

Asian Journal of Research in Agriculture and Forestry

5(2): 29-45, 2020; Article no.AJRAF.53958 ISSN: 2581-7418

Sensitivity of Genotype by Environment Interaction Models to Outlying Observations

S. Oluwafemi Oyamakin^{1*} and M. Olalekan Durojaiye¹

1 Biometry Unit, Department of Statistics, Faculty of Science, University of Ibadan, Nigeria.

Authors' contributions

This work was carried out in collaboration between both authors. Author SOO designed the study, performed the statistical analysis, wrote the protocol and wrote the first draft of the manuscript. Author MOD performed the statistical analysis, managed the analyses of the study and the literature searches. Both authors read and approved the final manuscript.

Article Information

DOI: 10.9734/AJRAF/2020/v5i230081 *Editor(s):* (1) Dr. Lucia Bortolini, University of Padova, Italy. *Reviewers:* (1) Joseph Adjebeng-Danquah, CSIR-Savanna Agricultural Research Institute, Ghana. (2) Jehad Mahmoud Hussein Ighbareyeh, Al-Quds Open University, Palestine. Complete Peer review History: http://www.sdiarticle4.com/review-history/53958

Review Article

Received 01 December 2019 Accepted 05 February 2020 Published 04 March 2020

ABSTRACT

___ Plant breeding program depends on its ability to provide farmers with genotypes with guaranteed superior performance (phenotype) in terms of yield and/or quality across a range of environmental conditions. To achieve this aim, it is necessary to have an understanding of the model suitable for or leading to a good phenotype. In this study, two cases of scenarios were considered to have a clearer view of the performance of Genotype by Environment Interaction on the following four models; Additive Main Effect and Multiplicative Interaction (AMMI), Finlay Wilkinson (FW), Genotype and Genotype by Environment Interaction (GGE) and Mixed Model. We experiment the inference behind the violation of the assumption of normal distribution by observing the data contamination of two case scenarios (Lowest and Highest outlying observations). It was observed on the two data Types of Balance and Unbalance designs with different Levels of generations. We achieved that by comparative performance of the data contamination techniques under the two case scenarios; Case I scenario was done for Lowest Outlying Observations where 50%, 100% and 500% data contamination on the First Quarter (P1), Mid quarter (P2) and Last Quarter (P3). We then deduced from the result of the model evaluation that, at each levels of data contamination for Balance and Unbalance design, Mixed model was the ideal model for $G \times E$ interaction. Case II scenario was done for Highest Outlying Observations where 50%, 100% and 500% data

contamination on the First Quarter (P1), Mid quarter (P2) and Last Quarter (P3) were examined on each levels of generations. We then observed from the result of the model evaluation that, at each levels of data contamination for Balance and Unbalance design, Mixed model also outperformed the other three models.

Keywords: Plant breeding; genotype-by-environment interaction; AMMI; FW; GGE and mixed model; simulation.

1. INTRODUCTION

Despite the enormous work done by plant breeders in developing high yielding crop varieties the changes in climate may limit the gains as this is tantamount to a shifting target. One of the remedies may be to breed new and adapted as well as the preservation of traditional locally adapted varieties that can tolerate climate variability and are suitable for changed climatic conditions [1]. Legumes in particular have many advantages. In addition to being a rich source of soil proteins, they enrich soil through biological nitrogen fixation. Studies have shown that the rotation of chickpea and pigeon pea reduces the use of chemical fertilizers and also enhances the output of paddy and wheat significantly [2].

1.1 Phenotype, Genotype and Environment

Genotype is the part (DNA sequencing) of genetic make-up of a cell and therefore of an organism or individual, which determines a specific characteristic (i.e. Phenotypes) of that organism or individual. Organism frequently encounter different environmental conditions because the physiological and behavioral responses of these conditions depend on the genetic make-up of individuals [3].

Genotype generally remains constant from one environment to another, although occasional spontaneous mutations may occur which caused it to change. However, when the same genotype is subjected to different environments, it can produce a wide range of phenotypes.

1.2 Genotype-By-Environment Interaction $(G \times E)$

In 1902, Garrod was one of the first scientists to note that the effect of genes on phenotype could be modified by the environment. Tumuhimbise et al. [4] also demonstrated that the development of a plant is often influenced by its surrounding. However, many years ago, the prediction of phenotypes from high-density marker information was recognized as a potential game changer in animal breeding [5].

In plant breeding, the current development of high throughput genotyping techniques alongside with similar techniques for phenotyping and envirotyping [6,7] provides opportunities for large scale phenotypic predictions of new, and therefore, untested genotypes in new environments under a wide spectrum of genotype-by-environment interaction $(G \times E)$ scenarios [8,9,10,11,12]. The increased volumes of phenotypic, genotypic and environmental data give a stimulus to the development of new statistical approaches for more precise description and prediction of $G \times E$ phenomena. A better modeling of $G \times E$ will undeniably contribute to a higher efficiency of breeding programs. In the light of the current developments, it is obvious that the study of $G \times E$ will become even more important in the near future than it was already in the past.

Future generations of students and researchers in plant breeding will require substantial training in the statistical aspects of $G \times E$. Environmental factors such as rainfall amount, soil structure, temperature and altitude may impact genotypes positively or negatively.

When genotypes performance rank differently in different environments this is what is known as genotype by environmental interaction [13]. It is the association between the environment and the phenotypic expression of a genotype that is known as $G \times E$ interaction. This interaction that determines whether a genotype is adapted to a whole range of environmental conditions or separate genotypes is manifested through growing the genotypes in different subenvironments.

Nigeria with its very diverse climatic conditions and soil types escalates the problem of $G \times E$ even further. To overcome this problem, the universal practice of scientists in most crops when selecting genotypes, is to plant them in yield (performance) trials over several environments and years to ensure that the selected genotypes have a high and stable performance over a wide range of environments. The assessment of genotype performance in genotype by environment and by year experiments is often difficult because of the presence of environment by year interaction (i.e. environmental effects).

The general aim of this study is to determine which of these models' best suits $G \times E$ interaction using Monte Carlo simulated data and also discuss certain statistical and biological limitations.

The specific objectives are:

- i. To compare the various statistical methods of analysis to determine the most suitable parametric procedure to evaluate and describe genotype performance under multi-location trials,
- ii. To determine the efficiency of each method (AMMI, Finlay-Wilkinson, GGE and Mixed model) in detecting $G \times E$ and
- iii. Also, to determine the adaptability and specificities of the methods.

2. MATERIALS AND METHODS

A combined analysis of variance procedure is the most common method used to identify the existence of $G \times E$ from replicated multi-location trials. If the $G \times E$ variance is found to be significant, one or more of the various methods for measuring the stability of genotypes can be used to identify the stable genotype(s). A wide range of methods is available for the analysis of $G \times E$ and can be broadly classified into four groups: the analysis of components of variance, stability analysis, multivariate methods and qualitative methods.

The methods to be adopted in this study are suitable for the plant breeders in estimating Genotype by Environment Interaction $(G \times E)$ parameters. The methods are as follows;

A. Additive Main Effect and Multiplicative Interaction (AMMI) Model

The AMMI model combines the features of ANOVA and SVD as follows: first, the ANOVA estimates the additive main effects of the twoway data table; then the SVD is applied to the residuals from the additive ANOVA model, estimating $N \le \min(I-1, J-1)$ interaction principal components (IPCs). The model can be written as [16]:

$$
y_{ijk} = \mu + \alpha_i + \beta_j + \sum_{n=1}^{N} \lambda_n \gamma_{n,i} \delta_{n,j} + \rho_{i,j} + e_{ijk}
$$
 (1)

where y_{ijk} is the phenotypic trait (yield or some other quantitative trait of interest) of the ith genotype in the jth environment for replicate k; model

 μ is the grand mean;

 α_i are the genotype deviations from μ ;

 β are the environment deviations from μ ;

 λ_n is the singular value of the IPC analysis axis n;

 $\gamma_{n,i}$ and $\delta_{n,i}$ are the ith and jth genotype and environment IPC scores (i.e. the left and right singular vectors, scaled as unit vectors) for axis n, respectively;

 $\rho_{i,j}$ is the residual containing all multiplicative terms not included in the model;

 e_{ijk} is the experimental error; and N is the number of principal components retained in the model.

In matrix formulation the AMMI model can be written as:

$$
Y = \mathbf{1}_I \mathbf{1}_j^T \mu + \alpha_I \mathbf{1}_J^T + \mathbf{1}_I \beta_J^T + UDV^T + \varepsilon
$$
 (2)

where Y is the $(I \times J)$ two-way table of genotypic means across environments. The interaction part of the model $Y^* = Y - I_i^T \mu - \alpha_i I_i^T - I_i \beta_i^T$ is approximated by the product of matrices UDV^T , with U an $(I \times N)$ matrix whose columns contain the left singular vectors interactions of n, D a $(N \times N)$ diagonal matrix containing the singular values of Y^* , and V a $(J \times N)$ matrix whose columns contain the right singular vectors of Y^*

B. Finlay-Wilkinson Model

A more attractive alternative is to extend the additive model: $y_{ii} = \mu + \alpha_i + \beta_i + e_{ii}$ (3)

by incorporating terms that explain as much as possible of the $G \times E$. A popular strategy in plant breeding is that proposed by [14], which describes $G \times E$ as a regression line on the environmental quality. In the absence of explicit environmental information, the biological quality of an environment can be reflected in the average performance of all genotypes in that environment. The $G \times E$ part is then described by genotype-specific regression slopes on the environmental quality, and the model can be written in the following equivalent ways:

$$
y_{ij} = \mu + \alpha_i + \beta_j + b_i \beta_j + e_{ij}
$$
 (4)

$$
y_{ij} = \alpha_i + b_i' \beta_j + e_{ij}
$$
 (5)

Model (5) follows from model (4) by taking $\mu + \alpha_i = \alpha_i^{\dagger}$ and $\beta_j + b_i \beta_j = (1 + b_i) \beta_j = b_i^{\dagger} \beta_j$. Model (5) is easier to interpret because it looks as a set of regression lines; each genotype has a linear reaction norm with intercept α , and slope

 $b_{i}^{'}$. The explanatory environmental variable in these reaction norms is simply the environmental main effect β , Model (4) shows more clearly how $G \times E$ is captured by a regression on the environmental main effect, with the hope that as much as possible of the GEI signal will be retained by the term $b_i \beta_i$. Note that in model (5)

the average value of b is 1, meaning that $b > 1$ for genotypes with a higher than average sensitivity, and $b' < 1$ for genotypes that are less sensitive than average.

C. Genotype and Genotype by Environment Interaction (GGE) Model

Plant breeders are interested in the total genetic variation and not exclusively in the $G \times E$ part. For that reason, it is useful to have a modification of model (1) that considers the joint effects of the genotypic main effect and the $G \times E$ as a sum of interpretation procedures hold as for model (1). Because genotypic scores now describe genotypic main effects G and GE together, this type of model is also known as the "GGE model" and the Biplots are called "GGE Biplots" [15]. The model reads:

$$
y_{ij} = \mu + \beta_j + \sum_{n=1}^{N} \lambda_n \gamma_{n,i} \delta_{n,j} + \rho_{i,j} + e_{ij}
$$
 (6)

In GGE, the result of SVD is often presented in a "Biplot illustration". Its approximate overall performance (G + GE).

D. Mixed Model

The REML/BLUP method allows the consideration of different structures of variance and covariance for the genotypes by environments effects, which makes the model more realistic. For the $G \times E$ evaluation by mixed model, the following statistical model was used:

$$
y = Z\alpha + X\beta + W\eta + \varepsilon \tag{7}
$$

Where, y is the vector of observed data; α is the vector of genotype effects (assumed as random); β is the vector of block effects within each environment (assumed as fixed); η is the vector of GEI effect (assumed as random); and ε is the error vector (random). The uppercase letters represent the matrices of incidence for the referred effects. See below the distribution of the random effects:

$$
\alpha \mid \sigma_{\alpha}^{2} \sim N(0, I\sigma_{\alpha}^{2}), \eta \mid \sigma_{\eta}^{2} \sim N(0, I\sigma_{\eta}^{2})
$$

and
$$
\varepsilon \mid \sigma_{\varepsilon}^{2} \sim N(0, I\sigma_{\varepsilon}^{2})
$$

E. Simulation study

We simulate two-way data tables for balanced and unbalanced design with 3 replications each, where the interaction is explained by two multiplicative terms (i.e. two IPCs; $k = 2$ components to be retained). Without loss of generality, the two-way data tables are simulated in the following way:

1. Balance and Unbalance Design

Create a matrix X with $(n \times p)$ data design;

- (3×3) data design, where n = 3 rows (Genotypes) and p = 3 columns (Environments)
- (7×7) data design, where n = 7 rows (Genotypes) and p = 7 columns (Environments).
- (10×10) data design, where n = 10 rows (Genotypes) and p = 10 columns (Environments).
- (3×7) data design, where n = 3 rows (Genotypes) and p = 7 columns (Environments)
- (7×3) data design, where n = 7 rows (Genotypes) and p = 3 columns (Environments).
- (7×10) data design, where n = 7 rows (Genotypes) and p = 10 columns (Environments).
- (10×7) data design, where n = 10 rows (Genotypes) and p = 7 columns (Environments).
- with observations drawn from a Unif^[0, 0.5] distribution. Do the SVD of X and obtain the matrices U, V and D, containing, respectively, the left and right singular vectors and the singular values of X;
- Simulate the grand mean, the genotypic
and environmental main effects. and environmental main effects, considering: $\mu \sim N(15, 3)$, $\alpha \sim N(5, 1)$ and $\beta \sim N(8, 2)$ *(Rodrigues et al.(2015)).*

2. Data Contamination Experiment

The procedure is to experiment the inference behind the violation of the assumption of normal distribution by observing the data contamination of two case scenarios (Lowest and Highest outlying observations). And it would be observed on the two data Types of Balance and Unbalance designs as generated above.

Case I scenario (Lowest Outlying Observation):

The data type of balance and Unbalance design for each of the levels considered in the study $(3 \times 3, 7 \times 7, 10 \times 10$ and $3 \times 7, 7 \times 3, 7 \times 10$,

(i.) Type: Structure of the Design

$$
10 \times 7
$$
) would be contaminated at 50%, 100% and 500% on each data point of the first (P_1) , mid (P_2) and last (P_3) quarters of the data design. At every point of contamination, we would subtract the value from the original to lowering the outliers of each of the quarters observed in the process and take the inference. We would then proceed to the next step by returning the original quarter's data point and move to the next quarter for contamination. The rigorous procedure would continue till we exhausted all the data levels.

Case II scenario (Highest Outlying Observation):

The data type of balance and Unbalance design for each of the levels considered in the study (3×3 , 7×7 , 10×10 and 3×7 , 7×3 , 7×10 , 10×7) would be contaminated at 50%, 100% and 500% on each data point of the first (P_1) , mid (P_2) and last (P_3) quarters of the data design. At every point of contamination, we would add the value to the original data point to increase the outliers of each of the quarters observed in the process and take the inference. We would then proceed to the next step by returning the original quarter's data point and move to the next quarter for the same process of contamination. The rigorous procedure would continue till we exhausted all the data levels in the study. The cases structure summarized below;

$$
Design \begin{cases} Type\ I:\ Balance \\ Type\ II:\ Unbalance \end{cases}
$$

(ii.) Level: Dimension of the Design

3. Assessment of Genetic Stability

Various techniques are available for evaluation of adaptation and genetic stability of genotypes in different environments. These include regression [14], AMMI [16], GGE-biplot [15] and Mixed model.

Stability evaluation using Finlay-Wilkinson model

The regression of each genotype in an experiment on an environmental index and a function of the squared deviations from this regression would provide estimates of the desired stability parameters. Parameters are defined with the following model [14].

$$
\hat{y}_{ij} = \hat{\alpha}_i + \hat{b}_i \beta_j
$$

where \hat{y}_i is the phenotypic trait of the *i*th genotype in the *j*th environment, each genotype has a linear reaction norm with intercept $\hat{\alpha}_i$ and slope $\hat{b_i}$ and β_j are the environment deviations from μ

a. Computation of Environmental Index (β ,) **was done as follows:**

The environmental index β_i obtained as the mean of all yields at the *j*th environment minus the grand mean.

$$
\beta_j = \left(\frac{\sum_i Y_{ij}}{n_i}\right) - \left(\frac{\sum_i \sum_j Y_{ij}}{n_i}\right)
$$

Where $\sum_i \beta_i = 0$

b. Computation of Regression Coefficient (*ⁱ b* **)**

The first stability parameter is a regression coefficient estimated as follows:

$$
b_i = \frac{\sum_j Y_{ij} \beta_j}{\sum_j \beta_j^2}
$$

where, $\sum_i Y_i \beta_i$ is the sum of products of the environmental index with the corresponding mean of that genotype at each location and $\sum_i \beta_i^2$ is the sum of squares.

c. Computation of Mean Square deviation $(S^2_{d_i})$

$$
S_{d_i}^2 = \frac{\sum_j \delta_{ij}^2}{n_j - 2} - S_e^2 / r
$$

Where $\sum_j \delta_{ij}^2$ is pooled deviation, n_j is the number of environments and S_e^2/r is the mean sum of squares from ANOVA.

Stability evaluation using AMMI model

AMMI stability value (ASV) was calculated for each genotype according to the relative contributions of the principal component axis scores (IPCA1 and IPCA2) to the interaction sum of squares. Stability analysis was then done placing the seven varieties as genotypes in the Additive Main Effects and Multiplicative Interaction effects of principal components analysis (PCA) model. Genotype means and scores were generated and used for ranking genotype stability. The AMMI model does not make provision for a quantitative stability measure, such a measure is essential in order to quantify and rank genotypes according their yield stability, The AMMI stability value (ASV) as described by [17] was calculated as follows:

$$
ASV
$$

= $\sqrt{\left[\frac{[PCA1SumofSquares]}{[PCA2SumofSquares]}(IPCA1SCORE)\right]^2 + [IPCA1Score]^2}$

In effect the ASV is the distance from zero in a two-dimensional scatter gram of IPCA1 (Interaction Principal Component Analysis axis 1) scores against IPCA2 scores. Since the IPCA1 score contributes more to G×E sum of squares, it has to be weighted by the proportional difference between IPCA1 and IPCA2 scores to compensate for the relative contribution of IPCA1 and IPCA2 total G×E sum of squares. The distance from zero is then determined by using the theorem of Pythagoras.

AMMI Stability Value (ASV) refers to a statistical measure to determine the Stability and Adaptability of the various genotype. The larger ASV indicates that a genotype is adapted specifically to a certain environment while a smaller ASV indicates that a genotype is more stable across environment [18]. Yield stability index was also calculated using the sum of the ranking based on yield and ranking based on the AMMI stability value.

YSI = RASV + RY

where

RASV is the rank of the genotypes based on the AMMI stability value;

RY is the rank of the genotypes based on yield across environments (RY).

YSI incorporates both mean yield and stability in a single criterion. Low values of both parameters show desirable genotypes with high mean yield and stability [19]

Stability evaluation using GGE biplots

Comparison biplots, scatter biplot, and ranking biplots would be generated for genotype ranking.

Stability evaluation using Mixed Model

Environmental variance for each genotype would be used to determine the stability evaluation of mixed model as outlined in [20], following his published SAS code. Each unique combination of environment and design level would be considered.

Furthermore, again for computational feasibility, the genotype means dataset would be split into two subsets and then analyzed. The SAS PROC MIXED code [20] specifies an unstructured variance-covariance R matrix, which provides a unique environmental variance estimate for each genotype [21]. This allowed subsets of the data to be processed without influencing environmental variance estimates because of the presence or absence of certain genotypes in a specific data subset. All environmental variance estimates would be inspected to determine if they appeared to be reasonable as compared to a median value for the dataset.

3. RESULTS AND DISCUSSION

The analysis an interpretation of the simulated data was presented. We intentionally concentrated on the interface of the four models considered in the methodology, basically for checking the robustness of the models.

Understanding the implication of Genotype-by-Environment (G×E) interaction structure is an important consideration in plant breeding programs. Traditional statistical analyses of yield trials provide little or no insight into the particular pattern or structure of the G×E interaction. In this study, efforts were made to solve these problems under different level of data occurrence. We employed the simulation process of Monte Carlo in generating the data, since adoption of real-life data might pose a serious difficulty.

Fig. 1. Case 1 - Rank Performance

Fig. 2. Case 2 - Rank performance

3.1 Data Contamination

In this phase, we were able to justified our inference on the simulated data using data contamination approach and model evaluation criterion to check the robustness of the outlying observations. We therefore examined different cases of data contamination.

- **Case 1 - Lowest Outlying Observation:** Case 1 scenario revealed the lowest outlying observations of the type balance and unbalance design with (50%, 100% and 500%) different levels of contamination. Table 1 for Type I explained that among all the four models, mixed model is more sensitive to outlying observations due to the minimum values presented by the Root Mean Square, Mean Square Error and Absolute Bias. Similarly, Table 2 for Type II also reported the robustness of Mixed model.
- **Case 2 - Highest Outlying Observation:** Case 2 scenario revealed the highest outlying observations of the type balance and unbalance design with (50%, 100% and 500%) different levels of contamination. Table 3 and Table 4 shown the model evaluation for the Type I and II at different levels of contamination and the mixed model still prove to be more influential.

4. SUMMARY OF FINDINGS

In the findings, we simulated for two data Types of Balance and Unbalance designs with different Levels of generations (3×3, 7×7, 10×10 and 3×7, 7×3, 7×10, 10×7 respectively). We therefore check the performance of G×E interaction on four different models (AMMI, FW, GGE and Mixed model), and also their stability and adaptability. The findings revealed that, when the assumption was maintained, AMMI outperformed Finlay-Wilkinson model, GGE Biplot model and Mixed model.

However, in this same study, two case scenarios were considered to have a clearer view of the performance of Genotype by Environment Interaction on the four models. We experiment the inference behind the violation of the assumption of normal distribution by observing the data contamination of two case scenarios (Lowest and Highest outlying observations). It was observed on the two data Types of Balance and Unbalance designs with different Levels of generations. To achieve that, comparative performance of the data contamination techniques under the two case scenarios were assessed as follow;

Case I scenario was done for Lowest Outlying Observations where 50%, 100% and 500% data contamination on the First Quarter (P_1) , Mid quarter (P_2) and Last Quarter (P_3) were examined on each levels of generations. We then deduced from the result of the model evaluation that, at each levels of data contamination for Balance and Unbalance design, Mixed model was the ideal model for G×E interaction.

							LEVEL I						
3×3 Design				RMSE				MSE				Abs. Bias	
Contamination		AMMI	FW	GGE	Mixed model	AMMI	FW	GGE	Mixed model	GGE AMMI FW			Mixed model
	$P_{\scriptscriptstyle 1}$	3.9776	6.2548	3.6947	2.5305	2.5965	3.0219	4.4741	1.7210	2.6147	3.2535	4.8600	1.5760
50%	P ₂	4.4662	0.7659	2.3932	0.1986	2.1370	1.3478	3.4552	0.4613	2.9730	1.4564	0.2559	0.1322
	P_{3}	0.6618	5.3054	2.1046	0.6250	3.9005	3.3206	1.0023	0.1564	3.4524	0.1712	.4008	0.0987
100%	P_{1}	0.8302	0.4066	0.5978	0.0089	3.6864	1.8896	2.7429	1.1382	2.5447	2.6244	0.7718	0.5997
	P_{2}	3.5122	3.1019	4.5554	1.0005	0.5478	4.1373	1.9141	0.1382	2.3389	3.7763	5.7263	0.4829
	P_3	1.3398	2.0190	2.3720	1.4642	4.4850	2.3743	1.1382	0.0189	0.2779	0.3519	3.7643	1.5014
500%	P_{1}	3.3975	4.0813	2.1924	1.5235	1.3683	0.7436	4.4258	0.5622	0.9487	2.9123	2.2130	0.2806
	P_{2}	3.2508	1.5987	1.2958	0.5324	0.7822	2.7938	2.3045	0.3631	1.1823	3.7015	2.0555	1.1224
	P_3	3.5083	2.9215	4.4020	0.0023	2.1172	3.0717	3.0969	1.9522	2.8421	1.8497	2.2907	1.0119
							EVEL II						

CASE 1 - Table 1. Type I: Balance design

							LEVEL III							
10×10 Design			RMSE					MSE		Abs. Bias				
Contamination		AMMI	FW	GGE	Mixed model	AMMI	FW	GGE	Mixed model	AMMI	FW	GGE	Mixed model	
	P_{1}	2.6331	6.8466	3.3917	2.7384	2.3036	4.1953	5.1651	1.6818	3.7856	0.4989	1.5320	0.2384	
50%	P ₂	1.7304	3.7265	1.6823	0.2257	1.3571	3.6416	0.7979	0.1904	3.4438	3.2548	1.0151	1.0084	
	P_3	1.8348	1.9396	3.2759	1.3484	1.4645	1.0673	0.9434	1.0441	1.1193	4.2652	4.9487	1.0718	
100%	P_{1}	3.4321	3.9010	2.9753	1.6135	2.6769	1.2546	3.4052	0.5478	2.5582	3.4722	5.7779	0.4021	
	P_{2}	3.8016	0.8231	6.9347	0.2874	0.3907	5.6913	4.5928	0.2178	0.1965	1.9985	1.1985	1.1069	
	P_3	3.1948	1.3203	0.2859	0.7834	5.6026	2.6734	5.2994	2.0747	1.7705	3.3839	1.6280	1.4539	
500%	P_{1}	.8984	0.3656	3.4270	0.8695	2.6337	1.8211	1.0579	0.6981	2.9803	3.5371	2.6059	2.3114	
	P_{2}	2.6904	6.4308	0.8105	0.1153	1.1072	0.7594	2.4406	0.1604	3.0310	1.8189	1.0580	0.8345	
	P_3	5.2746	4.5194	2.6493	0.8888	6.9412	0.7436	1.4775	0.1334	4.8001	0.9099	1.3112	0.3113	

CASE 1 - Table 2. Type II: Unbalance design

Oyamakin and Durojaiye; AJRAF, 5(2): 29-45, 2020; Article no.AJRAF.53958

39

1.8914 3.6685 5.2327 0.1050 2.1975 5.2735 1.0182 0.7774 3.2362 2.5087 1.4944 1.4829

0.7860 2.7948 2.1190 0.6122 3.1051 1.7604 3.7557 0.6933 2.4068 1.3888 4.7508 0.8761

2.3873 0.1125 0.9368 0.1063 2.1870 6.1950 3.5534 1.3832 1.0771 3.8293 4.0075 0.6893

500%

 P_1

*P*₂

*P*3

Oyamakin and Durojaiye; AJRAF, 5(2): 29-45, 2020; Article no.AJRAF.53958

	LEVEL IV												
7×10 Design			RMSE MSE						Abs. Bias				
Contamination		AMMI	FW	GGE	Mixed model	AMMI	FW	GGE	Mixed model	AMMI	FW	GGE	Mixed model
	D	0.2579	1.3002	1.8793	0.9268	7.7057	1.9438	1.0843	0.9430	3.3818		2.7349 3.4363	1.8128
50%	Р,		1.4885 1.4120	1.5012	0.6238		2.3528 2.7359	3.1290	1.9584	2.7117	1.6961	1.3069	0.6420
	P_3		4.6320 0.7134	1.7653	0.0081		3.2045 2.3589	3.0521	2.3572	1.6095	3.2579 0.1977		0.0928
	P_{1}		1.1259 4.0580	0.6756	0.4793		1.0366 1.9033	0.8564	0.5197	3.9285	5.8000	1.6967	0.5478
100%	P_{γ}		0.7850 1.5907	0.4744	0.1925		3.5802 2.9806	1.1144 0.2105		1.0918		2.7243 1.0612	0.6818
	P_{3}	0.9111	5.1454	0.5803	0.3725		4.2133 4.8490	2.0656	1.5703	4.1112		2.3778 2.0362	1.4701
	P_{1}		0.5309 4.6632 2.1003		0.1413		2.6229 4.3176	1.9380	0.3578	1.0617		4.3839 1.5758	0.9996
500%	P_{2}		1.0500 4.3233	0.6974	0.2318		0.5599 2.9307	0.0105	0.1539	3.3370		2.9522 2.7589	0.4961
	P_{3}		0.5378 2.1718	1.3371	0.2078		5.3866 3.2484	1.2911	0.4443	0.3455	2.2293 7.0051		0.3439

CASE 2 - Table 3. Type I: Balance design

LEVEL III

CASE 2 - Table 4. Type II: Unbalance design

Oyamakin and Durojaiye; AJRAF, 5(2): 29-45, 2020; Article no.AJRAF.53958

Case II scenario was done for Highest Outlying Observations where 50%, 100% and 500% data contamination on the First Quarter (P_1) , Mid quarter P_2) and Last Quarter (P_3) were examined on each levels of generations. We then observed from the result of the model evaluation that, at each levels of data contamination for Balance and Unbalance design, Mixed model outperformed the other three models.

5. CONCLUSION

- The study has clearly shown that the four models considered detects the G×E interaction effect in a different way. We were able to evaluate and described G×E interaction performance by their stability and adaptability using multi-location trials.
- Also, this study confirmed the suitability of AMMI in detecting G×E when the assumptions are maintained or kept. That is, when outlier is not influential, AMMI can be used. More so, we were able to deduced from the study that even though AMMI has been adopted in the literature to summarize patterns and relationships of G×E successfully, it is not an appropriate model in a situation where outlying observations are influential but rather the Mixed Model which is often neglected by the plant breeders and agronomists. Therefore, the findings revealed that AMMI, FW and GGE are not robust to outlying observations.

6. SUGGESTION FOR FURTHER STUDY

The following area may be considered in further research;

- Modification of AMMI model in order to be sensitive to violation of model sensitive to assumptions.
- Our choice in G×E methods for mixed models with bilinear models' terms for G×E seems to be a good reflection of the current trends in the literature. Therefore, a comparative study should also be check more on Mixed models, Crop growth models and Bayesian models.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

- 1. Borron S. Building resilience for an unpredictable future: How organic agriculture can help farmers adapt to climate change. Food and Agriculture Organization of the United Nations, Rome; 2006.
- 2. Kumar S, Bourai VA. Economic analysis of pulses production, their benefits and constraints: A case study of sample villages of Assan Valley of Uttarakhand, India. IOSR Journal of Humanities and Social Science. 2012;41-53.
- 3. Keim P, Pearson T, Okinaka AR. Microbial Forensics: DNA fingerprinting of *Bacillus anthracis* (anthrax). Analytical Chemistry. 2008;80(13):4791-4800.
- 4. Tumuhimbise R, Melis R, Shanahan P, Kawuki R. Genotype environment interaction effects on early fresh storage root yield and related traits in cassava. The Crop Journal. 2014;2(5): 329–337.

View at: Publisher Site | Google Scholar.

- 5. Meuwissen THE, Hayes BJ, Goddard ME. Prediction of total genetic value using genome-wide dense marker maps. Genetics. 2001;157(4):1819-1829.
- 6. Cobb JN, Declerck G, Greenberg A, Clark R, McCouch S. Nextgeneration phenotyping: Requirements and strategies for enhancing our understanding of genotype-phenotype relationships and its relevance to crop improvement. Theor. Appl. Genet. 2013;126(4):867-887. DOI: 10.1007/s00122-013-2066-0
- 7. Cooper M, Messina CD, Podlich D, Totir LR, Baumgarten A, Hausmann NJ, Wright D, Graham G. Predicting the future of plant breeding: Complementing empirical evaluation with genetic prediction. Crop; 2014b.
- 8. Dawson JC, Endelman JB, Heslot N, Crossa J, Poland J, Dreisigacker S, Manes Y, Sorrells ME, Jannink J-L. The use of unbalanced historical data for genomic selection in an international wheat breeding program. Field Crops Res. 2013; 154:12-22.

DOI:10.1016/j.fcr.2013.07.020PastureSci.6 5(4):311-336

DOI:10.1071/CP14007

9. Heslot N, Akdemir D, Sorrells M, Jannink J-L. Integrating environmental covariates

and crop modeling into the genomic selection framework to predict genotype by environment interactions. Theor. Appl. Genet. 2013;127:1-18.

- 10. Jarquin D, Crossa J, Lacaze X, Cheyron P, Daucourt J, Lorgeou J, Piraux F, Guerreiro L, Perez P, Calus M, Burgue~no J, Campos G. A reaction norm model for genomic selection using high-dimensional genomic and environmental data. Theor. Appl. Genet. 2013;3:1-13.
- 11. Bustos-Korts D, Malosetti M, Chapman S, FA van Eeuwijk. Modelling of genotype by environment interaction and prediction of complex traits across multiple environments as a synthesis of crop growth modelling, genetics and statistics. Springer, New York. 2016;55-82. DOI:10.1007/978-3-319-20562-552
- 12. Malosetti M, Bustos-Korts D, Boer MP, van Eeuwijk FA. Multi environment genomic prediction: Issues in relation to genotype by environment interaction. Crop Sci; 2016.

DOI:10.2135/cropsci2015.05.0311

- 13. Hammami H, Rekik B, Gengler N. Genotype by Environment Interaction in Dairy Cattle. Biotechnologie, Agronomie, Societeet Environnement. 2009;13(1).
- 14. Finlay KW, Wilkinson GN. The analysis of adaptation in a plant breeding program. 1963;742-54.
- 15. Yan W, Tinker NA. Biplot Analysis of Multi-Environment Trial Data: Principles and

Applications. Canadian Journal of Plant Science. 2006;86:623-645.

DOI: http://dx.doi.org/10.4141/P05-169

- 16. Gauch HG. Model selection and validation for yield trials with interaction. Biometrics. 1988;705-715.
- 17. Purchase JL, Hatting H, Van Deventer CS. Genotype environment interaction of wheat in South Africa: Stability analysis of yield performance. South African Journal of Plant and Soil. 2000;17(3):101–107. View at: Publisher Site | Google Scholar.
- 18. Farshadfar E, Mahmodi N, Yaghotipoor A. AMMI stability value and simultaneous estimation of yield and yield stability in bread wheat (Triticum aestivum L.). Australian Journal of Crop Science. 2011; 5(13):1837–1844. View at: Google Scholar
- 19. Bose LK, Jambhulkar NN, Pande K, Singh ON. Use of AMMI and other stability statistics in the simultaneous selection of rice genotypes for yield and stability under direct-seeded conditions, Chilean Journal of Agricultural Research. 2014;74(1):1–9. View at: Publisher Site | Google Scholar
- 20. Piepho H-P. Stability analysis using the SAS system. Agron. J. 1999;91:154-160. Crossref
- 21. SAS Institute. Table 46.7 in the online document; 2015. Available:http:==support:sas:com=onlined oc=913=getDoc=en=statug:hlp=mixedsect 19:htm [Accessed 11 June 2015].

 $_$, and the set of th *© 2020 Oyamakin and Durojaiye; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.*

> *Peer-review history: The peer review history for this paper can be accessed here: http://www.sdiarticle4.com/review-history/53958*