

# Applying Multivariate Analysis to Assess Phenotypic Diversity in Rice Cultivars Grown at Fogera and Pawe, Northwest Ethiopia

Taddesse Lakew

Ethiopian Institute of Agricultural Research, Fogera Rice Research and Training Centre  
P.O. Box 1937, Bahir Dar, Ethiopia

## Abstract

Field experiment was conducted to assess genetic variation and traits relationship in rice cultivars from Ethiopia based on quantitative traits. Sixty rice cultivars comprising improved varieties and collected accessions were evaluated at Fogera and Pawe research stations. The experiment was arranged using alpha lattice design of three replications. Correlation studies revealed panicle length, plant height, filled grains per panicle, spikelet fertility rate, biomass yield, harvest index and thousand seeds weight having a significant positive association with grain yield both at Pawe and Fogera. However, days to heading and days to maturity exhibited a significant negative association with grain yield at Fogera but positive association at Pawe. Principal component analysis showed that the first four components explained a cumulative variance of 81.07, 74.95%, and 80.63% at Fogera, Pawe and combined data, respectively. Days to heading, days to maturity, plant height, panicle length, filled grains per panicle and grain yield were the most discriminate traits for explaining the total variation. Cluster analysis classified sixty rice cultivars into four different groups. The first and second clusters comprised the largest number of cultivars. About 48% of improved cultivars containing the desired traits adapted predominantly to upland production system belonged to Cluster I, having the lowest intra-cluster distance. Clusters II and IV included both upland and lowland cultivars and the highest intra-cluster distance in both confirmed their heterogeneous composition while Cluster III contained exclusively of upland cultivars comprising improved and landrace types. The highest inter-cluster distance was observed between clusters II and IV, followed by clusters III and IV but the lowest between clusters I and II. Thus, future crossing program between cultivars of different cluster groups could possibly result in better heterosis in our rice breeding program.

**Keywords:** Genetic variation, rice, cluster, principal component, Ethiopia

**DOI:** 10.7176/JBAH/11-15-03

**Publication date:** August 31<sup>st</sup> 2021

## 1. Introduction

Rice (*Oryza sativa* L.) is the most important cereal crop cultivated and consumed globally to meet the daily calorie needs of the increasing world population mainly in Asia and Sub-Saharan Africa (Anyaocha *et al.* 2018). This important cereal is cultivated and consumed across Africa, where domestic production is unable to meet local demands for rice resulting in huge annual imports. In Ethiopia, since its introduction, the crop is increasingly expanding and its consumption has sharply increased. Rainfed-upland and lowland are predominant ecosystems in Ethiopia covering about 85,288.87 ha of land and cultivated by 230,496 rice farmers and contributing about 268, 223 tones of paddy rice to the annual cereal production (CSA 2020). Considering the potential, which is about 5million ha highly suitable for rice cultivation, the national share of rice crop in terms of area coverage is yet insignificant.

Despite the huge potential in the country, the national average yield of the crop is about 3.15 t/ha. In fact, it is the second high yielding crop, next to maize among major cereal crops in the country. This national average is significantly lower compared to major rice producing countries ( $\geq 5$  t/ha) as well as that of global average of 4.66 t/ha (FAOSTAT 2020). Lower productivity in rice could be attributed to low yielding varieties and improper agronomic practices, among other things. Thus, improvement of grain yield and yield attributing traits of rice is necessary for different production ecosystems in Ethiopia. The development of varieties is a continuous process and the success of the plant breeding programme which aimed at identifying high yielding, better quality, fertilizer responsive, disease and insect resistance varieties depends upon the selection of suitable parents to be utilized in breeding (Pathak *et al.* 2020).

The effectiveness of selection depends primarily upon the magnitude of genetic variability in the breeding materials available (Qamar *et al.* 2012; Maji and Shaibu 2012). Although Ethiopia has some locally known rice cultivars and more than 30 improved rice varieties recommended for various production systems, they are not yet effectively used as parents in rice breeding programs. Thus, precise evaluation, classification and identification among cultivars in Ethiopia are required for further breeding program. Such evidence is particularly useful to assess the potential of heterotic combinations before attempt crossing and hence saving time and resources (Hallauer and Miranda 1988). Phenotypic diversity and yield performance of some selected rice accessions of Ethiopia have been reported (Bitew *et al.* 2016, 2018; Girma *et al.* 2018). However, there has been no report comprising local and improved cultivars that are adapted to different localities in Ethiopia. To investigate the

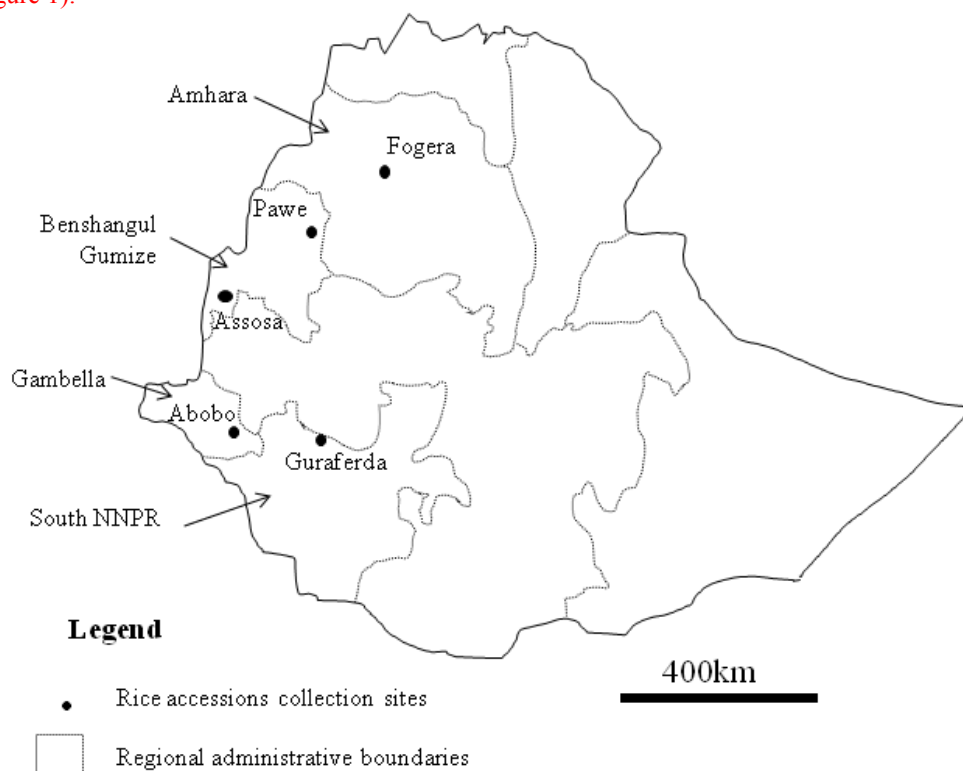
extent of phenotypic diversity and traits relationship among improved cultivars and collected accessions, it is important to evaluate these materials together under the same experimental procedure. In this study, we focused on some selected quantitative traits under two environments and analyze the pattern of diversity and relationship among different traits.

Different statistical methods can be applied to assess diversity and traits relationships in plant materials. Multivariate statistical tools are extensively used in summarizing and describing the inherent variations among crop genotypes and helps plant breeders to formulate their selection approaches for improving traits of their interest (Riaz *et al.* 2018). Principal component analysis (PCA), as a multivariate statistical technique, analyses data consisting of several inter-correlated quantitative dependent variables as observations and used to explore the variation among genotypes (Riaz *et al.* 2018; Mahendran *et al.* 2015). PCA helps to identify the traits with the highest variability, while correlation analysis reveals the strength of relationships of the identified traits with yield and other traits, which makes the development of new varieties efficient and effective (Chakravorty *et al.* 2013). Moreover, the cluster analysis using Euclidian distance provides a useful statistical tool for measuring the genetic diversity in rice gemplasms with respect to traits considered and thus hybridization programme involving genetically diverse parents belonging to different clusters would provide the opportunity of maximum heterosis for different traits of interest to improve and high variability in subsequent generations (Tejaswini *et al.* 2016). The present study, therefore, aims to determine the relationship among rice cultivars and to estimate the relative contribution of different traits and thereby guide the selection of promising parents for use in rice variety improvement programme.

## 2. Materials and methods

### 2.1 Plant materials

This investigation was carried out using a total of sixty rice cultivars comprising twenty-nine released varieties, four candidate varieties and twenty-seven collected accessions. Rice accessions were collected from five different districts of four regional states in Ethiopia; Fogera in Amhara region, Pawe and Assosa in Benshangul Gumize region, Guraferda in South Nations Nationalities and Peoples Region (SNNPR), and Abobo in Gambella region (Figure 1).

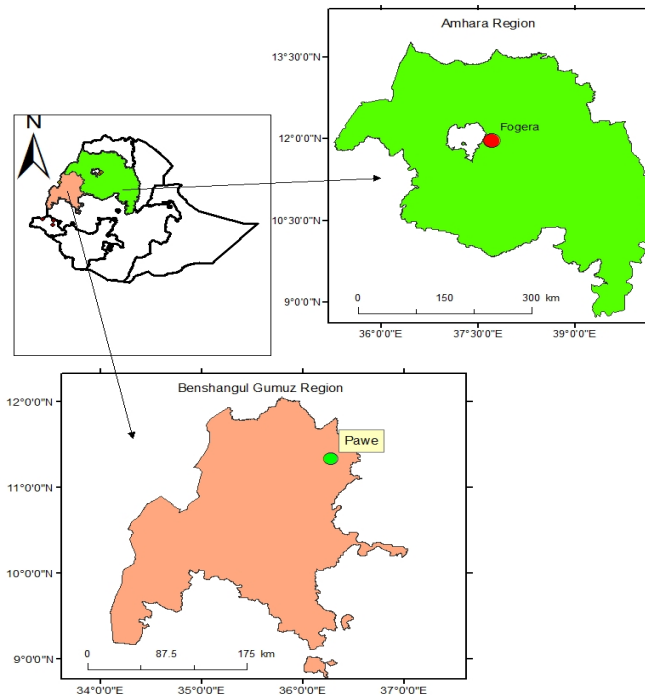


**Figure 1:** Rice accessions collection sites in Ethiopia

These accessions were collected with kind cooperation of rice researchers at Fogera, Pawe, Assosa, Bonga, and Gambella research centers. Improved rice varieties included the study were released by different research centers in Ethiopia. NERICA-10 was not approved for release in Ethiopia but released in other African countries. They are from diverse production systems; rain fed upland and lowland as well as intermittent irrigated types. Improved varieties included NERICA varieties that are adapted to rain fed upland and irrigated conditions.

## 2.2 Description of experimental sites

The field experiment was carried out at two rice research stations, Fogera and Pawe (Figure 2). Fogera is found in Amhara Region and Pawe in Benshangul Gumuz Region, both in Northwest Ethiopia. Fogera station is positioned at  $11^{\circ}58'N$ , and  $37^{\circ}41'E$  at elevation of 1810 meter above sea level. Predominantly lowland rain fed rice is grown. It is characterized by high elevation and low temperature. Fogera areas frequently face terminal moisture stress. The soil is vertisol with slightly acidic pH of 5.90 (Tilahun *et al.* 2013). Rainfall of the area is unimodal, mainly from June to October, amounting to 1234.5mm. Important rain fall months at Fogera ranges from mid-June to early October which covers more than 90% of annual rainfall. Pawe station is located at  $11^{\circ}19'N$  and  $34^{\circ}24'E$ , at elevation of 1120 meter above sea level. Upland rain fed is the major production domain in Pawe areas. This station has relatively high temperature and long rainy season. Rainfall of the areas is unimodal, mainly from April to October amounting to 1570.3mm. Important rain fall months at Pawe ranges from late April to September which covers more than 90% of annual rainfall.



**Figure 2:** Field experiment sites in Amhara and Benshangul Gumuz Regions, Northwest Ethiopia.

## 2.3 Experimental design

At each experimental site, the trial was laid down in a  $15 \times 4$  alpha lattice design of three replications, each replication with four blocks. Each cultivar was sown in six rows of plot size  $7.5 \text{ m}^2$  ( $5\text{m} \times 1.5\text{m}$ ). Spacing between replications, blocks, plots, and rows was 1.5m, 1m, 50cm, and 25cm, respectively. Fertilizers (Urea and DAP) were applied as per to the local recommendation. DAP was applied all at planting and Urea in three splits (planting, tillering and panicle initiative) Weeding and other field management operations were effective accordingly uniformly across experimental plots as required.

## 2.4 Data collection and statistical analysis

Data were collected from the middle four rows excluding two border rows. Ten quantitative traits were assessed and data recorded for each cultivar per block per replication: Days to 50% heading (DTH, days), days to 85% maturity (DTM, days), plant height (PH, cm), panicle length (PL, cm), filled grains per panicle (FGP, count), thousand seed weight (TSW, g), grain yield (Gy, t/ha), biomass yield (By, t/ha), and harvest index (HI, %), following the guidelines developed by IIRI (1996). Except for DTH, DTM, Gy, By, HI, and TSW, for the other traits data were collected from ten randomly selected plants per plot. Grain yield and biomass yield per plot was sampled from the middle four rows by cutting all plants from the bottom and sundried within bags. Weight of dried sample for each plot was measured to obtain above ground biomass and converted in tonnes per hectare (t/ha). For grain yield, each sample was threshed and weighed and adjusted at 14% moisture content in t/ha. Harvest index (HI, %) was calculated a ratio of grain yield to total above ground biomass yield.

Statistical analysis was carried out for each site and for the combined data. All collected data were subjected to analysis of variance using SAS software version 9.0 (SAS 2002). Phenotypic correlations coefficients were computed to elucidate relationships between quantitative traits by using PROC CORR

procedure in SAS. Principal component analysis (PCA) was performed using GenStat software version 16 (GenStat 2013). PCA was employed to identify the different quantitative traits that contributed the most to total variance in the measured variables. In PCA, the raw data for each variable were standardized and the distance matrix using the correlation matrix was computed and the proportion of variance criterion was used to identify the different principal components that contributed to the total variance in the dataset. Ward's minimum variance method was applied for cluster analysis of rice cultivars (Ward 1963).

### 3. Results and discussion

#### 3.1 Correlation analysis

The degree of correlation among different traits is very important to deal with complex traits such as grain yield (Anyaocha *et al.* 2018). It is often reported that yield in rice is determined by indirect traits like plant height, growth period, tillering ability, panicle length, seed length, seed setting rate, and grains per panicle as well as direct traits like panicle number per unit area and/or per plant, filled grains per panicle and 1000 grain-weight (Sakamoto and Matsuoka 2008; Huang *et al.* 2013; Anyaocha *et al.* 2018; Li *et al.* 2019). Correlation results of 10 quantitative agro-morphological traits are presented in Table 1 and Table 2. The correlation between days to heading and days to maturity was found to be significantly positive at Fogera ( $r = 0.73$ ), Pawe ( $r = 0.71$ ) and in combined data ( $r = 0.69$ ), but both traits tended to show significant negative correlation with all other traits considered at Fogera and in combined data. On the other hand, at Pawe, days to heading and days to maturity revealed positive significant association with filled grains per panicle ( $r = 0.37, 0.37$ ) and grain yield ( $r = 0.27, 0.29$ ), respectively while they showed significant negative association with thousand seed weight ( $r = -0.23, -0.24$ ), but weekly correlated with panicle length, plant height and biomass yield (Table 1). The positive significant association of days to heading and maturity to grain yield and yield attributed traits at Pawe suggested that relatively late and high yielding varieties can be selected for Pawe due to the presence of long rainy season with optimum temperature. Earliness is not farmers' top priority at Pawe (personal informal interview). Most farmers preferred high yielding and medium to late type varieties. Farmers explained that early type varieties are subjected to birds attack and get sprout in the long rainy season and hence not preferred.

**Table 1:** Correlation coefficient analysis of yield and yield attributing traits in 60 rice cultivars evaluated at Fogera (bellow diagonal) and Pawe (above diagonal)

Traits	DTH	DTM	PL	PH	FGP	FR	Gy	TSW	By	HI
DTH		0.71***	0.11	0.17	0.37*	0.12	0.27*	-0.23*	0.19	-0.02
DTM	0.73***		0.12	0.13	0.37*	0.21*	0.29*	-0.24*	0.19	0.03
PL	-0.34*	-0.31*		0.49**	0.45**	0.01	0.24*	-0.19	0.20*	0.05
PH	-0.24*	-0.21*	0.60**		0.1	0.03	0.34*	-0.06	0.48**	-0.07
FGP	-0.34*	-0.30*	0.41**	0.34*		0.16	0.04	-0.08	-0.02	0.15
FR	-0.46**	-0.44**	0.31*	0.37*	0.53**		0.10	-0.09	0.14	0.02
Gy	-0.34*	-0.38*	0.41**	0.45**	0.25*	0.30*		0.22*	0.59**	0.16
TSW	-0.25*	-0.27*	0.09	0.25*	0.05	0.13	0.16		0.28*	0.04
By	-0.25*	-0.25*	0.41**	0.52**	0.29*	0.30*	0.81***	0.07		-0.31*
HI	-0.51**	-0.57**	0.36*	0.29*	0.23*	0.34*	0.66**	0.17	-0.27*	

\*, \*\*, \*\*\*: Significant at 0.05, 0.01, and 0.001 levels, respectively. DTH: days to heading, DTM: days to maturity, PL: panicle length (cm), PH: plant height (cm), FGP: filled grains per panicle (no.), Gy: grain yield (t/ha), TSW: 1000 seed weight (g), By: biomass yield (t/ha), HI: harvest index (%).

However, at Fogera early maturing and high yielding rice varieties are preferred by all farmers due to recurrent terminal moisture stress. In addition, early type varieties in Fogera areas create favorable condition for double cropping using residual moisture, a common practice after rice. Traits such as panicle length ( $r = 0.41$ ), plant height ( $r = 0.45$ ), filled grains per panicle ( $r = 0.25$ ), fertility rate ( $r = 0.137$ ) and harvest index ( $r = 0.66$ ) showed significant positive correlations with the grain yield at Fogera with similar trend at Pawe and in combined data (Tables 1, 2). Such positive significant correlations between grain yield and other important traits panicle length, plant height, filled grains per panicle, fertility rate and biomass yield suggested that selection in favor of these traits may lead to positive indirect selection for high grain yield (Ahmadikhah 2010). Thousand seed weight was poorly correlated with other traits except for plant height ( $r = 0.25$ ) and grain yield ( $r = 0.16$ ) at Fogera and for grain yield ( $r = 0.22$ ) at Pawe.

The trait, biomass yield was found to have a significant positive correlation with panicle length, plant height, filled grains per panicle, fertility rate and grain yield both at Fogera and in combined data (Tables 1, 2). At Pawe, biomass yield was weakly but positively correlated with days to heading, days to maturity and fertility rate while it exhibited significant positive association with panicle length ( $r = 0.20$ ), plant height ( $r = 0.48$ ), grain yield ( $r = 0.59$ ) and thousand seed weight ( $r = 0.28$ ) (Table 1). Harvest index exhibited significant positive correlation with panicle length ( $r = 0.36$ ), plant height ( $r = 0.29$ ), filled grains per panicle ( $r = 0.23$ ), fertility rate ( $r = 0.34$ ), and

grain yield ( $r= 0.66$ ) but negatively correlated with biomass yield ( $r= -0.27$ ) at Fogera but at Pawe and in combined data it was poorly correlated with most other quantitative traits except for biomass yield which had significant negative association ( $r= -0.31, -0.25$ ) (Tables 1, 2). Similar result was reported by Lakshmi *et al.* (2019) who evaluated 31 rice accessions and found that grain yield per plant had significant positive correlation with days to flowering (0.479) and spikelets per panicle (0.497), while a negative association was found with 1000 seed weight (-0.294). They further explained that 1000 seed weight was found to have a significant negative association with days to flowering and a negative relation with the remaining traits considered. Konate *et al.* (2016) also reported that grain yield/ plant showed a positive significant association with day to 50% flowering ( $r=0.3997$ ), plant height ( $r=0.2794$ ), panicle weight ( $r=0.112$ ) and biomass ( $r=0.9291$ ) and they also found that plant height had positive correlation with all the characters except with number of fertile spikelets and in turn number of fertile spikelets had negative correlation with almost all the characters except 1000 grains weight.

**Table 2:** Correlation coefficient analysis of yield and yield attributing traits in 60 rice cultivars combined across two sites

Traits	DTH	DTM	PL	PH	FGP	FR	Gy	TSW	By
DTM	0.69**								
PL	-0.20*	-0.33*							
PH	-0.15	-0.33*	0.67**						
FGP	-0.07	-0.33*	0.54**	0.43**					
FR	-0.35*	-0.49**	0.35*	0.42**	0.50**				
Gy	-0.07	-0.08	0.28*	0.32*	0.054	0.17			
TSW	-0.23*	-0.19	-0.06	0.07	-0.05	0.04	0.02		
By	-0.06	-0.22*	0.37*	0.57**	0.28*	0.32*	0.57**	0.15	
HI	-0.13	-0.06	0.06	-0.04	0.03	0.04	0.30*	0.08	-0.25*

\*, \*\*, \*\*\* significant at 0.05, 0.01, and 0.001 levels, respectively. DTH: days to heading, DTM: days to maturity, PH: plant height(cm), PL: panicle length(cm), FGP: filled grains per panicle, FR: fertility rate(%), Gy: grain yield(t/ha), By: biomass yield(t/ha), TSW: thousand seed weight(g), and HI: harvest index(%).

### 3.2 Principal component analysis

Principal Component Analysis (PCA) was computed based on 10 quantitative agro-morphological traits of 60 rice cultivars at each site and in the combined data (Table 3). Principal components were computed from the correlation matrix and genotypic scores obtained for the first component and succeeding components with Eigen values greater than unity (Jeger *et al.* 1983). The PCA measures the contribution of each component or independent impact of a particular trait to the total variance observed in a given population in relation to the traits of interest to the breeder (Anyaocha *et al.* 2018). In the present study, only the first four principal components (PC1 to PC4) had Eigen values greater than one in all the three data sets such as Fogera, Pawe and in combined data set. The first four principle components (PCs) were responsible to explain a cumulative variation of 81.07%, 74.95% and 80.63% at Fogera, Pawe and in combined data, respectively (Table 3). In previous study, Chakravorty *et al.* (2013) found that the first six principal components in the PCA analysis with Eigen values greater than one contributed 75.9% of the variability among 51 rice landraces evaluated for different agro-morphological traits in which traits such as leaf length (0.625), plant height (0.751), culm diameter (0.780), culm number (0.802) and panicle length (0.697) accounted for most of the observed variability in PC1 (23.47%). On the other hand, Lakshmi *et al.* (2019) reported that only the first two principal components with Eigen values of 2.643 and 1.189, respectively, together explained 54.752% of the total variance for all the characters in 31 rice accessions and characters such as yield (0.809), days to 50% flowering (0.764) and number of spikelets per panicle (0.669) had higher contributions in PC1 (37.763%), while the PC2 accounted for 16.988% of the total variation with panicle length (0.541) giving the highest contribution.

At Fogera, principal component one (PC1), with Eigen value of 4.804, contributed 43.67% of the total variation and traits that positively contributed to the variation comprised plant height (0.278), panicle length (0.327), filled grains per panicle (0.373), fertility rate (0.355), grain yield (0.335) and biomass yield (0.304), while negatively contributing traits included days to heading (-0.334) and days to maturity (-0.349). PC2 with Eigen value of 1.858 accounted for 16.89% of total variation with greatly contributing traits including days to heading (0.286), days to maturity (0.315), plant height (0.279), panicle length (0.264), thousand seed weight (-0.285), biomass yield (0.299) and harvest index (-0.523). PC3 and PC4 also had Eigen values of 1.217 and 1.038, respectively and contributed 11.07% and 9.44% to the observed variation, respectively where traits such as plant height, grain yield, and harvest index contributed to both PCs while fertility rate (0.258) and thousand seed weight (0.521) only to PC3 and days to heading (0.352), filled grains per panicle (-0.314), and biomass yield (0.360) to PC4 (Table 3).

At Pawe, PC1 explained 31.30% of the total variance with Eigen values of 3.44 and only four quantitative traits such as days to heading (0.347), days to maturity (0.349),

**Table 3:** Principal component analysis of 10 quantitative traits for 60 rice cultivars evaluated at Fogera and Pawe

Traits	Fogera				Pawe				Combined			
	PC 1	PC 2	PC 3	PC 4	PC 1	PC 2	PC 3	PC 4	PC 1	PC 2	PC 3	PC 4
Eigen value	4.804	1.858	1.217	1.038	3.44	1.96	1.73	1.11	3.186	2.579	1.848	1.255
Variation (%)	43.67	16.89	11.07	9.44	31.30	17.83	15.73	10.09	28.97	23.45	16.8	11.41
Cumulative (%)	43.67	60.56	71.63	81.07	31.30	49.13	64.86	74.95	28.97	52.42	69.22	80.63
Eigen vector												
Days to heading	<b>-0.334</b>	<b>0.286</b>	-0.105	<b>0.352</b>	<b>0.347</b>	0.028	<b>-0.512</b>	0.043	<b>-0.462</b>	0.107	0.218	<b>0.257</b>
Days to maturity	<b>-0.349</b>	<b>0.315</b>	-0.050	0.235	<b>0.349</b>	0.066	<b>-0.457</b>	0.015	<b>-0.470</b>	0.111	0.194	0.225
Plant height(cm)	<b>0.278</b>	<b>0.279</b>	<b>0.411</b>	<b>0.319</b>	0.089	<b>0.510</b>	0.161	0.035	0.238	<b>0.282</b>	<b>0.401</b>	-0.185
Panicle length(cm)	<b>0.327</b>	<b>0.264</b>	0.104	0.152	0.174	<b>0.261</b>	<b>0.453</b>	<b>-0.413</b>	0.248	<b>0.460</b>	0.033	-0.147
Filled grains/panicle	<b>0.373</b>	0.243	-0.240	<b>-0.314</b>	<b>0.476</b>	-0.128	0.181	-0.107	-0.145	<b>0.485</b>	<b>-0.335</b>	-0.246
Fertility rate (%)	<b>0.355</b>	-0.039	<b>0.258</b>	-0.054	0.217	<b>-0.273</b>	-0.197	0.075	<b>0.413</b>	0.004	-0.057	0.005
Grain yield (t/ha)	<b>0.335</b>	-0.242	<b>-0.286</b>	<b>0.455</b>	<b>0.395</b>	0.109	0.186	<b>0.352</b>	<b>0.301</b>	<b>0.291</b>	-0.007	<b>0.578</b>
1000 grain weight (g)	0.125	<b>-0.285</b>	<b>0.521</b>	-0.176	-0.062	-0.051	0.254	<b>0.787</b>	0.218	-0.237	0.073	<b>-0.374</b>
Biomass yield (t/ha)	<b>0.304</b>	<b>0.299</b>	0.019	<b>0.360</b>	0.163	<b>0.570</b>	-0.075	0.212	0.178	<b>0.299</b>	<b>0.513</b>	0.165
Harvest index (%)	0.178	<b>-0.523</b>	<b>-0.366</b>	<b>0.297</b>	0.204	<b>-0.476</b>	<b>0.275</b>	0.077	0.207	-0.006	<b>-0.519</b>	<b>0.451</b>

filled grains per panicle (0.476), and grain yield (0.395) exhibited higher contribution to the total variation while the remaining traits tended to show lower contribution. PC2 with Eigen value of 1.96 accounted for 17.83% of the variation and three traits; plant height (0.510), panicle length (0.261) and biomass yield (0.570) contributed positively to the overall variation while fertility rate (-0.273) and harvest index (-0.476) had a negative contribution. According to PC3, with Eigen value of 1.73, traits such as days to heading (-0.512), days to maturity (-0.457), panicle length (0.453), and harvest index (0.275) had significantly higher contributions to the total phenotypic variability of 15.73%, while the PC4 accounted for 10.09% of the total variation with panicle length (-0.413), grain yield (0.352) and particularly that of thousand seed weight (0.787) had the highest contribution to the overall variance as observed among 60 rice cultivars (Table 3).

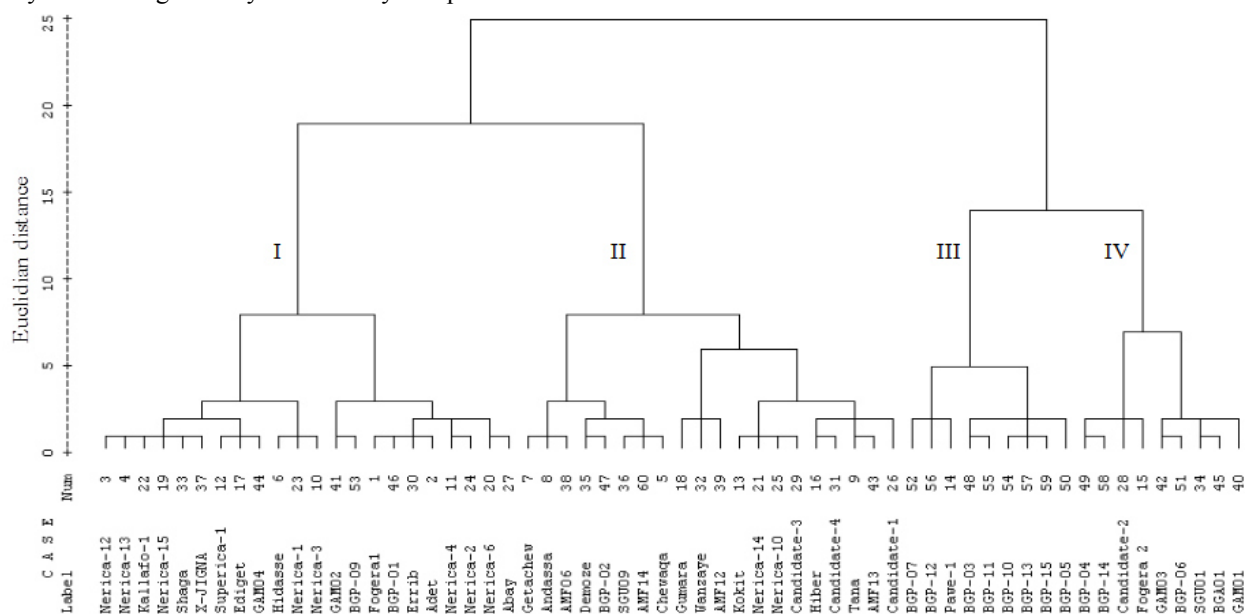
In combined data, the first principal component (PC1) with Eigen value of 3.186 explained 28.97% of the total variance, while PC2 with Eigen value of 2.579 accounted for 23.45% of total variability. Four traits that contributed the highest to the overall variation in PC1 included days to heading (-0.462), days to maturity (-0.470), fertility rate (0.413) and grain yield (0.301), while plant height (0.282), panicle length (0.460), filled grains per panicle (0.485), grain yield (0.291) and biomass yield (0.299) contributed the highest in PC2. PC3 had Eigen value of 1.848 and contributed 16.80% to the overall variability. Similarly, PC4 had contributed 11.41% of the total variability with Eigen value of 1.255. While plant height (0.401), filled grains per panicle (-0.335), biomass (0.513) and harvest index (0.451) contributed the highest to the variation in PC3, the other four quantitative traits such as days to heading (0.257), grain yield (0.578), thousand seed weight (-0.374) and harvest index (0.451) contributed largely to the variation reflected in PC4 (Table 3).

### 3.3 Cluster analysis

An attempt was made using Ward's minimum variance method of hierarchical clustering based on eleven quantitative traits combined across two sites (Fogera and Pawe) to visualize the patterns of grouping of 60 rice cultivars. Accordingly, the sixty cultivars were grouped into four different clusters as illustrated in Figure 3. Cluster I was the largest cluster which included 22 (36.7%) cultivars, followed by Cluster II which comprised 20 (33.3%) cultivars. Clusters III and IV included nine (15%) cultivars each dominated by collected accessions (Figure 3 and Table 4). About 77% of cultivars in cluster I comprised improved varieties including Nericas. X-Jigna and other four accessions also belonged to this cluster. Days to heading, days to maturity, plant height, filled grains per panicle, grain yield and biomass yield contributed highly to the genetic diversity which should be given importance during hybridization and selection in segregating populations by choosing plants with early duration (Fogera areas), medium to late duration (Pawe areas) and more grain yield per plant. Previous study reported by Lakshmi *et al.* (2019) classified 31 rice accessions based on their Euclidean distances into six clusters with the first cluster included the maximum number of accessions and they explained that panicle length

and days to 50% flowering contributed high variation. Similar results of grouping the rice accessions into six clusters was reported by Pathak *et al.* (2020) while studying 32 rice accessions using Euclidian distance analysis and the highest inter-cluster distance was observed between clusters III and VI.

Cultivars in cluster I were the earliest in days to heading and days to maturity with intermediate mean values for plant height, panicle length, grain yield, and biomass yield but the highest mean values for filled grains per panicle and harvest index (Table 5). Cultivars in cluster II were also relatively early in days to heading and days to maturity with higher mean values for thousand seed weight, grain yield and biomass yield while intermediate for other traits including plant height, panicle length, filled grains per panicle, fertility rate and harvest index (Table 5). Cluster III comprised rice cultivars entirely from Pawe area while cluster II and IV were composed from accessions and improved cultivars of upland and lowland types. Cultivars belonging to cluster III were late in days to heading and days to maturity, and they had the lowest mean values for plant height, panicle length, fertility rate, grain yield, biomass yield and harvest index but intermediate for filled grains per panicle and thousand seed weight (Table 5). Cluster IV included cultivars with intermediate mean values of harvest index, biomass yield and fertility rate but higher mean values for plant height, panicle length, filled grains per panicle, and grain yield and the lowest for thousand seed weight. Cultivars in this cluster were late in days to heading and days to maturity compared to those in clusters I and II.



**Figure 3:** Hierarchical clustering of 60 rice cultivars using Ward’s method constructed based on Euclidean distance estimated from 10 quantitative traits evaluated at Fogera and Pawe.

The average intra-cluster and inter-cluster distance observed in the present clustering procedure of 60 rice cultivars is shown in Table 6. The result demonstrated that the maximum intra-cluster distance was observed in cluster II (20.51), followed by cluster IV (20.18) and intermediate within that of cluster III (17.94). The cluster with maximum number of cultivars, cluster I, showed intra-cluster distance of 14.22 which was the lowest of all intra-cluster distances observed (Table 6). This low intra-cluster distance indicated that the cultivars contained in the cluster were closely related (Sharma *et al.* 2020).

**Table 4:** Distribution of 60 rice cultivars by clusters evaluated at Fogera and Pawe

Cluster	No. of cultivars	Proportion (%)	Cultivars
I	22	36.7	Fogera1, Adet, Nerica-12, Nerica-13, Hidasse, Nerica-3, Nerica-4, Superica-1, Ediget, Nerica-15, Nerica-6, Kallafo-1, Nerica-1, Nerica-2, Abay, Errib, Shaga, X-Jigna, GAM02, GAM04, BGP-01, BGP-09
II	20	33.3	Chewaqa, Getachew, Andassa, Tana, Kokit, Hiber, Gumara, Nerica-14, Nerica-10, Candidate-1, Candidate-3, Candidate-4, Wanzaye, Demoze, SGU09, AMF06, AMF12, AMF13, BGP-02, AMF14
III	9	15.0	Pawe-1, BGP-03, BGP-05, BGP-07, BGP-10, BGP-11, BGP-12, BGP-13, BGP-15
IV	9	15.0	Fogera 2, Candidate-2, SGU01, GAM01, GAM03, BGA01, BGP-04, BGP-06, BGP-14

On the other hand, the largest intra-cluster distance observed in clusters II and IV, regardless of the number of cultivars, could be attributed to the heterogeneous nature of the cultivars within each cluster. The inter-cluster distance was found to be maximum between clusters II and IV (45.79), followed by clusters III and IV (35.70) and between clusters I and III (35.62). Minimum inter-cluster distance was obtained between cluster I and cluster II (25.33) while it was intermediate distance between II and III (30.95), and between clusters I and IV (30.92). In all the cases, the inter-cluster distances observed was larger than that of intra-cluster distances (Table 6). Therefore, crossing between cultivars in clusters II and IV could result higher heterosis, followed by between clusters III and IV for particular traits of interest targeted. Similar result was reported by Pathak *et al.* (2020) indicating the highest inter-cluster distance between clusters III and VI and hence, hybridization between the genotypes of these clusters would yield desirable segregates with the accumulation of favorable genes in the segregating generations

**Table 5:** Intra-cluster mean values for 10 quantitative traits of 60 rice cultivars evaluated at Fogera and pawe

Traits	Cluster means				Mean
	I	II	III	IV	
Days to heading	79.2	81.8	99.8	97.3	89.5
Days to maturity	121.5	124.0	141.6	140.4	131.9
Plant height (cm)	90.3	93.5	79.7	96.1	89.9
Panicle length (cm)	19.5	19.5	17.4	20.2	19.2
Filled grain/panicle	105.9	89.3	92.9	115.5	100.9
Fertility rate (%)	92.0	92.5	86.9	90.6	90.5
1000 grain weight(g)	27.0	27.3	26.1	24.9	26.3
Grain yield (t/ha)	2.9	3.1	2.3	3.1	2.8
Biomass yield (t/ha)	8.0	9.1	7.4	8.8	8.3
Harvest index (%)	37.9	36.2	32.9	34.7	35.4

**Table 6:** Average intra-cluster (bold diagonal) and inter-cluster distance among 60 rice cultivars evaluated at Fogera and Pawe.

Cluster	I	II	III	IV
I	<b>14.22</b>	25.33	35.62	30.92
II		<b>20.51</b>	30.95	45.79
III			<b>17.94</b>	35.70
IV				<b>20.18</b>

#### 4. Conclusion

The current results obtained from sixty rice cultivars grown at two sites showed that grain yield was significantly and positively correlated with panicle length, plant height, filled grains per panicle, spikelets fertility rate, biomass yield, harvest index and thousand seeds weight at Pawe and Fogera. Principal component analysis identified the traits days to heading, days to maturity, panicle length, filled grains per panicle, grain yield, biomass yield and harvest index contributing towards the maximum divergence among sixty rice cultivars evaluated. Cluster analysis also classified rice cultivars into four different groups attributed to different traits.



Cultivars included in cluster I were early maturing and with good panicle length, large number of filled grains per panicle, good fertility rate, high thousand seed weight and with high grain yield as well as largest harvest index where some of the cultivars in clusters III and IV showed low performance of these important traits. Selection of contrasting parents for hybridization program should be on the basis of the magnitude of genetic distance, contribution of different traits towards the total variance, patterns of clustering and magnitude of cluster means for different traits performance. Therefore crosses between selected cultivars of clusters I and III, I and IV, II and III, II and IV, III and IV could result in high heterosis as well as wide range of genetic variation in the subsequent generations of the crossing. Overall, results from principal component and clustering analysis of this study demonstrated presence of genetic variation among Ethiopian rice cultivars for quantitative traits to be exploited differently and /or concurrently for the improvement of rice varieties at the two contrasting environments.

### Acknowledgements

The author greatly acknowledges the support provided by Ethiopian Institute of Agricultural Research, Fogera Rice Research and Training Center for providing the resources for conducting the experiment. Rice breeding team at Fogera Rice Research and Training Center is also highly appreciated for supporting the execution of the field experiment.

### References

- Ahmadikhah A (2010). Study on selection effect, genetic advance and genetic parameters in rice. *Annals of Biological Research*, 4: 45-51.
- Anyaocha C., Adegbehingbe F., Uba U., Popoola B., Gracen V., Mande S., Onotugoma E., Fofana M (2018). Genetic diversity of selected upland rice genotypes (*O. sativa* L.) for grain yield and related traits. *International Journal of plant & soil science*, 22(5): 1-9.
- Bitew J.M., Mekbib F., Assefa A (2018). Correlation coefficient and path analysis among yield and yield related traits in upland rice (*Oryza sativa* L. and *Oryza glaberrima* Steud) genotypes in Northwestern Ethiopia. *Int. J. Plant Breed. and Crop Sci.* 5: 429-436.
- Bitew J.M., Mekbib F., Assefa A (2016). Genetic variability among yield and yield related traits in selected upland rice (*Oryza sativa* L. and *Oryza glaberrima* Steud) genotypes in Northwestern Ethiopia. *World Sci. News*, 47: 62-74.
- CSA (Central Statistical Agency Vol 1) (2020). Report on area and production of major crops (private peasant holdings, meher season). The Federal Democratic Republic of Ethiopia, ADDIS ABABA, Ethiopia.
- Chakravorty A., Ghosh P.D., Sahu P.K (2013). Multivariate Analysis of Phenotypic Diversity of Landraces of Rice of West Bengal. *American Journal of Experimental Agriculture*, 3(1):110-123.
- FAOSTAT (2020). Food and Agriculture Organization of the United Nations (FAO). [www.fao.org/faostat/en/#data](http://www.fao.org/faostat/en/#data).
- GenStat 2013. GenStat Statistical software. GenStat for Windows 16th Edition. VSN International, UK.
- Girma B.T., Kitil M.A., Banje D.J., Biru H.M., Serbessa T.B (2018). Genetic variability study of yield and yield related traits in rice (*Oryza sativa* L.) genotypes. *Adv Crop Sci and Tech*, 6: 381. doi:10.4172/2329-8863.1000381.
- Hallauer A.R., Miranda J.B (1988). Quantitative genetics in maize breeding. (2<sup>nd</sup> Ed.), Iowa State University Press, Ames, I.A.
- Huang R., Jiang L., Zheng J., Wang T., Wang H., Huang Y (2013). Genetic bases of rice grain shape: so many genes, so little known. *Trends Plant Sci.* 18, 218-226, doi: 10.1016/j.tplants.2012.11.001.
- IRRI (International Rice Research Institute) (1996). Standard Evaluation System for Rice (SES), Manila, Philippines.
- Jeger M.I., Garethojones D., Griffiths E (1983). Genetic divergence in wheat. *Indian J. Genet.* 43: 63-67.
- Konate A.K., Zongo A., Kam H., Sanni A., Audebert A (2016). Genetic variability and correlation analysis of rice (*Oryza sativa* L.) inbred lines based on agro- morphological traits. *African Journal of Agricultural Research*, 11(35): 3340-3346, DOI: 10.5897/AJAR2016.11415.
- Lakshmi V.G.I., Sreedhar M., Vanisri S., Anantha M.S., Subba Rao L.V., Gireesh C (2019). Multivariate analysis and selection criteria for identification of African rice (*Oryza glaberrima*) for genetic improvement of indica rice cultivars. *Plant Genetic Resources: Characterization and Utilization*, 1-7, doi:10.1017/S1479262119000327.
- Li R., Li M., Ashraf U., Liu S., Zhang J (2019). Exploring the Relationships Between Yield and Yield-Related Traits for Rice Varieties Released in China From 1978 to 2017. *Front. Plant Sci.* 10:543. doi: 10.3389/fpls.2019.00543.
- Mahendran R., Veerabathiran P., Robin S., Raveendran M (2015). Principal component analysis of rice germplasm accessions under high temperature stress. *International Journal of Agricultural Science and*

- Research*, 5 (3):355-360.
- Maji A., Shaibu A.A (2012). Application of principal component analysis for rice germplasm characterization and evaluation. *Journal of Plant Breeding and Crop Science*, 4(6): 87-93, DOI: 10.5897/JPBCS11.093.
- Pathak P., Singh S.K., Korada M., Habde S., Singh D.K., Khaire A., Majhi P.K (2020). Genetic characterization of local rice (*Oryza sativa* L.) genotypes at morphological and molecular level using ssr markers. *Journal of Experimental Biology and Agricultural Sciences*, 8(2):148-156.
- Qamar Z.U., Akhtar J., Ashraf M., Akram M., Hameed A (2012). A multivariate analysis of rice genetic resources. *Pak. J. Bot*, 44(4): 1335-1340.
- Riaz A., Bibi T., Raza Q., Sabar M., Akhter M. (2018). Evaluation of rice (*Oryza sativa* L.) advance uniform lines using multivariate analysis. *African Journal of Plant Science*, 12(11): 284-289, DOI: 10.5897/AJPS2018.1665.
- Sakamoto T., Matsuoka M (2008). Identifying and exploiting grain yield genes in rice. *Curr. Opin. Plant Biol.* 11, 209-214.
- SAS (2002) Statistical Analysis Software. Version 9.0 SAS Institute INC. Cary, North Carolina, USA.
- Sharma S., Pokhrel A., Dhakal A., Poudel A (2020). Agromorphological characterization of rice (*Oryza sativa* L.) landraces of Lamjung and Tanahun district, Nepal,” *Annals of Plant Sciences*, 9 (2): 3731-3741.
- Tejaswini K.L.Y., Manukonda S., Ramana R. P.V., Kumar S.R.R., Mohammad L.A., Raju S.K (2016). Cluster analysis studies in rice (*Oryza sativa* L.) using wards minimum variance method. *Journal of Agricultural and Crop Research*, 4(9): 129-139.
- Tilahun T., Nigussie D., Wondimu B., Setegn G (2013). Effect of Hydro-Priming and Pre-Germinating Rice Seed on the Yield and Terminal Moisture Stress Mitigation of Rain-Fed Lowland Rice. *Agriculture, Forestry and Fisheries* 2: 89-97.
- Ward J.H (1963). Hierarchical Grouping to Optimize an Objective Function. *J. Am. Stat. Assoc.* 58: 236-244.