



Statistical Research Issues in Crop and Cropping System Experiments for Enhancing Food Security Support

Murari Singh

International Center for Agricultural Research in the Dry Areas (ICARDA), Amman, Jordan

SUMMARY

Food security is essential to maintain the human population and the environment. The strategies for food security must take into consideration the human population dynamics and dynamics of natural resources, plant and animal populations on which human survival depends. It is essential to look into effective options for search and conservation of crop and animal germplasm, along with necessary genetic modifications to develop new seeds and suitable embedding of the new seeds in cropping systems for high and sustainable productivity. At almost every significant stage of technological development, statistical designs to gather valid and efficient evidence (data) and methods of assessment of the evidence are needed to form decisions on the technologies. This presentation is restricted to the crop and cropping system aspects of experiments. The statistical issues related to the following would be discussed: (1) search and maintenance of diverse crop germplasm, (2) evaluation of crop germplasm, (3) cropping systems related to conservation agriculture and crop rotations, and (4) need of crop actuary. In the context of germplasm conservation two aspects of maintaining the germplasm, mini-core selection and focused identification of germplasm strategy will be discussed. The issues related to field experimentation for evaluating moderate to large number of genotypes and methods of statistical data analysis with a view to exploit genotype – environment interaction will be presented. The challenges of design and analysis of cropping systems and evaluating the data in terms of productivity and sustainability indicators will be discussed. The role of a crop actuary to mitigate riskiness due to the effect of weather uncertainty on crop production and productivity will also be discussed in brief.

Keywords: Field experimentation, Crop diversity, Productivity, Sustainability, Riskness.

1. INTRODUCTION

In order to feed the growing human population with constraint on availability of arable lands, further enhancement of sustainable productivity of various food and feed crops— cereals, legumes, oil seeds, fruits and vegetables— is essential. The key research areas that support crop productivity include: (1) search a suitable genetic resource base containing genes of interest, (2) plant breeding techniques and tools for gene transfer followed by evaluation of new seeds in multi-environment trials, (3) identify the suitable production practices using the new seeds for short-term intensified

production options and for long-term sustainable production with due conservation of the biophysical environment, and (4) need for a crop actuary in uncertain climatic changes.

Some of the statistical issues related to the above four components of agricultural research are highlighted and discussed in this presentation.

2. SEARCH FOR GENETIC RESOURCES

Conservation of bio-diversity has been a practice of all concerned professionals including farmers since time immemorial. However, the changes in

environmental conditions, population pressure and technological advancement have resulted in genetic modification, including replacement of landraces, and erosion of genetic diversity. To minimize genetic erosion and contain the loss of diversity, Convention on Biological Diversity (Articles 8 and 9), Agenda 21 (Chapters 14 and 15) (<http://www.iisd.org/rio+5/agenda/agenda21.htm>) and the Global Biodiversity Strategy (WRI *et al.* 1992) have stressed collection and conservation of biodiversity using several mechanisms such as *in situ*, *ex situ* and *in vitro*. A number of references dealing with various aspects can be found in Guarino *et al.* (1995). Technical guidelines for germplasm exploration and collection, including planning, methods and procedures illustrated with real germplasm collection missions are given in Engels *et al.* (1995). Planning and execution of a genetic resource collection are given in Bennette (1970), Chang (1985), Damania (1989) and Kameswara Rao and Bramel (2000).

At ICARDA, the Genetic Resources Unit undertakes the germplasm collection-missions and maintains the germplasm *in-situ* and *ex-situ* environments. The germplasm collections are large in size, thus sampling becomes a practical necessity, and therefore, devising efficient statistical methods is a challenge.

2.1 Genetic Diversity in a Population and Sampling Strategy for a Mini-core Collection

One practical way for preserving the species and genetic diversity is to collect and conserve (and regenerate) germplasm accessions with maximum diversity in species and genomic information in relation to the environment or region. The sampling strategy would depend on the population structure, distribution of the traits and genes, a statistical measure of the diversity captured in the sample and the precision required. On the one hand, the sample (sampling fraction) should be large enough to contain maximum diversity, but on the other hand, should be small enough to be curated — collected and maintained — within the limited time and resources available in practice. Optimum sampling strategies for genetic conservation of crop plants under threat of extinction have been discussed by Marshall and Brown (1975), Weir (1990) and Brown and Marshall (1995), among others. The optimum sample depends on the genetic variation

within and between the set of populations under investigation in terms of genotype and allele frequencies or allelic richness, gene diversity, heterozygosity levels, disequilibrium coefficients etc. Often in practice, allelic richness for a single population is measured as the average number of alleles for a large number of markers. In case of sampling from several populations, the sampling strategy depends on the extent of genetic divergence among populations (*e.g.* in terms of number of alleles that attain appreciable frequencies in individual populations) and spread in the level of genetic variation (*e.g.* in the distribution of number of alleles per locus). Upadhyaya *et al.* (2010) suggested approach of mini-core, as a miniaturization of larger germplasm collection with almost full representation of genetic diversity, which helps in identifying the population structure, association mapping and mining targeted genes (see also Upadhyaya and Ortiz 2001).

Statistical methods are needed for developing optimum sampling strategy in terms of 1) sampling fraction, 2) scheme of selecting the accessions to form a core or mini core, 3) types of phenotypic traits as well as which traits, 4) number of genetic markers, and 5) diversity measures — number of alleles and indices. The commonly used approach is to select 5-10% of available accessions in large genebanks to form a core of accessions for evaluation and removal of duplicates and apply clustering techniques to form clusters of accessions with maximum diversity, and compare them in relation to the core.

2.2 FIGS (Focussed Identification of Germplasm Strategy)

While core sampling strategy is desirable for mining alleles for all the traits, it is highly improbable to achieve in practice. Another strategy practiced at ICARDA (2013) is of FIGS, “focused identification of germplasm strategy”, where germplasm samples are first used to identify traits responsible for biotic (*e.g.*, diseases and pests: Powdery Mildew, Sunn Pest, virulent wheat stem rust strain, Ug99 and Russian Wheat Aphid) and abiotic stress (*e.g.*, drought and heat), climate change responsiveness and tolerance. The ICARDA approach develops an association/relationship between the desired trait characteristics and the environmental parameters and predicts the characteristics for other geographic locations of interest.

The prediction may help identify hot-spots for presence of accessions with desired characteristic. For wheat stem rust, see Bari *et al.* (2012).

Statistical methods are needed to address: 1) use of geo-reference which is available at some cost, 2) use of environmental variables, which are normally measured with errors, 3) identification of traits of interest for FIGS approach, which are generally binary or qualitative, *e.g.*, dealing with resistance to biotic and abiotic stresses, 4) preparation of GIS maps for sources of resistance to various stresses, 5) while FIGS is efficient in identifying the sources of resistance for forming a smaller set of accessions, assessment of the genetic diversity for other loci, and 6) having identified traits for key stress factors, identification of genotypes responsive/tolerant to climate change for a given trait. **The challenge lies in developing statistical techniques using weather variables that could be used to model the genotypic response, *e.g.* developing effective indices of weather variables including their distribution over the cropping cycle, in terms of principal components.**

2.3 Determination of Sample Size and Number of Loci for Diversity Assessment

Determining the number of accessions for estimating genetic diversity using Simpson's index (based on heterozygosity) is of interest to Genetic Resources scientists. For a population in Hardy-Weinberg equilibrium (when allelic frequencies follow the binomial distribution), the number of individuals sampled can be determined using the variance expression. However, the inbred species are never in Hardy-Weinberg equilibrium. **The statistical issues are to estimate/determine the sample size/number of accessions as well as the number of markers/loci that may be used to capture the diversity.** In an evaluation of a set of wild types of wheat using AFLP (amplified fragment length polymorphism) marker data at ICARDA, the sample size was obtained using a distribution-free resampling method. This approach led to exploration of an empirical relationship between sample size and precision of diversity. Such an approach is planned for SSR marker data (Singh *et al.* 2008, Nelson and Anderson 2013). The resampling method was also applied to the determination of sample size required for developing interrelationships or clustering of populations based on a criterion such as

the least squares scaling of stress (LSS) derived from the similarity matrix (Singh and Chabane 2010).

3. CROP IMPROVEMENT—NEW SEEDS

3.1 Crop Variety Trials – Experimental Designs

Crop improvement is a long process of selection-crossing-selection cycle passing through selection of parents, crossing of the parents producing progenies which are then selected for desirable trait and further crossings may be followed by further selections and so on. There are various crossing-selection strategies followed in different crops: pedigree methods, repeated back-crossing, bulking or single seed decent and their variations (Wricke and Weber 1986, Rajaram *et al.* 1995, Allard 1999). At advanced generation of genetic materials, *e.g.* F_7 in lentils, the evaluation of genetic material needs to be carried out in field trials in multi-environments. At a preliminary stage of evaluation, limited amount of seed restricts the experiments to be implemented in only un-replicated trials, while seeds become available for replicated trials only at further later generations. In this context, there are a number of issues on designs and analysis of crop variety trials including: 1) Accommodating the availability of variable replications – *e.g.* a single replication of materials of early generation, two or more replications of advanced generation materials, and even more replication of the checks. 2) In the process of development of new germplasms – retaining or exclusion of existing genotypes provides a sound basis for assuming a random effect model for the test genotypes. Therefore, it is desirable that criterion for developing experimental design must incorporate this feature, and thus maximize the genetic advance or heritability. 3) There is also a need for designs to account for spatial variability. If the genetic material comprises diverse groups of genotypes, *e.g.*, having different maturity groups or from diverse crosses, or arising from the different common-ancestors, the designs would be needed to account for variation in the genotypic variability across the groups. Search for efficient designs incorporating all these features, might require efficient iterative algorithms with convergence in high dimension. Australian biometricians have developed innovative computing algorithm and R-packages to generate experimental designs with many of the above features and such designs have gained popularity in Australian agricultural research (Cullis *et*

al. 2006). **There is a need to evaluate these designs with due modifications suited to the requirements of agricultural research in other institutions.**

3.2 Crop Variety Trials - Statistical Analysis

The classical statistical analysis is based on modeling observed data in terms of design factors and independent and homogenous plot errors. While spatial variability has been seen as reality, its accountability in terms of plot-error covariances is yet to become a routine. Further, the genotypes often share pedigree and thus have common ancestors. However, due to lack of this information, the estimates of genotypic performance do not get their due precision. Often a trial is seen as well-run as long as the coefficient of variation in the field is estimated to be below 10%, even though scope exists to explore ways to increase the precision and genetic advance further. One such situation calls for examining the spread of unstructured heterogeneity in the field and/or associated with certain group of genotypes. The heterogeneity may arise due to residual effects from the preceding seasons, unstructured spread of pests and disease-pathogens on the host plants in the field. Such heterogeneity in general is less likely to be modeled by transformation or the use of a link function in generalized linear/non-linear model. Some attempts have been made at ICARDA to account for unstructured heterogeneity between clusters of plots for uncorrelated/correlated plots errors (Singh *et al.* 2010, 2012) and between clusters of genotypes (Singh *et al.* 2013). These approaches include a level of similarity to form the clusters and a level of significance to control Type I error on heterogeneity of variances. **In both the cases however, improved methods are needed for cluster formation and detecting heterogeneity with a control on Type I error.**

3.3 Genetic/Molecular Marker Technology - High Dimensional Data

With the availability of markers over the whole genome the MAS (marker assisted selection) and GS (genomic Selection) techniques seek to build regression models to predict the breeding value of the individuals. **The challenge lies in dealing with very large number of variables, and much larger than the number of individuals on which the trait values are obtained, incorporating the kinship relationship among the individuals, and building a multi-trait model**

(Meuwissen *et al.* 2001, Crossa *et al.* 2010, Nakaya and Isobe 2012, Segura 2012).

3.4 Integrated Pest and Disease Management

In the IPDM (Integrated Pest and Disease Management) areas, the objective often is to screen genotypes which are resistant to applicable insects or diseases. The response variables, called the scores, are in general ordered categorical or ordinal variables. Table 1 shows the complex, non-linear, scoring system for recording the Ascochyta blight, Chocolate spot or rust damages in legumes such as faba bean/chickpea. In the IPDM field experiments, normally two check cultivars are used, one is susceptible and the other is resistant. The limitation of the experimental design is

Table 1. Disease scores, disease class and description of the response symptoms in Faba bean.

Score	Class [§]	Description
1	R	No diseases or slight attack without any development
2	R	Slight attack with moving from one leaflet to another
3	R	Small aggressiveness in the leaves of lower nodes with less than 1mm of diameter in the top of the plant
4	MR	Development of spores aggressively and displaying to the leaflet of the 1/3 of the plant
5	MS	Development of spores on the stem and dropping it for the 1/3 of the plants
6	S	Diseases developed on the pods
7	S	Death at least 30% of the plant
8	S	Death at 50% to 80% of the plant
9	S	Total death of the plant

[§]R = Resistant, MR = Moderately resistant, MS = Moderately susceptible, S = Susceptible

that the test entries have none or two replications, however, the checks are to be more frequently repeated or replicated over the layout to provide an even coverage of the layout. The main purpose of the susceptible check is to ensure sufficiently high infestation level in the screening area/layout which is essential for searching resistant lines. In case the infestation could be artificially developed, one could examine the success of screening with the help of (non-

spatial) distribution of infestation as observed in terms of susceptible check(s) scores. If the infestation is natural, the methods are needed that are capable of making spatial adjustment to the score which are ordinal in nature.

In common practice, the underlying statistical model assumes additive effects and normal errors, and is validated by the behaviour of the residuals on checks. ICARDA GenStat® programs are available for spatial adjustment for quantitative response. These programs are also applied to the data with scores on 1 to 9 scales. Alternatively, nonparametric methods based on ranks are also used. **However, there are limitations on their availability for incomplete blocks and spatially dependent data.** This situation needs further improvement. **The challenge lies in building suitable models for evaluation of genotypes using score-based variables.**

Furthermore, recording of the scores is often subject to error of measurement/misclassification with probability varying with the score. Probability of misclassification of the extreme scores may be negligible while it could be maximum in the middle (*i.e.* around score 5). Analysis of scores is very routinely done. **Suitable statistical methods which could incorporate misclassification errors would be worthwhile. Bayesian approach might be an appropriate way for analysis of score datasets.**

3.5 Examining the Nature of $G \times E$ Interaction using Climatic Information

$G \times E$ (genotype \times environment) interaction studies are conducted to select the best performing genotypes in the sense of high yield and stability over a range of stress factors. Such analyses are often limited to partitioning of the environment variation into location and year components of variation, but not any further, *e.g.* partitioning using bio-physical measurements of the environments. Current advances in ICT (Information and Communication Technology), GPS (Geographical Positioning Systems), and remote sensing satellites technologies have led to collection of affordable space-time reference data on temperature, rainfall, soil moisture at the experimental plot level as well as the crop growth parameters with time. However, their use in evaluation of genotypic response to environments has been very limited. **This is one area which needs to be looked at for building models for genotypic response**

and selecting crop genotypes for specific adaptation to regions of various climatic conditions.

During the recent decades, climate change has been blamed for creating noticeable effects leading to degradation of human existential resources — crops, livestock, marine life, forests, biodiversity, water, land, and other resources — and various aspects of life. To cope with implications of climate change, CGIAR has recently reorganized its research agenda (<http://www.cgiar.org/our-research/cgiar-research-programs/>) with distinct emphasis on CRP Dryland Systems lead by ICARDA and CCFSA (Climate Change, Agricultural and Food Security). One may identify the barley and durum wheat land races for responsiveness and/or tolerance to changes in temperature and rainfall profile based on monthly data by utilizing principal components scores and PLS (Partial least squares) components scores in order to extract the climatic information with using less number of variables. With much more richer datasets on climatic variables, there may be several other ways to look into further possibilities.

Predicting crop response for the climate change: Where crop response model using weather variables are available, generation of geo-spatial and temporal scenarios will help guide work on various Food Security options specific to the target regions.

4. CROPPING SYSTEMS FOR SUSTAINABLE FOOD PRODUCTION

The need for statistical research in relation to the experimentation dealing with conservation agriculture and cropping systems for enhanced and sustainable productivity is highlighted in the following.

4.1 Conservation Agriculture

Conservation agriculture (CA) is an improved management and application of three key principles— minimum soil disturbance, permanent soil cover and diversified crop associations and rotations (http://www.fao.org/ag/ca/doc/FAO_REOSA_Technical_Brief1.pdf). An important component of sustainable agriculture is conservation agriculture. CA stands for the cropping system which creates minimum soil disturbance, maintains the soil cover with organic matter and is associated with crop diversification—horizontally (same season) and/or vertically (over time).

A challenge in designing experiments for CA is to evaluate genetic and agronomic input factors, some of which may require plots of relatively larger sizes (main plots) while the other factors require plots of relatively smaller sizes. Often we practice split-plot design in complete block structure, while situation may require use of incomplete blocks at every level of the plot sizes.

4.2 Crop Productivity and Sustainability

Improved germplasm, finally released as varieties, are introduced in the crop production systems, either as monocrops, or intercrops, or in rotation with other crops and fallow land. Crop rotation and intercropping systems diversify the crops and make more effective use of the soil moisture and nutrients. Cropping systems research is concerned with developing suitable integration of crops and crop management systems which yield high productivity in the short term and contribute favourably to build stable resource base for sustainable production in the long run.

In a crop rotation, different crops are grown in a sequence on the same piece of land (*i.e.* one after the other) during the season appropriate to the component crops. Experimentation is undertaken to identify the rotation which provides better control over factors that limit crop production, such as pests, diseases and soil degradation. Therefore, crop rotation serves as an effective mechanism for sustainable crop production.

4.3 Crop Rotations Trials

In order to efficiently capture year-to-year variability in climatic profile, a requirement of long-term rotation trial (LRT) design is that each phase in the rotation appears each year. This results in multiple plots under the same rotation within a replicate. There may also be other agronomic input factors that may be specific to the component crops. **Challenge is to search for an efficient design, likely to be a new-named design but specific to the experiment along the following points:** 1) it is not known a-priori how many rotations would be needed to establish the differences in the average rotation responses, 2) while the experimental design should be resource optimal, it should also be flexible or robust enough to administer changes in treatments in future. The essential and basic considerations of issues on designs of long-term rotation trials can be found in Cochran (1939), Yates

(1949, 1954), Patterson (1964), and Cady and Mason (1964). **The frequently asked questions include termination of trials, changes in the cropping systems in on-going experiments and unavailability of data.**

Statistical analysis of crop rotation trials deals with partitioning of between plot totals and within plot-totals. While the routine analyses are carried out using either independent plot errors over years or with constant covariance, it is obvious that more is needed to model covariance error structures, particularly when the time series gets longer. Ignoring dependence of plot-errors is likely to affect precision and tests of significance of model parameters. Further, not only are individual plot errors correlated over time but there is also an underlying spatial variability in the field layout. Thus, the plot-error correlations should be evaluated across three directions- years, rows and columns of the layout and effort would be needed to model the covariances (Singh and Jones 2002, 2008, Singh and Pala 2004). **LRT data raise challenges for modelling of covariance structure, in time and space dimension, and association between productivity and sustainability parameters.**

While the differences between plot-totals provide information on the main effects and interactions of experimental factors (such as crop management methods), the differences within-plots give information on changes in the treatment effects and their interactions with time. **Thus, the challenge lies in assessing these effects and their interactions under the selected covariance structures of plot errors. For short-series of data over years, use of first order auto-correlations has been found useful in analysis of data from barley-legume rotations** (Singh and Jones 1997, Jones and Singh 2000) and Singh and Pala (2004) for three-course wheat rotation.

Another question that runs with every repeat of a LRT is: when should be the last cycle or year of the rotation? The insight on such a question may vary with the trial objective, *e.g.*, it could be based on detecting consistent significant differences in treatment effects in productivity, or estimating significant time-trend. In both the scenarios, we evaluate sustainability of the cropping system. If it takes too long to detect significant difference in the productivity or a time-trend, the system implies unsustainable or impractical. For the first order auto-regressive models for plot-errors over years or

cycles, with accountability of linear or quadratic contribution of rainfall, the number of cycles or years have been obtained at ICARDA (Singh and Jones 1997, Jones and Singh 2000) for continuous cropping and two-course rotations respectively. However, ICARDA results are for short series, but as a series gets longer, any other interesting features of the series need to be captured *e.g.* in terms of ARIMA (autoregressive integrated moving average) models for plot errors and incorporating seasonal variation in the weather variables. **Examining statistical properties, in terms of confidence limits, of the time or cycles to detect significant time-trend would be helpful in terminating or replacing a treatment or the trial itself.**

In crop rotation trials, statistical analysis of data on the test crop alone or the component crops one at a time is the common practice. The other issues involve assessing residual effects of the previous crop and joint modelling of all the crop yields, to account for multi-way interactions and crop correlations. **An effectively integrated model will provide a more realistic assessment of productivity and sustainability of the cropping systems.**

Furthermore, data with experimental details are the key resources to generate knowledge. They must be stored for possible sharing with potential users to enable them develop newer approaches to modelling and to provide insight for establishing new long-term trials in search of sustainable technology for the future. Therefore, it is also a challenge to statisticians to help design a database structure, identify the data variables at the plot level and at the environment level, set up tools for data verification and validation, storage, retrieval and reports. **The structured databases can preserve the intellectual property and memory of the institution while actual teams of scientists, which were involved in generating data, have in fact left the institution.**

4.4 Intercropping and Alley Cropping

Intercropping and alley cropping are another crop production mechanisms that are attributed to the sustainability of crop production due to their more effective resource sharing among the diverse crops. Experimental designs are needed to address competition or interaction between and within various species of trees, shrubs, crops or grasses. The issues involve taking

into account the nature of the crops and their treatments, bi-variate or multivariate nature of the response, and covariance structures. Developing suitable stability criteria is also needed in the context of multi-environment experiments (Singh and Gilliver 1988, Singh *et al.* 1988, Ghassali *et al.* 2011).

5. CROP ACTUARY

An appropriately chosen cropping system package for an agro-ecology could be highly productive and provide long-term sustainability. However, productivity depends on the rainfall and temperature distribution during the crop growing period. The unpredictability of rainfall, drought frequency, and frost severely affects the production/productivity of the system in the short run, and their repeated occurrences may further make a crop production package unsustainable, even resulting in suicides by some farmers. Using probabilistic, mathematical and financial theories and tools to predict risk of failures, actuaries develop insurance policies for the health and industry sectors (<http://www.actuariesindia.org/index.aspx>; <http://harvest.cals.ncsu.edu/ncars/index.cfm?pageID=5249&printable=1>; Raju and Chand 2008, Chen and Miranda 2008, Beach *et al.* 2010, Wang *et al.* 2010). **Statisticians have a major role in helping develop crop insurance policies for the agricultural production sector.**

5.1 Vedic Verses

The following invocations have appeared in *Atharva Veda* and *Śrīmad Bhāgavatam* in relation to barley crop which is known to sustain the harsh climatic conditions and has in turn been an existential factor for human and the animals by providing food and feed since time immemorial.

Atharva Veda VI, 142. *For increase of barley*

- “Spring up, become fair, be distended, O barley, with your own increase! Burst all vessels designed to contain you! May lightning not smite you in that place where we make our appeal to you” (*Atharva Veda* VI, 142, 1)
- “In response, divine barley, to our invocation, rise up there tall as the sky, inexhaustible as the boundless sea!” (*Atharva Veda* VI, 142, 2)

- “May those who tend you prove inexhaustible, inexhaustible their barns, inexhaustible those who offer you in sacrifice and those who consume you!” (Atharva Veda VI, 142, 3)

Atharva Veda VI, 79. *For abundance at home*

- “May the Lord of the clouds protect our stores, piled high in our homes!” (Atharva Veda VI, 79, 1)
- “May the Lord of the clouds give us vitality in our homes, granting goods and riches!” (Atharva Veda VI, 79, 2)

Śrīmad Bhāgavatam, Canto 11, Chapter 16: *The Lord's Opulence*, Mantra 21

“...and among the plants I am the ones bearing grains [barley].”

6. CONCLUSION

Food security is dependent on perpetual availability of diverse genetic resources from diverse agro-ecologies and hot-spots, suitable tools for genetic modification to fit the changing scenario of climate change, growing population vis-à-vis limited biophysical resources, implementing the crops in cropping system for high and sustainable productivity, and of course a mechanism to economically sustain should the weather gods miss to bestow their blessings. **This presentation highlights a limited number of statistical issues, with emphasis on design and analysis of experiments, which need to be addressed to support research conducted by agricultural scientists in various disciplines. These are some of the ways through which agricultural statisticians could make a significant contribution to the Food Security.**

ACKNOWLEDGEMENTS

I am grateful to Professor Sudhir Gupta, Northern Illinois University, DeKalb, USA and Dr. Shiv Agrawal, ICARDA, Rabat, Morocco for their critical reviews, suggestions to enhance the statistical information content and adding value to the overall presentation of an earlier version of the manuscript.

REFERENCES

- Allard, R.W. (1999). *Principles of Plant Breeding*. John Wiley and Sons, New York.
- Bari, A., Street, K., Mackay, M., Endresen, D.T.F., De Pauw, E. and Amri, A. (2012). Focused identification of germplasm strategy (FIGS) detects wheat stem rust resistance linked to environmental variables. *Genet. Resour. Crop Evol.*, **59**, 1465-1481.
- Beach *et al.* (2010). *Climate Change Impacts on Crop Insurance*. Diane Publishing Co., Darby, PA.
- Bennett, E. (1970). Tactics of plant exploration. In: *Genetic Resources in Plants – their Exploration and Conservation*. (Eds. Frankel, O.H. And Bennett, E.) Blackwell, Oxford, 157-1888.
- Brown, A.H.D. and Marshall, D.R. (1995). A basic sampling strategy: theory and practice, pp 76-91. In: Guarino, L., V. Ramanatha Rao and R. Reid (eds.) *Collecting Plant Genetic Diversity: Technical guidelines*. Wallingford, UK: CAB International.
- Cady, F.B. and Mason, D.D. (1964). Comparison of fertility treatments in a crop rotation experiment. *Agronomy J.*, **81**, 476-479.
- Chang, T.T. (1985). Collection of crop germplasm. *Iowa State J. Res.*, **59**, 349-364.
- Chen and Miranda (2008). Modeling Texas Dryland cotton yields, with application to crop insurance actuarial rating. *J. Agri. Appl. Eco.*, **40(1)**, 239-252.
- Cochran, W.G. (1939). Long-term agricultural experiments (with Discussion). *J. Roy. Statist. Soc.*, **6[Suppl.]**, 104-148.
- Crossa J, Campos G, de L, Pe´rez, P., *et al.* (2010). Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. *Genetics*, **186**, 713-724.
- Cullis, B.R., Smith, A.B. and Coombes, N.E. (2006). On the design of early generation variety trials with correlated data. *J. Agri. Bio. Environ. Statist.*, **11**, 381-393.
- Damania, A.B. (1987). Sampling cereal diversity in Morocco. *Plant Genet. Resour. Newsletter*, **72**, 29-30.
- Engels, J.M.M., Arora, R.K. and Guarino, L. (1995). An introduction to plant germplasm exploration and collecting: planning, methods and procedures, follow-up, pp 31-63. In: Guarino, L., V. Ramanatha Rao and R. Reid (eds.) *Collecting Plant Genetic Diversity: Technical guidelines*. Wallingford, UK: CAB International.

- Ghassali, F., Osman, A.E., Singh, M., Norton, B., Louhaichi, M. and Tiedeman, J. (2011). Potential use of Mediterranean saltbush (*Atriplex halimus* L.) in alley cropping in the low rainfall-cropping zone of northwest Syria. *Range Mgmt. and Agroforestry*, **32(1)**, 1-8.
- Guarino, L.V Ramanatha, Rao and Reid, R. (editors) (1995). *Collecting Plant Genetic Diversity: Technical Guidelines*. CAB International, Wallingford, UK
- ICARDA (2013). *Research to Action 3 - FIGS - Focused Identification of Germplasm Strategy*. ICARDA (International Center for Agricultural Research in the Dry Areas), Beirut, Lebanon.
- Jones, M.J. and Singh, M. (2000). Time-trends of yield in long-term trials. *Exp. Agric.*, **36**, 165-179.
- Kameswara Rao, N. and Bramel, P. J. (eds.) (2000). *Manual of Genebank Operations and Procedures*. Technical Manual no. 6. Patancheru 502 324, Andhra Pradesh. India: International Crops Research Institute for the Semi-Arid Tropics.
- Marshall, D.R. and Brown, A.H.D. (1975). Optimum sampling strategies in genetic conservation. In: Falk, D.A. and Holsinger, K.E. (eds.) *Conservation of Rare Plants: Biology and Genetics*. Oxford University Press, Oxford, UK.
- Meuwissen, T.H.E., Hayes, B.J. and Goddard, M.E. (2001). Prediction of total genetic value using genome wide dense marker maps. *Genetics*, **157**, 1819-1829.
- Nakaya, A. and Isobe, S.N. (2012). Will genomic selection be a practical method for plant breeding? *Ann. Botany*, **110**, 1303-1316. doi:10.1093/aob/mcs109.
- Nelson, M.F. and Anderson, N.O. (2013). How many marker loci are necessary? Analysis of dominant marker data sets using two popular population genetic algorithms. *Ecol. Evol.*, **3(10)**, 3455-3470. doi: 10.1002/ece3.725.
- [www.actuariesindia.org/.../CROP% 20INSURANCE% 20PRODUCT.docx](http://www.actuariesindia.org/.../CROP%20INSURANCE%20PRODUCT.docx)
- <http://harvest.cals.ncsu.edu/ncars/index.cfm? pageID= 5249 &printable=1>
- Rajaram, S., van Ginkel, M. and Fischer, R.A. (1995). CIMMYT's wheat breeding mega-environments (ME). *Proceedings of the 8th International Wheat Genetics Symposium*, Jul. 19-24 1993. Beijing, China (<http://www.fao.org/docrep/006/y4011e/y4011e09.htm>)
- Raju, S.S. and Chand, R. (2008). *Agricultural Insurance in India Problems and Prospects*. NCAP Working Paper No. 8. National Centre for Agricultural Economics and Policy Research (Indian Council of Agricultural Research), New Delhi.
- Segura, V., Vilhjálmsson, B.J., Platt, A., Korte, A., Seren, Ü., Long, Q., and Nordborg, M. (2012). An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. *Nature Genetics*, **44**, 825-830. doi:10.1038/ng.2314
- Singh, M. and Chabane, K. (2008). Optimum sample size for clustering wild wheat populations using AFLP markers. *Stat. Appl.*, **6**, 159-172.
- Singh, M., Chabane, K., Valkoun, J. and Blake, T. (2006). Optimum sample size for estimating gene diversity in wild wheat using AFLP markers. *Genet. Resour. Crop Evol.*, **53(1)**, 23-33.
- Singh, M., Chaubey, Y.P., Sarker A. and Sen, D. (2010). Modelling unstructured heterogeneity along with spatially correlated errors in field trials. *J. Ind. Soc. Agril. Statist.*, **64(2)**, 313-321.
- Singh, M., Gilliver, B. and Rao, M. R. (1988). Stability of Genotype in intercropping. *Biometrics*, **44**, 561-570.
- Singh, M. and Jones, M.J. (1997). Estimating time to detect time-trends in continuous cropping. *J. Appl. Statist.*, **24**, 659-669.
- Singh, M. and Jones, M.J. (2002). Modelling yield sustainability of different rotations in long-term barley trials. *J. Agric., Bio. Environ. Statist.*, **7(4)**, 525-535.
- Singh, M. and Jones, M. (2008). Modelling spatial-temporal covariance structures in monocropping barley trials. *J. Appl. Statist.*, **35**, 321-333.
- Singh, M. and Pala, M. (2004). Use of covariance structures for temporal errors in the analysis of a three-course wheat rotation and tillage trial. *J. Agric. Sci., Cambridge*, **142**, 193-201.
- Singh, M., Tadesse, W., Sarker, A., Maalouf, F., Imtiaz, M., Capettini, F., and Nachit, M. (2013). Capturing the heterogeneity of the error variances of a group of genotypes in crop cultivar trials. *Crop Sci.*, **53(3)**, 811-818. doi: 10.2135/cropsci2012.11.0637.
- Singh, M., van Ginkel, M., Sarker, A., Malhotra, R.S., Imtiaz, M. and Kumar, S. (2012). Increasing precision of even otherwise well-run trials by capturing heterogeneity of plot error variances. *Agric. Res.*, **1**, 285-294.
- Patterson, H.D. (1964). The theory of cyclic rotation experiments (with Discussion). *J. Roy. Statist. Soc.*, **B26**, 1-45.

- Upadhyaya, H.D. and Ortiz, R. (2001). A mini core subset for capturing diversity and promoting utilization of chickpea genetic resources in crop improvement. *Theo. Appl. Genet.*, **102**, 1292-1298.
- Upadhyaya, H.D., Yadav, D.N., Gowda, C.L.L. and Singh, S. (2010). Mini core germplasm collections for infusing genetic diversity in plant breeding programs. *Elect. J. Plant Breed.*, **1(4)**, 1294-1309 (July 2010)
- Wang, E., Yua, Y., Little, B.B., and Li, Z. (2010). Crop Insurance Premium Design Based on Survival Analysis Model. *Agriculture and Agricultural Science Procedia* **1**, 67-75.
- Weir, B.S. (1990). *Genetic Data Analysis*. Sinauer Associate Inc., Sunderland.
- WRI, IUCN and UNEP (1992). Global Biodiversity Strategy. Guidelines for Action to Save, Study and Use Earth's Biotic Wealth Sustainability. WRI, IUCN and UNEP, Washington DC.
- Wricke, G. and Weber, E. (1986). Quantitative Genetics and Selection in Plant Breeding. Walter de Gruyter and Co., Berlin, Germany.
- World wide web references for CBD, Agenda 21.
<http://www.iisd.org/rio+5/agenda/agenda21.htm> (accessed 31 October 2013).
<http://www.cbd.int/convention/> (accessed 31 October 2013).
- Yates, F. (1949). The design of rotation experiments. In: Technical Communication No 46, *Proceedings of the First Commonwealth Conference on Tropical and Sub-Tropical Soils*, 1948, 142-155. Harpenden, England: Commonwealth Bureau of Soil Science.
- Yates, F. (1954). The analysis of experiments containing different crop rotations. *Biometrics*, **10**, 324-346.