Electronic Journal of Plant Breeding

Research Article

Multi trait genotype- ideotype distance index (MGIDI) for early seedling vigour and yield related traits to identify elite lines in rice (*Oryza sativa* L.)

M. Pallavi[1*], B. P. Maruthi Prasad[1], P. Shanthi[2], V.L.N. Reddy[3] and A.R. Nirmal Kumar[4]

[1]Department of Genetics and Plant Breeding, University of Agricultural Sciences, GKVK, Bengaluru, Karnataka, India.
[2]Department of Genetics and Plant Breeding, S.V. Agricultural College, Tirupati, A.P., India.
[3]Department of Molecular Biology and Biotechnology, S.V. Agricultural College, Tirupati, A.P., India.
[4]Department of Crop Physiology, S.V. Agricultural College, Tirupati, A.P., India.
*E-Mail: pallavimmuniraja@gmail.com

Abstract
Rice (*Oryza sativa* L.) stands as a important cereal sustaining over half of the world's population. This study delves into the challenges confronting breeders in the realm of crop improvement, specifically focusing on the intricate task of designing an ideotype—a genotype amalgamating diverse attributes for optimal performance. Traditional methodologies, exemplified by the Smith–Hazel (SH) index, grapple with issues such as multicollinearity and the complexities of economic weighting decisions. In response to these challenges, the Multi-Trait Genotype–Ideotype Distance Index (MGIDI), conceptualized by Olivoto and Nardino (2021), emerges as a ground breaking approach. Principal Component Analysis (PCA) aids in the reduction of trait dimensionality, revealing four key factors that collectively contribute to 79.444% of total variability. The Scree plot guides factor selection, ensuring a targeted analysis. The MGIDI index computation yields a total genetic gain of 273.025%, with specific traits like spikelet fertility and seedling dry weight exhibiting significant gains. Six high-performing rice accessions—SM227, NLR33892, MTU3626, 239(3), SMB3, and 405C3 were identified through MGIDI. These identified genotypes serve as valuable resources for developing recombinant populations, aligning with sustainable and effective crop improvement strategies. Additionally, these promising varieties exhibit strengths across various traits, offering potential for simultaneous trait improvement in future breeding programmes. The efficiency of MGIDI is highlighted through its innovative application in simultaneous trait selection, underscoring its significance across a wide range of crops.

Keywords: MGIDI, multi-trait selection, Scree plot, Ideotype, Rice

INTRODUCTION

Rice (*Oryza sativa* L.) is the second most important cereal crop after wheat and serves as a primary food source for over 50% of the global population. About 90 % of rice production and consumption is concentrated in Asia. and remaining 10% of rice cultivation is distributed worldwide. With the world's population anticipated to reach 8 billion by 2025, there is a growing demand for increased rice production and cultivation area. Based on the current consumption pattern, about 35 to 40% increase in rice production is projected by 2025 to meet the needs of the expanding population.

Genetic gain is pivotal in plant breeding, shaping the trajectory of breeding programs. Relying solely on a few traits for selection is deemed inadequate, as it neglects potential gains in other crucial traits. Thus, breeders strive to amalgamate diverse desirable traits into a single genotype to enhance overall performance. So, in crop

improvement initiatives, breeders often aim to develop an ideotype, a genotype combining various attributes for optimal performance (Donald, 1968). The ideotype design's objective is to enhance crop performance by simultaneously considering multiple attributes in genotype selection (Olivoto, and Nardino, 2021). Various selection indices have emerged for this purpose, yet the challenge lies in assigning realistic economic weightings to express the economic value of traits effectively. This hurdle often constrains breeders in identifying the optimal genotypes. The Smith–Hazel (SH) index, a linear selection index frequently utilized in multi-trait selection by breeders, relies on the phenotypic and genotypic covariance matrices. Additionally, a vector of economic weights is employed to determine the selection of index coefficients, maximizing the correlation between unknown genetic values and phenotypic values (Smith, 1936; Hazel, 1943).

However, the computation of the SH index involves inverting a phenotypic covariance matrix, as noted by Smith (1936). When traits are highly correlated leads to presence of multicollinearity, which commonly occurs when assessing multiple traits, can result in poorly conditioned matrices and biased index coefficients. This, in turn, affects the accuracy of genetic gain estimates (Graham, 2003 and Olivoto *et al*., 2017). The challenge posed by multicollinearity is rooted in the suboptimal conditioning of the phenotypic covariance matrix. Furthermore, if the matrix is not positive definite, the SH index may encounter an infinite number of solutions (Graham, 2003 and Olivoto *et al*., 2017). In addition to the multicollinearity challenge, breeders also grapple with determining realistic economic weightings for traits, a critical aspect of expressing their economic value (Bizari *et al*., 2017). This decision-making process further adds to the complexities faced by breeders in optimizing the SH index for effective crop improvement strategies.

In response to the challenges associated with the Smith–Hazel (SH) index, a pioneering approach has been introduced in the form of the multi-trait genotype–ideotype distance index (MGIDI), as developed by Olivoto and Nardino (2021). This innovative multivariate selection index has been specifically designed to address the shortcomings of traditional indices, particularly the issues stemming from multicollinearity. Unlike conventional methods, the MGIDI considers the inherent correlation among traits and effectively selects all variables in the evaluation process. By doing so, it aims to overcome the limitations posed by poorly conditioned matrices and biased index coefficients encountered by the SH index while selecting superior genotypes. The emphasis on the genotype–ideotype distance in the MGIDI aligns with the ideotype breeding concept, facilitating a more comprehensive and effective approach to crop improvement. Olivoto and Nardino (2021) assert that the application of MGIDI results in significant genetic gain, marking a promising advancement in the field of multivariate selection indices by evaluating the strengths

and weakness of the tested genotype. This introduction sets the stage for a deeper exploration of the MGIDI and its potential to revolutionize breeding strategies in the face of complex trait interactions and economic considerations. The main objective of this work is to identify rice accessions with higher value for yield traits and early seedling vigour for future genetic improvement initiatives using multi-trait genotype ideotype distance index (MGIDI).

MATERIALS AND METHODS
Evaluation of Yield and Its Component traits under Field Conditions was conducted at wetland farm of S. V. Agricultural College, Tirupati, Andhra Pradesh (13° 37' 28.092'' N, 79° 22' 45.912'' E), situated at an altitude of 182.9 meters above sea level within the Southern Agro-climatic Zone. The experiment was carried out during *Kharif* 2020-21 in a Randomized Complete Block (RCB) design with three replications. Forty-two rice genotypes were evaluated under field conditions **(Table 1)**. Observations were recorded on yield and yield attributing traits namely, Days to 50% flowering, plant height, panicle length, number of panicles per plant, 1000-grain weight, number of filled grains per panicle, number of grains per panicle, spikelet fertility, and grain yield per plant.

Early seedling vigour related traits were evaluated under laboratory conditions in the Dept. of Genetics and Plant Breeding, Sri Venkateswara Agricultural College, Tirupati, Andhra Pradesh. Each genotype was assessed for ESV related traits by following Completely Randomized Design (CRD) using standard germination test (Paper towel method) as per International Seed Testing Association (ISTA, 2015). Seeds showing 2 mm of radicle length were considered as germinated. Germination per cent was determined as per ISTA rules for seed testing. Observations were recorded on representative sample taken from randomly selected five plants in each replication for each entry on components related to early seedling vigor. encompassing first count of germination, final count of germination, germination index, germination rate, seedling fresh and dry weight, shoot length, root length, seedling height, root to shoot ratio, and two seedling vigor indices.

Statistical analysis for the Multi-Trait Genotype–Ideotype Distance Index (MGIDI) utilized the R Package 'metan' version 1.18.0, developed by Olivoto and Lúcio in 2020 (https://github.com/TiagoOlivoto/metan). The analysis was carried out in R version 4.3.1 where construction of the MGIDI index involves four main steps which are as follows

Rescaling Traits (Normalization)

- Adjust each trait's scale so that they all range from 0 to 100.
- This ensures a uniform comparison and interpretation of trait values.

Table 1. The designation and parentage of 42 rice genotypes utilized in the current investigation

S. No.	Designation	Parentage
1	54-3	BPT5204 x NLR33892
2	150-3	BPT5204 x NLR33892
3	187-1	BPT5204 x NLR33892
4	187-3	BPT5204 x NLR33892
5	239-3	BPT5204 x NLR33892
6	419-1	BPT5204 x NLR33892
7	419-3	BPT5204 x NLR33892
8	538-3	BPT5204 x NLR33892
9	BPT5204	(GEB24 x TN1) x MAHSURI
10	NLR33892	NLR 27999 x MTU 4870
11	NDP3	NLR34449 x IR64 Drt1
12	ND13	NLR34449 x IR64 Drt1
13	ND42	NLR34449 x IR64 Drt1
14	ND44	NLR34449 x IR64 Drt1
15	ND60	NLR34449 x IR64 Drt1
16	NLR34449	IR 72 x BPT5204
17	MDP2	MTU1010 x IR64 Drt1
18	MDP3	MTU1010 x IR64 Drt1
19	MDP4	MTU1010 x IR64 Drt1
20	MD4	MTU1010 x IR64 Drt1
21	MD5	MTU1010 x IR64 Drt1
22	MDP6	MTU1010 x IR64 Drt1
23	MTU1010	KRISHNAVENI x IR 64
24	28C1	N22 x IR64
25	46C2	N22 x IR64
26	76C2	N22 x IR64
27	81C	N22 x IR64
28	130C	N22 x IR64
29	153C1	N22 x IR64
30	221C3	N22 x IR64
31	273C2	N22 x IR64
32	405C3	N22 x IR64
33	425C3	N22 x IR64
34	457C2	N22 x IR64
35	476C3	N22 x IR64
36	488C	N22 x IR64
37	NAGINA 22	LANDRACE
38	IR64	IR5657-33-2-1 x IR2061-465-1-5-5
39	SM227	SWARNA MUTANT
40	MTU3626	IR B x MTU 3
41	MTU1121	BPT5204 x MTU BB 8-24-1
42	SMB3	SWARNA MUTANT

Let X_{ij} be a two-way table with **i** rows/ genotypes/ treatments and j columns/traits. The rescaled value for the **ith** row and jth column ($_rX_{ij}$) is given by:

$$rX_{ij} = \frac{\eta_{nj} - \varphi_{nj}}{\eta_{oj} - \varphi_{oj}} \times (\theta_{ij} - \eta_{oj}) + \eta_{nj}$$

Where,

η_{nj} and φ_{nj} are the new maximum and minimum values for the trait *j* after rescaling, respectively; η_{oj} and φ_{oj} are the original maximum and minimum values for the trait *j*. θ_{ij} is the original value for the *jth* trait of the *ith* genotype.

For traits in which negative gains are desired: η_{nj} =0 and φ_{nj} =100 is considered

For traits in which positive gains are desired: η_{nj} =100 and φ_{nj} =0 is considered

Factor Analysis: In the computation of the Multi-Trait Genotype–Ideotype Distance Index (MGIDI), a factor analysis approach was employed, akin to the FAI-BLUP index methodology proposed by Rocha *et al*. (2018). This analytical technique addressed the inherent correlation structure among traits, providing a means to reduce the dimensionality of the original trait set and elucidate relationships within latent variables. The application of factor analysis facilitated a simplified interpretation for practitioners, as the orthogonal axes among final factors mitigated issues of multicollinearity in genotype or treatment scores. This will helps to identify underlying factors that contribute to observed correlations, simplifying the complexity of the data.

$$X = \mu + Lf + \varepsilon$$

Where, X is a p × 1 vector of rescaled observations**;** μ is a p × 1 vector of standardized means; L is a p × f matrix of factorial loadings**;** f is a p × 1 vector of common factors**;** ε is a p × 1 vector of residuals, being p and f, the number of traits and common factors retained, respectively. The eigenvalues and eigenvectors are obtained from the correlation matrix of $_rX_{ij}$. The initial loadings are obtained considering the screeplot values.

Then final loadings are obtained after varimax rotation (Kaiser, 1958). The scores obtained as follows:

$$F = Z(A^TR^{-1})^T$$

Where, F is a g × f matrix with the factorial scores; Z is a g × p matrix with the standardized means; A is a p × f matrix of canonical loadings, and R is p × p correlation matrix between traits. g, f and p are genotypes, factors retained and traits analysed respectively.

Ideotype planning
• Define an ideotype by specifying target values for traits based on known or desired characteristics.

• This step involves determining the ideal combination of trait values for a genotype.

Computing Genotype Distance to Ideotype
• Calculate the distance between each genotype and the planned ideotype.
• This distance measurement quantifies how closely a genotype aligns with the desired trait values set in the ideotype.

$$MGIDI_i = [\sum_{j=1}^{f} (\gamma_{ij} - \gamma_j)^2]^{0.5}$$

Where, MGIDI$_i$ is the distance index for the ith genotype; γ_{ij} Score of the ith genotype in the jth factor (where i = 1, 2, 3, … g; j = 1, 2, …,f). γ_j is the jth score of the ideotype.

Proportion of the MGIDI index of the ith genotype explained by the jth factor (ω_{ij}) is used to shoe the strength and weakness of genotypes.

$$\omega_{ij} = \frac{\sqrt{D_{ij}^2}}{\sum_{j=1}^{f} \sqrt{D_{ij}^2}}$$

Where, D$_{ij}$ is the distance between the **ith** genotype and the ideotype for the **jth** factor. Low contributions of a factor indicate that the traits within such a factor are close to the ideotype.

RESULTS AND DISCUSSION
The likelihood ratio tests conducted on the analyzed traits, using the BLUP method with random effects of genotype and fixed effects of replication, revealed substantial genotype effects across multiple variables. Notably, plant height (PH), panicle length (PL), number of panicles per plant (NPP), thousand-grain weight (TGW), number of filled grains per panicle (TFG), Number grains per panicle (TGP), spikelet fertility (SF), first count of germination (FIRC), final count of germination (FINC), root to shoot ratio (RS), and both Seedling Vigour Index I and II (SVI I and SVI II) exhibited significant genotype effects (**Fig. 1**). These findings underscore the importance of genetic factors in influencing the observed variation in these traits. In the context of the Multi-Trait Genotype–Ideotype Distance Index (MGIDI), these significant genotype effects are crucial as they contribute to the genetic diversity that the MGIDI aims to capture.

Multicollinearity, characterized by high correlations among independent variables, presents a substantial challenge in Multivariate Analysis of Variance (MANOVA) models (Hair *et al*., 1998). In our study, a Condition Number Analysis using the kappa statistic on the MANOVA model revealed a condition number of 71.8466, indicating potential multicollinearity issues (Kutner *et al*., 2005). This elevated condition number underscores the necessity
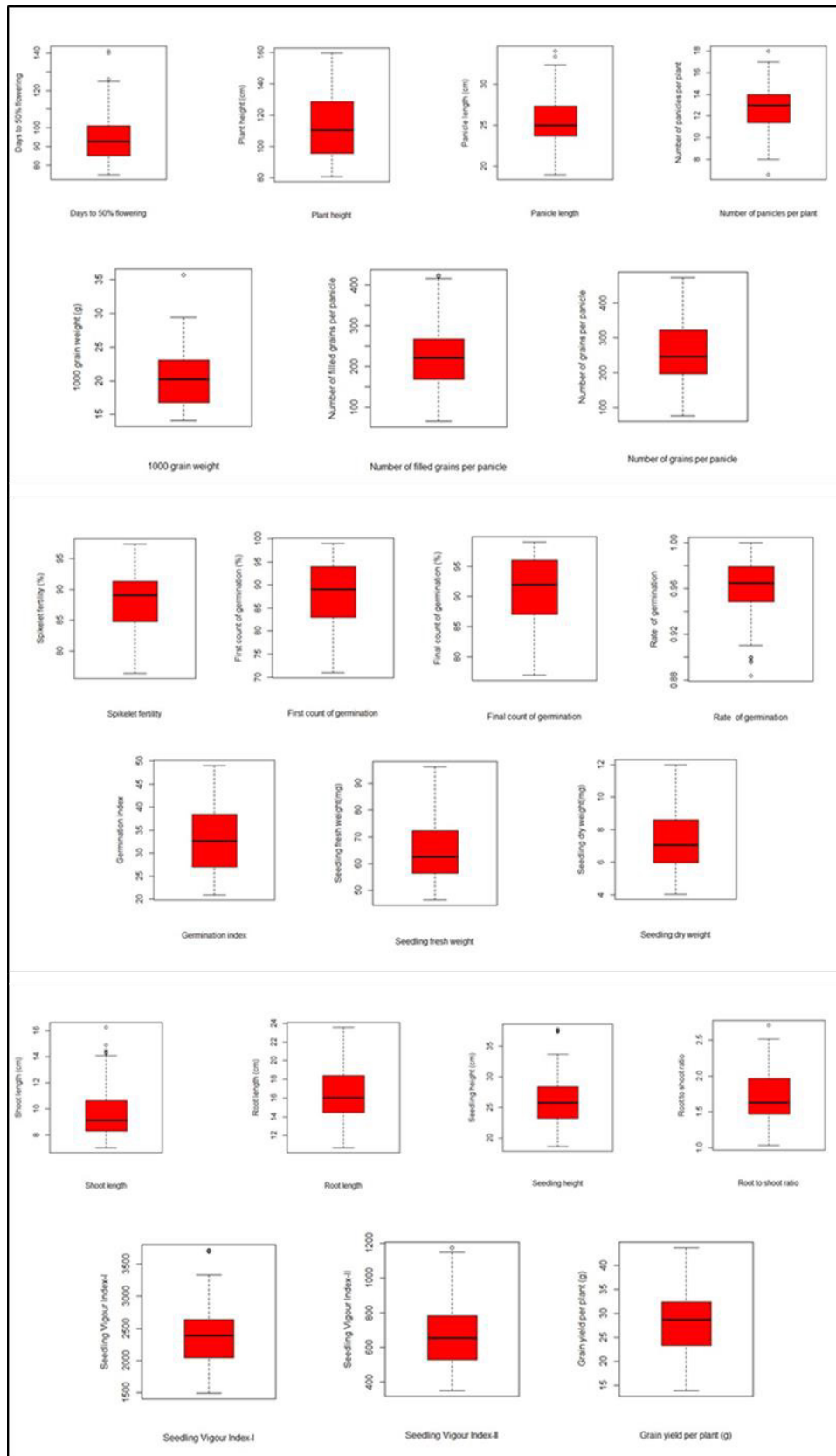
Fig. 1. Boxplot analysis of variability for 21 traits of 42 rice genotypes

for addressing multicollinearity to ensure the stability of parameter estimates in the MANOVA model. To mitigate this concern, we explored strategic approaches such as examining variable correlations through the correlation matrix, applying Factor Analysis to reduce dimensionality (Costello and Osborne, 2005), and considering MGIDI as an alternative method (Olivoto and Lúcio, 2020).

The correlation matrix reveals significant relationships among the variables in the study. Number of filled grains per panicle and number of grains per panicle exhibited nearly perfect correlation (0.990), while traits like Seedling Vigour Index-II with seedling dry weight (0.969), seedling height with Root length (0.911), seedling Vigour Index-I with seedling height (0.950) and many other traits exhibited a strong positive correlation (**Fig. 2)**. These correlations could potentially lead to multicollinearity issues.

In addressing the challenge of multicollinearity within our dataset, a systematic approach through correlation analysis was employed to scrutinize the interrelationships among independent variables (Hair *et al*., 1998). Calculating the correlation matrix allowed for the identification of pairs of variables exhibiting high correlation coefficients, indicative of potential multicollinearity (Kutner *et al*., 2005). Subsequently, strategic measures were taken, such as variable selection based on theoretical significance, dimensionality reduction using techniques like Principal Component Analysis (Costello and Osborne, 2005), and variable transformations to alleviate the observed correlations.

Principal Component Analysis and Factors selection based on Scree plot: The study incorporated 42 rice genotypes for Principal Component Analysis (PCA) to analyze 21 traits related to early seedling vigour, yield, and yield attributing characteristics. The resulting Principal Components (PCs) revealed that only four PCs exhibited eigenvalues surpassing 1.000, collectively

explaining approximately 79.444% of the total variability (**Table 2**). PC1, with an eigenvalue of 8.350, accounted for 39.761% of the variability, followed by PC2 (eigenvalue 3.936, 18.743% variability), PC3 (eigenvalue 2.876, 13.696% variability), and PC4 (eigenvalue 1.521, 7.245% variability). Subsequent PCs exhibited a gradual decline in variability. The first principal component, representing a linear combination of the original predictor variables, captured the maximum variance in the dataset. Its selection is crucial, as it determines the direction of the highest variability. In this context, the substantial variability captured by PC1, with a 39.761% contribution, underscores its importance for guiding the selection of lines, making them particularly desirable for further breeding. Similar results were observed by Prasad *et al.* (2023).

Cattell (1996) introduced a visual diagnostic tool known as the Scree plot for determining the optimal number of factors in factor analysis. The plot depicts eigenvalue magnitude on the vertical axis against the corresponding factor numbers on the horizontal axis. Dots representing eigenvalues are connected by a line, and the extraction of factors is halted at the point where an 'elbow' or levelling of the plot occurs. This graphical test aids in identifying the ideal number of factors to extract before the influence of unique variance surpasses the common variance structure (Hair *et al*., 1998; Cattel, 1973). In **Fig. 3**, the Scree test graphically presents eigenvalues against the ten components in the order of extraction. Larger factors with higher eigenvalues are initially extracted, followed by smaller ones. The Scree plot serves as a decisive tool for discerning the number of factors to retain (Shrestha, 2021). Notably, the Scree plot in this study (**Fig. 3**) indicates the presence of four factors with eigenvalues greater than one, collectively explaining a substantial portion of the total variance. The remaining factors contribute minimally to the overall variability and are considered less critical. Consequently, only these four influential factors will
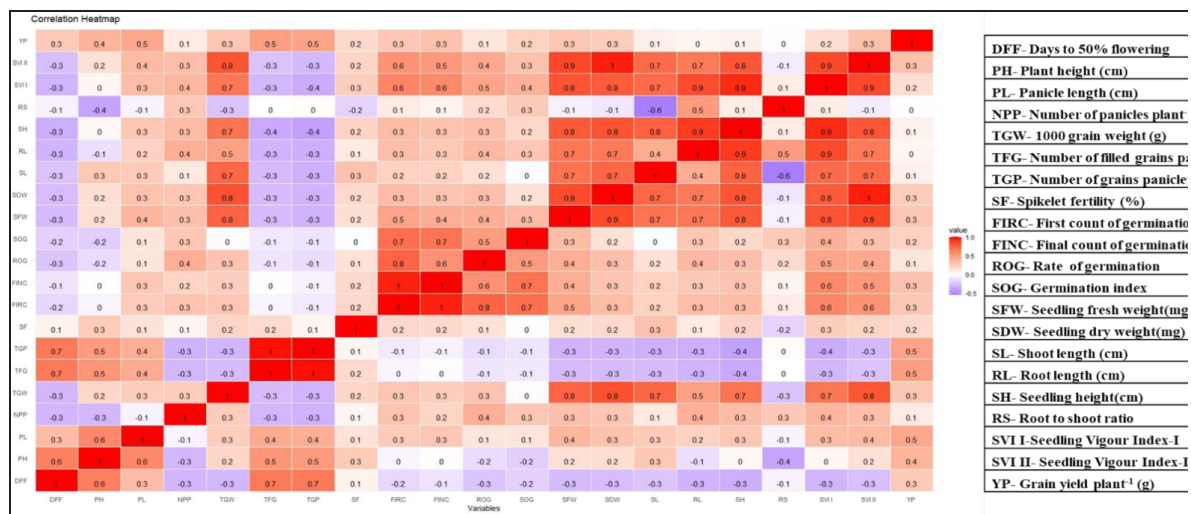

Fig. 2. Heatmap of correlation coefficients of 21 traits of rice genotypes

Table 2. Eigenvalues, % variance and cumulative eigenvalues of Rice genotypes.

| PC	Eigenvalues	Variance (%)	Cum. variance (%)
PC1	8.350	39.761	39.761
PC2	3.936	18.743	58.504
PC3	2.876	13.696	72.199
PC4	1.521	7.245	79.444
PC5	0.935	4.451	83.895
PC6	0.861	4.100	87.995
PC7	0.555	2.645	90.640
PC8	0.484	2.305	92.945
PC9	0.390	1.856	94.801
PC10	0.312	1.486	96.287
PC11	0.275	1.312	97.598
PC12	0.166	0.790	98.388
PC13	0.139	0.663	99.051
PC14	0.109	0.519	99.570
PC15	0.082	0.391	99.961
PC16	0.006	0.028	99.989
PC17	0.002	0.008	99.997
PC18	0.000	0.002	99.999
PC19	0.000	0.001	100.000
PC20	0.000	0.000	100.000
PC21	0.000	0.000	100.000
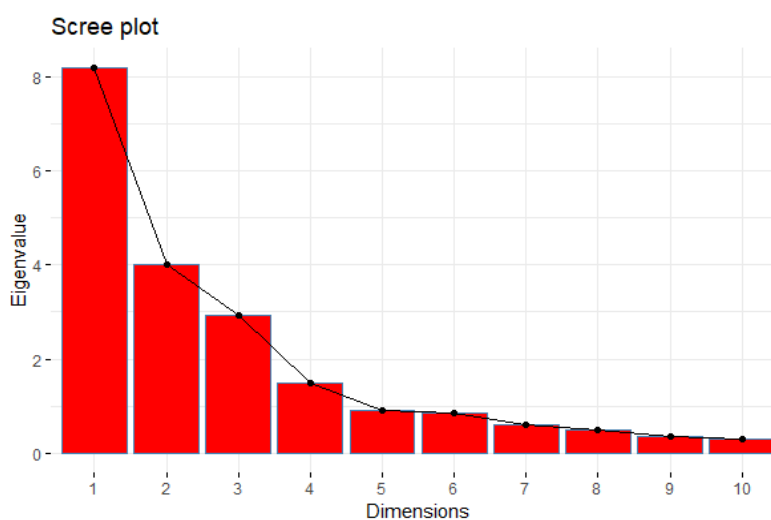
Fig. 3. Scree plot

be further considered in the calculation of the MGIDI index. This selective focus ensures a more targeted and meaningful analysis, aligning with the identified dimensions of maximum variance in the dataset.

Considering four factors from the initial set of 21 traits, which collectively contribute to 79.444 percent of the total variation among traits, analysis revealed distinct trait groupings. These factors provide a more streamlined perspective on the interrelatedness of traits. Factor 1 encompasses traits such as 1000 grain weight (g), seedling fresh weight (mg), seedling dry weight (mg), shoot length (cm), root length (cm), seedling height (cm), seedling vigour index-I, and seedling vigour index-II. Factor 2 is associated with traits like days to 50% flowering, plant height (cm), panicle length (cm), number of filled grains per panicle, number of grains per panicle, and grain yield per plant (g). Factor 3 is linked to traits

including first count of germination (%), final count of germination (%), rate of germination, and germination index. Lastly, Factor 4 is associated with traits like number of panicles per plant, spikelet fertility (%), and root to shoot ratio (**Table 3**). The average communality and uniqueness accounted for 79.4% and 20.6%, respectively, of all the genetic variability in the dataset, as outlined in **Table 3**. This reaffirms the efficacy of factor analysis in establishing an index that optimally selects traits (Olivoto and Nardino, 2021). Evaluation of the MGIDI index demonstrated desired genetic gains across all 21 traits, with a total genetic gain of 273.025 percent. Noteworthy traits such as spikelet fertility, seedling dry weight, panicle length, number of grains per panicle, and 1000 grain weight exhibited substantial percent selection gains of 25.062, 23.472, 22.609, 19.716, and 19.543, respectively (**Table 3**). This emphasizes the effectiveness of the MGIDI index in facilitating targeted and favourable trait selection for enhanced crop improvement strategies.

Selection of Genotypes Using MGIDI: In this study, 21 traits were employed to assess variations among the 42 rice genotypes. Among the 42 rice genotypes evaluated, the Multi-Trait Genotype–Ideotype Distance Index (MGIDI) pinpointed six accessions as high-performing for multiple traits, offering significant potential for the simultaneous improvement of twenty-one measured traits in rice breeding programs (**Fig. 4. and Table 4**). The genotypes identified by the MGIDI index—SM227, NLR33892, MTU3626, 239(3), SMB3, and 405C3—emerge as promising candidates with exceptional characteristics (**Fig. 4**). Notably, 405C3, situated near the cut point indicated by the red line, suggests intriguing features that warrant further investigation. Researchers are encouraged to scrutinize genotypes in close proximity to the cutpoint, as emphasized by Olivoto and Nardino (2021). The MGIDI index efficiently selected accessions SM227, NLR33892, MTU3626, 239(3), SMB3, and 405C3 as promising candidates for direct-seeded rice improvement programs. The versatility of the MGIDI

Table 3. Factorial loadings, communalities, uniqueness, and predicted genetic gains (PSG) based on the multi-trait genotype–ideotype distance index (Bold values represent traits with high contribution to each component)

VAR	FA1	FA2	FA3	FA4	Communality	Uniquenesses	PSG (%)	sense	goal
Days to 50% flowering	0.272	**-0.758**	0.192	-0.053	0.689	0.311	15.759	increase	100.000
Plant height (cm)	-0.224	**-0.763**	0.179	-0.413	0.835	0.165	16.181	increase	100.000
Panicle length (cm)	-0.382	**-0.708**	-0.139	-0.014	0.667	0.333	22.609	increase	100.000
Number of panicles plant[-1]	-0.310	0.246	-0.333	**0.411**	0.436	0.564	11.782	increase	100.000
1000 grain weight (g)	**-0.852**	0.037	-0.101	-0.236	0.793	0.207	19.543	increase	100.000
Number of filled grains panicle[-1]	0.315	**-0.889**	0.010	-0.011	0.891	0.109	16.981	increase	100.000
Number of grains panicle[-1]	0.355	**-0.874**	0.049	0.024	0.893	0.107	19.716	increase	100.000
Spikelet fertility (%)	-0.246	-0.229	-0.182	**-0.307**	0.240	0.760	25.602	increase	100.000
First count of germination (%)	-0.256	-0.055	**-0.944**	-0.024	0.960	0.040	10.618	increase	100.000
Final count of germination (%)	-0.226	-0.102	**-0.912**	-0.080	0.899	0.101	17.725	increase	100.000
Rate of germination	-0.266	0.079	**-0.790**	0.119	0.716	0.284	14.191	increase	100.000
Germination index	-0.055	0.033	**-0.825**	0.215	0.731	0.269	17.626	increase	100.000
Seedling fresh weight(mg)	**-0.880**	-0.006	-0.293	-0.014	0.860	0.140	13.997	increase	100.000
Seedling dry weight(mg)	**-0.940**	0.011	-0.141	-0.053	0.906	0.094	23.472	increase	100.000
Shoot length (cm)	**-0.851**	0.070	-0.018	-0.445	0.928	0.072	4.039	increase	100.000
Root length (cm)	**-0.776**	0.092	-0.157	0.554	0.941	0.059	2.722	increase	100.000
Seedling height(cm)	**-0.938**	0.097	-0.118	0.182	0.936	0.064	1.070	increase	100.000
Root to shoot ratio	0.094	0.007	-0.146	**0.951**	0.934	0.066	5.246	increase	100.000
Seedling Vigour Index-I	**-0.870**	0.043	-0.394	0.126	0.929	0.071	2.591	increase	100.000
Seedling Vigour Index-II	**-0.898**	-0.021	-0.354	-0.068	0.937	0.063	3.330	increase	100.000
Grain yield plant[-1] (g)	-0.199	**-0.659**	-0.296	-0.018	0.562	0.438	8.225	increase	100.000
Average					0.794	0.206			

Table 4. The multi-trait genotype-ideotype distance index for Rice genotypes

| S.No.	Genotype	MGIDI	S.No.	Genotype	MGIDI
1	SM227	3.144	22	28C1	5.372
2	NLR33892	3.287	23	187(1)	5.389
3	MTU3626	3.598	24	488C	5.450
4	239(3)	4.018	25	419(1)	5.502
5	SMB3	4.092	26	ND60	5.507
6	405C3	4.229	27	419(3)	5.526
7	MTU1121	4.284	28	ND44	5.607
8	221C3	4.390	29	MDP6	5.667
9	457C2	4.452	30	273C2	5.681
10	NDP3	4.486	31	ND42	5.741
11	76C2	4.663	32	MD4	5.872
12	ND13	4.742	33	476C3	5.895
13	IR64	4.770	34	N22 C3	5.974
14	425C3	4.846	35	187(3)	6.179
15	150(3)	5.006	36	54(3)	6.200
16	130C	5.036	37	MDP2	6.276
17	81C	5.084	38	NLR34449	6.426
18	538(3)	5.152	39	MDP4	6.467
19	153C1	5.205	40	MDP3	6.492
20	MD5	5.207	41	MTU1010	6.841
21	46C2	5.227	42	BPT5204	7.210
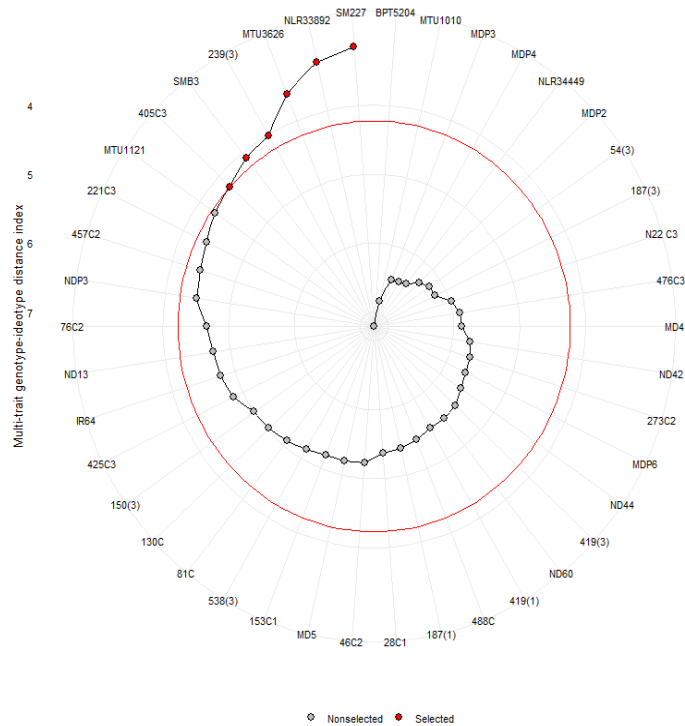
Fig. 4. Rice accession rankings showing selected accessions using the multi-trait genotype– ideotype index (MGIDI). The selected accessions are shown as red dots, while the unselected accessions are shown as black dots. The red circle represents the cut point according to the selection pressure.

model is underscored by its successful application in assessing ideal yield and yield-related traits in various crops such as Bush Yam (Adewumi *et al.,* 2023), Maize (Palaniyappan *et al.,* 2023), Wheat (Meier *et al.,* 2021), Eggplant (Uddin *et al.,* 2021), Guar (*Cymopsis tetragonoloba*) (Benakanahalli *et al.,* 2021), soybean (Woyann *et al.,* 2019, Volpato *et al.,* 2019). These diverse studies collectively highlight the efficiency of multivariate selection indices for simultaneous trait selection. Moreover, Olivoto and Nardino (2021) have indicated that MGIDI stands out as the most efficient index for selecting genotypes with desired characteristics, further reinforcing its applicability and effectiveness in crop improvement strategies.

Jalalifar *et al*. (2023) shed light on the promising prospects of selected rice genotypes through MGIDI in their study, emphasizing their careful selection as a valuable resource for breeding programs. These chosen genotypes serve as a foundation for creating recombinant populations through strategic crosses, fostering maximum genetic diversity for the development of new rice lines. This innovative approach aligns with sustainable and effective crop improvement strategies.

Strengths and Weaknesses of Genotypes Based on MGIDI Factors: **Fig. 5**. provides a comprehensive overview of the strengths and weaknesses exhibited by different

genotypes, delineated by the contribution of each factor to the Multi-Trait Genotype–Ideotype Distance Index (MGIDI). Accessions associated with Factor 1 (FA1), such as SMB3 and MTU3626, demonstrate particular strengths in traits such as 1000 grain weight (g), seedling fresh weight (mg), seedling dry weight (mg), shoot length (cm), root length (cm), seedling height (cm), seedling vigour index-I, and seedling vigour index-II (**Fig. 5. and Table 3**). On the other hand, accessions like 239(3) and NLR33892, linked to FA2, showcase strength in traits like days to 50% flowering, plant height (cm), panicle length (cm), number of filled grains per panicle, number of grains per panicle, and grain yield per plant (g). Furthermore, MTU3626 and 405C3, associated with FA3, exhibit strength in traits like first count of germination (%), final count of germination (%), rate of germination, and germination index. Lastly, Factor 4 (FA4) with accessions like NLR33892 and SM227 demonstrates strength in traits like number of panicles per plant, spikelet fertility (%), and root to shoot ratio.

These insights into the strengths and weaknesses of genotypes can serve as valuable guidance for selecting parents in future breeding programmes. Mamun *et al*. (2022) in their study underscored the significance of ideal rice mutants identified through MGIDI, emphasizing their potential for improved quantitative traits. The rice lines identified, including those highlighted in this study, emerge as optimal genotypes for future rice breeding programs,
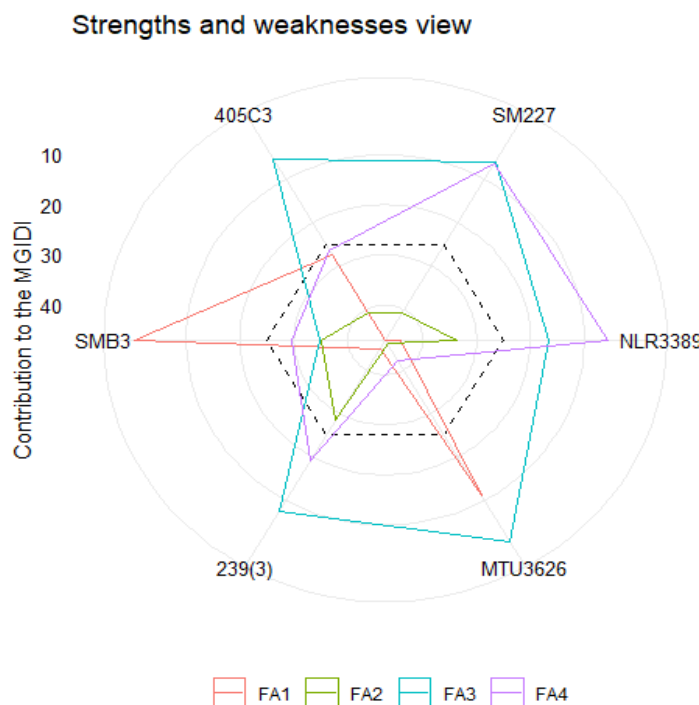

Fig. 5. The strengths and weaknesses of the selected genotypes are shown as the proportion of each factor on the computed multi-trait genotype–ideotype index (MGIDI). The smaller the proportion explained by a factor (closer to the external edge), the closer the traits within that factor are to the ideotype. The black broken circle at the center shows the theoretical value if all the factors contributed equally

playing a pivotal role in enhancing overall crop quality. This strategic utilization of factors and traits contributes to the development of resilient and high-performing rice varieties.

The Multi-Trait Genotype–Ideotype Distance Index (MGIDI) demonstrated remarkable efficiency in selecting superior rice genotypes, showcasing desirable gains across multiple traits. The identified genotypes, including SM227, NLR33892, MTU3626, 239(3), SMB3, and 405C3, through MGIDI underscores their potential for commercial release or utilization as key breeding materials in direct-seeded rice improvement programs. The comprehensive assessment of strengths and weaknesses provided valuable insights, emphasizing the significance of an ideal rice genotype with improved quantitative traits. These identified genotypes stand out as prime candidates for prospective breeding programs, positioning MGIDI as a transformative tool for enhancing rice varieties with enhanced early seedling vigor and yield traits. The adoption of MGIDI presents a promising and innovative pathway for advancing rice breeding strategies, contributing to the development of resilient and high-performing rice varieties poised for sustainable agricultural practices.

REFERENCES

Adewumi, A.S., Asare, P.A., Adejumobi, I.I., Adu, M.O., Taah, K.J., Adewale, S., Mondo, J.M. and Agre, P.A. 2023. Multi-trait selection index for superior agronomic and tuber quality traits in bush yam (*Dioscorea praehensilis* Benth.). *Agronomy*, **13**(3): 682. [Cross Ref]

Benakanahalli, N.K., Sridhara, S., Ramesh, N., Olivoto, T., Sreekantappa, G., Tamam, N., Abdelbacki, A.M.M., Elansary, H.O. and Abdelmohsen, A.M. 2021. A framework for identification of stable genotypes basedon MTSI and MGDII indexes: An example in guar (*Cymopsis tetragonoloba* L.). *Agronomy*, **11**: 1221. [Cross Ref]

Bizari, E.H., Val, B.H.P., Pereira, E.D.M., Mauro, A.O.D. and Unêda-Trevisoli, S.H. 2017. Selection indices for agronomic traits in segregating populations of soybean1. *Revista Ciência Agronômica*, **48**: 110-117. [Cross Ref]

Cattel, R.B. 1973. Factor analysis, Greenwood Press, Westport, CT.

Cattell, R.B. 1966. "The scree test for the number of factors," Multivariate Behavioural Research, **1:** 245-276. [Cross Ref]

Costello, A.B. and Osborne, J. 2005. Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical assessment, research, and evaluation*, **10**(1): 7.

Donald, C.M. 1968. The breeding of crop ideotypes. *Euphytica*, **17**: 385–403. [Cross Ref]

Graham, M.H. 2003. Confronting multicollinearity in ecological multiple regression. *Ecology*, **84**(11): 2809-2815. [Cross Ref]

Hair, J.F., Anderson, R.E., Tatham, R.L. and Black, W.C. 1998. Multivariate data analysis (5th ed.), N J: Prentice-Hall, Upper Saddle River.

Hazel, L.N. 1943. The genetic basis for constructing selection indexes. *Genetics*, **28**: 476–490. [Cross Ref]

ISTA-International Seed Testing Association. 2015. ISTA rules full issue. *International Rules for Seed Testing*. 1-276. [Cross Ref]

Jalalifar, R., Sabouri, A., Mousanejad, S. and Dadras, A.R. 2023. Estimation of genetic parameters and identification of leaf blast-resistant rice RILs using cluster analysis and MGIDI. *Agronomy*, **13**(11): 2730. [Cross Ref]

Kaiser, H.F. 1958. The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, **23**: 187–200. [Cross Ref]

Kutner, M.H., Nachtsheim, C.J., Neter, J. and Li, W. 2005. *Applied linear statistical models*. McGraw-hill.

Mamun, A.A., Islam, M.M., Adhikary, S.K. and Sultana, M.S. 2022. Resolution of genetic variability and selection of novel genotypes in EMS induced rice mutants based on quantitative traits through MGIDI. *Intl J Agric Biol,* **28**:100–112.

Meier, C., Marchioro, V.S., Meira, D., Olivoto, T. and Klein, L.A. 2021. Genetic parameters and multiple-trait selection in wheat genotypes. *Pesquisa Agropecuária Tropical*, **51**. [Cross Ref]

Olivoto, T. and Lúcio, A.D.C. 2020. Metan: An R package for multi-environment trial analysis. *Methods in Ecology and Evolution*, **11**(6): 783-789. [Cross Ref]

Olivoto, T., de Souza, V.Q., Nardino, M., Carvalho, I.R., Ferrari, M., de Pelegrin, A.J., Szareski, V.J. and Schmidt, D. 2017. Multicollinearity in path analysis: a simple method to reduce its effects. *Agronomy journal*, **109**(1): 131-142. [Cross Ref]

Olivoto, T. and Nardino, M. 2021. MGIDI: A novel multi-trait index for genotype selection in plant breeding. *Bioinformatics*, **37**: 1383–1389. [Cross Ref]

Palaniyappan, S., Ganesan, K.N., Manivannan, N., Ravichandran, V. and Senthil, N. 2023. Multi trait genotype-ideotype distance index-A tool for identification of elite parental inbreds for developing

heterotic hybrids of fodder maize (*Zea mays* L.). *Electronic Journal of Plant Breeding*, **14**(3): 841-849. [Cross Ref]

Prasad, B.P.M., Patil, B.R., Geeta, D. and Matiwade, P.S. 2023. Principal component analysis (PCA) and hierarchial clustering in tobacco (*Nicotiana tabacum* L.) for yield and yield attributing traits. *Electronic Journal of Plant Breeding*, **14** (2):737-746. [Cross Ref]

Rocha, J.R.D.A.S.D.C., Machado, J.C. and Carneiro, P.C.S., 2018. Multi trait index based on factor analysis and ideotype-design: Proposal and application on elephant grass breeding for bioenergy. *Gcb Bioenergy*, **10**(1): 52-60. [Cross Ref]

Shrestha, N. 2021. Factor analysis as a tool for survey analysis. *American Journal of Applied Mathematics and Statistics*, **9**(1): 4-11. [Cross Ref]

Smith, H.A. 1936. Discriminant function for plant selection. *Ann. Eugen.* **7**: 240–250. [Cross Ref]

Uddin, M.S., Billah, M., Afroz, R., Rahman, S., Jahan, N., Hossain, M.G., Bagum, S.A., Uddin, M.S., Khaldun, A.B.M., Azam, M.G. and Hossain, N. 2021. Evaluation of 130 Eggplant (*Solanum melongena* L.) genotypes for future breeding program based on qualitative and quantitative traits, and various genetic parameters. *Horticulturae*, **7**(10): 376. [Cross Ref]

Volpato, L., Rocha, J.R.A.S.C, Alves, R.S., Ludke, W.H., Borém, A. and Silva, F.A. 2019. Inference of population effect and progeny selection via a multi-trait index in soybean breeding. *Acta Scientiarum,* **43**: e44623. [Cross Ref]

Woyann, L.G., Meira, D., Matei, G., Zdziarski, A.D., Dallacorte, L.V., Madella, L.A. and Benin G. 2019. Selection indexes based on linear-bilinear models applied to soybean breeding. *Agronomy Journal,* **112**: 1-8. [Cross Ref]