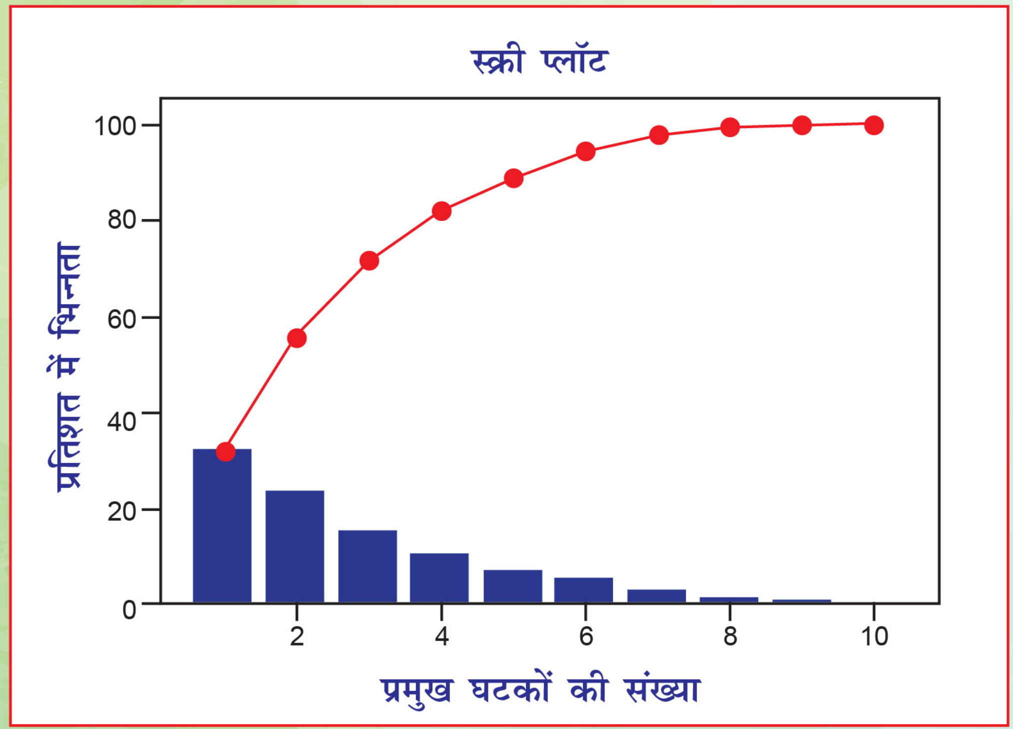
सांख्यिकी-विमर्श

2023

अंक

19



भा.कृ.अनु.प.-भारतीय कृषि सांख्यिकी अनुसंधान संस्थान
लाइब्रेरी एवेन्यू, पूसा, नई दिल्ली - 110012
https://iasri.icar.gov.in

सांख्यिकी-विमर्श

2023

अंक

19

संपादक मंडल	
प्रवीण आर्य	अध्यक्ष
मो. समीर फारुकी	सदस्य
सुशील कुमार सरकार	
सुकान्त दाश	
राजीव रंजन कुमार	
मनोज कुमार	
देवेन्द्र कुमार	

भाकृअनुप
ICAR

भा.कृ.अनु.प. - भारतीय कृषि सांख्यिकी अनुसंधान संस्थान

लाइब्रेरी एवेन्यू, पूसा, नई दिल्ली – 110012

https://iasri.icar.gov.in

सांख्यिकी–विमर्श

2023

अंक

19

द्वारा प्रकाशित

निदेशक

भा.कृ.अनु.प.–भारतीय कृषि सांख्यिकी अनुसंधान संस्थान

लाइब्रेरी एवेन्यू, पूसा, नई दिल्ली – 110012

मुद्रण : जुलाई 2024

रूपरेखा एवं मुद्रणः नक्षत्र आर्ट, ए 31, बेसमेंट, नारायणा इण्डस्ट्रिल एरिया फेस–2, नई दिल्ली–110028 फोन. +91-11-45636883

निदेशक की कलम से ....

भा.कृ.अनु.प.– भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, कृषि के तीन महत्वपूर्ण विषयोंः कृषि सांख्यिकी, संगणक अनुप्रयोग एवं कृषि जैव सूचना में अनुसंधान के साथ–साथ शिक्षण एवं प्रशिक्षण में भी अपना बहुमूल्य योगदान दे रहा है। संस्थान हिन्दी भाषा के प्रोत्साहन में भी अग्रणी भूमिका निभाता रहा है और राजभाषा की नीतियों को सुचारु रूप से कार्यान्वित करता रहा है। इसी क्रम में प्रस्तुत पत्रिका सांख्यिकी विमर्श के प्रकाशन का प्रारंभ, वर्ष 2005–06 में किया गया था। इस प्रकार सांख्यिकी विमर्श का यह उन्नीसवाँ अंक है। इस अंक को आप सुधी पाठकों के समक्ष प्रस्तुत करते हुए मुझे अपार प्रसन्नता हो रही है। बीते वर्षों के साथ इस पत्रिका की गुणवत्ता में लगातार बढ़ोतरी हो रही है, यह हमारे लिए गौरव का विषय है। यह पत्रिका सांख्यिकी/विज्ञान जैसे कठिन विषय को हिन्दी भाषा में प्रस्तुत करके उसकी जटिलता को सुगम बनाने का एक सशक्त माध्यम है।

पत्रिका के इस अंक मे हर वर्ष की भांति इस वर्ष भी संस्थान के कीर्ति स्तम्भ, संस्थान में चल रहे राजभाषा से संबंधित कार्यों की जानकारी के अतिरिक्त कृषि सांख्यिकी, कृषि में संगणक के अनुप्रयोग एवं कृषि जैव सूचना से संबंधित लेख सम्मिलित है। इन लेखों में विज्ञान के अति आधुनिक तकनीक यथा कृषि में ड्रोन का उपयोग से संबंधित लेखों को भी शामिल किया गया है। इस पत्रिका में हिन्दी साहित्य को ध्यान में रखते हुए कुछ मौलिक रोचक कविताओं को भी सम्मिलित किया गया है।

मैं पत्रिका के प्रकाशन के लिए उन सभी लेखकों की सराहना करता हूँ जिन्होंने इस पत्रिका के प्रकाशन हेतु अपने लेख देकर हमारे इस प्रयास को सफल बनाने में मदद की। इसके अतिरिक्त मैं संपादक मंडल के अध्यक्ष डॉ. प्रवीण आर्य, एवं अन्य सदस्यों डॉ. मो. समीर फारूकी, डॉ. सुशील कुमार सरकार, डॉ. सुकान्त दाश, डॉ. राजीव रंजन कुमार, श्री मनोज कुमार तथा श्री देवेन्द्र कुमार का आभार व्यक्त करता हूँ जिनके अथक प्रयासों से यह पत्रिका इस रूप में आपके समक्ष प्रस्तुत है। आशा ही नहीं अपितु मुझे पूर्ण विश्वास है कि पत्रिका का यह अंक पाठकों के लिए रोचक, उपयोगी एवं ज्ञानवर्धक साबित होगा। इस पत्रिका के माध्यम से पाठकगण सांख्यिकी, संगणक एवं जैव सूचना जैसे कठिन विषय को आसानी से समझ पाएंगे। इस पत्रिका के आगामी अंकों में सुधार के लिए आपके सुझाव वांछनीय है।

राजेन्द्र प्रसाद
निदेशक

सम्पादकीय

सांख्यिकी वह विज्ञान है जो डेटा के संग्रह, विश्लेषण, व्याख्या, और प्रस्तुति से संबंधित है। यह संख्यात्मक तथ्यों और सूचनाओं को समझने और विश्लेषण करने की क्षमता प्रदान करता है। आँकड़े हमारे दैनिक जीवन में विभिन्न क्षेत्रों में महत्वपूर्ण भूमिका निभाते हैं, जैसे कि अर्थशास्त्र, समाजशास्त्र, चिकित्सा, खेल, और व्यवसाय। आँकड़ों की मदद से वैज्ञानिक विभिन्न परिकल्पनाओं का परीक्षण कर सकते हैं और प्राप्त परिणामों का विश्लेषण कर सकते हैं, और भविष्य की संभावनाओं का अनुमान लगा सकते हैं। इसके अलावा, आँकड़े नीतिगत निर्णय लेने में सहायक होते हैं, जिससे समाज और संगठनों की प्रगति और विकास सुनिश्चित किया जा सकता है। इस प्रकार, आँकड़े एक महत्वपूर्ण उपकरण हैं जो जानकारी को व्यवस्थित और उपयोगी बनाते हैं।

पत्रिका सांख्यिकी विमर्श का उन्नीसवाँ अंक प्रस्तुत करते हुए हमें अत्यंत हर्ष हो रहा है। इस सफलता के पीछे हमारे पाठकों और सभी लेखकों का योगदान और कौशल निहित है। हमें यह जानकर प्रसन्नता होती है कि सांख्यिकी विमर्श के लेखों को व्यापक सराहना मिल रही है। हालांकि, यह हमारे लिए केवल एक शुरुआत है, क्योंकि हम सांख्यिकी विमर्श को और भी ऊँचाइयों पर ले जाना चाहते हैं। इस दिशा में हमारे प्रयास निरंतर जारी हैं, और हमें पूर्ण विश्वास है कि आने वाले अंकों में यह पत्रिका और भी अधिक रोचक और ज्ञानवर्धक सिद्ध होगी।

इस अंक में सांख्यिकी एवं कृषि साहित्य के विभिन्न आयामों पर आधारित लेख प्रस्तुत किए गए हैं जिसमे

सुडोकू वर्ग अभिकल्पनाएं एवं उनके अनुप्रयोग, प्रतिदर्श सर्वेक्षण ड्रोन का कृषि में योगदान, गूगल अर्थ इंजन की सहायता से सतह के तापमान का स्थानिक मानचित्रण, प्रोटीन संरचना प्रिडिक्शन, बेजीयन फ्रेमवर्क का उपयोग एवं अन्य आयामों से संबंधित लेख प्रस्तुत है। हमें आशा ही नहीं अपितु पूर्ण विश्वास है कि सांख्यिकी विमर्श का यह अंक पाठकों के लिए अत्यंत ज्ञानवर्धक और रोचक सिद्ध होगा।

सांख्यिकी विमर्श का मुख्य खंड अनुसंधान खंड है जिसमें विज्ञान की जटिलताओं को सरल भाषा में लेखों के माध्यम से समझाया गया है। इसके अलावा पत्रिका में संस्थान के कीर्ति स्तंभ, संस्थान की राज्य भाषा यात्रा, कविताएं भी शामिल हैं। इस प्रकार हम कह सकते हैं कि यह पत्रिका विज्ञान एवं साहित्य का एक अनुपम संगम है जो इसको एक मौलिक अतुलनीय पत्रिका बनाती है। संपादक मंडल "सांख्यिकी विमर्श" के सभी लेखकों के योगदान के लिए हृदय से आभारी है एवं आशा करता है कि भविष्य में भी उनका अमूल्य सहयोग जारी रहेगा।

इस पत्रिका को समृद्ध एवं रुचिकर बनाने के लिए पाठकों के सुझाव अपेक्षित हैं। आग्रह है कि कृपया अपनी प्रतिक्रिया हमें अवश्य भेजें ताकि हम इस पत्रिका को और अधिक प्रभावशाली ज्ञानवर्धक एवं रोचक बना सकें। इस पत्रिका के निरंतर विकास की कामना के साथ।

प्रवीण आर्य
अध्यक्ष, संपादक मंडल

अनुक्रमणिका

संस्थान के कीर्ति स्तम्भ • प्रोफेसर आलोक डे	1
अनुसंधान खण्ड	
मल्टीस्ट्रैटम प्रतिक्रिया सतह अभिकल्पनाओं पर एक अवलोकन • राहुल बनर्जी, भारती, दीपक सिंह, पंकज दास, अंकुर विश्वास, कौस्तब आदित्य एवं राजू कुमार	5
प्रतिदर्श सर्वेक्षणों में परिमित समष्टि औसत के अनुमानकों का एक बेहतर वर्ग • दीपक सिंह, राजू कुमार, राहुल बनर्जी, पंकज दास, भारती, अंकुर बिश्वास एवं कौस्तव आदित्य	13
ड्रोन प्रौद्योगिकीः कृषि में उच्च उत्कृष्टि का संवर्धन • पंकज दास, भारती, राहुल बनर्जी, दीपक सिंह एवं राजू कुमार	19
कृषि आवश्यकताओं के व्यख्यान हेतु किसान कॉल सेंटर के किसानों के प्रश्नों का गहन विश्लेषण • समर्थ गोदारा, राम स्वरूप बाना, श्रुति गोदारा, राजेंद्र प्रसाद, सुदीप मारवाहा राजू कुमार एवं दीपक सिंह	24
न्यूट्रोसोफ्रिक आंकड़ों हेतु सांख्यिकीय तकनीकें • आशुतोष दलाल, सिनी वर्गीस, मोहम्मद हारुन एवं देवेन्द्र कुमार	33
अभिकल्पित परीक्षणों का गैर–प्राचलिक विश्लेषण • नीतू आर एस, सिनी वर्गीज, मो. हारून, आशुतोष दलाल, अनिंदिता दत्ता एवं देवेंद्र कुमार	39
महाराष्ट्र के पुणे जिले के लिए गूगल अर्थ (Google Earth) इंजन का इस्तेमाल करके भूमि की सतह के तापमान का स्थानिक मानचित्रण • नोबिन चन्द्र पॉल, नव्याश्री पोन्नगंटि, राहुल बनर्जी, धनन्जय डी. नांगारे एवं तौकीर अहमद	48
दीर्घावधि उर्वरक परीक्षणः एक अवलोकन • सुशील कुमार सरकार, अनिंदिता दत्ता एवं रवि वंजारी	51
सुडोकू वर्ग अभिकल्पनाएँ एवं उनके अनुप्रयोग • नेहाताई डब्ल्यू. अगाशे, सिनी वर्गीस, मो० हारून, आशुतोष दलाल एवं देवेंद्र कुमार	59
प्रोटीन संरचना प्रेडिक्शनः टूल्स और तकनीकें • सुधीर श्रीवास्तव, स्नेहा मुर्मू, दीपा भट्ट, मयंक रश्मि, कृष्ण कुमार चतुर्वेदी, मोहम्मद समीर फारूकी एवं गिरीश कुमार झा	65
बेसियन फ्रेमवर्क का उपयोग करके सिस और ट्रांस–एक्सप्रेशन क्वांटिटेटिव ट्रेट लोसाइ (ईक्यूटीएल) की पहचान • हिमाद्री शेखर रॉय, अमृत कुमार पॉल, रंजित कुमार पॉल, प्रकाश कुमार, एमडी यासीन, एवं एस.पी.सिंह	72

समय श्रृंखला पूर्वानुमान के लिए इंट्यूशनिस्टिक फ़ज़ी आधारित ऑटोरेग्रेसिव इंटीग्रेटेड मूविंग एवरेज (फरिमा) मॉडल • अनिता सरकार, मो. यासीन, रंजित कुमार पॉल, हिमाद्री शेखर रॉय, प्रकाश कुमार, अंकित कुमार सिंह, अमृत कुमार पॉल, एवं एस. पी. सिंह	83
जैव सूचना विज्ञान में बृहत आंकड़ों विश्लेषण • प्रकाश कुमार, रंजीत कुमार पॉल, हिमाद्री शेखर रॉय, मो. येसीन, राजू कुमार, दीपक सिंह, एवं अमृत कुमार पॉल	88
विभिन्न स्थितियों में संशोधित प्रसरण विश्लेषण द्वारा आनुवंशिकता का आकलन • रंजित कुमार पॉल, मो. यासीन, हिमाद्री शेखर रॉय, प्रकाश कुमार, अमृत कुमार पॉल एवं एस.पी.सिंह	99
राजभाषा खण्ड	
संस्थान की राजभाषा यात्राः वर्ष 2023 हिन्दी एकक	109
कविताएं	112
हिन्दी कार्यशाला – 2023	118
हिन्दी पखवाड़ा – 2023	122

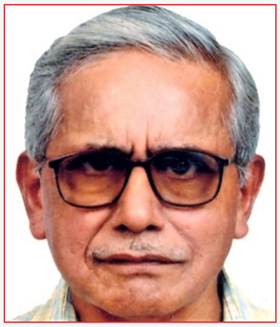
संस्थान के कीर्ति स्तम्भ

प्रोफेसर आलोक डे

प्रोफेसर आलोक डे का जन्म 12 सितंबर 1945 को हुआ था। कहावत है कि पूत के पाँव पालने मे नजर आ जाते है, इसी कहावत को शुरुआत से प्रोफेसर आलोक डे ने चरितार्थ किया। प्रोफेसर डे का शैक्षणिक कार्यकलाप बहुत शानदार रहा है। उन्होंने 1964 में गणित में स्नातकोत्तर की उपाधि प्राप्त की, 1966 में स्नातकोत्तर की उपाधि प्रथम स्थान के साथ एवं 1969 में विद्या वाचस्पति (पीएचडी) की उपाधि भा.कृ.अनु.प.–भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली से प्राप्त की। परीक्षण अभिकल्पना पर उनकी विद्या वाचस्पति की उपाधि आज भी अत्यन्त प्रासंगिक है।

प्रोफेसर आलोक डे ने अपने व्यावसायिक जीवन की शुरुआत उत्तर प्रदेश कृषि विश्वविद्यालय में एक संक्षिप्त कार्यकाल के साथ की थी, प्रोफेसर डे 1970 नें एक संकाय सदस्य के रूप में भा.कृ.अनु.प.–भारतीय कृषि सांख्यिकी अनुसंधान संस्थान में कार्य ग्रहण किया और 1977 में महज बत्तीस वर्ष की आयु में वरिष्ठ प्रोफेसर बन गए। इसके बाद, 1989 में, वह एक प्रोफेसर के रूप में भारतीय सांख्यिकी संस्थान चले गए, और 2007 में अपनी औपचारिक सेवानिवृत्ति तक इस पद पर बने रहे। उन्होंने संयुक्त राज्य अमेरिका, कनाडा और ताइवान सहित विदेशों में संस्थानों में वरिष्ठ शैक्षणिक पदों पर भी कार्य किया। वह भारतीय राष्ट्रीय विज्ञान अकादमी के वरिष्ठ वैज्ञानिक के रूप में भारतीय सांख्यिकी संस्थान से भी जुड रहे।

प्रोफेसर आलोक डे विभिन्न भारतीय अकादमिक संस्थाओं में मजबूत स्तम्भ के रूप में रहे, वे भारतीय राष्ट्रीय विज्ञान अकादमी (आईएनएसए), नेशनल एकेडमी ऑफ साइंसेज के फेलो सदस्य भी रहे और अंतर्राष्ट्रीय सांख्यिकी संस्थान के निर्वाचित सदस्य भी रहे। सांख्यिकी के विभिन्न क्षेत्रों में मौलिक अनुसंधान में एक वैश्विक रत्न थे, सुरुचिपूर्ण ढंग से लिखित पुस्तकों के माध्यम से विचारों के उत्कृष्ट प्रसार के लिए और प्रभावशाली संपादकीय कार्य के लिए भी सतत प्रयत्नशील रहे।



प्रोफेसर डे का शोध कार्य, पिछले पाँच दशकों की अथक मेहनत है, जो ना केवल सांख्यिकी बल्कि गणित से संबंधित कई क्षेत्रों को प्रभावित करता है, जो उनकी अद्‌भुत बहुमुखी प्रतिभा और विषय में समझ की गहराई को दर्शाता है। उनका कार्यक्षेत्र में विशेषतः परीक्षण अभिकल्पना, सर्वेक्षण प्रतिदर्श, सांयोगिक सिद्धांत और रैखिक बीजगणित शामिल हैं। प्रोफेसर डे के शोध का विशिष्ट पहलू सबसे चुनौतीपूर्ण समस्याओं को सरल समाधानों के साथ प्रस्तुत करने में है, जो हमारी समझ को काफी हद तक बढ़ाती है। उनके शोध की व्यापकता और गहराई को इस संक्षिप्त लेख में समाना असंभव है।

इसलिए, केवल कुछ उदाहरणों के साथ इसका वर्णन प्रस्तुत हैः

परीक्षण अभिकल्पना में, प्रोफेसर डे ने फैक्टोरियल अभिकल्पना, वैराइटल ब्लॉक जैसे विविध क्षेत्रों में पथ–प्रदर्शक योगदान दिया है। पंक्ति–स्तंभ अभिकल्पना, वेइंग अभिकल्पना, प्रतिक्रिया सतह अभिकल्पना, क्रॉसओवर अभिकल्पना, जैविक परख और डाइ–एलिल क्रॉस के लिए अभिकल्पना, और इसी तरह, व्यावहारिक रूप से महत्वपूर्ण लेकिन गणितीय रूप से कठिन असममित ऑर्थोगोनल भिन्नात्मक तथ्यात्मक अभिकल्पनाओं और संबंधित ऑर्थोगोनल सारणियों पर उनका शोध,

तत्काल प्रयोज्यता के साथ विशेष रूप से औद्योगिक परीक्षण और गुणवत्ता नियंत्रण क्षेत्र में सैद्धांतिक लालित्य को मिश्रित करता है।

विगत वर्षों में, उन्होंने आंशिक तथ्यात्मक अभिकल्पनाओं को प्राप्त करने की कठिन समस्या पर काम किया जब कुछ अन्योन्यक्रियाएं (इंटरैक्शन) महत्वपूर्ण होती हैं और परिमित प्रोजेक्टिव ज्यामिति से उपकरणों के उपयोग के माध्यम से एक सरल समाधान प्रस्तुत किया। प्रोफेसर डे की एक और खोज, नॉनऑर्थोगोनल ब्लॉकिंग के तहत इष्टतम मुख्य प्रभाव अभिकल्पनाओं पर एक नये क्षेत्र का मार्ग प्रशस्त किया है। परीक्षण अभिकल्पना में उनके अन्य योगदानों में नेस्टेड संरचना के साथ अपूर्ण ब्लॉक परीक्षण अभिकल्पनाओं का एक नया वर्ग शामिल है, जिसका अध्ययन सांख्यिकीय और सांयोगिक दृष्टि से बड़े पैमाने पर किया गया है। कुछ पंक्ति–स्तंभ अभिकल्पनाओं की सार्वभौमिक अनुकूलता और गैर–अनुकूलता पर उनका काम अपने विपरीत निष्कर्षों के लिए प्रसिद्ध है, जबकि जैविक परख और इष्टतम वेइंग अभिकल्पना के लिए इष्टतम अभिकल्पनाओं पर उनके शुरुआती परिणाम अब संबंधित क्षेत्रों में उदाहरणीय हैं।

प्रोफेसर डे ने अन्य क्षेत्रों में भी बहुत उल्लेखनीय योगदान दिया जैसे कि असमान संभाव्यता प्रतिदर्श सर्वेक्षण, सशर्त अपेक्षाओं के माध्यम से लक्षण वर्णन की समस्याएं, सामरिक विन्यास, तिरछे रेंज प्रमुख आव्यूह, और इसी तरह, आव्यूह बीजगणित में कुछ परिणामों के उनके सांख्यिकीय प्रमाण सांख्यिकीय समुदाय के लिए वरदान हैं। उनके निष्कर्षों की गहराई को देखते हुए, यह कोई आश्चर्य नहीं है कि प्रोफेसर डे ने सांख्यिकी की सबसे अच्छी पत्रिकाओं में बड़े पैमाने पर प्रकाशित किया गया है और उनके पत्रों को कई उद्धरण प्राप्त हुए हैं। प्रोफेसर डे के काम का कोई भी वर्णन उनकी पुस्तकों के विशेष उल्लेख के बिना पूरा नहीं हो सकता है, जो वास्तविक ग्रंथ हैं। भिन्नात्मक तथ्यात्मक अभिकल्पनाओं पर उनकी पुस्तक पाठ प्रदर्शक का काम करती है। प्रोफेसर डे ने क्रॉसओवर अभिकल्पनाओं के कठिन लेकिन व्यावहारिक रूप से उपयोगी विषय पर एक पुस्तक लेखन भी किया।

प्रोफेसर डे ने 2002–05 के दौरान संख्या, इंडियन जर्नल ऑफ स्टैटिस्टिक्स के संपादन की जिम्मेदारी भी निभाई। उनके प्रख्यात नेतृत्व में और उनके श्रमसाध्य प्रयासों के माध्यम से, पत्रिका ने नई ऊंचाइयां प्राप्त की है। वह सांख्यिकी और अनुप्रयोग के अध्यक्ष संपादक और सांख्यिकी, कंप्यूटर और अनुप्रयोग सोसायटी के कार्यकारी अध्यक्ष रहे हैं।

एक उत्कृष्ट शोधकर्ता होने के बावजूद, प्रोफेसर डे हमेशा एक शिक्षक के रूप में अपनी जिम्मेदारियों के प्रति सजग रहे हैं। छात्रों की लगातार पीढ़ियों ने उनके अंदर के महान शिक्षक के संरक्षण में खुद को लाभान्वित किया। एक व्यक्ति के रूप में, प्रोफेसर डे शास्त्रीय भारतीय मूल्यों के प्रतीक हैं। उनके सहयोगी हमेशा न केवल उनके ज्ञान बल्कि उनकी ईमानदारी और अखंडता से भी मंत्रमुग्ध रहे हैं । वह विशेष रूप से युवा सहयोगियों के प्रति दयालू रहे है।

प्रोफेसर डे द्वारा 129 शोध पत्रों एवं पाँच पुस्तकोंः ए. डे (1985), ऑर्थोगोनल फ्रैक्शनल फैक्टोरियल डिजाइन; जॉन विली; ए. डे (1986), ब्लॉक डिजाइन का सिद्धांत; जॉन विली/हेलस्टेड प्रेस; ए. डे एवं आर. मुखर्जी (1999), भिन्नात्मक तथ्यात्मक योजनाएं; जॉन विली; एम. बोस एवं ए. डे (2009), इष्टतम क्रॉसओवर डिजाइन, विश्व वैज्ञानिक; ए. डे (2010) एवं अपूर्ण ब्लॉक डिजाइन, हिंदुस्तान बुक एजेंसी/वर्ल्ड साइंटिफिक का लेखन किया गया।

प्रोफेसर डे को 2010 में सांख्यिकी के क्षेत्र में आजीवन योगदान के लिए सांख्यिकी एवं कार्यक्रम कार्यान्वयक मंत्रालय, भारत सरकार द्वारा प्रतिष्ठित प्रोफेसर पी.वी. सुखात्मे राष्ट्रीय पुरस्कार से सम्मानित किया गया। प्रोफेसर डे का निधन 10 फरवरी 2020 को नई दिल्ली में हुआ। सांख्यिकी के क्षेत्र में उनका महत्वपूर्ण योगदान सदैव स्मरण रहेगा।

सांख्यिकी-विमर्श
2023

अंक
19

अनुसंधान खण्ड

मल्टीस्ट्रैटम प्रतिक्रिया सतह अभिकल्पनाओं पर एक अवलोकन

राहुल बनर्जी, भारती, दीपक सिंह, पंकज दास, अंकुर विश्वास, कौस्तब आदित्य एवं राजू कुमार
भा.कृ.अनु.प. – भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली

सार

प्रायोगिक अभिकल्पनाओं का मूल सिद्धांत यह है कि उपचारों को प्रायोगिक इकाइयों पर पूरी तरह से यादृच्छिक रूप से वितरित किया जाना चाहिए। हालांकि, व्यावहारिक रूप में इकाइयों में अक्सर कोई संरचना होती है, जिससे ब्लॉकिंग की आवश्यकता प्रकट होती है। कई औद्योगिक प्रयोगों में कुछ ऐसे कारक शामिल होते हैं जिनके स्तर दूसरों की तुलना में निर्धारित करना कठिन होता है। । अगर इसके अतिरिक्त, कुछ कारक इकाइयों के बीच बदलना अन्य कारकों की तुलना में अधिक कठिन होता है, तो एक बहु–स्तरीय संरचना स्वाभाविक रूप से उत्पन्न होती है। इनका सामना करने का सबसे अच्छा तरीका यह है कि प्रयोग को ध्यानपूर्वक एक विभाजित–खंड (split-plot) या अधिक सामान्य रूप से एक मल्टी–स्ट्रेटम (multi-stratum) अभिकल्पना के रूप में योजना बनाया जाए। 2001 के बाद साहित्य में विभिन्न विभाजित–खंड प्रकार के प्रतिक्रिया सतह अभिकल्पना को निर्माण करने के लिए कई विभिन्न उपाय प्रस्तावित किए गए हैं, जिससे प्रयोगकर्ताओं को सामग्रियों का बेहतर उपयोग करने की सुविधा मिली है, जैसे कि क्लासिकल संतुलित अभिकल्पना से अधिक कुशल अभिकल्पना। इनमें से एक उपाय है, स्तर–द्वारा–स्तर रणनीति। सबसे उच्च स्तर से शुरू होकर, डिजाइन को स्तर–द्वारा–स्तर निर्मित किया जाता है। प्रत्येक स्तर में, उस स्तर में लागू किए जाने वाले कारकों के लिए एक पूर्व–निर्धारित उपचार सेट को उस स्तर में उन इकाइयों के लिए लगभग अपर्याप्त किया जाता है, जो उच्च स्तरों में होती हैं, जिससे सभी प्रभावों को अनुमानित किया जा सकता है। इस विधि की उपयोगीता को दिखाने के लिए तीन उदाहरण दिए गए हैं एवं और इन्हें अंतिम डिजाइन के संबंध की जांच करने के लिए भी उपयोग किया गया है।

मुख्य शब्दः ब्लॉक संरचना, फैक्टोरियल डिजाइन, औद्योगिक प्रयोग, अवशिष्ट अधिकतम संभावना (आरईएमएल), स्प्लिट–प्लॉट डिजाइन।

परिचय

प्रतिक्रिया सतह अध्ययन आमतौर पर दो कारणों से किया जाता है।

- प्रतिगमन मॉडल के माध्यम से प्रतिक्रिया और इनपुट कारकों के बीच वास्तविक संबंध का अनुमान लगाना।
- इनपुट कारकों के लिए सेटिंग्स का पता लगाने जिनके लिए प्रतिक्रिया इष्टतम (अधिकतम या न्यूनतम) है।

प्रतिक्रिया सतह पद्धति (रेस्पॉन्स सरफेस मेथडोलॉजी) एक ऐसी रणनीति है जिसका उपयोग इन उद्देश्यों को प्राप्त करने के लिए प्रभावी ढंग से किया जा सकता है। प्रतिक्रिया सतह पद्धति प्रक्रियाओं के विकास, सुधार और अनुकूलन के लिए उपयोगी सांख्यिकीय और गणितीय तकनीकों का एक संग्रह है, नए उत्पादों के डिजाइन, विकास एवं निर्माण के साथ–साथ मौजूदा उत्पाद डिजाइनों के सुधार में भी इसका महत्वपूर्ण अनुप्रयोग है। प्रतिक्रिया सतह पद्धति अ मात्रात्मक कारकों के साथ

तथ्यात्मक प्रयोगों के लिए उपयोगी है, जिन्हें कारकों के स्तर के संयोजन को निर्धारित करने के लिए लिया जाता है, जिस पर कुछ अर्थों में प्रतिक्रिया को अनुकूलित करने के लिए प्रत्येक मात्रात्मक कारक को सेट किया जाना चाहिए। मान लीजिए कि अ स्वतंत्र चर निरूपित हैं, तथा प्रतिक्रिया चर ल है, और छ अवलोकन हैं। प्रतिक्रिया को इनपुट कारकों के एक फंक्शन के रूप में पोस्ट किया गया है, जैसे,

$$y_u = \varphi(x_{1u}, x_{2u}, ..., x_{vu}) + e_u,\ u = 1,2,\ldots,N$$

y_u - u^{th} उपचार संयोजन से प्रतिक्रिया, x_{iu} - u^{th} उपचार संयोजन में आईटीएच कारक का स्तर (i= 1,2,…,v)। N अंक द्वारा ऐसे निरूपित किया जाता है:

$$\begin{matrix} x_{11} & x_{21} & \cdots & x_{i1} & \cdots & x_{v1} \\ x_{12} & x_{22} & \cdots & x_{i2} & \cdots & x_{v2} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ .. & . & . & . & . & . \\ x_{u1} & x_{2u} & \cdots & x_{iu} & \cdots & x_{vu} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ x_{1N} & x_{2N} & \cdots & x_{iN} & \cdots & x_{vN} \end{matrix}$$

फंक्शन φ उस रूप का वर्णन करता है जिसमें प्रतिक्रिया और इनपुट चर संबंधित हैं और e_u, u^{th} अवलोकन से जुड़ी यादृच्छिक त्रुटि है। फंक्शन को अक्सर प्रयोगात्मक क्षेत्र के भीतर चर में उपयुक्त डिग्री के बहुपद द्वारा अनुमानित किया जाता है। विचरण के सामान्य विश्लेषण के माध्यम से फिट किए गए बहुपद की पर्याप्तता का परीक्षण किया जाता है। बहुपद जो वास्तविक खुराक–प्रतिक्रिया संबंध का पर्याप्त रूप से प्रतिनिधित्व करते हैं, प्रतिक्रिया सतह कहलाते हैं। वे अभिलापनाएँ जो प्रतिक्रिया सतहों की फिटिंग की अनुमति देते हैं और उनकी पर्याप्तता के परीक्षण के लिए एक उपाय प्रदान करते हैं, प्रतिक्रिया सतह अभिकल्पना कहलाते हैं।

पहला ऑर्डर आर एसः

$$y_u = \beta_0 + \beta_1 x_{1u} + \beta_2 x_{2u} + \ldots + \beta_v x_{vu} + e_u$$

द्वितीय ऑर्डर आर एसः

$$y_u = \beta_0 + \sum_{i=1}^{v} \beta_i x_{iu} + \sum_{i=1}^{v} \beta_i x_{iu}^2 + \sum_{i=1}^{v-1} \sum_{i'=i+1}^{v} \beta_{ii'} x_{iu} x_{i'u} + e_u$$

मल्टीस्ट्रेटम प्रयोग

हाल के वर्षों में यह तेजी से पहचाना गया है कि कई, शायद अधिकांश, औद्योगिक प्रयोगों एवं कई प्रयोगशाला प्रयोगों में कुछ ऐसे कारक शामिल होते हैं जिनका स्तर दूसरों की तुलना में निर्धारित करना कठिन होता है। यह स्पष्ट है कि ऐसे स्थितियों से निपटने का सबसे अच्छा तरीका यह है कि जब प्रयोग को डिजाइन किया जाए तो कठिन–से–सेट कारकों को एक संरचित तरीके से रखा जाए, ताकि उनके स्तरों को प्रत्येक रन के लिए सेट करने की आवश्यकता न हो, बल्कि केवल कम बार ही सेट करना पड़े।। यदि केवल हार्ड–टू–सेट और आसान–से–सेट कारक हैं, तो यह एक (आमतौर पर नॉनऑर्थोगोनल) स्प्लिट–प्लॉट संरचना होता है। यदि बहुत–हार्ड–टू–सेट (वी.एच.एस), फेयरली–हार्ड–टू–सेट (एच.एस.) और आसान–से–सेट (ई.एस.) कारक हैं, तो यह एक स्प्लिट–स्प्लिट–प्लॉट संरचना होते हैं। आम तौर पर, कारकों में कठोरता के प्रत्येक स्तर को डिजाइन में ध्यान में रखा जाता है, एक स्ट्रेटम को परिभाषित करता है, जैसा कि अवरुद्ध के प्रत्येक स्तर तथा कम से कम दो स्तरों में कारकों के साथ ऐसे डिजाइनों को मल्टीस्ट्रैटम डिजाइन कहा जाता है।

मल्टीस्ट्रैटम प्रयोग वे हैं जिनमें उपचार कई कारकों के स्तरों के संयोजन होते हैं और प्रयोगात्मक इकाइयों के सेट होते हैं, जिनमें सभी में कम से कम एक कारक का समान स्तर होना चाहिए। सबसे आम उदाहरण स्प्लिट–प्लॉट डिजाइन है, लेकिन अधिक सामान्यीकृत संरचनाओं को तीन तरीकों से माना जा सकता है:

अधिक स्तर हो सकते हैं (जैसे स्प्लिट–स्प्लिट–प्लॉट्स, आदि)।

उपचार डिजाइन एक (दूसरे क्रम) बहुपद प्रतिक्रिया

सतह मॉडल फिट करने के लिए उपयुक्त होगा।

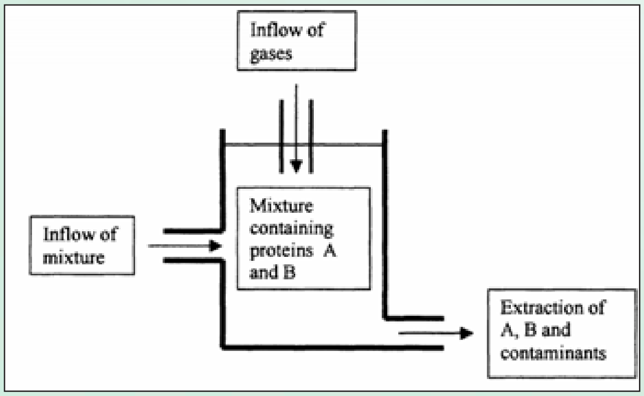
डिजाइन में प्रत्येक पूरे भूखंड में सबप्लॉट कारकों के संयोजन का पूरा सेट नहीं होगा।

एटकिंसन और डोनेव (1989) तथा कुक और नचत्सहेम (1989) ने डी–ऑप्टिमल ब्लॉकिंग के लिए एल्गोरिदम दिया। ट्रिंका और गिल्मर (2000) ने भारित माध्य दक्षता कारक मानदंड का उपयोग करके ब्लॉकिंग के लिए एक एल्गोरिदम दिया। गिल्मर और ट्रिंका (2000) ने अवरुद्ध आरएस डिजाइन से डेटा के विश्लेषण और प्रयोग के डिजाइन के निहितार्थ पर चर्चा की। लेटसिंगर एट अल. (1996) को ''द्वि–यादृच्छिकीकरण'' संरचना कहा जाता है, जो एक विभाजित कथानक संरचना है, लेकिन असंतुलित है। इन लेखकों ने, डेटा विश्लेषण के तरीकों की समीक्षा करने के बाद, इन दो–स्तरित संरचनाओं में कई मानक आरएस डिजाइनों की तुलना की। ये डिजाइन कृषि प्रयोगों के लिए बहुत उपयोगी हैं, लेकिन उपलब्ध डिजाइन अनुप्रयोग के कुछ अन्य क्षेत्रों के लिए अपर्याप्त रूप से लचीले हैं।

व्यावहारिक उदाहरण (गिल्मर एट अल. 2001)

एक बायोकेमिकल इंजीनियर ने चित्र 1 में दर्शाए गए निरंतर प्रोटीन–निष्कर्षण प्रक्रिया की समझ को बेहतर बनाने तथा अनुकूलित करने के लिए एक प्रयोग करने की योजना बनाई।

अन्य घटकों के अलावा, दो मूल्यवान प्रोटीन युक्त मिश्रण, पहले किण्वन और शुद्धिकरण प्रक्रियाओं से प्राप्त



चित्र.1: प्रोटीन निष्कर्षण प्रक्रिया का योजनाबद्ध प्रतिनिधित्व।

होता है। अध्ययन के तहत प्रक्रिया का उद्देश्य मिश्रण से दो प्रोटीनों को अलग करना है, और उद्देश्य दो प्रोटीनों की पैदावार तथा शुद्धता को अधिकतम करना है। ये उद्देश्य कुछ हद तक विरोधाभासी हैं, क्योंकि उपज को अधिकतम करना आम तौर पर शुद्धता को कम करता है और इसके विपरीत। इसलिए, डिजाइन चरण में, एक डिजाइन प्राप्त करने पर ध्यान केंद्रित करें जो दूसरे क्रम के बहुपद और उसके उप–मॉडल के मापदंडों का अच्छा अनुमान देगा। प्रयोगकर्ता ने पांच कारकों की पहचान की थी जिनसे प्रोटीन निष्कर्षण को प्रभावित करने की उम्मीद थी – मिश्रण के प्रवाह के लिए फीड स्थिति, फीड प्रवाह दर, गैस प्रवाह दर, प्रोटीन ए की सांद्रता, और प्रोटीन बी की सांद्रता। यह तय किया गया था कि प्रत्येक कारक के तीन स्तरों का उपयोग किया जाएगा। लगभग 20 दिनों तक प्रयोग किया जा सकता था, लेकिन फीड पोजिशन सेट करना उपकरण को खोलने और फिर से असेंबल करने की आवश्यकता होती थी, इसलिए इसे केवल एक बार प्रतिदिन किया जा सकता था। हालांकि, फीड पोजिशन को स्थिर रखकर, प्रतिदिन दो रन (अर्थात, अन्य कारकों की दो सेटिंग्स) किए जा सकते थे।

गैर-आर्थोगोनल मल्टीस्ट्रेटम संरचनाओं से डेटा का विश्लेषण

गैर–ऑर्थोगोनल मल्टीस्ट्रेटम डिजाइन के विश्लेषण का सामान्य रूप रैखिक मिश्रित मॉडल को फिट करते हैं, जिसमें अवशिष्ट अधिकतम संभावना (आरईएमएल) द्वारा यादृच्छिक प्रभावों और सामान्यीकृत न्यूनतम वर्गों द्वारा निश्चित प्रभावों का अनुमान लगाया जाता है। इस विश्लेषण को कई मानक सांख्यिकीय पैकेजों का उपयोग करके कार्यान्वित किया जा सकता है। मॉडल को इस प्रकार लिखा जा सकता है:

$$y = \sum_{i=1}^{s} Z_i \alpha_i + X\beta$$

जहां, y प्रतिक्रियाओं का एक $n \times 1$ वेक्टर है, α_i यादृच्छिक प्रभावों का मिश्रण1 वेक्टर है, $m_i = \coprod_{j=1}^{i} n_j$ Z_i के लिए एक

n×mi मैट्रिक्स संकेतक u_i है, x पी पैरामीटर रिस्पॉन्स सरफेस मॉडल के लिए ऑर्डर एनएक्सपी का डिजाइन मैट्रिक्स है तथा β पैरामीटरों का P×1 वेक्टर है और ऐसा माना जाता है $\boldsymbol{\alpha}_i \sim N(\mathbf{0}, \sigma_i^2 \mathbf{I}_{m_i})$ और α_i तथा α_j के लिए असंबद्ध $i \neq j$.

प्रतिक्रिया सतह प्रयोग में मुख्य रुचि प्रतिक्रिया सतह मॉडल यानी निश्चित प्रभाव मापदंडों β का आकलन करने में है, β का सामान्यीकृत लीस्ट स्क्वायर अनुमानक है ,

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}\mathbf{y}$$

जहाँ, **Σ**, y का प्रसरण–सहप्रसरण मैट्रिक्स है:

$$\Sigma = \sum_{i=1}^{s} \mathbf{Z}_i \mathbf{Z}'_i \sigma_i^2$$

$\hat{\boldsymbol{\beta}}$ का विचरण मैट्रिक्स है, $V(\hat{\boldsymbol{\beta}}) = (\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}$

परिणाम यह मानकर चलता है कि $\sigma = [\sigma_1, \ldots \sigma_s]$ ज्ञात है। में, $\hat{\boldsymbol{\beta}}$ अनुमानक में, σ आमतौर पर इसके REML अनुमानक σ द्वारा प्रतिस्थापित किया जाता है।

स्ट्रैटम दर स्ट्रैटम अभिकल्पना निर्माण पद्धति

अभिकल्पनाओं का निर्माण उच्चतम स्तर से निम्नतम स्तर तक किया जाता है। प्रत्येक स्ट्रैटम में, U_i, जिस पर कारक लागू होते हैं, चार चरणों का आयोजन किया जाता है, जो कि निम्नानुसार है:

1. सामान्य तरीके से स्ट्रेटम U_i में लागू कारकों के लिए एक उपचार अभिकल्पना चुनें। यह इन कारकों पर प्रक्षेपण के लिए एक अच्छा अभिकल्पना सुनिश्चित करता है।
2. 1 में चुने गए उपचार डिजाइन को बनाए रखते हुए, U_{i-1} में इकाइयों को ब्लॉकों के रूप में मानें और डिजाइन को सामान्य तरीके से ब्लॉकों में व्यवस्थित करें। यह सुनिश्चित करता है कि इस परत में लागू कारकों के प्रभावों का आकलन उच्चतर परत में इकाइयों से लगभग लंबवत रूप में किया जाए।
3. चयनित ब्लॉक डिजाइन को बनाए रखते हुए, ब्लॉकों को पुनः व्यवस्थित करें ताकि इस परत में लागू कारकों और उच्चतर परतों में लागू कारकों के बीच अंतःक्रियाओं का सही आकलन सुनिश्चित हो सके।
4. 3 में प्राप्त डिजाइन को बनाए रखते हुए, ब्लॉकों को पुनः व्यवस्थित करें ताकि जब ये U_{i2} में इकाइयों के रूप में हों, तो प्रभावों का आकलन लगभग लंबवत रूप में U_{i-2} की इकाइयों से अनुमानित हो सके। यह सुनिश्चित करता है कि इस स्ट्रेटम में अनुमानित प्रभाव उच्चतर स्ट्रेटम इकाइयों से लगभग लंबवत हैं यदि। $\sigma^2_{i-1} = 0$। इसे $U_{i-3}, \ldots, U_1$ स्ट्रेटा के लिए भी दोहराएं।

पूर्ववर्ती विवरण में, यह सुझाव दिया गया है कि कोई सामान्य तरीके से उपचार डिजाइन और ब्लॉक डिजाइन चुन सकता है। उपयोग की जाने वाली विधियां कोई भी हो सकती हैं जो उचित हैं, लेकिन चरण 2–4 में आमतौर पर कंप्यूटर–आधारित एल्गोरिथ्म की आवश्यकता होती है।

अनुप्रयोग

उदाहरण 1. चार कारकों के लिए एक प्रयोग, प्रत्येक तीन स्तरों पर, तीन मुख्य इकाइयों में से प्रत्येक में पांच ब्लॉकों में चलाया जाना है, जिनमें से प्रत्येक में तीन सबयूनिट हैं। यह एक तीन–स्ट्रेटम डिजाइन $(s = 3)$ है, जहां कारक X_1, और X_2 को दूसरे स्ट्रेटम में इकाइयों पर लागू किया जाना चाहिए और X_3 और X_4 को तीसरे स्ट्रेटम में इकाइयों पर लागू किया जाना चाहिए।

इस स्थिति के लिए डिजाइन का निर्माण करने के लिए निम्नलिखित विधि का पालन किया जाता है:

1. U2 के लिए कोई X1 एवं X2 के लिए उपचार सेट का चयन कर सकता है। इस स्ट्रेटम में 15 इकाइयाँ हैं, इसलिए एक केंद्रीय समग्र डिजाइन (सीसीडी) का उपयोग किया जा सकता है (जो कि 32 फैक्टोरियल डिजाइन भी है) जिसमें फैक्टोरियल भाग डुप्लिकेट और तीन केंद्र बिंदु होते हैं। इस उपचार सेट को अब तीन इकाइयों के पांच ब्लॉकों

में व्यवस्थित किया जाना चाहिए। एक ऑर्थोगोनल अवरुद्ध डिजाइन मौजूद नहीं है, और डिजाइन का चयन करने के लिए एक अवरुद्ध एल्गोरिदम का उपयोग किया जा सकता है। रैखिक प्रभावों को प्राथमिकता देते हुए (वजन पैटर्नः क्रमशः रैखिक, द्विघात और अंतःक्रिया प्रभावों के लिए 2, 1, और 1), भारित माध्य दक्षता कारक मानदंड के संबंध में X1 और X2 के लिए सबसे अच्छा अवरुद्ध डिजाइन प्राप्त किया जाता है। यह डिजाइन रैखिक और द्विघात प्रभावों के लिए 100% कुशल है और इंटरैक्शन प्रभाव के लिए 41-67% कुशल है,

Stratum U_2									
Block 1		**Block 2**		**Block 3**		**Block 4**		**Block 5**	
X_1	X_2	X_1	X_2	X_1	X_2	X_1	X_2	X_1	X_2
-1	0	-1	-1	-1	1	-1	-1	-1	0
0	1	0	0	0	-1	0	0	0	1
1	-1	1	1	1	0	1	1	1	-1

2. U3 के लिए कोई X3 और X4 के लिए उपचार सेट का चयन कर सकता है। इस स्तर में 45 इकाइयाँ हैं और कोई भी 5 बार दोहराए गए 32 फैक्टोरियल डिजाइन का उपयोग कर सकता है। इस उपचार सेट को 3 के 15 ब्लॉकों में व्यवस्थित किया जाना चाहिए। फिर से एक ऑर्थोगोनल डिजाइन मौजूद नहीं है, और एक ब्लॉकिंग एल्गोरिदम का उपयोग किया जा सकता है। रैखिक प्रभावों को प्राथमिकता देते हुए (1 के समान वजन पैटर्न का उपयोग करके) डिजाइन प्राप्त किया जाता है। यह डिजाइन रैखिक और द्विघात प्रभावों के लिए 100% दक्षता कारक और अंतःक्रिया प्रभाव के लिए 50% दक्षता कारक देता हैः
3. स्तर U2 और U3 में अवरुद्ध डिजाइनों का चयन करने के बाद, X1 और X2 के स्तरों के संयोजन के लिए X3 और X4 के लिए ब्लॉक आवंटित करें। ऐसा करने के लिए किसी को X3 तथा X4 के

Stratum U_2									
Block 1		**Block 2**		**Block 3**		**Block 4**		**Block 5**	
X_1	X_2	X_1	X_2	X_1	X_2	X_1	X_2	X_1	X_2
-1	0	-1	-1	-1	1	-1	-1	-1	0
0	1	0	0	0	-1	0	0	0	1
1	-1	1	1	1	0	1	1	1	-1
Block 6		**Block 7**		**Block 8**		**Block 9**		**Block 10**	
X1	X2	X1	X2	X1	X2	X1	X2	X1	X2
-1	-1	-1	0	-1	1	-1	0	-1	0
0	1	0	1	0	-1	0	-1	0	1
1	0	1	-1	1	0	1	1	1	-1
Block 11		**Block 12**		**Block 13**		**Block 14**		**Block 15**	
X1	X2	X1	X2	X1	X2	X1	X2	X1	X2
-1	1	-1	1	-1	-1	-1	1	-1	1
0	0	0	0	0	0	0	0	0	-1
1	-1	1	-1	1	1	1	-1	1	0

लिए अन्य शब्दों के अनुमान को नुकसान पहुंचाए बिना Xl, X2 और X3, X4 के बीच की बातचीत को ध्यान में रखना होगा। चूँकि इन अंतःक्रियाओं का अनुमान मुख्य रूप से स्ट्रेटम U3 के भीतर की जानकारी से लगाया जाता है, डिजाइन का सबसे अच्छा समामेलन स्ट्रेटम U2 में इकाइयों को ब्लॉक के रूप में मानकर चुना जाता है – यानी, तीन इकाइयों में से प्रत्येक में 15 ब्लॉक। X1 और X2 के स्तरों के विभिन्न संयोजनों वाले ब्लॉकों के बीच X3 और X4 के लिए 15 समूहों को इंटरचेंज करके सर्वोत्तम व्यवस्था की तलाश की जाती है। अब उपयोग किया जाने वाला डिजाइन मानदंड Au है क्योंकि Xl, X2, X3 और X4 के लिए उपचार सेट निश्चित नहीं है।

4. स्ट्रेटम U1 में अवरोध के संबंध में 3 में डिजाइन में अभी भी सुधार किया जा सकता है। कोई व्यक्ति

Table 1: Design at stage 3 for Example 1

Block 1				Block 2				Block 3				Block 4				Block 5			
X_1	X_2	X_3	X_4	X_1	X_2	X_3	X_4	X_1	X_2	X_3	X_4	X_1	X_2	X_3	X_4	X_1	X_2	X_3	X_4
-1	-1	-1	-1	-1	-1	-1	1	-1	0	-1	-1	-1	1	-1	1	-1	1	-1	0
-1	-1	0	0	-1	-1	0	0	-1	0	0	1	-1	1	0	-1	-1	1	0	-1
-1	-1	1	1	-1	-1	1	-1	-1	0	1	0	-1	1	1	0	-1	1	1	1
0	0	-1	-1	0	0	-1	-1	0	1	-1	0	0	-1	-1	0	0	0	-1	0
0	0	0	0	0	0	0	1	0	1	0	1	0	-1	0	1	0	0	0	-1
0	0	1	1	0	0	1	0	0	1	1	-1	0	-1	1	-1	0	0	1	1
1	1	-1	-1	1	1	-1	-1	1	-1	-1	1	1	0	-1	-1	1	-1	-1	0
1	1	0	0	1	1	0	0	1	-1	0	-1	1	0	0	1	1	-1	0	-1
1	1	1	1	1	1	1	1	1	-1	1	0	1	0	1	0	1	-1	1	1

स्ट्रेटम U1 में पांच ब्लॉकों के बीच X3 और X4 के लिए तीन उपचार संयोजनों के समूहों का आदान–प्रदान कर सकता है। हालाँकि, अब केवल कारकों X1 और X2 के स्तरों के समान संयोजन वाली इकाइयों के बीच आदान–प्रदान की अनुमति है अन्यथा चरण 3 में पाया गया डिजाइन खो जाएगा। ब्लॉक 3 में तीन उपचारों के अंतिम समूह को ब्लॉक 5 में अंतिम समूह के साथ बदल दिया गया था और ब्लॉक 4 में पहले समूह को ब्लॉक 5 में पहले समूह के साथ बदल दिया गया था।

निष्कर्ष

कई औद्योगिक प्रयोगों में कुछ ऐसे कारक शामिल होते हैं जिनका स्तर दूसरों की तुलना में निर्धारित करना कठिन होता है। इनसे निपटने का सबसे अच्छा तरीका यह है कि प्रयोग को एक स्प्लिट–प्लॉट, या अधिक सामान्यतः मल्टी–स्ट्रैटम, डिजाइन के रूप में सावधानीपूर्वक योजनाबद्ध किया जाए। डिजाइन का निर्माण स्तर दर स्तर किया जाता है, जो उच्चतम स्तर से शुरू होता है। प्रत्येक स्तर में उस स्तर में लागू कारकों के लिए एक पूर्व–निर्दिष्ट उपचार सेट को उच्च स्तर की इकाइयों

Table 2: Final Design for Example 1

Block 1				Block 2				Block 3				Block 4				Block 5			
X_1	X_2	X_3	X_4	X_1	X_2	X_3	X_4	X_1	X_2	X_3	X_4	X_1	X_2	X_3	X_4	X_1	X_2	X_3	X_4
-1	-1	-1	-1	-1	-1	-1	1	-1	0	-1	-1	-1	1	-1	0	-1	1	-1	1
-1	-1	0	0	-1	-1	0	0	-1	0	0	1	-1	1	0	-1	-1	1	0	-1
-1	-1	1	1	-1	-1	1	-1	-1	0	1	0	-1	1	1	1	-1	1	1	0
0	0	-1	1	0	0	-1	-1	0	1	-1	0	0	-1	-1	0	0	0	-1	0
0	0	0	0	0	0	0	1	0	1	0	1	0	-1	0	1	0	0	0	-1
0	0	1	-1	0	0	1	0	0	1	1	-1	0	-1	1	-1	0	0	1	1
1	1	-1	1	1	1	-1	-1	1	-1	-1	0	1	0	-1	-1	1	-1	-1	1
1	1	0	0	1	1	0	0	1	-1	0	-1	1	0	0	1	1	-1	0	-1
1	1	1	-1	1	1	1	1	1	-1	1	1	1	0	1	0	1	-1	1	0

Table 3: Expected Variances for the Parameter Estimates for the Final Design for Example 1: ($\sigma^2_3 = 1$)

σ^2_1	1	4		16		
σ^2_2	1	1	4	1	4	16
$\hat{\beta}_1$	0.13	0.13	0.43	0.13	0.43	1.63
$\hat{\beta}_2$	0.13	0.13	0.43	0.13	0.43	1.63
$\hat{\beta}_{11}$	0.49	0.49	1.56	0.49	1.56	5.84
$\hat{\beta}_{22}$	0.49	0.49	1.56	0.49	1.56	5.84
$\hat{\beta}_{12}$	0.28	0.35	0.95	0.39	1.17	3.62
$\hat{\beta}_3$	0.03	0.03	0.03	0.03	0.03	0.03
$\hat{\beta}_4$	0.03	0.03	0.03	0.03	0.03	0.03
$\hat{\beta}_{33}$	0.10	0.10	0.10	0.10	0.10	0.10
$\hat{\beta}_{44}$	0.10	0.10	0.10	0.10	0.10	0.10
$\hat{\beta}_{13}$	0.05	0.05	0.05	0.05	0.05	0.05
$\hat{\beta}_{14}$	0.05	0.05	0.05	0.05	0.05	0.05
$\hat{\beta}_{23}$	0.05	0.05	0.05	0.05	0.05	0.05
$\hat{\beta}_{24}$	0.05	0.05	0.05	0.05	0.05	0.05
$\hat{\beta}_{34}$	0.08	0.08	0.09	0.08	0.09	0.10

के लिए लगभग ऑर्थोगोनल बनाने की व्यवस्था की जाती है, जिससे सभी प्रभावों का अनुमान लगाया जा सके। तीन उदाहरण विधि की प्रयोज्यता दिखाते हैं और उपचार सेट की पसंद के साथ अंतिम डिजाइन के संबंध की जांच करने के लिए भी उपयोग किए जाते हैं। ट्रिंका और गिल्मर के मल्टीस्ट्रैटम एल्गोरिदम का सामान्य विचार यह है कि प्रत्येक स्ट्रैटम के लिए यादृच्छिक कारकों के लिए एक डिजाइन अलग से चुना जाता है और फिर उन्हें इस तरह से संयोजित किया जाता है जो मल्टी–स्ट्रैटम संरचना को ध्यान में रखता है।

संदर्भ

एटकिंसन, ए.सी., एवं डोनेव, ए.एन. (1989)। द कंस्ट्रक्शन ऑफ एगजेकट डी – ओपटीमम एक्सपेरिमेंटल डीजाइनस विथ एप्लीकेशन टू ब्लॉकिंग रिस्पांस सरफेस डिजाइनस। बिओमेट्रिका, 76, 515–526।

एडमंडसन, आर.एन. (1991)। एग्रीकल्चरल रिस्पांस सरफेस एक्सपेरिमेंट्स ऑन फोर लेवल फैक्टोरियल डिजाइनस। बायोमेट्रिक्स, 47, 1435–1448।

कुक, आर.डी., एवं नचटशाइम, सी.जे. (1989)। कंप्यूटर–एडेड ब्लॉकिंग ऑफ फैक्टोरियल एंड रिस्पांस सरफेस डिजाइनस। टेक्नोमेट्रिक्स, 31, 339–346।

गिल्मर, एस.जी., एवं ट्रिनका, एल.ए. (2000)। सम प्रैक्टिकल एडवाइस ऑन पोलीनोमिअल रिग्रेशन एनालिसिस फ्रॉम ब्लॉक्ड रिस्पांस सरफेस डिजाइनस, कम्युनिकेशन इन स्टेटिस्टिक्स – थ्योरी एंड मेथड्स, 29, 2157–2183।

ट्रिनका, एल.ए. एवं गिल्मर, एस.जी. (2001)। मल्टीस्ट्रैटम रिस्पांस सरफेस डिजाइनस। टेक्नोमेट्रिक्स, 43 (1), 25–33।

ड्रापर, एन.आर., एवं जॉन, जे.ए. (1998)। रिस्पांस सरफेस डिजाइनस वेयर लेवल्स ऑफ सम फैक्टर्स आर डिफिकल्ट टू चेंज। ऑस्ट्रेलियन एंड न्यूजीलैंड जर्नल ऑफ स्टैटिस्टिक्स, 40, 487–495।

पियर्स, एस.सी. (1974)। ऑप्टीमेलटी ऑफ डिजाइनस इन पॉट एक्सपेरिमेंट्स, प्रोसीडिंग्स ऑफ द 8थ इंटरनेशनल बायोमेट्रिक कांफ्रेंस। वाशिंगटन डी.सी.: इंटरनेशनल बायोमेट्रिक सोसाइटी।

बिंघम, डी., एवं सिटर, आर.आर. (1999 ए)। मिनिमम ऐबररेशन टू लेवल फ्रैक्शनल फैक्ट्रोरिअल स्प्लिट प्लाट डिजाइनस। टेक्नोमेट्रिक्स, 41, 62–70।

बिंघम, डी., एवं सिटर, आर.आर. (1999 बी)। सम थेरोरेटिकल रिजल्ट्स फॉर फ्रैक्शनल फैक्टोरियल स्प्लिट–प्लॉट डिजाइनस। द एनल्स ऑफ स्टैटिस्टिक्स, 27, 1240–1255।

लेटसिंगर, जे.डी., मायर्स, आर.एच., एवं लेंटनर, एम. (1996)। रिस्पांस सरफेस मेथड्स फॉर बाइ–रैंडमाइजेसन स्ट्रक्टर्स। जर्नल ऑफ क्वालिटी टेक्नोलॉजी, 28, 381–397।

शॉन, ई.डी. (1999)। डिजाइनिंग फ्रैक्शनल टू लेवल्स एक्सपेरिमेंट्स विथ नेस्टेड एरर स्ट्रक्चर। जर्नल ऑफ एप्लाइड स्टैटिस्टिक्स, 26, 495–508।

हुआंग, पी., चेन, डी., एवं वोएलकेल, जे.ओ. (1998)। मिनिमम– ऐबररेशन टू लेवल स्प्लिट–प्लॉट डिजाइनस। टेक्नोमेट्रिक्स, 40, 314–326।

प्रतिदर्श सर्वेक्षणों में परिमित समष्टि औसत के अनुमानकों का एक बेहतर वर्ग

दीपक सिंह, राजू कुमार, राहुल बनर्जी, पंकज दास, भारती, अंकुर बिश्वास एवं कौस्तव आदित्य

भा.कृ.अनु.प. – भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली

सार

इस अध्ययन में प्रतिदर्श सर्वेक्षणों में सुलभ सहायक जानकारी का उपयोग करने वाली समष्टि के जनसंख्या विचरण का मूल्यांकन करने के लिए अनुमानकों के एक बेहतर वर्ग का प्रस्ताव किया गया है। आकलक के प्रस्तावित वर्ग के पूर्वाग्रह और माध्य वर्गीय की त्रुटि की जांच बड़े प्रतिदर्श सन्निकटन के तहत की गई है। अनुभवजन्य अध्ययन (empirical study) के तहत, यह प्रदर्शित किया गया है कि अनुमानकों का प्रस्तावित वर्ग कुछ वर्तमान अनुमानकों से बेहतर प्रदर्शन कर रहा है।

परिचय

सर्वेक्षण नमूने में, यह अच्छी तरह से स्थापित है कि यदि अनुमान चरण में चर x पर सहायक जानकारी का उपयोग किया जाए तो अध्ययनकर्ता y के चरित्र के अनुमानक की सटीकता को बढ़ाया जा सकता है। जब सहायक चर x के पैरामीटर, जैसे की तिरछापन का गुणांक $\beta_1(x)$, कर्टोसिस का गुणांक $\beta_2(x)$, भिन्नता का गुणांक C_x, x और y के बीच सहसंबंध गुणांक ρ इत्यादि ज्ञात होते हैं, फिर जनसंख्या माध्य और जनसंख्या विचरण के लिए अनुपात और उत्पाद अनुमानकों के कई सामान्यीकरणों को परिभाषित किया जा सकता है।

इस अध्ययन में प्रतिदर्श सर्वेक्षणों में सुलभ सहायक जानकारी का उपयोग करने वाली समष्टि के विचरण का मूल्यांकन करने के लिए अनुमानकों के एक बेहतर वर्ग का प्रस्ताव किया गया है। आकलक के प्रस्तावित वर्ग के पूर्वाग्रह और माध्य वर्गीय की त्रुटि की जांच बड़े प्रतिदर्श सन्निकटन के तहत की गई है। उन संक्षेपणों की पहचान की गई है जिनके तहत अनुमानकों का प्रस्तावित वर्ग कुछ स्थापित अनुमानकों से बेहतर प्रदर्शन कर रहा है। अनुभवजन्य अध्ययन (empirical study) के तहत, यह प्रदर्शित किया गया है कि अनुमानकों का प्रस्तावित वर्ग कुछ वर्तमान अनुमानकों से बेहतर प्रदर्शन कर रहा है।

जनसंख्या विचरण के लिए कुशल अनुमानकों के निर्माण की समस्या पर विभिन्न लेखकों द्वारा चर्चा की गई है जैसे की दास और त्रिपाठी (1978), इसाकी (1983), सिंह एट अल. (1988), अग्रवाल एवं स्टापिट (1995), गार्सिया एवं सेब्रेन (1997), उपाध्याय एवं सिंह (1999), आर्कोस एट अल. (2005), कादिलर एवं सिंगी (2006, 2006इ), सिंह एट अल. (2009), यादव एट अल. (2013), इत्यादि। अध्ययन के लिए, हम परिभाषित करते हैं

$$f=(n/N),$$

$$h=\mu_{22}/(\mu_{20}\mu_{02}), \qquad C=(h-1)/\{\beta_2(x)-1\}$$

$$S_x^2=\{1/(N-1)\}\sum_{i=1}^{N}\left(x_i-\bar{X}\right)^2 \quad \text{और}$$

$$\mu_{rs}=(1/N)\sum_{i=1}^{N}\left(y_i-\bar{Y}\right)^r\left(x_i-\bar{X}\right)^s;(r,s)$$

गैर नकारात्मक पूर्णांक हैं।

अध्ययन में, सीमित जनसंख्या सुधार को नजरअंदाज कर दिया गया है। सबसे आम ज्ञात S^2_y का निष्पक्ष अनुमानक $t_o = S^2_y$ है, जिसका प्रथम–डिग्री सन्निकटन विचरण निम्न प्रकार से है

$$V(t_0) = MSE(t_0) = \left(S_y^4/n\right)\left[\left\{\beta_2(y)\text{-}1\right\}\right] \qquad (1)$$

इसाकी (1983) ने जनसंख्या विचरण का एक अनुमानक प्रस्तावित किया है जब सहायक चर x का जनसंख्या विचरण ज्ञात होता है जिसे इस प्रकार परिभाषित किया जाता है

$$t_1 = s_y^2\left(\frac{S_x^2}{s_x^2}\right)$$

$$B(t_1) = \left(S_y^2/n\right)\left\{\beta_2(x)\text{-}1\right\}(1\text{-}C)$$

$$MSE(t_1) = \left(S_y^4/n\right)\left[\left\{\beta_2(y)\text{-}1\right\}+\left\{\beta_2(x)\text{-}1\right\}(1\text{-}2C)\right] \qquad (2)$$

अनुमानकर्ताओं का सुझाया गया प्रस्तावित समूह

हमने उपाध्याय एट अल. (2011) और यादव एवं कादिलर (2013) द्वारा प्रेरित हो कर $T^{p,q}_{\omega_1,\omega_2\beta}$, जनसंख्या विचरण S^2_y का अनुमानक का सुझाव दिया है:

$$T^{p,q}_{\omega_1,\omega_2\beta} = s_y^2\left[\omega_1\exp\left\{\frac{p\left(S_x^2\text{-}s_x^2\right)}{S_x^2+(a\text{-}1)s_x^2}\right\}+\omega_2\exp\left\{\frac{q\left(s_x^2-S_x^2\right)}{S_x^2+(a\text{-}1)s_x^2}\right\}\right] \qquad (3)$$

इसमें ω_1 और ω_2 उपयुक्त रूप से चुना गया स्थिरांक है जिस पर MSE न्यूनतम है जबकि 'a' एक धनात्मक वास्तविक अदिश राशि है। अदिश 'a' के विभिन्न मान रखने पर, हमें अलग–अलग अनुमानक प्राप्त होते हैं:

माना कि $S^2_y = (1+e_0)\, S^2_y$ और $S^2_x = (1+e_1)\, S^2_x$ है, जिसमें कि $E(e_0) = E(e_1) = 0$ और परिमित जनसंख्या सुधार (एफ.पी.सी.) की अनदेखी करते हुए, हमारे पास है $E(e^2_0) = \{\beta_2(y)\text{-}1\}/n$, $E(e^2_1) = \{\beta_2(x)\text{-}1\}/n$, और $E(e_0 e_1) = (h\text{-}1)/n$, है। जब $i>2$ है तो $E(e^i_o)$ और $E(e^i_1)$ का योगदान और जब $(i+j)>2$ है तब $E(e^i_o e^j_1)$ का योगदान नगण्य है। समीकरण (3) को त्रुटि रूप में व्यक्त पर हमें समीकरण (4) मिलती है।

$$T^{p,q}_{\omega_1,\omega_2,a} = S_y^2\left[\begin{array}{l}(\omega_1+\omega_2)(1+e_0)+\dfrac{(-\omega_1 p+\omega_2 q)}{a}(e_1+e_0e_1)+\\ \left\{\begin{array}{l}(\omega_1 p-\omega_2 q)(a-1)+\\ \left(\dfrac{\omega_1 p^2+\omega_2 q^2}{2}\right)\end{array}\right\}\dfrac{e_1^{\,2}}{a^2}\end{array}\right] \qquad (4)$$

e_1 की श्रृंखला का विस्तार करने के बाद, हम समीकरण (5) प्राप्त करते हैं।

$$T^{p,q}_{\omega_1,\omega_2,a} - S_y^2 = S_y^2\left[\begin{array}{l}(\omega_1+\omega_2-1)+(\omega_1+\omega_2)e_0+\\ \dfrac{(-\omega_1 p+\omega_2 q)}{a}(e_1+e_0e_1)+\\ \left\{(\omega_1 p-\omega_2 q)(a-1)+\left(\dfrac{\omega_1 p^2+\omega_2 q^2}{2}\right)\right\}\dfrac{e_1^2}{a^2}\end{array}\right] \qquad (5)$$

समीकरण के दोनों ओर अपेक्षाएं रखते हुए हम $T^{p,q}_{\omega_1,\omega_2,a}$ का पूर्वाग्रह (Bias) समीकरण (6) के रूप में प्राप्त करते हैं।

$$Bias\left(T^{p,q}_{\omega_1,\omega_2,a}\right) = S_y^2\left[\begin{array}{l}(\omega_1+\omega_2-1)+\dfrac{(-\omega_1 p+\omega_2 q)}{a}\dfrac{(h\text{-}1)}{n}+\\ \left\{(\omega_1 p-\omega_2 q)(a-1)+\left(\dfrac{\omega_1 p^2+\omega_2 q^2}{2}\right)\right\}\dfrac{\{\beta_2(x)\text{-}1\}}{a^2 n}\end{array}\right] \qquad (6)$$

समीकरण (4) के दोनों ओर वर्ग और अपेक्षाएँ लेने पर माध्य वर्गीय त्रुटि प्राप्त होती है।

$$MSE\left(T^{p,q}_{\omega_1,\omega_2,a}\right) = S_y^4\left[1+\omega_1^{\,2}A_1+\omega_2^{\,2}A_1+2\omega_1\omega_2A_1-2\omega_1A_1-2\omega_2A_1\right] \qquad (7)$$

जहाँ कि

$$A_1 = 1-\frac{4p}{a}\frac{(h\text{-}1)}{n}+\frac{\{\beta_2(y)\text{-}1\}}{n}+2\left(p^2+p(a-1)\right)\frac{\{\beta_2(x)\text{-}1\}}{a^2n}$$

$$A_2 = 1+\frac{4q}{a}\frac{(h\text{-}1)}{n}+\frac{\{\beta_2(y)\text{-}1\}}{n}+2\left(q^2-q(a-1)\right)\frac{\{\beta_2(x)\text{-}1\}}{a^2n}$$

$$A_3 = 1+\frac{2(-p+q)}{a}\frac{(h\text{-}1)}{n}+\frac{\{\beta_2(y)\text{-}1\}}{n}+\left\{-pq+(p-q)(a-1)+\frac{\left(p^2+q^2\right)}{2}\right\}\frac{\{\beta_2(x)\text{-}1\}}{a^2n}$$

$$A_4 = 1-\frac{p}{a}\frac{(h\text{-}1)}{n}+\left(p(a-1)+\frac{p^2}{2}\right)\frac{\{\beta_2(x)\text{-}1\}}{a^2n}$$

$$A_5 = 1+\frac{q}{a}\frac{(h\text{-}1)}{n}+\left(-q(a-1)+\frac{q^2}{2}\right)\frac{\{\beta_2(x)\text{-}1\}}{a^2n}$$

है। समीकरण (7) से हमें ω_1 और ω_2 का इष्टतम मान मिलता है।

$$\omega_1 = \left(\frac{A_2 A_4 - A_3 A_5}{A_1 A_2 - A_3^2}\right) = \omega_1',\ \omega =_2 \left(\frac{A_1 A_5 - A_3 A_4}{A_1 A_2 - A_3}{}_2\right) = \omega_2' \Big\} \tag{8}$$

ω_1 और ω_2 के इष्टतम मान पर हमें न्यूनतम माध्य वर्गीय त्रुटि प्राप्त होती है।

$$MSE_{\min}\left(T^{p,q}_{\omega_1,\omega_2,a}\right) = S_y^4 \left\{ 1 - \frac{\left(A_2 A_4^2 + A_1 A_5^2 - 2 A_3 A_4 A_5\right)}{A_1 A_2 - A^2{}_3} \right\} \tag{9}$$

विशेष स्थिति

यदि हम समीकरण संख्या (3) में $(\omega_1+\omega_2)$ का प्रतिबंध करते हैं, तो हमें रूपांतरित अनुमानक समीकरण (10) के रूप में मिलता है।

$$T^{p,q}_{\omega,1-\omega,\beta} = s_y^2 \left[\omega \exp\left\{ \frac{p\left(S_x^2 - s_x^2\right)}{S_x^2 + (a-1)s_x^2} \right\} + (1-\omega)\exp\left\{ \frac{q\left(s_x^2 - S_x^2\right)}{S_x^2 + (a-1)s_x^2} \right\} \right] \tag{10}$$

$T^{p,q}_{\omega,1-\omega,\beta}$ अनुमापक पूर्वाग्रह और माध्य वर्गीय त्रुटि निम्न प्रकार से है।

$$Bias(T^{p,q}_{\omega,1-\omega\beta}) = S_y^2 \left[\begin{array}{l} \{q - \omega(p+q)\}\dfrac{(h-1)}{n} + \\ \left\{ \omega(p+q)\left(\dfrac{(p-q)}{2} + (\beta-1)\right) + q\left(\dfrac{q}{2} - (\beta-1)\right) \right\} \dfrac{\{\beta_2(x)-1\}}{\beta^2 n} \end{array} \right] \tag{11}$$

$$MSE\left(T^{p,q}_{\omega,1-\omega,\beta}\right) = \frac{S_y^4}{\beta n} \left[\begin{array}{l} \omega^2 \dfrac{\{\beta_2(x)-1\}(p+q)^2}{\beta} - 2\omega(p+q)\left\{(h-1) + \dfrac{q\{\beta_2(x)-1\}}{\beta}\right\} \\ +2q(h-1) + \beta\{\beta_2(y)-1\} + \dfrac{q^2\{\beta_2(x)-1\}}{\beta} \end{array} \right]$$

या

$$MSE\left(T^{p,q}_{\omega,1-\omega,a}\right) = S_y^4\left[(1 + D_2 - 2D_5) + \alpha^2(D_1 + D_2 - 2D_3) - 2\alpha(D_2 - D_3 + D_4 - D_5)\right]\Big] \tag{12}$$

समीकरण संख्या (10) के लिए न्यूनतम माध्य वर्गीय त्रुटि

$$\omega_{opt} = \frac{\beta C + q}{p+q} = \frac{(A_2 - A_3 + A_4 - A_5)}{(A_1 + A_2 - 2A_3)},\ \left(C = \frac{(h-1)}{\{\beta_2(x)-1\}}\right)$$

पर हैं।

$$MSE_{\min}\left(T^{p,q}_{\omega,1-\omega,\beta}\right) = \frac{S_y^4}{n}(\beta_2(y)-1)\left[1 - \frac{(\beta_2(x)-1)C^2}{(\beta_2(y)-1)}\right]$$

$$= \bar{Y}^2\left[1 + Z_2 - 2Z_5 - \frac{2(Z_2 - Z_3 + Z_4 - Z_5)^2}{Z_1 + Z_2 - 2Z_3}\right]$$

$$= \frac{S_y^4}{n}(\beta_2(y)-1)(1-\rho^{*2}) = MSE\left(s^2_{y\,LR}\right) \tag{13}$$

यहां s^2_{yLR} जनसंख्या विचरण का रैखिक प्रतिगमन अनुमानक है $s^2_{yLR} = s^2_y + \hat{\beta}\left(S_x^2 - s_x^2\right)$।

दक्षता तुलना

समीकरण (12) की तुलना क्रमशः (1) और (2) से करने पर हमें मिलता है कि अनुमानकों का प्रस्तावित परिवार t_0 और t_1 से अधिक कुशल है।

$$MSE(t_0) - MSE_{\min}\left(T^{p,q}_{\omega,1-\omega,a}\right) = \frac{S_y^4 (h-1)^2}{n\{\beta_2(x)-1\}} \tag{14}$$

$$MSE(t_1) - MSE_{\min}\left(T^{p,q}_{\omega,1-\omega,\beta}\right) = \frac{S_y^4 (\beta_2(x)-h)^2}{n\{\beta_2(x)-1\}} \tag{15}$$

अनुमानकों के प्रस्तावित परिवार की दक्षताओं की उसके उपपरिवार के साथ तुलना करने पर, हम पाते हैं कि अनुमानकों का प्रस्तावित परिवार अपने अनुमानकों के उपपरिवार की तुलना में अधिक कुशल है।

$$MSE\left(T_{\omega,1-\omega,a}\right) - MSE_{\min}\left(T_{\omega_1,\omega_2,a}\right) = S_y^4 \frac{\left[A_1(A_2 - A_5) + A_4(A_3 - A_2) + A_3(A_5 - A_3)\right]^2}{(A_1 + A_2 - 2A_3)(A_1 A_2 - A_3^2)} \geq 0 \tag{16}$$

निष्कर्ष

सामान्य निष्पक्ष अनुमानक की तुलना में अनुमानकों के प्रस्तावित परिवार के प्रदर्शन को देखने के लिए हमने निम्नलिखित जनसंख्या डेटा सेट लिए:

यह तुलना करने के लिए कि हमारे आकलनकर्ताओं का नया सेट सामान्य निष्पक्ष आकलनकर्ता की तुलना में कितना अच्छा प्रदर्शन करता है, हमने ''प्रतिशत सापेक्ष दक्षता'' (पी.आर.ई.) की गणना की। इससे हमें यह देखने में मदद मिलती है कि हमारे नए अनुमानक सामान्य अनुमानक की तुलना में कितने कुशल हैं। हमने

तालिका 1: जनसंख्या I और II का विवरण

जनसंख्या	स्रोत	अध्ययन चर (Y)	सहायक चर (x)
1	मूर्ति (1967, पृ. 226)	उत्पादन	श्रमिकों की संख्या
2	गुजराती (2004, पृ. 433)	औसत (मील प्रति गैलन)	शीर्ष गति (मील प्रति घंटा)

तालिका 2: जनसंख्या I और II के मापदंडों का विवरण

पैरामीटर	जनसंख्या I	जनसंख्या II
N	80	39
n	25	14
$\bar{X}$	11.2646	106.2
$\bar{y}$	51.8264	26.8433
Cy	0.3542	0.2326
Cx	0.7507	0.1049
ρ	0.9413	0.96682
C	0.44413	2.14378
$\beta_2(y)$	2.2384	2.4032
$\beta_2(x)$	2.8306	2.993
h	2.193	2.4882
S_y	18.3569	6.2442
S_x	8.45636	11.1413
ρ^*	0.52624	0.53215
f	0.3125	0.35897

निम्नलिखित सूत्र का उपयोग करके ऐसा किया।

$$PRE\left(T^{p,q}_{\omega_1,\omega_2,a}, s_y^2\right) = \frac{\{\beta_2(y)-1\}}{n\left\{1-\frac{\left(Z_2Z_4^2+Z_1Z_5^2-2Z_3Z_4Z_5\right)}{Z_1Z_2-Z_3^2}\right\}} \quad (17)$$

$$PRE\left(T^{p,q}_{\omega_1,\omega_2,a}, T^{p,q}_{\omega,1-\omega,a}\right) = \frac{\left[\{\beta_2(y)-1\}\left(1-\rho^{*2}\right)\right]}{n\left\{1-\frac{\left(Z_2Z_4^2+Z_1Z_5^2-2Z_3Z_4Z_5\right)}{Z_1Z_2-Z_3^2}\right\}} \quad (18)$$

तालिका 3 के निष्कर्षों से हमें पता चलता है कि अनुमानकों के प्रस्तावित परिवार का एम.एस.ई. इसके अनुमानकों के उपपरिवार, जनसंख्या भिन्नता के निष्पक्ष अनुमानक (t_0), इसाकी (1983) (t_1) और जनसंख्या भिन्नता के रैखिक प्रतिगमन अनुमानक s^2_{yLR} की तुलना में कम है।

इससे साबित होता है कि अनुमानकों का प्रस्तावित परिवार उपर्युक्त अनुमानकों और उन सभी अनुमानकों के लिए कुशल है जिनकी दक्षता s^2_{yLR} से कम या बराबर है।

तालिका 3: जनसंख्या I और II के मापदंडों का विवरण

संदर्भ

दास, ए.के. एवं त्रिपाठी, टी.पी. (1978)। यूज ऑफ ऑक्सीलियरी इन एस्टिमेटिंग द फाइनिट पॉपुलेशन वैरिअन्स। सांख्य, 40, 139–148।

इसाकी, सी.टी. (1983)। वैरिअन्स एस्टिमेशन यूसिंग ऑक्सीलियरी इनफार्मेशन। जर्नल ऑफ द अमेरिकन स्टैटिस्टिकल एसोसिएशन, 78, 117–123।

सिंह, एच.पी., उपाध्याय, एल.एन. एवं नामजोशी, यू.डी. (1988)। एस्टिमेशन ऑफ फाइनिट पॉपुलेशन वैरिअन्स, 57, 1331–1334।

अग्रवाल, एम.सी. एवं स्टैपिट, ए.बी. (1995)। अनबाएसड रेश्यो–टाइप वैरिअन्स एस्टिमेशन। सांख्यिकी और संभाव्यता पत्र, 25, 361–364।

गार्सिया, एम.आर. एवं सेब्रेन, ए.ए. (1997)। वैरिअन्स एस्टिमेश यूसिंग ऑक्सीलियरी इनफार्मेशनः ऐन ऑलमोस्ट अनबाएसड मल्टीवेरियेट रेश्यो एस्टिमेटर । मेट्रिका, 45, 171–178।

उपाध्याय, एल.एन. एवं सिंह, एच.पी. (1999ए)। ऐन एस्टिमेटर फॉर पापुलेशन वैरिअन्स डेट युटीलिसेस द कर्तोसिस ऑफ ऍन ऑक्सीलियरी वेरिएबल इन सैंपल सर्वेस। विक्रम मैथमेटिकल जर्नल, 19, 14–17।

उपाध्याय, एल.एन. एवं सिंह, एच.पी. (1999बी)। यूज ऑफ ट्रांसफोर्मेड ऑक्सीलियरी वेरिएबल इन एस्टिमेटिंग

तालिका 3: जनसंख्या I और II के मापदंडों का विवरण

P और q का मान		जनसंख्या (I) की प्रतिशत सापेक्ष दक्षता'' (पीआरई)				जनसंख्या (II) की प्रतिशत सापेक्ष दक्षता'' (पीआरई)			
P	Q	a=1		a=2		a=1		a=2	
		w.r.t. $T^{p,q}_{\omega,1-\omega,a}$	w.r.t. s^2_y	w.r.t. $T^{p,q}_{\omega,1-\omega,a}$	w.r.t. s^2_y	w.r.t. $T^{p,q}_{\omega,1-\omega,a}$	w.r.t. s^2_y	w.r.t. $T^{p,q}_{\omega,1-\omega,a}$	w.r.t. s^2_y
1	1	103.22	277.32	102.25	274.72	161.13	175.47	125.06	136.19
ρ^*	ρ^*	102.19	274.56	102.03	274.12	126.32	137.57	119.57	130.21
C_x	C_x	100.36	269.64	101.45	272.58	116.45	126.82	117.29	127.73
C_y	C_y	100.38	269.70	100.83	270.91	109.90	119.68	115.40	125.67
C	C	100.17	269.14	100.92	271.15	110.33	120.15	115.54	125.82
f	f	100.48	269.98	117.88	316.71	111.29	121.19	199.27	217.01
$1-f$	$1-f$	100.12	269.01	101.31	272.20	129.80	141.36	120.26	130.97
$1+f$	$1+f$	115.09	309.21	103.97	279.33	252.99	275.51	132.38	144.16
$1+2f$	$1+2f$	152.04	408.51	106.90	287.23	1192.77	1298.94	143.45	156.22
$f/1-f$	$f/1-f$	100.15	269.08	100.93	271.18	112.88	122.93	116.31	126.66
$2f/1-f$	$2f/1-f$	101.73	273.34	101.91	273.82	126.91	138.20	119.69	130.34
$1-2f/1-f$	$1-2f/1-f$	100.02	268.73	151.58	407.26	123.83	134.85	1246.92	1357.91
$2f/1-2f$	$2f/1-2f$	161.56	434.06	100.72	270.62	177.87	193.70	115.49	125.77
$1/1+f$	$1/1+f$	100.42	269.79	101.48	272.66	134.21	146.16	121.08	131.86
$f/1+f$	$f/1+f$	100.66	270.45	100.75	270.69	110.61	120.45	115.63	125.92
$2f/1+f$	$2f/1+f$	100.11	268.97	100.96	271.25	114.33	124.51	116.72	127.11
$1-2f/1+f$	$1-2f/1+f$	100.55	270.15	133.50	358.68	113.60	123.71	347.36	378.28
$1/1+2f$	$1/1+2f$	100.01	268.70	101.17	271.82	124.32	135.38	119.14	129.75
$f/1+2f$	$f/1+2f$	100.75	270.70	100.72	270.63	110.26	120.07	115.52	125.80
$2f/1+2f$	$2f/1+2f$	100.31	269.50	100.86	270.98	112.71	122.74	116.26	126.61
$1-2f/1+2f$	$1-2f/1+2f$	100.67	270.49	114.96	308.86	112.23	122.22	201.01	218.90

द फाइनिट पॉपुलेशन मीन। बायोमेट्रिकल जर्नल, 41(5), 627–636।

आर्कोस, ए., रुएडा, एम., मार्टिनेज, एम.डी., गोंजालेज, एस. एवं रोमन, वाई. (2005)। इनकॉर्पोरेटिंग द ऑक्सीलियरी इनफार्मेशन अवेलेबल इन वैरिअन्स एस्टिमेशन। एप्लाइड मैथेमेटिक्स एंड कम्प्यूटेशन, 160, 387–399।

कादिलर, सी. एवं सिंगी, एच. (2006ए)। सरल और स्तरीकृत यादृच्छिक नमूने में जनसंख्या भिन्नता के लिए अनुपात अनुमानक। एप्लाइड मैथेमेटिक्स एंड कम्प्यूटेशन, 173, 1047–1059।

कादिलर, सी. एवं सिंगी, एच. (2006बी)। इम्प्रूवमेंट इन वैरिअन्स एस्टिमेशन इन सिंपल रैंडम सैंपलिंग । सांख्यिकी–सिद्धांत और विधियों में संचार, 36, 2075–2081।

सिंह, आर., चौहान, पी., सावन, एन. एवं स्मरंडाचे, एफ. (2009)। दो सहायक चर का उपयोग करके जनसंख्या विचरण के लिए बेहतर घातीय अनुमानक। ऑक्टोगोन मैथेमेटिकल जर्नल, 17(2), 667–674।

यादव, एस.के. एवं कादिलर, सी. (2013)। इंप्रोवेद रेश्यो एंड प्रोडक्ट एक्सपोनेंशियल टाइप एस्टिमेटर्स। रेविस्टा कोलंबियाना डी एस्टाडिस्टिका, 36(1), 145–152।

ड्रोन प्रौद्योगिकी: कृषि में उच्च उत्कृष्टि का संवर्धन

पंकज दास, भारती, राहुल बनर्जी, दीपक सिंह एवं राजू कुमार
भा.कृ.अनु.प. – भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली

ड्रोन क्या है?

ड्रोन्स, जिन्हें Unmanned Aerial Vehicles (UAVs) भी कहा जाता है, एक आत्मचालित और स्वयं संचालित उड़ाने वाला यंत्र होता है जिसे दूरस्थ स्थानों से निर्देशित किया जा सकता है। यद्यपि ड्रोन के प्रारंभिक उपयोग प्रमुखतः सैन्य थे, लेकिन आजकल ड्रोन विज्ञान, संचार, और कृषि आदि कई क्षेत्रों में व्यापक रूप से उपयोग किये जाते हैं। उत्कृष्ट आकाशीय फोटोग्राफी से लेकर जीवन बचाने वाली खोज और बचाव मिशन तक, ड्रोन्स अत्यंत महत्वपूर्ण सिद्ध हो रहे हैं।

ड्रोन के आवश्यक भाग: ड्रोन के कुछ मुख्य भाग निम्नलिखित हैं:

मुख्य भाग: फ्रेम एवं आर्म्स

मूल रूप से, फ्रेम ड्रोन के शरीर का संरचन होता है, जो सभी अन्य घटकों के लिए एक मजबूत और हल्का आधार प्रदान करता है। कार्बन फाइबर या उच्च प्रभाव प्लास्टिक जैसे पदार्थ मजबूती और वजन की प्रभावकारिता के कारण ड्रोन के लिए सामान्य विकल्प होते हैं। आर्म्स बाजू फ्रेम से बाहर निकलते हैं, जो मोटर्स और प्रोपेलर्स को धारण करते हैं। आर्म्स की संख्या ड्रोन के आकार और विन्यास पर निर्भर करती है। क्वाडकॉप्टर्स (चार आर्म्स) ड्रोन का सबसे सामान्य प्रकार हैं।

उड़ान के इंजन: मोटर्स और प्रोपेलर

ड्रोन की मोटर्स जैसे कि ब्रशलेस डीसी इलेक्ट्रिक मोटर्स मुख्य रूप से मांसपेशियों की भूमिका निभाते हैं। ये मोटर्स प्रोपेलर्स को कुशलतापूर्वक घुमाते हैं, जिससे ड्रोन को हवा में उड़ाने और आगे बढ़ाने के लिए आवश्यक उच्चता और प्रोपल्शन उत्पन्न होता है। प्रोपेलर्स विविध

चित्र 1: ड्रोन के घटक (Source: www-hindibriefingpedia-com)

डिजाइनों में उपलब्ध होते हैं, जो गति, चुस्ती, या भार धारण क्षमता जैसी विशेष उड़ान की गुणवत्ता के लिए अनुकूलित होते हैं।

मस्तिष्कः उड़ान नियंत्रक

उड़ान नियंत्रक (flight controller) ड्रोन के मस्तिष्क की भूमिका निभाता है, रिमोट कंट्रोल से पायलट कमांड को समझाता है, और सिग्नल्स को मोटर्स के लिए रूपांतरित करता है। यह स्थिरता बनाए रखने और सुचारू उड़ान सुनिश्चित करने के लिए जाइरोस्कोप और एक्सेलेरोमीटर जैसे ऑनबोर्ड सेंसर का भी उपयोग करता है। कुछ उन्नत उड़ान नियंत्रक पूर्व–निर्धारित उड़ान मार्गों का उपयोग करके संचालन कार्य भी कर सकते हैं।

कैमरा और सेंसर

कई ड्रोनों में कैमरे लगे होते हैं, जो आकाशीय फोटोग्राफी और वीडियोग्राफी के लिए उपयुक्त होते हैं। ये कैमरे सरल हाई–डेफिनिशन मॉडल से लेकर सटीक कृषि में उपयोग किए जाने वाले परिष्कृत मल्टीस्पेक्ट्रल सेंसर तक हो सकते हैं। अतिरिक्त सेंसर, जैसे नेविगेशन के लिए जीपीएस और ऊंचाई माप के लिए बैरोमीटर, ड्रोन की क्षमताओं को और बढ़ाते हैं।

शक्ति का स्रोतः बैटरी

बैटरी ड्रोन को उसकी आवाजाही के लिए आवश्यक ऊर्जा प्रदान करती है, जिससे वह हवा में उड़ान भर सकता है। लिथियम पॉलिमर (LiPo) बैटरी, उनकी उच्च ऊर्जा घनत्व और हल्के डिजाइन के कारण, ड्रोन उपयोगकर्ताओं के बीच सबसे लोकप्रिय हैं। उड़ान की अवधि बैटरी क्षमता के आधार पर निर्धारित होती है, और बड़ी बैटरी लंबे मिशन के लिए अनुमति देती है।

रिमोट कंट्रोलः पायलट का कमांड सेंटर

रिमोट कंट्रोल ड्रोन के पायलट के संचालन साधन के रूप में कार्य करता है। इसमें आमतौर पर जॉयस्टिक जैसे यंत्रणा नियंत्रण के लिए बटन, टेक–ऑफ, लैंडिंग, और अन्य कार्यों के लिए बटन, और टेलीमेट्री डेटा जैसे बैटरी स्तर और सिग्नल शक्ति को प्रदर्शित करने के लिए एक स्क्रीन शामिल होती है।

अतिरिक्त घटक

विशिष्ट अनुप्रयोग के आधार पर, ड्रोन में अतिरिक्त घटक भी शामिल हो सकते हैं जैसे सुचारू टचडाउन के लिए लैंडिंग गियर, रात्रि उड़ान के लिए लाइटिंग सिस्टम, या विशिष्ट भार को ले जाने के लिए पेलोड बे इत्यादि। निर्बाध रूप से एक साथ काम करके, ये घटक

चित्र 2ः स्प्रेइंग ड्रोन (Source: https%//www-indiamart-com)

ड्रोन को आसमान में चढ़ने, उद्योगों में क्रांति लाने और भविष्य के लिए रोमांचक संभावनाएं प्रदान करने में सक्षम बनाते हैं।

कृषि ड्रोन क्या हैं?

कृषि ड्रोन, जिन्हें मानव रहित हवाई वाहन (UAVs) के रूप में भी जाना जाता है, दूरस्थ नियंत्रण वाली उड़ानी मशीनें होती हैं जिनमें विभिन्न सेंसर और कैमरे लगे होते हैं। ये बहुउपयोगी उपकरण स्वतंत्र रूप से या पायलट कंट्रोल के तहत कृषि क्षेत्रों में नेविगेट कर सकते हैं, मूल्यवान डेटा जुटा सकते हैं और विभिन्न कृषि कार्यों को सुविधाजनक बना सकते हैं। कृषि में दो प्रकार के ड्रोन का उपयोग ज्यादातर किया जाता है। वे हैं–

1. स्प्रेइंग ड्रोनः एक छिड़काव ड्रोन, जिसे कृषि ड्रोन या क्रॉप स्प्रेयर ड्रोन के रूप में भी जाना जाता है, एक प्रकार का मानव रहित हवाई वाहन (यूएवी) है जिसे विशेष रूप से फसलों में पानी, उर्वरक आदि जैसे तरल पदार्थ डालने के लिए डिजाइन किया गया है। वे आमतौर पर मल्टी–रोटर ड्रोन (क्वाडकॉप्टर) होते हैं जो विशेष छिड़काव टैंक और बूम से सुसज्जित होते हैं।

2. मल्टीकॉप्टर ड्रोन (कैमरा सेंसर आधारित ड्रोन)ः ये सबसे लोकप्रिय प्रकार के कृषि ड्रोन हैं, जो अपनी गतिशीलता और स्थिरता के लिए जाने जाते हैं। वे विभिन्न आकारों में आते हैं, जिनमें क्वाडकॉप्टर (चार रोटार) सबसे आम हैं। मल्टी–रोटर ड्रोन सटीक होवरिंग और कम ऊंचाई वाली उड़ान की आवश्यकता वाले कार्यों में उत्कृष्ट होते हैं, जिससे वे फसल स्वास्थ्य निगरानी और स्काउटिंग के लिए आदर्श बन जाते हैं।

कृषि में ड्रोन प्रौद्योगिकी के अनुप्रयोग

ड्रोन विभिन्न प्रकार के अनुप्रयोगों द्वारा कृषि परिदृश्य को बदल रहे हैं:

- **फील्ड मैपिंग और विश्लेषणः** उच्च–रिजॉल्यूशन कैमरों से सुसज्जित ड्रोन खेतों के विस्तृत मानचित्र बना सकते हैं, जिससे किसानों को मिट्टी की संरचना, ऊंचाई और जल निकासी पैटर्न में भिन्नता की पहचान करने की सुविधा मिलती है। यह जानकारी बुनाई की रणनीतियों और संसाधन वितरण को बेहतर बनाने के लिए महत्वपूर्ण है।
- **फसल की निगरानी और स्वास्थ्य मूल्यांकनः** मल्टीस्पेक्ट्रल कैमरे दिखाई नहीं देने वाली विशेष डेटा को कैप्चर कर सकते हैं, जो फसल के स्वास्थ्य के बारे में पता लगाने में मदद करते हैं। प्रकाश की परावर्तन की विभिन्नताओं का विश्लेषण करके, किसान प्रारंभिक रोग के संकेत, पोषक तत्वों की कमी, या कीटों की लगातारी की पहचान कर सकते हैं, जिससे उन्हें समय पर सुधारात्मक कार्रवाई करने की सुविधा मिलती है।
- **सटीक छिड़कावः** कीटनाशकों, शाकनाशियों और उर्वरकों के लक्षित अनुप्रयोग के लिए ड्रोन को विशेष

चित्र 3ः कैमरा सेंसर आधारित ड्रोन (Source: www.xboom.in)

स्प्रेयर से सुसज्जित किया जा सकता है। यह सटीक दृष्टिकोण अपशिष्ट को कम करता है, पर्यावरणीय प्रभाव को कम करता है, और यह सुनिश्चित करता है कि केवल आवश्यक क्षेत्रों में छिड़काव किया जाए।

- **सिंचाई प्रबंधनः** ड्रोन पर लगे थर्मल इमेजिंग कैमरे मिट्टी में अलग–अलग नमी के स्तर वाले क्षेत्रों की पहचान कर सकते हैं। यह डेटा लक्षित सिंचाई, पानी की बर्बादी को कम करने और संसाधन उपयोग को अनुकूलित करता है।
- **पशुधन प्रबंधनः** विशाल चराई क्षेत्रों में पशुधन के स्वास्थ्य और स्थान की निगरानी के लिये ड्रोन का उपयोग किया जा सकता है। यह बड़े झुंडों का प्रबंधन करने वाले पशुपालकों या मुश्किल–से–पहुंच वाले इलाकों के लिए विशेष रूप से फायदेमंद है। ड्रोन विशाल चराई क्षेत्रों में बाड़ की अखंडता का जल्दी और कुशलता से निरीक्षण कर सकते हैं, जिससे समय और श्रम की बचत होती है।
- **रोपण और बीजारोपणः** ड्रोन–माउंटेड सीडिंग सिस्टम पूर्व–निर्धारित पैटर्न में बोने बीज पौधों को, विशेष रूप से अनियमित भूमि या वृक्षारोपण प्रयासों के लिए उपयोगी हो सकते हैं।
- **फ्रॉस्ट संरक्षणः** ड्रोन अप्रत्याशित तापमान की बूंदों से फसलों को बचाने के लिए खेतों में विशेष ठंढ संरक्षण सामग्री तैनात कर सकते हैं।
- **कटाई के बाद का आकलनः** ड्रोन का उपयोग फसल के बाद फसल की उपज का आकलन करने के लिए किया जा सकता है, भविष्य की योजना और संसाधन आवंटन के लिए मूल्यवान डेटा प्रदान करता है।

ड्रोन प्रौद्योगिकी के लाभ

कृषि पद्धतियों में ड्रोन प्रौद्योगिकी का एकीकरण कई लाभ प्रदान करता हैः

- **बढ़ी हुई दक्षताः** ड्रोन डेटा एकत्र कर सकते हैं और पारंपरिक तरीकों की तुलना में कार्य को बहुत तेजी से कर सकते हैं। यह किसानों को अधिक क्षेत्र कवर करने में मदद करता है और डेटा के आधार पर तेजी से निर्णय लेने में मदद करता है।
- **बेहतर परिशुद्धताः** लक्षित डेटा संग्रहण और अनुप्रयोग तकनीकें संसाधनों के प्रयोग में अधिक सटीकता लाती हैं, अपशिष्ट को कम करती हैं और उत्पादकता को अधिकतम करती हैं।
- **खेत प्रबंधन का उन्नतीकरणः** ड्रोन द्वारा प्राप्त डेटा किसानों को उनकी खेतों का व्यापक अवलोकन प्रदान करता है, जिससे सुचित निर्णय लेने और खेती के प्रबंधन प्रथाओं में सुधार किया जा सकता है।
- **पर्यावरण संरक्षणः** संसाधनों के सटीक अनुप्रयोग और उन्नत निगरानी क्षमताएँ अधिक सतत कृषि प्रथाओं के लिए योगदान करती हैं।
- **कृषि में ड्रोन प्रौद्योगिकी का भविष्यः** कृषि में ड्रोन प्रौद्योगिकी का भविष्य बड़ी उन्नतियों का वादा करता है। निम्नलिखित कुछ रोमांचक संभावनाएं हैंः
- **कृत्रिम बुद्धिमत्ता (आर्टिफिशियल इंटेलिजेंस) के साथ एकीकरणः** एआई–शक्तियुक्त ड्रोन वास्तविक समय में डेटा का विश्लेषण कर सकेंगे, जिससे वे स्वतः समस्याओं की पहचान करने और सुधारात्मक कार्रवाई लेने में सक्षम होंगे।
- **उन्नत स्वचालनः** पूरी तरह स्वायत्त उड़ान पैटर्न और कार्य पूरा करने की क्षमता वाले ड्रोन अधिक कुशलता प्रदान करेंगे, जिससे मैनुअल संचालन पर निर्भरता को कम किया जा सकेगा।
- **नई सेंसर प्रौद्योगिकियांः** उन्नत सेंसर तकनीक कृषि की सेहत, मिट्टी की स्थिति, और पर्यावरणीय कारकों में गहरी समझ प्रदान करेंगे।

हालांकि, आगे आने वाली चुनौतियों को मानना महत्वपूर्ण है। नियामक ढांचे को कृषि में ड्रोन के बढ़ते प्रयोग को समायोजित करने के लिए अनुकूल बनाया जाना चाहिए। वैसे ही, डेटा गोपनीयता और सुरक्षा को सुनिश्चित करना एक महत्वपूर्ण विचार है। इसके अतिरिक्त, डेटा गोपनीयता और सुरक्षा सुनिश्चित करना महत्वपूर्ण है।

निष्कर्ष

ड्रोन प्रौद्योगिकी कृषि को तेजी से परिवर्तित कर रही है, किसानों के लिए अभियांत्रिक प्रयोगों को अनुकूलित करने, उत्पादकता में सुधार करने, और पर्यावरण संवेदनशीलता को बढ़ावा देने का एक शक्तिशाली उपकरण प्रदान कर रही है। जैसे–जैसे प्रौद्योगिकी विकसित होती रहेगी, कृषि में ड्रोन का उपयोग की संभावनाएं असीमित होंगी। एक विहंगम दृश्य और अनगिनत डेटा के साथ, किसान अधिक आत्मविश्वास और कुशलता के साथ कृषि के भविष्य के रास्ते का समन्वय करने के लिए तैयार हैं।

संदर्भ

बजयार, पी. एवं शैदाई, ई. (2021)। ऍन ओवरव्यू ऑफ मैपिंग एंड स्प्रेइंग ड्रोन सिस्टम्स यूजिंग इन प्रिसिजन एग्रीकल्चर।

गणेशकुमार, सी., डेविड, ए., शंकर, जेजी, एवं सगिनाला, एम. (2023)। एप्लीकेशन ऑफ ड्रोन टेक्नोलॉजी इन एग्रीकल्चरः ए प्रेडिक्टिव फोरकास्टिंग ऑफ पेस्ट एंड डिजीज इन्सिडेन्स। इन बुकः अप्लाईइंग ड्रोन टेक्नोलॉजीज एंड रोबोटिक्स फॉर एग्रीकल्चरल सस्टेनेबिलिटी।

रेनर, के. एवं विल्सन, एम. (2024)। डिस्टेंस एग्जामिनेशन ऑफ लाइवस्टॉक विथ ड्रोन्स – ऍन इफेक्टिव मेथड फॉर अस्सेस्सिन्ग हेल्थ एंड वेलफेयर। ऑस्ट्रेलियन वेटरनरी जर्नल।

साहनी आर., कुमार एसपी, थोराट डी., राजवाड़े वाई., ज्योति बी., रंजन जे., एवं आनंद आर. (2024)। ड्रोन स्प्रेइंग सिस्टम फॉर एफिसिएंट एग्रोकेमिकल एप्लीकेशन इन प्रिसिजन एग्रीकल्चर। इन बुकः एप्लिकेशन्स ऑफ कंप्यूटर विजन एंड ड्रोन टेक्नोलॉजी इन एग्रीकल्चर।

कृषि आवश्यकताओं के व्यख्यान हेतु किसान कॉल सेंटर के किसानों के प्रश्नों का गहन विश्लेषण

समर्थ गोदारा[1], राम स्वरूप बाना[2], श्रुति गोदारा[3], राजेंद्र प्रसाद[1], सुदीप मारवाहा[1]
राजू कुमार एवं दीपक सिंह

[1]भा.कृ.अनु.प. –भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली

[2]भा.कृ.अनु.प. –भारतीय कृषि अनुसंधान संस्थान, नई दिल्ली

[3]भा.कृ.अनु.प. –भारतीय वानिकी अनुसंधान एवं शिक्षा परिषद, देहरादून, उत्तराखंड

सार

भारत में विस्तार सेवाओं को प्रासंगिक और प्रभावशाली बनाए रखने के लिए तेजी से बदलते कृषि क्षेत्र में मांग–संचालित दृष्टिकोण अपनाना महत्वपूर्ण है। इस दिशा मे यह लेख जनवरी 2009 से अगस्त 2023 तक पंजाब, भारत में किसान कॉल सेंटर पर किए गए बीस लाख से अधिक किसान क्वेरी कॉल का व्यापक विश्लेषण प्रस्तुत करता है। ये पूर्व–संसाधित कॉल लॉग किसानों और कृषि सहायता सेवाओं के बीच एक महत्वपूर्ण कड़ी हैं एवं कृषि समुदाय की चुनौतियों और आवश्यकताओं के बारे में बहुमूल्य जानकारी प्रदान करता है । हमारा विश्लेषण किसान क्वेरी कॉल में अस्थायी रुझानों की जांच करके शुरू होता है, जो साल–दर–साल आंकड़े पेश करता है तथा जो समय के साथ कॉल वॉल्यूम के विकास का खुलासा करता है। इसके अलावा, हम इन प्रश्नों की मौसमी जानकारी पर प्रकाश डालते हुए, महीने–वार अंतर्दृष्टि प्रदान करने के लिए डेटासेट का विस्तृत विश्लेषण किया गया है ।कॉल गतिविधि के चरम महीनों की पहचान करके, हम उन महत्वपूर्ण अवधियों को इंगित कर सकते हैं जब किसानों को सहायता की सबसे अधिक आवश्यकता होती है। इसके अलावा, जिलेवार विश्लेषण इन कॉलों के भौगोलिक वितरण को मैप करने में सहायता करता है, जिससे नीति निर्माताओं और कृषि अधिकारियों को अनुरूप हस्तक्षेप के साथ विशिष्ट क्षेत्रों को लक्षित करने में सक्षम बनाया जाता है। फसल श्रेणी और क्वेरी प्रकार के आधार पर हमारा विश्लेषण किसानों की चिंताओं पर एक विस्तृत परिप्रेक्ष्य भी प्रदान करता है। फसलों और क्वेरी प्रकारों के आधार पर कॉलों को वर्गीकृत करके, हम किसानों के सामने आने वाली विशिष्ट चुनौतियों के बारे में बहुमूल्य जानकारी प्राप्त कर सकते हैं। यह जानकारी विभिन्न फसल श्रेणियों और क्वेरी प्रकारों की विशिष्ट आवश्यकताओं को संबोधित करने के लिए तैयार की गई कृषि नीतियों, विस्तार सेवाओं और समर्थन कार्यक्रमों के विकास का मार्गदर्शन कर सकती है। अंततः, यह अध्ययन भारत के कृषक समुदाय की दीर्घकालिक स्थिरता और समृद्धि सुनिश्चित करने के लिए कृषि सहायता प्रणालियों को बढ़ाने के लिए डेटा–संचालित अंतर्दृष्टि का उपयोग करने के महत्व को रेखांकित करता है।

मुख्य शब्दः कृषि प्रश्न, किसान कॉल सेंटर, पंजाब कृषि, किसान सहायता, फसल स्वास्थ्य प्रबंधन।

परिचय

कृषि सदियों से भारत की अर्थव्यवस्था की रीढ़ रही है, जो आबादी के एक महत्वपूर्ण हिस्से को आजीविका प्रदान करती है और आर्थिक विकास के एक महत्वपूर्ण चालक के रूप में कार्य करती है (लौहार एट अल.,

2020)। पंजाब में, जिसे अक्सर ''भारत का अन्न भंडार'' कहा जाता है, कृषि ने क्षेत्र के सामाजिक–आर्थिक परिदृश्य को आकार देने में विशेष रूप से महत्वपूर्ण भूमिका निभाई है। हालाँकि, पंजाब में किसानों के सामने आने वाली चुनौतियाँ विकसित हो गई हैं, जिससे इस महत्वपूर्ण क्षेत्र की निरंतर समृद्धि सुनिश्चित करने के लिए अनुकूली और उत्तरदायी सहायता प्रणालियों की आवश्यकता हो गई है।

भारत के पंजाब में किसान कॉल सेंटर (केसीसी), किसानों और कृषि सहायता सेवाओं के बीच एक महत्वपूर्ण इंटरफेस का प्रतिनिधित्व करता है। पिछले कुछ वर्षों में, यह कृषि सहायता की आधारशिला बन गया है, जो किसानों को उनकी महत्वपूर्ण कृषि चिंताओं (गोदारा और तोशनीवाल, 2022; गोदारा और तोशनीवाल, 2020) के लिए मार्गदर्शन, जानकारी और समाधान प्राप्त करने के लिए एक मंच प्रदान करता है। पंजाब सरकार द्वारा संचालित केसीसी, किसान क्वेरी कॉल के व्यापक डेटासेट को जो फसल की खेती की तकनीक से लेकर कीट प्रबंधन रणनीतियों और सरकारी योजनाओं तक विभिन्न विषयों को रिकॉर्ड और प्रबंधन करता है चाचरा एट अल., 2020; कौर एट अल., 2020)।

यह शोध पत्र भारत के पंजाब में केसीसी द्वारा प्राप्त किसान क्वेरी कॉल के व्यापक विश्लेषण पर आधारित है, जिसमें 2009 से लेकर आज तक के डेटासेट पर ध्यान केंद्रित किया गया है, जिसमें कुल 2,403,592 क्वेरी कॉल लॉग शामिल हैं। इस अध्ययन के उद्देश्य एवं लक्ष्य है:

1. यह देखें कि हर साल कितने किसान कॉल आते हैं। इससे हमें यह देखने में मदद मिलती है कि किसानों को किन समस्याओं का सामना करना पड़ रहा है और हमें उनकी मदद के लिए बेहतर योजनाएँ बनाने में मदद मिलती है।
2. यह देखें कि किसान किन महीनों में सबसे अधिक कॉल करते हैं ताकि यह समझ सकें कि उन्हें सहायता की सबसे अधिक आवश्यकता कब होती है। इससे हमें यह जानने में मदद मिलती है कि कब सहायता देनी है।
3. यह जानने के लिए जांचें कि संसाधन कहां भेजने हैं। किन क्षेत्रों में सबसे अधिक कॉल आती हैं। इससे हमें वहां सहायता देने में मदद मिलती है जहां इसकी सबसे अधिक आवश्यकता होती है।
4. कॉल को इस आधार पर क्रमबद्ध करें कि वे किस फसल से संबंधित हैं और उन्हें किस प्रकार की समस्या है। इससे हमें यह समझने में मदद मिलती है कि प्रत्येक किसान को क्या चाहिए और उनकी बेहतर मदद के लिए योजनाएँ बनाते हैं।

हम यह समझने के लिए अध्ययन करना चाहते हैं कि भारतीय खेती को कैसे बेहतर बना सकते है। हमारे शोध का उद्देश्य नीति निर्माताओं को सबूतों के आधार पर अच्छे निर्णय लेने में मदद करना और किसानों को उनकी बदलती जरूरतों के साथ मदद करना है। हमें उम्मीद है कि इससे किसानों की खेती मजबूत रहेगी, ताकि यह भोजन उपलब्ध करा सके और देश के विकास को समर्थन दे सके।

कार्यप्रणाली

इस शोध में नियोजित पद्धति में पंजाब, भारत में किसान कॉल सेंटर (केसीसी) से किसान क्वेरी कॉल डेटासेट का विश्लेषण करने के लिए संरचित चरणों की एक श्रृंखला शामिल थी। ये चरण इस प्रकार बताए गए हैं:

1. **डेटा संग्रहः** इस अध्ययन के लिए प्राथमिक डेटा स्रोत केसीसी डेटासेट था, जो पंजाब राज्य सरकार के डेटा सर्वर से प्राप्त किया गया था। इस डेटासेट में किसान क्वेरी कॉल के बारे में विस्तृत जानकारी शामिल है, जिसमें कॉल की तारीख और समय, कॉल करने वाले का स्थान (जिला), चर्चा की गई फसल , क्वेरी प्रकार और कॉल के दौरान प्रदान किया गया कोई भी प्रासंगिक विवरण शामिल है।
2. **डेटा प्रीप्रोसेसिंगः** डेटासेट की गुणवत्ता और स्थिरता सुनिश्चित करने के लिए डेटा प्रीप्रोसेसिंग आवश्यक

थी। इसमें लापता मानों, डुप्लिकेट और आउटलायर्स को संभालकर डेटा को साफ करना शामिल था। इसके अतिरिक्त, दिनांक प्रारूपों को मानकीकृत किया गया था, और डेटा को आगे के विश्लेषण के लिए विशिष्ट फसल और क्वेरी प्रकारों में वर्गीकृत किया गया था। विभिन्न समय–सीमाओं (उदाहरण के लिए, वर्ष–वार, माह–वार) और स्थानिक पैमाने (उदाहरण के लिए, जिला–वार) के लिए सबसेट बनाने के लिए डेटा एकत्र किया गया था।

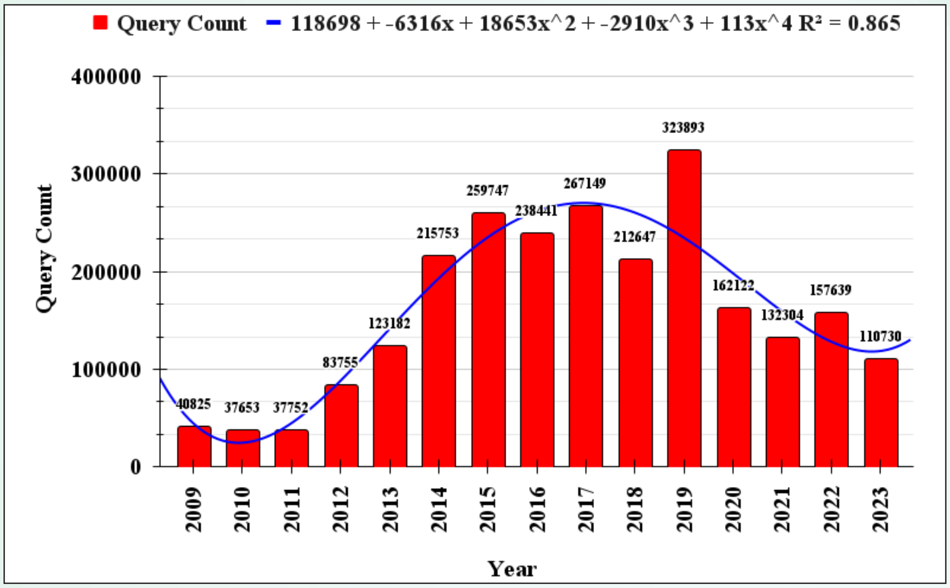
3. **अस्थायी विश्लेषण**: वर्ष–वार और माह–वार रुझानों का विश्लेषण करने के लिए, हमने प्रत्येक वर्ष और माह के लिए क्वेरी कॉल की कुल संख्या की गणना की। इससे यह जानकारी मिली कि कॉल की मात्रा कैसे विकसित हुई है और कॉल गतिविधि के चरम महीनों की पहचान की गई।
4. **भू–स्थानिक विश्लेषण**: प्रत्येक जिले की कुल कॉलों की गणना करके क्वेरी कॉल के जिले–वार वितरण का विश्लेषण किया गया था। इससे हमें उच्च कॉल वॉल्यूम वाले क्षेत्रों की पहचान करने में सहायता मिली।
5. **श्रेणीबद्ध विश्लेषण**: कॉलों को विभिन्न फसल और क्वेरी प्रकारों में वर्गीकृत किया गया था , और प्रत्येक श्रेणी के लिए कॉल की संख्या की गणना की गई थी। इस विश्लेषण ने विभिन्न फसलों और क्वेरी प्रकारों के बारे में किसानों के सामने आने वाली विशिष्ट चुनौतियों के बारे में जानकारी प्रदान की।
6. **डेटा विजुअलाइजेशन**: चार्ट और ग्राफ जैसे दृश्य प्रतिनिधित्व, निष्कर्षों को चित्रित करने के लिए बनाए गए थे, जिससे दर्शकों के लिए परिणामों को समझना और व्याख्या करना आसान हो गया।
7. **व्याख्या और अनुमान**: विश्लेषण से प्राप्त परिणामों की व्याख्या किसान क्वेरी कॉल में अस्थायी, भौगोलिक और श्रेणीबद्ध रुझानों के संबंध में सार्थक निष्कर्ष निकालने के लिए की गई थी।
8. **नीति अनुशंसाएँ**: शोध निष्कर्षों का उपयोग पंजाब में किसानों के सामने आने वाली चुनौतियों की जानकारी प्राप्त करने के लिए किया गया था। इन जानकारियों ने पहचाने गए मुद्दों के समाधान के लिए संभावित नीतिगत निहितार्थों और हस्तक्षेपों की जानकारी दी।

संक्षेप में, इस शोध में डेटा संग्रह, प्रीप्रोसेसिंग, विश्लेषण और व्याख्या को शामिल करते हुए एक संरचित पद्धति का पालन किया गया। जिसमे डेटासेट से अंतर्दृष्टि प्राप्त करने के लिए गणितीय गणनाओं को नियोजित किया। पंजाब में किसान क्वेरी कॉल की गतिशीलता के बारे में बहुमूल्य जानकारी प्रदान की और कृषि नीति और सहायता सेवाओं में साक्ष्य–आधारित निर्णय लेने की नींव के रूप में कार्य किया।

परिणाम

भारत के पंजाब में 2009 से 2023 तक किसान क्वेरी कॉल के वर्ष–वार विश्लेषण से उल्लेखनीय उतार–चढ़ाव और रुझान का पता चलता है (चित्र 1)। डेटा 2009 से 2012 तक घटती प्रवृत्ति को दर्शाता है, जो कृषि पद्धतियों और ज्ञान में सुधार का सुझाव देता है। हालाँकि, 2013 के बाद से, बढ़ती कृषि चुनौतियों को दर्शाते हुए, क्वेरी कॉल में लगातार वृद्धि हुई है। 2019 में महत्वपूर्ण शिखर का श्रेय विशिष्ट कृषि मुद्दों या सरकारी नीतियों को दिया जा सकता है। COVID–19 महामारी, सूचना और संचार प्रौद्योगिकी का विकास, अवधि के दौरान मौसम की घटनाएं और आर्थिक स्थिति जैसे कारक भी कॉल वॉल्यूम को प्रभावित कर सकते हैं । कुल मिलाकर, डेटा पंजाब के किसानों की गतिशील जरूरतों को पूरा करने के लिए अनुकूली और उत्तरदायी कृषि सहायता प्रणालियों की आवश्यकता को रेखांकित करता है।

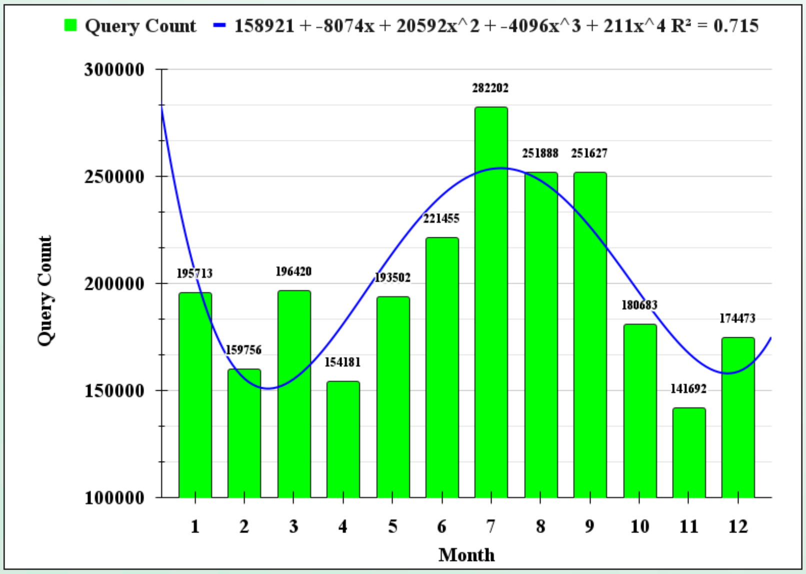
चित्र 2 पंजाब, भारत में किसान क्वेरी कॉल का माहवार विश्लेषण प्रदर्शित करता है। यह क्वेरी वॉल्यूम में विशिष्ट मौसमी पैटर्न और रुझानों को प्रकट करता है। जून से सितंबर (मानसून और मानसून के बाद की शुरुआती अवधि) के महीनों में लगातार उच्चतम कॉल



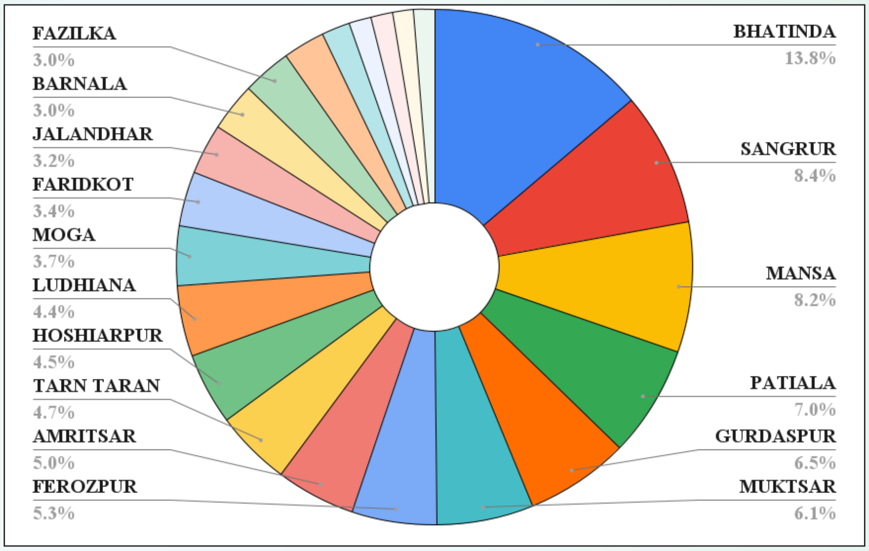
चित्र 1. पंजाब, भारत में 2009 से 2023 तक किसान कॉल की वर्ष–वार क्वेरी संख्या

वॉल्यूम दर्ज किया जाता है, जुलाई में 2.82 लाख कॉल अपने चरम पर होती हैं। यह रोपण अवधि और बढ़ते मौसम के दौरान मौसम संबंधी चिंताओं की महत्वपूर्ण भूमिका को दर्शाता है (बाना एट अल., 2022)। इसके अलावा, शुरुआती सीजन में खरपतवार का खतरा, बीज उपचार और पोषक तत्व प्रबंधन संबंधी प्रश्न भी इस अवधि के दौरान चिंता का विषय बने हुए हैं (बाना एट अल., 2020; अरोड़ा एट अल., 2022)।

इसके विपरीत, नवंबर से फरवरी तक के महीनों में क्वेरी की संख्या कम होती है, जो सर्दियों के दौरान खेती की गतिविधियों में कमी का संकेत देता है। जनवरी में प्रश्नों में थोड़ी वृद्धि देखी गई है, संभवतः आगामी फसल मौसम की तैयारी या रबी फसलों के प्रजनन चरणों में कीट–पतंगों और बीमारियों के संक्रमण के कारण। कुल मिलाकर, डेटा किसानों के प्रश्नों पर कृषि मौसम और मौसम की स्थिति के मजबूत प्रभाव को उजागर करता है,



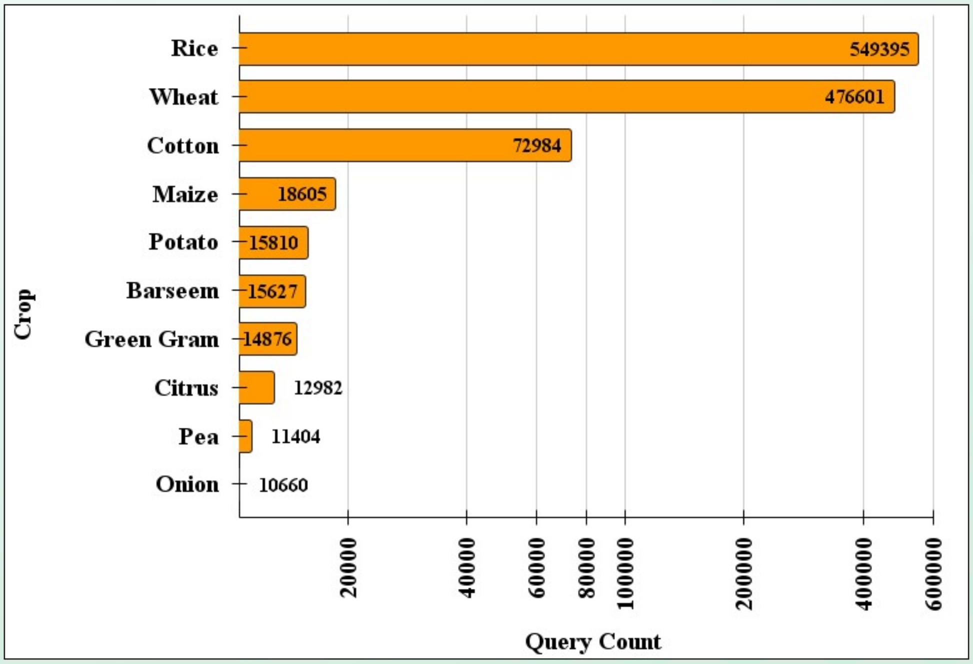
चित्र 2. पंजाब, भारत में 2009 से 2023 तक किसान कॉल की क्वेरी संख्या का माहवार विश्लेषण



चित्र 3. पंजाब, भारत में 2009 से 2023 तक किसान कॉल की क्वेरी संख्या का जिला–वार विश्लेषण

समय पर और क्षेत्र–विशिष्ट समर्थन और सूचना प्रसार की आवश्यकता पर जोर देता है।

जिलेवार विश्लेषण से विभिन्न क्षेत्रों में कॉल वॉल्यूम में उल्लेखनीय भिन्नता का पता चलता है (चित्र 3)। सबसे अधिक 3.32 लाख क्वेरी के साथ भटिंडा महत्वपूर्ण कृषि चिंताओं वाले जिले के रूप में सामने आया है। संगरूर और मनसा में भी बड़ी संख्या में कॉल की सूचना मिलती है, जो इन क्षेत्रों में कृषि संबंधी मुद्दों की व्यापकता का संकेत देता है। किसानों की विविध कृषि चुनौतियों पर जोर देते हुए, पटियाला और गुरदासपुर बारीकी से अनुसरण करते हैं। इसके विपरीत, कपूरथला और साहिबजादा जैसे जिले अजीत सिंह नगर की रिपोर्ट में कम क्वेरी संख्या बताई गई है, जो इन क्षेत्रों में अपेक्षाकृत कम तात्कालिक कृषि चिंताओं का संकेत



चित्र 4. पंजाब, भारत की शीर्ष–10 फसलों में किसानों की कृषि प्राथमिकताएँ और चिंताएँ

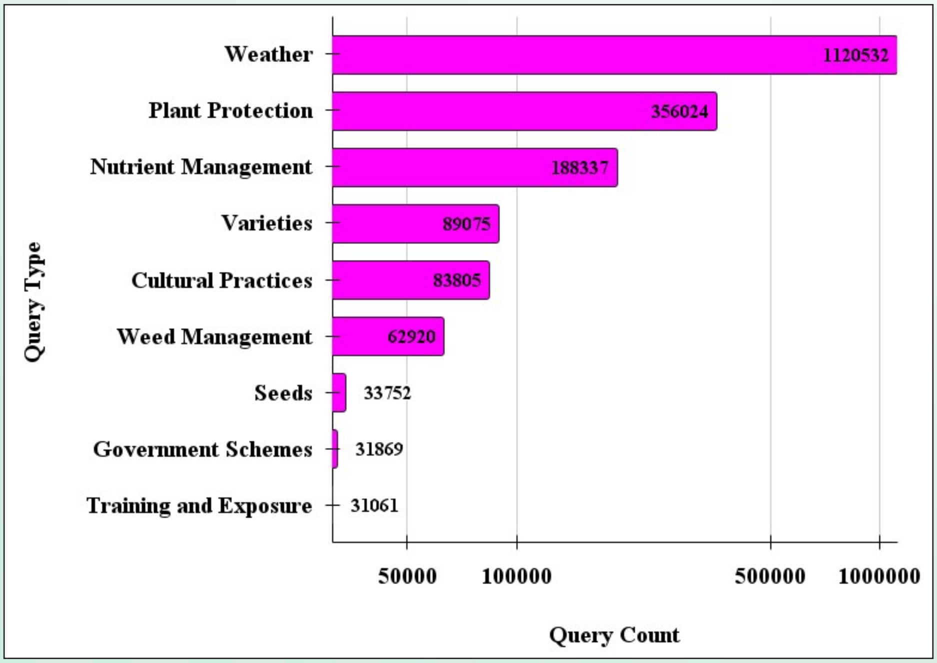
देती है। ये विविधताएँ विभिन्न जिलों में किसानों की विशिष्ट आवश्यकताओं को पूरा करने के लिए समर्थन और हस्तक्षेप के महत्व को रेखांकित करती हैं।

चित्र 4 में शीर्ष 10 फसल क्वेरी विश्लेषण, पंजाब में किसानों की कृषि प्राथमिकताओं और चिंताओं के बारे में महत्वपूर्ण अंतर्दृष्टि प्रकट करता है। सूची में चावल (धान) और गेहूं का दबदबा है, जो राज्य की कृषि में उनकी महत्वपूर्ण भूमिका को रेखांकित करता है। कपास (कपास) का बारीकी से अनुसरण किया जाता है, जो राज्य में संभावित विविधीकरण विकल्प के रूप में फसल के महत्व और कपड़ा उद्योग में इसकी भूमिका को दर्शाता है। मक्का (मक्का) की फसल चौथे स्थान पर होने के साथ, पंजाब राज्य की फसल प्रणाली विविधीकरण प्रयासों में इसकी क्षमता को उजागर करती है (बम्बोरिया एट अल., 2022)। आलू, बरसीम और हरा चना फसल विविधता का प्रदर्शन, खेती की जाने वाली विभिन्न फसलों पर प्रकाश डालता है। साइट्रस, मटर (सब्जी), और प्याज बागवानी और फलीदार फसलों की प्रासंगिकता को दर्शाता है। ये क्वेरी गणना क्षेत्र की कृषि नीतियों और विस्तार सेवाओं को प्रभावित करने वाली इन फसलों से संबंधित जानकारी और समर्थन की आवश्यकता पर जोर देती हैं।

क्वेरी प्रकार–वार विश्लेषण पंजाब में किसानों के विशिष्ट विषयों का एक व्यापक अवलोकन प्रदान करता है (चित्र 5)। 1, 120, 532 की महत्वपूर्ण क्वेरी संख्या के साथ ''मौसम'' सबसे अधिक बार पूछे जाने वाले विषय के रूप में उभरा है, जो खेती के निर्णय लेने में मौसम से संबंधित जानकारी की महत्वपूर्ण भूमिका को उजागर करता है। ''पौधा संरक्षण'' और ''पोषक तत्व प्रबंधन'' का अनुसरण किया जाता है, जो कीट और रोग नियंत्रण और पोषक तत्वों के उपयोग को अनुकूलित करने में गहरी रुचि का संकेत देता है।

''सांस्कृतिक प्रथाएं'' और ''खरपतवार प्रबंधन'' फसल की पैदावार बढ़ाने में कृषि संबंधी प्रथाओं के महत्व को दर्शाते हैं। ''बीज'' से संबंधित प्रश्न फसल बीज की गुणवत्ता और चयन पर ध्यान केंद्रित करने का सुझाव देते हैं।

इसके अलावा, सरकारी योजनाएं, प्रशिक्षण एवं एक्सपोजर, उस व्यापक समर्थन और संसाधनों का प्रतिनिधित्व करते हैं जो किसान सरकारी कार्यक्रमों और



चित्र 5. पंजाब, भारत में 2009 से 2023 तक किसान कॉलों का प्रश्न–प्रकार–वार विश्लेषण

प्रशिक्षण पहलों से चाहते हैं। यह विश्लेषण कृषि संबंधी प्रश्नों की बहुमुखी प्रकृति को रेखांकित करता है, पंजाब में कृषि उत्पादकता और स्थिरता को बढ़ाने के लिए इन महत्वपूर्ण क्षेत्रों में अनुरूप विस्तार सेवाओं, समय पर सूचना प्रसार और समर्थन की आवश्यकता पर जोर देता है ।

चर्चा

किसान क्वेरी कॉल के मासिक विश्लेषण के आधार पर कई नीतिगत सिफारिशें सामने आती हैं। कॉल वॉल्यूम में विशिष्ट मौसमी पैटर्न को संबोधित करने के लिए, नीति निर्माताओं को पीक महीनों (जून से सितंबर) के दौरान समय पर सूचना प्रसार को प्राथमिकता देनी चाहिए जब मौसम संबंधी चिंताएं प्रचलित होती हैं। निचले क्वेरी महीनों (नवंबर से फरवरी) के दौरान, किसानों को आगामी फसल सीजन के लिए तैयार करने के लिए प्रशिक्षण और क्षमता निर्माण पहल सहित ऑफ–सीजन समर्थन की ओर प्रयास करना चाहिए। विशिष्ट प्रश्नों के लिए डेटा संग्रह तंत्र को मजबूत करना और साल भर सूचना पहुंच प्रदान करने के लिए डिजिटल प्लेटफॉर्म का लाभ उठाना महत्वपूर्ण है। सार्वजनिक–निजी भागीदारी और जागरूकता अभियान अधिक संवेदनशील और किसान–केंद्रित दृष्टिकोण सुनिश्चित करते हुए कृषि सहायता सेवाओं की प्रभावशीलता को और बढ़ा सकते हैं (गोदारा एट अल., 2022)।

जिलावार विश्लेषण लक्षित और क्षेत्र–विशिष्ट नीतिगत हस्तक्षेप की आवश्यकता पर प्रकाश डालता है। उच्च क्वेरी संख्या के साथ, भटिंडा, संगरूर और मनसा जैसे जिलों को प्रचलित कृषि चुनौतियों का समाधान करने के लिए केंद्रित समर्थन मिलना चाहिए। इसके विपरीत, कपूरथला और साहिबजादा जैसे कम क्वेरी संख्या वाले जिले अजीत सिंह नागर, सक्रिय ज्ञान हस्तांतरण कार्यक्रमों और कृषि विस्तार सेवाओं से लाभ उठा सकते हैं। नीति निर्माताओं को इस विश्लेषण से प्राप्त डेटा–संचालित अंतर्दृष्टि का लाभ उठाकर अपने निर्णयों और संसाधन आवंटन को प्रत्येक जिले की विशिष्ट आवश्यकताओं के अनुरूप बनाना चाहिए, अंततः टिकाऊ कृषि को बढ़ावा देना चाहिए और पूरे पंजाब में किसानों की भलाई को बढ़ाना चाहिए।

किसान क्वेरी कॉल में चर्चा की गई शीर्ष 10 फसलों का विश्लेषण पंजाब की कृषि में विशिष्ट फसलों के महत्व को रेखांकित करता है। नीति निर्माताओं को धान, गेहूँ और कपास के लिए संसाधनों और अनुसंधान प्रयासों को प्राथमिकता देनी चाहिए, जो सूची में प्रमुख हैं और राज्य की अर्थव्यवस्था के लिए महत्वपूर्ण हैं। फसल विविधीकरण को प्रोत्साहित करने और कुछ फसलों पर निर्भरता कम करने के लिए आलू, मक्का और बरसीम जैसी वैकल्पिक फसलों के लिए लक्षित समर्थन और प्रोत्साहन लागू किया जाना चाहिए बाना एट अल., 2015)। इसके अतिरिक्त, इन फसलों के लिए ज्ञान प्रसार और विस्तार सेवाओं को बढ़ावा देने से कृषि उत्पादकता और लचीलापन बढ़ सकता है, जिससे अंततः कृषक समुदाय और कृषि क्षेत्र को लाभ होगा।

क्वेरी प्रकार–वार विश्लेषण पंजाब में किसानों की महत्वपूर्ण सूचनात्मक आवश्यकताओं पर प्रकाश डालता है, कृषि प्रथाओं और उत्पादकता को बढ़ाने के लिए विशिष्ट नीति निर्देशों का सुझाव देता है। ''मौसम'' प्रश्नों के महत्व को देखते हुए, नीति निर्माताओं को सटीक और समय पर मौसम पूर्वानुमान और सलाह की पहुंच को प्राथमिकता देनी चाहिए। ''पौधा संरक्षण'' और ''पोषक तत्व प्रबंधन'' चिंताओं को संबोधित करने के लिए एकीकृत कीट प्रबंधन प्रथाओं को बढ़ावा देने और टिकाऊ विकल्पों (योगी एट अल., 2023) सहित पोषक तत्वों के उपयोग को अनुकूलित करने की आवश्यकता है ।

इसके अलावा, ''उर्वरक उपयोग'' और ''किस्में'' जैसे विषयों के लिए, किसानों को उर्वरकों और फसल की किस्मों के बारे में सूचित निर्णय लेने में सहायता करने के लिए अनुरूप मार्गदर्शन और सूचना प्रसार पर जोर दिया जाना चाहिए। इसके अतिरिक्त, विस्तार सेवाओं और प्रशिक्षण के माध्यम से ''सांस्कृतिक प्रथाओं'' और ''खरपतवार प्रबंधन'' का समर्थन करने से फसल

प्रबंधन तकनीकों को बढ़ाया जा सकता है (गांधी एट अल., 2019; बर्नार्ड एट अल., 2019)। अंत में, ''सरकारी योजनाओं'' और ''प्रशिक्षण और एक्सपोजर'' अवसरों के बारे में जागरूकता और पहुंच को बढ़ावा देने से किसानों को मूल्यवान संसाधनों और ज्ञान के साथ सशक्त बनाया जा सकता है, जिससे अंततः पंजाब के कृषि क्षेत्र की लचीलापन और स्थिरता को बढ़ावा मिलेगा।

निष्कर्ष

निष्कर्षतः, भारत के पंजाब में किसान कॉल सेंटर (केसीसी) में किसान क्वेरी कॉल के एक दशक से अधिक समय के इस व्यापक अध्ययन ने क्षेत्र में कृषि के गतिशील परिदृश्य में मूल्यवान अंतर्दृष्टि प्रदान की है। अस्थायी विश्लेषण से पिछले कुछ वर्षों में क्वेरी कॉल वॉल्यूम में उतार–चढ़ाव का पता चला, जिसमें हाल के दिनों में उल्लेखनीय वृद्धि हुई है। ये रुझान किसानों की बढ़ती जरूरतों और चुनौतियों को रेखांकित करते हैं और उत्तरदायी कृषि सहायता प्रणालियों की महत्वपूर्ण भूमिका को उजागर करते हैं। जिलावार विश्लेषण पंजाब के विभिन्न क्षेत्रों में अलग–अलग कृषि संबंधी चिंताओं पर प्रकाश डालता है। इसने कृषक समुदायों की अद्वितीय कृषि प्राथमिकताओं को पहचानते हुए, प्रत्येक जिले की विशिष्ट आवश्यकताओं को पूरा करने के लिए सिलाई समर्थन और हस्तक्षेप के महत्व पर जोर दिया । शीर्ष 10 फसलों और क्वेरी प्रकारों की जांच ने किसानों के दिमाग में विशिष्ट फसलों और विषयों के महत्व को रेखांकित किया। इसने इन महत्वपूर्ण फसलों और कृषि विषयों से संबंधित लक्षित जानकारी और संसाधन प्रदान करने के महत्व को सुदृढ़ किया। कुल मिलाकर, यह अध्ययन पंजाब में नीति निर्माताओं, कृषि विशेषज्ञों और हितधारकों के लिए एक मूल्यवान संसाधन है। यह डेटा–संचालित निर्णय लेने की आवश्यकता और कृषक समुदाय की उभरती जरूरतों के लिए कृषि सहायता सेवाओं को अपनाने के महत्व को रेखांकित करता है। इस शोध के निष्कर्षों के साथ नीतियों और पहलों को जोड़कर, पंजाब की कृषि बदलती चुनौतियों का सामना कर सकती है और राज्य की अर्थव्यवस्था और खाद्य सुरक्षा में महत्वपूर्ण योगदान देना जारी रख सकती है।

संदर्भ

अरोड़ा, के., बाना, आर. एस. एवं सीपत, एस. (2022)। पोटैशियम मैनेजमेंट एंड रेसिडुयु रिसाइकलिंग इफेक्ट ओन व्हीट (ट्रिटिकम एस्टीवम) अंडर मेज (जिया मेयस) व्हीट रोटेशन। इंडियन जर्नल ऑफ एग्रीकल्चरल साइंसेज, 92 (12): 1517–1519।

बम्बोरिया, एस.डी., बाना, आर.एस., कुरी, बी.आर., कुमार, वी., बम्बोरिया, एस. एवं मीना, आर.पी (2022)। अचीविंग हायर प्रोडक्शन फ्रॉम लो इनपुट उसींग सिनर्जिस्टिक क्रॉप इंटरेक्शन्स अंडर मेज बेस्ड पोलीकल्चर सिस्टम्स। एन्वाइरोमेन्टल सस्टनेबिलिटी, 5, 145–159।

बाना, आर.एस., बंबोरिया, एस.डी., पडरिया, आर. एन., धाकड़, आर.के., खसवान, एस.एल., चौधरी, आरएल., एवं बंबोरिया, जे.एस. (2022)। प्लाटिंग पीरियड इफेक्ट्स ऑन व्हीट प्रोडक्टिविटी एंड वॉटर फुटप्रिंट्सः इनसाइड थ्रू अडाप्टिव ट्रायल्स एंड अपसिम सिम्युलेशन्स, एग्रोनोमी, 12, 226।

सी., गोयम, सिंह, ह., सरकार, म. और जैन अ. (2020)। डी इम्पैक्ट ऑफ किसान कॉल सेन्टर ऑन फार्मिंग सेक्टर। "एन्वाइरोमेन्टल एंड एग्रीकल्चरल इन्फार्मेटिक्स, कॉन्सेप्ट, मेथोडोलोजिज टूल्स एंड ऍप्लिकेशन्स"। आईजीआई ग्लोबल, 66–78।

गांधी, वी.पी. एवं जॉनसन, नि. (2019)। डिसीजन ओरिएंटेड इनफार्मेशन सिस्टम्स फॉर फार्मर्सः ए स्टडी ऑफ किसान कॉल सेंटर्स (के.सी.सी), किसान नौलेगे मैनेजमेंट सिस्टम (के. के. ऍम. स), फार्मर्स पोर्टल एंड एम– किसान पोर्टल। एग्रीकल्चरल सिचुएशन इन इंडिया, 76(6): 29–32।

गोदारा, एस. एवं तोशनीवाल, डी. (2022)। डीप लर्निंग बेस्ड क्वेरी काउंट फोरकास्टिंग उसिंग फार्मर्स हैल्प लाइन डाटा। कंप्यूटर एंड इलेक्ट्रॉनिक्स इन एग्रीकल्चरल, 196: 106875।

गोदारा, एस. एवं तोशनीवाल, डी. (2020)। सीकवुहँश्ल

पैटर्न माइनिंग कमबाइंड मल्टीक्रिटेरिआ डिसिशन मेकिंग फॉर फार्मर्स क्वेरीज करक्टराइजेशन। कंप्यूटर एंड इलेक्ट्रॉनिक्स इन एग्रीकल्चरल, 173, 105448।

कौर, ए., एट अल. (2020)। इम्पैक्ट ऑफ किसान कॉल सेंटर इन पंजाब। इंडियन जर्नल ऑफ एक्सटेंशन एजुकेशन, 56(3), 68–74।

बर्नार्ड, जेएनएफ (2019)। ए स्टडी ऑफ प्रोफॉर्मन्स ऑफ किसान कॉल सेंटर्स। थिंक इंडिया जर्नल 22(10): 8411–8417।

न्यूट्रोसोफिक आंकड़ों हेतु सांख्यिकीय तकनीकें

आशुतोष दलाल, सिनी वर्गीस, मोहम्मद हारुन एवं देवेन्द्र कुमार
भा.कृ.अनु.प. – भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली

सार

कई बार, आंकड़ों में कुछ हद तक अनिश्चितता (अस्पष्ट, धुंधला, आंशिक रूप से अज्ञात, विरोधाभासी, अपूर्ण, आदि) होती है। सीमा या अस्पष्ट मूल्यों को नजरअंदाज करना और मध्य मूल्य लेना सदैव उचित नहीं है। अतः, यहां अस्पष्ट तर्क अवधारणा आती है जो आगे चलकर न्यूट्रोसोफिक तर्क में सामान्यीकृत हो जाती है। न्यूट्रोसोफिक सांख्यिकी, शास्त्रीय सांख्यिकी का एक विस्तार है जो स्पष्ट मूल्यों के बजाय निर्धारित मूल्यों से संबंधित है। यह दर्शन और गणित की एक नई शाखा का वर्णन करता है जो अनिश्चित और धुंधली अवधारणाओं के अध्ययन से संबंधित है, जहां किसी कथन की सत्यता या असत्यता को सटीकता से निर्धारित नहीं किया जा सकता है। अस्पष्टता या संदिग्ध आंकड़े प्रबंधन को छोड़कर यह शास्त्रीय आंकड़ों से बहुत अलग नहीं है। यहां, न्यूट्रोसोफिक तर्क के एक परिचय द्वारा यह दिखाने के लिए चर्चा की गई है कि न्यूट्रोसोफिक सांख्यिकी कैसे अनिश्चितता के साथ स्थितियों को संभालने में सक्षम है। न्यूट्रोसोफिक तर्क में, एक चर **X** को तीन मूल्यों अर्थात., $X = (t,i,f)$, द्वारा वर्णित किया जाता है, जहां **t** 'सत्यता' की कोटि है, **f** 'असत्यता' की कोटि है और **i** 'अनिश्चितता' का स्तर है। किसी न्यूट्रोसोफिक आंकड़े **N** को $\mathbf{N} = d + i$ के रूप में व्यक्त किया जा सकता है, जहां d, **N** का निर्धारित (निश्चित) भाग है, और *i*, **N** का अनिर्धारित (अनिश्चित) भाग है ।

मुख्य शब्दः सटीक आंकड़े, अस्पष्ट तर्क, न्यूट्रोसोफिक सांख्यिकी, विचरण का विश्लेषण, यादृच्छिक पूर्ण खण्ड अभिकल्पना।

परिचय

शब्द ''न्यूट्रोसोफिक'' ग्रीक शब्द ''न्यूट्रसोफिया'' से लिया गया है, जो ''न्यूट्रस'' (जिसका अर्थ है तटस्थ) और ''सोफिया'' (जिसका अर्थ है ज्ञान) को जोड़ता है। रोमानियाई–अमेरिकी गणितज्ञ, दार्शनिक और लेखक फ्लोरेंटिन स्मारांडचे ने 1990 के दशक के अंत में इस शब्द का आरंभ किया। न्यूट्रोसोफिक तर्क सांख्यिकी, मशीन लर्निंग, कृत्रिम बुद्धिमत्ता, संचालन अनुसंधान, कतार सिद्धांत, संभाव्यता सिद्धांत, दूरसंचार प्रणाली, गतिशील प्रणाली, अमूर्त बीजगणित आदि सहित विज्ञान के कई क्षेत्रों में आंकड़ों में अनिश्चितता को नजरअंदाज नहीं करता है। यह अस्पष्ट, धुंधला, अंतर्ज्ञानवादी तर्क का एक विस्तारध्संयोजन है। न्यूट्रोसोफिक तर्क में, आसान तरीके से, प्रत्येक तार्किक चर **X** को एक क्रमित तीन मूल्यों, $\mathbf{X = (t, i, f)}$ द्वारा वर्णित किया जाता हैय जहां **t** सत्य की कोटि है, **f** असत्य की कोटि है और **i** अनिश्चितता का स्तर है । अस्पष्ट, धुंधला तर्क में केवल सदस्यता फलन सम्मिलित हैय अंतर्ज्ञानवादी तर्क, सदस्यता के साथ–साथ गैर–सदस्यता कार्य पर भी विचार करता है। दूसरी ओर, न्यूट्रोसोफिक तर्क में सदस्यता, गैर–सदस्यता कार्य और अनिश्चितता सम्मिलित है। इसलिए, यह अस्पष्ट, धुंधला तर्क का सामान्यीकरण है।

न्यूट्रोसोफिक सांख्यिकी शास्त्रीय सांख्यिकी का विस्तार है और यह स्पष्ट मूल्यों के बजाय निर्धारित मूल्यों से संबंधित है। अधिकांश शास्त्रीय सांख्यिकी समीकरणों और सूत्रों में, कई संख्याओं को समुच्चय से बदल देता है। परिणामस्वरूप, संख्याओं के साथ संचालन के बजाय,

कोई समुच्चय के साथ संचालन का उपयोग करता है। सामान्यतः अनिश्चित (अशुद्ध, अनिश्चित और यहां तक कि पूरी तरह से अज्ञात) प्राचल को प्रतिस्थापित करता है। इसीलिए यह परिपाटी बनाई गई है कि किसी भी संख्या **a** को एक समुच्चय **aN** से प्रतिस्थापित कर दिया जाता है, जिसका अर्थ है न्यूट्रोसोफिक **a,** या अस्पष्ट, अनिश्चित **a** । **aN, a** का पड़ोसी हो सकता है, एक अंतराल हो सकता है जिसमें **a** सम्मिलित है, और सामान्य तौर पर यह कोई भी समुच्चय हो सकता है जो **a** का अनुमान लगाता है। जब **a** से संबंधित कोई अनिश्चितता न हो, तो **aN ¾ a** । जबकि शास्त्रीय सांख्यिकी में आंकड़े ज्ञात होते है, जो स्पष्ट संख्याओं से बनते है, न्यूट्रोसोफिक सांख्यिकी में आंकड़ों में कुछ अनिश्चितता होती है।

न्यूट्रोसोफिक आंकड़ों में, आंकड़े अस्पष्ट, धुंधले, सटीक, अधूरे, यहां तक कि अज्ञात भी हो सकते हैं। शास्त्रीय आँकड़ों में उपयोग की जाने वाली स्पष्ट संख्याओं के बजाय, न्यूट्रोसोफिक आँकड़ों में समुच्चय (जो क्रमशः इन स्पष्ट संख्याओं का अनुमान लगाते हैं) का उपयोग करते हैं। शास्त्रीय सांख्यिकी में, सभी आंकड़े निर्धारित किए जाते हैंय यह न्यूट्रोसोफिक सांख्यिकी और शास्त्रीय सांख्यिकी के बीच अंतर है। कई मामलों में, जब अनिश्चितता शून्य होती है, तो न्यूट्रोसोफिक आँकड़े शास्त्रीय आँकड़ों से मेल खाते हैं।

न्यूट्रोसोफिक आंकड़ों के प्रकार

एक न्यूट्रोसोपिकल सांख्यिकीय संख्या **N,** $N + i$ का रूप है, जहां **d, N** का निश्चित भाग है, और **i, N** का अनिश्चित भाग है ।

न्यूट्रोसोफिक आँकड़ों को प के आधार पर असतत एवं निरंतर प्रकार और मात्रात्मक (संख्यात्मक) एवं गुणात्मक (श्रेणीबद्ध) न्यूट्रोसोफिक आँकड़ों में वर्गीकृत किया जा सकता है।

परिस्थिति

मिट्टी की 50% की नमी माप को (50%, 20%, 30%) के रूप में दर्शाया जा सकता है, जहां 50% सत्य की कोटि का प्रतिनिधित्व करता है, 20% अनिश्चितता की कोटि का प्रतिनिधित्व करता है, और 30% असत्यता की कोटि का प्रतिनिधित्व करता है। अनिश्चितता की कोटि, माप त्रुटि या मिट्टी की संरचना में परिवर्तनशीलता जैसे कारकों के कारण माप से जुड़ी अनिश्चितता का प्रतिनिधित्व करती है, जबकि असत्यता की कोटि गलत अंशांकन या दोषपूर्ण उपकरण जैसे कारकों के कारण माप गलत होने की संभावना का प्रतिनिधित्व करती है।

न्यूट्रोसोफिक संख्याओं पर संचालन

मान लीजिए $N_1 \in a, b], N_2 \in [d, c]$ दो न्यूट्रोसोफिक संख्या हैं, तो संचालन निम्नलिखित विधि से किया जा सकता है:

$$N_1 + N_2 = [a + d, b + c] \qquad N_1 - N_2 = [a - c, b - d]$$

$$N_1 * N_2 = [\min\{a*c, b*d, a*b, b*d,\} \max\{a*, c\ b*d, a*b, b*d\}]$$

$$\frac{N_1}{N_2} = \left[\frac{a}{c}, \frac{b}{d}\right] \qquad \sqrt{N_1} = \left[\sqrt{a}, \sqrt{b}\right];\ a, b \geq 0$$

न्यूट्रोसोफिक परिकल्पना

शास्त्रीय (सांख्यिकी) परिकल्पना और न्यूट्रोसोफिक परिकल्पना के बीच अंतर यह है कि न्यूट्रोसोफिक आंकड़ों में जनसंख्या विशेषताओं का वर्णन करने वाले चर न्यूट्रोसोफिक हैं या उन मूल्यों के लिए जिनकी हम तुलना करते हैं, जनसंख्या विशेषताओं में से कम से कम एक न्यूट्रोसोफिक है ।

इसी प्रकार, शास्त्रीय आँकड़ों के लिए NH_0 द्वारा निरूपित एक न्यूट्रोसोफिक शून्य परिकल्पना, वह कथन है जिसे प्रारम्भ में सत्य माना जाता है। जबकि NH_a द्वारा निरूपित न्यूट्रोसोफिक वैकल्पिक परिकल्पना, दूसरी परिकल्पना है । सामान्य रूप से वितरित यादृच्छिक चर **x** को इस प्रकार मानकीकृत किया जा सकता है, $z = \frac{x - \mu}{\sigma}$, जहां $\mu = x$ का माध्य और $\sigma = x$ का मानक वितरण है । यदि चर **x** के बारे में न्यूट्रोसोफिक शून्य परिकल्पनाः $NH_O: \mu \in [a, b]$ है, जहां [a, b], $a \leq b$ के साथ, परिकल्पित अंतराल है, तो न्यूट्रोसोफिक परीक्षण

संख्यिकीः $z=\frac{\bar{x}-[a,b]}{s/\sqrt{n}}$ है, जहां x प्रतिदर्श माध्य है, s प्रतिदर्श मानक विचलन है और द प्रतिदर्श आकार है ।

वैकल्पिक परिकल्पनाएँ

NHa: $\mu > b$; अस्वीकार करें H_0 यदि z > z न्यूनतम महत्वपूर्ण मान (ऊपरी–पूंछ परीक्षण)य

NHa: $\mu > a$; अस्वीकार करें H_0 यदि z < –z अधिकतम महत्वपूर्ण मान (निचली–पूंछ परीक्षण);

NHa: $\mu \notin [a, b]$; य अस्वीकार करें H_0 यदि या तो z > z न्यूनतम महत्वपूर्ण मान, या z < – z अधिकतम महत्वपूर्ण मान (द्वि–पूंछ परीक्षण)।

मान लीजिए कि किसी ने महत्व α के विशेष स्तर पर न्यूट्रोसोफिक च–मान की गणना की है, जहां α एक स्पष्ट धनात्मक संख्या है ।

1. यदि अधिकतम {न्यूट्रोसोफिक P–मान} ≤ α, तो NH_0 को, α के स्तर पर अस्वीकार करें।
2. यदि न्यूनतम {न्यूट्रोसोफिक P–मान} > α, तो NH_0 को, α के स्तर पर अस्वीकार न करें।
3. यदि न्यूनतम {न्यूट्रोसोफिक P–मान} < α < यदि अधिकतम {न्यूट्रोसोफिक P–मान} तब एक अनिश्चितता होती है।

इस प्रकार, α – न्यूनतम { न्यूट्रोसोफिक P–मान }

अधिकतम {न्यूट्रोसोफिक P–मान}– न्यूनतम {न्यूट्रोसोफिक P–मान}

α स्तर पर H_O को अस्वीकार करने की संभावना है, और

अधिकतम {न्यूट्रोसोफिक P–मान} – α

अधिकतम {न्यूट्रोसोफिक P–मान} – न्यूनतम {न्यूट्रोसोफिक P–मान}

α स्तर पर H_O को अस्वीकार न करने की संभावना है। α^N को महत्व का न्यूट्रोसोफिक स्तर मान लीजिये।

4. यदि अधिकतम {न्यूट्रोसोफिक P–मान} ≤ न्यूनतम $\{\alpha^N\}$, तो NH_0 को, α^N के स्तर पर अस्वीकार करें
5. यदि न्यूनतम {न्यूट्रोसोफिक P–मान } झ अधिकतम $\{\alpha^N\}$, तो NH_0 को, α^N के स्तर पर अस्वीकार न करें।

उदाहरण

आइए विचार करें कि एक अमेरिकी कॉलेज के छात्रों के नमूने के लिए परीक्षा–चिंता स्कोर निम्नलिखित थेः

$n = 64$, $\bar{x} = [48.0, 50.0]$, और s = 25. तब μ = वास्तविक, परीक्षा–चिंता माध्य।

NH_0: $\mu \in [40.0, 41.0]$ और NH_a: $\mu > 41.0$ ।

न्यूट्रोसोफिक परीक्षण सांख्यिकी इस प्रकार हैः

$$z=\frac{[48.0,50.0]-[40.0,41.0]}{25/\sqrt{64}}=\frac{[48.0-41.0,50.0-40.0]}{25/8}$$

$$=\frac{[56.0,80.0]}{25}=\left[\frac{56.0}{25},\frac{80.0}{25}\right]=[2.24,3.20]$$

α = 0.10 के लिए पिछली तालिका से संबंधित एक–पूंछ वाला z महत्वपूर्ण मान 1-28 है। इसलिए NH_O को अस्वीकार कर दिया गया है क्योंकि z = [2-24, 3-20] > 1-28 । निष्कर्षतः, औसत परीक्षा–चिंता स्कोर 41-0 से अधिक है ।

महत्व के स्तर पर α1 = 0.10, अधिकतम से NH_O को अस्वीकार करें [0.0007, 0.0125] = 0.0125 < 0.10.

महत्व के स्तर पर α2 = 0.0005, अधिकतम से NH_O को अस्वीकार न करें [0.0007, 0.0125] = 0.0125 > 0.0005. महत्व के स्तर पर α_3 = 0.01, तब से अनिश्चितता है 0.01 ϵ [0.0007, 0.0125]; इसलिएः स्तर α3 पर NH0 को अस्वीकार करने की संभावना

$$= 0.01,\ \frac{0.01-0.0007}{0.0125-0.0007}=\frac{0.0093}{0.0118}\ 79\%$$ है ।

एवं स्तर α3 पर NH0 को अस्वीकार न करने की संभावना = 0.01, $\frac{0.0125-0.01}{0.0125-0.07}=\frac{0.0025}{0.0118}$ 21% है ।

यादृच्छिक पूर्ण ब्लॉक अभिकल्पना में न्यूट्रोसोफिक

न्यूट्रोसोफिक टू–वे एनोवा एक न्यूट्रोसोफिक रैखिक मॉडल परीक्षण है चाहे विभिन्न उपचारों के बीच महत्वपूर्ण

अंतर है अथवा विभिन्न क्षेत्रों के बीच महत्वपूर्ण अंतर है। इस सांख्यिकीय परीक्षण के विज्ञान के कई क्षेत्रों में कई अनुप्रयोग हैं, जैसे चिकित्सा क्षेत्र में जहां किसी व्यक्ति विशेष के लिए ग्लूकोज का स्तर तय नहीं होता है। यादृच्छिक पूर्ण ब्लॉक अभिकल्पना का मॉडल नीचे दिया गया है:

$$y_{ij}^{N} = \mu^{N} + \tau_i^{N} + \beta_j^{N} + \varepsilon_{ij} \qquad (1)$$

$y^N_{ij'}$ रवे ब्लॉक से i वे उपचार की न्यूट्रोसोफिक प्रतिक्रिया है, μ^N सामान्य न्यूट्रोसोफिक माध्य प्रभाव है, τ^N_i, i वा न्यूट्रोसोफिक उपचार प्रभाव है, $\beta^N_{j'}$ jवा न्यूट्रोसोफिक ब्लॉक प्रभाव है और ϵ^i_j यादृच्छिक त्रुटि है जो 0 माध्य और नियत विचरण σ^2_N के साथ न्यूट्रोसोफिक सामान्य वितरण का अनुसरण करती है। उपचारों के बीच कोई महत्वपूर्ण अंतर नहीं है ।

$NH_0 1$: उपचारों के बीच कोई सार्थक अंतर नहीं है।

$NH_0 2$: ब्लॉकों के बीच कोई सार्थक अंतर नहीं है।

इसका परीक्षण दो–तरफा एनोवा के माध्यम से किया जा सकता है क्योंकि हमारे पास विचरण, उपचार और सेक्टर के दो स्रोत हैं। मॉडल (1) को आव्यूह संकेतों का उपयोग करके पुनः इस प्रकार लिखा गया है

$$\mathbf{Y}^{N} = \boldsymbol{\mu}^{N}\mathbf{1} + \boldsymbol{\Delta}'\boldsymbol{\tau}^{N} + \mathbf{D}'\boldsymbol{\beta}^{N} + \boldsymbol{\varepsilon} \qquad (2)$$

एनोवा में सम्मिलित गणनाएं न्यूट्रोसोफिक अंकगणितीय परिचालनों का उपयोग करके भी की जाती हैं।

सेक्टर और उपचारों के लिए न्यूट्रोसोफिक F परीक्षण सांख्यिकी क्रमशः $F^N_1 = \frac{MSB^N}{MSE^N}$ एवं $F^N_2 = \frac{MST^N}{MSE^N}$ हैं।

दी गई परिकल्पनाओं के बारे में यह देखते हुए निर्णय लेना कि: $F^N \epsilon [F^L, F^U]$

यदि $F^L > F^*$ तब हम शून्य परिकल्पना को अस्वीकार करते हैं ।

यदि $F^U > F^*$ तब हम शून्य परिकल्पना को स्वीकार करते हैं।

यदि $F^L > F^*$ और $F^U > F^*$ तब हमारे पास कोई निर्णय नहीं है (हम अनिर्णीत हैं)।

परिणामों को न्यूट्रोसोफिक एनोवा (NANOVA) तालिका नामक तालिका में निम्नानुसार प्रस्तुत किया जा सकता है:

काल्पनिक आंकड़ों का उपयोग करके चित्रण

18 मधुमेह रोगियों के एक प्रतिदर्श का इलाज 6 प्रकार की दवाओं अर्थात A, B, C, D, E और F से

	A	B	C	D	E	F
आयु समूह 1	[110, 111]	[125, 128]	[135, 137]	[129, 134]	[128, 129]	[149, 150]
आयु समूह 2	[115, 131]	[135, 136]	[130, 132]	[112, 130]	[125, 129]	[139, 141]
आयु समूह 2	[112, 124]	[140, 144]	[141, 144]	[111, 128]	[124, 126]	[126, 136]
आयु समूह 4	[111, 112]	[121, 122]	[129, 133]	[115, 117]	[110, 119]	[120, 140]

तालिका 3: न्यूट्रोसोफिक एनोवा (NANOVA) तालिका

विचरण स्रोत	स्वतन्त्रता की कोटि (df)	वर्गों का योग (SS)	वर्ग माध्य (MS)	F	p- मान
उपचार	5	[1693.3, 1282.7]	[256.5, 338.7]	[6.31, 6.41]	[0.0022, 0.0024]
ब्लॉक	3	[441.3, 375.4]	[125.1, 147.1]	[2.74, 3.13]	[0.0570, 0.0798]
त्रुटि	15	[804.7, 599.8]	[40.1, 53.6]		
योग	23				

तालिका 4: एनोवा (NANOVA) तालिका

विचरण स्रोत	स्वतन्त्रता की कोटि	वर्गों का योग (SS)	वर्ग माध्य (MS)	F	p- मान
उपचार	5	1431.9	286.4	7.567	0.001
ब्लॉक	3	378.4	126.1	3.33	0.048
त्रुटि	15	567.7	37.8		
योग	23				

किया गया। मरीजों को विभिन्न आयु समूहों के अनुसार विभाजित किया गया था। प्रत्येक रोगी के लिए ग्लूकोज माप की एक परास ली गई है जो माप को अधिक विश्वसनीय बनाती है। आँकड़े इस प्रकार हैं:

यहाँ, पहला क्रांतिक मान F^*_1, 3.287 है और हम देख सकते हैं कि F^N_1 [2.74, 3.13] है। $F^U_1 = 3.13 < 3.287$ इसलिए, हम शून्य परिकल्पना को स्वीकार करते हैं ताकि विभिन्न आयु समूहों के बीच ग्लूकोज में कोई महत्वपूर्ण अंतर न हो। दूसरा क्रांतिक मान F^N_2 , 2.90 है और यह ध्यान किया जा सकता है कि F^N_2, [6.31, 6.41,। $F^L_2 = 6.31 > 2.90$। अतः इस स्थिति में शून्य परिकल्पना अस्वीकार की जाती है। यदि हम माध्य मान लेते हैं तो एक अलग एनोवा तालिका दिखाई देगी जो नीचे दी गई है:

उपरोक्त एनोवा तालिका से हम यह निष्कर्ष निकाल सकते हैं कि दोनों मामलों में 5: महत्व के स्तर पर p–मान 0.05 से कम है। इसलिए, उपचार के साथ–साथ ब्लॉक प्रभावों के बीच महत्वपूर्ण अंतर है जो न्यूट्रोसोफिक एनोवा तालिका से अलग है।

इसी प्रकार, इसे अन्य अभिकल्पनाओं जैसे पूर्ण यादृच्छिक अभिकल्पना, लैटिन स्क्वायर अभिकल्पना, स्प्लिट–प्लॉट अभिकल्पना आदि में भी बढ़ाया जा सकता है।

निष्कर्ष

न्यूट्रोसोफिक तर्कय अस्पष्ट, धुंधला तर्क का विस्तार है और इसमें विभिन्न पहलुओं अर्थात चिकित्सा अध्ययन, कृषि, शारीरिक अध्ययन आदि में विशाल अनुप्रयोग क्षमता है। इसका उपयोग शास्त्रीय सांख्यिकी की तरह ही किया जाता है, केवल अंतर यह है कि इसे अस्पष्ट, धुंधले आंकडों पर लागू किया जाता है। यहां, न्यूट्रोसोफिक का कुछ परिचय प्रदान किया गया है जिसमें बुनियादी संचालन, परिकल्पना परीक्षण, न्यूट्रोसोफिक सामान्य वितरण सम्मिलित है और यादृच्छिक पूर्ण ब्लॉक अभिकल्पना के माध्यम से परीक्षणों की अभिकल्पना में इसका अनुप्रयोग भी प्रदान किया गया है। इस तकनीक का लाभ निर्णयों को सुरक्षित और वास्तविक निर्णय के निकट बनाना है ।

संदर्भ

आदिश, ए.ए., एस्रे, एस.ए., ग्योरकोस, टी.डब्ल्यू., जीन–बैप्टिस, जे., रोझानी, ए. (1999)। इफैक्ट ऑफ कंजंप्शन ऑफ फूड कुक्ड इन आइरन पोट्स ऑन आइरन स्टेटस एंड ग्रोथ ऑफ यंग चिल्ड्रन। ए रैनडमाइज्ड ट्रायल। लैंसेट, 353, 712– 716।

स्मरांडाचे, एफ. (2007)। न्यूट्रोसोफिक सेट, न्यूट्रोसोफिक प्रोबेबिलिटी एवं स्टेटिस्टिक्स। इन्फो लर्न क्वेस्ट, सिक्स्थ एडीशन। https://digitalrepository.unm.edu/cgi/viewcontent.cgi?article=1109&conte=math_fsp।

स्मरांडाचे, एफ. (2014)। इंट्रोडक्शन टू न्यूट्रोसोफिक स्टेटिस्टिक्स। साइटेक एंड पब्लिशिंग। https://digitalrepository.unm.edu/cgi/viewcontent-cgi?article=1022&conte=math_fsp।

स्मरांडाचे, एफ. (2019)। न्यूट्रोसोफिक स्टेटिस्टिक्स

वर्सेस क्लासिकल स्टेटिस्टिक्स। निडस आइडियारम / सुपरल्यूमिनल फिजिक्स, 7,थर्ड एडीशन। http:// fs-unm-edu/NidusIdearum7&ed3-pdf।

मिआरी, एम., आनन, एम.टी., एवं जीना, एम.बी. (2022)। न्यूट्रोसोफिक टू वे एनोवा। इंटरनेशनल जर्नल ऑफ न्यूट्रोसोफिक साइंस, 18(3), 73–83।

अलआइता, ए., एवं असलम, एम. (2022)। एनालिसिस ऑफ कोवेरीयन्स अंडर न्यूट्रोसोफिक स्टेटिस्टिक्स। जर्नल ऑफ स्टेटिस्टिकल कम्प्यूटेशन एंड सिमुलेशन। 93(3), 397–415।

अलआइता, ए., तालेबी, एच., असलम, एम., एवं अल सुल्तान, के. (2023)। न्यूट्रोसोफिक स्टेटिस्टिकल एनालिसिस ऑफ स्प्लिट–प्लॉट डिजाइन। सॉफ्ट कम्प्यूटिंग। https://doi-org/10.1007/s00500-023-08025-y।

कुमारी, एस., अजरुद्दीन एस. एवं जेम्स, जे. (2022)। ए स्टडी ऑन न्यूट्रोसोफिक कंप्लीटली रेण्डमाइज्ड डिजाइन। मेथेमेटिकल स्टेटिस्टिशियन एंड इंजीनियरिंग एप्लीकेशंस, 71(4), 3738–3747।

अभिकल्पित परीक्षणों का गैर-प्राचलिक विश्लेषण

नीतू आर एस, सिनी वर्गीज, मो. हारून, आशुतोष दलाल, अनिंदिता दत्ता एवं देवेंद्र कुमार
भा.कृ.अनु.प. – भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली

परिचय

आधुनिक सांख्यिकीय पद्धतियों के विकास में, जनसंख्या के बारे में एक अच्छा अनुमान प्राप्त करने की पहली तकनीक यह थी कि आंकड़ों को एक विशेष वितरण, सामान्यतः सामान्य वितरण के बाद जनसंख्या से लिया गया माना जाता था। लेकिन ऐसी कई स्थितियाँ हैं जिनमें उस जनसंख्या के बारे में कोई कठोर धारणा नहीं बनाई गई है जिससे आंकड़ों का प्रतिदर्श लिया गया है। ऐसी वितरण मुक्त तकनीकें या गैर–प्राचलिक तकनीकें एक मॉडल पर आधारित होती हैं जो केवल बहुत सामान्य स्थितियों को निर्दिष्ट करती हैं और वितरण के विशिष्ट रूप के संबंध में कुछ भी निर्दिष्ट नहीं करती हैं। कुछ धारणाएँ अधिकांश गैर–प्राचलिक सांख्यिकीय परीक्षण ों, जैसे कि यादृच्छिकता एवं अवलोकन की स्वतंत्रता से संबद्ध होती हैं लेकिन ये धारणाएँ प्राचलिक परीक्षण ों से संबद्ध मान्यताओं की तुलना में कम और कमजोर हैं। परीक्षणात्मक आंकड़ों के विश्लेषण के लिए गैर–प्राचलिक परीक्षण विकसित करने पर बहुत ध्यान दिया जा रहा है। इनमें से अधिकांश गैर–प्राचलिक परीक्षण प्रक्रियाएँ रैंक आँकड़ों पर आधारित हैं। यहां एकल–पथ पूर्ण यादृच्छिक अभिकल्पना (one-way CRD), द्वि–पथ यादृच्छिक पूर्ण खण्ड अभिकल्पना (two-way RCBD), संतुलित अपूर्ण खण्ड अभिकल्पना (BIBD), असंतुलित अपूर्ण खण्ड अभिकल्पना, प्रति प्लॉट एकाधिक अवलोकन और क्रॉस–ओवर परीक्षणों के विश्लेषण के लिए गैर–प्राचलिक परीक्षण प्रक्रिया सम्मिलित हैं।

परीक्षणात्मक आंकड़ों के विश्लेषण हेतु गैर-प्राचलिक प्रक्रिया

परीक्षणात्मक स्थितियों की प्रकृति के आधार पर विभिन्न गैर–प्राचलिक परीक्षणों का उपयोग किया जाता है ।

पूर्ण यादृच्छिक अभिकल्पना के माधयम से उत्पन्न आँकड़ों के लिए परीक्षण प्रक्रिया

पूर्ण यादृच्छिक अभिकल्पना सेट–अप के लिए समान्यतः उपयोग किया जाने वाला गैर–प्राचलिक परीक्षण क्रुस्कल–वालिस एच परीक्षण है, जो कई स्वतंत्र प्रतिदर्शों की समानता का परीक्षण करता है (क्रुस्कल, 1952)। K-W सांख्यिकी, 2 स्वतंत्र प्रतिदर्शों के लिए मैन–व्हिटनी–विलकॉक्सन परीक्षण का एक गैर–प्राचलिक विस्तार है। N अवलोकनों को 1 से सबसे छोटे अवलोकन और N से सबसे बड़े अवलोकन तक रैंक देकर रैंक करें। मान लीजिए R_{ij}, y_{ij} को दी गई रैंक को दर्शाता है जहां $i=1, 2, ..., v$ और $j = 1, 2, ..., ri$ है।

i^{th} उपचार से संबंधित अवलोकनों की रैंक का योग R_i है, जहां $R_i = \sum_{j=1}^{ri} R_{ij}$

i^{th} उपचार से संबंधित अवलोकनों की रैंक का माध्य R_i, $R_i = \sum_{j=1}^{ri} R_{ij}$ होगा

सभी R_i का माध्य R होगा।

$$H = \frac{12}{N(N+1)}\sum_{i=1}^{V}\frac{R_i^2}{r_i} - 3(N+1) \sim \chi^2_{v-1} \qquad (1)$$

यदि संबंध होते हैं तो मध्य रैंक दी जाती है और एच सांख्यिकी

$$H = \frac{\frac{12}{N(N+1)}\sum_{i=1}^{V}\frac{R_i^2}{r_i} - 3(N+1)}{1-\frac{\sum_{s=1}^{g}(t_s^3-t_s)}{N^3-N}} \qquad (2)$$

बन जाती हैं, जहां विभिन्न बंधे हुए रैंकों के बंधे

हुए उपचारों की संख्या को g द्वारा दर्शाया गया है, s^{th} उपचार में बंधे हुए रैंकों की संख्या को जत्रे द्वारा दर्शाया गया है।

यदि के–डब्ल्यू परीक्षण (K-W test) के परिणामस्वरूप उपचार प्रभावों की समानता की शून्य परिकल्पना को अस्वीकार कर दिया जाता है तो उपचार युग्मों का पता लगाने के लिए कई तुलना परीक्षण किए जाते हैं जो दूसरों के बीच महत्वपूर्ण अंतर दिखाते हैं ।

C_{ij} द्वारा दर्शाए गए उपचार युग्मों के बीच न्यूनतम महत्व का अंतर

$$C_{ij}= Z_p\sqrt{\frac{N(N+1)}{12}\left(\frac{1}{r_i}+\frac{1}{r_j}\right)} \qquad (3)$$

यदि $|\overline{R_i} - \overline{R_j}| \geq C_{ij}$ तो i^{th} और j^{th} उपचार प्रभावों के बीच अंतर α% महत्व के स्तर पर महत्वपूर्ण है और Z_p, $p = \frac{\alpha}{v(v-1)}$ के साथ मानक सामान्य वितरण के तहत क्रम 1-p की मात्रा को इंगित करता है।

यादृच्छिक पूर्ण खण्ड अभिकल्पना के माधयम से उत्पन्न आँकड़ों के लिए परीक्षण प्रक्रिया

यादृच्छिक पूर्ण खण्ड अभिकल्पना में भी सामान्य न होने की समस्या देखी जा सकती है. फ्रीडमैन परीक्षण सामान्यतः द्वि–पथ एनोवा (फ्रीडमैन, 1937) के मामले में गैर–प्राचलिक परीक्षण का उपयोग किया जाता है।

$$T = \frac{12}{bv(v+1)}\sum_{i=1}^{v} R_i^2 - 3b(v+1) \sim \chi^2_{v-1} \qquad (4)$$

जहां, $R_i = \sum_{j=1}^{b} R_{ij}$

यदि एक ही खण्ड में बंधी हुई रैंक हैं तो T इस प्रकार हो जाता है,

$$T = \frac{12\sum_{i=1}^{v} R_i^2 - b^2 v(v+1)^2}{bv(v+1) + \frac{bv - \sum_{j=1}^{b}\sum_{s=1}^{g}(t_{js}^3)}{v-1}} \sim \chi^2_{v-1} \qquad (5)$$

जहां, g_j = jवें खण्ड में बंधे रैंकों के सेट की संख्या, t_{js} = jवें खण्ड में में बंधी रैंकों के s वें सेट का आकार। परीक्षण प्रक्रिया द्वारा युग्मवार तुलना करने के लिए महत्वपूर्ण मूल्य को इस प्रकार परिभाषित किया गया है,

$$c = Z_p\sqrt{\frac{bv(v+1)}{6}} \qquad (6)$$

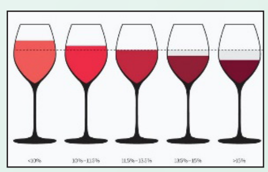
यदि $|\overline{R_i} - \overline{R_k}| \geq Z_p\sqrt{\frac{bv(v+1)}{6}}$ तो i वे और k वे उपचार प्रभावों के बीच अंतर α% महत्व के स्तर पर काफी भिन्न होता है जिसमें Z_p, $p = \frac{\alpha}{v(v-1)}$ के साथ मानक सामान्य वितरण के तहत 1-p क्रम की मात्रा को इंगित करता है। बीआईबी आधारित रैंक आंकड़ों का व्यापक रूप से कुछ कृषि संबंधी किस्म परीक्षणों, संवेदी मूल्यांकन परीक्षणों, नैदानिक परीक्षणों, ऑटोमोबाइल परीक्षण के परीक्षणों आदि में उपयोग किया जाता है।

संतुलित अपूर्ण खण्ड अभिकल्पना के माधयम से उत्पन्न आँकड़ों के लिए परीक्षण प्रक्रिया

संतुलित अपूर्ण खण्ड आधारित रैंक आंकड़ों का व्यापक रूप से कुछ कृषि संबंधी किस्म परीक्षणों, संवेदी मूल्यांकन परीक्षणों, नैदानिक परीक्षणों, ऑटोमोबाइल परीक्षण के परीक्षणों आदि में उपयोग किया जाता है।

आइसक्रीम चखने का परीक्षण (कोनोवर, 1999)

डार्क चॉकलेट की किस्मों की रूपरेखा (ह्यूसन एवं पेजेस, 2003)



रस की गुणवत्ता की संभावना कटाई के बाद के दौरान पकना (मार्टिनेज, 2014)

संतुलित अपूर्ण खण्ड आधारित रैंक आंकड़ों के पारंपरिक गैर–प्राचलिक विश्लेषण में डर्बिन परीक्षण (डर्बिन, 1951) सम्मिलित है। यह काई–स्क्वायर पर आधारित एक परीक्षण है और इसे समान्यतः उपचार वितरण की समानता के परीक्षण के रूप में चित्रित किया जाता है।

डर्बिन सांख्यिकी इस प्रकार दी गई है:

$$D = \frac{12(v-1)}{rv(k-1)(k+1)}\sum_{i=1}^{V} R_i^2 - \frac{3r(v-1)(k+1)}{(k-1)} \sim \chi^2_{v-1}$$

R_{ij}, एक ब्लॉक के भीतर 1 से k तक अवलोकनों की रैंक y_{ij} को निरूपित करता है। यदि D कट–ऑफ अंक से अधिक है तो परीक्षण शून्य परिकल्पना को अस्वीकार कर देता है। कट–ऑफ अंक v-1 स्वतंत्रता की कोटि के साथ काई–स्क्वायर वितरण तालिका का संदर्भ लेकर प्राप्त किया जाता है। यदि α महत्व के स्तर पर $|\overline{R_i} - \overline{R_j}| \geq c$ हो तो i, j उपचार को काफी भिन्न घोषित किया जाता है। यहां c, i और j उपचार के बीच सबसे कम महत्वपूर्ण अंतर मान है ।

$$c = Z_p\sqrt{\frac{vr(k+1)(k-1)}{6(v-1)}}$$

फॉसेट एवं साल्टर (1987) ने पाया कि D सांख्यिकी का प्रदर्शन सभी संतुलित अपूर्ण खण्ड अभिकल्पना, विशेषकर छोटी संतुलित अपूर्ण खण्ड अभिकल्पना के लिए संतोषजनक नहीं होगा। शून्य परिकल्पना की अस्वीकृति की दर कम होगी भले ही उपचारों के बीच अंतर के बड़े साक्ष्य उपस्थित हों। उन्होंने पाया कि D सांख्यिकी का काई–स्क्वायर सन्निकटन बहुत धीमी गति से इसके एसिम्प्टोटिक काई–स्क्वायर वितरण में परिवर्तित होने के कारण खराब था। इसलिए अभिसरण की दर आम तौर पर λ के आकार से प्रभावित होती है, तदुपरान्त संतुलित अपूर्ण खण्ड अभिकल्पना में r और b से प्रभावित होती है। D सांख्यिकी के लिए p–मान प्राप्त करने के लिए क्रमपरिवर्तन (यादृच्छिकीकरण) परीक्षण का उपयोग किया जाना चाहिए। यद्यपि, क्रमपरिवर्तन परीक्षण के माध्यम से D के लिए p– मान की गणना करने वाला सॉफ्टवेयर सर्वदा आसानी से उपलब्ध नहीं हो सकता है

कॉनओवर (1999), रैंकों पर विचरण का विश्लेषण (ANOVA) करने और उपचार अंतर के लिए F परीक्षण का उपयोग करने का सुझाव देता है। बेस्ट एवं रेनर (2014) ने डर्बिन परीक्षण आधारित काई–स्क्वायर परीक्षण पर F परीक्षण की उपयुक्तता की जांच करने के लिए विभिन्न संतुलित अपूर्ण खण्ड अभिकल्पनाओं पर 100,000 प्रतिदर्शों के आधार पर एक सिमुलेशन अध्ययन किया और इस निष्कर्ष पर पहुंचे कि आकार और शक्ति मूल्य, F परीक्षण पर आधारित गैर–प्राचलिक ANOVA के लिए बेहतर था। अतः, F परीक्षण के उपयोग की अनुशंसा इसके परीक्षण आकारों के डर्बिन परीक्षण आकारों की तुलना में नाममात्र के निकट होने के आधार पर की जा सकती है।

कॉनओवर के F मान की गणना इस सूत्र का उपयोग करके की जा सकती है,

$$F_A = \frac{SS_{adj\ trt}/v\text{-}1}{SS_E/edf}$$

$$= \left\{\frac{edf}{v-1}\right\}\left[\frac{\frac{k}{\lambda t}\sum_{i=1}^{t}(R_i - \frac{r(k+1)}{2})^2}{\sum_{i,j} r_{ij}^2 - \frac{bk(k+1)^2}{4} - \frac{k}{\lambda t}\sum_{i=1}^{t}(R_i - \frac{r(k+1)}{2})^2}\right] \quad (7)$$

जहाँ,

$SS_{total} = \sum_{i,j} r_{ij}^2 - \frac{r_{..}^2}{bk}$, जहाँ, $r_{..} = \frac{bk(k+1)}{2}$

$SS_{trt} = \sum_{i,j} r_{ij}^2 - \frac{bk(k+1)^2}{4}$

$SS_{adjtrt} = \frac{k}{\lambda t}\sum_{i=1}^{t}$ (R_i उन खण्डों का खण्ड माध्य जिनमें i^{th} उपचार सम्मिलित है$)^2$

$$= \frac{k}{\lambda t}\sum_{i=1}^{t}(R_i - \frac{r(k+1)}{2})^2$$

$SS_E = SS_{total} - SS_{adjtrt}$

लिविंगस्टन और रेनर ने 2022 में डर्बिन परीक्षण में एक बंध समायोजन किया जिसे कॉनओवर के F परीक्षण के समकक्ष परीक्षण माना जा सकता है।

यदि डेटा सेट में संबंध होते हैं और मध्य रैंक आवंटित की जाती है ताकि रैंक योग प्रभावित न हो तो एक समायोजित डर्बिन परीक्षण सांख्यिकी इस प्रकार विकसित की जाती है,

$$D_A = \frac{(v-1)\left\{\sum_{i=1}^{v} R_i^2 - \frac{rbk(k+1)^2}{4}\right\}}{\sum_{i,j} r_{ij}^2 - \frac{bk(k+1)^2}{4}} \quad (8)$$

जहाँ, $r_{ij} = j$ खण्ड पर i उपचार की रैंक, $R_i = i$ उपचार की रैंक का योग

$i = 1, 2, ..., v$ एवं $j = 1, 2, ..., k$.

यदि कोई संबंध नहीं है तो j खण्ड पर रैंक $1, 2,$

..., k हैं। तब,

$$\sum_{i,j} r_{ij}^2 = \frac{bk(k+1)(2k+1)}{6} \quad (9)$$

और DA में हर $\frac{b(k-1)k(k+1)}{12}$ हो जाता है जो सरलीकरण पर D सांख्यिकी में कम हो जाता है। D_A, D का बंध समायोजित मूल्य है।

$$D_A = D/c \text{ जहाँ}, \; c = 1 - \frac{\sum_{g,j}(v_{g,j}^3 - v_{g,j})}{b(k-1)k(k+1)}$$

प्राचलिक और गैर–प्राचलिक ANOVA के बीच शक्ति अंतर 0-2 से अधिक हो सकता है, इसलिए त्रुटि वितरण पर सावधानीपूर्वक विचार करना महत्वपूर्ण है। भले ही समायोजित डर्बिन और कॉनओवर का F परीक्षण, समतुल्य परीक्षण हैं, समायोजित डर्बिन परीक्षण का आकार आम तौर पर 0.05 से थोड़ा कम होता है जबकि ANOVA F परीक्षण का आकार 0.05 से थोड़ा अधिक होता है। इस आकार के लाभ के कारण, सामान्यतः समायोजित डर्बिन की तुलना में F परीक्षण की सिफारिश की जाती है। स्थिति के आधार पर कोई यह निर्णय ले सकता है कि किसका उपयोग करना है। F परीक्षण की तुलना में समायोजित डर्बिन के लिए संगणकीय सुगमता अधिक पाई जाती है। इसलिए यदि डेटा सेट बड़ी परिवर्तनशीलता दिखाता है तो दोनों परीक्षण शून्य परिकल्पना को अस्वीकार कर देंगे। उस स्थिति में F परीक्षण के स्थान पर समायोजित डर्बिन का उपयोग किया जा सकता है। कॉनओवर के F परीक्षण का उपयोग करने के लिए उच्च सटीक अनुमान की आवश्यकता होती है।

उदाहरण 1: कुटनर इत्यादि (2005) एक स्वाद–परीक्षण पर विचार करें जिसमें पांच नाश्ता अनाज (t = 5) को दस निर्णायकों (b = 10) द्वारा दस बिंदु हेडोनिक पैमाने पर एक समय में तीन (k = 3) द्वारा स्कोर किया गया था । प्रत्येक अनाज को छह बार चखा गया (r = 6) । रैंक किए गए आंकड़ों को तालिका 1 में दिखाया गया है (ध्यान दें कि रैंक बंधी हुई हैं)।

तालिका 1: अनाज चखने के परीक्षण के सापेक्ष रैंक किया गए आंकड़े

निर्णायक	अनाज				
1	A	B	C	D	E
2	1.5	1.5	3	-	-
3	1	2.5	-	2.5	-
4	1.5	2	-	-	1.5
5	1	-	2	3	-
6	1.5	-	3	-	1.5
7	1	-	-	3	2
8	-	2	3	1	-
9	-	2.5	2.5	-	1
10	-	3	-	2	1
योग	-	-	3	2	1
	7.5	15	17	14	8

विश्लेषण परिणाम

प्राचल (v=5, b=10, k=3, r=6).

$\sum_{i,j} r_{ij}^2 = 137.5$

$\sum_{i=1}^{v} R_i^2 = 785$

$\sum_{g,j}(v_{g,j}^3 - v_{g,j}) = 30$

तालिका 2: प्राप्त डर्बिन मान एवं F मान

सांख्यिकी	मान	P मान
D	7.4	0.116
D_A	11.84	0.018
F	5.8	0.004

तीनों परीक्षण आंकड़ों से संकेत मिलता है कि अनाज की रैंकिंग में महत्वपूर्ण अंतर था (बेस्ट एवं रेनर, 2014)। DA मान D मान से अधिक पाया गया जो काई–स्क्वायर सन्निकटन के प्रति उच्च अभिसरण को दर्शाता है जिससे F आंकड़ों के समान शून्य परिकल्पना को खारिज करने की अधिक संभावना मिलेगी।

क्रासओवर अभिकल्पना के माधयम से उत्पन्न आँकड़ों के लिए परीक्षण प्रक्रिया

यदि आँकड़े गैर सामान्य दिखाई देते हैं तो छोटे और बड़े क्रॉसओवर अध्ययनों के मामले में गैर प्राचलिक दृष्टिकोण उपयुक्त हैं। विलकॉक्सन रैंक सम परीक्षण सामान्यतः 2×2 क्रॉसओवर परीक्षण के लिए NP परीक्षण का उपयोग किया जाता है।

एक नैदानिक परीक्षण पर विचार करें

मान लीजिए कि lवीं अवधि में kवें उपचार पर iवें क्रम से jवें विषय के लिए परिणाम y_{ijkl} है तब

$$y_{ijkl} = \mu + \tau_k + \eta_i + \pi_l + \lambda_{k'l-1} + \varepsilon_{j(i)kl} \qquad (10)$$

जहां μ सामान्य माध्य है, η_i, l^{th} अनुक्रम का प्रभाव है, π_l, lवीं अवधि का माध्य अतिरिक्त प्रभाव है, $\lambda_{k'l-1}$ l^{th} अवधि में (l–1) वां एवं $\varepsilon_{j(i)klm}$ यादृच्छिक त्रुटि घटक प्रशासित k^{th} उपचार का माध्य अतिरिक्त कैरीओवर प्रभाव है।

विलकॉक्सन रैंक सम सांख्यिकी W इस प्रकार दी गई है, (कोच, 1972)

$$W = \frac{\sqrt{12}\left(R_1 - \frac{n_1(n_1+n_2+1)}{2}\right)}{\sqrt{n_1 * n_2(n_1+n_2+1)}} \qquad (11)$$

जहां, $R_1 = \sum_{j=1}^{n_1} D_{1j}(y'_{ij})$ एवं $y'_{ij} = y_{ijkl} - y_{ijk'2}$

उदाहरण 2: क्रेग इत्यादि, 1998. ने एक नैदानिक परीक्षण किया जिसमें एलर्जिक राइनाइटिस से पीड़ित रोगियों में दिन के समय नींद की माप के आधार पर नाक के स्टेरॉयड (उपचार A) और प्लेसिबो (उपचार B) की तुलना की गई। पांच रोगियों को या तो उपचार A या B के लिए यादृच्छिक किया गया और 8 सप्ताह के लिए दिन में दो बार स्व–प्रशासित किया गया। 8वें सप्ताह में प्रत्येक रोगी को बिना वाशआउट के दूसरे उपचार में भेजा गया।

विश्लेषण परिणाम

चार रोगियों में अवधियों के बीच आईडीएस में कोई अंतर नहीं था, जबकि शेष में प्लेसबो की तुलना में स्टेरॉयड पर आईडीएस बेहतर था। चार रोगियों की रैंक से पता चलता है कि अवधियों के बीच कोई अंतर नहीं है, उन्हें मध्य रैंक पर 5.5 के रूप में सेट किया गया है, जिससे परीक्षण आंकड़ों के 16 अद्वितीय मान प्राप्त होते हैं। विलकॉक्सन W मान 0.0407 का दो तरफा स्पर्शोन्मुख p मान उत्पन्न करता है जो प्रभावों के बीच महत्वपूर्ण अंतर का संकेत देता है। युग्मित तुलना के संबंध में एक विचार प्राप्त करने के लिए HL अनुमान का उपयोग किया जाता है। HL अनुमान ने संकेत दिया कि स्टेरॉयड ने 95% सटीक आत्मविश्वास अंतराल (0, 1,794) के साथ दिन के समय तंद्रा में 0.642 इकाइयों तक सुधार किया (पुट एवं चिंचिली, 2004)।

तालिका 3: AB और BA अनुक्रम पर व्यक्तिगत रोगियों के लिए कमजोर आईडीएस माध्य स्कोर।

	AB अनुक्रम		BA अनुक्रम	
रोगी	अवधि 1	अवधि 2	अवधि 1	अवधि 2
1	2	1.29	4	4
2	1.83	0	0	0.86
3	0	0	0	0
4	0.89	-	1.14	2.14
5	3	3	0	2.29

तालिका 4: दो अनुक्रमों के सापेक्ष अवधियों के आधार पर कमजोर आईडीएस स्कोर एवं अवधियों के बीच का अंतर

अनुक्रम		अवधि 1	अवधि 2	अंतर
AB	1	2	1.29	0.71
	2	1.83	0	1.83
	3	0	0	0
	4	0.89	NA	NA
	5	3	3	0
BA	6	4	4	0
	7	0	0.86	-0.86
	8	0	0	0
	9	1.14	2.14	-1
	10	0	2.29	-2.29

असंतुलित अपूर्ण खण्ड अभिकल्पना के माधयम से उत्पन्न आँकड़ों के लिए परीक्षण प्रक्रिया

असंतुलित अभिकल्पना के मामले में स्किलिंग और मैक परीक्षण (S-M test) का उपयोग किया जाता है। यह स्किलिंग्स और मैक (1981) द्वारा प्रस्तावित एक फ्रीडमैन–प्रकार का परीक्षण सांख्यिकी है। परीक्षण आँकड़े प्राप्त करने के लिए रैंक किए गए आंकड़ों के अनुरूप समायोजित उपचार योग का उपयोग किया गया था। अवलोकनों को 1 से k_j तक रैंक करें, जहाँ kj, jवा खण्ड आकार है। यदि अवलोकन उपस्थित है तो मान लें कि R_{ij}, y_{ij} की रैंक है, अन्यथा, मान लें कि R_{ij}, $(k_j+1)/2$ है। मान लीजिए कि A_i, iवे उपचार के लिए समायोजित उपचार योग को दर्शाता है।

$$A_i = \sum_{j=1}^{b}\sqrt{\frac{12}{k_j+1}}\left[\frac{R_{ij}-(k_j+1)}{2}\right]$$

मान लीजिए आव्यूह $\sum$, जो यादृच्छिक सदिश $A' = (A_1, \ldots, A_v)$ का सहप्रसरण आव्यूह है, जहां ऑफ–विकर्ण तत्व $(-\lambda_{ii'})$ हैं, और उस पंक्ति में विकर्ण तत्व, ऑफ –विकर्ण तत्वों के योग के ऋणात्मक हैं। $H_0: \tau_1 = \tau_2 = \tau_3 = \ldots = \tau_v = \tau$, के अंतर्गत j खण्ड में परीक्षण आँकड़ों को इस रूप में आकार दिया जा सकता है

$$T = A'\sum^{-}A$$

चूंकि सहप्रसरण आव्यूह $\sum$ एकवचन है, क्योंकि पंक्तियों (स्तंभों) का योग सर्वदा शून्य होता है। किसी भी जुड़ी खण्ड अभिकल्पना में, $\sum$ की रैंक v-1 होगी। तो $\sum^{-}$, $\sum$ का सामान्यीकृत व्युत्क्रम है। T, $\sum$ की रैंक के रूप में स्वतंत्रता की कोटि के साथ लगभग काई स्कवायर–वितरण का अनुसरण करता है । इस सांख्यिकी को एच्छिक ढंग से गायब आँकड़े संरचना वाली सभी खण्ड अभिकल्पनाओं के लिए एक सामान्य आँकड़ा माना जा सकता है, अर्थात, अधूरा खण्ड अभिकल्पना जैसी अभिकल्पना से गायब होना, या यादृच्छिक रूप से पूरी तरह से गायब होना। सामान्यतः उपयोग किए जाने वाली फ्रीडमैन परीक्षण सांख्यिकी और डर्बिन परीक्षण सांख्यिकी इस S-M परीक्षण के विशेष मामले हैं ।

उदाहरण 3: एक इंजीनियर 4 प्रकार के गैसोलीन एडिटिव्स की माइलेज प्रदर्शन विशेषताओं का अध्ययन कर रहा है। सड़क परीक्षण में वह कारों को खण्ड (9) के रूप में उपयोग करना चाहता है। (डिजाइन संसाधन सर्वर)।

तालिका 5: 9 कारों में गैसोलीन एडिटिव्स का माइलेज प्रदर्शन

गैसोलीन एडिटिव्स	कार (खण्ड)								
	1	2	3	4	5	6	7	8	9
A	3.2	3.1	4.3	3.5	3.6	4.5	-	4.3	3.5
B	4.1	3.9	3.5	3.6	4.2	4.7	4.2	4.6	-
C	3.8	3.4	4.6	3.9	3.7	3.7	3.4	4.4	3.7
D	4.2	4	4.8	4	3.9	-	-	4.9	3.9

तालिका 6: 9 कारों में गैसोलीन एडिटिव्स के माइलेज प्रदर्शन को दी गई रैंक

गैसोलीन एडिटिव्स	कार (खण्ड)								
	1	2	3	4	5	6	7	8	9
A	1	1	2	1	1	2	1.5*	1	1
B	3	3	1	2	4	3	2	3	2*
C	2	2	3	3	2	1	1	2	2
D	4	4	4	4	3	2*	1.5*	4	3

तालिका 7: परिकलित समायोजित उपचार राशियाँ (Ai)

	खण्ड									समायोजित उपचार राशियाँ (A_i)
	1	2	3	4	5	6	7	8	9	
A	-2.324	-2.324	-0.775	-2.324	-2.324	0.000	0.000	-2.324	-1.732	-14.126
B	0.775	0.775	-2.324	-0.775	2.324	1.732	1.000	0.775	0.000	4.281
C	-0.775	-0.775	0.775	0.775	-0.775	-1.732	-1.000	-0.775	0.000	-4.281
D	2.324	2.324	2.324	2.324	0.775	0.000	0.000	2.324	1.732	14.126

A_i = (-14.1256, 4.2812, -4.2812, 14.1256)'

विश्लेषण परिणाम

प्रत्येक कार के भीतर आँकड़ों को रैंक करें। यदि अवलोकन नहीं है तो हम Rij = (kj+1)/2 का उपयोग करते हैं।

सहप्रसरण आव्यूह ∑ उपचार युग्म के एक साथ होने की संख्या की गणना करके प्राप्त किया जाता है। स्वतंत्रता की तीसरी कोटि पर x^2 का सारणीबद्ध मान 7-815 है और परिकलित मान 15-49 है। इसलिए 5% महत्व के स्तर पर उपचार काफी भिन्न हैं।

प्रति प्लॉट एकाधिक अवलोकन स्थिति के लिए परीक्षण प्रक्रिया

गोर (1975) ने परीक्षण सांख्यिकी विकसित की जिनका उपयोग उन स्थितियों में किया जा सकता है जिनमें खण्ड अभिकल्पना में प्रत्येक सेल से एक से अधिक अवलोकन उपलब्ध होते हैं।

गोर परीक्षण सांख्यिकी इस प्रकार है :

$$T = \frac{12N}{v^2}\left[\sum_{i=1}^{v}\frac{(u_i-\frac{(v-1)b}{2})^2}{q_{i.}} - \frac{\left\{\sum_{i=1}^{v}\frac{u_i-\frac{(v-1)b}{2}}{q_{i.}}\right\}^2}{q_{..}}\right] \qquad (12)$$

जहाँ, $N = \sum_i^v \sum_j^b n_{ij}$, $u_i = \sum_{i'\neq i=1}^{v}\sum_{j=1}^{b} u_{i'ij}$,

$$p_{ij} = \frac{n_{ij}}{N}, q_{ij} = \frac{1}{p_{ij}}, q_{i.} = \sum_j^b q_{ij}, q_{..} = \sum_i^v \frac{1}{q_{ij}}$$

यदि T स्वतंत्रता की (v-1) कोटि के साथ काई–स्क्वायर वितरण के ऊपरी कट ऑफ बिंदु से अधिक है तो परीक्षण H_0 को अस्वीकार कर देता है।

उदाहरण 4: मिट्टी की दो स्थितियों में उगाई गई अनाज की फसल की तीन किस्मों के पकने के दिनों की संख्या को हल्की और भारी मिट्टी के प्रकार के तहत विभिन्न प्रकार के प्रदर्शन तक पहुंचने के लिए माना जाता था। (डिजाइन संसाधन सर्वर)। सेल (1, 1) में प्रत्येक अवलोकन लें और सेल (2, 1) में 4 अवलोकनों के साथ इसकी तुलना करें, इसलिए, यू 1,2,1, यहां कुल मिलाकर 16 (4×4) तुलनाएं हैं। अवलोकन 130 सेल (2, 1) के सभी 4 मानों से बड़ा है। इसलिए, यह 4 का योगदान देता है। इसी प्रकार 115 का योगदान 3 है, 123 का योगदान 3 है और 142 का योगदान 4 है।

तालिका 8: अनाज की फसल की तीन किस्मों के पकने के दिनों की संख्या

किस्म	मिट्टी के प्रकार	
	हल्की	भारी
A	130,115,123,142	117,125,139
B	108,114,124,106	91,111,110
C	155,146,151,165	97,108

विश्लेषण परिणाम

इस डाटा सेट में n_{11}=4, n_{12}=3, n_{21}=4, n_{22}=3, n_{31}=4, n_{32}=2 and v= 3, b= 2, N=20

तालिका 9: आँकड़ों के अनुरूप p और q मान

मिट्टी के प्रकार	किस्म	p	q
हल्की	A	0.2	5
	B	0.2	5
	C	0.2	5
भारी	A	0.15	6.67
	B	0.15	6.67
	C	0.1	10

q_1= 5+6.67 = 11.67; q2 = 5+6.67=11.67; q3. = 5+10 = 15; q.. = 0.2380

सेल (1, 1) में प्रत्येक अवलोकन लें और सेल (2, 1) में 4 अवलोकनों के साथ इसकी तुलना करें, इसलिए, $u_{1,2,1}$, यहां कुल मिलाकर 16 (4×4) तुलनाएं हैं। अवलोकन 130, सेल (2, 1) के सभी 4 मानों से बड़ा है। इसलिए, यह 4 का योगदान देता है। इसी प्रकार 115 का योगदान 3 है, 123 का योगदान 3 है और 142 का योगदान 4 है। कुल योग14 है। तो, $u_{1,2,1}$, 14/16 है।

इसी प्रकार, $u_{1,2,2}=9/9; u_{1,3,1}=0; u_{1,3,2}=1; u_{2,3,1}=0; u_{2,3,2}=4/6; u_{2,1,1}=2/16; u_{2,1,2}=0; u_{3,1,1}=1; u_{3,1,2}=0; u_{3,2,1}=1; u_{3,2,2}=2/6$; u_1= 46/16, u_2 = 76/96, u_3 = 14/6

अतः, T = 5-28- स्वतंत्रता की दूसरी कोटि पर x^2 का सारणीबद्ध मूल्य 5-991 है और गणना मूल्य 5.28 है, इसलिए अनाज की फसल की किस्में 5% महत्व के स्तर पर महत्वपूर्ण रूप से भिन्न नहीं हैं। यहां x^2 से अधिक 0-0713 संभावना है।

निष्कर्ष

यहां वर्णित विधियां रैंकों के संबंध में जानकारी का उपयोग करती हैं और अवलोकनों के मात्रात्मक मूल्यों का कोई उपयोग नहीं करती हैं। इसलिए इन गैर–प्राचलिक पद्धति का व्यापक अनुप्रयोग था क्योंकि डेटा सेट के संबंध में निष्कर्ष पर पहुंचने के लिए कुछ अंतर्निहित धारणाओं को संतुष्ट किया जाना चाहिए। इन तकनीकों का सुझाव उन स्थितियों पर काबू पाने के लिए दिया जा सकता है जब त्रुटि प्रसरण सामान्यता और समरूपता का उल्लंघन करती है ।

संदर्भ

बेस्ट, डी.जे., एवं रेनेर, जे.सी.डब्ल्यू. (2014)। कॉनओवर्स एफ टेस्ट एज एन आल्टरनेटिव टु डर्बिंस टेस्ट। जर्नल ऑफ मॉडर्न अप्लाइड स्टेटिस्टिकल मेथड्स, 13, 76–83।

कोनोवर, डब्ल्यू.जे. (1999)। प्रक्टिकल नॉन पैरामीट्रिक स्टेटिस्टिक्स। थर्ड एडी, जॉन विली एंड संस, न्यूयॉर्क। कुटनर, एम., नचत्शेम, सी., नेटर, जे. और ली, डब्ल्यू. (2005)। अनुप्रयुक्त रैखिक सांख्यिकीय मॉडल 5वीं शिक्षा, मैकग्रा–हिल इरविन, बोस्टन।

क्रेग, टी.जे., टीट्स, एस., लेहमैन, ई.बी., चिंचिली, वी. एम. एवं ज़्विलिच, सी. (1998)। नसल कंजेसन सेकण्डरी टु एलर्जिक राइनाइटिस एज ए कॉज ऑफ स्लीप डिस्टर्बेंस एंड डे टाइम फटिग एंड रेस्पोंस टु टोपिकल नसल कॉर्टिकोस्टेरॉइड्स। जर्नल ऑफ एलर्जी और क्लीनिकल इम्यूनोलॉजी, 101, 633–637।

डर्बिन, जे. (1951)। इंकंप्लीट ब्लाक्स इन रैनकिंग एक्सपेरीमेंट्स। ब्रिटिश जर्नल ऑफ मेथेमेटिकल एंड स्टेटिस्टिकल साइकोलोजी, 4, 84–90।

फॉसेट, आर. एवं साल्टर, के. (1987)। डिस्ट्रिब्यूशनल स्टडीज एंड द कंप्यूटरः एन अनालिसिस ऑफ डर्बिन्स रैंक टेस्ट। द अमेरिकन स्टेटिस्टिसियन, 41, 81–83।

फ्रीडमैन, एम. (1937)। द यूज ऑफ रेंक्स टु अवॉइड द अजम्पशन ऑफ नोर्मलिटी इंपलीसिट इन द अनालिसिस ऑफ वेरियन्स। जर्नल ऑफ अमेरिकन स्टैटिस्टिकल एसोसिएशन, 32, 675–701।

गोर, ए.पी. (1975)। सम नॉन–पैरामीट्रिक टेस्ट्स एंड सेलेक्शन प्रोसीजर फॉर मैन इफेक्ट्स इन टु–वेज लेआउट्स। अन्नल्स ऑफ द इंस्टीट्यूट ऑफ स्टैटिस्टिकल मेथेमेटिक्स, 27, 487–500।

कोच, जी. (1972)। द यूज ऑफ नॉनपैरामीट्रिक मेथड्स

इन द स्टैटिस्टिकल अनालिसिस ऑफ द टू–पीरियड चेंज–ओवर डिजाइन। बायोमेट्रिक्स, 28, 577–584.

क्रुस्कल, डब्ल्यू.एच. (1952)। ए नॉनपैरामीट्रिक टेस्ट फॉर द सेवेरल सेंपल प्रोब्लेम। अन्नल्स ऑफ मेथेमेटिकल स्टेटिस्टिक्स, 23, 525–540।

कुटनर, एम., नचत्शेम, सी., नेटर, जे. एवं ली, डब्ल्यू. (2005)। अप्लाइड लिनियर स्टैटिस्टिकल मॉडल्स 5जी एड., मैकग्रा–हिल इरविन, बोस्टन ।

लिविंगस्टन, जी.सी., रेनेर, जे.सी.डब्ल्यू. (2022)। नॉनपैरामीट्रिक अनालिसिस ऑफ बैलेन्स्ड इंकम्प्लीट ब्लॉक रैंक डेटा। जर्नल ऑफ स्टैटिस्टिकल थ्योरी एंड प्रैक्टिस, 16, 66–81।

पुट, एम.ई. एवं चिंचिल्ली, वी.एम. (2000)। ए रोबस्ट अनालिसिस ऑफ क्रॉसओवर डिजाइन्स यूजिंग मल्टीसेंपल जनरलाइज्ड एल– स्टेटिस्टिक्स। जर्नल ऑफ अमेरिकन स्टैटिस्टिकल एसोसिएशन, 95, 256–1262।

स्किलिंग्स, जे.एच. एवं मैक, जी.ए. (1981)। ऑन द यूज ऑफ ए फ्रीडमैन टाइप स्टेटिस्टिक्स इन बैलेन्स्ड एंड अनबैलेन्स्ड ब्लॉक डिजाइन्स। टेक्नोमेट्रिक्स, 23, 171–177।

डिजाइन रेसोर्स सर्वर. इंडियन एग्रीकल्चरल स्टेटिस्टिक्स रिसर्च इंस्टीट्यूट (आईसीएआर), न्यू दिल्ली–110 012, इंडिया। "http://www.iasri.res.in/design"www.drs.icar.gov.in।

महाराष्ट्र के पुणे जिले के लिए गूगल अर्थ (Google Earth) इंजन का इस्तेमाल करके भूमि की सतह के तापमान का स्थानिक मानचित्रण

नोबिन चन्द्र पॉल[1], नव्याश्री पोन्नगंटि[1], राहुल बनर्जी[2], धनन्जय डी. नांगारे[1] एवं तौकीर अहमद

[1]भा.कृ.अनु.प. – राष्ट्रीय अजैविक स्ट्रैस प्रबंधन संस्थान, बारामती, महाराष्ट्र

[2]भा.कृ.अनु.प. – भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली

सार

इस अध्ययन में, हमने महाराष्ट्र के पुणे जिले में भूमि की सतह के तापमान को स्थानिक रूप से मैप करने के लिए Google Earth इंजन का इस्तेमाल किया हैं। गूगल अर्थ इंजन में उपलब्ध उपग्रह इमेजरी के विशाल संग्रह का उपयोग करते हुए, हमने अध्ययन क्षेत्र में एलएसटी के स्थानिक पैटर्न का विश्लेषण किया हैं। प्राप्त परिणाम विभिन्न भूमि के प्रकारों में एलएसटी में महत्वपूर्ण भिन्नताओं को प्रकट करते हैं। यह अध्ययन क्षेत्रीय स्तर पर पर्यावरणीय चर की निगरानी और विश्लेषण के लिए एक शक्तिशाली मंच के रूप में गूगल अर्थ इंजन की उपयोगिता को प्रदर्शित करता है।

मुख्य शब्दः पर्यावरणीय सहसंयोजक, GEE (जी), एलएसटी (LST), स्थानिक मानचित्रण।

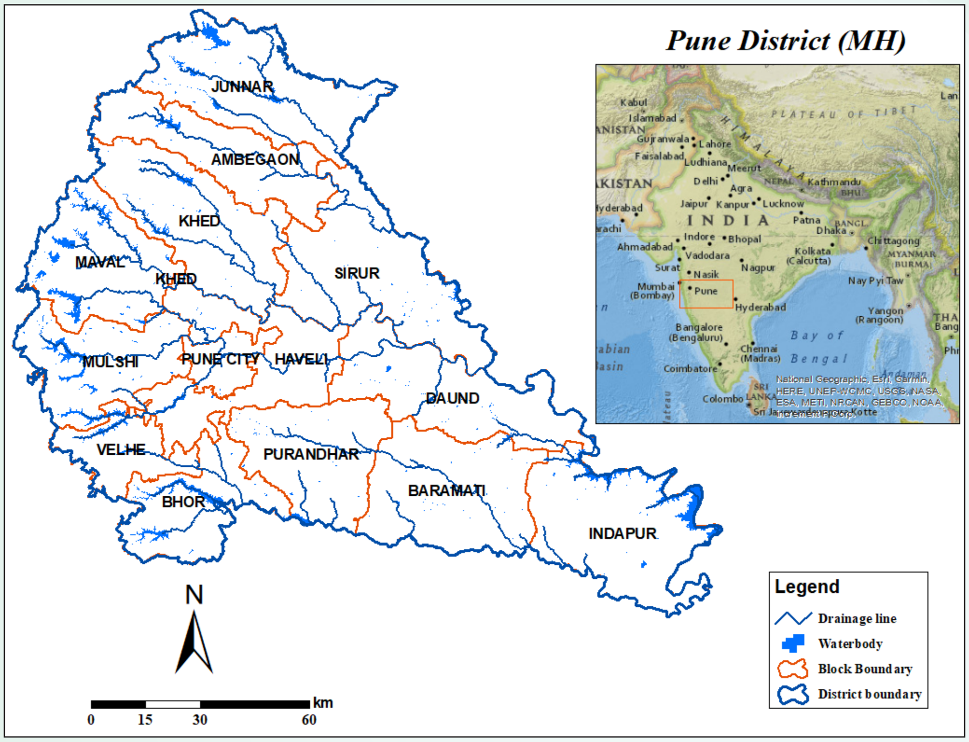
परिचय

इस लेख में, हमने पुणे जिले पर ध्यान केंद्रित किया हैं, जो भारत के महाराष्ट्र में तेजी से शहरीकरण होने वाले क्षेत्रों में से है। पुणे जिले की विशेषता विविध भूमि आवरण प्रकार और भूमि उपयोग पैटर्न है। भाकृअनुप–राष्ट्रीय मृदा सर्वेक्षण एवं भूमि उपयोग योजना ब्यूरो, पुणे द्वारा वर्गीकृत दक्कन पठार अर्ध–शुष्क कृषि–पारिस्थितिक उप–क्षेत्र के भीतर स्थित, इसमें 14 तहसीलें शामिल हैं (चित्र 1)। इस अध्ययन का मुख्य उद्देश्य गूगल अर्थ इंजन (GEE) का उपयोग करके पुणे जिले में भूमि की सतह के तापमान (LST) का स्थानिक मानचित्रण करना था, जो एक क्लाउड–आधारित प्लेटफॉर्म है, जो उपग्रह इमेजरी और प्रसंस्करण क्षमताओं के व्यापक संग्रह तक अभिगम प्रदान करता है। भूमि सतह तापमान (एलएसटी) की तीव्रता और वितरण किसी क्षेत्र की वनस्पति स्थिति से जटिल रूप से जुड़ा हुआ है (संदीप एट अल. 2021)। इस अध्ययन में, 2000 से 2022 तक फैले MODIS वैश्विक LST डेटा (MOD11A2) को गूगल अर्थ इंजन (GEE) का उपयोग करके हासिल किया गया था। (मालव एट. अल. 2022; यादव एट अल 2023)

औसत वार्षिक एलएसटी की गणना प्रत्येक संबंधित वर्ष के लिए सभी छवियों के औसत से की गई थी। इसके बाद, सभी वर्षों में फैले इन औसत वार्षिक डेटासेट का उपयोग दीर्घकालिक (22 वर्ष) औसत एलएसटी की गणना करने के लिए किया गया था। समीकरण (1) का उपयोग करके, डेटा को डिग्री सेल्सियस (°C) में बदल दिया गयाः

$$LST = (0\text{-}02 \times DN - 273.15) \quad (1)$$

जहां, एलएसटी भूमि की सतह का तापमान है, 0.02 स्केल कारक है एवं डीएन पिक्सेल का डिजिटल नंबर है।



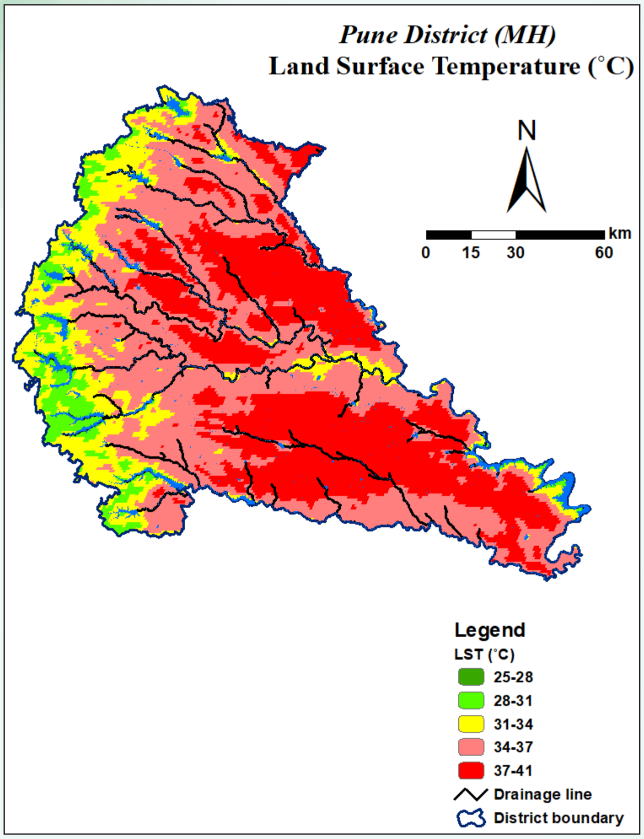
चित्र 1: पुणे जिले का अध्ययन क्षेत्र मानचित्र

पुणे जिले की एलएसटी मैपिंग

एलएसटी मिट्टी की नमी की उपलब्धता से जुड़े एक पैरामीटर के रूप में कार्य करता है, जो बदले में एक क्षेत्र की वनस्पति स्थिति को प्रभावित करता है (फरामार्जी एट.अल. 2018; दास एट.अल. 2023)। कम वनस्पति आवरण वाले क्षेत्र आमतौर पर उच्च एलएसटी प्रदर्शित करते हैं, जो वनस्पति में तनाव की अधिक मात्रा का संकेत देता है। नतीजतन, एलएसटी अप्रत्यक्ष रूप से पौधों द्वारा अनुभव किए गए अजैविक तनाव को दर्शाता है, उच्च एलएसटी मूल्यों को प्रदर्शित करने वाले क्षेत्रों में तनाव के उच्च स्तर देखे गए हैं (पॉल एट. अल. 2024)। हमारे अध्ययन से पुणे जिले में एलएसटी में स्थानिक भिन्नता का पता चलता है, जिसमें तापमान 25 से 41 डिग्री सेल्सियस (चित्र 2) तक होता है। इन तापमानों को पाँच वर्गों में वर्गीकृत किया गया था: 25–28, 28–31, 31–34, 34–37, और 37–41 डिग्री सेल्सियस। पुरंदर, बारामती और दौण्ड के कुछ क्षेत्रों सहित जिले के दक्षिणी और दक्षिण–पूर्वी हिस्सों में स्थित ब्लॉकों ने उच्च एलएसटी मूल्यों का प्रदर्शन किया। विशेष रूप से, उच्च तापमान का स्तर मुख्य रूप से जिले के दक्षिण–पश्चिमी और दक्षिणी क्षेत्रों में देखा गया।

निष्कर्ष

यह अध्ययन पुणे जिले में भूमि की सतह के तापमान के मानचित्रण में गूगल अर्थ इंजन GEE की प्रभावकारिता को दर्शाता है। रिमोट सेंसिंग डेटा और भू–स्थानिक विश्लेषण तकनीकों का एकीकरण क्षेत्र की थर्मल विशेषताओं में मूल्यवान अंतर्दृष्टि प्रदान करता है। एलएसटी के देखे गए स्थानिक पैटर्न शहरी नियोजन और जलवायु नम्यता प्रयासों के लिए महत्वपूर्ण निहितार्थ रखते हैं। कुल मिलाकर, यह अध्ययन पर्यावरणीय चर की रिमोट सेंसिंग–आधारित निगरानी पर शोध में योगदान देता है और स्थानिक विश्लेषण और निगरानी के लिए एक मूल्यवान उपकरण के रूप में जीईई की क्षमता पर प्रकाश डालता है।



चित्र 2: पुणे जिले का एलएसटी मानचित्र

संदर्भ

फरामार्जी, एम., हेइदरीजादी, जेड., मोहम्मदी, ए. एवं हेदरी, एम. (2018)। डिटेक्शन ऑफ वेजटेशन चंगेज इन रिलेशन टू नोर्मालाईज्ड डिफरेंस वेजटेशन इन्डेक्स (NDVI) इन सेमि एरिड रेंजलैंड इन वेस्टर्न ईरान, जर्नल ऑफ एग्रीकल्चर, साइंस एंड टेक्नोलॉजी, 20, 51–60।

दास, बी., देसाई, एस., दारिपा, ए. एट अल. (2023)। लैंड डीग्रेडेशन वल्नेरेबिलिटी मैपिंग इन ए वेस्ट कोस्ट रीवर बेसिन ऑफ इंडिया उसींग एनालिटिकल हायरार्की प्रोसेस कंबाइंड मशीन लर्निंग मॉडल्स, एनवायर्नमेंटल साइंस एंड पोल्लुशण रिसर्च, 30, 83975–83990।

मालव, एल.सी., यादव, बी., टेलर, बी.एल., पटनायक, एस., सिंह, एस.वी., कुमार, एन., रेड्डी, जी.पी.ओ., मीना, बी.एल., द्विवेदी, बी.एस. एवं झा, पी.के. (2022)। मैपिंग ऑफ लैंड डीग्रेडेशन वल्नेरेबिलिटी इन दी सेमि एरिड वाटरशेड ऑफ राजस्थान, इंडिया, सस्टेनेबिलिटी, 14(16): 10198।

पॉल, एन.सी., गायकवाड, बी.एच., नंगारे, डी.डी. एवं रेड्डी, के.एस. (2024)। एन्हान्सिंग क्रॉप रेसिलिएंसः हरनेसिंग मशीन लर्निंग मॉडल्स फॉर एबीओटिक स्ट्रैस मैनेजमेंट, जस्ट एग्रीकल्चर, 4(6), 79–83।

संदीप, पी., रेड्डी, जी.पी.ओ., जेगनकुमार, आर. एवं कुमार, के.ए. (2021)। मॉनिटरिंग ऑफ एग्रीकल्चरल द्रोह्त इन सेमि एरिड एकोसिस्टम ऑफ पेनिन्सुलर इंडिया थ्रू इंडिसेस डेरिवेद फ्रॉम टाइम सीरीज CHIRPS एंड MODIS डाटासेट, इकोलॉजिकल इंडीकेटर्स,121:107033।

यादव, बी., मालव, एल.सी., जिमेनेज–बैलेस्टा, आर., कुमावत, सी., पात्रा, ए., पटेल, ए., जांगिड़, ए., नोगिया, एम., मीना, आर.एल., मोहराना, पी.सी., एट अल. (2023)। मॉडलिंग एंड अस्सेस्मेंट ऑफ लैंड डीग्रेडेशन वल्नेरेबिलिटी इन एरिड इकोसिस्टम ऑफ राजस्थान उसिंग एनालिटिकल हायरार्की प्रोसेस एंड जोस्पेटियल टेक्निक्स, लैण्ड, 12(1), 106।

दीर्घावधि उर्वरक परीक्षणः एक अवलोकन

सुशील कुमार सरकार[1], अनिंदिता दत्ता[1] एवं रवि वंजारी[2]
[1]भा.कृ.अनु.प. – भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली
[2]भा.कृ.अनु.प. – भारतीय मृदा विज्ञान संस्थान, भोपाल

सार

दीर्घावधि उर्वरक परीक्षण, फसल उत्पादकता, मिट्टी के गुणों और प्रक्रियाओं में परिवर्तन, उभरते पोषक असंतुलन और कमियों की पहचान करने और मिट्टी की गुणवत्ता और स्वास्थ्य पर समग्र प्रभाव का अध्ययन करने के लिए सर्वोत्तम संभव मंच प्रदान करता हैं। इस परियोजना का मुख्यालय भा.कृ.अनु.प. –भारतीय मृदा विज्ञान संस्थान, भोपाल में है। इसके देशभर में 18 केंद्र हैं, जिनपर यह परीक्षण नियोजित किया जाता है। दीर्घावधि उर्वरक परीक्षणों से यह निष्कर्ष निकाला गया है कि, उर्वरकों का सही तरीके से इस्तेमाल के बिना मृदा की उत्पादकता को बनाए रखना संभव नही होगा।

मुख्य शब्दः दीर्घावधि उर्वरक परीक्षण, मिट्टी की गुणवत्ता, पोषक तत्व निष्कासन, सूक्ष्म पोषक तत्व, उत्पादकता।

परिचय

मृदा की उत्पादकता और मिट्टी की गुणवत्ता पर रासायनिक उर्वरक के प्रभाव का अध्ययन करने के लिए, भारतीय कृषि अनुसंधान परिषद ने सितंबर 1970 में 11 केंद्रों पर "दीर्घावधि उर्वरक परीक्षण पर अखिल भारतीय समन्वित अनुसंधान परियोजना (एआईसीआरपी–एलटीएफई)" शुरू करने का निर्णय लिया। 1997 में पंचवार्षिक समीक्षा समिति द्वारा दीर्घावधि उर्वरक परीक्षण के विभिन्न केंद्रों पर किए गए कार्यों की समीक्षा की गई और परियोजना के अधिदेश और उद्देश्यों को बढ़ाने की सिफारिश की गई और शीर्षक को ''मिट्टी की गुणवत्ता, फसल उत्पादकता और स्थिरता में परिवर्तन का अध्ययन करने के लिए दीर्घकालिक उर्वरक प्रयोग'' पर अखिल भारतीय समन्वित अनुसंधान परियोजना (एआईसीआरपी) के रूप में बदल दिया गया"। परियोजना के उद्देश्य निम्नवत हैः

- फसल की उपज, पोषक तत्वों की संरचना और एकाधिक फसल प्रणालियों में ग्रहण पर द्वितीयक और सूक्ष्म पोषक तत्वों (आवश्यकता के अनुसार) सहित कार्बनिक और अकार्बनिक रूपों में पौधों के पोषक तत्वों के एकल और संयोजन में निरंतर अनुप्रयोग के प्रभाव का अध्ययन करना।
- फसल की पैदावार पर माध्यमिक और सूक्ष्म पोषक तत्वों (आवश्यकतानुसार) के अनुप्रयोग के प्रभाव का अध्ययन करना और गहन फसल कार्यक्रम के तहत इन तत्वों की आवश्यकता का आकलन करना।
- फसलों द्वारा पोषक तत्व निष्कासन की मात्रा ज्ञात करना।
- इसकी उत्पादकता के संबंध में मिट्टी की भौतिक, रासायनिक और सूक्ष्मजीवविज्ञानी विशेषताओं के संबंध में निरंतर खाद और फसल के परिणामस्वरूप मिट्टी के गुणों में परिवर्तन की निगरानी करना।
- अवशेष निर्माण और मिट्टी की उत्पादकता पर जैवनाशी रसायनों (खरपतवारनाशी और कीटनाशकों) के उपयोग के प्रभाव की गहन जांच करना।
- प्रस्तावित खाद और फसल कार्यक्रम के तहत मिट्टी जनित बीमारियों की घटनाओं और कीटों और रोगजनकों में परिवर्तन का आकलन करना।

- फसल उत्पादकता और मिट्टी के स्वास्थ्य पर बदलते जलवायु के प्रभाव का आकलन करना ।

प्रणाली

महत्वपूर्ण फसल प्रणालियों के साथ विभिन्न कृषि पारिस्थितिक क्षेत्रों (एईजेड) में निश्चित स्थलों पर दीर्घावधि उर्वरक परीक्षण का संचालन करने का उद्देश्य न केवल उर्वरकों और जैविक के माध्यम से पौधों के पोषक तत्वों के आदान–प्रदान के निरंतर अनुप्रयोग के कारण मिट्टी के गुणों और उपज प्रतिक्रियाओं और मिट्टी के वातावरण में परिवर्तन की निगरानी करना था। स्रोत, बल्कि मिट्टी की गुणवत्ता में सुधार और पर्यावरणीय क्षरण को कम करने के लिए उर्वरकों के तर्कसंगत उपयोग और प्रबंधन के लिए रणनीतियों और नीतियों को विकसित करना भी है। इस प्रकार, दीर्घावधि उर्वरक परीक्षण पर अखिल भारतीय समन्वित अनुसंधान परियोजना का जोर उत्पादकता, स्थिरता और पर्यावरण सुरक्षा पर है।

उपचार विवरण

उपचार–1: 50% इष्टतम एनपीके
उपचार–2: 100% इष्टतम एनपीके
उपचार–3: 150% इष्टतम एनपीके
उपचार–4: 100% इष्टतम एनपीके + हाथ से निराई
उपचार–5: 100% इष्टतम एनपीके + जिंक या चूना
उपचार–6: 100% इष्टतम एनपी
उपचार–7: 100% इष्टतम एन
उपचार–8: 100% इष्टतम एनपीके + जैविक खाद
उपचार–9: 100% इष्टतम एनपीके (सल्फर मुक्त स्रोत)
उपचार–10: अप्रशिक्षित (नियंत्रण)

अनुपचारित

अखिल भारतीय समन्वित अनुसंधान परियोजना – दीर्घावधि उर्वरक परीक्षण के समन्वयन केंद्रः

आईसीएआर–भारतीय मृदा विज्ञान संस्थान, भोपाल का मुख्यालय है और यह एआईसीआरपी–एलटीएफई

तालिका 1: एआईसीआरपी, एलटीएफई के विभिन्न स्थानों पर फसल प्रणाली और उर्वरक दरें

स्थान	प्रारम्भ वर्ष	मृदा प्रकार	फसल प्रणाली	प्रारंभिक मिट्टी परीक्षण के आधार पर 100% एनपीके पर मृदा प्रकार फसल प्रणाली उर्वरक दर		
बैरकपुर (पश्चिम बंगाल)	1971	इन्सेप्टिसोल	चावल	120	26	50
			गेहूं	120	26	50
			जूट	60	13	50
लुधियाना (पंजाब)	1971	इन्सेप्टिसोल	मक्का	120	26	25
			गेहूं	120	26	25
नई–दिल्ली (दिल्ली)	1971	इन्सेप्टिसोल	मक्का	120	26	33
			गेहूं	120	26	33
कोयंबटूर (तमिलनाडू)	1971	इन्सेप्टिसोल	रागी	90	20	14.5
			मक्का	135	29.5	29
जबलपुर (म.प्र.)	1972	वर्टीसोल	सोयाबीन	20	35	17
			गेहूं	120	35	33
बेंगलुरु (कर्नाटक)	1986	अल्फिसोल्स	रागी	100	22	41.5
			मक्का	100	33	83
जगतियाल (तेलंगाना)	2000	इन्सेप्टिसोल	चावल (खरीफ)	120	26	33
			चावल (रबी)	120	26	33

रांची (झारखंड)	1972	अल्फिसोल्स	सोयाबीन	25	26	33
			गेहूं	80	26	33
भुवनेश्वर (उड़ीसा)	2002	इन्सेप्टिसोल	चावल	80	17.5	50
			चावल	80	17.5	50
पालमपुर (हि. प्र.)	1972	अल्फिसोल्स	मक्का	120	26	33
			गेहूं	120	26	25
पंतनगर (उत्तरांचल)	1971	मोल्लिसोल	चावल	120	26	37
			गेहूं	120	26	33
जूनागढ़ (गुजरात)	1997	वर्टीसोल	मूँगफली	12.5	11	0
			गेहूं	120	26	50
अकोला (महाराष्ट्र)	2005	वर्टीसोल	ज्वार	100	22	33
			गेहूं	120	26	50
पट्टाम्बी (केरल)	1997	अल्फिसोल्स	चावल	90	20	37
			चावल	90	20	37
रायपुर (छत्तीसगढ़)	1997	वर्टीसोल	चावल	100	26	33
			गेहूं	100	26	33
उदयपुर (राजस्थान)	1997	इन्सेप्टिसोल	मक्का	120	26	25
			गेहूं	120	26	25
परभणी (महाराष्ट्र)	1997	वर्टीसोल	सोयाबीन	30	26	25
			कुसुम	60	17.5	0

के सहयोगी केंद्रों के लिए परियोजना समन्वय इकाई के रूप में कार्य करता है। सहयोगी केंद्रों की सूची, उनकी फसल प्रणालियों, शुरुआत की तारीख और 100ः एनपीके खुराक के साथ तालिका 1 में दी गई है।

प्रमुख निष्कर्ष

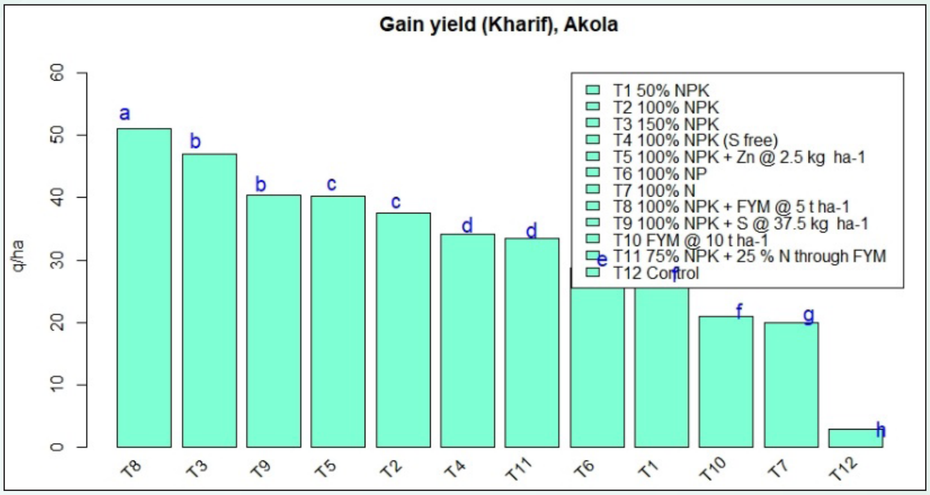
देश भर में किए गए दीर्घावधि उर्वरक परीक्षण (एलटीएफई) से स्पष्ट रूप से पता चला कि संतुलित और एकीकृत पोषक तत्व प्रबंधन ने फसल उत्पादकता, मिट्टी की गुणवत्ता और समग्र मिट्टी के स्वास्थ्य को बनाए रखा है।

राँची केंद्र के दीर्घावधि उर्वरक परीक्षण का विश्लेषण

उपचार माध्यों का तुलनात्मक अध्ययन

विश्लेषण के केंद्रवार परिणाम नीचे दिए गए हैं।

खरीफ फसल ज्वार के लिए अकोला के अनाज उपज डेटा के 8 वर्षों (2007–08 से 2014–15) के विश्लेषण से पता चलता है कि उपचार–8 (5 टन/हेक्टेयर पर 100% एनपीके + जैविक खाद) से 51.12 क्विंटलध्हेक्टेयर की उच्चतम अनाज उपज मिलती है। उपचार उपचार–3 (150% एनपीके) द्वारा जो 46.95 क्विंटल प्रति हेक्टेयर देता है। उपचार उपचार–10 (नियंत्रण) के लिए सबसे कम अनाज उपज (2.93 क्विंटल/हेक्टेयर) दर्ज की गई है। इसी प्रकार, रबी फसल गेहूं के लिए अकोला केंद्र के 7 वर्षों के अनाज उपज डेटा का संयुक्त विश्लेषण किया गया। उपचार–8 (100ः एनपीके + जैविकखाद 5 टन/हेक्ट.) के लिए 30.52 q/ha की उच्चतम अनाज उपज देखी गई, इसके बाद उपचार–3 (150% एनपीके) के लिए 28.04 q/ha दर्ज की गई। उपचार–10 (नियंत्रण) के लिए सबसे कम अनाज उपज 0.534 क्विंटल/हेक्टेयर देखी गई। इससे स्पष्ट रूप से पता चलता है कि

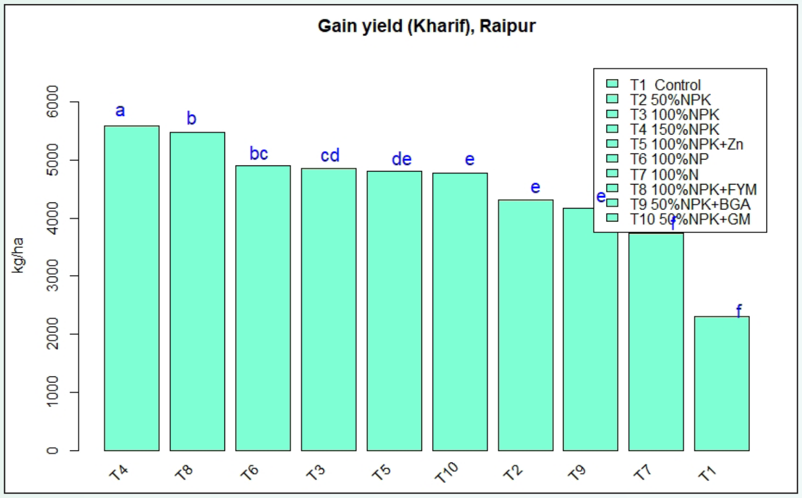


अकोला क्षेत्र में ज्वार और गेहूं दोनों में अनाज की उपज स्थिरता के लिए 5 टन/हेक्टेयर पर 100% एनपीके + जैविक खाद सर्वोत्तम है। शून्य उर्वरक का प्रयोग किसी भी फसल में अनाज की पैदावार को कायम नहीं रख सकता।

पालमपुर केंद्र के 45 वर्षों के आंकड़ों (1973–74 से 2017–18) के आधार पर खरीफ और रबी फसलों के लिए अनाज की उपज का विश्लेषण किया गया। यह पाया गया कि उपचार–8 (100ः एनपीके + जैविक खाद) उच्चतम अनाज उपज (31 क्विंटलध्हेक्टेयर) देता है और उपचार–10 (100% एनपीके + चूना) भी सांख्यिकीय रूप से गेहूं (रबी) फसल के लिए उपचार–8 के बराबर है। गेहूं (रबी) की फसल के लिए उपचार–7 (100% एन) के लिए सबसे कम उपचार औसत (1.83 क्विंटल/हेक्टेयर) देखा गया। मक्का (खरीफ) के लिए, फिर से उपचार–8 (100% एनपीके + जैविक खाद) उच्चतम अनाज उपज (48.26 कु./हेक्ट.) देता है, उसके बाद उपचार–10 (100% एनपीके + चूना) 43.23 कु./हेक्ट. है और ये सांख्यिकीय रूप से बराबर थे। मक्का (खरीफ) के लिए उपचार–7 (100ः एन) के लिए सबसे कम अनाज उपज (3.5 क्विंटल/हेक्टेयर) देखी गई। विश्लेषण से पता चलता है कि 100% एन अकेले पालमपुर की मिट्टी के प्रकार में मक्का और गेहूं में अनाज की उपज को बनाए नहीं रख सकता है। इसी प्रकार, शून्य उर्वरक प्रयोग भी पालमपुर की मिट्टी के प्रकार में मक्का और गेहूं में अनाज की उपज को बनाए नहीं रख सकता है।

रायपुर केंद्र के 5 वर्षों (2012–13 से 2016–17) के आंकड़ों के आधार पर खरीफ और रबी फसलों के लिए अनाज की उपज का विश्लेषण किया गया। यह देखा गया कि उपचार–4 (150% एनपीके) सबसे अधिक अनाज उपज (55.84 क्विंटल/हेक्टेयर) देता है, उसके बाद उपचार–8 (100% एनपीके + जैविकखाद) होता है जो 54.73 क्विंटल/हेक्टेयर देता है। चावल (खरीफ) की फसल के लिए उपचार4 और उपचार8 दोनों सांख्यिकीय रूप से एक दूसरे के बराबर हैं। चावल की फसल के लिए उपचार1 (नियंत्रण) के लिए अनाज की सबसे कम उपज 23 क्विंटल/हेक्टेयर देखी गई। गेहूं (रबी) के लिए, फिर से उपचार–4 (150% एनपीके) ने उच्चतम अनाज उपज (36.13 क्विंटलध्हेक्टेयर) दी, उसके बाद उपचार8 (100% एनपीके + जैविकखाद) ने 34.04 क्विंटल/हेक्टेयर दिया और ये सांख्यिकीय रूप से बराबर थे। गेहूं (रबी) के लिए उपचार1 (नियंत्रण) में सबसे कम अनाज उपज (10.44 क्विंटल/हेक्टेयर) देखी गई। विश्लेषण से पता चलता है कि अकेले शून्य उर्वरक प्रयोग से रायपुर की मिट्टी के प्रकार में चावल और गेहूं में अनाज की पैदावार को बरकरार नहीं रखा जा सकता है।

देश भर में किए गए दीर्घावधि उर्वरक परीक्षण से स्पष्ट रूप से पता चला कि संतुलित और एकीकृत



पोषक तत्व प्रबंधन ने फसल उत्पादकता, मिट्टी की गुणवत्ता और समग्र मिट्टी के स्वास्थ्य को बनाए रखा। एआईसीआरपी एलटीएफई से निकले निष्कर्षों पर यहां प्रकाश डाला गया है:

दीर्घावधि उर्वरक परीक्षण (एलटीएफई) में फसल उपज का रुझान 100% एनपीके + जैविक खाद > 150% एनपीके > 100% एनपीके + जिंक / चूना >100% एनपीके > 50% एनपीके > 100% एनपी > 100% के क्रम में है। एन>अधिकांश एलटीएफई स्थानों पर नियंत्रण। संतुलित उर्वरक उपयोग ने लगभग सभी स्थानों पर फसल उत्पादकता को बनाए रखा। एकीकृत पोषक तत्व प्रबंधन (100% एनपीके + जैविक खाद) ने एलटीएफई के तहत फसल उत्पादकता और मिट्टी की गुणवत्ता में और सुधार किया ।

संतुलित उर्वरक उपयोग के साथ–साथ एकीकृत पोषक तत्व प्रबंधन (100% एनपीके+जैविक खाद) से मिट्टी की भौतिक, रासायनिक और जैविक पारस्थितिकी में सुधार हुआ।

पालमपुर, रांची और बेंगलुरु के अल्फिसोल्स के फसल उपज के आंकड़ों से संकेत मिलता है कि जहां तक फसल उत्पादकता का सवाल है, जैविक खाद के साथ संशोधित मिट्टी चूने से बेहतर पाई गई है। जैविक खाद का प्रयोग पोषक तत्वों की आपूर्ति के अलावा मिट्टी की स्थिति को नियंत्रित करने में मदद करता है जबकि चूना केवल मिट्टी के पीएच को बढ़ाता है।

पोषक तत्वों (उर्वरकों और खाद) के संतुलित उपयोग के परिणामस्वरूप मिट्टी के सूक्ष्मजीवों की आबादी में वृद्धि हुई। पोषक तत्वों के अनुप्रयोग में 100 से 150% की वृद्धि से मिट्टी में सूक्ष्मजीवों की संख्या और किण्वक (एंजाइम) गतिविधियों पर भी सकारात्मक प्रभाव पड़ा।

रायपुर के वर्टिसोल में, चावल और गेहूं में उर्वरक पोषक तत्वों के संतुलित उपयोग के साथ–साथ एकीकृत पोषक तत्व प्रबंधन (आईएनएम) के साथ उपज को बनाए रखा जा सकता है ।

मिट्टी और फसल की परवाह किए बिना, जैविक खाद या हरी खाद को शामिल करने से न केवल उत्पादकता में वृद्धि हुई, बल्कि माइक्रोबियल गिनती और उनकी एंजाइमेटिक गतिविधियों में भी वृद्धि हुई।

संतुलित उर्वरकों के साथ लगातार जैविक खाद के प्रयोग से असंतुलन और अखाद्य उपचारों की तुलना में मिट्टी के थोक घनत्व में कमी पाई गई। इसी प्रकार, एलटीएफई के तहत संतुलन और आईएनएम (100% एनपीके़जैविक खाद) प्रथाओं के तहत हाइड्रोलिक चालकता, औसत वजन व्यास, जल स्थिर समुच्चय, घुसपैठ आदि में उल्लेखनीय सुधार हुआ।

मिट्टी के जैविक मापदंडों जैसे माइक्रोबियल गिनती,

डीहाइड्रोजनेज गतिविधि, बायोमास सी और एन, यूरिया गतिविधि, फॉस्फेट एंजाइम को एलटीएफई के तहत मिट्टी के प्रकारों में संतुलित और आईएनएम (100% एनपीके + जैविक खाद) पोषक तत्वों के विकल्प के साथ प्रोत्साहित किया गया है।

एलटीएफई साइटों पर किए गए पोषक तत्वों की गतिशीलता के अध्ययन से पता चलता है कि मैक्रो और सूक्ष्म पोषक तत्व दोनों अंश 100% एनपीके + जैविक खाद में उच्चतम पाए गए। अधिकांश एलटीएफई में संतुलन पोषक तत्व अनुप्रयोग और आईएनएम उपचार में पानी में घुलनशील, विनिमेय और गैर–विनिमेय K को उच्च दर्ज किया गया।

पोषक तत्वों के असंतुलित अनुप्रयोग के परिणामस्वरूप एलटीएफई में प्रमुख फसलों की कम उपज हुई। इसके विपरीत, पोषक तत्वों का संतुलित उपयोग फसल उत्पादकता के साथ–साथ पोषक तत्वों के अवशोषण को भी बनाए रखता है। INM ने 100% NPK+FYM के माध्यम से पोषक तत्व ग्रहण में और सुधार किया।

परिणामों से पता चला कि भूखंडोंध्खेतों में पी की मात्रा घटकर आधी हो गई, जिसमें लुधियाना (पंजाब), बैंगलोर (कर्नाटक) और जबलपुर (मध्य प्रदेश) की मिट्टी के मामले में पी जमा हुआ, जिससे फसल उत्पादकता पर कोई प्रतिकूल प्रभाव नहीं पड़ा। इस प्रकार, यह अनुमान लगाया जा सकता है कि वर्षों से संचित पी का इन क्षेत्रों में पुनः उपयोग किया जा सकता है।

उर्वरक के संतुलित उपयोग ने मिट्टी में डिहाइड्रोजनेज गतिविधि जैसी एंजाइमेटिक गतिविधियों को उत्प्रेरित किया जिसके परिणामस्वरूप अधिक CO_2 का विकास हुआ। इस प्रकार, अवलोकनों से स्पष्ट रूप से पता चला कि उर्वरक का संतुलित उपयोग न केवल फसल उत्पादकता को बनाए रखने के लिए बल्कि मिट्टी में सूक्ष्मजीवों की आबादी को बनाए रखने या बढ़ाने के लिए भी आवश्यक है।

पंतनगर (मोलिसोल्स) में चावल की उपज से संकेत मिलता है कि Zn और S के संयुक्त अनुप्रयोग का योगात्मक प्रभाव था। इस प्रकार, उत्तराखंड के मोलिसोल्स में फसल उत्पादकता को बढ़ानेध्बनाए रखने के लिए Zn और S को एक साथ आपूर्ति करने की आवश्यकता है।

संतुलित उर्वरक के उपयोग से उत्पादकता में वृद्धि के परिणामस्वरूप मिट्टी में कार्बन और माइक्रोबियल आबादी में वृद्धि हुई और इस प्रकार इस धारणा को खारिज कर दिया गया कि रासायनिक उर्वरक मिट्टी के कार्बन को खराब करते हैं और सूक्ष्मजीवों के विकास पर प्रतिकूल प्रभाव डालते हैं।

कार्बन प्रबंधन सूचकांक (सीएमआई) 100% एनपीके+जैविक खाद के साथ अधिकतम था। उच्च सीएमआई मान वाली मिट्टी को बेहतर प्रबंधित मिट्टी माना जाता है। संतुलित या असंतुलित पोषक तत्व अनुप्रयोग की तुलना में सीएमआई में 100% एनपीके़जैविक खाद द्वारा उल्लेखनीय वृद्धि हुई है।

जनजातीय उपयोजना (टीएसपी) और अनुसूचित जाति उपयोजना (एससीएसपी) के तहत किसानों के खेत पर प्रौद्योगिकी प्रदर्शन ने असंतुलन और किसानों पर पोषक तत्वों की संतुलित और एकीकृत पोषक तत्व प्रबंधन (आईएनएम) खुराक अपनाने पर उपज लाभ और मौद्रिक लाभ को स्पष्ट रूप से प्रदर्शित किया। 'अभ्यास।

संतुलित पोषक तत्व अनुप्रयोग अधिक टिकाऊ उपज सूचकांक (एसवाईआई) देता है और एकीकृत पोषक तत्व प्रबंधन (आईएनएम) ने एसवाईआई में और सुधार किया है। अल्फिसोल्स में, अकेले एन के अनुप्रयोग ने संतुलन और आईएनएम उपचार की तुलना में एसवाईआई के कम मूल्य दिए।

मृदा गुणवत्ता सूचकांक (एसक्यूआई) 30–35 मिट्टी मापदंडों (रासायनिक, भौतिक और जैविक) को ध्यान में रखकर निकाला गया था और यह पाया गया कि संतुलित पोषक तत्व अनुप्रयोग में एसक्यूआई अधिक था। इसके अलावा, जैविक खाद / लाइम के साथ एनपीके के अनुप्रयोग से एसक्यूआई में सुधार हुआ।

भारत में दीर्घावधि उर्वरक परीक्षण के तहत वर्षों से मिट्टी के साथ–साथ फसल पर डेटासेट का उपयोग

करके आईसीएआर–भारतीय मृदा विज्ञान संस्थान, भोपाल द्वारा एक मृदा कार्बन (सी) और नाइट्रोजन (एन) टर्नओवर मॉडल विकसित किया गया है।

पालमपुर, रांची और बेंगलुरु जहां की मृदा प्रकार अल्फिसोल्स है की फसल उत्पादकता जैविक खाद के साथ संशोधित मिट्टी का अनुप्रयोग, केवल चूने के अनुप्रयोग से बेहतर पाई गई है। जैविक खाद का प्रयोग पोषक तत्वों की आपूर्ति के अलावा मिट्टी की पारिस्थितिकी को नियंत्रित करने में मदद करता है जबकि चूना केवल मिट्टी के पीएच को बढ़ाता है।

पोषक तत्वों (उर्वरकों और खाद) के संतुलित उपयोग के परिणामस्वरूप मिट्टी के सूक्ष्मजीवों की आबादी में वृद्धि हुई । पोषक तत्वों के अनुप्रयोग में 100 से 150% की वृद्धि से मिट्टी में सूक्ष्मजीवों की संख्या और किण्वक (एंजाइम) गतिविधियों पर भी सकारात्मक प्रभाव पड़ा।

रायपुर के वर्टिसोल मृदा में, चावल और गेहूं की फसल में उर्वरक पोषक तत्वों के संतुलित उपयोग के साथ–साथ एकीकृत पोषक तत्व प्रबंधन (आईएनएम) के साथ उपज को बनाए रखा जा सकता है।

मिट्टी और फसल की परवाह किए बिना, जैविक खाद या हरी खाद को शामिल करने से न केवल उत्पादकता में वृद्धि हुई, बल्कि माइक्रोबियल गिनती और उनकी एंजाइमेटिक गतिविधियों में भी वृद्धि हुई।

संतुलित उर्वरकों के साथ लगातार जैविक खाद के प्रयोग से असंतुलन और अखाद्य उपचारों की तुलना में मिट्टी के थोक घनत्व में कमी पाई गई। इसी प्रकार, एलटीएफई के तहत संतुलन और आईएनएम (100% एनपीके+जैविक खाद) प्रथाओं के तहत हाइड्रोलिक चालकता, औसत वजन व्यास, जल स्थिर समुच्चय, घुसपैठ आदि में उल्लेखनीय सुधार हुआ।

मिट्टी के जैविक मापदंडों जैसे माइक्रोबियल गिनती, डीहाइड्रोजनेज गतिविधि, बायोमास सी और एन, यूरिया गतिविधि, फॉस्फेट एंजाइम को एलटीएफई के तहत मिट्टी के प्रकारों में संतुलित और आईएनएम (100% एनपीके + जैविक खाद) पोषक तत्वों के विकल्प के साथ प्रोत्साहित किया गया है।

एलटीएफई साइटों पर किए गए पोषक तत्वों की गतिशीलता के अध्ययन से पता चलता है कि मैक्रो और सूक्ष्म पोषक तत्व दोनों अंश 100% एनपीके + जैविक खाद में उच्चतम पाए गए। अधिकांश एलटीएफई में संतुलन पोषक तत्व अनुप्रयोग और आईएनएम उपचार में पानी में घुलनशील, विनिमेय और गैर–विनिमेय K को उच्च दर्ज किया गया।

पोषक तत्वों के असंतुलित अनुप्रयोग के परिणामस्वरूप एलटीएफई में प्रमुख फसलों की कम उपज हुई। इसके विपरीत, पोषक तत्वों का संतुलित उपयोग फसल उत्पादकता के साथ–साथ पोषक तत्वों के अवशोषण को भी बनाए रखता है। प्छड ने 100% NPK+FYM के माध्यम से पोषक तत्व ग्रहण में और सुधार किया।

लुधियाना (पंजाब), बैंगलोर (कर्नाटक) और जबलपुर (मध्य प्रदेश) के परिणामों से यह भी पता चला कि भूखंडों/खेतों में जहां मिट्टी में फास्फोरस काफी मात्रा में संचित है, यदि अनुसंशित मात्रा का आधा फास्फोरस का अनुप्रयोग किया जाए तो भी फसल उत्पादकता पर कोई प्रतिकूल प्रभाव नहीं पड़ता है। इस प्रकार, यह अनुमान लगाया जा सकता है कि वर्षों से संचित फास्फोरस का इन क्षेत्रों में पुनः उपयोग किया जा सकता है।

उर्वरक के संतुलित उपयोग ने मिट्टी में डिहाइड्रोजनेज गतिविधि जैसी एंजाइमेटिक गतिविधियों को उत्प्रेरित किया जिसके परिणामस्वरूप अधिक CO_2 का विकास हुआ। इस प्रकार, अवलोकनों से स्पष्ट रूप से पता चला कि उर्वरक का संतुलित उपयोग न केवल फसल उत्पादकता को बनाए रखने के लिए बल्कि मिट्टी में सूक्ष्मजीवों की आबादी को बनाए रखने या बढ़ाने के लिए भी आवश्यक है।

पंतनगर (मोलिसोल्स) में चावल की उपज से संकेत मिलता है कि जिंक (Zn) और सल्फर (S) के संयुक्त अनुप्रयोग का योगात्मक प्रभाव था। इस प्रकार, उत्तराखंड

के मोलिसोल्स में फसल उत्पादकता को बढ़ाने/बनाए रखने के लिए जिंक (Zn) और सल्फर (S) की एक साथ आपूर्ति करने की आवश्यकता है।

संतुलित उर्वरक के उपयोग से उत्पादकता में वृद्धि के साथ–साथ मिट्टी में कार्बन और सूक्ष्मजीवों की आबादी में वृद्धि हुई और इस प्रकार इस धारणा को खारिज कर दिया गया कि रासायनिक उर्वरक मिट्टी के कार्बन को खराब करते हैं और सूक्ष्मजीवों के विकास पर प्रतिकूल प्रभाव डालते हैं।

कार्बन प्रबंधन सूचकांक (सीएमआई) 100% एनपीके्जैविक खाद के साथ अधिकतम था। उच्च सीएमआई मान वाली मिट्टी को बेहतर प्रबंधित मिट्टी माना जाता है। संतुलित या असंतुलित पोषक तत्व अनुप्रयोग की तुलना में सीएमआई में 100% एनपीके+जैविक खाद द्वारा उल्लेखनीय वृद्धि हुई है।

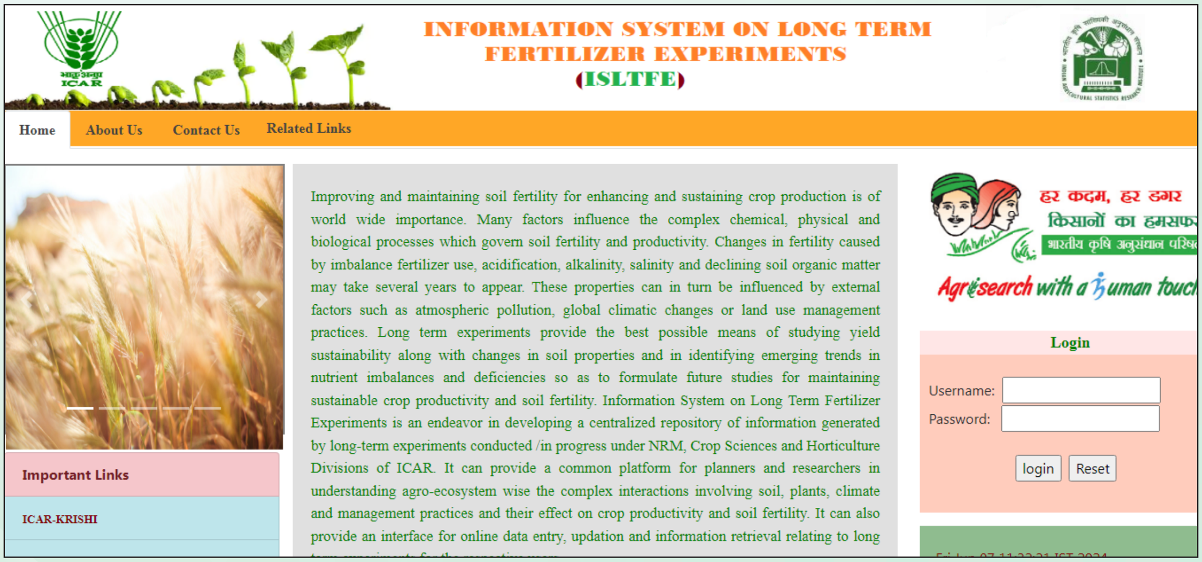
जनजातीय उपयोजना (टीएसपी) और अनुसूचित जाति उपयोजना (एससीएसपी) के तहत किसानों के खेत पर प्रौद्योगिकी प्रदर्शन ने असंतुलन और किसानों पर पोषक तत्वों की संतुलित और एकीकृत पोषक तत्व प्रबंधन (आईएनएम) खुराक अपनाने पर उपज लाभ और मौद्रिक लाभ को स्पष्ट रूप से प्रदर्शित किया। 'अभ्यास।

संतुलित पोषक तत्व अनुप्रयोग, टिकाऊ उपज सूचकांक (एसवाईआई) देता है और एकीकृत पोषक तत्व प्रबंधन (आईएनएम) ने टिकाऊ उपज सूचकांक में और सुधार किया है। अल्फिसोल्स मृदा में, अकेले नत्रजन (एन) के अनुप्रयोग ने संतुलित और एकीकृत पोषक तत्व प्रबंधन की तुलना में टिकाऊ उपज सूचकांक के सापेक्ष काफी कम है ।

मृदा गुणवत्ता सूचकांक (एसक्यूआई) 30–35 मिट्टी मापदंडों (रासायनिक, भौतिक और जैविक) को ध्यान में रखकर निकाला गया था और यह पाया गया कि संतुलित पोषक तत्व अनुप्रयोग में एसक्यूआई अधिक था। इसके अलावा, जैविक खाद / चूना के साथ एनपीके के अनुप्रयोग से एसक्यूआई में सुधार हुआ।

भारतीय कृषि सांख्यिकी अनुसंधान संस्थान (ICAR-IASRI), नई दिल्ली द्वारा। ICRP LTFE डेटाबेस यानी दीर्घावधि उर्वरक परीक्षण पर सूचना प्रणाली (ISLTFE) विकसित और अद्यतन किया गया। कृषि परियोजना के सहयोग से विभिन्न केंद्रों के लिए ISLTFE पर डेटा अपलोड किया गया है। डेटा अपडेट का काम चल रहा है।

भारत में दीर्घावधि उर्वरक परीक्षण के तहत वर्षों से मिट्टी के साथ–साथ फसल पर डेटासेट का उपयोग करके आईसीएआर–भारतीय मृदा विज्ञान संस्थान, भोपाल द्वारा एक मृदा कार्बन (सी) और नाइट्रोजन (एन) टर्नओवर मॉडल विकसित किया गया है।



सुडोकू वर्ग अभिकल्पनाएँ एवं उनके अनुप्रयोग

नेहाताई डब्ल्यू. अगाशे, सिनी वर्गीस, मो० हारून, आशुतोष दलाल एवं देवेंद्र कुमार
भा.कृ.अनु.प. – भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली

सार

सुडोकू दुनिया भर में लाखों लोगों द्वारा खेला जाता है। सुडोकू वर्ग अभिकल्पनाओं की पुरःस्थापना का संयोजन अभिकल्पना सिद्धांतों पर एक महत्वपूर्ण प्रभाव है, क्योंकि वे तीन स्रोतों वाली भिन्नता से विषमता को हटाने में सक्षम हैं। यहां सुडोकू वर्ग अभिकल्पना और प्रयोगों के अभिकल्पन में उनके कुछ अनुप्रयोगों पर चर्चा की गई है। आंशिक रूप से संतुलित अपूर्ण ब्लॉक अभिकल्पना हेतु एक श्रृंखला की निर्माण विधि का वर्णन किया गया है। गुणात्मक और मात्रात्मक कारकों के लिए सुडोकू आधारित स्थान–पूर्ति (space-filling) अभिकल्पना नामक एक नई प्रकार की अभिकल्पना के निर्माण हेतु, एक नए दृष्टिकोण को कंप्यूटर परीक्षणों में विशेष अनुप्रयोग के साथ समझाया भी गया है। इसके अतिरिक्त, इन अभिकल्पनाओं को उप–अभिकल्पना के समूहों में इस प्रकार विभाजित किया जा सकता है कि, पूर्ण अभिकल्पना और प्रत्येक उप–अभिकल्पना अविभाज्य एवं द्विचर उपांत में अधिकतम एकरूपता प्राप्त कर सके। इन अभिकल्पनाओं की लक्षण–वर्णन विशेषताओं का अध्ययन करने के लिए R कोड भी लिखा गया है।

मुख्य शब्दः आंशिक संतुलित अपूर्ण खण्ड (PBIB) अभिकल्पना, कंप्यूटर परीक्षण, स्थान–पूर्ति अभिकल्पना, सुडोकू वर्ग अभिकल्पना।

परिचय

सुडोकू एक पहेली है जिसमें वृद्धिशील बुद्धि सम्मिलित है और यह दुनिया भर में बेहद लोकप्रिय हो गई है। सुडोकू अब दुनिया भर के कई समाचार पत्रों और प्रकाशनों में एक नियमित विशेषता है। यह अपनी आकर्षक और व्यसन जैसी संयोजन समस्या को हल करने के लिए जीवन के सभी क्षेत्रों से पाठकों को आकर्षित करता है। सुडोकू का सबसे सामान्य पैटर्न 1 से 9 तक की संख्याओं को एक वर्ग जिसमें 9 पंक्तियों को 9 स्तंभ और उप–विभाजित 9 बॉक्स होते हैं, में इस प्रकार व्यवस्थित करना है, ताकि प्रत्येक अंक प्रत्येक पंक्ति, स्तंभ और बॉक्स में 1 से 9 तक की संख्याएं केवल एक बार दिखाई दे। प्रत्येक सुडोकू वर्ग एक विशेष प्रकार का लैटिन वर्ग है (जहां 1 से n तक की संख्याओं को एक $n \times n$ सारणी में इस प्रकार व्यवस्थित किया जाता है कि प्रत्येक पंक्ति और प्रत्येक स्तंभ में प्रत्येक संख्या ठीक एक बार होती है)। वास्तव में, सुडोकू वर्ग समान क्रम के लैटिन वर्गों का एक छोटा अनुपात बनाते हैं। एक सुडोकू वर्ग अभिकल्पना का उपयोग चार व्याख्यात्मक चरों का अध्ययन करने के लिए किया जा सकता है। विभिन्नताओं के तीन स्रोत उपचार–उपचार, स्तंभ–स्तंभ और पंक्ति–पंक्ति विविधताएं हैं। भिन्नता के चौथे स्रोत का अध्ययन विशेष क्षेत्रों से आंतरिक ब्लॉकों (बक्सों) के निर्माण के माध्यम से किया जा सकता है जिन्हें आंतरिक ब्लॉक वर्गीकरण कहा जाता है। जब भी एक परीक्षणात्मक अभिकल्पना के रूप में एक लैटिन वर्ग उपयोग करने योग्य होता है, तो इसके बजाय परीक्षणात्मक अभिकल्पना के समान क्रम का एक सुडोकू वर्ग चुन सकते हैं। यह परीक्षणात्मक इकाइयों की संख्या में वृद्धि किए बिना चौथे व्याख्यात्मक चर (बॉक्स/आंतरिक ब्लॉक) को सम्मिलित करने की अनुमति देगा, और इस प्रकार एक अतिरिक्त व्यावहारिक लाभ प्रदान करेगा।

परीक्षण अभिकल्पना में, सुडोकू वर्ग विशेष प्रकार की

अभिकल्पना बनाता है जिन्हें सुडोकू अभिकल्पना के रूप में जाना जाता है। परीक्षणात्मक अभिकल्पना के रूप में सुडोकू वर्ग बेली इत्यादि (2008) द्वारा अध्ययन किया गया है। सुब्रमणि एवं पोन्नुस्वामी (2009) ने m^2 क्रम के सुडोकू अभिकल्पनाओं के एक वर्ग की संरचना पर चर्चा की। उन्होंने ऐसी अभिकल्पना से प्राप्त आंकड़ों के विश्लेषण के लिए चार अलग–अलग मॉडल प्रस्तुत किए। जू इत्यादि (2011) ने सुडोकू आधारित स्थान–पूर्ति अभिकल्पनाओं का अध्ययन किया। सबा एवं सिन्हा (2014) ने सुडोकू और पारस्परिक रूप से लांबिक सुडोकू वर्ग पर आधारित उन अभिकल्पनाओं के लिए अंतर्निहित प्रसरण विश्लेषण (ANOVA) तालिका के साथ एक विस्तृत सांख्यिकीय विश्लेषण दिया। कुमार एवं अन्य (2015) ने अनियमित क्लम्प्स के लिए तिराहा खण्ड (three way blocking) विन्यास के तहत अनियमित सुडोकू प्रकार की अभिकल्पना की संरचना की चार विधियाँ विकसित कीं जिनमें दो विधियाँ खाली नोड्स वाले क्लम्प्स के साथ अभिकल्पना देती हैं। कौर एवं गर्ग (2020) ने सुडोकू वर्ग के अवयवों के मध्य संबंध का अध्ययन कर एक संबंध योजना को परिभाषित किया और अपूर्ण सुडोकू वर्ग अभिकल्पनाओं की संरचना की ।

वास्तविक जीवन में, ऐसी स्थितियाँ आ सकती हैं जहाँ पूर्ण पंक्तियाँ या स्तंभ या दोनों एक सुडोकू वर्ग अभिकल्पना में वांछनीय नहीं हों, उदाहरण के लिए, किसी को भिन्नता के चार स्रोतों का अध्ययन करने के लिए एक अभिकल्पना की आवश्यकता हो सकती है, जिसमें मान लीजिए कि, छह पूर्ण स्तंभों, नौ अपूर्ण पंक्तियाँ और छह उप–आव्यूह की आवश्यकता है। ऐसे में एक अधूरा सुडोकू वर्ग अभिकल्पना बहुत काम का हो सकता है। कंप्यूटर परीक्षण सामान्यतः उत्पादों की अभिकल्पना एवं विकास के लिए उपयोग किए जाते हैं। कंप्यूटर परीक्षणों में, उत्पाद पर भौतिक रूप से परीक्षण करने के बजाय, उत्पाद के प्रदर्शन का वर्णन करने वाले गणितीय मॉडल इंजीनियरिंग/भौतिकी नियमों का उपयोग करके विकसित किए जाते हैं और कंप्यूटर पर परिमित तत्व विधि जैसे संख्यात्मक पद्धतियों के माध्यम से हल किए जाते हैं। क्योंकि नियतात्मक मॉडल परीक्षणों के लिए उपयोग किए जाते हैं, कंप्यूटर परीक्षण का आउटपुट यादृच्छिक विविधताओं के अधीन नहीं होता है, जो कंप्यूटर परीक्षणों की अभिकल्पना को भौतिक परीक्षणों से अलग बनाता है। उदाहरण के लिए, प्रतिकृति की आवश्यकता नहीं है। इस प्रकार, केवल इन कुछ महत्वपूर्ण कारकों का उपयोग करके एक अच्छा मॉडल फिट किया जा सकता है। यह सुडोकू आधारित स्थान–पूर्ति अभिकल्पना, जू इत्यादि, (2011) का उपयोग करके प्राप्त किया जा सकता है।

खंड 2, सुडोकू वर्ग अभिकल्पनाओं की कुछ परीक्षणात्मक स्थितियों पर चर्चा की गई है। खंड 3 में, अधूरी सुडोकू वर्ग अभिकल्पनाओं का उपयोग करके आंशिक संतुलित अपूर्ण खण्ड (PBIB) अभिकल्पना की एक श्रृंखला विकसित की गई है। खंड 4 सुडोकू आधारित स्थान–पूर्ति अभिकल्पनाओं के अनुप्रयोग से संबंधित है।

2. प्रारंभिक प्रबन्ध

2.1 कुछ परीक्षणात्मक स्थितियां

स्थिति 1. एक कृषि क्षेत्र के परीक्षण में जब किसी को बीज की विभिन्न किस्मों के प्रभावों का अध्ययन करना होता है, तो भिन्नता का एक स्रोत पानी की गुणवत्ता के विभिन्न स्तर हो सकते हैं, भिन्नता का दूसरा स्रोत विभिन्न प्रकार के उर्वरक हो सकते हैं और भिन्नता का तीसरा स्रोत मिट्टी की उर्वरता के विभिन्न स्तर हो सकते हैं। तीन तरह से अवरुद्ध संरचना के तहत परीक्षणात्मक अभिकल्पना में विषमता के इन तीन स्रोतों को एक साथ लिया जा सकता है।

स्थिति 2. जानवरों पर अलग–अलग चारों के प्रभावों का अध्ययन करने के लिए एक पशु परीक्षण में, परिवर्तनशीलता का एक स्रोत जानवरों की विभिन्न नस्लें हो सकती हैं, परिवर्तनशीलता का दूसरा स्रोत जानवरों के विभिन्न आयु समूह हो सकते हैं और परिवर्तनशीलता का तीसरा स्रोत जानवरों की अलग–अलग आवास

स्थितियां हो सकती हैं । विषमता के इन तीन स्रोतों को तीन–तरफा अवरोधक संरचना के तहत परीक्षणात्मक अभिकल्पना में एक साथ लिया जा सकता है ।

अपूर्ण सुडोकू वर्ग अभिकल्पनाओं का निर्माण चार व्याख्यात्मक चरों का अध्ययन करने में सक्षम है, अपूर्ण सुडोकू अभिकल्पना के एक घटक अभिकल्पना के रूप में निर्मित आंशिक संतुलित अपूर्ण खण्ड (PBIB) अभिकल्पना–उच्च साहचर्य वर्ग अंतर्निहित हैं। इस दिशा में, m–साहचर्य श्रेणी एसोसिएशन योजना पर चर्चा की गई है जो अपूर्ण सुडोकू वर्ग अभिकल्पनाओं पर आधारित है, कुछ नए निर्मित अपूर्ण सुडोकू वर्ग अभिकल्पना हैं, जो विषमता के दो–तरफा उन्मूलन के लिए अभिकल्पित होते हैं और सुडोकू वर्गों का उपयोग कर आंशिक संतुलित अपूर्ण खण्ड (PBIB) अभिकल्पनाओं की कुछ श्रृंखला भी प्राप्त की हैं।

2.2 सुडोकू वर्ग की मूल अवधारणा

यदि एक n पंक्ति × n स्तंभ वर्ग को समान रूप एवं आकार के साथ n बॉक्स (आंतरिक खण्डों) में उप–विभाजित किया जा सकता है और प्रत्येक अंक को प्रत्येक पंक्ति, स्तंभ और बॉक्स में एक बार और केवल एक बार प्रदर्शित किया जा सकता है, जब प्राकृतिक संख्या 1– n भरी जाती है, वर्ग को सुडोकू वर्ग कहा जाता है। परीक्षणात्मक अभिकल्पना के संदर्भ में, n^2 परीक्षणात्मक इकाइयों (भूखंडों) को n पंक्तियों, n स्तंभों और n बक्सों में विभाजित किया गया है और प्रत्येक में 1– n उपचार के साथ n परीक्षणात्मक अभिकल्पना हैं। इस प्रकार की अभिकल्पना को सुडोकू वर्ग अभिकल्पना कहा जाता है, जिसमें प्रत्येक उपचार में n प्रतिकृति होती है।

2.3 सुडोकू वर्ग की संरचना के लिए आवश्यक शर्त

$n \times n$ वर्ग एक ऐसा सुडोकू वर्ग बन जाता है, कि केवल वर्ग को n बक्सों में विभाजित किया जा सकता है, अर्थात, $n = pq$। हालाँकि, यदि p या q, 1 के बराबर है तो बॉक्स मौजूद नहीं होगा। इसलिए, सुडोकू वर्ग के निर्माण के लिए आवश्यक शर्त n - pq और p]q ≥ 2 है । प्राकृतिक संख्याओं में, अभाज्य संख्याएँ, जैसे 2, 3, 5, 7, 11, 13, को केवल कारक 1 और स्वयं द्वारा समान रूप से विभाजित किया जा सकता है, लेकिन गैर–अभाज्य संख्याओं को कम से कम 1 और स्वयं के अतिरिक्त ≥2 के किसी अन्य कारक द्वारा समान रूप से विभाजित किया जा सकता है। इस प्रकार, सभी गैर अभाज्य संख्या n सुडोकू वर्ग बनाती है और उनमें से कुछ एक से अधिक का गठन कर सकते हैं, और एक अभाज्य संख्या n नहीं हो सकती।

1	2	3	4	5	6	7	8	9
4	5	6	7	8	9	1	2	3
7	8	9	1	2	3	4	5	6
2	3	4	5	6	7	8	9	1
5	6	7	8	9	1	2	3	4
8	9	1	2	3	4	5	6	7
3	4	5	6	7	8	9	1	2
6	7	8	9	1	2	3	4	5
9	1	2	3	4	5	6	7	8

चित्र 1. पूर्ण सुडोकू ग्रिड

3. आंशिक संतुलित अपूर्ण खण्ड (PBIB) अभिकल्पनाओं का निर्माण

3.1. अपूर्ण सुडोकू वर्ग अभिकल्पना

अपूर्ण सुडोकू वर्ग एक $(n^2\text{-}l)\times(n^2\text{-}m)$ लैटिन आयत है जिसे $n^2\times n^2$ सुडोकू वर्ग से पहली या अंतिम l पंक्तियों को हटाकर प्राप्त किया जाता है । अपूर्ण सुडोकू वर्ग अभिकल्पना एक $(n^2 - l)\times(n^2 - m)$; $n = 3,4,\ldots$; $l = np$; $m = nk$ $(l\leq p, k\leq n - l)$ अपूर्ण सुडोकू वर्ग इस प्रकार से है कि यह हमें पंक्तियों, स्तंभों, लागू किए गए उपचारों और आंतरिक ब्लॉकों के रूप में चार व्याख्यात्मक चरों का अध्ययन करने की अनुमति देता है और परीक्षणात्मक अभिकल्पनाओं के अस्तित्व के लिए विभिन्न आवश्यक शर्तों को भी पूरा करता है (व्याख्या भाग 2 में)।

3.2. साहचर्य योजना

साहचर्य योजना को परिभाषित करने के लिए, पहली संख्या पर विचार करें और v उपचारों को एक वृत में, दक्षिणावर्त या वामावर्त दिशा में, उनके बढ़ते परिमाण के क्रम में व्यवस्थित करें। v उपचारों की इस वृत्ताकार व्यवस्था में, किसी भी उपचार के iवें – साहचर्य (i=1,2,…, m) α कहें, वे उपचार हैं जो इससे दक्षिणावर्त और वामावर्त दिशा में iवें स्थान पर स्थित हैं।

$m = \frac{v}{2}$; यदि v सम है और $v \geq 4$

$m = \frac{v-1}{2}$; यदि v सम है और $v \geq 5$

$n_1 = n_2 = n_{m-1} = 1$ (जब v सम है)

$n_1 = n_2 = n_m = 2$ (जब v विषम है)

3.3 संरचना पद्धति

3.3.1. पद्धति

किसी $v \times v$ सुडोकू वर्ग के लिए, जहाँ $v = R^2$ किसी भी $R \geq 2$ के लिए, यदि हम पहले स्तंभ $t = pR$ $(1 \leq p \leq R - 1)$, को छोड़ देंय परिणामी संरचना हमें एक अपूर्ण सुडोकू स्क्वायर वर्ग अभिकल्पना देती है जिसके स्तंभ पूर्ण हैं लेकिन पंक्तियाँ अपूर्ण हैं। अब, यदि इस अपूर्ण सुडोकू वर्ग की अधूरी पंक्तियों को खण्डों के रूप में लिया जाता है, तो परिणामी खण्ड का समुच्चय निम्नलिखित मापदंडों के साथ साहचर्य योजना 3.3 के आधार पर एक आंशिक संतुलित अपूर्ण खण्ड अभिकल्पना का गठन करेगाः $v = R^2 = b; r = R^2 - t = k, ki = r - i; i = 1; 2,\ldots,t.$ यदि $t < m$; तब $\lambda_{t+j} = \lambda t$ $(j = 1,2,\ldots, m-t)$

3.3.1.1 चित्रण

16×16 सुडोकू वर्ग के पहले 8 स्तंभों को छोड़ कर, एक 16×8 अपूर्ण सुडोकू वर्ग अभिकल्पना प्राप्त की जा सकती है जिसकी अपूर्ण पंक्तियाँ खण्डों के निम्नलिखित समुच्चय बनाती हैं:

खण्डों के उपरोक्त समुच्चय वाली आंशिक संतुलित अपूर्ण खण्ड अभिकल्पना के प्राचल इस प्रकार हैं:

B1	9	10	11	12	13	14	15	16
B2	13	14	15	16	1	2	3	4
B3	1	2	3	4	5	6	7	8
B4	5	6	7	8	9	10	11	12
B5	10	11	12	13	14	15	16	1
B6	14	15	16	1	2	3	4	5
B7	2	3	4	5	6	7	8	9
B8	6	7	8	9	10	11	12	13
B9	11	12	13	14	15	16	1	2
B10	15	16	1	2	3	4	5	6
B11	3	4	5	6	7	8	9	10
B12	7	8	9	10	11	12	13	14
B13	12	13	14	15	16	1	2	3
B14	16	1	2	3	4	5	6	7
B15	4	5	6	7	8	9	10	11
B16	8	9	10	11	12	13	14	15

$v = 16 = b, r = 8 = k, \lambda_1 = 7, \lambda_2 = 6, \lambda_3 = 5, \lambda_4 = 4, \lambda_5 = 3, \lambda_6 = 2, \lambda_7 = 1, \lambda_8 = 0$

4. सुडोकू आधारित स्थान-पूर्ति अभिकल्पना

इस खंड में, एक नयी विधि द्वारा सुडोकू की तीन प्रकार की एकरूपता का उपयोग करके एक दोगुना लांबिक सुडोकू लैटिन वर्गों का एक समुच्चय उत्पन्न करते हैं जिसमें पूर्ण वर्ग लांबिक होते हैं और उपयुक्त स्तर–समाप्त होने के बाद उप–वर्ग लांबिक होते हैं। वर्गों के इस समुच्चय का दूसरा उपयोग सुडोकू आधारित स्थान–पूर्ति अभिकल्पना बनाने के लिए है। इस तरह की अभिकल्पना में दो वांछनीय गुण हैं:

1) पूर्ण अभिकल्पना अविभाजित और द्विभाजित दोनों मार्जिन में अधिकतम एकरूपता प्राप्त करता है।
2) इसे किसी भी एक या दो आयामी प्रक्षेपण में से किसी के भी अधिकतम स्तरीकरण वाले उप–वर्ग के समूहों में विभाजित किया जा सकता है।

सुडोकू आधारित स्थान–पूर्ति अभिकल्पनाओं की संरचना के लिए पारस्परिक लांबिकता की तुलना में अधिक विस्तृत संरचना वाले सुडोकू लैटिन वर्गों की

आवश्यकता होती है । यहां दोगुने लांबिक सुडोकू लैटिन वर्गों का निर्माण इस प्रकार करना है कि पूरे वर्ग पारस्परिक रूप से लांबिक हों और एक ही स्थान पर उप–वर्ग उपयुक्त स्तर–समाप्त होने के बाद पारस्परिक रूप से लांबिक हों ।

4.1 सुडोकू आधारित स्थान-पूर्ति अभिकल्पनाओं की संरचना

सुडोकू आधारित स्थान–पूर्ति अभिकल्पना बनाने के लिए दोगुने लांबिक सुडोकू लैटिन वर्ग S के समुच्चय का उपयोग करने के लिए यह दो–चरणीय प्रक्रिया है। पहला चरण, S से एक विशेष कटा हुआ लांबिक आव्यूह B प्राप्त करता है। एक लांबिक सारणी $OA(n,d,s,t)$ एक $n \times d$ आव्यूह है जिसमें s प्रतीकों के एक समुच्चय से प्रविष्टियां होती हैं जैसे कि प्रत्येक $n \times t$ उप–आव्यूह के लिए, सभी स्तर संयोजन समान रूप से संयोजित होते हैं। एक कटा हुआ लांबिक सारणी एक ऐसा लांबिक सारणी है जिसे कई उप– सारणी में विभाजित किया जा सकता है, जिनमें से प्रत्येक कुछ स्तर–समाप्त के बाद एक छोटा लांबिक सारणी बन जाता है। दूसरा चरण कुछ विस्तृत योजना के माध्यम से सुडोकू आधारित स्थान–पूर्ति अभिकल्पना बनाने के लिए B का उपयोग करता है।

4.2 सुडोकू अभिकल्पनाओं के लक्षण वर्णन के अधययन हेतु R कोड

सुडोकू अभिकल्पनाओं के गुणों का अध्ययन करने हेतु R कोड लिखा गया है। यह लांबिक अभिकल्पना

```
> c_matrix
     [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    5   -1   -1   -1   -1   -1
[2,]   -1    5   -1   -1   -1   -1
[3,]   -1   -1    5   -1   -1   -1
[4,]   -1   -1   -1    5   -1   -1
[5,]   -1   -1   -1   -1    5   -1
[6,]   -1   -1   -1   -1   -1    5
> cannonical_efficiency
[1] 1
```

चित्र 2. आउटपुट

की तुलना में सुडोकू अभिकल्पना के उपचार प्रभाव और विहित दक्षता कारक से संबंधित सूचना आव्यूह देता ह। R कोड के स्नैपशॉट और आउटपुट यहाँ चित्र 1–2 में दिए गए हैं।

5. सारांश और निष्कर्ष

अपूर्ण सुडोकू वर्ग का उपयोग आंशिक रूप से संतुलित अपूर्ण खण्ड अभिकल्पनाओं की संरचना के लिए किया जा सकता है। सुडोकू स्थान–पूर्ति अभिकल्पनाओं को उप–अभिकल्पनाओं के समूहों में विभाजित किया जा सकता है और अविभाज्य एवं द्विचर दोनों मार्जिन में अधिकतम एकरूपता प्राप्त की जा सकती है। यह परीक्षणों की अभिकल्पना के सिद्धांत में सुडोकू वर्गों के महत्व को दर्शाता है और इस प्रकार के अध्ययन से सभी वर्तमान और भविष्य के शोधकर्ताओं को नए प्रकार की अभिकल्पनाओं की संरचना के लिए सुडोकू वर्गों के संयोजी गुणों एवं संरचना का अध्ययन करने की प्रेरणा मिलेगी। ऐसी अभिकल्पनाओं के गुणों की जांच के लिए R कोड का उपयोग किया जा सकता है।

```
a<-matrix(c(1,2,3,4,5,6,6,1,2,3,4,5,5,6,1,2,
3,4,4,5,6,1,2,3,3,4,5,6,1,2,2,3,4,5,6,1),nrow=6,byrow=T)
b<-matrix(c(1,1,1,1,1,3,1,2,2,2,2,3,2,2,
3,3,3,3,4,4,4,4,5,5,4,5,5,5,5,6,4,6,6,6,6,6),nrow=6,byrow=T)
N<-length(a)
v<-max(a)
x_mat1<-matrix(0,nrow=N,ncol=(v))
```

चित्र 1. R कोड

संदर्भ

बेली, आर.ए., कैमरून, पी.जे. एवं कोनेली, आर. (2008)। सुडोकू गेरेचटे डिजाइन्स, रेजोल्यूशन्स, एफाइन स्पेस, स्प्रेड्स, रेगुली एंड हैमिंग कोड्स। द अमेरिकन मेथेमैटिकल मंथली, 115(5), 383–404।

कौर, पी. एवं कुमार जी.डी. (2020)। कंस्ट्रक्शन ऑफ इंकम्प्लीट सुडोकू स्क्वायर एंड पार्शियली बैलेन्स्ड इंकम्प्लीट ब्लॉक डिजाइन। कम्युनिकेशंस इन स्टैटिस्टिक्स–थ्योरी एंड मेथड्स, 49(6), 1462–1474।

कुमार, ए., वर्गीज, सी., वर्गीज, ई. एवं जग्गी, एस. (2015)। ऑन द कंस्ट्रक्शन ऑफ डिजाइन्स विद थ्री–वे ब्लॉकिंग। मॉडल असिस्टेड स्टैटिस्टिक्स एंड ऐप्लीकेशंस, 10(1), 43–52।

सबा, एफ.ए. एवं सिन्हा, बी.के. (2014)। सुडोकू एज एन एक्सपेरिमेंटल डिजाइन – बियोंड द ट्रेडिशनल लैटिन स्क्वायर डिजाइन। स्टैटिस्टिक्स एंड ऐप्लीकेशंस, 12(1–2), 15–20 ।

सुब्रमणि, जे. एवं पोन्नुस्वामी, के.एन. (2009)। कंस्ट्रक्शन एंड एनालिसिस ऑफ सुडोकू डिजाइन्स। मॉडल असिस्टेड स्टैटिस्टिक्स एंड ऐप्लीकेशंस, 4(4), 287–301।

जू, एक्स., हैलैंड, बी.ई.एन. एवं कियान, पी.जेड. (2011)। सुडोकू–बेस्ड स्पेस–फिलिंग वाले डिजाइन्स। बायोमेट्रिका, 98(3), 711–720।

प्रोटीन संरचना प्रेडिक्शनः टूल्स और तकनीकें

सुधीर श्रीवास्तव, स्नेहा मुर्मू, दीपा भट्ट, मयंक रश्मि, कृष्ण कुमार चतुर्वेदी, मोहम्मद समीर फारूकी एवं गिरीश कुमार झा
भा.कृ.अनु.प. – भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली

सारांश

प्रोटीन संरचना प्रेडिक्शन कम्प्यूटेशनल जीव विज्ञान का एक महत्वपूर्ण क्षेत्र है, जो उनके अमीनो एसिड अनुक्रम के आधार पर प्रोटीन की 3डी संरचना की प्रेडिक्शन करता है। यह आलेख प्रोटीन संरचना प्रेडिक्शन के लिए उपयोग किए जाने वाले विभिन्न टूल्स और तकनीकों का एक अवलोकन प्रदान करता है, जिसमें होमोलॉजी–आधारित मॉडलिंग, फोल्ड रिकग्निशन और एब–इनिशियो विधियां शामिल हैं। प्रोटीन संरचना प्रेडिक्शन के विभिन्न तरीकों के प्रदर्शन का परीक्षण करने के लिए युग्मानुसार तुलनात्मक विश्लेषण भी किया गया है।

मुख्य शब्दः प्रोटीन, अमीनो एसिड, 3डी संरचना, होमोलॉजी मॉडलिंग, द्वितीयक संरचना एलिमेंट, डीप लर्निंग, मशीन लर्निंग एल्गोरिदम।

परिचय

कम्प्यूटेशनल जीवविज्ञान के क्षेत्र में संरचनात्मक जैव सूचना विज्ञान प्रमखु अनुसंधान क्षेत्रों में से एक है। संरचनात्मक जैव सूचना विज्ञान प्रोटीन, आर.एन.ए. और डी.एन.ए. जैसे जैविक अणु की 3–डी (त्रि–आयामी) संरचनाओं के विश्लेषण और प्रेडिक्शन को दर्शाता है। प्रोटीन संरचनाएं मुख्य रूप से एक्स–रे क्रिस्टलोग्राफी, इलेक्ट्रान माइक्रोस्कोपी या परमाणु चुंबकीय अनुनाद (एन.एम.आर.) स्पेक्ट्रोस्कोपी का उपयोग करके निर्धारित की जाती हैं। लेकिन ये विधियां समय लेने वाली, महंगी और सभी प्रोटीन के लिए संभव नहीं हैं। संरचनात्मक जैव सूचना विज्ञान में प्रोटीन 3–डी संरचनाओं का प्रेडिक्शन, मुख्य शोध समस्यायों में से एक है।

सार्वजनिक डेटाबेस में प्रोटीन एन्कोडिंग जीन के अनुक्रम और संरचनात्मक आँकड़े भारी मात्रा में जमा किये जा रहें हैं। डेटा आँकड़े में एन्कोड किए गए संरचनात्मक, इवोल्यूशनरी और कार्यात्मक संकेतों को समझना आणविक जीवविज्ञानी के लिए एक चुनौतीपूर्ण कार्य है। प्रोटीन की संरचनाओं और कार्यों को समझने के लिए विभिन्न विधियाँ और तकनीकें की आवश्यकता होती है, जिसमें अनुक्रम विश्लेषण, संरचना प्रेडिक्शन, कार्यात्मक जुड़ाव और इवोल्यूशनरी संबद्धता की पहचान शामिल हैं। प्रोटीन संरचना प्रेडिक्शन के लिए कई अलग–अलग दृष्टिकोण विकसित किए गए हैं। इनमें किसी अज्ञात के अनुक्रम से मेल खाने के लिए एक हल की गई संरचना को अपनाने से लेकर, पहले सिद्धांतों से एक प्रोटीन को मोड़ने (फोल्ड) के प्रयास तक का दायरा शामिल है। प्रोटीन संरचना प्रेडिक्शन के प्रत्येक तरीके के लिए सफलता दर भिन्न होती है। इस लेख में, हमने प्रोटीन संरचना प्रेडिक्शन के लिए कई तरीकों और टूल्स पर चर्चा की है।

प्रोटीन संरचना प्रेडिक्शन

प्रोटीन की 3–डी संरचना, उसके कार्यों और अन्य अणुओं के साथ इंटरेक्शन की पहचान करने में महत्वपूर्ण है। अमीनो एसिड अनुक्रम के आधार पर प्रोटीन संरचना की प्रेडिक्शन करना संरचनात्मक जीवविज्ञान का एक प्रमुख शोध क्षेत्र है, और पांच दशकों से अधिक समय से इसकी खोज की जा रही है। हालाँकि एक बड़ी चुनौती यह है कि सैद्धांतिक रूप से प्रोटीन के मुड़ने

(प्रोटीन फोल्डिंग) के तरीकों की संख्या खगोलीय है। शोधकर्ता उन्नत कंप्यूटर सिमुलेशन का उपयोग करके पूर्वानुमानित फोल्डिंग पथों (पाथवे) को मॉडल करने का प्रयास कर रहे हैं जो प्रकृति प्रोटीन की मूल स्थिति तक पहुंचने के लिए उपयोग में आती है तथा फोल्डिंग प्रक्रिया के अंतर्निहित भौतिक इंटरैक्शन का अनुमान लगाती है। संरचनात्मक जैव सूचना विज्ञान समुदाय इस समस्या को सबसे बड़ी समस्या मानता है। इस अनुभाग में, अत्याधुनिक तकनीकों का उपयोग करके प्रोटीन तृतीयक संरचना की प्रेडिक्शन करने के तरीकों पर चर्चा की गई है।

जीवित प्रणाली में होने वाले जैविक कार्यों, मार्गों तथा कई आणविक तंत्रों को समझने के लिए प्रोटीन की त्रिआयामी संरचना का निर्धारण महत्वपूर्ण है। आमतौर पर, प्रोटीन की विश्वसनीय 3डी संरचना एक्स–रे एवं एन.एम.आर. तकनीक जैसी प्रायोगिक तकनीकों द्वारा निर्धारित की जाती है। ये तकनीकें समय लेने वाली और अत्यधिक महंगी हैं। इसके अलावा, कई प्रोटीन इन तरीकों से निर्धारित नहीं किए जा सकते हैं और सफलता दर 20% से भी कम है। सार्वजनिक डोमेन में अनुक्रम और 3डी संरचना के जानकारी की उपलब्धता के बीच एक बड़ा अंतर है। इसलिए, 3–डी संरचनाओं की प्रेडिक्शन करने के लिए कई इन–सिलिको विधियां विकसित की गयी हैं। निम्नलिखित तीन विधियों का उपयोग करके प्रोटीन की तृतीयक संरचना की प्रेडिक्शन या इन–सिलिको मॉडलिंग की जा सकती हैः

होमोलॉजी मॉडलिंगः होमोलॉजी–आधारित मॉडलिंग, जिसे तुलनात्मक मॉडलिंग के रूप में भी जाना जाता है। यह एक समजात प्रोटीन की ज्ञात संरचना के आधार पर प्रोटीन की त्रिआयामी संरचना के प्रेडिक्शन की एक कम्प्यूटेशनल विधि है। विधि मानती है कि विकास के दौरान प्रोटीन संरचना के संरक्षण के कारण समान अमीनो एसिड अनुक्रम वाले प्रोटीन की संरचना समान होती है। होमोलॉजी मॉडलिंग में सात चरण शामिल हैं, जिसमें टेम्पलेट चयन और संरेखण, बैकबोन और साइड–चेन मॉडलिंग, तथा मॉडल अनुकूलन और सत्यापन शामिल हैं। अनुमानित संरचना की सटीकता संरेखण की गुणवत्ता, टारगेट और टेम्पलेट प्रोटीन के बीच समानता पर निर्भर करती है। सीमाओं के बावजूद, प्रोटीन संरचनाओं की प्रेडिक्शन करने के लिए होमोलॉजी–आधारित मॉडलिंग एक शक्तिशाली उपकरण है, खासकर जब प्रयोगात्मक तरीके संभव नहीं हैं। कई प्रोटीन संरचनाओं की प्रेडिक्शन करने के लिए इसका सफलतापूर्वक उपयोग किया गया है।

फोल्ड रिकग्निशन या थ्रेडिंगः फोल्ड रिकग्निशन या थ्रेडिंग एक कम्प्यूटेशनल विधि है जो अमीनो एसिड अनुक्रम के आधार पर प्रोटीन की त्रि–आयामी संरचना का प्रेडिक्शन करता है, जिसके लिए ज्ञात संरचना के साथ एक समरूप प्रोटीन की आवश्यकता नहीं होती है। इसमें ज्ञात प्रोटीन संरचनाओं की लाइब्रेरी में समान संरचनात्मक रूपांकनों की पहचान करना और टेम्पलेट प्रोटीन संरचना की रीढ़ पर टारगेट प्रोटीन के अमीनो एसिड अनुक्रम को फैलाना शामिल है। टारगेट प्रोटीन के लिए एक संरचनात्मक मॉडल उत्पन्न करने के लिए थ्रेडेड मॉडल को अनुकूलित और परिष्कृत किया जाता है। ज्ञात संरचनाओं के लिए कम अनुक्रम समानता वाले प्रोटीन की संरचना की प्रेडिक्शन करने के लिए फोल्ड पहचान विधियां उपयोगी हैं, लेकिन सटीकता संरचनात्मक लाइब्रेरी और अनुकूलन एल्गोरिदम की गुणवत्ता पर निर्भर करती है।

एब–इनिशियो विधिः एब–इनिशियो विधि प्रयोगात्मक आँकड़ों या ज्ञात संरचनाओं से किसी भी पूर्व जानकारी का उपयोग किए बिना अपने अमीनो एसिड अनुक्रम से प्रोटीन की 3–डी संरचना की प्रेडिक्शन करने के लिए एक कम्प्यूटेशनल विधि है। इसमें विभिन्न शब्दों से ऊर्जा योगदान का वर्णन करने के लिए फोर्स फील्ड का उपयोग करके भौतिकी और रसायन विज्ञान सिद्धांतों के आधार पर प्रोटीन संरचना की संभावित ऊर्जा की गणना करना शामिल है। विधि में न्यूनतम ऊर्जा संरचना की पहचान करने के लिए खंड संयोजन, संरचना अनुकूलन

और ऊर्जा न्यूनीकरण शामिल है। हालांकि, समजातीय संरचनाओं या छोटे प्रोटीन एवं पेप्टाइड्स के बिना प्रोटीन संरचना प्रेडिक्शन करना चुनौतीपूर्ण है। 3–डी संरचना प्रेडिक्शन के लिए उपयोग किए जाने वाले कुछ टूल्स तालिका 1 में सूचीबद्ध हैं।

प्रोटीन संरचना प्रेडिक्शन के विधियों का तुलनात्मक विश्लेषण

ई.एस.एस. परिवार ग्लूटामेटः Na+स्यूडोक्सैन्थोमोनस 3HH-4 सहवाहक का फास्टा अनुक्रम, यूनिप्रोट आईडी A0A543DRZ4 dks https://www.uniprot.org/

तालिका 1: प्रोटीन तृतीयक संरचना प्रेडिक्शन के लिए उपयोग किए जाने वाले टूल्स की सूची

नाम	विवरण	लिंक	संदर्भ
I. होमोलॉजी मॉडलिंग			
पी.एच.वाई.आ.रई.2	हिडन मार्कोव मॉडल (एच.एम.ए.म) का उपयोग करके संरेखण उत्पन्न करता है	http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index	केली एट अल. 2015
सी.पी.एच. मॉडल	अमीनो एसिड के Cα परमाणुओं के बीच संपर्क की प्रेडिक्शन करने के लिए तंत्रिका नेटवर्क का उपयोग करता है	https://services.healthtech.dtu.dk/service.php?CPHmodels-3.2	अमीनो एसिड के Cα परमाणुओं के बीच संपर्क की प्रेडिक्शन करने के लिए तंत्रिका नेटवर्क का उपयोग करता है
स्विस मॉडल	एक स्वचालित सर्वर	https://swissmodel.expasy.org/	किफर एट अल. 2009
मॉडेलर	स्टैंडअलोनय स्थानिक बाधाओं को संतुष्ट करता है	https://modbase.compbio.ucsf.edu/modweb/	वेब और साली 2014
रैप्टर.एक्स	दूरस्थ समरूपता का पता लगाने के लिए उपयोग किया जाता है	http://raptorx.uchicago.edu/	पेंग और क्सयू, 2011
II. फोल्ड रिकग्निशन			
आई–टस्सर (इटरेटिव थ्रेडिंग असेंबली रिफाइनमेंट)	एकाधिक थ्रेडिंग और टेम्प्लेट–आधारित अंशों को पुनरावृत्त रूप से असेंबल करके संरचना की प्रेडिक्शन करता है	https://zhanggroup.org/I-TASSER/	रॉय एट अल. 2010
मस्टर (मल्टी–सोर्स थ्रेड ई.आर.)	अनुक्रम प्रोफाइल–प्रोफाइल संरेखण और गतिशील प्रोग्रामिंग पर आधारित	https://zhanggroup.org/MUSTER/	वू और झांग 2008
एच.एच.प्रेड.	दूरस्थ होमोलॉजी का पता लगाने के लिए एच.एम.एम. का उपयोग करता ह		सोडिंग एट अल. 2005
III. एब–इनिशियो			
रोबेटा	खंड आधारित असेंबली	http://new.robetta.org/	किम एट अल. 2004
क्वार्क	प्रतिकृति–विनिमय मोंटे कार्लो सिमुलेशन पर आधारित	https://zhanggroup.org/QUARK/	जू और झांग एट अल. 2012

नाम	विवरण	लिंक	संदर्भ
अल्फाफोल्ड	गहन शिक्षण एल्गोरिदम पर आधारित जो एब–इनिशियो मॉडलिंग के सिद्धांतों पर बहुत अधिक निर्भर करता है	https://github.com/deepmind/alphafold	जम्पर एट अल. 2021

uniprotkb/A0A543DRZ4/entry से पुनर्प्राप्त किया गया था। पी.डी.बी. में समान अनुक्रमों की उपलब्धता की जांच करने के लिए ब्लास्टपी टूल का उपयोग किया गया है। ई–वैल्यू 0.01, गैप के साथ ब्लोसम 62 मैट्रिक्स पर पी.डी.बी. डेटाबेस के विरुद्ध कोई हिट नहीं पाया गया (जॉनसन एट अल, 2008य डब्ल्यूडब्ल्यू पीडीबी, कंसोर्टियम, 2019)। इस अध्ययन में विभिन्न प्रोटीन 3डी संरचना मॉडलिंग विधियों की तुलना की गई (तालिका 2)। जिसमें अल्फाफोल्ड विधि को संदर्भ के रूप में चुना गया है तथा अन्य विधियों की तुलना अल्फाफोल्ड से की गई है।

अल्फाफोल्ड द्वारा तैयार की गई संरचना को अन्य मॉडलिंग विधियों की तुलना करने के लिए संदर्भ के रूप में लिया गया है (चित्र 1)। रोसेटाफोल्ड और आई–टस्सर 396 अमीनो एसिड के साथ पूर्ण 3डी संरचनाएँ बनाने में सक्षम हैं। आर.एम.एस.डी. वैल्यू, टी.एम.–स्कोर और समतुल्य रेसिड्यूस के आधार पर, रोसेटाफोल्ड अन्य विधियों की तुलना में बेहतर प्रदर्शन कर रही है।

निष्कर्ष

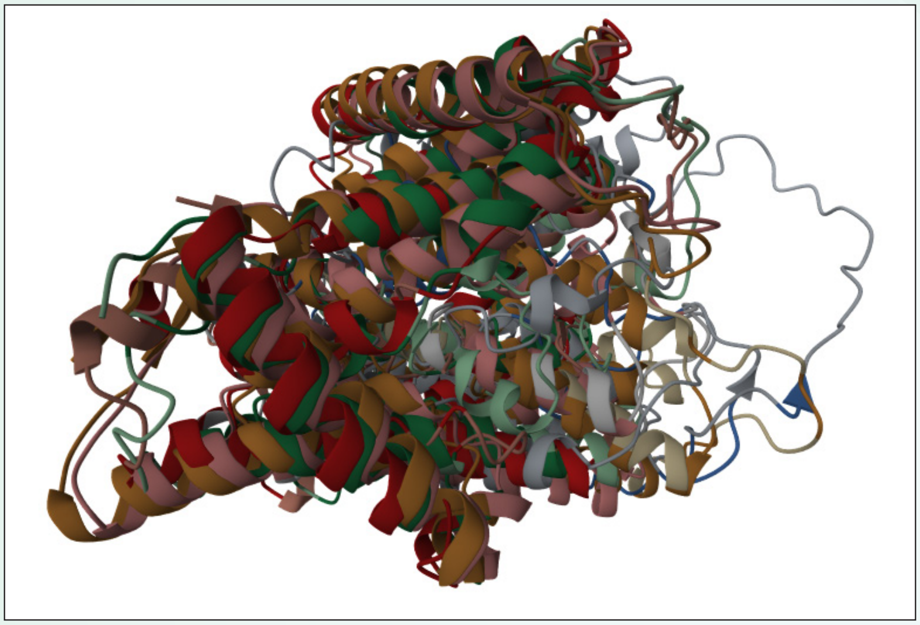
प्रोटीन संरचना की टूल्स एक चुनौतीपूर्ण लेकिन महत्वपूर्ण कार्य है, जिसका दवाओं की खोज, जैव प्रौद्योगिकी और जैव रसायन जैसे क्षेत्रों में कई अनुप्रयोग हैं। प्रोटीन संरचना प्रेडिक्शन के लिए उपयोग की जाने वाली तीन मुख्य कम्प्यूटेशनल विधियों में होमोलॉजी–आधारित मॉडलिंग, फोल्ड रिकग्निशन या थ्रेडिंग, और एब–इनिशियो विधि शामिल हैं। प्रत्येक विधि की अपनी ताकत और कमजोरियां होती हैं तथा विधि का

तालिका 2: प्रोटीन 3–डी संरचना मॉडलिंग की विधियां

क्र.सं.	सर्वर का नाम	विधियाँ	संदर्भ
1	अल्फाफोल्ड	हाइब्रिड	जम्पर एट अल, 2021
2	ई.वी. कपलिंग	हाइब्रिड	शेरिडन एट अल, 2015
3	रोसेटाफोल्ड	एब–इनिशियो	सिमंस एट अल। 1997; दास एंड बेकर, 2008; कॉफमैन एट अल., 2010
4	स्विस मॉडल	होमोलॉजी	श्वेडे एट अल, 2003; कोप्प और श्वेडे, 2004
5	आई–टस्सर	थ्रेडिंग	रॉय, 2010; यांग और झांग, 2015

तालिका 3: प्रतिरूपित प्रोटीन संरचनाओं के संरेखण का सारांश

संदर्भ टूल का नाम	तुलनात्मक टूल का नाम	आर. एम. एस. डी.	टी.एम.–स्कोर	पहचान (%)	समतुल्य रेसिड्यू	अनुक्रम की लंबाइ	मॉडल्ड रेसिड्यू
अल्फाफोल्ड	ई.वी. कपलिंग	7.08	0.28	61	75	389	389
	आई–टस्सर	4.84	0.65	56	243	396	396
	रोसेटाफोल्ड	2.98	0.81	54	335	396	396
	स्विस मॉडल	4.47	0.57	48	227	351	351



चित्र 1. अल्फाफोल्ड (पीला), ई.वी. कपलिंग (स्लेटी), आई–टस्सर (हरा), रोबेटाफोल्ड (गुलाबी) और स्विस मॉडल (लाल) द्वारा प्रतिरूपित संरचनाओं का संरेखण

चुनाव प्रयोगात्मक आँकड़ों की उपलब्धता, टारगेट और ज्ञात प्रोटीन के बीच समानता की डिग्री और उपलब्ध कम्प्यूटेशनल संसाधनों पर निर्भर करता है। प्रोटीन संरचना प्रेडिक्शन का भविष्य, मशीन लर्निंग एल्गोरिदम के विकास और कम्प्यूटेशनल तथा प्रायोगिक तरीकों के एकीकरण पर बहुत अधिक निर्भर होने की उम्मीद है। इन दृष्टिकोणों के संयोजन से अधिक सटीक और कुशल प्रिडिक्शन्स हो सकता हैं, जिसका दवाओं की खोज और अन्य जैव प्रौद्योगिकी अनुप्रयोगों पर महत्वपूर्ण प्रभाव पड़ेगा। जैसे–जैसे क्षेत्र का विकास जारी रहेगा, नए टूल्स और तकनीकें सामने आएंगी, जो शोधकर्ताओं को प्रोटीन संरचना और कार्य की जटिलता को दुनिया में और भी अधिक अंतर्दृष्टि प्रदान करेंगी।

सन्दर्भ

चाउ, पी.वाई., एवं फास्मान, जी.डी (1974)। प्रोटीन संरचना की प्रेडिक्शन बायोकेमिस्ट्री। 13(2), 222–245।

गार्नियर, जे., एवं रॉबसन, बी. (1989)। प्रोटीन में द्वितीयक संरचनाओं की प्रेडिक्शन के लिए गौर विधि। प्रेडिक्शन ऑफ प्रोटीन स्ट्रक्चर एंड दी प्रिंसिपल्स ऑफ प्रोटीन कन्फोर्मेशन, 417–465।

केली, एल.ए., मेजुलिस, एस., येट्स, सी.एम., वास, एम. एन., एवं स्टर्नबर्ग, एम .जे (2015)। प्रोटीन मॉडलिंग, प्रेडिक्शन और विश्लेषण के लिए पीएचवाईआरई2 वेब पोर्टल। नेचर प्रोटोकॉल, 10(6), 845–858।

नील्सन, एम., लुंडेगार्ड, सी., लुंड, ओ., एवं पीटरसन, टी.एन. (2010)। सी पी एच मॉडल–3.0–संरचना–निर्देशित अनुक्रम प्रोफाइल का उपयोग करके दूरस्थ होमोलॉजी मॉडलिंग। न्यूक्लिक एसिड रिसर्च, 38, डब्ल्यू576–डब्ल्यू581।

किफर, एफ., अर्नोल्ड, के., कुंजली, एम., बोर्डोली, एल., एवं श्वेडे, टी. (2009)। स्विस–मॉडल रिपॉजिटरी और संबंधित संसाधन. न्यूक्लिक एसिड अनुसंधान, 37, डी387–डी392।

वेब, बी., एवं साली, ए. (2016)। मॉडेलर का उपयोग करके तुलनात्मक प्रोटीन संरचना मॉडलिंग। जैव सूचना विज्ञान में वर्तमान प्रोटोकॉल, 54(1), 5–6।

पेंग, जे., एवं जू, जे. (2011)। रैप्टर एक्सः सांख्यिकीय अनुमान द्वारा प्रोटीन संरेखण के लिए संरचना जानकारी का दोहन। प्रोटीन संरचनाः कार्य और

जैव सूचना विज्ञान, 79(एस10), 161–171।

रॉय, ए., कुकुकुरल, ए., एवं झांग, वाई. (2010)। आई– टस्सरः स्वचालित प्रोटीन संरचना और कार्य भविष्यवाणी के लिए एक एकीकृत मंच। प्रकृति प्रोटोकॉल, 5(4), 725–738।

वू, एस., एवं झांग, वाई (2008)। मस्टरः संरचना जानकारी के कई स्रोतों का उपयोग करके प्रोटीन अनुक्रम प्रोफाइल–प्रोफाइल संरेखण में सुधार करना। प्रोटीनः संरचना कार्य और जैव सूचना विज्ञान, 72(2), 547–556।

सोडिंग, जे., बीगर्ट, ए., एवं लुपस, ए.एन. (2005)। प्रोटीन होमोलॉजी का पता लगाने और संरचना की प्रेडिक्शन के लिए एच एचप्रेड इंटरैक्टिव सर्वर। न्यूक्लिक एसिड अनुसंधान, 33, डब्ल्यू244– डब्ल्यू248।

किम, डी.ई., चिवियन, डी., एवं बेकर, डी. (2004)। रोबेटा सर्वर का उपयोग करके प्रोटीन संरचना की प्रेडिक्शन और विश्लेषण। न्यूक्लिक एसिड अनुसंधान, 32, डब्ल्यू526– डब्ल्यू531।

जू, डी., एवं झांग, वाई. (2012)। निरंतर संरचना के टुकड़ों और अनुकूलित ज्ञान आधारित बल क्षेत्र का उपयोग करके एब इनिटियो प्रोटीन संरचना असेंबली. प्रोटीन। संरचना, कार्य और जैव सूचना विज्ञान, 80(7), 1715–1735।

जम्पर, जे., इवांस, आर., प्रित्जेल, ए., ग्रीन, टी., फिगरनोव, एम., रोनेबर्गर, ओ., एवं हस्साबिस. (2021)। अल्फाफोल्ड के साथ अत्यधिक सटीक प्रोटीन संरचना प्रेडिक्शन। प्रकृति, 596(7873), 583–589।

जॉनसन एम., जेरेत्सकाया आई., रेत्सेलिस वाई., मेरेजुक वाई., मैकगिनिस एस., मैडेन टीएल. (2008)। एनसीबीआई ब्लास्टः एक बेहतर वेब इंटरफेस। न्यूक्लिक एसिड रिसर्च, 36: वेब सर्वर समस्या W5–W9. doi:10-1093/nar/gkn201।

शेरिडन आर., फील्डहाउस आर. जे., हयात एस., सन वाई., एंटीपिन वाई., यांग एल., हॉफ टी., मार्क्स डी. एस., सैंडर सी. (2015)। ईवी फोल्ड. संगठनः विकासवादी युग्मन और प्रोटीन 3डी संरचना प्रेडिक्शन। बायोरेक्सिव प्रीप्रिंट. डीओआईः https://doi.org/10.1101/021022।

सिमंस के. टी., कूपरबर्ग सी., हुआंग ई., बेकर डी. (1997)। सिम्युलेटेड एनीलिंग और बायेसियन स्कोरिंग फंक्शंस का उपयोग करके समान स्थानीय अनुक्रमों के साथ टुकड़ों से प्रोटीन तृतीयक संरचनाओं का संयोजन। जर्नल ऑफ मॉलिक्यूलर बायोलॉजी, 268: 209–225. डीओआईः 10.1006/जेएमबीआई.1997.0959।

दास आर., बेकर डी. (2007)। देशी–जैसी आरएनए तृतीयक संरचनाओं की स्वचालित डे नोवो प्रेडिक्शन। राष्ट्रीय विज्ञान अकादमी की कार्यवाही, 104:14664–14669. डीओआईः 10.1073/पीएनएएस. 0703836104।

कॉफमैन केडब्ल्यू., लेमन जी. एच., डीलुका एस. एल., शीहान जे. एच., मीलर जे. (2010)। व्यावहारिक रूप से उपयोगीः रोसेटा प्रोटीन मॉडलिंग सूट आपके लिए क्या कर सकता ह। जैव रसायन, 49: 2987–2998. कवपः 10-1021/bi902153g।

श्वेडे टी., कोप्प जे., गुएक्स एन., पीट्स्च एम. सी. (2003)। स्विस–मॉडलः एक स्वचालित प्रोटीन होमोलॉजी–मॉडलिंग सर्वर। न्यूक्लिक एसिड अनुसंधान, 31: 3381–3385. डीओआईः 10.1093/एनएआर/जीकेजी520।

कोप्प जे., श्वेडे टी. (2004)। एनोटेटेड त्रि–आयामी प्रोटीन संरचना होमोलॉजी मॉडल का स्विस–मॉडल रिपोजिटरी। न्यूक्लिक एसिड अनुसंधान, 32: डी230–डी234. कवपः 10.1093/nar/gkh008।

कल्बर्ग एम., वांग एच., वांग एस., पेंग जे., वांग जेड., लू एच., जू जे. (2012)। रैप्टरएक्स वेब सर्वर का उपयोग करके टेम्पलेट–आधारित प्रोटीन संरचना मॉडलिंग। प्रकृति प्रोटोकॉल, 7:1511–1522. डीओआईः 10.1038/एनप्रोट.2012.085।

यांग जे., झांग वाई. (2015)। आई– टस्सर सर्वरः

प्रोटीन संरचना और कार्य प्रेडिक्शन के लिए नया विकास। न्यूक्लिक एसिड अनुसंधान, 43(डब्ल्यू1): डब्ल्यू174–डब्ल्यू181. डीओआईः 10.1093/एनएआर/जीकेवी342।

चेंग जे., रान्डेल एज., स्वेरेडोस्की एम. जे., बाल्दी पी. (2005)। स्क्रैचः एक प्रोटीन संरचना और संरचनात्मक सुविधा प्रेडिक्शन सर्वर. न्यूक्लिक एसिड अनुसंधान, 33: 72–76।

wwPDB कंसोर्टियम. (2019)। प्रोटीन डेटा बैंकः 3डी मैक्रोमोलेक्यूलर संरचना आँकड़ों के लिए एकल वैश्विक संग्रह। न्यूक्लिक एसिड रेस. 47 (डी1) 520–528।

बेसियन फ्रेमवर्क का उपयोग करके सिस और ट्रांस-एक्सप्रेशन क्वांटिटेटिव ट्रेट लोसाइ (ईक्यूटीएल) की पहचान

हिमाद्री शेखर रॉय, अमृत कुमार पॉल, रंजित कुमार पॉल, प्रकाश कुमार, एमडी यासीन, एवं एस.पी.सिंह

भा.कृ.अनु.प. – भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली

सार

जीन एक्सप्रेशन जैसी जैविक विशेषताओं के लिए एक्सप्रेशन क्वांटिटेटिव ट्रेट लोसाइ (ईक्यूटीएल) की पहचान जीनोमिक्स में महत्वपूर्ण है। सिस और ट्रांस–ईक्यूटीएल का अस्तित्व वांछित गुणों के लिए ईक्यूटीएल के संचयी महत्व को स्थापित करने के लिए महत्वपूर्ण है। जीनोमिक्स में एक महत्वपूर्ण विचार सिस और ट्रांस ईक्यूटीएल की पहचान करना है जो दूर के जीन की अभिव्यक्ति में पर्याप्त परिवर्तन को पकड़ते हैं। इस अध्ययन का उद्देश्य उन सिस और ट्रांस–क्यूटीएलएस को खोजने के लिए अभिव्यक्ति मात्रात्मक लोसाइ के लिए एक एकीकृत पदानुक्रमित बेसियन मॉडल का उपयोग करना है। क्यूटीएल या ईक्यूटीएल की खोज होने पर केवल उम्मीदवार जीन को वर्गीकृत करने के लिए आणविक दृष्टिकोण का उपयोग किया जाता है। जीन के अंदर या उसके निकट भिन्नताएं अनुवांशिक भिन्नताओं को निर्धारित करने के लिए परिकल्पना की जाती हैं जो प्रतिलेख स्तर को दर्शाती हैं। ईक्यूटीएल की खोज ने हमें जीन नियमन और जटिल लक्षण विश्लेषण को बेहतर ढंग से समझने में मदद की है। वर्तमान अध्ययन जौ की फसल पर केंद्रित है, और जिसमें केवल सिस – ईक्यूटीएल की खोज की गई है, जहाँ कोई अतिरिक्त ईक्यूटीएल हॉटस्पॉट नहीं खोजा गया था। चूहा का जीन अभिव्यक्तियों का उपयोग ट्रांस ईक्यूटीएल की जांच के लिए किया गया है, और पर्याप्त सिस और ट्रांस ईक्यूटीएल के साथ–साथ चार ईक्यूटीएल हॉटस्पॉट की खोज की गई है।

मुख्य शब्दः बेसियन मॉडल सिस – ईक्यूटीएल, ईक्यूटीएल, हॉटस्पॉट, मात्रात्मक गुण वाले लोसाई, ट्रांस–ईक्यूटीएल।

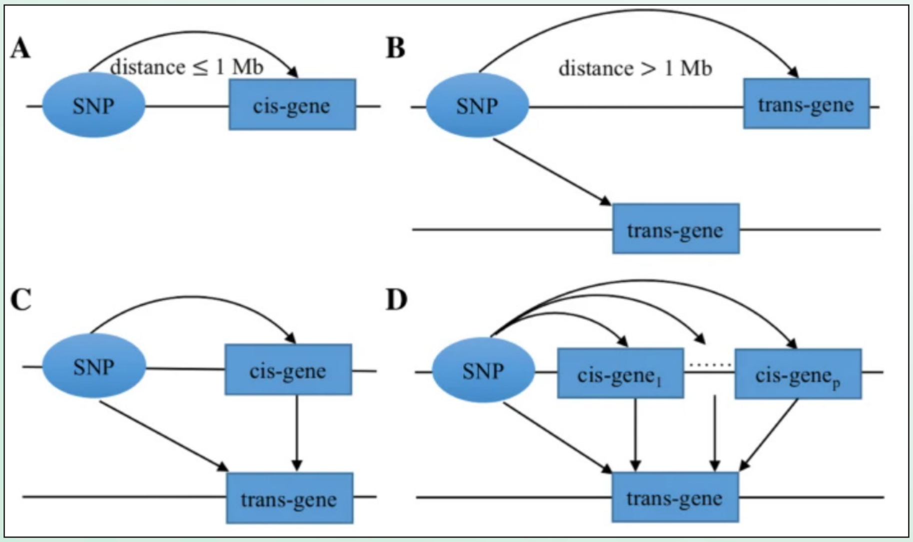
1 परिचय

एक मात्रात्मक विशेषता को एक जैविक अनुसंधान में मात्रात्मक फेनोटाइप के रूप में वर्णित किया गया है, जैसे उपज, रोग प्रतिरोध, और इसी तरह, जो संख्यात्मक रूप से भिन्न होता है और कई जीनों और पर्यावरण की सामूहिक गतिविधियों पर निर्भर होता है। क्वांटिटेटिव कैरेक्टरिस्टिक लोसाइ (क्यूटीएल) मैपिंग (ड्रूका एट अल., 2010) में प्रारंभिक चरण एक निश्चित मात्रात्मक विशेषता से जुड़े क्रोमोसोमल क्षेत्रों का पता लगाने के लिए संभावित जीन की पहचान करना है। हाल ही में, एक्सप्रेशन क्वांटिटेटिव ट्रेट लोसाइ (ईक्यूटीएल) (गेलफोंड एट अल., 2007) मैपिंग ट्रांसक्रिप्टोम और जीनोटाइपिंग डेटा का उपयोग करके जीन के लिए नियामक क्षेत्रों का पता लगाने के लिए एक लोसाइप्रिय और सफल तरीका बन गया है। ये जीनोमिक क्षेत्र हैं जो एमआरएनए या प्रोटीन अभिव्यक्ति को नियंत्रित करते

हैं। सबसे महत्वपूर्ण आणविक फेनोटाइप जीन एक्सप्रेशन हैं, जो मात्रात्मक गुणों के रूप में काम करते हैं जो आनुवंशिक विविधता को फेनोटाइपिक विचरण से जोड़ते हैं। जीन अभिव्यक्ति में ये अंतर फेनोटाइपिक अंतर और प्रजातियों के विकास में महत्वपूर्ण होने की उम्मीद है। परिणामस्वरूप, ईक्यूटीएल की पहचान करना जैविक अनुसंधान का एक महत्वपूर्ण तत्व बन गया है (पोटोकिना एट अल., 2008)। पारंपरिक क्यूटीएल मैपिंग में अक्सर व्यापक क्षेत्र मिलते हैं। क्यूटीएल या ईक्यूटीएल की खोज के बाद, केवल उम्मीदवार जीन को वर्गीकृत करने के लिए आणविक विधियों का उपयोग किया जाता है। माना जाता है कि अनुवांशिक अंतर जो प्रतिलेख स्तर को दर्शाते हैं, जीन के अंदर या उसके करीब भिन्नताओं द्वारा निर्धारित किए जाते हैं। ईक्यूटीएल की खोज ने जीन विनियमन और जटिल विशेषता विश्लेषण (ब्रेम एट अल., 2002, गोरिंग एट अल., 2007) की हमारी समझ का विस्तार किया है। डीएनए माइक्रोएरे का उपयोग करके जीनोम में लगभग सभी जीनों के लिए ईक्यूटीएल को मैप करना अब संभव है। अभिव्यक्ति क्यूटीएल को निष्पक्ष रूप से दो प्रकारों में वर्गीकृत किया जाता है: सीस और ट्रांस। सीस – ईक्यूटीएल के मामले में, ईक्यूटीएल की प्रेक्षित स्थिति जीन की साइट (चित्र 1(a)) के समान है, लेकिन ट्रांस– ईक्यूटीएल के मामले में, ईक्यूटीएल का प्रेक्षित स्थान जीन के स्थान के समान (चित्र 1 (b)) नहीं है। एक उदाहरण के रूप में, एक ट्रांस–ईक्यूटीएल एक ट्रांसक्रिप्शन कारक की स्थिति को अकेले प्रतिबिंबित कर सकता है जो लोसाइ की अभिव्यक्ति को नियंत्रित करता है।

ईक्यूटीएल (फ्लुट्रे एट अल., 2010) की खोज हमें कई प्रजातियों में आनुवंशिकी और जीन अभिव्यक्ति नियंत्रण प्रणाली को बेहतर ढंग से समझने में मदद करती है। एसएनपी–जीन संयोजनों की भारी संख्या के कारण ट्रांस–ईक्यूटीएल पहचान सिस–ईक्यूटीएल पहचान की तुलना में काफी अधिक चुनौतीपूर्ण है, जिसकी ट्रांस–एसोसिएशन मैपिंग के लिए जांच की जानी चाहिए। ट्रांस–ईक्यूटीएल विश्लेषण को समान शक्ति प्राप्त करने के लिए सिस–ईक्यूटीएल विश्लेषण की तुलना में अधिक नमूना की आवश्यकता होती है। हालाँकि, ट्रांस–ईक्यूटीएल, सिस–ईक्यूटीएल की तुलना में कम प्रभाव दिखाते हैं (पियर्स एट अल., 2014)। ट्रांस–क्यूटीएल की वास्तविक संख्या निर्धारित करने के लिए हाल ही में कई दृष्टिकोण विकसित किए गए हैं।

कई शोधकर्ताओं ने तथाकथित ईक्यूटीएल हॉटस्पॉट में ट्रांस–ईक्यूटीएल (एक ही स्थान से जुड़े कई जीन)



चित्र 1. ईक्यूटीएल का चित्रमय चित्रण दिखाता है। a) सिस–ईक्यूटीएल, b) ट्रांस–क्यूटीएल, c) ट्रांस–क्यूटीएल एक एकल सिस – मध्यस्थ द्वारा मध्यस्थता, और d) ट्रांस–क्यूटीएल कई सिस –मध्यस्थों द्वारा विनियमित (शान एट अल., 2019)।

की पर्याप्त क्लस्टरिंग पाई है, यह दर्शाता है कि इन जीनोमिक क्षेत्रों में वेरिएंट होते हैं जो ट्रांसक्रिप्शनल विनियमन के गतिशील और वैश्विक चरित्र (झू एट अल., 2016) को बदलते हैं। ईक्यूटीएल मैपिंग अध्ययनों के अनुसार, कुछ जीनोमिक क्षेत्रों को विनियामक हॉटस्पॉट के रूप में जाना जाता है, जो कई जीनों के अभिव्यक्ति स्तर को बदल सकते हैं।

क्योंकि ईक्यूटीएल और नियमित क्यूटीएल जांच के बीच मुख्य अंतर फेनोटाइप की संख्या है, यह अप्रत्याशित नहीं है कि पारंपरिक क्यूटीएल दृष्टिकोण को एक समय में ईक्यूटीएल एक जीन खोजने के लिए नियोजित किया गया है। यह ''एक समय में एक जीन'' रणनीति, दूसरी ओर, कई आवश्यक दहनशील प्रभावों और जीन इंटरैक्शन की उपेक्षा करती है। एक अनुभवजन्य बेयस विधि के माध्यम से, केंड्जिओर्स्की एट अल., (2006) ने मार्कर और ट्रांसक्रिप्ट दोनों में सूचना साझा करने को बढ़ावा देने के लिए मिक्सचर ओवर मार्कर (एमओएम) मॉडलिंग तकनीक का सुझाव दिया था।

वेन, (2016) ने कई परीक्षण परिकल्पनाओं के एक साथ परीक्षण के लिए एक सरल विधि विकसित की जो गलत खोज दर नियंत्रण (एफडीआर) के मामले में दृढ़ है। एकल और बहु–ऊत्तक ईक्यूटीएल मैपिंग के अनुप्रयोगों में, सुझाई गई तकनीक कम्प्यूटेशनल रूप से कुशल है। ईक्यूटीएल जांच में अत्यधिक बहुलता समस्या से निपटने के लिए बेसियन मॉडल का हाल ही में व्यापक रूप से उपयोग किया गया है। स्पार्स बेसियन रिग्रेशन (एसबीआर) मॉडलिंग पर आधारित कई दृष्टिकोण क्यूटीएल अनुसंधान के लिए स्पष्ट रूप से विकसित किए गए हैं, जैसे कि आर–क्यूटीएलबीआईएम (क्यूटीएल–बेसियन इंटरवल मैपिंग) पैकेज (यी एट अल., (2008)। बाद में, बनर्जी एट अल., (2008) ने एक साथ कई गुणों (जीन) को संभालने के लिए इस दृष्टिकोण में सुधार किया। कम्प्यूटेशनल चुनौती के कारण, उन्होंने लक्षणों की संख्या को पांच की सीमा तक सीमित कर दिया है। इस पत्र के तहत, एक एकीकृत पदानुक्रमित बेसियन मॉडल (स्कॉट–बॉयर एट अल., 2012) को बड़ी मात्रा में एसएनपी–जीन जोड़ी अध्ययन को हल करने के लिए प्राथमिकता दी गई है।

यह पेपर निम्नानुसार व्यवस्थित है। खंड 2, सिस और ट्रांस–एक्सप्रेशन क्वांटिटेटिव ट्रेट लोसाइी (ईक्यूटीएल) की पहचान करने के लिए दो सांख्यिकीय दृष्टिकोणों का वर्णन करता है। खंड 3 सिस और ट्रांस–ईक्यूटीएल की सही संख्या की पहचान करने के लिए एक वास्तविक डेटासेट के लिए प्रस्तावित कार्यप्रणाली के परिणामों पर चर्चा करता है। पेपर निष्कर्ष पर अंतिम खंड के साथ समाप्त होता है।

2 सामग्री और प्रणालियां

2.1 सांख्यिकीय पद्धति

व्यक्तियों में जीन अभिव्यक्ति माप का प्रतिनिधित्व करने के लिए निम्नलिखित मॉडल (1) का उपयोग किया जा सकता है:

$$y_{ig} = \mu_g + \sum_{j=1}^{S} x_{ij\gamma_{jg}\beta_{jg}} + \varepsilon_{ig} \qquad (1)$$

यहाँ, $g = 1, \ldots, G$ एक विशेष जीन या विशेषता को इंगित करता है, $i = 1, \ldots, n$ एक विशिष्ट इन्डविजूअल को दर्शाता है, और $j = 1, \ldots, S$ एक विशिष्ट एसएनपी को दर्शाता है, और है y_{ig} iवें व्यक्ति के लिए जीन जी के लिए जीन अभिव्यक्ति स्तर है, जीन g (सभी व्यक्तियों में) की जीन अभिव्यक्ति का समग्र माध्यम है; μg एक योज्य, प्रभावी, या अप्रभावी आनुवंशिक मॉडल के तहत x_{ij} iवें व्यक्ति के लिए लोसाइस र पर जीनोटाइप का प्रतिनिधित्व करता है; β_{jg} जीन ह पर एसएनपी j का प्रभाव आकार है।

क्योंकि केवल कुछ मार्कर सीधे जीन अभिव्यक्तियों को प्रभावित करते हैं, कई β लगभग शून्य हैं, बाइनरी इंडिकेटर चर मॉडल में जोड़े जाते हैं ताकि यह निर्दिष्ट किया जा सके कि मॉडल में कौन से मार्कर शामिल किए जाने चाहिए, यानी यदि एसएनपी j को जीन के लिए γ_{jg} मॉडल में शामिल किया जाना चाहिए g और $\gamma_{jg} = 0$। ε_{ig} जीन–विशिष्ट विचरण के साथ एक गाऊसी त्रुटि शब्द है।

लचीले जीनोम–व्यापी पूर्व वितरण प्रतिनिधित्व का उपयोग करके इस संभावना का अनुमान लगाने के लिए जीनों में उधार ली गई स्ट्रेंन्थ के साथ जीन/मार्कर–विशिष्ट क्यूटीएल संभावना को प्राथमिकता w_{jg} $P(\gamma_{jg}=1)$ के रूप में माना जाता है। मॉडल की पदानुक्रमित संरचना कई जीनों से जुड़े ईक्यूटीएल का पता लगाने में सहायता करती है। यह भी माना जाता है कि यादृच्छिक त्रुटि शब्द अर्थात स्वतंत्र और समान रूप से वितरित (ई.आई.डी) हैं ताकि जीन सभी मॉडल मापदंडों को देखते हुए सशर्त रूप से स्वतंत्र हों। लेकिन, जीन के बीच निर्भरता को पहले γ_{jg} एक विनिमेय का उपयोग करके प्रस्तुत किया जाता है, जो कम्प्यूटेशनल रूप से समझने में आसान है।

ईक्यूटीएल जांच में हजारों जीन अभिव्यक्ति स्तर मात्रात्मक फेनोटाइप के रूप में कार्यरत है। इस प्रकार के डेटा का विश्लेषण करने के लिए प्रत्येक जीन अभिव्यक्ति प्रोफाइल के लिए यूनीवेरिएट क्यूटीएल विश्लेषण की अक्सर आवश्यकता होती है। ऊपर उल्लिखित मॉडल को सबसे आवश्यक तत्वों पर लागू किया गया था – 1) अधिकांश ईक्यूटीएल का जीन अभिव्यक्ति पर कई प्रभाव पड़ता है। 2) एक ही कार्यात्मक जीनों के आपस में जुड़े होने की संभावना अधिक होती है। नतीजतन, अधिक व्यावहारिक निष्कर्ष निकालने के लिए सैकड़ों या हजारों जीन अभिव्यक्ति सुविधाओं से जानकारी एकत्र करने की संभावना है।

2.2 पूर्व वितरण

मॉडल के विभिन्न पूर्व वितरणों को निम्नानुसार वर्णित किया गया है– पैरामीटर सेटअप $\theta = (\mu_g,\sigma^g{}_2,\gamma_{jg},\beta_{jg},w_{jg},p_j,a_j,b_j)$ द्वारा दिया गया है। पूर्वगामी के सेटअप को निम्नानुसार परिभाषित किया गया है–

जीन g के लिए मॉडल में एसएनपी j को शामिल करने की पूर्व संभावना γ_{jg}~Bernoulli (w_{jg}) के द्वारा दी गई है, $P(\gamma_{jg}=1) = w_{jg}$ जहां एक अज्ञात पैरामीटर है।

एसएनपी की केवल छोटी संख्या एक जीन अभिव्यक्ति के निर्धारक के रूप में कार्य करती है, समावेशन संभावना पैरामीटर w_{jg} अधिकांश समय मान 0 को प्राथमिकता देते हैं। जब w_{jg} 0 नहीं होता है, तो इसे बीटा वितरण Beta(a_j,b_j) से माना जाता है। इसे 0 पर डायराक द्रव्यमान के मिश्रण और वजन p_j के साथ बीटा वितरण के रूप में और $1-p_j$ निम्नानुसार व्यक्त किया जा सकता है

$$w_{jg} \sim p_j\, \delta_o\,(w_{jg}) + (1-p_j)\mathrm{Beta}(a_j,b_j)\,(w_{jg}) \quad (2)$$

पैरामीटर pj (जहाँ w_{jg} कि संभावना 0 है) सभी जीनों के लिए समान है जो एक एसएनपी का पता लगाने में मदद करता है जो कई जीन एक्सप्रेशन के साथ कमजोर रूप से जुड़ा हुआ है (लोसाइडास एट अल., 2006)। अब, हाइपरपैरामीटर a_0 और b_0 के साथ p_j के लिए, एक सामान्य संयुग्म बीटा पूर्व $p_j \sim \mathrm{Beta}(a_0,b_0)$ का उपयोग किया जाता है– a_j और $b_j\lambda_{_a}$ and $\lambda_b{:}a_j \sim \exp(\lambda_a)$ $b_j \sim \exp(\lambda_b)$ अब, $\mu_g \sim N(m_g,\tau_g{}^2)$ को हाइपरपैरामीटर के साथ घातीय वितरण का पालन करने के लिए माना जाता है और जहां m_g और τ_g अनुभवजन्य माध्य और जीन अभिव्यक्ति g का विचरण है।

$\beta_{jg} = 0$ if $\gamma_{jg} = 0$ and $\beta_{jg} \sim N(0,v_g{}^2$ if $\gamma_{jg} = 1$, with $v_{jg}2 = c(x_j{}^T x_j)\text{-}1\ \sigma_g{}^2$ जहाँ c स्केलिंग फैक्टर पैरामीटर है और $(x_j^T x_j)^{-1} = (\sum_{i=1}^{n} x_{ij}^2)^{-1}$ प्रतिगामी विचरण की नकल करता है, जो जेलनर (1986) के प्रसिद्ध g – पूर्वगामी के रूप में जाना जाता है।

यहाँ यी एट अल., (2008) के दृष्टिकोण का पालन किया जाता है और c को एसएनपी की संख्या के बराबर निरंतर माना जाता है। बॉटलो और रिचर्डसन (2010) ने जेलनर और सिओ (1980) पूर्व पर आधारित एक व्युत्क्रम–गामा पूर्व $c \sim$ व्युत्क्रम–गामा $(\frac{1}{2},\frac{n}{2})$ पर विचार किया। पेट्रेटो एट अल., (2010), लियांग एट अल., (2008) के पूर्व के साथ सभी जीनों के लिए एक सामान्य c माना जाता है। $c \sim \frac{1}{1+c}$ अंतराल $(0, M)$ में, जहां समापन बिंदु M है, $M = \max(n, S^2)$। जहाँ, $v_{jg}{}^2$ का समग्र विचरण $\sigma_g{}^2$ है पैरामीटर मॉडल में एक उप

पैरामीटर σ_g^2 है और इसे एकीकृत किया जा सकता है, जहाँ, $\sigma_g^2 \sim$ उलटा–गामा $\frac{1}{2},\frac{1}{2}$।

2.3 पैरामीटर अनुमान

पश्च वितरण मार्कोव चेन मोंटे कार्लो (एमसीएमएसी) विधियों का उपयोग करके उत्पन्न किया जाता है (गेलफोन्ड और स्मिथ, 1990)। सभी अद्यतनों के लिए गिब्स नमूनाकरण का उपयोग किया गया था, और a_j और b_j के लिए पूर्ण सशर्त के लिए कोई बंद प्रपत्र नहीं हैं। अनुकूली अस्वीकृति नमूनाकरण का उपयोग इन दो मापदंडों (गिलक्स एंड वाइल्ड, 1995) को अद्यतन करने के लिए किया जाता है। पुनरावृत्तियों की संख्या राफटेरी और लुईस के दृष्टिकोण (राफटेरी एट अल., 1992य लुईस एट अल., 1996) का उपयोग करके निर्धारित की गई है। यहां संकलित किए गए दो डेटासेट के लिए मानक पश्च मूल्यों का अनुमान लगाने के लिए, 50,000 बर्न–इन पुनरावृत्तियों के साथ 10,000 पुनरावृत्तियों को नियोजित किया गया है। अंत में, मॉडल चार हाइपरपरमेटर्स पर आधारित है जो पूर्व निर्धारित हैं। इन मूल्यों की प्राथमिकता की गणना करने के लिए ई–क्यूटीएल की अपेक्षित संख्या और उनके फैलाव का उपयोग किया जाता है। इस पद्धति के तहत जो कम ईक्यूटीएल वाले मॉडल का समर्थन करता है, चुना जाता है,। अंतिम में जीन/एसएनपी संबंधों को खोजना है, जो हमारे मॉडल से पैरामीटर अनुमानों की सहायता से किया जा सकता है। यदि एसएनपी j में जीन जी के जुड़ाव की पश्च संभावना 0.80 से अधिक है, तो एक ईक्यूटीएल की पहचान की जाती है

3 परिणाम और चर्चा

यहाँ, दो वास्तविक डेटासेट का विश्लेषण किया गया है ताकि ईक्यूटीएल के प्रकार और हॉटस्पॉट की पहचान का पता लगाया जा सके।

डेटासेट 1: St/Mx जनसंख्या में जौ पत्ती रतुआ रोगजनक पी. होर्डेई के लिए मात्रात्मक प्रतिरोध की विशेषता के लिए एक प्रयोग किया गया है। Agilent जौ कस्टम माइक्रोएरे का उपयोग पी. होर्डेई के साथ चुनौती दी गई St/Mx आबादी की 139 डीएच लाइनों में प्रतिलेख बहुतायत का आकलन करने के लिए किया गया है। डेटासेट पहले रोस्टोक्स एट अल., (2005) द्वारा उत्पन्न किए गए थे। और चेन एट अल. (2010) में क्रमशः 4286 एसएनपी और 595754 भाव है।

तालिका 1: पीपीए > 0.80 के साथ पता लगाए गए ईक्यूटीएल

एस. एन.	जीन	एसएनपी	पीपीए	प्रकार
1	Contig12563_s_at	Contig20996_10	0.81	सिस
2	Contig10533_at	Contig12729_5	1.00	सिस
3	Contig1031_at	Contig5754_9	0.80	सिस

डेटासेट 2: हमने सार्वजनिक रूप से उपलब्ध ईक्यूटीएल डेटा का विश्लेषण करने के लिए ऊपर चर्चित एफडीआर दृष्टिकोण लागू किया है। डेटा क्यूटीएल आर्काइव में उपलब्ध है, जो अब माउस फेनोम डेटाबेस का हिस्सा है, http://phenome.jax.org/db/q? rtn=projects/projdet&reqprojid=532. जीन अभिव्यक्ति को मापने के लिए कस्टम एगिलेंट दो–रंग इंक–जेट माइक्रोएरे का उपयोग किया गया है। एफीमेट्रिक्स 5K जीन चिप का उपयोग चूहों जीनोटाइप करने के लिए किया गया है। अंत में, डेटासेट में 28 नमूने, 918 एसएनपी और 23,238 अभिव्यक्ति मूल्य हैं। ऊपर दिए गए डेटासेट में पहचाने गए ईक्यूटीएल की संख्या 4243 है। इस मामले में पता लगाने का तंत्र कनेक्शन (पीपीए) की पश्चगामी संभावना पर आधारित है। महत्वपूर्ण एसएनपी–जीन कनेक्शन का मूल्यांकन किया गया है यदि पीपीए 0.80 से अधिक है, जो विधि का कटऑफ मान है। सभी पीपीए मूल्यों के साथ–साथ जीन–एसएनपी संबंध तालिका 2 में बताए गए हैं। तालिका 3 में खोजे गए सिस और ट्रांस ईक्यूटीएल सूचीबद्ध है।

तालिका 2: पीपीए 0.8 के कट ऑफ मूल्य से परे और एफडीआर स्तर 10% के साथ विभिन्न जीन और एसएनपी के साथ जुड़ाव की पश्च संभावना

एस. एन.	जीन	एसएनपी	पीपीए
1	1422462_at	rs6263067	1
2	1438426_at	rs6263067	0.97
3	1438852_x_at	rs6263067	0.93
4	1450813_a_at	rs6263067	0.88
5	1416647_at	rs4222763	0.87
6	1428844_a_at	rs4222763	0.88
7	1436955_at	rs4222763	0.9
8	1441568_at	rs4222763	0.88
9	1448604_at	rs4222763	0.97
10	1448958_at	rs4222763	0.98
11	1427711_a_at	rs13476267	1
12	1444320_at	mCV23574676	1
13	1439075_at	rs6331493	1
14	1424811_at	rs3142215	0.92
15	1415673_at	rs13479058	0.92
16	1415677_at	rs13479058	0.89
17	1415703_at	rs13479058	0.9
18	1415713_a_at	rs13479058	0.88
19	1415715_at	rs13479058	0.9
20	1450003_at	rs3693435	0.93

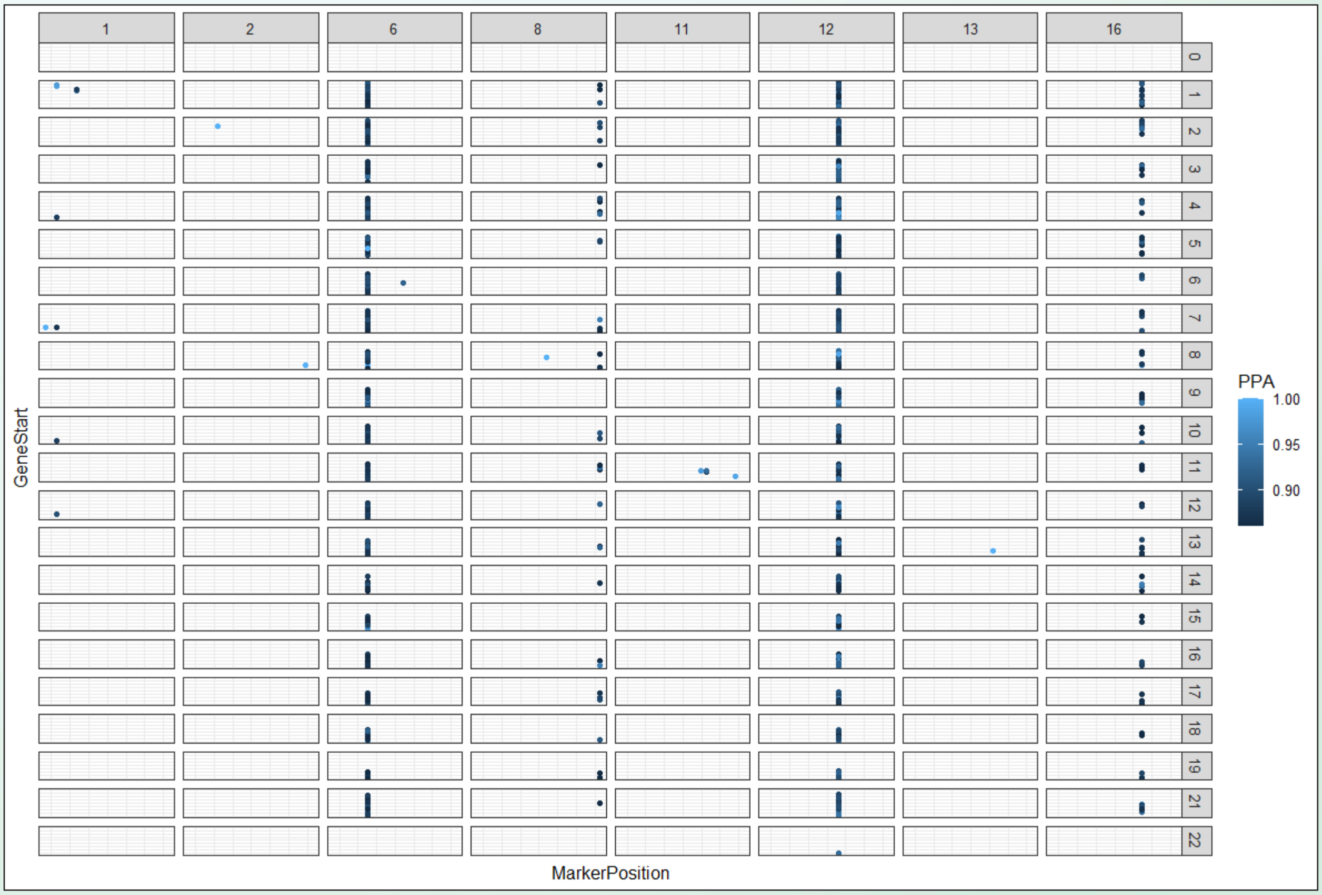
ईक्यूटीएल हॉटस्पॉट पहचान

ईक्यूटीएल और जीन के बीच जुड़ाव (पीपीए) की पश्च संभावना के आधार पर, चार हॉटस्पॉट पाए गए, और चार हॉटस्पॉट की सूची चित्र 2 में प्रस्तुत की गई है। चित्र 2 से पता चलता है कि ट्रांस–ईक्यूटीएल की उच्च संख्या में मार्कर गुणसूत्र 6, 8, 12, और 16 के साथ समान स्थान में है। तालिका 4 से पता चलता है कि एसएनपी rs13479058 एक महत्वपूर्ण में संख्या जीन से जुड़ा हुआ है और गुणसूत्र 6 (मार्कर गुणसूत्र 6) पर समान स्थानों में पाया गया है। नतीजतन, इन जीनों को हॉटस्पॉट के रूप में जाना जाता है। इसी तरह, तालिका 6 से पता चलता है कि मार्कर क्रोमोसोम 16 का उपयोग करते हुए, एसएनपी rs3693435 जीन की एक महत्वपूर्ण संख्या के साथ जुड़ा हुआ है और समान स्थान में है। तालिका 6 से पता चलता है कि एसएनपी rs6170344 जीन की एक महत्वपूर्ण संख्या से जुड़ा हुआ है और क्रोमोसोम 12 (मार्कर क्रोमोसोम 12) पर एक समान स्थान में है। एसएनपी rs6343961 क्रोमोसोम 8 पर जीन और साझा साइटों की एक महत्वपूर्ण संख्या से जुड़ा हुआ है, जो तालिका 7 में मार्कर क्रोमोसोम 8 हॉटस्पॉट के लिए ईक्यूटीएल का एक उदाहरण प्रदर्शित करता है।

तालिका 3: कुछ पहचाने गए सिस और ट्रांस–ईक्यूटीएल को पीपीए के साथ प्रस्तुतिकरण

जीन	एसएनपी	पीपीए	जीन गुणसूत्र	जीन प्रारंभ	जीन अंत	क्रोम मार्कर	मार्कर स्थान	प्रकार
1448958_at	rs4222763	0.98	1	1.58E+08	1.58E+08	1	168084791	ट्रांस
1427711_a_at	rs13476267	1	7	26246721	26262644	1	183253943	ट्रांस
1444320_at	mCV23574676	1	8	26835796	26865068	2	26699898	ट्रांस
1439075_at	rs6331493	1	2	1.44E+08	1.44E+08	2	144676277	ट्रांस
1424811_at	rs3142215	0.92	6	85767215	85770966	6	89630672	ट्रांस
1415673_at	rs13479058	0.92	5	1.3E+08	1.3E+08	6	137525036	ट्रांस
1415677_at	rs13479058	0.89	14	56357857	56364527	6	137525036	ट्रांस
1415703_at	rs13479058	0.9	21	1.48E+08	1.48E+08	6	137525036	ट्रांस
1415713_a_at	rs13479058	0.88	12	1.05E+08	1.05E+08	6	137525036	ट्रांस
1415715_at	rs13479058	0.9	5	33995865	34000626	6	137525036	ट्रांस
1415734_at	rs13479058	0.86	6	87949100	87995264	6	137525036	ट्रांस

जीन	एसएनपी	पीपीए	जीन गुणसूत्र	जीन प्रारंभ	जीन अंत	क्रोम मार्कर	मार्कर स्थान	प्रकार
1415746_at	rs13479058	0.9	7	26052723	26079178	6	137525036	ट्रांस
1415749_a_at	rs13479058	0.96	4	1.24E+08	1.24E+08	6	137525036	ट्रांस
1415754_at	rs13479058	0.89	15	78971797	78982197	6	137525036	ट्रांस
1415758_at	rs13479058	0.87	5	73411430	73647857	6	137525036	ट्रांस
1415771_at	rs13479058	0.9	1	88241294	88256030	6	137525036	ट्रांस
1415772_at	rs13479058	0.88	1	88241294	88256030	6	137525036	ट्रांस
1415790_at	rs13479058	0.9	11	44268073	44284000	6	137525036	ट्रांस
1415816_at	rs13479058	0.86	6	85402067	85418466	6	137525036	ट्रांस
1415830_at	rs13479058	0.87	5	21992303	22056247	6	137525036	ट्रांस



चित्र 2: एक्स–अक्ष जीनोम के साथ प्रत्येक ईक्यूटीएल की स्थिति का प्रतिनिधित्व करता है, जबकि वाई–अक्ष जांच सेट लोसाइष्य की स्थिति का प्रतिनिधित्व करता है। क्रोमोसोमल बॉर्डर ग्रे लाइनों द्वारा दिखाए जाते हैं। सिस–क्यूटीएल द्वारा एक विकर्ण रेखा बनाई जाती है। एकल ट्रांस–क्यूटीएल से संबंधित प्रतिलेख समूह लंबवत बैंड द्वारा दिखाए जाते हैं।

तालिका 4: मार्कर क्रोमोसोम 6 के हॉटस्पॉट के लिए ईक्यूटीएल का एक नमूना

जीन	एसएनपी	पीपीए	जीन क्रोम	जीन स्टार्ट	जीनएंड	मार्कर क्रोम	मार्कर स्थिति	प्रकार
1415673_at	rs13479058	0.92	5	1.3E+08	1.3E+08	6	1.38E+08	ट्रांस
1415677_at	rs13479058	0.89	14	56357857	56364527	6	1.38E+08	ट्रांस
1415703_at	rs13479058	0.9	21	1.48E+08	1.48E+08	6	1.38E+08	ट्रांस
1415713_a_at	rs13479058	0.88	12	1.05E+08	1.05E+08	6	1.38E+08	ट्रांस
1415715_at	rs13479058	0.9	5	33995865	34000626	6	1.38E+08	ट्रांस
1415734_at	rs13479058	0.86	6	87949100	87995264	6	1.38E+08	ट्रांस
1415746_at	rs13479058	0.9	7	26052723	26079178	6	1.38E+08	ट्रांस
1415749_a_at	rs13479058	0.96	4	1.24E+08	1.24E+08	6	1.38E+08	ट्रांस
1415754_at	rs13479058	0.89	15	78971797	78982197	6	1.38E+08	ट्रांस
1415758_at	rs13479058	0.87	5	73411430	73647857	6	1.38E+08	ट्रांस
1415771_at	rs13479058	0.9	1	88241294	88256030	6	1.38E+08	ट्रांस
1415772_at	rs13479058	0.88	1	88241294	88256030	6	1.38E+08	ट्रांस
1415790_at	rs13479058	0.9	11	44268073	44284000	6	1.38E+08	ट्रांस
1415816_at	rs13479058	0.86	6	85402067	85418466	6	1.38E+08	ट्रांस

तालिका 5: मार्कर क्रोमोसोम 16 के ईक्यूटीएल हॉटस्पॉट का एक नमूना

जीन	एसएनपी	पीपीए	जीन क्रोम	जीन स्टार्ट	जीनएंड	मार्कर क्रोम	मार्कर स्थिति	प्रकार
1415771_at	rs3693435	0.9	1	88241294	88256030	16	63508672	ट्रांस
1415830_at	rs3693435	0.87	5	21992303	22056247	16	63508672	ट्रांस
1415927_at	rs3693435	0.88	2	1.14E+08	1.14E+08	16	63508672	ट्रांस
1416066_at	rs3693435	0.88	6	1.25E+08	1.25E+08	16	63508672	ट्रांस
1416237_at	rs3693435	0.88	9	44850508	44862098	16	63508672	ट्रांस
1416270_at	rs3693435	0.87	19	8867619	8873047	16	63508672	ट्रांस
1416575_at	rs3693435	0.86	16	18780540	18812080	16	63508672	ट्रांस
1416687_at	rs3693435	0.86	9	92437061	92503266	16	63508672	ट्रांस
1416979_at	rs3693435	0.88	5	1.49E+08	1.49E+08	16	63508672	ट्रांस
1417160_s_at	rs3693435	0.93	11	83522517	83524850	16	63508672	ट्रांस
1417294_at	rs3693435	0.87	4	1.39E+08	1.39E+08	16	63508672	ट्रांस

तालिका 6: मार्कर क्रोमोसोम 12 के हॉटस्पॉट के लिए ईक्यूटीएल का एक नमूना

जीन	एसएनपी	पीपीए	जीन क्रोम	जीन स्टार्ट	जीनएंड	मार्कर क्रोम	मार्कर स्थिति	प्रकार
1415683_at	rs6170344	0.91	11	1.03E+08	1.03E+08	12	83729204	ट्रांस
1415684_at	rs6170344	0.94	10	43988164	44084097	12	83729204	ट्रांस
1415690_at	rs6170344	0.94	11	94515081	94521403	12	83729204	ट्रांस
1415693_at	rs6170344	0.95	15	57701057	57723973	12	83729204	ट्रांस
1415695_at	rs6170344	0.95	7	1.21E+08	1.21E+08	12	83729204	ट्रांस
1415697_at	rs6170344	0.92	5	92481172	92512761	12	83729204	ट्रांस
1415703_at	rs6170344	0.92	21	1.48E+08	1.48E+08	12	83729204	ट्रांस
1415718_at	rs6170344	0.86	11	57615139	57623719	12	83729204	ट्रांस
1415735_at	rs6170344	0.86	19	10680115	10704312	12	83729204	ट्रांस
1415742_at	rs6170344	0.92	6	83004647	83007674	12	83729204	ट्रांस

तालिका 7: मार्कर क्रोमोसोम 8 के हॉटस्पॉट के लिए ईक्यूटीएल का एक नमूना

जीन	एसएनपी	पीपीए	जीन क्रोम	जीन स्टार्ट	जीनएंड	मार्कर क्रोम	मार्कर स्थिति	प्रकार
1418669_at	rs6343961	0.86	4	1.37E+08	1.37E+08	8	18316422	ट्रांस
1419009_at	rs6343961	0.86	4	56756285	56757797	8	18316422	ट्रांस
1420351_at	rs6343961	0.9	4	1.55E+08	1.55E+08	8	18316422	ट्रांस
1420716_at	rs6343961	0.86	1	1.68E+08	1.68E+08	8	18316422	ट्रांस
1422628_at	rs6343961	0.9	19	12620230	12664258	8	18316422	ट्रांस
1423419_at	rs6343961	0.87	11	82594610	82616364	8	18316422	ट्रांस
1423774_a_at	rs6343961	0.95	7	87439403	87461145	8	18316422	ट्रांस
1424187_at	rs6343961	0.86	16	45093515	45127778	8	18316422	ट्रांस
1424923_at	rs6343961	0.92	12	1.05E+08	1.05E+08	8	18316422	ट्रांस
1426190_at	rs6343961	0.9	5	1.14E+08	1.14E+08	8	18316422	ट्रांस
1426638_at	rs6343961	0.87	17	86012948	86031153	8	18316422	ट्रांस
1427324_at	rs6343961	0.89	5	1.24E+08	1.24E+08	8	18316422	ट्रांस
1427786_at	rs6343961	0.94	11	96874708	96886048	8	18316422	ट्रांस

निष्कर्ष

जीनोमिक्स का एक महत्वपूर्ण लक्ष्य जीन एक्सप्रेशन जैसे लक्षणों के लिए एक्सप्रेशन क्वांटिटेटिव ट्रेट लोसाइ (ईक्यूटीएल) के स्थान का पता लगाना और उसकी पहचान करना है। सिस और ट्रांस–ईक्यूटीएल की उपस्थिति का पता लगाना लक्षित विशेषताओं के लिए ईक्यूटीएल के संचयी महत्व को निर्धारित करने के लिए महत्वपूर्ण है। दूरस्थ जीनों की अभिव्यक्ति में बड़े बदलावों को पकड़ने वाले सिस और ट्रांस ईक्यूटीएल की पहचान जीनोमिक्स में एक महत्वपूर्ण अवधारणा है। वर्तमान पत्र में, अभिव्यक्ति मात्रात्मक लोसाइी के लिए

एकीकृत पदानुक्रमित बेसियन मॉडल का उपयोग करके उन सिस और ट्रांस–ईक्यूटीएलएस की पहचान करने पर प्रमुख ध्यान दिया गया है। वर्तमान जांच मुख्य रूप से जौ की फसल के लिए की गई है, जिसमें केवल सिस – ईक्यूटीएल का पता चला, कोई अन्य ईक्यूटीएल हॉटस्पॉट नहीं पाए गए। ट्रांस एक्टिंग ईक्यूटीएल के अध्ययन के लिए चूहे का जीन एक्सप्रेशंस लिए गए जिसमें सिस ट्रांस ईक्यूटीएल दोनों महत्वपूर्ण पाए गए और चार ईक्यूटीएल हॉटस्पॉट भी पाए गए, जो ऊपर सूचीबद्ध किया गया है।

संदर्भ

बनर्जी, एस., यांडेल, बी.एस., एवं यी, एन. (2008)। बेसियन क्वांटिटेटिव ट्रेट लोसाइी मैपिंग फॉर मल्टीपल ट्रेट्स। जेनेटिक्स, 179(4), 2275–2289।

ब्रेम, आर.बी., यवर्ट, जी., क्लिंटन, आर., एवं क्रुगल्याक, एल. (2002)। जेनेटिक डाइसेक्शन ऑफ ट्रांसक्रिप्शनल रेग्यलैशन इन बडिंग यीस्ट। साइंस, 296(5568), 752–755।

बॉटलो, एल., एवं रिचर्डसन, एस. (2010)। इवोल्यूशनरी स्टोकेस्टिक सर्च फॉर बेसियन मॉडल एक्सप्लोरेशन। बेसियन एनालिसिस, 5(3), 583–618।

चेन, एक्स., हैकेट, सी.ए., निक, आर.ई., हेडली, पी.ई., बूथ, सी., ड्रुका, ए., एवं वॉ, आर. (2010)। एन ईक्यूटीएल एनालिसिस ऑफ पार्शल रिजिस्टन्स टू पुकिनिया होर्डेई इन बार्ली। प्लोस वन, 5(1), ई8598।

ड्रुका, ए., पोटोकिना, ई., लुओ, जेड., जियांग, एन., चेन, एक्स., कर्सी, एम., एवं वॉ, आर. (2010)। इक्स्प्रेशन क्वान्टिटेटिव ट्रेट लोसाइी एनालिसिस इन प्लांट। प्लांट बायोटेक्नोलॉजी जर्नल, 8(1), 10–27।

फ्लुट्रे, टी., वेन, एक्स., प्रिचर्ड, जे., एवं स्टीफंस, एम. (2013)। कई ऊतकों में संयुक्त ईक्यूटीएल विश्लेषण के लिए सांख्यिकीय ढांचा। प्लस जेनेट., 2013, 9(5), ई1003486।

गेलफोन्ड, जे.ए., इब्राहिम, जे.जी., एवं जू, एफ. (2007)। प्राक्सिमिटी मॉडल फॉर इक्स्प्रेशन क्वान्टिटेटिव ट्रेट लोसाइी (ईक्यूटीएल) डिटेक्शन। बायोमेट्रिक्स, 63(4), 1108–1116।

गिलक्स, डब्ल्यू.आर., रिचर्डसन, एस., एवं स्पीगलहेल्टर, डी.जे. (1995)। इन्ट्रोडूसिंग मार्कोव चेन मोंटे मार्कोव चेन मोंटे कार्लो इन प्रैक्टिस, (पहला संस्करण), चौपमैन एवं हॉल / सीआरसी।

गोरिंग, एच.एच., क्यूरन, जे.ई., जॉनसन, एम.पी., डायर, टी.डी., चार्ल्सवर्थ, जे., कोल, एस.ए., एवं ब्लैंगेरो, जे. (2007)। डिस्कवरी ऑफ इक्स्प्रेशन इक्यूटीएल यूजिंग लार्ज स्केल ट्रांसक्रिप्शनल प्रोफाइलिंग इन ह्यूमन लिम्फोसाइटों। नेचर जेनेटिक्स, 39(10), 1208–1216।

केंदजियोर्स्की, सी.एम., चेन, एम., युआन, एम., लैन, एच., एवं एटी, ए.डी. (2006)। स्टटिस्टिकल मेथड्स फॉर एक्सप्रेशन क्वांटिटेटिव ट्रेट लोसाइी (इक्यूटीएल) मैपिंग। बायोमेट्रिक्स, 62(1), 19–27।

लियांग, एफ., पाउलो, आर., मोलिना, जी., क्लाइड, एम.ए., एवं बर्जर, जे.ओ. (2008)। मिक्स्चर ऑफ जी प्रायर्स फॉर बेसियन वेरीअबल सिलोसाइज़न। जर्नल ऑफ द अमेरिकन स्टैटिस्टिकल एसोसिएशन, 103(481), 410–423।

मार्सेल, टी.सी., वार्ष्णेय, आर.के., बारबेरी, एम., जाफरी, एच., डी कॉक, एम.जे.डी., ग्रेनर, ए., एवं निक्स, आर.ई. (2007)। ए हाइ डेन्सिटी कन्सेन्सस मैप ऑफ टू कम्पेर द डिस्ट्रब्यूशन ऑफ क्यूटीएल फॉर पार्शल रिजिस्टन्स टू पुकिनिया होर्डेई इन बार्ली ऑफ डिफेन्स जीन होमोलॉग्स। थ्योरेटिकल एंड एप्लाइड जेनेटिक्स, 114(3), 487–500।

पेट्रेटो, ई., बोटोलो, एल., लैंगली, एसआर, हेनिग, एम., मैकडरमॉट–रो, सी., सरवर, आर., एवं रिचर्डसन, एस. (2010)। न्यू इन्साइट इनटू द जेनेटिक कन्ट्रोल ऑफ जीन इक्स्प्रेशन यूजिंग ए बेसियन मल्टी–टिशू अप्रोच। प्लोस कम्प्यूटेशनल बायोलाजी, 6(4), इ1000737।

पियर्स, बी.एल, टोंग, एल., चेन, एल.एस., रहमान, आर., आर्गोस, एम., जैस्मीन, एफ., एंड अहसान,

एच. (2014)। मीडीएशन एनालिसिस डेमन्स्ट्रेट देट ट्रांस–ईक्यूटीएल आर ऑफन इक्स्प्लेन बाइ सिस–मीडीएशनः ए जीनोम– बाइड एनालिसिस अमंग 1,800 साउथ एशियन्स। पीएलओएस जेनेटिक्स, 10(12), ई1004818।

पोटोकिना, ई., ड्रुका, ए., लुओ, जेड., वाइज, आर., वॉ, आर., एंड केर्सी, एम. (2008)। जीन एक्सप्रेशन क्वांटिटेटिव ट्रेट लोसाइस एनालिसिस ऑफ 16000 बार्ली जीन रिवील ए काम्प्लोसाइस पैटर्न ऑफ जीनोम–वाइड ट्रांसक्रिप्शनल रेगुलेशन। द प्लांट जर्नल, 53(1), 90–101।

राफटेरी, ए.ई., लुईस, एस., बर्नार्डो, जे.एम., बर्जर, जे.ओ., डेविड, ए.पी., एंड स्मिथ, ए.एफ.एम. (1992)। बेसियन स्टोकेस्टिक्स। ऑक्सफोर्ड साइंस पब्लोसाइेशन, 323–349।

रोस्टोक्स, एन., मुडी, एस., कार्डल, एल., रसेल, जे., रामसे, एल., बूथ, ए., एवं वॉ, आर. (2005)। जीनोम–वाइड एसएनपी डिस्कवरी एंड लोसाइेज एनालिसिस इन बार्ली बेस्ड ऑन जीन रिस्पान्सिव टू एबाइआटिक स्ट्रेस। मोलोसाइयलर जेनेटिक्स एंड जीनोमिक्स, 274(5), 515–527।

स्कॉट–बोयर, एम.पी., इम्होल्ते, जी.सी., तैयब, ए., लब्बे, ए., डेसचेपर, सी.एफ., एंड गोटार्डो, आर. (2012)। इन्टग्रेटिड हाइअरार्किकल बेसियन मॉडल ईक्यूटीएल मैपिंग। मल्टवेरीइट स्टैटिस्टिकल ऐप्लोसाइेशन इन जेनेटिक्स मोलोसाइयलर बाइआलजी, 11(4)।

शबालिन, ए. ए. (2012)। मैट्रिक्स ईक्यूटीएलः अल्ट्रा फास्ट ईक्यूटीएल वाया लार्ज मैट्रिक्स ऑपरेशन। बायोइनफॉरमेटिक्स, 28(10), 1353–1358।

शान, एन., वांग, जेड., एंड हौ, एल. (2019)। आइडेन्टफकेशन ऑफ ट्रांस–ईक्यूटीएल यूजिंग मीडीएशन एनालिसिस विद मल्टपल मीडीएटर। बीएमसी बायोइनफॉरमेटिक्स, 20(3), 87–97।

वेन, एक्स. (2016)। मोलोसाइयलर क्यूटीएल डिस्कवरी इन्कॉर्परेट जीनोमिक एनोटेशन यूजिंग बेसियन फाल्स डिस्कवरी कन्ट्रोल रेट। द ऐनल्ज ऑफ अप्लाइड स्टोकेस्टिक्स, 10(3), 1619–1638।

यी, एन., एवं श्राइनर, डी. (2008)। ऐड्वैन्स इन बेसियन मल्टीपल क्वांटिटेटिव ट्रेट लोसाइी मैपिंग इन इक्स्पेरमेन्टल क्रॉस। हरेडिटी, 100(3), 240–252।

जेलनर, ए. (1986। आन असेसिंग प्रायर डिस्ट्रब्यूशन एंड बेसियन रिग्रेशन एनालिसिस विद जी प्रायर डिस्ट्रब्यूशन। इन इन्फ्रेन्स एंड डिसीजन टेक्नीकः एसे इन हॉनर ऑफ ब्रूनो डी फिनेटी।

जेलनर ए., सिओ ए. (1980)। पास्टीरीअर ऑड्स रेशो फॉर सलोसाइिटड रिग्रेशन हाइपाथिसिस। इनः बर्नार्डो जेएम, डी ग्रोट एमएच, लिंडले डीवी, स्मिथ एएफएम, एडिटर। बेसियन स्टैटिस्टिक्स प्रोक फर्स्ट इंटरनेशनल मीटिंग, 585–603।

झू, जेड., झांग, एफ., हू, एच., बख्शी, ए., रॉबिन्सन, एम.आर., पॉवेल, जेई., एवं यांग, जे. (2016)। इन्टग्रेशन ऑफ समरी डेटा फ्राम गवासी एंड ईक्यूटीएल स्टडीज प्रिडिक्ट काम्प्लोसाइस ट्रेट जीन टारगेट। नेचर जेनेटिक्स, 48(5), 481–487।

समय श्रृंखला पूर्वानुमान के लिए इंटड्यूशनिस्टिक फ़ज़ी आधाारित ऑटोरेग्रेसिव इंटीग्रेटेड मूविंग एवरेज (फरिमा) मॉडल

अनिता सरकार, मो. यासीन, रंजित कुमार पॉल, हिमाद्री शेखर रॉय, प्रकाश कुमार, अंकित कुमार सिंह, अमृत कुमार पॉल, एवं एस. पी. सिंह

भा.कृ.अनु.प.– भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली

सारांश

समय श्रृंखला पूर्वानुमान में भविष्य की भविष्यवाणियों को लेकर अंतर्निहित अनिश्चितता के कारण, फजी तकनीकों का उपयोग जटिल और अनिश्चित डेटा के प्रबंधन के लिए एक शक्तिशाली उपकरण के रूप में उभरा है। इस शोध में, पूर्वानुमान के लिए इंट्यूशनिस्टिक फजी आधारित ऑटोरेग्रेसिव मूविंग एवरेज (आईफरिमा) मॉडल का प्रस्ताव करने के लिए पारंपरिक अरिमा मॉडल के साथ फजी लॉजिक को एकीकृत करके एक नया दृष्टिकोण पेश किया गया है। हम यह मूल्यांकन करना चाहते हैं कि जनवरी 2010 से दिसंबर 2022 तक उपजे कुसुम के मासिक थोक मूल्यों (रुपयेध्क्यू) पर विशेष रूप से ध्यान केंद्रित करते हुए, यह एकीकरण पूर्वानुमान सटीकता को कैसे बढ़ाता है। यह विश्लेषण पूर्वानुमान प्रदर्शन में 50% से अधिक की पर्याप्त वृद्धि का खुलासा करता है, जो प्रभावशीलता पर प्रकाश डालता है। फजी लॉजिक को अरिमा फ्रेमवर्क में एकीकृत करता है।

मुख्य शब्दः फ़ज़ी लॉजिक, फ़ज़ी सी– का अर्थ है क्लस्टरिंग, टाइम सीरीज पूर्वानुमान, अस्थिरता।

परिचय

अर्थशास्त्र, कृषि, जीव विज्ञान, इंजीनियरिंग, चिकित्सा और मौसम विज्ञान जैसे कई क्षेत्रों में समय श्रृंखला विश्लेषण और पूर्वानुमान विधियां महत्वपूर्ण हैं। वे भविष्य के मूल्यों की भविष्यवाणी करने के लिए पिछले डेटा रुझानों की व्याख्या को शामिल करते हैं। हाल के दिनों में, कृषि कीमतों का पूर्वानुमान अध्ययन का एक उल्लेखनीय आशाजनक क्षेत्र बन गया है (मित्रा और पॉल, 2017)। रैखिक समय श्रृंखला मॉडल के दायरे में, बॉक्स जेनकिंस (बॉक्स एट अल., 2007) ऑटोरेग्रेसिव इंटीग्रेटेड मूविंग एवरेज (अरिमा) पद्धति अपनी मजबूत सांख्यिकीय विशेषताओं के कारण एक महत्वपूर्ण और व्यापक रूप से नियोजित तकनीक के रूप में सामने आती है। कृषि वस्तु मूल्य डेटा में निहित पर्याप्त अनिश्चितता और अस्थिरता के कारण, डेटा में मौजूद अनिश्चितता या अस्पष्टता को संबोधित करने पर अधिक ध्यान देने की आवश्यकता है क्योंकि अरिमा मॉडल डेटा में इन चीजों की उपस्थिति को कैप्चर करने में विफल रहता है। इन अनिश्चितताओं और अस्पष्टताओं को कम करने के लिए, जादेह (1965) ने फजी लॉजिक और मेंबरशिप फंक्शन का उपयोग करके फजी सेट्स (एफएस) नामक एक अभूतपूर्व अवधारणा पेश की, जो अनिश्चितता और अस्पष्टता को प्रभावी ढंग से मॉडलिंग करने में सक्षम है। सॉन्ग और चिसोम (1993) ने मॉडल विकसित करके और इसे अलबामा विश्वविद्यालय के नामांकन डेटा पर लागू करके फजी टाइम सीरीज की अवधारणा का बीड़ा उठाया। क्रैसिमिर अटानासोव (1986) ने इन्टूइशनिस्टक

फजी सेट सिद्धांत की शुरुआत की, जिसमें अनिश्चितता और हेजिटन्सी के आगे के स्तरों को शामिल करने के लिए 1965 से लोटफी जादेह के फजी सेटों का विस्तार किया गया। यह सिद्धांत न केवल मेंबरशिप डिग्री बल्कि गैर– मेंबरशिप डिग्री और हेजिटेशन डिग्री का प्रतिनिधित्व करने में सक्षम बनाता है, जिससे वास्तविक दुनिया के परिदृश्यों में अक्सर आने वाली अनिश्चितता और अनिर्णय को कैप्चरिंग किया जाता है। इस अस्पष्ट तर्क का उपयोग करते हुए कई शोधकर्ताओं (घोष एट अल, 2016य द्विवेदी एट अल, 2023) ने भविष्य के मूल्यों की प्रभावी भविष्यवाणी के लिए कई मॉडल विकसित किए। त्सेंग एट अल (2001) ने एक फजी अरिमा (फरिमा) मॉडल विकसित किया, जिसमें अरिमा (p, d, q) मॉडल को फजी रिग्रेशन मॉडल के साथ मिला दिया गया। यह एक्सटेंशन अंतराल मापदंडों के साथ अंतराल मॉडल को समायोजित करता है, और फरिमा मॉडल को भविष्य के मूल्यों के लिए संभाव्यता वितरण प्रदान करने के लिए तैयार किया गया है। लाह एट अल. (2022) ने अरिमा का उपयोग करते हुए और डेटा प्रीप्रोसेसिंग के दौरान फजी डेटा तैयारी तकनीकों को शामिल करते हुए एक तापमान पूर्वानुमान दृष्टिकोण पेश किया। कार्यप्रणाली में फजी त्रिकोण बनाने, फजी डेटा को प्रभावी ढंग से प्रबंधित करने के लिए मानक विचलन विधियों को नियोजित करना शामिल था।

इस शोध में, एक इन्टूइशनिस्टक फजी आधारित ऑटोरेग्रेसिव इंटीग्रेटेड मूविंग एवरेज (फरिमा) मॉडल प्रस्तावित किया गया है जो कृषि वस्तुओं की कीमतों के पूर्वानुमान में अनिश्चितता और अस्थिरता से निपटने के लिए अरिमा मॉडल के साथ गहरे फजी लॉजिक को एकीकृत करता है। इस लेख का प्राथमिक उद्देश्य भारत के विभिन्न बाजारों में कृषि वस्तुओं की कुल कीमतों का अनुमान लगाने के लिए एक कुशल फजी–आधारित अरिमा मॉडल की पूर्वानुमानित क्षमताओं का मूल्यांकन और तुलना करना है। प्रस्तावित मॉडल के मूल्यांकन के लिए गडग, बेल्लारी, गुलबर्गा, लातूर आदि जैसे विभिन्न बाजारों से कुसुम मूल्य डेटा (आरएस/क्यू) का उपयोग किया जाता है। आइफरिमा मॉडल की तुलना छह सटीकता उपायों का उपयोग करके पारंपरिक अरिमा मॉडल से की गई है।

अरिमा मॉडल

अरिमा (ऑटोरिग्रेसिव इंटीग्रेटेड मूविंग एवरेज) मॉडल, जिसका उपयोग समय श्रृंखला पूर्वानुमान और विश्लेषण में बड़े पैमाने पर किया जाता है, गैर–स्थिर समय श्रृंखला डेटा में पूर्वानुमान में सुधार करने के लिए भिन्नता के साथ–साथ ऑटोरेग्रेसिव (AR) और मूविंग एवरेज (MA) घटकों को जोड़ता है। अरिमा (p, d, q) के रूप में दर्शाया गया है, जहां च ऑटोरेग्रेसिव ऑर्डर का प्रतिनिधित्व करता है, क भिन्न डिग्री को इंगित करता है, और q मूविंग औसत ऑर्डर को दर्शाता है। मॉडल मापदंडों का अनुमान लगाने में ऑटोसहसंबंध और आंशिक ऑटोसहसंबंध विश्लेषण जैसी सांख्यिकीय तकनीकों को नियोजित करना शामिल है। अरिमा मॉडल को दिए गए समीकरण (1) के रूप में दर्शाया जा सकता है।

$$x_t = a+\varphi_1 x_{1-1} + \varphi_2 x_{t-2} + \ldots + \varphi_p x_{t-p} + \varepsilon t - \beta_1 \varepsilon_{t-1} - \beta_2 \varepsilon_{t-2} \ldots - \beta_q \varepsilon_{t-q} \quad (1)$$

जहां, x_t समय $t(\in N)$ पर दिया गया समय श्रृंखला डेटा हैय a स्थिर है; ε_t शून्य माध्य और स्थिर विचरण वाला अवशिष्ट पद है; $\varphi_1 x_{t-1} + \varphi_2 x_{t-2} + \cdots + \varphi_p x_{t-p}$ एर भाग हैय और $\varepsilon_t - \beta_1 \varepsilon_{t-1} - \beta_2 \varepsilon_{t-2} - \ldots \beta_q \varepsilon_{t-q}$ एमए भाग है।

प्रस्तावित फ़रिमा मॉडल

इंट्यूशनिस्टिक फजी आधारित ऑटोरेग्रेसिव इंटीग्रेटेड मूविंग एवरेज (आइफरिमा) मॉडल समय श्रृंखला पूर्वानुमान के लिए पारंपरिक अरिमा मॉडल के साथ गहरे फजी लॉजिक को जोड़ता है। यह एकीकरण फजी सदस्यता और गैर–सदस्यता कार्यों के माध्यम से डेटा अनिश्चितता से निपटने के लिए फजी सेट सिद्धांत का

उपयोग करता है। यह समामेलन समय श्रृंखला डेटा में जटिल पैटर्न को पकड़ने और वास्तविक दुनिया की स्थितियों में भविष्यवाणियाँ करने के लिए अरिमा मॉडल की पूर्वानुमान क्षमता में सुधार करता है। यह एकीकरण फजी सदस्यता और गैर–सदस्यता कार्यों के माध्यम से डेटा अनिश्चितता से निपटने के लिए फजी सेट सिद्धांत का उपयोग करता है। यह समामेलन समय श्रृंखला डेटा में जटिल पैटर्न को पकड़ने और महत्वपूर्ण अनिश्चितता की विशेषता वाली वास्तविक दुनिया की स्थितियों में भविष्यवाणियाँ करने के लिए अरिमा मॉडल की पूर्वानुमान क्षमता में सुधार करता है। दिया गया समीकरण (2) आइफरिमा मॉडल का प्रतिनिधित्व करता है।

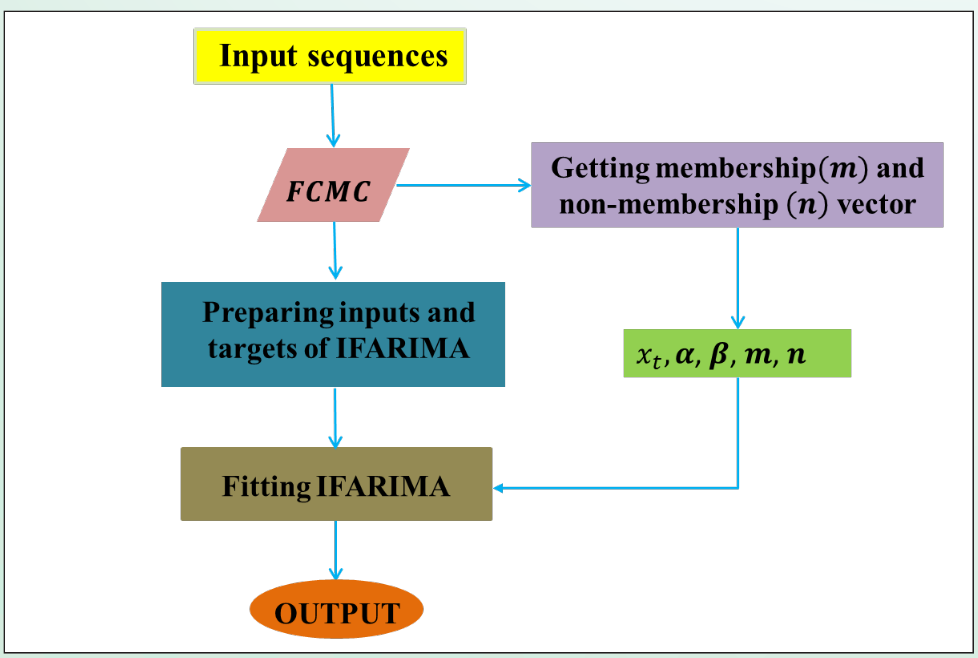
$$x_t = a+\varphi_1 x_{t-1} + \varphi_2 x_{t-2}½+ \cdots +\varphi_p x_{t-p} +\varepsilon_t - \beta_1\varepsilon_{t-1} - \beta_2\varepsilon_{t-2} - \cdots - \beta_{-q}\,\varepsilon_{t-q} + m\alpha + n\beta \;...... \quad (2)$$

जहां, m और n क्रमशः सदस्यता और गैर–सदस्यता वेक्टर हैं, जिनकी गणना फजी C – मतलब क्लस्टरिंग (एफसीएमसी) विधि द्वारा की जाती है और α और β संबंधित गुणांक वेक्टर हैं। फरिमा मॉडल में, एफसीएमसी को लागू करने से पहले क्रिस्प इनपुट अनुक्रमों को फजी अनुक्रमों में बदल दिया जाता है और अनुमानित मूल्यों को प्राप्त करने के लिए सदस्यता और गैर–सदस्यता मूल्यों के साथ अंतराल मूल्यों को प्रस्तावित मॉडल के माध्यम से फिट किया गया है।

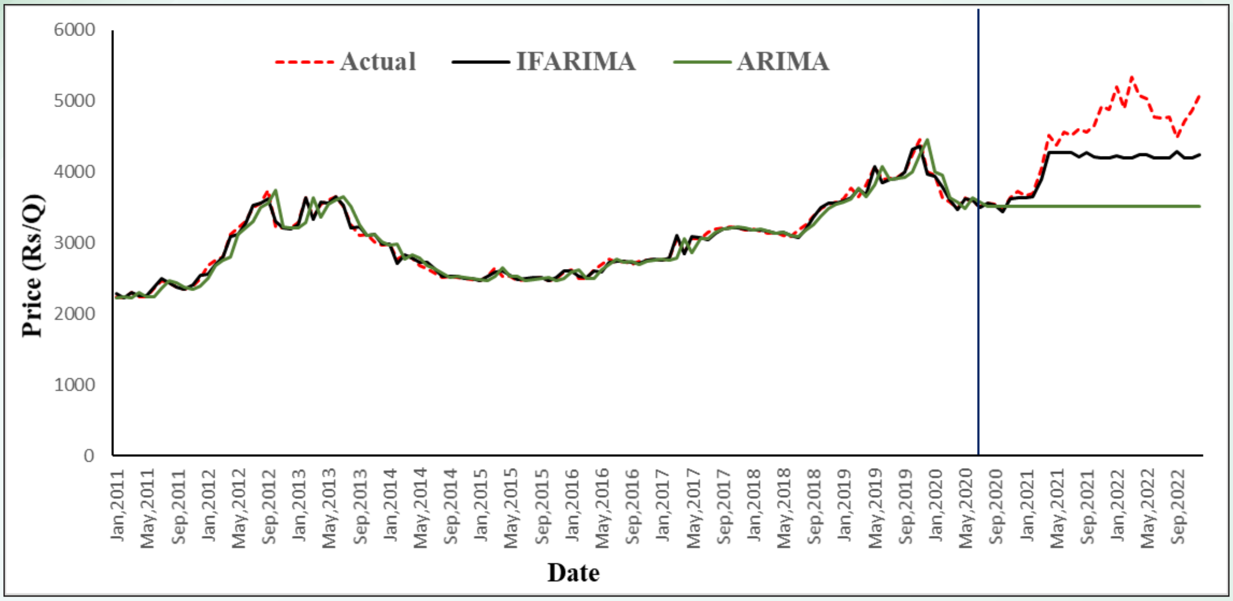
परिणाम और चर्चा

भारत के विभिन्न बाजारों से कुसुम की मासिक थोक कीमतें (रुपये प्रति क्विंटल में) एगमार्कनेट पोर्टल (https://agmarknet.gov.in/) से जनवरी 2010 से दिसंबर 2022 तक एकत्रित की गई हैं। इस डेटासेट में 156 अवलोकन शामिल हैं, जो है 80:20 के अनुपात में प्रशिक्षण और परीक्षण सेटों में विभाजित करना। प्रशिक्षण सेट में पहले 126 महीने के अवलोकन शामिल हैं, जिनका उपयोग मॉडल निर्माण के लिए किया जाता है, जबकि शेष 30 महीने प्रस्तावित मॉडल की सटीकता का आकलन करने के लिए आरक्षित हैं। मॉडल तुलना के लिए छह त्रुटि मैट्रिक्स जैसे रूट मीन स्क्वायर एरर (आरएमएसई), मीन एब्सोल्यूट परसेंटेज एरर (एमएपीई), मीन एब्सोल्यूट एरर (एमएई), एब्सोल्यूट परसेंटेज एरर (एपीई), बायस और R^2 की गणना की जाती है जो नीचे दी गई तालिका में दी गई है।

तालिका 1 से यह स्पष्ट है कि आईफरिमा मॉडल सभी



चित्र 1. प्रस्तावित इफरिमा मॉडल का प्रवाह चार्ट



चित्र 2. आईफरिमा और अरिमा मॉडल के वास्तविक और अनुमानित प्लॉट क्रमश

तालिका 1. कुसुम के लिए प्रस्तावित मॉडलों की प्रदर्शन तुलना।

शुद्ध मैट्रिसेस	प्रशिक्षण सेट		परीक्षण सेट	
	आईफारिमा	अरिमा	आईफारिमा	अरिमा
आरएमई	47.67	128.89	529.86	1108.60
एमएपीई	0.01	0.03	0.09	0.20
एमएई	32.47	88.51	419.13	964.88
एपीई	0.09	0.25	0.96	2.01
बायस	17.45	58.92	394.29	1081.68
R^2	0.99	0.94	0.81	0.71

छह सटीकता उपायों के आधार पर पारंपरिक अरिमा की तुलना में बेहतर प्रदर्शन करता है। इसलिए, प्रस्तावित मॉडल अरिमा मॉडल से बेहतर है क्योंकि यह पारंपरिक मॉडल की तुलना में अनिश्चितता और अस्थिरता को अधिक प्रभावी ढंग से संभाल सकता है।

आईफरिमा और अरिमा मॉडल का वास्तविक प्लॉट और अनुमानित प्लॉट नीचे दिए गए चित्र में दिखाया गया है।

निष्कर्ष

समय श्रृंखला डेटा में अंतर्निहित अनिश्चितता और अप्रत्याशितता सटीक पूर्वानुमान के लिए प्रयास करने वाले शोधकर्ताओं के लिए पर्याप्त बाधाएं पेश करती है। नतीजतन, अरिमा अक्सर ऐसे परिदृश्यों में संतोषजनक परिणाम देने में विफल रहता है। इन चुनौतियों से निपटने के लिए, यह अध्ययन एक अस्पष्ट–आधारित अरिमा अर्थात् आईफरिमा मॉडल पेश करता है जिसे विशेष रूप से अत्यधिक अनिश्चित और अस्थिर समय श्रृंखला डेटा की जटिलताओं को संभालने के लिए तैयार किया गया है। कुसुम मूल्य डेटा का विश्लेषण करके, अध्ययन पूर्वानुमान सटीकता बढ़ाने में आईफरिमा मॉडल की प्रभावकारिता का आकलन करता है।

संदर्भ

अटानासोव, के. टी. (1986)। इन्टूइश्निस्टक फजी सेट. फजी सेट्स ऐन्ड सिस्टम, 20, 87–96।

बॉक्स, जी.ई.पी., जेनकिंस, जी.एम., एवं रीन्सेल जी.सी. (2007)। टाइम सीरीज अनैलिसिसः फॉर्कैस्टिंग ऐन्ड

कन्ट्रोल, थर्ड अडिशन। पियर्सन एजुकेशन, इन्डिया।

द्विवेदी, ए.के., कलियापेरुमल सुब्रमण्यम, यू., कुरुविला, जे., थॉमस, ए., शांति, डी., एवं हल्दोराई, ए. (2023)। टाइम सीरीज डेटा प्रीडिक्शन प्राब्लम अनैलिसिस थ्रू मल्टीलेयरड इन्टूइश्निस्टक फजी सेट। सॉफ्ट कंप्यूटिंग, 27(3), 1663—1671।

घोष, एच., चौधरी, एस., एवं प्रजनेशु। (2016)। एन इम्प्रूव्ड फजी सेट टाइम—सीरीज मेथड ऑफ फॉर्कैस्टिंग बैस्ट आन एल—आर फजी सेट ऐन्ड इट्स ऐप्लकैशन। जर्नल ऑफ एप्लाइड स्टैटिस्टिक्स, 43(6), 1128—1139।

लाह, एम.एस.सी., अरबई, एन., हासिम, वाई.एम. एम., लिन, पी.सी., एवं याकूब, एस.बी. (2022)। फजी—ऑटोरेग्रेसिव इंटीग्रेटेड मूविंग एवरेज (फारिमा) मॉडल टू इम्प्रूव। इन इंट्रेस्टेशनल कान्फरन्स आन सॉफ्ट कंप्यूटिंग ऐन्ड डेटा माइनिंग, 46—55।

एल. ए. जादेह (1965)। फजी सेट. इन्फॉर्मैशन ऐन्ड कन्ट्रोल, 8(3), 338—353।

मित्रा, डी., एवं पॉल, आर.के. (2017)। हाइब्रिड टाइम—सीरीज मॉडल फॉर फॉर्कैस्टिंग ऐग्रिकल्चरल कमाडिटी प्राइस। मॉडल असिस्टिड स्टैटिस्टिक्स ऐन्ड ऐप्लकैशन, 12(3), 255—264।

गीत, प्र.य चिसोम, बी.एस. फॉर्कैस्टिंग एन्रोल्मन्ट विद फजी टाइम सीरीज – पार्ट I. फ़ज़ी सेट्स सिस्टम। 54, 1—9।

त्सेंग, एफ.एम., त्जेंग, जी.एच., यू, एच.सी., एवं युआन, बी.जे. (2001)। फजी अरिमा मॉडल फॉर फॉर्कैस्टिंग द फॉरिन इक्स्चेन्ज मार्किट। फजी सेट ऐन्ड सिस्टम, 118(1), 9—19।

जैव सूचना विज्ञान में बृहत आंकड़ों विश्लेषण

प्रकाश कुमार, रंजीत कुमार पॉल, हिमाद्री शेखर रॉय, मो. येसीन, राजू कुमार, दीपक सिंह, एवं अमृत कुमार पॉल

भा.कृ.अनु.प.–भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली

सार

उच्च परिचरण तकनीकों, जीनोमिक्स, ट्रांसक्रिप्टोमिक्स, और मेटाबोलोमिक्स में प्रगति, जीवविज्ञान जानकारी प्राप्त करने में महत्वपूर्ण भूमिका निभाती हैं। गणनात्मक जीवविज्ञान और जैवसूचना के क्षेत्र में उच्च–परिचरण तकनीकों और अन्य उच्च–परिचरण तकनीकों के आगमन के साथ महत्वपूर्ण विकास का किया है। इसके परिणाामस्वरूप उत्पन्न होने वाले बृहत आंकड़ों ने विश्लेषण के लिए संभावनाएं और चुनौतियां दोनों प्रस्तुत की हैं। बृहत आंकड़ों के विश्लेषण से सार्थक अंशों को प्राप्त करने का आवश्यक माध्यम है । इस लेख में गणनात्मक जीवविज्ञान और जैवसूचना में बृहत आंकड़ों विश्लेषण की वर्तमान स्थिति का अवलोकन प्रदान करेंगे। हम बड़े आंकड़ों विश्लेषण के विभिन्न पहलुओं की चर्चा करेंगे, जिसमें आंकड़ों प्राप्ति, भंडारण, प्रसंस्करण, और विश्लेषण शामिल हैं। हम इस क्षेत्र में बड़े आंकड़ों विश्लेषण की कुछ चुनौतियों और अवसरों को भी उजागर करेंगे। चुनौतियों के बावजूद, बड़े आंकड़ों विश्लेषण महत्वपूर्ण अवसर प्रस्तुत करता है, जैसे कि जीववैज्ञानिक प्रक्रियाओं की समझ में आगे बढ़ाने के लिए कुशल और तेज गणना ऍल्गोरिदम्स का विकास, प्रजनन अनुसंधान और विकास के लिए नए बायोमार्कर्स की पहचान, बीमारी की पूर्वानुमान, और औषधि विकास कार्यक्रम के लिए संभावित औषधि लक्ष्यों की पहचान आदि शामिल है।

मुख्य शब्द: गणनात्मक जीवविज्ञान, बैश शैल–स्क्रिप्ट, हैडूप, पायथन, आर–भाषा, बड़े आंकड़ों, मशीन लर्निंग।

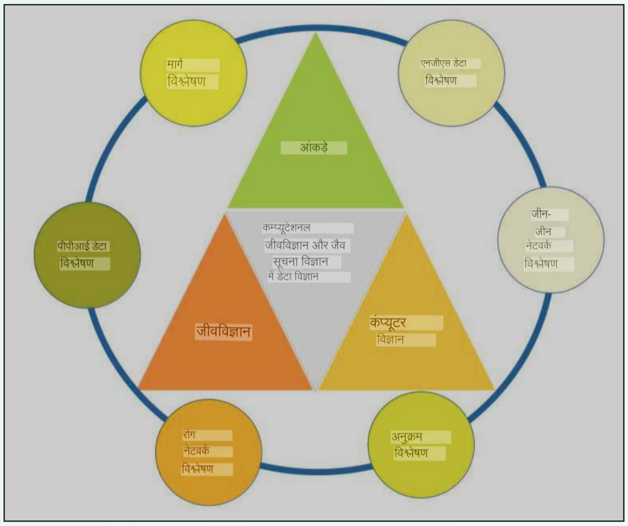
परिचय

उच्च–गति प्रौद्योगिकियों के प्रस्तुतिकरण के कारण, जिसने सूक्ष्म स्तर पर जीवविज्ञान अध्ययन को गति प्रदान की है। जीवविज्ञान और जीनोमिक्स में आंकड़े बहुत तेजी से उत्पन्न है, इस बृहत आंकड़ों से सार्थक अंशों को प्राप्त करने के लिए बड़े आंकड़ों विश्लेषण तकनीकों का उपयोग करना अत्यंत आवश्यक हो गया है। गणनात्मक जीवविज्ञान और जैवसूचना में, बड़े आंकड़ों विश्लेषण का व्यापक रूप से उपयोग जीनोमिक्स, ट्रांसक्रिप्टोमिक्स, प्रोटीओमिक्स, और अन्य उच्च–परिचरण आंकड़ों के विश्लेषण के लिए किया जाता है। प्रस्तुत अध्ययन में गणनात्मक जीवविज्ञान और जैवसूचना में बृहत आंकड़ों विश्लेषण की वर्तमान स्थिति का एक अवलोकन प्रदान करने का उद्देश्य रखता है। यह बृहत आंकड़ों विश्लेषण के विभिन्न पहलुओं, सहित आंकड़ों प्राप्ति, भंडारण, प्रसंस्करण, और विश्लेषण को शामिल करेगा। इस अध्याय में इस क्षेत्र में बृहत आंकड़ों विश्लेषण की कुछ चुनौतियों और अवसरों को भी प्रोत्साहित किया जाएगा।

जैवसूचना में बड़े आंकड़ों के बृहत् उपयोग, चित्र–1 में वर्णित किया गया है।

1. आंकड़ों प्राप्ति और भंडारण

जीनोम, ट्रांसक्रिपटोम, या प्रोटिओम सीक्वेंसिंग से बड़े मात्रा में आंकड़ों बायोइन्फॉर्मेटिक्स में उपयोग हो रहा है। आंकड़ों प्राप्ति और भंडारण गणनात्मक जीवविज्ञान और जैवसूचना में बड़े आंकड़ों विश्लेषण के लिए महत्वपूर्ण घटक हैं। हाई–थ्रूपुट सीक्वेंसिंग विधियों के आगमन के बाद आनुवांछित संख्या में जीनेटिक आंकड़ों तेजी से बढ़ गया है। उसी तरह, हाई–थ्रूपुट प्रोटियोमिक्स



चित्र 1. जैव सूचना विज्ञान में प्रयुक्त बृहत आंकड़ों की भूमिका का अवलोकन

और ट्रांसक्रिप्टोमिक्स विधियों के आगमन ने आंकड़ों उत्पन्न करने में वृद्धि की है। इस विशाल आंकड़ों के वॉल्यूम का सामना करने के लिए, बायोलॉजिकल आंकड़ों को संग्रहित करने और आदान–प्रदान करने के लिए विभिन्न आंकड़ोंबेस स्थापित किए गए हैं। सबसे लोकप्रिय आंकड़ोंबेस में राष्ट्रीय जैव प्रौद्योगिकी सूचना केंद्र (NCBI), यूरोपीय बायोइन्फॉर्मेटिक्स संस्थान (EBI), और जापान डीएनए डाटा बैंक (DDBJ) शामिल हैं। इन डेटाबेसों के अलावा, जैसे कि प्रोटीन सीक्वेंस और कार्यात्मक जानकारी के लिए यूनिप्रोट, वर्टीब्रेट जीनोम सीक्वेंस विविधता और प्राकृतिक नियामक अध्ययन के लिए एन्सेम्बल, मानव जीनोम में कार्यात्मक तत्वों के लिए एंसोड, रेफसेक, ऑनलाइन मेंडेलियन इन्हेरिटेंस इन मैन (OMIM) और अन्य भी महत्वपूर्ण हैं।

2. आंकड़ों प्रसंस्करण और विश्लेषण

गणनात्मक जीवविज्ञान और जैवसूचना में बड़े आंकड़ों का प्रसंस्करण और विश्लेषण कई कदमों को शामिल करता है, जिसमें गुणवत्ता नियंत्रण, पूर्वप्रोसेसिंग, असेम्बली, मैपिंग, समरेखितकरण और सांख्यिकीय विश्लेषण शामिल हैं। इस क्षेत्र में बड़े आंकड़ों विश्लेषण के लिए कुछ सबसे आम उपयोग किए जाने वाले उपकरण और तकनीकों में निम्नलिखित शामिल हैं।

- **गुणवत्ता नियंत्रण:** आंकड़ों की विश्वसनीयता और सटीकता की सुनिश्चित करने के लिए गुणवत्ता नियंत्रण अत्यंत महत्वपूर्ण है। FastQC और Trimmomatic जैसे उपकरण गुणवत्ता नियंत्रण और आंकड़ों फिल्टरिंग के लिए उपयोग होते हैं।
- **पूर्वप्रोसेसिंग:** पूर्वप्रोसेसिंग में आंकड़ों को डाउनस्ट्रीम विश्लेषण के लिए तैयार किया जाता है। इसमें आंकड़ों समानायन, विशेषता चयन, और आयाममूल्य संक्षेप शामिल हैं। DESeq2 और EdgeR जैसे उपकरण RNA-seq आंकड़ों के पूर्वप्रोसेसिंग के लिए सामान्यतः उपयोग होते हैं।
- **समरेखितकरण:** समरेखितकरण रीड्स को संदर्भ जीनोम या ट्रांसक्रिपटोम के साथ मैप करने का प्रक्रिया है। यह हाई–थ्रूपुट सीक्वेंसिंग आंकड़ों के विश्लेषण में एक महत्वपूर्ण कदम है। Bowtie, HISAT2, और STAR जैसे लोकप्रिय समरेखितकरण उपकरण हैं।
- **सांख्यिकीय विश्लेषण:** सांख्यिकीय विश्लेषण का उपयोग विभिन्नता प्रकट करने, म्युटेशन्स, या अन्य जीनोमिक विविधताओं की पहचान के लिए किया जाता है। DESeq2, Limma, और edgeR जैसे लोकप्रिय सांख्यिकीय विश्लेषण उपकरण हैं।

3. जैवसूचना में बड़े आंकड़ों विश्लेषण को संचालित करने के लिए उपकरण

लिनक्स (Linux) शैल स्क्रिप्टः लिनक्स शैल स्क्रिप्ट्स को गणनात्मक जीवविज्ञान और जैवसूचना में बड़े आंकड़ों विश्लेषण कार्यों को संचालित करने के लिए सामान्यतः इस्तेमाल किया जाता है। यहां कुछ मौखिक कदम हैं जो एक सामान्य शैल स्क्रिप्ट में शामिल हो सकते हैं:

पायथन (Python): पायथन एक बहुपरकारी प्रोग्रामिंग भाषा है जिसे गणनात्मक जीवविज्ञान और जैवसूचना में व्यापकता से इस्तेमाल किया जाता है। इसमें NumPy, Pandas, और Biopython जैसे लाइब्रेरीज हैं जो आंकड़ों मैनिपुलेशन, विश्लेषण, और दृश्यीकरण में मदद करते हैं।

आर–भाषा (R-language): आर एक सांख्यिकी प्रोग्रामिंग भाषा है जिसे आंकड़ों विश्लेषण और दृश्यीकरण के लिए सामान्यतः इस्तेमाल किया जाता है। इसमें Bioconductor जैसे कई पैकेजेज हैं जो जैवसूचना के लिए डिजाइन किए गए हैं।

बायोकॉन्डक्टर (Bioconductor): बायोकॉन्डक्टर एक ओपन–सोर्स परियोजना है जो हाई–थ्रूपुट जेनोमिक आंकड़ों के विश्लेषण और समझाने के लिए टूल प्रदान करता है। यह विशेषतः आर प्रोग्रामिंग भाषा के लिए डिजाइन किया गया है।

BLAST (Basic Local Alignment Search Tool): ठस्।ैज एक बायोइन्फॉर्मेटिक्स टूल है जिसका उपयोग जैव शृंग की तुलना के लिए किया जाता है। इससे बड़े आंकड़ोंबेस में समानांतर शृंगों की पहचान में मदद होती है।

SAMtools: SAMtools हाई–थ्रूपुट सीक्वेंसिंग आंकड़ों को सैम (सीक्वेंस अलाइनमेंट/मैप) फॉर्मेट में संवाद करने और प्रसंस्कृत करने के लिए उपयोगी उपकरण प्रदान करता है।

4. लिनक्स शैल स्क्रिप्ट्स

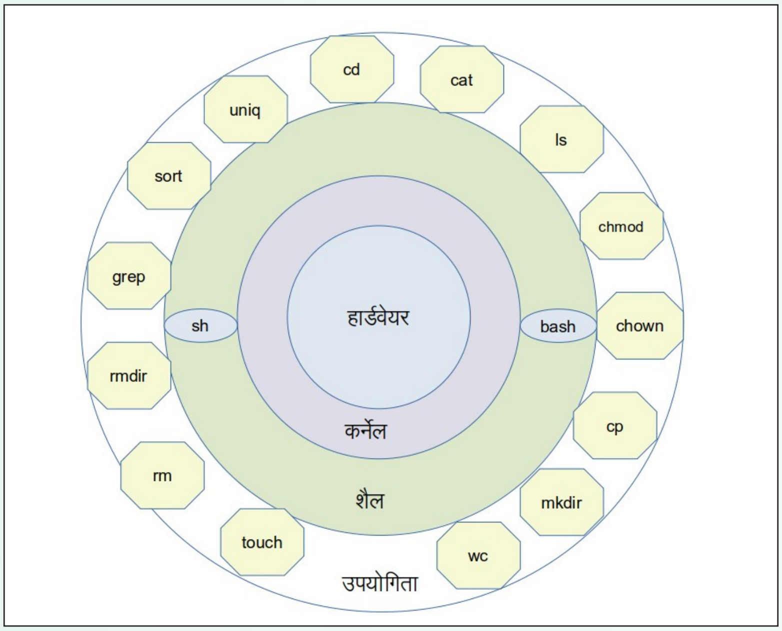
लिनक्स शैल स्क्रिप्ट्स का उपयोग बृहत आंकड़ों विश्लेषण कार्यों को संचालित करने के लिए किया जाता है जो गणनात्मक जीवविज्ञान और जैवसूचना में किए जा सकते हैं। ये शैल स्क्रिप्ट्स आंकड़ों प्रोसेसिंग, फाइल संचार, और अन्य कार्यों को सरल बनाने के लिए उपयोग होते हैं। कुछ आम शैल स्क्रिप्ट्स, उसके उपयोग एवं लिखने का तरीके निचे दी गई है हैं:

डाटा प्राप्तिः

```
# फाइल से आंकड़ों पढ़ने के लिए
cat filename-txt
# डाटा डायरेक्टरी पढ़ने के लिए
ls /path/to/directory
आंकड़ों संरचनाः
# डाटा को सॉर्ट करने के लिए
sort data-txt
# डाटा को फिल्टर करने के लिए
grep keyword data-txt
प्रक्रियाः
# एक स्क्रिप्ट द्वारा आंकड़ों प्रसंस्कृत करें
# उदाहरण के लिएः sh script-sh
सांख्यिकीय विश्लेषणः
# नुमेरिक आंकड़ों का योग
awk {sum += $1} END {print sum}' data-txt
# फाइल में पंक्तियों की गणना
wc -l filename-txt
```

जैवसूचना में बृहत आंकड़ों विश्लेषण के लिए एक सामान्य शैल स्क्रिप्ट में कई कदम हो सकते हैं, जो पूर्वानुमान और विश्लेषण की लगभगी प्रक्रिया को संचालित कर सकते हैं। निर्दिष्ट आंकड़ों का विश्लेषण और विश्लेषण के उद्देश्यों पर निर्भर करता है कि कौन से कदम और आदेश उपयोग किए जाएंगे, लेकिन ऊपर बताए गए सामान्य प्रक्रिया से ऐसे एक स्क्रिप्ट का डिजाइन करने के लिए एक उपयुक्त रूपरेखा प्रदान करता है।

यह स्क्रिप्ट बड़े आंकड़ों को संग्रहित करने, प्रसंस्कृत करने, और विश्लेषण करने के लिए विशेषज्ञ सॉफ्टवेयर



चित्र 2. बायोइन्फॉर्मेटिक्स और गणनात्मक जीवविज्ञान में यूबंटू शैल स्क्रिप्ट का उपयोग क्षेत्र का अवलोकन।

उपकरणों का उपयोग कर सकता है। उपयोगकर्ता द्वारा Ubuntu शैल स्क्रिप्ट का बड़े आंकड़ों बायोइन्फॉर्मेटिक्स और गणनात्मक जीवविज्ञान में विभिन्न उपयोग चित्र–2 में विवरणित हैं।

उदाहरण के लिए टर्मिनल में निम्नलिखित तरीके से आप ''bowtie'' को स्थापित कर सकते हैं:

sudo apt&get install bowtie

"sudo" कमांड एडमिनिस्ट्रेटिव अधिकारों के साथ कमांड को चलाने के लिए होता है, तो इसे बिना सावधानी के न चलाएं। उदाहरणतः RefSeq प्रोटीन्स आंकड़ों को डाउनलोड और प्रीप्रोसेस करने के लिए, आप निम्नलिखित तरीके का पालन कर सकते हैं:

आप एनसीबीआई वेबसाइट से मानव के लिए RefSeq प्रोटीन्स आंकड़ों को डाउनलोड कर सकते हैं। आमतौर पर आपको एनसीबीआई एंट्रेज एपीआई या एफटीपी सेवा का उपयोग करके आंकड़ों डाउनलोड करने की आवश्यकता होगी। यहां आपको फाइलें डाउनलोड करने के लिए तरीके दिए गए हैं:

wget https://ftp-ncbi-nlm-nih-gov/refseq/H_sapiens/annotation/GRCh38_latest/refseq_identifiers/GRCh38_latest_protein-faa-gz

tar &zÛvf GRCh38_latest_protein-faa-gz

आवश्यक फाइलों को डाउनलोड करने योग्य लिंक लिखें और उन्हें डाउनलोड करें

5. आंकड़ों को प्रीप्रोसेस

प्रीप्रोसेसिंग चरण आपकी विशेष आवश्यकताओं पर निर्भर कर सकते हैं, लेकिन सामान्य कार्यों में शामिल हो सकते हैं:

डाउनलोड की गई फाइलों को पार्स करके आवश्यक जानकारी निकालना।

अनचाहे आंकड़ों या रिकॉर्ड को फिल्टर करना।

आंकड़ों को आपके विश्लेषण या अनुप्रयोग के अनुसार सूचीबद्ध करना।

प्रोसेस किए गए आंकड़ों को सहेजें:

जब आप आंकड़ों को प्रीप्रोसेस कर लेते हैं, तो आप इसे अपने विश्लेषण या अनुप्रयोग के लिए उपयुक्त फॉर्मेट में सहेज सकते हैं। यह CSV फाइल, आंकड़ोंबेस, या

आपके लिए किसी भी अन्य फॉर्मेट हो सकता है।

यहां एक बुनियादी उदाहरण है कि आप पायथन का उपयोग करके इन चरणों को पूरा कर सकते हैं:

```
import pandas as pd
```

 # डाउनलोड किए गए आंकड़ों को पढ़ें (यह फाइल FASTA प्रारूप में है माना गया है)

```
def read_fasta=file_path):
    sequences = []
with open(file_path] 'r') as f:
    sequence = ''
        for line in f:
            if line-startswith('>'):
                if sequence:
                    sequences-append(sequence)
                sequence = {header}: line-strip()[1%]]
'sequence': ''}
        else:
                sequence['sequence'] += line-strip()
        if sequence:
        sequences-append(sequence)
    return sequences
```

अपने डाउनलोड किए गए FASTA फाइल के पथ के साथ 'file_path' को बदलें

data = read_fasta('path_to_downloaded_file-fasta')

आंकड़ों को DataFrame में परिवर्तित करें (आगे की प्रोसेसिंग या विश्लेषण के लिए)

df = pd-DataFrame(data)

प्रीप्रोसेसिंग चरण DataFrame 'df' पर किया जा सकता है। जिसके लिए pandas लाइब्रेरी का उपयोग किया जाता है।

प्रोसेस किए गए आंकड़ों को CSV फाइल में सहेजें

df-to_csv('processed_data-csv'] index=False)

#तालिका1 में दिया गया उदाहरण, जो जीन अभिव्यक्ति आंकड़ों फाइल की संरचना को प्रकट करता है:

तालिका 1: जीन अभिव्यक्ति आंकड़ों फाइल

सैंपल आईडी	जीन प्रतीक	जीन आईडी	अभिव्यक्ति मूल्य
S1	ABC	123456	5.6
S2	DEF	789012	3.2
S3	GHI	345678	6.8
S4	JKL	901234	2.5

उदाहरण के लिए

यहाँ उपरोक्त उदाहरण में:

''सैंपल आईडी'' सैंपल को पहचानने के लिए एक अद्वितीय पहचानकर्ता है।

''जीन प्रतीक'' प्रत्येक जीन के आधिकारिक प्रतीक को दर्शाता है।

''जीन आईडी'' प्रत्येक जीन का आधिकारिक पहचानकर्ता है।

''अभिव्यक्ति मूल्य'' प्रत्येक सैंपल में प्रत्येक जीन की अभिव्यक्ति मूल्य को दर्शाता है।

यह फाइल विभिन्न जीनों के अभिव्यक्ति स्तर को विभिन्न जीवविज्ञानीय सैंपलों में प्रकट करने के लिए उपयोग होती है।

फिर इस आंकड़ों पर दो प्रकार के विश्लेषण किया जा सकता है। पहला विश्लेषण प्रत्येक जीन के अभिव्यक्ति स्तरों की माध्य और वेरियंस की गणना करना, साथ ही यह पहचानना कि किस जीन के अभिव्यक्ति स्तर दो समूहों के सैंपलों के बीच में सार्थक रूप से भिन्न हैं। यह उपयुक्त आंकड़ों को बाहर निकालने और विश्लेषण करने के लिए tail, awk, cut, और system आदि कमांड का उपयोग किया जा सकता है।

तालिका 2 जो एक FASTQ फाइल का उदाहरण दर्शाता है और इसे ''sequences-fastq'' नामक फाइल में सहेजा गया है:

तालिका 2: FASTQ फाइल का उदाहरण

```
@SEQ_ID_1
GATTTGGGGTTCAAAGCAGTATCGATCAAATAGTAAATCCATTTGTTCAACTCA
CAGTTT
+
!"*((((***+))%%%++)(%%%%).1***-
+*"))**55CCF>>>>>>CCCCCCC65%%%%).1***-+*%%%
@SEQ_ID_2
GGGCAATGTCTGCTCTTGATGGCCAGGGAAGCGCCACCTGACTGGATGGTGT
TGTTCTTG
+
>>>>>!"*((((***+))%%%++)(%%%%).1***-+*"))**55CCF>>>>>>CCCCCCC65***-
+*"))**55CC
```

यहाँ, ''@'' चिह्न से शुरू होने वाली प्रत्येक पंक्ति एक सैंपल को निर्दिष्ट करती है। इसके बाद का पाठ वास्तविक DNA सिक्वेंस को दर्शाता है। ''+'' चिह्न के बाद का पाठ क्वालिटी स्कोरिंग के लिए है।

#फाइल FASTQ को FASTA में रूपांतरित करने के लिए आप यूबंटू टर्मिनल में निम्नलिखित कमांड उपयोग कर सकते हैं:

cat fastq_file-fq | paste & & & & | cut &f 1]2 | sed ‘s/‘@/>/‘ | tr ‘‘\t‘‘‘‘\n‘‘> fasta_file_name-fa

or

awk ‘NR:4==1 {print ‘‘>‘‘+1} NR:4==2 {print}‘ sequences-fastq > sequences-fasta

यह आपके ''sequences-fastq'' फाइल को ''sequences-fasta'' में रूपांतरित करेगा। यह कमांड AWK नामक टेक्स्ट प्रोसेसिंग उपकरण का उपयोग करता है। यह प्रत्येक FASTQ रेकर्ड की पहली पंक्ति को FASTA फाइल में अनुकरण करता है और दूसरी पंक्ति को प्रिंट करता है। तालिका 3, जो एक FASTA फाइल का उदाहरण दर्शाता है और इसे ''sequences-fasta'' नामक फाइल में सहेजा गया है:

6. बायोइंफोर्मेटिक्स और गणनात्मक जीवविज्ञान में R-साफ्टवेयर की भूमिका

R–सॉफ्टवेयर बड़े आंकड़ों जैवसूचना और गणनात्मक जीवविज्ञान में विशेष भूमिका निभाता है। यहाँ कुछ मुख्य कारण हैं:

सामर्थ्यपूर्ण पैकेजेस: R के ग्राफिक्स और आंकड़े प्रसंस्करण के लिए विशेष पैकेजेस उपलब्ध होते हैं, जो बायोलॉजिकल आंकड़ोंसेट्स के विस्तृत विश्लेषण और विजुअलाइजेशन को सरल और समर्थन करते हैं।

सांख्यिकीय विश्लेषण: R विशेषज्ञात्मक सांख्यिकीय

तालिका 3: FASTA फाइल का उदाहरण

```
>Seq1

ATCATGATCGGTGGTTTCGGAAACTGACTAGTCCCACTCATAAT
>Seq2
GGTAGGTACCGCCCTAAGNCTCCTAATCCGAGCAGACTTTGACCCAGCAGGA
GGAGGAGACCCAGTACTATACCAGCACCTATTCTGATTCTTAAAAAA
```

विश्लेषण के लिए विशेष पैकेजेस प्रदान करता है जो जीनोमिक्स और प्रोटीमिक्स आंकड़ों के साथ काम करते हैं।

मशीन लर्निंग: R मशीन लर्निंग एल्गोरिथ्मों को अनुमानित मॉडलिंग, डाटा क्लस्टरिंग और विभाजन के लिए उपयोग करता है, जिससे बड़े आकार के बायोलॉजिकल आंकड़ोंसेट्स का विश्लेषण किया जा सकता है।

विजुअलाइजेशन: R विभिन्न विजुअलाइजेशन टूल्स प्रदान करता है जो जीनोमिक्स, प्रोटीमिक्स, और अन्य बायोलॉजिकल आंकड़ों को ग्राफिक्स के माध्यम से समझने में मदद करते हैं।

7. पायथन प्रोग्रामिंग भाषा

आज का युग पायथन का है, जिसका उपयोग कई बड़ी कंपनियों जैसे Google, टास्क स्वचालन के लिए, एमेजॅन, टास्क स्वचालन और मशीन लर्निंग के लिए, फेसबुक, नेटफ्लिक्स, और इंस्टाग्राम में उपयोग किया जाता है। पायथन एक सामान्य उद्देश्यक एक प्रोग्रामिंग भाषा है जो व्यापक रूप से जियोमिक्स में बड़े पैमाने पर आंकड़ों विश्लेषण के लिए प्रयोग किया जाता है। पायथन विभिन्न लाइब्रेरी प्रदान करता है, जैसे Pandas, Numpy, और Scikit–learn, जो आंकड़ों मेनिपुलेशन और मशीन लर्निंग के लिए व्यापक रूप से प्रयोग किए जाते हैं। गणनात्मक जीवविज्ञान और बायोइंफोर्मेटिक्स में बड़े आंकड़ोंसेट को संभालने के लिए विशेषित सॉफ्टवेयर उपकरण और लाइब्रेरी का उपयोग करना सामान्य होता है।

पायथन लाइब्रेरी उदाहरण:

Pandas (पांडास): आंकड़ों विश्लेषण और मैनिपुलेशन के लिए।

NumPy (नम्पाई): अंकित गणित के लिए, विशेष रूप से मान्यांकन, सरणियों, और टेन्सर के साथ काम करने के लिए।

Matplotlib (मैटप्लॉटलिब): आंकड़ों को विजुअलाइज करने के लिए।

Scikit-learn (साइकिट–लर्न): मशीन लर्निंग और आंकड़ों मॉडेलिंग के लिए।# डाउनलोड किए गए आंकड़ों को पढ़ें (यह फाइल FASTA प्रारूप में है माना गया है)

TensorFlow (टेंसरफ्लो): डीप लर्निंग के लिए एक प्रमुख पुस्तकालय।

Keras (केरस): उच्च स्तरीय डीप लर्निंग लाइब्रेरी।

PyTorch (पायटॉर्च): डीप लर्निंग के लिए एक और प्रमुख पुस्तकालय।

NLTK (एनएलटीके): प्राकृतिक भाषा प्रोसेसिंग के लिए।

Beautiful Soup (ब्यूटिफुल सूप): वेब स्क्रैपिंग और आंकड़ों पार्सिंग के लिए।

Django (ड्जांगो): वेब एप्लिकेशन डेवलपमेंट के लिए एक प्रमुख फ्रेमवर्क।

यहाँ एक उदाहरण कोड है जो इन उपकरणों का उपयोग करता है ताकि बड़े आंकड़ों विश्लेषण किया जा सके:

1. जीन अभिव्यक्ति आंकड़ों

```
# आवश्यक पुस्तकालयों को आयात करें
import numpy as np
import pandas as pd
from Bio import SeqIO
from scipy-stats import ttest_ind
# आंकड़ों लोड करें
gene_expression_data = pd-read_csv('gene_expression_data-csv')
fasta_sequences = SeqIO-parse('sequences-fasta'] 'fasta')
# जीन अभिव्यक्ति विश्लेषण करें
mean_expression = gene_expression_data-mean()
variance_expression = gene_expression_data-var()
```

significant_genes = gene_eÛpression_data-loc[:] (test_ind(gene_expression_data-iloc[:] :10]] gene_eÛpression_data-iloc[:] 10:]] equal_var=False)[1] < 0-05)]

अनुक्रम विश्लेषण करें

sequence_lengths = [len(sequence-seq) वित sequence पद fasta_sequences,

mean_length = np-mean(sequence_lengths)

max_length = np-max(ssssequence_lengths)

आउटपुट परिणाम

print('माध्य अभिव्यक्तिः', mean_expression)

print('विचरण अभिव्यक्तिः', variance_expression)

print('महत्वपूर्ण जीनः', significant_genes)

print('माध्य अनुक्रम लंबाईः', mean_length)

print('अधिकतम अनुक्रम लंबाईः', max_length)

इस कोड में, हम पांडास (Pandas), नम्पाई (Numpy), और साइकिट–लर्न (Scikit-learn) का उपयोग करके एक आंकड़ों सेट को लोड करते हैं, उसे प्रसंस्कृत करते हैं, और एक रैंडम फॉरेस्ट क्लासिफायर मॉडल को प्रशिक्षित करने के लिए उपयोग करते हैं। अंत में, हम मॉडल की एक्यूरेसी को मूल्यांकन करते हैं।

मशीन लर्निंग

import pandas as pd

import numpy as np

from sklearn-model_selection import train_test_split

from sklearn-ensemble import RandomForestClassifier

from sklearn-metrics import accuracy_score

आंकड़ों लोड करें

data = pd-read_csv('आंकड़ोंसेट.csv')

आंकड़ों शैप की जाँच करें

print(''आंकड़ों आकारः'', data-shape)

डटा संगठन और प्रसंस्करण

X = data-drop('लेबल', axis=1)

y = data[लेबल]

ट्रेन और टेस्ट आंकड़ों में विभाजन

X_train, X_test, y_train, y_test = train_test_split(X, y] test_size=0-2, random_state=42)

रैंडम वन फोरेस्ट क्लासिफायर तैयार करें

model = RandomForestClassifier()

मॉडल को प्रशिक्षित करें

model-fit(X_train, y_train)

परिणाम का मूल्यांकन

y_pred = model-predict(X_test)

accuracy = ccuracy_score(y_test, y_pred)

print(''एक्यूरेसीः'', accuracy)

इस कोड में, हम पांडास (Pandas), नम्पाई (Numpy), और साइकिट–लर्न (Scikit–learn) का उपयोग करके एक आंकड़ों सेट को लोड करते हैं, उसे प्रसंस्कृत करते हैं, और एक रैंडम फॉरेस्ट क्लासिफायर मॉडल को प्रशिक्षित करने के लिए उपयोग करते हैं। अंत में, हम मॉडल की एक्यूरेसी को मूल्यांकन करते हैं।

साहित्य की समीक्षा

बड़े आंकड़ों ने बृहद स्तर पर बायोइंफोर्मेटिक्स के क्षेत्र को क्रांति ला दी है, जिसके लिए बड़े पैमाने पर जीवविज्ञानी आंकड़ों सेट को प्रसंस्करण और विश्लेषण के लिए शक्तिशाली उपकरण प्रदान किए गए हैं। इस साहित्य समीक्षा में बड़े आंकड़ों के आवेदन पर ध्यान केंद्रित है, जो जीनोमिक्स, सिस्टम्स बायोलॉजी, गणनात्मक जीवविज्ञान, और व्यक्तिगत चिकित्सा जैसे विषयों को शामिल करता है। बड़े आंकड़ों गणनात्मक जीवविज्ञान के क्षेत्र में प्रसिद्ध हो गया है, जो आंकड़ों प्रसंस्करण और विश्लेषण के लिए शक्तिशाली उपकरण प्रदान करता है। यह साहित्य समीक्षा हाल के उन्नतियों और बड़े आंकड़ों विश्लेषण के आवेदन में आने वाली चुनौतियों को शामिल करती है।

कई अध्ययनों में बड़े आंकड़ों का उपयोग करने के

लिए गणनात्मक जीवविज्ञान में चुनौतियों और अवसरों पर चर्चा की गई है। Ching et l- [1] ने जीवविज्ञान और चिकित्सा में डीप लर्निंग के लिए अवसरों और बाधाओं पर चर्चा की, Yang et al- [2] ने बड़े आंकड़ों विश्लेषण में एल्गोरिदम, विधियों, और अनुप्रयोगों की समीक्षा की, जबकि Yang et al- [3] ने बायोलॉजी में बड़े आंकड़ों विश्लेषण की समीक्षा की एवं कंप्यूटेशनल चुनौतियों पर जोर दिया। मशीन लर्निंग और कृत्रिम बुद्धिमत्ता का एकीकरण गणनात्मक जीवविज्ञान में बड़े आंकड़ों विश्लेषण के आवेदन में महत्वपूर्ण रहा है।

बड़े आंकड़ों विश्लेषण ने जटिल जीवविज्ञानी तंत्रों को समझने में भी एक महत्वपूर्ण भूमिका निभाई है। Kim और Kim [4] ने प्रणालियों बायोलॉजी में बड़े आंकड़ों विश्लेषण के अनुप्रयोग पर चर्चा की, जिसमें जटिल जीवविज्ञानी तंत्रों को बेहतर रूप से समझने के लिए विभिन्न आंकड़ों प्रकारों के एकीकरण पर जोर दिया गया है। Ljosa et al. [5] ने स्मॉल–मोलेक्यूल उपचार के कोशिकीय रूपी आकारिक प्रतिक्रियाओं के चित्र–आधारित प्रोफाइलिंग के लिए विभिन्न विधियों की तुलना की। बड़े आंकड़ों विश्लेषण, व्यक्तिगत चिकित्सा में महत्वपूर्ण रहा है। Khan et al. [6] ने स्वास्थ्य और बायोइंफोर्मेटिक्स में बड़े आंकड़ों विश्लेषण की समीक्षा की, जिसमें व्यक्तिगत स्वास्थ्य सेवाओं में बड़े आंकड़ों की भूमिका को जोर दिया। Chen et al. [7] ने जीनोमिक्स में बड़े आंकड़ों विश्लेषण की भूमिका को हाइलाइट किया गया है और फिर से Chen et al. [8] ने बड़े आंकड़ों की गणनात्मक जीवविज्ञान में भूमिका की समीक्षा की, व्यक्तिगत चिकित्सा के लिए संभावित उपयोग की चर्चा की गई है। कई अध्ययनों ने जीनोमिक्स में बड़े आंकड़ों को प्रसंस्करण और विश्लेषण के लिए कुशल उपकरणों की आवश्यकता को जोर दिया है। Peng और Tang [9] ने कैंसर अनुसंधान में बड़े आंकड़ों विश्लेषण का उपयोग की चर्चा की। Sharma और Menon [10] ने जीनोमिक्स में बड़े आंकड़ों विश्लेषण के आवेदन की समीक्षा की, जिसमें मशीन लर्निंग और डीप लर्निंग की भूमिका को जोर दिया गया है।

बड़े आंकड़ों को जीव विज्ञान में एकीकरण को भी ध्यान मिला है। Oulas et al. [11] ने बड़े ओमिक्स आंकड़ों का एकीकरण के लिए बेहतर मानव स्वास्थ्य की समीक्षा की। Nobile और Sealfon [12] ने न्यूरोसाइंस में बड़े आंकड़ों और मशीन लर्निंग का उपयोग पर चर्चा की। गणनात्मक जीवविज्ञान में, कई अध्ययनों ने बड़े आंकड़ों विश्लेषण के अनुप्रयोग पर ध्यान केंद्रित किया है। Alqurashi और Mavromatis [13] ने गणनात्मक जीवविज्ञान में बड़े आंकड़ों विश्लेषण का उपयोग की समीक्षा की और गणनात्मक जीवविज्ञान में बड़े आंकड़ों की भूमिका पर चर्चा की। Leung et al. [14] ने जेनोमिक मेडिसिन में मशीन लर्निंग की समीक्षा की, जिसमें कंप्यूटेशनल चुनौतियों और उपलब्ध आंकड़ों सेट्स पर जोर दिया गया। व्यक्तिगत चिकित्सा भी बड़े आंकड़ों विश्लेषण के आवेदन से लाभान्वित हुई है। मशीन लर्निंग एक प्रकार की कृत्रिम बुद्धिमत्ता है जो कंप्यूटरों को सीखने और अनुभव से सुधारने की सक्षमता प्रदान करती है। यह बड़े आंकड़ों के विशाल मात्रा में नमूने और संबंधों की पहचान करने के लिए एल्गोरिदम का उपयोग करती है, और फिर इस ज्ञान का उपयोग करके भविष्यवाणियाँ करने या कार्रवाई लेने में योग्यता प्रदान करती है। पायथन आंकड़ों साइंस के लिए एक लोकप्रिय प्रोग्रामिंग भाषा है और इसमें प्राकृतिक भाषा प्रसंस्करण (एनएलपी), आंकड़ों विजुअलाइजेशन, और नेटवर्क विश्लेषण के लिए कई library हैं। उदाहरणतः Matplotlib, और NetworkX जो विजुअलाइजेशन और ग्राफ सिद्धांत के लिए प्रयोग किए जाते हैं। बड़े आंकड़ों विश्लेषण उपकरणों और तकनीकों के उपयोग पर एक समीक्षा लेख कृषि विकास और कृषि के प्रमुख रुझानों में मूल्यवान अनुशासन प्रदान कर सकती है, किसानों और नीति निर्माताओं को भूमि के उपयोग, फसल प्रबंधन, और कृषि के अन्य महत्वपूर्ण पहलुओं के बारे में बेहतर निर्णय

लेने में मदद करने के लिए की चर्चा और एक समीक्षा लेख बड़े आंकड़ों के उपयोग पर लिखा गया है ।नउंत et al. [15]।

इस अध्ययन में समीक्षित साहित्य ने बायोइंफोर्मेटिक्स में बड़े आंकड़ों विश्लेषण की महत्वपूर्ण भूमिका को हाइलाइट किया है, जो विशाल पैमाने पर जीवविज्ञानी आंकड़ों सेट को प्रसंस्करण और विश्लेषण के लिए शक्तिशाली उपकरण प्रदान करता है। बड़े आंकड़ों के उपयोग ने जीनोमिक्स, प्रणालियों बायोलॉजी, गणनात्मक जीवविज्ञान, और व्यक्तिगत चिकित्सा में हमारे जीवाणु प्रक्रियाओं को समझने में महत्वपूर्ण सुधार किया है और रोग का निदान, उपचार, और रोकथाम में समर्थन करने की क्षमता बढ़ाई है।

चुनौतियाँ और अवसर

गणनात्मक जीवविज्ञान और बायोइनफोर्मेटिक्स में बड़े आंकड़ों का विश्लेषण कई चुनौतियों और अवसरों को प्रस्तुत करता है। कुछ चुनौतियों और अवसरों की विवरण नीचे दिया गया हैः

चुनौतियाँ

आंकड़ों संग्रहण और संरचनाः बड़े आंकड़ों सेट्स को संग्रहित करना और उन्हें संरचित करना एक चुनौती है, विशेषतः जब आंकड़ों की मात्रा बहुत बड़ी हो।

आंकड़ों का प्रसंस्करणः बड़े आंकड़ों सेट्स का प्रसंस्करण, स्पष्टीकरण और साफ–सफाई करना अन्य चुनौती है, जिसमें आंकड़ों की गुणवत्ता और संभावित त्रुटियों का समर्थन किया जाना होता है।

एल्गोरिदम विकासः बड़े आंकड़ों सेट्स के लिए उपयुक्त एल्गोरिदमों का विकास करना एक और महत्वपूर्ण चुनौती है।

संबंधितता का पता लगानाः आंकड़ों के बीच के संबंधों को पता लगाना और प्रासंगिक सूचनाओं का पता लगाना भी एक मुश्किल चुनौती है।

अवसर

नई ज्ञान का अनुसंधानः बड़े आंकड़ों सेट्स का विश्लेषण करके नई ज्ञान और पैटर्न्स का अनुसंधान करने का अवसर होता है, जो बायोलॉजी में महत्वपूर्ण हो सकता है।

नई उपयोग के लिए आंकड़ोंः बड़े आंकड़ों सेट्स को उपयोग करके नए उपाय और तकनीकों का अविष्कार करने का अवसर होता है जो जीवविज्ञान के डोमेन में नवीनीकरण को प्रोत्साहित कर सकता है।

नई उपयोगी उपकरणः बड़े आंकड़ों का विश्लेषण करने के लिए नए और उपयोगी उपकरणों और प्रोग्रामों का विकास करने का अवसर हो सकता है जो जीवविज्ञानी समस्याओं को हल करने में सहायक हो सकते हैं।

निष्कर्ष

आजकल, गणनात्मक जीवविज्ञान और जैवसूचना में बड़े आंकड़ों विश्लेषण एक महत्वपूर्ण अनुसंधान और विकास क्षेत्र है । उच्च–परिचरण प्रौद्योगिकी जीववैज्ञानिक आंकड़ों के विकास को तेजी से बढ़ाती है, इसलिए यह महत्वपूर्ण हो गया है कि इसे इन विशाल और जटिल आंकड़ों को संग्रहित, प्रसंस्कृत, पुनः प्राप्त, सांख्यिकीय विश्लेषण, और दृश्यीकरण करने के लिए नए उपाय, एल्गोरिदम, और उपकरण बनाने की आवश्यकता है। जीनोमिक्स में सिक्वेंसिंग आंकड़ों जैसे विभिन्न प्रकार के बायोलॉजिकल आंकड़ों की बड़ी मात्रा में होने के कारण, हमारे जीववैज्ञानिक प्रक्रियाओं और रोग अनुसंधान के ज्ञान में महत्वपूर्ण सुधार हो सकता है। यह अब भी महत्वपूर्ण चुनौतियों को सामने कर रहा है जो जीवविज्ञानिकों, कंप्यूटर वैज्ञानिकों, सांख्यिकीविदों, और अन्य विशेषज्ञों के बीच बहुविद्यालयी सहयोग की आवश्यकता को प्रामाणिक बनाते हैं। और अधिक अनुसंधान और विकास के साथ, हम गणनात्मक जीवविज्ञान और बायोइनफोर्मेटिक्स के क्षेत्रों में बड़े आंकड़ों विश्लेषण की पूर्ण संभावनाओं को पूरा कर सकते हैं, जिससे इन

क्षेत्रों में नए दृष्टिकोण और सफलताएं प्राप्त हो सकती हैं। बड़े आंकड़ों विश्लेषण गणनात्मक जीवविज्ञान और बायोइनफोर्मेटिक्स का एक महत्वपूर्ण हिस्सा बन गया है। जीनोमिक्स, ट्रांसक्रिप्टोमिक्स और प्रोटियॉमिक्स में इतने सारे सिक्वेंसिंग आंकड़ों उत्पन्न हो रहे हैं, इसलिए इसे कुशल आंकड़ों संग्रहण, संग्रहण, प्रसंस्करण, और विश्लेषण उपकरणों और एल्गोरिदमों का विकास करना महत्वपूर्ण है। हालांकि बड़े आंकड़ों विश्लेषण अनेक चुनौतियों का सामना कराता है, लेकिन यह जीववैज्ञानिक प्रक्रियाओं को समझने में हमारी विशेष तरह से मदद करने का एक अच्छा मौका भी है।

संदर्भ

चिंग टी, हिम्मेलस्टीन डीएस, बोलियू–जोन्स बीके, कालिनिन एए, डू बीटी, वे जीपी, अलीपनाही बी (2018)। जीवविज्ञान और चिकित्सा में गहरी सीखने के अवसर और बाधाएँ। द रॉयल सोसायटी इंटरफेस जर्नल 15(141):20170387।

यांग जे, ज्जांग एक्स, लियू सी (2017)। बायोइनफोर्मेटिक्स में बड़े आंकड़ों विश्लेषणः एल्गोरिदम, विधियाँ, और अनुप्रयोग। इंटरनेशनल जर्नल ऑफ आंकड़ों माइनिंग और बायोइनफोर्मेटिक्स 17(2):105–124।

यांग सीटी, क्रिस्टियानी ई, लिओंग वाईके, चांग जे एस (2023)। बड़े आंकड़ों और मशीन लर्निंग ड्राइवन बायोप्रोसेसिंग – हाल के प्रवृत्तियाँ और महत्वपूर्ण विश्लेषण। बायोरिसोर्स टेक्नोलॉजी 372:128625।

किम जेएच, किम एस (2019)। प्रणालियों बायोलॉजी में बड़े आंकड़ों विश्लेषण। जर्नल ऑफ माइक्रोबायोलॉजी एंड बायोटेक्नोलॉजी 29(2):171–180।

ल्जोसा वी, काइ पीडी, टेर होर्स्ट आर, सोकोलनिकी के एल, जेंकिन्स ईएल, डाया एस, कारपेंटर एई (2013)। छोटे–अणु उपचार के सेलुलर रूप के छवि–आधारित प्रोफाइलिंग के लिए विधियों का तुलनात्मक मुल्यांकन। जर्नल ऑफ बायोमोलेक्युलर स्क्रीनिंग 18(10):1321–1329।

खान एस, खान हू, नजीर एस (2022)। प्रभावी देखभाल और रोग निदान के लिए हेल्थकेयर बड़े आंकड़ों विश्लेषण की व्यवस्थित विश्लेषण। सायं रिपोर्ट्स 12:22377।

चेन वाई, पालकाल एम, जो बी (2017)। जीनोमिक्स में बड़े आंकड़ों विश्लेषण। जर्नल ऑफ बायमेडिकल इनफॉर्मेटिक्स 67:1–3।

चेन वाई, लियू वाई, यु सी (2019)। गणनात्मक जीवविज्ञान में बड़े आंकड़ोंः एक समीक्षा। बायोमेडिसिन एंड फार्माकोथेरेपी 110:524–532।

पेंग वाई, टैंग एच (2016)। कैंसर अनुसंधान में बड़े आंकड़ों विश्लेषण। करंट फार्माकोलॉजी रिपोर्ट्स 2(6):305–313।

शर्मा ए, मेनन आर (2019)। जीनोमिक्स में बड़े आंकड़ों विश्लेषणः एक समीक्षा। जीनोमिक्स 111(1):43–50।

ओलस ए, कराथानासिस एन, लोउलोपी ए, पाव्लोपुलोस जीए, पोइराजी पी, इलिओपोलस आई (2019)। बेहतर मानव स्वास्थ्य के लिए बड़े ओमिक्स आंकड़ों का एकीकरण। हेल्थकेयर में बड़े आंकड़ों विश्लेषण 211–232।

नोबिल मएस, सीलफोन एससी (2017)। न्यूरोसाइंस में बड़े आंकड़ों और मशीन लर्निंग। न्यूरोइम्यून फार्माकोलॉजी जर्नल 12(1):1–2।

अलकुरैशी एम, माव्रोमैटिस सी (2018)। गणनात्मक जीवविज्ञान में बिग आंकड़ों एनालिटिक्सः एक समीक्षा। करंट बायोइन्फॉर्मेटिक्स 13(5):452–462।

लियुंग एमके, डेलोंग ए, अलिपनाही बी, फ्रे बीजे (2019)। जीनोमिक चिकित्सा में मशीन लर्निंगः कम्प्यूटेशनल समस्याओं और आंकड़ों सेट्स की समीक्षा। प्रोसीडिंग्स ऑफ द आईईईई 107(3):512–532।

कुमार पी, कुमार ए, पंवार एस, डैश एस, सिन्हा के, चौधरी वीके, रे एम (2018)। कृषि में बिग आंकड़ों की भूमिका – एक सांख्यिकी परिप्रेक्ष्य। एन्न एग्रीक रिस 39(2):210–215।

विभिन्न स्थितियों में संशोधित प्रसरण विश्लेषण द्वारा आनुवंशिकता का आकलन

रंजित कुमार पॉल, मो. यासीन, हिमाद्री शेखर रॉय, प्रकाश कुमार, अमृत कुमार पॉल एवं एस.पी.सिंह
भा.कृ.अनु.प.–भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली

सारांश

संशोधित प्रसरण विश्लेषण, जो प्रसरण घटकों के प्राचलों पर प्रतिबंध लगाती है, का उपयोग स्वीकार्य सीमा में आनुवंशिकता का अनुमान प्राप्त करने के लिए किया गया है। असंतुलित एवं मानक संदूषण की उपस्थिति में इस पद्धति के प्रदर्शन की तुलना अन्य पारंपरिक तरीकों जैसे प्रसरण विश्लेषण, अधिकतम संभावना (एमएल), अधिकतम संभावना (आरईएमएल) अनुमानक और एमआईवीक्यू (0) के साथ करने के लिए मोंटेकार्लो अध्ययन किया गया था। यह देखा गया कि संशोधित प्रसरण विश्लेषण अनुमानक में विभिन्न प्रतिदर्श आकारों के तहत आनुवंशिकता के सभी प्राचलों के लिए प्रसरण विश्लेषण की तुलना में कम माध्य वर्ग त्रुटि है।

परिचय

पिछले कई दशकों के दौरान यादृच्छिक प्रभाव प्रारूप में प्रसरण घटकों के आकलन की समस्या पर व्यापक ध्यान दिया गया है। प्रसरण विश्लेषण विधि का उपयोग आम तौर पर प्रसरण घटकों और उनके अनुपातों का अनुमान लगाने के लिए किया जाता है। प्रसरण विश्लेषण से प्राप्त अनुमान नकारात्मक हो सकते हैं, ऐसी स्थिति में भी प्रसरण घटक सकारात्मक माने जाते हैं। अनुचित और अस्पष्ट नकारात्मक अनुमानों से बचने के लिए, कई अनुमानकों का सुझाव दिया गया है, उदाहरण के लिए, अधिकतम संभावना (एमएल) अनुमानक (हर्बाक, 1959), प्रतिबंधित अधिकतम संभावना (आरईएमएल) अनुमानक (थॉम्पसन, 1962) जो कि प्राप्त अनुमानकों को नगण्य मानकर प्राप्त किए जाते हैं। एमएल उन मूल्यों के आधार पर प्रसरण घटकों का अनुमान लगाता है जो प्राचल स्पेस पर पूर्ण संभावना फंक्शन को अधिकतम करते हैं और आरईएमएल संभावना को दो भागों में विभाजित करता है, जिनमें से एक निश्चित प्रभावों से मुक्त है और संभावना के केवल उस हिस्से को अधिकतम करता है। एमएल और आरईएमएल की सीमा भिन्नता अलग–अलग है।

घटकों का अनुमान आमतौर पर पुनरावृत्त रूप से प्राप्त किया जाना चाहिए, और आवश्यक कंप्यूटिंग कठिन हो सकती है और शायद बड़े डेटा सेट के लिए असंभव हो सकती है। प्रसरण घटकों के आकलन पर किए गए काफी शोध के बावजूद, बहुत सारे अनुमानक उपलब्ध नहीं हैं, जो गणनात्मक दृष्टिकोण से लागू करना आसान है और साथ ही यह सुनिश्चित करते हैं कि अनुमानकों की अप्राचलिक साहित्य में उपलब्ध है।

खत्री (1998) ने प्राचल पर प्रतिबंध को ध्यान में रखते हुए, प्रसरण घटकों के समाधान प्राप्त करने की एक विधि के रूप में हेंडरसन की प्रसरण पद्धति में संशोधन का सुझाव दिया। इस संशोधित प्रसरण विश्लेषण का अभ्यास करना सरल है और प्राप्त प्रसरण घटकों के अनुमान हमेशा गैर नकारात्मक और अद्वितीय होते हैं। तो, हम स्वीकार्य सीमा में आनुवंशिकता का अनुमान प्राप्त कर सकते हैं। वर्तमान जांच में, इस पद्धति का उपयोग आनुवंशिकता के अनुमान पर असंतुलन के प्रभाव और पैमाने पर संदूषण की भूमिका का अध्ययन करने के लिए किया गया है। डेटा सेट की विभिन्न स्थितियों में प्रसरण

विधि, एमएल, आरईएमएल और एमआईवीक्यू (0) जैसे अन्य पारंपरिक तरीकों के साथ इस विधि के प्रदर्शन की तुलना करने के लिए एक मोंटेकार्लो तुलना की गई थी।

कार्यप्रणाली:

मान लीजिए $(m = (m_1, m_2,..., m_p)'$ कि हेंडरसन की प्रसरण विधि का उपयोग करके प्राप्त $p\times 1$ माध्य वर्ग त्रुटि का m_p प्रतिनिधित्व करने वाला माध्य वर्गों का एक वेक्टर है। माना कि $(E(m) = A\sigma)$ त्रुटि प्रसरण होने के साथ जहाँ $\sigma = (\sigma^2_1, \sigma^2_2,.....,\sigma^2_p$ प्रसरण का वेक्टर है σ^2_p और केवल विशेष डेटा संरचना पर निर्भर करता है और $p\times p$ σ तत्वों के ज्ञात गुणांक का वर्ग मैट्रिक्स है।

प्रसरण विश्लेषण में व्यवहार्य क्षेत्र को परिभाषित करने वाले $\sigma\geq 0$ प्रतिबंध की अनदेखी करते हुए σ रैखिक समीकरणों $A\sigma = m$ का समाधान ढूंढना शामिल है।

$$A = \begin{bmatrix}\lambda & 1\\ 0 & 1\end{bmatrix} \quad \boldsymbol{\sigma} = \begin{bmatrix}\sigma^2_{\mathbf{s}}\\ \sigma^2_e\end{bmatrix} \text{ and } \mathbf{m} = \begin{bmatrix}m_1\\ m_2\end{bmatrix}$$

प्रसरण विधि द्वारा प्राप्त प्रसरण घटकों के अनुमान नकारात्मक हो सकते हैं, क्योंकि एक गैर–नकारात्मक प्राचल की व्याख्या स्पष्ट रूप से एक समस्या है। खत्री (1998) के अनुसार, प्राचल स्पेस पर प्रतिबंध रखते हुए, माध्य वर्गों को कुछ सार्थक अर्थों में उनके अपेक्षित मूल्यों के सबसे करीब लाकर भिन्न घटकों का समाधान प्राप्त करने के लिए संशोधित हेंडरसन विधि पर काम किया जा सकता है।

गणितीय रूप से, इसे एक $\tilde{\boldsymbol{\sigma}}$ समाधान प्राप्त करने के रूप में कहा जा सकता है ताकि $\|\mathbf{A}\sigma - \mathbf{m}\|$ बाधा के विषय को कम किया जा सके, जहां $\mathbf{m}$ माध्य वर्गों का वेक्टर है, σ प्रसरण घटकों का वेक्टर है और A σ तत्वों के ज्ञात गुणांक का मैट्रिक्स है।

$\|x\| = (x'x)^{1/2}$, यूक्लिडियन मानदंड के तहत, $\|\mathbf{A}\sigma - \mathbf{m}\|$ न्यूनतम करने की समस्या $\boldsymbol{\sigma} \geq \mathbf{0}$ का विषय है

$\frac{1}{2}\boldsymbol{\sigma}'\mathbf{A}'\mathbf{A}\boldsymbol{\sigma} - \boldsymbol{\sigma}'\mathbf{A}'\mathbf{m}$ के बराबर है

$\boldsymbol{\sigma} \geq \mathbf{0}$ न्यूनतम करना, के अधीन।

इसे इस प्रकार लिखा जा सकता है, $\frac{1}{2}\boldsymbol{\sigma}'\mathbf{H}\boldsymbol{\sigma} - \boldsymbol{\sigma}'\mathbf{c}$ न्यूनतम करना, $\boldsymbol{\sigma} \geq \mathbf{0}$ अधीन करना

जहां $\mathbf{H = A'A'}$ और $\mathbf{c = A'm}$

जो रैखिक बाधाओं के साथ एक विशिष्ट द्विघात प्रोग्रामिंग समस्या है। चूँकि A एकवचन नहीं है (हेंडरसन की विधि इसे सुनिश्चित करती है), (H) की रैंक = (A'A) की रैंक = (A) की रैंक = परिणामस्वरूप, H सकारात्मक निश्चित है। चूँकि H सकारात्मक निश्चित है, रैखिकता स्थितियों के अधीन गैर–रेखीय बाधित समस्या के स्थिर बिंदुओं की पहचान करने के लिए कुह्न–टकर स्थितियों का उपयोग द्विघात प्रोग्रामिंग समस्या को रैखिक पूरक प्रणाली में परिवर्तित करने के लिए किया जा सकता है। $g(x)$ बाधाओं के साथ किसी भी फंक्शन $f(x)$ को छोटा करने के लिए कुह्न–टकर स्थितियाँ हैं

$$\nabla\mathbf{f(x)} - \mathbf{u}\nabla\mathbf{g(x)} = \mathbf{0} \quad \mathbf{u'g(x)} = \mathbf{0} \quad \mathbf{u} \geq \mathbf{0} \quad \mathbf{g(x)} \geq \mathbf{0}$$

जहां $\mathbf{u}$ लैग्रेंज गुणक है.

यहाँ पर, कुह्न–टकर स्थितियाँ हैं

$$\mathbf{H}\tilde{\boldsymbol{\sigma}} - \mathbf{c} - \mathbf{u} = \mathbf{0} \quad \mathbf{u}'\tilde{\boldsymbol{\sigma}} = \mathbf{0} \quad \mathbf{u} \geq \mathbf{0} \quad \tilde{\boldsymbol{\sigma}} \geq \mathbf{0}$$

इन समीकरणों को इस प्रकार लिखा जा सकता है

$$\mathbf{u} = \mathbf{H}\tilde{\sigma} - \mathbf{c} \quad \mathbf{u}'\tilde{\boldsymbol{\sigma}} = \mathbf{0} \quad \mathbf{u} \geq \mathbf{0} \quad \tilde{\boldsymbol{\sigma}} \geq \mathbf{0}$$

यह एक रैखिक संपूरकता समस्या है, जिसके माध्यम से द्विघातीय समीकरण (QP) समाधान प्राप्त किया जा सकता है।

एसएएस आईएमएल में एलसीपी सबरूटीन का उपयोग समाधान खोजने के लिए इस रैखिक संपूरकता समस्या $\tilde{\boldsymbol{\sigma}}$ को हल करने के लिए किया गया है। इस विधि द्वारा प्राप्त अनुमान गैर–नकारात्मक, अद्वितीय और हेंडरसन के अनुमान के समान होते हैं जब भी हेंडरसन गैर–नकारात्मक होते हैं।

परिणाम और चर्चा

अन्य पारंपरिक तरीकों के साथ संशोधित प्रसरण विश्लेषण के प्रदर्शन की अनुभवजन्य तुलना करने के लिए, आनुवंशिकता के विभिन्न पैरामीट्रिक मानों (h^2 0-1, 0-4, 0-8) के लिए डेटा का अनुकरण किया

गया था। मुख्य प्रसरण इस तरह से लिया गया था कि जिस जनसंख्या से प्रतिदर्श संबंधित है, उसमें आनुवंशिकता के अलग–अलग मान हों। अनुकरण करते समय, विभिन्न पारिवारिक संरचनाओं, असंतुलितता की डिग्री और पैमाने पर संदूषण को भी ध्यान में रखा गया। सिम्युलेटेड डेटा को विचरण घटकों के आकलन और बाद में का h^2 अनुमान लगाने के लिए प्रसरण विधि, संशोधित प्रसरण विधि, एमएल, आरईएमएल और एमआईवीक्यू तरीकों के अधीन किया गया था।

सामान्य वितरण और असंतुलितता की विभिन्न डिग्री (पैमाने पर संदूषण के बिना) के साथ अनुरूपित प्रतिदर्शों से प्राप्त अनुमानों का माध्य, पूर्वाग्रह और एमएसई तालिका–1 और तालिका–2 में प्रस्तुत किए गए हैं। तालिका में प्रत्येक प्रविष्टि की गणना 500 प्रतिदर्शों से की जाती है। संतुलित डेटा के साथ और नकारात्मक अनुमानों की उपेक्षा करते हुए, एमएल को छोड़कर सभी तकनीकें समान अनुमान उत्पन्न करती हैं। असंतुलित डेटा के साथ प्रत्येक तकनीक प्रसरण घटक अनुमानों का एक अलग सेट और परिणामस्वरूप अलग–अलग आनुवंशिकता अनुमान उत्पन्न करती है। असंतुलितता की विभिन्न डिग्री के तहत आनुवंशिकता के सभी पैरामीट्रिक मूल्यों के लिए, संशोधित प्रसरण विधि अनुमान आरईएमएल को छोड़कर अन्य अनुमानों से बेहतर पाए जाते हैं, जो विभिन्न स्थितियों में आशाजनक पाए जाते हैं।

एमएल विधि द्वारा प्राप्त अनुमान नीचे की ओर पक्षपाती हैं जो कि सिद्धांत के अनुसार बहुत अच्छी तरह से है (एमएल निश्चित प्रभावों के आकलन से जुड़ी स्वतंत्रता की डिग्री में हानि का ध्यान रखने में विफल रहता है, यानी एमएल एस के बजाय एस का उपयोग करता है) स्वतंत्रता की समूह डिग्री के बीच के रूप में। परिणामस्वरूप, एमएल अनुमानक आम तौर पर नीचे की ओर पक्षपाती होता है, खासकर तब जब s छोटा होता है)। गैर–नकारात्मकता इन सभी अनुमानकों में ऊपर की ओर पूर्वाग्रह उत्पन्न करती है, जो आनुवंशिकता के निम्न मूल्यों के मामले में अधिक स्पष्ट है, जो तब होता है जब मुख्य विचरण के नकारात्मक अनुमान उत्पन्न होने की सबसे अधिक संभावना होती है। एमएल अनुमानक के लिए, यह ऊपर की ओर पूर्वाग्रह प्रभाव नीचे की ओर पूर्वाग्रह से कम हो जाता है। यह भी देखा जा सकता है कि जैसे–जैसे नर की संख्या बढ़ती है, अनुमानों का पूर्वाग्रह कम होता जाता है।

परिणाम

तालिका–1 और तालिका–2 से यह स्पष्ट है कि संशोधित प्रसरण विधि अनुमानक को आनुवंशिकता के सभी प्राचलिक मूल्यों के लिए छोटे एमएसई होने के कारण प्रसरण विधि पर अपना लाभ मिला है। हालाँकि, एमएल और आरईएमएल में संशोधित प्रसरण विधि अनुमानक की तुलना में छोटा एमएसई है। जब डेटा संतुलित होता है, तो एमआईवीक्यू और प्रसरण विधि अनुमानक समान होते हैं जो सिद्धांत के अनुसार बहुत अच्छी तरह से है। असंतुलित डेटा के मामले में प्रसरण विधि का एमएसई एमआईवीक्यू से कम है और विचलन की सीमा असंतुलन के स्तर पर निर्भर करती है। एमआईवीक्यू केवल आनुवंशिकता के निम्न मूल्यों के लिए अच्छा प्रदर्शन करता है और इसका उपयोग केवल तभी किया जाना चाहिए जब किसी को विश्वास हो कि आनुवंशिकता कम है। तालिकाओं से यह देखा जा सकता है कि सभी अनुमानकों के लिए जैसे–जैसे असंतुलन की डिग्री बढ़ती है, एमएसई बढ़ता है, लेकिन जैसे–जैसे संतों की संख्या बढ़ती है, एमएसई घटता जाता है।

स्केल संदूषण के साथ अनुरूपित नमूनों से प्राप्त आनुवंशिकता के अनुमान का माध्य, पूर्वाग्रह और एमएसई तालिका–3 और तालिका–4 में प्रस्तुत किए गए हैं। तालिकाओं से यह देखा गया है कि सभी तरीकों से प्राप्त अनुमान नीचे की ओर पक्षपाती हैं (संदूषण के उच्च अनुपात के कारण हो सकता है)। इसके अलावा, स्केल संदूषण को ध्यान में रखते हुए इन अनुमानकों के प्रदर्शन ने स्केल संदूषण के बिना इन अनुमानकों के प्रदर्शन के

तालिका 1: सामान्य के लिए विभिन्न पैरामीट्रिक मानों के साथ आनुवंशिकता के विभिन्न अनुमानों की तुलना तीन सायर्स वाली पाप्यलेशन, जिनमें से प्रत्येक की अलग–अलग डॉटर्स हैं।

h^2 के पैरामीट्रिक मानों की		0.1	0.1	0.1	0.4	0.4	0.4	0.8	0.8	0.8
विधि डॉटर्स की संख्या		50,50,50	30,50,70	10, 50, 90	50,50,50	30,50,70	10, 50, 90	50,50,50	30,50,70	10, 50, 90
प्रसरण विधि	माध्य बायस	0.0978	0.0802	0.0851	0.3600	0.3528	0.3512	0.7143	0.6857	0.6552
	एमएसई	-0.0022	-0.0198	-0.0149	-0.0400	-0.0472	-0.0488	-0.0857	-0.1143	-0.1448
		0.0177	0.0233	0.0308	0.1532	0.1433	0.1817	0.3481	0.3347	0.4045
एमओडी. प्रसरण विधि	माध्य बायस	0.1082	0.0973	0.1068	0.3689	0.3590	0.3625	0.7165	0.6886	0.6602
	एमएसई	0.0082	-0.0027	0.0068	-0.0309	-0.0410	-0.0375	-0.0834	-0.1114	-0.1398
		0.0150	0.0189	0.0249	0.1452	0.1380	0.1719	0.3442	0.3299	0.3961
एमएल	माध्य बायस	0.0696	0.0504	0.0484	0.2406	0.2335	0.2258	0.5135	0.4883	0.4464
	एमएसई	-0.0303	-0.0496	-0.0515	-0.1594	-0.1664	-0.1742	-0.2864	-0.3117	-0.3536
		0.0089	0.0102	0.0145	0.1022	0.1033	0.1169	0.2978	0.2894	0.3422
आरईएमएल	माध्य बायस	0.1082	0.0979	0.1148	0.3689	0.3589	0.3667	0.7165	0.6897	0.6739
	एमएसइ	0.0082	-0.0021	0.0149	-0.0309	-0.0411	-0.0333	-0.0834	-0.1103	-0.1261
		0.0150	0.0191	0.0353	0.1452	0.1375	0.1675	0.3442	0.3233	0.3511
मिवक्यू	माध्य बायस	0.0978	0.0959	0.0966	0.3600	0.3593	0.3572	0.7143	0.6835	0.6407
	एमएसइ	-0.0022	-0.0041	-0.0034	-0.0400	-0.0406	-0.0428	-0.0857	-0.1165	-0.1593
		0.0177	0.0205	0.0268	0.1532	0.1484	0.2042	0.3481	0.3658	0.4965

तालिका 2: सामान्य के लिए विभिन्न पैरामीट्रिक मानों के साथ आनुवंशिकता के विभिन्न अनुमानों की तुलना चार सायर्स वाली पाप्यलेशन, जिनमें से प्रत्येक की अलग–अलग डॉटर्स हैं।

h^2 के पैरामीट्रिक मानों की विधि डॉटर्स की संख्या		0.1	0.1	0.1	0.4	0.4	0.4	0.8	0.8	0.8
		50,50,50, 50	30,50,50, 70	10,10, 50, 90	50,50,50, 50	30,50,50, 70	10,10, 50, 90	50,50,50, 50	30,50,50, 70	10,10, 50, 90
प्रसरण विधि	माध्य बायस एमएसई	0.0960	0.1045	0.0925	0.3841	0.3808	0.3504	0.6954	0.6928	0.6707
		-0.0040	0.0046	-0.0074	-0.0159	-0.0192	-0.0495	-0.1046	-0.1071	-0.1292
		0.0277	0.0190	0.0310	0.1098	0.0977	0.1425	0.2572	0.2291	0.3396
एमओडी. प्रसरण विधि	माध्य बायस एमएसई	0.1106	0.1133	0.1161	0.3861	0.3832	0.3584	0.6960	0.6941	0.6746
		0.0106	0.0133	0.0160	-0.0138	-0.0168	-0.0413	-0.1040	-0.1059	-0.1253
		0.0239	0.0171	0.0243	0.1081	0.0956	0.1354	0.2563	0.2271	0.3331
एमएल	माध्य बायस एमएसई	0.0612	0.0716	0.0566	0.2867	0.2821	0.2351	0.5447	0.5461	0.5077
		-0.0388	-0.0284	-0.0433	-0.1132	-0.1179	-0.1648	-0.2552	-0.2539	-0.2920
		0.0119	0.0100	0.0138	0.0838	0.0745	0.1090	0.2418	0.2199	0.2964
आरईएमएल	माध्य बायस एमएसइ	0.1106	0.1137	0.1240	0.3861	0.3837	0.3566	0.6960	0.7010	0.6900
		0.0106	0.0137	0.0240	-0.0138	-0.0163	-0.0433	-0.1040	-0.0990	-0.1099
		0.0239	0.0170	0.0327	0.1081	0.0942	0.1309	0.2563	0.2294	0.2886
मिवक्यू	माध्य बायस एमएसइ	0.0960	0.1125	0.1087	0.3841	0.3818	0.3582	0.6954	0.6832	0.6415
		-0.0040	0.0125	0.0087	-0.0159	-0.0182	-0.0417	-0.1046	-0.1188	-0.1584
		0.0277	0.0185	0.0291	0.1098	0.1061	0.1851	0.2572	0.2452	0.4716

तालिका 3: अन्बैलन्स्ट डाटा एन्ड स्केल कन्टैमनेशन के लिए विभिन्न पैरामीट्रिक मानों के साथ आनुवंशिकता के विभिन्न अनुमानों की तुलना तीन सायर्स वाली पाप्यलेशन, जिनमें से डॉटर स्ट्रक्चर्स अलग–अलग हैं।

h^2 के पैरामीट्रिक मानों की विधि डॉटर्स की संख्या		0.1	0.1	0.1	0.4	0.4	0.4	0.8	0.8	0.8
		50,50,50	30,50,70	10, 50, 90	50,50,50	30,50,70	10, 50, 90	50,50,50	30,50,70	10, 50, 90
प्रसरण विधि	माध्य बायस	0.0941	0.0870	0.0882	0.3836	0.3213	0.3459	0.5848	0.5789	0.5409
		-0.0058	-0.0129	-0.0118	-0.0363	-0.0786	-0.0540	-0.2652	-0.2210	-0.2590
	एमएसई	0.0309	0.0313	0.0359	0.1519	0.1393	0.1921	0.4783	0.5523	0.5006
एमओडी. प्रसरण विधि	माध्य बायस	0.1113	0.1048	0.1132	0.3706	0.3283	0.3572	0.5910	0.5851	0.5492
		0.0113	0.0047	0.0133	-0.0294	-0.0716	-0.0428	-0.2090	-0.2148	-0.2508
	एमएसई	0.0265	0.0667	0.0291	0.1460	0.1330	0.1823	0.4682	0.5421	0.4868
एमएल	माध्य बायस	0.0625	0.0522	0.0542	0.2458	0.2152	0.2152	0.4273	0.4304	0.4304
		-0.0374	-0.0457	-0.0477	-0.1542	-0.1847	-0.1827	-0.3727	-0.3696	-0.4136
	एमएसई	0.0130	0.0126	0.0144	0.1034	0.1068	0.1309	0.4329	0.4953	0.4757
आरईएमएल	माध्य बायस	0.1113	0.1033	0.1204	0.3706	0.3392	0.3628	0.5910	0.5905	0.5854
		0.0113	0.0033	0.0204	-0.0294	-0.0660	-0.0372	-0.2090	-0.2094	-0.2346
	एमएसइ	0.0265	0.0236	0.0366	0.1460	0.1363	0.1800	0.4682	0.5425	0.4971
मिवक्यू	माध्य बायस	0.0941	0.1073	0.1089	0.3836	0.3215	0.3502	0.5848	0.5745	0.5226
		-0.0058	0.0073	0.0089	-0.0363	-0.0784	-0.0497	-0.2652	-0.2254	-0.2773
	एमएसइ	0.0309	0.0324	0.0332	0.1519	0.1384	0.2132	0.4783	0.5680	0.5419

तालिका 4: अन्बैलन्स्ट डाटा एन्ड स्केल कन्टैमनेशन के लिए विभिन्न पैरामीट्रिक मानों के साथ आनुवंशिकता के विभिन्न अनुमानों की तुलना चार सायर्स वाली पाप्यलेशन, जिनमें से डॉटर स्ट्रक्चर्स अलग–अलग हैं।

h^2 के पैरामीट्रिक मानों की विधि डॉटर्स की संख्या		0.1	0.1	0.1	0.4	0.4	0.4	0.8	0.8	0.8
		50,50,50, 50	30,50,50, 70	10,10, 50, 90	50,50,50, 50	30,50,50, 70	10,10, 50, 90	50,50,50, 50	30,50,50, 70	10,10, 50, 90
प्रसरण विधि	माध्य	0.1036	0.1004	0.0944	0.3812	0.3591	0.3468	0.5767	0.5732	0.5667
	बायस	0.0036	0.0004	-0.0055	-0.0187	-0.0409	-0.0531	-0.2832	-0.2267	-0.2318
	एमएसई	0.0216	0.0242	0.0370	0.1036	0.0956	0.1528	0.4290	0.4492	0.5437
एमओडी. प्रसरण विधि	माध्य	0.1137	0.1114	0.1175	0.3839	0.3610	0.3550	0.5192	0.5753	0.5739
	बायस	0.0137	0.0114	0.0175	-0.0161	-0.0389	-0.0449	-0.2807	-0.2247	-0.2280
	एमएसई	0.0192	0.0214	0.0305	0.1014	0.0940	0.1457	0.4250	0.4459	0.5040
एमएल	माध्य	0.0738	0.0724	0.0647	0.2846	0.2644	0.2436	0.4074	0.4579	0.4627
	बायस	-0.0261	-0.0275	-0.0352	-0.1153	-0.1355	-0.1564	-0.3926	-0.3423	-0.3373
	एमएसई	0.0112	0.0129	0.0192	0.0793	0.0723	0.1176	0.4271	0.4226	0.4992
आरईएमएल	माध्य	0.1137	0.1122	0.1228	0.3839	0.3611	0.3651	0.5192	0.5799	0.6108
	बायस	0.0137	-0.0123	0.0228	-0.0161	-0.0389	-0.0349	-0.2807	-0.2201	-0.1894
	एमएसइ	0.0192	0.0222	0.0396	0.1014	0.0929	0.1465	0.4250	0.4446	0.5258
मिवक्यू	माध्य	0.1036	0.1111	0.1101	0.3812	0.3509	0.3418	0.5767	0.5678	0.5165
	बायस	0.0036	0.0111	0.0101	-0.0187	-0.0391	-0.0582	-0.2832	-0.2322	-0.2835
	एमएसइ	0.0216	0.0220	0.0360	0.1036	0.1012	0.1916	0.4290	0.4732	0.5924

समान प्रवृत्ति दिखाई। पूर्वाग्रह और एमएसई के मामले में संशोधित प्रसरण विधि प्रसरण विधि से बेहतर है। हालाँकि आरईएमएल पूर्वाग्रह के संदर्भ में सुसंगत पाया गया है। लेकिन एमएसई के मामले में यह संशोधित प्रसरण विधि के बाद दूसरे स्थान पर है। यह देखा जा सकता है कि नमूना आकार बढ़ाने से प्रदर्शन में सुधार हुआ है।

संदर्भ

हर्बाच, एल. एच. (1959)। प्रापर्टीज आफ टाइप ।। अनैलिसिस आफ वेरीअन्स टेस्ट। ऐन. मैथ। स्टेट., 30: 939–959।

खत्री, आर. (1998) सम प्रैक्टिकल एस्टमेशन प्रोसीजर फॉर वेरीअन्स कम्पोनन्ट्स। कम्प्यूटेशनल स्टटिस्टिक्स ऐन्ड डेटा अनैलिसिस, 28: 1–32।

थॉम्पसन, डब्ल्यू.ए. जूनियर (1962)। द प्राब्लम आफ नेगटिव एस्टमेट आफ वेरीअन्स कम्पोनन्ट्स। ऐन. मैथ स्टेटिस्ट., 33: 273–289।

सांख्यिकी-विमर्श
2023

अंक
19

राजभाषा खण्ड

भा.कृ.अनु.प.-भारतीय कृषि सांख्यिकी अनुसंधान संस्थान

हिन्दी के प्रगामी प्रयोग की रिपोर्ट

भा.कृ.अनु.प.–भारतीय कृषि सांख्यिकी अनुसंधान संस्थान में दिन प्रतिदिन 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 निर्धारित लक्ष्यों को संस्थान में लगभग पूरा कर लिया गया है। संस्थान द्वारा समस्त प्रशासनिक कार्य शत–प्रतिशत हिन्दी में किया जाता है तथा धारा 3(3) का भी पूर्ण रूप से अनुपालन किया जा रहा है।

संस्थान में राजभाषा हिन्दी की प्रगति का जायजा लेने के लिए उपमहानिदेशक (कृषि शिक्षा), भारतीय कृषि अनुसंधान परिषद् मुख्यालय द्वारा समय–समय पर संस्थान का राजभाषा संबंधी निरीक्षण किया गया। उपमहानिदेशक (कृषि शिक्षा), ने निरीक्षण रिपोर्ट में संस्थान में हिन्दी में हो रहे कार्यों की प्रगति पर संतोष व्यक्त करते हुए संस्थान की सराहना की। संस्थान के अलग–अलग प्रभागों एवं अनुभागों में हिन्दी में किए जा रहे कार्यों की समीक्षा करने के लिए हिन्दी एकक के अधिकारियों द्वारा कुल 20 प्रभागों एवं अनुभागों का निरीक्षण किया गया।

संस्थान में प्रशासनिक कार्य के साथ – साथ वैज्ञानिक प्रकृति के कार्यों में भी हिन्दी के उपयोग को प्रोत्साहित किया जाता है। संस्थान के वैज्ञानिक प्रभागों द्वारा आयोजित प्रशिक्षण कार्यक्रमों की ई–संदर्भ पुस्तिकाओं में आमुख एवं आवरण पृष्ठ द्विभाषी रूप में प्रस्तुत करने के साथ–साथ परियोजना रिपोर्टों के आवरण पृष्ठ, आमुख, एवं सारांश द्विभाषी रूप में प्रस्तुत करने के साथ –साथ कुछ हिन्दी के व्याख्यान भी शामिल किए। संस्थान के वैज्ञानिकों द्वारा हिन्दी में वैज्ञानिक विषयों पर हिन्दी कार्यशालाओं का भी आयोजन किया गया। इसके अतिरिक्त, संस्थान में एम.एससी. तथा पीएच.डी के विद्यार्थियों ने अपने शोध–प्रबन्धों के सार द्विभाषी रूप में प्रस्तुत किए। वैज्ञानिकों एवं तकनीकी कर्मियों द्वारा कुछ शोध–पत्र भी हिन्दी में प्रकाशित किए गए।

प्रतिवेदनाधीन अवधि के दौरान संस्थान में राजभाषा कार्यान्वयन समिति की 04 बैठकें क्रमशः मार्च 29, 2023; जून 23, 2023; सितम्बर 27, 2023 एवं दिसम्बर 28, 2023 को आयोजित की गई। इन बैठकों में राजभाषा नियम एवं अधिनियम को कारगर ढंग से लागू करने तथा इसमें दिए गए प्रावधानों के अनुसार वार्षिक कार्यक्रम में निर्धारित लक्ष्यों की प्राप्ति हेतु चर्चा की गई तथा आवश्यक कदम उठाए गए।

संस्थान के समस्त कर्मियों को 02 वर्ष की अवधि में कम से कम एक बार हिन्दी कार्यशाला में सहभागिता करने का अवसर मिले, इस अनिवार्यता के संबंध में भारत सरकार, गृह मंत्रालय, राजभाषा विभाग द्वारा समय – समय पर जारी कार्यालय ज्ञापन द्वारा निर्धारित लक्ष्य की प्राप्ति के लिए, प्रोत्साहित किया गया।

प्रतिवेदनाधीन अवधि के दौरान संस्थान में विभिन्न वर्गों के कर्मियों एवं राष्ट्रीय कृषि अनुसंधान प्रणाली के अन्य संस्थानों के वैज्ञानिकों के लिए पाँच हिन्दी कार्यशालाएं आयोजित की गई। पहली कार्यशाला संस्थान में वैज्ञानिक एवं तकनीकी वर्ग के कर्मियों के लिए परीक्षण अभिकल्पना प्रभाग के वैज्ञानिकों, डॉ. सुकान्त दाश, एवं डॉ. सुशील कुमार सरकार द्वारा मार्च 28–29, 2023 (02 दिवसीय) के दौरान ''परीक्षण अभिकल्पना के अनुप्रयोग'' विषय पर ऑन–लाइन आयोजित की गई, जिसमें 06 वक्ताओं द्वारा विषय से संबंधित 09 उप–विषयों पर व्याख्यान दिए गए। इस कार्यशाला में कुल 14 प्रतिभागियों (संस्थान के 06 वैज्ञानिकों, भारतीय कृषि अनुसंधान परिषद के अधीनस्थ अन्य संस्थानों से 08 वैज्ञानिकों) ने सहभागिता की।

दूसरी कार्यशाला संस्थान में प्रशासनिक वर्ग के कर्मियों के लिए हिन्दी एकक द्वारा मई 30, 2023 को ''सरकारी काम काज को हिन्दी में करने के लिए ई–टूल्स एवं

राजभाषा नियम अधिनियम की जानकारी'' विषय पर ऑफ–लाइन आयोजित की गई, जिसमें 02 वक्ताओं द्वारा विषय से संबंधित 02 उप–विषयों पर व्याख्यान दिए गए। इस कार्यशाला में कुल 28 प्रतिभागियों (संस्थान के 07 अधिकारी एवं 21 कर्मचारी) ने सहभागिता की।

तीसरी कार्यशाला संस्थान में वैज्ञानिक एवं तकनीकी वर्ग के कर्मियों के लिए सांख्यिकी आनुवंशिकी प्रभाग के वैज्ञानिकों, डॉ. कांचन सिन्हा, डॉ. मृन्मय राय एवं डॉ. राजीव रंजन कुमार द्वारा जून 06–12, 2023 (07 दिवसीय) के दौरान ''कृषि में सांख्यिकीय और मशीन लर्निंग तकनीक का परिचय'' विषय पर ऑन लाइन आयोजित की गयी जिसमें 16 वक्ताओं द्वारा विषय से संबंधित 19 उप–विषयों पर व्याख्यान दिए गए। इस कार्यशाला में कुल 25 प्रतिभागियों (संस्थान के 01 वैज्ञानिक, 02 तकनीकी अधिकारी एवं भा.कृ.अनु.प. के बाह्य संस्थानों से 12 वैज्ञानिक तथा 06 तकनीकी अधिकारी के अलावा 04 तकनीकी सहायकों) ने सहभागिता की।

चौथी कार्यशाला संस्थान में वैज्ञानिक एवं तकनीकी वर्ग के कर्मियों के लिए प्रतिदर्श सर्वेक्षण प्रभाग के वैज्ञानिकों, डॉ. पंकज दास, डॉ. राहुल बनर्जी एवं डॉ. भारती द्वारा सितम्बर 06–13, 2023 (08 दिवसीय) के दौरान ''कृषि एवं प्रतिदर्श आँकड़ो का विश्लेषण हेतु सांख्यिकीय सॉफ्टवेयर का अनुप्रयोग'' विषय पर ऑन–लाइन आयोजित की गई। जिसमें 11 वक्ताओं द्वारा विषय से संबंधित 15 उप–विषयों पर व्याख्यान दिए गए। इस कार्यशाला में कुल 42 प्रतिभागियों (संस्थान के 04 वैज्ञानिक तथा 01 तकनीकी सहायक तथा भा.कृ.अनु.प. के बाह्य संस्थानों से 32 वैज्ञानिक और 01 तकनीकी अधिकारी के अलावा 04 तकनीकी सहायक) ने सहभागिता की।

पाँचवी कार्यशाला संस्थान में वैज्ञानिक एवं तकनीकी वर्ग के कर्मियों के लिए पूर्वानुमान एवं कृषि प्रणाली मॉडलिंग प्रभाग के वैज्ञानिकों, डॉ. कांचन सिन्हा, डॉ. मृन्मय राय एवं डॉ. राजीव रंजन कुमार द्वारा दिसम्बर 07–13, 2023 (07 दिवसीय) के दौरान ''सांख्यिकीय एवं मशीन लर्निंग तकनीक के माध्यम से समय श्रंखला का पूर्वानुमान'' विषय पर ऑन–लाइन आयोजित की गई। जिसमें 17 वक्ताओं द्वारा विषय से संबंधित 19 उप–विषयों पर व्याख्यान दिए गए। इस कार्यशाला में कुल 15 प्रतिभागियों (संस्थान के 08 वैज्ञानिक, 01 तकनीकी अधिकारी तथा 01 तकनीकी सहायक और भा.कृ.अनु.प. के बाह्य संस्थानों से 04 वैज्ञानिक और 01 तकनीकी अधिकारी) ने सहभागिता की।

राजभाषा विभाग द्वारा जारी वार्षिक कार्यक्रम में निहित लक्ष्यों को पूरा करते हुए संस्थान के अधिकारियों/कर्मचारियों द्वारा समस्त पत्राचार हिन्दी में अथवा द्वि भाषी रूप में किया गया। संस्थान के विभिन्न वैज्ञानिक प्रभागों तथा प्रशासनिक अनुभागों द्वारा आयोजित बैठकों की कार्यसूची तथा कार्यवृत्त हिन्दी अथवा द्विभाषी रूप में जारी किए गए। संस्थान में अपना कार्य शत–प्रतिशत हिन्दी में करने के लिए 12 अनुभागों को विनिर्दिष्ट किया गया है। गृह मंत्रालय, राजभाषा विभाग द्वारा जारी विभिन्न नकद पुरस्कार योजनाएँ संस्थान में लागू हैं तथा संस्थान के कर्मियों ने इन योजनाओं में बढ़–चढ़कर भाग लिया।

संस्थान में कार्यरत सभी हिन्दीतर अधिकारियों/कर्मचारियों द्वारा हिन्दी ज्ञान संबंधी प्रशिक्षण पूरा किया जा चुका है। आज तक की स्थिति के अनुसार, संस्थान में अब कोई ऐसा हिन्दीतर अधिकारी/कर्मचारी शेष नहीं रह गया है जिसे हिन्दी ज्ञान संबंधी प्रशिक्षण दिया जाना शेष हो।

संस्थान की वेबसाइट पर 'हिन्दी सेवा लिंक' उपलब्ध है। जिसमें 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 उपलब्ध हैं। संस्थान के कर्मियों द्वारा अपना दैनिक कार्य हिन्दी में सरलता से करने के लिए इस सेवा का उपयोग किया जाता है।

संस्थान द्वारा प्रकाशित वार्षिक हिन्दी पत्रिकाः 'सांख्यिकी–विमर्श' का नियमित प्रकाशन किया जा रहा है तथा 18वें अंक का प्रकाशन प्रतिवेदनाधीन अवधि के दौरान किया गया।

संस्थान में 14 से 29 सितम्बर, 2023 के दौरान

हिन्दी पखवाड़े का आयोजन किया गया। इस वर्ष हिन्दी पखवाड़े का शुभारम्भ अर्थात उदघाटन हिन्दी दिवस एवं तृतीय अखिल भारतीय राजभाषा सम्मेलन के अवसर पर 14 सितम्बर, 2023 को पुणे (महाराष्ट्र) में श्री अजय कुमार मिश्रा, माननीय गृह राजी मंत्री जी द्वारा किया गया था। हिन्दी पखवाड़े का आयोजन एवं इससे संबन्धित प्रतियोगताएं संस्थान में दिनांक 14 से 29 सितम्बर, 2023 के दौरान आयोजित की गयी। दिनांक 14 सितम्बर, 2023 को काव्य–पाठ प्रतियोगिता का आयोजन किया गया। हिन्दी पखवाड़ा के दौरान डॉ दरोगा सिंह स्मृति व्याख्यान के साथ–साथ वैज्ञानिक प्रभागों में हिन्दी में सर्वाधिक वैज्ञानिक कार्य करने के लिए प्रभागीय चल–शील्ड, डिजिटल हिन्दी शोध–पत्र प्रस्तुतिकरण प्रतियोगिता (वैज्ञानिक वर्ग के कर्मियों के लिए), डिजिटल हिन्दी पोस्टर प्रस्तुति प्रतियोगिता (प्रशासनिक वर्ग के कर्मियों के लिए), हिन्दी श्रुतलेख प्रतियोगिता, हिंदीतर कर्मियों के लिए शब्दार्थ लेखन प्रतियोगिता, अंताक्षरी प्रतियोगिता तथा प्रश्न मंच प्रतियोगिता भी आयोजित की गईं। सभी प्रतियोगिताओं में संस्थान के विभिन्न वर्ग के कर्मियों ने बढ़–चढ़कर हिस्सा लिया। संस्थान में प्रत्येक वर्ष हिन्दी दिवस के अवसर पर डॉ दरोगा सिंह स्मृति व्याख्यान का आयोजन किया जाता है। इस वर्ष इस कड़ी का बत्तीसवाँ व्याख्यान भारतीय विश्वविद्यालय संघ की महासचिव, डॉ. (श्रीमती) पंकज मित्तल द्वारा दिया गया और इस कार्यक्रम की अध्यक्षता संस्थान के निदेशक महोदय द्वारा की गई। दिनांक 29 सितम्बर, 2023 को हिन्दी पखवाड़ा के समापन समारोह के अवसर पर इस दौरान आयोजित प्रतियोगिताओं के सफल प्रतियोगियों के साथ–साथ वर्ष 2022–23 के लिए सरकारी काम–काज मूल रूप से हिन्दी में करने के लिए प्रोत्साहन योजना के अंतर्गत भी नकद पुरस्कारों की घोषणा की गयी। इसके अतिरिक्त इस अवसर पर अक्टूबर, 2022 से अगस्त, 2023 तक की अवधि के दौरान संस्थान में आयोजित हिन्दी कार्यशालाओं के वक्ताओं को प्रशस्ति–पत्र प्रदान किए जाने की भी घोषणा की गयी।

हिन्दी पखवाड़ा के समापन दिवस पर (श्रीमती) पंकज मित्तल, महासचिव भारतीय विश्वविद्यालय संघ संस्थान के प्रांगण में पौधारोपण एवं बत्तीसवाँ डॉ दरोगा सिंह स्मृति व्याख्यान देते हुए।

डेटा साइंस से कृषि विकास

प्रकाश कुमार, वैज्ञानिक

कृषि क्षेत्र में तेजी से आगे बढ़ रहे हैं हम,
डेटा साइंस से कृषि विकास में ला रहे है उत्कृष्ट परिवर्तन।
कृषि आँकड़े का विश्लेषण हो, या हो मौसम की पूर्वानुमान,
डेटा साइंस से हो रहा है भारतीय कृषि प्रगतिमान ।
डेटा साइंस का महत्व, है अनमोल,
खेती में आये आधुनिकता, डेटा साइंटिस्ट की बस है यही भावना।
खेतों में आये लहलहाती फसल, हम सब की है यही कामना,
डेटा साइंस से हो रहा है खेती का नित नये सवेरा,
किसानों को मिले खुशी, छट जाये उनकी अँध्यारा।
जब भी बारिश है आई, या तूफान है लाई,
डेटा साइंस ने हमें हमेशा सच्चाई की राह दिखाई ।
समय समय पर पूर्वानुमान कर, फसल को बर्बाद होने से बचाई,
डेटा साइंस का है यह आधुनिकतम सहारा,
जिससे किसानों की आय बढ़ पाई।
डेटा साइंस का महत्व, है अनमोल,
खेती में आये आधुनिकता, डेटा साइंटिस्ट की ये भावना।
खेतों में आये लहलहाती फसल, हम सब की है यही कामना,

डेटा साइंस से हो रहा है खेती का नया सवेरा,
किसानों को मिले खुशी, छट जाये उनकी अँध्यारा।
डेटा साइंस के जादू से,
खेती की नई दशा और दिशा मिली,
किसानों की मेहनत को सलामी, जिससे परिवार की खुशहाली बढ़ाई ।
खेती का आधुनिकतम तकनीक अपनाकर, अपनी आय अधिक कर पाई ।
डेटा साइंस का महत्व, है अनमोल,
खेती में आये आधुनिकता, डेटा साइंटिस्ट की ये भावना।
खेतों में आये लहलहाती फसल, हम सब की है यही कामना,
डेटा साइंस से हो रहा है खेती का नया सवेरा,
किसानों को मिले खुशी, छट जाये उनकी अँध्यारा।।
डेटा साइंस से हो रहा है कृषि शोध में नित नए उत्थान,
हर खेती को मिल रहा है नया दिशा,
डेटा साइंस के सहारे बढ़ाये हर किसानों का मान
यही संकल्प हमारा, जिसमे आगे रहा हमेसा
भारतीय कृषि सांख्यिकी अनुसंधान संस्थान।

मेरे देश की विचारधारा

राखी सोनी, प्रवर लिपिक

चलो एक काम करते है, आज फिर एक, मंस्जिद को अपने नाम करते है
चल ऐ दोस्त एक काम करते है, आज फिर एक राम के बंदे का जीना हराम करते है
उंच नीच, जात पात और धर्म के नाम पर बंटवारा ही, तो जरूरते है हमारी
चल आज फिर भाई चारे, प्यार महोब्बत की इस दुकान को जला कर खाक करते है
आज मॅहगाई, पढाई लिखाई और रोजगार मुददा नही हमारा,
चल आज फिर रोजगार के नाम पर, फिर किसी गली मोहल्ले मे जाकर, दंगे फसाद करते है
धर्म तो कभी जात पात के नाम पर,
आज फिर एक बेटी को भरे बाजार नीलाम करते है
चल आज अस्तमत के बाजर मे,
फिर किसी बहन बेटी के वस्त्र तार तार करते है
हमेशा की तरह सिर झुका कर खोमशी से,
उन हैवानियतो की गलियो से निकल जाते है
चल हर बार की तरह, अपनी अंतरआत्मा को मार,
सब कुछ नजरअंदाज करते है
चल आज कलयुग की बिसात पर गंदी राजनिति की चौपड़ बिछाते है
क्योंकि आज श्री कृष्ण न अर्जुन को गीता का ज्ञान देने आयेगे,
नाही किसी और द्रोपदी की लाज बचाने, वो महाभारत दोहरायगे.
तो चल आज फिर एक मंस्जिद को अपने नाम करते है
आज फिर एक राम के बंदे का जीना हराम करते है
पर चल आज, क्यो न एक काम करे,
मिलकर नफरत के इन अधियारे गलियारे मे प्यार का दिया/दीप जलाते है
चांद और सूरज तो, अब छू लिये हमने
तो चल ऐ दोस्त उठ, अब मिलकर, अपने देश की संस्कृति और उसके संस्कारो को बचाते है.

हिंदी का गुणगान और विशेषताएं

केवल कुमार शर्मा, मुख्य वित्त एवं लेखा अधिकारी

देवभाषा संस्कृत की हूँ बड़ी बेटी,
आधुनिक भारत के जन जण की में बोली,
सम्पूर्ण राष्ट्र की एकता की में हूँ जननी,
सरल मधु हूँ और लगती मुझमे एक बिंदी,
पूजा अर्चना करती दुनिया मेरी, नाम है मेरा हिंदी,
में हूँ भेद भाव से परे की बोली ,
मुझसे ही आशा और जिज्ञासा में भरत की धरती डोली,
ववहार कुशल है, विशाल जनसमूह की में वाणी,
वास्तु, क्रिया और भाव, हर अभिवकीयति की में साँची,
विश्व में चौथे नंबर पे बोली जाती ,
११ स्वर और ३३ वायजन से सुसज्जित,
।५ लाख सब्द विद्यमान मुझमे, १८ बोलियों, ५ उपभाषाओ के मालकियत मेरी,
२५ देशो के ८० करोड़ जन जन की नाती,
१७५ विश्विद्यालओं में मेरी शिक्षा जारी,
देवनागरी कंप्यूटर की भी उपुक्त विज्ञानं लिपि,
हर रिश्ते के अलग अक्षर हैं, गूंगे अक्षर मुझमे नाही,
अन्य भासाओ के ग्रहण करने में मुझमें न कोई संकोची ,
देवभाषा संस्कृत की हूँ बड़ी बेटी,
आधुनिक भारत के जन जण की में बोली,
सम्पूर्ण राष्ट्र की एकता की में हूँ जननी,
सरल मधु हूँ और लगती मुझमे एक बिंदी,
पूजा अर्चना करती दुनिया मेरी, नाम है मेरा हिंदी,

20 कवि

केवल कुमार शर्मा, मुख्य वित्त एवं लेखा अधिकारी

वेद व्यास की में गीता वाणी,
वाल्मीकि की रामायण गाथा सारी
सूरदास के वात्सल्य रस को दर्शताती
तुलसी दास के चिन्मय सूरसागर में डूब जाती,
कबीर के दोहो की में वांदी,
मीरा भजन की में मीठासी,
प्रेम पीर धनानद के सुजान सागर बनाती,
रस खान के रासो में खो जाती,
आमिर खुसरो की पहली और मुकरियों से विस्मय लाती,
सूर्यकांत की अनयामिका मितवाल।
पंत की गुंजन पल्लव बन जाती,
रवीन्द्रनाथ की में गीतांजनली बन आती,
भारतेन्दु का रूप महानी,
नागार्जुन की युगधारी,
हरिवसं की मधुशाला में झूमत जाती,
दिनकर के अंदर तूफान बन उमड़ जाती,
सुब्द्ररा के ओझे गीत बन जाती,
म्हदेवी की में विरहा पियासी,
हार न मानेगी कभी अटल की वाणी ,
मेरी तरपान विश्वास ले आती ,
इन बीसियों कविओं के कारन ही आन बाण और शान बन जाती,
अब तुम्ही पहचान वचन दो , तभी राष्ट्र गौरव का मान और प्राण बन जाती,

आज की चुनातियाँ

केवल कुमार शर्मा, मुख्य वित्त एवं लेखा अधिकारी

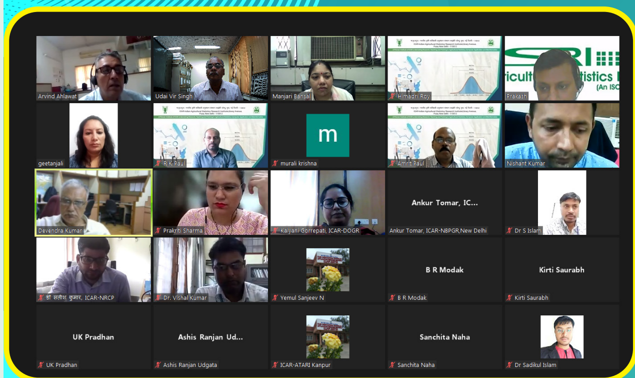
अगर छोड़ा हिंदी का साथ, तो तुम खुद से छूट जाओगे,
भाषा मरी तो तो तुम भी न बच पाओगे,
बोये बाबुल तो आम कहाँ से खायोगे,
हो करोड़ो गुण तुम में, कुछ न रच पाओगे,
महज भाष नहीं है हिंदी, है ये माँ हमारी,
आज जान लो तुम सभी अपनी जिम्मेदारी,
हिंदी है माँ हमारी,
हिंदी है माँ हमारी,
जय हिंदी, जय भारत,
जय हिन्दी, जय हिन्दुस्तान,

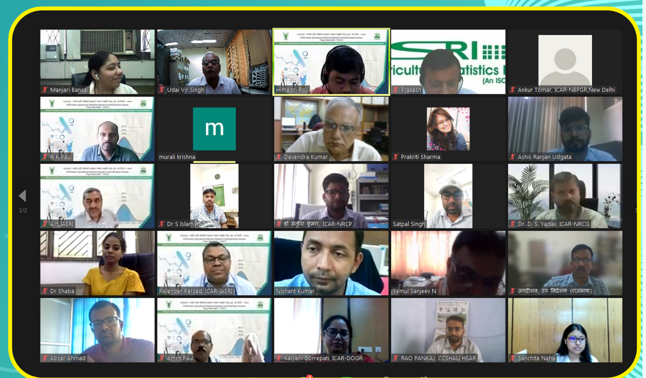
हिन्दी कार्यशाला - 2023

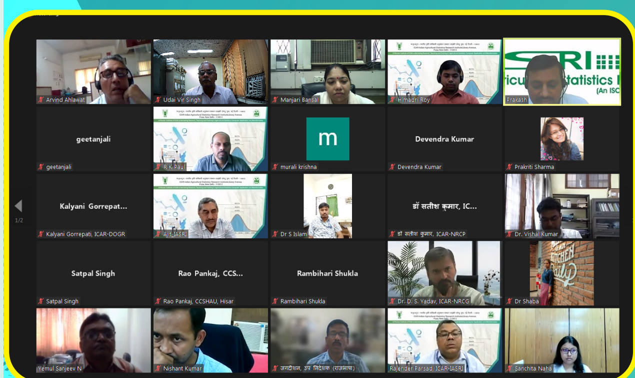
सरकारी काम काज को करने के लिये ई–टूल्स तथा राजभाषा नियम, अधिनियम की जानकारी" (समन्वयक: अभिषेक श्रीवास्तव सह– समन्वयक: मनोज कुमार)

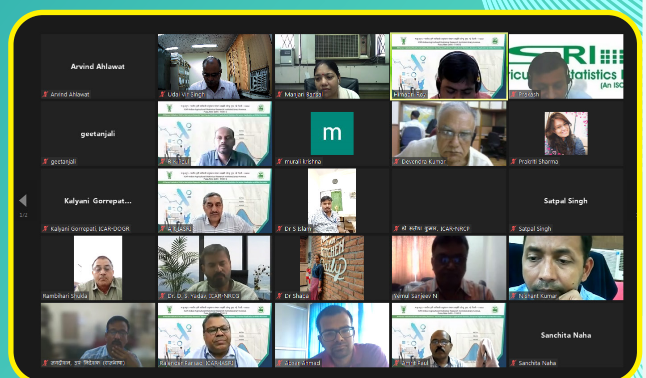
भाकृअनुप – भारतीय कृषि सांख्यिकी अनुसंधान संस्थान,
नई दिल्ली मई 30, 2023

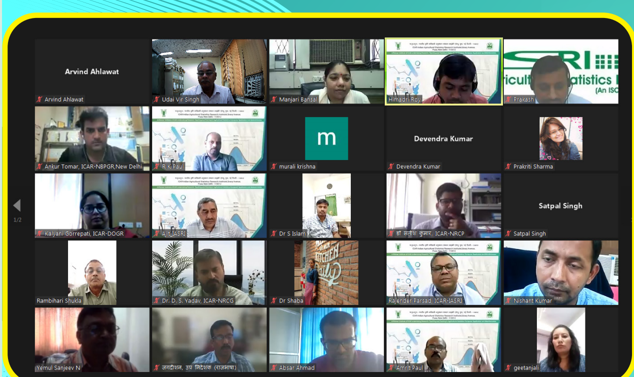
21–27 जून, 2023

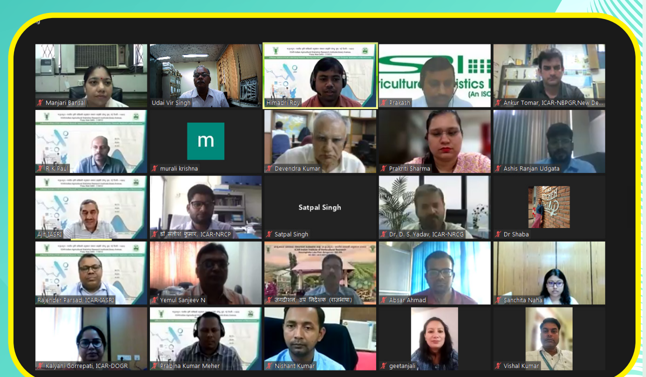




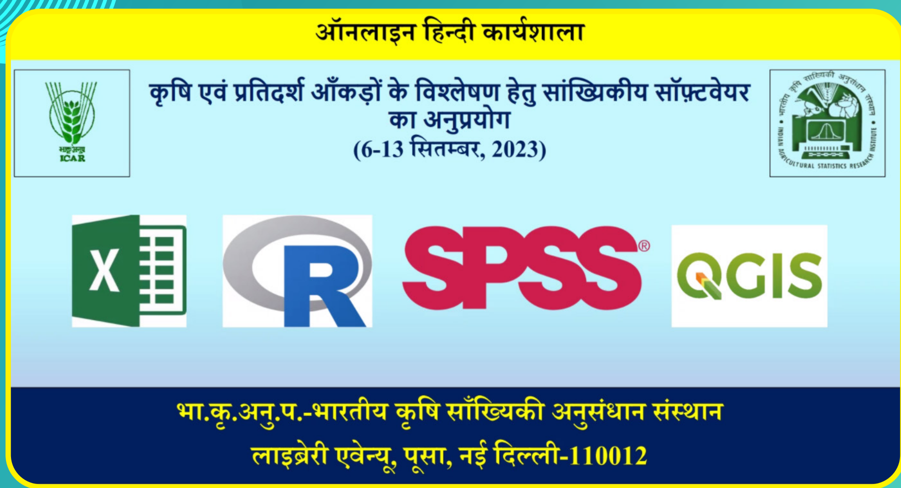


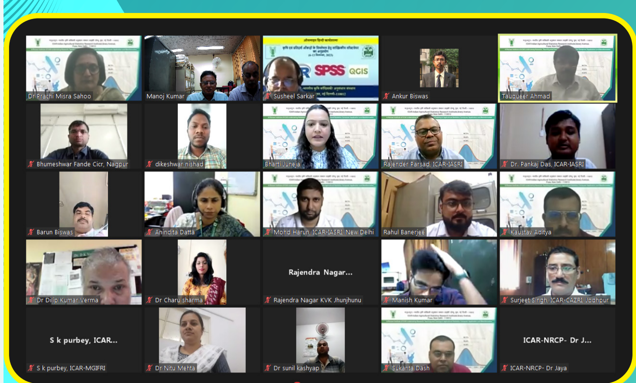


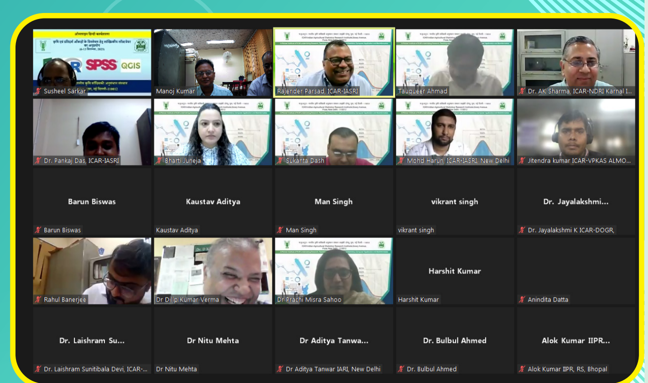


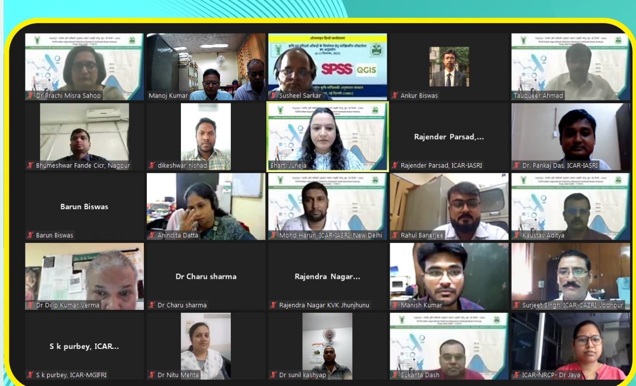


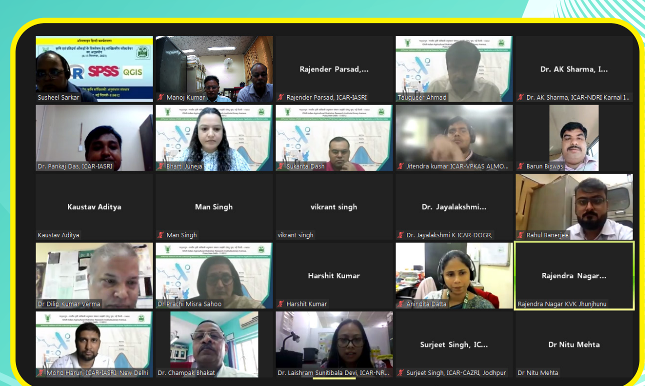
6–13 सितम्बर, 2023 (ऑन–लाइन) हिन्दी कार्यशाला



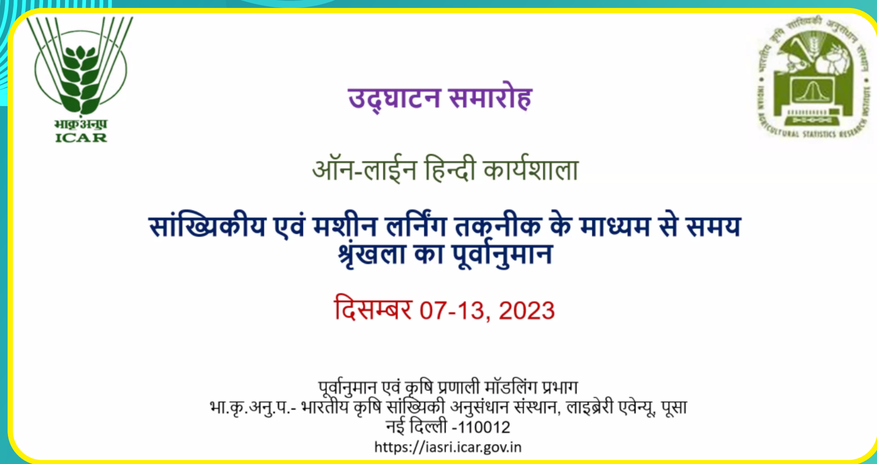


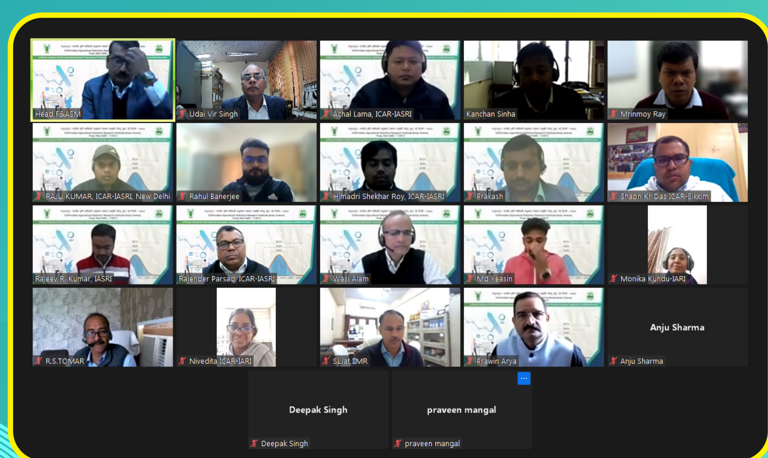


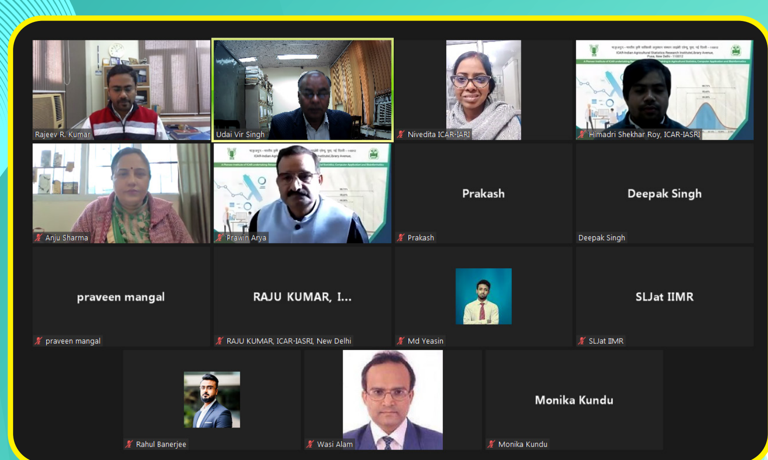




07–13 दिसम्बर, 2023







हिन्दी पखवाड़ा - 2023

संस्थान में हिन्दी पखवाड़े का आयोजन दिनांक 14–29 सितंबर, 2023 के दौरान किया गया। हिन्दी पखवाड़े के अंतर्गत काव्यपाठ, अंताक्षरी, प्रश्नमंच, शोधपत्र पोस्टर प्रस्तुति, चल–शील्ड प्रतियोगिता का आयोजन किया गया। पखवाड़े के अंतर्गत डॉ. दरोगा सिंह स्मृति व्याख्यान डॉ. पंकज मित्तल, महासचिव, एसोसिएशन ऑफ इंडियन यूनिवर्सिटीज के द्वारा दिया गया। पखवाड़े से संबंधित कुछ छायाचित्र निम्नवत हैः

ऑर्गेनोग्राम

अनुसंधान सलाहकार समिति ⟷ **निदेशक** ⟷ **संस्थान प्रबंधन समिति**

प्रभागाध्यक्ष
- परीक्षण अभिकल्पना
- सांख्यिकी आनुवंशिकी
- पूर्वानुमान एवं कृषि प्रणाली मॉडलिंग
- प्रतिदर्श सर्वेक्षण
- संगणक अनुप्रयोग
- कृषि जैव–सूचना
- आईटी एकक

अधिकारी प्रभारी
- पुस्तकालय एवं प्रलेखन
- उन्नत संकाय प्रशिक्षण केंद्र
- संस्थान प्रौद्योगिकी प्रबंधन एकक
- प्राथमिकीकरण, निगरानी एवं मूल्यांकन प्रकोष्ठ
- कृषिज्ञान प्रबंधन प्रकोष्ठ
- एनएएसएम

प्रोफेसर (कृषि संख्यिकी)
प्रोफेसर (संगणक अनप्रयोग)
प्रोफेसर (जैवसूचना विज्ञान)
वार्डन
- प्रशिक्षण प्रशासन प्रकोष्ठ

सतर्कता अधिकारी
- ओएमवी एकक

मुख्य प्रशासनिक अधिकारी
- प्रशासन I अनुभाग
- प्रशासन II अनुभाग
- रोकड़ अनुभाग
- केंद्रीय खरीद एवं उपकरण रखरखाव
- पावती एवं प्रेषण अनुभाग
- कार्य एवं रखरखाव अनुभाग
- रखरखाव अनुभाग
- भण्डार
- हिन्दी एकक

मुख्य वित्त एवं लेखाधिकारी
- लेखा परीक्षा अनुभाग
- लेखा अनुभाग

NOTES

भा.कृ.अनु.प.-भारतीय कृषि सांख्यिकी अनुसंधान संस्थान

लाइब्रेरी एवेन्यू, पूसा, नई दिल्ली – 110012

https://iasri.icar.gov.in