J Bangladesh Agril Univ 22(1): **17-28**, 2024 https://doi.org/10.5455/JBAU.174200

ISSN 1810-3030 (Print) 2408-8684 (Online)

Journal of Bangladesh Agricultural University

Journal home page: http://baures.bau.edu.bd/jbau

## Research Article

# Identifying Key Determinants of Rice Yield Potential through Multiple Statistical Techniques

**Md. Asif Rahman[1], Md. Abdullah Al Mamun[2], Mohammad Rafiqul Islam[1], Md. Ruhul Quddus[3], Naznin Akter Munna[4], Md. Abdul Azim[5], Md. Sakhawat Hosen Galib[6], Khandakar Md Iftekharuddaula[1] and Md. Abu Syed[1]✉**

[1]Plant Breeding Division, Bangladesh Rice Research Institute, Gazipur 1701, Bangladesh
[2]Agricultural Statistics Division, Bangladesh Rice Research Institute, Gazipur 1701, Bangladesh
[3]Hybrid Rice Division, Bangladesh Rice Research Institute, Gazipur 1701, Bangladesh
[4]Breeding Division, Bangladesh Sugarcrop Research Institute, Pabna 6620, Bangladesh
[5]Biotechnology Division, Bangladesh Sugarcrop Research Institute, Pabna 6620, Bangladesh
[6]Department of Genetics and Plant Breeding, Bangabandhu Sheikh Mujibur Rahman Agricultural University, Salna, Gazipur 1706, Bangladesh

**ARTICLE INFO**

**Article history**
Received: 26 October 2023
Accepted: 25 March 2024
Published: 31 March 2024

**Keywords**
Rice yield components, Stepwise regression, Bayesian linear regression, Exploratory factor analysis, Path analysis

**Correspondence**
Md. Abu Syed
✉: msyedso@yahoo.com



**ABSTRACT**

Grain yield is complex traits, influenced by genetics and the environment, posing challenges for prediction. The goal of the present study was to identify key traits that contribute to rice yield by applying seven statistical techniques. The study examined twenty-four rice genotypes following a randomized complete block design (RCBD) with three replications. Pearson's correlation analysis revealed several traits exhibiting significant positive correlations with grain yield (r≥0.60), including thousand-seed weight (r = 0.63), filled seeds/panicle (r = 0.74), and number of panicles/m$^2$ (r = 0.70). Multiple linear regression identified significant predictors as the number of panicles/m$^2$ (R=0.01) and thousand-seed weight (R=0.12). Stepwise linear regression suggested key yield indicators *viz.* the number of panicles/m$^2$ (R=0.01), filled seeds/panicle (R=0.02), thousand-seed weight (R=0.11), and panicle length (R=0.13). Bayesian linear regression found similar traits with a Bayes factor of 995.2 and an R-squared value of 0.77. Exploratory factor analysis showed grain number/panicle, filled seeds percentage, filled seeds/panicle, thousand-seed weight, grain width, and number of panicles/m$^2$ highly influenced grain yield and explained 61.9% total variance. Principal components analysis revealed the first three components explaining 79.1% yield variation, with thousand-seed weight, grain width, and number of panicle /m$^2$. Path analysis demonstrated the number of panicles/m$^2$, filled seeds/panicle, thousand-seed weight, and panicle length have large and significant positive direct effects on grain yield. These findings strongly suggest that selecting breeding materials with traits like high panicle density/m$^2$, larger panicle size, more filled seeds/panicle, and higher thousand-seed weight, can significantly increase the rice yield.



## Introduction

Rice (*Oryza sativa*) is a vital staple food for over half of the global population (Khush, 2005). Farmers favour its extensive cultivation due to its adaptability to diverse ecosystems and lower risks than other crops. However, as the global population grows, the projected need to produce 14,886 million tons (MT) of food by 2050 highlights the importance of increasing rice production (Islam and Karim, 2019). Currently, worldwide rice production stands at 503.27 MT, with China leading as the largest producer (USDA, 2022).

Rice holds a central position in the lives of Bengalis, permeating every aspect of their culture, politics, and economy. It is not merely a staple food but a symbol of identity and significance in people's lives. Despite the changes in Bengali life, rice continues to shine brightly, retaining its distinctive glory as staple food (Hassan, 2021). Agriculture contributes about 11.66% (FY 2021-22) to Bangladesh's total Gross Domestic Product (GDP), while rice is playing a significant role within the agricultural sector (BBS, 2022). Additionally, rice represents 70% of the agricultural GDP and one-sixth of

**Cite This Article**
Rahman, M.A., Mamun, M.A.A., Islam, M.R., Quddus, M.R., Munna, N.A., Azim, M.A., Galib, M.S.H., Iftekharuddaula, K.M. and Syed, M.A. 2024. Identifying Key Determinants of Rice Yield Potential through Multiple Statistical Techniques. *Journal of Bangladesh Agricultural University,* 22(1): 17-28. https://doi.org/10.5455/JBAU.174200

the national income (Sayeed and Yunus, 2018). Rice is cultivated year-round in Bangladesh across three seasons: Aus, Aman, and Boro. Since independence, rice production has tripled, increasing from around 11 million tonnes in 1971–72 to approximately 36.6 million tonnes in 2019–20 (BBS, 2020).

The diversity in yield-attributing traits, such as tillering capacity, panicle size, and grain size, provides options for breeders to select and combine traits that contribute to higher grain yield (Rashid et al., 2017 and Saha et al., 2019). Understanding the genetic connection between grain yield and its components enhances efficiency of breeding program by identifying optimal selection indices, revealing the significance of component traits in rice breeding. The selection indices method selects traits based on their importance for breeding objectives, considering both linear and nonlinear combinations, focusing on the best traits rather than individual ones (Hazel and Lush, 1942). The tiller number per unit area, grain number per panicle (GN/P), and thousand-seed weight (TSW) are crucial determinants of rice grain yield. According to Rahman et al., (2023), tiller number exerts the most significant influence on rice yield, GN/P plays a significant role, as stated by Hu et al., (2021), and TSW, is influenced by grain length (GL), grain width (GW), and grain thickness, has emphasized by Zuo et al., (2021).

Various statistical techniques, such as Pearson's correlation, multiple linear regression, stepwise regression, Bayesian linear regression, factor analysis, principal component analysis, and path analysis., have been employed to explain rice yield determinants. These methods facilitate a comprehensive understanding of the relationships and factors influencing rice yield. The correlation coefficient helps identify traits closely related to yield and facilitates breeders to select those traits for yield improvement (Shrestha et al., 2018). The direct and indirect contribution of the yield variables on yield can be measured through the path analysis approach (Dewey and Lu, 1959). Factor analysis (Walton, 1972) helps identify key components of variation in large datasets that can be used in rice (Ahmadikhah et al., 2008). As Nasr and Leilah (1993) and Mohamed (1999) demonstrated, stepwise multiple linear regression is more effective than full model regression for determining yield prediction equations. Bayesian linear regression, on the other hand, calculates posterior probabilities of coefficients to establish accurate predictive yield equations. Principal components analysis (PCA) simplifies complex data sets by transforming correlated variables into a smaller collection of meaningful principal components. These components are linear combinations of the original variables, as established by Dunn (1992). Considering the statistical methods discussed above for explaining rice yield, the objective of the study was to apply diverse analytical methods to comprehensively understand the relationships between various rice traits and grain yield. This would give plant breeders insights of using suitable statistical methods for selecting optimal traits when developing new high-yielding varieties

## Materials and Methods

The experiment was carried out at the BRRI Regional Station Satkhira in Boro season from December 2021 to April 2022, geographically located at 22°45'11.7"N latitude and 89°06'29.4"E longitude. Twenty-three breeding lines and one check variety were evaluated in RCB design with three replications (Table 1 and Fig. 1). Thirty-five days old seedlings were transplanted at 20 cm × 20 cm spacing using single seedling/hill following RCBD with three replications. The unit plot size was 7.5 $m^2$. Fertilizer management, weeding, pest control measures, and other essential cultural practices were methodically implemented to optimize crop growth and yield (Biswas et al., 2016 and Debsharma et al. 2023).

**Table 1. List of 24 rice genotypes evaluated for thirteen characters at BRRI Regional Station, Satkhira, during Boro 2021-22**

| SL | Name | SL | Name | SL | Name | SL | Name |
|---|---|---|---|---|---|---|---|
| G1 | SAT001 | G7 | SAT011 | G13 | SAT020 | G19 | SAT036 |
| G2 | SAT002 | G8 | SAT012 | G14 | SAT031 | G20 | SAT037 |
| G3 | SAT003 | G9 | SAT013 | G15 | SAT032 | G21 | SAT038 |
| G4 | SAT008 | G10 | SAT014 | G16 | SAT033 | G22 | SAT039 |
| G5 | SAT009 | G11 | SAT018 | G17 | SAT034 | G23 | SAT040 |
| G6 | SAT010 | G12 | SAT019 | G18 | SAT035 | G24 | BRRI dhan28 |

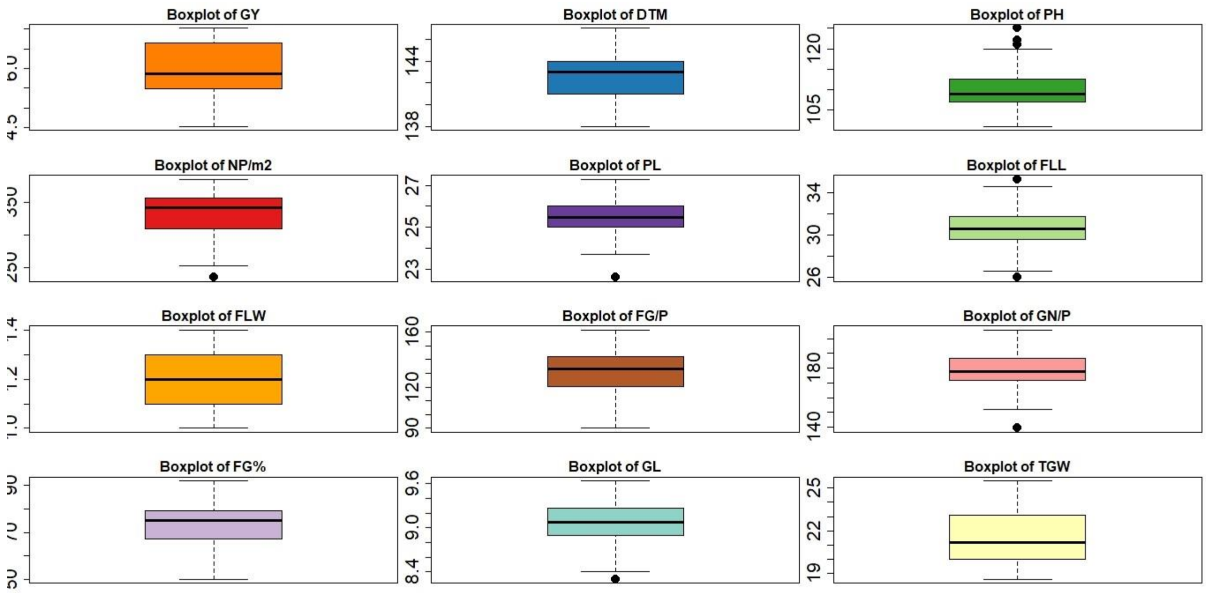

Fig. 1. Boxplot of estimated rice variables, providing a visual representation of their distribution.

DTM - Days to 80% maturity (days), PH - Plant height (cm), NP/m² - Number of panicles per m², PL - Panicle length (cm), FLL - Flag leaf length (cm), FLW - Flag leaf width (cm), FS/P - Filled grain per panicle, GN/P - Number of grains per panicle, FS (%) - Filled seeds percentage, GL - Grain length (mm), GW - Grain width (mm), TSW - Thousand-seed weight (gm) GY - Grain yield (t/ha).

### *Data collection*

Ten plants from each replication were used to collect data on plant height (cm), number of panicles per square meter, flag leaf length (cm), flag leaf wide (cm), and panicle length (cm). Days to 80% maturity were measured when 80% of the grains in the panicles were matured. Grain yield (t/ha) was calculated using the grain weight of whole plots adjusted to 14% moisture. The fertile seed percentage, filled seeds, thousand-seed weight, seed length, and seed breath were also measured.

### *The statistical techniques employed in the analysis*

In this study, seven different statistical techniques namely Pearson's correlation, Multiple linear regression, Stepwise multiple linear regression, Bayesian Linear Regression, Exploratory Factor Analysis, Principal components analysis, and Path analysis were applied.

### *Pearson's correlation*

The Pearson correlation coefficient quantifies the strength and direction of the linear relationship between two continuous variables (Asuero et al., 2006). Pearson's correlation was estimated between yield and yield components.

### *Multiple linear regression*

Multiple linear regression is a statistical method used to predict the value of a single dependent variable based on the values of two or more independent variables (Eberly, 2007). This method was used to estimate the contribution of each yield component and create a prediction model for grain yield (GY).

### *Stepwise multiple linear regression*

Stepwise multiple linear regression is an iterative statistical technique that adds or removes independent variables from a regression model based on their statistical significance (Steyerberg et al., 1999). The stepwise program performs a sequential multiple linear regression analysis by adding one variable at each step. The selected variable is the one that contributes the most to reducing the sum of squares error. Moreover, it exhibits the strongest partial correlation with the dependent variable, accounting for the constant values of the previously included variables.

### *Bayesian linear regression*

Bayesian linear regression is a statistical approach that estimates linear regression coefficients through Bayesian inference rather than frequentist methods (Gelman et al., 1995). It incorporates prior information about the coefficients and updates this prior distribution based on the observed data to determine a posterior distribution. Bayesian linear regression allows for probabilistic estimates of the regression parameters and can account for uncertainty distinct from frequentist linear regression.

### *Exploratory factor analysis*

Exploratory Factor Analysis (EFA) is a statistical approach employed to uncover the fundamental structure of a relatively large set of variables (Exploratory Factor Analysis - Wikipedia, 2012). The loading of the first factor was calculated and then repeated on the remaining matrix for additional factors. The process halted when a factor's contribution to the total trace percentage fell below 10%. Following the extraction process, the factor loading matrix underwent a varimax orthogonal rotation, a statistical method introduced by Kaiser (1958). This technique simplifies factor interpretation by maximizing variance, aiding in identifying meaningful and distinct factors. Factor analysis reveals groupings and contribution percentages to the total dependence structure. Factor analysis, uniqueness refers to the proportion of variance in a variable not accounted for by the extracted factors.

*Principal components analysis*

Principal components analysis classifies variables into key components based on their overall variation. According to Dunn (1992), the first principal component captures the most variation in the data, whereas the subsequent components explain the least amount of variation.

*Path analysis*

Path analysis is a statistical tool for exploring relationships among variables, assessing direct and indirect effects, and examining causal relationships in a model. Path coefficient analysis used a phenotypic correlations coefficient, with grain yield as the effect and other estimated traits as the cause. It was used to calculate the effects of component characters on grain yield directly and indirectly (Dewey and Lu, 1959).

*Data analysis*

The basic statistical operations, alongside multiple linear regression analysis, exploratory factor analysis, and principal components analysis, were conducted using the Jamovi (ver. 2.3) software (The Jamovi Project, 2022). We employed the JASP, Jeffrey's Amazing Statistics Program (ver. 0.16.3) software (JASP Team, 2022) for executing stepwise multiple linear regression and Bayesian Linear Regression analysis. Pearson's correlation and Path analysis were performed with the assistance of R (Ver. 4.1) Statistical Software (R Core Team, 2021).

## Results and Discussion

An overview of the estimated rice variables is provided, including the minimum and maximum values, arithmetic mean, and standard deviation. Additionally, coefficient of variation (CV) of each variable's was calculated, ranging from 1.7% for days to 80% maturity to 12.4% for filled seeds per panicle (Table 2). Figure 1 represents the distribution of variables.

**Table 2. Basis statistics, such as minimum and maximum values, mean, standard deviation (SD), and CV (%) for the estimated variables of rice**

| Variable | Minimum | Maximum | Mean | SD | CV (%) |
|---|---|---|---|---|---|
| DTM | 138.0 | 147.0 | 142.6 | 2.4 | 1.7 |
| PH | 101.0 | 125.0 | 110.5 | 5.0 | 4.5 |
| NP/m$^2$ | 235.0 | 386.0 | 332.6 | 31.4 | 9.4 |
| PL | 22.6 | 27.3 | 25.4 | 0.9 | 3.5 |
| FLL | 26.0 | 35.3 | 30.7 | 2.0 | 6.4 |
| FLW | 1.0 | 1.4 | 1.2 | 0.1 | 10.8 |
| FS/P | 90.0 | 161.6 | 130.4 | 16.2 | 12.4 |
| GN/P | 139.1 | 205.9 | 177.6 | 12.5 | 7.0 |
| FS (%) | 50.1 | 92.0 | 73.6 | 8.7 | 11.9 |
| GL | 8.3 | 9.6 | 9.1 | 0.3 | 2.8 |
| GW | 1.8 | 2.4 | 2.0 | 0.2 | 8.0 |
| TSW | 18.6 | 25.5 | 21.5 | 1.8 | 8.2 |
| GY | 4.5 | 7.0 | 5.9 | 0.7 | 11.8 |

DTM - Days to 80% maturity (days), PH - Plant height (cm), NP/m$^2$ - Number of panicles per m$^2$, PL - Panicle length (cm), FLL - Flag leaf length (cm), FLW - Flag leaf width (cm), FS/P - Filled seeds per panicle, GN/P - Number of grains per panicle, FS (%) - Filled seeds percentage, GL - Grain length (mm), GW - Grain width (mm), TSW - Thousand-seed weight (gm) GY - Grain yield (t/ha).

*Correlation analysis*

All variables in the study showed significant positive correlations with grain yield, except plant height (r=-0.32). Grain yield was positively linked with thousand-seed weight (r=0.63), grain width (r =0.54), grain length (r =0.41), fertile seed percentage (r=0.58), Filled seeds per panicle (r =0.74), flag leaf wide (r =0.35), panicle length (r =0.34), number of panicles per m$^2$ (r =0.70), days to 80% maturity (r=0.43) in Fig. 2 (Pearson's correlation among estimated variables/traits of rice.)

Several studies have investigated the correlation between various morphological traits and grain yield. Ogunbayo et al., (2014) discovered a significant positive association between grain yield and several characteristics, including panicle length, flag leaf width, grain length, number of panicles per square meter, and thousand-seed weight. This suggests that increasing these traits will likely result in higher grain yield. Tripathy and Sahoo, (2021) observed a negative correlation between grain yield and plant height. Kumar et al., (2022) focused on grain length and breadth and found a significant positive correlation between these traits and yield. This suggests that larger grains are associated with higher yield. They found grain yield significantly and positively correlated with filled grain number per panicle and filled grain percentage. Rahman et al., (2013) found a positive correlation between flag leaf length and yield. Chandra et al., (2009) focused on the number of grains per panicle and found a positive connection with yield. Their findings suggest a higher grain count per panicle is associated with greater grain yield. Lakshmi et al., (2014) explored the correlation between days to maturity and yield and found a positive relationship.

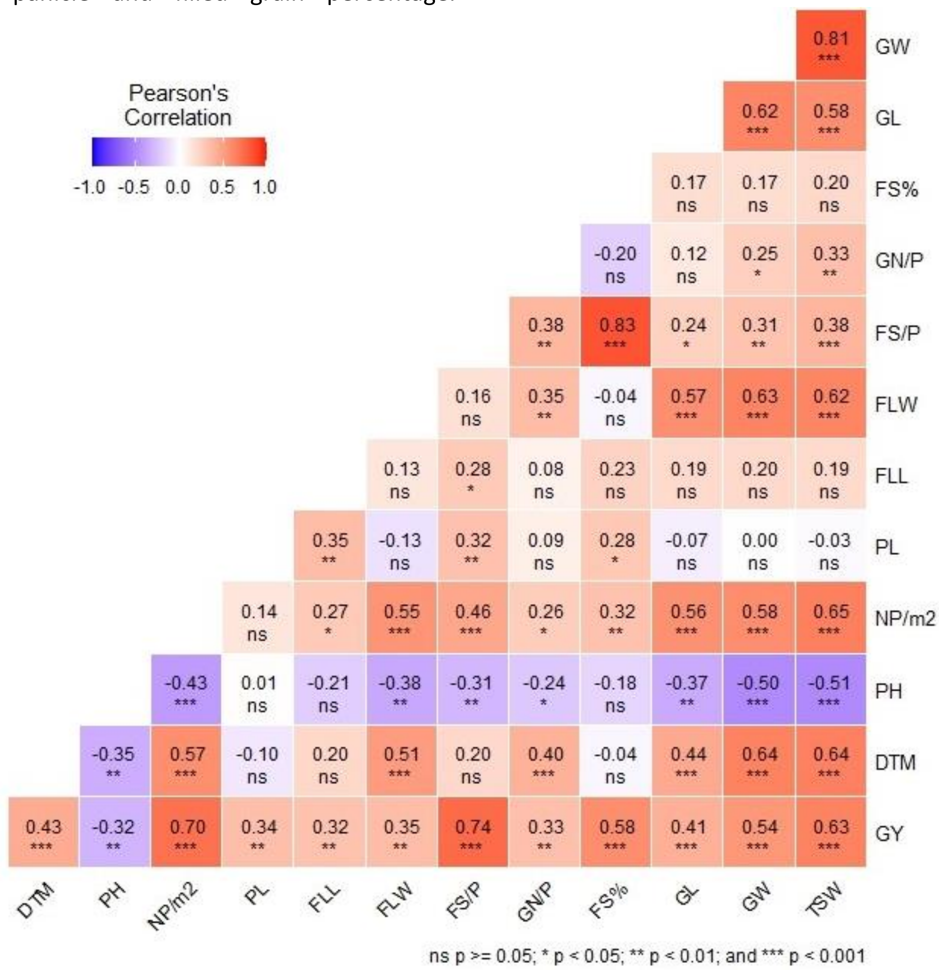


Fig. 2. Pearson's correlation among estimated variables/traits of rice

DTM - Days to 80% maturity (days), PH - Plant height (cm), NP/m² - Number of panicles per m², PL - Panicle length (cm), FLL - Flag leaf length (cm), FLW - Flag leaf width (cm), FS/P - Filled seeds per panicle, GN/P - Number of grains per panicle, FS (%) - Filled seeds percentage, GL - Grain length (mm), GW - Grain width (mm), TSW - Thousand-seed weight (gm) GY - Grain yield (t/ha).

### *Multiple linear regression analysis*

Regression coefficients and the probability of estimated variables predicting rice grain yield are displayed in Table 3. The model summary table shows that the R-squared value for the model was 0.78, which directs that the model was able to determine 78% of the variation in the dependent variable. An adjusted $R^2$ value of 0.74 suggests that the model accounts for 74% of the variance, considering the number of predictors and the potential overfitting. The Root MSE was 0.36, indicating an average prediction deviation of around 0.36 units from the actual values.

**Table 3. The model coefficient, standard error (SE), T-value and probability of the estimated variables in predicting rice grain yield by the multiple linear regression analysis**

| Predictor | Estimate | SE | t | p |
|---|---|---|---|---|
| Intercept | -7.76 | 7.62 | -1.02 | $0.31^{NS}$ |
| DTM | 0.01 | 0.03 | 0.17 | $0.87^{NS}$ |

| | | | | |
|---|---|---|---|---|
| PH | 0.02 | 0.01 | 1.65 | $0.11^{NS}$ |
| $NP/m^2$ | 0.01 | 0.00 | 3.48 | $9.525\times10^{-4}$*** |
| PL | 0.11 | 0.06 | 1.95 | $0.06^{NS}$ |
| FLL | 0.01 | 0.03 | 0.26 | $0.79^{NS}$ |
| FLW | -0.38 | 0.50 | -0.77 | $0.45^{NS}$ |
| FS/P | 0.01 | 0.04 | 0.30 | $0.77^{NS}$ |
| GN/P | 0.004 | 0.03 | 0.13 | $0.90^{NS}$ |
| FS% | 0.01 | 0.08 | 0.14 | $0.89^{NS}$ |
| GL | -0.03 | 0.24 | -0.13 | $0.90^{NS}$ |
| GW | 0.22 | 0.53 | 0.41 | $0.68^{NS}$ |
| TSW | 0.12 | 0.05 | 2.62 | 0.01*** |

** and *** means significant at 5% & 10%, NS; Not Significant, $R^2$=0.78, Adjusted $R^2$=0.74, Root MSE=0.36, DTM - Days to 80% maturity (days), PH - Plant height (cm), $NP/m^2$ - Number of panicles per $m^2$, PL - Panicle length (cm), FLL - Flag leaf length (cm), FLW - Flag leaf width (cm), FS/P - Filled seeds per panicle, GN/P - Number of grains per panicle, FS (%) - Filled seeds percentage, GL - Grain length (mm), GW - Grain width (mm), TSW - Thousand-seed weight (gm) GY - Grain yield (t/ha).

The coefficients table revealed that the independent variables $NP/m^2$ and TSW were significant predictors for the dependent variable. These results suggest that the number of panicles/$m^2$ and thousand-seed weight are important traits to consider for rice selection in breeding programs, as they positively impacted the predicted value of grain yield. These findings were also supported by Ogunbayo et al., (2014). Xing and Zhang, (2010) proposed that grain yield in rice is a complex trait, determined multiplicatively by three component traits: the number of panicles (*i.e.*, the number of effective tillers per plant), the number of grains per panicle, and the grain weight (measured as thousand-seed weight). Our study aligned with the findings of Xing and Zhang, (2010) except for the number of grains per panicle through this model.

Therefore, the predicted equation of grain yield,

$$Y = -7.76 + 0.01*DTM + 0.02*PH + 0.01*NP/m^2 + 0.11*PL + 0.01*FLL - 0.38*FLW + 0.01*FS/P + 0.004*GN/P + 0.01*FS\% - 0.03*GL + 0.22*GW + 0.12*TSW$$

### *Stepwise linear regression analysis*

The Model Summary table in Table 4 (a) presents the outcomes of five regression models. Model 5 seemed the best, having the highest R-squared (0.77) and adjusted R-squared (0.76). It also had the lowest RMSE value (0.35), indicating the highest accuracy among the models. According to the study's findings, Table 4 (a & b) indicates that Model 5 was the most accurate or optimal model for predicting grain yield.

**Table 4 (a). Model Summary – Yield (Stepwise Linear Regression)**

| Model | R | $R^2$ | Adjusted $R^2$ | RMSE |
|---|---|---|---|---|
| 1 | 0.00 | 0.00 | 0.00 | 0.70 |
| 2 | 0.74 | 0.55 | 0.54 | 0.47 |
| 3 | 0.85 | 0.72 | 0.71 | 0.38 |
| 4 | 0.86 | 0.75 | 0.74 | 0.36 |
| 5 | 0.88 | 0.77 | 0.76 | 0.35 |

R -Correlation Coefficient, $R^2$ -Coefficient of Determination, RMSE -Root Mean Square Error

**Table 4 (b). The model estimate, standard error (SE), T-value and probability of the estimated variables in predicting rice grain yield by stepwise linear regression analysis**

| Model | Predictor | Estimate | Standard Error | t | p |
|---|---|---|---|---|---|
| 1 | (Intercept) | 5.92 | 0.08 | 71.89 | $4.594\times10^{-68}$ |
| 2 | (Intercept) | 1.76 | 0.46 | 3.86 | $2.525\times10^{-4}$ |
| | FS/P | 0.03 | 0.00 | 9.18 | $1.244\times10^{-13}$ |
| 3 | (Intercept) | -0.48 | 0.50 | -0.95 | 0.345 |
| | FS/P | 0.02 | 0.00 | 7.31 | $3.683\times10^{-10}$ |
| | $NP/m^2$ | 0.01 | 0.00 | 6.42 | $1.446\times10^{-8}$ |
| 4 | (Intercept) | -1.30 | 0.56 | -2.34 | 0.022 |
| | FS/P | 0.02 | 0.00 | 7.25 | $4.984\times10^{-10}$ |
| | NP/m2 | 0.01 | 0.00 | 3.89 | $2.343\times10^{-4}$ |
| | TSW | 0.09 | 0.03 | 2.90 | 0.005 |
| 5 | (Intercept) | -4.40 | 1.34 | -3.29 | 0.002 |
| | FS/P | 0.02 | 0.00 | 6.33 | $2.391\times10^{-8}$ |
| | $NP/m^2$ | 0.01 | 0.00 | 3.75 | $3.687\times10^{-4}$ |
| | TSW | 0.11 | 0.03 | 3.46 | $9.552\times10^{-4}$ |

| | | | | |
|---|---|---|---|---|
| PL | 0.13 | 0.05 | 2.53 | 0.014 |

*Note.* The following covariates were considered but not included: DTM, PH, FLL, FLW, GN/P, FS%, GL, GW.
FS/P - Filled seeds per panicle, NP/$m^2$ - Number of panicles per $m^2$, TSW - Thousand-seed weight (gm), PL - Panicle length (cm),

Overall, the results of the stepwise linear regression analysis suggest that the variables number of panicles per square meter (NP/$m^2$), filled seeds per panicle (FS/P), thousand-seed weight (TSW), and panicle length (PL) are all strongly associated with grain yield of rice. Our findings were supported by an investigation conducted by Luzikihup (1988); the number of filled seeds per panicle, the number of panicles per plant, and the 1000-seed weight were identified as important factors influencing grain yield except for panicle length.
The predicted equation of grain yield,

GY = -4.40 + 0.01*NP/$m^2$ + 0.02*FS/P + 0.11*TSW + 0.13*PL

### *Bayesian linear regression*

The NP/$m^2$ + PL + FS/P + TSW model demonstrated the highest posterior probability (0.13), Bayes factor (995.2), and R-squared value (0.77) among the tested models in Table 5 (a). These results strongly suggest that the NP/$m^2$ + PL + FS/P + TSW model was the most optimal choice. The NP/$m^2$ + FS/P + TSW model ranked as the second-best model based on the findings. The posterior summaries of coefficients are presented in Table 5 (b).

**Table 5 (a). Model Comparison – Yield (Best five models) by Bayesian Linear Regression**

| Models | P(M) | P(M\|data) | $BF_M$ | $BF_{10}$ | $R^2$ |
|---|---|---|---|---|---|
| NP/m2 + PL + FS/P + TSW | $1.55\times10^{-4}$ | 0.13 | 995.27 | 1.00 | 0.77 |
| NP/m2 + FS/P + TSW | $3.50\times10^{-4}$ | 0.12 | 384.66 | 0.39 | 0.75 |
| NP/m2 + FS/P | 0.001 | 0.07 | 65.35 | 0.07 | 0.72 |
| PH + NP/m2 + PL + FS/P + TSW | $9.71\times10^{-5}$ | 0.05 | 538.05 | 0.59 | 0.78 |
| PH + NP/m2 + FS/P + TSW | $1.55\times10^{-4}$ | 0.03 | 224.25 | 0.25 | 0.76 |

The table presents a model comparison for yield using Bayesian Linear Regression. It includes various models with corresponding probabilities (P(M)), conditional probabilities (P(M|data)), Bayes Factors for the model ($BF_M$), Bayes Factor in favour of the alternative hypothesis ($BF_{10}$), and the coefficient of determination ($R^2$).

**Table 5 (b). Posterior Summaries of Coefficients (Bayesian Linear Regression)**

| | | | | | | 95% Credible Interval | |
|---|---|---|---|---|---|---|---|
| Coefficient | Mean | SD | P(incl) | P(incl\|data) | $BF_{inclusion}$ | Lower | Upper |
| Intercept | 5.92 | 0.04 | 1 | 1 | 1 | 5.84 | 6 |
| NP/m2 | 0.01 | 0 | 0.5 | 0.99 | 87.04 | 0 | 0.01 |
| PL | 0.12 | 0.05 | 0.5 | 0.68 | 2.11 | 0.03 | 0.22 |
| FS/P | 0.02 | 0 | 0.5 | 0.9 | 9.21 | 0.01 | 0.02 |
| TSW | 0.11 | 0.03 | 0.5 | 0.91 | 9.54 | 0.04 | 0.17 |

Table 5 (b) presents the posterior summaries of coefficients for Bayesian Linear Regression. It includes the mean, standard deviation (SD), probabilities of inclusion (P(incl)), posterior probabilities of inclusion given the data (P(incl|data)), Bayes Factor for inclusion (BFinclusion), and the 95% credible interval for each coefficient.

Our findings were also supported by Xing and Zhang, (2010), who found that rice grain yield is based on three major traits: grain weight or thousand-seed weight, grain number per panicle, and effective tiller number. These traits are also known as the yield components of rice. In Figure 2, the data revealed a positive correlation (r = 0.32) between the length of the panicle and the number of filled seeds per panicle. This suggests that larger panicles tend to contain a greater quantity of filled seeds. Laza et al., (2004) found that large panicle size or higher panicle length can also contribute to rice grain yield positively.

Therefore, the predicted equation of grain yield,
Y = 5.92 + 0.01*NP/m2 + 0.12*PL + 0.02*FS/P + 0.11*TSW

### *Exploratory factor analysis*

The exploratory factor analysis (EFA) results showed that the first three factors explain 61.9% of the variance in the data in Table 6 (b). Factor 1 was the most important, explaining 32.6% of the variation in the data. Factor 1 was positively correlated (factor loading > 0.50) with GW, TSW, FLW, GL, NP/$m^2$ and DTM and negatively correlated with PH in Table 6 (a). The suggested name for this factor is thousand-seed weight or grain width.

Factor 2 explains 19.9% of the variation in the data. Factor 2 was positively correlated (factor loading > 0.50) with GY, FS/P & FS%. The suggested name for this factor is filled seeds/panicle or Filled seeds percentage. Factor 3 explains 9.4% of the variation in the data. Factor 3 was positively correlated (factor loading > 0.50) with GN/P.

The exploratory factor analysis determined that the traits of grain number per panicle, filled seeds percentage, filled seeds per panicle, thousand-seed weight, grain width, and number of panicles per square meter made major contributions to explain the variance in rice grain yield.

Among the variables assessed, filled seeds/panicle, filled seeds percentage, thousand-seed weight, grain width, and number of panicles per $m^2$, exhibited the lowest uniqueness, as indicated by the data in Table 6 (a) and therefore, these variables contribute relatively highly to rice grain yield. Rice grain yield is determined by the number of panicles, the number of grains per panicle, and grain weight, all of which are typical quantitative traits (Xing and Zhang, 2010). Mohamed (1999) used factor analysis to classify wheat traits into key components accounting for a large proportion of variance; this study on rice revealed that factor analysis effectively categorized the rice variables into three major factors explaining 61.9% of the total variance in grain yield.

**Table 6 (a). Rotated (Varimax rotation) factor loadings and uniqueness for the estimated variables of rice**

| | Factor | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | Uniqueness |
| GW | 0.86 | | | 0.24 |
| TSW | 0.86 | | | 0.22 |
| FLW | 0.74 | | | 0.43 |
| GL | 0.71 | | | 0.48 |
| NP/m2 | 0.71 | 0.36 | | 0.36 |
| DTM | 0.7 | | | 0.45 |
| PH | -0.52 | | | 0.7 |
| FS/P | | 0.9 | | 0.11 |
| FS% | | 0.88 | -0.34 | 0.1 |
| PL | | 0.45 | | 0.78 |
| FLL | | 0.32 | | 0.86 |
| GN/P | | | 0.96 | 0 |
| GY | 0.52 | 0.69 | | 0.22 |

*Note.* 'Minimum residual' extraction method was used in combination with a 'varimax' rotation
DTM - Days to 80% maturity (days), PH - Plant height (cm), NP/$m^2$ - Number of panicles per $m^2$, PL - Panicle length (cm), FLL - Flag leaf length (cm), FLW - Flag leaf width (cm), FS/P - Filled seeds per panicle, GN/P - Number of grains per panicle, FS (%) - Filled seeds percentage, GL - Grain length (mm), GW - Grain width (mm), TSW - Thousand-seed weight (gm) GY - Grain yield (t/ha).

**Table 6 (b). Summary of factors loading for the estimated variables of rice**

| Factor | SS Loadings | % Of Variance | Cumulative % |
|---|---|---|---|
| 1 | 4.2 | 32.6 | 32.6 |
| 2 | 2.6 | 19.9 | 52.5 |
| 3 | 1.2 | 9.4 | 61.9 |

*Principal component analysis (PCA)*

Table 7 and Fig. 3 demonstrate that as the number of components increases, eigenvalue decreases. The trend reaches its maximum at three factors. This suggests that the principal components analysis successfully grouped the rice variables into three main components, explaining 79.1% of the total variation in grain yield. The first principal component (PC1) captures 51.8% of the total variation in the data.

**Table 7. Eigenvalue, factor scores and contribution of the first four principal component axes to variation in rice genotypes**

| Variables | PC1 | PC2 | PC3 | PC4 |
|---|---|---|---|---|
| DTM | 0.31 | 0.16 | -0.20 | -0.18 |
| PH | -0.28 | -0.18 | -0.11 | -0.08 |
| NP/$m^2$ | 0.35 | 0.02 | 0.03 | -0.01 |
| PL | 0.06 | -0.51 | -0.39 | -0.06 |
| FLL | 0.15 | -0.23 | -0.10 | -0.86 |
| FLW | 0.31 | 0.26 | -0.06 | 0.10 |
| FS/P | 0.26 | -0.41 | -0.01 | 0.30 |

| | | | | |
|---|---|---|---|---|
| GN/P | 0.18 | 0.12 | -0.69 | 0.25 |
| FS% | 0.16 | -0.50 | 0.40 | 0.19 |
| GL | 0.30 | 0.19 | 0.35 | -0.12 |
| GW | 0.36 | 0.12 | 0.14 | -0.03 |
| TSW | 0.36 | 0.09 | 0.00 | 0.00 |
| GY | 0.33 | -0.27 | -0.03 | 0.10 |
| Proportion of Variance | 51.8 | 18.0 | 9.3 | 6.6 |
| Cumulative Proportion | 51.8 | 69.8 | 79.1 | 85.7 |
| Eigen Values | 6.73 | 2.34 | 1.21 | 0.86 |

DTM - Days to 80% maturity (days), PH - Plant height (cm), NP/$m^2$ - Number of panicles per $m^2$, PL - Panicle length (cm), FLL - Flag leaf length (cm), FLW - Flag leaf width (cm), FS/P - Filled seeds per panicle, GN/P - Number of grains per panicle, FS (%) - Filled seeds percentage, GL - Grain length (mm), GW - Grain width (mm), TSW - Thousand-seed weight (gm) GY - Grain yield (t/ha).

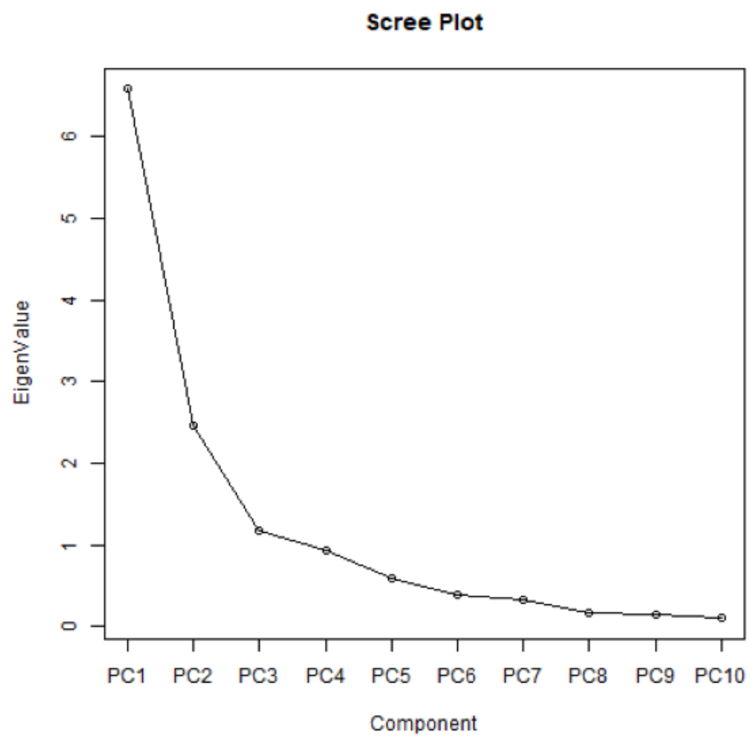


Fig. 3. Scree plot displaying eigenvalues in relation to the number of components for rice estimated variables.

Variables (factor scores > 0.3) such as TSW, GW, NP/$m^2$, FLW and DTM had the highest positive associations with PC1. PH had a negative relationship with PC1. Variables (factor scores > 0.15) such as FLW, GL & DTM had positive associations with PC2. PC2 explains 18.0% of the total variance (Table 7), contributing to a cumulative proportion of 69.8% when combined with PC1. PC3 explains 9.3% of the total variance, contributing to a cumulative proportion of 79.1% when combined with PC1 and PC2. The variables (factor scores > 0.30), such as FS% and GL had positive associations with PC3. A high correlation between PC1 and a variable indicates a strong association with the primary direction of variation in the dataset. The rice yield components comprise panicles per unit area, spikelets per panicle, spikelet weight, and spikelet sterility or filled spikelet (Fageria, 2007).

### *Path coefficient analysis*

The results of the path coefficient analysis provide valuable insights into the direct effects of twelve morphological traits of rice genotypes on yield (Table 8). Days to maturity (DTM) and taller plants (PH) had a slight positive effect on yield, with correlation coefficients of 0.03 and 0.12, respectively. Increasing the number of panicles/$m^2$ (NP/$m^2$) had the highest significant positive impact on yield (coefficient: 0.33). Longer panicles (PL) had a moderate positive influence on yield (coefficient: 0.13), while longer flag leaves (FLL) contributed slightly (coefficient: 0.03). However, wider flag leaves (FLW) had a slightly negative effect (coefficient: -0.06). A higher number of filled seeds/panicle (FS/P) positively impacts yield (coefficient: 0.32). More grains per panicle (GN/P) and filled seeds percentage (FS%) had strong positive effects (coefficients: 0.32 and 0.07, respectively). Longer grains (GL) and wider grains (GW) had negative and positive effects (coefficients: -0.01 and 0.06, respectively). Higher thousand-seed weight (TSW) significantly improves yield (coefficient: 0.30).

The results indicate that the number of panicles/m$^2$ (NP/m$^2$), the number of filled seeds per panicle (FS/P), the thousand-seed weight (TSW), and the panicle length (PL) had a positive impact on yield and contribute significantly more than other variables. Mohamed et al., (2012) revealed that factors such as the number of filled seeds per panicle, the number of panicles per square meter, and the thousand grains weight had a significant direct positive impact on grain yield through path analysis. In another investigation, panicle length also exhibited a positive direct effect on yield (Babu et al., 2012).

The statistical analysis of rice grain yield traits identified number of panicles per square meter (NP/m$^2$), thousand-seed weight (TSW), panicle length (PL), and filled grains per panicle (FS/P) as the most influential factors (Table 9). The results suggest that rice grain yield is most heavily dependent on number of panicles per square meter, thousand-seed weight, filled grain, and panicle length. Focusing improvement efforts on these vital traits will likely have the greatest impact on enhancing rice yields. Additional multi-location trials are needed to validate these results across diverse environments and genotypes.

**Table 8. Partitioning of direct and indirect effects of twelve morphological traits of 24 rice genotypes at the phenotypic level by path coefficient analysis**

| Variables | DTM | PH | NP/m2 | PL | FLL | FLW | FS/P | GN/P | FS% | GL | GW | TSW | GY |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| DTM | **0.03** | -0.04 | 0.20 | -0.01 | 0.00 | -0.03 | 0.08 | 0.03 | 0.00 | -0.01 | 0.04 | 0.20 | 0.43 |
| PH | -0.01 | **0.12** | -0.14 | 0.00 | -0.01 | 0.02 | -0.10 | -0.01 | -0.02 | 0.00 | -0.03 | -0.16 | -0.32 |
| NP/m2 | 0.02 | -0.05 | **0.33** | 0.01 | 0.01 | -0.04 | 0.15 | 0.02 | 0.04 | -0.01 | 0.03 | 0.20 | 0.70 |
| PL | 0.00 | 0.00 | 0.03 | **0.13** | 0.01 | 0.01 | 0.10 | 0.00 | 0.03 | 0.00 | 0.00 | 0.00 | 0.34 |
| FLL | 0.00 | -0.02 | 0.09 | 0.05 | **0.03** | -0.01 | 0.10 | 0.00 | 0.03 | 0.00 | 0.01 | 0.06 | 0.32 |
| FLW | 0.01 | -0.05 | 0.19 | -0.01 | 0.00 | **-0.06** | 0.06 | 0.02 | 0.00 | -0.01 | 0.04 | 0.19 | 0.35 |
| FS/P | 0.01 | -0.04 | 0.15 | 0.04 | 0.01 | -0.01 | **0.32** | 0.03 | 0.10 | 0.00 | 0.02 | 0.11 | 0.74 |
| GN/P | 0.01 | -0.03 | 0.08 | 0.01 | 0.00 | -0.02 | 0.12 | **0.07** | -0.02 | 0.00 | 0.01 | 0.11 | 0.33 |
| FS% | 0.00 | -0.02 | 0.11 | 0.04 | 0.01 | 0.00 | 0.26 | -0.01 | **0.12** | 0.00 | 0.01 | 0.06 | 0.58 |
| GL | 0.01 | -0.05 | 0.19 | -0.01 | 0.00 | -0.03 | 0.08 | 0.01 | 0.02 | **-0.01** | 0.04 | 0.17 | 0.41 |
| GW | 0.02 | -0.06 | 0.19 | 0.00 | 0.01 | -0.04 | 0.10 | 0.02 | 0.02 | -0.01 | **0.06** | 0.24 | 0.54 |
| TSW | 0.02 | -0.07 | 0.22 | 0.00 | 0.01 | -0.04 | 0.12 | 0.02 | 0.02 | -0.01 | 0.05 | **0.30** | **0.63** |
| Residual | 0.22 | | | | | | | | | | | | |

DTM - Days to 80% maturity (days), PH - Plant height (cm), NP/m$^2$ - Number of panicles per m$^2$, PL - Panicle length (cm), FLL - Flag leaf length (cm), FLW - Flag leaf width (cm), FS/P - Filled seeds per panicle, GN/P - Number of grains per panicle, FS (%) - Filled seeds percentage, GL - Grain length (mm), GW - Grain width (mm), TSW - Thousand-seed weight (gm) GY - Grain yield (t/ha).

**Table 9. Key traits identified that significantly impact rice grain yield using multiple statistical techniques**

| Variables | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Count | Frequency (%) | Rank |
|---|---|---|---|---|---|---|---|---|---|---|
| DTM | | | | | | | | 0 | 0% | |
| PH | | | | | | | | 0 | 0% | |
| NP/m2 | ■ | ■ | ■ | ■ | ■ | ■ | ■ | 7 | 100% | 1 |
| PL | | | ■ | ■ | | | ■ | 3 | 42.86% | 3 |
| FLL | | | | | | | | 0 | 0% | |
| FLW | | | | | | | | 0 | 0% | |
| FS/P | ■ | | ■ | ■ | ■ | | ■ | 5 | 71.43% | 2 |
| GN/P | | | | | ■ | | | 1 | 14.29% | 5 |
| FS% | | | | | ■ | | | 1 | 14.29% | 5 |
| GL | | | | | | | | 0 | 0% | |
| GW | | | | | ■ | ■ | | 2 | 28.57% | 4 |
| TSW | ■ | ■ | ■ | ■ | ■ | ■ | ■ | 7 | 100% | 1 |

1= Pearson's correlation, 2=Multiple linear regression, 3= Stepwise multiple linear regression, 4= Bayesian Linear Regression, 5=Exploratory Factor Analysis, 6=Principal components analysis, 7= Path analysis (phenotypic level). NP/m$^2$ and TSW were recognized by all 7 techniques as critical, while PL and FS/P were identified by over 40% of methods as important. Traits like plant height, flag leaf dimensions, and grain length were not found to have a major impact.

DTM - Days to 80% maturity (days), PH - Plant height (cm), NP/$m^2$ - Number of panicles per $m^2$, PL - Panicle length (cm), FLL - Flag leaf length (cm), FLW - Flag leaf width (cm), FS/P - Filled seeds per panicle, GN/P - Number of grains per panicle, FS (%) - Filled seeds percentage, GL - Grain length (mm), GW - Grain width (mm), TSW - Thousand-seed weight (gm) GY - Grain yield (t/ha).

## Conclusion

This statistical analysis identifies the most important traits for improving rice grain yield. Focusing on increasing panicle number, seed weight, filled grain number, and panicle length will help develop higher-yielding rice varieties. This knowledge provides rice breeders with clear targets to breed better rice. Improving these key traits will boost rice productivity to meet growing food demands, especially in rice-dependent regions. By using this information to guide breeding programs, scientists can create better rice varieties.

## Acknowledgements

The authors would like to thank Bangladesh Rice Research Institute, BRRI for technical support. We would like to thank the technical assistant staff of BRRI Regional Station, Satkhira for their support in conducting research activities.

## References

Asuero, A.G., Sayago, A., González, A.G. 2006. The correlation coefficient: An overview. Critical reviews in analytical chemistry, 36(1): 41-59. https://doi.org/10.1080/10408340500526766

Ahmadikhah, A., Nasrollanejad, S., & Alishah, O. 2008. Quantitative studies for investigating variation and its effect on heterosis of rice. *International Journal of Plant Production*, 2(4): 297-308. https://doi.org/10.22069/ijpp.2012.621

Babu, V.R., Shreya, K., Dangi, K.S., Usharani, G., Shankar, A.S. 2012. Correlation and path analysis studies in popular rice hybrids of India. *International Journal of Scientific and Research Publications,* 2(3): 1-5.

Bangladesh Bureau of Statistics (BBS). 2022. Statistical Yearbook of Bangladesh, Bangladesh Bureau of Statistics (BBS).

Bangladesh Bureau of Statistics (BBS). 2020. Statistical Yearbook of Bangladesh, Bangladesh Bureau of Statistics (BBS).

Biswas, J. K., Kabir, M. S., Ali, M. A., Kashem, M. A. 2016. Modern Rice Cultivation (Adhunik Dhaner Chash in Bangla version). Bangladesh Rice Research Institute (BRRI), Gazipur, Bangladesh.

Chandra, B.S., Reddy, T.D., Ansari, N.A., Kumar, S.S. 2009.Correlation and path analysis for yield and yield components in rice (*Oryza sativa* L.). Agricultural Science Digest, 29(1): 45-47.

Debsharma, S. K., Syed, M. A., Ali, M. H., Maniruzzaman, S., Roy, P. R., Brestic, M., ... & Hossain, A. (2022). Harnessing on genetic variability and diversity of rice (Oryza sativa L.) genotypes based on quantitative and qualitative traits for desirable crossing materials. *Genes*, 14(1), 10. https://doi.org/10.3390/genes14010010

Dewey, D.R., Lu, K. 1959. A correlation and path-coefficient analysis of components of crested wheatgrass seed production 1. *Agronomy journal*, 51(9):515-518. https://doi.org/10.2134/agronj1959.00021962005100090002x

Dunn, G. 1992. Review papers: design and analysis of reliability studies. Statistical Methods in Medical Research, 1(2): 123-157. https://doi.org/10.1177/096228029200100202

Eberly, L. E. 2007. Multiple linear regression. Methods in molecular biology, 404:165-187. https://doi.org/10.1007/978-1-59745-530-5_9

Fageria, N.K. 2007. Yield physiology of rice. *Journal of plant nutrition*, 30(6): 843-879. https://doi.org/10.1080/15226510701374831

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B. 1995. Bayesian Data Analysis. Chapman and Hall/CRC. https://doi.org/10.1201/9780429258411

Hassan, S. 2021. The rice economy. The Daily Star [cited 2023 Aug 5]. Available from: https://www.thedailystar.net/business/economy/news/the-rice-economy 2114957

Hazel, L.N., Lush, J.L. 1942. The efficiency of three methods of selection. *Journal of Heredity*, 33(11): 393-399. https://doi.org/10.1093/oxfordjournals.jhered.a105102

Hu, L., Chen, W., Yang, W., Li, X., Zhang, C., Zhang, X., Zheng, L., Zhu, X., Yin, J., Qin, P. and Wang, Y. 2021. OsSPL9 regulates grain number and grain yield in rice. Frontiers in Plant Science, 12: 682018. https://doi.org/10.3389/fpls.2021.682018

Islam, S.M., Karim, Z. 2019. World's demand for food and water: The consequences of climate change. Desalination-challenges and opportunities, 1-27. https://doi.org/10.5772/intechopen.85919

Kaiser, H.F. 1958.The varimax criterion for analytic rotation in factor analysis. Psychometrika, 23:187-200. https://doi.org/10.1007/BF02289233

Khush, G.S. 2005. What it will take to feed 5.0 billion rice consumers in 2030. Plant Molecular Biology, 59:1-6. https://doi.org/10.1007/s11103-005-2159-5

Kumar, P., Singh, G., Prasad, B. K., Kumar, R. 2022. Correlation analysis for yield and quality contributing characters involved in rice (*Oryza sativa* L.) genotypes. *International Journal of Applied Research*, 8(1): 135-141.

Lakshmi, M. V., Suneetha, Y., Yugandhar, G., Lakshmi, N. V. 2014. Correlation studies in rice (Oryza sativa L.). *International Journal of Genetic Engineering and Biotechnology*, 5(2): 121-126.

Laza, M.R., Peng, S., Akita, S., Saka, H. 2004. Effect of panicle size on grain yield of IRRI-released indica rice cultivars in the wet season. Plant Production Science, 7(3): 271-276. https://doi.org/10.1626/pps.7.271

Luzikihupi, A. 1998. Interrelationship between yield and some selected agronomic characters in rice. *Africa Crop Science Journal*, 6(3): 323-328. https://doi.org/10.4314/acsj.v6i3.27805

Mohamed, N.A. 1999. Some statistical procedures for evaluation of the relative contribution for yield components in wheat. *Zagazig Journal of Agricultural Research*, 26(2): 281-290.

Mohamed, K.A., Idris, A.E., Mohammed H.I., Adam K.A. 2012. Ranking rice (*Oryza sativa* L.) genotypes using multi-criteria decision making, correlation and path coefficient analysis. *British Biotechnology Journal,* 2(4): 211-228. https://doi.org/10.9734/BBJ/2012/1821

Nasr, S.M., Leilah, A.A. 1993. Integrated analysis of the relative contribution for some variables [leaves-roots] in sugar beet using some statistical techniques. Bulletin of Faculty of Agriculture, Cairo Univ.(Egypt).

Ogunbayo, S.A., Si&e, M., Ojo, D.K., Sanni, K.A., Akinwale, M.G., Toulou, B., Shittu, A., Idehen, E.O., Popoola, A.R., Daniel, I.O. and Gregorio, G.B. 2014. Genetic variation and heritability of yield and related traits in promising rice genotypes (*Oryza sativa* L.). *Journal of Plant Breeding and Crop Science*, 6(11): 153-159. https://doi.org/10.5897/JPBCS2014.0457.

Rahman, M. A., Islam, M. R., Islam, A. K. M. S., Munna, N. A., Razu, M. A. U. Rana, M. M. (2023). Genetic analyses of advanced breeding lines of rice (Oryza sativa L.) based on morphological traits. *Journal of Bioscience and Agriculture Research*, 30(2): 2559-2569.

R Core Team. 2021. R: A Language and environment for statistical computing (Version 4.1) [Computer software]. Retrieved from https://cran.r-project.org.

Rahman, M.A., Haque, M.E., Sikdar, B., Islam, M.A., Matin, M.N. 2013. Correlation analysis of flag leaf with yield in several rice cultivars. *Journal of Life and Earth Science*, 8: 49-54. https://doi.org/10.3329/jles.v8i0.20139.

Rashid, M., Nuruzzaman, M., Hassan, L., & Begum, S. (2017). Genetic variability analysis for various yield attributing traits in rice genotypes. *Journal of the Bangladesh Agricultural University,* 15(1): 15–19. https://doi.org/10.3329/jbau.v15i1.33525.

Saha, S. R., Hassan, L., Haque, M. A., Islam, M. M., & Rasel, M. 2019. Genetic variability, heritability, correlation and path analyses of yield components in traditional rice (Oryza sativa L.) landraces. *Journal of the Bangladesh Agricultural University*, 17(1): 26–32. https://doi.org/10.3329/jbau.v17i1.40659

Sayeed, K.A., Yunus, M.M. 2018. Rice prices and growth, and poverty reduction in Bangladesh Food and Agriculture Organization of the United Nations, Rome; 45.

Shrestha, N., Poudel, A., Sharma, S., Parajuli, A., Budhathoki, S., Shrestha, K. 2018. Correlation coefficient and path analysis of advance rice genotypes in central mid-hills of Nepal. *International Journal of Research in Agricultural Sciences,* 5(3): 2348-3997.

Steyerberg, E.W., Eijkemans, M.J., Habbema, J.D. 1999. Stepwise selection in small data sets: a simulation study of bias in logistic regression analysis. *Journal of Clinical Epidemiology*, 52(10): 935-942. https://doi.org/10.1016/S0895-4356(99)00103-1

Team, J. A. S. P. 2022. JASP (Version 0.16.3) [Computer software]. Retrieved from https://jasp-stats.org.

The Jamovi project. 2022. Jamovi (Version 2.3) [Computer Software]. Retrieved from https://www.jamovi.org.

Tripathy, S.K., Sahoo, B. 2021.Phenotyping and Association Analysis of Zinc Biofortified Rice Varieties for Grain Yield and Quality Traits. *International Journal of Plant and Environment* 7(3): 208-212. http://dx.doi.org/10.18811/ijpen.v7i03.4

USDA (United State of Department of Agriculture). 2022.World Rice Production 2022/2023. Available: http://www.worldagriculturalproduction.com/crops/rice.aspx

Walton, P.D. 1972. Factor Analysis of Yield in Spring Wheat (*Triticumaestivum* L.) 1. Crop Science, 12(6): 731-733. https://doi.org/10.2135/cropsci1972.0011183X001200060003x

Wikipedia Contributors. Exploratory factor analysis [Internet]. 2012. Wikipedia. Wikimedia Foundation; 2023 [cited 2023 Aug 5]. Available from: https://en.wikipedia.org/wiki/Exploratory_factor_analysis

Xing, Y., Zhang, Q. 2010. Genetic and molecular bases of rice yield. Annual Review of Plant Biology, 61: 421-442. https://doi.org/10.1146/annurev-arplant-042809-112209

Zhao, H., Mo, Z., Lin, Q., Pan, S., Duan, M., Tian, H., Wang, S. and Tang, X. 2020. Relationships between grain yield and agronomic traits of rice in southern China. *Chilean Journal of Agricultural Research*, 80(1): 72-79. http://dx.doi.org/10.4067/S0718-58392020000100072

Zuo, Z.W., Zhang, Z.H., Huang, D.R., Fan, Y.Y., Yu, S.B., Zhuang, J.Y. and Zhu, Y.J. 2021. Control of thousand-grain weight by OsMADS56 in rice. *International Journal of Molecular Sciences*, 23(1): 125. https://doi.org/10.3390/ijms23010125