Assessing The Representative And Discriminative Ability Of Test Environments For Rice Breeding In Malaysia Using GGE Biplot

Yusuff Oladosu, M.Y. Rafii, Usman Magaji, Norhani Abdullah, Asfaliza Ramli, Ghazali Hussin

Abstract: Identification of outstanding rice genotype for target environments is complicated by genotype × environment interactions. Using genotype main effect plus genotype by environment interaction (GGE) Biplot software, fifteen rice genotypes were evaluated at five locations representing the major rice producing areas in peninsula Malaysia in two cropping seasons to (i) identify ideal test environment for selecting superior rice genotype, and (ii) identify discriminative and representative ability of test locations. Genotypes, locations, years, and genotypes by environment interaction effect revealed high significant difference ($P < 0.01$) for number of tillers per hill, grains per panicle, grain weight per hill, and yield per hectare. Grain yield per hectare had a non-repeatable crossover pattern that formed a complex and single mega-environment. Based on the crossover pattern, a set of cultivars were selected for the whole region on the merit of mean performance and their stability analysis. The tested environments were divided into two mega-environments. An ideal test environment that measures the discriminative and representative ability of test location reveal that environment Sekinchan SC is the best environment, while Kedah KD and Penang PN can also be considered as favorable environment whereas Serdang SS and Tanjung Karang TK were the poorest locations for selecting genotypes adapted to the whole region. This study serves a reference for genotypes evaluation as well as identification of test locations for rice breeding in Malaysia.

Keywords: Genotype by Environment interaction, stability analysis, GGE biplot, test location, grain yield.

1 INTRODUCTION

Rice (Oryza sativa L.) is an important food crop that serve as staple foods for over 3.5 billion of the world's population [1]. Appro ximately, 90% of world's rice is produced in Asia largely in China and India. In Malaysia, rice is regarded as security crop, where the government encourages commercialized rice production [2]. The total harvested paddy in Malaysia in 2016 was estimated at 2.6 million tonnes. At present, 70-73% of rice is produced locally while the remaining 30-27% (1.2 million tonnes) shortfall under importation with Thailand and Vietnam supplying more than 60% of the imported rice, while the remaining 40% was imported from Pakistan, India, and Cambodia [3].

- *Yusuff Oladosu: PhD, Institute of Tropical Agriculture and Food Security, Universiti Putra Malaysia (UPM), 43400 Serdang, Selangor, Malaysia. oladosuy@gmail.com*
- *M.Y. Rafii: PhD, Institute of Tropical Agriculture and Food Security, Universiti Putra Malaysia (UPM), 43400 Serdang, Selangor, Malaysia mrafii@upm.edu.my*
- *Magaji G. Usman: PhD, Institute of Tropical Agriculture and Food Security, Universiti Putra Malaysia (UPM), 43400 Serdang, Selangor, Malaysia magajiusman0@gmail.com*
- *Norhani Abdullah: PhD, Institute of Tropical Agriculture and Food Security, Universiti Putra Malaysia (UPM), 43400 Serdang, Selangor, Malaysia norhani.biotech@gmail.com*
- *Asfaliza Ramli: PhD, Malaysian Agricultural Research and Development Institute (MARDI), 43400 Serdang, Selangor, Malaysia aliza@mardi.gov.my*
- *Ghazali Hussin: PhD, Malaysian Agricultural Research and Development Institute (MARDI), 43400 Serdang, Selangor, Malaysia gbh@mardi.gov.my*
- *Corresponding author: oladosuy@gmail.com*

As Malaysia is determined to reduce her dependency on rice importation from other countries, the development of stable and high yielding rice varieties for local cultivation is of prime important for enhancing self-sufficiency. Breeders are interested in developing genotypes that are adapted to a wide range of environments. However, changes in environment or location affect both yield and yield component traits due to significant variation caused by genotype × environment interactions (GEI). This variation in relation to genotypic performance in diverse locations complicates selection and testing of superior genotypes, therefore reducing genetic progress [4]. The hindrances in selecting stable varieties associated with G×E interactions includes; difficulty in defining the target population of environments (TPE), choosing suitable test locations representative of the target population, and platforms for precise multi-location testing of a large number of varieties in a defined target population of environments [5]. An effective test location should have a discriminating ability so that genotypes differences can be detected easily using fewer replications. Also, it should be a representative of tested environments so that cultivars with desirable adaptation can be identified prior to release. According to Yan et al. [5], location representative for a target environment should be repetitive in such a way that selected superior genotype should maintain a relative performance in future years. Hence, knowledge of target environment is essential for breeding towards locally adapted genotypes and it is also required in dividing target environment into different mega-environment. Several researchers have proposed a different definition for mega-environment, among which Yan and Tinker [6] define mega-environment as a set of locations whereby a genotype performed consistently across the year. Gauch and Zobel [7] defined mega-environment as part of growing region of crop species where similar genotypes perform equally in a fairly homogenous environment. Braun et al. [8] defined a mega-environment as an extensive, non-contiguous area, occurring in different countries or even transcontinental, having similar biotic and abiotic stresses, consumer preferences, and cropping system. Understanding the concept of mega-environment is very important for research work dealing with

G×E interaction, environmental clustering, wide or narrow adaptation and yield stability [7]. Peterson and Pfeiffer [9], suggested that subdivision of growing area into different mega-environment helps in distribution of breeding program resources. Analysis of multi-environment yield trial data on a set of different varieties is generally used for the identification of mega-environment. Understanding the pattern of genotype by location interaction (GLI) in a target region and exploring the feasibility of dividing it into meaningful mega-environments are the main objectives of mega-environment analysis. Methods proposed for multi-environment trial data includes GGE biplot model, additive main effects and multiplicative interaction (AMMI) model, cluster analysis and correlation among environments [10]. Among these methods, GGE biplot is the most common and recently model for GEI analysis. The GGE biplots model is a graphical display of two-way data matrix (i.e. genotype-by-environment) that allows the visualization of genotypes by environments interactions [11]. Genotype main effect and GEI effect are the two major source of variation in GGE biplot that is applicable to identification mega-environment and genotype evaluation [11]. Therefore, this study was undertaken to analyze multi-environment trial data on rice yield with the objectives of (i) measure the correlation among locations and identify redundant test locations to improve efficiency; (ii) identify locations having high discriminating and representative ability. The outcome of this study will serve as a reference point for genotype evaluation and test locations for identification of rice cultivation in Malaysia.

2 MATERIALS AND METHODS

2.1 Test location and plant materials

Fifteen rice genotypes were evaluated for two years in five locations across the Peninsular Malaysia. The selected locations represent the major rice producing area in Malaysia which covered a wide range of climatic conditions, including Serdang (SS; 3° 02'N 101° 42'E; altitude, 32m; average temperature Min-Max, 24 ^{o}C – 33 ^{o}C; average humidity, 89%; rainfall, 934.7; soil texture, clay loam), Tanjung Karang (TK; 3°25' 0N 101° 10'E; altitude, 3m; average temperature Min-Max, 23 ^{o}C – 31 ^{o}C; average humidity, 83%; rainfall, 782.4; soil texture, clay loam), Penang (PN; 05°25'N 100°15'E; altitude, 3m; average temperature Min-Max, 22 ^{o}C – 30 ^{o}C; average humidity, 88%; rainfall, 934.7; soil texture, silty clay loam), Kedah (KD; 5° 59'N 100° 24'E; altitude, 18m; average temperature Min-Max, 22 ^{o}C – 33 ^{o}C; average humidity, 91%; rainfall, 550.6; soil texture, loam), Sekinchan (SC; 3.77° N, 100.9° E; altitude 41m; average temperature Min-Max 23 ^{o}C – 31 ^{o}C; average humidity 87%; rainfall 782.4; soil texture, loam). The fifteen genotypes consist of six advance mutant lines (ML4, ML6, ML9, ML10, ML21, and ML24) that were promoted from a preliminary study of ion beam irradiation [12], these mutant lines were evaluated in this study for the purpose of seeking registration support. Three mutant varieties (VN001, VN121, and VN124) were collect from Vietnam Atomic Energy Institute, three mutant varieties (Iratom, Binadhan4, and Binadhan7) were collect from Bangladesh Institute of Nuclear Agriculture. Three commercial varieties (MR219, MR220, and MR253) developed by Malaysian agricultural research and development institute were used as a control to validate the performance of the mutant lines.

2.2 Experimental layout and data collection

The experiment was design in a randomized complete block design with three replications across the entire locations. Twenty-one days old seedlings were transplanted to the plot size of 38 by 9 m^2, with subplot size of 2 by 2 m^2 units for each genotype in each replication. Optimum date for transplanting at each location was followed according to the farmer's schedule, other management practices such as fertilizer application were followed based on the standard practices prescribed by Malaysian Agricultural Research and Development Institute, (MARDI). Five plants were sample for each genotype in each replication for the number of tillers per hill, grains per panicle, and grain weight per hill. Yield per hectare was estimated from threshed weight of all panicle in 1.5×1.5 m^2 excluding border rows. The threshed weight was then converted to productivity ton per hectare (t/ha) and subjected to statistical analysis.

2.3 Data analysis

An analysis of variance (ANOVA) was calculated for genotypes, locations, years, and genotypes by locations by years (genotype by environment interaction) for all the observed characters using SASG×E code developed by Dia et al. [13] of SAS program version 9.4 to determine the level of variation. Genotypes, replications, years, and environment (combinations location and year) were considered as random effects. The Rstudio statistical software [14] was used to compute for test location evaluation and mega-environment identification. The graphical display of GGE biplot analysis was computed using 'GGEBiplotGUI' package of Rstudio statistical software [15] for identification of mega-environments, discriminative and representative ability of the test locations, correlation and redundancy among test locations. The graphical display of GGE biplot was constructed using the first and second principal components (PC1 and PC2) derived from subjecting the trait means of environment-centred to singular value decomposition (SVD). For the mega-environment analysis, trait means for each location were subjected to the biplot analysis. For the discriminative and representative of the test locations analyses, the trait mean for each location in each year was used. The biplots were based on standard deviation-standardised (scaling = 0), environment-centred (centering = 2), transformed (Transform = 0), and singular-value partitioning = 2.

3 RESULTS AND DISCUSSION

3.1 Analysis of Variance

The combined analysis of variance (ANOVA) revealed highly significant difference ($P \leq 0.01$) for genotypes, locations, years and genotype by locations by year's (genotype by environment interaction) (Table 1). The estimates of variance components are presented in Table 1. The proportion of the total variation measured from the total sum of square revealed that tillers per hill were controlled to a large extent by locations at 19% followed by the locations by years at 18%, and genotype by locations by years at 16%. Variation as a result of the year accounted for a little portion of variation at 0.25% while genotype accounted for 7%. Similarly, numbers grain per panicle was explained largely by year, locations by years, locations, genotype and genotype by locations by years at 19%, 18%, 16%, 15% and 11% variation respectively. For grain weight per hill largest variation was due to locations

which accounted for 30%, followed by the year at 29% and locations by years at 12%. However, genotypes by years, genotypes, genotypes by locations and genotypes by locations by years accounted for lowest variation at 2%, 3%, 4% and 5% respectively. Yield per hectare recorded the highest variation for genotypes by locations followed by years, genotypes, locations by years, locations and genotypes by locations by years at 19%, 16%, 15%, 14%, 13%, and 11% respectively of the percentage total sum of square (Table 1). Contrarily, genotypes by years show the smallest variation at 3 %. The observed differences in genotype response to varying environmental conditions establish the major constraint for identification of superior rice genotype for wide or narrow adaptation. In this present study, the significant mean square detected for genotypes for most traits indicates different responses of the genotypes to locations by year's interaction. This, however, call for the needs to identify stable and high-yielding genotypes across the test locations [16]. Highly significant interaction in yield per hectare between for genotype by location by year demand the need for an extensive evaluation of genotypes in multi-locations over the years before releasing or recommendations to the farmers. This also confirms the need for breeders to give genotype by environment interaction a serious attention in evaluating genotypes across Peninsular Malaysia and to have an estimate of its magnitude, relative to the magnitude of genotype and environment effects which affect grain yield. Furthermore, Examination of the percentage total sum of squares revealed that the locations and years accounted for most of the variation for all the traits except for the number of tillers. Hence, it revealed the environmental main effects. This finding is in corroboration with previous research findings on multi-environment trials, [16, 17, 18]. Therefore, it is very important to have a better understanding of the target test locations used for the evaluation of rice in Malaysia to determine if it could be subdivided into different mega-environments to facilitate a more meaningful genotypes evaluations and recommendations.

3.2 Yield Performance and Relative Magnitude of Genotype, Location, and G × E Interaction

According to Fan *et al.* [19], if there is a significant difference for location by year interaction, then, there is a need for a separated ANOVA for locations in each year. The ANOVA result for the yearly data that gave an overview of genotype, location, and genotype by location interaction variance was presented in Table 2. For the year 1 yield trials, genotype ML6 produced the highest grain yield of 8.96 t/ha, which was 22% higher than the control MR253 (Table 3). ML4 ranked the second at 8.62 t/ha, 19% higher than MR253. The third rank was ML24 at 8.21 t/ha, 15% higher than MR253. For the year 2 yield trials, ML9 topped the list with grain yield of 8.31 t/ha, which was 17% higher than the control. ML10 ranked the second (7.89 t/ha) and yielded 13% more than the control. ML4 was ranked third in grain yield with an average of 7.70 t/ha which as 10% higher than the control variety. Large variation was attributed to locations for year1 and year2 accounting for 39% and 34%, 28% and 62%, 41% and 64%, 38% and 36% for tillers per hill, grains per panicle, grain weight per hill and yield per hectare respectively (table 2). The magnitude of interaction between genotype and location suggested the possibility of different mega-environment existence. However, despite the level of diversity among the locations and genotypes, the relative contribution of genotype by location is small as compared with location and genotype main effect for tillers per hill, grains per panicle and grain weight per hill (Table 2). However, Annicchiarico [20], suggested that breeding rice for location adaptation for tillers per hill, grains per panicle and grain weight per hill will not be advantageous. Hence, for yield per hectare, genotype by location must be exploited for the identification of genotype that performs better in a specific location, or across location and to identify mega-environment. The significant genotypes by locations interactions for yield per hectare justified the use of the GGE biplot. Yan *et al.* [21] suggested the use GGE biplot model for analyzing the multi-location yield trial data.

Table 1. *ANOVA of 15 rice genotype tested in 5 locations in two growing years*

		NT		GPP		GWH		YLD	
SOV	DF	MS	%TSS	MS	%TSS	MS	%TSS	MS	%TSS
Location (L)	4	826.54**	19.13	36533.43**	16.03	23025.48**	29.82	74.84**	12.78
Year (Y)	1	42.63*	0.25	176901.85**	19.41	90627.56**	29.34	376.26**	16.06
L×Y	4	783.77**	18.14	41617.27**	18.26	10015.34**	12.97	19.84**	14.39
Replication (L×Y)	20	41.19**	4.77	666.56**	1.46	163.97^{ns}	1.06	1.75^{ns}	1.49
Genotype (G)	14	83.55**	6.77	9646.28**	14.81	587.41***	2.66	24.84**	14.84
G×L	56	40.67**	13.18	1591.28**	9.78	191.81*	3.48	4.10**	18.8
G×Y	14	73.64**	5.97	2081.48**	3.2	436.07***	1.98	5.04**	3.01
G×L×Y	56	50.14**	16.25	1792.44**	11.01	272.05**	4.93	4.70**	11.23
Pooled error	280	9.59	15.54	196.62	6.04	151.90	13.77	2.29	7.41

Note: **highly significant at 0.01 level, *Significant at 0.05 level, ns non-significant, S.O.V Source of variation, DF Degree of freedom, MS Mean square, %TSS Percentage total sum of squares, NT Number of tillers per hill, GPP grains per panicle, GWH grain weight per hill, YLD Yield in t/ha.

Table 2. *ANOVA by season of 15 rice genotype tested in 5 locations during the first and second year*

Source	Year	DF	NT		GPH		GWH		YLD	
			MS	TSS	MS	TSS	MS	TSS	MS	TSS
Genotype (G)	First	14	114.75**	14.12	7584.85**	40.74	654.75**	26.07	18.49**	27.30
Location (L)		4	1107.36**	38.93	5276.51**	28.10	3582.85**	40.75	79.01**	28.23
G × L		56	74.69**	36.77	2202.48**	27.32	150.05**	23.90	2.82**	35.12
Replication (L)		10	19.77**	1.74	50.18^{ns}	0.19	14.46*	0.41	0.41*	0.49
Pooled error		140	6.86	8.44	67.77	3.64	22.29	8.87	0.64	8.86
Genotype (G)	Second	14	42.44**	10.14	4142.91**	12.24	368.73^{ns}	2.82	11.39**	13.99
Location (L)		4	502.95**	34.33	72874.19**	61.49	29457.96**	64.35	15.67**	25.50
G × L		56	16.11*	15.40	1181.24**	13.95	313.80*	9.60	5.98*	39.35
Replication (L)		10	62.62**	10.69	1282.94**	2.71	313.48^{ns}	1.71	3.08^{ns}	2.71
Pooled error		140	12.33	29.45	325.48	9.61	281.51	21.52	3.95	18.46

Note: **highly significant at 0.01 level, *Significant at 0.05 level, ns non-significant, S.O.V Source of variation, DF Degree of freedom, MS Mean square, %TSS Percentage total sum of squares, NT Number of tillers per hill, GPP grains per panicle, GWH grain weight per hill, YLD Yield in t/ha.

Table 3. *Grain yield of cultivars at five test location*

Year	Genotype	SS	SC	TN	PN	KD	Mean	Ranking
First	Binadhan4	5.46 ± 0.07	5.44 ± 0.18	6.76 ± 1.39	5.11 ± 1.68	6.33 ± 0.94	5.82 ± 0.52	8
	Binadhan7	3.2 ± 0.18	4.54 ± 1.08	4.47 ± 1.13	5.72 ± 0.65	6.32 ± 0.54	4.85 ± 0.51	11
	Iratom38	3.4 ± 0.22	3.42 ± 0.22	3.51 ± 0.96	4.12 ± 1.88	4.35 ± 0.84	3.76 ± 0.51	15
	ML10	6.71 ± 0.05	5.73 ± 0.83	8.18 ± 0.72	10.86 ± 1.00	8.09 ± 2.57	7.91 ± 0.85	5
	ML21	6.55 ± 0.71	5.27 ± 0.08	8.68 ± 0.31	9.53 ± 1.32	7.78 ± 1.62	7.56 ± 0.56	6
	ML24	4.62 ± 0.51	5.73 ± 0.15	10.68 ± 0.80	10.33 ± 1.59	9.71 ± 1.53	8.21 ± 0.82	3
	ML4	6.94 ± 0.35	6.42 ± 0.28	9.28 ± 1.63	9.82 ± 0.91	10.62 ± 1.31	8.62 ± 0.53	2
	ML6	7.02 ± 0.41	5.66 ± 0.31	10.81 ± 1.13	9.77 ± 2.10	11.52 ± 0.43	8.96 ± 0.66	1
	ML9	7.32 ± 0.13	7.11 ± 0.42	9.94 ± 0.06	7.47 ± 1.61	8.89 ± 1.45	8.15 ± 0.50	4
	MR219	4.87 ± 0.32	5.77 ± 0.62	4.58 ± 0.81	3.89 ± 1.01	6.77 ± 0.37	5.18 ± 0.30	10
	MR220	4.59 ± 0.11	5.42 ± 0.30	4.6 ± 0.74	4.38 ± 1.58	4.11 ± 0.53	4.62 ± 0.43	14
	MR253	5.46 ± 0.06	4.94 ± 0.14	6.19 ± 0.41	7.71 ± 1.31	10.2 ± 0.45	6.90 ± 0.49	7
	VN121	4.52 ± 0.51	4.76 ± 0.29	5.15 ± 1.23	4.08 ± 1.07	5.42 ± 1.08	4.79 ± 0.52	12
	VN124	4.24 ± 0.19	3.24 ± 0.24	4.73 ± 1.34	4.51 ± 0.57	6.92 ± 0.59	4.73 ± 0.57	13
	VN001	3.58 ± 1.05	5.84 ± 0.37	6.19 ± 1.00	5.73 ± 1.52	5.48 ± 1.53	5.36 ± 0.65	9
	Sites mean	5.23 ± 0.20	5.29 ± 0.24	6.92 ± 0.37	6.87 ± 0.40	7.50 ± 0.36	6.36 ± 0.16	
Second	Binadhan4	5.5 ± 0.12	5.64 ± 0.91	5.4 ± 1.41	6.31 ± 0.52	6.34 ± 0.94	5.84 ± 0.41	10
	Binadhan7	3.38 ± 0.42	3.46 ± 0.11	6.78 ± 0.52	6.44 ± 1.48	5.11 ± 0.53	5.03 ± 0.60	11
	Iratom38	3.61 ± 0.84	3.27 ± 1.24	4.21 ± 0.35	2.89 ± 1.01	3.36 ± 0.84	3.47 ± 0.48	15
	ML10	6.73 ± 1.04	6.39 ± 0.43	7.56 ± 0.74	10.44 ± 1.01	8.35 ± 0.59	7.89 ± 0.37	2
	ML21	6.54 ± 1.63	6.51 ± 0.56	7.71 ± 0.09	9.14 ± 0.47	5.95 ± 1.12	7.17 ± 0.43	4
	ML24	4.66 ± 0.63	4.82 ± 0.12	5.88 ± 0.41	3.74 ± 0.49	5.88 ± 1.15	5.00 ± 0.35	12
	ML4	6.96 ± 1.35	7.22 ± 0.72	6.7 ± 0.38	8.68 ± 0.40	8.94 ± 0.69	7.70 ± 0.47	3
	ML6	6.33 ± 1.64	6.64 ± 0.19	6.63 ± 1.09	9.22 ± 0.77	6.79 ± 0.74	7.12 ± 0.63	5
	ML9	7.33 ± 1.52	7.43 ± 0.41	8.34 ± 1.02	11.33 ± 1.86	7.1 ± 1.06	8.31 ± 0.61	1
	MR219	5.28 ± 0.41	5.14 ± 0.65	4.09 ± 0.39	4.4 ± 0.43	5.52 ± 1.50	4.89 ± 0.40	14
	MR220	5.00 ± 0.53	4.16 ± 0.62	7.43 ± 0.40	7.09 ± 1.04	7.31 ± 0.81	6.20 ± 0.45	9
	MR253	4.91 ± 1.69	4.73 ± 0.36	6.74 ± 0.55	6.75 ± 0.38	10.67 ± 0.40	6.76 ± 0.57	7
	VN121	3.98 ± 1.31	4.21 ± 0.34	6.51 ± 0.67	6.76 ± 0.54	9.97 ± 1.23	6.29 ± 0.67	8
	VN124	4.5 ± 1.08	4.85 ± 0.15	7.39 ± 0.82	8.48 ± 1.24	9.08 ± 0.88	6.86 ± 0.62	6
	VN001	3.65 ± 1.21	3.58 ± 0.22	6.36 ± 0.41	5.89 ± 0.53	5.12 ± 0.86	4.92 ± 0.70	13
	Sites mean	5.22 ± 0.19	5.20 ± 0.19	6.52 ± 0.21	7.17 ± 0.32	7.03 ± 0.34	6.23 ± 0.14	

Note: Serdang (SS), Tanjung Karang (TK), Penang (PN), Kedah (KD), Sekinchan (SC).

3.3 Mega-environment evaluation

Upon pooling the yield data from all evaluation locations, the winning genotype that performs best can be identified visually by looking at genotypes “point angle” in each artificial area from a GGE biplot. To facilitate visualization of genetic correlations among environments, a two-dimensional polygon view of GGE-biplot shown in figure 1 represent the multiple-environment trial data of 2 years of 15 genotypes in 5 locations

that were based on environment-standardized data and environment-focused singular value partition. The polygon view aids in revealing the best genotype(s) in a group of the environments or a specific environment. The polygon was constructed in a way that all genotypes markers that are far from biplot origin are contained within the polygon. The rays that originated from the biplot origin perpendicularly bisect the polygon divides the polygon into different sectors [22]. The main objective of the mega-environment analysis is to understand the pattern of interaction between genotype by location (GL) within a target region. This evaluation is necessary for exploring the possibility of dividing target region into mega-environments, whereby the cause of genotype adaptation to a specific location can be exploited through GL. Hence, selection response within a mega-environment can be minimized and overall yield within a target environment can be increased. The polygon view of the GGE biplot in figure 1 explained 83.21% of the GGE variation for yield per hectare. This percentage indicated that the biplots of first two principal components adequately displayed the GGE patterns for yield per hectare. In figure 1 (Panel A), rays 1 perpendicularly bisect the polygon at the side that connected G8 and G9, ray 2 passes through the side between G9 and G12, while ray 3 bisect the side between G12 and G3 and so on. The seven rays that originated from the biplot origin divide the polygon into sectors, each sector having its own winning genotype at the vertex. If all environment markers fall into one sector, the genotype at the vertex of that sector is the best genotype across the environment. Contrarily, if environment marker falls into more than one sector, genotypes at the vertex of each sector won in all the environment of such sector. Therefore the existence of crossover pattern or genotype by environment interaction can be revealed. This interaction shows that target environments can be divided into different mega-environments. Therefore, figure 1 (Panel A) shows two mega-environment for yield per hectare. The first mega-environment consists of environment Sekinchan SC, Penang PN, and Kedah KD, while the second mega-environment comprise of environment Tanjung Karang TK and Serdang SS with genotypes G8 (ML6) and G9 (ML9) as winning cultivar respectively in each of the mega-environment.

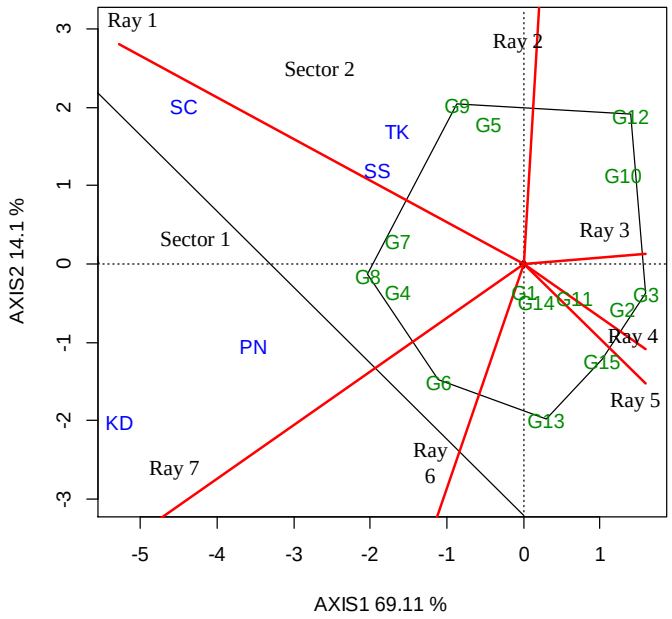

Figure 1. *The polygon (mega-environment) view of GGE biplot of 15 rice genotypes tested in 2 years and five locations for yield per hectare.* ***NOTE****: G1=Binadhan4, G2=Binadhan7, G3=Iratom38, G4=ML10, G5=ML21, G6=ML24, G7=ML4, G8=ML6, G9=ML9, G10=MR219, G11=MR220, G12=MR253, G13=VN121, G14=VN124, G15=VN001,Serdang (SS), Tanjung Karang (TK), Penang (PN), Kedah (KD), Sekinchan (SC).*

As mega-environment could be defined as a set of environments that constantly shares the best set of genotypes across the years [23], Yan and Tinker [6] suggested that the tested environments should be divided into separate mega-environments to examine if there are any repeatable crossover patterns across years [6]. Therefore, genotypes plus genotypes by locations interaction effect (GGL) biplots for individual years was constructed (Figure 2 and 3). The straight line originated from the biplot origin and perpendicular bisects the polygon divides the biplots into eight and six sectors respectively for year1 and year2. These sectors divide test locations into two groups which indicate the presence of different mega-environments. For a simple mega-environment, one or a few test locations are sufficient to identify the best genotypes that can be recommended everywhere within the mega-environment. On the contrary, for a complex mega-environment, multiple-environment trials are essential and genotype recommendation must be based on both mean and stability [11]. Also, deploying different genotypes in different mega-environments and classification of target environments into different mega-environments is the best way to deal with GEI [7]. Test locations were grouped based on the polygon (which-won-where) view of GGL biplots for individual year revealed that ML9 and ML4 are considered as winning genotypes in the two mega-environments across the year. Visualization of which-won-where patterns revealed the existence of different mega-environments of rice growing area in Malaysia. However, Yan and Kang [11] suggested that in drawing a conclusive conclusion on the existence of mega-environment, it should not be based on a merely repeatable of the environment-grouping pattern rather on repeatable which-won-where pattern. The GGL biplot by year revealed that test

locations had different winning genotypes that were not repeatable across years as shown in figure 2 and 3. Therefore, the G×E that causes the crossovers among winning genotypes cannot be converted into G [24]. This implies that the target environment consists of a complex mega-environment. Therefore, for a complex mega-environment, breeders can only select broadly adapted genotypes across the environment based on mean and stability analysis.

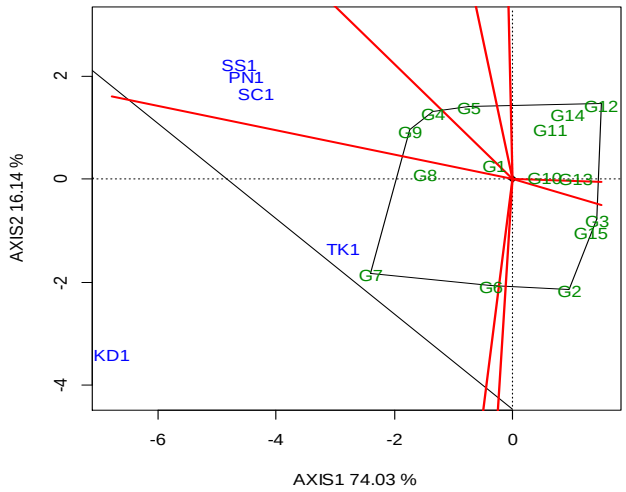

Figure 2. *The polygon (mega-environment) view of GGE biplot of 15 rice genotypes tested in years 1 at five locations for yield per hectare.*

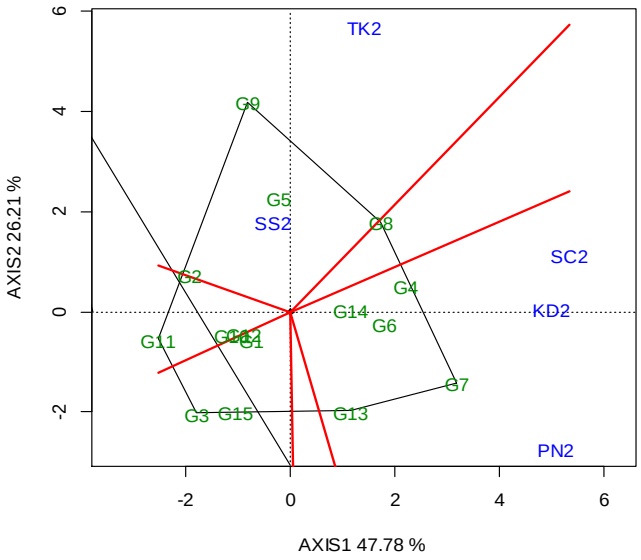

Figure 3. *The polygon (mega-environment) view of GGE biplot of 15 rice genotypes tested in years 2 at five locations for yield per hectare.*

3.4 Environment evaluation

Test location evaluation is the next step after identification of mega-environment to determine the location discriminative, representativeness ability, and redundant among the locations using GGE biplot graph. The discriminative and representativeness of the test locations was presented in figure 4. The biplot accounted for 83.2% of G and GE interaction of yield per hectare in the environments that could be interpreted by PC1 (69.11%) and PC2 (14.2.8%) on the biplot (Figure 4). The graphical display of test location in the biplot revealed that the environmental locations are group into three categories depending on environment vectors angle with the AEC abscissa. The first category of the environmental marker is those with short vector closer to the biplot origin (environment SS and TK), this indicates that all genotype perform equally in this environment and cannot be used because it provides little or no information about genotype discrimination or genotype differences. The second type of environment marker is those with long vector and large angle from the AEC abscissa (environment KD and PN), although this environment cannot be used to select superior genotype but can be used to in culling unstable genotypes. The third type of environment marker is those with long vector and a small angle closer to the AEC abscissa (environment SC). This is regarded as an ideal environment [22] and it is required in evaluating test environment because it does represent many environments in proximate therefore can be a representative of a mega-environment [24]. According to Lin and Binns [25], the environmental effect on genotypes is greatly influenced by two factors which include predictable and unpredictable elements (i.e. soil and weather). The soil is a predictable element as it is persistent from year to year and it can be regarded as fixed factors. Contrarily, the weather is a complex element which has a predictable part defined by the general climatic zone while the unpredictable part occurs as a result of variation due to time (year to year). Therefore, Lin & Binns [25] suggested that in a genotype × location × year experiment, the averaged mean of genotype × location across the year is equivalent to genotype × predictable variation while year within a location is equivalent to genotype × unpredictable variation.

3.5 Discriminative and representative ability of locations

According to Yan and Kang [11], the discriminating ability of test location is measured by the length of location vectors which is approximately to the standard deviation within each location. The longer the location vector, the higher the discriminative ability of the location. Therefore, among the five locations in this study, Kedah KD and Sekinchan SC were the most informative (discriminating) while Tanjung Karang TK, Serdang SS, and Penang PN were least discriminating (figure 4).

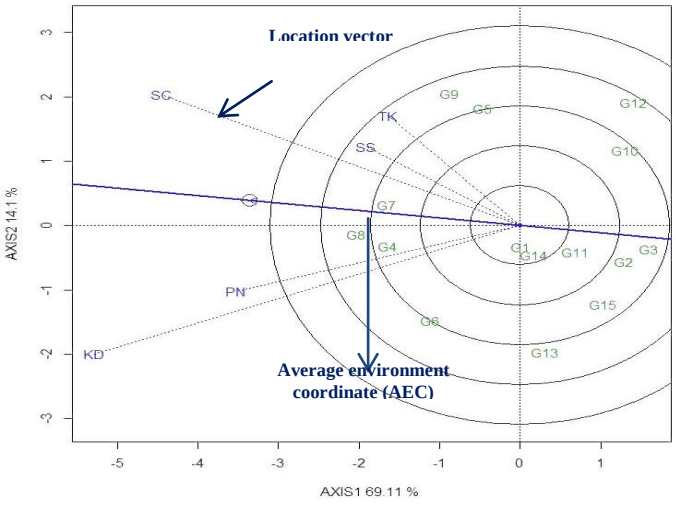

Figure 4. *The vector view of GGE biplot of 15 rice genotypes tested in 2 years and five locations for yield per hectare.*

The representativeness of a test location referred to the consistency of a targeted location when compared with other locations or the mean of all test locations. Yan and Kang [11] proposed the use of biplot for determining the representativeness of location that makes use of the average environment as the benchmark. Average environment coordinate (AEC) abscissa was constructed by the mean of PC1 and PC2 of all environments. The representativeness of individual location is determined by the proximate angle with AEC, the smaller the angle between location vector and the AEC, the more representative of the tested location. Tested locations were classified into two different mega-environment based on repeatable of which-won-where pattern across the years (GGL biplots). Therefore, representative view of GGE biplots were constructed for each of the mega-environment as shown in figure 5 and 6. Location Serdang SS and Kedah KD were the most representative location in each of the two mega-environments as indicated in figure 5 and 6 respectively. The ideal environment which shows the environment with the most discriminative and representative ability was presented in figure 7. The biplot is based on G+GE tester-centered table without scaling defines an ideal test environment which is often located in the center of the concentric circles. Sekinchan SC is closest to this point and is, therefore the best location. Location Kedah KD and Penang PN can also be considered as favorable environment whereas Serdang SS and Tanjung Karang TK were the poorest locations for selecting cultivars adapted to the whole region.

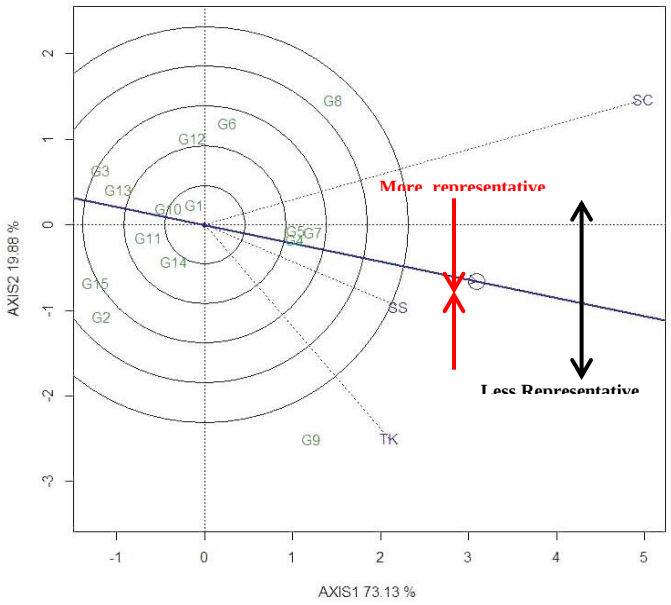

Figure 5. *The vector view of GGE biplot of 15 rice genotypes tested in 2 years and five locations for yield per hectare.*

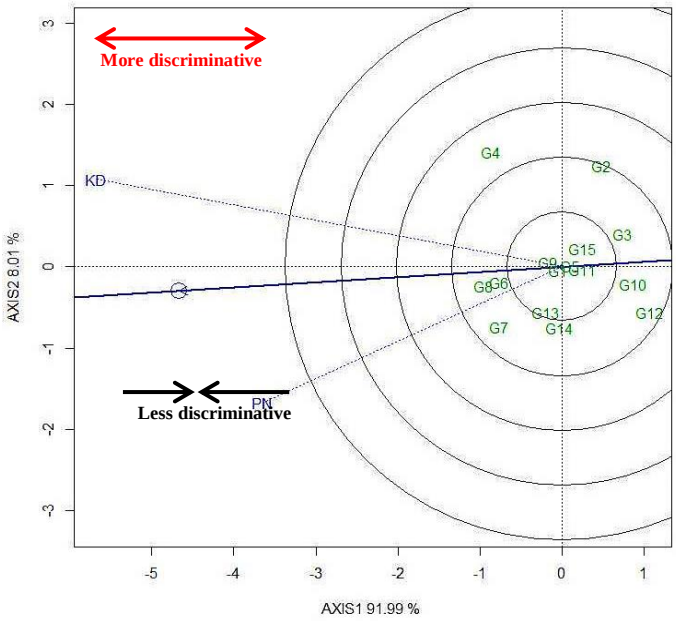

Figure 6. *The vector view of GGE biplot of 15 rice genotypes tested in 2 years and five locations for yield per hectare.*

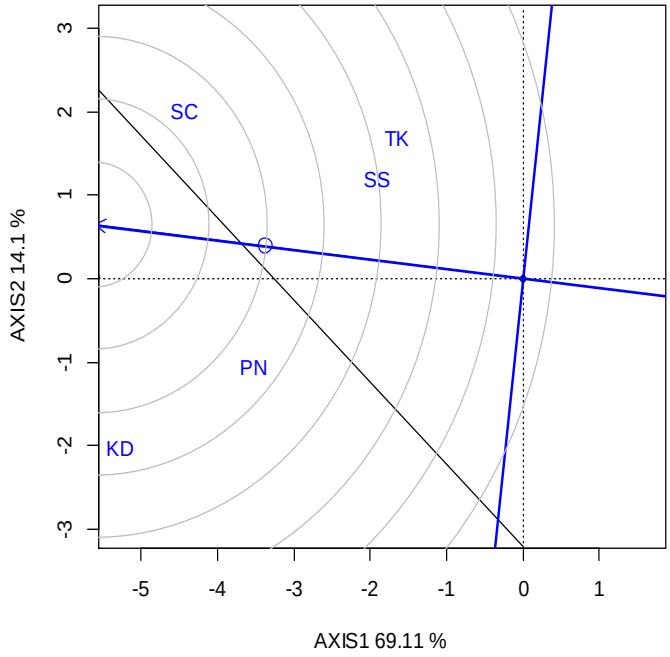

Figure 7. *The vector view of GGE biplot of 15 rice genotypes tested in 2 years and five locations for yield per hectare.*

4 CONCLUSIONS

The main objective of multi-environmental yield trials is to evaluate genotypes based on average performance from the whole region and to identify elite cultivars. Culturally, distribution and genotypes recommendation are based on average performance on yield, quality, and disease resistance. Less attention has been paid to stability and adaptability of the genotypes across different locations and years. Genotypic adaptability to a specific production area has rarely been considered, except for the determination of suitable regions for

released cultivars [26]. Theoretically, an ideal rice genotype should be high yielding, high stability and suitable for wide range of environments. Our results demonstrate that ML6, ML4, and ML9 were closer to the ideal rice genotype. Regional distribution of a cultivar specifically suitable to corresponding ecological conditions is a strategy to increase large-scale rice production, even though the general stability of that cultivar may be low across different regions. Multi-location tests of rice genotypes may also provide information about cultivar suitability to particular ecological zones. When a cultivar is recommended for production, its response to genotype-environment interaction should be considered. However, some cultivars may be missed in identification due to their average performance in a large area and some neighboring sites [27]. The GGE-Biplot method can overcome this problem by displaying both high yielding ability and stability of cultivars across the test locations. Our results showed that Serdang SS and Kedah KD had a better representativeness in terms of yield than that of other test sites. On the other hand, Kedah KD and Sekinchan SC had better discrimination ability than other test sites. Low discrimination ability at a test site could be due to environmental or human effect. In conclusion, the entire test locations were grouped into two mega-environments for cultivating rice based on grain yield using GGE biplot. Identification of these mega-environments has several implications for plant breeding in Malaysia. First, crossover pattern of genotype by location can be minimized through evaluation and selection of genotype based on genotype main effect or general adaptation. Secondly, stable and high yielding genotypes should be cultivated in complex mega-environments to achieve maximum yield. Lastly, evaluation in any of the highly correlated environment is sufficient since the environment gave a similar result. Therefore, any of these locations can be drop in further evaluation; Kedah KD or Penang PN, and Serdang SS and Tanjung Karang TK.

ACKNOWLEDGMENT

This work was supported by the [International Atomic Energy Agency] under Grant [CRP: D2.30.30] for enhancing sustainable rice production in Malaysia.

REFERENCES

[1] Oladosu, Y., Rafii, M. Y., Abdullah, N., Hussin, G., Ramli, A., Rahim, H. A., & Usman, M. (2016). Principle and application of plant mutagenesis in crop improvement: a review. *Biotechnology & Biotechnological Equipment*, *30*(1), 1-16.

[2] Oladosu Y, Rafii MY, Abdullah N, Magaji U, Miah G, Hussin G, Ramli A, 2017. Genotype × Environment interaction and stability analyses of yield and yield components of established and mutant rice genotypes tested in multiple locations in Malaysia. Acta Agric. Scand., Sect. *B*, 1-17.

[3] Najim MMM, Lee TS, Haque MA, Esham M, 2007. Sustainability of rice production: A Malaysian perspective. JAS. 3:1−12.

[4] Cooper M, Byth DE, 1996. Understanding plant adaptation to achieve systematic applied crop improvement–a fundamental challenge. M. Cooper, G.L. Hammer (Eds.), Plant Adaptation and Crop Improvement, CABI Publishing, Wallingford, pp. 5−23.

[5] Yan W, Pageau D, Frégeau-Reid J, Durand J., 2011. Assessing the representativeness and repeatability of test locations for genotype evaluation. Crop Sci. 51:1603−1610.

[6] Yan W, Tinker NA, 2006. Biplot analysis of multi-environment trial data: Principles and applications. Can *J Plant* Sci. 86:623−645.

[7] Gauch H, Zobel RW. 1997. Identifying mega-environments and targeting genotypes. Crop Sci.37: 311−326.

[8] Braun HJ, Rajaram S, Ginkel M, 1996. CIMMYT's approach to breeding for wide adaptation. Euphytica. 92:175−183.

[9] Peterson CJ, Pfeiffer WH, 1989. International winter wheat evaluation: Relationships among test sites based on cultivar performance. Crop Sci. 29:276−282.

[10] Malosetti M, Ribaut JM, Van Eeuwijk FA, 2013. The statistical analysis of multi-environment data: modeling genotype-by-environment interaction and its genetic basis. Front. Physiol. 4:44.

[11] Yan W, Kang MS, 2002. GGE biplot analysis: A graphical tool for breeders, geneticists, and agronomists. CRC press.

[12] Oladosu Y, Rafii MY, Abdullah N, Malek MA, Rahim HA, Hussin G, Kareem I, 2015. Genetic variability and diversity of mutant rice revealed by quantitative traits and molecular markers. Agrociencia. 49:249−266.

[13] Dia M, Wehner TC, Arellano C, 2015. Analysis of genotype × environment interaction (GxE) using SAS programming. http://cuke.hort.ncsu. edu/cucurbit/wehner/software.html (accessed 27 May. 2017).

[14] R Core Team 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/

[15] RStudio. 2014. RStudio: Integrated development environment for R (Computer software v0.98.1074). http://www.rstudio.com/ (accessed 30 Dec. 2016).

[16] Badu-Apraku B, Akinwale RO, Obeng-Antwi K, Haruna A, Kanton R, Usman I, Oyekunle M, 2013. Assessing the representativeness and repeatability of testing sites for drought-tolerant maize in West Africa. Can J Plant Sci. 93:699−714.

[17] Dia M, Wehner TC, Hassell R, Price DS, Boyhan GE, Olson S, Tolla GE, 2016. Genotype × environment interaction and stability analysis for watermelon fruit yield in the United States. Crop Sci. 56:1645−1661.

[18] Oladosu Y, Rafii MY, Abdullah N, Abdul Malek M, Rahim HA, Hussin G, Kareem I, 2014. Genetic variability and selection criteria in rice mutant lines as revealed by quantitative traits. Scientific World J. http://dx.doi.org/10.1155/2014/190531

[19] Fan XM, Kang MS, Chen H, Zhang Y, Tan J, Xu C, 2007. Yield stability of maize hybrids evaluated in multi-environment trials in Yunnan, China. Agron J. 99:220−228.

[20] Annicchiarico P, 2002. Genotype × environment interactions: Challenges and opportunities for plant breeding and cultivar recommendations. FAO Plant Production and Protection Paper No. 174. FAO, Rome

[21] Yan W, Hunt LA, Sheng Q, Szlavnics Z, 2000. Cultivar evaluation and mega-environment investigation based on the GGE biplot. Crop Sci. 40:597−605.

[22] Yan W, 2002. Singular-value partitioning in biplot analysis of multi environment trial data. Agron J. 94:990−996.

[23] Dia M, Wehner TC, Hassell R, Price DS, Boyhan GE, Olson S, Juarez B, 2016. Value of Locations for Representing Mega-Environments and for Discriminating Yield of Watermelon in the US. Crop Sci. 56:1726−1735.

[24] Yan W, Kang MS, Ma B, Woods S, Cornelius PL, 2007. GGE biplot vs. AMMI analysis of genotype-by-environment data. Crop Sci. 47:643−653.

[25] Lin CS, Binns MR, 1994. Concepts and methods for analyzing regional trial data for cultivar and location selection. Plant Breed Rev. 12:271−297.

[26] Luo J, Pan YB, Xu L, Zhang H, Yuan Z, Deng Z, Que Y, 2014. Cultivar evaluation and essential test locations identification for sugarcane breeding in China. Scientific World J. http://dx.doi.org/10.1155/2014/302753

[27] Luo J, Zhang H, Deng ZH, Xu LP, Xu LN, Yuan ZN, Que YX, 2013. Analysis of yield and quality traits in sugarcane cultivars (lines) with GGE-Biplot. Acta Agron Sin. 39:142−152.