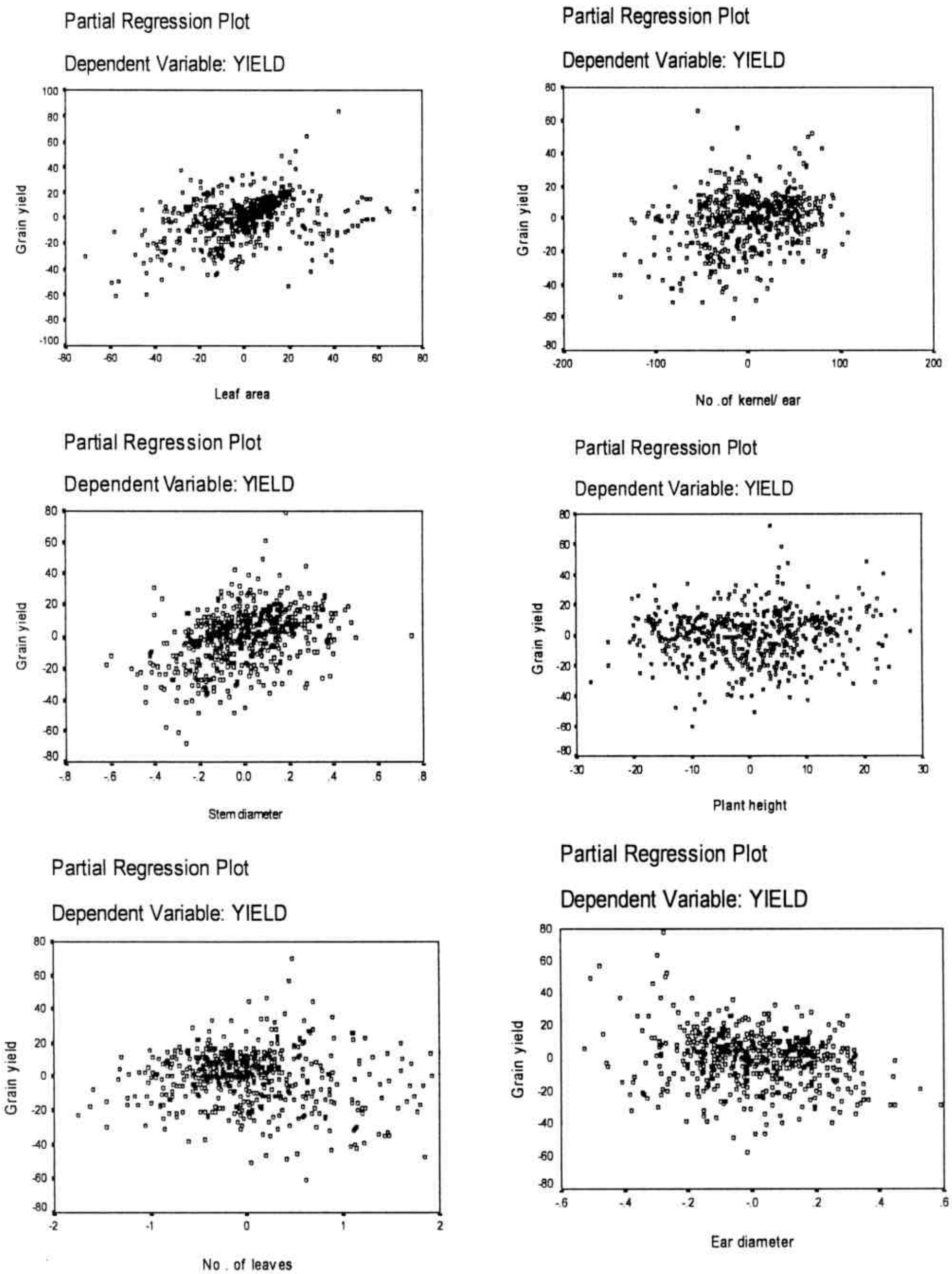


Fig 32: Partial regression plot for leaf area, No. of kernels/ear, stem diameter, plant height, No. of leaves/ plant and ear diameter with grain yield at sample size of 540.

Figure of leaf area, stem diameter and plant height which will look like the scatter in Figure (1e) called time – based dependence.

Figure of ear diameter which will look like the scatter in Figure (1h) called nonlinearity and heteroscedasticity.

Figure of number of leaves which will look like the scatter in Figure (1d) called heteroscedasticity.

Factor analysis:

To evaluate factor analysis we used communality value (h^2) because when an indicator variable has a low communality, the factor model is not working well for that indicator and possibly it should be removed from the model. However, communalities must be interpreted in relation to the interpretability of the factors. For this reason, the researcher used communality value (h^2) for evaluating the efficiency of factor analysis model with estimation its value which was affected by sample size. Communality value (h^2) as affected by sample size is presented in Tables 16, 17 and 18in 2004, 2005 and 2006 seasons, respectively.

The results indicate that increasing sample size increased h^2 value of all characters. A communality value of 0.6 seems high and its effect is greater. When communality value less than 0.6 seems low and had no effect. When communality is high, plays a greater role in the interpretation of the factor. Also, noted that number of row/ear, number of kernels/row, number of kernels/ear and weight of 100 kernels are the most variables which were effective in independent structure at small sample size which less than 50 observations. Hence, these characters are

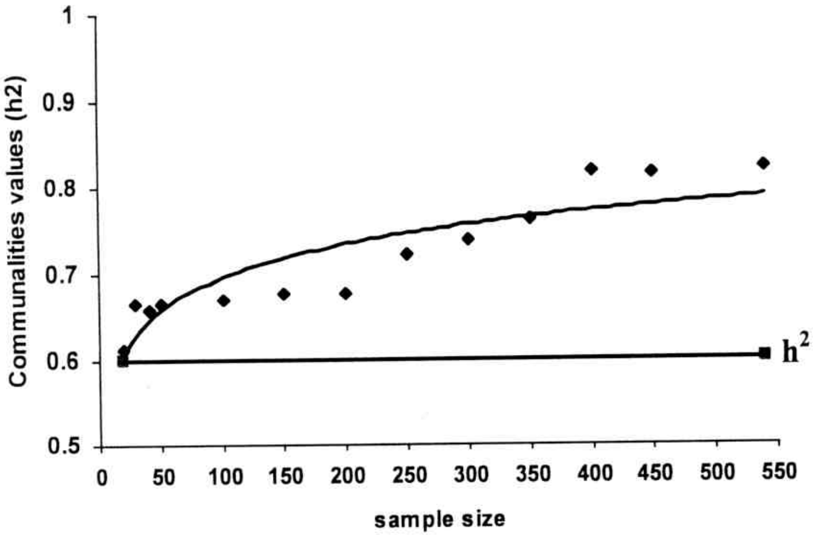


Fig (33): Effect of sample size on communalities values (h^2) of No. of rows/ ear in maize as mean of three seasons of experimentation.

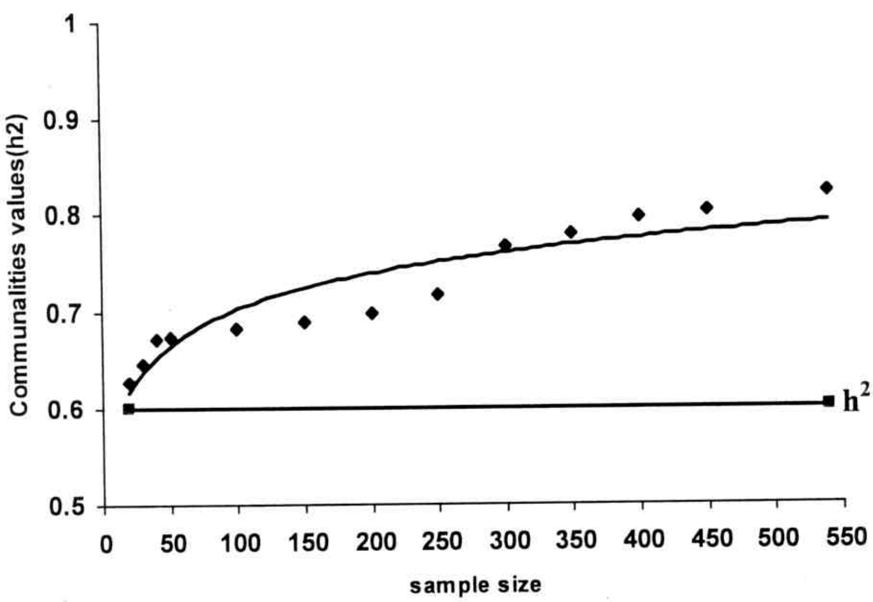


Fig (34): Effect of sample size on communalities values (h^2) of No. of kernels/row in maize as mean of three seasons of experimentation.

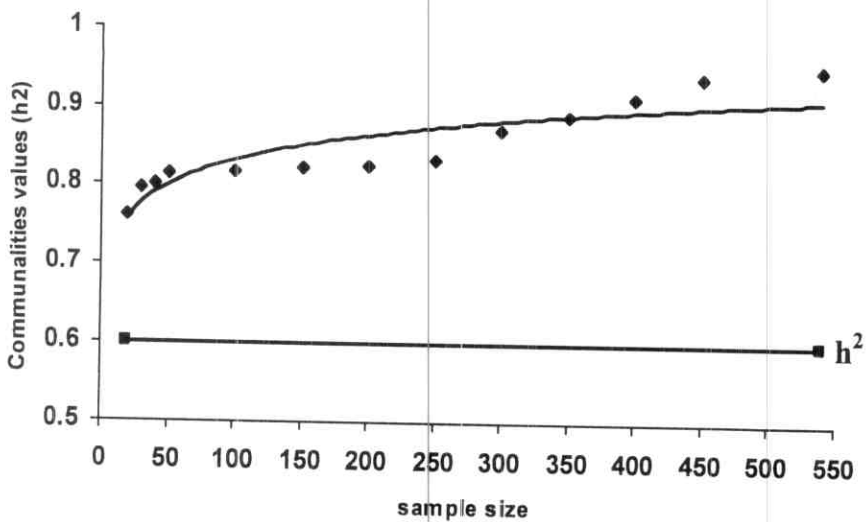


Fig (35): Effect of sample size on communalities values (h^2) of No. of kernels/ear in maize as mean of three seasons of experimentation

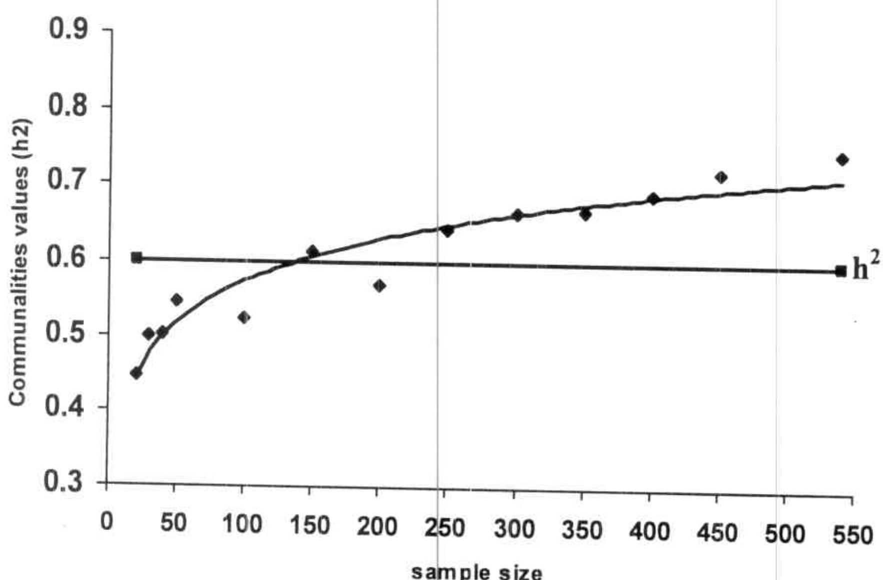


Fig (36): Effect of sample size on communalities values (h^2) of ear diameter (cm) in maize as mean of three seasons of experimentation

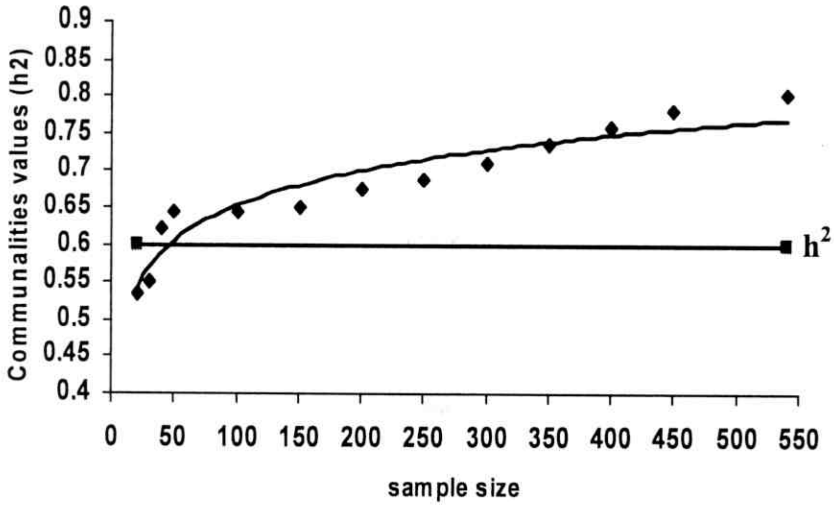


Fig (37): Effect of sample size on communalities values (h^2) of weight of 100 kernels in maize as mean of three seasons of experimentation

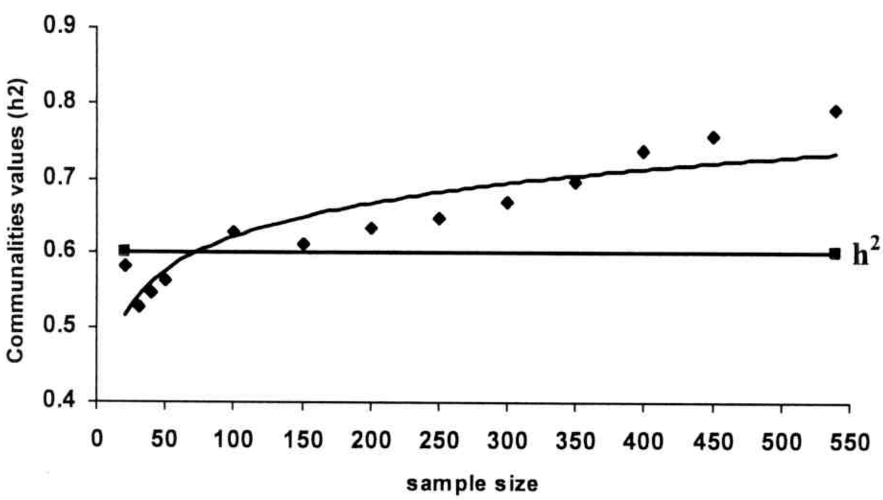


Fig (38): Effect of sample size on communalities values (h^2) of No. of leaves in maize as mean of three seasons of experimentation

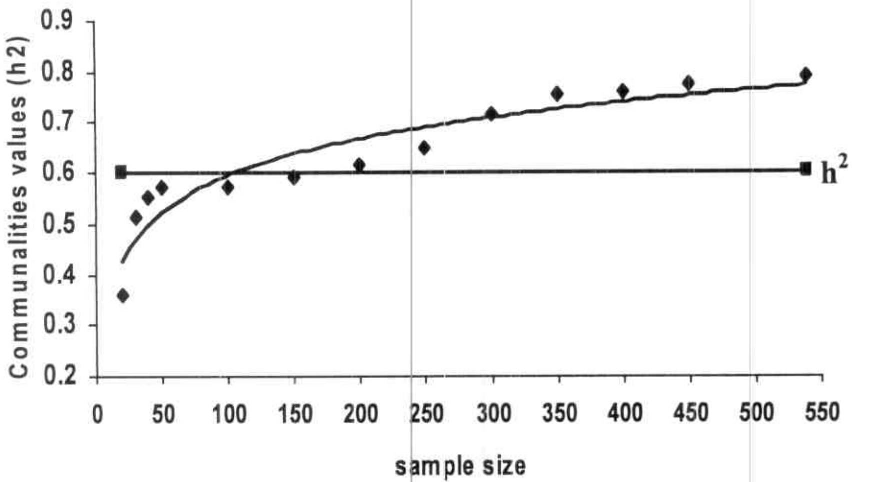


Fig (39): **Effect of sample size on communalities values (h^2) of plant height (cm) in maize as mean of three seasons of experimentation**

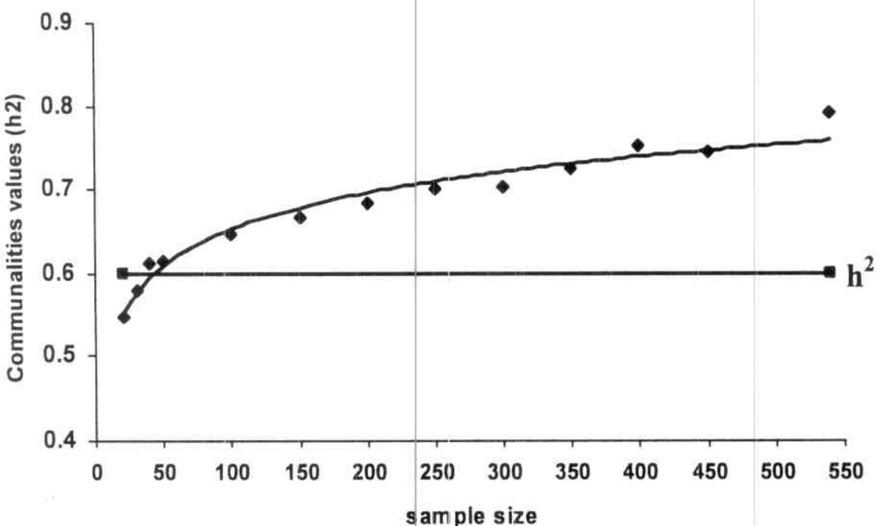


Fig (40): **Effect of sample size on communalities values (h^2) of ear length (cm) in maize as mean of three seasons of experimentation**

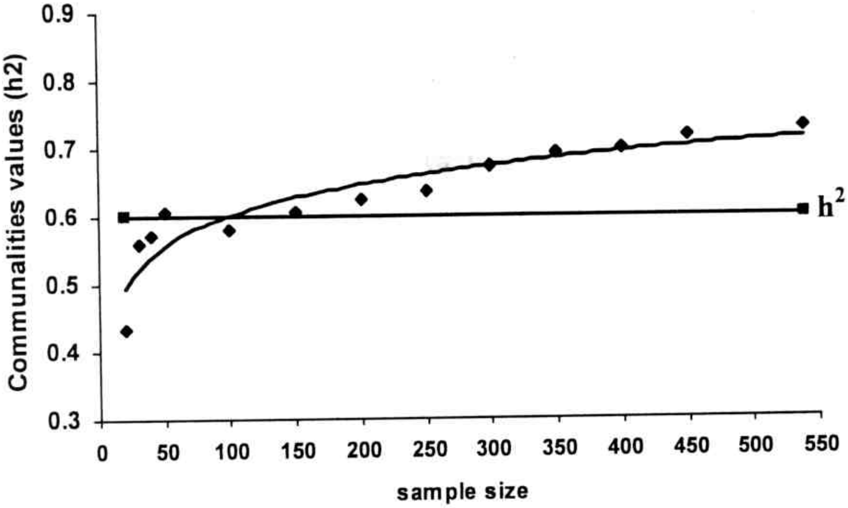


Fig (41): Effect of sample size on communalities values (h^2) of stem diameter in maize as mean of three seasons of experimentation

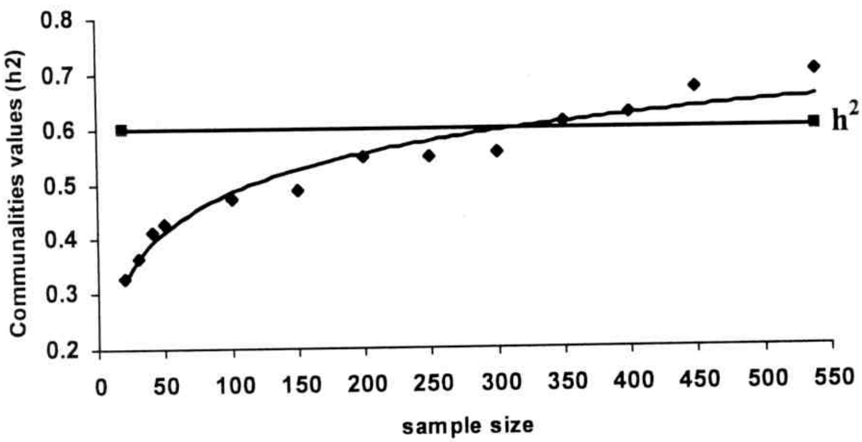


Fig (42): Effect of sample size on communalities values (h^2) of leaf area (cm^2) in maize as mean of three seasons of experimentation be achieved with relatively small samples. However, when communalities are low, recovery of population factors is difficult to achieve unless sample size is extremely large **(MacCallum *et al*, 1999 and Velicer and Fava, 1998).** This result is in agreement with Mac **Callum *et al* (2001),** who reported that when communalities are lower, much larger samples are needed, with this phenomenon being amplified by poorly over determined factors. Indeed, sample sizes may have to be much larger than typically recommended when communalities are low and factors are not highly over determined. It is not possible to make blanket recommendations regarding this issue without considering other important aspects of design.

Another parameter was used to evaluate factor analysis was eigen value. The eigen value for a given factor measures the variance in all the variables which is accounted for by that factor. The ratio of eigen values is the ratio of explanatory importance of the factors with respect to the variables. Eigen value and variance ratio as affected by sample size are presented in Tables 19, 20 and 21 in 2004. 2005 and 2006 seasons, respectively.

According to results in Tables 19, 20 and 21, increasing sample size increased eigen value and variance ratio. Also, noted that the values of variance ratio are greater than the values of eigen value. On the other hand, when sample size is greater than 400 the values of eigen value and variance ratio became fixed .Sample size in factor analysis should be at least 400 cases **(Garson, 2008).**

Factor analysis was constructed using the principal factor analysis technique to establish the dependent relationship

Table 19: Eigen value and percentage of variance as affected by Sample sizes in factor analysis in maize in 2004 season.

Sample size	Eigen value	percentage of variance
20	5.431	54.313
30	5.585	55.853
40	5.638	56.383
50	5.663	56.631
100	5.710	57.097
150	5.822	58.222
200	5.926	59.260
250	6.164	61.639
300	6.526	65.259
350	6.622	66.295
400	6.662	66.618
450	7.046	70.461
540	7.093	70.931

Table 20: Eigen value and percentage of variance as affected by Sample sizes in factor analysis in maize in 2005 season.

Sample size	Eigen value	percentage of variance
20	6.353	63.526
30	6.588	65.876
40	6.823	68.229
50	6.959	69.585
100	6.923	69.232
150	7.018	70.177
200	7.062	70.624
250	7.188	71.877
300	7.531	75.313
350	7.631	76.314
400	8.073	80.732
450	8.215	82.140
540	8.296	82.960

Table 21: Eigen value and percentage of variance as affected by Sample sizes in factor analysis in maize in 2006 season.

Sample size	Eigen value	percentage of variance
20	6.262	62.621
30	6.305	63.049
40	6.316	63.158
50	6.361	63.596
100	6.438	64.381
150	6.472	64.722
200	6.531	65.299
250	6.543	65.431
300	6.721	67.211
350	6.794	67.935
400	7.012	70.104
450	7.123	70.474
540	7.198	71.975

between morphological characteristics and yield components. Factor analysis is one that can be used successfully for large amount of multivariate data, and should be applied more frequently in field experiments **(Joseph, *et al*, 1992 and Hammed, 1993).** This study describes one application of factor analysis that further explains the multivariate structure. Use for factor analysis by plant breeders has the potential of increasing the comprehension of the causal relationships of variables, and can help to determine the nature and sequence of trials to be selected in a breeding program.

The results are in agreement with that of **Walton (1972),** who stated that to make equation useful of predicative purposes we should include model many independent variables. The reliable fitted values could be determined by minimizing the costs for getting large number of observations. Also, the results are in agreement with those of **Guadagnoli and Velicer (1988), Mulaik (1990) and Widaman (1993), Nasr and Leilah (1993a), Velicer and Fava (1998) and MacCallum *et al* (1999).** Also, the results are in agreement with those of **MacCallum *et al* (2001).** They found that samples some what smaller than traditionally recommended are likely sufficient when communalities are high. When communalities are low, much larger samples are needed, with this phenomenon being amplified by poorly over determined factors. Indeed, sample sizes may have much larger than typically recommended when communalities are low and factors are not highly over determined. They also agree with the findings of **Jason and Costello (2004), Guo (2004)** who reported that large sample gave factor loadings, are more precise estimates of population loadings and are more stable. Results agree with **Osborne and Costello (2004),** who mentioned that the most valid conclusion regarding big sample size. Also the results agree with Costello and **Osborne (2005).**

Four hundred optimum sample size gave the more accuracy, small value of standard error and most replicable results in factor analysis.

The results of factor analysis are presented in Table 22. Factor analysis divided the ten variables into two main factors. Factor loading greater than 0.5 were considered important for purpose of interpretation. A summary of the composition of variables of the two factors with loading is given in Table 22. The two factors explained 64.722 % of the total variability in the dependence structure.

Factor 1 included six variables (number of rows/ ear, number of kernels/ row, number of kernels/ ear, ear length, weight of 100 kernels and ear diameter) which accounted for 40.137 % of the total variance. The six variables were positively correlated with factor I. Thus, this factor may be regarded as ear factor.

Factor 2 was called a growth factor because it consisted of number of leaves, stem diameter, plant height and leaf area. It accounted only 24.585 % of the total variance in the dependence structure. All variables loading of factor 2 were positive. The sign of the loading indicates the direction of the relationship between the factor and variable.

The ear factor (Factorl) had high loadings for six variables. The correlation between these variables and factor 1

Table 22 . Summary of factor loading for 10 characters of maize

Characters	Loading	% Total communality	Percentage of variance	Latent root	Suggested factor name
Factor 1 :			40.137	4.014	
1- No . of rows / ear	0.712	0.738			
2- No . kernels / row	0.903	0.631			Yield factor
3-- No. kernels / ear	0.950	0.515			
4- Ear Length (cm)	0.820	0.284			
5- Weight of 100 kernels	0.596	0.143			
6- Ear diameter (cm)	0.628	0.028			
Factor 2 :			24.585	2.458	
7- No .of Leaves	0.542	5.283			
8- Stem diameter	0.842	1.189			Growth factor
9- Plant height (cm)	0.786	0.834			
10- Leaf area (cm^2)	0.586	0.356			
Cumulative variance			64.722		

are given by the appropriate factor loading. Based on the investigated genotypes, selection for plant type with more (number of rows / ear, number of kernels / row, number of kernels/ear, ear length, weight of 100 kernels and ear diameter) will enable breeders to realize desired gains in grain yield of maize.

Factor 2 (growth factor) contained four variable (number of leaves / plant, stem diameter, plant height and leaf area) usually associated with grain yield on a biological yield. Also, the high simple correlation between these four variables would indicate that selection for rather variable will be a determined to the other. Use of factor analysis by plant breeders has the potential of increasing the comprehension of the causal relationships of variables and can help to determine the nature and sequence of traits to be selected in a breeding program. This result is confirmed by the findings of **El- Kalla and El- Rayes (1984), Ashmawy (1989), Mohamed and Sedhom (1993), Ashmawy (1994), Mohamed (2004) and Atia and Mahmond (2006).** So, it can be recommended from the previous results that: It is essential to detect characters having the greatest influence on yield and their relative contribution to yield variation. This is useful in designing and evaluating breeding programs and agronomic systems. The important characters over all statistical procedures used were ear length, ear diameter and stem diameter. The results of statistical procedures indicated that, factor analysis would seem to be more suitable and efficient than other procedures.