Dr. Rajendra Prasad Memorial Lecture	Dr. VG Panse Memorial Lecture
Dr. GP Samanta Secretary, MOSPI and Chief Statistician of India Government of India, New Delhi **Topic:** Statistics, AI/ML and Big Data Analystics	Dr. RC Agrawal DDG (Agricultural Education), ICAR, New Delhi **Topic:** Paradigm Shift in Agricultural Education to meet Agriculture Revolution 4

IP 01: Dr. Daroga Singh Centenary Celebration Session
Chairperson: Dr. P.S. Pandey, Vice Chancellor, DRPCAU, PUSA, Samastipur
Convener: Dr. B.V.S. Sisodia
Speakers:

1. **Dr. Padam Singh**
   Statistics for Sustainable Development

2. **Dr. Murari Singh (ONLINE)**
   A Short Review on Bayesian Estimation of a Common Coefficient of Variation from Inverse Gaussian Distributions

3. **Dr. Pranesh Kumar (ONLINE)**
   Neutrosophic Single Factor Experimental Designs and Neutrosophic Data Analysis

IP 02: Dr. Lalmohan Bhar Memorial Session
Chairperson: Dr. RC Agrawal, DDG (Agricultural Education), ICAR, New Delhi
Conveners: Dr. B.N. Mandal/ Dr. Ranjit Kumar Paul

Speakers:

4. **Dr. A. Dhandapani* (ONLINE)**
   Hadamard R–R package to generate Hadamard Matrices

1. **Dr. Ananta Sarkar**
   Gender Disaggregated Data in Agriculture and Data Issues

5. **Dr. Anurup Majumder* (ONLINE)**
   Outliers: A Review on Detection of Outliers in Field Experiments and Some Case Studies

TEA

IP 03: Dr. Hukum Chandra Memorial Session
Chairperson: Dr. Tauqueer Ahmad and Dr. S.A. Mir
Conveners: Dr. Dalip Singh/Dr. Sanghamitra Pal

Speakers:

1. **Dr. Girish Chandra**
   Weighted Ranked Set Sampling for Skew Distributions

2. **Dr. Santanu Pramanik**
   Statistical validation of a large-scale web survey during the COVID-19 pandemic in India

3. **Dr. Kaustav Aditya**
   Food and Nutrition in Indo Gangetic Plain Region -A Disaggregate Level Analysis

4. **Dr. Saurav Guha**
   Localised Estimates and Spatial Mapping of Income Inequality: Evidence from Labour Force Survey Data of India

IP 04: Big Data Analytics, Machine Learning, Artificial Intelligence and their Applications in Agriculture (01)
Chairperson: Dr. Anil Rai
Speakers:

1. **Dr. S Ravichandran**
   Data Science in Indian Agriculture - Scope, Status and the Road Ahead

6. **Dr. Alka Arora (ONLINE)**
   Machine Learning based Approach for Measuring Senescence in Wheat Crop

2. **D. Mukesh Kumar**
   Mobile Applications for Dissemination of Knowledge in Livestock Farming

IP 04: Big Data Analytics, Machine Learning, Artificial Intelligence and their Applications in Agriculture (02)
Chairperson: Dr. A.R. Rao
Speakers:

3. **Dr. A.R. Rao**
   Integration of Artificial Intelligence (AI) and –omics for Crop and Animal Improvement

7. **Dr. Amrit Kaur Mahal* (ONLINE)**
   Role of Statistical Software in Data Analysis

6. **Dr. Shakeel Ahmad Mir**
   Artificial Intelligence –Importance and Application in Agriculture

8. **Dr. K.K. Chaturvedi* (ONLINE)**
   Significance of Big Data's vs. in Digital Agriculture

CONCURRENT TECHNICAL SESSIONS

	……Hall	**…. Hall**
1100-1300	**IP 05: Online Agricultural Education and Extension** **Chairperson: Dr. Padam Singh/Dean, SKUASTK**	**IP 06: Statistical Modelling and Forecasting in Agriculture (01)** **Chairperson: Dr. K.K. Tyagi**
	Speakers: **Dr. Sudeep** Modernizing Agriculture Education through IT interventions- Steps Undertaken **Dr. Rajni Jain (ONLINE)** Factors affecting ICT Usage in Agricultural Higher Education	**Speakers:** **Dr. K.N. Singh** Drought Prediction using Machine Learning Technique **Dr. Ramasubramanian V** Development of classification tree enhanced by genetic algorithm for forecasting in agricultural ergonomics **Dr. Girish K Jha**

	Dr. Anshu Bharadwaj* (ONLINE) Virtual Classroom and Agri-DIKSHA: Embracing the future of Digital Learning in Agriculture Higher Education **Dr. Shashi Dahiya (ONLINE)** Online Initiative for Ranking of Green and Clean Agricultural University Campus	AI-enabled Price Forecasting Model for Agricultural Commodities **IP 06: Statistical Modelling and Forecasting in Agriculture (02)** **Chairperson: Dr. S. Ravichandran and Dr. Radha A Ashrit (ONLINE)** **Dr. J Jayasankar** Marine Fishery Resource Modeling- Trends, Travails and Talismanic tools and Takeaways **Dr. DK Panda** Spatio-temporal Assessment of Water Resources in India under Changing Climate: Application of Statistical Tools **Dr. Ranjit Kumar Paul** Decomposition based Machine Learning Techniques for Forecasting Agricultural Prices
1300-1400	**LUNCH**	
1400-1530	**IP 07: Basic Research in Statistical Sciences (01)** **Chairperson: /Dr. J. Jayasankar** **Speakers:** **Dr. Radha A Ashrit (ONLINE)** Advances in agricultural knowledge and the use of innovative approaches among farmers in four different districts of Andhra Pradesh, India - A baseline and endline biotech intervention assessment conducted as part of the Biotech KISAN Mission **Dr. M.R. Verma** Allocation Proportional to Strata Total and Exponential Phase Effect under Cost Constraints **Dr. Rajesh Kumar** Efficiency of Statistical Design at Advance Level of Sugarcane Varietal Trials **Dr. Cini Varghese* (ONLINE)** 2-part Designs for Agricultural Research **IP 07: Basic Research in Statistical Sciences (02)** **Chairperson: Dr. Sanjeev Panwar** **Speakers:** **Dr. Sheela Misra** Emerging Need of Composite Indices of Evaluation with respect to HDI and SDG **Dr. BK Hooda* (ONLINE)** Genetic Algorithm Based Cluster Analysis	**IP 08: Remote Sensing, UAV, IoT and GIS and their Applications in Crop Yield Estimation in Agriculture** **Chairperson: Dr. B.V.S. Sisodia** **Speakers:** **Dr. Tauqueer Ahmad** **Dr. B Sailaja** Smart Precision Models for Rice Yield Estimation **Dr. V. Bhushana Babu** IoT Application for Analysis and Estimation of Agricultural Accidents Survey Data

	Dr. Shashi Bhushan On Some Improved Imputation Methods Under MCAR Approach **Dr. Showqat Maqbool** Estimation of Finite Population Mean under Predictive Modeling Approach	
1530-1700		

* Online presentation

Abstracts of Invited Papers

I.P. 1: Dr. Daroga Singh Centenary Celebration Session

Statistics and Sustainable Development Goals

Padam Singh

Medanta Institute of Education and Research, Sector 38, Gurugram

India is a signatory to goals set by United Nations. Among 8 Millennium Development Goals (MDGs) the first goal was "Eradicate extreme poverty and hunger". Similarly, among 17 Sustainable Development Goals (SDGs) the first two are "No poverty" and "Zero Hunger".

Global Hunger Index (GHI) is calculated and disseminated annually since 2006. India is ranked poorly below Pakistan, Bangladesh and Nepal based on this index. This is contrary to the fact that India ranks 5th in world economy. Further, India is one of the top 5 food producing countries in the world.

To deliberate on this issue ICMR constituted a Technical Committee on Global Hunger Index. A white paper entitled "Global Hunger Index does not really measure hunger - An Indian perspective" based on the recommendations of the committee has been published in Indian Journal of Medical research. Further, the findings of the committee "Global Hunger Index is a misnomer" has been accepted by NITI Aayog.

The presentation will focus on choice of indicators used and statistical/ methodological issues involved in measuring hunger.

A Short Review on Bayesian Estimation of a Common Coefficient of Variation from Inverse Gaussian Distributions

Murari Singh, Yogendra P Choubey, Debaraj Sen and Ashutosh Sarker
University of Toronto, Department of Statistical Sciences, Toronto, Canada

Coefficient of variation (CV) has found several applications in diverse contexts, for example in agricultural experimentation, it is used to measure field heterogeneity and stability of a genotype across environments. Statistical inference methods for CV, from frequentist and Bayesian perspectives, have been widely studied in theory and practice for distributions where the variable of interest is non-negative. In this presentation, I will review three papers jointly prepared with Professors YP Chaubey and D. Sen, Concordia University, and Dr. A. Sarker, ICARDA. The review includes papers on: 1) estimation of a common CV from inverse Gaussian distributions and testing of hypotheses under frequentist approach, 2) exact and simulation based results on

posterior mean and HPD intervals for the CV under the two sets of priors, and 3) posterior distribution of a common CV for a set of three selected lentil genotypes based on prior information obtained from multi-environment trials data on the remaining genotypes of the trials. The results of these papers will be highlighted. Exploring the inverse Gaussian distribution in routine analysis of agronomic traits and comparison with standard normal distribution based analysis will be emphasized.

Neutrosophic Single Factor Experimental Designs and Neutrosophic Data Analysis

Pranesh Kumar
Department of Mathematics and Statistics, University of Northern British Columbia. Prince George, BC V2N 4Z9, Canada
Email: Pranesh.kumar@unbc.ca

We collect data in studies, for example, in the fields of agriculture, health, clinical trials, biology, marketing, and industrial production, by conducting either experiments or by observing it in surveys. Experimental design and analysis is one of the major branches of statistics that deals with the designing experiments statistically and analyzing data from these experiments. However, it is not uncommon to see sometime that data values are vague, incomplete, or imprecise for known or unknown reasons. To analyze data set with such imprecise observations, researchers have proposed various emerging approaches such as fuzzy, intuitionistic fuzzy and neutrosophic logic and analysis, which provide better understanding and interpretation of the data set. Neutrosophic logic is an extension of fuzzy logic where a variable x is described by triplet values, i.e.,$x = (t, i, f)$, where t is the degree of "truth", f is the degree of "false" and i is the level of "indeterminacy". A neutrosophic data$x$ can be expressed as$x = d + i$, where d is the determinate (sure) part of x, and i is the indeterminate (unsure) part of x. In this talk, we consider three of the widely used experimental designs, i.e., the completely randomized design, the randomized block design, and the factorial design, and data analysis using the neutrosophic statistics.

IP2: Dr. Lalmohan Bhar Memorial Session

Hadamard R – R package to generate Hadamard Matrices

Appavoo Dhandapani
ICAR-National Academy of Agricultural Research Management, Hyderabad

Late Dr. Lalmohan Bhar, a close friend, was highly interested in utilizing the results from Matrix Algebra to solve problems in the area of Outliers and Robustness in Design of Experiments. I feel, it would be a tribute to the departed soul to talk on Hadamard Matrices and their generation.

Hadamard Matrices was introduced by Sylvester in 1867 before Hadamard in 1893 studied in detail. An Hadamard matrix, H, is a square matrix with entries 1 or -1 with the property that HH'

= nI, where H' is the transpose of H and I is the identity of matrix and n is the order of the matrix H. It is known that such matrices exist if n = 1, 2 or a multiple of 4. However, whether H exists for all multiples of 4 is still an open problem. The smallest order for which there is no construction method available is 668. Hadamard Matrices are useful in many fields such as construction of symmetric Balanced Incomplete Block Designs, Group Divisible Designs, fractional factorial designs, asymmetrical factorial designs, supersaturated designs etc. Another important area of use of Hadamard matrices is in constructing half-replicates in Balanced Repeated Replications. Hadamard Matrices are also used in Telecommunications, Error Correcting Codes etc.

The construction problem of Hadamard Matrices is studied mainly by mathematicians and the users are mostly statisticians and engineers. There are several construction methods proposed in the literature, each creating same or different orders of Hadamard Matrices. There were several attempts to provide tools for generating Hadamard Matrices in software such as SAS, Matlab and even in R. However, there is no single tool which can generate Hadamard Matrices of different orders.

HadamardR is a R package, available in CRAN (https://cran.r-project.org/web/packages/HadamardR) which is useful in generation of Hadamard Matrices. The R Package can Hadamard Matrices of order < 5000. Out of 1250 possible Hadamard Matrices of order < 5000 (ignoring trivial orders 1 and 2), the construction methods are not known for 45 orders; 1158 orders are possible using Hadamard_Matrix() function. Only 47 Hadamard matrices not available in this package as of now. The package provides convenient way of generating Hadamard Matrices by either specifying the order required or by specifying the order and method.

In this talk, various features available in the package will be explained and how it can be used for getting addition/multiplication tables under Galois Fields for prime and prime power and their possible uses in teaching/other construction methods in Design of Experiments.

Gender disaggregated data in agriculture and data issues

Ananta Sarkar, Anil Kumar, Neetish Kumar, Shaji, A. and Tanuja S.
ICAR-Central Institute for Women in Agriculture, Bhubaneswar

Understanding gender related issues and constraints for making gender based policies in agriculture lack sufficient gender disaggregated data as need based comprehensive information on important indicators related to men and women in the country is not available at one place. Major data sources for the same include Census, Ag-census, MOSPI, AISHE, NFHS etc. which provides information on some indicators related to gender and agriculture. List of important indicators include work participation as workers, cultivators and agricultural labourers, operated land holdings, male and female headed households, literacy, enrolment of students in agriculture, health & nutrition, income, wage rate, scientific staff in ICAR, SHGs, migration, and selected

empowerment indicators at national level. Synchronized information on various dimensions of gender in agriculture/ socio-economic indicators are not readily available. For example, gender disaggregated data on fisheries sector is missing from the census. This leads to misrepresentation or under representation of contribution of women in fisheries. Further, the information on above indicators also vary with time of survey. Efforts made by the ICAR-CIWA, Bhubaneswar attempted to bring the available information in its Gender Knowledge System (GKS) in Agriculture portal. Comparative scenario on selected indictors has been built using numbers, percentages, ratios and gender gaps and depicted through tables, charts and graphs. Gender gap has narrowed down over the years for few indicators, considerable gender gap and data gap still exist in respect of different socio-economic indicators of development. Women in India contribute significantly in agriculture and at the same time, they face numerous gender based constraints, categorization of activities in which women are involved need to be penned down to further collect data on the obstacle or constraints faced by them in their occupation. Therefore, the Institute is focusing on generating gender disaggregated data in key areas of their socio-economic involvement for effective planning of gender based strategies/ policies. Assessing gender mainstreaming of activities of agricultural institutions using concept of gender factsheets and issues related to gender disaggregated data have been discussed.

Construction of Patterson Type Balanced and Partially Balanced Cross over Designs

Anurup Majumder, Moupiya Roy and Md M. Hasan Middya
Bidhan Chandra Krishi Viswavidyalaya, Nadia,

When the experimental units in an experiment are human beings or animals, mere grouping or blocking of those units is not sufficient for controlling or reducing the error variation. Such units are used in the experiment in such a manner that they are repeatedly exposed to a sequence of different treatments over times or periods. Such designs are known as repeated measurement designs or crossover designs (COD). Pioneer work on such designs had done by Patterson (1952) in Biometrika. Several binary balanced (direct and carryover) and universal optimal crossover designs are developed and listed in the article.

In the present article, some new balanced (direct and carryover), binary and universal optimal (direct and carryover) crossover designs are developed for the parameters v, n, p e.g. (4,12,3); (7,21,3); (7,42,3); (6,30,3) and (6,30,5). The designs are new in the sense that they are not yet listed in available literature. The properties of balanced and universal optimality of the above designs are established by C matrices of direct and carryover effects.

Availability of Balanced Cross over designs is restricted due to parametric relations of the designs. The concept of Partially Balanced Cross over Designs (PBCOD) was introduced by Blaisdell and Raghavrao (1980). Present article also provides some new PBCODs based on two associate class group divisible (GD) and cyclic PBIB designs. The developed designs along with efficiency factors are listed in the article.

IP 03: Dr. Hukum Chandra Memorial Session

Weighted Ranked Set Sampling For Skew Distributions

Dinesh S. Bhoj[1] and Girish Chandra[2]
[1]Rutgers University, Camden, New Jersey, USA
[2]Indian Council of Forestry Research and Education, Dehradun

Ranked Set Sampling (RSS) is a useful technique for improving the estimator of population mean when the sampling units in a study can be easily ranked than the actual measurement. RSS performs better than simple random sampling (SRS) when the mean of units corresponding to each rank is used. The performance of RSS can be increased further by assigning weights to the ranked observations. In this paper, we propose weighted RSS procedures to estimate the population mean of positively skew distributions. It is shown that the gain in the relative precisions of the population mean for chosen distributions are uniformly higher than those based on RSS. The gains in relative precisions are substantially high. Further, the relative precisions of our estimator are slightly higher than the ones based on Neyman's optimal allocation model for small sample sizes. Moreover, it is shown that, the performance of the proposed estimator increases as the skewness increases by using the example of lognormal family of distributions.

Statistical validation of a large-scale web survey during the COVID-19 pandemic in India

Santanu Pramanik
National Council of Applied Economic Research,

Face-to-face surveys as a method of data collection were commonplace prior to the COVID-19 crisis, particularly in developing countries. However, fear of contracting infection and non-pharmaceutical interventions such as physical distancing and mobility restrictions to contain the spread of infection made it infeasible to continue data collection with this mode during the pandemic. At the same time, there was an overwhelming demand for data to respond to economic and health emergencies, this forced the remote modes of data collection such as mobile and web surveys to come to the forefront. The primary concerns with remote mode surveys are undercoverage of target population and self-selection of the survey respondents resulting in biased estimates. Using data from COVID-19 Trends and Impact Survey (CTIS) from India, the largest public health web survey, we examine the bias in the estimates of vaccine uptake, a population measure which changes rapidly with time. Our findings show that the bias in CTIS estimates of vaccine uptake is not constant over time, rather it increases up to a certain point of time and then decreases. Our findings are explained by the fact that the variability in the outcome of interest in the population first increases with time and then goes downward after more than 50% of the population are vaccinated. The validation of CTIS vaccine uptake estimates was possible as it is one of the rare situations where reliable gold standard was available. For another key indicator from CTIS, COVID-like illness (CLI) constructed based on

self-reporting of symptoms, it is not trivial to assess the bias in the outcome as the quality of gold standard is questionable. Since absence of independently verified 'ground truth' or 'gold standard' for assessing bias in surveys is well acknowledged, we discuss the need for statistical representativeness of web surveys with respect to key demographic characteristics of respondents which are often correlated with many outcome variables.

Food and Nutrition in Indo Gangetic Plain Region - A Disaggregate Level Analysis

Hukum Chandra, Kaustav Aditya*, Swati Gupta, Saurav Guha and Bhanu Verma
ICAR-Indian Agricultural Statistics Research Institute, New Delhi

India is agriculture based country and has experienced an enormous change in food and nutrition utilization design since the financial change in mid 1990s. Agribusiness is considered as the backbone for Indian economy, therefore Indo Gangetic Plain (IGP) holds vast agricultural importance contributing to major portion to our national income. The welfare of an expanding economy are not shared equally as the country is still home to one-third of the world's poor. Hunger in India is considered as a genuine imprint on its development and food security has now evolved as a principal issue. Presently, interest in agriculture, nutrition, and dietary security is a prime worry for the country to accomplish the target of enrichment. An expansive area of Indian population is experiencing lack of healthy sustenance and deficiency of nourishment grains. This paper demonstrates nourishment utilization design across selected social and economic groups in the states coming under IGP region of India which includes West Bengal, Bihar, Uttar Pradesh Punjab and Haryana. The analysis helps in distinguishing the disparities among calorie, protein and fat consumption in IGP region. An attempt has also been made in recognizing socio-economic groups suffering from deficiencies in nutrition consumption.

Localised Estimates and Spatial Mapping of Income Inequality: Evidence from Labour Force Survey Data of India

Saurav Guha
ICAR-Indian Agricultural Statistics Research Institute, New Delhi

The economy of India is growing continuously with its gross domestic product increasing rapidly than most of the developing countries. Nonetheless, an increase in national gross domestic product does not reveal the income parity at micro level in the country. The income inequality in India has unfavorably obstructed underprivileged in accessing elementary needs like health and education. Periodic labour force survey conducted by National Statistical Office of India generates estimates on income status at national and state level for both rural and urban sectors separately. However, due to small sample size problem, these surveys cannot generate reliable estimates at micro-level viz. district or block. Thus, owing to unavailability of district-level estimates, analysis of income inequality is restricted to the national and the state level. Therefore, the existing variability in disaggregate-level income distribution often goes unnoticed. This article describes multivariate small area estimation method to generate precise and representative district-level estimate of income distribution in rural and urban areas of the Indian State of Bihar by linking Periodic labour force survey data of 2018-19 and 2011 Population Census data of

India. Theil index is used to measure the degree of income inequality by district in rural and urban areas. These disaggregate-level estimates and spatial mapping of income inequality are essential for measuring and monitoring the goal of reduced inequalities related to the sustainable development of 2030 agenda.
They expected to offer key insights to decision-makers and policy experts for identifying the areas demanding more attention.

IP 04(A): Big Data Analytics, Machine Learning, Artificial Intelligence and their Applications in Agriculture

Data Science in Indian Agriculture – Scope, Status and the Road Ahead

S.Ravichandran
ICAR-National Academy of Agricultural Research Management, Hyderabad

Agriculture, the backbone of Indian economy is highly dependent on climate, soil, irrigation, cultivation, harvesting, pesticides, rainfall, and many other factors, viz. crop area, good soil, temperature, and related data collected and maintained by various agencies. However, India grapples with many challenges associated with agriculture, such as climate change, extreme weather conditions like floods and droughts, scarcity of groundwater resources, etc. since agriculture is practiced in the entire country by cultivating all crops. India's agriculture contributes to only 16% of the GDP of the country whereas nearly 40% of the workforce is involved in agriculture either directly or indirectly out of the 70% of the rural population. Indian agriculture still depends upon the human labour for most of the agricultural activities when compared to the developing countries. Contribution of agriculture to the GDP of the developing economies is also quite high. World bank group developed a country classification involving five stages based on the role of agriculture in GDP and employment. Based on the study, country segmentation has five stages for associated characteristics. Most of the developed nations such as USA, UK, France, Germany, etc. employ less than 10% of total employment in agriculture and the GDP contribution is also less than 10% whereas more Indian population is engaged in agriculture to generate more GDP and hence India should concentrate more on agricultural GDP contribution by reducing employment in agriculture. This is possible only by way of mechanization and by making use of the latest Internet of things (IOT) enabled farm operations. The present paper concentrates on the use of modern technologies such as applications of data science, a powerful analytical tool for achieving higher productivity by engaging very few people in farm-related activities, which can transform agriculture, and will ultimately transform the Indian economy. Growing expectation is also that stakeholders are realizing how crucial, data is to the economy. But India, still a long, long way from reaching that stage because most of the data collected from Indian agriculture is highly unstructured and difficult to analyze to find patterns.

Machine Learning based Approach for Measuring Senescence in Wheat Crop

Alka Arora and Mohit Kumar
ICAR-Indian Agricultural Statistics Research Institute, New Delhi

Leaf senescence is an integral response of leaf cells to the regular ageing process but also due to unfavourable conditions. Many physiological, biochemical, and molecular studies of leaf senescence have shown that during senescence, leaf cell sunder go highly coordinated changes in cell structure, metabolism and gene expression. The earliest and most significant change is the breakdown of chloroplasts; leaf senescence leads to the degradation of photosynthetic pigments such as chlorophyll, with the degradation manifested in observable leaf colour changes from the usual deep green to pale green, to yellow and finally to brown.

Current methods for phenotypically measuring senescence are completely visual and labour-intensive. Using high definition, high resolution RGB images, one could attribute plant image colour into a few categories of senescence and green portion. Based on pixel values of each category, it is possible to train machine learning algorithms that can classify each pixel belongs to either green or senescence portion. Once pixels are classified, senescence portion is segmented from green portion and they can be quantified. In this study wheat plant image data was taken from Nanaji Deshmukh Plant Phenomics Centre ICAR-IARI. Machine Learning based algorithms were trained on five specified classes; light green, dark green, yellow, pale yellow and finally brown for wheat plant. Among all the trained classifiers, ANN outperformed with 97.28% test accuracy.

Open-Source Big Data Databases

S.B. Lal, K. K. Chaturvedi, Anu Sharma and Md. Samir Farooqi
ICAR-Indian agricultural Statistics research Institute

The current trend of rapid growth in data generation has become the global workforce. It is very necessary to transform vast amounts of organized and unstructured data into insightful and useful knowledge for progress in different sectors. As a result, there are many different big data tools to handle and store data available on the present global market. Data has no relevance until it is transformed into knowledge and information that can help in managing the business. Big data innovation provides a business with a limitless array of features that deal with insight and forecasting to maximize efficiency and save costs. International Data Corporation has predicted that by the year 2025, the amount of data created worldwide is going to exceed 163 Zettabytes. This will be more than ten times as much data generated at present. Open-source database solutions are widely used by commercial organizations all over the world for handling and analyzing their data. Due to open source's adaptability and the opportunity to contribute to the platform's evolution, many organizations are showing inclination towards open source tools. Because of fast pace of changing world, a firm must make significant investments in data analytics. Businesses all around the world now have a platform to create new database models employing large-scale analytics, thanks to the fast-paced proliferation of information and technological improvements. Around the world, artificial intelligence is at the forefront of significant innovation. Few important Open-Source Big Data Databases that are being used by the various organizations these days will be discussed.

Mobile Applications for Dissemination of Knowledge in Livestock Farming

Mukesh Kumar, Soumen Pal and Sudeep
ICAR-Indian Agricultural Statistics Research Institute, New Delhi

Livestock is a key source of income for landless and marginal farmers and plays a major role for upliftment of rural economy which contributes significantly towards total economy of the nation. To enhance the growth of rural income and higher public investment in livestock farming, the farmers need to adopt latest technology and animal husbandry management practices. Livestock occupation is not a linear system and it requires constant inputs at every stage related to livestock health and animal reproduction. Up to date, relevant and quality information on livestock is crucial for efficient management and development of the livestock sector.To bridge the information gap between the farmer/Livestock owner and livestock production technology, IT interventions are required. Mobile technology, in present era, have created new channels to communicate with various stakeholders in a well-managed way. Government is also taking steps and encouraging development of mobile applications which can enable the farmers across the country to get information on time. A number of mobile apps viz. Animal Reproduction, Pig farming, Vaccination Guide, Dairy Manager, Artificial Insemination, Pig Ration which are related to livestock farming have been developed at ICAR-IASRI in collaboration with ICAR-IVRI. The mobile apps are available in Google Play Store and Mobile Gallery of ICAR KRISHI Portal freely. These apps have transformed the information sharing in many ways in livestock sector. The present paper highlights on these mobile apps which provide desired knowledge and skills to Graduating Veterinarians, Field Veterinary Officers, Developmental Organisations and Entrepreneurs for promoting livestock farming.

IP 05: Online Agricultural Education and Extension

Modernizing Agriculture Education through IT interventions- Steps Undertaken

Sudeep, Alka Aroa, Anshu Bharadwaj, Shashi Dhaiya, S.N. Islam, Chandan Kumar Deb, Ashraful Haque, Sanchita Naha

ICAR-Indian Agricultural Statistics Research Institute

Indian Council of Agricultural Research (ICAR) - Indian Agriculture Statistical Research Institute (IASRI) aims at creating state-of-the-art digital infrastructure and various IT initiatives to digitally transform the knowledge dissemination, management and governance of the

agriculture universities as well as the accreditation and ranking of these universities by ensuring digital services, digital access, digital inclusion, digital empowerment and by bridging the digital divide. The Division of Computer Application, ICAR-IASRI is engaged in designing, developing and implementing various systems under the guidance of Agriculture Education Division from last three decades. Over the years, very fast and drastic changes in digital technologies had happened and the division is engaged in transforming these systems as per the need of the hour. In this paper, various digital initiatives to strengthen the agriculture ecosystem are described as below:

1. **Academic Management System (AMS):** AMS is developed to automate all the academic activities in the Agricultural Universities and enhance the efficiency of the overall academic system by reducing time and efforts involved in manual processes. The system is functional over 55 Ag. Universities and is leading the digital transformation of the agriculture universities. The system learns the best practices from the implementing universities and enables other universities to adopt these best practices. One common academic system across all agricultural universities makes the dream of standardization of academic processes across universities, a reality. This also helps in implementation of NEP 2020 and empowers the universities to implement student and faculty exchange with lot more ease keeping all the student and faculty records unambiguous and correct.
2. **Agricultural Universities- Project Information Management System (AU-PIMS)** : AU-PIMS is developed to manage research projects in Agricultural Universities by creating, identifying, collecting, organising, sharing, adapting and using project information. It essentially creates a repository of research projects along with the results and findings thereby enabling all researchers across universities to search and review these projects resulting in better future research proposals as well as reducing duplicity of research efforts.
3. **Agriculture University Ranking System (AURS):** AURS is developed to capture information from the Agricultural Universities on the required parameters such as teaching resources, outcomes, research, extension, outreach and peer recognition thereby, generating rankings of the agricultural universities across India. The system makes the ranking more transparent, robust and fast by capturing most of the data from the source and existing systems.
4. **Accreditation Portal for Higher Agricultural Educational Institutions:** Accreditation Portal is developed with a vision to strengthen the agricultural education system by allowing both government and private Higher Agricultural Educational Institutions (HAEIs) to apply for ICAR accreditation. The system increases the speed of accreditation process thereby providing timely accreditation in both new as well as renewal of accreditations.
5. **Krishi Vishwavidyalaya Chhatr Alumni Network (KVC ALNET)**: KVC ALNET is developed to provide an exclusive online platform to connect all the alumnus of

Agricultural Universities across India. All alumni can use this system to effectively to connect with the existing students and university officials.

6. **E- Learning Portal:** E-Learning portal is developed with an objective to strengthen the Agricultural Higher Education in India by developing and disseminating the e-courses for undergraduate and postgraduate Agricultural Courses. Multiple calls were issued to faculty of agriculture universities to engage them in e-content development process in collaborative way. The system allowed one course to be developed by many faculty members who are affiliated to multiple universities.
7. **Education Portal:** Education portal provides vital information about all the Agricultural Universities across the country at a central platform. This has been developed as an initiative of ICAR under the Digital India Programme to provide single window interface to all students and faculty. The features like Unique Student ID and Unique Faculty ID enables the authentication of the registered students and faculty thereby providing a unifying mechanism across all portals related to agriculture education.
8. **Krishi Megh:** A state-of-the-art IT infrastructure is established to strengthen the overall Agricultural Higher Education System ensuring high availability cloud ready infrastructure for Agricultural Universities with multi-layer physical & data security infrastructure and meet the growing IT needs of the NARES system. Krishi Megh is an enabler environment to host all the developed systems and makes the Council future ready.Krishi Megh is distributed across two data centers viz. ICAR- Data Centre (ICAR-DC) and ICAR-Disaster Recovery Center (ICAR-DRC). These are established to provide Email, Unified Communication and Web Hosting solution to all the ICAR institutions and now enabling Agricultural Universities to utilize most of these services. ICAR DC meets the international standard of Data Centre Tier II specifications with adherence to ISO 27001, ITIL and TIA 942 standards
9. **Kritagya Portal:** Kritagya Portal is developed to automate end-to-end modules of national level ag-tech hackathons. Using this portal, two national level hackthons have been organized in virtual mode resulting in multiple good ideas and products.
10. **Project Monitoring and Tracking System (PMTS):** The system is developed to provide a digital platform to NAHEP participating agricultural universities with an automated solution to project tools required for monitoring and tracking of projects, ensuring projects to complete within estimated timelines.
11. **Grievance Redressal Mechanism System (GRM):** GRM is developed to have fair and transparent system for NAHEP participating agricultural universities to report and resolve grievances with respect to procurement, social, environmental and other issues.
12. **Direct Benefit Transfer (DBT) DARE MIS:** DBT DARE MIS is developed to capture records of individual beneficiary and transactions under DBT applicable schemes of

DARE. Out of 20 schemes, 18 schemes are from the agriculture education division of the Council.

13. **Capacity Building Portal (CBP):** CBP facilitates online management of all training programs under Centre for Advanced Faculty Training (CAFT), Summer- Winter Schools (SWS), and short courses sponsored by Agricultural Division of ICAR. The system end-to-end automation of all the processes involved in advertisement of the courses, filling of proposals by various faculty for organization of the courses, registration of course applicants and their selection in various courses, uploading of time table, generation of e-books, library of photographs and the feedback of the participants.
14. **Agricultural Experts Information System (AEIS):** AEIS is developed to create a network of subject matter experts in agriculture sector and build a repository of the CVs of these experts across various fields of agriculture. The system also enables the top management to engage these experts in various high level committees as well as develop the repository of recommendations made by these committees.
15. **Virtual Classrooms and Agri Diksha Web Education Channel:** The infrastructure has been established to provide seamless teaching – learning experience to faculties and students. This supports Agricultural Universities in making Agricultural Higher Education System more resilient and relevant. A network of 76 virtual classrooms across all Agricultural Universities and IASRI has been developed. All these classrooms are centrally connected to Krishi Megh, and a huge repository of video classrooms is being created which is freely accessible to all students of the country thereby empowering them with latest knowledge and inclusive digital growth a reality.
16. **AI Mobile Application for Pest and Disease Diagnostics (NIBPP):** The mobile application is developed in collaboration with Agriculture Universities to identify the diseases in crops by analysing the picture of the infected part using image analysis, deep learning and related Artificial Intelligence technologies. A huge corpus of 3 lakhs images have been developed to enable the modelling of the identification of major diseases of major crops.

Amid the changing landscape, the need for bringing in digital transformation in the agriculture ecosystem cannot be emphasized enough to make it stronger and more resilient. The new paradigm of digital agriculture encompasses disruptive technologies aimed at enhancing the efficiency of the overall agriculture ecosystem by unlocking productivity and value through transformative cloud infrastructure, applications and technologies, thereby unleashing a new wave of innovation.

The digital initiatives listed above aim at strengthening the agricultural education system by improving the quality, learning outcomes and enhancing the access to high quality education through appropriate, effective and interactive digital learning channels. These initiatives have been designed using a holistic approach for stakeholders beyond students viz. faculties, farmers, ICAR officials to create an integrated agricultural education platform to address

educational competency gaps and information asymmetry. They have been envisaged with an objective to transform the functionality of agricultural universities and build a greater understanding of concepts and applications through emerging frontline digital technologies, encompassing autonomous systems, cognitive systems, single window platforms, new digital content delivery and learning methods and tools, educational e-resources, alumni and experts connect, up-skilling of faculties on a long term and continuous basis, thereby improving the learning outcomes of not only the students but also the lifelong learners. All these initiatives are inculcating a digital change in the academic arena and ICAR-IASRI is further striving to improve the system and make it more robust for academic usage across agricultural universities.

Factors affecting ICT Usage in Agricultural Higher Education

Rajni Jain

ICAR-National Institute of Agricultural Economics and Policy Research, New Delhi

Information and Communication Technology (ICT) helps to widen access to tertiary education, improves student engagement, enhance academic performance and create opportunity for better employment. Although several efforts have been made by the Government to digitalize the agricultural education, its status of use in agricultural higher education is not much studied. This study identifies use, availability, awareness and determinants of ICT usage level among the agricultural students through a primary survey. The study found that 73 per cent students use social media for academic purpose. It was estimated that 65 per cent of the respondents belong to high or medium category pertaining to use of academic ICT resources. The results highlight the direct relationship of ICT use with its availability and awareness. Further, the use of ICT resources is influenced by uninterrupted internet connectivity, quality of ICT initiatives, coverage of contents and related trainings. Regression estimates indicate that one per cent increase in awareness level will increase the ICT use by 0.4 per cent. Besides, course, gender, social category and family background are significant factors affecting the ICT use. Thus, ICT use gap among agricultural students should be reduced by imparting training and improving the quality of available ICT resources.

Virtual Classroom and Agri-DIKSHA: Embracing the future of Digital Learning in Agriculture Higher Education

Anshu Bharadwaj[1],Sudeep[1], Alka Arora[1], Mukesh Kumar[1], Shashi Dahiya[1], S.N. Islam[1], Soumen Pal[1], Rajender Parsad[1], Anuradha Agrawal[2], R.C. Agrawal[2]

[1]ICAR-Indian Agricultural Statistics Research Institute, New Delhi

[2]NAHEP- PIU, ICAR, New Delhi

The National Education Policy (NEP), 2020, has paved the path towards revolutionizing the outlook of the education landscape in the country in general. The NEP, 2020 pushes for acceleration in the development of smart classrooms for using digital pedagogy and thereby enriching the teaching-learning process with online resources and collaborations.

Agricultural higher education is also undergoing a digital transformation across India. Both offline and online education modes have grown in leaps and bounds in the past few years. The COVID 19 pandemic has further accelerated the transition to a more fluid, student centric teaching learning methodology. The use of new technology platforms and technology-aided learning tools is transforming the traditional teacher-class based teaching to digital learning in agricultural universities.

One such digital learning initiative aligned with ICAR's vision of "Strengthening and Development of Higher Agricultural Education in India and Improving Quality of Agricultural Education" under National Agricultural Higher Education Project's (NAHEP) Component II project 'Investment in ICAR Leadership in Agricultural Higher Education', is setting up of "Virtual Classrooms", to enhance the teaching-learning experience in agricultural universities.

To improve the quality of education and widen the access of education for students as well as up-skilling of teachers across the country, this initiative come out as a boon. These Virtual Classrooms are equipped with sophisticated, state-of-the-art physical infrastructure with various hardware equipment. The virtual classroom facility is bundled with Agri-DIKSHA web channel (https://agridiksha.krishimegh.in) which is an interactive portal for facilitating teachers to develop and broadcast virtual learning modules. The virtual classrooms are part of the blended learning method that combines on-line and in-person teaching/learning wherein quizzes, video lectures and other learning materials can be embedded in virtual learning modules. It combines entrepreneurial training, collaborative teaching and the latest technological teaching tools to create modern and effective education service environment in education setting leading to a resilient and sustainable Agriculture Higher Education system.

Online Initiative for Ranking of Green and Clean Agricultural University Campus

Shashi Dahiya[1], Sudeep Marwaha[2], P. Ramasundaram[2], Anshu Bharadwaj[1]

[1]ICAR-IASRI, New Delhi

[2]National Coordinator, NAHEP-IDP, New Delhi

Indian Council of Agricultural Research is embarking upon an ambitious step in further strengthening the National Agricultural Education System in the country through National Agricultural Higher Education Project with financial assistance of the World Bank by investing on infrastructure, competency and commitment of faculty, and attracting talented students to agriculture. A Green Campus is a place where environmental friendly practices and education combine to promote sustainable and eco-friendly practices. The green campus concept offers an institution an opportunity to take a lead in redefining its environmental concern and develop new paradigms by creating sustainable solutions to environmental and social needs of the inhabitants. For ideal educational institutional buildings, it is imperative to put in place sustainable environmental management practices. Greening the campus is all about turning around wasteful inefficiencies and using conventional sources of energies for its daily energy needs, correct disposal handling, procurement of environment friendly supplies and effective recycling program. Universities have to work out the time bound strategies to implement green campus initiatives. These strategies need to be incorporated into the institutional planning and budgeting processes with the aim of developing a clean and green campus.

An online portal is developed to showcase the clean and green initiatives undertaken by agricultural universities in India in order to promote sustainable living. The universities are evaluated on pre-defined parameters of a clean campus, greenery to provide pollution free air and carbon sink, minimize waste and consumption of water and energy, adoption and deployment of environment friendly activities and Impact of use of digital technology and management to reduce consumption of natural/non-renewable resources - paper, gas, water, energy etc. The portal provides automated ranking of universities based on the validation of these parameters.

IP 06(A): Statistical Modelling and Forecasting in Agriculture
Chairperson: Dr. K.K. Tyagi

Drought prediction using Machine Learning Techniques

K.N. Singh, Rajeev R. Kumar and Mrinmoy Ray
ICAR-Indian Agricultural Statistics Research Institute, New Delhi

Drought is a multi face ted meteorological phenomenon of very less rainfall regions that significantly impact agriculture, water resource sustainability, climate and environmental management. Drought forecasting is an excellent way to determine how drought affects water resources and agriculture. Data-driven machine learning forecasting techniques are promising approaches for drought forecasting since they involve less development time, fewer inputs, and are less sophisticated than dynamic or physical models. Machine learning models for drought forecasting use drought indices that are more operational than raw climatic variables. In this study, we evaluate the forecasting performance of machine learning models like support vector

machine (SVM), extreme learning machine (ELM), artificial neural network (ANN), extreme learning machine (W-ELM) models, Multiple Kernel extreme learning machine (MK-ELM) and wavelet Multiple Kernel extreme learning machine (W-MK-ELM). Evaluation criteria like root mean square error (RMSE) and mean absolute percentage error (MAPE) has been used to evaluate the forecasting performance of the models.

Development of classification tree enhanced by genetic algorithm for forecasting in agricultural ergonomics

Ramasubramanian V., Mrinmoy Ray and Md. Wasi Alam
ICAR-Indian Agricultural Statistics Research Institute, New Delhi

Decision tree based classification methods are quite popular for their easy comprehension and less stringent assumptions on data on which they are built. In this study, development of an alternative and improved classification tree (CT) modelling procedure enhanced by Genetic Algorithm (GA) optimization has been done. In the conventional exhaustive search method in the Classification And Regression Tree (CART) methodology, it has been seen that simultaneously selecting the split variables as well as the split points introduces some selection bias for the variables with greater number of possible splits. To reduce this bias, a better way is to separate these two procedures which option had been tried earlier as well.In the present study, the same option has been attempted albeit in a different way. For split variable selection at each node, testing of significance of association of each predictor variable with the categorical i.e. binary response variable, with chi-square test for categorical predictors and one way ANOVA for quantitative predictors were done (which was also followed in the aforesaid option of earlier study as well). Thereafter, the predictor variable giving minimum probability value for the test statistic has been selected at each node. For selecting the split point for the selected split variable, a certain procedure from the stratified sampling perspective also involving GA optimization has been employed in a novel and in a quite different manner to divide the observations present at a node into two groups of the binary response variable, each being more homogeneous with respect to the two classes of it. It can be viewed as stratifying a heterogeneous population (the parent node) into two more homogeneous sub-populations (the daughter nodes) enhanced by GA optimization in the CT tree. Ergonomics data with study variable as presence/ absence of discomfort for agricultural labourers during farm operation have been considered. The associated variables are load given to farm machinery, modes of operation and percent aerobic capacity of the farm labourers (all these three are qualitative) and difference between working and resting heart rates and oxygen consumption at the time of farm operation (which are quantitative). The data was divided into training and test sets for model building and validation respectively. Three (Training: Test) per cent data sets as (60:40); (70:30); (80:20) and three independent random sets under each combination yielded nine different datasets for fitting models separately. Results on applications of the developed GA based CT models for classification have been found better

when compared with the existing methods. Thus proposed GA based CT model is a better alternative for classification problems as have been demonstrated by practical application in the agricultural domain.

AI-enabled Price Forecasting Model for Agricultural Commodities

Girish Kumar Jha
ICAR-Indian Agricultural Research Institute, New Delhi

Timely and accurate price forecasts of agricultural commodities can provide crucial and useful information for farmers, consumers, policy planners and agro-based industries. However, due to the confluent influence of seasonality of production of agricultural commodities, the derived nature of their demand, market imperfections, globalization, speculative trading, etc., agricultural commodity prices possess highly complex characteristics of non-stationary, nonlinearity and chaos, which pose a great challenge for time series forecasting. Traditional statistical models fail to capture the complexities of these price series due to their inherent assumptions and pre-specified model driven approaches. On the other hand, artificial intelligence (AI) based models, a data driven, self-adaptive with flexible operational designs, have recently emerged as a promising alternative for time series forecasting. Deep learning methods namely multilayer perceptron (MLP), convolutional neural networks (CNN) and long short-term memory (LSTM) networks, can be used for automatic learning of temporal dependence as well as temporal structures such as trend and seasonality for complex agricultural price prediction problems. Recently, we have developed a deep long short-term memory (DLSTM) based model for the accurate forecasting of a nonstationary and nonlinear agricultural prices series. DLSTM model is a type of deep neural network which is advantageous in capturing the nonlinear and volatile patterns by utilizing both the recurrent architecture and deep learning methodologies together. We compared the price forecasting ability of the developed DLSTM model with MLP and ARIMA models using international monthly price series of maize and palm oil. The empirical results demonstrated the superiority of the developed DLSTM model over other models in terms of various forecasting evaluation criteria like root mean square error, mean absolute percentage error and mean absolute deviation. The DLSTM model also showed dominance over other models in predicting the turning points of those monthly price series.

IP 07(A): Basic Research in Statistical Sciences

Advances in Agricultural Knowledge and the use of Innovative Approaches among Farmers in Four Different Districts of Andhra Pradesh, India - A Baseline and Endline Biotech Intervention Assessment Conducted as part of the Biotech KISAN Mission

[1]Radha R Ashrit and Shipra Joshi
[1]NITI Aayog, Government of India
[2]Department of Biotechnology, Government of India

The majority of India's consumption requirements for pulses and oilseeds are being met through the country's extensive network of international imports. Andhra Pradesh is a leading state in the country that contributes significantly to the production of crops, as mentioned above, and it is one of the top states in the country overall. The low output and productivity of pulses and oilseeds in India can address through the implementation of tried and tested technological interventions, which can provide novel solutions. This study aimed to determine the differences in knowledge and adoption of Sustainable Agricultural Practices among farmers of Andhra Pradesh who received biotech intervention training at the baseline and end line of the study to increase the productivity of these crops. This was done to validate the same, and the purpose of this study was to determine those differences. In addition to this, it investigated the various probable aspects connected to knowledge and adoption habits.

During the study, 240 farmers from communities with inadequate technological improvements and low pulses and groundnut yield participated. After the training, the results showed that social and behavioural elements had greatly improved, significantly impacting farmers' understanding and adoption behaviours.

Allocation Proportional to Strata Total and Exponential Phase Effect under Cost Constraints

Med Ram Verma

ICAR- Indian Veterinary Research Institute, Izatnagar

Impact assessment of the development programmes is one of the important aspects in socio-economic surveys. Selection of the proper sample size is essential for conducting the surveys for assessment of the impact of the development programmes. In the present paper we have proposed sample allocation procedures when sample size in each stratum is proportional to the strata total and exponential phase effect. We have also considered the situation when the cost of observing the units varies from stratum to stratum. The proposed sample allocation procedures

are general sample allocation procedures for assessment of development programmes. The paper concludes with empirical illustrations.

Efficiency of Statistical Designs at Advance level of Sugarcane varietal trials

Rajesh Kumar, Rajendra Gupta, A.D. Pathak and Atul Sachan

ICAR - Indian Institute of Sugarcane Research, Lucknow

Two experiments were conducted according to randomized complete block experiment replicated twice with 21 cultivars and alpha lattice design with incomplete blocks with two replications, 21 cultivars, 3 blocks within a replicate and 7 plots per block in each replication. The data of each parameter were subjected to statistically analyzed according to the technique of analysis of variance (ANOVA) for the alpha lattice design developed by Patterson and Williams (1976) and RCBD through the computer software SAS 9.3. The error mean squares from each analysis were used to estimate the relative efficiency of an alpha lattice design (ALD) compared with Randomized Complete Block Design(RCBD) according to the following equation:

Efficiency of Alpha Lattice Designs (ALD) as compared to Randomized Complete Block Design (RCBD) is

$$\text{Efficiency of ALD as compared to RCBD} = \frac{\text{Amount of information in ALD}}{\text{Amount of information in RCBD}} 100$$

$$\text{Relative Efficiency of ALD as compared to RCBD} = \frac{\text{MSe(RCBD)}}{\text{MSe(Alpha)}} 100 * \text{CF}$$

where

$$\text{CF} = \frac{(e_1 + 1)(e_2 + 3)}{(e_1 + 3)(e_2 + 1)}$$

e1 = df of the MSe in a RCBD

and e2 = df of the MSe in a ALD

Relative efficiency of Alpha Lattice as compared to RCBD in field were found better for germination (30 Days), germination (60 Days), tiller (90 days), tiller (180 days), tiller (210 days), sucrose(%) at 10 months, height at harvest, girth at harvest, internode at harvest, total weight at harvest. Out of fourteen character, gain in efficiency (%) of nine biometrical characterfor Alpha Lattice Design (Field) vs RBD in Alpha Lattice Design (I) were found as for germination (30 Days) (99.10%), germination (60 Days) (98.09%), tiller (90 days) (23.25%),

tiller (180 days) (27.47), sucrose(%) at 10 month (1.53%), height at harvest (13.48%), girth at harvest (21.80%), internode at harvest (18.43%), total weight at harvest (0.43%).

Where as in qualitative characters, Alpha Lattice Design was inferior over Randomized Complete Block Design. Because qualitative characters in sugarcane mainly depends upon weather factors like low temperature and it is also does not depends upon the fertility of the soil and fertility of the experimental field. Least square mean of different verities conducted in Alpha Lattice Design and RCBD for ranking of the varieties were also discussed. Out of twenty-one varieties, among top four varsities, three are same in Alpha Lattice Design and RCBD for tillers at 90 and 180 days and yield, It indicated that ranking of the varieties are not different for cane yield, if we change the design. Most of the varsities are superior than the tested varieties. Seventy five percent ranks of the varieties are same in both the design.

Overall, Alpha Lattice Design was found superior over Randomized Complete Block Design for detecting genotypic difference of mean in advance level of trials of sugarcane specially conducted under Varietal Improvement Programme of All India coordinated Research Project (AICRP) on Sugarcane.Hence, we can occlude that alpha lattice design is more efficient than complete randomized block. Whereas alpha lattice design slightly more efficient than complete randomized block design of alpha lattice for most of the important biometrical characters. Whereas for few qualitative characters, both the RCBD were found superior than alpha lattice design. It may be due less variation among the varieties at month of observations.

2-part Designs for Agricultural Research

Cini Varghese
ICAR-Indian Agricultural Statistics Research Institute, New Delhi

Many a times, row-column designs are not readily available when the number of treatments is more than the levels of row and column blocking factors. One of the means to obtain new row-column designs for specific situations is through the amalgamation of two block designs in a systematic manner. Row-column designs obtained through the statistical fusion of two incomplete block designs will result in 2-part row-column designs. Algorithmic approach has also been explored for constructing series of row-column designs with incomplete rows and columns. A wide range of incomplete block designs, viz., balanced incomplete block designs/ partially balanced incomplete block designs/ t-designs, are available in the literature, which can be selected as input designs to construct such designs. To avoid the complexity involved in the construction algorithm, an R package “iRoCoDe” is also developed for the generation of these designs. Further, fusion of two incomplete block designs can also yield 2-part row-column designs for investigating a set of treatment combinations. Again, structurally incomplete 2-part row-column designs have been developed for situations where certain experimental units are not available for the application of treatments.

IP 04(B): Big Data Analytics, Machine Learning, Artificial Intelligence and their Applications in Agriculture

Integration of Artificial Intelligence (AI) and -omics for crop and animal improvement

[1]A.R. Rao, [2]T.K. Sahu, [3]P.K. Meher and [3]S. Sahu

[1]Indian Council of Agricultural Research, Krishi Bhawan, New Delhi
[2]ICAR-National Bureau of Plant Genetic Resources, New Delhi
[3]ICAR-Indian Agricultural Statistics Research Institute, New Delhi

With the advent of recent and less expensive next generation sequencing technologies, huge facets of genomic data are generated and shared in public domain. On the other hand, phenotypic data is also available with the plant and animal breeders. Now, the challenge is to effectively mine such huge amount of both genomic and phenotypic data to understand complex phenomena of plant and animal improvement. However, in the event of complexity becoming exponentially high, it may be difficult to solve the problems by human mind in which cognitive functions such as learning and problem solving are associated. In such cases, the cognitive functions are mimicked by machines, especially computer systems, through Artificial Intelligence (AI) with traits:reasoning, problem solving, knowledge representation, planning, learning, natural language processing (NLP), perception, motion and manipulation, social intelligence, etc.

Even though several factors can be ascribed for trait improvement in plant and animals, only a few of them are discussed here where machine learning, deep learning and predictive analytics such as artificial neural networks, support vector machines, random forest have been applied in -omic sciences. The present talk covers prediction of splice junctions; prediction of antimicrobial peptides; classification of coding and non-coding RNAs; identification of nitrogen fixation genes, herbicide resistant genes, insecticide resistant proteins, and heat shock proteins; DNA barcode based identification of microbial species; prediction of multiple sub-cellular localization of genes; abiotic stress responsive miRNA prediction; Spike recognition and counting from visual imaging; Genomic selection and prediction of genomic estimated breeding values; Association of non-coding RNAs in FMD in cloven hoofed animals; Epitope prediction for vaccine designing.

As nutritional security and climate change are two major challenges of Indian agriculture, there is an urgent need to focus on(i) biofortified plant and animal products and (ii) plant varieties and animal breeds with special traits. The horticultural crops grown and animals reared in the Jammu & Kashmir and Ladakh regions provide an opportunity to study the genomic insights by AI that may help understand the expression of traits in extreme environments.

Role of Statistical Software's in Data Analysis

Amrit Kaur Mahal
Punjab Agricultural University, Ludhiana

In research results, the role of data analysis techniques is quite predominant. Applications of statistical software's becomes crucial part of data analysis with the growth of quantitative research. The emergence of statistical software's in the twenty first century has benefitted researchers in the physical and social sciences to improve the quality of research from manual analysis with paper to more efficient computer packages. Statistical software packages are available both as free open-source software and proprietary/licensed software. All the statistical software's yield almost same results, although it follows different algorithm in each case. In addition, some more information is provided in each package compared to others. Each software has its own advantages and disadvantages. Thus according to requirement, users can choose their package carefully. In this study an attempt is tried to describe strengths, weaknesses and uses of different statistical software packages.

Artificial Intelligence –Importance and Application in Agriculture

S.A. Mir and Mehnaz Shakeel

Sher-e-Kashmir University of Agricultural Sciences and Technology of Kashmir, Srinagar

Agriculture has a key role in supplying food to the world's rising population. According to ICAR forecasts, food grain demand would rise from 192 MT in 2000 to 345 MT in 2030. Agriculture is beset by issues ranging from land preparation to product marketing.To overcome these issues, technology will be required, which may include the usage of Artificial Intelligence. Artificial intelligence (AI) refers to the emulation of human intellect in robots that are programmed to think and act like humans (Salmon et al.,2021). The capacity to handle huge quantities of data, form inferences depending on the circumstances, analyse and solve complicated issues, and learn based on historical patterns, expert input, and feedback loops are all general properties of AI (Canhoto& Clear, 2020). Seasonal forecasting, pest and disease control, agricultural yield

enhancement, IoT device data enhancement, improved crop selection, chatbots for farmers, and agriculture robotics are all areas where AI is being employed. Machine learning, deep learning, robotics, expert systems, fuzzy logic, and natural language processing are some of the AI topics on which we may build our competence. With sensors, drones, robots, and intelligent monitoring systems, artificial intelligence (AI) and machine learning have a wide range of applications in agriculture. In horticulture and field crops, computer vision-based phenotyping of plant stress, diagnostics, and severity evaluation of plant diseases have gained traction. Early detection and prediction of plant diseases, are being conducted using the Internet of Things, which is based on networking sensors. Unmanned aerial drones are used to phenotype orchards in order to apply plant protection agents precisely. Smartphone-based field diagnostics are gaining traction throughout the world, particularly in distant places where disease diagnosis in the lab is challenging(Prabha., 2021).

Significance of Big Data's Vs in Digital Agriculture

K.K. Chaturvedi
ICAR - Indian Agricultural Statistics Research Institute, New Delhi

The big data initially begins with 3Vs i.e., Volume, Variety and Velocity. Later, Veracity and Value have also been added to make it 5Vs in Big Data. Subsequently, another few Vs are also added in this series such as volatility, vagueness, validity, visualization and many more. The data is available in three different forms i.e., structured, unstructured and semi-structured. The further inclusion of Vs in Data has more significance in analyzing these data to draw important and meaningful knowledge from the raw as well as derived data with the help of knowledge discovery in databases workflow. Recent advances in Information and Communication Technologies enable researchers in transforming the traditional agriculture into digital agriculture. The application of Artificial Intelligence and Machine Learning techniques enables farm practitioners to implement automated way to handle various operations of agriculture in smarter way. The special V as visualization in Big Data will play important role in displaying various information in different ways through slicing, dicing, roll-up, drill – down and drill through across various dimensions with the help of Dimensional Modelling and OLAP technologies of Data Warehousing in storing, integrating and aggregating the information.

IP 06(B): Statistical Modelling and Forecasting in Agriculture

Marine Fishery Resource Modeling - Trends, Travails and Talismanic Tools and Takeaways

J. Jayasankar
ICAR-Central Marine Fisheries Research Institute, Kochi

Marine fishery resource modeling has always been wrought with conundrums and caveats. Amongst natural resources this is one that needed high dimensional data analytics right from early days of analytics. Modeling fish stocks and making them to yield to predictions with biotic and abiotic extraneous factor is as much a biologists challenge as that of statisticians. Numerous frequentist and Bayesian approaches have been put to use to unravel the sets of causes and the relevance of sets of traits that are of biological and socioeconomic relevance. Similar to quantitative genetic parameters, the functionals and indices that make up a comprehensive study of population dynamics of fish stocks is extremely convoluted and is a combination of sub-parameters that can only be plugged, with no formal possibility of deriving exact posteriors. Deep learning algorithms too have an unusual multi-level applicability at such situations, while contriving layers of latent relationships. An attempt would be made to apprise, evaluate and evolve convolutional neural networks based modeling of marine fish stocks leading to their management strategies for better governance to ensure sustainability and livelihoods.

Spatio-temporal Assessment of Water Resources in India Under Changing Climate: Application of Statistical Tools

D. K. Panda
ICAR-Indian Institute of Water Management, Bhubaneswar

Currently, climate change is one of the major challenges that humanity witnesses in the form of recurrent drought, flood, heat wave, cyclone, earthquake, forest fire and their compounding impacts. In the global comparison, India's challenges are unique and the scale of risk is also enormous, owing to the inextricable linkage between biophysical and socio-economic systems. This is for the primary reason: nearly 60% of its huge population over 1.3 billion relies on natural resources for their own livelihood that untimely resurrect the national economy. Now, India stands fifth among the most vulnerable countries of the world.

Especially, the monsoon rainfall during June-September contributes irrefutably to the socio-economic stabilization of the nation. For example, a deficit of only 19% monsoon rainfall in 2002 exacted ~1% loss of gross domestic production (GDP). At the same time surplus rainfall lead to evacuation of millions people and damages the infrastructure particularly in the densely populated low-lying areas along the 5423 km mainland coast line. A record extreme rainfall event of 944 mm/day (i.e. on 26 July 2005) in Mumbai led to considerable loss of property and human lives. Similarly, the 2008 flash flood in northern Bihar affected 2.3 million people,

attracted global attention and continues to be a topic of discussion in the scientific community. In 2015, for the first time in India, water train was introduced to carry portable water to the southern Peninsular India, reeling under drought and drinking water crisis. In 2019, like the Cape Town in South Africa, Chennai declared 'Day Zero', coined to connote severe water crisis, as a result of drought-induced drying of reservoirs that supply water to mega cities.

Under this backdrop, it is important to critically assess the water resources and their response to climate variability, lest there would be an existential crisis. However, the spatio-temporal data requirement is a major concern, given the inherent diversity of soil, climate and topography of the country that ranges from wettest to desert domains of the world. Generate a representative dataset is a costly as well as time consuming. This necessitates application of statistical tools and techniques to draw precise conclusions. For irrigation scheduling and contingency planning following drought or flood, for example, require soil moisture data. But it is difficult and requires huge expenditure to measure routinely soil moisture at the national scale. Similarly, the terrestrial water storage (TWS) that provides a rough estimate of total water stored below and above the Earth's surface. Equally crucial is also the land water storage from global hydrological models (e.g., WGHM, PCR-GLOBWB, GLDAS NOAH, MOSAIC, VIC, CLM, and CLSM).This talk focuses on "How statistical tools and techniques can be leveraged for sustainable management of water resources of India", given the looming water crisis that affect a large part of the society.

Decomposition based Machine Learning Techniques for Forecasting Agricultural Prices

Ranjit Kumar Paul
ICAR-Indian Agricultural Statistics Research Institute, New Delhi, India

Presence of nonlinearity, non-stationarity, non-normality and heteroscedasticity are frequently observed in most of the agricultural commodity price data. These features hinder the application of usual linear and nonlinear parametric models like Autoregressive integrated moving average (ARIMA), Generalized autoregressive conditional heteroscedastic (GARCH) and their component models to represent the series satisfactorily. It is well known that every time series is consisting of actual signal with noise. It is difficult to extract actual signal from noisy time series observations. In this regard, nonparametric wavelet technique has the advantage of denoising the series to extract the actual signal. Optimizing level of decomposition and choosing appropriate wavelet filter can represent the series with high chaotic nature and sophisticated nonlinear structure more effectively. The decomposition can describe the useful pattern of the series from both global as well as local perspective. The denoised components can be modeled using Machine Learning (ML) techniques like Artificial Neural Network (ANN), Support vector regression (SVR), Random Forest (RF) etc. to result in wavelet-based hybrid models and eventually, inverse wavelet transform is carried out to obtain the prediction of original series. The above algorithms have been applied in number of agricultural commodity prices with different level of decomposition and wavelet filters. The prediction accuracy of the hybrid model

is compared empirically with that of ARIMA, GARCH, ANN, SVR, RF technique and it is observed that proposed hybrid algorithm outperforms the other models.

IP 07(B): Basic Research in Statistical Sciences /Statistical Inference and Multivariate Analysis

Emerging need of Composite indices of evaluation with respect to HDI and SDG

Sheela Misra
University of Lucknow, Lucknow

Statistics is a science of making decisions in the face of uncertainty and randomness. Indices help us in doing the same. Despite of many indices doing good at individual levels in different dimensions of developments around human life we do not see desired status of holistic well being on the planet for the human and nature both. This talk explores various dimensions of the need of appropriate composite indices with reference to HDI and SDG.

Genetic Algorithm based Cluster Analysis

B.K. Hooda
CCS Haryana Agricultural University, Hisar

Conventional clustering techniques such as k-means and Fuzzy c-means employ a greedy search method over the search space to optimize the compactness of the clusters. These techniques are popular among applied scientists due to their availability in statistical packages and computational efficiency. However, depending on the choice of the initial cluster centers these algorithms generally get stuck at some local optimum solution. Also, there is a need to define the number of clusters a priori. On the other hand, Genetic Algorithms are based on the principle of natural genetics and follow the process of evolution. An optimization approach based on Genetic Algorithm finds a near-global or global optimal solution for a given fitness function. In the present talk, the performance of the Genetic Algorithm based clustering method has been studied and empirically compared with conventional clustering methods such as k-means and Ward's clustering methods. The problem of determining the optimal number of clusters in a data set has also been discussed in addition to a hybrid clustering approach for getting improved cluster solutions.

On Some Improved Imputation Methods Under Mcar Approach

Shashi Bhushan
University of Lucknow, Lucknow

Missing data is an inevitable phenomenon in most of the sample surveys and when it is not handled cautiously at beginning of the study, outcomes may result to substantial biases in the survey estimates. Retaining such evidences into cognizance, this paper offers few possible improved imputation methods to figure out the problem of missing data at the starting of the study and proffers computation procedure of the population mean using simple random sampling. The characteristics of the offered imputation methods are ascertained till first order approximation pursued by a computational study carried out using several existent and hypothetical data sets. The findings are turned out rather promising showing dominance over the existing imputation methods.

Estimation of Finite Population Mean under Predictive Modeling Approach

S Maqbool, M S Pukhta T A Raja and Immad A Shah

SKUAST of Kashmir, Srinagar

Estimation part is important objective of the sampling theory and every researcher's interest is to enhance that estimation part using auxiliary information to increase the precision of parameters of interest to be estimated. In this paper, we address the problem of estimating the finite population mean under predictive modeling approach. Modified ratio type predictive estimator of finite population mean using linear combination of auxiliary variable is proposed. The expression of bias and mean square error are obtained. Theoretical results are also supported by numerical illustration.

IP 08: Remote Sensing, UAV, IoT and GIS and their Applications in Crop Yield Estimation in Agriculture

Integrated sampling methodology for crop yield estimation using remote sensing, GIS, Geo-statistics and field surveys for Crop Insurance

Tauqueer Ahmad, Prachi Misra Sahoo and Ankur Biswas

ICAR-Indian Agricultural Statistics Research Institute, New Delhi

Pradhan Mantri Fasal Bima Yojana (PMFBY) was launched by the honourable Prime Minister of India, Shri Narendra Modi on 18 February 2016, which is a yield based Scheme. With the introduction of PMFBY, number of Crop Cutting Experiments (CCEs) has increased many fold i.e. about one crore CCEs at national level for crop yield estimation from the existing about 13.00 lakh CCEs for crop yield estimation under General Crop Estimation Surveys (GCES) Scheme which has not only become unmanageable at states level with the diminishing manpower but also results in significant increase in non-sampling errors. Therefore, in order to reduce and optimize the number of CCEs conducted across the country, an integrated methodology for crop yield estimation in the context of crop insurance has been developed. Under this study, cotton crop was chosen in Buldana district of Maharashtra for Kharif 2018-19, wheat and mustard in Barabanki district of Uttar Pradesh and Morena district of Madhya Pradesh for Rabi 2018-19 respectively. The different bands of satellite data were stack layered and classified using the latitude and longitude points of the CCE plots of the study area. Deep stratification using various vegetative indices like NDVI, NDWI, SAVI etc. and neighbouring GPs (administrative boundaries) was done. GP level estimates along with percentage standard error were obtained within stratum using traditional sampling techniques. GP level crop yield estimates along with percentage standard error were obtained and compared with the actual CCEs based estimates. The results were found to be encouraging and showed that the number of CCEs can be reduced significantly (around 30% or even lesser) with less than 10% standard error at GP level. The developed methodology is not yet at implementation stage as it needs further research using drones, advanced sample survey techniques e.g. Geographically Weighted Regression (GWR) in survey sampling, Calibration approach in survey sampling, Machine Learning approaches in survey sampling, Small area estimation etc. besides, Remote Sensing, Geo-statistics and crop yield modelling using satellite data, weather data and CCE yield data etc.

Smart Precision Models for Rice Yield Estimation

B. Sailaja, S. Gayathri, D. Subrahmanyam, K. Surekha, Santosha Rathod, R. Nagarjuna Kumar and R.M. Sundaram

ICAR-Indian Institute of Rice Research, Hyderabad

Current advances in ICT are making smart farms as DATA has become key element in modern agriculture to help growers in critical decision making. Remote Sensing and Proximity Sensing (Internet of Things i.e. sensors and other devices) drives agriculture 4.0 generates a big amount of valuable precise information. This ever growing data available at field level should be analysed for extracting meaningful data with big data analytics. The Next Era is Agriculture 5.0 i.e. Artificial Intelligent platforms with machine/deep learning algorithms to help farmers by giving intelligent decisions providing detailed information on soil, crop status, and environmental conditions to allow precise applications of phytosanitary products, resulting in a

reduced use of herbicides and pesticides, improved water use efficiency and increased crop yield. Rice is the major food crop in India. The average productivity of rice is still low (~2.7 t/ha) because of the diversity in its growing environments. Management levels and production constraints vary spatially and temporally in these environments. Therefore, future improvement in rice productivity requires management technologies that are tailored to specific characteristics of individual farms / fields and environments through adoption of precision technologies. Hence, smart precision technologies like Internet of Things (IoTs), Remote sensing, GIS integrated with crop models are needed to improve the quality of decisions of farmers. A Web based Spatial Decision Support system has been developed at ICAR-IIRR by integrating ORYZA2000 model, RS/GIS layers and further linked with weather sensor data to generate optimum crop management advisories to get the maximum yield. Further past database of DSS will be analysed, logical patterns can be programmed using Artificial Intelligence and site specific choices on varieties, sowing dates, inputs etc. can be planned automatically by using the intelligent precision models.

IoT application for analysis and estimation of agricultural accidents survey data

V. Bhushana Babu, R.R. Potdar, M.B. Tamhankar, Nandini Thakur and K.N. Agrawal

ICAR-Central Institute of Agricultural Engineering, Bhopal

Collection of agricultural accidents data through sample surveys is a time consuming and tedious job. Accuracy of such data is always under question and depends on proper training to investigators and thorough verification for non-sampling errors before analysis for better estimation. An alternative to manual method is to make use of integrated app cum web-based IoT system was developed having capability to feed and upload data live from the spot of incidence or village that is assumed to reduce non sampling errors and improve accuracy of estimate due to agricultural accidents. The developed integrated system consists of mobile application at front end and a server-based software at back end. The multilingual application has multi-level security with ease to operate through any android mobile phone. The system can generate reports as per the access authentication. The system can generate reports on accidents type and nature, total number, farm machinery, hand-tools etc and capable of estimating accidents at district, state and region of choice. Android Studio, Adobe Dreamweaver, MySQL,J2EE ,CVS, Eclipse, Apache Tomcat were used for development of this integrated system. The developed system is assumed to be accurate, less time consuming, portable, easy to use, easy to monitory, cost effective, error free and thus help in providing better estimates of

agricultural accidents and the economic loss thereof for better planning to minimize agricultural accidents.