PRINCIPLE COMPONENT ANALYSIS (PCA) OF WHEAT GENOTYPES CONCERNING MORPHOLOGICAL RESPONSES TO HEAT STRESS

Abstract. This study was conducted at Dev Bhoomi Uttarakhand University in Dehradun, India, with three replications using the Randomized Block Design (RBD) pattern to assess morphological responses to stress conditions in eight wheat genotypes (PBW-550, PBW-343, PBW-292, RAJ-3765, HD-2967, HD-3086, and NABI MG {BLACK WHEAT}.During the investigation, wheat genotypes were subjected to Principle Component Analysis (PCA). The 1st principal component accounted for 47.2689% of the total variance (PC1). The second main component (PC2) contributed 24.6007% of the total variance. The 3rd main component was responsible for 16.3402% of the overall variance (PC3). The fourth main component accounted for 8.3215 percent of the total variance (PC4). The overall ratio between the four major components to overall variance was 96.5311%. In the1st principal component, FLA had the largest positive component loading (0.294), followed by PH (0.038).There are six variables with the largest negative connection with this component: PL (-0.117), NPT (-0.319), DTH (-0.469), DTF (-0.477), DTM (-0.479), and GY (-0.346). The genotypes utilized in the study exhibit significantly different morphological responses. For breeders who additionally aspire to create variation, this is important, and it is a beneficial approach to use such genotypes as genitors in breeding studies.

Commented [oI1]: stress here what you mean ? heat stress

Keywords:, Heat stress, wheat, principal component Analysis (PCA),Scree plot, Biplot

INTRODUCTION

Wheat is a grain crop that is essential to the human diet as a source of calories and protein. Developing high-yielding cultivars which are tolerant to both abiotic and biotic stressors is essential to ensure food security. Owing to its high nutritional richness, Wheat contributes 21 percent of the total calorie intake and 20 percent of the protein consumed to the nearly 5.5 billion individuals living in approx 100 nations that are less developed countries.(Braun et al., 2010). Carbohydrates account for around 70% of the wheat crop, followed by approx three and two percent mineral substances and fatty acids respectively two per cent fiber, vitamin B complex and riboflavin as well as minerals such as Fe, Zn, Se, and Mg (Sharma, 2004).

Grain yield is the most significant biometric characteristic in each crop and is influenced by genetics, environment, and the interaction of genetics and environment (Dia *et al.*, 2016). Crop output is limited by a number of biotic and abiotic variables, the most important of which are now the abiotic stresses—heat stress in particular is becoming more prevalent owing to climate change brought on by global warming. According to a UNEP report from 2019, if greenhouse gas emissions continue to rise, by the end of the century, 3.5°C of expected temperature increase is on the upper end of the range. Wheat requires a temperature of 21±3°C for most of its crop stages. Heat stress has a large influence on production and the standard affecting around 0.036 billion hectares in the temperate zones and 7.0 million hectares in economically backward regions (Reynolds *et al.,* 2001).In cool-season cereal species, heat stress induces physiological changes via decreasing chlorophyll content, which results in leaf senescence. Excessive temperatures have a direct impact on the length and pace of grain filling (Kumar *et al.*, 2012; Lobell 2012; Gourdji *et al.,* 2013), which reduces the quantity of grains and size of the kernels (Ferris et al., 1998).Physiological screening methods for heat tolerance have included green fluorescent chlorophyll (Moffatt *et al.,* 1990) and fall in canopy temperature (Reynolds et al., 1998). The ability of a plant to flourish and provide a profit under high temperatures is known as heat tolerance.(Hall, 2001). Plant breeders have concentrated mostly on grain yield since it is a complicated quantitative characteristic that is influenced by several genes and has poor heritability because it depends so much on environmental factors. (Khairnar *et al.,* 2018). Because of this, breeders will be able to improve wheat's heat tolerance through indirect yield selection, which involves screening for connected qualities. Direct yield selection will thus be less successful. Principal component analysis are examples of multivariate analytic techniques. Evaluate diversity in genes and calculate the relative contributions of each visual and metabolic specific to the overall wheat yield (Phougat, 2022).

The creation of new variety depends critically on the existence of genetic diversity and the application of that variability. To do this, the primary principal components and the primary attributes that have the largest positive or negative loadings in each component which determine how various genotypes cluster together are analyzed. Determining the critical components and characteristics that accounted for the majority of the variance and variety within breeding lines was the main objective of the current investigation.

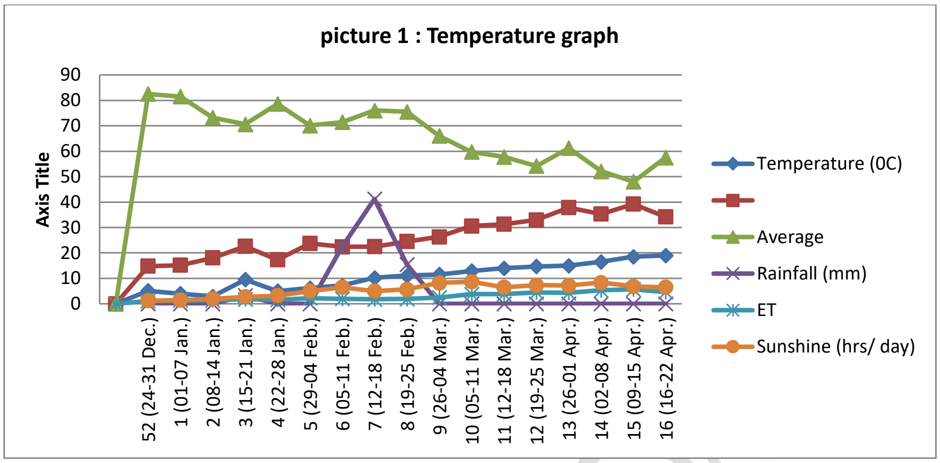
MATERIALS AND METHODS

Site Description

In Rabi 2023–24 an experiment conducted at the agricultural research farm of Dev Bhoomi Uttarakhand University, Dehradun, to establish the influence of thermal stress on the production of grains and the key causes of it. A total of eight types, PBW-550, PBW-343, PBW-292, RAJ-3765, HD-2967, HD-3086, and NABI MG (BLACK WHEAT), were sown on December 25, 2024, to maintain temperatures that are comparatively high during the reproductive period, particularly for grain filling.

Commented [oI2]: growing one year ??

Commented [oI3]: this wrong how sowing in December 25 2024



picture 1 : Temperature graph

The proposed study was conducted in Dev Bhoomi Uttarakhand University, Dehradun, Uttarakhand, India, situated at an altitude of about 640 meters (2,100 feet) above sea level, Dehradun has a subtropical environment with scorching summers and refreshing winters. Annual precipitation is around 2200 mm, with moderate to heavy rainfall falling during the monsoon season (usually July to September). Soils of many sorts, including clayey, sandy loam, and alluvial, are present in the area, and each has its own set of advantages and disadvantages when it comes to growing wheat. Laboratory space for biochemical analysis, chambers for controlled environment plant growth studies, and necessary infrastructure for field experiments are all provided by DBUU's advanced agricultural research facilities. Experimental plots, irrigation infrastructure, and data recording are all overseen by a group of agricultural experts and technicians at this location. Researchers and staff engaged in monitoring, data collecting, and other research operations find DBUU to be easily accessible by road and positioned near important transportation networks. Overall, the conducive climate, diversity of soil types, and state-of-the-art research facilities at DBUU, Dehradun make it an excellent site for wheat stress response and variety assessment studies.

Commented [oI4]: soil sample analysis put in table ?

List 1 : **Different Wheat Genotypes and their characteristics**

S.No.	Name of Genotype	Area/Institute released & Year	Charcateristics
1.	**PBW-550**	Punjab agriculture university (2017)	• A gene introgressed rust resistant version of popular variety PBW 550 • Medium duration variety suitable for mid-late planting • Average plant height: 86 cm • Days to maturity: 145 • Resistant to yellow and brown rust • Average grain yield: 23.0 q/acre
2.	**PBW-343**	Punjab agriculture university (2017)	• Unnat PBW 343 is a gene introgressed rust resistant version of mega-variety PBW 343 • First wheat variety developed through MABB (Marker Assisted Backcross Breeding) released and notified at National Level for North Western Plains Zone including Punjab • Average plant height: 100 cm • Days to maturity: 155 days • Average grain yield: 23.2 q/acre
3.	**PBW-292**		
4.	**RAJ-3765**	RAJASTHAN AGRICULTURAL RESEARCH INSTITUTE	• Raj 3765 have the tolerance to high temperature • And rusts and are suitable for normal to very late sowing conditions
5.	**HD-2967**	ICAR-IARI, New Delhi (2014)	• Plant height: 98 cm (range: 72-112) • Maturity range: Seeding to flowering: 99 Days (range: 84-108) • Seed to Seed: 143 days (range: 127-160) • Average yield- 5.456 t/ha
6.	**DBW-187**		
7.	**HD-3086**		
8.	**NABI MG (BLACK WHEAT)**		

Commented [oI5]: wher is the pedigree of wheat genotypes ?

Experiment

The varieties were planted using a Randomized Block Design (RBD) with 3- replications, and the prescribed set of measures were carried out to guarantee a healthy harvest. The mean of high and low temperatures during the reproductive stage were 30.700 and 15.120 degrees Celsius, respectively.

Measurements

Seven morphological traits were monitored, including plant height (cm), flag leaf area (cm2), and the number of productive tillers. Days to seventy five percent heading, days to seventy five percent flowering, days to seventy five percent maturity, and the amount of grain produced (per plant) were determined by selecting five plants at random from each replication line and computing the mean for different statistical features.

Principal Component Analysis, or PCA, is a statistical method that keeps crucial information while minimizing the number of variables in order to simplify complicated data sets. The conversion of the initial variables into a fresh set of independent variables known as principle components, aids in identifying of patterns and correlations within data, and the primary variables resulting the most to total variability were determined using a scree plot. All data were examined using the OPSTAT application.

Commented [oI6]: what the differnt between daye to heading and days flowering ? there are no differnt betwnn them ???

Commented [oI7]: ref?

RESULT

Principal component analysis :-

We employed PCA to compress the size of the analysed agricultural attributes dataset. The overall variation was determined using eight major component axes, Eigen values, Variability values (%), and Cumulative values (%).The first principal component was responsible for 47.2689% of the overall variance (PC 1). The second byprimary component(PC2) contributed 24.6007% of the total variance. The third main component (PC3) was responsible for 16.3402% of the overall variation. The 4th component accounted for 8.3215 percent of overall variation(PC4).The overall variance of the four major components was 96.5311%. The remaining components (PC5=1.8679%, PC6=1.5832, PC7=0.01755%,accounted for 3.4689% of the total variation. The PCA analysis yielded eight primary component axes, which contribute for all of the overall variance. The eight principal components accounted for 100% of the complete variance.

Commented [oI8]: added the PCA analysis for the wheat genotypes ?

	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Eigenvalues	3.782	1.968	1.307	1.307	0.149	0.127	0.001
% of Variance	47.2689	24.6007	16.3402	8.3215	1.8679	1.5832	0.01755
Cumulative (%)	47.2689	71.8697	88.2099	96.5313	98.3992	99.9825	100

Table-1 Eigen values, Variability and Cumulative Values

Scree Plot (Graphical representation of Eigenvalues)

Figure **1** depicted a scree plot (Graphical visualization of Eigen values). Eigen values for PC1 were 3.782 and 1.968 (PC2), 1.307 (PC3), 1.307 (PC4), 0.149 (PC5) and 0.127 (PC6), 0.001 (PC7), and zero (PC8), respectively. If the Eigen values exceed one, It indicates that the assessed main component scores are reliable (Mohammadi 2003). On the other hand, Iezzoni (1991) found that PCs with Eigen values larger than 1 are more informative than the original variable. Thus PC1 , PC2 , PC3 and PC4 has explained the major portion of genetic variation caused due to heat stress. Through this we can therefore categorise the traits that are strongly impacted by heat stress and traits which are weakely impacted by heat stress.

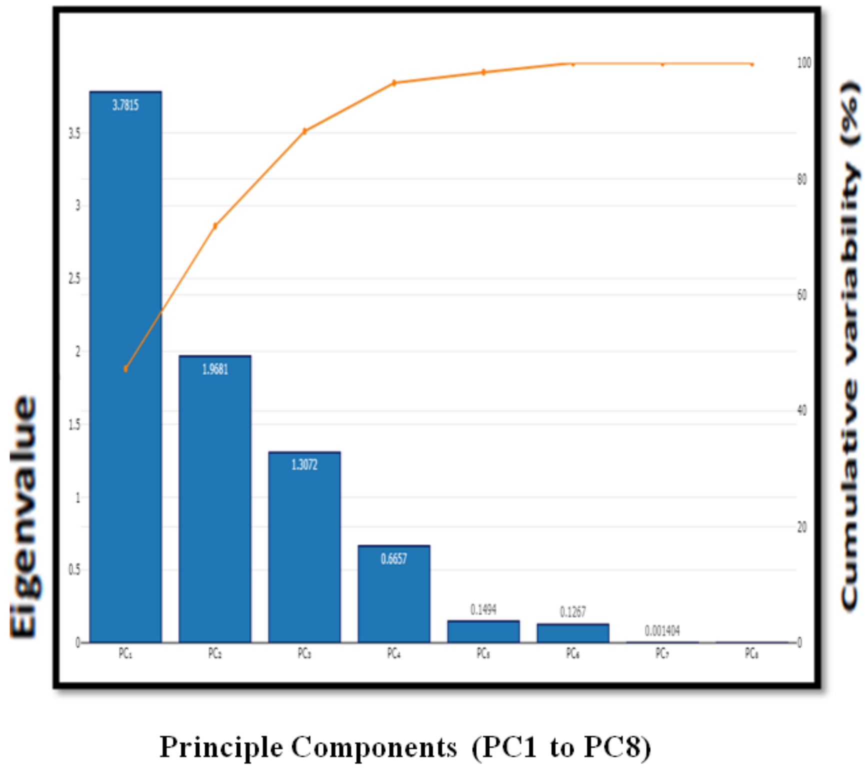


Figure-1. Graphical visualisation of Eigen values

Biplot

When viewing this biplot (Fig. 2), the narrow angle features reveal a positive relationship. Right-angle characteristics are unrelated to each other. Wide-angle characteristics have unfavourable interactions with one another. The biplot approach allows for the identification of correlations between parameters as well as a detailed examination of a heterogeneous data set. (Yan,2002).

Biplot (axes PC1 and PC2: 71.87 %)

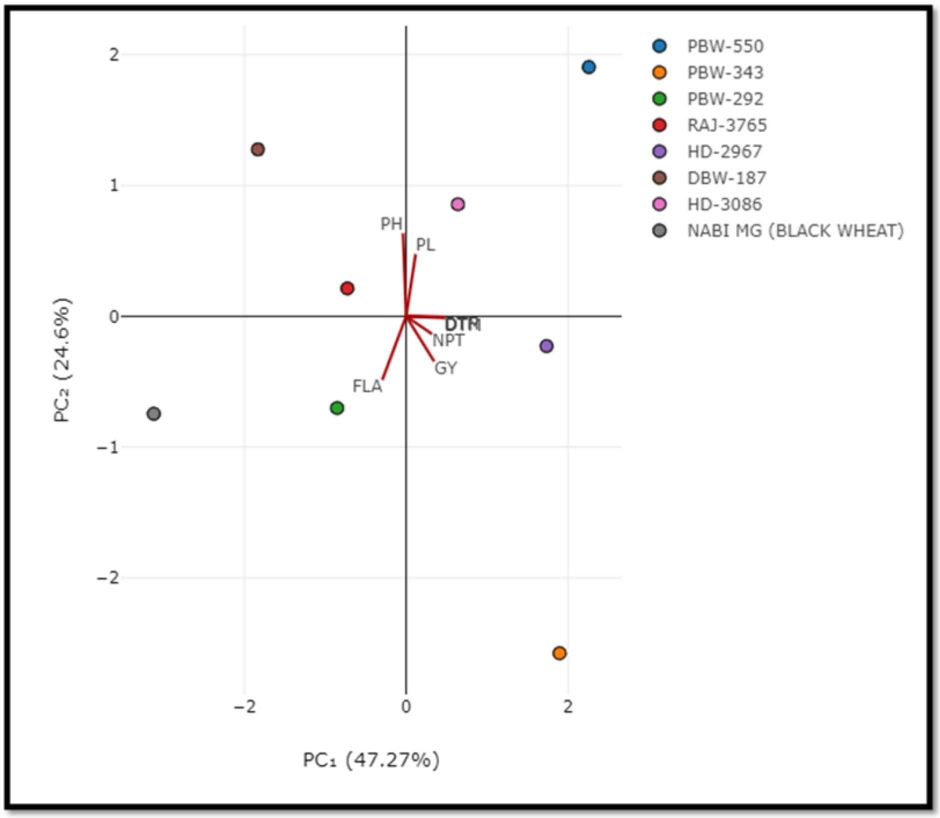


Figure.2. Biplot of the morphological responses to heat stress of 8-wheat varieties for the first two principle components.

Commented [oI9]: biplot for six traits ?

Table: 2 Factor loading of morphological character in relation to the major factor in wheat.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
PH	0.038	-0.634	0.348	0.023	0.327	0.478	0.337	-0.161
FLA	0.294	0.486	0.053	-0.500	0.505	-0.058	0.372	-0.168
PL	-0.117	-0.474	-0.580	-0.124	0.441	-0.463	-0.039	0.010
NPT	-0.319	0.135	-0.637	-0.102	-0.215	0.531	0.337	-0.149
DTH	-0.469	0.006	0.187	-0.426	0.073	0.097	-0.548	-0.498
DTF	-0.477	0.006	0.191	-0.367	0.085	0.063	0.088	0.763
DTM	-0.479	0.013	0.251	0.124	-0.207	-0.490	0.562	-0.304
GY	-0.346	0.345	0.000	0.626	0.585	0.131	-0.102	0.018

A eigenvector with reversed signs for all its components is still considered a valid solution!

PH: Plant height, FLA: Flag leaf area, PL: Peduncle length, NPT: No. of productive Tiller, DTH: Days to 75%Heading, DTF: Days to 75%flowering, DTM:Days to maturity, GY:Grain yield

Commented [oI10]: in M&M you studied r seven traits in table 8 traits ? how

DISCUSSION

The first main component had the largest positive component loading from FLA (0.294), followed by PH (0.038). There are six variables having PL (-0.117), NPT (-0.319), DTH (-0.469), DTF (-0.477), DTM (-0.479) and GY (-0.346) had the strongest negative connection. In the second primary component, FLA (0.486) explained the biggest positive contribution followed by GY (0.345) and NPT (0.135). Characters such as PH (-0.634) and PL (-0.474) showed the most negative relationship with this component. In 2^{nd}component, plant height and peduncle length were shown to contribute the most variability. In the third principal component PH (0.348), DTM

(0.251), DTF (0.191) had the highest coefficients, respectively and the NPT (-0.637) and PL (-0.580) having highest negative coefficients. For the fourth principal component GY (0.626) having highest coefficient and FLA (-0.500) having highest negative coefficients. For the fifth principal component GY (0.585) followed by FLA (0.505) having highest coefficient and NPT (-0.215) having highest negative coefficients. For the sixth principal component NPT (0.585) followed by PH (0.478) having highest coefficient and DTM (-0.490) having highest negative coefficients. For the seventh principal component DTM (0.562) and DTH (-0.548) having highest positive and negative values, correspondingly.

The Principal Component Analysis (PCA) conducted on our dataset revealed several key insights into the structure and relationships among the variables:

1. **First Principal Component (PC1)**:
   - **Positive Loadings**: The first principal component had the largest positive loading from Flag Leaf Area (FLA) (0.294), followed by Plant Height (PH) (0.038).
   - **Negative Loadings**: Six variables showed strong negative loadings: Peduncle Length (PL) (-0.117), Number of Productive Tillers (NPT) (-0.319), Days to Heading (DTH) (-0.469), Days to Flowering (DTF) (-0.477), Days to Maturity (DTM) (-0.479), and Grain Yield (GY) (-0.346).
   - **Interpretation**: PC1 captures a contrast between FLA and the other six variables, indicating a potential trade-off or inverse relationship between these groups of traits.
2. **Second Principal Component (PC2)**:
   - **Positive Loadings**: FLA (0.486) had the highest positive contribution, followed by GY (0.345) and NPT (0.135).
   - **Negative Loadings**: PH (-0.634) and PL (-0.474) showed the most negative relationships with PC2.
   - **Interpretation**: This component highlights the variability contributed primarily by plant height and peduncle length, with FLA and GY being the key positive contributors, suggesting distinct patterns of growth and productivity.
3. **Third Principal Component (PC3)**:
   - **Positive Loadings**: PH (0.348), DTM (0.251), and DTF (0.191) had the highest positive coefficients.
   - **Negative Loadings**: NPT (-0.637) and PL (-0.580) had the highest negative coefficients.

- **Interpretation**: PC3 differentiates between traits related to reproductive timing (DTM and DTF) and structural attributes (NPT and PL).

4. **Fourth Principal Component (PC4)**:
   - **Positive Loadings**: GY (0.626) had the highest positive coefficient.
   - **Negative Loadings**: FLA (-0.500) had the highest negative coefficient.
   - **Interpretation**: This component contrasts grain yield with flag leaf area, indicating a potential inverse relationship between these two important agronomic traits.
5. **Fifth Principal Component (PC5)**:
   - **Positive Loadings**: GY (0.585) and FLA (0.505) had the highest positive coefficients.
   - **Negative Loadings**: NPT (-0.215) had the highest negative coefficient.
   - **Interpretation**: PC5 emphasizes the positive association between grain yield and flag leaf area, contrasting with the number of productive tillers.
6. **Sixth Principal Component (PC6)**:
   - **Positive Loadings**: NPT (0.585) and PH (0.478) had the highest positive coefficients.
   - **Negative Loadings**: DTM (-0.490) had the highest negative coefficient.
   - **Interpretation**: This component shows a strong relationship between the number of productive tillers and plant height, against days to maturity.
7. **Seventh Principal Component (PC7)**:
   - **Positive Loadings**: DTM (0.562) had the highest positive coefficient.
   - **Negative Loadings**: DTH (-0.548) had the highest negative coefficient.
   - **Interpretation**: PC7 highlights the contrast between days to maturity and days to heading, which may reflect different aspects of plant developmental timing.

Correlation Coefficient Analysis

The correlation coefficient analysis (Pearson, 1900) revealed positive-significant associations between DTH and DTF, DTH and DTM, and DTF and DTM). This circumstance demonstrates that there is a close link between heading, blooming, and maturity.

Below mentioned result of correlation coefficient Analysis revealed that,GY is negatively correlated with PH, FLA and PL. As in heat stress condition if plant height, flag leaf area and peduncle length is more, due to more vegetative growth plant unable to complete their need at later stages and decline in grain yield. Here our result also justifies the same that PH, FLA and PL are negatively correlated with grain yield. On other hand, GY is positively correlated with NPT, DTH, DTF and DTM in the heat stress conditions. According to this conclusion comes that if the variety taking more time for heading, flowering and maturity it means that variety are stable for heat stress condition as most of the variety which are late sown takes more time for these above-mentioned traits, as they are late sown variety. In our research trial we also selected 8 different genotype and sown on the 25th December 2023, for finding the suitable variety which can easily tolerate heat stress and there is no effect on grain yield.

Table-3. Correlation matrix (Pearson (n)

	PH	FLA	PL	NPT	DTH	DTF	DTM	GY
PH	1							
FLA	-0.526	1						
PL	0.303	-0.546	1					
NPT	-0.483	-0.256	0.461	1				
DTH	0.013	-0.356	0.094	0.443	1			
DTF	0.013	-0.383	0.093	0.445	0.998	1		
DTM	-0.008	-0.557	0.014	0.338	0.867	0.891	1	
GY	-0.433	-0.220	-0.190	**0.456**	**0.448**	**0.484**	**0.661**	1

Values in bold are different from 0 with a significance level alpha=0.05.

PH: Plant height, FLA: Flag leaf area, PL: Peduncle length, NPT: No. of productive Tiller, DTH: Days to 75%Heading, DTF: Days to 75%flowering, DTM:Days to maturity, GY:Grain yield

CONCLUSION

The correlation coefficient analysis (Pearson, 1900) highlighted significant positive associations between days to heading (DTH), days to flowering (DTF), and days to maturity (DTM), indicating a strong interrelationship among these phenological traits. This close linkage suggests that the stages of heading, blooming, and maturity are synchronized, which is crucial for understanding plant development under varying environmental conditions. Additionally, the analysis revealed that grain yield (GY) is negatively correlated with plant height (PH), flag leaf area (FLA), and peduncle length (PL). This negative correlation implies that under heat stress conditions, increased vegetative growth—evidenced by taller plants, larger flag leaves, and longer peduncles—impedes the plant's ability to sustain reproductive growth, leading to a decline in grain yield. Conversely, GY showed positive correlations with the number of productive tillers (NPT), DTH, DTF, and DTM under heat stress. This finding suggests that varieties with longer periods for heading, flowering, and maturity are more stable under heat stress, particularly those sown later in the season. In our research trial, we evaluated eight different genotypes sown on December 25, 2023, to identify varieties that can effectively tolerate heat stress without compromising grain yield. The Principal Component Analysis (PCA) further refined our understanding by reducing the dimensionality of the dataset and identifying key components that account for the majority of the variation. The PCA results showed that the first principal component (PC1) explained 47.2689% of the total variance, primarily influenced by flag leaf area (FLA) and plant height (PH). The second principal component (PC2) accounted for 24.6007% of the variance, with FLA and GY contributing positively, and PH and PL contributing negatively. The first four principal components together explained 96.5311% of the total variance, underscoring their significance in capturing the primary sources of variation in the dataset. Overall, the PCA identified eight principal components that together accounted for 100% of the total variance, highlighting the complex interplay of traits affecting grain yield under heat stress. Notably, FLA, PH, and GY emerged as critical traits influencing the primary components. These findings provide valuable insights for breeding programs aimed at improving heat stress tolerance in crops. In conclusion, this study underscores the importance of understanding the relationships among key agronomic traits under heat stress conditions. By identifying varieties that exhibit stable performance through prolonged phenological stages and optimal trait combinations, we can enhance grain yield resilience in the face of increasing heat stress challenges. Future research should focus on exploring the genetic basis of these traits to develop heat-tolerant crop varieties with improved yield potential.

REFERENCES

Bhatti, N., Dalal, M. S., Phougat, D., & Devi, S. (2022). Multivariate analysis of morpho-physiological traits in bread wheat genotypes under terminal heat stress. *Electronic Journal of Plant Breeding*, *13*(4), 1180-1186.

Braun, H.J., Atlin, G. and Payne, T. 2010. Multi-location testing as a tool to identify plant response to global climate change. Climate change and crop production, 1:115-138.

Dia, M., Wehner, T.C. and Arellano, C. 2016. Analysis of genotype × environment interaction (G×E) using SAS programming.Agronomy Journal, 108(5):1838- 1852.

Gourdji, S.M., Mathews, K.L., Reynolds, M., Crossa, J. and Lobell, D.B. 2013. An assessment of wheat yield sensitivity and breeding gains in hot environments. Proceedings of the Royal Society B: Biological Sciences, 280(1752): 20122190.

Iezzoni, A. F., Pritts, M. P. (1991): Applications Of Principal Component Analysis To Horticultural Research. – HortScience 26(4): 334-338.

Kumar, S., Singh, R.G.M.S., Grover, M. and Singh, A.K. 2012. Terminal heat an emerging problem for wheat production. Biotechnol today, 2(2):7-9.

Lobell, D.B. and Gourdji, S.M. 2012. The influence of climate change on global crop productivity. Plant physiology, 160(4):1686-1697

Moffatt, J.M., Sears, R.G., Cox, T.S. and Paulsen, G.M. 1990. Wheat high temperature tolerance during reproductive growth. II. Genetic analysis of chlorophyll fluorescence. Crop science, 30(4):886- 889.

Mohammadi, S. A., Prasanna, B. M. (2003): Analysis of Genetic Diversity in Crop Plants Salient Statistical Tools and Considerations. – Crop Sci. 43(4): 1235-1248

Phougat, D. and Verma, A. 2022. Diversity in bread wheat genotypes (*Triticum aestivum* L.) deciphered by multivariate biplot analysis. Biological Forum – An International Journal, 14(1): 1610-1616.

Reynolds, M.P., Nagarajan, S., Razzaque, M.A. and Ageeb, O.A.A.2001. Heat tolerance. Application of physiology in wheat breeding, 124-135.

Reynolds, M.P., Singh, R.P., Ibrahim, A., Ageeb, O.A.A., Larque-Saavedra, A. and Quick, J.S. 1998. Evaluating physiological traits to complement empirical selection for wheat in, warm environments. Euphytica, 100(1):85-94.

Sharma, S.K. and Jain, N.K. 2004. Nutrient management in wheat-based cropping system in sub-humid Southern Zone of Rajasthan. Indian Journal of Agronomy, 59(1):26-33.

UNEP, 2019. Emissions Gap Report 2019. Executive Summary, United Nations Environment Programme, Nairobi.

Wang, X., Cai, J., Liu, F., Dai, T., Cao, W., Wollenweber, B. and Jiang, D. 2014. Multiple heat priming enhances thermo-tolerance to a later high temperature stress via improving sub cellular antioxidant activities in wheat seedlings. Plant Physiology and Biochemistry, 74: 185-192.

Yan, W., Rajcan, I. (2002): Biplot analysis of test sites and trait relations of soybean in Ontario. – Crop Science 42: 11-20