Arbeitsgemeinschaft
Tropische und Subtropische Agrarforschung (ATSAF) e.V.

Council for Tropical and Subtropical Agricultural Research

ATSAF - CGIAR++ Junior Scientists Program Final Report

Name of student: Srinivasa Reddy Mothukuri
University: Georg-August-University Göttingen
Supervisor at University: Dr. Stefanie Griebel, Prof. Dr. Timothy Beissinger
International Agricultural Research Center: International Maize and Wheat Improvement Center (CIMMYT)
Country: Mexico
Supervisor at IARC: Dr. Juan Burgueño
Start and end date of stay at IARC: -
Start and end date of remotely supervised project: 01 Jul 2021 - 31 Dec 2021
Title: Optimization of sparse phenotyping strategy in multi-environmental trials in maize

I am extremely honoured and humbled to have been selected as a recipient of the ATSAF-CGIAR ++ Junior Scientist, as it has been the highlight of my master's program. My CGIAR++ centre was International Maize and Wheat Improvement Center (CIMMYT) in Mexico. I have been working on my master thesis project for the last six months. This research was divided into two chapters: Chapter 1:- Traditional complete data sets of CIMMYT Global Maize Program (GMP) for mimicking sparse phenotyping strategies (Historical data). Chapter 2:- Real sparse phenotyping data sets from CIMMYT Global Maize Program.

Role and responsibility of the MS Thesis

This research project was designed into two chapters, as I mentioned above. Each chapter has different objectives to solve during my six-month research. The chapter-1 focused on optimizing sparse phenotyping strategies to increase genetic gains, which is crucial. The research's main objectives were A) To compare the different sparse designs to decide which sparse design reduces the evaluation cost while increasing the precision. B) How to allocate the lines into sets C) To identify which statistical measurements provide better prediction accuracy for optimization of the training set. D) To evaluate the accuracy of different sparse designs and allocation methods in order to decide the better sparse design. The chapter-2 focused on the to evaluate real sparse phenotyping MET results obtained by GMP in Africa. My role and responsibility were to solve all objectives during the six months to complete this project. I prepared one document with different steps. My responsibility was to update the work progress in our weekly meetings and complete all the steps in order to complete the project.

MS thesis in remotely:-

I was interested in working from CIMMYT, unfortunately not possible due to the Covid-19. CIMMYT was not allowing new visitors in earlier because of the covid-19 regulations at CIMMYT and Mexico. My master thesis was planned only for analysis, I got the existing data sets required for the analysis. I used these data sets as part of my master thesis study, which was supervised by CIMMYT and co-supervised by Göttingen University. Dr. Juan Burgueño explained the situation in CIMMYT at the beginning of my master thesis and started to work on it remotely. I completed my statistical analysis successful from remotely. I, Dr. Juan Burgueño, Dr. Stefanie Griebel saw all possibilities and permissions to travel to Mexico twice; unfortunately, it was impossible due to covid-19 situations in Mexico and Germany. During the planned to travel CIMMYT, Dr. Juan Burgueño and Dr. Stefanie

Griebel helped to collect the documents from the CIMMYT and other necessary documents for visa application. As Covid 19 spread, Dr. Juan Burgueño kept me regularly updated on the Mexico situation and current travel guidelines issued by the Mexican government and CIMMYT regulations. Finally, it was impossible to travel and work in CIMMYT.

Interaction with Division of Plant Breeding Methodology & Biometric and Statistics Unit (BSU):-

➢ Dr. Juan Burgueño and Dr. Stefanie Griebel co-supervised my master thesis. I have received outstanding support from both of them during my master thesis analysis.

➢ Dr. Juan Burgueño & Dr. Stefanie Griebel provided a one-hour weekly basis to update my research progress and receive feedback during the meetings.

➢ They were interested in teaching the knowledge about the research to the student for my future career; it has helped me gain good knowledge about the research.

➢ I never faced any problems with my advisors during their supervision; it was an excellent experience to learn so many new things while working with them.

➢ I was officially a master thesis student in the division of plant breeding methodology at Goettingen University. I have regularly attended maize group and weekly division meetings with the division of plant breeding methodology at the University of Goettingen.

➢ I gave Four formal 20 minutes presentations and 40 minutes questions and discussions with a division of plant breeding methodology. In one of the presentations, Dr. Juan Burgueño attended and gave feedback. It helped me to know how much knowledge I gained and where I needed to focus on my project.

➢ I have given the formal presentation in the Biometric and Statistical Unit in CIMMYT. It was approximately 35 to 40 minutes of the complete presentation about chapter-1 and chapter-2. Dr. Stefanie Griebel and Prof.Dr.Timothy Beissinger attended this presentation from the Goettingen University. I received the feedback of my research and discussion in reaming 20 to 25 minutes.

➢ As a result, all best communication a great desire to collaborate and work on this project jointly.

Skill development for future career

The abilities necessary to complete the four work packages, script-writing skills, loops and provide the intended deliverables included a significant amount of time spent in RStudio. I learned to write an R script, working on a Linux server to run the mixed models by using the ASReml-R package. ASReml-R is a paid package that was used to calculate the variance components, coefficients, model for prediction during my research. In the earlier master thesis, Dr. Stefanie Griebel contacted an animal breeding group in Goettingen university to get a license to use the ASReml-R package. The animal breeding group gave the license to access the ASReml-R, but it was expired at the end of November in 2021 after Dr. Juan Burgueño gave access to ASReml-R and I completed all analysis. I was able to learn the essential skills in R Studio to successfully finish all work in packages and produce all outputs on time. I learned more technical abilities in the last six months of my study than in the previous six years. R is the scientific language, and these six months have improved my ability to communicate in it, if not fluently. I learned to run the shell script and batch file in order to complete my analysis automatically. These technical abilities will benefit me significantly in the future.

Methodology

I have received sufficient data from the CIMMYT to analyse and complete my master thesis, which was collected in 2018 under the CIMMYTs' Global Maize Program in Africa. First, I started to understand phenotypic and genotypic data. After understanding the data sets, different sparse phenotyping strategies were designed to allocate the lines into the sets among the five different locations in terms of precision and cost. Sparse designs is a phenotyping strategy in which not all genotypes are evaluated in the field; some of them need to be predicted. For example, the data is divided into full and sparse sets. The full set was evaluated in all locations, while the five sparse sets were evaluated in four locations (Burgueño et al., 2012). A mixed model (Butler, D. G et al., 2017) was built to calculate the variance components and coefficients to calculate the adjusted values. Lines were allocated according to sparse designs randomly for the complete data set. Different statistical measurements (Akdemiret, 2017) were calculated to see the relationship between the full and sparse sets. Researchers have not used these statistics in sparse designs; that is the reason to see around 20 different statistical methods to optimize the training set. Sparse phenotyping strategies were analysed using genomic enabled prediction models (Jarquín, Kocak, et al., 2014) to predict not tested genotypes in sparse designs. I have evaluated the results for all sparse designs to calculate precision by using

Pearson correlation between the observed and predicted breeding values. I saw the common percentage of genotypes from observed and predicted breeding values as well as selection differential.

Conclusion:-

I believe I gained a lot of experience and knowledge from my supervisors Dr. Juan Burgueño at CIMMYT and Dr. Stefanie Griebel, and Prof. Dr. Timothy Beissinger from Georg-August-Universität Göttingen. I would like to express my gratitude to them for providing me with the opportunity to do this research under their supervision. Goal and objectives were successfully solved and found expected results. I am currently writing a research article as mentioned in our research proposal.

Being a research scholar with the ATSAF-CGIAR ++ Junior Scientist Program has been a wonderful experience. It has enabled me to acquire new abilities and contribute critical information to the research centre. My efforts and development of technical skills through the ATSAF-CGIAR ++ Junior Scientist Program will undoubtedly lay a solid foundation for my future. The Junior Scientist Program exposed me to this professional path and steered me toward applied research in the future.

Acknowledgements:-

I take this opportunity to express my gratitude to all the people who have helped, guided and encouraged me during my research work. I thank the almighty for his blessings and love, which helped me achieve my present goal.

At this moment of accomplishment, first of all, I would like to express my deep sense of gratitude and thank you a lot to my supervisor Dr. Juan Burgueño, who gave me the opportunity to work under his project. I would like to thank you so much for Dr. Stefanie Griebel for supervising the from Goettingen university as the main supervisor. I am thankful to Prof.Dr. Timothy Beissinger for supporting the division of plant breeding methodology. It has been my privilege to work under them as a team. I would always cherish the opportunity of working with Dr. Juan Burgueño, Dr. Stefanie Griebel and learning from their knowledge and experience. Thanks for the general support and showing that you care.

I am a self-funding master student and very interested in working with CIMMYT from my under-graduation; it was fulfilled with the support of the ATSAF -CGIAR ++ Junior Scientist Program. My master thesis was made possible with the generous financial support of the ATSAF -CGIAR ++ Junior Scientist Program. I am very thankful for the ATSAF -CGIAR ++ Junior Scientist Program for

- 6 -

providing the funding support and helping to complete my master thesis financially. Furthermore, my special gratitude goes to Dr. Alejandro Pieters (Coordinator ATSAF - CGIAR++ Junior Scientists Program) for support in queries related to funds and documents. I got very good help and supported from her.

References

Akdemir, D. (2017). STPGA: Selection of Training Populations by Genetic Algorithm. R package version 5.2.1.

Butler, D. G., Cullis, B.R., A. R. Gilmour, Gogel, B.G. and Thompson, R. 2017. ASReml-R Reference Manual Version 4. VSN International Ltd, Hemel Hempstead, HP1 1ES, UK.

Burgueño, J., Campos, G. de los, Weigel, K., & Crossa, J. (2012). Genomic Prediction of Breeding Values when Modeling Genotype × Environment Interaction using Pedigree and Dense Molecular Markers. Crop Science, 52(2), 707–719. https://doi.org/10.2135/cropsci2011.06.0299

Jarquín, D., Crossa, J., Lacaze, X., Du Cheyron, P., Daucourt, J., & Lorgeou, J. et al. (2014). A reaction norm model for genomic selection using high-dimensional genomic and environmental data. *Theoretical And Applied Genetics*, *127*(3), 595-607. doi: 10.1007/s00122-013-2243-1.

- 6 -