

2014-15

अंक  
10

# सांख्यिकी-विमर्श



भा.कृ.अनु.प.-भारतीय कृषि सांख्यिकी अनुसंधान संस्थान  
लाइब्रेरी एवेन्यू, पूसा, नई दिल्ली-110012



2014-15

अंक  
10

# सांख्यिकी-विमर्श

संपादक मंडल

अशोक कुमार गुप्ता

अनिल कुमार द्विजेश चन्द्र मिश्रा

सन्तोष कुमार सिंह विजय बिन्दल

संतोष कुमार ब्रह्मजीत गहलौत

ऊषा जैन

कलाकार

अमर रंजन पॉल



भा.कृ.अनु.प.-भारतीय कृषि सांख्यिकी अनुसंधान संस्थान  
लाइब्रेरी एवेन्यू, पूसा, नई दिल्ली-110012



मुद्रण : मार्च, 2015



निदेशक

भा.कृ.अनु.प.-भारतीय कृषि सांख्यिकी अनुसंधान संस्थान  
लाइब्रेरी एवेन्यू, पूसा, नई दिल्ली  
की ओर से प्रकाशित



## निदेशक की कलम से .....

वर्ष 2005-06 में संस्थान में हिन्दी पत्रिका 'सांख्यिकी-विमर्श' के प्रकाशन का शुभारम्भ हुआ और तब से अब तक प्रतिवर्ष इस पत्रिका का प्रकाशन किया जा रहा है। प्रस्तुत अंक इस पत्रिका का दसवाँ अंक है। संस्थान की हिन्दी पत्रिका "सांख्यिकी-विमर्श 2014-15" आपके समक्ष प्रस्तुत करते हुए मुझे हार्दिक प्रसन्नता की अनुभूति हो रही है।

इस अंक में संस्थान के कीर्तिस्तम्भ, संस्थान द्वारा किये गये अनुसंधानों व अन्य कार्यों के संक्षिप्त विवरण, संस्थान में राजभाषा से सम्बन्धित कार्यों आदि की जानकारी के साथ-साथ कृषि सांख्यिकी, कृषि में संगणक अनुप्रयोग एवं कृषि जैव-सूचना से सम्बन्धित विभिन्न लेखों एवं शोध-पत्रों को इस पत्रिका में सम्मिलित किया गया है। हमारे संस्थान में प्रतिवर्ष डॉ. दरोगा सिंह स्मृति व्याख्यानमाला के अन्तर्गत हिन्दी में एक वैज्ञानिक व्याख्यान का आयोजन किया जाता है। इस वर्ष यह व्याख्यान संस्थान की पूर्व प्रमुख वैज्ञानिक, डॉ. रंजना अग्रवाल जी द्वारा दिया गया था जिसे आमंत्रित ज्ञानवर्धक लेख के रूप में पत्रिका में सम्मिलित किया गया है। अन्त में पाठकों के हिन्दी ज्ञानवर्धन के लिए दैनिक स्मरणीय शब्द-शतक हिन्दी व अँग्रेजी में दिया गया है।

मैं पत्रिका के प्रकाशन के लिए उन सभी लेखकों का आभारी हूँ जिन्होंने इस पत्रिका में प्रकाशन हेतु अपने लेख देकर हमारे इस प्रयास को सफल बनाने में हमारा सहयोग किया। विशेष रूप से मैं डॉ. रंजना अग्रवाल जी का आभार व्यक्त करता हूँ जिन्होंने अत्यन्त व्यस्त होते हुए भी अपना अमूल्य समय देकर अपने डॉ. दरोगा सिंह स्मृति व्याख्यान को इस पत्रिका के लिए लेख के रूप में परिवर्तित किया। पत्रिका के प्रकाशन के लिए सम्पादक मंडल के अध्यक्ष, डॉ. अशोक कुमार गुप्ता एवं सदस्यों, डॉ. अनिल कुमार, डॉ. द्विजेश चन्द्र मिश्रा, श्री संतोष कुमार सिंह, सुश्री विजय बिन्दल, श्री संतोष कुमार, श्री ब्रह्मजीत गहलौत तथा सुश्री ऊषा जैन का मैं आभार व्यक्त करता हूँ जिनके अथक प्रयासों से यह पत्रिका इस रूप में आपके समक्ष आ सकी। पत्रिका का कवर-पेज पृष्ठ एवं लेआउट प्रदान करने के लिए संस्थान के पूर्व वरिष्ठ आर्टिस्ट, श्री अमर रंजन पाल धन्यवाद के पात्र हैं।

आशा है इस अंक की विषय-वस्तु पाठकों के लिए सूचनाप्रद एवं उपयोगी सिद्ध होगी और सांख्यिकी जैसे तकनीकी विषय में भी हिन्दी साहित्य का प्रयोग करके पाठकों का ज्ञानवर्धन करने में सहयोगी सिद्ध होगी। इसके भावी अंकों में सुधार के लिए आपके विचारों तथा सुझावों की प्रतीक्षा रहेगी।

(उमेश चन्दर सूद)  
निदेशक



## अनुक्रमणिका

संस्थान के कीर्तिस्तम्भ	1
प्रोफेसर बाल बी.पी.एस. गोयल – अशोक कुमार गुप्ता एवं विजय बिन्दल	
संस्थान में प्रगति के बढ़ते चरण	5
डॉ. दरोगा सिंह स्मृति व्याख्यान फसलों की उपज के कटाई-पूर्व अनुमान हेतु मौसम आधारित मॉडल – रंजना अग्रवाल	14
भा.कृ.अनु.प. में एकीकृत संचार एवं वेब होस्टिंग – एक अवलोकन – अंजनी कुमार चौबे, अलका अरोड़ा, सुदीप मरवाह, एन.एस. राव, मुकेश कुमार, राकेश सैनी, सुभाष चन्द एवं ऊषा जैन	20
सहसंबंध एवं समाश्रयण – अशोक कुमार गुप्ता, अंकुर बिस्वास एवं विजय बिन्दल	25
मछलियों के मौसमी विकास के लिए साइन वेब फंक्शन के साथ संशोधित गोम्पर्टज मॉडल – अमृत कुमार पाल, नाओरेम ओकेन्द्रो सिंह, रंजीत कुमार पॉल, बिशाल गुरुंग, वसी आलम एवं सत्यपाल सिंह	34
विषम चक्रीय आँकड़ों के पूर्वानुमान के लिए फूरियर स्वसमाश्रयी (एफ-एआर) गुणांक अरैखिक काल-श्रृंखला प्रतिमान-एक व्यष्टि अध्ययन – हिमाद्रि घोष, सविता वधवा एवं प्रज्ञेषु	39
सी.बी.पी. वोटल : भा.कृ.अनु.प. द्वारा प्रायोजित प्रशिक्षण कार्यक्रम के ऑनलाइन प्रबन्धन हेतु सूचना तंत्र – अलका अरोड़ा, सुदीप मरवाह, कुसुमाकर शर्मा, प्रदीप कुमार मल्होत्रा, ए.के.एम. सामिमुल आलम, पुष्पांजलि रावत एवं ऊषा जैन	43
भारत में सरसों की वस्तुओं में बाजार समाकलन – प्रवीण आर्य, धर्मराज सिंह, अनिल कुमार, कमलेश नारायण सिंह, रंजीत कुमार पॉल एवं उषा जैन	46
उपग्रह छाया चित्र में बादल आवरण के अन्तर्गत फसल क्षेत्र आकलन की कार्यप्रणाली विकसित करने के लिये अध्ययन – प्राची मिश्रा साहू, तौकीर अहमद, कमलेश नारायण सिंह, अशोक कुमार गुप्ता एवं नीलम चन्द्रा	52
एस.पी.एस.एस. का उपयोग कर मिश्रित मॉडल का विश्लेषण – सुकान्त दाश, सुशील कुमार सरकार, वैद्यनाथ मण्डल एवं देवेन्द्र कुमार	58

अपेक्षित उत्पादन के लिए प्रक्रिया चर के साथ मिश्रण प्रयोगों में अनुकूलतम स्थितियाँ	65
— उपेन्द्र कुमार प्रधान, कृष्ण लाल, विनोद कुमार गुप्ता, प्रबिण कुमार मेहेर एवं विजय पाल सिंह	
भारत में स्टिक—लाख के पूर्व अनुमान के लिये स्टोकास्टिक मॉडल	71
— बिशाल गुरुंग, अमृत कुमार पॉल, रंजीत कुमार पॉल एवं कमलेश नारायण सिंह	
भैंस में नए गुण—बिन्दुओं की खोज	75
— सारिका, मीर आसिफ इकबाल, ऊषा जैन, अनिल राय एवं दिनेश कुमार	
अरहर दाल के खुदरा मूल्य का पूर्वानुमान	78
— रंजीत कुमार पॉल, बिशाल गुरुंग, अमृत कुमार पॉल, वसी आलम एवं कंचन सिन्हा	
एफिड सरसों के प्रबंधन के लिये श्रेणीबद्ध परीक्षण प्रक्रियाएँ	80
— वसी आलम, अमृत कुमार पॉल, रंजीत कुमार पॉल, कंचन सिन्हा एवं राम नरेश	
प्रत्याशित—मान प्राचलों का उपयोग करते हुए आर्डर वन के स्वसमाश्रयी में फॉक्स मॉडल का संयोजन	83
— अमृत कुमार पॉल, नाओरेम ओकेन्द्रो सिंह, रंजीत पॉल, बिशाल गुरुंग एवं ऊषा जैन	
रन अनुक्रमों में न्यूनतम स्तर परिवर्तन सहित द्विस्तरीय बहु—उपादानाणी परीक्षण	89
— अर्पण भौमिक, एल्दो वरगीस, सीमा जग्गी, सिनी वरगीस, कादर अली सरकार एवं विजय बिन्दल	
कृत्रिम तंत्रिका नेटवर्क एवं पारंपरिक वर्गीकरण पद्धतियों का तुलनात्मक अध्ययन	94
— सुकान्त दाश, सुशील कुमार सरकार, देवेन्द्र कुमार एवं ज्ञान सिंह	
भारत के समुद्रीय और पशुधन के निर्यात उत्पादन में वाष्पशीलता और मुद्रीता	97
— बिशाल गुरुंग, कमलेश नारायण सिंह, रंजीत कुमार पॉल एवं लॉरेंस लेपचा	
भारत में पशुधन एवं डेरी उत्पादन की स्थिति	102
— वसी आलम, रंजीत कुमार पॉल, अमृत कुमार पॉल, बिशाल गुरुंग, कंचन सिन्हा, कमलेश नारायण सिंह, उषा जैन एवं राम नरेश	
गुणात्मक और मात्रात्मक आंकड़ों के मिश्रण पर आधारित जर्मप्लाज़म का एक कोर सेट के विकसित करने की विधि	106
— उदय प्रताप सिंह, आत्माकुरी रामाकृष्णा राव, सन्त दास वाही, प्रबिण मेहेर एवं रूपम कुमार सरकार	
अर्द्ध—सतत डाटा हेतु लघु क्षेत्र आकलन	117
— हुकुम चन्द्र एवं विजय बिन्दल	
संस्थान की राजभाषा यात्रा : 2014—15	126
— ऊषा जैन	
दैनिक स्मरणीय शब्द शतक	129

### आवश्यक सूचना

इस अंक में प्रकाशित रचनाओं में व्यक्त विचारों/आंकड़ों आदि के लिए लेखक स्वयं उत्तरदायी हैं।

## संस्थान के कीर्तिस्तम्भ

### प्रोफेसर बाल बी.पी.एस. गोयल अशोक कुमार गुप्ता एवं विजय बिन्दल



प्रोफेसर बाल बी.पी.एस. गोयल का जन्म २८ अगस्त, १९३७ को रोहतक, हरियाणा में हुआ था। आपने सन् १९५७ में दिल्ली विश्वविद्यालय से गणित विषय में स्नातक होनर्स (Honours) की उपाधि प्राप्त की। आपने सन् १९५६ में सांख्यिकी विषय में एवं सन् १९६४ में गणित विषय में दिल्ली विश्वविद्यालय से स्नातकोत्तर की उपाधियाँ प्राप्त की। आपने सन् १९७४ में विभागीय उम्मीदवार के रूप में भारतीय कृषि अनुसंधान संस्थान (पी.जी.) स्कूल, नई दिल्ली से कृषि सांख्यिकी में पीएच.डी. की उपाधि प्राप्त की जिसमें आपको स्वर्ण पदक (Gold Medal) प्रदान किया गया। बचपन से ही आप बहुत कुशाग्र बुद्धि के विद्यार्थी थे तथा अपने सहपाठियों की पढ़ाई में सहायता किया करते थे।

प्रोफेसर बाल बी.पी.एस. गोयल ने अपने व्यवसायिक जीवन में अनेक राष्ट्रीय एवं अन्तरराष्ट्रीय संगठनों में महत्त्वपूर्ण पदों पर कार्य किया। आपने सन् १९६० में भारतीय कृषि सांख्यिकी अनुसंधान संस्थान (भा.कृ.सां.अ.सं.), नई दिल्ली में जिसका उस समय नाम कृषि सांख्यिकी अनुसंधान संस्थान (Institute of Agricultural Research Statistics) था, में सांख्यिकीय सहायक के पद पर प्रतिदर्श सर्वेक्षण प्रभाग में कार्य भार गृहण किया। तदोपरान्त, दो वर्षों से भी कम समय में संघ लोक सेवा आयोग (UPSC) द्वारा संचालित परीक्षा में सांख्यिकी अन्वेषक (Statistical Investigator) के रूप में चयनित हुए। सन् १९६० से १९७१ के दौरान भा.कृ.सां.अ.सं., नई दिल्ली में सांख्यिकी विषय में विभिन्न अनुसंधानिक एवं अध्यापकीय पदों पर कार्य किया। प्रोफेसर गोयल ने भा.कृ.सां.अ.सं., नई दिल्ली में सन् १९७१ से १९७७

तक सांख्यिकीविद्-एसोसिएट प्रोफेसर (Statistician-Associate Professor) के पद पर तथा सन् १९७७ से १९८३ के दौरान वरिष्ठ वैज्ञानिक, कृषि सांख्यिकी के पद पर कार्य किया। आपने सन् १९८३ में योजना आयोग (Planning Commission) में उपसलाहकार, परिपेक्ष्य योजना (Deputy Advisor-Perspective Planning) के पद पर कार्य करने के उपरान्त, आप सन् १९८३-८४ में भा.कृ.सां.अ.सं., नई दिल्ली में संयुक्त निदेशक के पद पर तथा सन् १९६६-६७ के दौरान निदेशक के पद पर आसीन रहे। निदेशक के पद पर कार्य करते हुए प्रोफेसर गोयल ने केवल रु. १/- (Rupee One only) प्रति माह वेतन पर संस्थान की निःस्वार्थ सेवा की जो अति सराहनीय है।

अन्तरराष्ट्रीय स्तर पर प्रोफेसर गोयल ने सन् १९८४-८६ के दौरान काठमांडू, नेपाल में कृषि सांख्यिकीविद्, खाद्य एवं कृषि संगठन (Food and Agriculture Organization), संयुक्त राष्ट्र (United Nations) द्वारा संचालित फसल पूर्वानुमान, पूर्व चेतावनी एवं कृषि सांख्यिकी में सुधार परियोजना (Crop Forecasting, Early Warning and Improvement of Agricultural Statistics Project) पर कार्य किया तथा सन् १९८६-८७ में प्रभारी एवं सलाहकार, कृषि सांख्यिकी के पद पर कार्य किया। प्रोफेसर गोयल ने सन् १९८७-८८ में सलाहकार, कृषि सांख्यिकी, पी-४ (Advisor, Agricultural Statistics, P-4), सन् १९८८ से १९९५ के दौरान वरिष्ठ सलाहकार, कृषि सांख्यिकी, पी-५, स्टेप-८ (Senior Advisor, Agricultural Statistics, P-5, Step-8) के रूप में खाद्य एवं कृषि



संगठन, संयुक्त राष्ट्र द्वारा संचालित Southern Africa Development Community (एसएडीसी)/एफएओ परियोजना "खाद्य सुरक्षा के लिए क्षेत्रीय पूर्व चेतावनी प्रणाली" (Regional Early Warning System for Food Security) पर, जो डेनमार्क द्वारा वित्तपोषित थी तथा 92 देशों को कवर कर रही थी एवं जिसका मुख्यालय हरारे, जिम्बाबे में था, में भी कार्य किया। प्रोफेसर गोयल ने एसएडीसी देशों में खाद्य फसलों, फसल पूर्वानुमान, फसल उत्पादन के अनुमान से सम्बन्धित कृषि सांख्यिकी के तरीकों का शोधन (Refinement of Agricultural Statistics Methodologies relating to Food Crops, Crop Forecasts, Crop Production Estimates in SADC Countries) किया। आपने फसल उत्पादन के पूर्वानुमान और निगरानी के लिए डेटा संग्रह एवं विश्लेषण के लिए विकास और/या प्रक्रियाओं के समेकन पर (Development and/or consolidation of methodologies procedures) एसएडीसी देशों में राष्ट्रीय प्रारंभिक चेतावनी इकाइयों (National Early Warning Units) को सलाह दी। प्रोफेसर गोयल ने खाद्य सुरक्षा बुलेटिनों, उनके अद्यतन एवं रिपोर्ट लिखने में भी भाग लिया तथा सन् १९८८ से १९९५ के बीच त्रैमासिक खाद्य सुरक्षा बुलेटिनों के २० मुद्दों (Issues) में सह-लेखक के रूप में भी कार्य किया। आपने एसएडीसी देशों के तकनीकी कर्मचारियों के लिए लंबे समय तक प्रयोग के लिए खाद्य सुरक्षा से सम्बन्धित फसल पूर्वानुमान और पूर्व चेतावनी प्रणाली से सम्बन्धित पद्यति मैनुअल भी तैयार किए।

आपने भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली में निदेशक के पद को स्वीकार करने के लिए अपने मौद्रिक लाभ को छोड़कर खाद्य एवं कृषि संगठन से स्वेच्छा से सेवानिवृत्ति ली। भा.कृ. सां.अ.सं., नई दिल्ली में निदेशक के पद पर जनवरी १९९६ से अक्टूबर १९९७ तक २२ माह कार्य करते हुए प्रोफेसर गोयल ने केवल रु. १/- प्रति माह वेतन पर केवल रु. २२/- टोकन वेतन (Token Salary) लिया और संस्थान की निःस्वार्थ भाव से सेवा की जो अति सराहनीय है। निदेशक के सामान्य कर्तव्यों एवं जिम्मेदारियों के अलावा संस्थान की प्रतिष्ठा के सुधार

के लिए आपने कड़ी मेहनत की और उसमें सफलता भी प्राप्त की।

भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली के शिक्षण एवं प्रशिक्षण कार्यों में आपने विशेष रुचि लेते हुए वरिष्ठ प्रमाण पत्र कोर्स, भारतीय कृषि अनुसंधान संस्थान पी. जी. स्कूल, नई दिल्ली के एम.एससी. और पीएच. डी. के विद्यार्थियों को विभिन्न विषयों में शिक्षण कार्य किया। आपने संस्थान के अध्यापन कार्यों में रुचि लेते हुए दिये गये किसी भी विषय/कोर्स को पढ़ाने में कोई संकोच नहीं किया तथा पूरी निष्ठा एवं तन्मयता से शिक्षण कार्य किया। आपने अपने सेवाकाल के दौरान प्रतिचयन तकनीकों एवं कृषि में उनका अनुप्रयोग से सम्बन्धित कई विषयों पर शोध, अध्यापन एवं विद्यार्थियों का अनुसंधान मार्गदर्शन (Research Guidance) किया। प्रोफेसर गोयल ने लगभग ४ दशकों तक खाद्य फसलों एवं पशुधन पर पायलट एवं वृहत् स्तर पर प्रतिदर्श सर्वेक्षणों की योजना बनाने, निष्पादन करने, आंकड़ों का विश्लेषण करने तथा परिणामों की व्याख्या, प्रसार और रिपोर्ट तैयार करने का भी कार्य किया। प्रोफेसर गोयल ने भारत में पंचवर्षीय विकास योजनाओं के निर्माण एवं मूल्यांकन के लिए राष्ट्रीय एवं उपराष्ट्रीय कृषि सांख्यिकी का समन्वय किया। आपने भारतीय कृषि सांख्यिकी संस्था में विभिन्न पदों पर कार्य करते हुए अनेक सम्मेलनों का आयोजन एवं उनमें सहभागिता की। प्रोफेसर गोयल को राष्ट्रीय एवं अन्तरराष्ट्रीय परियोजनाओं में प्रशासनिक, वित्तीय एवं अनुसंधानिक प्रबन्धन का भी अनुभव है।

भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली में कार्य करते हुए प्रोफेसर गोयल को भारत के विशिष्ट इलाकों (Typical tracts) में गायों (Bovines) की विभिन्न नस्लों की हद (Extent of different breeds), भोजन और प्रबन्धन के तरीकों (Feeding and management practices) के अध्ययन के लिए पायलट प्रतिदर्श सर्वेक्षणों के कराने की जिम्मेदारी दी गई। इन सर्वेक्षणों के आधार पर विकसित की गई प्रणाली को अपनाने और नियमित आधार पर इन आंकड़ों के संग्रहण करने के लिए, राज्यों में पशुपालन विभागों को

सिफारिश (Recommendation) की गई। यह वैज्ञानिक आधार पर गायों के दूध उत्पादन के अनुमान देने का प्रथम प्रयास था। उसी समय में, इसी प्रकार के सर्वेक्षण जैसे अंडा उत्पादन और पोल्ट्री प्रथाओं (Practices) का अध्ययन, ऊन उत्पादन और भेड़ प्रथाओं एवं मांस उत्पादन/प्रथाओं के आकलन के लिए विभिन्न सर्वेक्षण किए गए। इन सर्वेक्षणों के परिणामों ने एक लम्बे समय से रह रही इनके आकलन की विधियों की आवश्यकता को पूर्ण कर दिया।

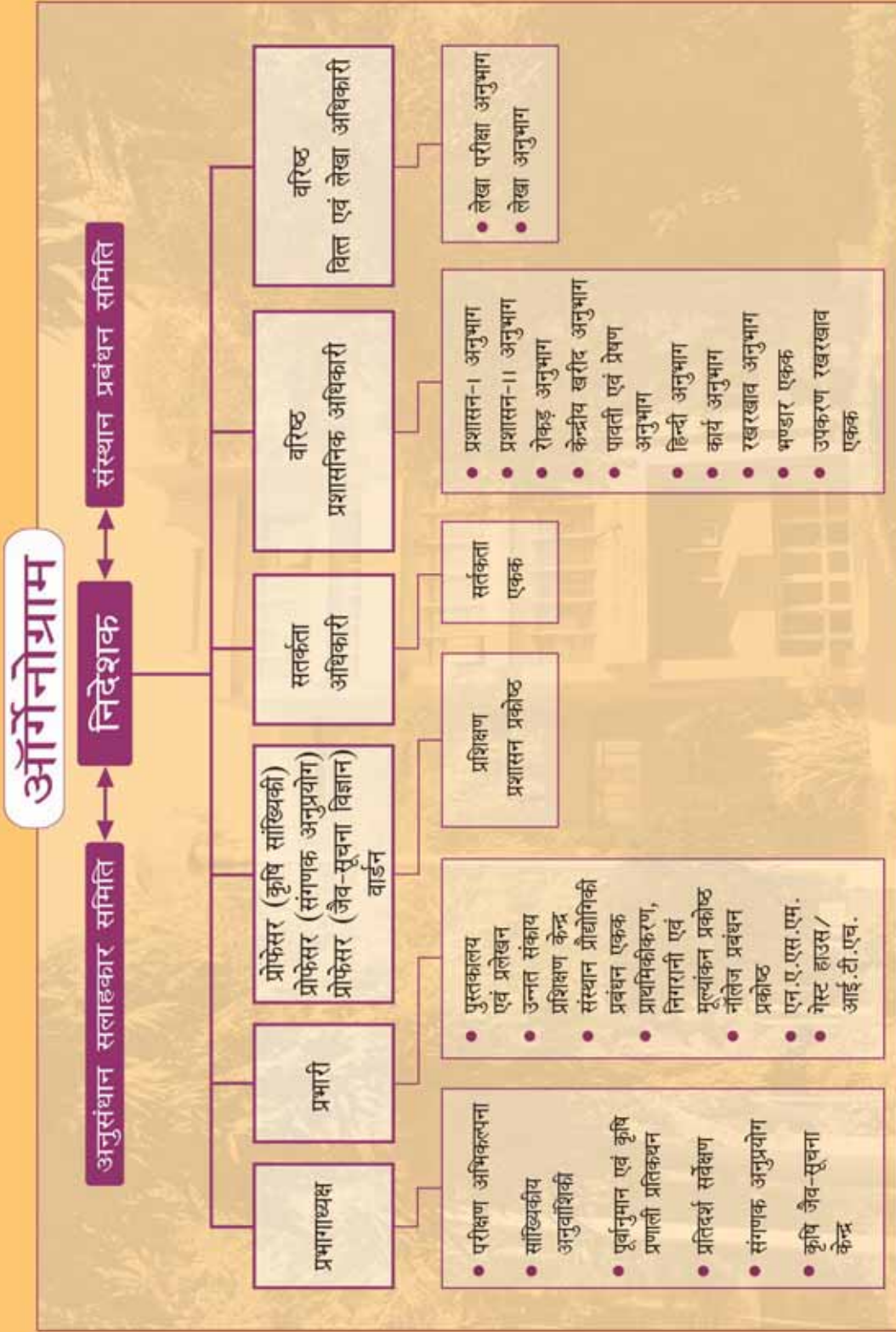
प्रोफेसर गोयल ने १०० से भी अधिक शोध-पत्र, पुस्तकें, मोनोग्राफ्स, परियोजना एवं शोध रिपोर्ट, लोकप्रिय लेख और खाद्य सुरक्षा बुलेटिन प्रकाशित किए। इनके द्वारा लिखे गए शोध-पत्र राष्ट्रीय एवं अन्तरराष्ट्रीय स्तर पर ख्याति प्राप्त जर्नलों में प्रकाशित हैं। प्रोफेसर गोयल ने भारत एवं विदेशों में आयोजित अनेकों सम्मेलनों, कार्यशालाओं और सेमिनारों में भी बहुत से शोध-पत्र प्रस्तुत किए हैं। आपने लगभग १० एम. एससी. और पीएच. डी. के विद्यार्थियों के अनुसंधान कार्य का पर्यवेक्षण किया। आपके द्वारा पर्यवेक्षण किए गए विद्यार्थियों ने अपना अनुसंधान कार्य प्रतिदर्श तकनीकों (Sampling techniques) में पूर्ण किया। आपने सेवानिवृत्ति से पहले और बाद में एफएओ तथा विकासशील 25-30 देशों में भ्रमण किया तथा कृषि सांख्यिकी की आवश्यकताओं के बारे में शोध करके भा.कृ.सां.अ.सं. का प्रचार किया। भा.कृ.सां.अ.सं., नई दिल्ली के कई वैज्ञानिकों ने इन देशों का भ्रमण कर उनके अधिकारियों को कृषि सांख्यिकी की विभिन्न तकनीकों से अवगत कराया तथा संस्थान का नाम विदेशों में रोशन किया। भा.कृ.सां.अ.सं. में इन देशों ने प्रशिक्षण के लिए अपने अधिकारियों एवं कर्मचारियों

को भेजा ताकि वे कृषि सांख्यिकी में विकसित की गई प्रविधियों का अपने देश में प्रयोग कर सकें।

प्रोफेसर गोयल भारतीय कृषि सांख्यिकी संस्था, नई दिल्ली के सन् १९७३ से १९८३ तक संयुक्त सचिव, सन् १९८३-८४ में कोषाध्यक्ष, सन् १९६६-६७ में अवैतनिक सचिव थे। आप भारतीय कृषि सांख्यिकी संस्था, नई दिल्ली के सन् १९६८ से सम्पादक मंडल के अध्यक्ष हैं। प्रोफेसर गोयल सन् १९६६-६८ के दौरान भारतीय कृषि विज्ञान जर्नल, नई दिल्ली के सम्पादक मंडल के सदस्य भी थे। आप भारतीय कृषि सांख्यिकी संस्था, नई दिल्ली एवं भारतीय विज्ञान कांग्रेस एसोसिएशन के आजीवन सदस्य भी हैं। प्रोफेसर गोयल सन् १९८५ से अन्तरराष्ट्रीय सांख्यिकी संस्थान, नीदरलैंड एवं गवर्निंग परिषद (Governing Council), सामाजिक-आर्थिक अनुसंधान केन्द्र, नई दिल्ली के सदस्य हैं।

इन सबके अतिरिक्त प्रोफेसर गोयल एक अत्यन्त सरल प्रवृत्ति वाले व्यक्ति हैं और सभी से अपनत्व से बात करते हैं। अपने विद्यार्थियों एवं स्टाफ के सदस्यों को वे उसी प्रकार स्नेह व संरक्षण देते थे जैसे एक पिता अपने बच्चों को देते हैं। आपको अपनी योग्यताओं एवं उपलब्धियों पर कभी भी अभिमान नहीं हुआ। संस्थान में आयोजित विभिन्न प्रकार की शोध गतिविधियों में आप आज भी भाग लेते हैं। अब भी आपका अधिकतर समय प्रतिदर्श तकनीकों के शोध कार्य में व्यतीत होता है। प्रोफेसर गोयल का व्यक्तित्व आज भी हम सबके लिए अनुकरणीय है। वास्तव में यह कहना अतिशयोक्ति नहीं होगा कि उनका कद उनके पद से ऊँचा है। वे सदैव सबके हितैषी और भला चाहने वाले रहे हैं और भविष्य में भी रहेंगे।

□



## संस्थान में प्रगति के बढ़ते चरण

भारतीय कृषि सांख्यिकी अनुसंधान संस्थान (भा.कृ.सां.अ.सं.) प्रारंभ से ही मौजूदा ज्ञान में अंतराल को कम करने के लिए कृषि सांख्यिकी में अनुसंधान कार्य के लिए मुख्य रूप से उत्तरदायी है। संस्थान ने सांख्यिकी का प्रयोग विज्ञान के रूप में किया है तथा इसके साथ सूचना विज्ञान का प्रासंगिक प्रयोग किया है और कृषि अनुसंधान की गुणवत्ता को और अधिक उन्नत बनाने में महत्वपूर्ण योगदान दिया है। संस्थान द्वारा देश में प्रशिक्षित मानवशक्ति प्रदान करने के लिए कृषि सांख्यिकी और सूचना विज्ञान के क्षेत्र में शिक्षा/प्रशिक्षण भी प्रदान किया जाता है। कृषि अनुसंधान में गुणवत्ता सुधारने तथा उभरते हुए नए क्षेत्रों की चुनौतियों से निपटने के लिए अनुसंधान एवं शिक्षा का प्रयोग किया जाता है। संस्थान को देश में भारतीय कृषि के लिए प्रथम उच्च संगणना हब 'अशोका' (कृषि में ओमिक्स नॉलेज के लिए उन्नत उच्च संगणना हब) की स्थापना कर अपार प्रसन्नता हो रही है, जिसे माननीय कृषि मंत्री द्वारा दिनांक 15 जनवरी, 2014 को राष्ट्र को समर्पित किया गया। इसके अतिरिक्त, भा.कृ.अनु.प. में एक ई-आरपी प्रणाली भी विकसित एवं कार्यान्वित की गई, जिसमें वित्तीय प्रबंधन प्रणाली (एपफएमएस) सहित एक रॉबस्ट एवं फ्रलेक्सीबल प्रबंधन सूचना प्रणाली (एमआईएस) सम्मिलित है। एनएआरएस के लिए सांख्यिकीय संगणना सुदृढीकरण परियोजना के माध्यम से एनएआरएस में बेहतर सांख्यिकी संगणना वातावरण सृजित किया गया। संस्थान आईएसओ – 9001:2008 प्रमाणित संस्थान है। संस्थान का थीम गीत तैयार किया गया।

अपने लक्ष्य और अधिदेश को प्राप्त करने के लिए प्रतिवेदनाधीन वर्ष के दौरान संस्थान में अनेक अनुसंधान परियोजनाएँ चलाई गईं। संस्थान में वुफल 77 अनुसंधान परियोजनाओं के अंतर्गत अनुसंधान कार्य किया गया (इनमें से 39 संस्थान द्वारा वित्त पोषित, 24 अन्य बाह्य एजेन्सियों द्वारा वित्तपोषित, 13 अन्य संस्थानों के सहयोग से चलाई गयीं परियोजनाएँ तथा 01 राष्ट्रीय प्रोपेफसर योजना है)। इस वर्ष 25 नई परियोजनाएँ पूर्ण हुईं तथा 20 परियोजनाएँ आरंभ की गईं।

संस्थान की अनुसंधानिक गतिविधियों की वुफछ विशेषताएँ निम्न प्रकार हैं :

- प्रजनन कार्यक्रमों के लिए उपयुक्त कंट्रोल लाइन के साथ टेस्ट लाइनों की तुलना करने तथा टेस्ट बनाम कंट्रोल तुलनाओं के लिए द्वि-पथीय ब्लॉकिंग सेटअप के अंतर्गत विशिष्ट संयोजन क्षमता (एससीए) के प्रभावों से मुक्त सामान्य संयोजन क्षमता (जीसीए) प्रभावों में तुलना करने में सहायता देने के लिए मैटिंग-पर्यावरणीय पंक्ति-स्तंभ (एमईआरसी) अभिकल्पनाएँ विकसित की गईं। द्वि-पथीय ब्लॉकिंग सेटअप के अंतर्गत डायलल (या आंशिक डायलल) क्रॉस परीक्षण के लिए जीसीए तथा एससीए प्रभावों से संबंधित प्रारंभिक व्यतिरेकों के आकलन हेतु सूचना आव्यूह के लाम्बिक विभाजन के लिए पद्धति विकसित की गई। एमईआरसी अभिकल्पनाओं के जनरेशन, यादृच्छिकीकरण तथा विश्लेषण के लिए एसएस सॉफ्टवेयर में

मैक्रो विकसित किए गए।

- पंक्ति-स्तंभ अभिकल्पना संरचना में व्यवहारिक दृष्टिकोण से एक स्तंभ में दो से अधिक परीक्षणात्मक इकाइयों का समायोजन करना संभव नहीं हो सकता। लाम्बिक प्राचलीकरण हेतु न्यूनतम रन (चक्र) संख्या में मुख्य प्रभावों तथा दो सरल अन्योन्यक्रियाओं के लाम्बिक आकलन के लिए दो पंक्तियों के साथ पंक्ति-स्तंभ संरचना अभिकल्पना की एक सामान्य प्रणाली विकसित की गई। कुछ परीक्षणात्मक स्थितियों में, जैसे 2-क्लर माइक्रोएरे परीक्षणों के लिए, जहाँ शून्य स्थिति अथवा आधार-रेखा विद्यमान हो सकती है, आधार-रेखा प्राचलीकरण के आधार पर मिश्रित स्तर बहुउपादानी परीक्षणों के लिए दो पंक्तियों में दक्ष डब्ल्यू-इष्टतम बहुउपादानी पंक्ति स्तंभ अभिकल्पनाएँ प्राप्त करने की एक सामान्य कार्यविधि विकसित की गई।
- अनेक अभिकल्पित परीक्षणों में परीक्षणात्मक इकाइयाँ समय और स्थान की तुलना में सहज उपनति दर्शाती हैं। बहुउपादानी परीक्षणों के संबंध में सभी मुख्य प्रभावों तथा निम्न स्तर की अन्योन्यक्रिया प्रभावों का लाम्बिक रूप से आकलन करने हेतु बहु-स्तरीय बहुउपादानी परीक्षणों के निर्माण की एक सामान्य प्रणाली विकसित की गई, जो मुख्य प्रभावों और कुछ निम्न स्तर अन्योन्यक्रिया प्रभावों के लिए रैखिक उपनति-मुक्त हैं।
- दो की दूरी पर प्रतिवेशी इकाइयों से आकाशीय अप्रत्यक्ष प्रभावों के साथ ब्लॉक मॉडल, जिसमें उपनति घटक शामिल है, का अध्ययन किया गया तथा ब्लॉक अभिकल्पना के लिए उपनति मुक्त स्थितियाँ प्राप्त की गईं। दो की दूरी पर प्रतिवेशी इकाइयों से आकाशीय प्रत्यक्ष प्रभावों के लिए संतुलित उपनति मुक्त ब्लॉक अभिकल्पनाओं की श्रृंखला प्राप्त की गई।
- भारत में संचालित अभिकल्पित परीक्षणों के लिए सूचना प्रणाली विकसित की गई जो <http://www.isari.res.in> पर उपलब्ध है। यह ऑन-लाइन सूचना

प्रणाली देश में आयोजित अभिकल्पित परीक्षणों, मुख्यतः कृषि फील्ड परीक्षण (परिशुद्ध किस्मगत परीक्षणों को छोड़कर), कृषि प्रणाली अनुसंधान परियोजना निदेशालय के अंतर्गत आयोजित ऑन-पफार्म और ऑन-स्टेशन परीक्षणों तथा दीर्घकालिक उर्वरक परीक्षणों पर अखिल भारतीय समन्वित अनुसंधान परियोजना के अंतर्गत दीर्घकालिक उर्वरक परीक्षणों का एक केंद्रीयकृत संग्रह (स्टोरेज) है।

- भारतीय एनएआरएस सांख्यिकी संगणना पोर्टल (<http://stat.iasri.res.in/sscnarsportal>) में क्रॉस ओवर अभिकल्पनाओं के मॉड्यूल और ब्लॉक अभिकल्पनाओं से आनुवंशिक प्रसरण-सहप्रसरण के आकलन जोड़कर उसका सुदृढीकरण किया गया। इस पोर्टल में उपलब्ध 24 विश्लेषण मॉड्यूलों को मूल सांख्यिकी, परीक्षणों की अभिकल्पना, बहुचर विश्लेषण तथा सांख्यिकीय आनुवंशिकी के रूप में चार बड़े वर्गों में वर्गीकृत किया गया।
- अनुसंधानिक आँकड़ों के संग्रह के सृजन तथा अखिल भारतीय समन्वित ज्वार पफसल सुधार परियोजना के आँकड़ों के विश्लेषण के लिए एनएएआरएम, हैदराबाद तथा आईएएसआरआई, नई दिल्ली द्वारा ज्वार अनुसंधान निदेशालय, हैदराबाद के सहयोग में एक प्रोटोटाइप विकसित किया गया। इस सिस्टम का प्रचालन किया गया और उसे [www.aicsip.naarm.org.in](http://www.aicsip.naarm.org.in) पर उपलब्ध किया गया।
- अनुसंधान में कुछ ऐसी स्थितियाँ भी देखी गईं जहाँ वांछित चर और सहचरों (उदाहरण के लिए, उपज और उर्वरक के बीच संबंध) के बीच संबंध अध्ययन काल तक स्थिर नहीं रहता, जिसे आकाशीय अनुप्रगामीयता कहते हैं। एक उपयुक्त लघु क्षेत्र आकलन (एसएई) कार्यप्रणाली विकसित की गई, जो आँकड़ों में आकाशीय अनुप्रगामीयता की उपस्थिति दर्शाती है। विशेष रूप से, मॉडल आकाशीय अनुप्रगामीयता से संबंधित भौगोलिक भारित समाश्रयण अवधारणा के साथ क्षेत्र स्तर

मॉडल के अंतर्गत भौगोलिक रूप से भारित स्यूडो आनुभविक बेस्ट रैखिक (अनभिनत) प्राग्वक्ता विकसित किया गया। इसके अलावा, एमएसई आकलन विधि विकसित की गई। उत्तर प्रदेश में धान की पफसल की उपज के जिला स्तरीय आकलन प्राप्त करने हेतु विकसित की गई एसई विधि का उपयोग किया गया। इस विधि से विश्वसनीय माइक्रो स्तरीय सांख्यिकी प्राप्त करने में कापफी उत्साहवर्धक परिणाम प्राप्त हुए।

- विषम डाटा के लिए, एक एसई तकनीक विकसित की गई जिसे अरैखिक रूपांतरण, विशेष रूप से लघुगणकीय रूपांतरण के बाद, रैखिक रूप से प्रतिरूपित किया जा सकता है। विषम डाटा के लिए लघु क्षेत्र मानों का एक उत्कृष्ट आकलन {जिसे आनुभविक प्राग्वक्ता (ईबीपी) कहते हैं} विकसित किया गया।
- हमारे देश में विभिन्न पफसलों के उपज दर आकलन विकसित करने हेतु क्रॉप कटिंग परीक्षण पद्धति का उपयोग किया जाता है। कपास एक बहुमुखी उपयोगी पफसल है। इसके लिए अपेक्षा की गई कि सीमित पिकिंग संख्या के आधार पर इसके उत्पादन आकलन के लिए सरल पद्धति विकसित की जाए। कपास के क्रॉप कटिंग परीक्षणों के सहायक डाटा विश्लेषण से यह जानकारी प्राप्त हुई कि दूसरी/तीसरी पिकिंग वुफल पिकिंग संख्या के आधार पर उपज से अत्यधिक सहसंबंधित हैं। सहायक चर के रूप में दूसरी/तीसरी पिकिंग तथा अध्ययन के रूप में वुफल पिकिंग आँकड़ों का उपयोग करते हुए दोहरा प्रतिचयन समाश्रयण आकलन के प्रयोग के द्वारा एक सरल पद्धति विकसित की गई।
- प्रतिदर्श सर्वेक्षण संबंधी समस्याओं में सहायक सूचना एक महत्वपूर्ण भूमिका निभाती है। प्रतिदर्श चयन में सर्वेक्षणों को उपयुक्त रूप से डिजाइन करने तथा यथार्थ आकलक विकसित करने हेतु उपलब्ध सहायक का उपयोग किया जा सकता है। आकलन स्तर पर सहायक सूचना

के समावेशन के लिए अंशशोधन (केलिब्रेशन) अप्रोच एक सुव्यवस्थित अप्रोच उपलब्ध करती है। केलिब्रेशन अप्रोच का प्रयोग करते हुए दो-स्तरीय प्रतिचयन अभिकल्पना के लिए परिमित समष्टि मीन/टोटल के यथार्थ आकलक भी विकसित किए गए। अनुकार अध्ययनों की सहायता से परिशुद्धता में प्राप्त उपलब्धियों को प्रदर्शित किया जाता है।

- बिहार राज्य के वैशाली जिले में पॉप्लर ट्री आधारित कृषिवानिकी की शुरुआत पर प्रभाव का अध्ययन किया गया। यह निष्कर्ष निकाला गया कि वैशाली जिले के किसानों की सामाजिक-आर्थिक स्थिति पर कृषिवानिकी का कापफी प्रभाव था और इसलिए यह सिपफारिश की जाती है कि बिहार के समस्त जिलों में विशेष रूप से तथा देश के सभी राज्यों में सामान्य रूप से पॉप्लर ट्री कृषिवानिकी मॉडल कार्यान्वित किया जाना चाहिए।
- अर्थशास्त्र एवं सांख्यिकी निदेशालय, कृषि मंत्रालय में पफसल आकलन सर्वेक्षण का गुणवत्ता संबंधी पहलू एक चिंता का विषय रहा है। कृषि सांख्यिकी में सुधार लाने के लिए गठित प्रोपेफसर वैद्यानाथन समिति ने विश्वसनीय पफसल क्षेत्र और पफसल उपज आकलन के लिए प्रतिदर्श आकारों की सिपफारिश की। आईसीएस (फसल सांख्यिकी का सुधार) का प्रयोग करते हुए अध्ययन में यह पाया गया कि संस्तुत प्रतिदर्श आकार धान तथा गेहूँ की पफसलों की औसत उपजों के आकलन करने में उपयुक्त थे। अन्य पफसलों के लिए, औसत उपज के आकलन हेतु संस्तुत प्रतिदर्श आकार उपयुक्त नहीं पाए गए। इसके अलावा, आईसीएस आधारित प्रतिदर्श आकार पफसल क्षेत्र आकलन के लिए उपयुक्त नहीं पाए गए।
- प्रतिदर्श सर्वेक्षणों में गैर-अनुक्रिया एक आम समस्या है, जो अभिनत (बायस्ड) आकलन देता है। हेनसेन एवं हरविट्ज़ (1946) ने मुख्य रूप से गैर-अनुक्रिया के समायोजन करने के लिए

एक तकनीक विकसित की। केलिब्रेशन अप्रोच का प्रयोग करते हुए हेनसेन एवं हुरविट्ज़ (1946) तकनीकी आधारित आकलक उस स्थिति के लिए विकसित किया गया जहाँ सहायक चर संबंधी सूचना को समस्त समष्टि इकाइयों के लिए जाना जाता है। प्रसरण तथा प्रसरण आकलकों के लिए व्यंजक (एक्सप्रेशन) विकसित किए गए। आनुभविक परिणामों में यह पाया गया कि प्रस्तावित केलिब्रेशन अप्रोच आधारित आकलक हेनसेन एवं हुरविट्ज़ (1946) के आकलक से कापफी ज्यादा बेहतर थे।

- कृषि अनुसंधान डाटा पुस्तक (एआरडीबी) 2013 तैयार की गई जो अपनी श्रृंखला में सौलहवीं है। इस डाटा पुस्तक में जून 2013 के अंत तक देश में विभिन्न स्रोतों से उपलब्ध मुख्य घटकों से संबंधित सूचना/कृषि अनुसंधान से संबंधित संकेतकों तथा संबद्ध पहलुओं व आयामों को एक साथ संग्रहीत किया गया है। राज्य-वार डाटा प्राप्त करने के लिए विषयपूरक मानचित्र बनाए गए हैं, जिसके लिए भौगोलिक सूचना प्रणाली (जीआईएस) का प्रयोग किया गया।
- अधिकतर जिंसों की अनिश्चित उपजों तथा आपेक्षिक रूप से माँग की कम मूल्य सापेक्षता के कारण कृषि उत्पादन जोखिमों और अनिश्चितताओं पर आधारित है। जिंसों की कीमतों में उतार-चढ़ाव कासमग्र मैक्रो-आर्थिक निष्पादन पर बड़ा प्रभाव पड़ता है, जिसके कारण प्याज की कीमत में कापफी तेजी आ जाती है। इसलिए जिंस-मूल्य पूर्वानुमान मैक्रो-आर्थिक नीति के नियोजन और सूत्रीकरण के लिए महत्वपूर्ण इनपुट हैं। पफसल आवक से पहले, विशेष रूप से जब पफसल अपर्याप्त हो, तथा प्याज के मूल्यों का पूर्वानुमान करने के लिए काल श्रृंखला डाटा का उपयोग करते हुए कृषि जिंसों के मूल्यों पर एक अध्ययन किया गया, जिसमें यह पाया गया कि दैनिक मूल्यों के आकलन के लिए सामान्यीकृत स्वस्माश्रयी शतप्रतिबंधित विषम विचालिता (गार्च) मॉडल स्वसमाश्रयी समेकित मूविंग एवरेज (ऐरीमैक्स)

मॉडल से बेहतर मॉडल है।

- एरिमा पद्धति अनेक अरैखिक काल-श्रृंखला डाटा, यथा ऑयल सारडाइन, मैकरेल तथा बम्बई डक के अवतरणों को उचित रूप से प्रतिरूपित करने में सक्षम नहीं हो सकता है। ऐसी स्थिति में, एक्सपोनेंशियल स्मूथ ट्रांजिशन आटोरिग्रेसिव (ईएसटीएआर) मॉडल का प्रयोग किया जा सकता है। पार्टिकल स्वार्म ऑप्टिमाइजेशन (पीएसओ) तकनीक का प्रयोग करते हुए ईस्टार के प्राचलों के आकलन के लिए कार्यविधि विकसित की गई। पिफ्ट किए गए मॉडल के निष्पादन की तुलना की गई और यह निष्कर्ष निकाला गया कि विचाराधीन डाटा सेटों के लिए ईस्टार मॉडलों का निष्पादन एरिमा मॉडल की तुलना में बेहतर था।
- समेकित विकास दर का आकलन करने के लिए एक प्रयोगकर्ता हितैषी सॉफ्टवेयर (बेबईसीजीआर) विकसित किया गया जो <http://iasri.res.in/cgr> पर उपलब्ध है।
- अमृतसर और भंटीडा बाजार में अरहर के थोक मूल्य पर तथा अखिल भारतीय स्तर पर अरहर के अधिकतम, न्यूनतम तथा मॉडल मूल्य पर लॉग मैमोरी की उपस्थिति टेस्ट की गई। सभी पाँचों श्रृंखलाओं में लॉग मैमोरी की उपस्थिति देखी गई। तदनुसार, उपरोक्त श्रृंखलाओं के थोक मूल्य की मॉडलिंग तथा पूर्वानुमान के लिए ऑटोरिग्रेसिव प्रेफक्शनली इंटीग्रेटेड मूविंग एवरेज (एआरएफआईएमए) मॉडल की पिफ्टिंग की गई। जीपीएच आकलक (जिवेक एवं पार्टर्स-ह्युडेक, 1983), गोसियन अर्ड प्राचलिक आकलक (रॉबिनसन, 1995) तथा वेवलेट पद्धति (जेनसेन, 1989) से लॉग मैमोरी का आकलन किया गया। न्यूनतम वर्ग माध्य त्रुटियों के आधार पर अनुकार विश्लेषण अध्ययन में वेवलेट विधि अन्य विधियों की तुलना में बेहतर पाई गई।
- अरैखिक सपोर्ट वेक्टर समाश्रयण (एनएलएसवीआर) तकनीक का प्रयोग करते हुए पफसल उपज पूर्वानुमान मॉडल विकसित किया गया तथा

उदाहरणस्वरूप मक्के की पफसल की उपज के पूर्वानुमान के लिए उसकी जाँच की गई। जाँच में अनुक्रिया चर के रूप में मक्का पफसल उपज डाटा तथा प्राग्वक्ता चरों के रूप में वुफल मानव श्रम, पफार्म ऊर्जा उर्वरक उपभोग तथा कीटनाशक उपभोग पर विचार किया गया। वर्ग माध्य मूल त्रुटि (आरएमएसई), माध्य-निरपेक्ष त्रुटि (एमएई) तथा माध्य निरपेक्ष पूर्वानुमान त्रुटि (एमएपीई) के आधार पर एक पिफटेड मॉडल के निष्पादन का निर्धारण किया गया। विचाराधीन डाटा के लिए कृत्रिम न्युरल नेटवर्क तकनीक की तुलना में एनएलएसवीआर तकनीक की श्रेष्ठता का प्रदर्शन किया गया। निष्कर्ष में यह पाया गया कि मॉडलिंग तथा पूर्वानुमान प्रयोजनों के लिए एनएलएसवीआर पद्धति कापफी सक्षम है।

- भा.कृ.अनु.प. में राष्ट्रीय कृषि जैवसूचना विज्ञान ग्रिड (एनएबीजी) की स्थापना की गई। एनएबीजी के अंतर्गत अनेक डाटाबेस और टूल्स विकसित किए गए तथा संगणनात्मक जीवविज्ञान एवं कृषि जैव सूचना विज्ञान के क्षेत्र में अनुसंधानकर्ताओं का सुग्राहीकरण करने और उन्हें प्रशिक्षित करने हेतु विभिन्न विषयों पर अनेक प्रशिक्षण कार्यक्रम/कार्यशालाएँ/बैठवेंफ आयोजित की गई।
- टमाटर माइक्रोसेटेलाइट डाटाबेस (टोमसेटडीबी), टमाटर के लिए पहला पूर्ण जिनोम आधारित माइक्रोसेटेलाइट डीएनए मार्वफर डाटाबेस, विकसित किया गया, जिसमें 146602 एसटीआर मार्वफर, माइन्ड इन सिलिको, प्रयोग करने के लिए माइक्रोसेटेलाइट (एमआईएसए) टूल शामिल हैं। वैट लैब की विशिष्ट आवश्यकताओं की पूर्ति हेतु इसमें स्वचालित प्राइमर डिजाइनिंग टूल को भी सम्मिलित किया गया है। डाटाबेस <http://webapp.cabgrid.res.in/tomsatddb> पर उपलब्ध किया गया है।
- एक वेब आधारित आपेक्षिक डाटाबेस विकसित किया गया, जिसमें बकरी के पूर्ण जिनोम अनुक्रमण में उपस्थित 865210 माइक्रोसेटेलाइट मार्वफर

सम्मिलित हैं। बहु प्राचलों का प्रयोग करते हुए गोसेटेडीबी माइक्रोसेटेलाइट सर्च में सुविधा देता है। यह <http://weapp.cabgrid.res.in/goat/> पर उपलब्ध है।

- माइक्रोसेटेलाइट डीएनए मार्वफरों का प्रयोग करते हुए प्रथम बकरी नस्ल पहचान सर्वर विकसित किया गया, जो <http://cabin.iasri.res.in/gomi/> पर उपलब्ध है।
- रेपफरेंस डाटा कायम रखने तथा नस्ल पहचान करने के लिए वेब आधारित पशु नस्ल पहचान सर्वर विकसित किया गया। पूर्वानुमान मॉडल के लिए उपयोग किए गए रेपफरेंस डाटा को 08 पशु नस्ल ओर 18 माइक्रोसेटेलाइट डीएनए मार्वफरों से प्राप्त किया गया, जिसमें 18000 एलील डाटा शामिल हैं। यह सर्वर <http://cabin.iasri.res.in/biscattle/> पर उपलब्ध है।
- राष्ट्रीय कृषि जैवसंगणना पोर्टल विकसित किया गया, जो उच्च निष्पादन संगणना (उचपीसी) के लिए प्रत्यक्ष पहुंच उपलब्ध कराता है। जैवसूचना विज्ञान संबंधी कार्यों के निष्पादन के लिए पोर्टल एक वातावरण उपलब्ध कराता है। पोर्टल प्रयोक्ता को अनुप्रयोग-विशिष्ट कार्यों (जॉब) के प्रबंधन और प्रस्तुत करने में सहायता करता है। पोर्टल के माध्यम से प्रस्तुत किए गए कार्यों की अनुसूची बनाई जाती है तथा रिसोर्स प्रबंधक के द्वारा संसाधनों का आवंटन किया जाता है। पोर्टल देश में जैवप्रौद्योगिकीय अनुसंधान के संगणनात्मक आवश्यकताओं में सहायता करता है। इससे जिनोमिक सूचना और ज्ञान के बीच, सांख्यिकी एवं संगणनात्मक विज्ञानों का उपयोग करने में तथा क्रॉस प्रजातियों के जिनोमिक संसाधनों से सूचना एवं ज्ञान को निष्कर्षित करने के लिए हाई-एंड संगणनात्मक ऊर्जा के साथ बड़े जिनोमिक डाटाबेसों, डाटा वेयरहाउस, सॉफ्टवेयर टूलों, एल्गोरिथ्म, जिनोमिक ब्राउसर्स स्थापित करने में अंतराल भरा जाएगा।
- स्टैंडर्ड डाटाबेस प्रबंधन सिद्धांतों का अनुसरण



- करते हुए बैकएंड डाटाबेस के साथ एक सुरक्षित जिनोम प्रस्तुतीकरण पोर्टल विकसित किया गया (<http://nabg.iasri.res.in>)। देश में घरेलू जिनोम डाटाबेस बनाने तथा प्लेटफार्म विश्लेषण करने के लिए यह पहल की गई। इस डाटाबेस से उच्च गतिवान सूचना के प्रसंस्करण तथा ज्ञान निष्कर्षण के लिए उन्नत हार्डवेयर संसाधन और समानांतर संगणना सुविधाएँ स्थापित की गई।
- कोशिका के भीतर ट्रांसक्रिप्शन पेफक्टर (टीएफ) तथा माइक्रोआरएनए मुख्य जीन रेग्युलेटर हैं। इन दो मुख्य रेग्युलेटरों की रेग्युलेटरी कार्यप्रणाली अजैव तथा जैव दबावों अंतर्दृष्टि का बोध कराती है। इस कार्यप्रणाली के बारे में यह स्पष्ट नहीं है कि माइक्रोआरएनए रेग्युलेटेड टीएफ लक्षित जीन किस प्रकार दबाव से प्रभावित कोशिका के भीतर समन्वित रेग्युलेशन को प्रभावित करते हैं। रेग्युलेटरी जीन अभिव्यंजकता चिन्हकों, जैसे माइक्रोआरएनए तथा टीएफ के लिए ट्रांसक्रिप्शनल मॉड्यूल की खोज करने हेतु एक इन सिलिको विश्लेषण पाइपलाइन का प्रस्ताव किया गया।
  - प्रोटीन-प्रोटीन अन्वोन्यक्रियाएँ प्रोटीन के बीच ऐसी अन्वोन्यक्रियाएँ हैं जो वुफछ रासायनिक अनुक्रियाओं तथा इलेक्ट्रोस्टेटिक दबावों के कारण उत्पन्न होती हैं। भौतिक-रासायनिक विशिष्टताओं, जैसे जलभीत, x-कोऑर्डिनेट (उपसहसंयोजकता), y-कोऑर्डिनेट, z-कोऑर्डिनेट, सरपेफस टेंशन चार्ज, अल्पफा हेलिक्स, -हेलिक्स, टर्न, वेन डर वाल्स, आण्विक वजन, घुलनशीलता आदि के साथ संरचनात्मक सूचना का प्रयोग करते हुए प्रोटीन-प्रोटीन अन्वोन्यक्रियाओं का पूर्वानुमान किया जाता है। सपोर्ट वेक्टर मशीन (एसवीएम) का प्रयोग करते हुए प्रोटीन-प्रोटीन अन्वोन्यक्रिया के पूर्वानुमान के लिए एल्गोरिथ्म विकसित की गयी।
  - पशु जिनोम पर भैंस जीनों के ऑर्थोलोगस एनोटेड डाटा प्राप्त करने के लिए एक वेब इंटरपेफस विकसित किया गया। पशु जिनोम पर भैंस के जीनों की मैपिंग के लिए हल्के वजन तथा ब्राउज़र से प्रयोग किए जाने वाला एक जिनोम व्युअर टूल विकसित किया गया।
  - गैर-जीवाण्विक पेप्टाइडों के वर्गीकरण के लिए एसवीएम का प्रयोग करते हुए गैर-जीवाण्विक पेप्टाइड पूर्वानुमान मॉडल विकसित किए गए और उनके निष्पादन के लिए मूल्यांकन तथा तुलना की गई। चूँकि एसवीएम मॉडलों ने एएनएन से बेहतर निष्पादन प्रदर्शित किया, इसलिए उन्हें एन-टर्मिनल, सी-टर्मिनल तथा पूर्ण अनुक्रमण के लिए सीजीआई -पर्ल, एचटीएमएल, पीएचपी का प्रयोग करते हुए वेब सर्वर में क्रियान्वित किया गया, जो <http://cabin.iasri.res.in/amp> पर उपलब्ध है।
  - गुण संयोजित जीनों के पूर्वानुमान की समस्या को पफीचर चयन समस्या के रूप में उल्लेखित किया जा सकता है, जहाँ माइक्रो एरै डाटा के जीनों को प्रफीचर्स के रूप में किया जा सकता है तथा चयनित प्रमुख जीन एक विशेषक व गुण को इंगित करते हैं। किसी विशेष गुण से संबंधित प्रमुख जीनों के पूर्वानुमान हेतु गैर-रैखिक पेनालाइज्ड स्पोर्ट वेक्टर मशीन (एसवीएम) का उपयोग किया जाता है और यह क्रॉस वैधीकरण तकनीक पर आधारित पूर्वानुमान यथार्थता की प्रक्रिया भी उपलब्ध कराता है।
  - जैव सूचना विज्ञान संबंधी वर्कफलो प्रणालीवर्कफलो का एक विशिष्ट स्वरूप है, जिसे संगणनात्मक श्रृंखलाओं या डाटा दोहन प्रक्रियाओं या विशिष्ट क्षेत्र में वर्कफलो कम्पोज और कार्यान्वित करने के लिए विशेष रूप से डिजाइन किया गया है। समानांतर संगणन टूलों और लाइब्रेरियों का प्रयोग करते हुए वर्कफलो के लिए अपेक्षित स्थानीय जिनोम डाटाबेस, अभिव्यंजित अनुक्रमण टैग (ईएसटी) से जीन की पहचान करने हेतु एक समानांतर पाइपलाइन, समानांतर आर्किटेक्चर पर जातिवृतीय विश्लेषण करने के लिए एक वर्कफलो

प्रणाली तथा एसएसआर मार्फरों और प्राइमर डिजाइनिंग की पहचान के लिए एक पाइपलाइन विकसित व निर्मित की गई। ऑनलाइन पादप आनुवंशिकी विश्लेषण के लिए भी सॉफ्टवेयर विकसित किया गया।

- भा.कृ.अनु.प. में वित्तीय प्रबंधन प्रणाली (एमएमएस) सहित एक रॉबस्ट एवं लोचनीय प्रबंधन सूचना प्रणाली (एमआईएस) विकसित की गई, जिसमें वित्तीय प्रबंधन, परियोजना प्रबंधन, सामग्री प्रबंधन, मानव संसाधन प्रबंधन तथा भा.कृ. अनु.प. में पे-रोल सम्मिलित हैं। इसके लिए एक भा.कृ.अनु.प.-ईआरपी सॉल्युशन विकसित किया गया। भाकृसांअसं तथा सहभागी संगठनों में भा.कृ.अनु.प.- ईआरपी सॉल्युशन के विभिन्न मॉड्यलों पर संस्थान के कार्मिकों को प्रशिक्षण प्रदान करने हेतु भाकृसांअसं में प्रयोक्ता स्वीकार्यता परीक्षण (यूएटी) कार्यशालाएँ आयोजित की गईं।
- कृषि से संबंधित सूचना के लिए एक ऑनलाइन नॉलेज संग्रह तथा नॉलेज प्रबंधन प्लेटफॉर्म, भाकृसांअसं एग्रोपीडिया विकसित किया गया। इसमें विभिन्न पफसलों पर भिन्न-भिन्न रूपों में व्यापक रूप से कृषि सूचना उपलब्ध की गई है। वर्तमान में एग्रोपीडिया पोर्टल को <http://www.agropedia.in> पर उपलब्ध किया गया है। इस संबंध में, तकनीकों के पूर्वानुमान के लिए एक नॉलेज मॉडल तैयार किया गया है। यह पूर्वानुमान विधियों, अनुप्रयोगों, अवस्थितियों तथा कुछ मॉडलों से संबंधित सामान्य सूचना उपलब्ध कराता है।
- संस्थान ने दिनांक 27-28 नवंबर 2013 के दौरान एनडीआरआई, करनाल में राष्ट्रीय डेयरी अनुसंधान संस्थान (एनडीआरआई) के साथ संयुक्त रूप से कृषि अनुसंधान सांख्यिकीविदों के 17वें राष्ट्रीय सम्मेलन का आयोजन किया गया, जहाँ कृषि सांख्यिकी एवं सूचना विज्ञान पर पाँच तकनीकी सत्र आयोजित किए गए। संस्थान ने अनुसंधान सलाहकार समिति (आरएसी), संस्थान प्रबंधन

समिति (आईएमसी) तथा संस्थान अनुसंधान समिति (आईआरसी) का नियमित रूप से आयोजन किया। रिजल्ट फ्रैमवर्क दस्तावेज, 2014-15 प्रस्तुत किया गया, जबकि वर्ष 2013-14 के लिए संस्थान ने 100 का स्कोर प्राप्त किया।

- संस्थान के वैज्ञानिकों ने राष्ट्रीय और अंतरराष्ट्रीय रेपफरीड जर्नलों में 105 शोध-पत्र, 29 लोकप्रिय लेख, 01 पुस्तक, 20 पुस्तक अध्याय, सम्मेलन कार्यवाहियों में 09 शोध पत्र और 55 परियोजना रिपोर्ट/तकनीकी बुलेटिन/संदर्भ मैनुअल/ब्रोशर प्रकाशित किए। इसके अतिरिक्त संस्थान की वेबसाइट पर उपलब्ध 23 मैक्रो/ई-रिसोर्सिस भी विकसित किए गए।

इस वर्ष 21 प्रशिक्षण कार्यक्रमों का आयोजन किया गया; जिनमें 374 प्रतिभागियों ने प्रशिक्षण प्राप्त किया।

- उन्नत संकाय प्रशिक्षण केन्द्र के अंतर्गत कृषि में सांख्यिकीय मॉडलिंग तकनीकों में नूतन उन्नतियों, पशुओं के परीक्षणों के लिए सांख्यिकी विधियों में उन्नतियाँ, कृषि में प्रौद्योगिकियों के विकास के लिए परीक्षणात्मक अभिकल्पनाओं में उन्नतियाँ तथा 'ओमिक्स' डाटा के लिए जैवसूचना विज्ञान में संगणनात्मक एवं सांख्यिकीय उन्नतियों पर पाँच (05) इक्कीस (21) दिवसीय प्रशिक्षण कार्यक्रमों का आयोजन किया गया
- फसलों में पूर्वानुमान मॉडलिंग पर तथा कृषि प्रबंधन के लिए बेव अनुप्रयोग के विकास पर दो ग्रीष्मकालीन/शीतकालीन स्कूलों का आयोजन किया गया।
- राष्ट्रीय कृषि नवोन्मेषी परियोजनाओं के अंतर्गत छः प्रशिक्षण कार्यक्रम आयोजित किए गए: एसएसएस का प्रयोग करते हुए आँकड़ा विश्लेषण, सीएलसी बायो सॉफ्टवेयर हेतु प्रयोक्ताओं के लिए उच्चतर प्रशिक्षण, एनएसजी आँकड़ा विश्लेषण के लिए संगणनात्मक पहलू: एनएबीजी के अंतर्गत ओम अनुसंधान सुविधा केंद्र में लैब टू फील्ड सॉजर्न तथा अजैव दबाव सहिष्णुता के लिए जीन एवं एलील माइनिंग (विकल्पी खनन) के जैव पूर्वक्षण

के अंतर्गत उन्नत विश्लेषण तकनीक।

- आईएसएस अभ्यर्थियों के लिए आँकड़ों का विश्लेषण एवं निर्वचन पर सांख्यिकी एवं कार्यक्रम कार्यान्वयन मंत्रालय द्वारा प्रायोजित तीन संसाधन सृजन प्रशिक्षण कार्यक्रम तथा समेकित सर्वेक्षण कार्यप्रणाली (पुनश्चर्या पाठ्यक्रम) पर पशुपालन, डेयरी एवं मात्स्यिकी विभाग, भारत सरकार द्वारा प्रायोजित दो अन्य कार्यक्रम आयोजित किए गए।
- भा.कृ.अनु.प. के तकनीकी कार्मिकों के लिए डाटा बेस प्रबंधन, एनएसएस के लिए आईएआईआर के कार्यों एवं क्रियाकलापों, अफगानिस्तान इस्लामिक गणराज्य के वैज्ञानिकों के लिए विषयपूरक प्रशिक्षण के सांख्यिकी मॉड्यूल तथा कृषि विज्ञान में स्नातकोत्तर शिक्षण पर अफगानिस्तान के नागरिकों के लिए अंतरराष्ट्रीय एम.एससी. कार्यक्रम के प्रतिभागियों के लिए कृषि सांख्यिकी अनुसंधान हेतु सांख्यिकी विधियों में मॉड्यूल पाठ्यक्रम पर पाँच अन्य प्रशिक्षण आयोजित किए गए। भारतीय सामाजिक विज्ञान अनुसंधान संस्थान द्वारा प्रायोजित 'पूरे यूरोप और भारत में विज्ञान की कल्चरल अथोरिटी की मैपिंग (एमएससीएस – ईयू एवं भारत)' के अंतर्गत संगणक समर्थित टेक्स्ट विश्लेषण पर एक प्रशिक्षण कार्यशाला आयोजित की गई।

संस्थान के वैज्ञानिकों ने विभिन्न एजेंसियों से पुरस्कार प्राप्त कर संस्थान का गौरव बढ़ाया है। डॉ. हुवुफम चन्द्र ने दिनांक 16 जुलाई 2013 को सामाजिक विज्ञान के क्षेत्र में उत्कृष्ट योगदान देने के लिए भा.कृ.अनु.प. को लाल बहादुर शास्त्री उत्कृष्ट युवा वैज्ञानिक पुरस्कार-2012 प्राप्त किया डॉ. दिनेश बुफमार ने जैवसूचना विज्ञान के क्षेत्र में उत्कृष्ट योगदान देने के लिए अनुप्रयुक्त जैवप्रौद्योगिकी संस्था से पैफलो पुरस्कार-2012 प्राप्त किया, डॉ. विशाल गुरफंग ने भारतीय कृषि सांख्यिकी संस्था (आईएसएस) से डॉ. जी आर सेठ स्मृति युवा वैज्ञानिक 2013 प्राप्त किया, डॉ. अर्पण भौमिक ने पीएच.डी. शोध के लिए आईएआरआई के पीजी स्क्वॉल के 52वें दीक्षांत समारोह के दौरान

आईएआरआई मेधावी पदक प्राप्त किया, डॉ. अनिल राय ने भारतीय कृषि अभियांत्रिकी संस्था से पभारत में मुख्य पफसलों तथा पशुधन उत्पाद के हार्वेस्ट तथा पोस्ट हार्वेस्ट हानियों के निर्धारण पर महत्वपूर्ण योगदान देने के लिए टीम पुरस्कार प्राप्त किया। डॉ. राजेन्द्र प्रसाद को आईएसएस पैफलोशिप की गई।

डॉ. यू सी सूद ने प्रतिचयन एवं अनुसंधानिक प्रणालियों पर परामर्श प्रदान करने हेतु कृषि सांख्यिकी (आईसीएस-VI) और भूटान पर छठे अंतरराष्ट्रीय सम्मेलन में पआकाशीय लघु क्षेत्र मॉडल के अंतर्गत जिला स्तरीय फसल उपज आकलन, शीर्षक पर पबंगाल, रियो-डे-जेनेरियो, ब्राजील में समेकित कृषि उत्पादन सांख्यिकी के दोहन एवं प्रसारण परियोजना से संबंधित कार्यशाला पर तीसरे मिशन के संबंध में बांग्लादेश की यात्रा की।

डॉ. हुकुम चन्द्र को बैंकॉक, थाइलैंड में लघु क्षेत्र आकलन पर पहले एशियन अंतरराष्ट्रीय सांख्यिकी संस्था सेटलाइट सम्मेलन में सहभागिता करने हेतु वैज्ञानिक समिति के सदस्य के रूप में भेजा गया। एफएओ परामर्शदाता के रूप में उन्होंने इथिओपिया की भी यात्रा की।

डॉ. प्राची मिश्रा साहू ने कृषि मंत्रालय एवं मात्स्यिकी, ओमान सलतनत को परामर्श देने हेतु ओमान का दौरा किया। डॉ. संजीव पंवार ने पफसल सुधार आंदोलन में जैवमितीय एवं जैवसूचना विज्ञान एवं जैव सूचना विज्ञान टूलों के अनुप्रयोग पर प्रशिक्षण में जीनोटाइप x एन्वायरन्मेंट विश्लेषण पर सत्र लेने हेतु एक रिसोर्स व्यक्ति के रूप में केन्या, नैरोबी का दौरा किया।

डॉ. एम ए इकबाल और डॉ. सारिका को आयोवा राज्य विश्वविद्यालय, एम्स, आयोवा, यूएसए में तीन माह के लिए फजैवसूचना विज्ञान, पर एनएआईपी वित्तपोषित अंतरराष्ट्रीय प्रशिक्षण में भाग लेने हेतु भेजा गया। मो. समीर पफारफकी को कृषि विज्ञान, आयोवा राज्य विश्वविद्यालय, आयोवा, यूएसए में जैवसूचना विज्ञान एवं तुलनात्मक जिनोमिक के क्षेत्र में अंतरराष्ट्रीय प्रशिक्षण में भाग लेने हेतु भेजा गया। श्री के के चतुर्वेदी को कोरनेल विश्वविद्यालय, इथाका, यूएसए में एनएआईपी

के तहत “जैवसूचना विज्ञान” पर प्रशिक्षण में भाग लेने हेतु भेजा गया। डॉ. डी सी मिश्रा को जैवसूचना विज्ञान एवं जैव सांख्यिकी विभाग, लुईसवाईल, केन्टकी, यूएसए में तीन माह के लिए संगणनात्मक जीवविज्ञान के क्षेत्र में प्रशिक्षण में भाग लेने हेतु प्रतिनियुक्त किया गया। डॉ. सुशील वुफमार सरकार जारागोजा, स्पेन में जारागोजा मेडिटेरियल कृषि विज्ञान संस्थान (आईएएमजेड) में जनरेशन चैलेंज कार्यक्रम – समेकित प्रजनन प्लेटफार्म (जीसीजी-आईबीपी) के अंतर्गत समेकित प्रजनन बहुवर्षीय पाठ्यक्रम (आईबी-एमवाईसी) पर प्रशिक्षण लेने हेतु प्रतिनियुक्त किया गया।

शिक्षा और प्रशिक्षण से संबंधित संस्थान की समस्त गतिविधियाँ स्नातकोत्तर शिक्षण कार्यक्रमों की योजना, संगठन तथा समन्वय, पी.जी. स्क्वॉल, आईएआरआई के सहयोग से संचालित की गयीं। प्रतिवेदनाधीन वर्ष के दौरान 16 छात्रों ने अपने डिग्री कार्यक्रम पूरे किए

{02 पीएच. डी. (कृषि सांख्यिकी), 08 एम. एससी. (कृषि सांख्यिकी), 05 एम. एससी. (संगणक अनुप्रयोग) तथा 01 एम.एससी. छात्र (जैव सूचना विज्ञान)}। 23 नए छात्रों को {05 पीएच. डी. (कृषि सांख्यिकी), 08 एम एससी (कृषि सांख्यिकी), 02 पीएच. डी. (संगणक अनुप्रयोग), 04 एम. एससी. (संगणक अनुप्रयोग) तथा 04 एम. एससी. (जैव सूचना विज्ञान)} में प्रवेश दिया गया।

कृषि सांख्यिकी और संगणना में एक वरिष्ठ प्रमाण-पत्र पाठ्यक्रम आयोजित किया गया। इस प्रमाण-पत्र पाठ्यक्रम में 02 पदाधिकारियों ने सहभागिता की।

(संस्थान की वार्षिक रिपोर्ट 2013-14 के  
'विशिष्ट सारांश' से उद्धृत)

□

देश की भाषाओं के बीच हिंदी सम्पर्क भाषा के रूप में कार्य करती है। इसे इसलिए यह मान्यता नहीं मिली है कि यह सबसे अधिक विकसित भाषा है, वरन् इसलिए की इसे गैर हिंदी भाषी लोगों ने अंगीकार किया है।

—इंदिरा गाँधी

## डॉ. दरोगा सिंह स्मृति व्याख्यान

### फ़सलों की उपज के कटाई—पूर्व अनुमान हेतु मौसम आधारित मॉडल

रंजना अग्रवाल\*

भविष्य जानने की जिज्ञासा मनुष्य में आरम्भ से ही रही है। परन्तु भविष्यवाणी अथवा पूर्वानुमान (forecast) केवल जिज्ञासा की बात ही नहीं है वरन् यदि समय से सटीक पूर्वानुमान प्राप्त हो सकें तो आगामी कार्यकलापों के लिए अत्यंत उपयोगी होंगे। फ़सल की उपज के संदर्भ में बात करें तो उपज के विश्वसनीय पूर्वानुमान भंडारण, वितरण, मूल्य निर्धारण, आयात—निर्यात आदि संबंधी विभिन्न नीतियों के निर्धारण में महत्वपूर्ण भूमिका निभाते हैं। फ़सल की उपज के पूर्वानुमान दीर्घकालिक हो सकते हैं, उदाहरण के लिए यदि हम जानना चाहें कि पाँच वर्ष बाद किस फ़सल की उपज कितनी होगी या अगले वर्ष गेहूँ की उपज कितनी होगी तो ये दीर्घकालिक पूर्वानुमान की श्रेणी में आयेंगे। इसके लिए प्रायः दीर्घकालिक समय श्रृंखला (long term time series) का विश्लेषण करके पिछले वर्षों की प्रवृत्ति (trend) के अनुसार पूर्वानुमान लगाया जाता है कि आगामी वर्षों में स्थिति क्या होगी अर्थात् इसके लिए समय श्रृंखला विश्लेषण (time series analysis) का उपयोग होता है।

कटाई—पूर्व अनुमान की स्थिति इस पारंपरिक पूर्वानुमान से थोड़ी भिन्न है। आगे बढ़ने से पूर्व यह जानना अपेक्षित होगा कि किसी स्थान पर विभिन्न वर्षों में उपज किन—किन कारकों (factors) से प्रभावित होती है। उपज प्रयुक्त बीज, खाद, पानी, कीट/रोगनाशकों तथा मौसम पर निर्भर होती है। कटाई—पूर्व

अनुमान वर्तमान वर्ष में खड़ी फ़सल की कटाई से एक—दो माह पूर्व उपज का अनुमान लगाने से सम्बन्धित है। यहाँ हमें प्रयुक्त बीज, खाद, पानी, कीटनाशकों आदि की जानकारी होती है, फ़सल की दशा भी ज्ञात होती है और पूर्वानुमान के समय तक मौसम भी ज्ञात होता है अतः इसमें पूरी तरह पिछली प्रवृत्ति (trend) पर निर्भर रहने की आवश्यकता नहीं होती वरन् इन विभिन्न वर्षों में विभिन्न कारकों तथा उपज के आँकड़ों के बीच मॉडल बनाकर उसके आधार पर पूर्वानुमान किया जा सकता है। प्रयुक्त बीज, खाद, पानी, कीटनाशकों आदि कृषि निवेशों (agricultural inputs) को संयुक्त रूप से प्रौद्योगिकीय कारक (technological factor) कहा जाता है। यह माना जाता है तथा पाया भी गया है कि उपज के पूर्वानुमान के लिए प्रौद्योगिकीय विकास के विभिन्न घटकों के प्रभावों का पृथक—पृथक आकलन करने की आवश्यकता नहीं है वरन् वर्ष—क्रमांक (year number) अथवा उसके उपयुक्त फलन (function) को मॉडल में स्वतन्त्र चर (independent variable) के रूप में लेना पर्याप्त है। इसके अतिरिक्त, विभिन्न वर्षों में उपज वर्षा, तापमान, आर्द्रता आदि मौसम चरों से प्रभावित होती है। इस प्रकार वर्ष—क्रमांक तथा मौसम—चरों पर आधारित मॉडल का विकास किया जा सकता है। अब प्रश्न यह उठता है कि मौसम चर किस अवधि के लिए जाएँ। मौसम—चरों का प्रभाव फ़सल के विभिन्न चरणों

\* पूर्व प्रमुख वैज्ञानिक, भा.कृ.अनु.प.—भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली

में अलग-अलग होता है। उदाहरण के लिए जिस समय फसल को आवश्यकता है उस समय वर्षा न हो और तब हो जाए जब वर्षा की आवश्यकता नहीं है तो वर्षा का कुल मान उपयुक्त (appropriate) होते हुए भी फसल के लिए हानिकारक होगा। अतः मौसम-चरों की पूरी अवधि में कुल मान (total magnitude) पर आधारित मॉडलों से सही पूर्वानुमान प्राप्त नहीं हो सकेंगे। इस समस्या के समाधान के लिए फसल की कुल अवधि को लघु समय-अन्तरालों (small time intervals) में विभाजित करके इन समय-अन्तरालों में फसल पर मौसम चरों के प्रभावों का अध्ययन करना अपेक्षित है। परन्तु इस प्रक्रिया में स्वतंत्र चरों की संख्या तथा इसके फलस्वरूप प्राचलों (parameters) की संख्या बहुत बढ़ जाती है और इतने अधिक प्राचलों के विश्वसनीय आकलन (reliable estimation) के लिए अत्यधिक आँकड़ों की आवश्यकता होती है जो प्रायः उपलब्ध नहीं होते। अतः किसी ऐसी विधि के विकास की आवश्यकता है जो लघु समय-अन्तरालों में मौसम-चरों के उपज पर प्रभाव को मॉडल में प्रदर्शित कर सके तथा साथ ही मॉडलों में प्राचलों की संख्या अधिक न बढ़े। इस संदर्भ में फिशर ने 1924 में एक मॉडल का विकास किया था जिसमें उन्होंने फसल की कुल अवधि को सप्ताहों में विभाजित किया। फिर यह माना कि विभिन्न सप्ताहों में फसल पर मौसम चरों का प्रभाव आकस्मिक (abrupt) या अनियमित (erratic) न होकर किसी नियम पर आधारित होगा जिसे गणितीय समीकरण द्वारा प्रदर्शित किया जा सकता है। उन्होंने माना कि प्रभाव सप्ताह-क्रमांक (week number) के पाँच घातीय लाम्बिक बहुपद (5<sup>th</sup> degree orthogonal polynomial) द्वारा प्रदर्शित किया जा सकता है। विभिन्न सप्ताहों में मौसम-चरों के मान को भी उन्होंने सप्ताह-क्रमांक के पाँच घातीय लाम्बिक बहुपद द्वारा प्रदर्शित किया। इन बहुपदों को उपज व साप्ताहिक मौसम चरों के बीच समाश्रयण समीकरण (regression equation) में लगा कर मॉडल प्राप्त किया जिसका उपयोग गेहूँ पर वर्षावृष्टि (rainfall) के प्रभाव के उदाहरण द्वारा प्रदर्शित किया। इस मॉडल में प्राचलों की संख्या सात

थी। हेंड्रिक्स एवं शॉल (1943) ने फिशर के मॉडल में संशोधन करते हुए विभिन्न सप्ताहों में मौसम-चरों के प्रभावों को प्रदर्शित करने के लिए पाँच घातीय लाम्बिक बहुपद के वास्तविक प्राप्त (observed) मानों का उपयोग करके समाश्रयण समीकरण का विकास किया। इस मॉडल में एक मौसम-चर के प्रभाव का अध्ययन करने के लिए केवल चार प्राचलों की आवश्यकता है। प्राचलों की संख्या फसल की कुल अवधि में समय-अंतरालों की संख्या पर निर्भर नहीं है। अतः फिशर तथा हेंड्रिक्स एवं शॉल द्वारा प्रस्तावित मॉडलों में प्राचलों की संख्या बढ़ाये बिना फसल की अवधि को लघु-अंतरालों में विभाजित करके अध्ययन किया जा सकता है। मॉडल इस पर निर्भर नहीं है कि फसल की कुल अवधि को कितने समय-अन्तरालों (time intervals) में विभाजित किया गया है। भारतीय कृषि सांख्यिकी अनुसंधान संस्थान में इस विधि में संशोधन करके मॉडलों का विकास किया गया। संशोधित विधि में विभिन्न सप्ताहों में मौसम-चर के फसल पर प्रभाव को सप्ताह क्रमांक (week number) के द्विघातीय बहुपद के स्थान पर उपज के साथ मौसम-चर के सहसम्बन्ध गुणांक (correlation coefficient) के उपयुक्त फलन को (आरम्भ में द्विघातीय तथा पुनः संशोधन करके रैखिक) लिया गया। इस संशोधन के साथ प्राप्त मॉडल हेंड्रिक्स एवं शॉल मॉडल से बेहतर पाया गया। कुछ अन्य संशोधन भी किये गये तथा कुछ चयनित फसलों तथा जिलों के लिए पूर्वानुमान मॉडल विकसित किये गये।

### जिला स्तर पर पूर्वानुमान मॉडल

प्रस्तुत (recommended) मॉडल का गणितीय स्वरूप इस प्रकार था

$$Y = A_0 + \sum_{i=1}^P \sum_{j=0}^1 a_{ij} Z_{ij} + \sum_{i \neq i'=1}^P \sum_{j=0}^1 a_{ii'j} Z_{ii'j} + c^T + e$$

जहाँ

$$Z_{ij} = \sum_{w=1}^m r_{iw}^j X_{iw}; Z_{ii'j} = \sum_{w=1}^m r_{ii'w}^j X_{iw} X_{i'w}$$

Y	उपज
$X_{i,w}$	समय अन्तराल 'w' में 'i' वें मौसम चर का मान
$Z_{ij}$	मौसम चर $X_i$ के सूचक
$Z_{ii'j}$	$X_i$ तथा $X_{i'}$ के संयुक्त सूचक
T	वर्ष-क्रमांक
$r_{i,w}$	वर्ष के प्रभाव के लिए समायोजित (adjusted) उपज तथा समय अंतराल 'w' में 'i' वें मौसम चर के बीच सहसम्बन्ध गुणांक
$r_{ii'w}$	वर्ष के प्रभाव के लिए समायोजित उपज का समय अंतराल 'w' में $X_i$ तथा $X_{i'}$ के गुणनफल के साथ सहसम्बन्ध गुणांक
m	पूर्वानुमान का समय
P	उपयोग किये गये मौसम-चरों की संख्या
e	यादृच्छिक त्रुटि
$A_o, a_{ij}, a_{ii'j}$	तथा c मॉडल के प्राचल

इस मॉडल में प्रत्येक मौसम-चर के लिए दो सूचक विकसित किये गये – एक मौसम-चर का कुल योग अर्थात् अभाषित (unweighted) सूचक तथा दूसरा विभिन्न समय-अंतरालों में मौसम-चरों के उपज के साथ सह-सम्बन्ध गुणांक का भार देकर भारित योग अर्थात् भारित (weighted) सूचक। इसी प्रकार दो मौसम चरों के संयुक्त प्रभाव का आकलन करने के लिए विभिन्न समय-अंतरालों में मौसम चरों के गुणनफल लेकर मौसम-सूचक विकसित किये गये। इन मौसम-सूचकों तथा वर्ष-क्रमांक को स्वतन्त्र चर लेकर पदशः समाश्रयण पद्धति (stepwise regression) का उपयोग करके मॉडल विकसित किये गये। ये मॉडल लघु अवधि फ़सलों (चावल के लिए कम वर्षा वाले वर्षापोषित (rainfed with deficit rainfall) जिले – रायपुर, पर्याप्त वर्षा वाले वर्षापोषित (rainfed with adequate rainfall) जिले – पुरी, गेहूँ के लिए सिंचित (irrigated) जिले – अमृतसर) दीर्घ अवधि फ़सल (गन्ना – कोल्हापुर) के लिए विकसित किये गये। गेहूँ तथा चावल के लिए मौसम-चरों के साप्ताहिक आँकड़ों

का उपयोग किया गया। बुआई के समय से पहले का मौसम भी महत्त्वपूर्ण होता है अतः फ़सल की बुआई से एक पखवाड़ा पहले से लेकर पूर्वानुमान समय तक के आँकड़ों का उपयोग किया गया। गन्ने के लिए मौसमचरों के पखवाड़ा-वार आँकड़ों का उपयोग किया गया। (अग्रवाल इत्यादि 1980, 83, 86; जैन इत्यादि 1980; मेहता इत्यादि 2000)।

परिणामों से पाया गया कि इन मॉडलों द्वारा आगामी वर्षों के गेहूँ व चावल की उपज के विश्वसनीय पूर्वानुमान (reliable forecasts) कटाई से लगभग दो माह पूर्व तथा गन्ने की उपज का पूर्वानुमान कटाई के लगभग 3 माह पूर्व प्राप्त किया जा सकता है। रायपुर जिले में (जहाँ मॉडलों के विकास हेतु 20 वर्ष के आँकड़े उपलब्ध थे) वास्तविक प्राप्त उपज से पूर्वानुमान का अन्तर 5% था जबकि अन्य जिलों में (जहाँ आँकड़े कम थे) यह अन्तर 5 से 10% तक पाया गया। इन मॉडलों में प्रयुक्त चर (variables used), पूर्वानुमान समय (time of forecast) तथा प्राप्त वास्तविक उपज (observed yield) से पूर्वानुमान (forecast) का अन्तर आदि का विवरण चित्र 1 में दिया गया है।

प्राप्त उपज से विचलन Deviation from observed yield			
10%	5%	10%	5-10%
पूर्वानुमान समय Time of Forecast			
कटाई से 2 1/2 माह पूर्व 2 1/2 months before harvest		बुआई से 8-9 माह बाद 8-9 months after planting	
प्रयुक्त चर Variables used			
<ul style="list-style-type: none"> <li>• माध्य तापमान</li> <li>• तापमान अन्तर</li> <li>• Mean Temperature</li> <li>• Temperature difference</li> </ul>	<ul style="list-style-type: none"> <li>• अधिकतम तापमान</li> <li>• सापेक्ष आर्द्रता (2 बचे)</li> <li>• वर्षा</li> <li>• Maximum Temperature</li> <li>• Relative Humidity (14hrs)</li> <li>• Rainfall</li> </ul>	<ul style="list-style-type: none"> <li>• अधिकतम तापमान</li> <li>• घट की घटे</li> <li>• सापेक्ष आर्द्रता</li> <li>• वर्षा की दिनों की संख्या</li> <li>• Maximum Temperature</li> <li>• Sunless Hours</li> <li>• Relative Humidity</li> <li>• No. of Rainy days</li> </ul>	<ul style="list-style-type: none"> <li>• अधिकतम तापमान</li> <li>• तापमान अन्तर</li> <li>• घट की घटे</li> <li>• सापेक्ष आर्द्रता</li> <li>• वर्षा की दिनों की संख्या</li> <li>• Maximum Temperature</li> <li>• Temperature difference</li> <li>• Relative Humidity</li> <li>• No. of Rainy days</li> </ul>
अमृतसर Amritsar	रायपुर Raipur	पुरी Puri	कोल्हापुर Kolhapur
गेहूँ Wheat	चावल Rice		गन्ना Sugarcane
मौसम सूचक आधारित पद्धति Weather Indices Based Methodology			

चित्र 1: जिला स्तर पूर्वानुमान

### कृषि जलवायवीय स्तर पर मॉडल

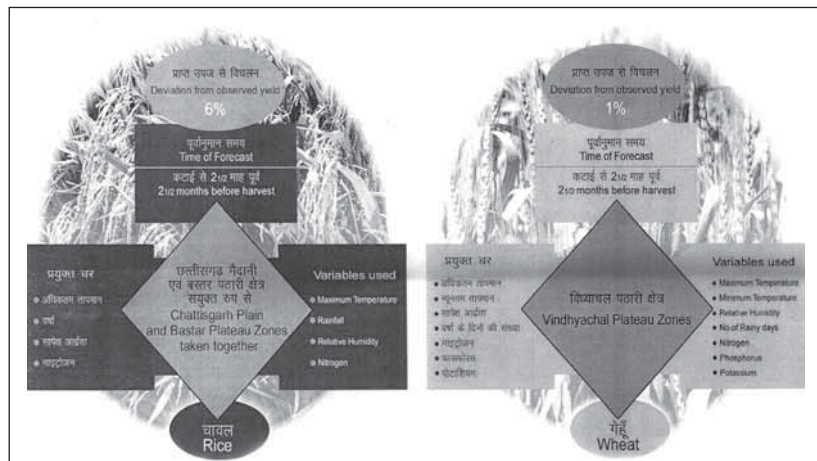
जिला स्तर पर किये गये अध्ययनों से यह भी निष्कर्ष निकाला कि इन मॉडलों से उपज के विश्वसनीय पूर्वानुमान प्राप्त करने के लिए लगभग 20-25 वर्षों के आँकड़ों की आवश्यकता होगी। प्रायः कई स्थानों के

लिए इतनी लम्बी अवधि के आँकड़े उपलब्ध नहीं होते हैं। अतः यह सोचा गया कि जिला स्तर के स्थान पर कृषि-जलवायवीय क्षेत्र (agro-climatic zone) स्तर पर मॉडलों का विकास किया जाए जिससे कि एक कृषि-जलवायवीय क्षेत्र के अंतर्गत विभिन्न जिलों के आँकड़ों का उपयोग करके कम अवधि में आँकड़ों की लम्बी श्रृंखला प्राप्त की जा सके। यह माना जा सकता है कि एक कृषि-जलवायवीय क्षेत्र के विभिन्न जिलों में फसल तथा मौसम-चरों के प्रभावों की स्थिति लगभग समान होगी, अतः इन आँकड़ों को एक श्रृंखला मान कर मॉडल का विकास किया जा सकता है। इस अवधारणा का परीक्षण करने के लिए गेहूँ व चावल के लिए अध्ययन किया गया (अग्रवाल इत्यादि 2001)। गेहूँ के लिए मध्य प्रदेश में विन्ध्याचल के पठारी क्षेत्र का चयन किया गया तथा चावल के लिए छत्तीसगढ़ का मैदानी व बस्तर का पठारी क्षेत्र संयुक्त रूप से लिया गया (बस्तर जिले का कुछ भाग छत्तीसगढ़ के मैदानी क्षेत्र में आता है तथा कुछ भाग बस्तर के पठारी क्षेत्र में और उपज के आँकड़े पूरे जिले के ही उपलब्ध थे अतः दोनों क्षेत्रों को संयुक्त रूप से लेकर मॉडल विकसित करना पड़ा)।

मॉडलों के विकास के लिए चयनित कृषि-जलवायवीय क्षेत्र के अंतर्गत आने वाले विभिन्न जिलों के लिए उपज के आँकड़े, बुवाई के दो सप्ताह पूर्व से पूर्वानुमान समय तक के साप्ताहिक मौसम-चरों के आँकड़े, प्रतिशत सिंचित क्षेत्रफल, अधिक उपज

वाली प्रजातियों (high yielding varieties) के अंतर्गत प्रतिशत क्षेत्रफल तथा उर्वरकों (नाइट्रोजन, फासफोरस, पोटैश) के आँकड़ों का उपयोग किया गया। जिला स्तर के मॉडल में सुधार करके कृषि-जलवायवीय क्षेत्र के स्तर पर पूर्वानुमान मॉडलों का विकास किया गया। कृषि-जलवायवीय क्षेत्र के अंतर्गत विभिन्न जिलों की उपज में विचलन के आकलन के लिए जिला स्तर पर कृषि-निवेशों (agricultural inputs), पिछले वर्ष की उपज तथा पिछले 2 से 5 वर्षों की उपज के गतिमान औसत (moving average) का उपयोग किया गया।

मौसम तथा कृषि निवेशों (agricultural inputs) के आँकड़ों पर आधारित मॉडल केवल मौसम आधारित मॉडल की अपेक्षा अधिक उपयुक्त पाया गया। इस मॉडल से विश्वसनीय पूर्वानुमान बुवाई के दो सप्ताह पहले से बुवाई के 12 सप्ताह बाद तक के आँकड़ों के आधार पर (कटाई से दो माह पूर्व) प्राप्त किये जा सकते हैं। जिला स्तर पर विश्वसनीय पूर्वानुमान के लिए 20-25 वर्षों की तुलना में कृषि जलवायवीय क्षेत्र के स्तर पर विश्वसनीय पूर्वानुमान के लिए 10-15 वर्षों के आँकड़ें पर्याप्त पाये गये। यह भी देखा गया कि यदि मॉडल के विकास अथवा पूर्वानुमान के समय कुछ जिलों के आँकड़े उपलब्ध न हों तो भी अन्य जिलों के उपलब्ध आँकड़ों के आधार पर कृषि-जलवायवीय क्षेत्र की उपज का विश्वसनीय पूर्वानुमान संभव है। गेहूँ के लिए वास्तविक प्राप्त उपज से पूर्वानुमान का अन्तर मात्र 1% पाया गया। चावल के लिए यह अन्तर 6%



चित्र 2: मौसम सूचक आधारित पद्धति



था। यह अपेक्षित भी था क्योंकि दो कृषि-जलवायवीय क्षेत्रों के लिए संयुक्त मॉडल का विकास किया गया था और दोनों क्षेत्रों की कृषि-जलवायवीय स्थितियाँ एक समान होने की सम्भावना नहीं है। मौसमचरों तथा कृषि निवेशों पर आधारित मॉडलों का विवरण चित्र 2 में दिया गया है।

### राज्य स्तर पूर्वानुमान

मौसम सूचक आधारित मॉडलों के परिणाम कुछ चयनित जिलों तथा कृषि जलवायवीय क्षेत्रों के लिए अत्यन्त उत्साहवर्धक पाये गये। अतः व्यापक स्तर पर इस विधि के उपयोग का अध्ययन करने के लिए उत्तरप्रदेश राज्य के विभिन्न जिलों / कृषि जलवायवीय क्षेत्रों के लिए अध्ययन किया गया (अग्रवाल इत्यादि 2005)। अध्ययन के लिए एक प्रमुख खरीफ फसल (चावल), एक प्रमुख रबी फसल (गेहूँ), दीर्घकालीन फसल (गन्ने) तथा एक सब्जी (आलू) का चयन किया गया। जिला स्तर (25 जिलों के लिए) तथा कृषि-जलवायवीय क्षेत्र स्तर (8 क्षेत्रों के लिए) पर मॉडलों का विकास किया गया। यह पाया गया कि केवल भारत मौसम सूचकों (weighted weather indices) का उपयोग करके विकसित किये गये मॉडल भारत (weighted) तथा अभारित (unweighted) दोनों प्रकार के सूचकों पर आधारित मॉडलों के समकक्ष थे अतः मॉडल पिछले मॉडलों की तुलना में अधिक सरल हो गये।

कृषि-जलवायवीय क्षेत्र के स्तर पर विकसित किये गये मॉडलों में विभिन्न चरों में जिलों की विभिन्नता (variability) हटाकर आँकड़ों की समांगी (homogeneous) श्रृंखला बनाने के लिए जिले के औसत से विचलन (deviations) लिए गये। इन मॉडलों का उपयोग करके आगामी वर्षों के लिए विभिन्न जिलों के लिए पूर्वानुमान निकाले गए और इन पूर्वानुमानों से फसल के अंतर्गत क्षेत्रफल का भार देकर भारत योग लेकर कृषि-जलवायवीय क्षेत्र तथा राज्य स्तर पर पूर्वानुमान प्राप्त किये गये। राज्य स्तर पर प्राप्त पूर्वानुमानों की तुलना दो प्रकार से प्राप्त उपज के साथ की गयी – एक उन जिलों पर आधारित उपज जिनके लिए पूर्वानुमान प्राप्त किये गये और दूसरी राज्य के

सभी सत्तर जिलों पर आधारित।

इन मॉडलों द्वारा प्राप्त निष्कर्षों से पहले किये गये अध्ययनों के निष्कर्षों की पुष्टि हुई अर्थात् चावल तथा गेहूँ के लिए लगभग 2 माह पूर्व उपज के विश्वसनीय पूर्वानुमान प्राप्त किये जा सकते हैं तथा गन्ने के लिए सितम्बर के मध्य में आलू के विश्वसनीय पूर्वानुमान चार सप्ताह पूर्व प्राप्त किये जा सकते हैं। गेहूँ तथा गन्ने के लिए पूर्वानुमान तथा वास्तविक प्राप्त उपज (उन जिलों पर आधारित जिनके लिए पूर्वानुमान विकसित किये गये) में अन्तर 1% से भी कम था। चावल, आलू तथा राज्य के सभी जिलों पर आधारित वास्तविक प्राप्त उपज से पूर्वानुमान में अन्तर लगभग 4-5% पाया गया। जिला स्तर पर विकसित मॉडलों के परिणाम कृषि जलवायवीय क्षेत्र स्तर पर विकसित मॉडलों के परिणामों से बेहतर पाये गये। अतः यदि समुचित आँकड़े उपलब्ध हों तो जिला स्तर पर मॉडलों का विकास अधिक उपयुक्त रहेगा। मॉडलों के परिणाम तालिका में दिये गये हैं। इस तकनीक का उपयोग अन्तरिक्ष अनुप्रयोग केन्द्र, अहमदाबाद द्वारा राष्ट्र-स्तर पर गेहूँ की उपज का पूर्वानुमान प्राप्त करने के लिए किया गया। उपज-पूर्वानुमान तथा वास्तविक प्राप्त उपज में मात्र तीन प्रतिशत का अन्तर पाया गया। वर्तमान में इस तकनीक का उपयोग FASAL (Forecasting Agricultural Production using Satellite, Agrometeorology and Land based observations) कार्यक्रम में हो रहा है।

### तालिका : राज्यस्तर पूर्वानुमान

वास्तविक मान से पूर्वानुमान का अन्तर (%)	चावल	गेहूँ	गन्ना	आलू
25 जिलों पर आधारित वास्तविक उपज	4.2	0.7	0.8	3.8
उ.प्र. के सारे जिलों पर आधारित वास्तविक उपज	3.3	4.3	4.5	4.7

प्रस्तावित पूर्वानुमान तकनीक सरल है तथा उपयोग में सुविधाजनक है। इसमें आँकड़ों के विस्तृत संकलन (detailed data collection) अथवा जटिल सांख्यिकीय विधियों (Complicated statistical techniques) के

उपयोग की आवश्यकता नहीं है तथा इस तकनीक से काफी अच्छे पूर्वानुमान प्राप्त होते हैं। यह तकनीक जिला, कृषि-जलवायवीय क्षेत्र के साथ-साथ राज्य स्तर के पूर्वानुमान के लिए भी विश्वसनीय पायी गयी है। संस्थान द्वारा विकसित इस तकनीक का विभिन्न संगठनों द्वारा सफलतापूर्वक उपयोग किया जा रहा है।

### संदर्भ

1. अग्रवाल, रंजना, जैन, आर.सी. एवं सिंह, डी. (1980): "फोरकास्टिंग ऑफ राइस यील्ड यूजिंग क्लाइमेटिक वेरिबिलिस", इंडियन जर्नल ऑफ एग्रीकल्चरल साइंस, 50(9) : 680-684।
2. अग्रवाल, रंजना, जैन, आर.सी. एवं झा, एम.पी. (1983): "ज्वाइंट इफेक्ट्स ऑफ वेदर वेरिबिलिस ऑन राइस यील्ड", मौसम, 34(2) : 189-194।
3. अग्रवाल, रंजना, जैन, आर.सी. एवं झा, एम.पी. (1986): "मॉडल्स फॉर स्टडीइंग राइस क्रॉप-वेदर रिलेशनशिप", मौसम, 37(1) : 67-70।
4. अग्रवाल, रंजना, जैन, आर.सी. एवं मेहता, एस.सी. (2001): "यील्ड फोरकास्ट बेस्ड ऑन वेदर वेरिबिलिस एण्ड एग्रीकल्चरल इनपुट्स ऑन एग्रो-क्लाइमेटिक जोन बेसिस", इंडियन जर्नल ऑफ एग्रीकल्चरल साइंस, 71(7) : 487-490।
5. अग्रवाल, रंजना, रामकृष्णा, वाई.एस., केशव राव, ए.वी.आर., कुमार, अमरेन्द, भर, लाल मोहन, मदन मोहन एवं सक्सेना, आशा (2001): "मॉडलिंग फॉर फोरकास्टिंग ऑफ क्रॉप यील्ड यूजिंग वेदर पैरामीटर्स एण्ड एग्रीकल्चरल इनपुट्स", ए.पी. सेस फण्ड परियोजना रिपोर्ट।
6. फिशर, आर. (1924): "दि इन्फ्लूएन्स ऑफ रेनफाल ऑन दि यील्ड ऑफ व्हीट ऐट रोदेमस्टेड", रॉयल सोसासटी (लंदन), फिल. ट्रांस. सीरीज बी., 213 : 89-142।
7. हेंड्रिक्स, डब्ल्यू.ए. एवं शॉल, जे.सी. (1943): "टेक्नीक इन मेजरिंग ज्वाइंट रिलेशनशिप. दि ज्वाइंट इफेक्ट्स ऑफ टेम्प्रेचर एण्ड प्रेसिपिटेशन ऑन क्रॉप यील्ड", नार्थ कैरोलिना एग्रीकल्चरल स्टेशन टेक्निकल बुलेटिन: 74।
8. जैन, आर.सी., अग्रवाल, रंजना, एवं झा, एम. पी. (1980): "इफेक्ट्स ऑफ वेदर क्लाइमेटिक वेरिबिलिस ऑन राइस यील्ड एण्ड इट्स फोरकास्ट", मौसम, 31(4): 591-596।
9. मेहता, एस.सी., अग्रवाल, रंजना एवं सिंह, वी. पी.एन. (2000): "स्ट्रेटेजीज फॉर कम्पोजिट फोरकास्ट", जर्नल ऑफ इंडियन सोसायटी ऑफ एग्रीकल्चरल स्टैटिस्टिक्स, 53(3): 262-272।

□

## भा.कृ.अनु.प. में एकीकृत संचार एवं वेब होस्टिंग – एक अवलोकन

अंजनी कुमार चौबे, अलका अरोड़ा, सुदीप मरवाह, एन.एस. राव,  
मुकेश कुमार, राकेश सैनी, सुभाष चन्द एवं ऊषा जैन

भारतीय कृषि अनुसंधान परिषद् (भा.कृ.अनु.प.) देश में कृषि के साथ-साथ बागवानी, मत्स्य पालन एवं पशु विज्ञान में अनुसंधान एवं शिक्षा के समन्वय, मार्गदर्शन एवं प्रबन्धन के लिए एक सर्वोच्च संस्था है। भारतीय कृषि अनुसंधान परिषद् ने अपने अनुसंधान एवं प्रौद्योगिकी के विकास के माध्यम से हरित क्रान्ति और भारत में कृषि के क्षेत्र में अनुवर्ती विकास को कायम रखने में एक अग्रणी भूमिका निभाई है, जिससे 1930-31 से देश में खाद्यान्नों के उत्पादन में 4 गुना, बागवानी फसलों में 6 गुना, मत्स्य उत्पादन में 9 गुना (समुद्री 5 गुना तथा अन्तरदेशीय 17 गुना), दूध 5 गुना तथा अण्डे का उत्पादन 27 गुना बढ़ा है, जिसका राष्ट्र की खाद्य एवं पोषण सुरक्षा पर स्पष्ट प्रभाव दिखाई पड़ता है<sup>(1)</sup>।

भा.कृ.अनु.प. के 100 से अधिक संस्थान एवं उनके क्षेत्रीय केन्द्र हैं जो देश भर में फैले हुए हैं। ये सभी संस्थान अनुसंधान, शिक्षा एवं प्रसार से सम्बन्धित कार्यों में एकीकृत संचार प्रौद्योगिकी [ICT (आई.सी.टी.)] का गहन उपयोग करते हैं। आई.सी.टी. का उपयोग परियोजनाओं के नियोजन, क्रियान्वयन एवं मॉनीटरिंग में भी किया जा रहा है। भारतीय कृषि सांख्यिकी अनुसंधान संस्थान (भा.कृ.सां.अ.सं.), भा.कृ.अनु.प. का एक अग्रणी संस्थान है जो कृषि सांख्यिकी, जैव-सूचना विज्ञान एवं संगणक अनुप्रयोग के क्षेत्र में अनुसंधान, अध्यापन एवं परीक्षण को बढ़ावा देता है। संगणक अनुप्रयोग प्रभाग, संस्थान का ऐसा प्रभाग है जो परिषद् में प्रमुख रूप से अनुसंधान के लिए कम्प्यूटर सॉफ्टवेयर, डाटाबेस

एवं आई.सी.टी. सक्षम करने में महत्वपूर्ण भूमिका निभा रहा है। डाटा केन्द्र के निर्माण से प्रभाग ने 24x7 डाटा केन्द्र के विकास एवं ऑपरेशन में विशेषज्ञता प्राप्त कर ली है<sup>(2)</sup>।

सूचना एवं संचार प्रौद्योगिकी, कृषक समुदाय को सामयिक एवं सही सूचना तथा सलाह के लिए ऐक्सेस उपलब्ध कराने के लिए कृषि अनुसंधान संस्थानों को सक्षम बना रही है। ऑन-लाइन सेवाओं के लिए उपभोक्ताओं की बढ़ती हुई अपेक्षाओं, सूचना की सुरक्षित ऐक्सेस, संगठनात्मक ई-मेल एवं सहयोग तथा परिषद् द्वारा क्रियान्वित किये जा रही ई-गवर्नेन्स परियोजनाओं की बढ़ती हुई संख्या से भारतीय कृषि अनुसंधान एवं शिक्षा प्रणाली में डाटा केन्द्र की आवश्यकताओं में वृद्धि हुई है।

वर्तमान में, भा.कृ.अनु.प. के संस्थानों द्वारा विभिन्न ई-मेल सुविधाओं का प्रयोग किया जा रहा है जो निजी सेवा प्रदाताओं द्वारा प्रदान की जाती हैं तथा एप्लीकेशन्स की वेब होस्टिंग NIC, ERNET के उपयोग द्वारा एवं निजी सेवा प्रदाताओं द्वारा की जा रही है। भा.कृ.अनु.प. के प्रत्येक कार्मिक के लिए भा.कृ.अनु.प. डोमेन के साथ विशिष्ट ई-मेल, एकीकृत संदेश (unified messaging) एवं वेब होस्टिंग समाधान उपलब्ध कराने के लिए कोई केन्द्रीयकृत सुविधा नहीं है।

इन आवश्यकताओं को पूरा करने के लिए, भारतीय कृषि अनुसंधान परिषद् (भा.कृ.अनु.प.) द्वारा एन.ए.आई. पी. परियोजना के अन्तर्गत निम्नलिखित उद्देश्यों से भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली

में स्टेट-ऑफ-दि-आर्ट डाटा केन्द्र का सृजन किया गया है:

1. भा.कृ.अनु.प. के कार्मिकों (25000) के बीच सहयोग उपलब्ध कराने के लिए ई-मेल एवं एकीकृत मैसेजिंग सोल्यूशन उपलब्ध कराना ।
2. भा.कृ.अनु.प. के संस्थानों द्वारा विकसित वेबसाइट्स एवं एप्लीकेशन्स की होस्टिंग के लिए आई.सी.टी. के बुनियादी ढाँचे का सृजन करना तथा कॉमन सर्विस डिलिवरी प्लेटफार्म का निर्माण करना ।

### डाटा केन्द्र की विशेषताएँ

- ◆ भा.कृ.अनु.प. के सभी कार्मिकों के लिए विशिष्ट आई.डी. (abc@icar.gov.in)
- ◆ उपलब्धता-99 % एस.एल.ए.
- ◆ विश्वसनीयता
- ◆ गतिशीलता
- ◆ स्केलेबिलिटी
- ◆ प्रबन्धन क्षमता
- ◆ BYOD किसी भी डिवाइस (Desktop, Laptop, Mobile) को नेटवर्क में लगाना ।
- ◆ भारत सरकार की आई.टी. पॉलिसी के अनुसार लॉग प्रबन्धन ।

### भा.कृ.सां.अ.सं. के डाटा केन्द्र की सेवाएँ

डाटा केन्द्र केन्द्रीयकृत समर्थित स्थान है जहाँ कोई भी संगठन डाटा एवं सूचना का भण्डारण, प्रबन्धन एवं प्रसार के लिए अधिकांश आई.सी.टी. बुनियादी सुविधाओं का प्रबन्धन करता है। इसमें सर्वर एवं भण्डारण (Storage) उपकरण होते हैं जो एप्लीकेशन सॉफ्टवेयर को रन करते हैं, प्रोसेस करतें हैं तथा डाटा एवं सामग्री को स्टोर करते हैं। डाटा केन्द्र के सिस्टम के निष्पादन एवं संचालन की विश्वसनीयता को सुनिश्चित करने के लिए तापमान एवं आर्द्रता के संदर्भ में वातावरण नियन्त्रित किया जाता है। आमतौर पर, इसमें पॉवर सप्लाय, बैकअप पॉवर, चिलर्स, कैबलिंग, आग एवं पानी का पता लगाने वाला सिस्टम, सुरक्षा नियंत्रण, इत्यादि सुविधाएँ शामिल होती हैं।

भा.कृ.सां.अ.सं. के संगणक अनुप्रयोग प्रभाग में स्थापित डाटा केन्द्र को चित्र-1 में दर्शाया गया है ।



चित्र 1 : भा.कृ.सां.अ.सं., नई दिल्ली में स्थापित डाटा केन्द्र

भा.कृ.सां.अ.सं. में सृजित डाटा केन्द्र आई.एस.ओ. 27001, आई.टी.आई.एल. एवं टी.आई.ए. 942 मानकों का पालन करते हुए डाटा केन्द्र के टायर-III के विनिर्देशों के अन्तरराष्ट्रीय मानक को पूरा करता है। कम्प्यूटिंग वातावरण में 150 TB भण्डारण स्पेस सहित 448 कोर कम्प्यूटिंग उपलब्ध है जो आँकड़ों की कम्प्यूटिंग/प्रोसेसिंग की उच्च गति को सक्षम बनाता है। मेल मैसेजिंग एवं वेबसाइट्स/एप्लीकेशन्स सेवाओं की होस्टिंग उपलब्ध कराने के लिए विभिन्न डाटाबेसेस (MYSQL, MS-SQL SERVER) के साथ माइक्रोसॉफ्ट एक्सचेन्ज मेल सर्वर, ऑडियो, वीडियो, वेब कॉन्फ्रेन्सिंग के लिए माइक्रोसॉफ्ट Lync,

WINDOWS एवं LINUX OS डाटा केन्द्र में उपलब्ध हैं। इस सॉल्यूशन में वर्चुअलाइजेशन प्लेटफार्म उपलब्ध है। डायरेक्ट्री सोल्यूशन, ई-मेल, Lync एवं एक्टिव डायरेक्टरी (AD) के लिए सिंगल साइन-ऑन की सुविधा सुरक्षित ऐक्सेस के माध्यम से उपलब्ध है।

यह सॉल्यूशन एकीकृत संदेश एवं वेब होस्टिंग के लिए भा.कृ.अनु.प. की आई.सी.टी. बुनियादी ढाँचे की समस्त आवश्यकताओं को समेकित करने में मदद करेगा।

## तकनीकी आर्किटेक्चर

### (i) बुनियादी ढाँचा

डाटा केन्द्र का बुनियादी ढाँचा डाटा केन्द्र के सर्वोत्तम मानकों के आधार पर एवं उपभोक्ताओं की जरूरतों को ध्यान में रखते हुए डिजाइन किया गया है।

डाटा केन्द्र के महत्वपूर्ण बुनियादी ढाँचे में निम्न सिस्टम शामिल हैं :

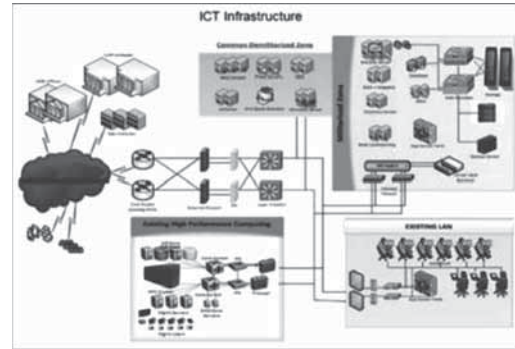
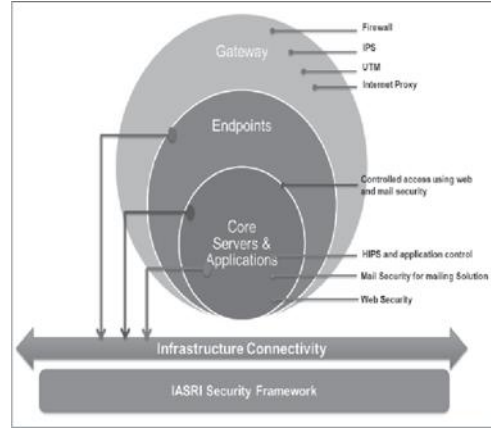
- ◆ सिविल एवं अन्तःप्रवेष्टा (interior) कार्य ।
- ◆ विद्युत – डी.जी., यू.पी.एस., पी.ए.सी. एवं रैक्स, जनरेटर सेट, पी.ए.सी. ।
- ◆ एयर चिलर (पानी पर आधारित) प्रणाली, प्रिसिजन एयर कंडीशनिंग एवं कम्फर्ट शीतलन प्रणाली ।
- ◆ सुरक्षा और निगरानी प्रणाली ।
- ◆ प्रवेश नियंत्रण एवं सीसीटीवी सिस्टम ।
- ◆ दोहरी (dual) पॉवर फीड ।
- ◆ एन + 1 150 KVA यू.पी.एस ।
- ◆ 150 KVA डीजल जनरेटर ।
- ◆ फायर संरक्षण, अग्नि-शमन प्रणाली ।
- ◆ पानी के रिसाव का पता लगाने के लिए अलार्म ।
- ◆ रोडेन्ट रेप्लेन्ट सिस्टम ।
- ◆ सार्वजनिक घोषणा प्रणाली ।

### (ii) आई.टी.

डाटा केन्द्र में राष्ट्रीय ज्ञान नेटवर्क (National Knowledge Network) द्वारा 100 mbps इन्टरनेट

कनेक्टिविटी उपलब्ध है। सुरक्षा डिवाइसेस के समस्त प्लेटफार्म, अर्थात् आई.पी.एस., पैरीमीटर फायरवॉल, आंतरिक (MZ) फायरवॉल, वेब प्रॉक्सी एवं लॉग प्रबन्धन के लिए दोहरे (dual) आधुनिक उपकरण उपलब्ध हैं।

एप्लीकेशन एवं उपकरणों को आवश्यकतानुसार मिलिटराइज एवं डिमिलिटराइज ज़ोन में रखा गया है। नेटवर्क सेवाओं को चित्र-2 में दर्शाया गया है।



चित्र 2 : नेटवर्क सेवाओं का आरेख

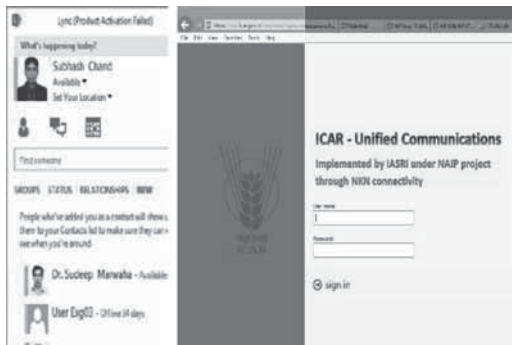
## एकीकृत संचार एवं वेब होस्टिंग (Unified Communication & Web Hosting)

### (i) एकीकृत संचार (Unified Communication):

यह सोल्यूशन एकीकृत संचार के अन्तर्गत डेस्कटॉप, लैपटॉप, टैबलेट, स्मार्ट फोन का उपयोग करके त्वरित संदेश, ई-मेल, वॉयस एवं वीडियो कॉल, वेब कॉन्फ्रेंसिंग आदि की एकीकृत सुविधा प्रदान करता है।

भा.कृ.अनु.प. के निदेशानुसार परिषद् एवं संस्थानों

के सभी कर्मचारियों के लिए ई-मेल आई.डी. "icar.gov.in" डोमेन के साथ बनाया जा रहा है। इस ई-मेल आई.डी. का प्रयोग सभी सरकारी संचार के लिए किया जाएगा। यह परिषद् एवं इसके सभी संस्थानों को एक संगठन के रूप में वैश्विक मान्यता प्रदान करेगा। इसके लिए सुरक्षित कनेक्शन (<https://mail.icar.gov.in>) एस.एस.एल. क्रियान्वित किया गया है। एकीकृत संचार के विभिन्न माध्यम चित्र 3 में दर्शाए गए हैं। एकीकृत संचार की सुविधा विभिन्न माध्यमों, जैसे वेब (<https://mail.icar.gov.in>) आउटलुक (Outlook) एवं माइक्रोसॉफ्ट Lync द्वारा एक्सेस की जा सकती है<sup>(3)</sup>। इसके लिए प्रयोक्ता पुस्तिका misfms की साइट (<https://www.iasri.res.in/misfms>) पर डॉक्स (Docs) लिंक पर एकीकृत संचार (Unified Communication) में उपलब्ध है<sup>(4)</sup>।



चित्र 3 : एकीकृत कम्युनिकेशन के लिए विभिन्न उपकरण एवं माध्यम

## (ii) वेब होस्टिंग (Web Hosting)

भा.कृ.अनु.प. के सभी संस्थानों को उनकी वेबसाइट/पोर्टल/एप्लीकेशन्स के लिए वेब होस्टिंग सेवाएँ उपलब्ध होंगी। डाटा केन्द्र सुरक्षित माध्यम के साथ-साथ, डाटाबेस ऑपरेटिंग सिस्टम एवं वर्चुअलाइजेशन प्लेटफार्म भी उपलब्ध कराएगा। डाटा केन्द्र में माइक्रोसॉफ्ट विण्डोज 2012 R2 एवं रेड हेट लाइनक्स 7, MS SQL- 11-0-3128 एवं MySQL - 5-6-20 उपलब्ध है।



चित्र 4 : डाटा केन्द्र में होस्टिड वेब एप्लीकेशन

होस्टिंग के लिए क्लाउड सर्वर n-tier आर्किटेक्चर का प्रयोग होगा इसमें सभी स्तरों (tiers) से सम्बन्धित प्रौद्योगिकियों को अलग-अलग सर्वरों पर रखा जाएगा ताकि उनको उपयुक्त सुरक्षा उपलब्ध करायी जा सके। संस्थानों को वेब होस्टिंग सुविधा उपलब्ध कराने के लिए नीति (Policy) विकास का कार्य किया जा रहा है।

## (iii) सुविधाएँ

- ई-मेल (वेब मेल, आउटलुक, मोबाइल)
- दस्तावेज का पूर्वालोकन (preview), वार्तालाप थ्रेड (Conversation thread), कैलेण्डर, सभागार/सम्मेलन कक्ष की उपलब्धता, तीन दिनों के तापमान का पूर्वानुमान इत्यादि।
- त्वरित संदेश (Instant Messaging)
- आवाज, वीडियो कॉलिंग एवं वेब कॉन्फ्रेन्सिंग
- प्रस्तुति (Presentation) शेयरिंग

- डेस्कटॉप शेयरिंग
- प्रोग्राम शेयरिंग
- व्हाइट बोर्ड
- बैठक निमंत्रण
- डायरेक्ट्री सर्विस
- वेब कॉन्फ़ेसिंग पोलिंग
- डेस्कटॉप, लैपटॉप, मोबाइल, टैबलेट पर मेल का समन्वय
- वेब होस्टिंग
- केन्द्रीयकृत डाटाबेस
- भण्डारण एवं आर्किवल

### निष्कर्ष

भा.कृ.सां.अ.सं., नई दिल्ली में स्थापित डाटा केन्द्र भा.कृ.अनु.प. को आई.टी. गतिविधियों में अग्रणी एवं आत्मनिर्भर बनाएगा। इससे संचालित सुविधाएँ परिषद् के लिए कृषि अनुसंधान एवं संचार समन्वय में काफी लाभप्रद होंगी। एकीकृत संचार सुविधा भा.कृ.अनु.प. के कार्मिकों को इन्स्टेन्ट मैसेजिंग, ई-मेल, ओडियो/वीडियो कॉल के माध्यम से एक-दूसरे के साथ जुड़े रहने में मदद करेगी। अनुसंधानिक/शैक्षणिक/प्रसार सम्बन्धी जानकारी को ऑनलाइन बैठक एवं प्रस्तुतियों के माध्यम से शीघ्र एवं प्रभावी तरीके से पहुँचाया जा सकता है। सभी के लिए एक समान ई-मेल आई.डी. परिषद् को विश्वस्तरीय पहचान देगा।

डाटा केन्द्र में एक केन्द्रीयकृत भण्डारण की

सुविधा से संस्थानों में विकसित डाटाबेस, ज्ञान-आधार (Knowledge base) को एक जगह पर रखा जा सकेगा जिससे इनमें इण्टर-ऑपरेटिबिलिटी की सुविधा उपलब्ध होगी तथा विभिन्न मॉडलिंग तकनीकों का प्रयोग करके पूर्वानुमान मॉडलों सहित ज्ञान प्रबन्धन, निर्णय सपोर्ट प्रणाली एवं नीति नियोजन में बहुमूल्य सहायता प्राप्त होगी। एकीकृत संचार से एक ओर संचार की गुणवत्ता एवं क्षमता प्रभाव आएगा वहीं दूसरी ओर इससे वीडियो कॉलिंग द्वारा यात्रा खर्च की लागत में कमी आएगी तथा समय की भी बचत होगी।

### आभार

लेखकगण, एन.ए.आई.पी. द्वारा इस कार्य के लिए उपलब्ध करायी गयी धनराशि एवं सहयोग के लिए आभार प्रकट करते हैं। लेखकगण, भा.कृ.सां.अ.सं. के निदेशक, डॉ. उमेश चन्दर सूद द्वारा उपलब्ध करायी गयी सुविधाओं, सहयोग एवं मार्गदर्शन के लिए आभार प्रकट करते हैं।

लेखकगण इस लेख के टंकण के लिए हिन्दी एकक के सहायक, श्री खुशियाल के आभारी हैं।

### संदर्भ

1. <https://www.icar.org.in>
2. <https://www.iasri.res.in>
3. <https://www.iasri.res.in/misfms/>
4. <https://mail.icar.gov.in>

□

## सहसंबंध एवं समाश्रयण

अशोक कुमार गुप्ता, अंकुर बिस्वास एवं विजय बिन्दल

वर्णनात्मक सांख्यिकी में, हमने केवल एक चर से संबंधित समस्याओं का अध्ययन किया है। अभ्यास में, हमे दो या दो से अधिक चरों के प्रयोग से जुड़ी बहुत सी समस्याओं का सामना करना पड़ता है। यदि दो चरों की स्थिति में, एक चर में आने वाले बदलाव दूसरे चर में आने वाले बदलाव से संबंधित होते हैं तो ये दोनों चर आपस में सह-संबंधित कहलाते हैं। उदाहरणार्थ, व्यक्तियों की ऊंचाई एवं भार के बीच, पति एवं पत्नी की आयु के बीच तथा वर्षा एवं चावल के उत्पादन के बीच कुछ संबंध होता है। अतः सहसंबंध एक सांख्यिकीय उपकरण है जो दो या दो से अधिक चरों के बीच होने वाले सहसंबंध की डिग्री ज्ञात करने में हमारी मदद करता है। वह मापक जिनकी सहायता से हम दो या दो से अधिक चरों के बीच संबंध की डिग्री का विश्लेषण करते हैं, सहसंबंध विश्लेषण कहलाते हैं तथा जो हमें सहसंबंध की दिशा एवं डिग्री ज्ञात कराने में सक्षम बनाते हैं उन्हें सहसंबंध गुणांक के रूप में जाना जाता है। ध्यान रहे कि दो चरों के बीच सहसंबंध विश्लेषण में किसी प्रकार के संबंध की आवश्यकता है जो प्रेक्षणों (observations) को युग्मों में (in pairs) साहचर्य प्रदान करता है।

व्यवहारिक जीवन में सहसंबंध अध्ययन अत्युपयोगी है क्योंकि अधिकतर चरों के बीच कुछ न कुछ संबंध होता ही है। एक बार यह ज्ञात हो जाए कि दो चर निकट रूप से (closely) संबंधित हैं तो हम समाश्रयण

विश्लेषण की सहायता से दूसरे चर का मान ज्ञात होने पर पहले चर के मान का आकलन कर सकते हैं। सहसंबंध विश्लेषण आर्थिकी व्यवहार को समझने के लिए योगदान देता है और उन पद्धतियों का सुझाव देता है जिनके द्वारा स्थिर (stabilizing) बल (forces) प्रभावी हो सकते हैं।

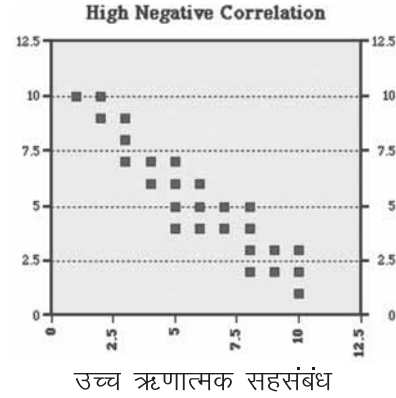
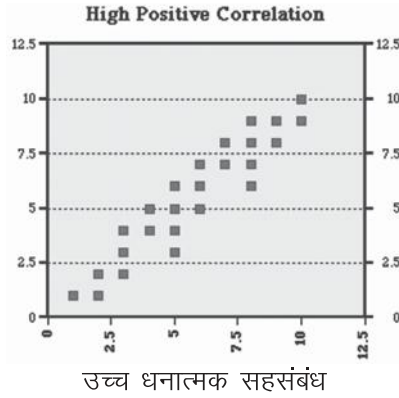
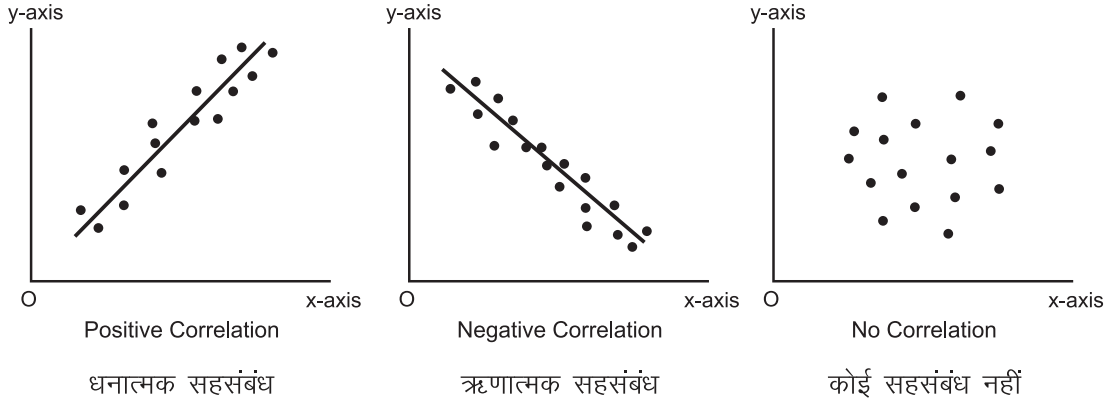
### सहसंबंध के प्रकार

सहसंबंध को कई विभिन्न तरीकों में वर्गीकृत किया जाता है लेकिन अति महत्वपूर्ण तरीके जिनके द्वारा सहसंबंध को वर्गीकृत किया जाता है, निम्न हैं :

### धनात्मक, ऋणात्मक एवं कोई सहसंबंध नहीं

यदि दोनों चर एक ही दिशा में गतिमान (varying) हैं अर्थात् यदि एक चर में वृद्धि के साथ दूसरे चर में भी वृद्धि हो रही है या एक चर में ह्रास होने की स्थिति में (decreasing) दूसरे चर में भी ह्रास हो रहा है तो सहसंबंध धनात्मक कहलाता है। सहसंबंध को ऋणात्मक कहा जाता है यदि दोनों चर विपरीत दिशाओं में गतिमान हैं अर्थात् यदि एक चर में वृद्धि के साथ दूसरे चर में ह्रास हो रहा है या एक चर में ह्रास के साथ दूसरे चर में वृद्धि हो रही है। यदि दो चरों के बीच कोई सहसंबंध नहीं है तो आलेखित (plotted) आलेख (graph) व्यापक रूप से छितरा (scattered) हुआ होगा।





- माना दो चरों के बीच सहसंबंध गुणांक को  $r$  द्वारा दर्शाया गया है। यदि  $x$  एवं  $y$  में प्रबल (strong) धनात्मक रैखिक सहसंबंध है तो  $r$  का मान  $+1$  के निकट होगा।  $r$  का पूर्णतः  $+1$  मान आदर्श/संपूर्ण धनात्मक फिट (perfect positive fit) को इंगित करता है।  $x$  एवं  $y$  चरों के बीच के संबंध के धनात्मक मान यह इंगित करते हैं कि यदि  $x$  के मान में वृद्धि होती है तो  $y$  के मान में भी वृद्धि होती है।
- यदि  $x$  और  $y$  में प्रबल ऋणात्मक रैखिक सहसंबंध है तो  $r$  का मान  $-1$  के निकट होगा।  $r$  का पूर्णतः  $-1$  मान एक आदर्श/संपूर्ण ऋणात्मक फिट (perfect negative fit) को इंगित करता है।  $x$  और  $y$  चरों के बीच के संबंध का ऋणात्मक मान यह इंगित करते हैं कि यदि  $x$  के मान में वृद्धि होती है तो  $y$  के मान में ह्रास (decrease) होता है।

- यदि दो चरों के बीच कोई रैखिक सहसंबंध नहीं है या बहुत न्यून रैखिक सहसंबंध है तो  $r$  का मान शून्य (0) के निकट होगा। शून्य के निकट मान का अभिप्राय है कि दो चरों के बीच सहसंबंध नहीं है।

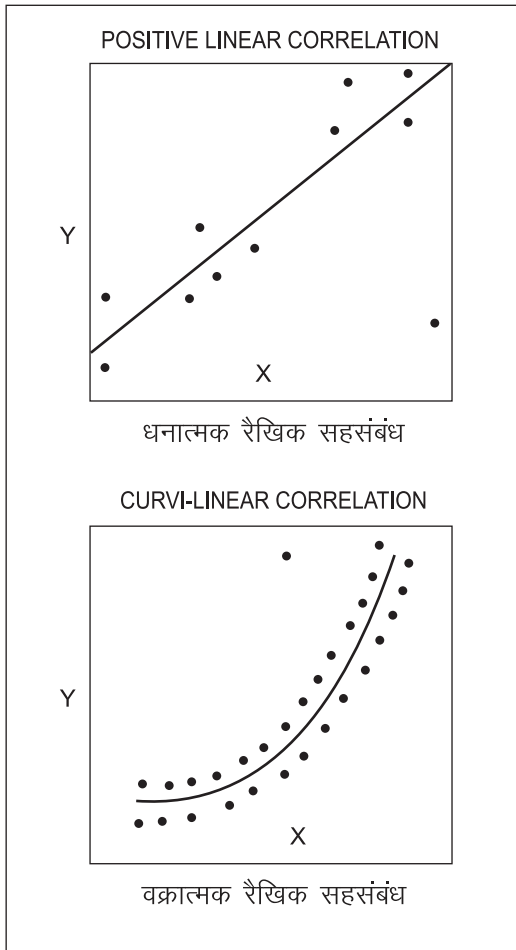
### सरल (simple), आंशिक (partial) एवं बहु (multiple) सहसंबंध

दो चरों के बीच होने वाला अध्ययन सरल सहसंबंध के अंतर्गत आता है। बहु सहसंबंध में, तीन या तीन से अधिक चरों का एक साथ अध्ययन किया जाता है। उदाहरणार्थ, यदि हम चावल की प्रति हेक्टेयर उपज, वर्षा की मात्रा एवं उर्वरक की मात्रा की बीच सहसंबंध का अध्ययन करते हैं तो यह बहु सहसंबंध के अंतर्गत होगा। आंशिक सहसंबंध में, दो से अधिक चरों का अध्ययन इस प्रकार किया जाता है कि केवल दो चर जो एक दूसरे को एक ही समय में प्रभावित कर रहे हैं, पर ध्यान केन्द्रित किया जाता है जबकि अन्य

प्रभावित चरों को स्थिरांक के रूप में माना जाता है। उदाहरणार्थ, यदि हम एक निश्चित औसत दैनिक तापमान के समय में होने वाली धान की प्रति हेक्टेयर उपज एवं वर्षा के बीच सहसंबंध विश्लेषण करना चाहते हैं तो यह आंशिक सहसंबंध के अंतर्गत आता है।

### रैखिक एवं अरैखिक (Linear and Non-linear) सहसंबंध

दो चरों के बीच सहसंबंध को रैखिक कहा जाता है यदि एक चर में होने वाले परिवर्तन की मात्रा एवं अन्य चर में होने वाले परिवर्तन की मात्रा के बीच स्थिर अनुपात (constant ratio) होता है। सहसंबंध को अरैखिक कहा जाता है यदि एक चर में होने वाले परिवर्तन की मात्रा एवं अन्य चर में होने वाले परिवर्तन की मात्रा के बीच स्थिर अनुपात नहीं होता है।



यदि सभी डाटा बिन्दु एक सीधी रेखा पर स्थित होते हैं तो दोनों चरों के बीच  $\pm 1$  का आदर्श/संपूर्ण (perfect) सहसंबंध होता है। यदि  $r = +1$  तो इस रेखा का ढाल (slope) धनात्मक होता है और यदि  $r = -1$  तो इस रेखा का ढाल ऋणात्मक होता है। आम तौर पर 0.8 से अधिक सहसंबंध को प्रबल (strong) सहसंबंध माना जाता है जबकि 0.5 के कम के सहसंबंध को प्रभावहीन (weak) सहसंबंध माना जाता है। ये मान अध्ययनाधीन डाटा के 'प्रकार' पर आधारित होते हैं। सामाजिक विज्ञान की तुलना में वैज्ञानिक डाटा पर आधारित अध्ययन के लिए एक प्रबल सहसंबंध की आवश्यकता होती है।

### सहसंबंध गुणांक

वह माप जिससे दो चरों की गतिविधियों के संबंध की डिग्री ज्ञात होती है, सहसंबंध गुणांक कहलाता है। सहसंबंध गुणांक की निम्न रूप से संगणना की जाती है:

$$r_{xy} = \frac{\text{Cov}(r_x, r_y)}{\sigma_x \sigma_y}$$

सहसंबंध गुणांक ( $r$ ) एक परिमाण-रहित (dimensionless) मात्रा है अर्थात् यह नियोजित इकाई पर निर्भर नहीं करता है।

सहसंबंध गुणांक दो चरों के बीच रैखिक संबंध की शक्ति (strength) का माप करता है। सहसंबंध गुणांक का मान -1 और +1 के बीच होता है। जो बिन्दु धनात्मक ढाल वाली एक सीधी सरल रेखा पर स्थित होते हैं उनका सहसंबंध +1 होता है तथा जो बिन्दु ऋणात्मक ढाल वाली सरल रेखा पर स्थित होते हैं उनका सहसंबंध -1 होता है। जिन बिन्दुओं का रैखिक संबंध नहीं होता है उनका सहसंबंध शून्य (0) होता है।

### कार्ल पियर्सन सहसंबंध गुणांक

रैखिक सहसंबंध गुणांक को विकसित करने वाले कार्ल पियर्सन के सम्मान में इस गुणांक को पियर्सन प्रोडक्ट मोमेंट सहसंबंध के रूप में संदर्भित किया जाता है। इसे 'r' से चिन्हित किया जाता है। 'r' को रैखिक सहसंबंध गुणांक कहा जाता है और यह दो चरों के

बीच एक रैखिक सहसंबंध की शक्ति (strength) एवं दिशा (direction) का माप करता है।

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

$$r = \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{n(\sum x^2) - (\sum x)^2} \sqrt{n(\sum y^2) - (\sum y)^2}}$$

जहां n डाटा के युग्मों (pairs) की संख्या है।

उपरोक्त सूत्र में सह-प्रसरण (covariance) एवं प्रसरण (variance) के प्रतिदर्श आधारित आकलों का प्रतिस्थापन करके हम r का मान प्राप्त कर सकते हैं।

### पियर्सन सहसंबंध गुणांक के लक्षण (properties)

- r का मान मापक की इकाई पर निर्भर नहीं करता है। उदाहरणार्थ, यदि किसी व्यक्ति के भार की इकाई x पाउंड में है तथा लंबाई y की इकाई इंच में है तो x एवं y के बीच सहसंबंध उस सहसंबंध के समान ही होगा जबकि x की इकाई किलोग्राम में और y की इकाई सेंटीमीटर में हो।
- r का मान चरों के लेबल पर निर्भर नहीं करता कि कौन से चर को x से चिन्हित किया गया है और कौन से चर को y से चिन्हित किया गया है अर्थात् x एवं y के बीच सहसंबंध = y एवं x के बीच सहसंबंध।
- $-1 \leq r \leq +1$ , r का धनात्मक मान, धनात्मक

रैखिक संबंध को दर्शाता है और ऋणात्मक मान, ऋणात्मक रैखिक संबंध को दर्शाता है।

- $r = +1$  एक आदर्श/संपूर्ण धनात्मक रैखिक संबंध को तथा  $r = -1$  एक आदर्श/संपूर्ण ऋणात्मक रैखिक संबंध को दर्शाता है। यह अत्यंत दुर्लभ है।
- r का निकटतम शून्य (0) मान का तात्पर्य है कि दो चरों के बीच कोई संबंध नहीं है। ध्यान रहे कि 'कोई रैखिक संबंध नहीं' का अर्थ यह नहीं है कि 'कोई संबंध मौजूद नहीं' है। जब r का निकटतम मान शून्य (0) हो तो भी एक उच्च स्तर का 'अरैखिक' संबंध मौजूद हो सकता है। हमेशा याद रहे कि r सिर्फ x एवं y के बीच रैखिक संबंध का माप ही करता है।

### सहसंबंध की शक्ति (strength)

हम सहसंबंध की शक्ति को निम्न रूप में परिभाषित कर सकते हैं

प्रबल (strong)  $|r| \geq 0.75$

संयत (moderate)  $0.5 \leq |r| < 0.75$

प्रभावहीन (weak)  $|r| < 0.5$

**नोट:** सहसंबंध सिर्फ एक रैखिक संबंध का माप करता है, इसका अभिप्राय यह नहीं है कि r का निकटतम शून्य मान x एवं y के बीच 'कोई संबंध नहीं है' को दर्शाता है।

**उदाहरण:** निम्न आंकड़ों से कार्ल पियर्सन सहसंबंध गुणांक प्राप्त करें :

(x)	68	71	62	75	58	60	67	68	71	69	68	67	63	62	60	63	65	67	63	61
(y)	4.1	4.6	3.8	4.4	3.2	3.1	3.8	4.1	4.3	3.7	3.5	3.2	3.7	3.3	3.4	4.0	4.1	3.8	3.4	3.6

व्याख्या :

(x)	(y)	x*y	x*x	y*y
68	4.1	278.8	4624	16.81
71	4.6	326.6	5041	21.16
62	3.8	235.6	3844	14.44
75	4.4	330.0	5625	19.36
58	3.2	185.6	3364	10.24

60	3.1	186.0	3600	09.61
67	3.8	254.6	4489	14.44
68	4.1	278.8	4624	16.81
71	4.3	305.3	5041	18.49
69	3.7	255.3	4761	13.69
68	3.5	238.0	4624	12.25
67	3.2	214.4	4489	10.24
63	3.7	233.1	3969	13.69
62	3.3	204.6	3844	10.89
60	3.4	204.0	3600	11.56
63	4.0	252.0	3969	16.00
65	4.1	266.5	4225	16.81
67	3.8	254.6	4489	14.44
63	3.4	214.2	3969	11.56
61	3.6	219.6	3721	12.96
<b>1308</b>	<b>75.1</b>	<b>4937.6</b>	<b>85912</b>	<b>285.45</b>

$$\begin{aligned}
 N &= 20 \\
 \Sigma xy &= 4937.6 \\
 \Sigma x &= 1308 \\
 \Sigma y &= 75.1 \\
 \Sigma x^2 &= 85912 \\
 \Sigma y^2 &= 285.45
 \end{aligned}$$

$$\begin{aligned}
 r &= \frac{20(4937.6) - (1308)(75.1)}{\sqrt{[20(85912) - (1308^2)] [20(285.45) - (75.1^2)]}} \\
 r &= \frac{98752 - 98230.8}{\sqrt{[1718240 - 1710864] [5709 - 5640.01]}} \\
 r &= \frac{521.2}{\sqrt{[7376] [68.99]}} \\
 r &= \frac{521.2}{\sqrt{508870.2}} \\
 r &= \frac{521.2}{713.3514} \\
 r &= 0.73
 \end{aligned}$$

अतः सहसंबंध गुणांक 0.73 है जो कि एक काफी प्रबल धनात्मक सहसंबंध को दर्शाता है।

### सहसंबंध गुणांक की सार्थकता (significance) परीक्षण (testing)

सहसंबंध गुणांक परिकलित करने के उपरान्त हम इस सम्भावना को निर्धारित कर सकते हैं कि क्या

प्रेक्षित (observed) सहसंबंध संयोग से ही प्राप्त हो गया है। इसके लिए हमें सार्थकता परीक्षण का प्रयोग करना पड़ेगा। अर्थात् सहसंबंध गुणांक की वास्तविकता (reality) की सम्भावना के लिए सार्थकता परीक्षण का प्रयोग करना होगा। इसके लिए परिकल्पना परीक्षण (testing of hypothesis) का प्रयोग करना होगा।

निराकरणीय या शुन्य परिकल्पना (Null Hypothesis)	$r = 0$
वैकल्पिक परिकल्पना (Alternative Hypothesis)	$r < > 0$

सभी परिकल्पना परीक्षणों में हम पहले सार्थकता स्तर निर्धारित करते हैं। सामान्यतः हम सार्थकता स्तर  $\alpha = 0.05$  लेते हैं। इसका अभिप्राय है कि परीक्षण से प्राप्त सहसंबंध गुणांक एक संयोग-मात्र है एवं इसकी सम्भावना 100 में से 5 से ज्यादा नहीं है। इस परिकल्पना परीक्षण का आसान तरीका सांख्यिकीय पुस्तकों में उपलब्ध  $r$  के क्रांतिक मानों (critical values) की तालिका है। तालिका से  $r$  के क्रांतिक मान देखने से पहले हमें स्वातंत्र्य कोटि (degrees of freedom - df) की गणना करनी आवश्यक है। स्वातंत्र्य

कोटि का मान  $(N - 2)$  के बराबर होता है। उपरोक्त उदाहरण में स्वातंत्र्य कोटि का मान  $(20 - 2 = 18)$  है। अंत में यह निश्चित करना है कि हमें वन-टेल परीक्षण करना है या टू-टेल। उपरोक्त उदाहरण में कोई दृढ़ पूर्व सिद्धान्त उपलब्ध नहीं है कि  $x$  एवं  $y$  के मध्य का सहसंबंध धनात्मक है या ऋणात्मक। इसके लिए हमें टू-टेल परीक्षण करना होगा। इन तीन सूचनाओं के साथ अर्थात् सार्थकता स्तर  $\alpha = 0.05$ , स्वातंत्र्य कोटि  $df = 18$  तथा परीक्षण का प्रकार = टू-टेल, की सहायता से हम सहसंबंध गुणांक का सार्थकता परीक्षण कर सकते हैं। पुस्तक से प्राप्त  $r$  का क्रांतिक मान = 0.4438। इसका अभिप्राय है कि यदि सहसंबंध गुणांक 0.4438 से अधिक है या कम है (याद रहे कि यह टू-टेल परीक्षण है) तो निश्कर्ष निकाल सकते हैं कि सहसंबंध गुणांक के संयोग-मात्र से आने की असमानता (odds) की सम्भावना 100 में से 5 से अधिक नहीं है परन्तु सहसंबंध गुणांक का आकलित मान 0.73 है जो कि काफी अधिक है। अतः यह निश्कर्ष निकलता है कि यह संयोग-मात्र नहीं है और सहसंबंध गुणांक सांख्यिकीय रूप से सार्थक है। अतः शुन्य परिकल्पना (Null Hypothesis) को अस्वीकार किया जा सकता है तथा वैकल्पिक परिकल्पना (Alternative Hypothesis) को स्वीकार कर सकते हैं।

### संभावित त्रुटि (Probable Error)

संभावित त्रुटि एक सीमा है जिसमें कार्ल पियर्सन सहसंबंध गुणांक का मान  $r$  के बीच स्थित है जिसकी प्रायिकता (probability) ज्ञात होती है। सहसंबंध गुणांक की संभावित त्रुटि इसके मान की व्याख्या करने में सहायता करती है। संभावित त्रुटि की सहायता से सहसंबंध गुणांक के मान की विश्वसनीयता को निर्धारित किया जा सकता है, जब तक यह यादृच्छिक प्रतिचयन के नियमों पर निर्भर है। संभावित त्रुटि निम्न प्रकार से प्राप्त की जा सकती है

$$\text{संभावित त्रुटि (P.E.)} = .6745 (1-r^2)/\sqrt{n}$$

1. सहसंबंध गुणांक की सीमा  $r \pm \text{P.E.}$  होनी चाहिए।
2. यदि  $r$  का मान संभावित त्रुटि से कम है तो सहसंबंध के कोई प्रमाण (evidence) नहीं है।

3. यदि  $r$  का मान संभाव्य त्रुटि के छः गुना से ज्यादा है तो सहसंबंध सार्थक (significant) है।
4. यदि संभाव्य त्रुटि ज्यादा नहीं है और यदि सहसंबंध गुणांक 0.5 या उससे ज्यादा है तो आमतौर पर सहसंबंध को सार्थक माना जाता है।
5. सहसंबंध गुणांक की व्याख्या के लिए संभाव्य त्रुटि का प्रयोग मापक के रूप में तब करना चाहिए जब प्रेक्षण युग्मों की संख्या अधिक हो। यदि  $n$  का मान छोटा होगा तो संभावित त्रुटि भ्रामक निष्कर्ष दे सकती है।

### स्पीयरमैन श्रेणी (Rank) सहसंबंध गुणांक

स्पीयरमैन श्रेणी सहसंबंध गुणांक या स्पीयरमैन रो (Rho), चार्ल्स स्पीयरमैन के नाम पर रखा गया है और इसे ग्रीक अक्षर  $\rho$  (rho) या  $r_s$  द्वारा दर्शाया जाता है। यह दो चरों के बीच सांख्यिकीय आश्रितता (dependence) का अप्राचलिक (non-parameteric) मापक है। यह निर्धारित करता है कि मोनोटोनिक फलन का प्रयोग करके दो चरों के बीच सहसंबंध कितनी अच्छी तरह वर्णित किये जा सकते हैं। यदि कोई पुनरावृत्त आँकड़ें नहीं हैं तो संपूर्ण (perfect) स्पीयरमैन सहसंबंध +1 या -1 होगा जब तक प्रत्येक चर, प्रत्येक दूसरे चर का एक संपूर्ण मोनोटोन फलन है।

स्पीयरमैन सहसंबंध गुणांक को श्रेणीबद्ध (ranked) चरों के बीच स्पीयरमैन सहसंबंध गुणांक के रूप में परिभाषित किया गया है। प्रतिदर्श आकार  $n$  के लिए  $n$  रॉ स्कोर (raw scores)  $x_i$  एवं  $y_i$  को श्रेणी (rank)  $x_i$  एवं श्रेणी (rank)  $y_i$  में परिवर्तित करके  $\rho$  का मान निम्न सूत्र से प्राप्त किया जाता है।

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

### रैखिक समाश्रयण (Linear Regression)

वक्र फिटिंग का मुख्य उद्देश्य एक चर (आश्रित चर) का दूसरे चर (चरों) (स्वतंत्र चरों) से आकलन प्राप्त करना है। आकलन की इस पद्धति को समाश्रयण द्वारा संदर्भित किया गया है। यदि  $Y$  को  $X$  से

किसी समीकरण द्वारा आकलित करना है तो इसको X पर Y का समाश्रयण समीकरण कहेंगे तथा तद्रूपी (corresponding) वक्र को X पर Y का समाश्रयण वक्र कहेंगे। समाश्रयण रेखा वह रेखा है जो अन्य चर के किसी विशिष्ट मान के लिए एक चर के मान का सर्वश्रेष्ठ आकल प्रदान करती है। अतः समाश्रयण रेखा "सर्वोत्तम फिट" रेखा है और इसे न्यूनतम वर्ग सिद्धान्त द्वारा प्राप्त किया गया है।

### न्यूनतम वर्ग विधि

सन्निकटन (approximation) वक्र के बारे में निर्णय (judgement) की संभावित त्रुटि से बचने के लिए एक सर्वोत्तम फिटिंग रेखा, एक सर्वोत्तम फिटिंग परवलय (parabola) इत्यादि की परिभाषा पर सहमत होना आवश्यक है। परिभाषा के लिए  $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$  आँकड़ा बिन्दुओं पर विचार करें। माना X के ज्ञात मान  $X_1$  के लिए  $Y_1$  तथा तद्रूपी (corresponding) Y के मानों के बीच का अन्तर  $d_1$  वक्र से प्राप्त होता है। इसे विचलन त्रुटि (deviation error) या अवशिष्ट (residual) के रूप में जाना जाता है और ये धनात्मक, ऋणात्मक या शून्य हो सकती है। इसी प्रकार के विचलन  $d_2, d_3, \dots, d_n$  तद्रूपी (corresponding)  $X_2, X_3, \dots, X_n$  मानों के लिए प्राप्त किये जा सकते हैं।

वक्र C के समंजन-सुष्ठुता (goodness of fit) के मापक को  $\sum d_i^2 = d_1^2 + d_2^2 + \dots + d_n^2$  के द्वारा प्राप्त किया जा सकता है। यदि  $\sum d_i^2 = d_1^2 + d_2^2 + \dots + d_n^2$  का मान न्यूनतम है तो फिट अच्छा है अन्यथा फिट खराब है। अतः यह नियम स्वीकार्य है कि सर्वोत्तम फिटिंग वक्र वह है जिसके लिए  $\sum d_i^2 = d_1^2 + d_2^2 + \dots + d_n^2$  न्यूनतम है। एक वक्र जिसमें यह गुण होते हैं उसे आँकड़ों के लिए न्यूनतम वर्ग फिटिड माना जाता है और उसे न्यूनतम वर्ग वक्र कहा जाता है। विशेष रूप से, इस गुण वाली सरल रेखा को न्यूनतम वर्ग रेखा (least square line) कहा जाता है।

यदि Y के स्थान पर X एक आश्रित चर है तो उर्ध्वाधर (vertical) विचलन के स्थान पर क्षैतिज (horizontal) विचलन लेते हुए इस परिभाषा को

संशोधित किया जाता है जो समन्वित अक्षों (coordinate axes) को परस्पर बदल कर प्राप्त किया जाता है। इन दोनों परिभाषाओं से सामान्य रूप से दो अलग-अलग न्यूनतम वर्ग वक्र प्राप्त होते हैं।

### सरल रेखा समाश्रयण का निर्धारण

यह समीकरण (1) में न्यूनतम वर्ग पद्धति द्वारा स्थिरांकों (constants) a एवं b का निर्धारण करके किया जाता है

$$F(a,b) = \sum d_i^2 = \sum (Y_i - a - bX_i)^2 \quad (1)$$

$$\text{न्यूनतम है तथा } \frac{\partial F}{\partial a} = 0 \text{ एवं } \frac{\partial F}{\partial b} = 0$$

इससे निम्न प्रसामान्य (normal) समीकरण प्राप्त होते हैं;

$$na + b \sum X = \sum Y$$

$$a \sum X + b \sum X^2 = \sum XY$$

उपरोक्त समीकरणों को हल करके a एवं b का मान ज्ञात किया जा सकता है;

$$\begin{aligned} b &= \frac{\sum (X - \bar{X})(Y - \bar{Y})}{\sum (X - \bar{X})^2} \\ &= \left[ \sum XY - \frac{(\sum X)(\sum Y)}{n} \right] / \left[ \sum X^2 - \frac{(\sum X)^2}{n} \right] \quad (2) \\ &= r \frac{\sigma_Y}{\sigma_X} \end{aligned}$$

$$a = \bar{Y} - b\bar{X}$$

जहां  $\bar{X} = \frac{1}{n} \sum X$ ,  $\bar{Y} = \frac{1}{n} \sum Y$  क्रमशः X एवं Y के माध्य हैं और r कार्ल पियर्सन सहसंबंध गुणांक है।

अतः X पर Y की वांछित समाश्रयण रेखा निम्न होगी

$$\begin{aligned} Y &= (\bar{Y} - b\bar{X}) + bX \\ &= \bar{Y} + b(X - \bar{X}) \\ &= \bar{Y} + r \frac{\sigma_Y}{\sigma_X} (X - \bar{X}) \quad (3) \end{aligned}$$

समाश्रयण रेखा (3) स्पष्ट रूप से  $(\bar{X}, \bar{Y})$  से निकलती है और स्थिरांक b समाश्रयण रेखा (3) का

ढाल (slope) है। गुणांक  $b$  से  $X$  में इकाई परिवर्तन के लिए  $Y$  में संभावित परिवर्तन के बारे में ज्ञात होता है।  $X$  पर  $Y$  की समाश्रयण रेखा को वैकल्पिक रूप से निम्न प्रकार से लिखा जा सकता है

$$Y = \alpha + bX \quad (4)$$

जहां  $\alpha = \bar{Y} - b\bar{X}$  एवं गुणांक  $b$  का मान उपरोक्त विधि से ज्ञात किया जा सकता है। इसी प्रकार  $Y$  पर  $X$  की समाश्रयण रेखा  $X = \bar{X} + r \frac{\sigma_X}{\sigma_Y} (Y - \bar{Y})$  होगी।

### समाश्रयण गुणांक

$X$  पर  $Y$  की समाश्रयण रेखा के ढाल (slope) “ $b$ ” को  $X$  पर  $Y$  का समाश्रयण गुणांक कहते हैं। यह स्वतन्त्र चर  $X$  के मान में इकाई परिवर्तन से होने वाले आश्रित चर  $Y$  के मान में परिवर्तन का निरूपण (depicts) करता है।  $X$  पर  $Y$  का समाश्रयण गुणांक निम्न समीकरण से ज्ञात होता है;

$$b_{YX} = r \frac{\sigma_X}{\sigma_Y}$$

इसी प्रकार  $Y$  पर  $X$  का समाश्रयण गुणांक निम्न होगा

$$b_{XY} = r \frac{\sigma_Y}{\sigma_X}$$

### समाश्रयण गुणांक के गुण (Properties)

- सहसंबंध गुणांक, समाश्रयण गुणांकों के मध्य गुणोत्तर माध्य अर्थात्  $r = \pm \sqrt{b_{XY} \times b_{YX}}$  होता है।
- यदि एक समाश्रयण गुणांक 1 से अधिक है तो दूसरा समाश्रयण गुणांक 1 से कम होगा।
- समाश्रयण गुणांकों के समान्तर माध्य (arithmetic mean) के मान का मापांक (modulus) सहसंबंध गुणांक  $r$  के मान के मापांक से कम नहीं हो सकता है अर्थात्  $|b_{XY} + b_{YX}| > |r|$
- समाश्रयण गुणांक मूल बिन्दु (origin) में होने वाले परिवर्तन से स्वतन्त्र है परन्तु पैमाने (scale) से नहीं।

### निरूपण गुणांक (Coefficient of Determination)

एक प्रतिदर्शज (statistic) जिसका व्यापक रूप

से प्रयोग करके यह निरूपण किया जाता है कि समाश्रयण रेखा ( $X, Y$ ) आंकड़ा युग्मों के समुच्चय को कितनी अच्छी प्रकार से फिट (fits) करती है उसे निरूपण गुणांक या बहु सहसंबंध गुणांक ( $R^2$ ) कहा जाता है। निरूपण गुणांक  $Y$  में होने वाले प्रतिशत परिवर्तनशीलता (variability) का प्रतिनिधित्व करता है जिसकी व्याख्या  $X$  एवं  $Y$  के बीच होने वाले रैखिक संबंध द्वारा की जा सकती है। दूसरे शब्दों में  $R^2$  से यह स्पष्ट होता है कि  $Y$  में होने वाली परिवर्तनशीलता की कितनी व्याख्या की जा सकती है जबकि यह ज्ञात है कि वह  $X$  से सह-संबंधित है (अर्थात् बिन्दु न्यूनतम वर्ग समाश्रयण रेखा के कितने करीब है)।

सरल रैखिक समाश्रयण में निरूपण गुणांक, पियर्सन सहसंबंध गुणांक के वर्ग के बराबर होता है अर्थात्

$$R^2 = (r)^2$$

निम्न तालिका के आंकड़ों को देखें एवं  $X$  पर  $Y$  की समाश्रयण रेखा और  $Y$  पर  $X$  की समाश्रयण रेखा, समाश्रयण गुणांक तथा  $r^2$  का मान ज्ञात करें।

X	1	3	4	6	8	9	11	14
Y	1	2	4	4	5	7	8	9

$$\sum X^2 = 1 + 9 + 16 + \dots + 156 = 524$$

$$\sum Y^2 = 1 + 4 + 16 + \dots + 81 = 256$$

$$\sum XY = (1 \times 1) + (3 \times 2) + \dots + (14 \times 9) = 364$$

$$\sum X = 56, \quad \sum Y = 40, \quad n = 8, \quad \bar{X} = 7, \quad \bar{Y} = 5$$

$$b = (\sum XY - n\bar{X}\bar{Y}) / (\sum X^2 - n\bar{X}^2)$$

$$= \frac{364 - 8 \times 7 \times 5}{(524 - 8 \times 49)} = \frac{84}{132} = \frac{7}{11}$$

$$a = 5 - \frac{7}{11} \times 7 = \frac{6}{11}$$

अतः  $X$  पर  $Y$  की समाश्रयण रेखा

$$Y = \frac{6}{11} + \frac{7}{11}X \quad \text{or} \quad Y - 5 = \frac{7}{11}(X - 7)$$

$$Y = 0.545 + 0.636 X$$

इसी प्रकार Y पर X की समाश्रयण रेखा

$$X = -0.50 + 1.5 Y$$

अब  $bb' = r^2 = 1.5 \times 0.636 = 0.954$  or  $r = 0.977$

सामान्यीकृत (Generalised) परिभाषा से

$$r^2 = \frac{b(\sum XY - n\bar{X}\bar{Y})}{(\sum Y^2 - n\bar{Y}^2)} = \left(\frac{7}{11} \times 84\right) / (256 - 200) = 0.954$$

इससे यह ज्ञात होता है कि स्वतंत्र चर से 95.4% के परिवर्तन होता है।

### संदर्भ

1. गुप्ता, एच सी, एवं कपूर, वी के (1994): "फन्डामेंटल ऑफ मैथमैटिकल स्टेटिस्टिक्स", नौवॉ संस्करण, सुल्तान चन्द ऐन्ड सन्स, नई दिल्ली।
2. गुप्ता, एस पी (2008): "स्टेटिस्टिकल मैथड्स", छत्तीसवॉ संस्करण, सुल्तान चन्द ऐन्ड सन्स, नई दिल्ली।

□

शिक्षक विद्यार्थियों की संस्कार रूपी जड़ों में खाद देते हैं और अपने श्रम से उन्हें सींचकर महाप्राण शक्तियाँ बनाते हैं।

—महर्षि अरविंद



## मछलियों के मौसमी विकास के लिए साइन वेब फंक्शन के साथ संशोधित गोम्पर्टज मॉडल

अमृत कुमार पाल, नाओरेम ओकेन्द्रो सिंह, रंजीत कुमार पॉल,  
बिशाल गुरुंग, वसी आलम एवं सत्यपाल सिंह

अत्याधिक दोलायमान या मौसमी वातावरण में मछली की वृद्धि वर्ष भर में समान दर से आगे नहीं बढ़ती है। तापमान और खाद्य आपूर्ति के लिये मुख्य रूप से उत्तार-चढ़ाव रहे हैं तो इसके रूप में मछली वृद्धि मौसमी दोलन को प्रदर्शित करती है। मछली अच्छी तरह से जीव विज्ञान के अग्रदूतों (फुलटन 901) के लिये जानी जाता थी। यद्यपि अधिकतर केंसों में मछली की वृद्धि का विकास वान-बर्टालानफी विकास मॉडल वीबीजी (VBG) के अनुसार माना जाता है। गैर मौसमी वृद्धि मॉडलों मत्स्यपालन में प्रजातियों के साथ अनुपयुक्त है इस धारणा के कारण जिसकी वृद्धि मौसमी है जोकि अत्यधिक समय के साथ वृद्धि अपरिवर्तनीय है। विकास मॉडल का एक मौसमी संस्करण का उपयोग करने की आवश्यकता पर कई लेखकों (स्पारे, 1990; पाली एट. एल.ए 1992) से विस्तार से चर्चा की गई है इसके बाद मौसमी मॉडल को विकसित किया गया। उदाहरण के लिये वीबीजी (VBG) मॉडल सोमर्स (1993) तथा होनिंग और हनुमारा (1993) के द्वारा विकसित वीबीजी (VBG) मॉडल के मौसमी परिवर्तन (पिचर तथा मैकडानाल्ड, 1973) मौसमी दोनलों के संस्करण को साइन-वेब फक्सन को जोड़कर समायोजन के द्वारा वीबीजी (VBG) मॉडल संशोधित किया गया। इसके अलावा सिंह एट एल.ए (2007) ने प्रेक्षित किया कि वान-बर्टालानफी मॉडल पन्तनगर उतराखण्ड में मोनोकल्चर तथा पालीकल्चर प्रणालियों को (माना 3 वर्ष से कम) टोर पुटीटोरा के विकास के प्रारम्भिक स्तरों की व्याख्या करने के लिये उपयुक्त पाया। टोर पुटीटोरा सबसे महत्वपूर्ण और लुप्तप्राय शीतजलीय

मछली प्रजातियों में से एक है। हालांकि, यह आवश्यक नहीं है कि वीबीजी (VBG) मॉडल का ही केवल संशोधित संस्करण वीबीजी (VBG) मॉडल के सभी प्रकार के मछली प्रजातियों के लिये या समान मछली प्रजातियों के लिये वीबीजी (VBG) मॉडल फिट होगा जब विभिन्न जलीय वातावरण मौसमी परिवर्तन या चक्रीय उतार-चढ़ाव शामिल है। इस पेपर में गोम्पर्टज विकास मॉडल एक संशोधित संस्करण प्रस्तावित है जो कि गोविन्दसागर जलाशय से टोर पुटीटोरा की डाटासेट लम्बाई- आयु पर सचित्र है।

### सामग्री और तरीके

मत्स्यपालन में सबसे अधिक इस्तेमाल किये जाने वाले विकास मॉडल निम्नलिखित है।

वान-वर्टैलनफी विकास मॉडल

$$L_t = L_\infty [1 - e^{-K(t-t_0)}], \quad (1)$$

गोम्पर्टज विकास मॉडल

$$L_t = L_\infty \exp[-\exp\{-K(t-t_0)\}], \quad (2)$$

जहां

$L_t$  t आयु पर मछली की लम्बाई;

$L_\infty$  अधिकतम मछली की लम्बाई;

K विकास गुणांक (प्रतिवर्ष);

t आयु (वर्षों में);

$T_0$  सैद्धान्तिक आयु (वर्षों में) जब मछली की लम्बाई शून्य थी

पिचर तथा मैक्डानाल्ड (1973) द्वारा पेश साइन वेव फंक्शन के साथ वान-वरटैलनफी मॉडल निम्नलिखित है।

$$L_t = L_\infty \left( 1 - e^{-\left[ C \sin\left(\frac{2\pi(t-S)}{52}\right) + K(t-t_0) \right]} \right), \quad (3)$$

जहां C एक गैर मौसमी या गैर अस्थिर विकास वक्र के आस पास विकास दौलनों की परिमाण का वर्णन करता है

जहां, S प्रारम्भिक बिन्दु (चरण से संबंधित है) और 52 सप्ताह में समय के पैमाने को इंगित करता है। P का प्रयोग, समीकरण 3 को 52 सप्ताह में निश्चित करने के बजाय यह मॉडल के स्कोप को विस्तृत करता है।

$$L_t = L_\infty \left( 1 - e^{-\left[ C \sin\left(\frac{2\pi(t-S)}{P}\right) + K(t-t_0) \right]} \right), \quad (4)$$

जहां C वृत्तीय आवर्तकाल है।

समान फैशन में, जब साइन वेव फक्सन समीकरण 2 के गोम्पर्टज मॉडल में जोड़ी जाती है निम्न प्रकार से हो जाती है।

$$L_t = L_\infty \exp\left[-\exp\left\{-\left(C \sin\left(\frac{2\pi(t-S)}{P}\right) + K(t-t_0)\right)\right\}\right] \quad (5)$$

### मॉडल फिटिंग

उपरोक्त समीकरण 1 से 5 गोविन्दसागर जलाशय से संग्रहित टोर पुटीटोरा के लिये आयु पर लम्बाई के प्रक्षणों को फिट किया गया (आगे के विवरण के लिये टन्डन तथा लोहल, 1983 को देखें)। उन्होंने 24.02.1981 पर गोविन्दसागर जलाशय से गिल नेट में 127.5 सेमी की कुल लम्बाई के आकार तथा 24.5 कि ग्रा वजन को पकड़ा तथा एक नमूना के आधार पर टोर पुटीटोरा के लिये आयु तथा विकास का अध्ययन किया। उन्होंने 1 से 17 आयु के वर्ष से जीवन के विभिन्न स्टेजो के लिये महसीर के पीछे से लम्बाई

गणना करके प्राप्त की। डाटा वर्तमान अध्ययन में आगे उपयोग किया गया है।

ऊपर विकास मॉडल पैरामीटर (एक) के एक फक्सन में विभिन्न मॉडलों कम से कम एक पैरामीटर के सम्बन्ध में व्युत्पन्न कम से कम एक के रूप में प्रकृति में अरैखिक है। विकास मॉडल का विवरण थोर्नली और फ्रांस (2007) के द्वारा दिया गया है।

एक अरैखिक प्रतिगमन मॉडल की अज्ञात मापदण्डों के आंकलन को प्राप्त करने को साहित्य में चार मुख्य विधियां (Seber and wild, 1989) उपलब्ध हैं। अर्थात् (i) गॉस-न्यूटन विधि (ii) स्टेपैस्ट डिसेन्ट विधि (iii) लेवेनवर्ग की मार्कक्वाडट तकनीकी तथा (iv) डू नॉट व्युत्पन्न (डी यू डी) विधि। लेवेनवर्ग की मार्कक्वाडट विधि अरैखिक न्यूनतम वर्ग आंकलन और वर्तमान अध्ययन में इस्तेमाल किया गया है कम्प्यूटिंग के लिये सबसे व्यापक रूप में इस्तेमाल किया गया और विश्वसनीय प्रक्रिया है।

मॉडल प्रदर्शन का परीक्षण करना, पूर्वानुमानित तथा प्रेक्षित चर कावेरी कैसे मापना समय में आवश्यक है। समय में covery कैसे का उपाय की जांच में लिये अरैखिक मॉडल के लिये, वर्ग माध्य मूल त्रुटि (RMSE) और माध्यनिरपेक्ष त्रुटि (एम ए ई) बहुधा आंकलन के लिये उपयोग किया जाता है।

पूर्णतया त्रुटि (MAE) आंकलन के लिये बहुधा प्रयोग की जाती है।

$$\text{वर्ग माध्य मूल RMSE} = \left[ \frac{1}{n} \sum_{t=1}^n (L_t - \hat{L}_t)^2 \right]^{1/2}$$

तथा

$$\text{MAE} = \frac{1}{n} \sum_{t=1}^n |(L_t - \hat{L}_t)|,$$

जहां

$\hat{L}_t$  आयु t पर पूर्वानुमान मछली की लम्बाई;

n पर्यवेक्षणों की संख्या, t = 1, 2, ..., n.

सबसे अच्छा मॉडल इन आंकडों के लिये सबसे कम संख्यात्मक मूल्यों का होगा। इसके अलावा, अवशेष के रेन्डमनेस कल्पना मॉडल की पर्याप्तता के

बारे में अग्रिम निर्णय लेने से पहले परीक्षण किया जाना चाहिये। उपरोक्त कार्य को प्राप्त करने के लिये, रन परीक्षण प्रक्रिया (राटकोवस्कीए 1990), उपलब्ध है। लेकिन सामान्य धारणा अरैखिक मॉडल के चयन के लिये अभावग्रस्त नहीं है क्योंकि उनके अवशेष सामान्य वितरण का पालन नहीं कर सकते।

एसपीएसएस (SPSS) 12.0 संस्करण पर अरैखिक समाश्रयण विकल्प का उपयोग करके मॉडल फिट किये गये। प्रारम्भिक पैरामीटर मान के विभिन्न अरैखिक मॉडल की सबसे अच्छी फिटिंग के लिये वेशविक अभिसरण कसौटी को पूरा करने की कोशिश कर रहे थे।

### परिणाम तथा चर्चा

गोम्पर्टज और वान-वर्टलानफी विकास मॉडल टोर पुटीटोरा के ऊपर आयु पर लम्बाई में डाटा के लिये फिट है।

पैरामीटर के आंकलन आरएमएसई, एमएई रन परीक्षण संख्यकी ( $|Z|$ ) मान तालिका 1 में दिये गये हैं।

गोम्पर्टज मॉडल VBG की तुलना में अच्छा प्रदर्शन करते हैं जब आरएमएसई तथा एमएई क्राइटेरिया उत्तम मॉडल की पहचान के लिये प्रयोग किया जाता है (तालिका-1)

अवशिष्ट के बारे में रेन्डमनैश धारणा सन्तुष्ट है क्योंकि रन परीक्षण ( $|Z|$ ) मान क्रिटिकल मान

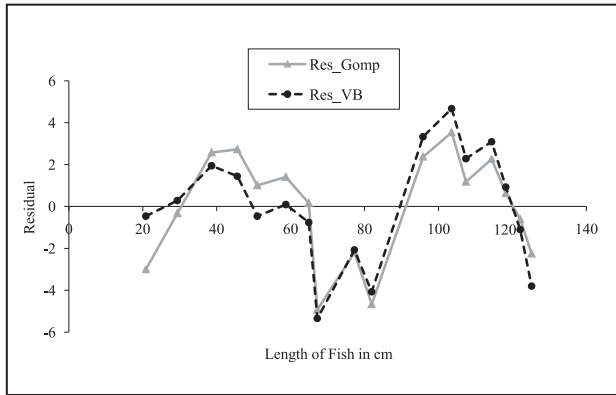
(सिगनीफिकेन्ट 5% के स्तर पर 1.96) के नीचे है। जब गोम्पर्टज तथा वीबीजी (VBG) मॉडल के अवशिष्ट अनुमानित लम्बाई के विपरीत फिट किया गया (चित्र 1), एक चक्रीय पैटर्न देखा। संभव समाधान के लिये साइन वेब फंक्शन मॉडल को शामिल करने का सुझाव दिया गया था। समीकरण (3) के वीबीजी (VBG) मॉडल साइन वेब फंक्शन से वर्तमान विचारणीय डाटा सेट के लिये आर्टीमल समाधान देने में असफल रहा। गोम्पर्टज तथा वीबीजी (VBG) मॉडल के संशोधित संस्करण उपरोक्त डाटा सेट तथा अनुमानित पैरामीटर के लिये पुनः फिट किये गये थे, तालिका-1 में दिया गया है। सरल गोम्पर्टज तथा वी बी जी (VBG) मॉडल की तुलना में आरएमएसई तथा एमएई के मान में काफी सुधार किया गया है। तथा संगत रन परीक्षण ( $|Z|$ ) मान 1.96 क्रिटिकल मान से भी कम है जिसका अर्थ रेन्डमनेस धारणा सन्तुष्ट है। गोम्पर्टज मॉडल का संशोधित संस्करण वर्तमान जलीय वातावरण के टोर पुटीटोरा के डाटा की व्याख्या करता है जोकि अन्य लोकप्रिय विकास मॉडल से बेहतर है, प्रेक्षित मानों के साथ चित्र 2 में दिखाया गया है। इसके अलावा, टोर पुटीटोरा की उपगामी लम्बाई, संशोधित गोम्पर्टज विकास मॉडल उपयोग करके लगभग 169 सेमी अनुमानित की गयी है। जोकि स्वीकृत है क्योंकि भारत में अधिकतम आकार 275 सेमी (झींग्राम, 1975) है तथा नेपाल में 180 सेमी (श्रेष्ठा, 1999)।

तालिका 1. मॉडलिंग फिटिंग के लिये सांराश आँकड़े

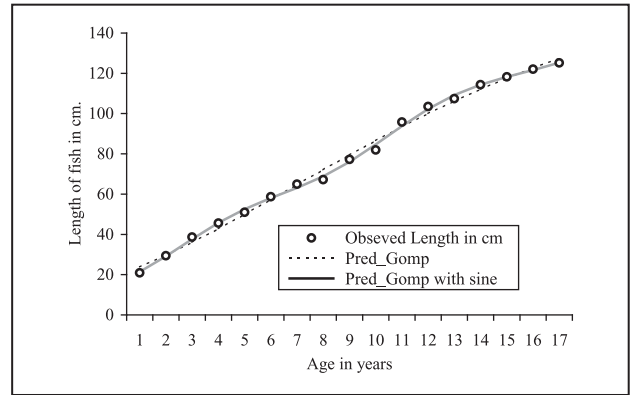
	गोम्पर्टज	साइन वेब फंक्शन गोम्पर्टज	वान-बर्टलानकी	सिने वेब फंक्शन के साथ वान-बर्टलानकी
<b>(1) पैरामीटर</b>				
$L_{\infty}$	172.928 (11.71)	169.444 (13.941)	423.500 (176.29)	304.893 (156.39)
K	0.117 (0.01)	0.121 (0.014)	0.019 (0.01)	0.030 (0.02)
$t_0$	6.849 (0.66)	6.701 (0.733)	-1.651 (0.45)	-1.032 (0.75)
S	-	1.902 (0.778)	-	0.443 (1.23)
C	-	0.059 (0.015)	-	0.016 (0.014)
P	-	8.992 (1.078)	-	10.169 (1.73)

	गोम्टर्ज	साइन वेब फंक्शन गोम्टर्ज	वान-बर्टलानकी	सिने वेब फंक्शन के साथ वान-बर्टलानकी
<b>(2) मॉडल पर्याप्तता</b>				
RMSE	2.516	1.290	2.665	1.449
MAE	2.108	1.014	2.126	1.187
<b>अवशिष्ट विश्लेषण</b>				
रन परीक्षण ( $ Z $ )	1.936	0.518	0.991	1.020
शापिरो-विल्क परीक्षण P-मूल्य	0.219	0.575	0.913	0.605

ब्रेकिट मान संगत उपगामी मानक त्रुटियाँ हैं।



चित्र 1. डाटासेट पर गोम्टर्ज तथा वान-बर्टलानकी मॉडल के फिटिंग के बाद शेष अवशिष्ट



चित्र 2. टोरपुटीटोरा की लम्बाई में प्रेक्षण तथा पूर्वानुमानित वृद्धि का चित्रमय प्रदर्शन

## आभार

प्रथम लेखक वास्तविक समर्थन और मार्गदर्शन के लिये डॉ. पी.सी. महंत, निदेशक, शीतजल मत्स्य अनुसंधान निदेशालय, भीमताल के आभारी हैं।

## संदर्भ

1. फुल्टन, टी.डब्ल्यू (1901): "कोड, हैडोक, वाहितिंग और नार्वे पौट की वृद्धि दर", स्कॉटलैंड के मत्स्य बोर्ड की 19 वीं वार्षिक रिपोर्ट, भाग III, 154-228।
2. होइनिंग, एन. एवं हनुमारा, आर. सी. (1990): "मौसमी विकास मॉडल की एक प्रायोगिक तुलना", फिसवाइट, 8(1), 32-34।

3. झींग्राम, वी.जी. (1975): "मछली और भारत के मत्स्यपालन", द्वितीय संस्करण, हिन्दुस्तान प्रकाशन निगम, दिल्ली।
4. पाल, डी. सोरयानो-वार्टज, मोर्यू, जे. एवं जारे टीकमान, ए. (1992): "मछलियों में वृद्धि की मौसमी समाप्ति के लिये एक नया मॉडल", लेखांकन/समुन्द्र और मीठे पानी में अनुसंधान का आस्ट्रेलियन जर्नल, 43, 1151-1156।
5. धड़ा, टी.जे. एवं मैकडोनाल्ड, पी.डी. एम (1973): "मौसमी वृद्धि के दो मॉडल पारिस्थितिकीय", एप्लाइड जर्नल, 10, 599-606।

6. राटकोसकी, संस्करण डिप्लामा और गैर रेखीय समाश्रयण मॉडल (1990): "हैंडबुक मार्सेल डेकर", न्यूयॉर्क।
7. सेबर, जी. ए. एफ. एवं वाइल्ड, सी.जे. (1989): "अरेखिक समाश्रयण मॉडल", जॉन विले ऐन्ड सन्स, न्यूयॉर्क।
8. श्रेष्ठा, जे. (1999): "शीतजल मछली और नेपाल में मछली पालन इन: मछली और मछली पालन अधिक ऊर्चाई", एशिया (स) टी. पेट्रू: एफ ए ओ, मत्स्य तकनीकी पेपर संख्या 385।
9. सिंह, एन. ओकेन्द्रो, पॉल, ए.के. एवं आलम, वसी, (2007): "मोनोकल्चर और पालीकल्चर सिस्टम के तहत टोरपुटीटोरा (हेमिल्टन) के विकास पैटर्न का करने के लिये अरेखिक मॉडल", भारतीय पशुविज्ञान जर्नल, 77(12), 1346–1348।
10. सोमर्स, आई एफ, (1988): "मौसम परिवर्तन पर विकास कार्य", फिसवाइट 6(1), 8–11।
11. स्पारे, पी. (1990): "क्या पारंपरिक मछली की लम्बाई का वजन स्टाक मूल्यांकन का उपयोग कर सकते जब विकास मौसमी है", फिशबाइट, 8, 29–32।
12. टंडन, के.के. एवं जोहल, एम. एस. (1983): "स्केल के द्वारा साक्ष्य के रूप में टोरपुटीटोरा (हेमिल्टन) के आयु तथा विकास पर अध्ययन", भारतीय मत्स्यपालन जर्नल, 30, 171–175।
13. थोर्नली, जे.एच.एम. एवं फ्रांस, जे. (2007): "पौधे, पशु और पर्यावरण विज्ञान के लिये कृषि मात्रात्मक तरीकों में गणितीय मॉडल", द्वितीय संस्करण, क्रामवेल प्रेस ट्रोब्रिज, कैम्ब्रिज संयुक्त राज्य अमेरिका, 136–163 पी.पी.।

□

हिंदी हमारे राष्ट्र की अभिव्यक्ति का स्रोत है।

—सुमित्रानंदन पंत

## विषम चक्रीय आँकड़ों के पूर्वानुमान के लिए फूरियर स्वसमाश्रयी (एफ-एआर) गुणांक अरैखिक काल-श्रृंखला प्रतिमान-एक व्यष्टि अध्ययन

हिमाद्रि घोष, सविता वधवा एवं प्रज्ञेषु

स्वसमाश्रयी (एआर) प्रतिमान आमतौर पर काल-श्रृंखला आँकड़ों के पूर्वानुमान के लिए प्रयोग किए जाते हैं और ये उस स्थिति में उत्कृष्ट परिणाम देते हैं जब चर का वर्तमान मूल्य अपने अतीत मूल्यों पर रैखिक रूप से निर्भर करता है (बॉक्स इत्यादि, 2009)। हांलाकि इस प्रतिमान के गुणांक स्थिरांक होना इसकी एक कमी है। लुडलो और एंडर्स (2000) ने एआर प्रतिमान क्रम 1 के लिए सामान्यीकरण का सुझाव दिया जिसमें स्वसमाश्रयी गुणांक काल निर्भर नित्यात्मक फलन  $\alpha_t$  हों। हांलाकि इसका फलन पूर्ण रूप से निर्दिष्ट नहीं है, परन्तु  $\alpha_t$  का व्यवहार आमतौर पर एक पर्याप्त लंबी फूरियर श्रृंखला से आंका जा सकता है (ब्लूमफील्ड, 2000)। उदाहरण के लिए, यदि  $\alpha_t$  सटीकता के किसी स्तर पर पूर्णएकीकृत फलन है तो इसे फूरियर-स्वसमाश्रयी (एफ-एआर) प्रतिमान के रूप में निम्न प्रकार से लिखा जा सकता है:

$$y_t = \alpha_t y_{t-1} + \varepsilon_t, \quad \alpha_t = A_0 + \sum_{k=1}^{T/4} \left[ A_k \sin \frac{2\pi kt}{T} + B_k \cos \frac{2\pi kt}{T} \right], \quad (1)$$

जहाँ  $T/4$ ,  $\alpha_t$  उत्पत्ति की प्रक्रिया में तिमाही काल-श्रृंखला आँकड़ों के प्रतिमानों के लिए अधिकतम आवृत्तियों की संख्या है और  $A_k, B_k, k=1,2,\dots, T/4$  फूरियर गुणांक है। रैखिक एआर प्रतिमान के विपरीत, एफ-एआर अरैखिक काल-श्रृंखला प्रतिमान है।

यह (फैन और याओ 2003, मिलास इत्यादि, 2006) विषम आँकड़ों का वर्णन करने में सक्षम है (अमेंडॉला

और स्टोरटी, 2002; कियानी, 2009) जहाँ प्रेक्षण की औसत संख्या ऊपरी चक्र में निचले चक्र से भिन्न है। इस प्रतिमान को पशु उत्पादों या दूध, मांस और ऊन उत्पादन आदि के निर्यात और आयात के लिए मात्स्यिकी और पशु विज्ञान के क्षेत्र में किसी भी विशम काल-श्रृंखला आँकड़ों के प्रतिमान मूल्यों का आकलन और पूर्वानुमान निर्धारित करने के लिए प्रयोग किया जा सकता है।

### एफ-एआर प्रतिमान की फिटिंग और पूर्वानुमान

एफ-एआर के प्राचलों का मान प्राप्त करने के लिए लीस्ट-वर्ग पद्धति का प्रयोग किया जाता है। इसके अलावा,  $H_{0k}: A_k = 0, B_k = 0$ , विपरीत  $H_{1k}$ : कम से कम एक  $A_k$  या  $B_k$  अशून्य है,  $k = 1, 2, \dots, T/4$  का परीक्षण टी-टेस्ट (T-2) डिग्रीज़ ऑफ फ्रीडम द्वारा किया गया जहाँ समीकरण (1) में साइन-कोसाइन के एक जोड़े का प्रयोग किया गया है। महत्वपूर्ण फूरियर गुणांकों का आकलन करने के लिए मॉटे कार्लो सिमुलेशन अध्ययन द्वारा महत्वपूर्ण मूल्यों की गणना की गई है। T सैम्पल के लिए  $\{y_t\}$  स्वतंत्र और समान मानक सामान्य बंटन है। सिमुलेशन को 2000 बार दोहराया गया है। सिमुलेटिड सैम्पल से टी-सांख्यिकी का मान प्राप्त करने के पश्चात् आयत-चित्र बनाया गया है जिसमें X-अक्ष पर क्रमबद्ध टी-सांख्यिकी और Y-अक्ष पर आवृत्तियों को दर्शाया गया है। महत्वपूर्ण मान 5% स्तर पर इस प्रकार आकलित किए गए हैं कि 5% प्रेक्षण निर्धारित स्तर से ज्यादा हों। टी-सांख्यिकी का मान निर्धारित मान से अधिक था और परीक्षण 5%

स्तर पर महत्वपूर्ण माना गया और महत्वपूर्ण चरों को प्रतिमान में सम्मिलित किया गया है। फिटेटेड प्रतिमान की गुणवत्ता की जाँच एकाई मापदंड (ए.आई.सी.) और बेज़ियन मापदंड (बी.आई.सी.) द्वारा की गई है।

परीक्षण का आकार प्राप्त करने के लिए, डबल बूटस्ट्रेप तकनीक (एफरॉन और तिब्बिरानी, 1993) का प्रयोग किया गया है। इसके लिए सर्वप्रथम आउटसैट टी-सांख्यिकी की गणना एफ-एआर प्रतिमान के आकलित गुणांकों का प्रयोग कर प्राचलिक बूटस्ट्रेप द्वारा की गई है। दूसरे चरण में सिमुलेटिड सांख्यिकी मानों की गणना शून्य (नल) बंटन पर आधारित सैम्पल द्वारा की गई। प्रथम चरण के प्रत्येक आउटसैट मान प्राप्त करने के लिए प्रथम चरण के निर्धारित मान पर टी-सांख्यिकी के मानों की गणना की गई और यदि वह आउटसैट टी-मान से ज्यादा है तो उसे 1 मूल्य दिया गया, अन्यथा शून्य मूल्य दिया गया है। आउटसैट टी-सांख्यिकी मान से अधिक मूल्य पर टी-सांख्यिकी अनुपात की गणना की गई। यदि यह मान पूर्वनिर्धारित स्तर 0.05 से कम है तो  $H_{0k}$  को 5% स्तर पर निर्धारित प्रथम चरण आउटसैट मान के लिए अस्वीकार कर दिया गया। प्रथम चरण के संबंध में इस प्रक्रिया को 500 बार दोहराते हुए 'H<sub>0k</sub> का अस्वीकरण जब यह सही है' के अनुपात की गणना की गई जिससे परीक्षण का आकार ज्ञात हुआ है।

परीक्षण के परिमाण की गणना करने के लिए प्रथम चरण की टी-सांख्यिकी के मानों की गणना की गई। वैकल्पिक परिकल्पना के अंतर्गत प्रथम चरण के नमूनों को कुल संख्या माना गया। दूसरे चरण में प्रत्येक प्रथम चरण के नमूने पर समाश्रयण किया गया। प्राचलिक बूटस्ट्रेप पद्धति का प्रयोग कर सिमुलेटिड सांख्यिकी मान प्राप्त किए गये हैं। उपरोक्त पद्धति को दोहराते हुए 'ज्यादा' को 'कम' और 'सही' को 'गलत' से प्रतिस्थापित किया गया। उपरोक्त कार्य के निष्पादन के लिए संबंधित संगणक प्रोग्राम एसएसएस (SAS) सॉफ्टवेयर के आईएमएल में विकसित किए गए। एफ-एआर और एआर प्रतिमानों की गुणवत्ता की तुलना करने के लिए फिटेटेड प्रतिमानों से एक-तिमाही आगे के पूर्वानुमान आकलित किए गए और माध्यवर्ग

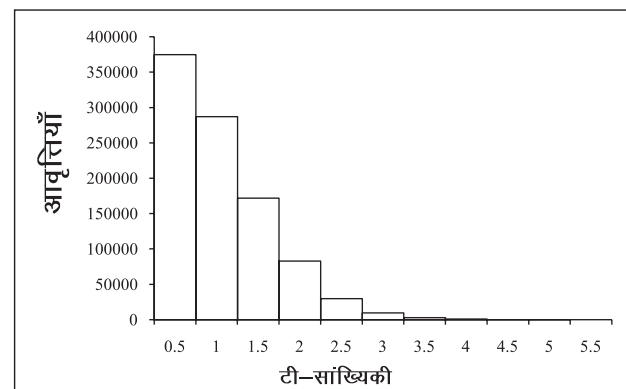
पूर्वानुमान त्रुटि की भी गणना की गई है।

## सांख्यिकी विश्लेषण

भारतीय तेल सारडीन (सारडीन लॉंगीसिप्स) भारत में सबसे महत्वपूर्ण व्यावसायिक मछलियों में से एक है और हिंद महासागर के उत्तरी क्षेत्रों में पाई जाती है। अध्ययन के लिए केन्द्रीय समुद्री मात्स्यिकी अनुसंधान संस्थान, कोच्चि के छब्बीस वर्ष (1985-2010) के तिमाही तेल सारडीन मछली के उत्पादन आँकड़ों (टन में) का उपयोग किया गया है। अध्ययन से आँकड़ों की चक्रीय प्रवृत्ति एवं विषमता का पता चलता है।

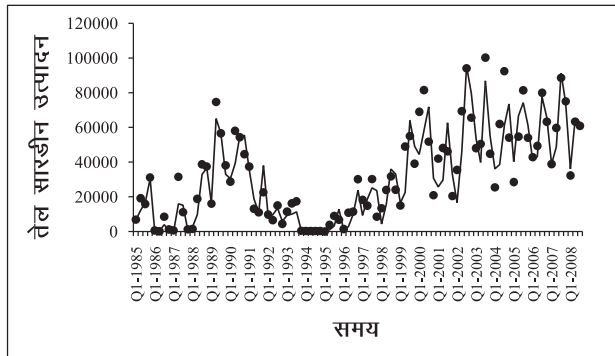
समीकरण (1) को फिट करने के लिए मोंटे कार्लो अनुकृति (सिमुलेशन) के माध्यम से 20,000 आँकड़ों की उत्पत्ति कर महत्वपूर्ण फूरियर गुणांक आकलित किए गए और टी-सांख्यिकी की गणना की गई, जिसे आयत-चित्र चित्र-1 में दर्शाया गया है। 5% स्तर पर महत्वपूर्ण मान 2.00 आँका गया अतः 5% स्तर पर महत्वपूर्ण फूरियर गुणांक  $A_1, B_1, A_2, B_3, B_6, A_{15}$  और  $B_{16}$  प्राप्त हुए एवं फिटेटेड एफ-एआर प्रतिमान निम्न प्रकार है;

$$y_t = -0.08 + 0.13 y_{t-1} + \left[ -0.87 \sin\left(2\pi \cdot 1 \frac{t}{96}\right) - 1.03 \cos\left(2\pi \cdot 1 \frac{t}{96}\right) - 1.79 \sin\left(2\pi \cdot 2 \frac{t}{96}\right) + 1.95 \cos\left(2\pi \cdot 3 \frac{t}{96}\right) - 1.61 \cos\left(2\pi \cdot 6 \frac{t}{96}\right) + 1.19 \sin\left(2\pi \cdot 15 \frac{t}{96}\right) + 0.68 \cos\left(2\pi \cdot 16 \frac{t}{96}\right) \right] y_{t-1} + \epsilon_t \quad (2)$$



चित्र-1: सिमुलेटिड क्रमबद्ध टी-सांख्यिकी का आयत-चित्र

ए.आई.सी. और बी.आई.सी. का मान क्रमशः 21.07 और 25.24 है। फिटेटेड एआर प्रतिमान के लिए ये मान क्रमशः 2157.359 और 2159.913 थे। विचाराधीन आँकड़ों के लिए दोनों सांख्यिकी के कम मान एआर प्रतिमान की तुलना में एफ-एआर प्रतिमान



चित्र-2: आँकड़ों के साथ फिटेटेड एफ-एआर प्रतिमान

की श्रेष्ठता को परिलक्षित करते हैं। एफ-एआर प्रतिमान के फिटेटेड मानों के साथ आँकड़ों का प्रदर्शन चित्र-2 में किया गया है।

इसके बाद परीक्षण के आकार की गणना की गई और महत्वपूर्ण परिणाम तालिका-1 में दर्शाए गए हैं। डबल बूटस्ट्रेप पद्धति का उपयोग करते हुए परीक्षण के वर्ग की गणना की गई और उसे भी तालिका-1 में दर्शाया गया है।

दो वर्षों के लिए तिमाही पूर्वानुमानों का आकलन किया गया और उन्हें तालिका-2 में दर्शाया गया है। अंत में, एफ-एआर प्रतिमान के लिए एमएसएफई का मान 41099124.32 प्राप्त हुआ जो कि एआर प्रतिमान के मान 995694711.90 से कम है। कम मान पूर्वानुमान के लिए एआर प्रतिमान की तुलना में एफ-एआर प्रतिमान की श्रेष्ठता को दर्शाता है।

तालिका-1: आकार की संगणना (5% स्तर पर) और परीक्षण का वर्ग

फूरियर गुणांक	परीक्षण सांख्यिकी	परीक्षण का आकार	परीक्षण का वर्ग
$A_1$	$t_{A_1}$	0.395	0.75
$B_1$	$t_{B_1}$	0.336	0.86
$A_2$	$t_{A_2}$	0.413	0.77
$B_3$	$t_{B_3}$	0.392	0.81
$B_6$	$t_{B_6}$	0.395	0.86
$A_{15}$	$t_{A_{15}}$	0.338	0.79
$B_{16}$	$t_{B_{16}}$	0.392	0.91

तालिका-2: एफ-एआर और एआर प्रतिमानों के लिए पूर्वानुमान संगणना

तिमाही / वर्ष	वास्तविक मूल्य (टन में)	एफ-एआर प्रतिमान के आधार पर पूर्वानुमान	एआर प्रतिमान के आधार पर पूर्वानुमान
1st / 2009	38635	44065.039	61321.740
2nd / 2009	14855	16168.642	61263.639
3rd / 2009	74225	71846.180	61272.648
4th / 2009	39712	43690.751	61271.251
1st / 2010	67470	66030.316	61271.647



तिमाही / वर्ष	वास्तविक मूल्य (टन में)	एफ-एआर प्रतिमान के आधार पर पूर्वानुमान	एआर प्रतिमान के आधार पर पूर्वानुमान
2nd / 2010	31303	32517.604	61271.434
3rd / 2010	41523	30713.824	61271.439
4th / 2010	119047	114673.314	61271.438

## सारांश

इस अध्ययन का उद्देश्य विषम चक्रीय आँकड़ों का वर्णन एवं पूर्वानुमान आकलित करने के लिए फूरियर स्वसमाश्रयी प्रतिमान एफ-एआर का उपयोग करना था। सांख्यिकीय विश्लेषण के लिए संगणक प्रोग्राम एसएसएस (SAS) सॉफ्टवेयर में विकसित किए गए हैं। भारत में स्थित केन्द्रीय समुद्री मात्स्यिकी अनुसंधान संस्थान, कोच्चि के छब्बीस वर्ष (1985-2010) के तिमाही तेल सारडीन मछली उत्पादन आँकड़ों (टन में) का उपयोग किया गया। दो वर्ष के तिमाही आँकड़ों का प्रयोग कर एक-तिमाही आगे के पूर्वानुमान विकसित कर एआर प्रतिमान की तुलना में एफ-एआर प्रतिमान की श्रेष्ठता को दर्शाया गया। इसकी क्षमता का उपयोग तेल सारडीन के लिए इष्टतम आयात और निर्यात नीतियों को विकसित करने के लिए किया जा सकता है। इस प्रकार की जानकारी संसाधनों का उपयोग कर मत्स्य पालन उद्योग को सक्षम करने में लाभप्रद सिद्ध होगी। नौकाओं और ट्रॉलर्स की इष्टतम संख्या को आबंटित करने के लिए कुशल तेल सारडीन प्रबंधन रणनीति को विकसित करने की आवश्यकता है, जिसमें यह अध्ययन उपयोगी सिद्ध होगा।

## संदर्भ

1. अमेंडॉला, ए. एवं स्टोरटी, जी. (2002): “ए नॉनलीनियर टाइम-सीरीज़ एप्रोच टु मॉडलिंग एसिमैट्री इन स्टॉक मार्केट इंडेक्सेस”, स्टैटिस्टिकल मेथड्स एंड एप्लिकेशन्स, 11: 201-216।
2. ब्लूमफील्ड, पी. (2000): “फूरियर एनालिसिस ऑफ टाइम-सीरीज़: ऐन इंट्रोडक्शन”, द्वितीय एडीशन, जॉन वाईले एण्ड सन्स, यू.एस.ए.।
3. बॉक्स, जी.ई.पी., जेनकिन्स जी. एम. एवं रेंसिल जी.सी. (2009): “टाइम-सीरीज़ एनालिसिस: फॉरकार्स्टिंग एण्ड कंट्रोल”, तृतीय एडीशन, पियर्सन एजुकेशन, सैन फ्रांसिस्को।
4. एफरॉन, बी. एवं तिब्बिरानी, जे. आर. (1993): “एन इंट्रोडक्शन टु दी बूटस्ट्रैप”, चैपमैन एण्ड हॉल, न्यूयार्क।
5. फैन, जे. एवं याओ, क्यू. (2003): “नॉनलीनियर टाइम-सीरीज़”, नॉन पैरामीट्रिक एण्ड पैरामीट्रिक मेथड्स, स्प्रिंगर, न्यूयार्क।
6. कियानी, के. एम. (2009): “एसिमैट्रीज़ इन मैक्रोइकनोमिक टाइम-सीरीज़ इन इलैविन एशियन इकोनॉमीज़”, इंटरनेशनल जर्नल ऑफ बिज़नेस एण्ड इकोनॉमिक्स, 8:37.54।
7. लुडलो, जे. एवं एंडर्स, डब्ल्यू. (2000): “एस्टिमेटिंग नॉनलीनियर अरमा मॉडल्स यूज़िंग फूरियर कोफिशिएंट्स”, इंटरनेशनल जर्नल ऑफ फोरकार्स्टिंग, 16: 333.347।
8. मिल्लास, सी., रोथमन, पी. एवं डिक, डी. (2006): “नॉनलीनियर टाइम-सीरीज़ एनालिसिस ऑफ बिज़नेस साइकिल्स”, एल्सेवियर बी.वी., एम्स्टर्डम।

□

## सी.बी.पी. वॉर्टल : भा.कृ.अनु.प. द्वारा प्रायोजित प्रशिक्षण कार्यक्रम के ऑनलाइन प्रबन्धन हेतु सूचना तंत्र

अलका अरोड़ा, सुदीप मरवाह, कुसुमाकर शर्मा, प्रदीप कुमार मल्होत्रा,  
ए.के.एम. सामिमुल आलम, पुष्पांजलि रावत एवं ऊषा जैन

कृषि के ज्ञान आधारित विकास में कृषि में सूचना संचार प्रौद्योगिकी के अनुप्रयोग में नवीन परिवर्तनों के संवर्धन एवं ज्ञान के प्रचार की अहम भूमिका है। अतः, अनुसंधान एवं शिक्षा को और अधिक प्रासंगिक बनाने के लिए वांछित गुणात्मक सुधार एवं समकालीन समस्याओं के आवश्यक अभिविन्यास के लिए अध्यापकों, शोधकर्ताओं एवं विस्तार विशेषज्ञों की विशिष्टता के क्षेत्र में नवीनतम ज्ञान एवं तकनीकों के व्यावसायिक कौशल को अद्यतन करने की आवश्यकता है। भा.कृ.अनु.प. की सामर्थ्य निर्माण नीति के अनुसार परिवर्तन को ध्यान में रखते हुए राष्ट्रीय कृषि अनुसंधान प्रणाली (एन.ए. आर.एस.) के शोधकर्ताओं एवं अध्यापकों का प्रशिक्षण जारी रहेगा तथा उनके व्यवसाय में अभिनव परिवर्तनों से उन्हें नियमित रूप से अवगत कराया जाएगा। सामर्थ्य फ्रेमवर्क का मूलभूत उद्देश्य एन.ए.आर.एस. के प्रोफेशनलों में केवल कार्य के लिए ही अपेक्षित योग्यता न हो बल्कि आवश्यकता आधारित अनुसंधान एवं शैक्षणिक कार्यक्रमों का आयोजन एवं विकास में सुधार लाने में सक्षम हों जिससे आजीविका सुरक्षा में वृद्धि होगी तथा सरल, सुलभ एवं लागत प्रभावी ज्ञान गहन सूचना प्रणाली का निर्माण होगा। वैज्ञानिकों/अध्यापकों के कैरियर की प्रगति सामर्थ्य निर्माण कार्यक्रमों से आवधिक ज्ञान से जुड़ी है। इसके लिए भा.कृ. अनु.प. उन्नत संकाय प्रशिक्षण केन्द्र (सी.ए.एफ.टी.) ग्रीष्मकालीन-शीतकालीन स्कूल (एस.डब्ल्यू.एस.) एवं अल्पकालीन पाठ्यक्रमों (21/10 दिवसीय अवधि के) को प्रायोजित कर समकालीन प्रसांगिकता एवं प्रत्याशित भविष्य के कमी वाले क्षेत्रों में सामर्थ्य निर्माण

को सम्बोधित करता है। कृषि एवं इसके संबद्ध विज्ञान के अत्याधुनिक क्षेत्रों में अध्यापन, अनुसंधान एवं विस्तार के संदर्भ में सामर्थ्य-निर्माण की माँग को पूरा करने के लिए ये कार्यक्रम एन.ए.आर.एस. के वैज्ञानिकों एवं संकाय सदस्यों की बहुत मदद कर रहे हैं।

भा.कृ.अनु.प. के कृषि शिक्षा प्रभाग द्वारा प्रायोजित प्रशिक्षण कार्यक्रमों के ऑनलाइन प्रबन्धन को सुविधाजनक बनाने हेतु सामर्थ्य निर्माण (सी.बी.पी.) के लिए एक वॉर्टल विकसित किया गया है। यह वॉर्टल भारतीय कृषि सांख्यिकी अनुसंधान संस्थान (भा.कृ.सां.अ.सं.) के संगणक अनुप्रयोग प्रभाग द्वारा भा.कृ.अनु.प. के कृषि शिक्षा प्रभाग के सहयोग से विकसित किया गया है। यह सिस्टम भा.कृ.सां.अ.सं. के सर्वर पर होस्ट किया गया है तथा इसे <http://iasri.res.in/cbp> से अथवा <http://icar.org.in> लिंक से एक्सेस किया जा सकता है।

### सी.पी.बी. की महत्वपूर्ण विशेषताएँ

1. भा.कृ.अनु.प. के कृषि शिक्षा प्रभाग द्वारा प्रायोजित समस्त प्रशिक्षण कार्यक्रमों से सम्बन्धित सूचना।
2. प्रशिक्षण कार्यक्रम हेतु आवेदन ऑनलाइन प्रस्तुत करना।
3. पाठ्यक्रम निदेशक द्वारा प्रशिक्षण प्रस्ताव की प्रस्तुति के लिए वर्क-फ्लो तथा भा.कृ.अनु.प. के कृषि शिक्षा प्रभाग द्वारा प्रस्ताव का मूल्यांकन।
4. प्रशिक्षण कार्यक्रमों की ई-पुस्तकों/लेक्चर नोट्स की ऑनलाइन उपलब्धता।

5. ऑनलाइन—हेल्प।
6. वॉर्टल के डिज़ाइन एवं विकास में मॉड्यूलर अप्रोच।
7. प्रभावशाली प्रशासनिक मॉड्यूल।
8. सभी वर्गों के प्रयोक्ताओं के लिए रिपोर्ट मॉड्यूल्स।

### सी.बी.पी. मॉड्यूल्स

सी.बी.पी. वॉर्टल विभिन्न प्रयोक्ताओं के लिए निम्नलिखित मॉड्यूल्स सहित डिज़ाइन कर विकसित किया गया है :

1. पाठ्यक्रम निदेशक मॉड्यूल
2. प्रतिभागी मॉड्यूल
3. सी.ए.एफ.टी. निदेशक मॉड्यूल
4. सहायक महानिदेशक मॉड्यूल
5. उप—महानिदेशक मॉड्यूल
6. प्रशासक मॉड्यूल
7. होम पेज

#### (i) पाठ्यक्रम निदेशक मॉड्यूल

यह मॉड्यूल पाठ्यक्रम निदेशक को प्रशिक्षण के प्रस्ताव की प्रस्तुति से लेकर रिपोर्ट प्रस्तुति तक प्रशिक्षण कार्यक्रम सम्बन्धी समस्त गतिविधियाँ निष्पादित करने की सुविधा प्रदान करता है। पाठ्यक्रम निदेशक मॉड्यूल की निम्नलिखित विशेषताएँ हैं :

- प्रशिक्षण प्रस्ताव की प्रस्तुति एवं मूल्यांकन।
- फाइलें, समय—सारणी एवं लेक्चर फाइलें अपलोड करना।
- प्रतिभागियों की स्वीकृति/मेलिंग सूची/अस्वीकृति।
- वॉर्टल के माध्यम से प्रतिभागियों के साथ मेल के माध्यम से समन्वय स्थापित करना।
- प्रतिभागी की ऑनलाइन अभिस्वीकृति।
- प्रतिभागी द्वारा प्रशिक्षण कार्यक्रम का ऑनलाइन मूल्यांकन।
- शिक्षा विभाग/प्रशासक को प्रशिक्षण के सम्बन्ध में फीड—बैक।
- प्रशिक्षण के लिए आवेदन करने वाले प्रतिभागियों,

चुने गए प्रतिभागियों की रिपोर्ट तथा प्रतिभागियों द्वारा प्रशिक्षण का मूल्यांकन।

#### (ii) प्रतिभागी मॉड्यूल

प्रतिभागी भा.कृ.अनु.प. के कृषि शिक्षा प्रभाग द्वारा प्रायोजित किसी भी प्रशिक्षण कार्यक्रम में सी.बी.पी. वॉर्टल पर लॉगिन अकाउन्ट बनाकर आवेदन कर सकते हैं। इस मॉड्यूल की निम्नलिखित विशेषताएँ हैं :

- व्यक्तिगत, शैक्षणिक, शोध—अनुभव एवं अन्य विवरण सहित ऑनलाइन प्रपत्र भरना। भरी हुई सूचना के आधार पर सिस्टम से अग्रिम आवेदन—प्रपत्र जनरेट होना।
- विधिवत अनुमोदित प्रपत्र सिस्टम पर अपलोड करना।

#### (iii) सी.ए.एफ.टी. निदेशक मॉड्यूल

यह मॉड्यूल सी.ए.एफ.टी. निदेशक मॉड्यूल को सी.ए.एफ.टी. केन्द्र से सम्बन्धित गतिविधियाँ निष्पादित करने की सुविधा प्रदान करता है।

- सी.ए.एफ.टी. केन्द्र की सूचना जोड़ना/अद्यतन करना।
- सी.ए.एफ.टी. केन्द्र के अन्तर्गत कर्मियों द्वारा प्रस्तुत प्रशिक्षण प्रस्ताव का अनुमोदन/अस्वीकृति।
- उच्च प्राधिकारियों को प्रशिक्षण प्रस्ताव की प्रस्तुति।
- ई—पुस्तकें/छायाचित्र (फोटो)/प्रशिक्षण रिपोर्टें एवं वार्षिक रिपोर्ट अपलोड करना।

#### (iv) सहायक महानिदेशक मॉड्यूल

कृषि शिक्षा प्रभाग के प्रबन्धकों को एक प्रशिक्षण से समस्त प्रशिक्षणों की संकलित रिपोर्टें एवं प्रतिभागियों तक की विभिन्न स्तरों की सूचना देखने का प्रावधान है। सहायक महानिदेशक मॉड्यूल की निम्नलिखित विशेषताएँ हैं :

- प्रशिक्षण प्रस्तावों का ऑनलाइन मूल्यांकन एवं स्वीकृति/अस्वीकृति।
- वॉर्टल के माध्यम से पाठ्यक्रम निदेशकों एवं सी.ए.एफ.टी. निदेशकों के साथ ऑनलाइन सम्प्रेक्षण।

- परिपत्रों/छायाचित्रों (फोटो) को अपलोड करने तथा सी.ए.एफ.टी. निदेशकों के विवरण को अद्यतन करने हेतु डाटा प्रबन्धन मॉड्यूल ।
- विभिन्न प्रशिक्षण कार्यक्रमों (सी.ए.एफ.टी. ग्रीष्मकालीन-शीतकालीन स्कूल, अल्पकालीन पाठ्यक्रम) की रिपोर्टें ।
- प्रशिक्षण कार्यक्रमों के प्रतिभागियों पर रिपोर्ट ।
- प्रशिक्षण कार्यक्रमों की ई-पुस्तकों/लेक्चर नोट्स को एक्सेस करना ।
- प्रशिक्षण सम्बन्धी दस्तावेज सिस्टम से जनरेट करना ।

#### (v) उप-महानिदेशक मॉड्यूल

इस स्तर पर प्रबन्धकों को प्रशिक्षण प्रस्तावों के मूल्यांकन एवं रिपोर्टों को एक्सेस करने के लिए निम्नलिखित सुविधाएँ उपलब्ध होंगी :

- प्रशिक्षण प्रस्तावों का ऑनलाइन मूल्यांकन ।
- विभिन्न प्रशिक्षण कार्यक्रमों (सी.ए.एफ.टी. ग्रीष्मकालीन-शीतकालीन स्कूल, अल्पकालीन पाठ्यक्रम) की रिपोर्टें ।
- प्रशिक्षण कार्यक्रमों की ई-पुस्तकों/लेक्चर नोट्स को एक्सेस करना ।

#### (vi) प्रशासन मॉड्यूल

इस मॉड्यूल में वॉर्टल के सम्पूर्ण क्रियाकलापों को नियन्त्रित करने की सुविधा उपलब्ध है :

- विभिन्न प्रकार के प्रयोक्ता खाते बनाना ।
- प्रशिक्षण कार्यक्रम से सम्बन्धित घोषणाओं को अपलोड करना ।
- डाटा प्रबन्धन एवं रिपोर्टों के लिए सभी स्तरों पर वॉर्टल मॉड्यूल को एक्सेस करना ।

#### (vii) होम पेज

वॉर्टल का होम पेज भा.कृ.अनु.प. के कृषि शिक्षा प्रभाग द्वारा आयोजित प्रशिक्षण कार्यक्रमों में रुचि रखने वाले प्रयोक्ताओं को जानकारी उपलब्ध कराता है । होम

पेज पर निम्नलिखित महत्वपूर्ण सूचना उपलब्ध है :

- कृषि शिक्षा प्रभाग द्वारा आयोजित प्रशिक्षण कार्यक्रमों (सी.ए.एफ.टी. ग्रीष्मकालीन-शीतकालीन स्कूल, अल्पकालीन पाठ्यक्रम) से सम्बन्धित सूचना ।
- प्रशिक्षण कैलेंडर ।
- प्रशिक्षण कार्यक्रम के लिए चुने गये अभ्यर्थियों के सम्बन्ध में सूचना ।
- प्रशिक्षण कार्यक्रम की ई-पुस्तकें/लेक्चर नोट्स ।
- शिक्षा विभाग के प्रपत्रों/रिपोर्टों/परिपत्रों/अधिसूचनाओं के सम्बन्ध में जानकारी ।
- प्रशिक्षण कार्यक्रम की फोटो गैलरी ।
- सी.ए.एफ.टी. केन्द्रों के सम्बन्ध में सूचना ।

#### निष्कर्ष

सी.बी.पी. वॉर्टल भा.कृ.अनु.प. के कृषि शिक्षा प्रभाग द्वारा प्रायोजित प्रशिक्षण कार्यक्रमों के प्रबन्धन एवं सूचना हेतु एक मंच उपलब्ध कराता है। वॉर्टल में उपलब्ध सूचना का किसी भी समय, उपभोक्ता की आवश्यकता के अनुरूप विश्लेषण किया जा सकता है। वॉर्टल में प्रशिक्षण कार्यक्रमों के लिए बनाए गए लेक्चर नोट्स/ई-पुस्तकें भी उपलब्ध हैं जो ई-लर्निंग का बहुमूल्य साधन हैं। भविष्य में सी.बी.पी. वॉर्टल के ज्ञान आधार (knowledge base) का उपयोग मानव संसाधन के विकास एवं प्रशिक्षण कार्यक्रमों को और अधिक प्रभावी बनाने के लिए किया जाएगा ।

#### आभार

लेखकगण, भारतीय कृषि अनुसंधान परिषद् के कृषि शिक्षा प्रभाग के उप-महानिदेशक, डॉ. अरविन्द कुमार का इस वॉर्टल के विकास के लिए उपलब्ध कराए गये वित्तीय सहयोग एवं मार्गदर्शन के लिए आभार प्रकट करते हैं। लेखकगण, भारतीय कृषि सांख्यिकी अनुसंधान संस्थान के निदेशक, डॉ. उमेश चन्दर सूद द्वारा सी.बी.पी. वॉर्टल के विकास के लिए उपलब्ध करायी गयीं सुविधाओं के लिए आभार व्यक्त करते हैं।

□

## भारत में सरसों की वस्तुओं में बाजार समाकलन

प्रवीण आर्य, धर्मराज सिंह, अनिल कुमार, कमलेश नारायण सिंह,  
रंजीत कुमार पॉल एवं उषा जैन

भारत विश्व में खाद्य तेलों का प्रमुख उत्पादक, आयातक एवं उपभोक्ता है। भारतीय वनस्पति तेल अर्थव्यवस्था का विश्व में चौथा स्थान है जिसका विश्व के तिलहन उत्पादन में लगभग 10 प्रतिशत का हिस्सा है। इसकी प्रति व्यक्ति न्यूनतम खपत (10.6 कि.ग्रा. प्रति वर्ष) के बावजूद भी भारत में 11.2 प्रतिशत वनस्पति तेल का आयात तथा 9.3 प्रतिशत खाद्य तेल की खपत होती है। देश के तिलहन उत्पादन में सोयाबीन, मूँगफली एवं सरसों का कुल मिलाकर लगभग 85 प्रतिशत योगदान है।

देश के उत्तरी एवं पूर्वी भागों में सरसों का तेल व्यापक रूप से खाद्य तेल के रूप में प्रयोग किया जाता है। स्वास्थ्य पर इसके प्रभावों की आशंका के बावजूद, जिस कारण यूरोप, अमेरिका एवं कनाडा जैसे विकसित देशों में खाना पकाने के लिए सरसों के तेल के प्रयोग पर प्रतिबन्ध लगा, समग्र खाद्य तेल खपत में इसका 13 प्रतिशत हिस्सा है एवं देश में इसे व्यापक लोकप्रियता एवं वरीयता प्राप्त है। सरसों की फसल के अन्तर्गत कुल क्षेत्र लगभग 6.45 मिलियन हेक्टेयर है और अनुमानित उत्पादन 7.11 मिलियन टन एवं उपज 1104 कि.ग्रा./हेक्टेयर (2012-13) है।

खाद्य आपूर्ति श्रृंखलाएँ गतिशील हैं तथा वे नीति परिवर्तन के प्रति अपने आपको शीघ्र समायोजित कर लेती हैं। भारत में, कृषि उत्पादों की आपूर्ति लगभग पूरी तरह से बाजारों द्वारा संचालित होती है। हालांकि, ये विदेशी प्रत्यक्ष निवेश, ग्रामीण भूमिहीन

श्रमिकों के लिए न्यूनतम गारन्टी योजना, एकाधिकार बीज कीमतों पर नियंत्रण अथवा सीलिंग, सब्सिडी का सीधा हस्तान्तरण, भविष्य बाजारों में चयनित कृषि वस्तुओं के व्यापार पर रोक अथवा नियंत्रण जैसी नीतियों में परिवर्तन से परोक्ष रूप से प्रभावित होते हैं। यह अध्ययन, भारत में प्रमुख बाजारों के लिए स्पेशियल एवं वर्टिकल इन्टीग्रेशन फ्रेमवर्क में सरसों की वस्तुओं अर्थात् बीज, तेल एवं केक की प्रकृति एवं विस्तार में अन्तर करने के लिए किया गया। यदि एक बाजार में कीमत संकेत एक दूसरे के साथ सांमजस्य में हो तो दो बाजारों को स्थानिक रूप से पृथक बाजारों में समाकलित कहा जाता है। इसे स्थानिक रूप से पृथक बाजारों में कीमत संकेतों एवं सूचना के निर्विघ्न हस्तान्तरण के रूप में परिभाषित किया गया है (गोलेटि, इत्यादि 1995)। बाजार समाकलन दक्षता का संकेत है और इसे दो स्थानों में माँग एवं आपूर्ति के विस्तार के माप के रूप में भी परिभाषित किया जा सकता है, (निगासा इत्यादि 2003)। इसी प्रकार, यदि वस्तुओं की कीमतें मूल्य श्रृंखला में वैसा व्यवहार दिखाते हैं तो बाजार लम्बवत (वर्टिकली) समाकलित हैं। बाजार समाकलन का स्पेशियली एवं वर्टिकली अध्ययन सरसों के तेल उत्पादन एवं विपणन में आपूर्ति श्रृंखला प्रबंधन सम्बंधी मुद्दों पर प्रकाश डालेगा तथा इससे घरेलू उत्पादन एवं प्रसंस्करण को प्रेरित करने के लिए उचित बाजार संकेतों के निर्देशन में भी मदद मिलेगी और इससे दक्षता हानियों (efficiency loss) में कमी होगी।

## डाटा एवं पद्धति

इस अध्ययन में सरसों की वस्तुओं के 1996 से 2008 तक की अवधि से सम्बन्धित काल-श्रृंखला आँकड़ों का प्रयोग किया गया है। सात बाजारों (दिल्ली गौंडा, हापुर, हाथरस, हिनडोन, कानपुर एवं कोलकाता) के सरसों के बीज तथा सरसो के तेल के लिए छः बाजारों (दिल्ली, हापुड, कानपुर, कोलकता, पटना एवं रोहतक) के मासिक थोक मूल्य आँकड़ों का चयन स्पेशियल समाकलन तथा वर्टिकल समाकलन के लिए किया गया। चयनित वस्तुओं की कीमतों में व्यवहार का अध्ययन करने के लिए दिल्ली एवं कानपुर का चयन किया गया।

सह-समाकलन परीक्षण में मूल आधार यह है कि इसमें दो या दो से अधिक काल श्रृंखला चरों के बीच दीर्घावधि सन्तुलन सम्बन्ध हो। ऐसे सम्बन्ध को बनाए रखने के लिए चरों को वैसी ही इन्टर टेम्पोरल विशेषताएं वहन करनी चाहिए। अतः न्यूनतम ए.आई.सी. (AIC) एकैक सूचना मानदण्ड का प्रयोग करते हुए संबंधित डिफि फ्यूलर ए.डी.एफ. (ADF) लागू कर चरों की स्थिरता के लिए जांच की जाती है। निम्नलिखित समाश्रयण को चलाकर ए.डी.एफ. परीक्षण लागू किया जाता है :

$$\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + \sum_{i=1}^m \alpha_i \Delta Y_{t-i} + \varepsilon_t$$

जहां पर  $Y_t$  एक सदिश है जिसका सह-समाकलन के लिए परीक्षण किया जाना है,  $t$  काल चर है तथा  $\varepsilon_t$  व्हाइट नोएस एरर टर्म है  $\beta$ 's,  $\delta$  एवं  $\alpha$ 's प्राचल हैं जिनका आकलन किया जाना है। यदि गुणांक  $\delta$  शून्य से सांख्यिकीय रूप से भिन्न नहीं है तो इसका तात्पर्य है कि कीमत श्रृंखला यूनिट रूट है तथा श्रृंखला गैर-स्थिर है।

सह-समाकलन समीकरणों की संख्या का आकलन करने के लिए गैर स्थिर श्रृंखलाओं पर जाहनसेन बहुचर मैक्सिमम लाइकलीहुड (maximum likelihood) को इन्टीग्रेशन तकनीक (ग्लारिया एवं स्टिवन 2001) पर आधारित अनुक्रमिक प्रक्रिया लागू की गयी।

सह-समाकलन का उद्देश्य यह निर्धारित करना है कि गैर-स्थिर श्रृंखलाओं का एक समूह सह-समाकलित है अथवा नहीं तथा वी.ई.सी. विनिर्देशन का आधार बनता है। P आर्डर के VAR पर आधारित जोहनसेन (1991, 1995) सह-समाकलन विधि को निम्नानुसार वर्णित किया जा सकता है :

$$y_t = A_1 y_{t-1} + \dots + A_p y_{t-p} + Bx_t + \varepsilon_t$$

जहां पर  $y_t$  गैर स्थिरता  $k$  चरों का वेक्टर है,  $x_t$  निर्धारणात्मक चरों का  $d$  वेक्टर है, तथा  $\varepsilon_t$  नवाचरों का एक वेक्टर है। हम VAR को निम्नानुसार पुनः लिख सकते हैं :

$$\Delta y_t = \Pi y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + Bx_t + \varepsilon_t$$

$$\Pi = \sum_{i=1}^p A_i - I, \Gamma_i = \sum_{j=i+1}^p A_j$$

जहां ग्रेन्जर की प्रतिनिधित्व प्रमेय के अनुसार, यदि गुणांक मैट्रिक्स  $\Pi$  में न्यूनीकृत रैंक  $r < k$  है, तो वहाँ  $k \times r$ ,  $\alpha$  और  $\beta$  मैट्रिक्स होगी एवं प्रत्येक का रैंक  $r$  होगा। इसी प्रकार कि  $\Pi = \alpha\beta'$  एवं  $\beta'Y_t$  स्थिर है अर्थात्  $I(0)$  है, भले ही  $Y_t$  स्वयं  $I(1)$  है (एंजल एवं ग्रेंजर, 1987)।  $|r$  सह-समाकलन सम्बन्धों सह-समाकलित रैंक की संख्या है तथा  $\beta$  का प्रत्येक कॉलम सह-समाकलित वेक्टर है। सह-समाकलन मैट्रिक्स मॉडल में दीर्घ अवधि सम्बन्धों का वर्णन करता है। वेक्टर त्रुटि सुधार मॉडल में  $\alpha$  तत्व समायोजन प्राचलों के रूप में जाने जाते हैं, जो लॉग-रन सम्बन्धों से विचरण पर विषिष्ट चरों के समायोजन की गति को स्पष्ट करते हैं। हालांकि, श्रृंखलाओं में सह-समाकलन स्थापित किया गया है, अलावधि में असंतुलन हो सकता है, अर्थात् बाजारों में मूल्य समायोजन तुरन्त न हो, बाजारों को समायोजन में समय लग सकता है। वेक्टर त्रुटि सुधार मॉडल ई.सी.एम चरों की त्रुटि अवधि में लॉग-रन सम्बन्धों से विचरण से मैग्नीट्यूड का तथा चरों की प्रतिक्रिया के मार्ग का सुझाव देता है। दूसरे शब्दों में यह सन्तुलन के प्रति चरों के समायोजन की गति अंकित करता है।

## परिणाम एवं परिचर्चा

सरसों की वस्तुओं, अर्थात् बीज एवं तेल के लिए चयनित बाजारों की कीमत श्रृंखलाओं में सहसमाकलन लागू किया गया, तालिका-1 में यूनिट रूट के ADF परीक्षण के परिणाम से पता चलता है कि सरसों के बीज के मामले में छः समाकलित वेक्टर उपस्थित थे, जिसका तात्पर्य है कि सभी चयनित सात बाजारों में आम प्रवृत्ति है तथा स्पेशियल रूप से सह-समाकलित हैं, जो इन बाजारों के बीच दीर्घ-अवधि के सन्तुलन सम्बन्ध की उपस्थिति को दर्शाता है। 5 प्रति सार्थकता स्तर पर छः समाकलन समीकरण दर्शाते हैं कि सात चर्चा बाजार एक कॉमन पाथ शेयर करते हैं। तालिका-2 में प्रस्तुत ई.सी.के संगत परिणाम सरसों के बीज के लिए सांख्यिकीय रूप से सर्व नकारात्मक मान ज्ञान हुआ, जहाँ हापुड़ एवं हिन्डौन में दीर्घ-अवधि सन्तुलन के प्रति गति का समायोजन तेज है तथा अन्य बाजार लगभग 2-3 सप्ताह धीमा है। दिल्ली में समायोजन में अधिक समय लगा होगा क्योंकि मान निम्न स्तर पर हैं। ऐसा प्रतीत होता है कि कोलकाता में कोई त्रुटि सुधार नहीं हो रहे हैं जिसका तात्पर्य है कि कीमतें एक बार सन्तुलन से विस्थापित होने पर पूर्व सन्तुलन स्तर पर वापस नहीं आ सकतीं। सरसों के बीज के मामले में, दीर्घ-अवधि के समायोजन पाथ अधिक स्थिरता के पक्ष में सबूत उपलब्ध कराते हैं। सरसों के बीज में क्षेत्रीय अधिकता है तथा राज्य स्तरीय कीमतें एम एस पी के अनुसार स्थानान्तरित पायीं जिससे कम विचरण सुनिश्चित होता है।

सरसों के तेल के मामले में, 1998 से 2007 तक की अवधि से सम्बन्धित बाजारों, नामतः दिल्ली, हापुड़, कानपुर, कोलकाता, पटना एवं रोहतक की कीमतों का प्रयोग किया गया। सरसों के तेल के चयनित सभी छः बाजार एक ही स्तर के पाए गए। जोहनसेन परीक्षण ने इन बाजारों में दीर्घ अवधि की ऐसोसिएशन की पुष्टि की और इसलिए एक आम प्रवृत्ति को शेयर करते हैं। तालिका-3 से यह स्पष्ट है कि अधिकतम आइजेन

मान परीक्षण द्वारा, सरसों के तेल में 5 सह-समाकलित समीकरण देखे गए। ये परिणाम सरसों के बीज के साथ सामंजस्य में हैं, जो यह दर्शाता है कि दीर्घ अवधि में विभिन्न बाजारों में ये दो वस्तुएँ स्पेशियली रूप से सह-समाकलित हैं।

**तालिका-1** : सरसों के बीज में स्पेशियल समाकलन का बहुचर विश्लेषण

स्थान	H <sub>0</sub>	अधिकतम आइजेन	CE(s) की संख्या	VAR का आर्डर
दिल्ली गोंडा हापुड़	r = 0	0.39	None *	1
	r ≤ 1	0.26	अधिकतम 1*	
	r ≤ 2	0.03	अधिकतम 2	
दिल्ली गोंडा हापुड़ हाथरस	r = 0	0.39	None *	1
	r ≤ 1	0.31	अधिकतम 1*	
	r ≤ 2	0.20	अधिकतम 2*	
	r ≤ 3	0.04	अधिकतम 3	
दिल्ली गोंडा हापुड़ हाथरस हिन्डोन कानपुर	r = 0	0.44	None *	1
	r ≤ 1	0.41	अधिकतम 1*	
	r ≤ 2	0.27	अधिकतम 2*	
	r ≤ 3	0.23	अधिकतम 3*	
	r ≤ 4	0.20	अधिकतम 4*	
	r ≤ 5	0.03	अधिकतम 5	
दिल्ली गोंडा हापुड़ हाथरस हिन्डोन कानपुर कोलकाता	r = 0	0.52	None *	7
	r ≤ 1	0.46	अधिकतम 1*	
	r ≤ 2	0.33	अधिकतम 2*	
	r ≤ 3	0.22	अधिकतम 3*	
	r ≤ 4	0.18	अधिकतम 4*	
	r ≤ 5	0.12	अधिकतम 5*	
r ≤ 6	0.03	अधिकतम 6		

\* 5 प्रतिशत सार्थकता स्तर पर परिकल्पना की अस्वीकृति को दर्शाता है।

**तालिका-2** : सरसों के बीज में स्पेशियल बाजारों में न्यूनीकृत VECM

दिल्ली	गोंडा	हापुड	हाथरस	हिन्डोन	कानपुर	कोलकाता
-0.018	-0.36	-1.03	-0.62	-1.67	-0.43	0.37
(-0.038)	(-2.13)	(-1.99)	(-2.62)	(-3.86)	(-1.48)	(0.77)

(कोष्ठक में आँकड़े t-स्टैटिस्टिक को दर्शाते हैं)

सरसों के तेल के लिए ECM के गुणांक नकारात्मक एवं 1 से कम हैं। सरसों के तेल के मामले में ये आकलन दिल्ली, कोलकाता एवं पटना के लिए उच्च थे जो सन्तुलन के प्रति त्वरित समायोजन का प्रतीक है। रोहतक एवं हापुड में कीमतों को समायोजन में अधिक समय, अर्थात् तीन सप्ताह से अधिक (तालिका-4) लगता है। सरसों के तेल के लिए भी अपेक्षाकृत उच्च मान दीर्घ अवधि के समायोजन पाथ में कम विचलन दर्शाते हैं।

**तालिका-3** : सरसों के तेल में स्पेशियल समाकलन का बहुचर विश्लेषण

स्थान	H <sub>0</sub>	अधिकतम आइजेन	CE(s) की संख्या	VAR का आर्डर
दिल्ली	r = 0	0.31	None *	1
हापुड	r ≤ 1	0.17	अधिकतम 1*	
कानपुर	r ≤ 2	0.04	अधिकतम 2	
दिल्ली	r = 0	0.36	None *	1
हापुड	r ≤ 1	0.27	अधिकतम 1*	
कानपुर	r ≤ 2	0.15	अधिकतम 2*	
कोलकाता	r ≤ 3	0.04	अधिकतम 3	
दिल्ली	r = 0	0.48	None *	1
हापुड	r ≤ 1	0.36	अधिकतम 1*	
कानपुर	r ≤ 2	0.22	अधिकतम 2*	
कोलकाता	r ≤ 3	0.16	अधिकतम 3*	
पटना	r ≤ 4	0.13	अधिकतम 4*	
रोहतक	r ≤ 5	0.04	अधिकतम 5	

\* 5 प्रतिशत सार्थकता स्तर पर परिकल्पना की अस्वीकृति को दर्शाता है।

**तालिका-4** : सरसों के तेल में स्पेशियल बाजारों में न्यूनीकृत VECM

दिल्ली	हापुड	कानपुर	कोलकाता	पटना	रोहतक
-0.88	-0.27	-0.37	-0.71	-0.41	-0.017
(-5.25)	(-1.55)	(-2.18)	(-3.65)	(-4.48)	(-0.13)

(कोष्ठक में आँकड़े t-स्टैटिस्टिक को दर्शाते हैं)

जोहनसेन बहुचर सहसमाकलन तकनीक का प्रयोग करते हुए दिल्ली एवं कानपुर के चयनित थोक बाजारों में सरसों की वस्तुओं (बीज, तेल एवं ऑयल केक) में वर्टिकल समाकलन के विस्तार का अध्ययन किया गया। दिल्ली एवं कानपुर के बाजारों में सरसों की प्रोसेसिंग के विभिन्न चरणों में वर्टिकल समाकलन का अध्ययन किया गया। तालिका-5 ने दर्शाती है कि सरसों के बीज, तेल तथा ऑयल केक बाजारों के बीच एक सहसमाकलन वेक्टर उपस्थित था, जो यह अंकित करता है कि बीज एवं तेल सह-समाकलित हैं जबकि ऑयल-केक नहीं था, यह इन उत्पादों के विभिन्न माँग एवं आपूर्ति फ़सलों के कारण हो सकता है।

तालिका-6 में VECM मान दर्शाते हैं कि भले ही दिल्ली और कानपुर दोनों बाजारों में बीज एवं तेल समाकलित हैं, कानपुर बाजार (-0.624) की तुलना में दिल्ली बाजार (-0.008) में समायोजन की गति धीमी है, जो दर्शाती है कि कानपुर, अल्पावधि के उतार-चढ़ाव दीर्घ अवधि के सन्तुलन के प्रतिकन्वर्जिंग है।



**तलिका-5:** सरसों के बीज, तेल एवं ऑयल केक के बीच वर्टिकल बाजार समाकलन हेतु जोहनसन अनुक्रमिक परीक्षण

स्थान	समाकलन	$H_0$	आइजेन value	CE(s) की संख्या	VAR का आर्डर
दिल्ली	बीज एवं तेल	$r = 0$	0.233	None *	1
		$r \leq 1$	0.029	अधिकतम 1	
	बीज, तेल एवं केक	$r = 0$	0.294	None *	1
		$r \leq 1$	0.055	अधिकतम 1	
कानपुर	बीज एवं तेल	$r = 0$	0.248	None *	1
		$r \leq 1$	0.040	अधिकतम 1	
	बीज, तेल एवं केक	$r = 0$	0.323	None *	2
		$r \leq 1$	0.091	अधिकतम 1	

\* 5 प्रतिशत सार्थकता स्तर पर परिकल्पना की अस्वीकृति को दर्शाता है ।

**तलिका-6 :** सरसों के बीज एवं तेल के लिए न्यूनीकृत फार्म VECM

स्थान	बीज	तेल
दिल्ली	-0.008	0.544
	(-0.082)	(4.355)
कानपुर	-0.623	-0.154
	(-4.520)	(-1.510)

(कोष्ठक में दिए गए आँकड़े t-स्टैटिस्टिक हैं)

हालांकि, सरसों के बीज एवं सरसों के तेल के बाजारों में स्पेशियल समाकलन की तरह वर्टिकल समाकलन में ऐसी ही प्रवृत्ति देखी गयी, वर्टिकल फ्रेमवर्क में, बीज एवं तेल सह-समाकलित हैं लेकिन ऑयल केक ने भिन्न व्यवहार दर्शाया।

स्पेशियल एवं वर्टिकल, दोनों फ्रेमवर्क में परिणामों ने बीज एवं तेल की मूल्य श्रृंखलाओं में सह-गतिविधि दर्शाया जो दीर्घ अवधि के सन्तुलन का संकेत है। इसके अतिरिक्त, अल्पावधि गतिशीलता भी सन्तुलन के प्रति अभिसरित है। ऑयल केक ने दो अन्य वस्तुओं के प्रति समाकलित न होने का भिन्न व्यवहार दर्शाया। यह इस तथ्य के कारण हो सकता है कि स्थान के आधार

पर बाजार-चलित तथा वैकल्पिक चारे की उपलब्धता के कारण ऑयल केक जैसे गौण-उत्पादों की कीमतों में बहुत अधिक उतार-चढ़ाव नहीं होता ।

### निष्कर्ष

सरसों के बीज एवं तेल के लिए चयनित बाजार समाकलित थे परन्तु कमजोर थे। कमजोर समाकलन बाजार अक्षमताओं के उच्च स्तर का पता चलता है। बाजार में अक्षमताओं को काबू करने के लिए उत्पादकताओं एवं मध्यस्थों को अपनी आपूर्ति चेन प्रबन्धन प्रक्रियाओं में सुधार करना चाहिए, जो उत्पादकों या उपभोक्ताओं अथवा दोनों की समृद्धि को कम करता है। तिलहन उत्पादन में तकनीकी अक्षमताएँ फार्म स्तर/प्रोसेसिंग यूनिट स्तर पर एक-चौथाई से एक-तिहाई तथा इससे भी अधिक पायी गयीं। तिलहन के अच्छे विनियमित बाजार, गुणवत्ता बीजों की सामयिक एवं सुनिश्चित आपूर्ति तथा परिवहन बाधाओं को दूर करने के साथ-साथ प्रोसेसिंग के लिए कच्चा माल तथा सकारात्मक नीति संकेत कुछ ऐसे महत्वपूर्ण कारक हैं जो भारत में सरसों के उद्योग के निष्पादन में सुधार कर सकते हैं।

## संदर्भ

1. एन्जल, आर.एफ एवं ग्रेंजर, सी. डब्ल्यू. जे. (1987): "कोइन्टीग्रेशन ऐण्ड एरर करेक्शन : रीप्रेसेन्टेशन, एस्टीमेशन ऐण्ड टेस्टिंग", इकनोमेट्रिका 55, 251-276 ।
2. ग्लोरिया, रिवेरा-गोनजेलेंज एवं स्टीवेन एम. हेल्प्पेण्ड (2001): "दि एक्स्टेन्ट, पैटर्न ऐण्ड डिग्री ऑफ़ मार्केट इन्टीग्रेशन : ए मल्टीवेरिएट एप्रोच फॉर दि ब्राजीलियन राइस मार्केट", अमेरिकन जर्नल ऑफ़ एग्रिकल्चरल इकनोमिक्स, 33, 3, 576-592 ।
3. गोलेटी, एफ., ए राएसुद्दीन एवं एन. फरीद (1995): "स्ट्रक्चरल डेट्रीमेन्ट्स ऑफ़ मार्केट इन्टीग्रेशन: दि केस ऑफ़ राइस मार्केट इन बंगलादेश", डिवेलिपिंग इकनोमिक्स, 33, 3, 185-202 ।
4. जोहनसेन, एस. (1991): "एस्टीमेशन ऐण्ड हाइपोथिसिस टेस्टिंग ऑफ़ कोइन्टीग्रेशन वेक्टर्स इन गाउसेन वेक्टर ऑटोरिग्रेसिव मॉडल्स", इकनोमेट्रिका, 59, 551-580 ।
5. जोहनसेन, एस. (1995): "लाइकलीहुड-बेस्ड इन्फ्रेंस इन कोइन्टीग्रेटेड वेक्टर-ऑटो -रिग्रेशनस", इन एडवान्स्ड टेक्स्ट्स इन इकनोमेट्रिक्स, ऑक्सफोर्ड, ऑक्सफोर्ड यूनिवर्सिटी प्रेस. ।
6. निगासा, ए., आर. मेयर्स एवं ई.जी. मालधीन (2003): "एनालाइजिंग दि मार्केट एफिशियेन्सी इन डिवेलिपिंग कन्ट्रीज : रिव्यू ऑफ़ एक्समैथड्स ऐण्ड एक्स्टेंशन्स टू दि पैरिटी बाउण्ड मॉडल, मार्केट ट्रेण्ड" इन्स्टीट्यूशनस डिवीजन, डिस्कशन पेपर : 63 ।

□

हिन्दी भाषा की शक्ति का प्रवाह बदला नहीं जा सकता ।

—मदन मोहन मालवीय

## उपग्रह छाया चित्र में बादल आवरण के अन्तर्गत फसल क्षेत्र आकलन की कार्यप्रणाली विकसित करने के लिये अध्ययन

प्राची मिश्रा साहू, तौकीर अहमद, कमलेश नारायण सिंह,  
अशोक कुमार गुप्ता एवं नीलम चन्द्रा

सुदूर संवेदन पद्धति द्वारा उपग्रह से प्राप्त आंकड़ों से फसल क्षेत्र का आकलन देश और विदेश में व्यापक रूप से किया जा रहा है। परन्तु यह खरीफ के मौसम में ही प्रचलित है। खरीफ के मौसम में अधिकतर बादल छाये रहते हैं। सुदूर संवेदन चित्रों द्वारा जानकारी एकत्रित करने के लिए उपग्रह चित्रों में बादल और इनकी छाया बहुत बड़ी बाधा है। भारत एक उष्णकटिबंधीय (tropical) देश है। अतः यहां काफी समय के लिये मानसून रहता है। मानसून के दौरान बादल रहित चित्र प्राप्त करना अत्यन्त कठिन है। बादलों की उपस्थिति के कारण हमें उपग्रह से आंकड़ें प्राप्त नहीं हो पाते हैं। इस स्थिति में विलुप्त आंकड़ों (missing data) की समस्या उत्पन्न होती है। देश के पहाड़ी क्षेत्र, मुख्य रूप से उत्तर पूर्वी पहाड़ी क्षेत्र भी अधिकतर बादलों से ढके रहते हैं। अतः इन क्षेत्रों के भी बादल रहित चित्र मिलना बहुत कठिन है क्योंकि यहां पर बादल रहित मौसम बहुत ही कम समय के लिए होता है। कई बार इतने कम समय सीमा में उपग्रह वहां से जा चुका होता है जिससे उस जगह के आंकड़ें प्राप्त नहीं हो पाते हैं। सामान्यतः मैदानी इलाकों में खरीफ की फसल क्षेत्र का आकलन कठिन है। विशेष रूप से उत्तर पूर्वी पहाड़ी क्षेत्र में खरीफ की फसल क्षेत्र का आकलन एक चुनौतीपूर्ण कार्य है।

उपरोक्त तथ्य ध्यान में रखते हुए इस शोध-पत्र में बादल और इसकी छाया को हटा कर बादल रहित चित्र उत्पन्न करने के लिए एक कार्यप्रणाली विकसित करने का प्रयास किया गया है। सुदूर संवेदन

तकनीक के आगमन से शोधकर्ताओं ने उपग्रह चित्रों में बादल आवरण और छाया की समस्या को दूर करने लिए विभिन्न तकनीकों का प्रयोग किया है। बादल आच्छादित चित्र के साथ बादल रहित चित्र का सम्मिश्रण प्रारंभिक दृष्टिकोण था लेकिन अब नई उन्नत प्रविधियों जैसे भू-सांख्यिकीय प्रविधियों (geo-statistical techniques) आदि के कारण बेहतर परिणाम निकाले जा रहे हैं। इसलिए यह अध्ययन उपग्रह चित्रों में बादल आवरण के अंतर्गत फसल क्षेत्र आकलन की एक कार्यप्रणाली विकसित करने के लिए किया गया। बादल और इसकी छाया को हटाने के लिए भू-सांख्यिकी प्रविधि प्रचलित की गयी और स्थानिक अध्यारोपण तकनीक (spatial imputation techniques) विकसित की गई। भू-सांख्यिकी और स्थानिक अध्यारोपण विधियों में सबसे कारगर विधि का चयन किया गया और जिला स्तर पर धान की फसल के अंतर्गत क्षेत्र का आकलन किया गया।

### अध्ययन क्षेत्र

यह अध्ययन मेघालय राज्य के रिबोई जिले में किया गया है। उत्तर पूर्वी राज्यों में मेघालय एक ऐसा राज्य है जो अधिकतर बादलों से घिरा रहता है। रिबोई जिला 4 जून, 1992 को उप प्रभागीय स्तर से पूर्ण जिला बनाया गया। नया जिला तत्कालीन पूर्वी खासी हिल्स से निकाला गया। रिबोई जिले की आबादी 2011 जनगणना के अनुसार 1,92,795 है और इसका क्षेत्रफल 2448 वर्ग किलोमीटर है। इस जिले का मुख्यालय नांगपोह है जो मेघालय की राजधानी शिलोंग से 53

किलोमीटर की दूरी पर है। रिबोई जिले के पथारकमाह में तीन सी और आर डी ब्लाक व एक प्रशासनिक एकक है। इस जिले के गांवों की संख्या 953 है। रिबोई जिले के जनसंख्या का घनत्व 79 प्रति वर्ग किलोमीटर है। यह जिला  $91^{\circ} 40' - 90^{\circ} 18'$  ई देशांतर और  $25^{\circ} 40' - 26^{\circ} 20'$  एन अक्षांश के बीच स्थित है। रिबोई जिला उत्तर में असम राज्य के कामरूप जिले से और पूर्व में जयंतिया हिल्स व कर्बी आंगलेंग जिले से एवं पश्चिम में खासी हिल्स जिले से घिरा है। रिबोई जिला बीहड़ और अनियमित भूमि की सतह द्वारा चिन्हित होता है। यह जिला पर्वत श्रेणियों की एक शृंखला है जिसका धीरे-धीरे उत्तर की तरफ ढलान है और अंततः ब्रह्मपुत्र घाटी में मिलता है। उमत्रेव, उमसियांग, उमरान और उमियम जैसी महत्वपूर्ण नदियां इस क्षेत्र से होकर बहती हैं।

### कार्यप्रणाली

भू-सांख्यिकी प्रविधि द्वारा बादल मुक्त चित्र की उत्पत्ति के लिए अपनायी गयी कार्यप्रणाली और बादल मुक्त चित्र की उत्पत्ति के लिए स्थानिक अध्यारोपण तकनीक विकसित करने की कार्यप्रणाली का इस खंड में वर्णन किया गया है। इस अध्ययन में सम्पूर्ण जिले का उपग्रह चित्र बहुत बड़ा है। अतः सरलता से गणना करने के लिए उपग्रह चित्र का एक छोटा भाग, जो एरिया आफ इन्टरेस्ट कहलाता है, लिया गया है और सभी आपरेशन चित्र के इस छोटे भाग पर किये गये हैं। इसके अतिरिक्त, अगर फसल क्षेत्र का आकलन उपग्रह आंकड़ों के उपयोग द्वारा किया जाता है तो तुलना के उद्देश्य से वास्तविक मूल्य उपलब्ध नहीं है। इसलिए वास्तविक मूल्य प्राप्त करने के लिये बादल मुक्त चित्र में यादृच्छिक रूप से बादल डाले गये। क्योंकि बादल आवरण और छाया एक यादृच्छिक घटना है, यह चित्र में कहीं भी दिखाई दे सकते हैं। प्रत्येक एरिया आफ इन्टरेस्ट में उत्पन्न बादल मुक्त चित्र का धान की फसल के तहत क्षेत्र के आकलन के लिये उपयोग किया गया। इनमें सबसे उचित प्रविधि की पहचान की गयी। इसके बाद जिला स्तर पर सबसे उचित प्रविधि का उपयोग कर फसल क्षेत्र का आकलन किया गया है।

### (i) भू-सांख्यिकी तकनीक के उपयोग द्वारा बादल मुक्त चित्र की उत्पत्ति

बादल मुक्त चित्र की उत्पत्ति के लिए निम्नलिखित चार विधियों का प्रयोग किया गया है –

1. क्रीगिंग
2. स्तरित क्रीगिंग
3. को-क्रीगिंग
4. स्तरित को-क्रीगिंग

### (ii) स्थानिक अध्यारोपण (Spatial Imputation)

अध्यारोपण एक सांख्यिकी प्रविधि है जो उपलब्ध जानकारी के आधार पर विलुप्त जानकारी का आकलन करती है। यहां औसत अध्यारोपण प्रविधि के अन्तर्गत उत्तरवर्ती (succeeding) या पूर्ववर्ती (preceding) इकाई या दोनों इकाईयों के औसत निकाले जाते हैं और इनको विलुप्त मूल्यों के स्थान पर रखा जाता है। इसके अतिरिक्त अनुपात अध्यारोपण और प्रतिगमन अध्यारोपण की विधि भी प्रचलित है। अतः सांख्यिकी अध्यारोपण प्रविधि के आधार पर विभिन्न स्थानिक अध्यारोपण प्रविधि विकसित करने का प्रयास किया गया है परन्तु स्थानिक आंकड़ों के लिए सामान्य सांख्यिकी अध्यारोपण प्रविधि लागू नहीं की जा सकती है। इस प्रकार विलुप्त आंकड़ों के लिए निम्नलिखित छः प्रविधियां विकसित की गई –

#### • पंक्तिवार पिक्सल पर आधारित स्थानिक अध्यारोपण (Spatial Imputation based on row-wise pixels)

इस प्रविधि में जिस पिक्सल का मूल्य विलुप्त है उसका अध्यारोपण पंक्तिवार पिक्सल द्वारा किया जाता है। इसी पंक्ति में विलुप्त मूल्य के ठीक पहले दो पिक्सल का औसत डी. एन. मूल्य निकाला जाता है। इस मूल्य को विलुप्त मूल्य के स्थान पर अध्यारोपित किया जाता है।

#### • स्तंभवार पिक्सल पर आधारित स्थानिक अध्यारोपण (Spatial Imputation based on column-wise pixels)

इस प्रविधि में जिस पिक्सल का मूल्य विलुप्त है उसका अध्यारोपण स्तंभवार पिक्सल द्वारा किया जाता है। उसी स्तम्भ में विलुप्त मूल्य के ठीक

पहले दो पिक्सल का औसत डी.एन. मूल्य निकाला जाता है। इस मूल्य को विलुप्त मूल्य के स्थान पर अध्यारोपित किया जाता है।

- **पंक्तिवार और स्तंभवार दोनों पिक्सल पर आधारित स्थानिक अध्यारोपण (Spatial Imputation based on both row-wise and column-wise pixels)**

इस प्रविधि में जिस पिक्सल का मूल्य विलुप्त है उसका अध्यारोपण पंक्तिवार और स्तंभवार दोनों पिक्सल द्वारा किया जाता है। इस विलुप्त पिक्सल के संलग्न पंक्ति और स्तम्भ में स्थित चारों मूल्यों का औसत निकाल कर इसको विलुप्त पिक्सल के स्थान पर अध्यारोपित किया जाता है।

- **निकटस्थ पिक्सल पर आधारित स्थानिक अध्यारोपण (Spatial Imputation based on neighbouring pixels)**

इस प्रविधि में जिस पिक्सल का मूल्य विलुप्त है उसका अध्यारोपण निकटस्थ पिक्सल द्वारा किया जाता है। इस तकनीक में विलुप्त पिक्सल की पहचान कर उसके चारों ओर के मूल्यों का औसत निकाला जाता है। इस औसत मूल्य को विलुप्त मूल्य के स्थान पर अध्यारोपित किया जाता है।

- **अनुपात दृष्टिकोण पिक्सल पर आधारित स्थानिक अध्यारोपण (Spatial imputation using ratio approach)**

इस प्रविधि में विलुप्त पिक्सल के मूल्य का अध्यारोपण करने के लिए अनुपात दृष्टिकोण लागू किया गया है। यहां नवम्बर, 1998 के चित्र को सहायक चर (auxiliary variable) के रूप में लिया गया है जो पहले का बादल रहित चित्र है। पिछले उपलब्ध आंकड़ों द्वारा चित्र में विलुप्त जानकारी की गणना के लिए अनुपात दृष्टिकोण का उपयोग किया गया है। यादृच्छिक संख्या जनरेटर एल्गोरिद्म (random number generator algorithm) द्वारा यादृच्छिक पचास अशून्य संख्या उत्पन्न की जाती है। दोनों चित्रों के लिए इन यादृच्छिक संख्याओं के पिक्सल मूल्य का औसत

निकाला जाता है। ये दोनों चित्र हैं – सितम्बर, 2002 जो वर्तमान वर्ष का बादल युक्त चित्र है और नवम्बर, 1998 जो पहले का बादल मुक्त चित्र है।

विलुप्त मूल्य पिक्सल के मूल्य की गणना निम्नलिखित प्रकार से की गयी है –

$$\hat{Y}_i = \frac{\bar{Y}}{\bar{X}} X_i$$

जहां

$\bar{Y}$  → इस साल के बादल युक्त चित्र के यादृच्छिक पचास चयनित पिक्सल के डी. एन. मूल्य का औसत

$\bar{X}$  → पिछले साल के बादल मुक्त चित्र के उसी (corresponding) 50 यादृच्छिक चयनित पिक्सल के डी. एन. मूल्य का औसत

$\hat{Y}_i$  → विलुप्त जानकारी वाले पिक्सल का अध्यारोपित डी. एन. मूल्य

$X_i$  → पिछले साल के बादल मुक्त चित्र के उसी पिक्सल का डी. एन. मूल्य

- **प्रतिगमन दृष्टिकोण पिक्सल पर आधारित स्थानिक अध्यारोपण (Spatial imputation using regression approach)**

इस पद्धति में विलुप्त मूल्यों की भविष्यवाणी के लिए प्रतिगमन समीकरण का उपयोग किया गया है। सहायक चर द्वारा अध्यारोपित मूल्य (imputed value) पहले से बताये जाते हैं। यहां नवम्बर, 1998 के चित्र को सहायक चर के रूप में लिया गया है जो पहले का बादल रहित चित्र है। पिछले उपलब्ध आंकड़ों द्वारा चित्र में विलुप्त जानकारी की गणना के लिए प्रतिगमन दृष्टिकोण का उपयोग किया गया है। यादृच्छिक संख्या जनरेटर एल्गोरिद्म द्वारा यादृच्छिक पचास अशून्य संख्या ली जाती है।

बादल युक्त पिक्सल के अध्यारोपित मूल्य निम्नलिखित प्रकार से निकाले गये हैं –

$$y_i = a + b \cdot x_i$$

जहां

$y_i \rightarrow$  शून्य बादल मूल्य (zero cloud value) के स्थान पर अध्यारोपित पिक्सल मूल्य

$x_i \rightarrow$  पिछले साल के बादल मुक्त आंकड़ों से प्राप्त उसी पिक्सल का डी एन मूल्य (DN value)

a और b स्थिरांक (constant) हैं जिसकी इस वर्ष और पिछले वर्ष के चित्रों से चयनित पिक्सल के डी एन मूल्य द्वारा गणना की जाती है ।

उपग्रह चित्र में विलुप्त जानकारी बादल और इसकी चित्र की उपस्थिति के कारण होती है। चित्र में तीन तरह के बैंड होते हैं। पहले व्यक्तिगत (individual) बैंड आंकड़े अलग किये जाते हैं। फिर हर एक बैंड में विलुप्त जानकारी को अध्यारोपित मूल्यों द्वारा स्थानित किया जाता है। इसके बाद मूल चित्र बनाने के लिए बैंड को संयुक्त किया जाता है जो स्थानिक अध्यारोपण के द्वारा बादल मुक्त चित्र बनाता है। इन आपरेशन के लिए सभी कोड मैट लैब R2010a (MATLAB R 2010a) में लिखा गया है।

### (iii) बादल मुक्त चित्रों की उत्पत्ति के लिए सबसे कुशल तकनीक की पहचान

उपरोक्त प्रविधि के द्वारा बादल मुक्त चित्र की उत्पत्ति के बाद बादल हटाने की सबसे कुशल तकनीक की पहचान की गयी है, जिसका उपयोग राज्य स्तर पर फसल क्षेत्र के आकलन के लिए किया गया। इन तकनीकों की तुलना करने के लिए निम्नलिखित गुणांक उपयोग में लाये गए –

#### 1. प्रतिशत सापेक्ष बायस (Percent relative bias)

प्रतिशत सापेक्ष बायस को निम्नलिखित प्रकार दिया गया है –

$$\text{प्रतिशत सापेक्ष बायस} = \left[ \frac{\text{आकलित मूल्य} - \text{वास्तविक मूल्य}}{\text{वास्तविक मूल्य}} \right] * 100$$

$$\% \text{ RB} = \frac{\text{Estimated Value} - \text{actual value}}{\text{actual value}} * 100$$

जहां आकलित मूल्य उत्पन्न बादल मुक्त चित्र

से आकलित फसल का क्षेत्र है और वास्तविक मूल्य, बादल मुक्त छवि से फसल का क्षेत्र है ।

#### 2. कापा गुणांक (Kappa coefficient)

कापा गुणांक ( $\hat{K}$ ) को निम्नलिखित प्रकार दिया गया है –

कापा गुणांक = (प्रेक्षित यथार्थता – संयोग समझौता) / (1- संयोग समझौता)

$$\hat{K} = \frac{\text{observed accuracy} - \text{change agreement}}{1 - \text{chance agreement}}$$

जहां प्रेक्षित यथार्थता (observed accuracy) त्रुटि आव्यूह (error matrix) में विकर्ण तत्वों (diagonal elements) द्वारा निर्धारित की जाती है और संयोग समझौता (chance agreement) उत्पन्न त्रुटि आव्यूह में बन्द विकर्ण तत्वों (off diagonal elements) द्वारा ।

बादल मुक्त चित्र की उत्पत्ति के लिए सबसे अच्छी प्रविधि कापा सांख्यिकी पर आधारित होती है और फसल के अन्तर्गत क्षेत्र का आकलन इस प्रविधि द्वारा किया जाता है।

#### 3. यथार्थ मूल्यांकन

चित्र के वर्गीकरण के बाद वर्गीकृत चित्र का यथार्थ मूल्यांकन किया गया। वर्गीकृत चित्र का यथार्थ मूल्यांकन एरडास इमैजिन साफटवेयर (ERDAS IMAGINE software) के द्वारा किया गया। इस प्रक्रिया में उपग्रह चित्र में यादृच्छिक अंक उत्पन्न किये गये। यथार्थता की गणना के लिए 100 यादृच्छिक अंक लिये गये। मूल्यांकन चार्ट में प्रत्येक यादृच्छिक बिन्दु के लिये संदर्भ मूल्यों को निविष्ट (input) किया गया जो चित्र की व्याख्या पुस्तिका (manual interpretation) पर आधारित थी।

यथार्थता की रिपोर्ट सही और गलत तरीके से वर्गीकृत पिक्सल की संख्या से पता चलती है जो कि हर वर्ग के लिए मिसक्लासिफाइड पिक्सल है। इसके अलावा यथार्थता रिपोर्ट में सही ढंग से वर्गीकृत पिक्सल का प्रतिशत दिखाया गया है। अतः यथार्थता आकलन रिपोर्ट से पता चलता है कि चित्र का वर्गीकरण कितना सही है।

#### 4. मिसक्लासिफाइड पिक्सल के कारण त्रुटि

यहां मिसक्लासिफाइड पिक्सल की संख्या द्वारा, संचालित वर्गीकरण में हुई त्रुटि को निर्धारित किया गया है। त्रुटि की गणना के लिए उपयोग किया गया सूत्र निम्नलिखित है –

$$\text{त्रुटि} = \frac{1}{n} \sum_{i=1}^n e_i^2$$

जहां n वर्ग की संख्या है, जिसमें वर्गीकरण किया गया है।

$e_i$ , प्रत्येक वर्ग में गलत ढंग से वर्गीकृत पिक्सल की संख्या है जो वर्गीकृत चित्र में यथार्थता मूल्यांकन के रूप में उत्पन्न त्रुटि आव्यूह से प्राप्त की जाती है।

#### 5. बादल के धब्बों के कारण त्रुटि का आकलन

सम्पूर्ण चित्र में बादल और इसकी छाया के पिक्सल का समूह है जिसको पैच या धब्बा कहते हैं। बादल मुक्त चित्र की उत्पत्ति के विभिन्न तरीकों की तुलना के लिए एवं सबसे अच्छी और सबसे कुशल विधि निकालने के हर तकनीक के लिए चित्र में पहचाने गये बादल के धब्बों के आधार पर त्रुटि की गणना की जाती है।  $j^{\text{th}}$  पैच के लिए त्रुटि की गणना निम्न प्रकार से की जाती है –

$$E_{pj} = \frac{1}{n_i} \sum_{i=1}^{n_i} (DN_p - DN_a)$$

जहां

$n_i$  →  $i^{\text{th}}$  पैच में पिक्सल की संख्या

$m$  → चित्र में बादल पैच की संख्या

$DN_p$  → भू-सांख्यिकी/स्थानिक अध्यारोपण तकनीक द्वारा उत्पन्न बादल मुक्त चित्र के पिक्सल का DN मूल्य है।

$DN_a$  → वास्तविक बादल मुक्त चित्र में समकक्ष पिक्सल का DN मूल्य

इस तरह प्रत्येक बादल पैच के लिए त्रुटि की गणना की जाती है और सूत्र के द्वारा उत्पन्न सभी बादल मुक्त चित्र के लिए औसत त्रुटि की गणना की

जाती है।

$t^{\text{th}}$  तकनीक के लिए त्रुटि निम्न प्रकार है –

$$E_t = \frac{1}{m} \sum_{j=1}^m E_{pj}$$

प्रत्येक बादल पैच मासकिंग के द्वारा नये चित्र की तरह उत्पन्न किये गये हैं जो कि एरिया आफ इन्टरेस्ट फाइल है।

#### परिणाम एवं निष्कर्ष

इस अध्ययन में बादल रहित चित्र की उत्पत्ति करने के लिए दस विधियों का उपयोग किया गया है जिसमें भू-सांख्यिकी तकनीकों एवं स्थानिक अध्यारोपण तकनीकों का समावेश किया गया है। बादल मुक्त चित्रों की उत्पत्ति के लिये विभिन्न स्थानिक अध्यारोपण तकनीकें विकसित की गयीं। प्रतिषट सापेक्ष पूर्वाग्रह, कापा गुणांक, मिसक्लासीफिकेशन के कारण त्रुटि और व्यक्तिगत (individual) धब्बे के कारण त्रुटि के आधार पर इन तकनीकों की तुलना की गयी व सबसे कुशल तकनीक का चयन किया गया। इन विधियों में प्रतिगमन दृष्टिकोण के उपयोग द्वारा स्थानिक अध्यारोपण सबसे अच्छा परिणाम देती है। अतः प्रतिगमन दृष्टिकोण के उपयोग द्वारा स्थानिक अध्यारोपण सबसे कारगर तकनीक सिद्ध हुई क्योंकि इस तकनीक में सबसे कम प्रतिशत सापेक्ष पूर्वाग्रह जो -2.235 है और 79% वर्गीकरण की उच्चतम यथार्थता है। गलत वर्गीकरण के कारण त्रुटि और बादल के धब्बों के कारण त्रुटि होने से पता चला कि कम संख्या वाले मिसक्लासीफाइड पिक्सल के साथ प्रतिगमन दृष्टिकोण के द्वारा भी स्थानिक अध्यारोपण सबसे कारगर विधि साबित हुई है। इसके उपरान्त स्तरित को-कीगिंग और को-कीगिंग तकनीकें भी उपयुक्त परिणाम देती हैं। सबसे कारगर विधि ज्ञात होने के उपरान्त इस का प्रयोग कर जिले का बादल रहित चित्र प्राप्त किया गया और सामान्य सुदूर संवेदन तकनीक से धान की फसल के अंतर्गत क्षेत्र का आकलन किया गया।

## संदर्भ

1. हेल्मर, ई. एच. एवं रिउफेन्वट, बी. (2005): “प्रतिगमन पेड़ और हिस्टोग्राम मैचिंग के साथ बादल मुक्त उपग्रह चित्र मोजैक”, फोटोग्रामेटिक इंजीनियरिंग एन्ड रिमोट सेंसिंग, 71, 1079–1089।
2. होन, एन. टी. एवं टैटेशी, आर. (2009): “नकली एएलओएस आंकड़ों के उपयोग द्वारा आप्टिकल चित्र का बादल हटाना”, द इन्टरनेशनल अरकाइव्स आफ फोटोग्रामेट्रिक, रिमोट सेंसिंग एन्ड स्पेशियल इनफार्मेशन साइंसेस 37 (बी 4), 379–384।
3. ली, एम., लिउ, एस. सी. एवं क्वा, एल. के. (2004): “बादल से घिरे हुए उपग्रह से बादल मुक्त और बादल की छाया से मुक्त चित्र मोजैक का स्वतः उत्पादन”, प्रोसीडिंग्स आफ इन्टरनेशनल जियोसाइंस एन्ड रिमोट सेंसिंग, 6, 3946–3948।
4. ली, वाइ. वाइ. एवं पारकर, ई. एल. (2008): “एक वायरलेस सेंसर नेटवर्क में विलुप्त आंकड़ों के साथ वर्गीकरण के लिए एक स्थानिक अस्थायी अध्यारोपण तकनीक”, प्रोसीडिंग्स आफ आई.ई.ई.ई. इन्टरनेशनल कानफ्रेंस आन इनटेलीजेंट रोबोट एन्ड सिस्टमस, नाइस, फ्रांस।
5. लिन, सी. एच., साई, पी. एच., लाई, के. एच. एवं चेन, जे. वाई. (2011): “सूचना क्लोनिंग के उपयोग द्वारा मल्टीटेम्पोरल उपग्रह चित्रों से बादल हटाना”, इन्सटीच्यूट आफ इलेक्ट्रिकल एन्ड इलेक्ट्रानिक्स इंजीनियर्स ट्रान्जैक्सनस आन जियोसाइंस एन्ड रिमोट सेंसिंग, 20(20), 1–10।
6. मेन्ग, क्यू., बार्डरस, बी. ई., सीजेवस्की, सी. जे. एवं मैडेन, एम. (2009): “बादल और बादल की छाया को हटाने के लिये निकटतम स्पेक्ट्रल फिट”, फोटोग्रामेट्रिक इंजीनियरिंग एन्ड रिमोट सेंसिंग, 75(5), 569–579।
7. साहू, टी. एवं पाठक, एस. (2009): “उपग्रह चित्र से स्वतः साहचर्य तंत्रिका नेटवर्क और स्टेशनरी तरंगिका रूपान्तरण के उपयोग द्वारा बादल हटाना”, इन्सटीच्यूट आफ इलेक्ट्रिकल एन्ड इलेक्ट्रानिक्स इंजीनियर्स, 978, 100–105।
8. स्टेन, एम. एल. (1999): “स्थानिक आंकड़ों का अन्तर्वेशन”, क्रीगिंग के कुछ सिद्धांत, स्प्रिंगर सीरीज इन स्टैटिस्टिक्स, स्प्रिंगर, न्यूयार्क।

□

निज भाषा उन्नति अहे, सब उन्नति को मूल

बिन निज भाषा ज्ञान के, मित्त न हिय को शूल।

—भारतेंदु हरिश्चंद्र



## एस.पी.एस.एस. का उपयोग कर मिश्रित मॉडल का विश्लेषण

सुकान्त दाश, सुशील कुमार सरकार, वैद्यनाथ मण्डल एवं देवेन्द्र कुमार

मिश्रित प्रभाव मॉडल, एक विशेष मॉडल का उल्लेख करता है जिसमें स्थिर एवं यादृच्छिक दोनों प्रभाव निहित हैं। स्थिर प्रभाव में चरों के संभावित मूल्य निश्चित हैं एवं यादृच्छिक प्रभाव उन चरों का उल्लेख करता है जिन में सम्भावित परिणामों को बदला जा सकता है। समष्टि को ध्यान में रखते हुए स्थिर प्रभावों को वे प्रभाव माने जा सकते हैं जिनके लिए समष्टि तत्व निश्चित है। यादृच्छ प्रभाव को ऐसा प्रभाव माना (सोचा) जा सकता है जिसके लिए समष्टि/तत्व बदल रहे हैं या बदले जा सकते हैं (यानी कि यादृच्छ चर)। उदाहरण के लिए यह कहा जा सकता है कि लिंग एक स्थिर प्रभाव चर है क्योंकि हम उस चर के मूल्यों (नर और मादा) को जानते हैं और वे मूल्य एक-दूसरे से स्वतंत्र हैं एवं वे (आमतौर पर) बदलते नहीं हैं। उच्च विद्यालय कक्षा को यादृच्छिक प्रभाव ले सकते हैं क्योंकि हम उन्हीं कुछ कक्षाओं का प्रतिदर्श ले सकते हैं जो कि अस्तित्व में हैं एवं उन विद्यार्थियों का उल्लेख नहीं होता जो प्रत्येक वर्ष उन कक्षाओं में आते हैं या जाते हैं।

यादृच्छिक प्रभाव कई प्रकार के हैं, जिसका एक और उदाहरण समूह अभिकल्पनाओं में देखा जा सकता है जहाँ उदाहरण के लिए छात्रों के उपलब्धि का मूल्यांकन कक्षाओं के भीतर समूह होते हैं और वे कक्षाएं स्कूलों के अंदर समूह होती हैं। कक्षाओं के अंतः-समूह के मूल्यांकन के लिए एक यादृच्छिक प्रभाव सहित श्रेणीबद्ध अभिकल्पना का और स्कूलों में समूह कक्षाओं हेतु दूसरा यादृच्छिक प्रभाव का उदाहरण होगा। समूह

आँकड़ें संरचना विभिन्न समूहों के बीच का संबंध होता है जैसे कि, एक ही कक्षा के सभी सदस्य समान माने जाते हैं और वे सभी सदस्य दूसरी कक्षा के सदस्यों से भिन्न हैं।

### रैखिक मिश्रित प्रभाव मॉडल

माना  $y = X\beta + \sum_{i=1}^m Z_i a_i + \varepsilon$  एक मॉडल है, जहाँ  $X$  एक ज्ञात  $n \times p$  आव्यूह है,  $Z_i$ 's ज्ञात  $n \times r_i$  पूर्ण रैंक आव्यूह हैं,  $\beta$  एक  $p \times 1$  अज्ञात प्राचलों का सदिश है,  $\varepsilon$  एक अज्ञात यादृच्छ सदिश इस प्रकार है कि  $E(\varepsilon) = 0$  तथा  $var(\varepsilon) = \sigma^2 I_n$  एवं  $a_i$ 's अज्ञात  $r_i \times 1$  यादृच्छ सदिश इस प्रकार है कि  $E(a_i) = 0$  तथा  $var(a_i) = \sigma_i^2 I_{r_i}$  इसके अतिरिक्त,  $var(a_i, a_j) = 0$  for  $i \neq j$ , के लिए जहाँ  $O, r_i \times r_j$  है एवं सभी प्रेक्षणों के लिए  $cov(a_i, \varepsilon) = 0$  जहाँ  $O, r_i \times n$  आव्यूह है। तब  $E(y) = X\beta$  एवं  $var(y) = \sum_{i=1}^m \sigma_i^2 Z_i Z_i' + \sigma^2 I_n$

उपरोक्त मॉडल के वर्णन के लिए उपनाम रैखिक मिश्रित मॉडल उपयुक्त प्रतीत होता है क्योंकि मॉडल में  $\beta$  में नियत प्राचलों के रैखिक कार्यों (फलन) एवं  $\varepsilon$  में यादृच्छ संख्याओं के रैखिक फलन का मिश्रण है।

**आँकड़ों की उत्पत्ति** (स्रोत: एचटीटीपी:// डब्ल्यूडब्ल्यूडब्ल्यू.यूएनटी.ईडीयू ([http:// www.unt.edu](http://www.unt.edu)))

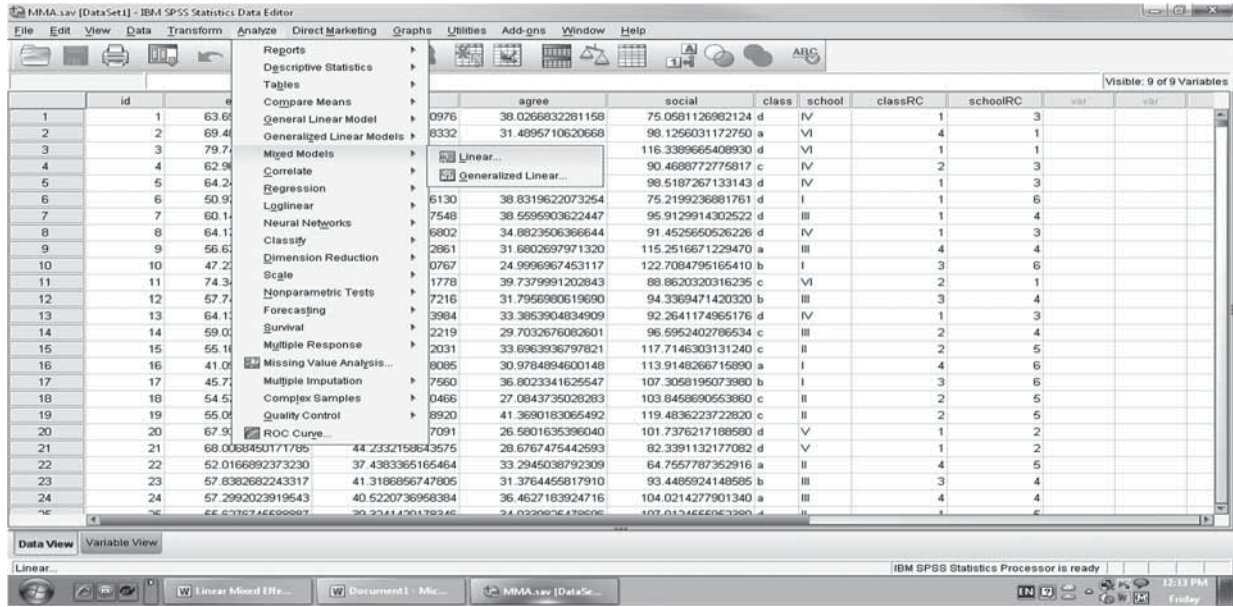
इस अध्ययन हेतु प्रयुक्त आँकड़े कृत्रिम हैं जहाँ अनुक्रिया चर अंतराल बहिर्मुखता (extro) जो कि

स्थिर प्रभाव का पूर्वानुमानित चर, नये अनुभवों के लिए जागरूकता (open), सहमतता (agree), सामाजिकता (social) और नामित श्रेणी (class RC) के साथ-साथ यादृच्छ चर जैसे कि class RC (within School RC) एवं स्कूल का यादृच्छ प्रभाव (school RC) निहित है।

इन आँकड़ों में 24 समूहों (6 स्कूलों के अंदर की 4 कक्षाएँ) में समान रूप से बंटे हुए 1200 मामले हैं।

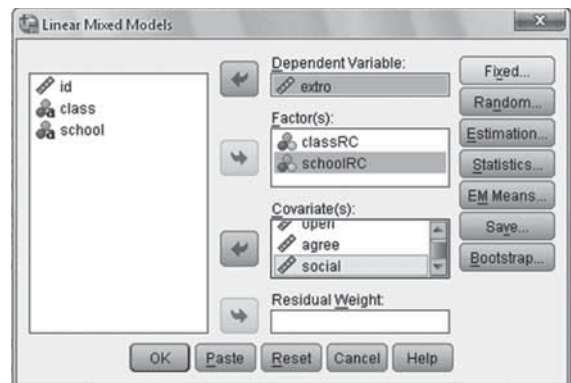
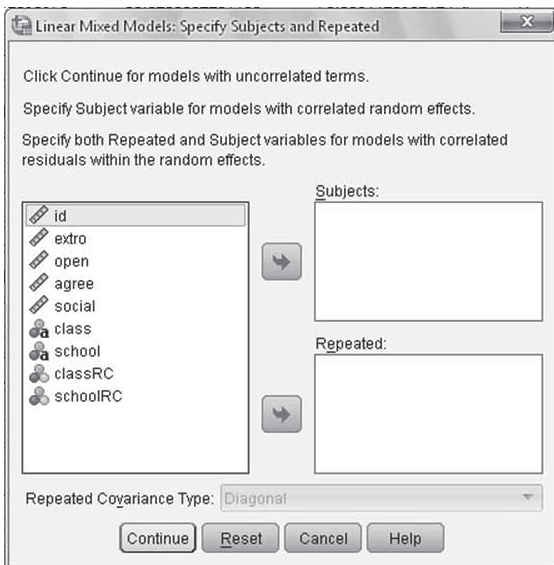
### विश्लेषण के चरण

Analyze → Mixed model → linear पर क्लिक कर के प्रारम्भ करें।

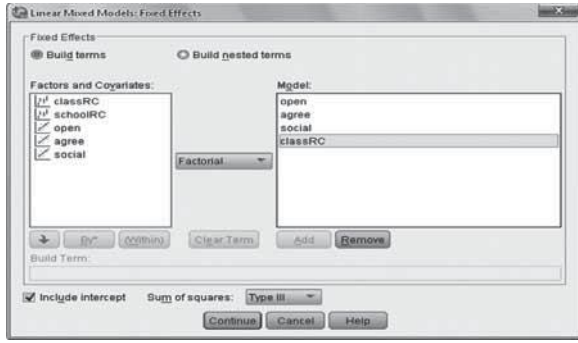


प्रारम्भिक संवाद बाक्स स्व-व्याख्यात्मक है लेकिन इस उदाहरण में प्रयुक्त नहीं होगा क्योंकि यह केवल आवर्ती मापों (repeated measure) के लिये प्रयुक्त होता है। अतः continue को क्लिक करें।

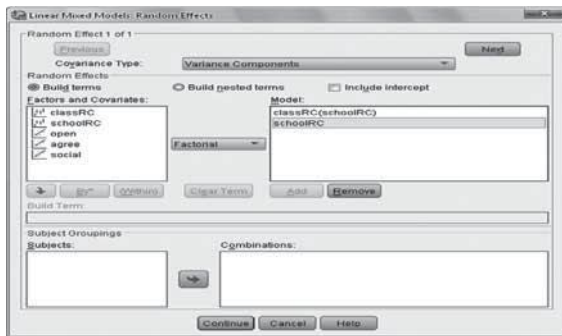
इसके बाद, मुख्य रैखिक मिश्रित मॉडल संवाद बाक्स है। मॉडल में जो चर हमें सम्मिलित करने हैं यहाँ स्पष्ट (निर्दिष्ट) करते हैं। निर्देशक (→) का उपयोग करते हुए extro को आश्रित चर क्षेत्र में भेजें, उसी प्रकार से class RC एवं school RC को factor क्षेत्र में एवं open, agree एवं social Covariate(s) क्षेत्र में क्रमानुसार भेजें। स्थिर प्रभाव को स्पष्ट करने के लिए Fixed बटन पर क्लिक करें।



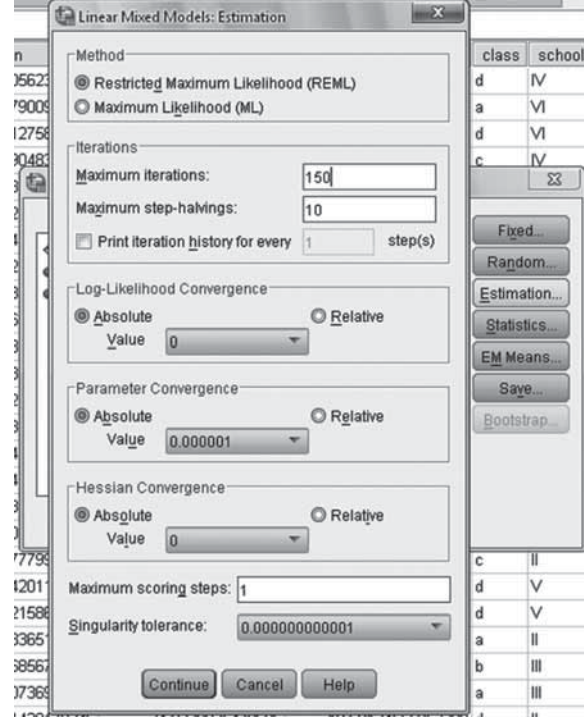
रैखिक मिश्रित प्रभाव मॉडल में स्थिर प्रभाव, पारंपरिक साधारण न्यूनतम वर्ग रेखीय होता है। स्थिर प्रभावों को स्पष्ट करने के लिए मॉडल बाक्स में open, agree, social एवं class RC तक पहुँचने के लिए Add बटन प्रयोग करें। ध्यान रहे हम इस मॉडल के लिए क्षेत्र में कोई भी अन्वय प्रभाव नहीं रहे। इसके उपरान्त continue बटन क्लिक करें।



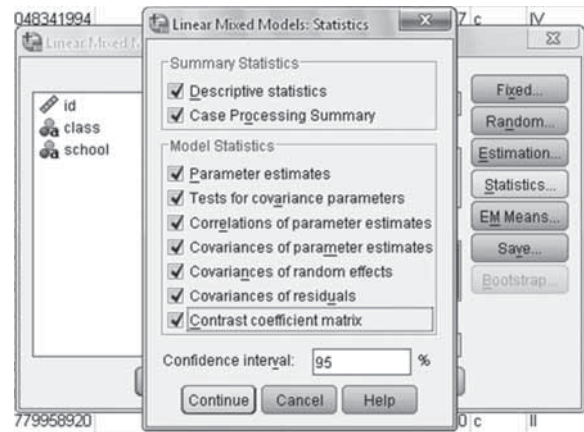
इसके बाद यादृच्छ प्रभाव को निर्दिष्ट करने के लिए Random पर क्लिक करे। सबसे पहले Build nested terms गोले पर क्लिक करे। इसके बाद classRC कारक को चिन्हांकित/चयन करें और Build terms बाक्स के classRC तक पहुँचने के लिए नीचे की ओर तीर बटन (↓) प्रयोग करें। इसके बाद (within) बटन क्लिक करें तथा School RC कारक चयन करें और नीचे की ओर तीर का पुनः उपयोग कर (within) बटन द्वारा बॉक्स के nested terms से बाहर आने के लिए Add बटन पर क्लिक करें। इसके बाद Build terms circle पर क्लिक करें। इसके बाद School RC कारक को चयन करें और मॉडल बाक्स के School RC तक भेजने के लिए Add बटन का प्रयोग करें। इसके उपरान्त संवाद बाक्स के सबसे नीचे के continue बटन पर क्लिक करें



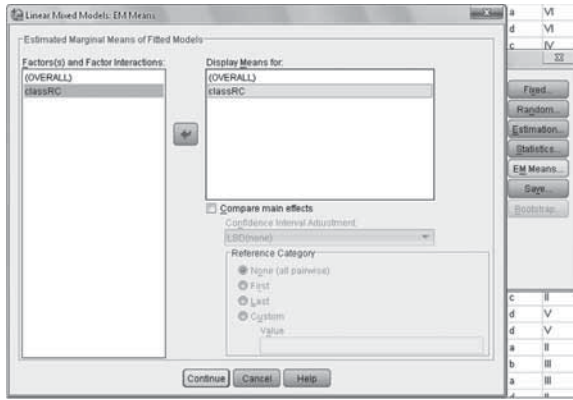
इसके आगे Estimation बटन पर क्लिक करें। इसके आगे अधिकतम पुनरावृत्ति को बकाया (100) से परिवर्तित करके 150 कर दें। यह कार्य तकनीकी रूप से आवश्यक नहीं है लेकिन, यह अनुमानित मूल्यों को निश्चित करता है। इसके उपरान्त continue बटन क्लिक करें।



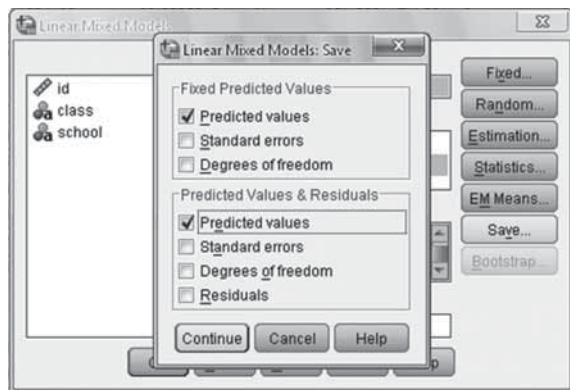
इसके आगे Statistics बटन पर क्लिक करें। हालांकि इसमें कुछ विकल्प आवश्यक नहीं है (Case Processing Summary), फिर भी सामान्यतया सभी को क्लिक कर दें।



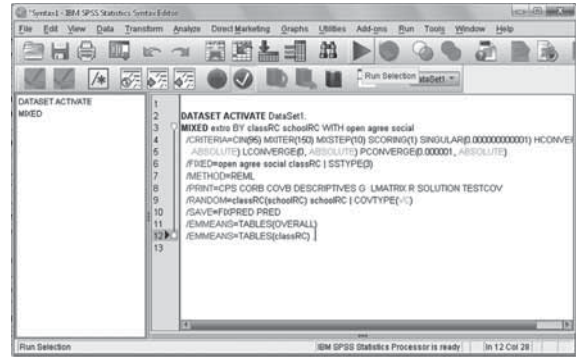
इसके आगे EM Means (Estimated Marginal Means) बटन पर क्लिक करें। जब (OVERALL) कारक Display Means for बाक्स में पहुंचता है तब समूह-औसत मिलेगा। classRC कारक यहां है (और Display Means for बाक्स में पहुंचता है) क्योंकि यह ही अकेला कारक (श्रेणीबद्ध चर) मॉडल में नियत प्रभाव के रूप में सम्मिलित है। दूसरे नियत प्रभाव श्रेणीबद्ध नहीं है और इसीलिए यहां दिखायी नहीं देते हैं। इसके आगे Continue बटन क्लिक करें।



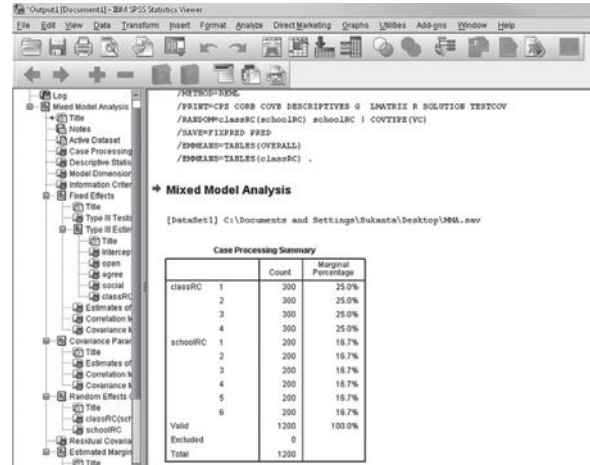
अब, save बटन क्लिक करें। सामान्यतया संभावित मूल्यों को सुरक्षित रखना अच्छा होता है। संभावित मूल्य जो कि संवाद बॉक्स के ऊपरी क्षेत्र में है वो संभावित मूल्य, स्थिर प्रभाव मॉडल का भाग है। संभावित मूल्य जो कि संवाद बॉक्स के नीचे के क्षेत्र में है वो संभावित मूल्य, पूर्ण प्रभाव मॉडल का संभावित मूल्य है। मॉडल के लिये नियत प्रभाव अंश पर पूर्ण आधारित नियत मूल्य होगा जबकि निम्न भविष्यवाणी मूल्य एवं अवशिष्ट भविष्यवाणी मूल्य पूर्ण मॉडल की भविष्यवाणी मूल्य होगा। इसके आगे continue बटन क्लिक करें।



फिर paste बटन क्लिक करें। प्रक्रिया क्रम नीचे दिया गया है। इस प्रक्रिया क्रम को उपयोग किया जा सकता है। इस प्रकार के मॉडलों में, इस प्रक्रिया के द्वारा आवश्यकतानुसार कोई भी रचनात्मक बदलाव किया जा सकता है। इसके आगे, प्रक्रिया क्रम में सारी विषयवस्तु को चयन करें और तब हरे रंग के 'रन' तीर को क्लिक करें।



### निष्कर्ष



उपरोक्त case processing summary यह दर्शाता है कि, किसी भी मामले को छोड़ा नहीं गया है। इसके आगे Descriptive Statistics की अपेक्षाकृत वृद्ध सारणी होती है। (नीचे केवल सारणी का उच्चतम भाग दिखाया गया है)

classRC	schoolRC		Count	Mean	Standard Deviation	Coefficient of Variation
1	1	extro	50	80.37807752	2.722230722	3.4%
		open	50	40.39493481	5.991433041	14.8%
		agree	50	36.24900085	6.167163053	17.0%
		social	50	101.3473505	17.62491141	17.4%
2	2	extro	50	68.35099481	.4383439954	.6%
		open	50	40.23107958	6.131491414	15.2%
		agree	50	34.42929885	5.656469706	16.4%
3	3	extro	50	98.41410772	15.47400680	15.7%
		open	50	63.78270073	.2813766361	.4%
		agree	50	40.44893006	5.181164905	12.8%
4	4	extro	50	35.51210349	4.703413381	13.2%
		open	50	100.3244996	14.29211931	14.2%
		agree	50	59.72533792	.2598249514	.4%
5	5	extro	50	38.35368536	5.697671139	14.9%
		open	50	36.48275880	4.752899366	13.0%
		agree	50	102.0887860	15.20937598	14.9%
6	6	extro	50	55.81179206	.2423683271	.4%
		open	50	40.25274418	5.700580131	14.2%
		agree	50	35.81782976	4.196563651	11.7%
Total	Total	extro	50	100.2169155	14.97269969	14.9%
		open	50	50.38830649	.6394010027	1.3%
		agree	50	40.10176951	4.745794123	11.8%
Total	Total	extro	50	34.53763317	4.450363639	12.9%
		open	50	95.47825880	18.32428567	19.2%
		agree	50	63.07286826	9.683215216	15.4%

नीचे दिये गये Model Dimension सारणी इस मॉडल के संबंध में है जिसमें चर (और उनके स्तरों की संख्या) स्थिर है और/या यादृच्छ प्रभाव है और प्राचल का आंकलन किया जा रहा है।

		Number of Levels	Covariance Structure	Number of Parameters
Fixed Effects	Intercept	1	Variance Components	1
	open	1		1
	agree	1		1
	social	1		1
	classRC	4		3
Random Effects	classRC(schoolRC) + schoolRC <sup>a</sup>	30		2
Residual				1
Total		38		10

a. As of version 11.5, the syntax rules for the RANDOM subcommand have changed. Your command syntax may yield results that differ from those produced by prior versions. If you are using version 11 syntax, please consult the current syntax reference guide for more information.  
b. Dependent Variable: extro.

आगे की सारणी योग्य सूचक को प्रदर्शित करती है। प्रत्येक सूचकांक हेतु संख्या जितनी कम है, आँकड़ों के लिये मॉडल उतना ही उपयोगी है।

-2 Restricted Log Likelihood	3528.106
Akaike's Information Criterion (AIC)	3534.106
Hurvich and Tsai's Criterion (AICC)	3534.126
Bozdogan's Criterion (CAIC)	3552.359
Schwarz's Bayesian Criterion (BIC)	3549.359

The information criteria are displayed in smaller-is-better forms.

a. Dependent Variable: extro.

आगे की सारणी में स्थिर प्रभाव परीक्षण के परिणाम दिये गये हैं। यहां पर हम intercept एवं class RC चरों का मुख्य योगदान प्रतीत होता हुआ देख सकते हैं।

### नियत प्रभाव

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	5.070	225.475	.000
open	1	1173.270	1.525	.217
agree	1	1173.203	1.843	.175
social	1	1173.359	.082	.774
classRC	3	15.001	11.997	.000

a. Dependent Variable: extro.

अगली 5 सारणियों से कोई बड़ी जानकारी नहीं मिलती है और केवल प्रत्येक प्राचल आकलन दर्शाती है (पांच में से केवल प्रथम और दूसरी सारणी का हिस्सा नीचे दर्शाया गया है)।

### टाइप III आकलन योग्य फलन

	L1
Fixed Effects Intercept	1
open	0
agree	0
social	0
[classRC=1]	.250
[classRC=2]	.250
[classRC=3]	.250
[classRC=4]	.250

	L2
Fixed Effects Intercept	0
open	1
agree	0
social	0
[classRC=1]	0
[classRC=2]	0
[classRC=3]	0
[classRC=4]	0

इसके आगे की "Estimates of Fixed Effects" सारणी (नीचे) बहुत महत्वपूर्ण है और मॉडल में उल्लेखित नियम प्रभावों का प्राचल आंकलन दर्शाती है। यह निश्चित होना चाहिए कि यह सारणी और इसकी व्याख्या ठीक वैसी ही है जैसी कि, पारंपरिक साधारण न्यूनतम वर्ग रेखीय समाश्रयण से अपेक्षा की जाती है। स्थिर चर हेतु SPSS किस प्रकार संदर्भ श्रेणी चुनता है यह ध्यान देने योग्य है। इस अध्ययन

उपरोक्त सारणी से यह निष्कर्ष निकाला जा सकता है कि जागरूकता (OPEN) कारक में एक इकाई की वृद्धि से बहिर्मुखता (EXTRO) में 0.006 की वृद्धि होती है। इसी प्रकार सहमतता (agree) कारक में एक इकाई वृद्धि से बहिर्मुखता (EXTRO) में 0.008 की कमी होती है। आगे आकलन में Class RC=3, का गुणांक 2.054798 है जिसका अर्थ है कि Class RC (3) के तृतीय समूह का बहिर्मुखता प्राप्तांक Class

**Estimates of Fixed Effects<sup>b</sup>**

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	57.383879	4.056589	5.299	14.146	.000	47.130810	67.636947
open	.006130	.004963	1173.270	1.235	.217	-.003608	.015868
agree	-.007736	.005698	1173.203	-1.358	.175	-.018916	.003444
social	.000531	.001852	1173.359	.287	.774	-.003103	.004165
[classRC=1]	5.665733	.983721	15.001	5.759	.000	3.568999	7.762468
[classRC=2]	3.704930	.983709	15.001	3.766	.002	1.608212	5.801648
[classRC=3]	2.054798	.983726	15.002	2.089	.054	-.041945	4.151541
[classRC=4]	0 <sup>a</sup>	0	.	.	.	.	.

a. This parameter is set to zero because it is redundant.

b. Dependent Variable: extro.

में आँकड़ों के समूह की class एवं school कारक की जगह Class RC एवं School RC का उपयोग किया गया है। यह जानना महत्वपूर्ण है कि SPSS चरों को क्रमबद्ध करता है। उदाहरणार्थ, यदि यही मॉडल मूल चर के साथ निर्दिष्ट किया जाता है तो स्थिर प्रभाव इण्टरसेप्ट 63.04 है। इस मॉडल का इण्टरसेप्ट 57.38 है जो संभावित मूल्य से भिन्न है।

RC (4) के अंतिम समूह के माध्य बहिर्मुखता प्राप्तांक 2.0547978 गुना अधिक है। क्योंकि Class RC (4) को संदर्भ श्रेणी के रूप में कूट किया गया है।

आगे की 2 तालिकाओं में स्थिर प्रभाव के आकलन के लिये सहसंबंध आव्यूह एवं सह-प्रसरण आव्यूह को दर्शाया गया है। इसमें बहुसंख्यता कोई मुद्दा नहीं है, क्योंकि उनके सहसंबंध (और सह-प्रसरण) काफी कम है।

**Correlation Matrix for Estimates of Fixed Effects<sup>b</sup>**

Parameter	Intercept	open	agree	social	[classRC=1]	[classRC=2]	[classRC=3]	[classRC=4]
Intercept	1	-.048	-.047	-.045	-.121	-.121	-.121	. <sup>a</sup>
open	-.048	1	-.012	-.006	.000	-.001	-.002	. <sup>a</sup>
agree	-.047	-.012	1	-.009	-.007	-.005	-.006	. <sup>a</sup>
social	-.045	-.006	-.009	1	.002	.001	.005	. <sup>a</sup>
[classRC=1]	-.121	.000	-.007	.002	1	.500	.500	. <sup>a</sup>
[classRC=2]	-.121	-.001	-.005	.001	.500	1	.500	. <sup>a</sup>
[classRC=3]	-.121	-.002	-.006	.005	.500	.500	1	. <sup>a</sup>
[classRC=4]	. <sup>a</sup>	. <sup>a</sup>	. <sup>a</sup>	. <sup>a</sup>	. <sup>a</sup>	. <sup>a</sup>	. <sup>a</sup>	. <sup>a</sup>

Covariance Matrix for Estimates of Fixed Effects<sup>b</sup>

Parameter	Intercept	open	agree	social	[classRC=1]	[classRC=2]	[classRC=3]	[classRC=4]
Intercept	16.455911	-.000967	-.001091	-.000339	-.482732	-.482820	-.483239	0 <sup>a</sup>
open	-.000967	2.463507E-5	-3.258247E-7	-5.116589E-8	-1.094142E-6	-2.505694E-6	-1.107368E-5	0 <sup>a</sup>
agree	-.001091	-3.258247E-7	3.247276E-5	-9.556020E-8	-3.941339E-5	-2.914959E-5	-3.195671E-5	0 <sup>a</sup>
social	-.000339	-5.116589E-8	-9.556020E-8	3.431005E-6	2.975101E-6	9.112723E-7	9.444991E-6	0 <sup>a</sup>
[classRC=1]	-.482732	-1.094142E-6	-3.941339E-5	2.975101E-6	.967706	.483864	.483875	0 <sup>a</sup>
[classRC=2]	-.482820	-2.505694E-6	-2.914959E-5	9.112723E-7	.483864	.967682	.483860	0 <sup>a</sup>
[classRC=3]	-.483239	-1.107368E-5	-3.195671E-5	9.444991E-6	.483875	.483860	.967718	0 <sup>a</sup>
[classRC=4]	0 <sup>a</sup>	0 <sup>a</sup>	0 <sup>a</sup>	0 <sup>a</sup>	0 <sup>a</sup>	0 <sup>a</sup>	0 <sup>a</sup>	0 <sup>a</sup>

a. The covariance is set to zero because it is associated with a redundant parameter.

b. Dependent Variable: extro.

यदि यादृच्छिक प्रभाव का राशि के प्रसरण में कोई सार्थक योगदान नहीं हो (यदि अवशिष्ट प्रसरण यादृच्छिक प्रभाव प्रसरण आकलन से अधिक है) तो यादृच्छिक प्रभाव को मॉडल में बाहर कर देना चाहिये और एक मानक सामान्य रेखीय मॉडल (या सामान्यीकृत रेखीय मॉडल) लगाना चाहिये।

### संदर्भ

1. रेन्शर, ए.सी. एवं शाल्ज़े, जी.बी. (2008): "लीनियर मॉडल्स इन स्टैटिस्टिक्स", सेकण्ड एडिशन, जॉन विले ऐन्ड संस, इंक, होबोकेन, न्यू जर्सी।
2. एचटीटीपी://डब्ल्यूडब्ल्यूडब्ल्यू.यूएनटी.ईडीयू (http://www.unt.edu)

□

हिन्दी भारत की अमर वाणी है।

—माखनलाल चतुर्वेदी

## अपेक्षित उत्पादन के लिए प्रक्रिया चर के साथ मिश्रण प्रयोगों में अनुकूलतम स्थितियाँ

उपेन्द्र कुमार प्रधान, कृष्ण लाल, विनोद कुमार गुप्ता,  
प्रबिण कुमार मेहेर एवं विजय पाल सिंह

साहित्य में प्रक्रिया चर के साथ मिश्रण प्रयोग के विश्लेषण की चर्चा की गयी है। सामान्यतः प्रयोगकर्ता की पैरामीटर के ऑकलन के साथ साथ उत्पादन की अनुकूलता में रुचि होती है। यदि प्रयोगकर्ता दो से अधिक चर उत्पादन रखता है तो दोनों परिवर्तनशीलता की उत्पादन की तकनीक का प्रयोग किया जा सकता है। इस अध्ययन में प्रक्रिया चर के साथ मिश्रण प्रयोग के लिए दोनों उत्पादन तकनीकों का विकास किया गया है। इस तकनीक की न्यूनतम मानक विचलन के साथ उत्पादन का अधिक माध्य करने के लिए एक डाटा सैट के साथ भी व्याख्या की गयी है।

एक मिश्रण प्रयोग वह प्रयोग है जिसमें उत्पादन मिश्रण में मौजूद इनग्रिडियन्स के समबंधित अनुपात पर निर्भर करती हैं ना कि मिश्रण की मात्रा पर। प्रक्रिया (Process) चर मिश्रण प्रयोग का वह कारक हैं जो कि मिश्रण का कोई भाग नहीं बनाते पर जिनका स्तर (level) अगर बदला जाए तो इनग्रिडियन की सम्मिश्रण गुणों को प्रभावित कर सकते हैं। जब मिश्रण प्रयोग प्रक्रिया चर के साथ किये जाते हैं तब उन प्रयोगों को प्रक्रिया चर के साथ मिश्रण प्रयोग कहा जाता है। इन प्रयोगों में प्रयोगकर्ता की रुचि न सिर्फ मिश्रण घटक की सम्मिश्रण गुणों का अध्ययन करना है बल्कि प्रक्रिया चर के स्तर के बदलाव के साथ सम्मिश्रण के व्यवहार को देखना भी है। यहाँ प्रक्रिया चर के साथ मिश्रण प्रयोग की प्रयोगात्मक स्थिति दर्शायी गयी हैं।

### प्रयोगात्मक स्थिति -I

एक प्रयोग कृषि क्षेत्र प्रयोग सूचना प्रणाली

(AFEIS) के तहत धान की फसल पर किया गया था। इसका उद्देश्य नाइट्रोजन की निश्चित मात्रा (30 kg.) की विभाजित मात्रा के प्रभाव का अध्ययन करना था। यूरिया के दो स्रोत, प्रिल्ड यूरिया (S1) और नीम कोटेड यूरिया (S2) से नाइट्रोजन प्राप्त होती हैं ट्रीटमेंट के विवरण तालिका-1 में दिये गये हैं।

**तालिका-1:** नाइट्रोजन का फसल के विभिन्न फसल वृद्धि स्टेज का उपयोग

ट्रीटमेंट	फसल वृद्धि स्टेज			यूरिया के स्रोत	उपज
	प्रारम्भिक स्टेज	अधिकतम जुताई स्टेज	पुष्प गुच्छ प्रारंभ		
T <sub>1</sub>	1	0	0	S <sub>1</sub>	2514
T <sub>2</sub>	1/2	1/2	0	S <sub>1</sub>	2528
T <sub>3</sub>	1/3	1/3	1/3	S <sub>1</sub>	2453
T <sub>4</sub>	1/4	1/2	1/4	S <sub>1</sub>	2587
T <sub>5</sub>	1/2	1/4	1/4	S <sub>1</sub>	2631
T <sub>1</sub>	1	0	0	S <sub>2</sub>	2556
T <sub>2</sub>	1/2	1/2	0	S <sub>2</sub>	2558
T <sub>3</sub>	1/3	1/3	1/3	S <sub>2</sub>	2452
T <sub>4</sub>	1/4	1/2	1/4	S <sub>2</sub>	2604
T <sub>5</sub>	1/2	1/4	1/4	S <sub>2</sub>	2707

**नोट :** इस प्रयोग के आँकड़े कृषि क्षेत्र प्रयोग सूचना प्रणाली के तहत किये गये प्रयोगों में से हैं।

प्रयोगकर्ता की रुचि तीनों स्टेजों पर प्रयुक्त नाइट्रोजन के उपयोग के अनुपाती संयोजन के साथ-साथ यूरिया के उस स्रोत का अध्ययन करना है



जिससे धान की उपज अधिकतम हो। उपरोक्त प्रयोग में मिश्रण नाइट्रोजन की तीन विभाजित मात्राओं से मिलकर बना है। क्योंकि कुल मात्रा (30 kg.) सभी ट्रीटमेंट में समान है अतः प्रयोग की उत्पादन विभाजित मात्राओं के अनुपात पर निर्भर करेगी। प्रिंल्ड यूरिया (S1) तथा नीम कोटेड यूरिया (S2) वह दूसरे कारक के स्तर हैं जो कि मिश्रण के घटक नहीं बनाते लेकिन धान की उपज को प्रभावित कर सकते हैं। इस प्रयोग को प्रक्रिया चर के साथ मिश्रण प्रयोग कहा जाता है।

उत्पादन सतह अभिकल्पना में एक प्रतिकृति की स्थिति में उत्पादन के अनुकूलतम बनाने वाली एक तकनीक का अध्ययन किया गया है। कृषि क्षेत्र या उद्योग में एक प्रयोगकर्ता का उद्देश्य वह आपरेटिंग परिस्थितियाँ प्राप्त करना है जो न्यूनतम प्रक्रिया परिवर्तनशीलता के साथ अपेक्षित उत्पादन के लिए वांछित लक्ष्य को प्राप्त करें। विनिंग तथा मायर्स (1990) ने दोहरी सतह उत्पादन को तागुची (1959, 1987) की मजबूत पैरामीटर अभिकल्पना को एक विकल्प के रूप में सुझाया। एनकेनमैन तथा डीन (2003) ने तागुची के मजबूत अभिकल्पना और दोहरी उत्पादन सतह अनुकूलतम पर उत्कृष्ट समीक्षा की। अग्रवाल इत्यादि (2006) ने डेल कैस्टलो इत्यादि (1997, 1999) द्वारा दी गई ग्लोबल दोहरी उत्पादन सतह अनुकूलतम तकनीक का प्रयोग अभिकल्पना चर के एक सेट जिसमें गुणात्मक तथा मात्रात्मक कारक शामिल हो, के लिए अनुकूल सेटिंग प्राप्त की है।

प्रक्रिया चर के साथ मिश्रण प्रयोग के अनुकूल संयोजन का निर्धारण कृषि के क्षेत्र में उपज के रूप के उत्पादन की वांछनीय विशेषताओं के संबंध में सबसे अच्छा आपरेटिंग संरचना को प्राप्त करने के शोध को दर्शाता है। इस शोध पत्र में प्रक्रिया चर जिसमें उत्पादन की संख्या दो या अधिक होने पर परिवर्तनशीलता कम से कम हो, के साथ मिश्रण प्रयोग में अनुकूलतम के लिए दोहरी उत्पादन कार्य पद्धति के उपयोग से संबंधित है। अग्रवाल (2006) द्वारा दी गई ग्लोबल दोहरी उत्पादन सतह अनुकूलतम की कार्य पद्धति जो मात्रात्मक तथा गुणात्मक कारकों को शामिल करती है। उसे प्रक्रिया चर के साथ मिश्रण प्रयोगों के लिए

सुधारा गया है।

## प्रक्रिया चर के साथ मिश्रण प्रयोगों की अनुकूलतम स्थितियाँ प्राप्त करने के लिए कार्य पद्धति

यह संभाग प्रक्रिया चर के साथ मिश्रण प्रयोग के अनुकूलतम से संबंधित रखता है। यहाँ डेल केसटीलो इत्यादि (1997) और अग्रवाल इत्यादि (2006) द्वारा दी गयी दोहरी उत्पादन सतह तकनीको का प्रयोग एक प्रक्रिया चर के साथ मिश्रण प्रयोग के अनुकूलतम प्रवर्तन स्थितियों को प्राप्त करने के लिए किया गया है। दोहरी उत्पादन तकनीक के प्रयोग के लिए हमें न्यूनतम परिवर्तनशीलता के साथ दो या उससे अधिक प्रतिकृति की आवश्यकता है।

माना एक प्रयोग में यहाँ उत्पादन मॉडल के लिए N उपयुक्त अभिकल्पना बिन्दु है। और ये N बिन्दु  $r \geq 2$  बार प्रतिकृति बनाते हैं। यह भी मानते हैं।  $y_m$  तथा  $y_s$  क्रमशः इन अभिकल्पना बिन्दु का नमूना मध्य और नमूना मानक विचलन है। निम्नलिखित पद्धति ग्लोबल अनुकूलतम उपाय प्राप्त करने के लिए अपनायी गई है।

एक प्रक्रिया चर के साथ q घटकों के लिए द्वितीय आदेश (quadratic) मिश्रण मॉडल का कैनालिनकल प्रारूप निम्न है।

$$E(y, z) = \sum_{i=1}^q \beta_i x_i + \sum_{i=1}^{q-1} \sum_{i < j}^q \beta_{ij} x_i x_j + \delta_0 z + \sum_{i=1}^q \delta_i x_i z \quad (1)$$

जहाँ  $\delta_0$  गुणात्मक कारक z के कारण का प्रभाव है।  $\beta_i$  प्रत्यावर्तन गुणांक है।  $\delta_i$  गुणात्मक कारक z तथा मात्रात्मक कारक  $i^{\text{th}}$  के बीच पारस्परिक प्रभाव है।  $x_i$   $i^{\text{th}}$  मात्रात्मक कारकों के मान है।  $\beta_{ij}$ ,  $i^{\text{th}}$  एवं  $j^{\text{th}}$  मात्रात्मक कारकों के संयुक्त प्रभाव है। विस्तृत जानकारी के लिए कार्नेल (2002) का उल्लेख कर सकते हैं यहाँ यह माना गया है कि प्रक्रिया चर के मिश्रण घटकों के उत्पादन अवधि के साथ कोई पारस्परिक प्रभाव नहीं है।

मॉडल (1) का कैनालिनकल प्रारूप प्रतिबन्ध  $\sum_{i=1}^q x_i = 1$  का प्रयोग कर निम्नलिखित रूप में रिपैरामेटराईज किया है:—

$$E(y, z) = \beta_0 + \sum_{i=1}^{q-1} \beta_i x_i + \sum_{i=1}^{q-1} \beta_{ii} x_i^2 + \sum_{i=1}^{q-2} \sum_{j=i+1}^{q-1} \beta_{ij} x_i x_j + \delta_0 z + \sum_{i=1}^{q-1} \delta_i x_i z \quad (2)$$

मिश्रण प्रयोग के इस मॉडल को (q-1) कारक स्थान पर अन्वेषण किया गया।

यहाँ हम दोहरी उत्पादन अनुकूलतम तकनीक का प्रयोग करते हैं जो मानक विचलन के न्यूनतम के साथ वांछित उत्पादन के लिए अनुपात के अनुकूलतम इष्टतम संयोजन को प्राप्त करने में मदद करता है।

मान लेते हैं कि N अभिकल्पना बिन्दु कम से कम दो बार प्रतिकृति है (i.e.  $r \geq 2$ ) मैट्रिक्स नोटेशन में मॉडल (2) को इस प्रकार लिख सकते हैं।

$$y(\mathbf{x}, z) = \mathbf{a} + \mathbf{x}'\mathbf{b} + \mathbf{x}'\mathbf{B}\mathbf{x} + \delta z + \mathbf{x}'\mathbf{c}z + \mu \quad (3)$$

यहाँ y अवलोकन का कालम वैक्टर है। x घटक (q-1) के एक कालम वैक्टर है। Z एक प्रक्रिया चर है। b पैरामीटर के वैक्टर है। B अज्ञात (q-1) × (q-1) आर्डर की मैट्रिक्स है, C (q-1) आर्डर का कालम वैक्टर है। तथा ε त्रुटि के कालम वैक्टर है जो  $N(0, \sigma^2 I)$  को अनुकरण करता है। जहाँ I आइडेन्टिटी मैट्रिक्स है।

यह मानते हैं कि  $y_m$  तथा  $y_s$  इन अभिकल्पना बिन्दुओं के नमूना मध्य तथा नमूना मानक विचलन के वैक्टर को दर्शाता है। पहले हम माध्यम  $y_m$  तथा मानक विचलन  $y_s$  के लिए एक द्वितीय आर्डर उत्पादन चर की तरह फिट करते हैं। उत्पादन चर के मध्य तथा मानक विचलन के लिए मॉडल है:

$$E[y_m(\mathbf{x}, z)] = \mathbf{a}_m + \mathbf{x}'\mathbf{b}_m + \mathbf{x}'\mathbf{B}_m\mathbf{x} + \delta_m z + \mathbf{x}'\mathbf{c}_m z \quad (4)$$

$$E[y_s(\mathbf{x}, z)] = \mathbf{a}_s + \mathbf{x}'\mathbf{b}_s + \mathbf{x}'\mathbf{B}_s\mathbf{x} + \delta_s z + \mathbf{x}'\mathbf{c}_s z \quad (5)$$

हम गुणात्मक कारक z ( $z = +1, z = -1$ ), के प्रत्येक स्तर के लिए उत्पादन मध्य तथा इसके नमूना मानक विचलन की गणना कर सकते हैं।

अब दोहरी अनुकूलतम समाधान प्राप्त करने के लिए जो प्रक्रिया अपनाई गयी है वह x का मान इस तरह प्राप्त करती है:

1.  $\hat{y}_s(\mathbf{x}, z)$  न्यूनतम/लक्ष्य है।
2.  $\hat{y}_m(\mathbf{x}, z)$  अधिकतम/लक्ष्य है।
3. अनुकूलतम समाधान वह x है जिसके लिए  $\hat{y}_s(\mathbf{x}, z)$  न्यूनतम है।

लक्ष्य बनाते हुये उत्पादन मध्य और नमूना मानक विचलन को प्राप्त करने के लिए उत्पादन सतह कार्य पद्धति का प्रयोग किया गया। उपरोक्त कार्य पद्धति की व्याख्या परीणामों में उदाहरण के साथ की गयी हैं।

#### दृष्टान्त-1

चर मिश्रण घटकों  $x_1, x_2, x_3, x_4$  तथा एक प्रक्रिया चर (z) को शामिल करने वाले एक प्रक्रिया चर के साथ एक मिश्रण प्रयोग पर विचार किया है। प्रयोग का उद्देश्य उत्पादन चर के साथ मिश्रण प्रयोग का अनुकूलतम समाधान प्राप्त करना है। हम निर्धारित निम्नलिखित काल्पनिक डेटा सेट को एक प्रक्रिया चर वाले मिश्रण प्रयोग में फिटिंग करने के लिए विचार किया है जैसे कि तालिका 2 में दिया गया है।

तालिका-2: प्रक्रिया चर के साथ मिश्रण प्रयोग के लिए डेटा

$X_1$	$X_2$	$X_3$	$X_4$	Z	$Y_1$	$Y_2$	$Y_3$	$Y_m$	$Y_s$
0.083	0.417	0.083	0.417	1	134	110	128	124.000	12.49
0.083	0.083	0.417	0.417	1	144	178	188	170.000	23.07
0.417	0.083	0.417	0.083	1	90	122	129	113.667	20.79
0.250	0.250	0.250	0.250	1	322	350	350	340.667	16.17
0.167	0.167	0.500	0.166	1	354	345	350	349.667	4.51
0.083	0.417	0.417	0.083	1	311	360	328	333.000	24.88
0.167	0.167	0.166	0.500	1	234	268	267	256.333	19.35
0.500	0.167	0.167	0.166	1	290	263	253	268.667	19.14

X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	Z	Y <sub>1</sub>	Y <sub>2</sub>	Y <sub>3</sub>	Y <sub>m</sub>	Y <sub>s</sub>
0.417	0.083	0.083	0.417	1	110	160	192	154.000	41.33
0.167	0.500	0.166	0.167	1	269	362	392	341.000	64.13
0.417	0.417	0.083	0.083	1	328	294	345	322.333	25.97
0.500	0.167	0.167	0.166	-1	81	168	78	109.000	51.12
0.000	0.333	0.333	0.334	-1	538	489	482	503.000	30.51
0.166	0.500	0.167	0.167	-1	98	110	105	104.333	6.03
0.333	0.000	0.334	0.333	-1	118	117	116	117.000	1.00
0.167	0.166	0.500	0.167	-1	129	154	131	138.000	13.89
0.334	0.333	0.000	0.333	-1	159	155	163	159.000	4.00
0.166	0.167	0.167	0.500	-1	328	391	394	371.000	37.27
0.333	0.333	0.334	0.000	-1	285	217	359	287.000	71.02
0.083	0.417	0.083	0.417	-1	145	158	160	154.333	8.145
0.250	0.250	0.250	0.250	-1	556	490	525	523.667	33.02
0.417	0.083	0.083	0.417	-1	140	172	205	172.333	32.50

इस डेटा सेट के लिए अनुकूलतम प्रवर्तन स्थिति प्राप्त करने के लिए हमने एक प्रक्रिया चर वाले चार घटक मिश्रण मॉडल को तीन घटक मिश्रण मॉडल में रिपैरामिटराइज किया जिसका एक गुणात्मक कारक है। तीन घटक मिश्रण के लिए मॉडल इस प्रकार होंगे।

$$Y(X, Z) = \beta_0 + \beta_1'x_1 + \beta_2'x_2 + \beta_3'x_3 + \beta_1''x_1^2 + \beta_2''x_2^2 + \beta_3''x_3^2 + \beta_12'x_1x_2 + \beta_13'x_1x_3 + \beta_23'x_2x_3 + \delta_0Z + \delta_1'x_1Z + \delta_2'x_2Z + \delta_3'x_3Z$$

$$\text{स्थिति के साथ } \sum_{i=1}^4 x_i = 1$$

हम  $x_1, x_2, x_3$ , तथा  $z$  के लिए मध्य के लिए द्वितीय आर्डर उत्पादन सतह समीकरण को मध्य के लिए फिट करते हैं।

$$\hat{y}_m(x, z) = -479.67 + 1827.34 X_1 + 2346.70 X_2 + 3067.43 X_3 - 2361.36X_1^2 - 3784.65X_2^2 - 4132.32 X_3^2 - 512.74 X_1X_2 - 3048.62X_1X_3 - 436.86X_2X_3 - 251.36Z + 429.71 X_1Z + 382.5345X_2Z + 263.54X_3Z$$

$$R^2=81.15\%$$

### प्रसरण के विश्लेषण

स्रोत	DF	SS	MS	F	P
प्रत्यावर्तन के कारण	13	274598.00	21123.00	2.65	0.0857
अवशेष त्रुटि	8	63784.00	7972.98		
योग	21	338382.00			

मानक विचलन के लिए फिट किया गया उत्पादन फलन इस प्रकार है।

$$\hat{y}_s(x, z) = 169.35 - 470.32X_1 - 491.23X_2 - 341.37X_3 + 417.50X_1^2 + 322.04X_2^2 + 49.31X_3^2 + 689.18X_1X_2 + 560.44X_1X_3 + 855.47X_2X_3 + 11.30Z - 28.48X_1Z - 0.10X_2Z - 23.91X_3Z$$

$$R^2=54.46\%$$

### प्रसरण के विश्लेषण

स्रोत	DF	SS	MS	F	P
प्रत्यावर्तन के कारण	13	4041.91	310.92	0.74	0.7009
अवशेष त्रुटि	8	3380.04	422.50		
योग	21	7421.94			

$$\hat{y}_m(x, z=1) = -731.03 + 2257.05X_1 + 2779.23X_2 + 3330.97X_3 - 3094.39X_1^2 - 2370.39X_2^2 - 1891.91X_3^2 - 889.99X_1X_2 - 3170.68X_1X_3 - 937.81X_2X_3$$

$$\hat{y}_s(x, z=1) = 180.65 - 498.80X_1 - 491.13X_2 - 365.28X_3 + 421.01X_1^2 + 325.19X_2^2 + 45.97X_3^2 + 699.21X_1X_2 + 557.41X_1X_3 + 851.83X_2X_3$$

$z=-1$  के लिए फिट किये मानक विचलन और उत्पादन मध्य के लिए फलन इस प्रकार दिये गये हैं।

$$\hat{y}_m(x, z=-1) = -228.31 + 1397.63X_1 + 1964.18X_2 + 2803.89X_3 - 2371.26X_1^2 - 3981.96X_2^2 - 4201.24X_3^2 - 649.61X_1X_2 - 2930.79X_1X_3 - 694.81X_2X_3$$

$$\hat{y}_s(x, z=-1) = 158.055 - 443.20X_1 - 492.33X_2 - 317.46X_3 + 421.01X_1^2 + 325.19X_2^2 + 45.97X_3^2 + 699.21X_1X_2 + 557.41X_1X_3 + 851.83X_2X_3$$

दोहरी उत्पादन के अनुकूलतम के लिए हम मानक विचलन को न्यूनतम तथा मध्य उत्पादन को अधिकतम करने वाली कार्य पद्धति का प्रयोग करते हैं। SAS कोड उपयोग करके यह देखा गया है कि  $z=1$  के लिए मानक विचलन के न्यूनतम लक्ष्य मान 19.12 है। हम बिन्दुओं को प्रयोग करके प्रत्येक स्तर पर मध्य मान के उत्पादन की गणना करते हैं। हमें वह वांछनीय बिन्दुओं को प्राप्त करने हैं। जिसके लिए उत्पादन चर अधिकतम हैं।

$z=1$  तथा लक्ष्य मानक विचलन 19.12 के लिए गत्र (0.2476, 0.3463, 0.2948) पर मध्य उपज अधिकतम है।

$z=1$  तथा लक्ष्य मानक विचलन  $\hat{y}^*(x, z)=19.12$  के लिए अधिकतम उपज  $\hat{y}_m(x, z)=411.10$   $ij$   $x = (0.2476, 0.3463, 0.2948)$  प्राप्त की गयी है।

इस प्रकार  $z=1$  पर अनुकूलतम उत्पादन मध्य के लिए घटको  $x = (0.2476, 0.3463, 0.2948, 0.1113)$  पर है क्योंकि  $z=-1$  के लिए मानक विचलन बहुत अधिक (29.45) है। इसलिए  $z=-1$  के लिए उत्पादन मध्य के अनुकूलतम नहीं किया गया।

उत्पादन चर के साथ मिश्रण प्रयोगों के लिए दोहरी उत्पादन की तकनीक पर परिचर्चा की गयी है। जब अवलोकन दो या दो से अधिक बार की गयी हो तब दोहरी उत्पादन की तकनीक का प्रयोग किया जा सकता है। इसकी विशेषताये यह है कि हमें जब आवश्यकता हो, प्रतिकृति को कम करने की तथा मध्य उत्पादन को अधिकतम करने का प्रयत्न किया जा सकता हैं।

## भविष्य शोध

मिश्रण प्रयोगों पर कार्य उपलब्ध है। जब त्रुटियों को सामान्य की तरह आईडेन्टीकली तथा स्वतन्त्र रूप में वितरण माना जाता है। एवं त्रुटियाँ सहसम्बंधित हो, वहाँ डेटा विश्लेषण की तकनीक के विकास का अवसर है। प्रक्रिया चर के साथ मिश्रण प्रयोगों पर भी शोध के अवसर है। जब प्रक्रिया चर के वर्गीकरण लगातार है। इससे आगे एक से अधिक प्रक्रिया चर वाले मिश्रण प्रयोग के विश्लेषण की जाँच की जाने की आवश्यकता है।

## संदर्भ

1. अग्रवाल, एम.एल., चौधरी, एस.आर., बंसल, ए. एवं मित्तल, एन. (2006): "इफिशन्ट रिस्पान्स सरफेस डिज़ाइनस् विद् क्वान्टिटेटिव एण्ड क्वालिटेटिव फैक्टर्स", जे. इन्ड. सो. एग्रील. स्टेटिस्ट. 60 (2), 90-99।
2. एनकेनमैन, आर.ई. एवं डीन, ए.एम. (2003): "क्वालिटी इमप्रूवमेंट एण्ड रोबस्टनेस वाया डिज़ाइन ऑफ एक्सपरीमेंटस", हेंडबुक ऑफ स्टेटिस्टिक्स, 22: स्टेटिस्टिक्स इन इंडस्ट्री, पार्ट 1-स्टेटिस्टिक्स इन रिसर्च एण्ड डेवलपमेंट, चेप्टर 8, एलसीवियर साईंस।
3. कार्नेल, जे.ए. (2002): "एक्सपरीमेंटस विद् मिक्सचर्स: डिज़ाइनस, मॉडल्स, एण्ड द एनालिसिस ऑफ मिक्सचर डाटा", जॉन विले एण्ड सन्स, इन्क, न्यूयार्क, एन वाई।
4. डेल केसटिलो, ई., फेन, एस.के. एवं सेम्पल, जे.

- (1997): "द कम्प्यूटेशन ऑफ ग्लोबल ऑप्टीमा इन ड्यूअल रिस्पान्स सिस्टम्स", जे. क्वाल. टेक., 29(3), 347–353।
5. डेल केसटिलो, ई., फेन, एस.के. एवं सेम्पल, जे. (1999): "ऑप्टीमाइजेशन ऑफ ड्यूअल रिस्पान्स सिस्टमस: ए कम्प्रीहेन्सिव प्रोसिजर फॉर डीजनरेट एण्ड नॉनडिजनरेट प्रॉब्लमस", इयूरो. जे. ओपर रिस., 112, 174–186।
6. खूरी, ए.आई. एवं कॉरनैल, जे.ए.(1996): "रिस्पान्स सरफेसिस: डिज़ाइनस एण्ड एनालिसिस", सेकन्ड एडीशन, मार्सल डेक्कर, इनक., न्यूयार्क।
7. तागुची, जी.(1959): "इन्ट्रोडक्शन टू एक्सपेरीमेंटल डिज़ाइन", यूएनआईपीयूबी, व्हाइट प्लेनस्, न्यूयार्क।
8. तागुची, जी. (1987): "सिस्टम ऑफ एक्सपेरीमेंटल डिज़ाइन: इंजीनियरिंग मेथडस टू ऑप्टीमाइज़ क्वालिटी एण्ड मिनिमाइज़ कोस्ट", व्हाइट प्लेनस्, न्यूयार्क।
9. विनिंग, जी.जी. एवं मियर्स, आर.एच. (1990): "कम्बाइनिंग तागुची एण्ड रिस्पान्स सरफेस फिलोसफिस: ए ड्यूअल रिस्पान्स अप्रोच", जे. क्वाल. टेक., 22, 38–45।

□

जो कोई शिक्षक बनना चाहता है उसका सच्चा लक्ष्य होना चाहिए – अपने विचार किसी पर ना थोपे, बल्कि उसके मस्तिष्क को प्रदिप्त करे।

—एफ. डब्ल्यू. रॉबर्टसन

## भारत में स्टिक-लाख के पूर्व अनुमान के लिये स्टोकास्टिक मॉडल

बिशाल गुरुंग, अमृत कुमार पॉल, रंजीत कुमार पॉल एवं कमलेश नारायण सिंह

वर्तमान अध्ययन में हमारा लक्ष्य भारत की वार्षिक स्टिकलाख उत्पादक आंकड़ों के लिये सबसे उचित पूर्व अनुमान मॉडल की एकसपो ने आटो रिग्रेसिव मॉडल पर आधारित आंकड़ों की रचना करना है। सांख्यिकीय मॉडलिंग और काल-श्रृंखला आंकड़ों के पूर्व अनुमान के मुख्य अस्थिरांक के बीच संबंध को सम्मिलित करना और नीति निर्धारकों को सहयोग देना। तदनुसार इस पत्र में Exponential autoregressive (EXPAR) परिवारों के मॉडलों का उपयोग भारत की सालाना स्टिकलाख के उत्पादन के आंकड़ों के चक्रिय अस्थिरता का चित्रण करता है। यह देखा गया है कि फिटिड मॉडल आंकड़ों को संतोश जनक रूप में पकड़ता है। इसके अतिरिक्त फिटिड मॉडल का निष्पादन विविधगुडनेस आफ फिट के माप और पूर्व अनुमान के निष्पादन से तुलना की गई है। यह निष्कर्ष निकला कि EXPAR मॉडल का निष्पादन चक्रिय आंकड़ों के मॉडलिंग और पूर्व अनुमान के लिये पूर्णतया ठीक है।

लाख एक लाल रालदार कई किले की प्रजाती का स्रोत है। कई हजार छोटे कीट कई परपोशी वृक्षों की शाखाओं पर संघजीवी रालदार रंगद्रव्य स्राव करते हैं। आवृत शाखायें परपोशी वृक्षों को काटकर स्टिकलाख प्राप्त करते हैं। कटे हुए स्टिकलाख को पीसने और छान कर अशुद्धियों को दूर किया जाता है। छने हुए पदार्थ को बार-बार धो कर कीटों के भाग और अतिरिक्त घुलनशील सामग्री को दूर करते हैं। लाख उत्पादन में भारत में अग्रिणी राज्य झारखंड, छत्तीसगढ़,

पश्चिम बंगाल और महाराष्ट्र हैं। भारत और थाईलैंड लाख और उससे बनने वाले पदार्थों के मुख्य निर्यातक हैं एवं दोनों समान रूप से बाजार में हिस्सेदार हैं। भारत की वार्षिक स्टिकलाख की उत्पादन वर्ष 1950 के मध्य में 50,000 टन था जो घटकर जो वर्ष 1980 के अंत में 12,000 टन रह गया है।

कृषि में, एकत्रित आंकड़ों को श्रेणीबद्ध आकलन के लिए सांख्यिकीयविदों के यहाँ रैखिक काल श्रृंखला मॉडल पर आधारित विधियों का प्रयोग किया जाता है जिसे हम अरिमा स्वः प्रतिगामी समेकित चक्रिय औसतन विधि कहते हैं। रैखिक मॉडल काल-श्रृंखला आंकड़ों की बहुत परीक्षणों की आकृतियों महत्वपूर्ण होती है। इसका प्रयोग अधिकतर लोकप्रिय काल-श्रृंखला मॉडल में किया जाता है। अधिक अनुप्रयोगों में, अरैखिक मॉडल बाजार के भविष्य को अग्रसिव करना और शुद्धता से श्रेणियों को गतिमान करना है।

अरैखिक काल-श्रृंखला मॉडल का मापदण्ड परिवार जैसे कि एक गुणांक स्वः प्रतिगामी EXPAR है। मापदण्ड स्पेस को बदलते हुए विशेष रूप के जोन क्षेत्रों में विभिन्न प्रकार के सीमांत बंटनों के साथ काल-श्रृंखला को उत्पन्न करने के लिए योग्य है। ये आंकड़ें उस चक्रिय प्रसरण को दर्शाते हैं और काल-श्रृंखला के लक्षणों को अगोशियन मॉडल के भविष्य इस मॉडल की व्याख्या करते हैं। गुणांक स्वः प्रतिगामी का महत्व पूर्वानुमान उद्देश्य और मॉडलिंग के लिए अरिमा के साथ तुलनात्मक अध्ययन करना है।

## सामग्री और विधियों

### मॉडल का विवरण:

#### EXPAR मॉडल

$$X_{t+1} = \{\varphi_1 + \pi_1 \exp(-\gamma X_t^2)\}X_t + \dots + \{\varphi_p + \pi_p \exp(-\gamma X_t^2)\}X_{t-p+1} + \epsilon_{t+1}$$

$\gamma > 0$  मापक स्थिरांक और शून्य के साथ शांत स्वभाव और प्रसरण  $\epsilon_{t+1}$  है।  $\gamma$  एक रेंज के संपर्क से विभिन्नता जैसे चयनित मूल्यों  $\exp(-\gamma X_t^2)$  के होते हैं।

#### मापदण्डों का आकलन:

बरगान, इत्यादि (2002) ने मापदण्डों को आकलन करने के लिए प्रक्रिया का एक संक्षिप्त विवरण प्रस्तुत किया गया है। लघुगणक एक अंतराल (a,b) बताता है।  $a \geq 0$   $\gamma$  मूल्यों से पहले विशिष्टता होगी। Mउप-अंतरालों (sub intervals) में खण्डों के रूप में अंतराल होते हैं जो मूल्यों अभियर्था 'y' के साथ एक ग्रिड के रूप में बनाये गये हैं। यदि  $\delta = (b-a)/M$  और  $\gamma = a$ , तो M समय बार के लिए, कुछ कार्य फलन नीचे प्रदर्शित किये गये हैं।

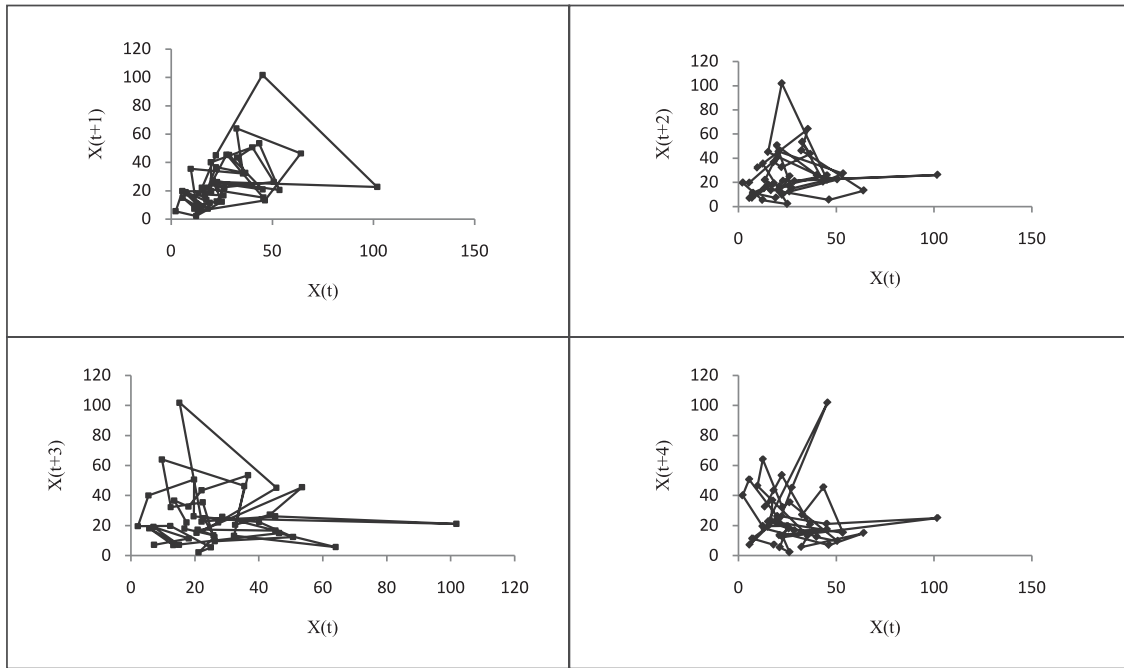
$$(1) \gamma = \gamma + \delta$$

(2) एन.ए.आई.सी. को गणित करना और दोहराना (ii)  $p=1, \dots, P$  जहाँ एक से अधिक, पहले विशिष्टता पूर्णांक है। मापदण्डों के अन्तिम आकलनों के प्राप्त न्यूनतम (NAIC) के द्वारा, इस प्रकार प्रदर्शित किया गया है।

$$NAIC = \{N \log(\hat{\sigma}^2) + 2(2p + 1)\} / (\text{Effective sample size})$$

#### परिणाम और चर्चा

स्टिक-लाख का उत्पादन 1930 से 2007 के दौरान भारत के लिए एक उदाहरण है जो [www.indiastat.com](http://www.indiastat.com) वेबसाइट पर उपलब्ध है। कुल 78 आंकड़ा बिन्दुओं पर 70 आंकड़ा बिन्दू मॉडल बनाने के लिए और 'शेष 8 आंकड़ा बिन्दू वैधता के उद्देश्य से बनाये गये हैं। एक्सपार (EXPAR) मॉडल को सुनियोजित करके आंकड़ा विश्लेषण की प्रारम्भिक छानबीन की गई है। आकृति न. 1 में विक्रत ग्राफ (scatter diagram) दर्शाया गया है जोकि संयुक्त बंटन के परीक्षणों में असीमित रूप से दिखाई गई है तथा संयुक्त बंटन नॉन-गासियन, जैसे द्विदिश सामान्य बंटन असीमित नहीं हो सकता है।



आकृति न. 1: भारत में स्टिक-लाख का कुल उत्पादन विक्रत ग्राफ द्वारा

प्रथम सांराश में अरिमा (ARIMA) मॉडल्स प्रक्रिया के अंतर्गत आंकड़ें विचाराधीन है। न्यूनतम मूल्यों के आधार पर अरिमा (1,1,0) मॉडल निम्नलिखित रूप से प्रयोग गया है :

$$X_{t+1} = 3.55 + 0.91X_t + \epsilon_{t+1}$$

$$\text{Var}\{\epsilon_t\} = 36.45$$

बारागोना (2002) के अनुसार प्रस्तावित लघुगणक मापदण्डों के आकलन के द्वारा प्रयोग किया गया है। न्यूनतम (NAIC) के उत्पत्ति के आधार पर निम्नलिखित (EXPAR) मॉडल प्रयोग किया गया है :

$$X_{t+1} = \{0.98 + 0.9 \exp(-1.01X_t^2)\}X_t + \epsilon_{t+1}$$

$$\text{Var}\{\epsilon_t\} = 30.64$$

एक यान्त्रिकीय सूझबूझ फिटिड मॉडल (EXPAR) की यह है कि  $\theta = \{0.98 + 0.9 \exp(-1.01X_t^2)\}$  बड़ा (छोटा) जबकि  $|X_{t+1}|$  (छोटा) बड़ा है। इसका अर्थ होता है कि भारत का कुल स्टिक-लेक उत्पादन का ट्रेंड एक निश्चित समय के पश्चात अनिश्चित भी होना है और मौखिक रूप से यह दृष्टता एक भावी चक्रीय जोकि परीक्षित आंकड़ा से सहमति रखती है। शब्दीय विचार प्राप्त करना है तो एक्सपार फिटिड का ग्राफ आकृति न. 2 प्रदर्शनी के साथ आंकड़ा बिन्दुओं मॉडल पर अंकित हैं एक दृष्टता निरीक्षण प्रदर्शित करता है कि फिटिड मॉडल्स उपस्थित प्रस्तुत आंकड़ा सेट में चक्रीय ज्वार-भाटा की तरह उचित रूप से पकड़ के योग्य है।



आकृति न: 2 आंकड़ा बिन्दुओं के साथ फिटिड अरैखिक काल-श्रृंखला मॉडल

तालिका 3: फिट मॉडल की स्वच्छता

मॉडल	अरिमा	एक्सपार
AIC	843.34	783.39
BIC	872.23	759.02
MSE	14873.42	10582.91

मॉडल्स (AIC), (BIC) और (MSE) एक्सपार (EXPAR) के साथ-2 अरिमा मॉडल्स तालिका न:1 में सूचनार्थ एवं गणित किया गया है। एक अवशेष सूचनार्थके वह एक्सपार मॉडल अरिमा के मद्देनज़र मॉडलिंग से बहुत अच्छा है।

पूर्वानुमान कार्यकुशलता का मूल्यांकन :- इस अनुभाग में, एक तुलनात्मक अध्ययन की मांग का मूल्यांकन अरिमा मॉडल और एक्सपार मॉडल्स की उपयोगिता की पूर्वानुमान का पता लगाया गया है। तथा कार्यकुशलता के फिटिड मॉडल्स तुलना के आधार पर एक स्टेप आगे भा.वर्ग.पूर्वा.त्र. (एम.एस.पी.ई.) से किया गया है तथा भा.वर्ग.पूर्वा.त्र. (एम.ए.पी.ई.) एवं सापेक्ष भा.प.पूर्वा.त्र. मॉडल में (7.83, 2.36 और 11.71) फलस्वरूप पाये गये हैं। जोकि अल्प एक सापेक्ष अरिमा मॉडल के लिए फिट किया जाता है। उदाहरण के लिए 14.98, 3.59 और 18.06 है। एक्सपार मॉडल का सर्वश्रेष्ठता अरिमा मॉडल के बाद आंकड़ा सेट के लिए पूर्वानुमान उद्देश्यों के लिए विचाराधीन है।

इस पेपर में, एक्सपार मॉडल अरैखिक काल-श्रृंखला मॉडल को ही चिन्हित एवं सर्वश्रेष्ठता मॉडल अरिमा मॉडल के साथ ही प्रदर्शित किया गया है और आशा की जाती है कि शोधक चक्रीय काल-श्रृंखला आंकड़ा के एक्सपार की शुरुआत करेंगे।

### संदर्भ

1. बारागोना, आर., बट्टागलिया, एफ., एवं कुसिना, डी. (2002): “ए नोट ऑन एस्टीमेटिंग ऑटोरिग्रेसिव एक्सपोनेनशियल मॉडल्स”, कुएड. डी. स्टेट., 4, 1-18 |
2. बॉक्स जी.ई.पी., जेनकिंस जी.एम. एवं रेनसिल जी.सी. (2008): “टाईम सीरीज़ एनालिसिस: फॉरकार्स्टिंग एण्ड कंट्रोल”, प्रेंटिस हाल |



3. काई जेड. फेन जे. एवं ली. आर (2000): “एफिशियेंट एस्टीमेशन ऐन्ड इंफ्रेंसिज् फार वेरिंग-कोएफिशियेंट मॉडल्स”, जे. एमिर स्टेटिस्ट एसोसि., 95, 888–902 ।
4. फेन, जे. एवं याओ क्यू (2003): “एनॉन-लीनियर टाईम-सीरीज़: नॉन-पेरामेट्रिक ऐन्ड पेरामेट्रिक मैथडस्”, स्प्रिंगर, न्यूयार्क ।
5. घोष, एच., गुरुंग. बी. एवं प्रज्ञेषु (2011): “मैथोडोलोजी फार कम्बाइनिंग लीनियर ऐन्ड नान-लीनियर टाईम-सीरीज़ मॉडल्स फार साईक्लिकल डाटा”, जे. इण्ड.सोस, 65(2), 249–256 ।
6. ओजाकी, टी. (1993): “ए. नान-गोसियन करेक्टरस्टिक्स आफ एक्सपोनेन्शियल ऑटोरिग्रेसिव प्रोसेसिज् इन डेवलेपमेंटस इन टाईम-सीरीज़ एनालिसिज”, एड. टी. सुब्बा राव. चैपमैन ऐन्ड हाल, 257–73 ।

□

निज भाषा उन्नति अहे, सब उन्नति को मूल

बिन निज भाषा ज्ञान के, मिटत न हिय को शूल ।

—भारतेंदु हरिश्चंद्र

## भैंस में नए गुण-बिन्दुओं की खोज

सारिका, मीर आसिफ इकबाल, ऊषा जैन,  
अनिल राय एवं दिनेश कुमार

दूध, मांस, खाल, उर्वरक, ईंधन एवं भारवाही पशु शक्ति के माध्यम से भारतीय महाद्वीप एवं दक्षिण पूर्वी एशियन देशों की कृषि अर्थव्यवस्था में भैंस (*Bubalus bubalis*) का बहुत योगदान है। जनसंख्या का एक बहुत बड़ा हिस्सा विश्व की किसी अन्य पशुधन प्रजाति की अपेक्षा इस प्रजाति पर निर्भर है (शर्फ, 2000)। भैंस Bovidae परिवार से हैं जिसमें गुणसूत्र (chromosomes) की संख्या 48 (स्वैम्प भैंस में) तथा 50 (रिवराइन भैंस में) है (फिशर एवं अल्लिच, 1967)। गुण बिन्दुओं (gene) की पहचान तथा जीनोम एनोटेशन के लिए EST आँकड़े महत्वपूर्ण संसाधन हैं (लियांग इत्यादि, 2000)।

पशुओं में, विशेषकर नॉन-मॉडल पशुओं में, जिनके लिए वर्तमान में पूर्ण जीनोम अनुक्रम (sequences) उपलब्ध नहीं हैं, नए गुण बिन्दुओं की खोज के लिए अभिव्यक्त अनुक्रम टेग्स (ESTs) की सीक्वेंसिंग एवं विश्लेषण एक प्राथमिक टूल है। जैव प्रौद्योगिकीय तकनीक के विकास से ESTs में वृद्धि हुई है, जो वैज्ञानिक समुदाय को तुलनात्मक एवं कार्यात्मक जीनोमिक अध्ययन के लिए आनुवंशिक सूचना उपलब्ध कराने के साथ-साथ जीनोमिक सीक्वेंसेस की व्याख्या (एनोटेशन) में भी मदद करता है। भैंस के ESTs में SNPs (सिंगल न्यूक्लियोटाइड पॉलीमॉर्फिज्म) की खोज जीनोमिक डी.एन.ए.-आधारित मार्करों की डिजाइनिंग में मदद कर सकता है तथा इस प्रजाति के विकास में और अधिक जानकारी प्रदान कर सकता है। *Bubalus bubalis* के लिए सार्वजनिक डोमेन में उपलब्ध ESTs को SNPs एवं कार्यात्मक एनोटेशन का

पता लगाने हेतु संगणक द्वारा अध्ययन के लिए लिया गया है।

### सामग्री एवं विधियाँ

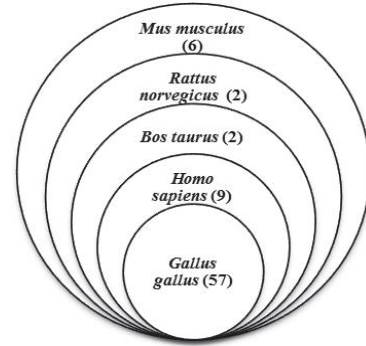
भैंस (*Bubalus bubalis*) के लिए सार्वजनिक डाटाबेस NCBI में मार्च, 2012 तक जमा किए गए कुल 1825 ESTs को EST क्लस्टरिंग, कार्यात्मक एनोटेशन, SNP माइनिंग एवं प्रोटीन डोमेन की पहचान के लिए FASTA फॉरमेट में डाउनलोड किया गया। भैंस के उपलब्ध ESTs को EGAsembler का प्रयोग करते हुए कॉन्टिग्स को नॉन-रिडन्डेंट (non-redundant) सेट में क्लस्टर किया गया (मसौदी इत्यादि, 2000)। ESTs के संरेखण से प्राप्त कॉन्टिग्स को मौजूदा गौ-पशु (*Bos taurus*) सीक्वेंस डाटाबेस में समरूपी खोजों के लिए विश्लेषित किया गया। एनोटेशन के लिए, छोड़े गये कॉन्टिग्स को नॉन-रिडन्डेंट डाटाबेस के प्रति BLASTX (एल्ट्राल इत्यादि, 1990) किया गया। अध्ययन से प्राप्त गुण बिन्दुओं को समकक्ष गुण बिन्दु (HomoloGene) (<http://www.ncbi.nlm.nih.gov/homologene>) का प्रयोग करते हुए मानव (*Homo sapiens*), घरेलू चूहा (*Mus musculus*), भूरे चूहे (*Rattus norvegicus*), गाय (*Bos taurus*) एवं मुर्गी (*Gallus gallus*) के गुण बिन्दुओं के साथ तुलना की गयी। सभी अपरिष्कृत EST अनुक्रमों को *Bos taurus unigene*, जो NCBI से बैच डाउनलोड किया गया था, के संदर्भ में SNPfinder का उपयोग करते हुए अनुमानित SNPs के लिए विश्लेषित किया गया (साँग इत्यादि, 2005)। इसके अतिरिक्त, प्राप्त प्रोटीन अनुक्रमों

के प्रोटीन डोमेन का पता लगाने के लिए भी विश्लेषण किया गया। कॉन्टिग्स की Batch Web CD-Search tool (मार्शलर-बाउर इत्यादि 2009) एवं Pfam डाटाबेस में तुलना की गयी।

### परिणाम एवं परिचर्चा

भैंस के उपलब्ध 1825 EST अनुक्रम जीन बैंक एवं dbESTs से डाउनलोड किये गए और 151 कॉन्टिग्स में कलस्टर किये गये। कम प्रचलित एवं अभिव्यक्त ट्रॉसक्रिप्ट्स, 1060 सिंगलटन के रूप में प्राप्त हुए। गाय (*Bos taurus*) गुण बिन्दु सूचकांकों का प्रयोग करते हुए भैंस (*Bubalus bubalis*) कॉन्टिग्स के साथ गाय (*Bos taurus*) के तुलनात्मक विश्लेषण ने समानता के उच्च स्तर का प्रदर्शन किया। इन एनोटेटेड गुण बिन्दुओं को ज्ञात प्रोटीन के लिए उनकी एकरूपता के आधार पर 15 विभिन्न कार्यात्मक श्रेणियों में वर्गीकृत किया गया। कुल कॉन्टिग्स में से 58.21 प्रतिशत, जैविक कार्य का वर्णन करते पाए गए। कॉन्टिग्स की अगली अधिकतम कार्यात्मक श्रेणी मेटाबॉलिज्म एवं ऊर्जा थे जिसमें कुल कॉन्टिग्स का 12.69 प्रतिशत विस्तृत था। यद्यपि, ऐसा विश्लेषण भैंस के मामले में निष्पादित नहीं किया जा सका, परन्तु ऊँट (*Camelus dromedarius*) में 16.83 प्रतिशत ट्रॉसक्रिप्ट्स सहित असेम्बल्ड EST सीक्वेंसेस मेटाबॉलिज्म एवं ऊर्जा प्रक्रिया में शामिल थे। इसके अतिरिक्त, दूसरी प्रमुख श्रेणी सैल्युलर ट्रॉसपोर्ट थी जिसमें कुल कॉन्टिग्स का 11.94 प्रतिशत शामिल था, जो ऊँट में भी रिपोर्ट किया गया है। संरचनात्मक/उत्प्रेरक प्रोटीन, प्रोटीन संश्लेषण (synthesis), ट्रॉसक्रिप्शन फैक्टर एवं कोशिका चक्र (cell cycle) तथा डी.एन.ए प्रोसेसिंग अन्य प्रमुख कार्यात्मक श्रेणियाँ हैं। विश्लेषण में प्रयुक्त अन्य प्रजातियों के संदर्भ में जिसमें आगे अध्ययन की आवश्यकता है, शेष 41.79 प्रतिशत भैंस विशिष्ट नॉवेल ESTs पायी गयीं। भैंस के EST अनुक्रम जिन्होंने घरेलू चूहे (*M. musculus*), भूरे चूहे (*R. norvegicus*), गाय (*B. taurus*), मानव (*H. sapiens*) एवं मुर्गी (*G. gallus*) के गुण बिन्दुओं के साथ समजातता (homology) दर्शाई है, वेन आरेख के माध्यम से आरेख-1 में प्रस्तुत हैं।

प्रस्तुत अध्ययन में, समस्त 151 कॉन्टिग्स में से 224 pSNP का पता चला, जिन्हें प्रजनन के लिए आण्विक (molecular) मार्कर्स के रूप में प्रयोग किया जा सकता है। समस्त प्रागुक्त (predicted) कॉन्टिग्स की Pfam डाटाबेस का प्रयोग करते हुए उनकी डोमेन संरचना के लिए आगे विश्लेषण किया गया। 151 कॉन्टिग्स में से संरक्षण डोमेन के PSSM-ID के साथ केवल 42 कॉन्टिग्स पाए गए। प्राप्त किये गए अधिकांश डोमेन भैंस (*Bos taurus*), मुर्गी (*Gallus gallus*), घरेलू चूहे (*Mus musculus*), भूरे चूहे (*Rattus norvegicus*) एवं मानव (*Homo sapiens*) के प्रति प्राप्त BLAST परिणामों के साथ पूरी तरह से समानता रखते थे।



आरेख-1 : शेयर्ड जींस का वेन आरेख

### निष्कर्ष

निष्कर्षतः, उपलब्ध ESTs का विस्तृत विश्लेषण भैंस एवं *Bovidae* के पूरे वंश के बेहतर ज्ञान के लिए भावी शोध के लिए बहुमूल्य साबित होगा। अनुक्रम स्तर पर मवेशी एवं भैंस के जीनोम की तुलना पशुधन प्रजनन कार्यक्रम में क्लोनिंग एवं उपयोग के लिए कुछ आर्थिक रूप से महत्वपूर्ण गुण बिन्दुओं का पता लगाने के लिए सुअवसर प्रदान करता है। भैंस की और अधिक जीनोमिक एवं ट्रॉसक्रिप्ट सूचना प्राप्त करने के लिए यह विश्व का पहला न्यूनतम आँकड़ा एवं प्रयास है। इसके अतिरिक्त, सूचित pSNPs उपयोगी जीन्स, आनुवंशिकी एवं भौतिक नक्शे (maps) तैयार करने के लिए, आनुवंशिक विविधता के अध्ययन, मार्कर-असिस्टेड चयन एवं पोजिशनल क्लोनिंग के

लिए उपयोगी होंगे। तुलनात्मक जीनोमिक्स अनुक्रम के विकास के अध्ययन एवं ऑर्थोलोगस जीन्स के कार्य तथा विविधता एवं अनुकूलन को समझने के लिए सुविधाजनक बनाते हैं। यह शोध ज्ञान भविष्य में भैंस की उत्पादकता बढ़ाने के लिए बहुत उपयोगी होगा।

### संदर्भ

1. शर्फ बी.डी. (2000) : "वर्ल्ड वाच लिस्ट फॉर डॉमेस्टिक एनिमल डायवर्सिटी", थर्ड एडि. फूड ऐण्ड एग्रिकल्चरल ऑर्गनाइजेशन ऑफ दि यूनाइटेड नेशन्स, कम्युनिकेशन डिविजन, एफ.ए. ओ., रोम, इटली ।
2. फिशर, एच. एवं अल्ब्रिच, एफ. (1967) : "क्रोमोसोम ऑफ दि मुराह बफैलो ऐण्ड इट्स क्रॉसब्रीड्स विद दि एशियाटिक स्वैम्प बफैलो (*Bubalus bubalis*)", Zeitschrift für Tierzüchtung und Züchtungsbiologie, 84 : 110-114 ।
3. लियांग, एफ., हॉल्ट, आई., पर्टिया, जी., कर्माईचेवा, एस., साल्जबर्ग, एस.एल. एवं क्वाकेनबश, जे. (2000) : "जीन इन्डेक्स एनालिसिस ऑफ दि ह्यूमन जीनोम एस्टीमेट्स ऐप्रोक्सीमेटली 1,20,000 जीन", नेचर जेनेटिक्स, 25 : 239-40 ।
4. मसौदी-नेजाद, ए., टोनोमुरा, के., क्वाशिमा, एस., मौर्या, वाई, सुजुकी, एम., इटोह, एम., कनेहिसा, एम., एण्डो, टी., गोदो, एस. एवं ई-गैसेम्बलर (2006) : "ऑनलाइन बायोइन्फॉरमेटिक्स सर्विस फॉर लार्ज स्केल प्रोसेसिंग, क्लस्टरिंग ऐण्ड असेम्बलिंग ESTs ऐण्ड जेनोमिक डी.एन.ए. फ्रैगमेंट्स", न्यूक्लिक एसिड्स रिसर्च, 34 : डब्ल्यू 462-डब्ल्यू 469 ।
5. एल्डशल, एस.एफ, गिश, डब्ल्यू, मिलर, डब्ल्यू, मायर्स, ई. डब्ल्यू, एवं लिपमैन, डी.जे (1990) : "बेसिक लोकल अलाइनमेंट सर्च टूल", जर्नल ऑफ मॉलीक्यूलर बायोलॉजी, 215 : 403-410 ।
6. साँग, जे., एक्सयू, वाई., व्हाइट, एस., मिलर, के.डब्ल्यू.पी., एवं वोल्निस्की, एम. (2005) : "एस.एन.पी. फाइन्डर - ए वेव-बेस्ड एप्लीकेशन फॉर जीनोम-वाइड डिस्कवरी ऑफ सिंगल न्यूक्लियोटाइड पालीमॉर्फिज्म इन माइक्रोबायल जीनोम्स", बायोइन्फॉरमेटिक्स, 21 : 2084 ।
7. मार्शलर-बाउर, ए., एन्डरसन, जे.बी., चिटसाज, एफ., डर्बीशायर, एम.के., स्कॉट, सी.डी., फॉग, जे. एच., गीयर, एल.वाई., गीयर, आर.सी., गोन्जल्स, एन.आर., ग्वाड्ज, एम., ही एस., हार्विट्ज, डी. आई., जैक्सन, जे.डी., की जेड., लैन्जिकी, सी.जे., लाइबर्ट, सी.ए., ल्यू, सी. ल्यू, एफ., ल्यू, एस., मार्शलर, जी.एच., मुल्लोकैन्डोव, एम., साँग, जे. एस., तासनीम, ए., थैन्की, एन., यामाशीटा, आर.ए., झाँग, डी., झाँग, एन. एवं ब्रयान्ट, एस.एच. (2009): "सी.डी.डी. : स्पेसिफिक फंक्शनल एनोटेशन विद दि कन्जर्ड डोमेन डाटाबेस", न्यूक्लिक एसिड्स रिसर्च, 37 : डी 205-डी 210 ।

□

## अरहर दाल के खुदरा मूल्य का पूर्वानुमान

रंजीत कुमार पॉल, बिशाल गुरुंग, अमृत कुमार पॉल,  
वसी आलम एवं कंचन सिन्हा

ग्राहकों की मांगों से खुदरा व्यापारी खरीद के नमूनों को अच्छी तरह से व्यय करने की क्षमता के लिए पूर्वानुमान की सहायता लेता है। सही पूर्वानुमान से, ग्राहकों की सहायता के लिए व्यय मांगों, अर्धमाह खुदरा व्यापारी, मालिकों और प्रबन्ध को अधिक लाभ दिलाना एक दीर्घकालिक समय के पश्चात् होता है। पूर्वानुमा अनुज्ञा मूल्य को समायोजित करता है तथा ग्राहक के व्यय नमूनों की वर्तमान स्तर के साथ समझने की क्षमता रखता है। इसलिए, पूर्वानुमान खुदरा मूल्य कृषि की चीजों पर उच्च सांख्यिकीय तकनीको के द्वारा अत्याधिक रूप में महत्वपूर्ण है। अतः दीर्घकालिक स्मृति मॉडल बढ़ते रूप में लोकप्रिय हो रहा है जैसे अर्थ काल श्रृंखला का वर्णात्मक रूप में एक टूल है। यद्यपि, दीर्घअर्थ काल श्रृंखला अस्थिर है। यह सत्य रूप से आवश्यक नहीं है कि प्रथम भिन्नवादों और एक अरिमा मॉडल ही एक अच्छा तरीका होगा। इन बॉक्स जेनकिंस विश्लेषण में, यह कल्पना की गई है कि यदि श्रृंखलाएं अस्थिर है तो भिन्नता श्रृंखलाएं बहुत अधिक तेजी से क्षीण होकर, स्वः संबंधों और मुक्त रूप से उनमें व्यवहारित होगी तथा अच्छा वर्णित एक अरिमा (ARIMA) मॉडल होगा। किन्तु यह उस परिस्थिति में नहीं होता है स्वः प्रतिगामी संकलित फलन चक्रीय औसत मॉडल (ARFIMA) एक नॉन-पूर्णांक 'd' मापदण्ड के लिए एक दीर्घकालिक स्मृति आंकड़ा ही दीर्घकालिक स्मृति को पहुंचाता है तथा नॉन-जीरो 'd' का अस्तित्व दीर्घकालिक स्मृति का एक संकेत और दीर्घकालिक स्मृति की शक्ति शून्य मापदण्डों से अलग हो जाती है। वर्तमान शोध पत्र में करनाल में तुर और

अरहर दाल के खुदरा व्यापारियों ने प्रतिदिन संरचना दीर्घकालिक स्मृति के शोध का प्रयास किया गया है जो अरहर और तूर के मॉडल पूर्वानुमान में फुटकर विक्रेता श्रृंखला आंकड़ों को विश्लेषित सांख्यिकीयविधों में एक नया विकास होगा।

अरफिमा मॉडल (ARFIMA MODEL) काल श्रृंखला में दीर्घकालिक स्मृति जैसे स्वः संबंध पर लौग लेग्स् (रॉबिनसन, 1995) ने दर्शाया है। वर्ष 2004 में जिन एवं फ्रेनचैटी (Jin and Frenchette) के अनुसार स्मृति का अर्थ होता है परीक्षण बढ़ते हुए घटनाक्रम में एक दूसरे पर परीक्षणों को प्रभावित करती है। काल-श्रृंखला में दीर्घकालिक स्मृति के वर्णन के लिए प्राथमिक वैचारिक फ्रेमवर्क के लिए फलन समेकित है। फलन समेकित पूर्णांक समेकित का एक सामान्यतः है। जो दीर्घकालिक श्रृंखला के अंतर्गत शून्य के आदेशानुसार अथवा एक को संभाविक प्रयोगिक को समेकित करना है। उदाहरणः के लिए एक स्वः प्रतिगामी चक्रीय औसत प्रक्रिया 'd' के आदेशानुसार अरफिमा (ARFIMA) को जैसे प्रतिदर्श कर सकते हैं।

$$(1-L)^d \phi(L)y_t = \theta(L)u_t$$

जहाँ  $u_t$  एक स्वतंत्र और सांकेतिक बंटन (i,i,d) याच्छदिक चर और स्थिर प्रसरण शून्य का अर्थ लेग ऑपरेटर (Lag operator) को प्रदर्शित करता है और  $\phi(L)$  और  $\theta(L)$  अनिश्चित बहुद्विपक्षीय को प्रदर्शित करते हैं जो कि लेग संचालक एक चक्रीय इकाई से बाहर की तरफ मूल रूप से होता है। यह परिवर्तन करता है कि  $-0.5 < d < 0.5$  के लिए प्रक्रिया  $y(t)$  और

चर एक स्थिर रूप से होती है। एक अरफिमा मॉडल का विस्तार (2003) में रॉबिन्सन से विस्तृत रूप में बताया है एवं दीर्घकालिक स्मृति मापदण्ड 'd' का आंकलन (1983) में गेवेक और पोर्टर हुड्डक ने किया था।

### परिणाम और चर्चा

इस शोध पत्र में करनाल में अरहर दाल के प्रतिदिन के खुदरा मूल्यों के लिए जनवरी 2011 से 31 जुलाई 2013 के दौरान एकत्रित आंकड़ों का प्रयोग किया गया है। ये आंकड़े भारत सरकार, उपभोक्ता मंत्रालय से प्राप्त किये गये हैं। स्थिरांक प्रयोग के लिए, दो प्रयोग ऑगमेंटेड डीके-फूल्लर (Augmented Dickey-Fuller) मूल इकाई प्रयोग में और फिलिप्स-पैरॉन (Philips-Peron) मूल इकाई प्रयोग में प्रचालित किया था, प्रयोग किया गया है। सांख्यिकीय प्रयोग शून्य परिकल्पना (null-hypothesis) मूल इकाई सांकेतिक की उपस्थिति में यह श्रेणी स्थिरांक है। अरफिमा मापदण्ड आंकलन का मापदण्ड मॉडल तालिका 1 में वर्णित किया गया है।

**तालिका 1:** अरफिमा मॉडल (ARFIMA MODEL)के मापदण्ड आंकलन

मार्केट	मापदण्ड	आंकलन	एस. ई.	जेड वेल्यू	प्राथमिकता
करनाल	d	0.030	0.008	3.407	< 0.001
	ए.आर. आई.	0.961	0.011	88.220	< 0.001

यह परिणाम प्रदर्शित करता है कि d मापदण्ड दीर्घकालिक स्मृति का मूल्य और अधिक महत्वपूर्ण सांकेतिक दीर्घकालिक स्मृति की उपस्थिति में प्रतिदिन करनाल में, अरहर और तुर दाल की मुख्य भूमिका हैं। 2013 के दौरान उपरोक्त मॉडल के फलस्वरूप गठित किया गया है।

शुद्धता को मापने के लिए आंकड़ों का काल-श्रृंखला मॉडल में फिट किया गया है। माध्य परम त्रुटि (Mean Absolute Error), माध्य परम प्रतिशत त्रुटि (Mean Absolute Percentage Error) और सापेक्ष माध्य परम पूर्वानुमान त्रुटि (Relative Mean Absolute Prediction Error) तालिका 2 में प्रदर्शित किये गये हैं।

**तालिका 2:** मॉडल्स की वैधता (Validation of Models)

एम.ए.पी.ई.	आर.एम.एस.पी.ई.	आर.एम.ए.पी.ई.(%)
0.798	0.891	1.23

उपरोक्त तालिका के अनुसार, सभी में मूल्य श्रेणी आंकड़ा आर.एम.ए.पी.ई मॉडल की शुद्धता को दर्शाते हुए, 2% से कम है।

### निष्कर्ष

दीर्घकालिक स्मृति अपनी क्षमता में दीर्घ काल-श्रृंखला आंकड़े इतिहास की दीर्घ घटनाक्रम को दोहराती है और अपने निर्णय की क्षमता को कुछ स्मृतियों के आधार पर निर्णय लेती है। एक धरोहर का मूल्य (अथवा वस्तुएं) एक पर्याप्त मार्केट में एक मार्टिंगल प्रक्रिया को ग्रहण करना चाहिए जो प्रत्येक इस मूल्य परिवर्तन के द्वारा अप्रभावित होता है। मॉडल एक अच्छे निष्कर्ष के लिए उसकी व्याख्यात चर और पूर्वानुमान शक्ति के लिए प्रदर्शित होता है। प्रस्तुत शोध में, शुद्धता पूर्वानुमान का प्रयोग करने की क्षमता के लिए सीधे समर्थकता जुटाती है और खुदरा व्यापारियों के साथ-2 ग्राहकों के लिए निर्णय लेने में सक्षम होती है।

### संदर्भ

1. जी.वे.क. जे. एवं पार्टर-हुड्डक, एस. (1983): "दी एस्टीमेशन एण्ड एप्लीकेशन ऑफ लॉग मेमोरी टाईम-सीरीज़ मॉडल्स", जरनल ऑफ टाइम-सीरीज़ ऐनालिसिस, 4, 221-238।
2. जिन, एच.जे. एवं फ्रेचचीटी, डी. (2004): "फ्रेक्शनल इन्टीग्रेशन इन एग्रीकल्चरल फ्यूचर्स प्राइस वोलैटिलिटीस. अमेरिकन", जरनल आफ एग्रीकल्चरल इकोनॉमिक्स, 86, 432-443।
3. रॉबिन्सन, पी.एम. (1995): "लॉग पीरियडओग्राम रिग्रेशन आफ टाईम-सीरीज़ दि लॉग-रेंज डिपेन्डेन्स", दी एनल्स ऑफ स्टेटिस्टिक्स, 23, 1048-1072।
4. रॉबिन्सन, पी.एम. (2003): "टाईम-सीरीज़ विद लॉग मेमोरिज़", आक्सफोर्ड यूनिवर्सिटी प्रेस, आक्सफोर्ड।

□

## एफिड सरसों के प्रबंधन के लिये श्रेणीबद्ध परीक्षण प्रक्रियाएँ

वसी आलम, अमृत कुमार पॉल, रंजीत कुमार पॉल, कंचन सिन्हा एवं राम नरेश

समेकित कीट प्रबंधन में, कीट घनत्व निर्धारित करने के क्रम में थ्रेशहोल्ड मूल्य से अधिक है या नहीं, श्रेणीबद्ध नमूना प्रक्रिया आमतौर पर प्रयोग की जाती है। यह पाया गया है कि मकई के रुटवर्म नमूनों के लिए निश्चित नमूना आकार प्रक्रिया से अधिक नमूने के आकार में 48–55% की कमी तथा 55% की नमूने समय में कमी है। प्रति नमूना यूनिट कीटों की संख्या सामान्यतः नकारात्मक द्विपद वितरण का अनुकरण करती है। कीट नियन्त्रण उपाय के प्रकोप की डिग्री आर्थिक थ्रेशहोल्ड मान है। श्रेणीबद्ध परीक्षण से अधिक कारगर है तथा आर्थिक तरीके को निर्धारित करना जब कम से कम समय प्रेक्षणों के आधार पर कीटनाशक प्रयोग किये जाते हैं। औसलेन एवं उनके सहयोगी (1987) ने पाया कि मकई के कम से कम दो वर्ष से विभिन्न मकई के खेतों से प्रथम वर्ष मकई के खेत के पैरामीटर मानों का प्रसरण किया। फाउलर एवं उनके सहयोगी (1987) के अनुसार कीट प्रबंधन में योजना नमूने के लिये (एस पी आर टी) श्रेणीबद्ध प्रायिकता अनुपात परीक्षण लागू किया जाना चाहिये। हुबर्ड तथा आलेन (1991) ने एक नकारात्मक द्विपद वितरण के माध्य के लिये (एस पी आर टी) रोबस्टनेस के बारे में प्रथम बार शायद अध्ययन किया जब प्रकरण पैरामीटर में एक परिवर्तन आया।

### पद्धति

पाटिल, जी.पी., 2002 के अनुसार पक्षीय आकार नकारात्मक द्विपद वितरण, भारित वितरक से ब्युतपन्न किया गया है। यादृच्छिक चर (rv)  $X$ ,  $\alpha$  एवं  $m$

मापदंडों से आकार पक्षीय नकारात्मक द्विपद वितरण का अनुपालन करता है तथा यह माना जाता है कि प्रायिकता द्रव्य फलन (pmf) के द्वारा दिया गया है।

$$P(X=x) = \binom{m+x-1}{x-1} \alpha^{x-1} (1-\alpha)^{m+1}; x=1,2,\dots \quad (1)$$

where  $0 < \alpha < 1, m > 0$ .

हम कल्पना करते हैं कि  $\alpha$  अज्ञात एवं  $m$  ज्ञात चर हैं। माना कि हम शून्य परिकल्पना  $H_0: \alpha = \alpha_0$  का परीक्षण साधारण वैकल्पिक परिकल्पना  $H_1: \alpha = \alpha_1$  ( $\alpha_1 > \alpha_0$ ) के संदर्भ में करना चाहते हैं तो समीकरण (1) में  $n$  आकार के लिए  $\underline{X} = (X_1, X_2, \dots, X_n)$  एक यादृच्छिक नमूना दिया गया है।

यदि  $(\alpha^*, \beta^*)$  (SPRT) की सुदृढ़ता (strength) है तो हम दो संख्याओं  $A \approx \frac{1-\beta^*}{\alpha^*}$  एवं  $B \approx \frac{\beta^*}{1-\alpha^*}$  का चयन करते हैं जैसा कि  $0 < B < 1 < A$ .  $n$ वें स्टेज के नमूने में स्वीकृत करते हैं if  $\sum_{i=1}^n Z_i \leq \ln B$  i.e.

$$\sum_{i=1}^n X_i < \frac{\log_e B}{\log_e \left( \frac{\alpha_1}{\alpha_0} \right)} + n \frac{\log_e \left[ \frac{\alpha_1 \{1-\alpha_0\}^{m+1}}{\alpha_0 \{1-\alpha_1\}} \right]}{\log_e \left( \frac{\alpha_1}{\alpha_0} \right)}$$

और  $H_0$  अस्वीकृत करते हैं यदि

$$\sum_{i=1}^n Z_i \geq \ln A \text{ i.e. } \sum_{i=1}^n X_i \geq \frac{\log_e A}{\log_e \left( \frac{\alpha_1}{\alpha_0} \right)} + n \frac{\log_e \left[ \frac{\alpha_1 \{1-\alpha_0\}^{m+1}}{\alpha_0 \{1-\alpha_1\}} \right]}{\log_e \left( \frac{\alpha_1}{\alpha_0} \right)}$$

तथा (n+1) वें परीक्षणों के नमूनों के लिए

$$\ln B < \sum_{i=1}^n Z_i < \ln A \text{ i.e.}$$

$$\frac{\log_e B}{\log_e \left( \frac{\alpha_1}{\alpha_0} \right)} + n \frac{\log_e \left[ \frac{\alpha_1}{\alpha_0} \left\{ \frac{1-\alpha_0}{1-\alpha_1} \right\}^{m+1} \right]}{\log_e \left( \frac{\alpha_1}{\alpha_0} \right)} < \sum_{i=1}^n X_i <$$

$$\frac{\log_e A}{\log_e \left( \frac{\alpha_1}{\alpha_0} \right)} + n \frac{\log_e \left[ \frac{\alpha_1}{\alpha_0} \left\{ \frac{1-\alpha_0}{1-\alpha_1} \right\}^{m+1} \right]}{\log_e \left( \frac{\alpha_1}{\alpha_0} \right)}$$

(SPRT) के फलन (OC) के आरम्भिक लक्षण दिये गये हैं के द्वारा

$$L(\theta) \approx \frac{A^h - 1}{A^h - B^h}, \quad (2)$$

जहाँ h समीकरण का अशून्य हल है।

$$E(e^{hZ_i}) = 1,$$

$$E \left[ \left( \frac{\alpha_1}{\alpha_0} \right)^{(\alpha_1 - 1)h} \left( \frac{1 - \alpha_1}{1 - \alpha_0} \right)^{(m+1)h} \right] = 1. \quad (3)$$

समीकरण (3) एवं (1) से हम प्राप्त करते हैं कि—

$$\left[ \frac{1 - \alpha}{1 - \alpha \left( \frac{\alpha_1}{\alpha_0} \right)^h} \right]^{m+1} \left( \frac{1 - \alpha_1}{1 - \alpha_0} \right)^{(m+1)h} = 1 \quad (4)$$

ASN फलन निम्न प्रकार दिया गया है

$$E_\alpha(N) \approx \frac{L(\alpha) \ln B + [1 - L(\alpha)] \ln A}{E(Z) = \frac{\alpha(m+1)}{1-\alpha} \ln \left( \frac{\alpha_1}{\alpha_0} \right) + \left[ \ln \left\{ \frac{\alpha_0(1-\alpha_1)^{m+1}}{\alpha_1(1-\alpha_0)^{m+1}} \right\} \right]} \quad (5)$$

### परिणाम एवं चर्चा

डाटासेट: एफिड सरसों पर आँकड़े भारतीय कृषि अनुसंधान संस्थान, नई दिल्ली के कीट विज्ञान प्रभाग

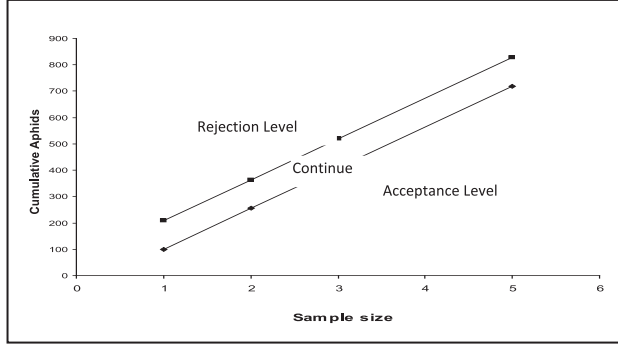
द्वारा विभिन्न फसलों पर प्रति पौधे के आँकड़े लिये गये हैं। हमने बुवाई के 100 दिनों के पश्चात् सरसों की फसल के 35 एफिड से प्रभावित पौधों पर विचार किया गया है।

उपरोक्त आँकड़ों का प्रयोग करके, हम उम्मीद करते हैं कि अज्ञात पैरामीटर के परीक्षण करने की कल्पना करके जब पक्षीय आकार (size-biased) के नकारात्मक द्विपद श्रेणीबद्ध प्रायिकता अनुपात परीक्षण बनाया जाता है तो ज्ञात पैरामीटर m का प्रयोग करते हैं। माना कि हम 25 से अधिक एफिड के प्रत्येक पौधे जो कि आर्थिक समर्थन स्तर है पूर्व वर्षों के आँकड़ों के ज्ञात मानों के आधार पर जबतक 24 पौधे प्राप्त नहीं करते हैं निरंतर परीक्षण करते हैं। आँकड़ों से हमारे पास m=11 पक्षीय आकार (size-biased) नकारात्मक द्विपद वितरण के लिए अज्ञात पैरामीटर  $\alpha$  की परिकल्पना को प्रकाशित करने के क्रम में, हम R-साफ्टवेयर का प्रयोग करके अज्ञात  $\alpha$  पैरामीटर के समान आकलन की अधिकतम गणना करते हैं। m के अज्ञात मान के आधार पर हम अधिकतम समान आकलन प्राप्त करते हैं। जैसे  $\alpha = 0.944$ , मानक त्रुटि = 0.00265 तथा  $-2 \log L = 1221.385$  हैं।

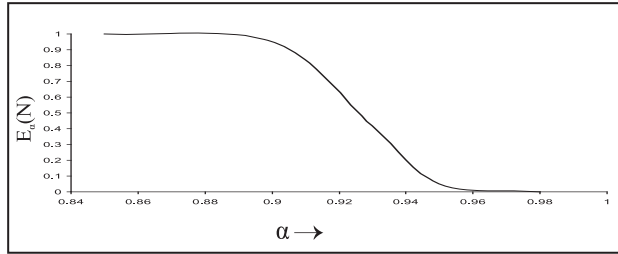
माना कि हम ज्ञात पैरामीटर 'm' = 11 से एफिड से प्रभावित सरसों के लिये अगले वर्ष समान प्लॉट पर श्रेणीबद्ध परीक्षण प्रक्रिया संचालित करना चाहते हैं (प्रति पौधा का ETL मान एफिड की संख्या से कम है)। हम परिकल्पना प्रकाशित करते हैं जोकि  $H_0: \alpha_0 = 0.90$  Vs  $H_1: \alpha_1 = 0.95$  के अन्तर्गत पक्षीय आकार (size-biased) नकारात्मक द्विपद वितरण के लिए संघटनात्मक श्रेणीबद्ध परीक्षण प्रक्रिया का प्रयोग करके हम पाते हैं कि 100 दिन की बुवाई के बाद एफिड सरसों के प्रति परिणाम रेखाएं हैं जो कि चित्र 1 में दिखाई गई हैं तथा संगत फलन OC तथा ASN चित्र 2 तथा चित्र 3 में दिये हैं। माना प्रथम नमूना पर एफिड की संख्या 500 है, तब चित्र 1 के अनुसार, हम शून्य परिकल्पना को रिजेक्ट (reject) करते हैं तथा एफिड सरसों के प्रति कीट नियंत्रक माप को लेते हैं। इसके विपरीत प्रथम नमूने पर एफिड की संख्या 40 है तब आकलित औसत नमूना संख्या के अनुसार निरंतर



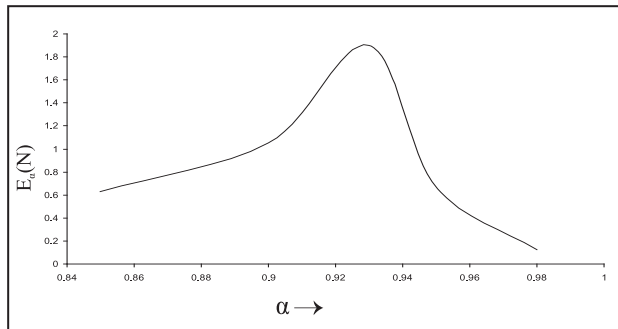
नमूने लेते हैं जैसा चित्र-1 एवं चित्र-3 में दिखाया गया है।



चित्र-1 : सरसों एफिड की बुवाई के 100 दिनों के उपरान्त आकार पक्षीय नकारात्मक द्विपद वितरण का उपयोग करके निर्णय रेखाएं।



चित्र-2: आकार पक्षीय नकारात्मक द्विपद वितरण का पैरामीटर के संबंध में परिकल्पना के परीक्षण के लिए श्रेणीबद्ध परीक्षण के ओ सी फलन।



चित्र-3: आकार पक्षीय नकारात्मक द्विपद वितरण का पैरामीटर के संबंध में परिकल्पना के परीक्षण के लिए श्रेणीबद्ध परीक्षण के ए एस एन ।

## संदर्भ

1. फोलर, जी. डब्ल्यू एवं लिंच, ए.एम. (1987): "कीट प्रबंधन में नमूना योजना वाल्ड की श्रेणीबद्ध प्रायिकता अनुपात परीक्षण पर आधारित है", पर्यावरण कीट विज्ञान, 16, 345-354।
2. हबर्ड, डी.जे. एवं एलेन, ओ.बी. (1991): "एस.पी. आर.टी. (SPRT) की सुदृढता का नकारात्मक द्विपद का मिसस्पेसीफिकेशन का प्रसारण मानक", बायोमेट्रिक्स, 47, 419-427।
3. मेक औसलेन, एच.जे., एलिस, सी.आर. एवं एलेन, ओ.बी. (1987): "दक्षिण ऑंटेरियो में वयस्क उत्तरी और पश्चिमी मकई जड़ के कीड़े की श्रेणीबद्ध नमूना", कनाडा के कीट विज्ञानी, 119, 577-585।
4. मीर, के.ए. (2009): "आकार पक्षीय नकारात्मक द्विपद वितरण एवं शून्य सूक्ष्म मामलों में इसके उपयोग", मापविज्ञान की समीक्षा, 9(1), 33-35।
5. पाटिल, जी.पी. (2002): "भारित वितरण", पर्यावरणीय विश्वकोष, 4, 2369-2377।
6. सुब्रमणियम, आर. एवं आरुमुगम, एन. (2006): "कपास एफिड के प्रबंधन में निर्णय लेने के लिए श्रेणीबद्ध नमूना तकनीक", कीट विज्ञान जर्नल, 3(3) : 254-260।

□

## प्रत्याशित—मान प्राचलों का उपयोग करते हुए आर्डर वन के स्वसमाश्रयी में फॉक्स मॉडल का संयोजन

अमृत कुमार पॉल, नाओरेम ओकेन्द्रो सिंह, रंजीत पॉल, बिशाल गुरुंग एवं ऊषा जैन

एक विशेष जलाशय (water body) के लिए मत्स्य स्टॉक की बहुतायत, पर्यावरणीय कारकों और मत्स्य स्टॉक गुणों के बीच अन्योन्यक्रिया का फलन है। स्टॉक में पर्यावरणीय दशाओं के एक विशेष सेट पर स्थिर होने की प्रवृत्ति होती है (गुलन्द, 1977)। जब अतिरेक (surplus) उत्पादन को हार्वेस्ट नहीं किया जाता, अधिकतम मत्स्य स्टॉक आकार के स्तर पर स्टॉक में अतिरिक्त विकास प्राकृतिक नश्वरता की क्षतिपूर्ति करने के लिए काफी पर्याप्त होती है तथा अतिरेक उत्पादन शून्य के बराबर होती है (हैडन, 2001)। इसका यह तात्पर्य हुआ कि मत्स्य योजनाओं को अतिरेक उत्पादन के रूप में अभिव्यक्त किया जा सकता है; वे बहुत लचीले हैं तथा उनमें विभिन्न विचरण होते हैं। यद्यपि, शेफर मॉडल सबसे अधिक प्रयुक्त अतिरेक उत्पादन मॉडल है, यह सममितीय पौराबोला (symmetrical parabola) वक्र उपलब्ध कराता है, जिसका प्राकृतिक फिनोमिना के प्रति हमेशा सही (true) होना आवश्यक नहीं है, क्योंकि पारिस्थितिक प्रणाली में कालिक (temporal) एवं स्पेशियल परिवर्तनों के आधार पर अधिकतम अक्षुण्ण उपज (MSY) में काफी परिवर्तन हो सकता है। फिर भी, फॉक्स मॉडल में मत्स्य की पकड़ बढ़ते प्रयास के साथ कुल बायोमास की ओर असममितीय रूप से बढ़ती है, जो शेफर मॉडल की अपेक्षा अधिक वास्तविक है, अतः, जीव-विज्ञानी फॉक्स मॉडल को वरीयता देते हैं। इसके अतिरिक्त, जब हम कैच-एफर्ट फिशरीज डाटा पर कार्य कर रहे होते हैं, जिसका सतत् समायावधि में अवलोकन किया गया है,

डाटा प्वाइन्ट्स आमतौर पर आपस में सहसम्बन्धित होते हैं। प्रज्ञेष्णु एवं रवीचन्द्रन (2003) ने प्राचल आकलकों के बीच अधिक सहसम्बन्ध को कम करने के लिए मत्स्य में अरैखिक फॉक्स मॉडल के संयोजन के लिए प्राचलों के पुनः प्राचलीकरण की महत्ता पर बल दिया है। अतः, प्रेक्षणों में स्वसहसम्बन्ध की उपस्थिति की जाँच द्वारा इस अध्ययन में प्रत्याशित—मान प्राचलों का प्रयोग करते हुए क्रमानुसार सहसम्बन्धित त्रुटि संरचना के साथ फॉक्स मॉडल के संयोजन की एक विधि का प्रस्ताव दिया गया है, जिसकी गोबिन्दसागर जलाशय, भारत के कैच-एफर्ट डाटा को लेकर उदाहरण के साथ व्याख्या की गयी है।

### सामग्री एवं विधियाँ

यह फॉक्स मॉडल (1970) MSY के आकलन तथा MSY ( $E_{MSY}$ ) की हार्वेस्टिंग के लिए इष्टतम मत्स्य प्रयास को सरल बनाता है। सन्तुलन फॉक्स मॉडल इस प्रकार दिया गया है :

$$C_t = KE_t \exp(-E_t/r), \quad (1)$$

जहाँ,  $t$  समय पर  $C_t$  एवं  $E_t$  कैच एवं प्रयास हैं,  $r$  मूलभूत विकास दर है;  $K$  ले जाने की क्षमता है तथा तदनुसार,  $MSY = rK/e$ ; ( $e=2.71828$ ) एवं  $E_{MSY} = r$ ।

चूँकि, हमारा सम्बन्ध काल-श्रेणी आँकड़ों से है, अतः त्रुटि टर्म की स्वतन्त्र अभिधारणा की जाँच द्वारा उपरोक्त मॉडल की वैधता की जाँच करना आवश्यक

है। इस उद्देश्य के लिए डरबिन-वॉटसन परीक्षण का प्रयोग किया जा सकता है, जो इस अभिधारणा पर आधारित है कि त्रुटियाँ ( $\varepsilon_t$ 's) आर्डर-वन के स्वसमाश्रयी का अनुसरण करती हैं। अनुवर्ती परीक्षण सांख्यिकी 'd' को निम्नानुसार परिभाषित किया गया है :

$$d = \frac{\sum_{t=2}^n (\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1}^n \varepsilon_t^2}, \quad 0 \leq d \leq 4. \quad (2)$$

सांख्यिकी 'd' का मान 0 एवं 4 के बीच होता है। 2 के करीब 'd' का मान बहुत कम स्वसहसम्बन्ध इंगित करता है; 0 की ओर मान सकारात्मक सहसम्बन्ध दर्शाता है जबकि 4 की ओर मान नकारात्मक सहसम्बन्ध दर्शाता है।

ऐसी स्थिति को हैण्डल करने के लिए जब स्वसहसम्बन्ध की उपस्थिति का प्रमाण होता है, आर्डर-वन की स्वसमाश्रयी (AR) त्रुटि टर्म  $\varepsilon_t$  उपरोक्त समीकरण (1) की दायें ओर जोड़ दिया जाता है :

$$\varepsilon_t = \Phi \varepsilon_{t-1} + u_t; \quad |\Phi| < 1, \quad (3)$$

जहाँ,  $u_t$  शून्य माध्य एवं सतत् चर के साथ स्वतन्त्र एवं सामान्य रूप से वितरित है तथा  $\Phi$  स्वसमाश्रयी प्राचलों को दर्शाता है। एक ए.आर (1) एडिटिव एरर संरचना शामिल करने से फॉक्स मॉडल निम्न प्रकार का हो जाता है :

$$C_t = KE_t \exp(-\frac{E_t}{r}) + \Phi \varepsilon_{t-1} + u_t. \quad (4)$$

### ए.आर(1) के साथ फॉक्स मॉडल का आंशिक प्राचलीकरण

जब हम प्राचलों के आकलन के लिए अरैखिक आकलन प्रक्रियाओं का प्रयोग करते हैं, उच्च प्राचल सहसम्बन्ध का पता चलता है, जो कि अवांछनीय है। रतकौस्की (1990, पृष्ठ 43) ने व्याख्यात्मक चर के मानों का चयन करने के लिए प्रत्याशित मान प्राचलों के उपयोग का सुझाव इस प्रकार दिया जिससे प्राचल सहसम्बन्ध को निम्न किया जा सके। प्रत्याशित

मान-प्राचल, डाटा की प्रेक्षित रेंज में होने चाहिए न कि डाटा रेंज के बाहर उपगामी अथवा बहिर्वेक्षण के प्रति अनुवर्ती, क्योंकि बाहर प्रेक्षित आँकड़ों की रेंज कम दक्ष है। इसके अतिरिक्त, प्रत्याशित-मान प्राचलों का प्रयोग आसानी से किया जा सकता है क्योंकि ये लगभग अनभिनत, प्रसामान्य रूप से वितरित तथा न्यूनतम चर आकलक होते हैं।

गणितीय नोटेशन्स के लिए सरल होने के लिए, उपरोक्त समीकरण (4) को निम्न रूप में पुनः लिखा जाता है :

$$C = KE \exp(-\frac{E}{r}) + \Phi \varepsilon. \quad (5)$$

मॉडल के सभी प्राचलों को प्रत्याशित-मान प्राचलों से प्रतिस्थापित करना हमेशा आवश्यक नहीं होता, तथापि, केवल आक्रामी (offending) प्राचलों को प्रतिस्थापित किया जा सकता है। अतः, उपरोक्त समीकरण (5) के  $r$  एवं  $K$  के लिए प्रत्याशित-मान प्राचल प्राप्त करने के लिए हमें  $E$  की प्रेक्षित रेंज में व्याख्यात्मक चर के  $E_1$  एवं  $E_2$  मान चुनने की आवश्यकता है तथा तदनुसार  $\varepsilon$  के  $\varepsilon_1$  एवं  $\varepsilon_2$  मानों का चयन किया जाना है। तब हमें समीकरण (5) से निम्नलिखित प्रत्याशित मान प्राप्त होते हैं :

$$C_1 = KE_1 \exp(-\frac{E_1}{r}) + \Phi' \varepsilon_1, \quad (6)$$

एवं

$$C_2 = KE_2 \exp(-\frac{E_2}{r}) + \Phi' \varepsilon_2. \quad (7)$$

समीकरण (6) एवं (7) का हल करने पर हमें प्राप्त होता है :

$$r = \frac{(E_2 - E_1)}{\log\{E_2(C_1 - \Phi' \varepsilon_1) / [E_1(C_2 - \Phi' \varepsilon_2)]\}}$$

एवं

$$K = \frac{(C_1 - \Phi' \varepsilon_1)}{E_1 \exp[-E_1 \log\{E_2(C_1 - \Phi' \varepsilon_1) / (E_1(C_2 - \Phi' \varepsilon_2))\}] / (E_2 - E_1)}$$

बशर्ते  $E_2 > E_1$ ,  $C_1 > \Phi' \varepsilon_1$  and  $E_2 (C_1 - \Phi' \varepsilon_1) > E_1 (C_2 - \Phi' \varepsilon_2)$ , 'r' एवं 'k' प्राचल सकारात्मक होने चाहिए।

मूल समीकरण (5) में पुनः प्रतिस्थापित करने पर, हमें प्राप्त होता है :

$$C = \{(C_1 - \Phi' \epsilon_1)/E_1\} E \cdot \exp[\log\{E_2(C_1 - \Phi' \epsilon_1)/(E_1(C_2 - \Phi' \epsilon_2))\} (E_1 - E)/(E_2 - E_1)] + \Phi \epsilon \quad (8)$$

जहाँ,  $\Phi'$ ,  $\Phi$  के प्रत्याशित मान हैं जो उपरोक्त समीकरण (4) के संयोजन से प्राप्त हुआ है। अब इस प्रक्रिया ने मूल प्राचलों  $r$  एवं  $K$  को समाप्त कर दिया है जो नए प्राचलों द्वारा प्रतिस्थापित हैं, जो प्रत्याशित-मान प्राचल होने के कारण है। एक ही मूल मॉडल के विभिन्न प्राचलीकरण एक ही संमजन सुष्ठुता (goodness-of-fit) एवं एक ही संयोजित मान उपलब्ध कराएँगे, हालांकि, उनके आकलन व्यवहार में बहुत अंतर हो सकता है (रतकौस्की, 1990, पृष्ठ 30)। उपरोक्त अरैखिक मॉडलों में अज्ञात प्राचलों एवं स्वसमाश्रयी प्राचल का आकलन लीवेनबर्ग-मारक्वार्डेट विधि का प्रयोग करते हुए किया गया है (सीबर एवं वील्ड, 1989)।

### मॉडल की पर्याप्तता के माप

प्रायः निम्नलिखित सांख्यिकी से इसका मूल्यांकन किया जाता है :

(i) वर्ग-माध्य-मूल त्रुटि, (Root Mean Square Error),

$$RMSE = \left[ \frac{\sum_{t=1}^n (C_t - \hat{C}_t)^2}{n} \right]^{1/2}$$

(ii) माध्य निरपेक्ष त्रुटि, (Mean Absolute Error)

$$MAE = \frac{\sum_{t=1}^n |C_t - \hat{C}_t|}{n}, \quad t = 1, 2, \dots, n,$$

जहाँ,  $n$  प्रेक्षणों की संख्या है;  $\hat{C}_t$ , समय  $t$  पर प्रागुक्त फिश-कैच है।

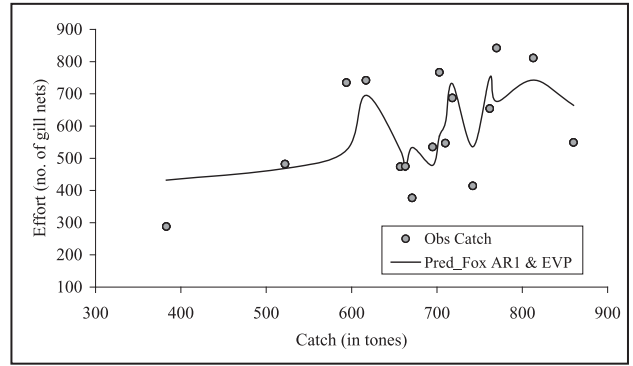
बेहतर संयोजित-मॉडल में उपरोक्त सांख्यिकी के न्यूनतम मान होंगे। इसके अतिरिक्त, विकसित किए जाने वाले मॉडल के लिए तैयार की गयी अभिधारणाओं की जाँच के लिए त्रुटि विश्लेषण की आवश्यकता है। अतः, त्रुटियों की स्वतंत्र अभिधारणाओं की जाँच की आवश्यकता है। त्रुटियों की स्वतंत्र अभिधारणा की जाँच के लिए साहित्य में रन-टेस्ट प्रक्रिया उपलब्ध है (रतकौस्की, 1990)। हालांकि, अरैखिक मॉडलों के चयन के लिए प्रसामान्यता अभिधारणा इतनी सख्त नहीं है क्योंकि उनकी त्रुटियाँ प्रसामान्य वितरण का अनुसरण न कर सकती हों।

### परिणाम एवं परिचर्चा

उपरोक्त पद्धति की व्याख्या के लिए कौशल, इत्यादि (2006) द्वारा प्रेक्षित 1974-75 से 1989-90 के दौरान गोबिन्दसागर जलाशय में फिश-कैच-एफर्ट और इसकी सहवर्ती उपज अथवा पकड़ के आँकड़ों को लिया गया है। उपरोक्त मॉडलों के संयोजन के लिए DCFR, भीमताल में उपलब्ध सामाजिक-विज्ञान के लिए सांख्यिकीय पैकेज 12.0 वर्जन का उपयोग किया गया है। अरैखिक मॉडलों के श्रेष्ठ संयोजन के लिए वैश्विक अभिसरण कसौटी को पूरा करने के लिए आरम्भिक प्राचल मानों के विभिन्न सेटों को डालने का प्रयास किया गया। MSY एवं फॉक्स मॉडल के विभिन्न रूपों के सहवर्ती इष्टतम प्रयासों की संगणना की गयी तथा संयोजित मॉडलों के परिणाम तालिका-3 में दिए गए हैं। फॉक्स मॉडल द्वारा आकलित MSY मान (टनों में) 974 है तथा इसके सहवर्ती इष्टतम प्रयास (गिल नेट्स की संख्या) 2252 है। यद्यपि मॉडल के लिए फिश-कैच में त्रुटि टर्म के संबंध में प्रसामान्यता अभिधारणा पूरी होती है क्योंकि संयोजित मॉडलों के शेपिरो-विल्क परीक्षण  $p$ -मान 0.125 है, हालांकि यादृच्छिकता अभिधारणा का अनुसरण नहीं होता, क्योंकि 5 प्रतिशत सार्थकता स्तर पर रन परीक्षण  $|Z|$  मान (2.303) प्रसामान्य वितरण के 1.96 से अधिक है। इसके अतिरिक्त, डरबिन-वॉटसन स्टैटिस्टिक की गणना की गयी है तथा स्टैटिस्टिक मान शून्य की ओर है और इसलिए सकारात्मक स्वसहसंबंध का अनुमान है।

ए.आर(1) त्रुटि संरचना को शामिल करते हुए फॉक्स मॉडल को पुनःसंयोजित किया गया तथा परिणाम तालिका 3 में दिए गए हैं। इसके अतिरिक्त, फिश-कैच की त्रुटियों में रन-परीक्षण एवं शेपिरो-विल्क परीक्षण लागू किए गए हैं। यहाँ पर पुनःसंयोजित मॉडल के रन-परीक्षण मान तथा शेपिरो-विल्क परीक्षण p-मान इंगित करते हैं कि यादृच्छिकता एवं प्रसामान्यता अभिधारणा, दोनों संतुष्ट हैं। पुनःसंयोजित मॉडल से परिकलित डरबिन-वॉटसन स्टैटिस्टिक मान 2 के बहुत करीब थे, अर्थात् स्वसहसंबंध की उपस्थिति नकारात्मक है। इसके अलावा, मूल फॉक्स मॉडल की तुलना में पुनः संयोजित मॉडल में RMSE एवं MAE के मानों में महत्वपूर्ण सुधार देखा गया। हालांकि, तालिका 1 में दी गयी सहसंबंध मैट्रिक्स दर्शाती है कि 'r' एवं 'K' प्राचलों के बीच सहसंबंध गुणांक के चरम मान अर्थात्  $\rho(r, K) = -0.991$  इंगित करते हैं कि दोनों प्राचल स्वतंत्र रूप से आकलित नहीं हैं, जबकि  $\rho(r, \Phi)$  एवं  $\rho(K, \Phi)$  के मान स्वीकार्य हैं। इसका संभावित हल प्राप्त करने के लिए समीकरण 8 संयोजित करने का प्रयास किया गया है, जो 'r' एवं 'K' के लिए प्रत्याशित मान प्राचलों सहित समीकरण 5 से प्राप्त की गयी है, चूँकि वे मॉडल के सम्भावित आक्रामी प्राचल हैं जबकि  $\Phi$  प्राचल को अपरिवर्तित रखा जा रहा है।  $\epsilon_1 = -62.0549$  एवं  $\epsilon_2 = 162.0726$  न्यूनतम सहसंबंध गुणांक के संदर्भ में बेहतर परिणाम देती है तथा 'r' एवं 'K' की सकारात्मक होने की अपेक्षित शर्तें भी पूरी करती है। व्यवहार में, E1 एवं E2 का चयन इस प्रकार किया जाता है कि वे एक दूसरे के करीब न हों।  $C_1$  एवं  $C_2$  का प्राचलों के अन्तिम आकलकों की गणना के लिए C के समवर्ती मानों अर्थात्  $C_1 = 735$  एवं  $C_2 = 811$  को आरम्भिक मानों के रूप में लिया जाता है। प्राचल-आकलक तालिका 3 में दिए गए हैं। अब, सहसंबंध गुणांक पूरी तरह से स्वीकार्य है क्योंकि प्राचलों के बीच सहसंबंध गुणांक बहुत कम हैं, जो तालिका 2 में प्रस्तुत है। अतः, हम लगभग कह सकते हैं कि प्राचल स्वतंत्र रूप से आकलित हैं। प्रत्याशित मान प्राचलों के समीकरण 8 सहित मूल मॉडल के समीकरण 4 तथा रूपांतरित मॉडल से परिकलित RMSE एवं MAE के मान वैसे

ही रहे, जैसाकि तालिका 3 में दिया गया है। प्रेक्षित फिश-कैच सहित संयोजित मॉडल का ग्राफ आरेख 1 में प्रस्तुत है। फॉक्स मॉडल के विभिन्न रूपों के लिए MSY के आकलनों के अध्ययन से पता चलता है कि ए.आर(1) सहित फॉक्स मॉडल के पुनः प्राचलीकरण द्वारा आकलित MSY 691 टन एवं इष्टतम प्रयास (1217 गिल नेट्स की संख्या) के मानों की तुलना में सरल फॉक्स मॉडल ने MSY एवं इष्टतम प्रयास मानों को अतिआकलित किया है।



**आरेख 1 :** प्रत्याशित-मान प्राचलों का प्रयोग करते हुए गोबिन्दसागर जलाशय से प्रेक्षित कैच-एफर्ट के डाटासेट में ए.आर.(1) के साथ फॉक्स मॉडल का संयोजन

यह लेख आर्डर-वन के स्वसमाश्रयी सहित त्रुटि टर्म शामिल कर प्रत्याशित-मान प्राचलों का प्रयोग करते हुए फॉक्स मॉडल के संयोजन को प्रस्तुत करता है। इसकी उदाहरण के साथ व्याख्या की गयी है, जिसमें गोबिन्दसागर जलाशय (reservoir) से प्राप्त क्रमानुसार सहसम्बन्धित कैच-एफर्ट डाटा को लिया गया है, जो इसे प्राप्त करने के लिए इष्टतम प्रयास सहित एम.एस.वाई. (MSY) के आकलन प्राप्त करने की ओर ले जाता है।

**तालिका 1 :** ए.आर.(1) के साथ फॉक्स मॉडल संयोजित करने के पश्चात प्राचल आकलकों का उपगामी सहसंबंध मैट्रिक्स

प्राचल	K	r	$\Phi$
K	1.000		
r	-0.991	1.000	
$\Phi$	0.279	-0.274	1.000

**तलिका 2 :** प्रत्याशित-मान प्राचलों का प्रयोग करते हुए ए.आर.(1) के साथ फॉक्स मॉडल संयोजित करने के पश्चात प्राचल आकलकों का उपगामी सहसंबंध मैट्रिक्स

प्राचल	$C_1$	$C_2$	$\Phi$
$C_1$	1.000		
$C_2$	-0.140	1.000	
$\Phi$	0.241	-0.175	1.000

**तलिका 3 :** गोबिन्दसागर जलाशय के प्रेक्षित कैच-एफर्ट आँकड़ों में फॉक्स मॉडल के विभिन्न रूपों के संयोजन हेतु सारांश सांख्यिकी

	फॉक्स मॉडल	ए.आर.(1) के साथ फॉक्स मॉडल	ए.आर.(1) के साथ फॉक्स मॉडल का प्राचलीकरण
<b>प्राचल</b>			
K	1.176 *(0.556)	1.544 (0.616)	-
r	2252.098 (3374.256)	1216.731 (837.606)	-
C1	-	-	525.568 (45.127)
C2	-	-	740.577 (53.964)
$\Phi$	-	0.614 (0.245)	0.614 (0.245)
<b>सांख्यिकी</b>			
MSY (टन में)	974	691	691
$E_{MSY}$ (गिल नेट्स की संख्या)	2252	1217	1217
डरबिन-वाट्सन सांख्यिकी	0.847	1.735	1.735

	फॉक्स मॉडल	ए.आर.(1) के साथ फॉक्स मॉडल	ए.आर.(1) के साथ फॉक्स मॉडल का प्राचलीकरण
<b>मॉडल की सटीकता</b>			
RMSE	140.412	115.264	115.264
MAE	123.696	97.795	97.795
<b>अवशिष्ट विश्लेषण</b>			
रन परीक्षण ( Z )	2.303	0.259	0.259
शेपिरो-विल्क परीक्षण p-मान	0.125	0.231	0.231

\*कोष्ठक में दिये गये मान उपगामी मानक त्रुटि के अनुरूप हैं ।

#### संदर्भ

1. फॉक्स, डब्ल्यू. एवं डब्ल्यू. जे.आर. (1970) : "एन एक्सपोनेन्शियल यील्ड मॉडल फॉर ऑप्टिमाइजिंग एक्स्प्लॉइटेड फिश पॉपुलेशन", ट्रॉस. एम. फिश सोसा., 99 : 80-88 ।
2. गुलन्द, जे.ए. (1977) : "दि प्रोब्लम्स ऑफ पॉपुलेशन डायनेमिक्स ऐण्ड कन्टम्पेरी फिशरिज़ मैनेजमेन्ट", फिश पॉपुलेशन डायनेमिक्स : दि इम्प्लीकेशन फॉर मैनेजमेन्ट, खण्ड XI (एड. जे. ए. गुलन्द), 383-406, जोहन विले, चीचेस्टर ।
3. हैडन, एम. (2001) : मॉडलिंग ऐण्ड क्वान्टिटेटिव मैथड्स इन फिशरिज़, चैपमैन ऐण्ड हॉल, बोसा रेटॉन, फ्ला ।
4. कौशल, डी.के., शर्मा, वी.के., कोहली, एम.पी.एस., झा, बी.सी. एवं सिंह एच.पी. (2006) : "प्रोडक्शन डायनेमिक्स ऑफ गोबिन्दसागर रिसर्वायर ऑन रिवर सतलुज, हिमाचल प्रदेश", फिशरीज ऐण्ड एक्वाकल्चर इन इन्डस रिवर रीजन (एड. एम.पी. सिंह कोहली), 100-113, इन्डियन सोसाइटी ऑफ फिशरीज प्रोफेशनल्स, मुम्बई, भारत ।

5. प्रज्ञेष्णु एवं रविचन्द्रन, एस. (2003) : "ए मैथड फॉर फिटिंग ऑफ नॉन-लीनियर फॉक्स मॉडल इन फिशरीज", इन्ड. जर्न. एनि. साई., 73(3): 329-331।
6. रतकौस्की, डी.ए. (1990) : "हैण्डबुक ऑफ नॉन-लीनियर रिग्रेशन मॉडल्स", मार्सेल डेक्कर, न्यूयार्क।
7. सीबर, जी.ए.एफ. एवं विल्ड, सी.जे (1989): "नॉन-लीनियर रिग्रेशन", जोहन विले ऐण्ड सन्स।

### परिशिष्ट

**व्युत्पत्ति : ए.आर (1) के साथ फॉक्स मॉडल के 'r' एवं 'K' के लिए प्रत्याशित-मान प्राचल**

आर्डर-वन के स्वसमाश्रयी सहित फॉक्स मॉडल :

$$C = KE \exp\left(-\frac{E}{r}\right) + \Phi \varepsilon. \quad (i)$$

उपरोक्त समीकरण (i) 'r' एवं 'K' के लिए प्रत्याशित-मान प्राचल प्राप्त करने के लिए E की प्रेक्षित रेंज में से घातांकी चर E के E1 एवं E2 मानों का चयन करना होगा तथा तदनुरूप,  $\varepsilon_1$  एवं  $\varepsilon_2$  के मान  $\varepsilon$  चुनना। तब, हम समीकरण (i) से निम्नानुसार प्रत्याशित-मान प्राप्त कर सकते हैं :

$$C_1 = KE_1 \exp\left(-\frac{E_1}{r}\right) + \Phi' \varepsilon_1, \quad (ii)$$

एवं

$$C_2 = KE_2 \exp\left(-\frac{E_2}{r}\right) + \Phi' \varepsilon_2, \quad (iii)$$

जहाँ,  $\Phi'$  का आकलित मान  $\Phi$  है जो पुनः प्राचलीकरण से पूर्व ए.आर.(1) सहित फॉक्स मॉडल संयोजित करने से प्राप्त हुआ है।

पुनः व्यवस्थित करने के पश्चात समीकरण (ii) को (iii) से विभाजित करने पर हमें प्राप्त होगा :

$$\begin{aligned} \left(\frac{C_1 - \Phi' \varepsilon_1}{C_2 - \Phi' \varepsilon_2}\right) &= \left(\frac{E_1}{E_2}\right) \exp\left(\frac{E_2 - E_1}{r}\right) \\ \Rightarrow \log\left(\frac{C_1 - \Phi' \varepsilon_1}{C_2 - \Phi' \varepsilon_2}\right) &= \left(\frac{E_2 - E_1}{r}\right) \log\left(\frac{E_1}{E_2}\right) \\ \Rightarrow r &= \frac{(E_2 - E_1)}{\log\left[\frac{E_2(C_1 - \Phi' \varepsilon_1)}{E_1(C_2 - \Phi' \varepsilon_2)}\right]} \end{aligned}$$

पुनः व्यवस्थित करने के पश्चात समीकरण (ii) में 'r' के मान को प्रतिस्थापित करने पर हमें प्राप्त होगा :

$$(C_1 - \Phi' \varepsilon_1) = KE_1 \exp\left[-E_1 \log\left\{\frac{E_2(C_1 - \Phi' \varepsilon_1)}{E_1(C_2 - \Phi' \varepsilon_2)}\right\} / (E_2 - E_1)\right]$$

$$\Rightarrow K = \frac{(C_1 - \Phi' \varepsilon_1)}{E_1 \exp\left[-E_1 \log\left\{\frac{E_2(C_1 - \Phi' \varepsilon_1)}{E_1(C_2 - \Phi' \varepsilon_2)}\right\} / (E_2 - E_1)\right]}$$

चूँकि प्राचल 'r' एवं 'K' गैर-नकारात्मक मान हैं, अतः शर्तें a)  $E_2 > E_1$ , (b)  $C_1 > \Phi' \varepsilon_1$  एवं c)  $E_2 (C_1 - \Phi' \varepsilon_1) > E_1 (C_2 - \Phi' \varepsilon_2)$  हर हालत में पूरी होनी चाहिए।

अब, समीकरण (i) एवं (ii) से हमें प्राप्त होता है :

$$\begin{aligned} C &= \frac{(C_1 - \Phi' \varepsilon_1)}{E_1 \exp(-E_1/r)} E \exp(-E/r) + \Phi \varepsilon \\ \Rightarrow C &= (C_1 - \Phi' \varepsilon_1) \frac{E}{E_1} \exp\left\{\frac{(E_1 - E)}{r}\right\} + \Phi \varepsilon. \quad (iv) \end{aligned}$$

उपरोक्त समीकरण (iv) में 'r' का मान प्रतिस्थापित करने पर हमें प्राप्त होता है :

$$\begin{aligned} C &= (C_1 - \Phi' \varepsilon_1) \frac{E}{E_1} \exp\left\{\frac{(E_1 - E)}{(E_2 - E_1)} \log\left[\frac{E_2(C_1 - \Phi' \varepsilon_1)}{E_1(C_2 - \Phi' \varepsilon_2)}\right]\right\} + \Phi \varepsilon \\ \Rightarrow C &= \left\{ (C_1 - \Phi' \varepsilon_1) / E_1 \right\} E \exp\left\{ \log\left[\frac{E_2(C_1 - \Phi' \varepsilon_1)}{E_1(C_2 - \Phi' \varepsilon_2)}\right] (E_1 - E) / (E_2 - E_1) \right\} + \Phi \varepsilon \quad (v) \end{aligned}$$

इस प्रकार, मूल प्राचलों 'r' एवं 'K' को नए प्राचलों से प्रतिस्थापित किया गया है, जो प्रत्याशित-मान प्राचल होने के आधार पर है। यद्यपि, उपरोक्त अभिव्यंजन (expressions) स्वरूप में मूल अभिव्यंजन की अपेक्षा अधिक बोझिल है, प्रत्याशित-मान प्राचल सहित प्राचलीकरण अनेक लाभ प्रदान करता है।

□

## रन अनुक्रमों में न्यूनतम स्तर परिवर्तन सहित द्विस्तरीय बहु-उपादानी परीक्षण

अर्पण भौमिक, एल्दो वरगीस, सीमा जग्गी, सिनी वरगीस,  
कादर अली सरकार एवं विजय बिन्दल

बहु-उपादानी परीक्षणों में दो या दो से अधिक कारकों का, प्रत्येक कारक दो या दो से अधिक स्तरों पर, एक साथ प्रयोग किया जाता है। इस तरह के परीक्षण प्रत्येक कारक के प्रभावों के साथ-साथ अनुक्रिया चर पर कारकों के बीच की अन्योन्यक्रियाओं के प्रभावों के अध्ययन की अनुमति देते हैं। ये परीक्षण पूर्ण या अपूर्ण हो सकते हैं। यदि एक पूर्ण बहु-उपादानी अभिकल्पना में संयोजनों की संख्या तार्किक रूप से व्यवहार्य होने के लिए बहुत अधिक होती है तो एक आंशिक बहु-उपादानी अभिकल्पना जिसमें से कुछ संभव संयोजनों को हटा दिया जाता है, को लिया जा सकता है। अतः आंशिक बहु-उपादानी परीक्षण अभिकल्पनाओं में पूर्ण बहु-उपादानी अभिकल्पना के परीक्षणात्मक रनों के सावधानीपूर्वक चुने हुए उप-समुच्चय निहित होते हैं। कृषि एवं औद्योगिक परीक्षणों में पूर्ण बहु-उपादानी एवं आंशिक बहु-उपादानी परीक्षणों के गहन अनुप्रयोग मिलते हैं। यह सदैव उपयुक्त है कि बहु-उपादानी एवं आंशिक बहु-उपादानी अभिकल्पनाओं के निष्पादन का क्रम यादृच्छिक होना चाहिए।

रन अनुक्रमों का यादृच्छिकीकरण एक तकनीक है जिसे बहु-उपादानी एवं आंशिक बहु-उपादानी अभिकल्पनाओं के प्रयोगकर्ता सामान्यतः समय प्रवृत्ति के परिणामस्वरूप रुचि अनुसार प्रभावों के आकलों में होने वाली अभिनति से बचने के लिए नियोजित करते हैं। अतः अपेक्षित है कि रन अनुक्रमों के यादृच्छिकीकरण करते हुए इस प्रभाव को सभी प्रभावों के बीच इस तरह से प्रसारित करना चाहिए कि यह प्रभाव उनमें से किसी को सार्थक रूप से प्रभावित नहीं करे। हालांकि

बहु-उपादानी परीक्षण में, यादृच्छिकीकरण कारक स्तरों में एक बड़ी मात्रा में परिवर्तन ला सकता है जिसके कारण परीक्षण करना अंततः महंगा, अधिक समय लेने वाला एवं कठिन हो जायेगा। बहुत से कृषि, फसलोत्तर एवं प्रसंस्करण, अभियांत्रिकी तथा औद्योगिक परीक्षणों में अधिक मात्रा में स्तर परिवर्तन प्रयोगकर्ता के लिए गंभीर चिंता का विषय है क्योंकि इन परीक्षणों में ऐसी स्थिति उत्पन्न हो सकती है जबकि कुछ कारकों के स्तरों में परिवर्तन करना प्राकृतिक रूप से अत्यधिक कठिन होता है। ऐसी स्थितियों में अनुशंसित अभ्यास यह है कि जहां किसी विशेष कारक के स्तर में परिवर्तन करना कठिन या लगभग असंभव हो, वहां उस विशेष कारक के स्तर को नियत करके अन्य कारकों के स्तरों (स्तरों के संयोजनों) का यादृच्छिकीकरण करना उचित होगा। इस पद्धति को कई परीक्षणात्मक परिस्थितियों में अपनाया गया है, जहां किसी एक कारक के स्तर में परिवर्तन करना प्राकृतिक रूप से लगभग असंभव होता है और विश्लेषण करने में फायदा लिया गया है। हालांकि यह पद्धति सर्वश्रेष्ठ पद्धति नहीं हो सकती है जबकि एक से अधिक कारकों के स्तरों में न्यूनतम मात्रा में परिवर्तनों की आवश्यकता हो और विशेषकर जबकि कुछ स्तर संयोजनों को मिलाया जा सकता है एवं विशेष क्रम का प्रयोग नहीं किया जा रहा है। क्योंकि यहां यादृच्छिकीकरण शेष बचे कारकों के स्तरों या स्तरों के संयोजनों पर किया जा रहा है अतः रन अनुक्रम प्राप्त करने की संभावना बहुत कम है जहां स्तर परिवर्तन की कुल मात्रा न्यूनतम हो। अतः ऐसी परिस्थितियों में जहां एक से अधिक कारकों के लिए



परिवर्तनों की मात्रा का न्यूनतम होना आवश्यक है वहां उपरोक्त पारम्परिक पद्धति के स्थान पर रन अनुक्रमों का निर्माण इस प्रकार करना चाहिए कि कारक के स्तर में परिवर्तन की मात्रा न्यूनतम हो। परीक्षात्मक रन अनुक्रमों, जो कारक स्तर में न्यूनतम मात्रा में परिवर्तन करते हैं एवं साथ ही अवांछनीय कारकों के संभावित प्रभाव को न्यूनतम करते हैं, की खोज कोई नई बात नहीं है। कॉक्स (1951) ने आवृत किस्म (replicated varietal) परीक्षणों के लिए पद्धतिबद्ध (systematic) अभिकल्पनाओं के अध्ययन की शुरुआत की। कारक स्तर में न्यूनतम मात्रा में परिवर्तन वाले रन अनुक्रम एक ऐसा उदाहरण हैं जो स्वचिंग सिद्धांत में ग्रे कोड के रूप में जाने जाते हैं। गिल्बर्ट (1958) द्वारा चक्रीय ग्रे कोड पर विचार-विमर्श दिया गया है। ड्रेपर एवं स्टोनमैन (1968) ने  $2^3$  बहु-उपादानी एवं विभिन्न आंशिक बहु-उपादानी अभिकल्पनाओं के लिए रन ऑर्डरों को तालिकाबद्ध किया है जिनमें 8 रनों की आवश्यकता होती है जिनमें मुख्य प्रभावों के आकलन के लिए आवश्यक समय प्रवृत्ति से मुक्त या लगभग मुक्त होते हैं एवं कारक स्तर परिवर्तनों की संख्या भी कम होती है। डिकिंसन (1974) ने ड्रेपर एवं स्टोनमैन (1968) के कार्य का विस्तार न्यूनतम लागत रन ऑर्डर तक प्रतिबंधित खोज के साथ  $2^4$  एवं  $2^5$  पूर्ण बहु-उपादानी योजनाओं के लिए किया। डी लियोन इत्यादि (2005) ने 8 या 16 रनों वाली अभिकल्पनाओं के लिए परीक्षात्मक ऑर्डरों को प्रस्तावित किया जो अज्ञात कारकों के प्रभाव के विरुद्ध कारक स्तरों में न्यूनतम संख्या में परिवर्तनों के साथ संरक्षण के उत्कृष्ट स्तर का संयोजन करते हैं। कोरिया इत्यादि (2012) ने 32 परीक्षणों वाली अभिकल्पनाओं के लिए उत्कृष्ट संभावित अनुक्रमों के साथ-साथ 64 एवं 128 परीक्षणों वाली अभिकल्पनाओं के लिए उत्कृष्ट लक्षण प्रदान करने वाले अनुक्रमों पर चर्चा की।

प्रस्तुत शोधपत्र में,  $2^k$  बहु-उपादानी परीक्षणों में न्यूनतम रूप से परिवर्तित रन अनुक्रमों को प्राप्त करने के लिए एक पद्धति पर विचार-विमर्श किया गया है। न्यूनतम रूप से परिवर्तित रन अनुक्रमों सहित विकसित  $2^k$  बहु-उपादानी परीक्षणों में कारक के अनुसार स्तर

परिवर्तनों की संख्या की गणना करने के लिए एक सामान्य सूत्र प्राप्त किया गया है। इसके अतिरिक्त, यह लेख न्यूनतम रूप से परिवर्तित रन अनुक्रमों सहित द्वि-स्तरीय आंशिक बहु-उपादानी परीक्षणों को प्राप्त करने के लिए एक पद्धति से भी संबंधित है।

### न्यूनतम रूप से परिवर्तित रन अनुक्रमों सहित $2^k$ बहु-उपादानी परीक्षणों में कारक स्तर परिवर्तनों की कुल संख्या

परीक्षण चलाने के लिए प्रयासों एवं लागत पर सार्थक प्रभाव डालने के लिए, कारक स्तरों पर निष्पादन करने के लिए आवश्यक परिवर्तनों की संख्या एक महत्वपूर्ण पहलू है। ऐसा तभी होता है जब स्तरों में परिवर्तन के लिए अत्यधिक मात्रा (उदाहरणार्थ मोल्ड परिवर्तन) या एक निश्चित समय (उदाहरणार्थ तापमान परिवर्तन) की प्रतीक्षा या किसी अन्य कारण (उदाहरणार्थ रिएक्टर सफाई) की आवश्यकता होती है। अतः परीक्षात्मक प्रक्रिया के लिए कारकों के स्तर परिवर्तनों की संख्या के लिए मापदंड एक अत्यन्त महत्वपूर्ण विचारणीय विषय है।

ऐसे परीक्षणों में, जिनमें कारक स्तरों में परिवर्तन करना मुश्किल होता है, रन अनुक्रमों का यादृच्छिकीकरण करने की बजाय कारक स्तरों में न्यूनतम संख्या में परिवर्तन वाले बहु-उपादानी परीक्षणों का प्रयोग एक संभावित विकल्प जाना जाता है। उदाहरणार्थ, मानक क्रम में, निम्न  $2^3$  बहु-उपादानी रन अनुक्रमों पर विचार करें।

रन अनुक्रम	कारक		
	A	B	C
1	-1	-1	-1
2	1	-1	-1
3	-1	1	-1
4	1	1	-1
5	-1	-1	1
6	1	-1	1
7	-1	1	1
8	1	1	1
कारक स्तर परिवर्तनों की संख्या	7	3	1

यहां कारक स्तर परिवर्तनों की कुल संख्या 11 है। इस मानक क्रम को, निम्न रन अनुक्रम जहां कारक स्तर परिवर्तनों की कुल संख्या 7 है, पुनःव्यवस्थित किया जा सकता है।

रन अनुक्रम	कारक		
	A	B	C
1	-1	-1	-1
2	1	-1	-1
6	1	-1	1
8	1	1	1
4	1	1	-1
3	-1	1	-1
7	-1	1	1
5	-1	-1	1
कारक स्तर परिवर्तनों की संख्या	2	2	3

अतः ऐसे परीक्षण जहां कारक स्तरों को परिवर्तित करना बहुत मुश्किल होता है वहां रन अनुक्रमों के मानक रूप के स्थान पर रन अनुक्रमों में कम कारक स्तर परिवर्तनों वाली बहु-उपादानी अभिकल्पना का प्रयोग किया जाना चाहिए।

कारक स्तर परिवर्तनों की कुल संख्या न्यूनतम होगी यदि सिर्फ कारक के स्तर को दो क्रमित परीक्षणात्मक ट्रायल्स पर बदल दिया जाता है या अभिकल्पना आव्यूह में एक पंक्ति से दूसरी में पारित करके सिर्फ एक चिन्ह में परिवर्तन होता है। अतः न्यूनतम परिवर्तित रन अनुक्रमों सहित  $2^k$  अभिकल्पना में कारक स्तर परिवर्तनों की कुल संख्या  $2^k - 1$  होती है।

### न्यूनतम परिवर्तित रन अनुक्रमों सहित $2^k$ बहु-उपादानी अभिकल्पनाएं विकसित करने के लिए पद्धति

जब कारक स्तर में परिवर्तन करना प्राकृतिक रूप से बहुत मुश्किल होता है तो यादृच्छिक रन अनुक्रमों के

प्रयोग के स्थान पर ऐसे रन अनुक्रमों का प्रयोग करना चाहिये जहां कारक स्तर परिवर्तन की कुल संख्या न्यूनतम हो। न्यूनतम परिवर्तन रन अनुक्रमों सहित  $2^k$  बहु-उपादानी परीक्षण प्राप्त करने के लिए निम्न पद्धति है :

यह पद्धति  $-1$  एवं  $1$  से चिन्हित दो स्तरों सहित एक एकल कारक से शुरू होती है जिनसे 2 पंक्तियां और एक एकल कॉलम बनते हैं। फिर प्रत्येक पंक्ति को लगातार दोहराने के परिणामस्वरूप 4 पंक्तियां एवं एकल कॉलम प्राप्त होते हैं। अन्य कारक के लिए  $-1$ ,  $1$  प्रविष्टियों के साथ शुरू करते हुए एक नया कॉलम जोड़ें और फिर उसका फोल्ड ऑवर (Fold Over) लें। इससे परिवर्तन की कुल संख्या सहित दो कारकों के समतुल्य 4 रन प्राप्त होंगे जो  $2^2$  बहु-उपादानी परीक्षणों के लिए न्यूनतम हैं। इसी प्रकार  $2^3$  बहु-उपादानी के लिए उपरोक्त  $2^2$  बहु-उपादानी पर विचार करें। उपरोक्त प्रक्रिया का अनुसरण करके अंतिम कॉलम अर्थात् फोल्ड ऑवर लेते हुए कॉलम की अंतिम प्रविष्टि तक तीसरा कारक जोड़ें। यह प्रक्रिया न्यूनतम परिवर्तित रन अनुक्रमों सहित  $2^{k-1}$  बहु-उपादानी से शुरू करके  $2^k$  बहु-उपादानी (जहां परिवर्तनों की कुल संख्या  $2^k - 1$  होगी) के लिए विस्तारित की जा सकती है। उपरोक्त पद्धति न्यूनतम परिवर्तित रन अनुक्रमों सहित  $2^k$  बहु-उपादानी परीक्षण प्रदान करेगी। इसके अलावा, विकसित  $2^k$  बहु-उपादानी परीक्षणों में कारक के अनुसार स्तर परिवर्तनों की संख्या प्राप्त करने के लिए सामान्य व्यंजक प्राप्त किया गया है। ज्ञात हुआ कि  $i$ -वें कारक ( $i=1, 2, \dots, k$ ) के लिए स्तर परिवर्तनों की संख्या  $2^{i-1}$  है।

**उदाहरण 1 :** एक  $2^2$  अभिकल्पना के लिए  $-1$  एवं  $1$  द्वारा चिन्हित 2 स्तरों सहित एक एकल कारक से शुरू करके 2 पंक्तियाँ एवं एक एकल कॉलम बनाएं। फिर प्रत्येक पंक्ति को लगातार दोहराते हुए और अन्य कारक के लिए  $-1$ ,  $1$  प्रविष्टियों से शुरू करते हुए एक नया कॉलम जोड़ें। फिर उसका फोल्ड ओवर लें। इस प्रकार न्यूनतम परिवर्तित रन अनुक्रमों सहित निम्न  $2^2$  बहु-उपादानी (जहां परिवर्तनों की कुल संख्या 3 होगी) प्राप्त कर सकते हैं:

न्यूनतम परिवर्तित रन अनुक्रमों सहित $2^2$ बहु-उपादानी	
-1	-1
-1	1
1	1
1	-1
1	2

यहां अंतिम पंक्ति कारक के अनुसार स्तर परिवर्तनों की संख्या को इंगित करती है। अब न्यूनतम परिवर्तित रन अनुक्रमों सहित उपरोक्त  $2^2$  बहु-उपादानी का प्रयोग करते हुए और उपरोक्त प्रक्रिया का पालन करते हुए न्यूनतम परिवर्तित रन अनुक्रमों सहित निम्न  $2^3$  बहु-उपादानी (जहां परिवर्तनों की कुल संख्या 3 होगी) प्राप्त किया जा सकता है।

न्यूनतम परिवर्तित रन अनुक्रमों सहित $2^3$ बहु-उपादानी		
-1	-1	-1
-1	-1	1
-1	1	1
-1	1	-1
1	1	-1
1	1	1
1	-1	1
1	-1	-1
1	2	4

यहां, अंतिम पंक्ति कारक के अनुसार स्तर परिवर्तनों की संख्या को इंगित करती है।

### न्यूनतम परिवर्तित रन अनुक्रमों सहित द्वि-स्तरीय आंशिक बहु-उपादानी

आंशिक बहु-उपादानी अभिकल्पनाओं में, परिवर्तनों की संख्या का मूल्यांकन करना इतना सरल नहीं है। यह ध्यान में रखना चाहिए कि अभिकल्पना आव्यूह में उपस्थित होने वाले कारकों को दो समूहों में विभाजित किया जा सकता है।

(अ) एक पूर्ण आव्यूह में शामिल कारक (उदाहरणार्थ,  $2^{5-2}$  में, एक पूर्ण  $2^3$  आव्यूह में कोई तीन कारक शामिल किये जा सकते हैं) एवं

(ब) पूर्व अन्योन्यक्रिया के आधार पर जेनरेट किए गए कारक

माना  $NC(X)$  एवं  $NC(Y)$  क्रमशः  $X$  एवं  $Y$  कारकों के परिवर्तनों की संख्या है जबकि अभिकल्पना आव्यूह को इस प्रकार क्रम दिया गया है कि परिवर्तनों की संख्या न्यूनतम हो। यदि नये कारक के चिन्हों के अनुक्रमों की पहचान  $XY$  की अन्योन्यक्रिया के साथ की जाती है तो इस कारक के परिवर्तनों की संख्या  $NC(XY)=NC(X)+NC(Y)$  होगी। यह इसलिए है क्योंकि यदि एक पंक्ति से दूसरी में  $X$  कॉलम का चिन्ह बदल जाता है तो  $XY$  कॉलम में भी इसी प्रकार होगा क्योंकि एक पंक्ति से दूसरी में जाते हुए पूर्ण आव्यूह केवल एक परिवर्तन के लिए ही प्रतिबंधित है। परिवर्तनों की न्यूनतम संख्या सहित किसी भी पूर्ण आव्यूह के लिए यह लक्षण वैद्य है।

इसी कारण, यदि अभिकल्पना आव्यूह में परिवर्तनों की न्यूनतम संख्या सहित तीन या अधिक कारकों की अन्योन्यक्रिया के द्वारा एक कारक जेनरेट होता है तो नये कारक के परिवर्तनों की संख्या उन कारकों के परिवर्तनों की संख्या के कुल योग के बराबर होगी जिन कारकों ने उन अन्योन्यक्रियाओं में हस्ताक्षेप किया है जिनके साथ वे अभिकल्पित है। (तालिका 1)

**तालिका 1 :**  $2^{7-4}$  अभिकल्पना के कारकों में परिवर्तनों की संख्या

A	B	C	D=AB	E=AC	F=BC	G=ABC
-1	-1	-1	1	1	1	-1
-1	-1	1	1	-1	-1	1
-1	1	1	-1	-1	1	-1
-1	1	-1	-1	1	-1	1
1	1	-1	1	-1	-1	-1
1	1	1	1	1	1	1
1	-1	1	-1	1	-1	-1
1	-1	-1	-1	-1	1	1
1	2	3	4	5	6	7

$2^{7-4}$  अभिकल्पना प्राप्त करने के लिए, पहले न्यूनतम परिवर्तित रन अनुक्रमों सहित  $2^3$  अभिकल्पना पर विचार करना होगा। तालिका 1 में, प्रारम्भिक तीन कॉलम उदाहरण 1 में प्राप्त न्यूनतम परिवर्तित रन अनुक्रमों सहित  $2^3$  अभिकल्पना का प्रतिनिधित्व करते हैं। फिर शेष 4 कॉलम जेनरेटर पद्धति द्वारा जेनरेट किये गये। तालिका 1 की अंतिम पंक्ति से यह देखा जा सकता है कि आंशिक बहु-उपादानी के नये कॉलमों में परिवर्तनों की संख्या, उन कारकों के परिवर्तनों की संख्या के कुल योग के बराबर होगी जिन कारकों ने उन अन्योन्यक्रियाओं में हस्ताक्षेप किया है जिनके साथ वे अभिकल्पित हैं।

### निष्कर्ष

बहु-उपादानी परीक्षण में, जहां कारकों के स्तरों का परिवर्तन करना कठिन होता है, वहां विकल्प के रूप में न्यूनतम परिवर्तित रन अनुक्रमों का प्रयोग किया जा सकता है क्योंकि रन अनुक्रमों का यादृच्छिकीकरण, कारकों के बीच स्तर परिवर्तनों की संख्या में वृद्धि कर देगा और इस कारण परीक्षण की लागत एवं समय में भी वृद्धि हो जाएगी। यहां, न्यूनतम परिवर्तित रन अनुक्रमों सहित  $2^k$  बहु-उपादानी प्राप्त करने के लिए एक पद्धति पर विचार विमर्श किया गया है। इस पद्धति पर आधारित, कारक अनुसार स्तर परिवर्तनों के लिए सामान्य व्यंजन प्राप्त किया गया है। इसके अलावा, द्वि-स्तरीय आंशिक बहु-उपादानी भी प्राप्त किये गये हैं।

### संदर्भ

1. कोरिया, ए ए, गरिमा, पी एवं टॉर्ट-मार्टोरेल, एक्स (2012): "एक्सपेरिमेन्टेशन ऑर्डर इन फैक्टोरियल डिजाइन्स: न्यू फाइन्डिंग्स", जर्नल ऑफ एप्लाइड स्टैटिस्टिक्स, 39(7), 1577-1591।
2. कॉक्स, डी आर (1951): "सम सिस्टैमैटिक एक्सपेरिमेंटल डिजाइन्स", बायोमैट्रिका, 38, 312-323।
3. डी लियोन, जी, गरिमा, पी एवं टॉर्ट-मार्टोरेल, एक्स (2005): "एक्सपेरिमेन्टेशन ऑर्डर इन फैक्टोरियल डिजाइन्स विद 8 या 16 रन्स", जर्नल ऑफ एप्लाइड स्टैटिस्टिक्स, 32(3), 297-313।
4. डिकिंसन, ए डब्ल्यू (1974): "सम ऑर्डर रिक्यूवायरिंग ए मिनिमम नम्बर ऑफ फैक्टर लैबल चेंजिस फॉर  $2^4$  एण्ड  $2^5$  मेन इफैक्ट प्लान्स", टैक्नोमैट्रिक्स, 16, 31-37।
5. ड्रेपर, एन एवं स्टोनमैन, डी (1968): "फैक्टर चेन्जिस एण्ड लीनियर ट्रैन्ड्स इन ऐट-रन टू-लैबल फैक्टोरियल डिजाइन्स", टैक्नोमैट्रिक्स, 10, 301-311।
6. गिल्बर्ड, ई एन (1958): "ग्रे कोड्स ऐण्ड पाथ्स ऑन दी एन-क्यूब", बैल सिस्टम टैक्निकल जर्नल, 37, 815-826।

□

## कृत्रिम तंत्रिका नेटवर्क एवं पारंपरिक वर्गीकरण पद्धतियों का तुलनात्मक अध्ययन

सुकान्त दाश, सुशील कुमार सरकार, देवेन्द्र कुमार एवं ज्ञान सिंह

वर्गीकरण के लिए पांच हाइरार्किकल विधियों (सिंगल लिंकेज, कम्पलीट लिंकेज, एवरेज बिटविन लिंकेज, एवरेज विदिन लिंकेज एवं वार्ड पद्धति) एवं एक नॉन-हाइरार्किकल विधि (K-माध्य पद्धति) साहित्य में उपलब्ध है। कृत्रिम तंत्रिका पर आधारित SOFM पद्धति का प्रयोग वर्गीकरण के लिए किया जा सकता है। प्रस्तुत पत्र में इन विधियों का पारंपरिक वर्गीकरण पद्धतियों के साथ तुलनात्मक अध्ययन किया गया है। वर्गीकरण के लिए पारंपरिक दूरता मापक पद्धतियों में से सिर्फ यूक्लिडिअन दूरी एवं सिटी ब्लॉक दूरी का प्रयोग किया गया है। इन विधियों को धान के कृत्रिम आँकड़ों के लिए प्रयुक्त किया गया है। इन सभी तरीकों में कृत्रिम तंत्रिका नेटवर्क वर्गीकरण पद्धति सर्वश्रेष्ठ पायी गयी है जिसकी, अवर्गीकरण औसत प्रतिशत प्रायिकता सबसे कम पायी गयी।

बहुभिन्न रूपी आँकड़ों की अधिक मात्रा का संक्षिप्तीकरण का कृषि विज्ञान की विभिन्न शाखाओं में उत्तरोत्तर अभ्यास किया जा रहा है। बहुभिन्न रूपी सांख्यिकीय तकनीकों के वर्गीकरण उद्देश्य हेतु बहुत सी विधियां क्रमशः गुच्छ विश्लेषण, प्रमुख घटक विश्लेषण, कारक विश्लेषण का व्यापक रूप से प्रयोग किया जा रहा है। पादप प्रजनकों की एक बुनियादी समस्या जीनोटाइप की वृहद संख्या का कम प्रबंधनीय सजातीय समूहों/गुच्छों में वर्गीकरण करना है। गुच्छ विश्लेषण वस्तु समूह का सजातीय समूहों की अज्ञात संख्या में वर्गीकरण हेतु एक तकनीक है। सजातीय समूह बनाने हेतु गुच्छ विधियां एवं विषमता उपाय

बड़ी संख्या में साहित्य में उपलब्ध हैं। प्रजनकों की एक मुख्य समस्या, विभिन्न गुच्छ विधियों एवं विषमता उपायों में से उपयुक्त विधि का चुनाव करना है क्योंकि, साहित्य में उपलब्ध विभिन्न गुच्छ विधियों एवं विषमता उपायों के प्रदर्शन के बारे में कोई सूचना नहीं है। इसके अतिरिक्त, हाल ही में वर्गीकरण प्रयोजनों हेतु कृत्रिम तंत्रिका नेटवर्क (ANN) का अनुप्रयोग बहुत अधिक पाया गया है। कोहोनेन नेटवर्क एक स्व-संगठित प्रणाली है जो स्वयं को इस प्रकार अपने अनुकूल बनाता है कि समान इनपुट, कृत्रिम तंत्रिका नेटवर्क (ANN) में सांस्थतीय (टोपोलोजिकल) रूप से समीप के न्यूरॉन्स से संबद्ध हो जाते हैं। अतः प्रतिष्ठित गुच्छन प्रणालियों की अपेक्षा कृत्रिम तंत्रिका नेटवर्क (ANN) का प्रदर्शन बहुत रुचिकर है। इसे ध्यान में रखते हुए प्रस्तुत शोध पत्र में कृत्रिम तंत्रिका नेटवर्क (ANN) समेत छः विभिन्न गुच्छन प्रणालियों की तुलना सजातीय समूहों में गुच्छ जीनोटाइप हेतु सामान्यतया प्रयुक्त दो विजातीय उपायों से की गयी है।

सामान्यतया प्रयुक्त मान्य गुच्छन विधि को श्रेणीबद्ध एवं गैर श्रेणीबद्ध दो सामान्य श्रेणियों में रखा जाता है। वर्तमान अध्ययन में निम्नलिखित पांच गुच्छन विधियां प्रयुक्त गयी हैं।

- (1) सिंगल लिंकेज (न्यूनतम दूरी या निकटतम पड़ोसी)
- (2) कम्पलीट लिंकेज (अधिकतम दूरी या दूरस्थ पड़ोसी)
- (3) एवरेज लिंकेज (औसत दूरी)
- (4) वार्ड पद्धति
- (5) नॉन-हाइरार्किकल विधि (K-माध्य पद्धति)

## कृत्रिम आँकड़ों की उत्पत्ति

धान के 75 जीनोटाइप जिनमें 25 लघु, 25 मध्यम एवं 25 वृहद पर बहुभिन्नरूपी आँकड़ों पर आधारित तीन सजातीय गुच्छ प्राप्त किये गये थे। ये गुच्छ धान के नौ गुणों जैसे :- प्रति पौधा बालियों की संख्या, पौधे की ऊचाई (सेमी.), पुष्प गुच्छों की लम्बाई (सेमी.), पुष्प गुच्छों का भार (ग्राम), 1000 दानों का भार (ग्राम), प्रति पौधा जैव द्रव्यमान (ग्राम), कटाई सूचकांक (प्रतिशत), प्रति पौधा भूसे की पैदावार (ग्राम), प्रति पौधा अन्न की पैदावार (ग्राम) पर आधारित है। ये आँकड़े भारतीय कृषि अनुसंधान संस्थान, नई दिल्ली के अनुवांशिकी प्रभाग से प्राप्त किये गये हैं। छद्म प्रयोजनों हेतु माध्य वैक्टर एवं प्रसरण-सह प्रसरण आव्यूह का जनसंख्या प्राचलों की तरह प्रयोग किया गया है।

## परिणाम एवं परिचर्चा

आम सहमति के आधार पर विभिन्न गुच्छन प्रक्रियाओं के परिणाम पर आधारित धान के 75 जीनोटाइप पर

बहुभिन्नरूपी आँकड़ों पर आधारित तीन सजातीय गुच्छ प्राप्त किये गये थे। छद्म प्रयोजनों हेतु माध्य वैक्टर एवं प्रसरण-सह प्रसरण आव्यूह का जनसंख्या प्राचलों की तरह प्रयोग किया गया है। सामग्री एवं पद्धतियों में दिये गये गुच्छन को पांच पारंपरिक पद्धतियों एवं कृत्रिम तंत्रिका नेटवर्क (ANN) का प्रयोग करते हुए विभिन्न छद्म बहुभिन्नरूपी आधारित सामान्य समष्टि की तुलना, दिये गये माध्य वैक्टर एवं प्रसरण-सह प्रसरण आव्यूह से की गयी है। इन विधियों का प्रदर्शन बीस विभिन्न छद्म प्रतिदर्श से प्राप्त त्रुटिपूर्ण वर्गीकरण की मान्य प्रायिकता के आधार पर तुलना की जाती है। इसके अतिरिक्त, इन विधियों की निरंतरता इन प्रायिकताओं की मानक त्रुटियों के आधार पर आंकी गयी। तीन गुच्छों लघु (30), मध्यम (60) एवं वृहद (150) प्रतिदर्श आकार के लिए बहुभिन्नरूपी आँकड़ों के कृत्रिम प्रतिदर्श लिये गये। गुच्छन की विभिन्न विधियों के लिए गैर-वर्गीकरण की माध्य प्रायिकता अपनी मानक त्रुटियों सहित तालिका-1 में दी गयी है।

**तालिका-1:** वर्गीय यूक्लिडियन दूरी एवं सिटी-ब्लॉक असमानता उपायों पर आधारित गुच्छन की विभिन्न विधियों हेतु त्रुटिपूर्ण वर्गीकरण की औसत प्रतिशत प्रायिकता (p) एवं इसकी मानक त्रुटि (S.E.) सिंगल लिंकेज माध्य वार्डस के माध्य कृत्रिम तंत्रिका नेटवर्क (ANN)

वर्गीकरण विधि								
प्रतिदर्श आकार	दूरी		सिंगल लिंकेज	कम्पलीट लिंकेज	एवरेज लिंकेज	वार्ड विधि	K-माध्य*	कृत्रिम तंत्रिका नेटवर्क (ANN) **
लघु	यूक्लिडियन दूरी	औसत प्रतिशत प्रायिकता (p)	16.501	11.500	11.001	9.333	9.667	5.666
		मानक त्रुटि (S.E.)	2.196	1.718	2.196	1.338	1.277	0.598
	सिटी ब्लॉक दूरी	औसत प्रतिशत प्रायिकता (p)	16.000	12.333	11.167	10.334		
		मानक त्रुटि (S.E.)	1.737	1.512	1.330	1.182		
मध्यम	यूक्लिडियन दूरी	औसत प्रतिशत प्रायिकता (p)	29.167	11.085	7.918	9.168	6.984	5.417
		मानक त्रुटि (S.E.)	3.613	1.264	0.879	0.909	0.663	0.399

वर्गीकरण विधि								
प्रतिदर्श आकार	दूरी		सिंगल लिंकेज	कम्पलीट लिंकेज	एवरेज लिंकेज	वार्ड विधि	K-माध्य*	कृत्रिम तंत्रिका नेटवर्क (ANN) **
	सिटी ब्लॉक दूरी	औसत प्रतिशत प्रायिकता (p)	35.417	11.476	8.499	9.698		
		मानक त्रुटि (S.E.)	3.030	1.28	0.813	0.938		
वृहद	यूक्लिडिअन दूरी	औसत प्रतिशत प्रायिकता (p)	47.416	8.500	7.801	9.201	6.899	4.534
		मानक त्रुटि (S.E.)	2.354	0.893	0.686	0.677	0.567	0.330
	सिटी ब्लॉक दूरी	औसत प्रतिशत प्रायिकता (p)	47.566	8.699	8.301	9.600		
		मानक त्रुटि (S.E.)	2.229	0.875	0.637	0.683		

\*K—माध्य विधि निकटतम केन्द्रक (माध्य) पर आधारित है।

\*\*कृत्रिम तंत्रिका नेटवर्क (ANN) विधि पर आधारित है।

कृत्रिम तंत्रिका नेटवर्क (ANN) विधि के लिए लघु, मध्यम एवं वृहद प्रतिदर्श आकारों के लिये त्रुटिपूर्ण वर्गीकरण की प्रतिशतता क्रमशः 5.666, 5.417 एवं 4.534 है जोकि, अन्य सभी विधियों की तुलना में कम है। इसी प्रकार से कृत्रिम तंत्रिका नेटवर्क (ANN) विधि के लिए लघु, मध्यम एवं वृहद प्रतिदर्श आकारों के लिये मानक त्रुटि 0.598, 0.399 एवं 0.330 है जोकि, अन्य सभी विधियों की तुलना में कम है। इससे यह निष्कर्ष निकलता है कि कृत्रिम तंत्रिका नेटवर्क (ANN) विधि सर्वश्रेष्ठ विधि है। लघु प्रतिदर्श आकार के लिये दूसरी उत्तम विधि वार्ड पद्धति है जिसमें त्रुटिपूर्ण वर्गीकरण की प्रतिशतता 9.333 है। परन्तु

मध्यम एवं वृहद प्रतिदर्श आकारों के लिये K-माध्य पद्धति दूसरी उत्तम पद्धति है जिसमें त्रुटिपूर्ण वर्गीकरण की प्रतिशतता 6.984 एवं 6.899 है। अतः उपरोक्त से यह निष्कर्ष निकाला गया कि, किसी भी प्रतिदर्श आकार या दूरी मापता विधि के लिए गुच्छन की छः विधियों में से कृत्रिम तंत्रिका नेटवर्क (ANN) विधि का प्रदर्शन सर्वोत्तम है।

### संदर्भ

1. जॉन्सन आर.ए. एवं विचर्न डी.डब्ल्यू. (2006): “अप्लाइड मल्टीवैरिएट स्टैटिस्टिकल एनालिसिस”, पाँचवां संस्करण, लन्दन, आईएनसी. पिअरसन प्रेन्टिस हॉल।

□

## भारत के समुद्रीय और पशुधन के निर्यात उत्पादन में वाष्पशीलता और मुद्रीता

बिशाल गुरुंग, कमलेश नारायण सिंह, रंजीत कुमार पॉल एवं लॉरेंस लेपचा

भूमण्डलीय मांगों के बढ़ते हुए, समुद्रीय उत्पादन और पशुधन के लिए एक आश्चर्य स्थिति में भूमण्डलीय मूल्यों की अधिकता हुई है तथा भारत से ये चीजे निर्यात में एक आश्चर्यजनक वृद्धि हुई है। आगे, भूमण्डलीय व्यापार संधि और अनेक प्रकार के घरेलू सुधार एवं विभिन्न प्रकार के पशुधन क्षेत्र के पहलू राष्ट्रीय और अंतर्राष्ट्रीय मार्केट में पर्याप्त सुधार हुआ है। मुख्य रूप से, पिछले समय के दौरान डब्ल्यू.टी.ओ. दौराने (कुमार 2010) ने वर्णित किया था। इसी के कारण ही, समुद्रीय उत्पादन और पशुधन का निर्यात अधिक फलन लक्षण का अचानक बढ़ना और गिरना दर्शाता है। यह क्या वाष्पशीलता कहलाती है। यह एक तरह का नमूना है जो प्रदर्शित किया जाता है। जबकि आजकल यह अंतर्राष्ट्रीय बाजार में एक अपवाद है। जै.की. (2005) के अनुसार, यह एक व्यापार संरचना पर देय हो सके।

जलवायु स्थिति और जिससे वह मूल्य परिवर्तन उत्पाद कर सकती है। बहुत सी कृषि चीजों का निर्यात एक उतार-चढ़ाव को दर्शाता है। क्योंकि अल्प उत्पादन का निर्णय के बीच मार्केट में उचित रूप से वितरण होता है। कुमार इत्यादि (2012) के अनुसार, भारत से भैंसों के मीट का कुशल निर्यात भी हुआ है तथा भारत से भैंसों की मीट की प्रतियोगिता भी पर्याप्त हुई है। यह कार्य यह प्रदर्शित करता है कि गत 02 दशक में भारत से भैंसों की मीट का निर्यात में एक स्मरणीय वृद्धि हुई है। सांख्यिकीय रूप से, वाष्पशीलता एक प्रायः माप न की गई है। जैसे नमूना

मानक प्रसरण, कभी-2 प्रसरण के एक प्रयोगिक भी है कि एक वाष्पशीलता मापन की गई है। कम या अधिक आर्क-एल.एम. (ARCH-LM) प्रयोग सांख्यिकी काल-श्रृंखला डेटा सेट्स वाष्पशीलताके लिए विकसित भी प्रयोग करना है। सामान्यता: काल-श्रृंखला आंकड़ा मॉडल वाष्पशीलता में स्वः प्रतिगामी विषम स्थिति गार्च परिवार का मापक अरैखिक काल-श्रृंखला मॉडल्स प्रस्तावित किया गया था। गार्च आकलन स्थिति षर्त रूप से अस्थिर वाष्पशीलता उत्पन्न प्रक्रिया एक उचित योजनायक नीति के निर्णय के लिए सहायक होती है। इस प्रकार अस्थिर आंकड़ों के विश्लेषण, जैसे कि समुद्रीय और पशुधन का निर्यात उत्पादन में समयबद्ध रूप से एक सामर्थ नीति को महत्वपूर्ण रूप में वांछनीय होती है। इस पेपर में, हम अस्थिरता का अध्ययन इसी के साथ-साथ इस आंकड़ा सेट में समुद्रीय उत्पादकों और पशुधन संबंध अंकड़ा में अस्थिरता होती है। तदानुसार, काल-श्रृंखला आंकड़ा गार्च मॉडल फिट करके भी प्रयोग के अन्तर्गत विचारणीय है। दूसरे महत्वपूर्ण काल-श्रृंखला आंकड़ा की वित आकृति मुद्रीता की उपस्थिति होती है। जहाँ प्रत्येक श्रेणीयाँ स्वतंत्र रूप से अस्थिर है। किन्तु एक अरैखिक संयोग इना अस्थिरता में होता है। किन्तु रैखिक संयोग एक स्थिर होता है। इस अध्ययन में, हम भी मुद्रा स्थिति के संयोग स्वरूप में द्विकाल-श्रृंखला आंकड़ा और वहीं पर सांख्यिकीय मापन के लिए जाते हैं।

### सामग्री और विधि

यह अध्ययन काल-श्रृंखला पर माहवार आधारित



समुद्रीय उत्पादक और पशुधन के निर्यात पर भारत में [www.indiastat.com](http://www.indiastat.com) पर उपलब्ध है। जो विचाराधीन फलनों की वाष्पशीलता को बढ़ाता है जैसे कि अस्थिर उत्पादकता को नीतिगत बनाना तथा योजनायक भविष्य की वाष्पशीलता के कारण मूल रूप से नीतियों को विकसित करना है।

परिमय विधियाँ जो एक स्टेप पर मूल/वर्तमान मूल्य के अनुपात में मानक विचलन का आकलन करती हैं तथा इसकी प्रति कदम स्थिति को अध्ययन में प्रयोग किया गया है तथा निर्यात अस्थिरता विधि के द्वारा आंकलित की गई है।

निर्यात अस्थिरता सूचकांक (Ie) = मानक विचलन

$$\text{का } \text{Log} \left( \frac{E_t}{E_{t-1}} \right) * 100$$

जहाँ, E निर्यात है और t महीने औ t-1 उसके लॉग (Log) हैं। बराबर ही यह अध्ययन कुमार और राज ने (2011) वाष्पशीलता और उसका झुकाव जानने के लिए भारत में पशुधन का उत्पाद घरेलू एवं अंतर्राष्ट्रीय उत्पादकों में किया गया था। मुख्य रूप से, आरंभिक स्थिति में काल-श्रृंखला का विश्लेषण एक मॉडल को जानने में किया गया था। जो कि बोकस एट आल 2008 ने माध्य व्यवहारिक आंकड़ें प्रभावित व्याख्या की थी। वर्तमान में, वाष्पशीलता से संबंधित आंकड़े उच्च रूप में प्रकट हुए हैं। क्योंकि वाष्पशीलता के नमूनों में वास्तविक रूप से आंकड़ा सेट होता है। वाष्पशीलता पूर्ण समय के पश्चात बदलती रहती है। इस प्रकार मॉडलिंग और पूर्वानुमान की वाष्पशीलता वित्त मार्केट के लिए विचित्र रहती है। मॉडलिंग वाष्पशीलता के लिए आंकड़ा सेट के साहित्य में पहुंचे उपलब्ध होती है। सामान्यता: स्व:प्रतिगामी की विषम स्थितियों में स्थिर पहुंच की विधियाँ अरैखिक काल-श्रृंखला मापक के परिवार विषमता होती है।

$$\varepsilon_t = \xi_t h_t^{1/2}$$

और

$$h_t = a_0 + \sum_{i=1}^q a_i \varepsilon_{t-i}^2 + \sum_{j=1}^p b_j h_{t-j}$$

जहाँ  $\xi_t \sim IID(0,1)$ ,  $a_0 > 0$ ,  $a_i \geq 0$ ,  $i = 1, 2, \dots, q$ ,  $b_j \geq 0$ ,  $j = 1, 2, \dots, p$

यहाँ गार्च के आकलन के मापदंड के लिए साहित्य में उपलब्ध विधियों हैं। एक विधि गोसिन मेक्सीमम आकलन के उपलब्धता में हैं। जोकि ई-व्यूज सॉफ्टवेयर की है जो हमार प्रयोगिक विश्लेषण में प्रयोग होती है। प्रयोगिक रूप से, दो या दो से अधिक अस्थिर काल श्रृंखला आंकड़ा में वित्त रूप से दो या दो से अधिक असिरि काल-श्रृंखला आंकड़ा में मुद्रिता की उपस्थित के बीच में होते हैं। इन जैसी स्थितियों में, परिणामों को सूचित किया जाता है। कि बहुत सी सीधी महत्वपूर्ण सापेक्षता अस्थिर चर-आर्थिक प्रस्तत रूप से आर्थिकी मॉडल्स अप्रमाणिक होगा। इसलिए मापदंड लाभप्रद एवं अर्थवान सांख्यिकी विश्लेषण के प्रमाणिक होता है।

इस पेपर में, हम परिकल्पना का प्रयोग करते हैं कि यहाँ एक सांख्यिकीय महत्व संबंध समुद्रीय और पशुधन के बीच दो श्रेणियों के मुद्रिता संयोग के एक अस्थित्व के लिए किया गया है।

यदि दो काल-श्रृंखला  $y_t$  और  $x_t$  आपस में मुद्रिता हैं तो एक रैखिक संयोग अस्थिर होता है। दूसरे शब्दों में

$$x_t - \alpha y_t = \varepsilon_t$$

जहाँ  $\varepsilon_t$  एक स्थिर है।

यह एंगल-ग्रेंगर टूविधि के द्वारा किया गया है। जहाँ साधारण अल्प वर्गों को प्रयोग करके आंकलित किया गया है। यह स्थिरता के लिए डिकी फूलर टेस्ट के द्वारा परीक्षण किया गया है।

### परिणाम एवं चर्चा

समुद्रीय उत्पाद और पशुधन की निर्यात अस्थिरता: भारत का माहवार समुद्रीय उत्पाद और पशुधन का निर्यात आंकड़ा जून 2006 से जून 2012 तक का विश्लेषण [www.indiastat.com](http://www.indiastat.com) पर है। जोकि प्रत्येक उत्पाद में अस्थिरता के लिए किया जाता है। दो पशुधन निर्यात आंकड़ा और समुद्रीय उत्पाद को बड़े अन्तराल में व्याखित किये गये हैं। फलस्वरूप, उच्च

मूल्य अस्थिरता का सूचकांक 22.31 और 2014 सरचित किया जा सके।

तालिका 1 में भारत की समुद्रीय उत्पाद एवं निर्यात पशुधन का प्रतिमाह अस्थिरता की तुलनात्मक आकृति प्रदर्शित की गई है। तालिका न. 1 में यह दर्शाता है कि समुद्रीय उत्पाद की तुलना में निर्यात पशुधन की चर अंतः माहवार निर्यात पशुधन उत्पाद में अधिक है। प्रतिमाह अस्थिरता का झुकाव भी वर्ष 2007 में इसकी के साथ पशुधन उत्पाद एवं समुद्रीय उत्पाद के लिए अधिक चर में दर्शाता है।

तालिका 1: अंत माह निर्यात आंकड़ा का प्रसरण

अंत माह वर्ष के लिए प्रसरण	पशुधन	समुद्रीय
2006	15.52	16.89
2007	31.82	30.74
2008	21.08	19.44
2009	15.66	16.07
2010	24.04	20.97
2011	19.65	18.45
2012	15.22	13.44

तालिका 2 : निर्यात आंकड़ा की वर्णात्मक सांख्यिकी

पशुधन		समुद्रीय	
माध्य	606.26	माध्य	832.60
माध्यिका	457.35	माध्यिका	724.11
अधिकतम	1643.60	अधिकतम	1997.75
न्यूनतम	176.83	न्यूनतम	365.11
मानक विचलन	358.03	मानक विचलन	370.95
विषमता	1.12	विषमता	1.17
जारके-बेरा	15.97 (0.0003)	जारके-बेरा	19.46 (0.00006)

उपरोक्त तालिका 2 को देखते हुए, हम देखते हैं कि अधिकतम मूल्य का निर्यात 10 गुना अधिक

पशुधन निर्यात के मामलों में अस्थिर रूप से मिलता है और 05 गुना अधिकतम रूप से समुद्रीय उत्पाद में निर्यात में अस्थिर रूप से होता है। इसलिए अरिमा मामलों में काल-श्रृंखलामॉडलिंग को प्रयोग नहीं कर सकते हैं कम या अधिक आंकड़ा का अनुशीलन बहुत बड़े काल-श्रृंखला पर उच्च वाष्पशीलता होती है। इसलिए दो श्रेणियों में वाष्पशीलता की पकड़ हम उपलब्ध सॉफ्टवेयर पैकेज, वर्जन-4 से अरैखिक की पकड़ हम उपलब्ध में गार्च (GARCH) में रखकर प्रयोग कर सकते हैं।

पशुधन आंकड़ा के लिए फिटिड मॉडल जून 2006 से दिसम्बर 2011 इन समीकरण में दिया गया है।

$$y_t = 1137.44 + 0.60y_{t-1} + 0.34y_{t-2} + \varepsilon_t$$

(661.50) (0.19) (0.18)

जहाँ  $\varepsilon_t = h_t^{1/2} \xi_t$ , और  $h_t$  प्रसरण समीकरण के संतुष्ट करता है।

$$h_t = 520.28 + 0.15\varepsilon_{t-1}^2 - 1.00h_{t-1}$$

(157.27) (0.07) (0.01)

समुद्रीय निर्यात आंकड़ा के लिए फिट मॉडल जून 2006 से दिसम्बर, 2011 दिया गया है।

$$y_t = 941.61 + 1.05y_{t-1} - 0.272y_{t-2} + \varepsilon_t$$

(128.52) (0.07) (0.15)

जहाँ  $\varepsilon_t = h_t^{1/2} \xi_t$ , और  $h_t$  प्रसरण समीकरण के संतुष्ट करता है।

$$h_t = 890.76 - 0.15\varepsilon_{t-1}^2 - 0.92h_{t-1}$$

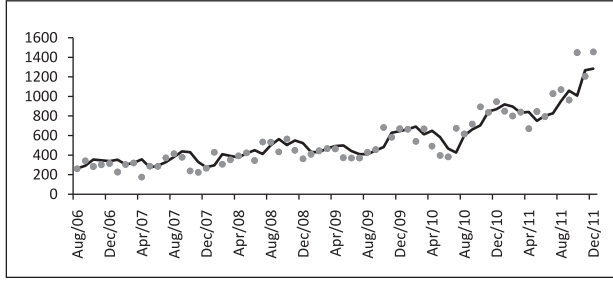
(277.73) (0.07) (0.05)

के मूल्यों में आकलन के मानक त्रुटियों के सापेक्षता में है।

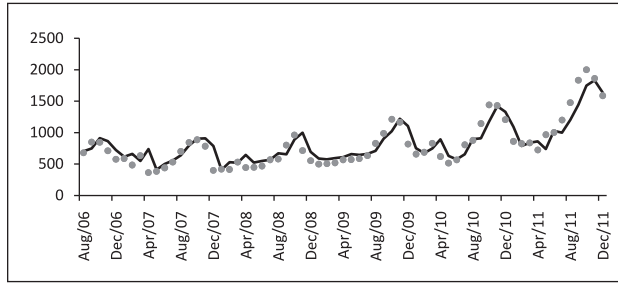
फिटिड मॉडल का mfp अध्ययन करना, स्वसंबंध फलन उपर्युक्त मानक और वर्गीय तथ्य यौगिक मिश्रित है। यह पता लगाया गया है कि दोनों परिस्थितियों में, स्वसंबंध फलन 5% पर स्तर की महत्ता अमहत्वपूर्ण

थी। इसलिए इसको निश्चित करने के लिए माध्य और प्रसरण समीकरणों को विशिष्ट किया है।

तालिका 1 एवं 2 में आंकड़ा बिन्दुओं के साथ मॉडल्स फिट को ग्राफ की व्याख्या की गई है। जो कि यह दर्शाता है कि फिट मॉडल में वाष्पशीलता की उपस्थिति को आंकड़ा को एक अच्छे तरीके से विस्तार किया है।



तालिका 1 में पशुधन के निर्यात आंकड़ा बिन्दुओं के साथ फिट गर्च (GARCH) मॉडल।



तालिका 2 समुद्रीय निर्यात आंकड़ा बिन्दुओं के साथ फिट गर्च (GARCH) मॉडल।

अचानक उतार-चढ़ाव के द्वारा निर्यात में विश्व-अर्थव्यवस्था की वृद्धि में चक्रीय मॉडल का स्वरूप जाना जा सकता है और इनकी स्थितियों के उत्पादन में भी आक्रामिक प्रभाव आता है। समुद्रीय उत्पाद और पशुधन दोनों से आय की विश्वसनीयता होती है। अप्रैल 2011 के पश्चात समुद्रीय उत्पाद और पशुधन के निर्यात में उभरते हुए झुकावों के लिए अचानक विदेशों से आयात बढ़ी है। जैसेकि चीन, और दक्षिण एशिया के कारण उनकी अर्थव्यवस्था में उच्च वृद्धि के कारण मांग भी बढ़ी है। समुद्रीय उत्पाद और पशुधन के लिए

उत्तरी अफ्रीका और मध्य एशिया में आयात की मांगों में मूल्य संबर्धन के लिए मुद्रास्फीति में तेल के दामों में अचानक बढ़ोतरी हुई है।

द्वि आंकड़ा जनित प्रक्रिया में मूल रूप की स्थिति का परीक्षण के लिए 5 प्रतिशत महत्वता का स्तर (डिके-फूलर) ने पता लगाया था। हम निष्कर्ष निकालते हैं कि पशुधन निर्यात के साथ-2 समुद्रीय निर्यात में अस्थिरता होती है जो कि मूल इकाई की उपस्थिति को शून्य परिकल्पना है। हम असफल होते हैं अब मुख्य समस्या है कि सांख्यिकीय निष्कर्ष एक स्थिर प्रक्रिया के साथ मुख्य समस्या रूप से वैध नहीं है। यदि काल-श्रृंखला की अस्थिर प्रक्रिया है तो काल-श्रृंखला की स्थिरता में विभिन्नता हो सके। अधिक सूचना को मिला करके श्रेणियों का स्तर विशेष सामान्यतः सैद्धांतिक आर्थिकी को चर स्तर के तुलना में विभिन्न का होता है।

इसके अंत में, मुद्रिता के बीच श्रेणियों की उपस्थिति में परीक्षण के लिए द्वि-पदरेजल-ग्रेंजर के लिए जाते हैं। यह पैकेज अर्थात 2007 ऐविकसित किया गया था और हम निष्कर्ष निकालते हैं कि द्वि-श्रेणियों मुद्रित की गई हैं। एक सामान्य रूप में स्टोस्टिक डिफ्ट में प्रतिगामी हैं। कम या अधिक प्रतिगामी गुणांक 0.587 था जबकि समुद्रीय उत्पाद एक पक्षीय चर लिया गया था और प्रतिगामी गुणांक जैसे 1.05 या अधिक रूप से पशुधन आंकड़ा अल्प  $p$ मूल्य दोनों मामलों में था। इसलिए यह बताना है कि यद्यपि द्वि-काल-श्रृंखला की अस्थिरता होती है। यहाँ रैखिक संयोग एक स्थिर है।

### पूर्वानुमान की कार्यकुशलता

जनवरी 2012 से जून 2012 तक समुद्रीय उत्पाद और पशुधन का निर्यात अखिल भारतीय महावार आंकड़ा के पत्राचार माध्यम के 06 बिन्दुओं के लिए पूर्वानुमान की कार्यकुशलता के आंकड़ों के लिए अध्ययन किया है। तालिका 3 में एक स्टेप आगे पशुधन निर्यात का पूर्वानुमान आंकड़ों और तालिका 4 में एक स्टेप आगे समुद्रीय उत्पाद के पूर्वानुमान आंकड़ों दर्शाए गए हैं।

**तालिका 3:** एक स्टेप आगे पशुधन निर्यात का पूर्वानुमान आंकड़ा (करोड़ में)

महीना	Actual	पूर्वानुमान
जनवरी-12	1643.6	1554.43
फरवरी-12	1284.7	1398.47
मार्च-12	1295.08	1283.57
अप्रैल-12	1173.62	1212.82
मई-12	1394.54	1306.82
जून-12	1331.18	1342.75

**तालिका 4:** एक स्टेप आगे समुद्रीय उत्पाद का पूर्वानुमान आंकड़ा (करोड़ में)

महीना	Actual	पूर्वानुमान
जनवरी-12	1299.42	1151.05
फरवरी-12	1089.06	1003.37
मार्च-12	1115.84	1087.94
अप्रैल-12	1082.86	1045.66
मई-12	1242.12	1224.11
जून-12	1326.38	1271.45

समुद्रीय उत्पादन और पशुधन निर्यात आंकड़ा उच्च स्थिरता को दर्शाता हैं। इसके उच्च स्थिर सूचकांक के द्वारा हम यह जान सकते हैं। यह भी उच्च एवं निम्न वाष्पशीलता को दर्शाता है। गर्च मॉडल के द्वारा ही पूर्वानुमान की वाष्पशीलता संतोशजनक थी। कम या अधिक रूप से यह देखा गया है कि दोनों निर्यात में काल-श्रृंखला आंकड़ा यद्यपि अस्थिर एवं मुद्रित थे। तथा अधिक समय में दो श्रृंखलों की सापेक्षता मॉडलिंग में मुद्रिता की काफी लाभप्रद रही होगी।

यह अध्ययन भारत के समुद्रीय और पशुधन के निर्यात उत्पादन में वाष्पशीलता, मुद्रिता और अस्थिरता का परीक्षण करने के लिए किया गया है। यह पता लगाया गया कि समुद्रीय निर्यात एक अंश अधिक स्थिर था और अल्पवाष्पशील पशुधन निर्यात उत्पादन की तुलना की गई है। मुद्रिता के बीच परीक्षण किये गये

जो निर्णायक थे कि दो निर्यात आंकड़ों-श्रेणीयाँ मुद्रित थी। इसका एक लाभ है कि अधिक समवय सापेक्षता एक सीधे प्रतिगामी चर स्तरों को सम्मिलित करती है। अन्त में, पूर्वानुमान भी प्राप्त किये गये थे जो बिल्कुल विश्वसनीय हैं।

### संदर्भ

1. बाक्स, जी.ई.पी, जेनकिंस जी एम एवं रेनसल जी.सी (2008): “टाईम-सीरीज़ एनालिसिस”, फोरकास्टिंग एण्ड कंट्रोल, एड 4, परेटिस हाल, यूएसए ।
2. एंगल, आर. एफ. एवं ग्रेनगर, सी. डब्ल्यू. जे. (1987): “को-इंटिग्रेशन एण्ड एरर-कोरेक्शन: रिप्रजेंटेशन, एस्टीमेशन एण्ड टेस्टिंग” इकनोमेट्रिक्स 55: 251-276 ।
3. जेफी, एस. (2005): “डिलिवरिंग एण्ड टेकिंग दी हीट”, इंडियन स्पार्डिसिस एण्ड इवोलिविंग प्रोडक्ट एण्ड प्रोसेस स्टेण्डर्ड्स. एग्रीकल्चर एण्ड रूरल डेवलेपमेंट डिस्कशन पेपर 19, दी वर्ल्ड बैंक, यू.एस.ए. ।
4. कुमार, ए. एण्ड राय, डी.सी. (2011): “ट्रेन्ड्स एण्ड वोलैटिलिटी इन डोमेस्टिक एण्ड इंटरनेशनल प्राइज़ आफ लाइवस्टॉक प्रोडक्ट्स इन इंडिया”, इंडियन जरनल आफ एनिमल साइंसिस 81 (12): 1248-1252 ।
5. कुमार, ए. जी. एस. एण्ड यादव, सी. (2012): “एक्सपर्ट आफ बुफेलो मीट फ्रोम इंडिया”, परफोरमेंस एण्ड प्रोसपेक्ट्स इंडियन जरनल आफ एनिमल साइंसिस 82(12): 1578-1583 ।
6. कुमार, ए. (2010): “एक्सपर्ट्स आफ लाइवस्टॉक प्रोडक्ट्स फ्रोम इंडिया”, परफोरमेंस, कम्पीटिटिवनेस एण्ड डेटरमिनेन्ट्स एग्रीकल्चरल इकनोमिक्स रिसर्च रिव्यू 23: 57-67 ।
7. पून, एस.एच. एण्ड ग्रेनगर, सी. डब्ल्यू.जे. (2003): “फोरकास्टिंग वोलैटिलिटी इन फाइनेनशियल मार्केट्स”, रिव्यू जरनल आफ इकनोमिक्स लिटरेचर 41: 478-539 ।

□

## भारत में पशुधन एवं डेरी उत्पादन की स्थिति

वसी आलम, रंजीत कुमार पॉल, अमृत कुमार पॉल, बिशाल गुरुंग,  
कंचन सिन्हा, कमलेश नारायन सिंह, उषा जैन एवं राम नरेश

भारत की ग्रामीण जनसंख्या के कल्याण में पशुधन क्षेत्र की महत्वपूर्ण भूमिका है। पशुधन पारिवारिक आय में, विशेषरूप से राष्ट्र के शुष्क एवं अर्ध-शुष्क क्षेत्रों में, स्थिरता उपलब्ध कराता है। गत 30 वर्षों में कुल सकल घरेलू उत्पाद (जी.डी.पी.) में कृषि के योगदान में सतत गिरावट के बावजूद भी पशुधन क्षेत्र, विशेषरूप से डेरी उत्पादन, के योगदान ने कुल जी.डी.पी. में सतत वृद्धि की प्रवृत्ति दर्शायी है। यह क्षेत्र भारतीय अर्थव्यवस्था के महत्वपूर्ण विकास के रूप में उभर रहा है। वर्ष 1998-99 में 74 मिलियन टन वार्षिक उत्पादन से वर्ष 2001 में 127.3 मिलियन टन की वृद्धि हुई है। यह वृद्धि उत्पादकता में सुधार तथा भैंसों एवं संकर मवेशियों के प्राथमिकता में बदलाव होने के कारण है। पशु प्रबन्धन एवं निर्णय लेने के लिए दूध उत्पादन का काल-श्रेणी विश्लेषण एक महत्वपूर्ण टूल रहा है क्योंकि इससे छिपी प्रवृत्तियों और मौसमीय पैटर्नों का पता चलता है। पशुधन प्रबन्धन में, साथ ही अन्य मामलों में भी, पूर्वानुमान नीतियाँ या तो विवरणनात्मक मॉडलों पर अथवा व्याख्यात्मक मॉडलों के विकास पर आधारित हैं। भारत में दूध उत्पादन के पूर्वानुमान के लिए सिंह इत्यादि (2007) ने सांख्यिकीय मॉडल लागू किये। पॉल एवं दास (2013) ने कृषि में मॉडल एरिमा (ARIMA) एप्लाइ किये हैं। एरिमा अप्रोच का एक लाभ यह है कि यह सिस्टम के लिए अच्छी समझ उपलब्ध कराने में सक्षम है। वर्तमान अध्ययन में, भारत में दूध उत्पादन की मॉडलिंग एवं पूर्वानुमान के लिए एरिमा मॉडल का प्रयोग किया गया है तथा गत तीन दशकों में उत्पादन

की प्रवृत्ति का अध्ययन किया गया है। भारत में कुल दूध उत्पादन में एक महत्वपूर्ण बढ़ती हुई रैखिक प्रवृत्ति पायी गयी है। वर्ष 2015 तक आउट-ऑफ-सैम्पल पूर्वानुमान हेतु विचाराधीन आँकड़ों के लिए पहचान किये गये उत्तम मॉडल का प्रयोग किया गया है।

### आँकड़ों का विवरण

वर्ष 1994 से 2004 तक के कृषि एवं सम्बद्ध क्षेत्रों के साथ-साथ पशुधन क्षेत्र की विकास दर के आँकड़े तथा 1979 से 2011 तक के भारत में दूध उत्पादन के काल-श्रेणी आँकड़े (मिलियन टन में) ग्यारहवीं पंचवर्षीय योजना (2007-2012) के लिए पशुपालन एवं डेरी पर कार्यशील समूह की रिपोर्ट तथा आंशिक आँकड़े पशुपालन, डेरी एवं मात्स्यिकी विभाग, कृषि मंत्रालय (<http://www.nddb.org/English/Statistics/Pages/Milk-Production.aspx>) से लिये गये हैं। वर्ष 1979 से 2007 के मिलियन टन में दूध उत्पादन के आँकड़े मॉडल के विकास के लिए तथा 2008 से 2011 के आँकड़े मॉडल के वैधीकरण के लिए प्रयोग किये गये हैं। आँकड़ों के विश्लेषण के लिए एस.ए.एस. 9.3 सांख्यिकीय सॉफ्टवेयर पैकेज का प्रयोग किया गया है।

### स्वसमाश्रयी एकीकृत गतिमान माध्य (ARIMA) अप्रोच

एरिमा (ARIMA) का तात्पर्य स्वसमाश्रयी एकीकृत गतिमान माध्य है। एरिमा पद्धति तीन चरणों में चलायी जाती है अर्थात् पहचान, आकलन एवं नैदानिक जाँच (diagnostic checking)। पहचान चरण पर अन्तरिम

रूप से चुने गए एरिमा मॉडल के प्राचलों का आकलन, आकलन चरण में किया जाता है तथा अन्तरिम रूप से चयनित मॉडल की उपयुक्तता की जाँच नैदानिक जाँच चरण पर की जाती है। यदि मॉडल अनुपयुक्त पाया जाता है तो तीनों चरणों की तब तब पुनरावृत्ति की जाती है जब तक विचाराधीन काल-श्रेणी के लिए संतोषजनक एरिमा मॉडल का चयन न हो जाए।

### परिणाम एवं परिचर्चा

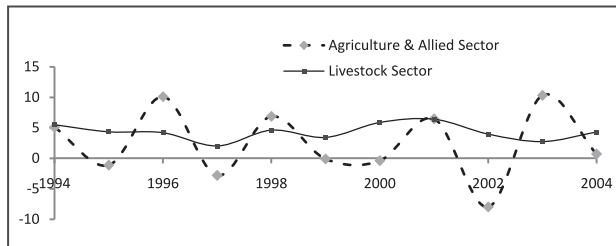
संयोजित प्रवृत्ति समीकरण निम्न रूप में प्राप्त होता है :

$$Y_t = 21.449 + 2.872 t$$

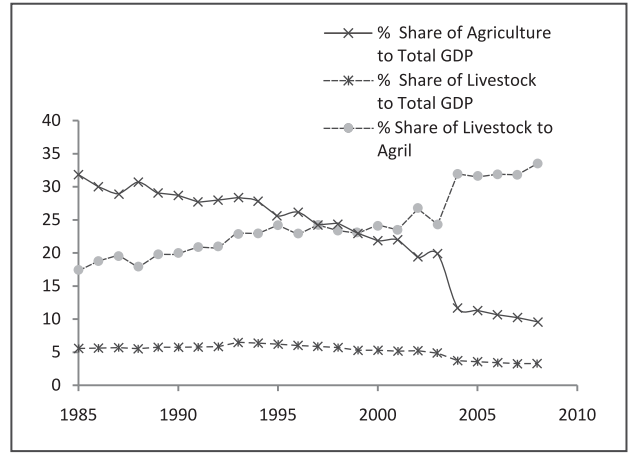
(1.553) (0.079)

जहाँ पर कोष्ठक () में दिये गये मान आकलनों की अनुवर्ती मानक त्रुटियों को दर्शाते हैं। सार्थकता के 1 प्रतिशत स्तर पर प्रवृत्ति सार्थक पायी गयी।

आरेख 1 से यह स्पष्ट है कि गत वर्षों के दौरान कृषि की विकास दर (1993-94 के मूल्यों के आधार पर जी.डी.पी. में प्रतिशत हिस्सा) स्थिर रही है तथा इसमें मानसून एवं अन्य जलवायवीय कारकों के आधार पर सार्थक उतार-चढ़ाव आया है। हाल में कृषि के विकास में मंदी आयी है। इसके विपरीत पशुधन क्षेत्र ने स्थिर विकास दर दर्शाई है तथा इससे सम्पूर्ण पारिवारिक आय में स्थिरता आयी है। आरेख 2 कुल जी.डी.पी. में कृषि एवं पशुधन के प्रतिशत हिस्से तथा कृषि में पशुधन के प्रतिशत हिस्से को भी दर्शाता है।



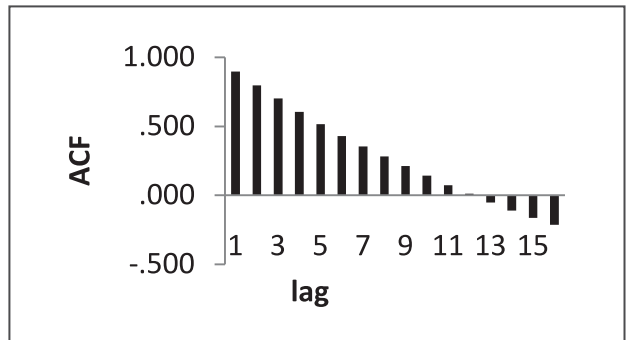
आरेख 1 : गत वर्षों के दौरान कृषि क्षेत्र एवं पशुधन क्षेत्र में विकास



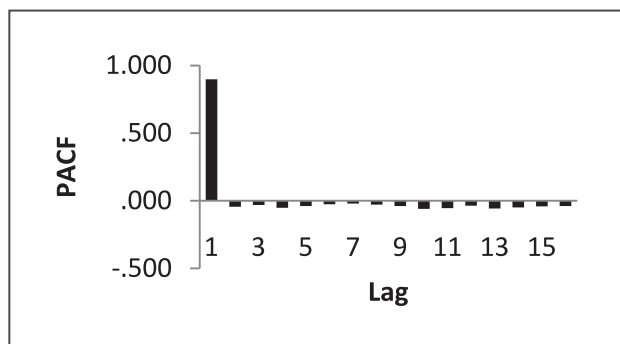
आरेख 2 : कुल जी.डी.पी. में कृषि एवं पशुधन का प्रतिशत हिस्सा तथा कृषि में पशुधन का प्रतिशत हिस्सा

### एरिमा (ARIMA) मॉडल का संयोजन

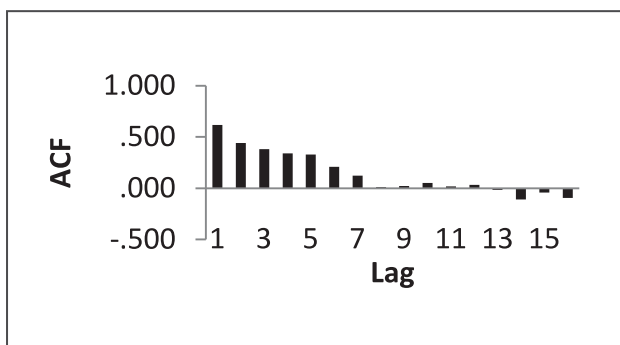
डिट्रेंडेड अवशिष्ट श्रृंखला  $\hat{\epsilon}_t$  गैर-स्थिर पायी गयी। स्थिरता प्राप्त करने के लिए विभेदन (differencing) किया जाता है। आरेख 2ए एवं 2डी से यह पाया गया कि यह बहुत धीरे-धीरे क्षीण (decays) होती है जिससे विभेदन की आवश्यकता होती है ताकि परिणामी श्रृंखला संभावित एरिमा मॉडलिंग के लिए एक पैटर्न दर्शा सके। एरिमा मॉडल के क्रम का चयन करने के लिए यूनिट रूट परीक्षण लागू किया गया, इससे यूनिट रूट की उपस्थिति ज्ञात हुई तथा इसके लिए विभेदन की आवश्यकता है। मूल श्रृंखला के प्रथम अन्तर को लेने पर यह देखा गया कि केवल कुछ ही acfs उच्च हैं, जिससे मॉडल के क्रम का चयन आसान हो जाता है।



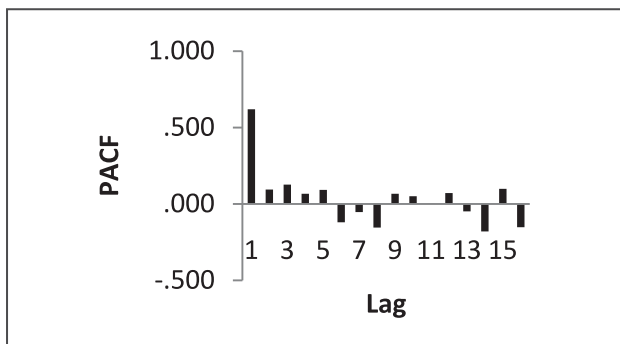
आरेख 2ए : मूल श्रृंखला के ACF



आरेख 2बी : मूल श्रृंखला के PACF



आरेख 2सी : विभेदित श्रृंखला के ACF



आरेख 2डी : विभेदित श्रृंखला के PACF

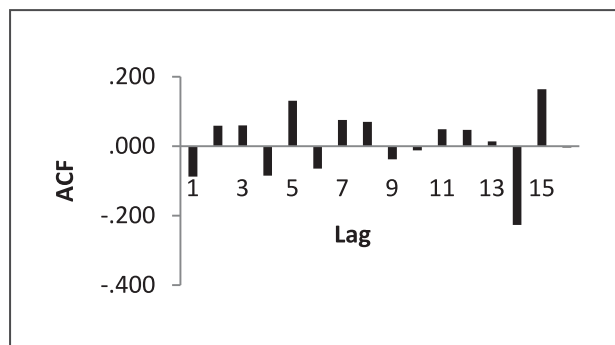
अतः, एक बार विभेदन के बाद, परिणामी श्रृंखला स्थिर बन जाती है। इसके स्व-सहसम्बन्ध फलन (acf) एवं आंशिक स्व-सहसम्बन्ध फलन (pacf) की जाँच करने पर एरिमा (1,1,0) मॉडल का चयन किया जाता है। संयोजित एरिमा (1,1,0) मॉडल के प्राचल आकलक एवं अनुवर्ती मानक त्रुटियाँ तालिका 2 में दी गयी हैं।

तालिका 2 : प्राचल आकलक एवं अनुवर्ती मानक त्रुटियाँ (एस.ई.)

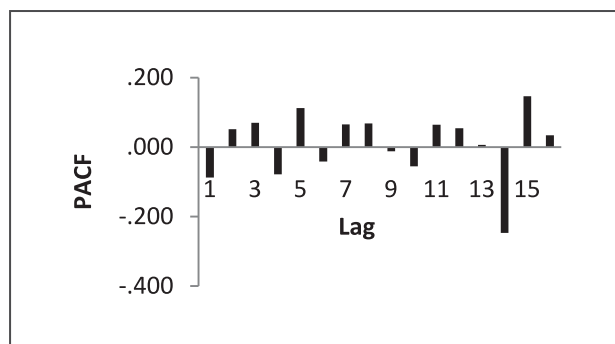
प्राचल	आकलक	एस.ई.	t-मान	सार्थकता
Constant	2.817	0.458	6.156	<0.001
AR1	0.612	0.178	3.433	0.002

### नैदानिक जाँच

मॉडल का सत्यापन मॉडल के अवशेष की जाँच से सम्बन्धित है, यह देखने के लिए कि कहीं उनमें व्यवस्थित पैटर्न निहित है, जिसे चयनित एरिमा में सुधार लाने के लिए अभी भी हटाया जा सकता है, जो विभिन्न क्रमों के अवशेषों की स्व-सहसम्बन्ध एवं आंशिक स्व-सहसम्बन्ध की जाँच के माध्यम से किया गया। अवशेषों के ACF एवं PACF आरेख 4 में दिये गये हैं जो मॉडल की समंजन सुष्ठुता को भी दर्शाते हैं।



आरेख 4ए : ACF



आरेख 4बी : PACF

आरेख 4 : अवशेष श्रृंखलाओं के ACF एवं PACF

**तालिका 3 :** भारत में दूध उत्पादन के पूर्वानुमान के लिए मॉडल का वैधीकरण

वर्ष	वास्तविक	ARIMA(1,1,0) द्वारा पूर्वानुमान	निम्न विश्वास्यता सीमा	उच्च विश्वास्यता सीमा
2008	112.2	112.24	110.25	114.23
2009	116.4	115.99	112.21	119.76
2010	121.8	119.37	113.90	124.84
2011	127.3	122.54	115.51	129.57

संयोजित काल-श्रृंखला मॉडल में उपयुक्तता मापने के लिए माध्य निरपेक्ष त्रुटि (MAE), माध्य निरपेक्ष प्रतिशत त्रुटि (MAPE) एवं सापेक्ष माध्य निरपेक्ष प्रागुक्ति त्रुटि (RMAPE) की गणना की गयी है। संयोजित एरिमा (1,1,0) मॉडल के लिए MAPE, RMAPE एवं RMSE मानों की क्रमशः 1.9%, 1.5% एवं 2.68% के रूप में गणना की गयी।

### पूर्वानुमान

वर्ष 2012–2015 के लिए भारत में कुल दूध उत्पादन के पूर्वानुमान के लिए श्रेष्ठ मॉडल अर्थात् एरिमा (1,1,0) मॉडल का प्रयोग किया गया है तथा यह पूर्वानुमान त्रुटि प्रसरण के साथ तालिका 4 में रिपोर्ट किया गया है।

**तालिका 4 :** भारत में दूध उत्पादन के पूर्वानुमान (मिलियन टन में)

वर्ष	पूर्वानुमान (मिलियन टन में)	निम्न विश्वास्यता सीमा	उच्च विश्वास्यता सीमा
2012	132.13	130.17	134.10
2013	136.48	132.57	140.40
2014	140.48	134.58	146.39
2015	144.23	136.39	152.07

भारत में दूध के कुल उत्पादन में महत्वपूर्ण बढ़ती हुई प्रवृत्ति पायी गयी है। एरिमा (1,1,0) मॉडल ने डाटा सेट में उपस्थित विचरण को काफी संतोषजनक तरीके से पता लगाया। व्याख्यात्मक परिवर्तनशीलता एवं प्रागुक्ति शक्ति के संदर्भ में इस मॉडल ने अच्छा निष्पादन किया है। प्रस्तुत अध्ययन के निष्कर्षों ने भारत में निर्णय लेने एवं पशुधन प्रबन्धन के लिए सटीक पूर्वानुमानों के संभावित उपयोग के लिए प्रत्यक्ष सहयोग उपलब्ध कराया है।

### संदर्भ

1. पॉल, आर.के. एवं दास, एम.के. (2013): “फोरकॉस्टिंग ऑफ एवरेज एनुअल फिश लैण्डिंग इन गंगा बेसिन”, फिशिंग चाइम्स, 33 (3), 51–54।
2. पॉल, आर.के., पंवार, एस., सरकार, एस.के., कुमार, ए., सिंह, के.एन., फारुकी, एस. एवं चौधरी, वी.के. (2013): “मॉडलिंग ऐण्ड फोरकॉस्टिंग ऑफ मीट एक्सपोर्ट फ्रॉम इण्डिया”, एग्रिकल्चरल इकोनॉमिक्स रिसर्च रिव्यू, 26 (2), 249–256।
3. सिंह, एस., रामसुब्रमण्यन, वी. एवं मेहता एस.सी. (2007): “स्टैटिस्टिकल मॉडल्स फॉर फोरकॉस्टिंग मिल्क प्रोडक्शन ऑफ इण्डिया”, जर्नल ऑफ इण्डियन सोसाइटी ऑफ एग्रिकल्चरल स्टैटिस्टिक्स, 61(2), 2007 : 80–83।

□



## गुणात्मक और मात्रात्मक आँकड़ों के मिश्रण पर आधारित जर्मप्लाज़म का एक कोर सेट के विकसित करने की विधि

उदय प्रताप सिंह, आत्माकुरि रामाकृष्णा राव, सन्त दास वाही,  
प्रबिण मेहेर एवं रूपम कुमार सरकार

फसल से संबंधित वैश्विक जर्मप्लाज़म का एक विशाल संग्रह पारंपरिक लैंड्रेसिस, आधुनिक खेती और जंगली किस्मों में शामिल हैं। हालांकि, इन जर्मप्लाज़म संग्रह में से केवल एक अंश की ही रक्षा जीन बैंकों द्वारा की जा सकती है। फ्रैंकल एवं ब्राउन (1984) ने एक बड़े जर्मप्लाज़म संग्रह की एक सबसेट के रूप में एक कोर संग्रह की अवधारणा शुरू की जो कि आनुवांशिक और प्ररूपी विविधता का प्रतिनिधित्व करता है। एक कोर सेट के विकास के लिए मौजूदा तरीकों या तो गुणात्मक या मात्रात्मक आँकड़ों पर आधारित है। कभी-कभी, परिवर्तनों का उपयोग मिश्रण आँकड़ों से निपटने में कठिनाई से बचने के लिए उन्हें गुणात्मक या ठीक इसके विपरीत मात्रात्मक आँकड़ा बनाने के लिए लागू किया जाता है। (किम एट एल. 2007)

विभिन्न क्लस्टरिंग के विधियाँ गुणात्मक या मात्रात्मक आँकड़ों पर आधारित सजातीय स्तर प्राप्त करने के लिए प्रयोग की जाती हैं। लेकिन, परिणाम अत्यधिक क्लस्टरिंग के तरीके पर निर्भर है, और ज्यादातर अनुमानी विधियाँ (किम एट एल, 2007) जर्मप्लाज़म संग्रह में सजातीय स्तर निर्धारित करने के लिए पालन किया गया है। एक कोर सेट के विकास के लिए विभिन्न कारकों को संबोधित किया जाना चाहिए जैसे कि कोर सेट के आधार, सजातीय तबके के गठन और नमूना रणनीति (वैन हिनटम और Th ,1999) शामिल हैं। अतीत में, कई अध्ययनों में जर्मप्लाज़म का एक बड़ा संग्रह से कोर सेट बनाया गया है, चावल जर्मप्लाज़म के संयुक्त राज्य अमेरिका कृषि विभाग (यू

एस डी ए) (यान एट अल., 2007) से नेलेकर एक कोर सेट के विकसित कर वर्णन किया गया और इस चावल की यू एस डी ए की कोर संग्रह से मिनी कोर सत्ता कोर सॉफ्टवेयर का उपयोग कर के बनाई गयी (आर डी ए जैनबैंक सूचना केन्द्र, [http://www.genebank.go.kr/eng/PowerCore/PowerCore\\_Software.zip](http://www.genebank.go.kr/eng/PowerCore/PowerCore_Software.zip)), (एगरोमा एट एल., 2009), (किम एट एल., 2007), गंगोपाध्याय एट एल (2010) द्वारा बैंगन जर्मप्लाज़म का एक कोर सेट विकसित करने के लिए प्रमुख घटक स्कोर रणनीति का प्रयोग किया गया। शर्मा एट एल (2010) ने अनाज ढाकना और कोमल फफूंदी प्रतिरोध के स्त्रोतों की पहचान करने के लिए विभिन्न परिग्रहण के प्रमुख संग्रह से विकसित मिनी कोर का मूल्यांकन किया। Yu एट एल (2012) ने मार्कर आँकड़ों की एक जीनोम वाइड कवरेज के साथ कपास जर्मप्लाज़म की एक कोर विकसित किया। वेन एट एल (2012) ने उष्णकटिबंधीय मक्का रेस टक्सपेनो जीनोम वाइड एकक न्यूक्लियोटाइड बहुरूपताओं (SNPs) का उपयोग भविष्य में मक्का सुधार में शोषण की जांच के लिए कैसे किया जा सकता है। जीबर्ट (Gibert) और कोर्ट्स (Cortes) (1997) ने गुण और क्लस्टर विश्लेषण के लिए गुणात्मक और मात्रात्मक चर भारों द्वारा प्राप्त दूरी मैट्रीसेस का ब्योरा प्रस्तुत किया हालांकि, भार मापदण्डों का उद्देश्य विकल्प अक्सर मुश्किल है क्योंकि विभिन्न स्त्रोतों से दूरी मैट्रीसेस के भार समस्याग्रस्त है। क्रोसा और फ्रैन्को (2004) ने मिश्रित वितरण मॉडल पर आधारित जीनोमिक वर्गीकरण तकनीकों के साथ-साथ सांख्यिकीय मॉडल

की भी समीक्षा की। डोरिंग एट एल (2004) ने मिश्रित आंकड़ों के लिए एक फज़ी क्लस्टरिंग प्रक्रिया का प्रस्ताव रखा। सरकार एट एल (2011) ने मिश्रित आंकड़ों के आधार पर विभिन्न क्लस्टरिंग प्रक्रियाओं के प्रदर्शन की तुलना की। हालांकि, एक कोर सेट के विकास के लिए मिश्रित आंकड़ों का पूर्ण उपयोग के साथ समूहों के इष्टतम संख्या की पहचान एक चुनौती बनी हुई है। हमारे ज्ञान के लिए, एक कोर सेट के विकास के लिए मौजूदा तरीकों में ज्यादातर या तो गुणात्मक या मात्रात्मक लक्षण पर आधारित हैं। इसके अलावा, एक उपयुक्त दूरी माप, क्लस्टरिंग कार्यप्रणाली, जर्मप्लाज़म का मिश्रित आंकड़ों पर आधारित एक कोर सेट के विकास के लिए, आवंटन रणनीति और मूल्यांकन मानदण्डों समूहों की संख्या की पहचान का अभी तक पूरी तरह से पता नहीं लगाया जा सका है। इसलिए वर्तमान अध्ययन में एक व्यवस्थित दृष्टिकोण मिश्रित आंकड़ों का उपयोग कर जर्मप्लाज़म का एक कोर सेट के विकास के लिए प्रस्तावित किया गया था। इस प्रकार विकसित दृष्टिकोण दोनों मात्रात्मक और गुणात्मक एस एन पी (SNP) जीनोटाइपिंग आंकड़ों का प्रयोग कर चावल जर्मप्लाज़म पर सचित्रीत किया गया है।

### सामग्री और विधियां

कोर संग्रह की पहचान एक दो कदम प्रक्रिया है। शुरु में परिग्रहण की एक उपयुक्त नमूना या आवंटन रणनीति का उपयोग करके मूल संग्रह के लिए चुना जाता है तो एक अंश सजातीय तबके में वर्गीकृत किया है और कर रहे है, प्रत्येक परत से विलय का गुणात्मक और मात्रात्मक आंकड़ों का एक मिश्रण का उपयोग करके क्लस्टरिंग तकनीक को सक्षम करने से पहले गुणात्मक और मात्रात्मक आंकड़ों पर दूरी प्रासंगिक उपयों का उपयोग के लिए अलग से गणना करते हैं। फिर, इन दूरी मैट्रीसेस के सीधे संयुक्त करके क्लस्टर विश्लेषण के लिए सामग्री के रूप में उपयोग किया जाता है।

इस अध्ययन में, जीनोम वाइड SNPs वाले 219 नमक-सहिष्णु चावल जर्मप्लाज़म परिग्रहणों को शामिल

किया गया था (प्रत्येक परिग्रहण के लिए एक विषम युग्मज के लिए 0 और 2 क्रमशः तथा समयुग्मज के लिए 1 के रूप में कोडित) 14 कृषि/प्रारूपी विशेषताओं तथा उपलब्धियों।

### दूरी माप

निम्न तीन दूरी मापों जो मात्रात्मक आंकड़ों के आधार पर एसेसन के बीच दूरी तय करते है का उपयोग किया है।

- (1) औसत सीमा मानकीकृत पूर्ण अंतर के आधार पर दूरी

$$A_1 = \frac{1}{p} \sum_{k=1}^p \frac{|x_{ik} - x_{jk}|}{r_k}$$

जहाँ  $x_{ik}$  और  $x_{jk}$ ,  $i^{\text{th}}$  और  $j^{\text{th}}$  परिग्रहण,  $k^{\text{th}}$  मात्रात्मक चर है,  $r_k$   $k^{\text{th}}$  चर की सीमा है और  $p$  मात्रात्मक चरों की कुल संख्या हैं (ग्रोवर, 1971)

- (2) पियर्सन सहसंबंध पर आधारित दूरी :

$$A_2 = (1 - r_{ij}^2)$$

जहाँ  $r_{ij}$ ,  $i^{\text{th}}$  और  $j^{\text{th}}$  परिग्रहणों के बीच सहसंबंध है। इस प्रकार विषमता = 1- समानता है।

- (3) मानकीकृत स्कोर पर आधारित रिस्केल्ड (rescaled) दूरी:

$$A_3 = \sum_{k=1}^p \frac{\left[ \frac{x_{ik} - x_{jk}}{\sigma_k} \right]^2}{\max(d_{ij}^*)}$$

जहाँ  $\sigma_k$ ,  $k^{\text{th}}$  चर का मानक विचलन है। और मैक्स ( $d_{ij}^*$ ) जो पूरे परिग्रहण आंकड़ा के सेट में दो परिग्रहणों के बीच की अधिकतम दूरी है।

गुणात्मक आंकड़ों के लिए निम्नलिखित दो अलग दूरी मापों द्वारा एसेसन के बीच दूरी तय करने पर विचार किया गया।

- (1) औसत बेमेल पर आधारित दूरी

$$B_1 = \frac{1}{m} \sum_{k=1}^m d_k$$

जहाँ  $d_k = 0$ , यदि  $y_{ik} = y_{jk}$ , जबकि  $d_k = 1$  (ग्रोवर, 1971)।

(2) औसत पूर्ण अन्तर के आधार पर रिस्केल्ड (rescaled) दूरी:

$$B_2 = \frac{3}{2} \times \frac{\frac{1}{m} \sum_{k=1}^m |y_{ik} - y_{jk}|}{1 + \frac{1}{m} \sum_{k=1}^m |y_{ik} - y_{jk}|}$$

जहाँ  $y_{ik}$  और  $y_{jk}$ ,  $i^{\text{th}}$  और  $j^{\text{th}}$  परिग्रहणों का  $k^{\text{th}}$  गुणात्मक चर है और  $m$  गुणात्मक चरों की कुल संख्या है। दूरी  $B_2$  मुनेक एट एल (2005) द्वारा दिया गया संशोधित माप है। इतने संशोधन किया जाता है कि  $B_2$  की मूल्य सीमा  $[0, 1]$  के बीच में निहित है।

इस प्रकार तीन मात्रात्मक दूरी मैट्रिसिस में तत्वों ( $A_1$ - $A_3$ ) के रेंज और दो गुणात्मक दूरी मैट्रिसिस ( $B_1$  और  $B_2$ ) जो 0 और 1 के बीच है। मिश्रित आंकड़ों के लिए विभिन्न संयुक्त दूरी मैट्रिसिस इसी जो गुणात्मक और मात्रात्मक आंकड़ों दूरी मैट्रिसिस को जोड़ कर इस रूप में परिभाषित किए गये हैं जो कि निम्न है:

$$A_1 B_1 = ((a_{1ij}) + (b_{1ij}))$$

$$A_2 B_1 = ((a_{2ij}) + (b_{1ij}))$$

$$A_3 B_1 = ((a_{3ij}) + (b_{1ij}))$$

$$A_1 B_2 = ((a_{1ij}) + (b_{2ij}))$$

$$A_2 B_2 = ((a_{2ij}) + (b_{2ij}))$$

$$A_3 B_2 = ((a_{3ij}) + (b_{2ij}))$$

जहाँ  $(a_{1ij}), (a_{2ij}), (a_{3ij}), (b_{1ij}), (b_{2ij})$ , और  $(b_{2ij})$ , मैट्रिसिस  $A_1, A_2, A_3, A_3, B_1, B_2$ , के परजी तत्वों और  $B_2$  का क्रमशः प्रतिनिधित्व करता है। इन गुणात्मक, मात्रात्मक और संयुक्त दूरी मैट्रिसिस के क्लस्टरिंग विश्लेषण की सामग्री के रूप में उपयोग किया गया है। इस अध्ययन में सजातीय समूहों की अधिकतम संख्या को खोजने के लिए अलग क्लस्टरिंग प्रक्रियाओं अर्थात् एक उठाना, पूरी लिंकेज, भारत औसत, वार्ड विधि, k-means और मेडोयड्स आसपास विभाजन माना जाता था मतलब गणित के साथ अनिर्धारित जोड़ी-समूह विधि, सात (पांच श्रेणीबद्ध और दो विभाजित) किया गया। यहाँ मॉटि एट एल (2003) के दृष्टिकोण बूटस्ट्रैपिंग द्वारा समूहों की स्थिरता का आकलन करने के लिए पीछा किया गया था और

1000 बूटस्ट्रैप नमूने क्लस्टर संख्या  $k = \{2, \dots, 10\}$  के प्रत्येक सेट के लिए दूरी के मैट्रिसिस को तैयार किया गया। एक आम सहमति क्लस्टरिंग परिणाम किन्हीं दो टिप्पणियों बूटस्ट्रैप नमूनों में एक साथ चुना जाता है कि बार की कुल संख्या के एक ही समूह में एक साथ पाए जाते हैं जो समय की संख्या के अनुपात लेने के द्वारा प्राप्त हुई थी। प्रत्येक क्लस्टरिंग प्रक्रिया व्यक्तियों की अलग क्लस्टरिंग सदस्यता दर्शाती है, के रूप में विभिन्न क्लस्टरिंग प्रक्रियाओं से प्राप्त सर्वसम्मति क्लस्टरिंग परिणाम एक विलय-आम सहमति क्लस्टरिंग परिणाम प्राप्त करने के लिए एक साथ विलय कर रहे हैं (सिम्पसन एट एल, 2010) विलय-आम सहमति क्लस्टरिंग परिणाम एक विशेष क्लस्टर संख्या  $k$  के लिए आम सहमति क्लस्टरिंग परिणामों के औसत लेने के द्वारा प्राप्त की है। इस कारण क्लस्टरिंग तर्ज पर किसी भी एक प्राथमिकताओं के बारे में जानकारी के अभाव में बराबर वजन प्रत्येक आम सहमति क्लस्टरिंग परिणाम को दिया जाता है, यानि बराबर महत्व क्लस्टरिंग प्रक्रिया से प्रत्येक के लिए दिया जाता है।

### समूहों की अधिकतम संख्या

ज्यादातर, पूर्व जानकारी (apriori) के आधार पर परिग्रहण वर्गीकृत करने के लिए समूहों की संख्या का निर्धारण के लिए प्रयोग किया जाता है, लेकिन इस तरह की जानकारी के अभाव में, यह समूहों की अधिकतम संख्या की पहचान के लिए भी उपयोगी है। इसके अलावा, समूहों का इष्टतम संख्या की पहचान सबसे चुनौतीपूर्ण मूद्दों में से एक प्रभावी और कुशल क्लस्टरिंग के लिए आवश्यक है (Everitt, 1979)। वह है जोकि समूहों ( $k$ ) का इष्टतम संख्या वह है जोकि  $k$  संभावित मूल्यों की एक सीमा के पार गणना की इष्टतम संख्या  $k$  का अनुमानित मूल्य प्राप्त करने के लिए  $k$  की संभावित मूल्यों की सीमा के अन्दर वह मूल्य होता है जिस पर संचयी घनत्व समारोह (सी डी एफ) ( $\Delta k$ ) के तहत क्षेत्र में परिवर्तन सबसे बड़ा है, जिस पर  $k$  के मूल्य के रूप में अनुमान लगाया गया है। माना कि  $M$  के विलय आम सहमति क्लस्टरिंग परिणाम

इंगित करता है। जिसका आर्डर  $N \times N$  है। अनुभवजन्य सी डी एफ द्वारा प्राप्त सीमा  $[0,1]$  के ऊपर परिभाषित किया जाता है।

$$CDF(c) = \frac{\sum_{i < j} I\{M(i, j) \leq c\}}{N(N-1)/2}$$

जहाँ  $1\{\dots\}$  के साथ एक सूचक समारोह,  $M(i, j)$  के साथ  $(i, j)$  विलय-आम सहमति मैट्रिक्स  $M$  का एक तत्व है।  $M$  की गणना सी डी एफ के तहत क्षेत्रफल का उपयोग का सूत्र इस प्रकार है:

$$AUC = \sum_{i=2}^m [x_i - x_{i-1}] CDF(x_i)$$

जहाँ  $\{x_1, x_2, \dots, x_m\}$  आर्डर सेट के तत्वों का मैट्रिक्स  $M$  विलय-आम सहमति का  $m = N(N-1)/2$  के साथ है, (मोंटे एट एल., 2003)।

### कलस्टर दृढ़ांगता

समूहों के इष्टतम संख्या निर्धारित करने के बाद, जर्मप्लाज़म का सबसे अच्छा-फिट कलस्टरिंग पैटर्न कलस्टर दृढ़ांगता के आधार पर निर्धारित किया जाता है। किसी भी कलस्टरिंग प्रक्रिया के तहत समूहों की दृढ़ांगता की गणना एक ही समूह में उन परिग्रहण का चयनित परिग्रहणों का विलय आम सहमति परिणाम की औसत लेने से गणना की है (सिम्पसन एट एल, 2010) जिसका सूत्र निम्न है।

$$m(k) = \frac{1}{N_k(N_k-1)} \sum_{i < j \in (E_k)} M(i, j)$$

औसत कलस्टर दृढ़ांगता मूल्य की गणना करने के लिए कलस्टरिंग कलन विधि का उपयोग के आधार पर सभी  $k$  समूहों भर में गणना की जाती है।

### आबंटन की विधि

एक कोर सेट के विकसित करने के लिए दूसरे और अंतिम चरण में सजातीय समूहों में से परिग्रहण चयन करने के लिए एक उपयुक्त नमूना या आवंटन रणनीति की आवश्यकता है। वैन Hintem एट एल (2000) और Hu एट एल (2000) खेतों में प्रयुक्त विभिन्न नमूने आनुपातिक आबंटन (पी रणनीति), लॉग

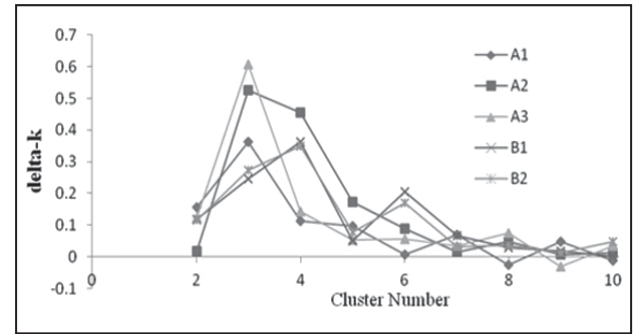
आवृत्ति आवंटन (एल रणनीति), स्थिर आबंटन (सी रणनीति) और सरल यादृच्छिक नमूना (आर रणनीति) रणनीतियों, एक कोर सेट की पहचान के लिए किया। इस चरण के दौरान, समूहों से नमूना परिग्रहण पहचान करके निम्नलिखित तीन अलग आबंटन के तरीकों का उपयोग किया है :

(1) आनुपातिक आबंटन

$$n_i = \left[ n \times \frac{N_i}{\sum_{i=1}^g N_i} \right]$$

(2) लॉग-आनुपातिक आबंटन

$$n_i = \left[ n \times \frac{\log(N_i)}{\sum_{i=1}^g \log(N_i)} \right]$$



चित्र 1: विभिन्न दूरी के उपायों का उपयोग कर गुणात्मक और मात्रात्मक आंकड़ों के लिए कलस्टर संख्या (k) के खिलाफ  $\Delta k$  की चित्रमय प्रतिनिधित्व

जहाँ  $n_i$ ,  $i^{\text{th}}$  कलस्टर सेट कोर के लिए चयनित परिग्रहण की संख्या है  $N_i$ ,  $i^{\text{th}}$  कलस्टर में विलय की संख्या है,  $n$  कोर सेट का आकार है,  $g$  समूहों की कुल संख्या और कोष्ठकों  $[\ ]$  के निकटतम पूर्णांक समारोह का प्रतिनिधित्व करते हैं।

(3) एकक प्रविष्टि की यादृच्छिक आवंटन (RASE)। यहाँ, समूहों की कोई इष्टतम संख्या निर्धारित की जाती है और परिग्रहण कोर सेट के आकार के बराबर होती है जो समूहों की संख्या में वर्गीकृत किया है। कलस्टर में से प्रत्येक एक एकल परिग्रहण कोर सेट के निर्माण के लिए चयन किया जाता है।

## कोर सेट का मूल्यांकन

मात्रात्मक आंकड़ों के लिए, एक कोर सेट की पहचान के लिए विधियों की दक्षता एवं विभिन्न सूचकांकों का उपयोग कर मूल्यांकन किया जाता है। मध्य अंतर (MD), विचरण अंतर (VD), चर दर (VR) और संयोग दर (CR) Hu एट एल, 2000) का गुणात्मक आंकड़ों के लिए, ऊपर उल्लिखित तरीके से सूचकांक औसत बहुरूपी जानकारी सामग्री का अंतर (APICD) के आधार पर मूल्यांकन कर रहे हैं। जोकि इस प्रकार है।

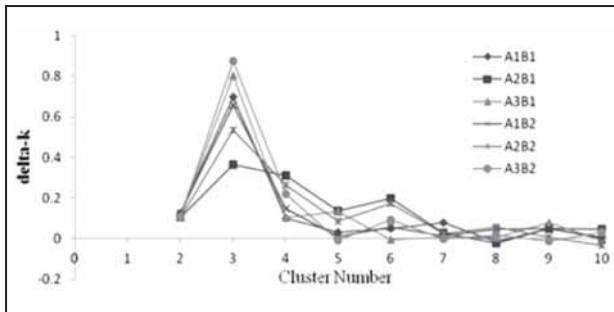
$$APICD(\%) = \frac{|\bar{P}_c - \bar{P}_c|}{\bar{P}_c} \times 100$$

जहाँ  $\bar{P}_c$  और  $\bar{P}_c$  क्रमशः कोर सेट और पूरे सेट सामग्री के परिग्रहणों की औसत बहुरूपी जानकारी है।

एक संयुक्त मूल्यांकन सूचकांक (CEI) मिश्रण आंकड़ों के आधार पर कोर सेट की विविधता का मूल्यांकन करने के लिए उपर्युक्त पांच सूचकांकों के संयोजन के द्वारा प्रस्तावित किया गया था। जोकि इस प्रकार है।

$$CEI = (w_1 M_1 + w_2 M_2) / (w_1 + w_2)$$

जहाँ,  $M_1 = [(100 - MD) + (100 - VD) + (100 - VRt) + CR]/4$ , साथ  $VRt = |100 - VR|$ ,  $M_2 = (100 - APICD)$ , और  $w_1 = (N_{quant}/N_T)$  और  $w_2 = (N_{qual}/N_T)$ ,  $N_{quant}$ ,  $N_{qual}$  और  $N_T = N_{quant} + N_{qual}$  क्रमशः मात्रात्मक, गुणात्मक और चर की कुल संख्या है साथ में  $w_1 + w_2 = 1$  और  $N_T = N_{quant} + N_{qual}$  कोर सेट और पूरे सेट के बीच समानता के प्रतिशत का प्रतिनिधित्व करता है। CEI के मूल्यों 0 से 100 के बीच हैं। इसके अलावा, 100 का



चित्र 2: मिश्रण आंकड़ों के लिए क्लस्टर संख्या (k) के खिलाफ ( $\Delta k$ ) की चित्रमय प्रतिनिधित्व संयुक्त दूरी उपायों का उपयोग

मूल्य दूरी आबादी का सबसे अच्छा प्रतिनिधि के तौर से मेल खाती है। आनुपातिक आवंटन के तहत CEI के बीच का अंतर, सभी दूरी के उपायों के लिए लॉग-आनुपातिक आवंटन और RASE तरीकों का बड़ा नमूना-2 z-परीक्षण का उपयोग कर परीक्षण किया है। सभी आवश्यक कोडिंग R साफ्टवेयर में किया जाता है। क्लस्टरिंग परिणाम आम सहमति और विलय-आम सहमति के लिए 'ClusterCons' पैकेज का इस्तेमाल किया गया था (सिम्पसन, 2010) और कोर सेट परिग्रहण नमूने के लिए 'ccChooser' पैकेज (Studnicki और Debski, 2012) में R साफ्टवेयर का इस्तेमाल किया गया था।

## परिणाम

आम सहमति द्वारा क्लस्टरिंग कार्यप्रणाली गुणात्मक और मात्रात्मक आंकड़ों के लिए अलग, उपयोग किया गया था, जिसमें मात्रात्मक आंकड़ों के लिए तीन दूरी उपायों ( $A_1$ - $A_3$ ) का और गुणात्मक आंकड़ों के लिए दो दूरी के उपायों ( $B_1$  और  $B_2$ ) का उपयोग लागू किया गया। चित्र 1 में  $\Delta k$  के मूल्यों के खिलाफ क्लस्टर संख्या (k) की चित्रित किया गया। चित्र 1 में दिखाया गया है, कि आंकड़ा मात्रात्मक और गुणात्मक दूरी के उपायों के आधार पर तीन और चार समूहों में क्रमशः शामिल हुआ। इस प्रकार, गुणात्मक या मात्रात्मक आंकड़ों का उपयोग अकेले व्यापक रूप से विभिन्न कोर सेटों में हो सकता है। इसके अलावा, छोड़ने या चर के किसी भी प्रकार बदलने (मात्रात्मक से गुणात्मक और इसके विपरीत) से विश्लेषण के सूचना का एक संभावित नुकसान हो सकता है। समूहों की संख्या के खिलाफ  $\Delta k$  मूल्यों की साजिश के रूप में संयुक्त गुणात्मक और मात्रात्मक आंकड़ों द्वारा चित्र 2 में दिखाया गया है। यह सब संयुक्त दूरी के उपायों के लिए, समूहों की संख्या के लिए के  $\Delta k$  शिखर मूल्य 3 के बराबर हो पाया था।

**तालिका 1:** औसत क्लस्टर मजबूती मूल्यों की विभिन्न क्लस्टरिंग के तरीके की विभिन्न संयुक्त दूरी उपायों का उपयोग

दूरी के उपाय	क्लस्टरिंग एल्गोरिथ्म						
	एकल	पूरा	औसत	भारित	वार्ड की विधि	k-माध्य	पाम (PAM)
$A_1B_1$	0.220	0.742	0.505	0.479	0.742	0.744	0.732
$A_2B_1$	0.258	0.860	0.850	0.848	0.824	0.845	0.844
$A_3B_1$	0.223	0.765	0.743	0.762	0.760	0.773	0.765
$A_1B_2$	0.474	0.708	0.453	0.698	0.712	0.727	0.715
$A_2B_2$	0.256	0.864	0.838	0.854	0.838	0.850	0.823
$A_3B_2$	0.244	0.848	0.244	0.820	0.823	0.828	0.792

जो देखा जा सकता है इसलिए, असहमति के मामलों में अकेले गुणात्मक या मात्रात्मक आंकड़ों का उपयोग करते हुए समूहों की संख्या को चुनने की समस्या, मिश्रण आंकड़ों के आधार पर संयुक्त दूरी के उपायों पर विचार करके हल किया जा सकता है।

अलग क्लस्टरिंग तरीकों और संयुक्त दूरी के उपायों के लिए औसत क्लस्टर मजबूती मूल्यों को तालिका 1 में दिया गया है। किसी दिए गए संयुक्त दूरी को मापने के लिए उच्चतम औसत क्लस्टर मजबूती मूल्य के साथ क्लस्टरिंग कार्यप्रणाली तो एक ओर सेट के विकास के लिए नमूना रणनीति अपनाने के लिए चुना गया था। यह पता करना कि k-माध्य क्लस्टरिंग एल्गोरिथ्म संयुक्त दूरी उपायों  $A_1B_1$ ,  $A_3B_1$  और  $A_1B_2$  पर आधारित जर्मप्लाज़म समूह के लिए उपयुक्त था जबकि  $A_2B_1$ ,  $A_2B_2$  और  $A_3B_3$  संयुक्त दूरी उपायों पर आधारित जर्मप्लाज़म समूह के लिए उपयुक्त पाया गया था।

परिग्रहण नमूने के लिए, तीन विभिन्न आवंटन के तरीकों अर्थात् आनुपातिक आवंटन लॉग-आनुपातिक आवंटन और आर ए एस ई को अपनाया गया। पहले दो आवंटन के तरीकों के लिए, परिग्रहण पूरे संग्रह से जर्मप्लाज़म के 20 प्रतिशत के साथ एक कोर सेट विकसित करने के लिए प्रत्येक संयुक्त दूरी को मापने के तहत पहचान, तीन समूहों से चयन किया गया था। इसके विपरीत, 44 समूहों (जर्मप्लाज़म की कुल संख्या का लगभग 20 प्रतिशत) में से प्रत्येक से एक भी प्रविष्टि की यादृच्छिक नमूना आर ए एस ई पद्धति का उपयोग समूहों के अधिकतम संख्या की अनदेखी द्वारा निर्धारित एक कोर सेट विकसित करने के लिए किया गया था। कोर सेट की पहचान करने के लिए प्रक्रियाओं की दक्षता का मूल्यांकन करने के लिए 500 स्वतंत्र कोर संग्रह प्रत्येक नमूने की रणनीति के तहत नकली थे।

**तालिका 2:** आनुपातिक आवंटन (CEI-P) लॉग आनुपातिक आवंटन (CEI-LP) और संयुक्त मूल्यांकन सूचकांक (CEI) से माध्य मूल्यों (कोष्ठक में उनके मानक त्रुटियों के साथ) और एकल प्रवेश की यादृच्छिक आवंटन (CEI-R) तरीकों और उनके जोड़ों में पूर्ण अंतर (AbsD)

	A1	A2	A3	B1	B2	A1B1	A2B1	A3B1	A1B2	A2B2	A3B2
CEI-P	93.47 (5.9)	94.84 (4.5)	93.83 (5.0)	98.05 (1.8)	97.95 (1.3)	96.83 (2.1)	96.53 (2.8)	96.95 (2.3)	97.81 (1.6)	95.74 (2.7)	97.05 (2.6)
CEI-LP	93.81 (4.8)	93.68 (5.5)	94.48 (4.5)	89.65 (2.5)	88.66 (2.9)	91.11 (2.8)	93.07 (4.1)	93.21 (4.1)	95.84 (2.3)	93.38 (4.4)	92.58 (4.1)

	A1	A2	A3	B1	B2	A1B1	A2B1	A3B1	A1B2	A2B2	A3B2
CEI-S	93.44 (5.3)	94.30 (5.5)	94.99 (4.1)	86.84 (0.3)	89.00 (0.3)	85.61 (0.5)	93.87 (1.0)	89.16 (0.9)	88.33 (0.6)	90.27 (0.4)	95.37 (1.6)
AbsD (P-LP)	0.34	1.16*	0.65	8.74*	9.29*	5.72*	3.46*	3.74*	1.97*	2.36*	4.47*
AbsD (P-S)	0.03	0.54	1.16*	11.21*	8.95*	11.22*	2.66*	7.79*	9.48*	5.47*	1.67*
AbsD (LP-S)	0.37	0.62	0.51	2.81*	0.34*	5.5*	0.80*	4.05*	7.51*	3.11*	2.79*

\*P< 0.05

CEI का माध्य मूल्यों, जो 500 से अधिक स्वतंत्र सिमुलेशन रन के तहत आनुपातिक, लॉग-आनुपातिक आवंटन और आर ए एस ई तरीकों को तालिका 2 में प्रस्तुत कर रहे हैं। आनुपातिक, लॉग आनुपातिक आवंटन और आर ए एस ई तरीकों के बीच CEI मूल्यों में पूर्ण अंतर का सांख्यिकीय परीक्षण किया गया जो तालिका में दिए गए हैं। तालिका 2 से, आनुपातिक और लॉग-आनुपातिक, तरीकों के बीच और आनुपातिक और आर ए एस ई के बीच CEI मूल्यों में अंतर काफी अधिक थे और यह स्पष्ट है कि इसलिए आनुपातिक आवंटन पद्धति की पहचान के लिए तीन आवंटन के तरीकों के बीच सबसे अच्छा इस्तेमाल किया गया दूरी उपायों की परवाह किए वगैर एक कोर सेट विधि का उपयोग किया था। इसके अलावा, संयुक्त दूरी उपायों के बीच इन CEI मूल्यों में मतभेद और गुणात्मक/ मात्रात्मक दूरी के उपायों के तहत आनुपातिक, लॉग-आनुपातिक और आर ए एस ई तरीकों को तालिका 3 में दिए गए हैं। आनुपातिक आवंटन के तरीकों के लिए, CEI का मूल्य संयुक्त दूरी के उपायों के बीच  $A_1B_2$  के लिए सबसे ज्यादा था। हालांकि, दूरी को मापने की  $A_1B_2$  की CEI मूल्यों  $A_3$  को  $A_1$  के उन लोगों से काफी अलग थे और एक ही समय में वे  $B_1$  और  $B_2$  की CEI मूल्यों से काफी अलग नहीं थे (तालिका 3)। इसके विपरीत, लॉग-आनुपातिक आवंटन पद्धति के लिए  $A_1B_2$  की CEI मूल्य सभी खगोलीय दूरी (तालिका 2) और बाकी हिस्सों से काफी अलग

की CEI मूल्यों में सबसे ज्यादा था (तालिका 3)। इसके अलावा, एक कोर सेट का निर्माण अनुमानी तरीकों के माध्यम से PowerCore का उपयोग कर के किया गया था (RDA-Genebank सूचना केन्द्र [http://www.genebank.go.kr/eng/PowerCore\\_software.zip](http://www.genebank.go.kr/eng/PowerCore_software.zip)) / PowerCore द्वारा निर्मित कोर सेट के CEI मूल्य आनुपातिक और लॉग-आनुपातिक आवंटन के तरीकों के तहत सभी संयुक्त दूरी उपायों में सबसे कम था जो 89.19, हो पाया था।

### विचार विमर्श

जर्मप्लाज़म संग्रह वर्गीकृत करने के लिए गुणात्मक और मात्रात्मक आँकड़ों का अलग-अलग उपयोग करने पर विभिन्न समूहों की संख्या और इसलिए, भिन्न समूह पैटर्न में प्रत्येक, क्लस्टरिंग कार्यप्रणाली के तहत अलग हो सकता है। इसलिए, गुणात्मक और मात्रात्मक आँकड़ों के विश्लेषण से प्राप्त अलग-अलग क्लस्टरिंग पैटर्न का सामान्यीकरण करना मुश्किल है। इसके अलावा, आँकड़ों छोड़ने या किसी भी प्रकार से बदलाव करने से उत्पन्न आँकड़ों की जानकारी का नुकसान हो सकता है। इसलिए, गुणात्मक और मात्रात्मक दूरी सूचकांकों द्वारा जानकारी के संयोजन में इस तरह के आँकड़ों को संभालने के लिए एक लाभदायक विधि है। वर्तमान अध्ययन में, छह अलग संयुक्त दूरी मापों का प्रस्ताव किया है और मूल्यांकन किया गया।

**तालिका 3:** आनुपातिक और लॉग-आनुपातिक आवंटन के तरीकों के तहत गुणात्मक/ मात्रात्मक दूरी उपायों और संयुक्त दूरी के उपायों के बीच संयुक्त मूल्यांकन सूचकांक में मतभेद का मान

	आनुपातिक आवंटन						लॉग-आनुपातिक आवंटन					
	$A_1B_1$	$A_2B_1$	$A_3B_1$	$A_1B_2$	$A_2B_2$	$A_3B_2$	$A_1B_1$	$A_2B_1$	$A_3B_1$	$A_1B_2$	$A_2B_2$	$A_3B_2$
$A_1$	3.36*	3.06*	3.48*	4.34*	2.27*	3.58*	2.70	0.74	0.60	2.03*	0.43	1.23
$A_2$	1.99*	1.69*	2.11*	2.97*	0.90*	2.21*	2.57	0.61	0.47	2.16*	0.30	1.10
$A_3$	3.00*	2.70*	3.12*	3.98*	1.91*	3.22*	3.37	1.41	1.27	1.36*	1.10	1.90
$B_1$	1.56*	1.86*	1.44	0.58	2.65*	1.34*	1.46*	3.42*	3.56*	6.19*	3.73*	2.93*
$B_2$	1.12*	1.42*	1.00*	0.14	2.21*	0.90*	2.45*	4.41*	4.55*	7.18*	4.72*	3.92*

\* $P < 0.05$

संयुक्त मापों को विकसित करते समय यह ध्यान रखा गया की दोनों गुणात्मक और मात्रात्मक आँकड़ों के दूरी मैट्रिसेस का गठबंधन किया गया है। गुणात्मक और मात्रात्मक दूरी मैट्रिसेस के संयोजन से पहले, इन मैट्रिसेस के प्रत्येक तत्वों को 0 और 1 बीच लेकर एक समान पैमाने में स्थापित किया गया है। कभी-कभी गुणात्मक और मात्रात्मक आँकड़ों के अलग-अलग विश्लेषण से प्राप्त क्लस्टरिंग पैटर्न के बीच पत्राचार की डिग्री के आधार पर कोर सेट पहचान की जाती है। इसके अलावा, आँकड़ों का वर्गीकरण समूह में इस्तेमाल क्लस्टरिंग कार्यप्रणाली पर निर्भर करता है। इसलिए, यह जरूरी है कि विश्लेषण के प्रारम्भिक चरण में गुणात्मक और मात्रात्मक आँकड़ा गठबंधन करने के महत्वपूर्ण है। इसके अलावा, समय के साथ कई क्लस्टरिंग एल्गोरिथम का उपयोग एक उपयुक्त आवंटन रणनीति को लागू करके एक कोर सेट उत्पन्न किया जाता है। Odong एट एल (2011) में मॉडल आधारित क्लस्टरिंग विधियों के ऊपर पारंपरिक क्लस्टरिंग विधियों के उपयोग की वकालत विशेष रूप से सरल अनुक्रम दोहराने वाले मार्कर आँकड़ा का उपयोग एक कोर सेट विकसित करने के लिए किया है। मॉडल-आधारित क्लस्टरिंग दृष्टिकोण के ऊपर हालांकि, एक कोर सेट के विकास के लिए प्ररूपी और एस एन पी जीनोटाइपिंग आँकड़ा पर आधारित एक साथ उपयुक्त प्रक्रिया अपना कर, एक उपयुक्त संयुक्त

दूरी माप लेकर, क्लस्टरिंग के तरीके, समूहों की संख्या, आवंटन पद्धति और मूल्यांकन रणनीति को शायद ही कभी प्रयुक्त किया है। इसलिए, वर्तमान अध्ययन में एक कोर सेट की पहचान का उपयुक्त अंतिम समाधान खोजने के लिए किया गया है।

वर्तमान अध्ययन में, आम सहमति और विलय-आम सहमति क्लस्टरिंग समूहों के परिणाम का अधिकतम संख्या की पहचान करने के लिए इस्तेमाल किया गया। मात्रात्मक और गुणात्मक आँकड़ों के आधार पर समूहों की अधिकतम संख्या (यानि तीन और चार) के बीच एक असहमति थी हालांकि, मिश्रण आँकड़ों के लिए समूहों की संख्या 3 ही पाया गया। सबसे अच्छा फिट वाले क्लस्टरिंग एल्गोरिथम के चयन के संबंध में संयुक्त दूरी मापों  $A_1B_1$ ,  $A_3B_1$  और  $A_1B_2$  के लिए उच्चतम औसत क्लस्टर दृढ़ता मूल्यों के आधारित क्लस्टरिंग k-माध्य (k-means) एल्गोरिथम पाया गया। इसके विपरीत, पूर्ण लिंकेज क्लस्टरिंग एल्गोरिथम में संयुक्त दूरी मापों  $A_2B_1$ ,  $A_2B_2$  और  $A_2B_2$  (तालिका 1) के लिए उच्चतम औसत क्लस्टर दृढ़ता मूल्य दी है।

Odong एट एल (2013) ने एक कोर सेट के मूल्यांकन के लिए विभिन्न परिस्थिति में अलग-अलग मानदंड की समीक्षा की। अक्सर, यह व्यक्तिगत सूचकांक पर आधारित एक कोर संग्रह की विविधता पर टिप्पणी करने के लिए मुश्किल हो सकता है। इस तरह के भ्रम की स्थिति से बचने के लिए एक संयुक्त माप



मान्य निष्कर्ष निकालने में मदद मिल सकती है। इस प्रकार MD, VD, VR, CR और APICD को शामिल कर कोर सेट में विविधता का मूल्यांकन करने के लिए एक CEI प्रयोग किया जाता है। सभी संयुक्त दूरी मापों के तहत CEI मूल्यों के बीच की तुलना लॉग-आनुपातिक और RASE तरीकों पर आनुपातिक आवंटन पद्धति की श्रेष्ठता को इंगित करता है। निम्नतम स्तर पर जर्मप्लाज़म वर्गीकृत कर यानि कोर सेट आकार के बराबर समूहों की संख्या लेकर परिग्रहण के चयन के लिए RASE विधि के उपयोग कर आनुपातिक आवंटन और लॉग-आनुपातिक आवंटन के तरीकों का सभी संयुक्त दूरी के मापों के लिए कोई लाभ प्रदान नहीं करता है केवल कुछ अपवादों को छोड़कर (तालिका 2)। इसके अलावा, कोर सेट के आकार के बराबर समूहों की संख्या के साथ जर्मप्लाज़म क्लस्टरिंग करना, क्लस्टरिंग कार्यप्रणाली में प्राकृतिक समूह का उल्लंघन होता है। यह जर्मप्लाज़म के चयन में पक्षपात बढ़ाता है। क्योंकि एक बड़ा क्लस्टर की तुलना में एक छोटे क्लस्टर से एक जर्मप्लाज़म चयन की संभावना अधिक है। इसलिए, पद्धति पर आधारित परिग्रहण के चयन के बाद आनुपातिक आवंटन सुदृढ़ सांख्यिकीय तकनीक पर आधारित समूहों की अधिकतम संख्या की पहचान एक विविध कोर सेट के विकास के लिए उचित है। इसके अलावा, तालिका 3, भले ही बहुमत मामलों में, आनुपातिक आवंटन पद्धति के तहत, संयुक्त दूरी को मापने के व्यक्तिगत उपायों पर बेहतर प्रदर्शन किया है। कुछ ऐसे मामले हैं जैसे  $A_2B_2$  बनाम  $B_2$ ,  $A_2B_2$  बनाम  $A_2$  और  $A_3B_2$  बनाम  $B_2$  जो संयुक्त दूरी मापों में व्यक्तिगत दूरी उपायों पर बेहतर प्रदर्शन नहीं किया था। इसी प्रकार, लॉग-आनुपातिक आवंटन पद्धति के तहत संयुक्त दूरी मापों  $A_2B_1$  और  $A_2B_2$  व्यक्तिगत उपाय  $A_2$  को मात नहीं कर पाये। इसलिए, यह निष्कर्ष निकाला है कि संयुक्त दूरी मापों से हमेशा अलग-अलग दूरी मापों की तुलना में बेहतर प्रदर्शन करेंगे, ऐसा नहीं किया जा सकता। हालांकि वर्तमान अध्ययन में, संयुक्त दूरी मापों  $A_1B_2$  ने बाकी के संयुक्त मापों में सबसे अच्छा प्रदर्शन किया। इसके अलावा, प्रस्तावित की तुलना में दृष्टिकोण की दक्षता PowerCore

के साथ एक तुलना द्वारा स्थापित की है। (RDA, Genebank सूचना केन्द्र [http://www.genebank.go.kr/eng/PowerCore/PowerCore\\_software.zip](http://www.genebank.go.kr/eng/PowerCore/PowerCore_software.zip))। इस मिश्रित आँकड़ों के लिए प्रस्तावित दृष्टिकोण का उपयोग कर के लाभ इंगित किया जा सकता है। इसलिए, परिग्रहण नमूने के लिए आनुपातिक आवंटन पद्धति के साथ k-माध्य क्लस्टरिंग एल्गोरिथम साधन का उपयोग कर संयुक्त माप  $A_1B_2$  चावल जर्मप्लाज़म का एक संग्रह से एक कोर सेट की पहचान के लिए पसंद किया जा सकता है।

### सारांश

जर्मप्लाज़म की एक बड़ी अत्याधिक संख्या में से कम से कम दो प्रतियों का परिग्रहण और अधिकतम विविधता के साथ एक विविध प्रतिनिधि कोर की फसल विकास सुधार कार्यक्रम में शामिल करना प्रजनन कर्ता के लिए बेहद जरूरी है। एक कोर सेट पहचान करने वाली मौजूदा तरीके में ज्यादातर या तो गुणात्मक या मात्रात्मक आँकड़ों पर आधारित है। इस अध्ययन में, गुणात्मक (एकक न्यकिलयोटाइड बहुरूपता जीनोटाइपिंग) तथा मात्रात्मक आँकड़ों के मिश्रित प्रतिक्रिया के आधार पर जर्मप्लाज़म में से एक कोर सेट की पहचान के लिए एक दृष्टिकोण का प्रस्ताव रखा है। इस उद्देश्य के लिए छह प्रस्ताविक अलग-अलग संयुक्त दूरी मापों, तीन मात्रात्मक आँकड़ों के लिए और दो गुणात्मक आँकड़ों के लिए मापों का मूल्यांकन किया गया। संयुक्त दूरी मैट्रीसेस सजातीय समूहों में जर्मप्लाज़म की जनसंख्या को वर्गीकृत करने के लिए सात अलग-अलग क्लस्टरिंग प्रक्रियाओं के लिए सामग्री के रूप में इस्तेमाल किया गया। इसके पश्चात, विभिन्न संयुक्त दूरी उपायों का उपयोग तथा सभी क्लस्टरिंग तरीकों के आधार, एक आम सहमति के आधार पर अधिकतम संख्या वाले समूहों की पहचान की गई। प्रत्येक क्लस्टरिंग कार्यप्रणाली के तहत समूहों के सभी की पहचान की अधिकतम संख्या पार औसत क्लस्टर मजबूती मूल्यों की गणना की गई। कुल मिलाकर, तीन अलग आवंटन के तरीकों उच्चतम औसत क्लस्टर मजबूती मूल्य एक कोर सेट तैयार करने के लिए इस्तेमाल किया जा रहा

है। जिसके साथ प्रत्येक क्लस्टरिंग कार्यप्रणाली के तहत पहचान समूहों से चयन किया गया था कि विलय के नमूने के लिए लागू किया गया। इसके अलावा, एक सूचकांक कोर सेट में विविधता के मूल्यांकन के लिए प्रस्तावित किया गया था। परिणामों से पता चलता है कि संयुक्त दूरी को मापने के लिए  $A_1, B_2$  मात्रात्मक आँकड़ों के लिए सीमा मानकीकृत पूर्ण अंतर के औसत के आधार पर इस दूरी आनुपातिक आबंटन पद्धति के साथ एल्गोरिथम क्लस्टरिंग k-means साधन का उपयोग करके ही पहचान की गई है कि तीन समूहों जिसमें से चावल-जर्मप्लाज़म का एक संग्रह एक कोर सेट की पहचान के लिए उपयुक्त था।

### संदर्भ

- अग्रामा एस ए, यान डब्ल्यू जी, ली एफ, फजेलस्ट्राम आर, चेन एम एच, जिया एम एवं एमसीकलुंग ए (2009): “आनुवांशिक आकलन का एक छोटा कोर सबसेट के यू एस डी ए चावल जीनबैंक”, फसल विज्ञान 49, 1336–1346 ।
- क्रोसा जे एवं फ्रेंको जे (2004): “जीनोटाइप वर्गीकृत करने के लिए सांख्यिकीय तरीका”, यूफीटीका 137, 19–37 ।
- डोरिंग सी, बोरगेल्ट सी एवं क्रुसे आर (2004): “मात्रात्मक और गुणात्मक डेटा के फ़ज़ी क्लस्टरिंग”, एन ए एफ आई पी एस की कार्यावाही में, Banff, अलबर्टा कनाडा, 84–89 ।
- एवरीट बी सी (1979): “अनसुलझे समस्याओं में क्लस्टर विश्लेषण”, बायोमीट्रिक्स 35, 168–181 ।
- फ्रेंकल ओ एच एवं ब्राउन ए एस डी (1984): “संयंत्र आनुवांशिक संसाधन आज: एक महत्वपूर्ण मूल्यांकन में: होल्डन जे एच डब्ल्यू और विलियम्स जेटी (एड्स) फसल आनुवांशिक संसाधन”, संरक्षण और मूल्यांकन, लंदन: जॉर्ज एलेन और अनविन लिमिटेड, 249–257 ।
- गंगोपाध्याय के के, महाजन आर के, कुमारजी, यादव एस के, मीना बी एल, पांडे सी, बिष्ट आई एस, मिश्रा एस के, शिवराज एन, गंभीर आर, शर्मा एस के एवं ढिल्लों बी एस (2010): “बैंगन में एक कोर सेट का विकास (सोलेनम मेलोंगेना एल)”, फसल विज्ञान 50: 755–762 ।
- जिबर्ट के एवं कोर्टस यू (1997): “क्लस्टरिंग तरीकों में मात्रात्मक और गुणात्मक चर भार”, मैथवेयर और सॉफ्ट कम्प्यूटिंग 4: 251–266 ।
- गोवर जे सी (1971): “एक सामान्य समानता के गुणांक और अपने गुण में से कुछ”, बायोमीट्रिक्स 27 : 857–874 ।
- हू जे, झू जे एवं जू एच एम (2000): “फसलों की जीनोटापिक मूल्यांकन पर आधारित तीन नमूने रणनीतियों के साथ चरणबद्ध क्लस्टरिंग द्वारा मुख्य संग्रह के निर्माण की विधियां”, सैद्धांतिक और व्यवहारिक आनुवांशिकी 101: 264–268 ।
- किम के डब्ल्यू, चुंग एच के, चो जी टी, मा के एच, चन्द्रबालन डी, ग्वाग जे जी, किम टी एस, चो ई जी एवं पार्क वाई जे (2007): “PowerCore: कोर सेट स्थापित करने के लिए एक अनुमानी खोज के साथ उन्नत M रणनीति को लागू करने के लिए एक कार्यक्रम”, जैव सूचना विज्ञान 23: 515–526 ।
- मोंटि एस मायो पी, मेसिरोव जे एवं गोलूड टी (2003): “आम सहमति क्लस्टरिंग कक्षा की खोज और जीन अभिव्यक्ति माइक्रोएरे डेटा के दृश्य के लिए एक रिसैम्पलिंग आधारित पद्धति”, मशीन लर्निंग 52: 91–118 ।
- म्युनैक बी, सिलेयुक के ए, सिमनसेन के एल, बीएविस डब्ल्यू डी एवं डुरिज आर डब्ल्यू (2005): “जीन अभिव्यक्ति क्लस्टरिंग को विश्वास जोड़ने”, आनुवांशिकी 170: 2003–2011 ।
- ओडोंग टी एल, वैन एच जे, जेन्सन जे, वैन एच टी जे एल एवं वैन ई एफ ए (2011): “जर्मप्लाज़म संग्रह की आनुवांशिकी संरचना के निर्धारण: परंपरागत पदानुक्रमित क्लस्टरिंग तरीकों आणविक मार्कर डेटा के लिए उपयुक्त है?”, सैद्धांतिक और व्यवहारिक आनुवांशिकी 123: 195–205 ।
- ओडोंग टी एल, जेन्सन जे, वैन ई एफ ए एवं वैन

- एच टी जे एल (2013): “आनुवांशिकी संसाधनों की समीक्षा, चर्चा और व्याख्या के प्रभावी उपयोग के लिए कोर संग्रह की गुणवत्ता”, सैद्धांतिक और व्यवहारिक आनुवांशिकी 126: 289–305 ।
15. सरकार आर के, राव ए आर, वाही एस डी एवं भट्ट के वी (2011): “गुणात्मक और मात्रात्मक डेटा का मिश्रण के लिए क्लस्टरिंग प्रक्रियाओं की एक तुलनात्मक निष्पादन-काले चने का एक अनुप्रयोग”, संयंत्र आनुवांशिक संसाधन: लक्षण वर्णन और उपयोग 9: 523–527 ।
16. शर्मा आर, राव वी पी, उपाध्याय एच डी, रेड्डी वी जी एवं ठाकुर आर पी (2010): “चारा जर्मप्लाज़म का एक छोटा-कोर संग्रह में अनाज ढालना और कोमल फफूंदी के लिए प्रतिरोध”, पौधा रोग 94: 439–444 ।
17. सिम्पसन टी आई (2010): “clusterCons: फिर से नमूना क्लस्टरिंग प्रयोगों से कई एल्गोरिदम और पैरामीटर का उपयोग करने के विकल्प के साथ आम सहमति क्लस्टरिंग परिणाम की गणना”, भार पैकेज संस्करण 3.0.2 <http://cran.r-project.org/src/contrib/Archive/clusterCons/> ।
18. सिम्पसन टी आई, आर्मस्ट्रांग जे डी एवं जरमैन ए पी (2010): “मर्ज किए गए आम सहमति क्लस्टरिंग का आकलन और माइक्रो एरे डेटा के साथ वर्ग खोज में सुधार करने के लिए”, बी एम सी जैव सूचना विज्ञान 11: 590 ।
19. स्टुडेनकि एम एवं डीबेस्की के (2012): “ccChooser: एक कोर संग्रह के विकास, आर पैकेज संस्करण 3.0.2”, <http://cran.r-project.org/package=ccChooser>. ।
20. वैन एच टी एवं टीएच जे एल (1999): “मुख्य चयनकर्ता, जर्मप्लाज़म परिग्रहण के प्रतिनिधि चयन उत्पन्न करने के लिए एक प्रणाली”, संयंत्र आनुवांशिक संसाधन न्यूज़लेटर 118 : 64–67 ।
21. वैन एच टी, ब्राउन ए एच डी, स्पाईलेन सीएवं होजकिन टी (2000): “संयंत्र आनुवांशिक संसाधनों की कोर संग्रह”, आई पी जी आर आई तकनीकी बुलेटिन न. 3, अंतरराष्ट्रीय संयंत्र आनुवांशिक संसाधन संस्थान, रोम, इटली ।
22. वैन डब्ल्यू, फ्रैंको जे, चावेज-टोवार वी एच, यान जे एवं टाबा एस (2012): “मक्का सुधार में आगे उपयोग के लिए एक उष्णकटिबंधीय मक्का दौड़ Tuxpeno की कोर सेट के लिए आनुवांशिक लक्षण वर्णन”, PLoS ONE 7: e32626 ।
23. यान डब्ल्यू, रूगेर जे एन, ब्रायंट आर जे, बोकेलमैन एच ई, फिल्लिस्ट्रोम आर जी, थॉमस एस सी, ताई एच एवं मैक-कल्युंग ए एम (2007): “यू एस डी ए चावल जर्मप्लाज़म संग्रह की एक कोर सबसेट के विकास और मूल्यांकन”, फसल विज्ञान 47: 869–876 ।
24. यू जे जेड, कोहेल आर जे, फेंग डी डी, चो जे, वैन डीएन्ज ए, यूल्लोआ एम, हॉफमैन एस एम, पेप्पर ए ई, स्टेली डी एम, जेनकींस जे एन, साहा एस, कुमपाटला एस पी, शाह एम आर, हयूजी डब्ल्यू वी और पर्सी, आर जी (2012): “एक उच्च-घनत्व सरल अनुक्रम दोहराने और टेट्राप्लोइड कपास जीनोम की एकक-न्यूक्लियोटाइड बहुरूपता का आनुवांशिक मानचित्र”, जीन जीनोम आनुवांशिकी 2: 43–58 ।

□

## अर्द्ध-सतत डाटा हेतु लघु क्षेत्र आकलन

हुकुम चन्द्र एवं विजय बिन्दल

व्यवसाय, कृषि, पर्यावरण, पारिस्थितिकी (ecological) एवं महामरी विज्ञान (epidemiological) सर्वेक्षणों से संबंधित कई चर स्वाभाविक रूप से अर्द्ध-सतत (semi continuous) होते हैं अर्थात् या तो वे एक नियत मान (आमतौर पर 0, शून्य) लेते हैं या उनका धनात्मक वास्तविक रेखा पर सतत, अक्सर वैषम्यिक (skewed) बंटन (distribution) होता है। यह लेख बहुधा (frequently) व्यावहारिक रूप में आने वाले एक विशेष प्रकार के शून्यों के मिश्रण एवं सतत वैषम्यिक बंटित धनात्मक मानों के अर्द्ध-सतत चर पर केन्द्रित है। एक अर्द्ध सतत चर सेंसर करने के लिए छोड़ दिए गए या छिन्न (truncated) कर दिए गए चरों से पूर्णतया भिन्न होते हैं क्योंकि शून्य वैध स्व-प्रतिनिधित्व डाटा मान होते हैं ना कि ऋणात्मक या लुप्त प्रतिक्रियाओं के लिए प्रतिपत्र (proxy)।

एक अर्द्ध सतत डाटा का दो प्रक्रियाओं के परिणाम के रूप में अवलोकन करना स्वाभाविक है, एक यह निर्धारण करना कि क्या अनुक्रिया (response) शून्य है और दूसरा यदि यह गैर-शून्य है तो वास्तविक स्तर का निर्धारण करना। उदाहरणार्थ, भारत के राष्ट्रीय प्रतिदर्श सर्वेक्षण कार्यालय के 59वें राऊंड सर्वेक्षण डाटा में बकाया ऋण की राशि।

अर्द्ध सतत चरों के लिए एक रैखिक मॉडल उपयुक्त नहीं है। इन चरों के लिए रैखिक मिश्रित मॉडलों पर आधारित लघु क्षेत्र आकलन के लिए सामान्यतः प्रयोग में लाई गई पद्धतियां उदाहरणार्थ अनुभवजन्य श्रेष्ठतम रैखिक अनभिन्नत पूर्वानुमानक (predictor)

(ई बी एल यू पी ) अक्षम हो सकते हैं (राव 2003)। चन्द्र एवं चैम्बर्स (2011 a) तथा बर्ग एवं चन्द्र (2012) ने वैषम्यिक (skewed) चरों के लिए लघु क्षेत्र आकलन की पद्धतियों का अन्वेषण किया जिनमें लघुगणक (लॉग) रूपान्तरण का प्रयोग करते हुए रैखिक मिश्रित मॉडल का अनुसरण करने वाले चरों को केन्द्रित किया गया। चन्द्र एवं चैम्बर्स (2011 a) ने वैषम्यिक चरों के लघु क्षेत्र आकलन के लिए दो पूर्वानुमानकों का वर्णन किया। पहला पूर्वानुमानक, मॉडल आधारित प्रत्यक्ष आकलन (एम बी डी ई ) को प्रतिदर्श इकाइयों के भारत योग के रूप में दर्शाया गया, जहां भारों को समष्टि माध्य के न्यूनतम माध्य वर्ग त्रुटि रैखिक पूर्वानुमानक द्वारा परिभाषित किया गया यदि लॉग स्तर पर रैखिक मिश्रित मॉडल के प्राचल ज्ञात हैं। दूसरा पूर्वानुमानक कार्लबर्ग (2000) की पूर्वानुमान आधारित पद्धति अर्थात् लॉग स्तर पर एक रैखिक मिश्रित मॉडल के अन्तर्गत अनुभवजन्य पूर्वानुमान पर आधारित है। यह अनुभवजन्य पूर्वानुमान, रैखिक मिश्रित मॉडल के अन्तर्गत सिंथेटिक आकलन के अनुरूप है। एम बी डी ई एक प्रत्यक्ष आकलन है और क्षेत्रों के बीच विषमांगता की उपस्थिति में अनभिन्नत है लेकिन प्रतिदर्श आकार छोटा होने की स्थिति में अस्थिर आकल दे सकता है। दूसरी ओर, सिंथेटिक प्रकार का अनुभवजन्य पूर्वानुमानक सिर्फ सहचरों के माध्यम से क्षेत्रीय विविधता दर्शाता है इसलिए क्षेत्रों के बीच विषमांगता होने की स्थिति में अनभिन्नत आकल दे सकता है। बर्ग एवं चन्द्र (2012) ने अनुभवजन्य सर्वश्रेष्ठ पूर्वानुमानक को अनभिन्नत पूर्वानुमानक वर्ग

में न्यूनतम माध्य वर्ग त्रुटि की उपस्थिति के रूप में वर्णित किया है। विशेषतया उन्होंने क्षेत्रीय विविधता के बीच मुद्दों को संबोधित किया है और लघु क्षेत्र प्रतिदर्श आकारों के कारण सभी क्षेत्रों से प्राप्त सूचना का प्रयोग करने की आवश्यकता को वर्णित किया है। हालांकि, लघु क्षेत्र आकलन की ये पद्धतियां वैषम्यिक चरों के लिए उपयुक्त हैं लेकिन इनका अनुप्रयोग सिर्फ और सिर्फ धनात्मक चरों तक ही सीमित है। इसलिए लघु क्षेत्र आकलन की इन पद्धतियों को अर्द्ध सतत चर के लिए अनुप्रयुक्त नहीं किया जा सकता।

सर्वेक्षण डाटा में अतिरिक्त शून्यों की उपस्थिति एक जानी पहचानी समस्या है और इसके समाधान के लिए बहुत सी पद्धतियों का सुझाव दिया गया है। हालांकि सर्वेक्षण डाटा का प्रयोग करके लघु क्षेत्र आकलन के बारे में बहुत कम ज्ञात है, यहां तक कि बृहत समग्र प्रतिदर्श की तुलना में लघु क्षेत्र में अतिरिक्त शून्यों की उपस्थिति स्पष्ट रूप में बहुत अधिक प्रभावशाली है। मिश्रित मॉडल के रूप में सदर्भित (फ्लेचर इत्यादि 2005) एक द्वि-खण्ड यादृच्छिक प्रभाव मॉडल (ऑलसेन एवं शैफर 2001) का शून्य-इंफ्लेटिड चरों सहित लघु क्षेत्र आकलन के लिए व्यापक रूप से प्रयोग किया जाता है। उदाहरणार्थ फौफरमैन इत्यादि (2008) तथा चन्द्र एवं सूद (2012) देखें। इस लेख में अर्द्ध-सतत चरों के लिए लघु क्षेत्र आकलन विधि का वर्णन द्वि-खण्ड यादृच्छिक प्रभाव मॉडल के अन्तर्गत किया गया है।

### रैखिक रूपान्तरण के अन्तर्गत लघु क्षेत्र आकलन

माना सीमित समष्टि  $U$  का आकार  $N$  है जिसमें  $U_i$ ; ( $i=1, \dots, D$ ),  $D$  गैर अतिव्यापी (non-overlapping) डोमेन हैं तथा इस समष्टि में से  $n$  आकार का एक प्रतिदर्श लिया गया है। इन डोमेन को हम लघु क्षेत्रों या सिर्फ क्षेत्रों से संदर्भित करेंगे। माना लघु क्षेत्र  $i$  में  $N_i$  ज्ञात समष्टि इकाइयां हैं जिसमें से  $n_i$  ( $i=1, \dots, D$ ) प्रतिदर्श लिए गए हैं। समष्टि में कुल  $N = \sum_{i=1}^D N_i$  इकाइयां हैं तथा कुल प्रतिदर्श आकार  $n = \sum_{i=1}^D n_i$  है। प्रतिदर्श में  $s$  इकाइयां हैं तथा  $i$  वे लघु क्षेत्र से  $s_i$  उपसमुच्चय लिए गए हैं अर्थात्  $|s_i| = n_i$ , लघु क्षेत्र  $i$

की इकाइयों के लिए  $j \in i$  तथा प्रतिदर्श  $s$  की इकाइयों के लिए  $j \in s$  का प्रयोग किया गया है। इसीप्रकार,  $i$  वे लघु क्षेत्र में  $j$  इकाइयों के समुच्चय को  $r_i$  से निरूपित किया गया है जो प्रतिदर्श में नहीं हैं जबकी  $|r_i| = N_i - n_i$  एवं  $U_i = s_i \cup r_i$  माना  $y_{ij}$ ,  $i$  वे क्षेत्र में  $j$  वे इकाई के लिए अध्ययनाधीन चर का  $Y_L$  का मान है तथा  $x_{ij}$ ,  $y_{ij}$  से संबंधित  $i$  वे क्षेत्र में सहायक चरों के मान का  $(m-1)$  सदिश है। अतः  $Y$  का लघु क्षेत्र माध्य

$$m_i = N_i^{-1} \sum_{j=1}^{N_i} y_{ij}$$

अध्ययनाधीन मात्रा (quantity) है।

हम उस स्थिति पर विचार करेंगे जहां अध्ययनाधीन चर, लॉग रूपान्तरण पर रैखिक मिश्रित मॉडल का अनुसरण करता है। अध्ययनाधीन चर  $y_{ij}$  के लिए एक लॉग स्तर रैखिक मिश्रित मॉडल निम्न होगा

$$\log(y_{ij}) = l_{ij} = Z_{ij}^T \beta + u_i + e_{ij} \quad (1)$$

जहां  $Z_{ij} = (1, \log(x_{ij}))$  एक उचित रूपान्तरित सहचरों का  $m \times 1$  सदिश है,  $\beta$  नियत प्रभावों का  $m \times 1$  सदिश है,  $u_i$ ,  $i$  वे क्षेत्र से संबंधित क्षेत्र-विशेष यादृच्छिक प्रभाव है, तथा  $e_{ij}$  लघु क्षेत्र  $i$  में  $j$  वे इकाई के लिए वैयक्तिक (individual) स्तर यादृच्छिक त्रुटियां हैं।

$$(u_i, e_{ij}) \sim N(0, \text{diag}(\sigma_u^2, \sigma_e^2))$$

माना  $\{(y_{ij}, z_{ij}); i=1, \dots, D; j \in s_i\}$  प्रेक्षण ज्ञात हैं। माना  $i$  क्षेत्र में  $U_i$  ( $i=1, \dots, D$ ) इकाइयों की समष्टि के लिए  $Z_{ij}$  ज्ञात है एवं  $\{y_{ij}; i=1, \dots, D; j \in s_i\} \cup \{z_{ij}; i=1, \dots, D; j \in s_i\}$  आंकड़े उपलब्ध हैं।

माना  $\emptyset = (\beta, \sigma_u^2, \sigma_e^2)^T$  मॉडल प्राचलों का सदिश है और  $\hat{\emptyset} = (\hat{\beta}, \hat{\sigma}_u^2, \hat{\sigma}_e^2)^T$ ,  $\emptyset$  का अधिकतम संभावित (एम एल) या अनुबंधित अधिकतम संभावित (आर ई एम एल) आकलन है। विशेषतया,  $\sigma^2 = (\sigma_u^2, \sigma_e^2)^T$  मॉडल का प्रसरण घटक है तथा  $\hat{\sigma}^2 = (\hat{\sigma}_u^2, \hat{\sigma}_e^2)^T$ ,  $\sigma^2$  का आकलक है। माना कि सहायक चरों के ज्ञात मानों के लिए गैर-सूचनाप्रद प्रतिदर्श हैं इसलिए प्रतिदर्श

आंकड़े, समष्टि मॉडल का अनुसरण करते हैं तथा माध्य एवं प्रसरण, मॉडल के अन्तर्गत हैं।

दिए गए प्रतिदर्श आंकड़ों के लिए, मॉडल (1) के अज्ञात प्राचलों (क्षेत्रीय प्रभाव सम्मिलित करके) का आकलन किया जा सकता है। अतः लॉग स्तर पूर्वानुमान  $\hat{y}_{ij} = \mathbf{z}_{ij}^T \hat{\beta} + \hat{u}_i$  के रूप में परिभाषित किये जा सकते हैं जहाँ  $\hat{\beta}$ ,  $\beta$  का आकल है,  $\hat{u}_i = \hat{\gamma}_i (\bar{y}_{is} - \bar{\mathbf{z}}_{is}^T \hat{\beta})$  यादृच्छिक क्षेत्रीय प्रभाव का पूर्वानुमानक है जहाँ  $\hat{\gamma}_i = \hat{\sigma}_u^2 (\hat{\sigma}_u^2 + n_i^{-1} \hat{\sigma}_e^2)^{-1}$  का आकलित संकुचित (shrinkage) प्रभाव है।

यहाँ  $\bar{y}_{is} = n_i^{-1} \sum_{j \in s_i} \log(y_{ij})$  एवं  $\bar{\mathbf{z}}_{is} = n_i^{-1} \sum_{j \in s_i} \mathbf{z}_{ij}$   $i$ वें क्षेत्र में सपर एवं  $\mathbf{z}_{ij}$  के क्रमशः प्रतिदर्श माध्य हैं। मॉडल (1) के अन्तर्गत कालबर्ग (2000) के समरूप एक पूर्वानुमानक आधारित पद्धति का प्रयोग करके चन्द्र एवं चैम्बर्स (2011a) में क्षेत्रीय माध्य  $m_i$  के लिए एक कृत्रिम (synthetic) प्रकार पूर्वानुमानक का वर्णन किया गया है। दूसरी ओर, चन्द्र एवं चैम्बर्स (2011a) ने एक मॉडल-आधारित प्रत्यक्ष आकलक  $\sum_{j \in s_i} w_{ij} y_{ij}$  जहाँ  $w_{ij}$  भार का आकलन है जो समष्टि माध्य का श्रेष्ठतम रैखिक अनभिन्नत आकलन देता है यदि मॉडल (1) के प्राचल ज्ञात हैं, को प्रस्तावित किया।

मॉडल (1) के अन्तर्गत, बर्ग एवं चन्द्र (2012) का अनुसरण करते हुए  $m_i$  के लिए न्यूनतम माध्य वर्ग त्रुटि (एम एम एस ई) पूर्वानुमानक निम्न होगा

$$\hat{m}_i^{EBP} = N_i^{-1} \left\{ \sum_{s_i} y_{ij} + \sum_{n_i} \hat{y}_{ij}^{EBP} \right\} \quad (2)$$

यहाँ  $\hat{y}_{ij}^{EBP} = \exp \left\{ \mathbf{z}_{ij}^T \hat{\beta} + \hat{\gamma}_i (\bar{y}_{is} - \bar{\mathbf{z}}_{is}^T \hat{\beta}) + 0.5 \hat{\sigma}_e^2 (1 + n_i^{-1} \hat{\gamma}_i) \right\}$

ज्ञात रहे कि (2) पूर्वानुमानक एक लघु क्षेत्र के अन्तर्गत परस्पर इकाई सहसंबंधों की अनुमति देता है। अतः यह एक अनुभवजन्य सर्वश्रेष्ठ पूर्वानुमानक (ई बी पी) है। फलस्वरूप, लॉग-प्रसामान्यता (log-normality) के अन्तर्गत एम एम एस ई पूर्वानुमानक के कृत्रिम (synthetic) प्रकार के पूर्वानुमानक पर हावी (dominate) होने की संभावना है। बर्ग एवं चन्द्र (2012) में दिए गए विवरण के अन्तर्गत एम एम एस ई पूर्वानुमानक (2) का संशोधित संस्करण (vession) निम्न है

$$\hat{m}_i^{EBP-BC} = N_i^{-1} \left\{ \sum_{s_i} y_{ij} + \sum_{n_i} \hat{y}_{ij}^{EBP-BC} \right\} \quad (3)$$

जहाँ

$$\hat{y}_{ij}^{EBP-BC} = \left( \hat{c}_{ij}^{EBP} \right)^{-1} \hat{y}_{ij}^{EBP}, \quad c_{ij}^{EBP} = \exp \left\{ \frac{1}{2} (\mathbf{a}_{ij} + \hat{c}_{i1} + \hat{c}_{i2} + 2\hat{c}_{i3}) \right\}$$

$$\mathbf{a}_{ij} = (\mathbf{z}_{ij}^T - \hat{\gamma}_i \bar{\mathbf{z}}_{is}^T)^T \hat{V}(\hat{\beta}) (\mathbf{z}_{ij}^T - \hat{\gamma}_i \bar{\mathbf{z}}_{is}^T)$$

$$\hat{d}_i = (\bar{y}_{is} - \bar{\mathbf{z}}_{is}^T \hat{\beta})$$

$$\hat{c}_{i1} = \left[ \left\{ 0.5 + 0.5 \frac{\hat{\gamma}_i^2}{n_i} - \frac{\hat{\gamma}_i^2 \hat{d}_i}{n_i \hat{\sigma}_u^2} \right\}^2 - \left\{ \frac{\hat{\gamma}_i^3}{n_i^2 \hat{\sigma}_u^2} - \frac{2\hat{\gamma}_i^2 \hat{d}_i}{n_i^2 \hat{\sigma}_u^4} \right\} \right] \hat{V}(\hat{\sigma}_e^2)$$

$$\hat{c}_{i2} = \left[ \left\{ \frac{(1-\hat{\gamma}_i)^2}{2} + \frac{\hat{\gamma}_i(1-\hat{\gamma}_i)\hat{d}_i}{\hat{\sigma}_u^2} \right\}^2 - \left\{ \frac{\hat{\gamma}_i(1-\hat{\gamma}_i)^2}{\hat{\sigma}_u^2} + \frac{2\hat{\gamma}_i^2(1-\hat{\gamma}_i)\hat{d}_i}{\hat{\sigma}_u^4} \right\} \right] \hat{V}(\hat{\sigma}_u^2)$$

$$\hat{c}_{i3} = \left[ \left\{ \frac{1}{2} + \frac{\hat{\gamma}_i^2}{2n_i} - \frac{\hat{\gamma}_i^2 \hat{d}_i}{n_i \hat{\sigma}_u^2} \right\} \left\{ \frac{(1-\hat{\gamma}_i)^2}{2} + \frac{\hat{\gamma}_i(1-\hat{\gamma}_i)\hat{d}_i}{\hat{\sigma}_u^2} \right\} + \left\{ \frac{\hat{\gamma}_i^2(1-\hat{\gamma}_i)}{n_i \hat{\sigma}_u^2} - \frac{\hat{\gamma}_i^2(1-2\hat{\gamma}_i)\hat{d}_i}{n_i \hat{\sigma}_u^4} \right\} \right] \hat{Cov}(\hat{\sigma}_e^2, \hat{\sigma}_u^2)$$

### मिश्रण मॉडल के अन्तर्गत लघु क्षेत्र आकलन

माना अनुक्रिया चर लपर एक अर्द्ध सतत चर है। विशेष रूप से, चर लपर या तो शून्य है या कुछ धनात्मक मान लेता है जिसका वितरण (distribution) वैषम्यिक (skewed) है। हमने इसप्रकार के चरों का मॉडल प्राप्त करने के लिए द्वि-खण्ड यादृच्छिक प्रभावों के मॉडल (मिश्रण मॉडल के रूप में संदर्भित) पर आधारित एक पद्धति का वर्णन किया है। इसप्रकार के मॉडलों के दो घटक होते हैं, पहला घटक अतिरिक्त शून्य मानों को समायोजित (accomodate) करता है तथा दूसरा घटक अनुक्रिया चरों के वैषम्यिक गैर-शून्य मानों का मॉडल तैयार करता है। ओल्सन एवं शैफर (2001), फ़ैफरमैन इत्यादि (2008), चन्द्र एवं चैम्बर्स (2011b) तथा चन्द्र एवं सूद (2012) द्वारा दिए गए दृष्टिकोणों का अनुसरण करते हुए हम एक अर्द्ध-सतत चर  $y_{ij} = \delta_{ij} \tilde{y}_{ij}$  को दो घटकों को गुणनफल के रूप में व्यक्त करते हैं। यहाँ, पहले घटक  $\tilde{y}_{ij}$  को  $y_{ij}$  के लॉग-रैखिक घटक के रूप में संदर्भित किया गया है तथा माना गया है कि यह लॉग-रूपान्तरण स्तर पर एक रैखिक मिश्रित मॉडल उदाहरणार्थ (1)

का अनुसरण करता है। दूसरा घटक  $\delta_{ij} = I(y_{ij} > 0)$  एक द्विआधारी (0/1) चर है और माना गया है कि यह लॉजिट लिंक फलन (ब्रेसलो एवं क्लेटन 1993) अर्थात् एक लॉजिस्टिक रैखिक मिश्रित मॉडल, के साथ एक व्यापकीकृत (generalized) रैखिक मिश्रित मॉडल (जी एल एम एम) का अनुसरण करता है जिसे  $y_{ij}$  के लॉजिस्टिक घटक के रूप में संदर्भित किया गया है। विशेषतया,  $s_{ij}, Y$  के धनात्मक मान के घटित होने के लिए एक बरनोली यादृच्छिक चर है अर्थात्  $s_{ij} = 1$  यदि  $y_{ij} > 0$  तथा  $s_{ij} = 0$  अन्यथा लॉजिस्टिक घटक  $\hat{y}_{ij}$  धनात्मक रूप से वैषमिक है तथा (1) के समरूप लॉग स्तर पर एक रैखिक मिश्रित मॉडल का अनुसरण करता है (फैफरमैन इत्यादि 2008 देखें)।

लघु क्षेत्र आकलन की प्रस्तावित मिश्रित मॉडल आधारित पद्धति को तीन चरणों में कार्यान्वित किया गया है। पहले चरण में एक रैखिक मिश्रित मॉडल को लघुगणक (logarithmic) परिवर्तित स्तर पर अनुक्रिया चर के धनात्मक (गैर-शून्य) वैषमिक (skewed) मानों के लिए फिट किया जाता है और दूसरे चरण में धनात्मक मानों की प्रायिकता के लिए एक लॉजिस्टिक रैखिक मिश्रित मॉडल फिट किया जाता है। अंत में, आकलन स्तर पर दोनों मॉडलों को संयुक्त कर दिया जाता है। चन्द्र एवं चैम्बर्स (2011b) ने शून्य-इन्फ्लेटेड वैषमिक आंकड़ों के लघु क्षेत्र आकलन के लिए एक इसी तरह के मिश्रित मॉडल का प्रयोग किया है। हालांकि, उन्होंने इस पद्धति को 'फिटिड मान' मॉडल के माध्यम से प्रतिदर्श भार प्राप्त करने के लिए तथा लघु क्षेत्र माध्यों के लिए एम बी डी ई आकल परिभाषित करने के लिए अपनाया (चन्द्र एवं चैम्बर्स, 2009)। उन्होंने प्रस्तावित एम बी डी ई आकल की माध्यम वर्ग त्रुटि आकलन पद्धति का भी वर्णन किया। विशेष रूप से, उन्होंने एम एस ई आकल के स्यूडो-रैखिकरण (pseudo-linearization) को परिभाषित किया (चैम्बर्स इत्यादि, 2011)। हालांकि, जैसा कि पहले से उल्लेखनीय है कि एम बी डी ई एक प्रत्यक्ष आकलक है और इसलिए जब क्षेत्र-विशेष प्रतिदर्श आकार बहुत छोटे होते हैं तो अस्थिर आकल प्रदान करता है। इसके विपरीत, लघु क्षेत्र आकलन की प्रस्तावित पद्धति एक अप्रत्यक्ष विधि है और इसलिए

ऐसे क्षेत्रों, जिनके लघु प्रतिदर्श आंकड़े ही उपलब्ध हैं, के लिए इस विधि के दक्ष होने की संभावना है।

उल्लेखनीय है कि द्वि-खण्ड यादृच्छिक प्रभाव मॉडल को फिट करना एक संगणनात्मक चुनौती है जो सामान्यीकृत (generalized) रैखिक मिश्रित मॉडल में पाई जाने वाली चुनौतियों के समरूप है। सामान्यतः निर्धारित गुणांक एवं प्रसरण घटकों के लिए अधिकतम संभावित आकल प्राप्त करने के लिए उच्च ऑर्डर लैप्लास सन्निकटनों (approximations) पर आधारित सन्निकट फिशर स्कोरिंग पद्धति का प्रयोग किया जाता है (ऑलसेन एवं शैफर (2001)।

फैफरमैन इत्यादि (2008) ने पूर्णतः द्वि-खण्ड यादृच्छिक प्रभाव मॉडल का प्रयोग किया जो मॉडल के दोनों खण्डों के यादृच्छिक प्रभावों के बीच सह-संबंधों को रखते हैं। हालांकि, उनके सिमुलेशन परिणामों से ज्ञात होता है कि इन सह-संबंध लघु क्षेत्र पूर्वकथक में कोई भी सार्थक सुधार नहीं जुड़ता है। इसके अतिरिक्त, इस सह-संबंध का प्रयोग मॉडल फिटिंग को संगणनात्मक रूप से गहन (intensive) बना देता है और हमेशा स्थिर (stable) नहीं रखता है। अभ्यास में, इन मॉडलों में सुविधा एवं संगणनात्मक व्यवहार्यता (feasibility) के लिए यादृच्छिक प्रभावों के बारे में अक्सर स्वतंत्र अभिधारणा की जाती है उदाहरणार्थ कार्लबर्ग (2000) एवं उसमें दिए गए संदर्भ देखें।

माना कि दोनों यादृच्छिक प्रभावों में सह-संबंध नहीं है अर्थात्-फैफरमैन इत्यादि (2008), चन्द्र एवं चैम्बर्स (2011b) तथा चन्द्र एवं सूद (2012) के संदर्भ में मिश्रित मॉडल के दो घटकों के बीच सह-संबंध नगण्य (negligible) है, कम से कम लघु क्षेत्र पूर्वकथक में उनका प्रभाव नगण्य है।

मिश्रित मॉडल के आकलन के लिए, लॉजिस्टिक घटक के लिए माना  $\delta_{ij}$ , जबकि  $x_{ij}$  ज्ञात है, स्वतंत्र बर्नोली यादृच्छिक चर हैं जिनकी प्रायिकता  $p_{ij} = P(y_{ij} > 0) = P(\delta_{ij} = 1)$  है। धनात्मक मानों वाली प्रायिकता  $p_{ij}$  को सहचरों के साथ जोड़ने (linking) से एक रैखिक मिश्रित मॉडल प्राप्त होगा जो  $i$  व  $j$  क्षेत्र के लिए निम्न प्रकार का होगा।

$$\begin{aligned} \text{logit}(p_{ij}) &= \ln \{p_{ij}/(1-p_{ij})\} \\ &= n_{ij} \\ &= x_{ij}^T \theta + v_i \quad (i = 1, \dots, D) \end{aligned} \quad (4)$$

जहां  $(p_{ij}) = \exp(n_{ij}) \{1 + \exp(n_{ij})\}^{-1}$   
 $= \exp(x_{ij}^T \theta + v_i) \{1 + \exp(x_{ij}^T \theta + v_i)\}^{-1}$

यहां  $\theta$  अज्ञात नियत (fixed) प्रभावों वाले प्राचलों का सादिश है तथा  $v_i$  ( $i = 1, \dots, D$ ),  $i$ वें क्षेत्र से संबंधित यादृच्छिक क्षेत्रीय प्रभाव है तथा माना इसका प्रसामान्य बंटन (distribution) है जिसका माध्य शून्य है तथा प्रसरण स्थिरांक (constant) है। लॉजिस्टिक घटक के अज्ञात प्राचलों का आकलन साई एवं चैम्बर्स (2003) तथा मैनेटेगा इत्यादि (2007) में वर्णित पद्धति का प्रयोग करके प्राप्त किया गया है। विशेष रूप से, हमने एक आइंटेरेटिव पद्धति का प्रयोग किया है जिसमें  $\theta$  एवं  $v_i$  के पैनेलाइज्ड क्वासी-संभाविता (पीक्यूएल) आकलन को प्रसरण घटक प्राचलों के आरईएमएल आकलन के साथ जोड़ा गया है। आकलन पद्धति को सांख्यिकीय सॉफ्टवेयर R में क्रियान्वित किया गया है। आकलित मानों का प्रयोग करके लॉजिस्टिक घटक की पूर्वानुमानित प्रायिकता निम्न होगी

$$\hat{p}_{ij} = \exp(x_{ij}^T \hat{\theta} + \hat{v}_i) \{1 + \exp(x_{ij}^T \hat{\theta} + \hat{v}_i)\}^{-1} \quad (5)$$

[साई एवं चैम्बर्स (2003) तथा मनटैगा इत्यादि (2007)] दूसरे घटके के प्राचलों का आकल अर्थात्  $y_{ij}$  का लॉग-रैखिक घटक प्राप्त करने के लिए माना  $s_+ = \{j \in s, y_{ij} > 0\}$  उस प्रतिदर्श का उपसमुच्चय है जिसके लिए अनुक्रिया चर गैर-शून्य है तथा जिसका वैष्मिक बंटन है तथा  $n_+ = \sum_{j \in s} \delta_j$  गैर-शून्य प्रतिदर्श इकाइयों की संख्या को दर्शाता है। तदनुसार,  $n_+$  आकार वाली गैर-शून्य प्रतिदर्श इकाइयों  $s_+$  से सम्बन्धित मात्रा को निरूपित करने के लिए हम सबस्क्रिप्ट '+' का प्रयोग करेंगे। प्रतिदर्श आंकड़े  $s_+$  का प्रयोग करके नियत प्रभाव प्राचलों का आकल तथा यादृच्छिक प्रभावों का पूर्वानुमान प्राप्त करने के लिए हम मॉडल (1) फिट करेंगे। इसके अतिरिक्त,  $n_+$  आकार वाले  $s_+$  प्रतिदर्श पर आधारित प्राचल आकल के लिए '+' का प्रयोग करेंगे। विशेष रूप से, जब  $n_+$  आकार वाले  $s_+$  गैर-शून्य प्रतिदर्श के लिए मॉडल

(1) फिट करने पर  $\beta$  का अनुभवजन्य सर्वश्रेष्ठ रैखिक अनभिनत आकलन (ई बी एल यू ई) निम्न होगा

$$\begin{aligned} \hat{\beta}^+ &= (\sum_{i=1}^D x_{is_+}^T \hat{v}_{iss_+}^{-1} x_{is_+})^{-1} (\sum_{i=1}^D x_{is_+}^T \hat{v}_{iss_+}^{-1} y_{s_+}) \\ \text{जहां} \\ E(\hat{\beta}^+ / \delta_j) &= \beta \text{ तथा } v(\hat{\beta}^+ / \delta_j) = (\sum_{i=1}^D x_{is_+}^T \hat{v}_{iss_+}^{-1} x_{is_+})^{-1} \\ \text{यहां} \\ \hat{v}_{iss_+} &= \text{diag.} \{ \hat{\sigma}_u^{2+} 1_{is_+} 1_{is_+}^T + \hat{\sigma}_e^2 I_{is}; i = 1, K, D \} \end{aligned}$$

जहां  $1_{is_+}$  एवं  $I_{is_+}$  क्रमशः 1 संख्या वाला सदिश एवं  $n_+$  ऑडर वाला आइडेंटिटी आव्यूह है। यहां सभी संभावित मान एवं प्रसरण, मॉडल के अन्तर्गत लिए गए हैं। यादृच्छिक क्षेत्र प्रभावों के लिए पूर्वकथकों की संगणना  $\hat{u}_i^+ = \hat{Y}_i^+ (\bar{l}_{is_+} + \bar{z}_{is_+}^T \hat{\beta}^+)$  एवं  $\hat{Y}_i^+ = \hat{\sigma}_u^{2+} (\hat{\sigma}_u^{2+} + n_i^{-1} \hat{\sigma}_e^2)^{-1}$  से की गई है।  $y_{ij}$  का पूर्वानुमित मान अर्थात्  $y_{ij}$  रैखिक घटक, (3) का प्रयोग करके प्राप्त किया जा सकता है। अनुभवजन्य सर्वश्रेष्ठ प्रकार का पूर्वकथक निम्न होगा

$$\begin{aligned} \hat{E}(y_{ij}^{EPB-BC+}) & \\ &= \hat{E}[(\hat{C}_{ij}^{EPB+})^{-1} \exp \{Z_{ij}^T \hat{\beta}^+ + \hat{\gamma}_i^+ (\bar{l}_{is} - Z_{ij}^T \hat{\beta}^+) \\ &+ \frac{1}{2} \hat{\sigma}_e^2 (1 + n_i^{-1} \hat{\gamma}_i^+)\}] \end{aligned} \quad (6)$$

एवं

$$\begin{aligned} \hat{C}_{ij}^{EPB+} &= \exp \left[ \frac{1}{2} \{ (Z_{ij}^T - \hat{\gamma}_i^+ \bar{z}_{is}^T)^T \hat{V}(\hat{\beta}^+) (Z_{ij}^T - \hat{\gamma}_i^+ \bar{z}_{is}^T) \right. \\ &+ \hat{C}_{i1}^+ + \hat{C}_{i2}^+ + 2\hat{C}_{i3}^+ \left. \right] \end{aligned}$$

जहां  $\hat{C}_{i1}^+$ ,  $\hat{C}_{i2}^+$  एवं  $\hat{C}_{i3}^+$  का मान क्रमशः  $\hat{C}_{i1}$ ,  $\hat{C}_{i2}$  एवं  $\hat{C}_{i3}$  में प्राचलों के आकल,  $(\hat{\beta}, \hat{\sigma}_e^2, \hat{\sigma}_u^2)$  के स्थान पर  $(\hat{\beta}^+, \hat{\sigma}_e^{2+}, \hat{\sigma}_u^{2+})$  का प्रयोग करके प्राप्त किये गये हैं। आगे हम देखते हैं कि

$$\begin{aligned} E(y_{ij}) &= E(y_{ij} / \delta_{ij} = 0) \Pr(\delta_{ij} = 0) \\ &+ E(y_{ij} / \delta_{ij} = 1) \Pr(\delta_{ij} = 1) = p_{ji} \mu_{ij} \quad (7) \end{aligned}$$

यहां  $\mu_{ij} = E(y_{ij} / \delta_{ij} = 1)$ । यह स्पष्ट है कि (7) के दाहिने तरफ के दो खण्डों का मॉडल निकालेंगे। पहला अनुक्रिया चर (5) के धनात्मक मानों की प्रायिकताओं के लिए तथा दूसरा धनात्मक मान वाले वैयक्तिक (individual) [(6) एवं (7)] के लिए। याद रहे कि



हमने मॉडल के दो खण्डों में यादृच्छिक प्रभावों के बीच सहसंबंध को नगण्य माना है। लघु क्षेत्र पूर्वानुमान समस्या के संबंध में, फ़ैफरमैन इत्यादि (2008), चन्द्र एवं चैम्बर्स (2011b) में दिये गये अनुभवजन्य साक्ष्यों (evidence) एवं दिए गए संदर्भों से स्पष्ट रूप से ज्ञात होता है कि यह एक उचित धारणा (assumption) है। परिणाम स्वरूप यह माना जा सकता है कि  $\mu_{ij}$  एवं  $p_{ij}$  में सहसंबंध नहीं है (या नगण्य सहसंबंध है)।

(7) में दिए गए प्राचलों के आकलित मानों का प्रयोग करके  $y_{ij}$  का पूर्वानुमानित मान निम्न होगा

$$\hat{E}(y_{ij}) = \hat{p}_{ij}\hat{\mu}_{ij} \quad (8)$$

जहां  $p_{ij}$  का मान (5) में दिया गया है जबकि  $\hat{\mu}_{ij} = \hat{E}(y_{ij}/\delta_{ij})$ ; (6) द्वारा परिभाषित किया गया है। हम देखते हैं कि

$$E\{\hat{E}(y_{ij})\} = p_{ij} \mu_{ij} = E(y_{ij}) \quad (9)$$

द्वि-खण्ड यादृच्छिक प्रभाव (या मिश्रण) मॉडल के अन्तर्गत लघु क्षेत्र  $i$  में  $Y$  के समष्टि के माध्य के लिए तीन विभिन्न आकलक निम्न होंगे

- कृत्रिम (synthetic) प्रकार का अनुभवजन्य पूर्वकथक अर्थात Mix EP
- एमबीडीई पूर्वकथक अर्थात Mix MBDE
- अनुभवजन्य सर्वश्रेष्ठ पूर्वकथक प्रकार का आकलक अर्थात Mix EBP

### अनुभवजन्य मूल्यांकन

इस अनुभाग में हम अनुभवजन्य मूल्यांकनों से प्राप्त परिणामों को देंगे जिनसे लघु क्षेत्र माध्यों के विभिन्न आकलकों के निष्पादन को वर्णित किया गया है। कृत्रिम समष्टि एवं प्रतिदर्श आंकड़ों को जेनरेट करने के लिए मॉडल आधारित सिमुलेशन का प्रयोग किया गया है। इसमें, प्रत्येक सिमुलेशन में मॉडल के अन्तर्गत समष्टि आंकड़े जेनरेट किए गए हैं और फिर सिमुलेटिड समष्टि में से, लघु क्षेत्र को स्तर के रूप में मानते हुए प्रतिस्थापन रहित स्तरीकृत सरल यादृच्छिक प्रतिचयन पद्धति का प्रयोग करते हुए, एक एकल प्रतिदर्श का चयन किया गया। इन आंकड़ों

का प्रयोग विभिन्न आकलकों के निष्पादन की तुलना करने के लिए किया गया। विभिन्न लघु क्षेत्र आकलन पद्धतियों, जिनका प्रयोग सिमुलेशन के लिए किया गया है, के सापेक्षिक (relative) निष्पादन के दो मापों का प्रयोग किया गया। एक आकलों की औसत प्रतिशत सापेक्षिक अभिनति तथा दूसरा आकलों की औसत प्रतिशत सापेक्षिक वर्ग माध्य मूल त्रुटि है।

**तालिका 1:** सिमुलेशन अध्ययनों में प्रयुक्त लघु क्षेत्र पूर्वानुमानकों की परिभाषाएं

आकलक	विवरण
<b>मिश्रित मॉडल आधारित पद्धति</b>	
मिक्स ईबीपी	अनुभवजन्य सर्वश्रेष्ठ पूर्वानुमानक प्रकार के आकलक
मिक्स ईपी	अनुभवजन्य पूर्वानुमानक प्रकार के आकलक
मिक्स एमबीडीई	एमबीडीई प्रकार के आकलक
<b>रॉ (Raw) स्तर की रैखिक मिश्रित मॉडल आधारित पद्धति</b>	
लिन ईबीएलयूपी	मानक ईबीएलयूपी

मॉडल आधारित सिमुलेशन, आध्ययनाधीन समष्टि की संरचना के बारे में संधारणाओं (assumptions) में विविधता के आकलन की पद्धति की संवेदनशीलता को उदाहरण देकर स्पष्ट करने के लिए एक सामान्य तरीका है। मॉडल आधारित सिमुलेशन अध्ययन, मॉडल (1) के अन्तर्गत जेनरेटिड समष्टि आंकड़ों पर आधारित हैं। माना समष्टि आकार  $N = 15000$  एवं प्रतिदर्श आकार  $n = 600$  है और जेनरेटिड लघु क्षेत्र समष्टि आकार  $N_d, d = 1, \dots, D; \sum_d N_d = N$  एवं प्रतिदर्श आकार  $n_d = N_d (n/N); \sum_d n_d = n$  है। लघु क्षेत्र समष्टि एवं प्रतिदर्श के औसत आकार क्रमशः 500 एवं 20 हैं जबकि कुल  $D = 30$  लघु क्षेत्र हैं। यह सभी सिमुलेशनों के लिए नियत (fixed) है। हमने मॉडल आधारित सिमुलेशन निम्न प्रकार से किया है। पहले गुणात्मक मॉडल  $y_{ai} = \beta_0 x_{ai}^\beta u_a e_{ai}$  जहां  $\beta_0 = 5.0$ , एवं  $\beta_1 = 0.5$ , में से  $y_{ai}$  ( $i = 1, \dots, Nd; d = 1, \dots, D$ ) के समष्टि

मान जेनरेट किए गए हैं और फिर यादृच्छिकीकरण के माध्यम से  $y_{di}$  के लिए शून्य मान (Zero values) बनाये गये हैं। इकाई स्तर यादृच्छिक त्रुटियां  $e_{di}$  स्वतंत्र रूप से लॉग-प्रसामान्य बंटन LN (0,  $O_e = 0.5$ ) से जेनरेट की गयी हैं। यादृच्छिक क्षेत्र प्रभाव  $ud$  ( $d = 1, \dots, D$ ), LN ( $0, O_u = 0.3$ ) से जेनरेट किए गए हैं। सहचर मान  $x_{di}$ , LN ( $2, O_x = 3$ ) से जेनरेट किए गए हैं। इस मॉडल से,  $y_{di}$  के मान (जिनमें शून्य मान भी शामिल हैं) लघु क्षेत्र  $D = 30$ , जिनका आकार  $n_d$  है, से जेनरेट किए गए तथा फिर प्रत्येक क्षेत्र से  $n_d$  आकार के यादृच्छिक प्रतिदर्शों का चयन किया गया। सिमुलेशन के अन्तर्गत हमने समष्टि स्तर पर सभी लघु क्षेत्रों के लिए  $p = 0.50, 0.70$  एवं  $0.90$  लेते हुए आंकड़ों का सृजन किया। यहां  $p$  धनात्मक मानों का अनुपात है जिसे समष्टि में धनात्मक मानों की कुल संख्या को, समष्टि में कुल मानों की संख्या से विभाजित करके परिभाषित किया गया। आगे  $n_d = 10$  आकार वाले क्षेत्र विशिष्ट प्रतिदर्श सहित  $n = 300$  आकार वाले लघु प्रतिदर्श वाले इन सिमुलेशन्स की पुनरावृत्ति की। सिमुलेशन्स रन्स की  $K = 1000$  बार पुनरावृत्ति की और प्रत्येक समुच्चय के लिए प्रत्येक सिमुलेशन में लघु क्षेत्र आकलों के मान तालिका 1 में वर्णित विभिन्न लघु क्षेत्र आकलन पद्धतियों का प्रयोग करके, परिकलित (calculated) किए। इन सिमुलेशनों के परिणाम तालिका 2 में दिए गए हैं।

**तालिका 2:** मॉडल आधारित सिमुलेशन्स में विभिन्न आकलों की प्रतिशत औसत सापेक्षित अभिनति (एव आर बी) एवं प्रतिशत औसत सापेक्षित आर एम एस ई (एव आर आर एम एस ई)

पी	मिक्स ईबीपी	मिक्स ई पी	मिक्स एमबीडीई	लिन ईबीएलयूपी
एव आर बी				
$n_i = 20$				
0.90	0.50	1.11	0.68	13.06
0.70	0.58	1.07	0.99	12.42
0.50	0.75	1.22	1.12	13.95

पी	मिक्स ईबीपी	मिक्स ई पी	मिक्स एमबीडीई	लिन ईबीएलयूपी
$n_i = 10$				
0.90	0.61	1.04	0.94	27.52
0.70	0.85	1.21	1.46	21.99
0.50	1.02	1.07	2.41	30.18
एव आर आर एम एस ई				
$n_i = 20$				
0.90	15.07	31.03	18.98	77.88
0.70	19.39	32.86	26.96	77.83
0.50	24.65	35.61	36.83	96.90
$n_i = 10$				
0.90	20.25	33.42	27.11	243.74
0.70	24.54	35.20	37.47	139.47
0.50	30.53	38.49	52.36	303.73

तालिका 2 में दिए गए प्रतिशत औसत सापेक्षित अभिनति (एव आर बी) के मानों से दो परिणाम सामने आते हैं, पहला रॉ पैमाने पर एक रैखिक मिश्रित मॉडल के अन्तर्गत लिन बी एल यू पी अत्यधिक अभिनत है तथा ये अभिनतियां सभी मिश्रित मॉडल आधारित लघु क्षेत्र आकलन पद्धतियों (अर्थात् मिक्स ई बी पी, मिक्स ई पी एवं मिक्स एम बी डी ई) की तुलना में सार्थक रूप से काफी अधिक हैं। इससे यह स्पष्ट होता है कि लिन ई बी एल यू पी; अर्द्ध सतत आंकड़ों के लिए उपयुक्त नहीं है। दूसरा मिश्रण मॉडल आधारित लघु क्षेत्र आकलन पद्धतियों में मिक्स एम बी डी ई एवं मिक्स ई पी की तुलना में मिक्स ई बी पी की अभिनतियां छोटी होती हैं। छोटे प्रतिदर्श आकार होने की स्थिति में मिक्स ई बी पी के अभिनति लाभ (gains) बड़े होते हैं। मध्यम (moderate) प्रतिदर्श आंकड़ों (अर्थात्  $n_i = 20$ ) के लिए अभिनतियों के संदर्भ में मिक्स ई पी की तुलना में मिक्स एम बी डी ई प्राबल्य (dominate) रखता है। लेकिन लघु प्रतिदर्श आकरों (अर्थात्  $n_i = 10$ ) के लिए ऐसा नहीं है। लघु क्षेत्र प्रतिदर्श आकरों को कम करने की स्थिति में या शून्य की अनुपातिक वृद्धि की स्थिति में (अर्थात् आंकड़े में शून्य मुद्रास्फीति (Zero inflation) का स्तर), सभी पद्धतियों

के लिए औसत सापेक्षित अभिनतियों में वृद्धि हो जाती है। तालिका 2 में दी गई प्रतिशत औसत सापेक्षित वर्ग माध्य मूल त्रुटियों (एव आर आर एम एस ई) से ज्ञात होता है कि लघु क्षेत्र विशिष्ट प्रतिदर्श आकार या वृहत क्षेत्र शून्य अनुपात की स्थिति में सभी पद्धतियों की प्रतिशत औसत सापेक्षित वर्ग माध्य मूल त्रुटियों में वृद्धि हो जाती है। हम देखते हैं कि मिश्रित मॉडल आधारित पद्धतियों की तुलना में लिन ई बी एल यू पी की सापेक्ष वर्ग माध्य मूल त्रुटियों के मान बहुत अधिक हैं।

शून्य मुद्रास्फीति (Zero inflation) वैषम्यिक आंकड़ों के अर्थात् लघु क्षेत्र आकलन के लिए लिन ई बी एल यू पी अस्थिर (unstable) हैं। मिश्रित मॉडल आधारित पद्धतियों में मिक्स ई बी पी अन्य पद्धतियों पर प्राबल्य रखता है। कुल मिलाकर, प्रस्तावित मिक्स ई बी पी में लघु अभिनतियां एवं उच्च दक्षता दोनों हैं। अध्ययनाधीन चर के अर्द्ध सतत होने की स्थिति में द्वि-खण्ड यादृच्छिक प्रभाव मॉडल (मिक्स ई बी पी) के अन्तर्गत प्रस्तावित पूर्वानुमानक, अन्य पूर्वानुमानकों की तुलना में प्रयाप्त अभिनति (substantial bias) एवं दक्षता लाभ (efficiency gains) प्रदान करता है। इस स्थिति में, लिन ई बी एल यू पी का प्रयोग करने की संस्तुति नहीं (not recommended) की जाती है क्योंकि यह अत्यधिक अभिनतियों सहित बहुत ही अस्थिर (unstable) है (देखें तालिका 2)।

इसके विपरीत प्रस्तावित मिश्रित मॉडल आधारित लघु क्षेत्र आकलन पद्धति विशेषतया मिक्स ई बी पी इस तरह के आंकड़ों के लिए एक बेहतर विकल्प के रूप में सिद्ध पायी गयी है।

### निष्कर्ष

इस शोधपत्र में हमने मिश्रित या द्वि-खण्ड यादृच्छिक प्रभाव मॉडल के अन्तर्गत अर्द्ध सतत चरों के लिए लघु क्षेत्र आकलन पद्धति का अन्वेषण (explore) किया है। लघु क्षेत्र माध्यों के लिए अनुभवजन्य सर्वश्रेष्ठ पूर्वानुमानक प्रकार का आकलक प्रस्तावित किया गया है। अनुभवजन्य परिणामों से ज्ञात होता है कि मिश्रित मॉडल आधारित अनुभवजन्य सर्वश्रेष्ठ

पूर्वानुमानक प्रकार का आकलक (मिक्स ई बी पी) कम अभिनत (less biased) है तथा आंकड़ों के वैषम्यिक होने की स्थिति में और पर्याप्त (substantial) अनुपात में शून्यों की उपस्थिति होने की स्थिति में, यह आकलक कृत्रिम (synthetic) प्रकार के आकलक (मिक्स ई बी पी) तथा समान मिश्रित मॉडल पर आधारित मॉडल आधारित प्रत्यक्ष प्रकार के आकलक (मिक्स एम बी डी ई) की तुलना में अधिक दक्ष हो सकता है। इसके अतिरिक्त, एक रैखिक मिश्रित मॉडल आधारित ई बी एल यू पी आकलक (लिन ई बी एल यू पी) अत्यधिक अभिनत (biased) है और बहुत अस्थिर (unstable) है, अतः यह इस प्रकार के आंकड़ों के लिए एक उपयुक्त पद्धति नहीं है।

### संदर्भ

1. ब्रेसलो, एन. ई. एवं क्लेटन, डी. जी (1993): "एपरोक्सिमेट इंफ्रेंस इन जेनेरेडाइज्ड लिनियर मिक्सड मॉडल", जर्नल ऑफ द अमेरिकन स्टैटिस्टिक्स एमोसिएशन, 88, 9-25।
2. बग, ई. एवं चन्द्र, एच. (2012): "स्मॉल एरिया प्रिडिक्शन फॉर ए यूनिट लेवल लॉगनॉर्मल मॉडल", प्रोसिडिंग्स ऑफ द 2012 फेडरल कमेटी ऑन स्टैटिस्टिकल मैथोडोलॉजी रिसर्च कॉन्फ्रेंस, वाशिंगटन, डी सी, यू एस ए, जनवरी 10-12, 2012।
3. चन्द्र एच. एवं सूद, यू.सी. (2012): "स्मॉल एरिया एस्टिमेशन फॉर जीरो-इंपलेटिड डाटा", कम्प्यूटेशनल इन स्टैटिस्टिक्स-सिमुलेशन एण्ड कम्प्यूटेशन, 41(5), 632-643।
4. चैम्बर्स, आर, चन्द्र, एच. एवं तविदिस, एन. (2011): "ऑन बायस-रॉबस्ट मीन स्क्वायर्ड एरर एस्टिमेशन फॉर लीनियर प्रिडिक्टर्स फॉर डोमेन", सर्वे मैथोडोलॉजी, 37(2), 1-17।
5. चन्द्र, एच. एवं चैम्बर्स, आर. (2011a): "स्मॉल एरिया एस्टिमेशन अण्डर ट्रांसफोरमेशन टू लीनियरिटी", सर्वे मैथोडोलॉजी, 37(1), 39-51।
6. चन्द्र, एच. एवं चैम्बर्स, आर. (2011b): "स्मॉल

- एरिया एस्टिमेशन फॉर स्क्वड डाटा इन प्रैजेन्स ऑफ जीरोज”, द बुलेटिन ऑफ कलकत्ता स्टैटिस्टिकल एसोसिएशन, 63, 249–252।
7. चन्द्र, एच. एवं चैम्बर्स, आर. (2009): “मल्टीपरपज बेटिंग फॉर स्मॉल एरिया एस्टिमेशन”, जर्नल ऑफ ऑफिशियल स्टैटिस्टिक्स, 25(3), 379–395।
  8. फ्लेचर, डी, मैकेंजी, डी. एवं विलोटा, ई. (2005): “मॉडलिंग स्क्वड डाटा विद मैनी जीरोज: ए सिम्पल एप्रोच कम्बाइनिंग ऑर्डिनरी ऐन्ड लॉजिस्टिक रिग्रेशन”, जर्नल ऑफ एन्वायरमेंटल एण्ड इकोलॉजिकल स्टैटिस्टिक्स, 12(1), 45–54।
  9. कालबर्ग, एफ. (2000): “पॉपुलेशन टोटल प्रिडिक्शन अण्डर ए लॉगनॉर्मल सुपर पॉपुलेशन मॉडल”, मेट्रॉन, LVIII, 53–80।
  10. मनटैगा, जी, डब्ल्यू, लोम्बार्डिया, एम. के., मोलिना, आई, मोरल्स, डी. एवं सन्तामारिया, एल (2007): “एस्टिमेशन ऑफ द मीन स्क्वायर्ड एरर ऑफ प्रिडिक्टरस ऑफ स्मॉल एरिया लीनियर पैरामीटरस अण्डर ए लॉजिस्टिक मिक्स्ड मॉडल”, कम्प्यूटेशनल स्टैटिस्टिक्स ऐन्ड डाटा अनेलिसिस, 51(5), 2720–2733।
  11. ओल्सन, एम. के. एवं शैफर; जे. एल. (2001): “ए टू-पार्ट रेण्डम-इफैक्ट मॉडल फॉर सेमीकॉन्टिन्यूज लॉन्गिट्यूडिनल डाटा”, जर्नल ऑफ द अमेरिकन स्टैटिस्टिकल एसोसिएशन, 96(454), 730–745।
  12. फ़ैफरमैन, डी., टैरी; बी एवं गॉरा, एल.ए.एस. (2008): “स्मॉल एरिया एस्टिमेशन अण्डर ए टू-पार्ट रेण्डम इफैक्ट्स मॉडल विद एप्लिकेशन टू एस्टिमेशन ऑफ लिटरेसी इन डेवलपिंग कंट्रीज”, सर्वे मैथोडोलॉजी, 34(2), 235–249।
  13. राव, जे. एन. के. (2003): “स्मॉल एरिया एस्टिमेशन वायले”, न्यूयॉर्क।
  14. साई, ए. एवं चैम्बर्स, आर. (2003): “स्मॉल एरिया एस्टिमेशन अण्डर लीनियर ऐण्ड जेनेरेलाइज्ड लीनियर मिक्सड मॉडल्स विद टाइम ऐन्ड एरिया इफैक्ट्स”, मैथोडोलॉजी वर्किंग पेपर न. M03/15, यूनिवर्सिटी ऑफ साउथम्पटन, यू. के. (www.s3ri.soton.ac.uk पर उपलब्ध)।

□

हिन्दी वह धागा है जो विभिन्न मातृभाषा रूपी फूलों को  
पिरोकर भारतमाता के लिए सुन्दर हार का सृजन करेगा।

— डॉ. जाकिर हुसैन

## संस्थान की राजभाषा यात्रा : 2014-15

ऊषा जैन

भारतीय कृषि सांख्यिकी अनुसंधान संस्थान में वर्ष-दर-वर्ष हिन्दी के प्रगामी प्रयोग में अभिवृद्धि हो रही है। राजभाषा नीति को संस्थान में सुचारु रूप से कार्यान्वित किया जा रहा है। भारत सरकार, गृह मंत्रालय, राजभाषा विभाग द्वारा जारी वार्षिक कार्यक्रम में निहित लक्ष्यों को इस संस्थान में लगभग पूरा कर लिया गया है। संस्थान द्वारा समस्त प्रशासनिक कार्य शत-प्रतिशत हिन्दी में और यथाआवश्यक द्विभाषी हो रहा है।

संस्थान में प्रशासनिक कार्य के अतिरिक्त वैज्ञानिक प्रकृति के कार्यों में भी हिन्दी के उपयोग को प्रोत्साहित किया जाता है। न केवल मात्रात्मक रूप में बल्कि हिन्दी के प्रयोग में गुणवत्ता की ओर भी ध्यान दिया जा रहा है। वैज्ञानिक प्रभागों द्वारा आयोजित प्रशिक्षण कार्यक्रमों की संदर्भ पुस्तिकाओं में आमुख एवं प्राक्कथन द्विभाषी रूप में प्रस्तुत करने के साथ-साथ वैज्ञानिकों ने अपनी परियोजना रिपोर्टों के कवर पेज, आमुख, प्राक्कथन एवं सारांश द्विभाषी रूप में प्रस्तुत किये तथा कुछ वैज्ञानिकों द्वारा अपनी परियोजना रिपोर्टों में तालिकाएँ भी द्विभाषी रूप में प्रस्तुत की गयीं। संस्थान के वैज्ञानिकों द्वारा हिन्दी में वैज्ञानिक विषय पर हिन्दी कार्यशाला का आयोजन किया गया। इसके अतिरिक्त, संस्थान में एम.एससी. तथा पीएच.डी. के विद्यार्थियों द्वारा अपने शोध-प्रबन्धों में सार द्विभाषी रूप में प्रस्तुत किये गये। वैज्ञानिकों एवं तकनीकी कर्मियों द्वारा शोध-पत्र हिन्दी में प्रकाशित किये गये।

राजभाषा विभाग द्वारा जारी वार्षिक कार्यक्रम

में निहित लक्ष्यों को पूरा करते हुए संस्थान के अधिकारियों/कर्मचारियों द्वारा समस्त पत्राचार हिन्दी में अथवा द्विभाषी रूप में किया गया। संस्थान के विभिन्न वैज्ञानिक प्रभागों तथा प्रशासनिक अनुभागों द्वारा आयोजित की जाने वाली बैठकों की कार्यसूची तथा कार्यवृत्त शत-प्रतिशत हिन्दी में अथवा द्विभाषी रूप में जारी किये गये। संस्थान में अपना कार्य शत-प्रतिशत हिन्दी में करने के लिए दस अनुभाग पहले से ही विनिर्दिष्ट हैं। गृह मंत्रालय, राजभाषा विभाग द्वारा जारी विभिन्न नकद पुरस्कार योजनाएँ संस्थान में लागू हैं तथा संस्थान के कर्मियों ने इन योजनाओं में भाग लिया।

प्रतिवदेनाधीन अवधि के दौरान संस्थान के विभिन्न वर्गों के कर्मियों के लिए चार हिन्दी कार्यशालाएँ आयोजित की गयीं। प्रथम कार्यशाला 24 मई 2014 को "एस.पी.एस.एस द्वारा आँकड़ों का विश्लेषण" जैसे तकनीकी विषय पर आयोजित की गयी। इस कार्यशाला में संस्थान के परीक्षण अभिकल्पना प्रभाग की प्रमुख वैज्ञानिक, डॉ. सीमा जग्गी, वरिष्ठ वैज्ञानिक, डॉ. सिनी वरगीस, वैज्ञानिक, डॉ. एल्दो वरगीस एवं डॉ. अर्पण भौमिक द्वारा उपरोक्त विषय से सम्बन्धित विभिन्न उप-विषयों पर हिन्दी भाषा में व्याख्यान दिये गये। कार्यशाला में वक्ताओं द्वारा प्रतिभागियों को व्याख्यान की सामग्री हिन्दी भाषा में उपलब्ध करायी गयी। संस्थान के विभिन्न वर्ग के अधिकारियों/कर्मचारियों के लिए 24 सितम्बर 2014, 12 दिसम्बर 2014 तथा 27 फरवरी 2015 को "हिन्दी मासिक प्रगति रिपोर्ट का प्रपत्र

भरने तथा डिस्पैच एवं डायरी रजिस्ट्रों का रखरखाव” विषय पर तीन हिन्दी कार्यशालाओं का आयोजन किया गया जिनमें संस्थान की वरिष्ठ तकनीकी अधिकारी, सुश्री ऊषा जैन ने प्रतिभागियों को सम्बन्धित विषय पर जानकारी उपलब्ध करायी/व्याख्यान दिया ।

प्रतिवेदनाधीन अवधि में संस्थान में राजभाषा कार्यान्वयन समिति की बैठकें आयोजित की गयीं। इन बैठकों में राजभाषा अधिनियम 1963 की धारा 3(3) के अनुपालन को सुनिश्चित करने, राजभाषा विभाग द्वारा जारी वार्षिक कार्यक्रम की विभिन्न मदों, राजभाषा विभाग एवं परिषद् मुख्यालय से समय-समय पर प्राप्त निदेशों का अनुपालन सुनिश्चित करने, हिन्दी पत्रिका के प्रकाशन, कार्यशालाओं के नियमित आयोजन, हिन्दी सप्ताह के आयोजन इत्यादि पर विस्तार से चर्चा हुई।

संस्थान में कार्यरत सभी हिन्दीतर भाषी अधिकारियों/कर्मचारियों द्वारा हिन्दी ज्ञान सम्बन्धी प्रशिक्षण पूरा किया जा चुका है। आज तक की स्थिति के अनुसार, संस्थान में अब कोई ऐसा हिन्दीतर भाषी अधिकारी/कर्मचारी शेष नहीं रह गया है जिसे हिन्दी ज्ञान सम्बन्धी प्रशिक्षण दिया जाना शेष हो। इसके अतिरिक्त, ‘हिन्दी शिक्षण योजना’ के अन्तर्गत हिन्दी आशुलिपि एवं हिन्दी टंकण के प्रशिक्षण का लक्ष्य अभी तक पूरा था परन्तु राजभाषा विभाग से प्राप्त दिशा-निर्देशों के अनुसरण में वर्ग ‘घ’ से वर्ग ‘ग’ में गये कर्मियों में से वर्ग ‘ग’ श्रेणी के लिए निर्धारित शैक्षिक योग्यता रखने वाले कर्मियों को रोस्टरबद्ध कर उन्हें हिन्दी टंकण का प्रशिक्षण दिये जाने सम्बन्धी कार्रवाई की जा रही है।

संस्थान की वेबसाइट द्विभाषी है जिसको समय-समय पर अद्यतन किया गया। संस्थान की वेबसाइट पर उपलब्ध ‘हिन्दी सेवा लिंक’ में सांख्यिकीय एवं प्रशासनिक शब्दावली के वर्ण क्रमानुसार कुछ शब्द, कुछ द्विभाषी प्रपत्र, दैनिक काम काज के प्रयोग में आने वाली कुछ टिप्पणियाँ, द्विभाषी पदनाम, वाक्यांश इत्यादि सामग्री उपलब्ध है तथा समय-समय पर इसमें सामग्री जोड़ी जाती है। अपना दैनिक कार्य हिन्दी में सरलता से करने के लिए संस्थान के कर्मियों द्वारा इस सेवा का उपयोग किया जाता है।

संस्थान द्वारा प्रकाशित हिन्दी पत्रिका, ‘सांख्यिकी-विमर्श’ के नौवें अंक का प्रकाशन मार्च 2014 में किया गया। इस पत्रिका में संस्थान के कीर्तिस्तम्भ, सम्बन्धित वर्ष में किये गये अनुसंधानों व अन्य कार्यों के संक्षिप्त विवरण, राजभाषा से सम्बन्धित कार्यों आदि की जानकारी के साथ-साथ कृषि सांख्यिकी, संगणक अनुप्रयोग एवं कृषि जैव-सूचना से सम्बन्धित विभिन्न लेखों एवं शोध-पत्रों को भी प्रस्तुत किया जाता है। हिन्दी दिवस के अवसर पर संस्थान में आयोजित डॉ. दरोगा सिंह स्मृति व्याख्यानमाला के अन्तर्गत प्रस्तुत वैज्ञानिक व्याख्यान को आमन्त्रित ज्ञानवर्धक लेख के रूप में इस पत्रिका में सम्मिलित किया जाता है। पाठकों के हिन्दी ज्ञानवर्धन के लिए दैनिक स्मरणीय शब्द-शतक हिन्दी व अँग्रेजी में दिया जाता है।

इस वर्ष संस्थान में 09 से 15 सितम्बर 2014 के दौरान हिन्दी सप्ताह का आयोजन किया गया। दिनांक 09 सितम्बर 2014 को हिन्दी सप्ताह का उद्घाटन संस्थान के निदेशक, डॉ. उमेश चन्दर सूद जी द्वारा किया गया। हिन्दी सप्ताह के उद्घाटन के तत्पश्चात प्रश्न-मंच का आयोजन किया गया। हिन्दी सप्ताह के दौरान डॉ. दरोगा सिंह स्मृति व्याख्यान के साथ-साथ वैज्ञानिक प्रभागों में हिन्दी में सर्वाधिक वैज्ञानिक कार्य करने के लिए प्रभागीय चल-शील्ड, काव्य-पाठ, प्रश्न-मंच, अन्ताक्षरी, हिन्दीतर कर्मियों के लिए हिन्दी श्रुतलेख एवं शब्दार्थ लेखन प्रतियोगिता आयोजित की गयी। अन्ताक्षरी प्रतियोगिता के संचालकों द्वारा इस प्रतियोगिता को ऑडियो विजुअल रूप में प्रस्तुत किया गया जिससे यह प्रतियोगिता अत्यन्त ही रोचक रही। सभी प्रतियोगिताओं में छात्रों सहित संस्थान के विभिन्न वर्गों के कर्मियों ने बढ़-चढ़कर सहभागिता की। संस्थान में प्रत्येक वर्ष हिन्दी दिवस के अवसर पर डॉ. दरोगा सिंह स्मृति व्याख्यान का आयोजन किया जाता है जिसमें किसी सुप्रसिद्ध वैज्ञानिक द्वारा किसी भी वैज्ञानिक विषय पर हिन्दी में व्याख्यान दिया जाता है। इस वर्ष इस कड़ी का तेइसवाँ व्याख्यान भारतीय कृषि सांख्यिकी अनुसंधान संस्थान की पूर्व प्रमुख वैज्ञानिक, डॉ. रंजना अग्रवाल जी द्वारा “फसलों के उपज के

कटाई-पूर्व अनुमान हेतु मौसम आधारित मॉडल" विषय पर दिया गया और इस कार्यक्रम की अध्यक्षता परिषद् के पूर्व सहायक महानिदेशक (एच.आर.डी.), डॉ. सुखदेव शर्मा जी द्वारा की गयी। दिनांक 15 सितम्बर, 2014 को हिन्दी सप्ताह के समापन समारोह के अवसर पर इस दौरान आयोजित प्रतियोगिताओं के सफल प्रतियोगियों को पुरस्कृत करने के साथ-साथ वर्ष 2013-14 के

दौरान "सरकारी कामकाज मूल रूप से हिन्दी में करने के लिए प्रोत्साहन योजना" के अन्तर्गत भी नकद पुरस्कार प्रदान किये गये। इसके अतिरिक्त, जुलाई 2013 से जून 2014 तक की अवधि के दौरान संस्थान में आयोजित कार्यशालाओं के वक्ताओं/प्रशिक्षकों को भी सम्मानित किया गया।

आभार : लेखिका टंकण हेतु श्री खुशियाल, सहायक की आभारी हैं ।

दुनिया को आकार देने वाली ताकतें हैं – विज्ञान, आत्मज्ञान एवं साहित्य। विज्ञान तथा आत्मज्ञान को जोड़ने वाला सेतु साहित्य है।

—विनोबा भावे

## दैनिक स्मरणीय शब्द शतक

1. Abnormal curve	अप्रसामान्य वक्र	30. Approximation	सन्निकटन
2. Absolute acceleration	निरपेक्ष त्वरण	31. Arable (land)	कृष्य, कृषिगत
3. Absolute coefficient	निरपेक्ष गुणांक	32. Area of influence	प्रभाव क्षेत्र
4. Absolute deviation	निरपेक्ष विचलन	33. Array	व्यूह
5. Absolutely unbiased estimator	निरपेक्ष अनभिनत आकलक	34. Artificial fertilizer	कृत्रिम उर्वरक
6. Acute symptom	तीव्र लक्षण	35. Artificial population	कृत्रिम समष्टि
7. Addition constant	अतिरिक्त स्थिरांक	36. Artificial vector	कृत्रिम सदिश
8. Adjacent	निकटवर्ती, संलग्न	37. Ascending order	आरोही क्रम
9. Adjustable	समायोज्य	38. Back coordination	पश्च समन्वयन
10. Admissible test	ब्राह्य परीक्षण	39. Ballistic test	प्राक्षेपिक परीक्षण
11. Agrarian	कृषि-भूमि सम्बन्धी	40. Characteristic equation	अभिलक्षणक —समीकरण
12. Agrestis	वन्य	41. Circular frequency	आवर्ती बारंबारता
13. Agricultural climatology	कृषि-जलवायु विज्ञान	42. Climate factor	जलवायवी कारक
14. Agricultural ecology	कृषि पारिस्थितिकी	43. Climate phenomena	जलवायवी घटक
15. Agricultural meteorological observatory	कृषि-मौसम प्रेक्षणशाला	44. Compatibility equation	सुसंगत समीकरण
16. Agricultural practice	कृषि-पद्धति	45. Compatible acceleration	सुसंगत त्वरण
17. Agrobiology	कृषि जैविकी	46. Complementary factor	पूरक कारक
18. Agroecology	कृषि पारिस्थितिकी	47. Complement set	पूरक समुच्चय
19. Agrostology	घास-विज्ञान	48. Complete orthogonal system	पूर्ण लांबिक निकाय
20. Algebraical equation	बीजीय समीकरण	49. Complex differentiation	सम्मिश्र अवकलन
21. Analogous	सदृश्य, अनुरूप, समवृत्ति	50. Complex function	सम्मिश्र फलन
22. Analogue computation	अनुरूप अभिकलन	51. Consistent estimate	संगत आकल
23. Analogue computer	अनुरूप कम्प्यूटर, अनुरूप अभिकलित्र	52. Degree of precision	परिशुद्धता की मात्रा
24. Analysis of variance	प्रसरण-विश्लेषण	53. Deterministic	निर्धारणात्मक
25. Analytical	विश्लेषणात्मक, विश्लेषिक	54. Disjoint	असंयुक्त
26. Analytical regression	विश्लेषिक समाश्रयण	55. Double correlation coefficient	द्वि-सहसंबंध गुणांक
27. Analytical vector	विश्लेषिक सदिश	56. Factorial design	बहु-उपादानी अभिकल्पना
28. Applied mathematics	अनुप्रयुक्त गणित	57. Factorial experiment	बहु-उपादानी प्रयोग
29. Applied statistics	अनुप्रयुक्त सांख्यिकी	58. Germplasm	जननद्रव्य
		59. Heteroscedasticity	विषम विचालिता



60. Identical eigen value	अभिन्न अभिलक्षणिक मान	79. Optimal order	इष्टतम क्रम
61. Identification	अभिनिर्धारण, पहचान	80. Order of stratification	संस्तरण—क्रम
62. Illusory association	आभासी साहचर्य	81. Parallelogram	समानान्तर चतुर्भुज
63. Implicit	अस्पष्ट	82. Parameter value	प्राचल मान
64. Incidental sampling	आनुसांगिक प्रतिचयन	83. Parametrization	प्राचलीकरण
65. Independent random variable	स्वतन्त्र यादृच्छिक चर	84. Percentage error	प्रतिशत त्रुटि
66. Inductive assumption	आगमनिक कल्पना	85. Point of observation	प्रेषण बिन्दु
67. Integrated Interval	समाकलित अन्तराल	86. Predicated variable	निर्धारक चर
68. Kinematical acceleration	शुद्धगतिकीय त्वरण	87. Predicted value	प्रागुक्त मान
69. Likelihood ratio test	संभावित अनुपात परीक्षण	88. Probability distribution	प्रायिकता बंटन
70. Linear approximation	रैखिक सन्निकटन	89. Quadratic response	द्विघाती अनुक्रिया
71. Macronutrient	बृहतपोषक	90. Random segregation	यादृच्छिक विसंयोजन
72. Mathematical analysis	गणितीय विश्लेषण	91. Ratio estimator	अनुपात आकलक
73. Micro-environment	सूक्ष्म—परिस्थिति, सूक्ष्म वातावरण	92. Significance level	सार्थकता स्तर
74. Non-dimensional	अविम, अविमीय	93. Significant correlation	सार्थक सहसम्बन्ध
75. Non-linear integral equation	अरैखिक समाकल समीकरण	94. Statistical population	सांख्यिकीय समष्टि
76. Non-variant	अचर	95. True distance	यथार्थ दूरी
77. Objective sampling	वस्तुनिष्ठ प्रतिचयन	96. Unsymmetrical	असममित
78. Observed regression	प्रेक्षित समाश्रयण	97. Variable acceleration	चर त्वरण
		98. Variance matrix	प्रसरण आव्यूह
		99. Weighting function	भारण फलन
		100. Zero dimensional set	शून्यविम समुच्चय

(संकलन : ब्रह्मजीत गहलौत एवं ऊषा जैन)

फार्म संबंधी सभी समस्याओं पर विशेषज्ञ की सलाह के लिये प्रातः 6:00 बजे से रात 10:00 बजे तक सभी सातों दिन टोल फ्री नम्बर 1551 पर डायल करें।



हर कदम, हर डगर

किसानों का हमसफर

भारतीय कृषि अनुसंधान परिषद

*Agrisearch with a human touch*