The research targets set by the Institute were implemented by six Divisions of the Institute, viz. Design of Experiments, Biometrics and Statistical Modelling, Forecasting and Econometrics Techniques, Sample Surveys, Computer Applications and Centre for Agricultural Bioinformatics. The basic, applied, adaptive and strategic research in Agricultural Statistics and Informatics is carried out under six broad programmes that cut across the boundaries of the Divisions and encourage interdisciplinary research. The six programmes are as under:

1. Development and analysis of experimental designs for agricultural system research
2. Forecasting and remote sensing techniques and statistical applications of GIS in agricultural systems
3. Development of techniques for planning and execution of surveys and analysis of data including economic problems of current interest
4. Modeling and simulation techniques in biological systems
5. Development of informatics in agricultural research
6. Teaching and training in Agricultural Statistics and Computer Application

Programme 1: DEVELOPMENT AND ANALYSIS OF EXPERIMENTAL DESIGNS FOR AGRICULTURAL SYSTEM RESEARCH

Efficient Multi-level $k$-circulant Supersaturated Design

A supersaturated design (SSD) is a fractional factorial design having insufficient run size for estimating all the main effects represented by the design matrix. SSDs have received considerable attention in literature because of potential application in factor screening experiments, computer experiments, software testing, medical, industrial, and engineering and biometrical experiments. An algorithm to construct and generate efficient balanced multi-level $k$-circulant supersaturated designs with $m$ factors and $n$ runs has been developed. Using the proposed algorithm many balanced multi-level supersaturated designs are constructed and catalogued. A list of many optimal and near optimal, multi-level supersaturated designs is also provided for $m \leq 60$, $n \leq 25$ and number of levels $q \leq 10$. The algorithm can be used to generate two-level $k$-circulant supersaturated designs also and some large optimal two-level supersaturated designs are presented. An upper bound to the number of factors in a balanced multi-level supersaturated design such that no two columns are fully aliased is also provided.

**Experimental Designs in Presence of Neighbour Effects**

In agricultural field experiments, in order to control heterogeneity and conserve resources, treatments are assessed using small adjacent plots. Under such situations, the treatment applied to one experimental plot may affect the response on neighbouring plots besides the response of the experimental plot to which it is applied. This may be due to spread of treatments such as fertilizer, irrigation or pesticide to adjacent plots causing neighbour effects. Thus, neighbour effects may
contribute to variability in experimental results and lead to substantial losses in efficiency. For precise comparison of treatment effects in presence of neighbour effects, neighbour balanced designs are useful. These designs ensure that no treatment is unduly disadvantaged by its neighbour(s).

- **Robust Neighbour Balanced Block Designs:** A block design with neighbour effect(s) is said to be neighbour balanced if every treatment has every other treatment appearing constant number of times as neighbour(s). However, there is a possibility that some of the observations could become unavailable for analysis. The robustness of neighbour balanced complete block designs has been examined when specific observations are missing. The information matrix for direct treatment effects of the resultant design (one-sided neighbour effects) after missing of an observation from a block is derived and the efficiency of resulting design is investigated. The efficiencies are found to be quite high indicating the designs to be fairly robust against missing observations.

- **Neighbour Balanced Row-column Designs:** Row-column designs are an important class of designs which are very useful in situations when the heterogeneity present in the experimental material is in two directions. Neighbour balanced row-column designs have been defined wherein every treatment has every other treatment appearing constant number of times as neighbour in rows and columns. Two types of models have been considered based on how the neighbour effects of treatments are taken into account viz., row-column model with non-directional neighbour effects having same effects from all the four sides and row-column model with directional neighbour effects having different effects from different sides. The methodology for estimating the direct and neighbour effects of treatments has been derived under both the models. Series of row-column designs balanced for neighbour(s) have been obtained and are found to be variance balanced for estimating direct and neighbour effects.

- **Response Surface Methodology Incorporating Neighbour Effects:** Response Surface Methodology (RSM) is used to explore the relationship between one or more response variables and a set of experimental variables or factors with an objective to optimize the response. It is generally assumed that the observations are independent and there is no effect of neighbouring units. But in field experiments, the neighbour effects from the treatments applied to adjacent neighbouring plots may arise. The response surface model incorporating neighbour effects from immediate left and right neighbouring units has been studied assuming that the units are arranged linearly with no gaps. Procedure has been developed to estimate the parameters of both the first order and second order response surface models with differential neighbour effects. The variance of estimated response has also been obtained and conditions for the model to be rotatable have been obtained. A method of obtaining designs satisfying the derived conditions has been developed. An illustration showing the impact of neighbour effects has been given using a simulated data set. Further, the problem of large runs has also been taken care of by giving a method of construction of response surface design incorporating neighbour effects using fractional factorials. The variation between the blocks in the experiment is accounted for by including block effects in the statistical model. For first order model, the condition for orthogonal estimation of the parameters of the model with block effects has been obtained. Numerical examples have been given for illustrating the experimental situation where experimental units are grouped into complete/incomplete blocks. The neighbour effect in terms of neighbour coefficient has also been estimated. Procedure has been developed to estimate the parameters of the first order response surface model when the units experience neighbour effects from adjacent units and also the observations are correlated. The variance of the estimated response has also been obtained. A method of obtaining designs satisfying the derived conditions has been developed.

- **Experimental Designs in the Presence of Indirect Effects of Treatments:** Indirect effects are effects which occur in an experiment due to the units which are adjacent (spatially or temporarily) to the unit being observed. Considering more than one relationship between observations on units over space, the methodology for estimating
the direct and spatial (neighbour) indirect effect has been developed under a block design setup with neighbour effect and incorporating trend component. Two series of linear trend free block (one complete and one incomplete) designs have been obtained that are totally balanced for estimating direct and spatial (neighbour) indirect effect of treatments. For ease accessibility of the designs by the experimenters, it is required that these designs are compiled and presented at one place. Considering the time period (residual) as the indirect effect, the module for web generation of William’s square designs for even number of treatments has been developed.

Analysis of Experimental Designs with t-Family of Error Distributions

The interpretation of experimental data based on analysis of variance is valid only when the assumptions like error are independently and identically distributed as normal with mean zero and constant variance and effects are additive in nature. In several data sets collected in agricultural experiments, these assumptions may not be satisfied. In the analysis of experimental data when errors follow t-family of symmetric distribution, the normal equations obtained from the derivative of log-likelihood function with respect to parameters do not yield explicit solutions for the parameters due to non-linearity of the function. Generally, these equations are not easy to solve by iterative method because of slow convergence, multiple roots, and convergence to incorrect values or even divergence. The theory of modified maximum likelihood estimation (MMLE) has an explicit solution of these equations and is asymptotically identical with MLE. It has been shown in the literature that modified maximum likelihood estimates (MMLEs) are almost as efficient as maximum likelihood estimates (MLEs).

The analytical procedures have been developed for the analysis of data generated from designs of one-way elimination of heterogeneity when the error follows the t-family of symmetric distribution. These procedures are based on the solution of modified maximum likelihood estimations. When the derivative of log-likelihood function with respect to parameters do not yield explicit solutions for the parameters due to non-linearity of the function, the non-linear function has been expanded using the Taylor’s expansion of the first order and by this procedure the function becomes linear and the equations become solvable. Further, the test statistics and their distributions have been worked out. Similar to the designs of one-way elimination of heterogeneity, the theory of MMLEs have been developed for the designs of two-way elimination of heterogeneity. Procedures for the analysis of designs of two-way elimination of heterogeneity have been developed. Also the test statistics for testing the null hypothesis of the effects of treatments, rows and columns for latin square have been developed. One of the most commonly used types of factorial designs is the $2^k$ factorial experiment. For the model of $2^k$ factorial experiments when the error follows the t-family of symmetric distribution, contrasts and sum of squares of contrasts for main effects and two factor interactions have been worked out. Variance of the error components has also been worked out. For testing hypothesis of the main effects and two factor interaction effects, test statistics have been developed. Further, this method of MMLE of $2^k$ factorial experiment has been extended for $2^2$ factorial experiments and have been generalized for the factorial experiments with $k$ factors each at 2 levels.

Robust Block Designs against Missing Data for Making all Possible Pairwise Treatment Comparisons

Robustness of incomplete block designs against the unavailability of data has been investigated in the literature in terms of average variance of all possible pairwise treatment comparisons in the design. But for examining the robustness of a design for the loss of observation(s) on the basis of individual pairwise treatment comparisons, loss of information of some of treatment comparisons may be more than that of the loss of information on the basis of average variance of the residual design. A design that is robust on the basis of overall efficiency may not be robust when the efficiency is worked out on the basis of individual pairwise treatment comparisons. Therefore, all the estimates of individual pairwise treatment contrasts for the loss of any number of observation(s) in a block for balanced incomplete block design and variance balanced block designs have been investigated. Designs that are robust on the basis of average variance but not on the basis of pairwise treatment contrasts have also been identified.

Efficient Designs for Drug Testing in Veterinary Trials

In veterinary trials, neither a specific intervention treatment (treatment that has not been tested earlier)
can be given continuously to animals, nor can these treatments be withdrawn after any period, for ethical reasons. An alternating treatments design (ATD) does not require treatment withdrawal and minimizes irreversibility problem of the treatments and enables to study sequence effects. It facilitates quick comparison of two or more experimental conditions with each other or baseline. Due to treatment surrounding rule, only a limited number of intervention treatments and baseline can be compared at a time. A general method for constructing variance balanced ATDs, suitable for making comparisons of two or more experimental conditions with each other or baseline has been developed.

Two series of row-column designs for comparing investigational products with an active control/placebo in veterinary trials have been obtained. The arrangement of investigational products in the design is such that they follow 3-associate class rectangular association scheme. Further, designs for making comparisons of investigational products with more than one active control have also been obtained. Two series of symmetric factorial row-column designs were obtained. Properties of these designs have been studied and it was observed that these are partially variance balanced.

**Agricultural Field Experiments Information System (AFEIS)**

AFEIS is a Web-enabled information system (http://js.iasri.res.in/afeis) wherein information relating to informed agricultural field experiments (excluding purely varietal trials) conducted in the country are stored and maintained on-line. Presently, the database contains information relating to 32,462 agricultural field experiments, out of which 1727 experiments have been entered on-line by regional officers and others during the current year. For experiments with manure, alone and in combination with other factors, it was observed that 59.71% experiments were on manure alone followed by 15.70% manure with varietal trials and 12.06% manure with cultural treatments. The number of replications used in an experiment affects the precision of inferences as well as the cost of experimentation. Distribution of replication adopted, has been seen to be three in 61.16% of the experiments. Out of 32,462 experiments, raw data is available for 15,561 experiments. The software has been modified to download randomized complete block design raw data from the information system into Excel spreadsheets. The distribution of 32,462 experiments cropwise and type wise is presented below:

**Experiments Planned ON STATIONS under the Project Directorate for Farming Systems Research**

Under the Project Directorate for Farming Systems Research, the experiments ON STATIONS are planned and conducted under four types of research programmes viz. Development of new cropping systems; Nutrient management in cropping systems; Development of system based management practices and Maximum yield research using randomized...
complete block (RCB) design, factorial RCB design, split plot designs, strip plot designs and reinforced $3^2 \times 2$ balanced confounded factorial experiments.

Data of 336 experiments conducted during the year 2010-11 have been received and analysis work for 220 experiments has been completed. Results have been tabulated in the form of summary tables and are being sent to the respective scientist-in-charge of the cooperating centres. The final tables of the results of the experiments have been prepared to be sent to PDFSR, Modipuram for inclusion in the project report of AICRP on IFS. The distribution of percent coefficient of variation (CV) for these 220 experiments is as follows:

<table>
<thead>
<tr>
<th>CV</th>
<th>Number of Experiments</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 5</td>
<td>47</td>
</tr>
<tr>
<td>5-10</td>
<td>59</td>
</tr>
<tr>
<td>10-15</td>
<td>46</td>
</tr>
<tr>
<td>15-20</td>
<td>37</td>
</tr>
<tr>
<td>≥ 20</td>
<td>31</td>
</tr>
</tbody>
</table>

The assumption of normality of residual was tested using Shapiro-Wilk test for 56 experiments of Permanent plot experiment on integrated nutrient supply system in a cereal based crop sequence conducted during 2010-11. The assumption of normality is satisfied in all the experiments. Homogeneity of error variances was tested using Bartlett's $\chi^2$ test. The error variances were found to be heterogeneous in 3 experiments, for these cases data were transformed by Box-Cox transformation procedure and were analysed using Skilling Mack non-parametric procedure for testing equality of means.

Module for online analysis of data pertaining to the experiment Intensification/Diversification of cropping sequence based on high value crops has been developed. Scripts for developing information system for experiment (Permanent plot experiment on integrated nutrient supply system in a cereal based crop sequence) have been prepared.

**ON FARM Research Experiments**

Three types of experiments viz. Response of nutrients; Diversification/Intensification of cropping system and Sustainable production system were planned and conducted during 2010-11 at 31 ON FARM centres. One more experiment “On-farm evaluation of farming system modules for improving the profitability of small and marginal households” has also been initiated during 2011-12 in all the On Farm Research Centres. Online data entry and analysis for Experiment Response of Nutrients conducted during 2009-10 and 15 experiments (482 trials) at 12 OFR centres has been carried out. The data of 69 experiments conducted at 1,326 farmers at 24 On Farm centres of two types (Intensification/ Diversification and Sustainable Production System) were also processed for statistical analysis.

For the experiment On farm evaluation of farming system modules for improving the profitability of small and marginal farmers suggested that it is an investigation in which interventions may help in improving the incomes of households. Also, as the holding size, crop sequences, animal size, family size etc. may not be alike for all farmers so the proposed treatments in the technical programme may be called interventions rather than the treatments. By utilizing the input and output of these interventions in the given situations, pair wise comparison can be made only by using paired t-test.

**Fertilizer Response Ratios for Various Crops and Crop Sequences**

Data of about 11,000 on farm trials conducted in different NARP zones with the recommended fertilizer in various crop sequences of the experiments Response of Nutrients during the period 1999-2000 to 2008-09 under AICRP on FSR has been utilized for evaluating the fertilizer response ratios (a measure of the increase in production per unit fertilizer use) of four crop sequences and different crops. Four fertilizer response ratios such as N, NP, NK and NPK over control have been obtained for different crop sequences and crops. The fertilizer response ratios of 4 major crop sequences (rice-rice, rice-wheat, maize-wheat and soybean-wheat) have been obtained and then grouped according to the groups formed on the basis of initial major nutrients (N, P, K) present in soil. The fertilizer response ratio of 15 crops (5 cereals, 4 pulses, 5 oilseeds and 1 fibre) has been obtained at NARP zones, states level and national level using suitable weights. These response ratios have been again evaluated in different soils and states.

The fertilizer responses ratios vary widely from crop to crop, state to state and also on available initial soil nutrients. FRR also observed to vary in different groups formed on the basis of available soil nutrient and pairwise comparison of groups shows the significant difference in FRR values in different soils for all the four crop sequences taken under study.

For rice-rice sequence, the initial soil test values of major nutrients were available in 710 trials. Groups of trials are formed for low N (< 280 kg/ha), high N, low P (< 10kg/ha), high P, low K (< 108 kg/ha) and high K by Muhr’s classification. The fertilizer response ratio of N
over control of rice (kharif) and rice (rabi) have been evaluated in low N and high N groups as below.

**Fertilizer response ratio of Rice-Rice sequence for groups on available initial major soil nutrients (Low and High) in experimental site**

<table>
<thead>
<tr>
<th>Groups</th>
<th>No. of trials</th>
<th>Rice (Kharif)</th>
<th>Rice (Rabi)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response ratio to N over control (kg/kg)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N at low level</td>
<td>391</td>
<td>5.80</td>
<td>6.13</td>
</tr>
<tr>
<td>N at high level</td>
<td>319</td>
<td>8.01</td>
<td>6.17</td>
</tr>
<tr>
<td>Significant</td>
<td>*</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Response ratio to P (kg/kg)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P at low level</td>
<td>131</td>
<td>11.99</td>
<td>13.20</td>
</tr>
<tr>
<td>P at high level</td>
<td>5/9</td>
<td>12.1/</td>
<td>12.15</td>
</tr>
<tr>
<td>Significant</td>
<td>NS</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Response ratio to K (kg/kg)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K at low level</td>
<td>100</td>
<td>18.10</td>
<td>19.00</td>
</tr>
<tr>
<td>K at high level</td>
<td>610</td>
<td>13.80</td>
<td>12.94</td>
</tr>
<tr>
<td>Significant</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

*Indicates significant at 5% level

**Fertilizer response ratios of different crop group (All India)**

<table>
<thead>
<tr>
<th>Crop Groups</th>
<th>Area ('000 Hectare)</th>
<th>No. of Trials</th>
<th>Average Control yield (kg/ha)</th>
<th>Response over control (kg/kg)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>N</td>
</tr>
<tr>
<td>Cereals</td>
<td>86863.1</td>
<td>9909</td>
<td>2056.2</td>
<td>9.51</td>
</tr>
<tr>
<td>Oilseeds</td>
<td>15677.6</td>
<td>1086</td>
<td>1082.5</td>
<td>9.38</td>
</tr>
<tr>
<td>Pulses</td>
<td>5580.1</td>
<td>197</td>
<td>679.2</td>
<td>7.77</td>
</tr>
<tr>
<td>Foodgrains</td>
<td>108120.8</td>
<td>11192</td>
<td>1843.9</td>
<td>9.40</td>
</tr>
</tbody>
</table>

Using Kruskal-Wallis test, the difference of response ratios of rice (kharif) between groups low N and high N is observed significant at 5% level whereas it is not significant for rice (rabi). For both rice (kharif) and rice (rabi) the fertilizer response ratios of K differs significantly between groups low K and high K.

The fertilizer response ratios at all India level for cereal group of crops are observed higher than those of oilseeds and pulses group. At national level the fertilizer response ratio of all nutrient combinations are high for rice crop whereas these values are moderate for wheat crop. The response ratio of jowar and bajra to various nutrients are low. In pulse crops, blackgram are low. Oilseeds observed response ratio to N over control as moderate and low for other nutrients.

Continuous use of fertilizer in farmer’s field may alter the control plot output and affect fertilizer response ratios. These fertilizer response ratios obtained in the present study may be used with caution as these have been evaluated from the trials conducted at farmer’s field with recommended dose of fertilizer.

**Planning, Designing and Analysis of Data Relating to Experiments Conducted under AICRP on Long-Term Fertilizer Experiments**

The data in respect of various crop wise characters viz. grain and straw yield, plant nutrients concentration/uptake and available soil nutrients after the completion of each crop cycle from cooperating centres for each season pertaining to the experiments were undertaken. The superimposed treatments data on various characters relating to bifurcated plots of original treatments from the centres viz. Ludhiana (4 bifurcated treatments, each with 3 superimposed treatments); Pantnagar (2 bifurcated treatments, each with 5 superimposed treatments); Ranchi (4 bifurcated treatments, each with 3 superimposed treatments); Coimbatore (2 bifurcated treatments, each with 4 superimposed treatments) and Bangalore (4 bifurcated treatments, each with 3 superimposed treatments) were statistically analysed using the nested model to explore the possibility to utilize the build up P and to overcome the decline in yield due to inadequate supply of K or other nutrients like Zn and S. Guidelines and layout plan for bifurcations of original treatments with new superimposed treatments in one of the replications to generate new information on managing/harnessing of soil nutrients from the on-going experiment were provided to the centre in-charge of the Barrackpore centre. The scrutinized results and summary tables were provided to them for preparing their individual annual reports.
Programme 2: FORECASTING AND REMOTE SENSING TECHNIQUES AND STATISTICAL APPLICATIONS OF GIS IN AGRICULTURAL SYSTEMS

Weather Based Forewarning of Mango Pests

Weather based forewarning models have been developed for weekly disease incidence assuming that disease incidence in a particular week is due to two reasons viz. natural disease growth pattern and weather. Therefore, the model has been developed in two stages, modeling natural growth pattern and relating the deviations (from natural pattern) to appropriate lagged weather variables.

For natural growth pattern, the appropriate model was

\[ Y_t = \frac{23.119}{1 + 226.222e^{-0.799Wk}} \]

where \( Y_t \) is per cent disease incidence in \( i^{th} \) week (averaged over years), \( Wk \) is week number.

Deviations of weekly disease incidence from natural growth pattern were related with previous week population and weather with appropriate lags. The form of the model was

\[ y_k = A_0 + \sum_{i=1}^{p} \sum_{j=0}^{n} a_{ij} z_{ij} + \sum_{i=1}^{p} \sum_{j=0}^{n} a_{ij} z_{ij} + \epsilon \]

where

\[ z_{ij} = \sum_{m=1}^{n} r_{im} x_{im}, \quad z_{ij} = \sum_{m=1}^{n} r_{im} x_{im} x_{im} \]

\( y_k \): Per cent disease incidence in \( k^{th} \) week (deviation from natural pattern)

\( y_{km} \): Disease incidence in \( m^{th} \) lag-week to \( k^{th} \) week (deviation from natural pattern)

\( x_m \): Value of \( i^{th} \) weather variable (deviation from average) in \( m^{th} \) week lag

\( r_{im} \): Correlation coefficient between \( y_k \) and \( x_m \)

\( r_{im} \): Correlation coefficient between \( y_k \) and product of \( x_m \) and \( x_m \)

\( p \): Number of weather variables

\( r_{km} \): Initial/Final week lag for which weather data was included in the model and

\( \epsilon \): random-error term

Models were developed taking weather variables (\( x, i=1, ... 6 \) represent maximum relative humidity, minimum relative humidity, maximum temperature, minimum temperature, wind velocity and rainfall) with different lags (1 to 4) and disease incidence of previous week. The model taking data of previous two weeks (lag 1 and 2) was found appropriate. The model was

\[ Y_t = 0.045 + 1.08 y_{t-1} + 5.836 z_{t-1} + 3.093 z_{t-2} + \epsilon \]

\( R^2 = 0.89 \). Using this model, forecasts of per cent disease incidence in different weeks for 2008-09 were obtained and are presented as

The results indicate that forecasts are quite close to the observed ones (deviations less than 10 per cent) except in the first week i.e. week of disease appearance. It may be due to the reason that in this year, disease appeared late as compared to years used for modeling. Time of first appearance can be obtained using the model reported last year. Therefore, it can be concluded that using this model, reliable forecasts for per cent disease incidence can be obtained using two weeks data up to preceding week.

Visioning, Policy Analysis and Gender (V-PAGE)
(Sub-Prog. III): Policy Analysis and Market Intelligence (NAIP Project)

Demand elasticities were estimated for major spices (turmeric, garlic, ginger, dry chilli and other spices) for rural and urban areas of different regions. The expenditure elasticities of demand for selected spices except ginger were found to be moderately inelastic and ranged from 0.65 to 1.0 in all the regions. These elasticities were lower in urban areas as compared to rural areas of north, west, south and east regions. On the other hand, the expenditure elasticities were higher in the urban areas than that of rural areas of north-east region. Further, the expenditure elasticities of demand
for ginger were higher in the urban areas than that of rural areas of all the regions. Direct demand for major spices was also projected for the year 2015 and 2020 under high and moderate growth scenarios. The projections showed that the household demand for spices under moderate and high growth scenarios would be 4.56 & 5.52 million tonnes during 2015 and 6.72 & 8.22 million tonnes during 2020, respectively. During 2020, the demand for garlic and ginger was projected to be 1.38 & 0.64 million tonnes under moderate and 1.78 & 0.81 million tonnes under high growth scenarios, respectively. The demand of spices would be the highest in the southern region followed by western region of India. The supply analysis was carried out for major wheat and rice producing states of India. The estimation procedure includes an estimation of yield and area response through econometric analysis and these models were estimated simultaneously using Zellner’s SUR method. The analysis showed that gross irrigated area, seed and lagged yield were significantly influencing the area sown and current yield realized. Own and competing crops relative prices were found to have significant influence on yield in all selected states. Fertilizer consumption had significantly increased the area under wheat crop in Uttar Pradesh.

The efficiency criterion for future and cash markets was examined for discovering better price in soybean trade. The vertical market integrations among wholesale prices of groundnut products, namely seed, oil and cake were studied using the sequential procedure of Johansen’s multivariate cointegration technique. The results of the sequential multivariate cointegration tests for Chennai markets showed that the seed and oil prices of groundnut were integrated while oilcake was out of the system. In case of Hyderabad, price series at all the processing levels were integrated with each other in the long run. The speed of adjustment to the equilibrium was also studied using Vector Error Correction Model. The results showed that even though the seed and oil are integrated in both Chennai and Hyderabad markets, the speed of adjustment is more in Chennai as compared to Hyderabad. The analysis of farmers’ participation in future markets showed that most of the potato growers were unaware about future trading and opined that future market is a speculative market (Satta Bazaar) and not for hedging against price risk. The warehouse owners can be used as agency for financing, providing reliable market intelligence and quality and quantity certification. The margin should be fixed for farmers for instilling the confidence among the farmers’ about participation in futures market. Based on information collected from potato farmers, warehouse owners and future traders, a model is being developed for potato farmers’ participation in futures market.

Asymmetry in Retail Wholesale Price Transmission for Selected Essential Commodities

The prevailing large difference between wholesale and retail price of gram in the important markets in the country indicated towards delayed or lack of information flow and not following the market efficiency criterion. The study of vertical and horizontal cointegration between wholesale and retail price of gram in the selected markets of Bhopal, Chittoor, Delhi and Ganganagar using test and eigenvalue statistics indicated that there existed cointegrating vectors and cointegrating equations which confirmed a long run relationship in the Gram markets. The value of error correction coefficient $\hat{a}$ was observed to be significantly higher (the speed of price adjustment) in Chittoor and Bhopal markets as compared to Ganganagar and Delhi markets. The value of long run multiplier $\gamma$ suggest that the equilibrium between wholesale and retail price of gram in Chittoor market takes minimum time of 4 days, Bhopal 7 days, Ganganagar 49 days and Delhi market takes 63 days to attain the equilibrium level between wholesale and retail prices.

Programme 3: DEVELOPMENT OF TECHNIQUES FOR PLANNING AND EXECUTION OF SURVEYS AND ANALYSIS OF DATA INCLUDING ECONOMIC PROBLEMS OF CURRENT INTEREST

Sampling Methodology for Estimation of Meat Production in Meghalaya

The study was planned to modify the existing sampling methodology for estimation of meat production by working out the ratios of meat production between slaughter/butcher shops in meat markets and households and to estimate the species-wise number of animals slaughtered and meat production from different sources at district level with reasonable degree of precision.

A sample survey was carried out in East Khasi Hills district of Meghalaya State. Here, the meat markets are the major source for capturing the information on meat production. The complete enumeration of butcher shops was carried out in all the weekly and daily meat markets once in a year. 13 daily and 12 weekly meat markets
were randomly selected for recording the data on species-wise animal slaughtered and meat yield. The study established that a substantial number of each kind of animals were slaughtered by the households in the villages of the East Khasi Hills district. Neglect in capturing the slaughtering in villages could be the reason of underestimate of animals slaughtered in the Meghalaya State. In north eastern hilly regions, there are difficulties to collect the data from villages due to the typical topography so it is suggested that the data on animal slaughtered and meat production may be collected at a regular interval of time from the villages and a correction factor of 8.6% to the estimates of meat production from meat markets should be added as the annual meat production from villages for estimating the total annual meat production in the years for which data is not collected from the villages.

**National Initiative on Climate Resilient Agriculture (NICRA)-Agroforestry Component**

For estimation of area under agroforestry using Remote Sensing and GIS techniques, two districts namely, Ludhiana from Punjab and Vaishali from Bihar States were identified. High resolution satellite imageries (LISS IV), digitized district maps with village boundaries and toposheets of the two districts under study namely, Ludhiana from Punjab and Vaishali from Bihar States were procured. Geometric correction of the satellite imageries was carried out with the help of scanned and geo-referenced Survey of India (SOI) toposheets. Digital image processing of the satellite data for both the districts was done using ERDAS Imagine software. Extensive ground truthing for image analysis was done in 20 villages of 7 blocks of Ludhiana district and in 22 villages of 8 blocks of Vaishali district.

For estimating area under agroforestry, the satellite imageries were first classified using unsupervised classification method in ERDAS Imagine software and then supervised classification method using maximum likelihood classifier. Another approach of classification i.e. masking of each layer and then mosaicing was also used. Land use land cover map was generated for both the districts. The total number of classes obtained in Ludhiana and Vaishali districts was nine and five respectively. Area under agroforestry for Ludhiana and Vaishali districts were obtained as 6039.80 hectares and 217.86 hectares respectively. Accuracy assessment, an essential part of remote sensing based mapping, was also done and overall classification accuracy for Ludhiana district was found to be 94.28%. Area estimated is a product of number of pixels classified under agroforestry and resolution of the pixels (5x5 sq.m.). Assuming one tree per pixel under agroforestry, the number of trees estimated in Ludhiana and Vaishali are 24,17,400 and 87,144 respectively.

**On Small Area Inference using Survey Weights**

In several cases population level auxiliary information is not available and sometimes this information exist but is not accessible for use in survey estimation. In some cases, population and sample data can have inconsistency with respect to definition of auxiliary information since they are collected by different agencies/sources. In such situations, it is not possible to make full use of auxiliary information available in the survey data. When small area estimation is used, this problem becomes more serious as auxiliary information is very crucial and important. Keeping this in view, a small area estimator for small area means has been developed for the situation when population level auxiliary information is not available. The developed small area estimator uses estimated population level auxiliary information using survey weights. Unbiasedness property of the proposed small area estimator has also been studied. Besides, Mean Square Error estimator of small area estimator has also been developed. The developed method has a wide scope for adaptation in real life survey data analysis.

**Spatial Nonstationarity in Small Area Estimation under Area Level Model**

In recent years, the focus of economic and social planning has shifted from the macro to the micro level, and small area estimation is now of considerable interest to both planners and researchers in the country. Commonly used method of small area estimation, for example, empirical best linear unbiased predictor (EBLUP) under a linear mixed model assumes that the relationship between the variable of interest and auxiliary information is same over the study space. However, in many data sets this is not true and the process is referred to as the spatial non-stationary. This type of phenomenon is very common in agricultural and environmental data where relationship, that is, rate of change in target variable and auxiliary information changes from location to location. A geographically weighted pseudo empirical best linear unbiased predictor (GWEBLUP) for small area means under area level model has been developed. The proposed
estimator of small area mean is based on geographical weighted regression approach to incorporate the spatial non-stationarity present in the data. The method captures the local variation (or relationship) between the variable of interest and available auxiliary information via location specific models in small area estimation. The micro level estimates generated by using the developed small area estimation method are more reliable than the existing methods when spatial non-stationarity is present in the data. A comparison of relative biases and relative root mean square errors of the proposed GWEBLUP (solid line, red) and the EBLUP method (dash line, blue) are shown below. The GWEBLUP estimates have both smaller biases and root mean square errors as compared to the EBLUP estimates.

**Determination of Optimum Sample Size for Crop Yield Estimation at the Gram Panchayat Level**

In this study data was procured from the Ministry of Agriculture. Analysis of data was made through R package. Codes were written for generation of all possible simple random without replacement sample from a population and working per cent CV as per the design of sample survey for yield estimation at the GP level. Sample sizes were determined as per the specified criterion.

**Optimum sample size required for yield estimation at GP level**

<table>
<thead>
<tr>
<th>State</th>
<th>Sample Size</th>
<th>Crop</th>
</tr>
</thead>
<tbody>
<tr>
<td>Madhya Pradesh</td>
<td>3</td>
<td>Paddy (un-irrigated and irrigated), Soybean (Raisen district), Arhar, Gram</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>Wheat (irrigated and un-irrigated), Soybean (Sehore district)</td>
</tr>
<tr>
<td>Andhra Pradesh</td>
<td>5</td>
<td>Maize</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>Maize, Paddy</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Groundnut</td>
</tr>
</tbody>
</table>

**District Level Poverty Incidence Estimation from NSSO Data using Small Area Estimation Technique**

Measurement of poverty and its estimation has always been an important activity of the planning process in every developing country. Poverty measurement is generally based on household consumption expenditure surveys. Household consumption expenditure data in India are collected through National Sample Surveys Office (NSSO). At present, reliable estimates of poverty related parameters are available only at the state level. In view of the problem of non-availability of reliable estimates of various poverty related parameters at the district level for micro-level planning, a study was undertaken with the specific objectives of developing reliable district level estimates of various poverty parameters using small area estimation technique. The poverty parameters considered in the study were head count ratio, income gap ratio, poverty gap ratio and squared poverty gap ratio. Besides, estimates were also obtained for other related parameters like the number of persons, number of poor persons, monthly per capita consumer expenditure (MPCE) at the district level as well as various holding category wise, amount of loan outstanding at the district level and number of persons.
in the various MPCE classes. The data of 59th and 61st round data of NSSO for the rural areas was used. The 59th round data pertained to UP and was used to estimate the parameter amount of loan outstanding at the district level. The 61st round data pertaining to UP, MP and Punjab was used to estimate number of persons, number of poor persons, poverty ratio, income gap ratio, poverty gap ratio, squared poverty gap ratio and number of persons falling in MPCE classes. Two types of estimates were developed. The first one was direct district level estimate which used only district specific sample size. The district level small area estimates were developed using an area level linear mixed model. The covariates used in the small area estimates were obtained from 2001 population census data. The results of the analysis clearly revealed that the small area estimation technique based estimates had better precision as compared to the direct district level estimates. Further, various small area diagnostic procedures were performed to validate the reliability of the model-based small area estimates versus direct survey estimates. The study clearly showed that the various poverty estimates were highest for MP state followed by UP while the minimum values were obtained for Punjab state.

**Agricultural Research Data Book (ARDB)**

The Agricultural Research Data Book 2011, which is the fourteenth in the series, is an attempt to put together main components/indicators of all the information pertaining to agriculture. The Data Book comprising 170 tables is organized, for the purpose of convenience of the users, into ten sections namely, Natural Resources; Agricultural Inputs; Animal Husbandry, Dairying and Fisheries; Horticulture; Production and Productivity; Agricultural Engineering and Produce Management; Export and Import; India’s Position in World Agriculture; Investment in Agricultural Research and Education; and Human Resources under National Agricultural Research System (NARS). It also contains at the end, a list of important National and International Agricultural Research Institutions associated with agricultural research and education along with their addresses, telephone numbers and e-mail addresses. This edition contains the latest information/data as available in the country at the end of May 2011. In ARDB 2011, some value editions like predicting the future year production of foodgrain crops etc., based on previous years data using statistical models, pictorial/graphical representations of data have been done. For depicting state-wise data, thematic maps have been prepared using Geographical Information System (GIS).

### Programme 4: MODELING AND SIMULATION TECHNIQUES IN BIOLOGICAL SYSTEMS

**Stochastic Process Modelling and Forecasting through Discrete Non-linear Time Series Approach**

Threshold Autoregressive (TAR) model is an important parametric family, which is capable of describing cyclical data. An important subclass of TAR model is Self-exciting TAR (SETAR) model, which is based on a “piecewise” linear approximation via partitioning a state-space into several subspaces. Self-exciting threshold autoregressive moving average model (SETARMA) two-regime model has been fitted which is capable of describing cyclical data. Real-coded genetic algorithm (RGA) is employed to estimate the parameters of this model. As an illustration, annual Mackerel catch data of Karnataka, India during the period 1961-2008 is considered. Before estimation of parameters an exploratory data analysis for justifying SETARMA model has been carried out. Using Naive approach and Monte Carlo simulation technique, out-of-sample forecast performance of SETARMA model, is examined on the basis of Mean Square Forecasting Error (MSFE). It is found that SETARMA model generally performs better. Finally, optimal out-of-sample forecasts up to three-steps ahead along with their forecast error variances are derived theoretically for fitted SETARMA model. It is observed that, for hold-out data, observed values are quite close to forecast values and estimated variances are near to theoretical values up to three steps ahead prediction.

**Development of Methodology for Estimation of Compound Growth Rate and its Web-based Solution**

Formulae for computation of compound growth rates on the basis of two four-parameter nonlinear growth models, viz. Richards and Mixed-influence model have been derived. Formulæe have also been derived for computation of compound growth rates in respect of non-monotonic situations for all the three possibilities, viz. Over-damped, critically damped, and under-damped.

Assuming the random variable, where $r_i$ represents $\log(1+r_{i+1})$ compound growth rate, to follow a nonstationary process, the model describing growth for time-series data is
\[ Z = \log Y_r = \log(1+r) = g \left( \frac{r}{k} \right) + \varepsilon, \]

Local linear smoothing approach has been employed to estimate \( f_i = g \left( \frac{x_i}{h} \right) \). Optimal bandwidth of the kernel used in local linear smoothing has been obtained by minimizing Mean Integrated Squared Error (MISE). Modified plug-in bandwidth estimation approach has been used to minimize MISE. Finally trend of growth has been estimated by the large sample approximation \( \hat{f}(r) = \exp \left( \frac{r}{h} \right) - 1 \). Local polynomial smoothing approach has been employed to estimate time-dependent compound growth rate. Optimal bandwidth of the kernel used in polynomial smoothing is obtained by minimizing MISE through bootstrap. The bias and variance are obtained by replacing the estimated local trend and bandwidth in asymptotic theoretical expression. Under dependent error set-up, the two kernel smoothers have been used. Modified plug-in bandwidth estimation approach has been used to minimize MISE which is a function of unknown \( g(.) \) and the scale and decay parameters. The optimal bandwidth of the kernel is obtained by minimizing MISE by iterative procedure where the scale and decay parameters of autocovariance function is estimated by regressing log periodogram of estimated error sequence on logarithm of frequencies. Code has been constructed in R language for estimation of compound growth rate using three-parameter nonlinear growth models, viz. Monomolecular, Logistic, Gompertz model and four-parameter nonlinear growth models, viz. Richards and Mixed-influence model. A web-based user interface has been developed to upload users’ data online and to provide initial parameter values for the selected growth model. Based on the uploaded data and values of the initial parameters, code has been constructed to compute the estimated compound growth rate through the web-interface. Code for computing various statistics like Estimation of parameters, Predicted and residuals values, Shapiro-wilk normality test for residuals have also been constructed and incorporated into the web application.

**Bio-prospecting of Genes and Allele Mining for Abiotic Stress Tolerance**

Characterization and utilization of bio-diversity that is available in India is essential to meet the challenges of biotic and abiotic stresses under changing climate. For meeting the challenges of characterization of biodiversity

- Rice Germplasm database has been populated with the phenotypic data provided by the consortium centres, viz., NRCPB, IARI-New Delhi and CRRI-Cuttack.
- A transcriptome databases on camel, goat and bacteria (Enterobacter aerogenes) species has been developed from the data provided by NBAGR, Karnal and CIFRI, Barrackpore.
- SNVs data of camel and goat species provided by NBAGR has been parsed and a database on the same has been developed.
- Key residues responsible for salt stress tolerance across species for six different gene families have been identified by using bioinformatics approaches.
- ESTs similar to salt tolerant genes were annotated from the sequence data provided by CIFRI and further submitted at NCBI.
- Anoxia tolerant protein models were predicted and submitted to Protein Model DataBase (PMDB). These protein structures were predicted to find the structural conservedness among the anoxia tolerant proteins.
- Gene sequences for drought, acidity and temperature (chilling, freezing, heat and cold) stresses have been collected across the species and studied for the functional and structural conservedness.

**Programme 5: DEVELOPMENT OF INFORMATICS IN AGRICULTURAL RESEARCH**

**Strengthening Statistical Computing for NARS**

Strengthening Statistical Computing for NARS (www.iasri.res.in/sscnars) targets at providing

- research guidance in statistical computing and creating sound and healthy statistical computing environment and
- providing advanced, versatile, innovative and state-of-the-art high end statistical packages for analysis of data so as to enable drawing meaningful and valid inferences and converting research output into knowledge

The efforts also involve design intelligent algorithms to implement statistical techniques particularly for analyzing massive data sets, simulation, bootstrap, etc. It also involves capacity building. Achievements, usage and impact is summarized in the sequel.
Capacity Building

- 776 researchers of NARS have been trained through 37 training programmes of one week duration each. With this the number of researchers trained has gone upto 1672 through a total of 80 training programmes. Out of these 37 training programmes in 2011-12, 07 were organized by IASRI, New Delhi and rest 30 by consortium partners. 13 of these training programmes were organized at doorsteps of users such as at Pt. Deen Dayal Upadhyaya Veterinary University and Gau-Anusandhan Sansthan, Mathura; NDUAT, Faizabad; PAU, Ludhiana; GBPUA&T, Panthagar; CIRG, Makhdoom; CSUA&T, Kanpur; OUAT, Bhubaneshwar; BCKV, Kalyani; IGBK, Raipur; ICAR Research Complex for NEH Region, Manipur centre, Imphal; CPCRI, Kasargod; Karnataka Veterinary, Animal and Fisheries Science University, Bidar; RARS Tirupati. 09 of these training programmes were on specific topics such as Data Analysis of Natural Resource management Research, Genetics/Genomics Data Analysis using SAS; Data Analysis in Social Sciences Research; Data Analysis and Interpretation of Farm Implements and Machinery Research, Data Mining Using SAS; Data Analysis in Dairy Science; Analysis of Veterinary Science Data and Multivariate Data Reduction and Analysis.
- 100 scientists have been sensitized on Data Analysis Using SAS through FOCARS by NAARM, Hyderabad (a total of 258 scientists were sensitized).
- WebEx sessions on JMP Genomics 5.1 were arranged.
- To discuss the progress made, lessons learnt and future course of action for Strengthening Statistical Computing for NARS, two Partners’ meet were organized. It was decided that Statistical Computing hubs should facilitate the installation of the software at regional stations of different NARS organizations located in the same/nearby cities that of Statistical Computing Hubs. Nodal Officers from other NARS organizations may also be requested to help in this endeavour.

Updates, Upgrades and Installation

- Updates and upgrades were received. To sort out implementation issues and refinements in installation process, handing over of updates and upgrades and face to face interaction with nodal officers, second Workshop-cum-Installation training programmes at 08 statistical computing hubs except ICAR RC NEHR Barapani were organized. Updates and upgrades have been handed over to 128 NARS organizations. The software is installed on 1623 computers across NARS out of which 653 installations were done during the year.

Strengthened Indian NARS Statistical Computing Portal

- For providing service oriented computing, Indian NARS Statistical Computing portal has been established which is available to NARS users through IP authentication at http://stat.iiasri.res.in/sscnarsportal. Any researcher from Indian NARS may obtain User name and password from Nodal Officers of their respective NARS organizations, list available at www.iiasri.res.in/sscnars. Analysis of
data generated from any block design (complete or incomplete), augmented block designs, split plot design and combined analysis of block designs is available on this portal. Some screen shots showing the analysis of augmented block designs through portal are shown in above snapshots.

Following can also be accessed through IP authenticated networks:

- **Web Report Studio**: http://stat.iasri.res.in/sscnarswebreportstudio
- **BI Dashboard**: http://stat.iasri.res.in/sscnarsbidashboard
- **Public Page**: http://stat.iasri.res.in/sscnarsportal/public
- **E-Miner 7.1**: http://stat.iasri.res.in/SASEnterpriseMinerJWS/Status
- **E-Miner 6.1**: http://sas.iasri.res.in:6401/AnalyticsPlatform
- **Web OLAP Viewer**: http://sas.iasri.res.in:8080/sscnarswebolapviewer

### Macros Developed for Customized Analysis

- For customized analysis, macros for analysis of data generated from Split-split plot design; Split Factorial (Main A, Sub B x C) designs and econometric analysis (diversity indices, instability index, compound growth rate, Garret scoring technique and Demand analysis using LA-AIDS model) have been developed and made available on the project website.

### Sensitization of Researchers

- Website of the project is being maintained and updated regularly. Website registered under google analytics on November 15, 2010. Till March 31, 2012, there were 10153 page views across 228 cities of 60 countries. During April 01, 2011 – March 31, 2012, there are 8214 page views across 205 cities of 56 countries. Average time on page is 4.12 minutes.

- To sensitize the researchers about the availability of this high end statistical package, more than 20 presentations were made in training programmes/ Workshops/ Conferences/ Special Session at different NARS organizations.

- Nodal officers have been requested to provide a link of the project websites on LAN of their respective organizations.

### Usage and Impact

The capacity building efforts have paved the way for publishing research papers in the high impact factor journals.

- Based on feedback received from 142 NARS organizations, 98 research reports, 100 research papers have been published/ accepted for publication (in journals like Animal Food Science and Technology, Field Crops Research, Journal of Food Engineering, Euphytica, Journal of Applied Polymer Science, etc.) by analysing the data using high end statistical computing facility; 60 students have used this in their dissertations; 984 students have used in their course work. The software is installed on more than 1623 computers across NARS. There are 692 new users out of 1892 total users of this high statistical computing facility.

- Nodal officer from CMFRI, Kochi has reported in saving of 20 man months in compilation of data related to Marine Fish Household Census 2010 consisting of 10 lakh households with 16 attributes, On the consortium website, there are hits across 228 cities of 60 countries.

- The e-manual developed has been cited in Journal of Doctoral Research in Economics of the Bucharest Academy of Economic Studies. The macro developed for augmented designs has been cited by Jennifer Kling, Oregon State University in Introduction to Augmented Designs.

- Number of hits on Indian NARS Statistical Computing Portal (outside IASRI) since April 2011: 4587 (on an average more than 12 hits per day).

- It has been included in FOCARS training programme by NAARM, Hyderabad.

### Establishment of National Agricultural Bioinformatics Grid in ICAR

A genomic portal for submission of genomic data has been developed and this would be used for storage of Nucleotide, Genes, Genome, EST, GSS, SNP, RNA etc. Apart from this, number of other databases related to biological data have been extended and populated Cattle Genome Resource Information System.

Following four important research studies have been initiated in collaboration with partner institutions:

- Identification and characterization of genomic sequences responsible for salinity-stress in cereal crops-rice, sorghum, maize and wheat.

Analysis/Assessment of synonymous codon usage of Cytochrome P450 mono-oxygenase in agriculturally important insects.

In-Silico identification of genes responsible for late blight disease in potato.

A detail review regarding genomic resources in the field of agriculturally important insect has been done.

A study on analysis and functional annotation of EST’s of Water Buffalo has been conducted. In this, EST including functional annotation, detection of SSRs, pSNPs, protein domains, signal peptides in Bubalus bubalis are performed for 1825 EST sequences obtained from public domain.

A phylogenetic analysis and secondary structure prediction of 15 drought tolerant Cap binding proteins from different plant species was carried out. On the rectangular cladogram Mirabilis jalapa was nearest to the origin and is placed separately with Riginus communis forming separate cluster with root distance 0.019090 and pair distance 0.10204 with Riginus communis.

A review article of anti-microbial peptide and its role in agricultural biotechnology has been written.

Numbers of other studies in different area of research in bioinformatics have been initiated such as:

- Genome-Wide Analysis for Identification of Salt-Responsive Genes in Oryza Sativa.
- Functional Analysis of Salt-Responsive ESTs in Oryza Sativa
- Synonymous Codon Usage of Cytochrome P450 Monoxygenase (Cyps) in Agriculturally Important insects
- Functional Analysis of Salt-Responsive ESTs in Vitis Vinifera (Grapes)
Frequency of genes in different functional categories *Bubalus bubalis* and Venn diagram for shared gene with different organism.
Risk Assessment and Insurance Products for Agriculture

In order to characterization and mapping of all the districts based on various socio-economic parameters four indices were developed i.e. infrastructural index, health and sanitation index, nutritional index, economic status index for all 500 districts of the country. Finally, all these indices were integrated using data driven weights to formulate a livelihood index. Bio-physical index has been developed for all 500 districts of the country for assessment of agricultural potential. It is based on long term weather parameters and soil conditions etc. in collaboration with CRIDA. In order to assess the income risk at household level, Logistic regression model after incorporating survey weights has been developed. Also, to assess the yield risk at district level, Weather index based models have been developed. Further, Classification and Regression Technique (CART) has been applied on different weather parameters in Tamil Nadu to get various thresholds for yields in rice crop. This will be useful for development of new customized insurance products at local (district) level. The timely dissemination of information to the farmers is closely linked to the agricultural development and well being of the rural communities. Quick information transfer between the researchers and the farmers has specific importance. Hence, a prototype of comprehensive information and Online Decision Support System is developed for effective knowledge delivery for farm entrepreneur related to risk assessment and insurance product. The purpose of this system is to provide to farmers, insurance companies and policy makers for risk mitigation against uncertain risks like climate risks, production risks, etc. The system is divided into four major modules

- Farmers Module
- Insurance Company Module: Description for various insurance schemes to help and to provide further assistance to reduce crop risks.
- Policy Maker Module: Information about different policies being associated to the crop risk and the solutions for the farmers.
- Administrator Module: Administer all the tasks for DSS e.g. Login Process with roles.
**Farmers Window:** In this, the farmer has to provide information on the basis of his household characteristics like lighting source, cooking source, ration card information etc. Farmer also needs to provide information on the basis of his land attributes like land owned, land type, land irrigated etc. which are among the most important attributes for the crop insurance. Other attributes are farm assets and financial attributes. The farmer belongs to small marginal or upper marginal is based on his financial status. On the basis of these
particulars of the farmer, the crop insurance products which are available for framer for a particular crop on the basis of his profile will be displayed. The farmer will have option to choose the products from the alternatives which suits his risk coverage and finances. Payment of the premium amount is done after selection of payment mode and specific bank name. After this the farmer is redirected to the bank’s site for further processing. Further, farmer can directly go for the second insurance by clicking on add new crop insurance or log out from the system. On the further login a user will see all the previous insurance details along with the scenario to go for new one. The data flow diagram of the system is provided in figure.

**Administrator Window:** All the databases, models related to risk assessments and users profiles are to be administered through the administration window. Administrator window shows various user profiles for activation/de-activation. Information about all the insurance companies is available in order to activate or de-activate user accounts. As soon as specific user is activated, he is allowed to login into the system. A mail is sent to the user’s e-mail for the login details and activation or deactivation information. All data extraction module and model building exercise is in control of administration window.

**Insurance Company Window**
The entire list of farmers applied for different products of insurance products offered by insurance company will be displayed after successful login. The insurance agent can now see the full detail information of the farmer including risk profile at different level calculated from statistical models. The company can grant insurance or reject the application of the farmer which will be communicated through e-mail to the farmer.

**Policy Maker Window:** Policy maker is not an authorized person in order to enter the DSS and view various types of reports until and unless the administrator of the system allows him by setting his account as active. Policy makers are allowed to view different kinds of reports on the basis of data from the system. A crop wise report is generated to show the details for premium and compensation amount with respect to district, year and insurance product.

**Software for Survey Data Analysis (SSDA) 2.0**
A web based software SSDA 2.0 has been developed for survey data analysis for stratified multistage sampling
design. The home page of SSDA 2.0 server is shown in figure.

Home page of SSDA 2.0

It has links such as contact us, upload file, download file, delete file, and download test data. Some of the important features of the software are: New user registration and editing user profile, individual data storage folder for imputation and analysis under name My Folder, Feedback, Help Manual, Extraction of NSSO Data, Calculation of Summary Statistics, Scrutiny and Editing of Outlier Data, Sample Selection, Imputation of Missing Data using mean zero and mean of neighboring unit methods, Sampling Weight Calculations and Estimates of Parameter. SSDA 2.0 is capable of extracting the NSSO data. This extraction program module takes the required text file as well as the meta-data defining the positions of relevant input variables.

SSDA 2.0 performs the survey weights up to stratified three stage sampling and its subsets. The sampling designs incorporated in the software are SRSWR and SRSWOR and systematic sampling under equal probability and PPSWR under unequal probability. Figure below shows the estimated results for mean, total and variances. It also includes estimates for ratio, sub-population and domain.

SSDA 2.0 computes the estimates of parameters for any sampling design if the survey weights are supplied by the user.

Half-Yearly Progress Monitoring System of the Scientists in ICAR (HYPM)

For Half-Yearly Progress Monitoring (HYPM) of the Scientists in ICAR, a web based software has been designed and developed to ensure more objective evaluation of the half-yearly performance of scientists.
Research Achievements

in ICAR. The HYPM system has been implemented from 01 April 2012 for online submitting the proposed targets by the scientists for the first half year period (01.04.2012 to 30.09.2012). It is launched from IASRI server and made available at http://hypm.iasri.res.in.

For effective implementation of HYPM from all the institutes of ICAR, the PME Cell I/Cs have been nominated as Nodal Officers of their respective institute. The Nodal officers are responsible for data management and customization of HYPM to maintain the website from their respective institute.

ARS (ICAR) scientists posted at any one of the ICAR institutes have access to HYPM software through his/her unique User-IId and Password. Facility has been provided to enter proposed targets for the coming half-year and achievements of the completed half-year independently with respect to Research, Teaching, Training, Extension and Other Prioritized Activities. Finally, the scientist can submit the proposed targets to the concerned Reporting Officer through the option Submit Targets as visible in the home page of scientist.

The Reporting Officer (Head of Division/Regional Station) has access to the Proposed Targets and Achievements details submitted by all concerned scientists. He/she may add his/her remarks and give recommendations on the basis of the progress reports/inputs submitted by the concerned scientists using the option as visible in home page of Reporting Officer. In addition he/she can also submit his/her own proposed targets.

Reviewing Officer has dual facilities as he/she may be the Reporting Officer for some scientists like Head of Divisions and Reviewing Officer for other scientists. The Reviewing Officer is able to add his/her own assessment remarks and final overall grading on the Proposed Targets and Achievements of all scientists.
As per requirement of HYPM, facility to the Director of the Institute to submit Institute Level General Information containing salient achievements of the institute for the completed half year period (as shown in the following image) has been provided.

For DG/SMD/ICAR level of monitoring progress of the scientists at different institutes, various reports are generated for the proposed targets status as submitted by the scientists and comments of the Reporting/Reviewing Officers. These reports include target submission status reports with facility to view individual level proposed targets of the scientist and other options like manpower status, research projects, and salient research achievements of the institutes through the options under Report Module.

For effective implementation of HYPM from all the institutes of ICAR, Nodal Officers level User-Id and Password have been issued to customize HYPM for implementation from their respective institutes. HYPM system has been made available to all the scientists w.e.f. 15 March 2012. UserId and Password to all SMDs have also been issued for monitoring the progress through HYPM.
Development of Web Enabled Statistical Package for Agricultural Research (SPAR 3.0)
SPAR 3.0 is a web enabled software package that has been designed and developed for the statistical analysis of experimental research data in Plant Breeding and Genetics. It has been developed using Microsoft .NET technology (ASP.NET with C#). It consists of all the modules of data analysis which are available in SPAR 2.0. It has modules on Descriptive Statistics, Estimation of Breeding Values (General Mean and Scaling Tests), Correlation and Regression Analysis and Path Analysis, Variance and Covariance Components Estimation, Stability Analysis, Multivariate Analysis (Cluster Analysis, Discriminant Analysis and Principal Component Analysis), Mating Designs (Complete Diallel, Partial Diallel, Line x Tester - with and without parents, Three way cross, Double cross and North Carolina Designs I, II, III). It has Complete Online Help with Contents, Index, Search and Favorites facilities.

National Information System on Agricultural Education Network in India (NISAGENET)
The NISAGENET web portal is being maintained at the Central Server of IASRI, New Delhi and is accessible at http://www.iasri.res.in/Nisagenet/. The database of this system contains the information on various aspects such as Academic data of the universities, Infrastructural facilities, Budget provision, Manpower employed, Faculty and R&D activities. Moreover, it has an exhaustive Query/Reports system to provide information at Country, State, University and college levels as well. Keeping in view the importance and utility of NISAGENET, ICAR has approved it to maintain as a regular ongoing activity of the Council. For maintaining up to date information, 19 Agricultural Universities established in the recent past have been added. The operational architecture has been modified to three tier web architecture and now it is possible to directly enter/update data from university/college(s). The HRD data with regard to students Admitted/Passed and Faculty/Administrative manpower for the years 2009-10 and 2010-11 uploaded by the universities is available in the form of reports. The NISAGENET acts as a single Window Information Delivery System and is an effective solution to check overlapping and parallel flow of information from the same university, but from different sources.

Management System for Post Graduate Education
This is a web enabled system for management of various activities of PG School of IARI. The system caters to the requirements of different users: Dean, Registrar, Professors, Heads, Guides, Faculty, Teachers, Students, Administrators and Officials for performing their assigned tasks. There are 5 modules: Courses Management, Faculty Management, Student Management, Administration Management and e-Learning.

Course Management module has various menu items which provide various facilities. The professor has access to add/delete/update courses, offer courses for each trimester, allocate courses to faculty, and allot students to guides. The course registration is accomplished by approval of students registered courses by course leader/instructors, guide, and professors. The Course Leader may declare class schedule, examination scheme and dates, and result. The Professor can suggest examiners for qualifying examination and thesis evaluation.

Faculty Module is meant to be used by teachers, guides, professors, heads, and dean. It provides them with the opportunity to perform various tasks as per their authorization such as approval of student courses, progress report, PPW and ORW and offering of trimester courses, allocation of courses to course leaders and instructors and submission of student examination grades etc. Dean can view current status of each activities of PG School and initiate actions/issue instructions for the pending tasks of guides/professors/heads. Dean may also approve various forms/results/reports after necessary approval by faculty/guides/professors/heads.
Administration Management module provides features for Administrator, Dean Office, and Registrar etc. The requests of new users, students, teachers, faculty, guides and professors are received by Administrator and approved after verifying the details. Administrator may also start and stop trimester registration, change student status as Pass Out/Current/Left Out. Through e-learning component, course instructors may attach, 5 files for Lecture Notes, Presentation Slides, Assignment, References and other resources for each topic.

The system has a strong reporting module to cater to the requirements of various functionaries. The system also has a provision of alert mechanism wherein the pending work is displayed on the homepage of the users. For important events emails are also sent automatically to the respective faculty. The system is available at http://pgs.iasri.res.in and is in use by PG School IARI since the year 2009-10. The system is ready for adoption by other deemed universities of ICAR for which requests have already been received.

**Project Information & Management System of ICAR (PIMS-ICAR)**

Project Information & Management System of ICAR (PIMS-ICAR) designed, developed and implemented at IASRI, New Delhi at http://pimsicar.iasri.res.in/ to help in taking decisions to check duplication in research projects both at divisional as well as inter divisional level of ICAR. PIMS-ICAR has also been integrated with Half Yearly Progress Monitoring of Scientists (HYPM) system developed and implemented for all the ICAR institutes. The integration has facilitated the visibility of Research Projects details of ongoing projects with respective PIs and Co-PIs in HYPM. As per the data entry status available in PIMS-ICAR, the ICAR institutes have initiated project data entry process for more than 5110 ongoing and 5150 completed projects into PIMS-ICAR from their respective institutes. Process for digitization and creation of repositories of RPFs-III is in progress. The RPF-III of 3740 projects has already been uploaded by institutes and is available in PIMS-ICAR.

**Phenomics of Moisture Deficit and Low Temperature Stress Tolerance in Rice**

Phenomics database for rice has been designed. The database is scalable and can act as model database for other crops also. The application has been designed to capture meta data and experimental data about different experiments. Utility has been developed to upload data from the excel files. The application is based on n-tier Web architecture and is using JAVA, JSP and my SQL database. In addition to this LIRE image analysis API has been studied and tested with crop disease images.