

सांख्यिकी-विमर्श

अंक
17

2021



KISAN SARATHI
Powered By IIDS



डिजिटल इंडिया
Power To Empower

किसान साथी

आई आई डी एस (IIDS) द्वारा संचालित
किसानों के लिए कृषि प्रौद्योगिकी सूचना प्रणाली
का स्वचालित हस्तांतरण

भारतीय कृषि अनुसंधान परिषद और डिजिटल इंडिया कॉर्पोरेशन का
संयुक्त प्रयास



Krishi Integrated Solution for Agri Apps Navigation



**एग्री ऐप्स नेविगेशन के लिए
कृषि एकीकृत समाधान
किसान 2.0**

विकसित और अनुरक्षित

© 2021 एनएचईपी, कम्पोनन्ट - 2, कंप्यूटर अनुप्रयोग प्रभाग, भाक-अनुप-भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली



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



सांख्यिकी-विमर्श

2021

अंक
17

संपादक मंडल	
अनिल कुमार	अध्यक्ष
प्रवीण आर्य	सदस्य
मो. समीर फारूकी	
सुशील कुमार सरकार	
द्विजेश चन्द्र मिश्र	
सुकान्त दाश	
उमेश चन्द्र बन्दूनी	
देवेन्द्र कुमार	
सविता वधवा	
नेहा नारंग	
उदय वीर सिंह	



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

लाइब्रेरी एवेन्यू, पूसा, नई दिल्ली-110012

<https://iasri.icar.gov.in>



सांख्यिकी-विमर्श

2021

अंक
17

<https://aicrponstation.icar.gov.in/aicrponstation/>

AICRP on IFS: On-Station Research
Online Data Entry and Analysis

Home About the Software Help Contact us

Home
Centre
Publications
Important Links
Login

The experiments on stations under the AICRP on IFS are planned and conducted under four types of research programmes viz., (i) development of new cropping systems (ii) nutrient management in cropping systems (iii) development of system based management practices and (iv) maximum yield research. These experiments were conducted using Randomized Complete Block (RCB) design, Factorial RCB design, split plot designs, strip plot designs and $3^2 \times 2$ balanced confounded factorial experiments. There are two objectives as (i) To statistically analyze the data of experiments conducted at 37 Cropping Systems Research Centres and to identify the suitable statistical designs for conduct of experiments as per the technical programme and to develop suitable methods of analysis for the identified designs. To consolidate the results based on the combined analysis of the data for the experiments.

द्वारा प्रकाशित
निदेशक

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

मुद्रण : मार्च, 2022



निदेशक की कलम से

भा.कृ.अनु.प.–भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, कृषि सांख्यिकी, संगणक अनुप्रयोग एवं जैव सूचना में अनुसंधान के साथ-साथ श्रेष्ठ शिक्षण-प्रशिक्षण प्रदान करके अपने दायित्व को पूर्णरूपेण निभा रहा है। हिन्दी के प्रयोग को और प्रोत्साहित करने के विचार से वर्ष 2005–06 में संस्थान में हिन्दी पत्रिका “सांख्यिकी-विमर्श” के प्रकाशन का शुभारम्भ हुआ और तब से अब तक प्रतिवर्ष इस पत्रिका का प्रकाशन किया जा रहा है। प्रस्तुत अंक इस पत्रिका का सत्रहवां अंक है। संस्थान की हिन्दी पत्रिका “सांख्यिकी-विमर्श 2021” आप के समक्ष प्रस्तुत करते हुए मुझे हार्दिक प्रसन्नता हो रही है।

भा.कृ.अनु.प.–भारतीय कृषि सांख्यिकी अनुसंधान संस्थान में वर्ष-दर-वर्ष हिन्दी के प्रगामी प्रयोग में अभिवृद्धि हो रही है। राजभाषा नीति को संस्थान में सुचारु रूप से कार्यान्वित किया जा रहा है। वैज्ञानिक कार्यों में भी हिन्दी के प्रयोग को प्रोत्साहित किया जाता है। न केवल मात्रात्मक रूप में बल्कि हिन्दी के प्रयोग में गुणवत्ता की ओर भी ध्यान दिया जा रहा है।

पत्रिका के इस अंक में संस्थान के कीर्तिस्तम्भ, संस्थान में राजभाषा से सम्बन्धित कार्यों आदि की जानकारी के साथ-साथ कृषि सांख्यिकी, कृषि में संगणक अनुप्रयोग एवं कृषि जैव सूचना से सम्बन्धित विभिन्न लेखों एवं शोध-पत्रों को सम्मिलित किया गया है। पत्रिका में पाठकों के हिन्दी ज्ञानवर्धन एवं दैनिक उपयोग के लिए दैनिक स्मरणीय शब्द शतक (सांख्यिकी एवं तकनीकी) हिन्दी व अंग्रेजी में दिये गये हैं।

मैं पत्रिका के प्रकाशन के लिए उन सभी लेखकों का आभारी हूँ जिन्होंने इस पत्रिका में प्रकाशन हेतु अपने लेख देकर हमारे इस प्रयास को सफल बनाया। पत्रिका के प्रकाशन के लिए सम्पादक मंडल के अध्यक्ष, डॉ. अनिल कुमार एवं सदस्यों, डॉ. प्रवीण आर्य, मो. समीर फारुकी, डॉ. सुशील कुमार सरकार, डॉ. द्विजेश मिश्र, डॉ. सुकान्त दाश, श्री उदयवीर सिंह, श्री देवेन्द्र कुमार, तथा सुश्री नेहा नारंग का आभार व्यक्त करता हूँ जिनके अथक प्रयासों से यह पत्रिका इस रूप में आप के समक्ष आ सकी।

आशा है इस अंक की विषय-वस्तु पाठकों के लिए सूचनाप्रद एवं उपयोगी सिद्ध होगी और सांख्यिकीय विज्ञान जैसे तकनीकी विषय में भी हिन्दी साहित्य का प्रयोग कर के पाठकों का ज्ञानवर्धन करने में सहयोगी सिद्ध होगी। इस पत्रिका के भावी अंकों में सुधार के लिए आप के विचारों एवं सुझावों की प्रतीक्षा रहेगी।

राजेन्द्र प्रसाद
निदेशक

सम्पादकीय

सांख्यिकीय सभ्यता की गति में अंकों का योगदान बड़ा ही महत्वपूर्ण रहा है, और अंक पद्धति के विकास का बहुत बड़ा श्रेय भारत को प्राप्त है। मनुष्य के ज्ञान की प्रत्येक शाखा अंकों का ऋणी है। सांख्यिकी विज्ञान भी अंकों का विज्ञान है, जिन्हें आँकड़े कहते हैं, परन्तु इन अंकों के विशिष्ट लक्षण होते हैं। व्यष्टि का अध्ययन न करके, समष्टि नाम से अभिहित समूह या समुदाय का अध्ययन करना सांख्यिकी विज्ञान की मौलिक धारणा है। जैसा कि आप सबको विदित है कि सांख्यिकी एक गणितीय विज्ञान है जिसमें किसी वस्तु/अवयव/तंत्र/समुदाय से सम्बन्धित आँकड़ों का संग्रह, विश्लेषण, व्याख्या या स्पष्टीकरण की प्रस्तुति की जाती है। सांख्यिकी विज्ञान एक ऐसी विधा है, जो विज्ञान के हर क्षेत्र चाहे वो चिकित्सा विज्ञान, सामाजिक विज्ञान, व्यवसायिक विज्ञान, प्रबंधन विज्ञान, मनोविज्ञान अथवा कृषि विज्ञान हो, सभी क्षेत्रों में अपनी अग्रणी भूमिका निभाती है।

उल्लेखनीय है कि पत्रिका सांख्यिकी विमर्श के अब तक सोलह अंक प्रकाशित हो चुके हैं और अब आपके सम्मुख सतरहवा अंक प्रस्तुत किया जा रहा है। सांख्यिकी विमर्श के इस निरंतर सफलता के पीछे पाठक वृन्द एवम् समस्त लेखकों का योगदान एवम् कौशल निहित है। यह अत्यन्त ही हर्ष का विषय है कि सांख्यिकी विमर्श लोकप्रिय हो रही है एवम् इसके लेखों को सराहा जा रहा है। किन्तु, हमारे लिए केवल इतना ही पर्याप्त नहीं है क्योंकि हम जानते हैं कि सांख्यिकी विमर्श को शीर्षतम उँचाइयों तक ले जाया जा सकता है। इस दिशा में हमारे प्रयास निरंतर जारी है और हम उम्मीद करते हैं कि आने वाले दिनों में यह पत्रिका और भी रुचिकर एवम् ज्ञानवर्धक सिद्ध होगी।

सांख्यिकी विमर्श के इस अंक में सांख्यिकी एवम् कृषि साहित्य के विभिन्न आयामों यथा ग्रीष्म ग्वार

का अध्ययन, कृषि वानिकी परीक्षण, समानता आव्यूह की समीक्षा, फसल कटाई परीक्षणों की संख्या, गाम्पर्टज प्रसंभाव्य, प्रसामान्य बंटन, समय श्रृंखला मॉडल, कृषि परीक्षण से संबंधित लेख, गैर रैखीय मिश्रित प्रभाव माडल, स्मार्ट खेती में आईओटी (IOT) की भूमिका, फसल जीन प्रारूप का वर्गीकरण, रेपिटेड मेजरमेंट में प्रेडिक्शन, प्रोटिओमिक्स डाटा विश्लेषण, स्थायित्व विश्लेषण, पूर्वानुमान विज्ञान से संबंधित लेख प्रस्तुत हैं। हमें उम्मीद ही नहीं अपितु पूर्ण विश्वास है कि सांख्यिकी विमर्श का यह अंक पाठको के लिए अत्यन्त ज्ञानवर्धक एवम् रोचक सिद्ध होगा।



उपरोक्त लेखों के अतिरिक्त साहित्यिक लेख जैसे कि दैनिक स्मरणीय शब्द शतक, संस्थान के कीर्ति स्तम्भ, एवम् अन्य स्तम्भ हैं। इस प्रकार हम कह सकते हैं कि यह पत्रिका विज्ञान एवम् साहित्य का अनुपम संगम है, जो इसको एक मौलिक एवम् अतुल्यनीय पत्रिका बनाती है। संपादक मंडल "सांख्यिकी विमर्श" के सभी लेखकों के योगदान के लिए हृदय से आभारी है एवम् उम्मीद करता है कि भविष्य में भी उनका अमूल्य सहयोग जारी रहेगा।

इस पत्रिका को समृद्ध एवम् रुचिकर बनाने के लिए, पाठको का सुझाव अपेक्षित है। सभी पाठको एवम् लेखकों से आग्रह है कि कृपया अपनी प्रतिक्रिया हमें अवश्य भेजें ताकि हम इस पत्रिका को और अधिक प्रभावशाली ज्ञानवर्धक एवम् रोचक बना सकें। इस पत्रिका के निरंतर विकास की कामना के साथ।

अनिल कुमार
अध्यक्ष, संपादक मंडल

अनुक्रमणिका

संस्थान के कीर्ति स्तम्भ	
• डॉ. लालमोहन भर	1
• डॉ. हुकुम चंद्र	3
डॉ. दरोगा सिंह स्मृति व्याख्यान: प्रसरण गुणांक के विविध दृष्टिकोण एक चयनात्मक पुनरावलोकन	5
• डॉ. मुरारी सिंह	
अनुसंधान खण्ड	
मिनी-लाइसीमीटर द्वारा सिंचाई के विभिन्न स्तरों पर ग्रीष्म ग्वार के वास्तविक वाष्पोत्सर्जन का अध्ययन	11
• हरिमोहन मीना, रंजय कुमार सिंह, प्रियव्रत सान्त्रा एवं सी.बी. पाण्डेय	
वृक्ष-तंत्र संतुलन से युक्त कृषि-वानिकी परीक्षण	16
• टी.बी. पीटर, सिनी वर्गीस, सीमा जग्गी, मोहम्मद हारून, एल्दो वर्गीस एवं देवेन्द्र कुमार	
विशेषता निष्कर्षण तकनीकों और समानता आव्यूह की समीक्षा: सामग्री-आधारित छवि पुनर्प्राप्ति	22
• मधु, सपना निगम, संचिता नाहा एवं चंदन कुमार देब	
फसल कटाई परीक्षणों की कम संख्या का उपयोग करके जिला स्तरीय प्रमुख फसल उपज अनुमान	28
• कौस्तव आदित्य, पंकज दास, भारती, अंकुर विश्वास, श्रीला दास एवं मंजू गौतम	
बहिर्जात चर के साथ गॉम्पर्ट्ज़ प्रसंभाव्य विभिन्नात्मक समीकरण पद्धति	34
• हिमाद्री घोष एवं सविता वधवा	
कुमारस्वामी बंटन: प्रसामान्य बंटन की एक नई श्रेणी	39
• पंकज दास, कौस्तव आदित्य और भारती	
दीर्घकाल अनुस्मरण समय श्रृंखला मॉडल द्वारा भारत में चावल की कीमतों का पूर्वानुमान	45
• रंजीत कुमार पॉल, अमृत कुमार पॉल, दीपांकर मित्रा एवं एस. पी. सिंह	
कृषि परीक्षण हेतु प्रचलन प्रतिरोधी लागत प्रभावी द्विस्तर घटकीय रन क्रम	52
• बिजोय चन्द, अर्पण भौमिक, सीमा जग्गी, एल्दो वर्गीस, सिनी वर्गीस, अनिदिता दत्ता एवं देवेन्द्र कुमार	
पशुओं में गैर-रैखिक मिश्रित प्रभाव मॉडल का उपयोग करके वृद्धि मापदंडों का अनुमान और स्थिर प्रभाव मॉडल के साथ तुलना	56
• पंकज दास, भारती, कौस्तव आदित्य, एवं अमृत कुमार पॉल	
स्मार्ट खेती में इन्टरनेट ऑफ थिंग्स का उपयोग	63
• सपना निगम, अक्षय धीरज, संचिता नाहा, मो. अशराफुल हक, सौमेन पाल एवं मधु	
स्वसंयोजित लक्षण प्रतिचित्र का उपयोग करते हुए फसल जीन प्रारूप का वर्गीकरण	67
• सुकान्त दाश, अनिल कुमार, बैद्यनाथ मंडल, सुशील कुमार सरकार एवं देवेन्द्र कुमार	
रिपीटेड मेज़रमेंट डेटा में प्रेडिक्शन दृष्टिकोण	73
• राहुल बनर्जी, संगमित्रा पाल, तौकीर अहमद एवं भारती	

प्रोटिओमिक्स आँकड़ों के विश्लेषण का संक्षिप्त विवरण <ul style="list-style-type: none"> सुधीर श्रीवास्तव, स्नेहा मुर्मू, मो. समीर फारुकी, नीरज बुढलाकोटी, द्विजेश चंद्र मिश्र, यू. बी. अंगडि एवं के. के. चतुर्वेदी 	78
स्थायित्व विश्लेषण— अप्राचालिक दृष्टिकोण <ul style="list-style-type: none"> प्रकाश कुमार, रंजीत कुमार पॉल, अमृत कुमार पॉल, राजू कुमार, राजीव रंजन कुमार, मृन्मय राय एवं मो. यासीन 	85
गार्च और एस वी आर के संयोजित मॉडल का उपयोग करके कृषि कमोडिटी की कीमतों में अस्थिरता का पूर्वानुमान <ul style="list-style-type: none"> रंजीत कुमार पॉल, तनिमा दास, अमृत कुमार पॉल, प्रकाश कुमार, मोहम्मद यासीन एवं एस.पी. सिंह 	92
फसल उपज के पूर्वानुमान के लिए उन्नत मौसम सूचकांक आधारित बेसिअन समाश्रयण मॉडल <ul style="list-style-type: none"> मोहम्मद यासीन, के.एन. सिंह, अचल. लामा, बिशाल. गुरुंग, रंजीत कुमार पॉल, प्रकाश कुमार, अमृत कुमार पॉल, हिमाद्रि शेखर रॉय एवं विशाल दिनकर 	96
मैग्नापोर्थे ओरेजे से एफेक्टर प्रोटीन और चावल से माइटोकॉन्ड्रिया संबद्ध प्रोटीन कॉम्प्लेक्स के बीच परस्पर प्रभाव की जांच <ul style="list-style-type: none"> भारती पाण्डेय, स्नेहा मुर्मू, सौम्या शर्मा, ऋत्विका दास, नीरज बुढलाकोटी, द्विजेश चंद्र मिश्र, दीपा भट्ट, सुधीर श्रीवास्तव और मोहम्मद समीर फारुकी 	104
आणविक गतिशीलता सिमुलेशन का उपयोग करके मंकीपॉक्स प्रोटीन के विरुद्ध फाइटोकेमिकल अवरोधकों की पहचान <ul style="list-style-type: none"> स्नेहा मुर्मू, सुधीर श्रीवास्तव, दीपा भट्ट, भारती पांडे, सौम्या शर्मा, ऋत्विका दास, के. के. चतुर्वेदी और मोहम्मद समीर फारुकी 	112
राजभाषा खण्ड	
संस्थान की राजभाषा यात्रा: वर्ष 2021 <ul style="list-style-type: none"> उदय वीर सिंह 	121
हिन्दी पखवाड़ा 2021	125
दैनिक स्मरणीय शब्द—शतक	127

संस्थान के कीर्ति स्तम्भ

डॉ. लालमोहन भर

डॉ. लालमोहन भर का जन्म 31 जनवरी 1967 को पश्चिम बंगाल के हुगली जिला के फुलिया गांव में हुआ था। डॉ. भर, सामान्य रूप से सैद्धांतिक और व्यावहारिक सांख्यिकी और विशेष रूप से कृषि सांख्यिकी के क्षेत्र में अपने शोध योगदान के लिए जाने जाते थे। डॉ. भर ने अपनी स्नातक (कृषि) की उपाधि, बिधान चंद्र कृषि विश्वविद्यालय से वर्ष 1990 में प्राप्त की। तदोपरान्त कृषि सांख्यिकी में एम.एससी. और पीएच.डी. की उपाधि क्रमशः 1993 और 1997 में भारतीय कृषि अनुसंधान संस्थान, नई दिल्ली से प्राप्त की। वह परीक्षण अभिकल्पना में विशेष रूप से आउटलायर्स के अध्ययन पर एक अंतरराष्ट्रीय स्तर पर जाने-माने और प्रसिद्ध शोधकर्ता थे। वह एक प्रख्यात वैज्ञानिक थे जिन्होंने विशेष रूप से भारत में सांख्यिकीय आनुवंशिकी की लोकप्रियकरण का प्रयास किया। डॉ. भर भारतीय कृषि सांख्यिकी अनुसंधान संस्थान के सचिव (जनवरी 2019 से अक्टूबर 2020) एवं सोसायटी ऑफ स्टेटिस्टिक्स, कम्प्यूटर एंड एप्लिकेशन के उपाध्यक्ष भी रहे। डॉ. भर एक प्रख्यात शोधकर्ता के साथ साथ उत्कृष्ट शिक्षक भी थे। उन्होंने चार पीएच.डी. और सात एम.एससी. छात्रों का कृषि सांख्यिकी के क्षेत्र में मार्गदर्शन भी किया। डॉ. भर को भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली में प्रमुख रूप से अनुसंधान, शिक्षण, प्रशिक्षण और परामर्श में 26 वर्षों का अनुभव था। उन्हें भारतीय कृषि सांख्यिकी संस्था (ISAS) में सर्वश्रेष्ठ पेपर पुरस्कार जैसे कई अन्य राष्ट्रीय पुरस्कारों से सम्मानित किया गया। उन्हें भारतीय कृषि सांख्यिकी संस्था और सोसाएटी ऑफ स्टेटिस्टिक्स, कम्प्यूटर एंड एप्लिकेशन जैसी दो अलग-अलग सोसाएटी के संयुक्त सचिव के रूप में कार्य किया। उन्होंने सांख्यिकी और अनुप्रयोग के प्रबंध

संपादक के रूप में भी काम किया। उन्हें भारत में 2019 में आयोजित कृषि सांख्यिकी पर 8वें अंतरराष्ट्रीय सम्मेलन के आयोजन समिति के सचिव के रूप में कार्य करने का अवसर प्राप्त हुआ। उन्होंने भारत में कई अंतर-मंत्रालयी और संस्थागत समितियों सहित विभिन्न राष्ट्रीय संगठनों तथा समितियों में सांख्यिकीय विशेषज्ञ तथा सलाहकार के रूप में अपना महत्वपूर्ण प्रभाव छोड़ा है। डॉ. भर ने कई शैक्षणिक गतिविधियों जैसे एसोसिएट संपादक, अतिथि संपादक और प्रसिद्ध वैज्ञानिक पत्रिकाओं के समीक्षक, वैज्ञानिक समितियों के सदस्य और विभिन्न राष्ट्रीय सम्मेलनों के संयोजक जैसी कई शैक्षणिक गतिविधियों में स्वेच्छा से योगदान दिया। उनके कार्यों में सांख्यिकीय पद्धति और अनुप्रयुक्त समस्याओं के विविध क्षेत्र, विशेष रूप से परीक्षणों की अभिकल्पना, सांख्यिकीय आनुवंशिकी और परीक्षण अभिकल्पनाओं में अन्य निदान संबंधी समस्याएं सम्मिलित हैं। डॉ. भर ने उच्च प्रभाव वाली अंतरराष्ट्रीय पत्रिकाओं में बड़े पैमाने पर प्रकाशित किया है। अपने कार्यकाल में उन्होंने 30 से अधिक राष्ट्रीय परियोजनाओं में अन्वेषक के रूप में काम किया। 100 से अधिक शोध पत्र, पुस्तकें, कई तकनीकी बुलेटिन, रिपोर्ट, पुस्तक अध्याय, कार्य पत्र, प्रशिक्षण और शिक्षण संदर्भ नियमावली प्रकाशित की। 31 जुलाई, 2021 को कैंसर से पीड़ित डॉ. भर सदा के लिए काल के गाल में समा गए। सामान्य रूप से सांख्यिकी और विशेष रूप से परीक्षणों की अभिकल्पना स्टैटिस्टिकल मॉडलिंग एवं आनुवंशिकी के प्रति उनका योगदान अविस्मरणीय है।



संस्थान के कीर्ति स्तम्भ

डॉ. हुकुम चंद्र

प्रतिभाशाली डॉ. हुकुम चंद्र, जो सामान्य रूप से सैद्धांतिक और व्यावहारिक सांख्यिकी दोनों एवं विशेष रूप से कृषि सांख्यिकी में अपने शोध योगदान के लिए जाने जाते हैं, का जन्म 07 नवम्बर 1972 को उत्तर प्रदेश में प्रतापगढ़ जिले के रायगढ़ गांव में हुआ था। डॉ. हुकुम ने अपनी स्नातक, स्नातकोत्तर (एम.एससी.) एवं एम.फिल., दिल्ली विश्वविद्यालय से क्रमशः 1993, 1995 और 1998 में स्वर्ण पदक विजेता के रूप में की। उन्होंने 2007 में संयुक्त गणराज्य (UK) के साउथेम्प्टन विश्वविद्यालय से सांख्यिकी में पीएच.डी. और तदुपरान्त ऑस्ट्रेलिया के वोलोंगोंग विश्वविद्यालय से पोस्ट डॉक्टरेट की। डॉ. हुकुम उत्कृष्ट शिक्षक भी थे। उन्होंने कृषि सांख्यिकी के क्षेत्र में दो पीएच.डी. और चार एम.एस.सी. छात्रों का मार्गदर्शन किया। डॉ. हुकुम को प्रमुख रूप से भा.कृ.अनु.प.—भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली में अनुसंधान, शिक्षण, प्रशिक्षण और परामर्श में 23 वर्षों से अधिक का अनुभव था। उन्हें अपने शैक्षणिक जीवन के दौरान इंटरनेशनल एसोसिएशन ऑफ सर्वे स्टैटिस्टिक्स एंड यंग रिसर्चर / अमेरिकन स्टैटिस्टिकल एसोसिएशन के छात्र पुरस्कार द्वारा कोकरन-हैनसेन पुरस्कार से सम्मानित किया गया था। शोध कार्यकाल में उनके उत्कृष्ट योगदान के लिए सांख्यिकी और कार्यक्रम कार्यान्वयन मंत्रालय, भारत सरकार द्वारा सांख्यिकी में राष्ट्रीय पुरस्कार, भारतीय कृषि अनुसंधान परिषद द्वारा राष्ट्रीय अध्येता और लाल बहादुर शास्त्री उत्कृष्ट युवा वैज्ञानिक पुरस्कार, राष्ट्रीय कृषि विज्ञान अकादमी कारिकोर्निसोन अवार्ड, एवं भारतीय कृषि सांख्यिकी संस्था (ISAS) द्वारा डॉ. डीएन लाल मेमोरियल अवार्ड, प्रोफेसर पी.वी. सुखात्मे गोल्ड मेडल अवार्ड आईएसएस फेलो अवार्ड एवं राष्ट्रीय कृषि विज्ञान अकादमी (NAAS) के अध्येता इत्यादि से सम्मानित किया गया। वह अंतर्राष्ट्रीय सांख्यिकी संस्थान, नीदरलैंड के निर्वाचित सदस्य रहे हैं। उन्होंने भारत में कई अंतर-मंत्रालयी

और संस्थागत समितियों, श्रीलंका, इथियोपिया, म्यांमार और इटली में संयुक्त राष्ट्र के खाद्य और कृषि संगठन सहित विभिन्न राष्ट्रीय और अंतर्राष्ट्रीय संगठनों और समितियों में सांख्यिकीय विशेषज्ञ और सलाहकार के रूप में अपना महत्वपूर्ण योगदान दिया। डॉ. हुकुम ने कई शैक्षणिक गतिविधियों जैसे जर्नल ऑफ स्टैटिस्टिकल थ्योरी एंड प्रैक्टिसेस एवं मॉडल असिस्टेड स्टैटिस्टिक्स एंड एप्लिकेशन की एसोसिएट संपादकों, जर्नल ऑफ भारतीय कृषि सांख्यिकी संस्था की अतिथि संपादकों और वैज्ञानिक समितियों के सदस्य और विभिन्न राष्ट्रीय एवं अंतर्राष्ट्रीय सम्मेलनों के संयोजक के रूप में स्वेच्छा से योगदान दिया। उन्होंने व्यावहारिक समस्याओं के विविध क्षेत्र; लघु क्षेत्र आकलन; बूटस्ट्रैप तरीके; कृषि, सामाजिक-आर्थिक और स्वास्थ्य संकेतकों का अलग-अलग स्तर का आकलन एवं विश्लेषण; कृषि एवं पशुधन आँकड़ों में सुधार के लिए सर्वेक्षण आंकड़े एवं सांख्यिकीय विधियों के लिए स्थानिक मॉडल सम्मिलित हैं। अपने कार्यकाल में उन्होंने 25 से अधिक राष्ट्रीय और अंतर्राष्ट्रीय परियोजनाओं में अन्वेषक के रूप में कार्य किया, 125 से अधिक शोध पत्र, 04 पुस्तकें, कई तकनीकी बुलेटिन, रिपोर्ट, पुस्तक अध्याय, वर्किंग पेपर, प्रशिक्षण और शिक्षण संदर्भ नियमावली प्रकाशित की। 26 अप्रैल, 2021 को भारत में कोविड-19 के महामारी के कारण डॉ. हुकुम चंद्र का आकस्मिक निधन हो गया। डॉ. हुकुम की स्मृति में इंटरनेशनल एसोसिएशन ऑफ सर्वे स्टैटिस्टिक्स ने 2021 से सभी मध्य-व्यवसायी शोधकर्ताओं के लिए प्रतिवर्ष 500 यूरो का हुकुम चंद्र मेमोरियल पुरस्कार प्रारम्भ किया। प्रतिदर्श सर्वेक्षण के सिद्धांत और अनुप्रयोगों में उनका महत्वपूर्ण योगदान सदैव स्मरण रहेगा।



डॉ. दरोगा सिंह स्मृति व्याख्यान

प्रसरण गुणांक के विविध दृष्टिकोण: एक चयनात्मक पुनरावलोकन

डॉ. मुरारी सिंह

सांख्यिकी विज्ञान विभाग, टोरंटो विश्वविद्यालय, कनाडा

प्रसरण गुणांक का परिचय स्नातक की प्रारम्भिक कक्षाओं में ही हो जाता है, और यह एक मौलिक शब्द है। इसका सांख्यिकी में कई जगह वर्णन आता है, इसका उपयोग कई जगह विभिन्न तरीकों से किया जाता है। प्रसरण गुणांक का उपयोग प्रसरण या विचलन ही नहीं बल्कि विभिन्नता को मापने के लिए भी होता है, अगर हम इसको व्युत्क्रम तरीके से देखें तो समरूपता के मापक के रूप में भी इसका प्रयोग किया जा सकता है।

प्रसरण गुणांक, औसत के सापेक्ष में विचलन का मापन करता है। प्रसरण गुणांक, परीक्षणात्मक क्षेत्र के समरूपता को मापने का एक माध्यम है। प्रसरण गुणांक, औसत के सापेक्ष में विचलन का मापन करता है। प्रसरण, कृषि परीक्षणात्मक क्षेत्र के समरूपता को मापने का एक माध्यम है। बीज की गुणवत्ता का परीक्षण यानी सीड क्वालिटी टेस्टिंग, में प्रसरण गुणांक का उपयोग शुद्धता, समरूपता एवं स्थायित्व (distinctness, uniformity and stability) के परीक्षण के लिए किया जाता है। किसी अनुवांशिक रूप (genotype) का विभिन्न वातावरणों में उसका प्रदर्शन कैसा है, यह प्रसरण गुणांक के जरिए जाना जा सकता है। किसी उद्यम की आय कितनी स्थिर/स्थायी है यह भी प्रसरण गुणांक के जरिए जाना जा सकता है।

परीक्षण अभिकल्पना में हमें प्रतिरूप की संख्या (number of replication) का निर्धारण करने की आवश्यकता पड़ती है। प्रतिदर्श की संख्या कितनी होनी चाहिए, उसमें भी हम प्रसरण गुणांक का उपयोग कर सकते हैं। प्रसरण गुणांक दो प्राचलों पर निर्भर रहता

है, पहला है मानक विचलन और दूसरा है औसत, अगर विभिन्न प्रसरण गुणांकों के बीच विचलन को देखें तो मानक विचलन एवं औसत के बीच स्थायित्व को देखने के लिए प्रसरण गुणांक ज्यादा दृढ़ सूचक है।

विभिन्न पर्यावरणों में कुछ ना कुछ अंतर तो होता ही है, चूंकि समय के साथ हम बहुत ज्यादा आँकड़े इकट्ठा करने लगे हैं ऐसे में प्रसरण गुणांक के पूर्वगामी बंटन के बारे में जान सकते हैं, और इस ज्ञान का उपयोग बेयसियन पद्धति से प्रसरण गुणांक के पश्च बंटन अथवा पश्च औसत आदि का आकलन लगा सकते हैं। भा.कृ.अनु.प.—भारतीय कृषि सांख्यिकीय अनुसन्धान संस्थान में भी इस विषय पर अनेक शोधकार्य हुए हैं।

प्रतिदर्श सर्वेक्षण में भी प्रसरण गुणांक का प्रयोग होता है। डॉ. दरोगा सिंह जी फसल कटाई परीक्षणों से जुड़े रहे थे। इस परिपेक्ष्य में शोधकार्य अभी भी जारी है, जैसे कि इष्टतम कटाई क्षेत्र कितना होना चाहिए, इसको जानने के लिए भी प्रसरण गुणांक का प्रयोग किया जाता है।

मुझको प्रोफेसर चौबे के साथ कोंकोर्डिया विश्वविद्यालय, मोंट्रिअल, कनाडा में काम करने का सुअवसर प्राप्त हुआ है। उनकी रुचि कई क्षेत्रों में है, और उनमें से एक प्रमुख क्षेत्र है प्रसरण गुणांक का प्रतिदर्श बंटन (sampling distribution of CV)। प्रसरण गुणांक के बंटन संबंधित शोध कार्य में उनके साथ जुड़ा रहा हूँ। इस विषय पर प्रकाशित 6 शोध पत्र, जिनके सारांश स्मृति व्याख्यान में प्रस्तुत किए थे, निम्न हैं।

1. विभिन्न बंटनों से प्रसरण गुणांक (CV) का प्रतिदर्श

बंटन (sampling distribution of CV)

2. एक सामान्य वितरण/समष्टि से प्रतिदर्श प्रसरण गुणांक के सममितीय परिवर्तन पर
3. व्युत्क्रम गाऊसी वितरण से प्रतिदर्श प्रसरण गुणांक के परिवर्तन को स्थिर करने वाला सममिति और विचरण
4. व्युत्क्रम गाऊसी वितरण से एक सामान्य प्रसरण गुणांक का अनुमान और परीक्षण
5. प्रसरण गुणांक पर महत्व के साथ व्युत्क्रम गाऊसी आँकड़ों के लिए बेयसियन अनुमान
6. फसल परीक्षणों में पुनरावृत्ति की संख्या का बेयसियन निर्धारण

इन छह शोधपत्रों के सार यहाँ प्रस्तुत हैं ।

प्रथम प्रश्न उठता है कि प्रसरण गुणांक दिखता कैसा है ?

यदि $X \sim (\mu, \sigma^2)$ एक यादृच्छिकृत (रैंडम) चर (वैरिएबल) है तो

$$\text{प्रसरण गुणांक: } \theta = \frac{\sigma}{\mu}$$

जहाँ θ , प्रसरण गुणांक, μ , औसत, एवं σ , मानक विचलन है।

यादृच्छिक प्रतिदर्श x_1, x_2, \dots, x_n , आकार d का प्रतिदर्श है, जिसका प्रसरण गुणांक (सीवी) $\hat{\theta} = \frac{s}{\bar{X}}$

जहाँ प्रसरण गुणांक $\hat{\theta}$, औसत $\bar{X} = \sum_{i=1}^n x_i$, एवं मानक विचलन $s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{X})^2}{(n-1)}}$ है।

1. विभिन्न बंटनों से प्रसरण गुणांक (सीवी) का प्रतिदर्श बंटन

चूंकि X एक सामान्य (नार्मल) बंटन का चर है, इसका औसत शून्य भी हो सकता है अथवा ऋणात्मक भी हो सकता है ऐसी स्थिति में प्रसरण गुणांक का आकलन सही नहीं होगा। इस स्थिति में प्रसरण गुणांक की जगह, इसके व्युत्क्रम का प्रयोग किया जाए तो यह उचित होगा। अतः प्रतिदर्श के व्युत्क्रम प्रसरण गुणांक

$$\theta_1 = \frac{\bar{X}}{s}$$

का बंटन सरल होगा। क्या चर X का रूपांतरण इस तरह से संभव है इससे प्राप्त होने वाले गुणांक

के प्रसरण में स्थायित्व आ जाए ? इसके लिए \sinh^{-1} रूपांतरण का उपयोग किया जा सकता है, जिसके अंतर्गत आठ समष्टियों (Normal, Logistic, Laplace, Lognormal, Chi-square, Exponential, Weibull और Pareto) का अध्ययन किया गया है तथा निम्न तीन सांख्यिकीय सूत्रों का आधार लिया गया है:

1. कोई सुधार न करें; $\hat{\theta} = \frac{s}{\bar{X}}$ ऐसे ही रहने दें।
2. पूर्वाग्रह उपचरित (bias corrected)
3. व्युत्क्रम \sinh (\sinh^{-1}) रूपांतरण का उपयोग

इन तीनों आंकलकों का प्रयोग उपरोक्त वर्णित आठ समष्टियों पर किया गया तथा उसका प्रभाव देखा गया।

गुणांक θ_1 के मोमेंट्स (moments) के रूप में पूर्वाग्रह (bias), औसत वर्ग त्रुटि (MSE), ढलवा (skewness), कर्टोसिस (kurtosis) प्रकट किया गया। इसकी निकटता की जांच कुछ आँकड़ों की उत्पत्ति करके की गयी है। इसके सैद्धान्तिक सूत्र n^2 तक सन्निकट है प्रतिदर्श की संख्या 20 एवं 200 इत्यादि ली गयी है। परिणाम मे यह पाया गया कि Pareto को छोड़कर बाकी सभी 7 समष्टियों में पूर्वाग्रह (bias) और औसत वर्ग त्रुटि (MSE) अपने अभिष्ट मानों के सन्निकट है। यह भी पाया गया है की ये जो व्युत्क्रम रूपांतरण प्रभावी है, की अनुशंसा की गयी।

2. एक सामान्य वितरण/समष्टि से प्रतिदर्श प्रसरण गुणांक के सममितीय परिवर्तन पर

प्रसरण गुणांक का बंटन, प्रसामान्य बंटन है, इसके लिए दो तरह का रूपांतरण प्रयुक्त किया गया:

1. प्रसरण स्थायित्व के लिए
2. सममित रूपांतरण (symmetrizing transformation)
सममित रूपांतरण (symmetrizing transformation) का अनुभव बड़ा जटिल रहा; पहले सूत्र व्युत्पन्न किया गया फिर कुछ आँकड़ों की उत्पत्ति कर उसकी जांच की गयी।

परिणाम यह पाया गया की यदि प्रतिदर्श का आकार 50 से ज्यादा है तो सममित रूपांतरण (symmetrizing transformation) का बंटन प्रसामान्य है। यदि प्रतिदर्श

का आकार 50 से कम है तो सममित रूपान्तरण (symmetrizing transformation) का बंटन प्रसामान्य एवं कार्ई-स्क्वेअर (Chi-square) का मिश्रित सूत्र है इसके बाद कृषि के आंकड़ों पर इसे प्रयुक्त किया गया।

3 व्युत्क्रम गाऊसी वितरण से प्रतिदर्श प्रसरण गुणांक के परिवर्तन को स्थिर करने वाला सममिति और विचरण

अब दूसरी समष्टि है व्युत्क्रम गौसीयन (inverse Gaussian), यह गौसीयन से ज्यादा उपयोगी है क्योंकि यह माना जाता है कि प्रसरण गुणांक का उपयोग केवल धनात्मक मापदंड के लिए किया जाता है ना कि ऋणात्मक आँकड़ों के लिए जबकि व्युत्क्रम गौसीयन धनात्मक रेंज में ही रहता है। वास्तविक परिस्थितियों से संबन्धित आँकड़े अक्सर धनात्मक होते हैं, अतः व्युत्क्रम गौसीयन काफी प्रयुक्त होने वाला बंटन है। हम लोग प्रसामान्य बंटन को बहुत ज्यादा प्रयोग करते हैं और व्युत्क्रम गौसीयन बंटन का प्रयोग कम करते हैं।

व्युत्क्रम गौसीयन बंटन इस तरह से दिखता है:

$$f(X|\mu, \lambda) = \left\{ \frac{\lambda}{2\pi x^3} \right\}^{1/2} \exp \left\{ -\frac{\lambda}{2\mu^2 x} (x - \mu)^2 \right\}; x > 0, \mu > 0, \lambda > 0$$

$$\text{प्रसरण गुणांक (CV)} \theta = \sqrt{\mu/\lambda}$$

मान लें व्युत्क्रम गौसीयन (inverse Gaussian) IG $(\mu, \phi = \mu/\lambda = \theta^2)$, से प्राप्त यादृच्छिक प्रतिदर्श x_1, x_2, \dots, x_n , आकार n का प्रतिदर्श है, प्रसरण गुणांक (सीवी) का आकलन

$$\hat{\phi} = \hat{\mu} \left(\frac{1}{\lambda} \right) = \bar{X}U$$

जहां प्रतिदर्श प्रसरण गुणांक $\hat{\theta} = \sqrt{\hat{\phi}}$, औसत $\bar{X} = \sum_{i=1}^n x_i$, एवं मानक विचलन $U = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{X})^2}{(n-1)}}$ है।

इस शोध पत्र में दो रूपान्तरणों (सममित रूपान्तरण एवं प्रसरण सममित रूपान्तरण (Variance Symmetrizing Transformation) के गुणों को देखा गया है। विशाल आकार के प्रतिदर्श से एक और उपयोगी संप्रयोग है कि, प्रसरण गुणांक के परीक्षण के लिए हमें एक अन्य मात्रा की आवश्यकता पड़ती है जिसका प्रतिपादक (power) के रूप में प्रयोग करके

निम्न राशि का बंटन प्राप्त कर लेते हैं :

$$Z = \frac{(\hat{\phi}/\phi_0)^{h_0} - \mu'_{1h_0}}{\sigma(h_0)} \sim N(0,1) \text{ under } H_0: \phi = \phi_0$$

यह रूपान्तरण बहुत उपयोगी है हम h_0 को आंकलित कर सकते हैं तो सांख्यिकी प्रसामान्य बंटन होगा।

4. व्युत्क्रम गाऊसी वितरण से एक सामान्य प्रसरण गुणांक का अनुमान और परीक्षण

इसके बाद हम आते हैं जीनोटाइप×वातावरण परस्पर क्रिया (G×E interaction) पर, जिस संदर्भ में बहुत से जीनोटाइप का बहुत से वातावरण में मूल्यांकन करते हैं तथा उपज हमेशा धनात्मक ही होती है। अतः व्युत्क्रम गौसीयन बंटन बहुत उपयोगी है। यहाँ यह देखते हैं कि प्रसरण गुणांक बराबर है कि नहीं। यहाँ दो प्रश्न हैं:

1. इसे किस विधि से जांचा जाए ?
2. यदि बराबर है, तो उभयनिष्ठ प्रसरण गुणांक क्या होगा ?

इसके लिए कुछ आँकड़ों की उत्पत्ति कर, कुछ निष्कर्ष निकाले गए हैं। पावर रूपान्तरण (Power Transformation) बहुत उपयोगी पाया गया है परंतु

- 1 अभी भी यह कार्य अधूरा है, इसमें 6 सांख्यिकीय सूत्र निकाले/खोजे गए थे, जिनमें से एक ठीक कार्य कर रहा है बाकी 5 को ठीक प्रारूप में लाये जाने की अभी आवश्यकता है।
- 2 व्युत्क्रम गौसीयन बंटन से भिन्न आँकड़े हो तो क्या करना होगा ?

5 प्रसरण गुणांक पर महत्व के साथ व्युत्क्रम गाऊसी डेटा के लिए बेयसियन अनुमान

फिर आता है विशाल आँकड़ों (big data) के संकलन जो कि वर्तमान में सभी संस्थाएं कर रही है। भा.कृ. अनु.प. – भारतीय कृषि सांख्यिकीय अनुसन्धान संस्थान बहुत आँकड़ों का संकलन कर कोष बनाती है। बेयसियन अनुमान यह दो प्राचलों के लिए किया गया है, एक है औसत दूसरा है प्रसरण गुणांक। औसत

और प्रचलन गुणांक के दो पूर्वगामी बंटन का अध्ययन किया गया है और उनका पश्च बंटन और पश्च औसत निकला गया है।

6 फसल परीक्षणों में प्रतिकृतियों की संख्या का बेयसियन निर्धारण

छठवा शोध पत्र मेरे शोध छात्र, सीराज ने किया है और यह पाया है कि यदि प्रसरण गुणांक 16% से कम है तो प्रतिकृति कि संख्या बेयसियन के सापेक्ष कम लगती है और यदि प्रसरण गुणांक 16% से ज्यादा है तो बेजीयन मे ज्यादा प्रतिकृति कि संख्या लगती है।

संदर्भ

- सिंह, एम. (1983). सैंपलिंग डिस्ट्रिब्यूशन ऑफ सीवी फ़्रोम सेवेरल डिस्ट्रिब्यूशनस। संख्या, सीरीज बी, 55: 65–76
- चौबे, वाई. पी., सिंह, एम. एवं सेन, डी. (2013). ऑन सीमेट्राइजिंग ट्रान्स्फ़ोर्मेशन ऑफ द सैंपल कोइफिशिएंट ऑफ वैरिएसन फ़्रोम ए नॉर्मल पोपुलेशन। कोम्युनिकेसन्स इन स्टैटिस्टिक्स – सिमुलेसन एंड कौंप्यूटेसन, 42, 2118–2134
- चौबे, वाई. पी., सिंह, एम. एवं सेन, डी. (2017). सीमेट्राइजिंग एंड वैरिएन्स स्टेबिलाइजिंग ट्रान्स्फ़ोर्मेशनस ऑफ सैंपल कोइफिशिएंट ऑफ वैरिएशन फ़्रोम एन इनवर्स गौसीयन डिस्ट्रिब्यूशन।

संख्या, सीरीज बी 79: 217–246.

- चौबे वाई. पी., सिंह एम. एवं सेन डी. (2021). बायेसियन इनफरेंस फॉर इनवर्स गाऊसी डेटा विथ एम्फैसिस ऑन द कोफीफिकेशन ऑफ वैरिएशन। इन: चौबे वाई. पी., लाहमीरी एस., नेबेबे एफ., सेन ए. (संस्करण) एप्लाइड स्टैटिस्टिक्स एंड डेटा साइंस। सीसीएस 2021। गणित और सांख्यिकी में सिंगर कार्यवाही, वॉल्यूम 375, सिंगर, चाम। <https://doi.org/10.1007/978-3-030-86133-9-4>
- सिंह एम., चौबे वाई. पी., सेन डी. एवं सरकार ए. (2021). एस्टिमेशन एंड टेस्टिंग ऑफ ए कॉमन कोइफिशिएंट ऑफ वैरिएसन फ़्रोम इनवर्स गौसीयन डिस्ट्रिब्यूशन, इन: चौबे वाई.पी., लाहमीरी एस., नेबेबे एफ., सेन ए. (संस्करण) एप्लाइड स्टैटिस्टिक्स एंड डेटा साइंस। सीसीएस 2021। गणित और सांख्यिकी में सिंगर कार्यवाही, वॉल्यूम 375, सिंगर। https://doi.org/10.1007/978-3-030-86133-9_5
- ओमर, एस. ओ., अब्दुल्ला ए. डब्ल्यू. एच., सरकार, ए. एवं सिंह, एम. (2014). बेसियन डिटेर्मिनेशन ऑफ द नंबर ऑफ रेप्लीकेसन्स इन क्रॉप ट्रायल्स। यूरोपियन जर्नल ऑफ़ एक्सपेरिमेंटल बायोलॉजी 4(6):129–133।

सांख्यिकी-विमर्श

2021

अंक
17

अनुसंधान
खण्ड

मिनी-लाइसीमीटर द्वारा सिंचाई के विभिन्न स्तरों पर ग्रीष्म ग्वार के वास्तविक वाष्पोत्सर्जन का अध्ययन

हरिमोहन मीना, रंजय कुमार सिंह, प्रियव्रत सान्त्रा एवं सी.बी. पाण्डेय
भा.कृ.अनु.प.—केन्द्रीय शुष्क क्षेत्र अनुसंधान संस्थान, जोधपुर—342003

ग्वार (सायमोप्सिस टेट्रागोनोलोबा) दलहनी परिवार की फसल है जो बहुत ही उच्च तापमान एवं सूखे को सहन कर सकता है। इसकी खेती किसी-न-किसी रूप में देश के विभिन्न भागों, खासकर शुष्क तथा अर्द्ध-शुष्क क्षेत्रों में की जाती है। पश्चिमी राजस्थान के शुष्क क्षेत्रों के किसानों की आय के साधन के रूप में यह एक महत्वपूर्ण फसल है। शुष्क क्षेत्रों में पानी की उपलब्धता में कमी के कारण कृषि उत्पादन के लिए जल के उपयोग एवं जल संसाधन का महत्व अधिक है। गर्म शुष्क क्षेत्रों में सतह ऊर्जा संतुलन, जल उपयोग की गणना और दैनिक तथा मौसमी जल बजट के मूल्यांकन के लिए महत्वपूर्ण है। शुष्क क्षेत्रों में जल के आदान-प्रदान की प्रक्रिया मृदा की सतह और वातावरण के बीच एक महत्वपूर्ण भूमिका निभाता है। मिनी-लाइसीमीटर का उपयोग फसल का वास्तविक वाष्पोत्सर्जन, जल निकासी, जल के अन्य संतुलन घटकों और मृदा नमी में छोटे से परिवर्तन की गणना में करते हैं। मिनी-लाइसीमीटर के द्वारा छोटे क्षेत्रों से वाष्पीकरणीय प्रवाह की माप होती है। स्थापना के दौरान पर्यावरण में कम अवरोध और सस्ता होना, इसके दूसरे लाभ हैं। मिनी-लाइसीमीटर लोड सेल अंशांकन की सुदृढ़ विधि के साथ वाष्पोत्सर्जन को उचित सटीकता से माप सकता है। उपरोक्त तथ्यों को ध्यान में रखते हुए इस प्रयोग में जल घटक अवयवों के प्रक्षेत्र मापन के लिए मिनी-लाइसीमीटर को विकसित किया गया। मिनी-लाइसीमीटर का आयाम $0.50 \times 0.50 \times 0.55$ मी. था। अध्ययन में सिंचाई के चार स्तर रखे गए जो 50 मिमी. संचयी पैन वाष्पीकरण के आधार पर 100, 80, 60 एवं 40% थे। भा.कृ.अनु.प.—केन्द्रीय शुष्क क्षेत्र अनुसंधान संस्थान जोधपुर के प्रायोगिक क्षेत्र

में 2015–2017 की गर्मियों के मौसम में विशेष रूप से विकसित मिनी-लाइसीमीटर के द्वारा विभिन्न सिंचाई स्तरों पर ग्रीष्मकालीन ग्वार फसल के लिए जल के वास्तविक मापन के लिए एक अध्ययन किया गया।

ग्वार सीड के वैश्विक उत्पादन में भारत का योगदान 80% है जहां इसकी खेती 3.40 मिलियन हेक्टेयर के क्षेत्र में की जाती है, जिसकी उत्पादकता 580 किलोग्राम / हेक्टेयर (NRAA, 2014) से बहुत कम है। ग्वारफली (क्लस्टरबिन) एक नकदी फसल है, जो राजस्थान, हरियाणा और गुजरात के अर्ध-शुष्क और शुष्क क्षेत्रों में बरसात (खरीफ) के मौसम में उगाई जाती है। अंतरराष्ट्रीय बाजार में इसकी मांग के कारण, इसे आंध्र प्रदेश, तमिलनाडु, कर्नाटक, महाराष्ट्र और छत्तीसगढ़ जैसे गैर-पारंपरिक क्षेत्रों में भी इसकी पैदावार आरम्भ की गयी है। इसके अलावा, इसकी खेती गर्मियों के दौरान सिंचित परिस्थितियों में भी की जा रही है (भट्ट एट अल, 2017; मंजुनाथ एट अल, 2018)। यह फसल पोषक तत्वों की भरपाई करने का एक स्रोत है और विशेष रूप से यह कम उर्वरता वाली मिट्टी की नाइट्रोजन और नमी में कमी का सामना कर सकती है। यह एक सहज फसल है (राज सिंह, 2014) और इसे गर्मी के मौसम में सिंचाई की सुविधा के साथ भी उगाया जा सकता है। हालांकि, राजन्ना एट अल (2016) ने बताया कि सिंचाई के साथ बेड प्लांटिंग (FIRBS) को अपनाया क्लस्टरबीन में 0.80 का संचयी पैन वाष्पीकरण पर उपज मापदंडों के साथ-साथ विकास मापदंडों में वृद्धि की ओर जाता है, लेकिन इस भाग में तापमान होने पर ग्रीष्मकालीन ग्वार पर कोई अध्ययन नहीं किया गया था। वर्तमान

अध्ययन का उद्देश्य मिनी—लाइसीमीटर प्रणाली के तहत मिनी—सिंप्रकलर के माध्यम से ग्रीष्मकालीन क्लस्टरबिन को प्रदान किए गए पानी के विभिन्न स्तरों पर फसल जल के उपयोग या वास्तविक वाष्पीकरण की मात्रा का लक्ष्य है।

मिनी—लाइसीमीटर का डिजाइन और निर्माण

मिनी—लाइसीमीटर का डिजाइन और निर्माण, $0.50 \times 0.50 \times 0.55$ मी. के आयाम के साथ इलेक्ट्रॉनिक एकल लोड सेल के आधार पर करने के उपरांत संस्थान के प्रयोगात्मक क्षेत्र में स्थापित किया। इलेक्ट्रॉनिक तौल—आधारित मिनी—लाइसीमीटर में तीन घटक थे—(I) मिनी—लाइसीमीटर टैंक (II) एकल लोड सेल आधारित इलेक्ट्रॉनिक वजन मापने वाली मशीन तथा (III) मिनी—लाइसीमीटर कक्ष। टैंक में मिनी—लाइसीमीटर टैंक, जल अपवाह प्लेट, जल अपवाह संग्रहण टैंक, जल अपवाह छिद्र, पैर और हैंडल शामिल थे। जल अपवाह प्लेट को $(0.49 \text{ मी.} \times 0.49 \text{ मी.} \times 4 \text{ मिमी. मोटाई})$ मिनी—लाइसीमीटर टैंक के अंदर रखा गया। यह हार्ड रोल्ड शीट की मदद से बना था जिसमें 49 छिद्र (प्रत्येक छिद्र का व्यास 6 मिमी.) (चित्र 1अ) थे। इलेक्ट्रॉनिक वजन तौलनेवाली मशीन (चित्र 1ब) मिनी—लाइसीमीटर का सबसे महत्वपूर्ण हिस्सा है। यह लोड सेल (मेक—एडीआइ कंपनी इंडियाय रेग्युलर 410) पर आधारित है। प्रक्षेत्र में 1 मी. \times 1 मी. \times 1 मी. आयाम के गड्डों को पी. सी. सी. के माध्यम से बनाया गया जिसका अंतिम आयाम जमीनी स्तर के 10 सेमी. ऊपर $0.62 \text{ मी} \times 0.62 \text{ मी.} \times 1 \text{ मी.}$ रखा गया (चित्र 1स) और इसमें 10 सेमी. व्यास का एक छिद्र रखा गया ताकि टैंक से पानी का समुचित निकास हो सके। सभी भागों को एक साथ जोड़कर इलेक्ट्रॉनिक वजनी एकल लोड सेल आधारित मिनी—लाइसीमीटर को संचालित किया गया (चित्र 2)। एक मिनी—लाइसीमीटर के निर्माण की लागत रु. 14,500/— था।

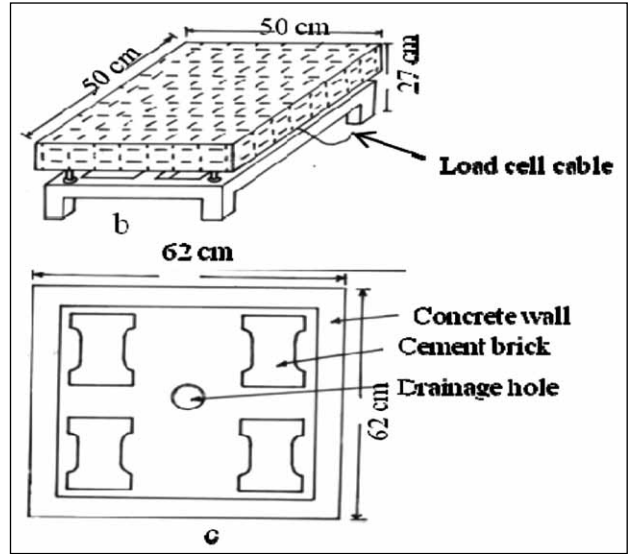
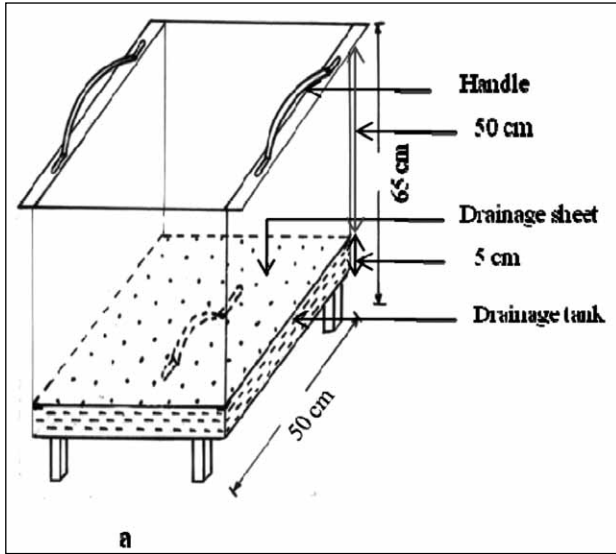
लोड सेल आधारित मिनी—लाइसीमीटर का सत्यापन

प्रक्षेत्र में स्थापित मिनी—लाइसीमीटर को कुल चार दिनों, 7 से 10 सितम्बर, 2014 के लिए, जिसमें, 9

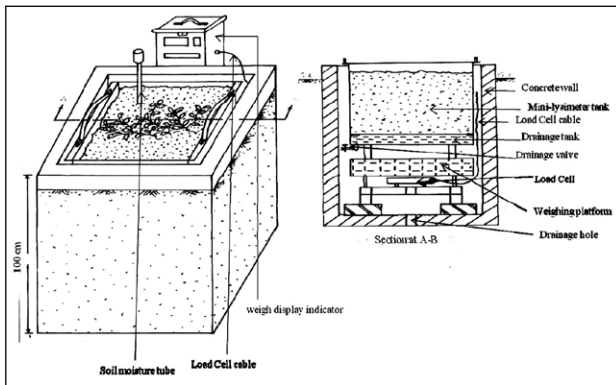
सितम्बर को 23.8 मिमी. वर्षा हुई थी, सत्यापित किया गया (तालिका 1)। इससे पहले, 2 से 4 सितम्बर, 2014 के दौरान 130 मिमी. का वर्षापात रिकॉर्ड किया गया था। लोड सेल द्वारा प्रदर्शित मिनी—लाइसीमीटर का वजन प्रत्येक दिन सुबह 8:30 बजे गणना कर मिनी—लाइसीमीटर के वजन में शुद्ध परिवर्तन दर्ज किया गया। 7 से 8 सितम्बर के दौरान, मिनी—लाइसीमीटर के वजन में शुद्ध परिवर्तन -1.70 किग्रा. प्रतिदिन पाया गया जोकि 6.8 मिमी. के समतुल्य है। चूँकि, इस अवधि में वर्षापात नहीं हुआ और जल अपवाह भी नहीं पाया गया, इसलिए, मिनी—लाइसीमीटर के वजन में परिवर्तन सिर्फ वाष्पोत्सर्जन की वजह से था। इसके लिए एक दिन के लिए फसल बायोमास वजन को नजरअंदाज किया गया। इसी प्रकार, 9 सितम्बर को, जब 23.8 मिमी. वर्षापात हुआ और जल अपवाह 2.73 मिमी. हुआ, मिनी—लाइसीमीटर के वजन में 4.80 किग्रा. का परिवर्तन पाया गया। इस प्रकार 21.7 मिमी. का प्रभावी वर्षा जोकि मिट्टी में नमी भंडारण वृद्धि और वाष्पोत्सर्जन में इस्तेमाल हुआ। मिट्टी की नमी भंडारण में वृद्धि (19.2 मिमी.) मिनी—लाइसीमीटर के वजन में रूपान्तरण कारक से गुणा करने के बाद गणना की गयी। इस प्रकार, 9 सितम्बर की सुबह से 10 सितम्बर की सुबह तक, जोकि वर्षावाला दिन था, वाष्पोत्सर्जन की अनुमानित राशि 2.05 मिमी. के रूप में दर्ज की गयी। विकसित मिनी—लाइसीमीटर का प्रक्षेत्र में अंशांकन कारक 0.2 मिमी. के विभेदन के साथ 1 किग्रा. = 4 मिमी. है। यह देखा गया कि एकल लोड सेल आधारित मिनी—लाइसीमीटर का निर्माण पारंपरिक लाइसीमीटर की तुलना में किफायती है। इसलिए, इसको प्रयोग क्षेत्र स्थिति में पुनरावृत्त और आसानी से स्थापित किया जा सकता है।

फसल वाष्पोत्सर्जन माप

इस अध्ययन में वास्तविक फसल वाष्पोत्सर्जन (ईटी) 100% सिंचाई (टी 1) स्तर पर अधिकतम और सबसे कम 40% सिंचाई (टी 4) करने पर पाया गया। तीन वर्षों के दौरान औसतन वाष्पोत्सर्जन (ईटी) टी 1 में (686 मिमी) और टी 4 में (340 मिमी) था। जैसा कि



चित्र 1: मिनी-लाइसीमीटर का प्रारूप (अ) मिनी-लाइसीमीटर टैंक (ब) एकल लोड सेल आधारित इलेक्ट्रॉनिक वजन मापने वाली मशीन तथा (स) मिनी-लाइसीमीटर कक्ष



चित्र 2: मिनी-लाइसीमीटर का त्रिविम दृश्य एवं पार अनुभाग

तालिका 2 से देखा जा सकता है कि जैसे जैसे सिंचाई के पानी की मात्रा कम की गयी है उसी के अनुसार फसल उत्पादन में कमी हुई है। फसल की खेती और फसल की वृद्धि का दृश्य (प्लेट 1)। यहाँ पर यह देखा गया कि अलग अलग सिंचाई के पानी की मात्रा पर फसल उत्पादन में कमी का अनुपात क्या रहा है। 20% सिंचाई जल कम करने पर फसल वाष्पोत्सर्जन में 19% की कमी हुई है जबकि फसल की उपज में 10% की कमी हुई है। 20% से ज्यादा सिंचाई जल कम करने पाया गया कि फसल की उपज में ज्यादा कमी हुई है। टी-3 (60% सिंचाई) ओर टी-4 (40% सिंचाई)

तालिका 1: विभिन्न अवसरों पर वर्षा की घटनाओं के दौरान लोड सेल का सत्यापन

दिनांक	मिनी-लाइसीमीटर भार (कि ग्रा)			मिनी-लाइसीमीटर भार में शुद्ध परिवर्तन (कि ग्रा)			वर्षा (मि मी.)	जल निकासी (मि मी.)
	1	2	3	1	2	3		
07-09-2014	192.60	18.45	18.75	—	—	—	0	0
08-09-2014	190.85	16.70	17.25	-1.75	-1.75	-1.50	0	0
09-09-2014	189.05	15.00	15.55	-1.80	-1.70	-1.70	0	0
10-09-2014	193.90	19.80	20.30	+4.85	+4.80	+4.75	23.8	2.73*

*जल निकास अवयव तीन मिनी- लाइसीमीटर के औसत को दर्शा रहा है।

तालिका 2 : ग्रीष्म ग्वार में वाष्पोत्सर्जन, उत्पादन एवं जल उपयोग

ट्रीटमेंट्स	वास्तविक वाष्पो. (मिमी.)				वास्तविक वाष्पो. में कमी (%)	उपज (किग्रा./है.)				उपज में कमी (%)	जल उपयोग दक्षता (किग्रा./है./मिमी.)
	2015	2016	2017	औसत		2015	2016	2017	औसत		
टी-1 (100% सिंचाई)	662	723	663	686		2101	2651	1694	2149		
टी-2 (80% सिंचाई)	574	577	512	554	19	2082	2184	1507	1924	10	0.31
टी-3 (60% सिंचाई)	442	425	494	454	34	1076	1516	764	1119	48	0.35
टी-4 (40% सिंचाई)	313	299	407	340	50	359	1086	680	708	67	0.21


प्लेट 1: प्रक्षेत्र में ग्रीष्म ग्वार की फसल

करने पर फसल की उपज में क्रमशः 48% और 67% की कमी हुई (तालिका 2)। सिंचाई के विभिन्न स्तरों पर जल-उत्पादकता का आकलन किया गया जो 0.21 से 0.35 किलो/घन मीटर तक पाया गया। सबसे ज्यादा जल-उत्पादकता 80% (टी 2) सिंचाई स्तर पाया गया। इस अध्ययन में यह पाया गया है कि 10% ग्वारफली की उपज को कम कर 20% सिंचाई जल बचाया जा सकता है जिससे शुष्क क्षेत्र में सिंचित क्षेत्र को बढ़ाया जा सकता है।

निष्कर्ष

शुष्क तथा अर्द्ध-शुष्क क्षेत्रों में कृषि उत्पादन प्रणालियों में मृदा जल की उपलब्धता एक प्रमुख सीमित कारक है। पानी की आपूर्ति, पूर्ण और सीमित, स्थानीय वातावरण में फसल की प्रतिक्रिया का ज्ञान जल प्रबंधन और फसल उत्पादकता में सुधार के लिए प्रभावी सिंचाई रणनीतियों के विकास में सहायता कर सकता है। इस अध्ययन से यह निष्कर्ष निकलता है कि 80% सिंचाई स्तर पर ग्वार उत्पादन में बहुत ज्यादा कमी (10%) नहीं होती है, जबकि 20% सिंचाई जल की बचत होती है। इसलिए, शुष्क क्षेत्रों में, जहाँ जल एक अत्यंत ही कीमती प्राकृतिक श्रोत है, 80% सिंचाई स्तर को अनुशंसित किया जा सकता है।

संदर्भ

- मंजुनाथ, एन., लोकेश, एच. एवं देशमान्या जाग्रथि बी. (2018). स्ट्रक्चरल चेंजेज इन दी परफोरमेंस ऑफ गम ग्वार इन इंडिया. इंडियन जर्नल ऑफ एग्रिकल्चरल रिसर्च, 52(3):336-338।
- एन.आर.ए.ए. (2014). पोर्टेशियल ऑफ रेनफेड ग्वार (क्लस्टरबीन) कल्टीवेशन, प्रोसेसिंग एंड एक्सपोर्ट इन इंडिया। पॉलिसी पेपर न.3। नेशनल रेनफेड एरिया अथॉरिटी, नास कॉम्प्लेक्स, डीपीएस मार्ग, न्यू दिल्ली, पीपी 109।
- रजन्ना जी. ए., धीवल, ए. एस. एवं श्रीहर्षा, वी. पी. (2016). परफोरमेंस ऑफ क्लस्टरबीन (स्योप्सिस टेटरगोनलोबा) अंडर वैरिएबल इरीगेशन शेड्यूल्स एण्ड क्रॉप इस्टैब्लिशमेंट टैक्नीक। इंडियन जर्नल ऑफ एग्रोनोमी, 61(2):223-230।
- सिंह, आर (2014). इम्प्रूड कल्टीवेशन प्रैक्टिस फॉर क्लस्टरबीन इन खरीफ एंड समर सीज़न। बुलेटिन ऑन : http://www.cazri.res.in/publications/Guar_RajSingh.pdf

“

आद्रा बरसे पुनर्वसुजाय, दीन अन्न कोऊ न खाय।।

अर्थ : यदि आद्रा नक्षत्र में वर्षा हो और पुनर्वसु नक्षत्र में पानी न बरसे तो ऐसी फसल होगी कि कोई दिया हुआ अन्न भी नहीं खाएगा।

-घाघ

”

वृक्ष-तंत्र संतुलन से युक्त कृषि-वानिकी परीक्षण

टी.बी. पीटर⁰, सिनी वर्गीस¹, सीमा जग्गी², मोहम्मद हारून¹, एल्दो वर्गीस³ एवं देवेन्द्र कुमार¹

⁰पूर्व पीएच.डी. विद्यार्थी, भा.कृ.अनु.प.—भारतीय कृषि सांख्यिकी अनुसंधान संस्थान,
लाइब्रेरी एवेन्यू, नई दिल्ली—110012

¹भा.कृ.अनु.प.—भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, लाइब्रेरी एवेन्यू, नई दिल्ली—110012

²कृषि शिक्षा विभाग, भा.कृ.अनु.प., नई दिल्ली—110012

³भा.कृ.अनु.प.—केन्द्रीय समुद्री मात्स्यिकी अनुसंधान संस्थान, कोची, केरल—682018

1. परिचय

कृषि, खाद्य फसलों, पशुधन एवं उपभोक्ताओं की बढ़ती माँग पूरी करने के लिए कच्चे माल के उत्पादन के लिए संपूर्ण विश्व में किया जाने वाला बहुत पुराना व्यवसाय है। लघु एवं सीमांत किसान सामान्यतः जोखिम कम करने की रणनीति के तहत एक ही भूमि पर एक से अधिक फसल उगाते हैं (कहन, 2008), और वे अपने खेतों में उग रहे वृक्षों को सहर्ष नहीं काटते। किसान अनजाने में ही फलों की व्यवस्था, भूमि उपादेयता बनाए रखने एवं फसलों पर अनूकूल प्रभावों को बनाए रखने के लिए फसल प्रणाली में वृक्षों का उपयोग करते हैं। अतः वृक्षों द्वारा लिए जा रहे लाभ सर्वोत्तम हैं जो कि कृषि उपादेय भूदृश्य में निरंतर लगाए जा रहे वृक्षों द्वारा हैं और इस व्यवसाय को कृषि-वानिकी के नाम से जाना जाता है।

‘कृषि-वानिकी’ शब्द का उपयोग भूमि-उपयोग प्रणाली एवं प्रौद्योगिकियों, जिनमें जंगली सदाबहार वृक्षों, फल वृक्षों, बाँस, झाड़ियों इत्यादि सहित एक ही भूमि पर जानबूझकर फसलों और कभी-कभी पशुधन के साथ-साथ स्थान व्यवस्था अथवा अस्थायी अनुक्रम के किसी रूप में (लंग्रेन एवं रेंट्री, 1982) उगाये जाने के लिए एक संयुक्त नाम के लिए किया गया है।

कृषि-वानिकी कार्य का सरोकार वृक्षों एवं फसलों, वृक्षों एवं पशुधन, अथवा वृक्षों, फसलों एवं पशुधन का सुविचारित संयोजन से है, जिसमें एक समेकित कृषि-पारितंत्र की तरह घटकों के मध्य अंतः क्रिया का

सघन प्रबंधन सम्मिलित होता है (ध्यानी एवं हाण्डा, 2014)। सम्मिलित घटकों के प्रकार पर आधारित कृषि-वानिकी प्रणाली के तीन मुख्य प्रकारों के नाम एग्रीसिलवोकल्चर (वृक्षों एवं फसलों का संयोजन), सिलवोपासचरल (वृक्षों, चारा एवं पशुधन का संयोजन) तथा एग्रीसिलवोपासचरल (वृक्षों, फसलों एवं पशुधन का संयोजन) हैं (डी बेट्स इत्यादि, 2007)।

2. कृषि-वानिकी का महत्व

कृषि-वानिकी कार्य में, केंद्र बिन्दु फार्मों पर एवं ग्रामीण भूदृश्य में उगे हरे वृक्षों के विस्तृत क्षेत्र हैं जो अधिक विविध निर्माण, उपादेय, लाभप्रदता, स्वस्थ एवं दीर्घकालिक भूमि उपयोग प्रणालियों के माध्यम से बहुत सारे लाभ दिला सकते हैं। बहुघटकों (वृक्षों, फसलों, पशुधन इत्यादि) की अन्तः क्रिया एवं व्यवस्था निश्चित रूप से प्रणाली की प्राप्तियों को प्रभावित करती है।

यद्यपि यहाँ पर प्रयोजन यह है कि वास्तव में परंपरागत कृषि कार्यों की अपेक्षा कृषि-वानिकी प्रणाली कितनी उपादेय है। कृषि-वानिकी में मुख्य मुद्दा यह नहीं है कि किसी दी गई भूमि की इकाई को बहुत सारे कृषि-वानिकी घटकों से भरना है अपितु यह जानना है की घटक कितनी सावधानी से चुने जायें और उचित व्यवस्था और प्रबंधन द्वारा प्राप्तियाँ इष्टतम हों। अतएव, कृषि-वानिकी परीक्षण में समुचित अभिकल्पनाओं का उपयोग महत्वपूर्ण है। बहुत से घटकों के निहित होने के कारण, कृषि-वानिकी परीक्षणों के सभी प्रकारों हेतु किसी विशेष अभिकल्पना की अनुशंसा करना बहुत ही

चुनौतीपूर्ण है।

सहज विचारणीय बिन्दु में, अभिकल्पना का विकल्प कृषि-वानिकी प्रणाली के प्रकार, निहित घटकों की संख्या एवं परीक्षण के उद्देश्य का फलन है। नायर (1993) के अनुसार कृषि-वानिकी में परीक्षणों का वृहद वर्गीकरण इस प्रकार किया जा सकता है :

- (i) **बहुउद्देश्य वृक्ष छाँटना एवं चयन परीक्षण:** इसमें बहुत सारे विश्वस्त जर्मप्लाज्म छाँटना निहित है (अधिकतर बहुत सारी प्रजातियाँ, किन्तु कभी-कभी एक या दो प्रजातियों की किस्में)।
- (ii) **घटक एवं प्रणाली प्रबंधन परीक्षण:** इस श्रेणी का सरोकार विशिष्ट कृषि-वानिकी प्रौद्योगिकियों के सुधार से है।
- (iii) **घटकों की अन्तः क्रिया अध्ययन:** इसमें भूमि के नीचे अथवा ऊपर के संसाधनों की साझेदारी अन्वेषण के लिए कृषि-वानिकी प्रणालियों (अधिकतर वृक्ष-फसल अन्तः क्रियाओं) के घटकों के मध्य अंतरसंबंधों को मात्रात्मक रूप देना और इसकी समझ करवाने से संबंधित अध्ययन निहित है।
- (iv) **प्रतिमान मूल्यांकन परीक्षण:** इनका उपयोग आदर्श खेत स्थितियों के अंतर्गत कृषि-वानिकी प्रौद्योगिकियों के विशिष्ट पैकेजों के मूल्यांकन करने हेतु किया जाता है। कृषि-वानिकी प्रणालियों की समझ केवल तभी हो सकती है जब प्रत्येक प्रणाली का अध्ययन सावधानीपूर्वक हो। प्रणालियों का अध्ययन उचित परीक्षणों के माध्यम से वैज्ञानिक ढंग से किया जाये।

3. कृषि-वानिकी परीक्षणों में उपयोग की जाने वाली अभिकल्पनाएँ

कृषि-वानिकी में अन्वेषण के संचालन में बहुत सी परीक्षण अभिकल्पनाएँ उपयोग की गयी हैं। इनमें सबसे पुरानी में फ़ैन अभिकल्पनाएँ, ल-अभिकल्पनाएँ, स्टार-अभिकल्पना एवं बीहाइव अभिकल्पनाएँ जो कि विशिष्ट कृषि-वानिकी परीक्षणों के लिए हैं, सम्मिलित हैं। इन सभी में सर्वाधिक अभयनिष्ट सम्पूर्ण यादृच्छिक अभिकल्पना (CRD) है। वृहद रूप से उपयोग

किया जाने वाला एक और अभिकल्पनाओं का वर्ग यादृच्छिक सम्पूर्ण खण्ड अभिकल्पना (RCBD) है। RCBD का एक उन्नत किन्तु कठिन रूप स्प्लिट-प्लॉट अभिकल्पना (SPD) है जिसे कृषि-वानिकी परीक्षणों में बहुत उपयोग किया गया है।

जग्गी इत्यादि (2007) ने कृषि-वानिकी हेतु कुछ निकटवासी संतुलित खण्ड अभिकल्पनाएँ (NBBDD) प्रस्तावित की जिनमें T वृक्ष प्रजातियाँ T स्तरों के साथ, C फसल प्रजातियाँ c स्तरों के साथ, एवं अभिकल्पनाएँ वृक्षों के स्तरों हेतु संतुलित हैं। जब भी परीक्षात्मक सामाग्री में विषमता दो स्रोतों तक पता लगाने योग्य समझी जाती है, परीक्षात्मक प्लॉटों को कतार-स्तम्भ सारणी में व्यवस्थित किया जा सकता है, जो कतार-स्तम्भ अभिकल्पनाओं के नाम से जानी जाती हैं।

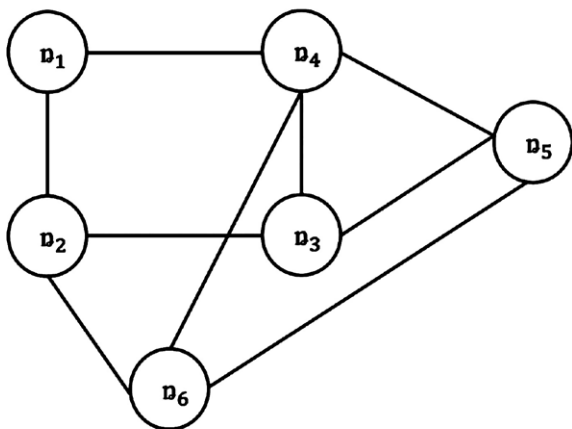
कृषि-वानिकी प्रणाली को समझने एवं सुधारने के उद्देश्य से यह महत्वपूर्ण है कि विचरण के स्रोतों की उचित विश्वसनीयता के माध्यम से रुचि के प्रभावों के अधिक सटीक आंकलन को सुनिश्चित करने के लिए उठाए गए कदम उपयुक्त अभिकल्पना उपलब्ध करती है। अतः प्रस्तुत अध्ययन, रैखिक तंत्र प्रभाव मॉडल का उपयोग कृषि-वानिकी परीक्षणों हेतु अभिकल्पनाएँ विकसित करने के लिए किया गया है।

4. तंत्र परिकल्पना

तंत्र परिकल्पना लेखाचित्र सिद्धांत से ली गयी है। लेखाचित्र किसी तंत्र अथवा विषयों/वस्तुओं के अंतर्संबद्ध संग्रह को निरूपित करने का औपचारिक तरीका है। लेखाचित्रों का प्रचलित अनुप्रयोग परिवहन, संगणक विज्ञान एवं तंत्र, विद्युत अभियांत्रिकी, भौगोलिक तंत्र जैसे कि खेत में वृक्षों की व्यवस्था, अप्रत्यक्ष ऑनलाइन समूह जैसे कि फ़ेसबुक, लिंकडइन इत्यादि के साथ-साथ वर्तमान महामारी Covid-19 में संपर्क दूढ़ने के लिए किया जा सकता है। चार्टरेंड एवं झांग (2012) के अनुसार लेखाचित्र का संबंध शीर्षों के एक निश्चित अरिक्त समुच्चय(N) से है जो कि रेखाओं के एक समुच्चय जिन्हें छोर (E) कहा जाता है, से सम्बद्ध है। गणित में इसे $G = (N, E)$ के रूप में लिखा जा

सकता है, जहां पर $N(G) = \{n_1, n_2, \dots, n_n\}$ n शीर्षों, जिन्हें ग्रंथि (Node) कहते हैं, का एक समुच्चय है, $E(G) = \{e_1, e_2, \dots, e_p\}$ p छोरों अथवा कड़ियों का एक समुच्चय है तथा G , n क्रम एवं p आकार का एक लेखाचित्र है। G में कड़ियों के माध्यम से छोरों की विशेष व्यवस्था एक तंत्र की रचना करती है। यदि n_i एवं n_k ग्रंथियां, एक छोर द्वारा सीधे तौर पर सम्बद्ध हैं तो n_i एवं n_k को G में एक दूसरे के साथ लगे हुए अथवा निकटवासी कहे जाते हैं।

समान्यतः ग्रंथियों के समुच्चय (G) को एक $n \times n$ आव्यूह द्वारा वर्णित किया जाता है, जिसे निकटवर्ती आव्यूह $A = \{a_{ik}\}$ कहा जाता है, जहां पर $a_{ik} = 1$, यदि n_i एवं n_k सीधे तौर पर संबद्ध हैं एवं $a_{ik} = 0$ अन्यथा ($i, k = 1, 2, \dots, n$)। एक सरल अनिर्देशित लेखाचित्र हेतु ग्रंथियों के स्व निकटवासी (छल्ले नहीं) नहीं हैं अतः $a_{ii} = 0$ एवं $a_{ik} = a_{ki}$ । एक सरल अनिर्देशित लेखाचित्र



चित्र 4.1 6 ग्रंथियों वाला एक सरल लेखाचित्र निकटवर्ती आव्यूह

$$A = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix}$$

निकटवर्ती आव्यूह

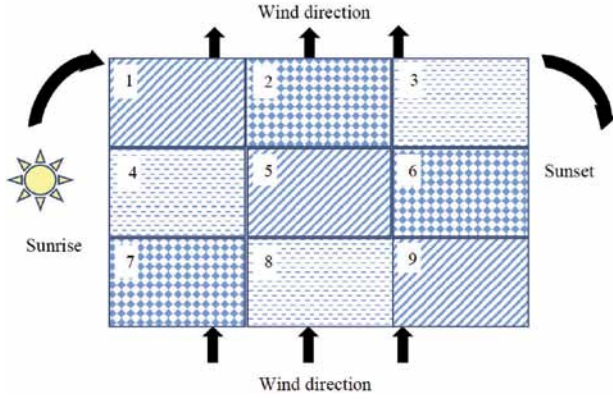
जिसमें 6 (n_1 से n_6) ग्रंथियां हैं, परिकल्पना को दर्शाने के लिए चित्र 4.1 में दिया गया है। लेखाचित्र का निकटवर्ती 6×6 आव्यूह A भी नीचे दिया जा रहा है:

पार्कर इत्यादि (2017) के अनुसार $G = (N, E)$ तंत्र एक अनिर्देशित लेखाचित्र, N ग्रंथियों एवं $E \subseteq (N \times N)$ छोर का एक संग्रह है जहां ग्रंथियां, विषयों अथवा परीक्षात्मक इकाईयां जिन पर कुछ उपचार अनुप्रयुक्त किया है, को निरूपित करती हैं। तब एक $|N| = n$ सामाजिक परीक्षण में विषय, जिन पर v उपचार अनुप्रयुक्त किए गए हों, निहित हैं और विषयों के मध्य संबंध, $n \times n$ निकटवर्ती आव्यूह A द्वारा दर्शाया गया है

5. कृषि-वानिकी परीक्षणों में तंत्र संतुलन

कृषि-वानिकी परीक्षणों के संदर्भ में तंत्र परिकल्पना परिवर्तित है, हालांकि अवकलन/गणना जटिल है। कृषि-वानिकी के लिए परीक्षण अभिकल्पना करना एकल फसल प्रणाली की अपेक्षा अधिक जटिल है। इसका एक कारण यह है कि एकल फसल प्रणाली में निकटवर्ती प्लॉटों से प्रभावों की दिशा आसानी से पहचानी जा सकती है जबकि वृक्षों के गुणों के कारण कृषि-वानिकी प्रणाली में निकटवासी प्रभावों की दिशा को पहचानना अत्यंत कठिन है। मान लीजिये एक कृषि-वानिकी क्षेत्र G, n वर्गाकार प्लॉटों में विभक्त है जिनके मध्य कोई खाली स्थान नहीं है और इन प्लॉटों में बहुत सी मिश्रित वृक्ष-फसलें लगाई गई हैं। इस अवस्था के लिए, यदि परीक्षण प्लॉट इस प्रकार बनाए गए हैं कि हवा एवं सूर्योदय-सूर्यास्त की दिशाएँ चित्र 5.1 में दिये अनुसार लम्बवत्त हैं तब दिये गए प्लॉट पर निकटवर्ती प्रभाव (प्लॉट 5, कहिए) निकटवर्ती प्लॉटों 4 एवं 6 (छाया के कारण) एवं 8 (वायु के कारण) के निकटवर्ती प्रभावों का योग होगा।

यहाँ पर, प्लॉट लचीले नहीं हैं क्योंकि क्षेत्र (G) में भूमि पर उन्हें नियत स्थिति में ही बनाया जा सकता है, अतः प्लॉट में संबंध केवल निकटवर्ती प्लॉटों, जो कि एक-दूसरे से लगे हुए हैं तक प्रतिबंधित है। यह सुनिश्चित करने के लिए सभी प्लॉटों के निकटवर्ती प्लॉटों की संख्या समान रहे, G को m प्लॉटों में इस प्रकार विभक्त किया जाता है कि n प्लॉट अंदर रहें



चित्र 5.1: 9 प्लॉटों में विभक्त किया गया कृषि-वानिकी क्षेत्र

चित्र 5.1 से कृषि-वानिकी क्षेत्र G का निकटवर्ती आव्यूह नीचे दिया गया है

$$A = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

और $m-n$ छोर प्लॉट हों (मेंड की तरह कार्य करने के लिए)। अतः अंदर वाले प्लॉट के 4 निकटवर्ती प्लॉटों (ऊपर, नीचे, दायें एवं बायें) सीधे इससे मिले हुए होंगे।

6. तंत्र प्रभाव मॉडल

अब कृषि-वानिकी परीक्षण का एक ऐसा विन्यास लीजिए जहां n प्लॉटों में एक ही फसल में v विभिन्न वृक्ष प्रजातियाँ लगाई गयी हैं, प्रत्येक प्लॉट में एक वृक्ष की एक ही प्रजाति है। यह अपेक्षा की जाती है कि दिये गये प्लॉट पर लगे वृक्ष का प्रभाव सभी निकटवर्ती प्लॉटों पर अनुभव किया जा सके। मुख्य प्लॉट के वृक्षों के प्रभाव को 'वृक्ष सीधा प्रभाव' और निकटवर्ती प्लॉटों पर वृक्षों के प्रभाव को 'वृक्ष तंत्र प्रभाव' से उद्धृत करते

हैं। प्लॉटों (वृक्षों के कारण) के मध्य संयोजन विन्यास आसानी से निर्दिष्ट किया जा सकता है यदि सरल लेखाचित्र में प्लॉटों को ग्रंथियों की तरह और मध्य संयोजनों को छोर की तरह लिया जाए। यह मानते हुए कि Y_{ij} अनुक्रिया (फसल से मापी गयी है), फसल उपज का परिणाम है, (जो कि सभी प्लॉटों के लिए समान है) 'वृक्ष सीधा प्रभाव' (τ_j) प्लॉट i से जिसमें j वृक्ष प्रजातियाँ है व 'वृक्ष तंत्र प्रभाव' (δ_{lk}) यदि l वृक्ष प्रजातियाँ, निकटवर्ती संयोजित प्लॉट k पर लगाए गए हैं, जहां k वाँ प्लॉट, i वे प्लॉट के बायें, दायें, ऊपर और नीचे के 4 निकटवर्ती प्लॉटों में से ही एक है। मान लीजिये $A_{n \times n} = \{a_{ik}\}$, इस परीक्षण के लिए एक समरूप निकटवर्ती आव्यूह है, तब तंत्र प्रभाव मॉडल इस प्रकार दिया गया है

$$Y_{ij} = \mu + \tau_j + \sum_{l=1}^v \sum_{k=1}^n a_{lk} \delta_{lk} + \varepsilon_{ij} \quad (5.1)$$

$$i = 1, 2, \dots, n; k = 1, 2, \dots, n; i \neq k; j = 1, 2, \dots,$$

$$v; l = 1, 2, \dots, v$$

जहां पर μ सामान्य माध्य है और ε_i को (माध्य 0 और σ^2 नियत प्रसरण के साथ) समरूप, स्वतंत्र एवं समान्यतः वितरित ($iidN$) माना गया है।

मान लीजिये u_j संकेतक सदिश है जिसमें j वां अवयव 1 है जब j वृक्ष प्रजातियाँ i प्लॉट में लगाई गयी हैं अन्यथा 0 हैं, ताकि हम आव्यूह $U = [u_1 \ u_2 \ \dots \ u_j \ \dots \ u_v]$ लिख सकते हैं, तब (5.1) में तंत्र प्रभाव मॉडल को आव्यूह चिन्हों में इस प्रकार अभिव्यक्त करते हैं:

$$Y = \mu \mathbf{1} + U\tau + AU\delta + \varepsilon \quad (5.2a)$$

$$Y = [\mathbf{1} \ U \ N] \begin{bmatrix} \mu \\ \tau \\ \delta \end{bmatrix} + \varepsilon = X\theta + \varepsilon \quad (5.2b)$$

जहाँ पर, Y - n प्लॉटों से मापे गए प्रेक्षणों का $n \times 1$ सदिश है, μ - कुल माध्य है, $\mathbf{1}$ -इकाईयों का $n \times 1$ सदिश है, U - प्रेक्षणों बनाम प्रत्यक्ष वृक्ष प्रभाव का $n \times v$ अभिकल्पना आव्यूह है, τ - प्रत्यक्ष वृक्ष प्रभावों का $v \times 1$ सदिश है, A - प्लॉटों के निकटवर्ती का $n \times n$ आव्यूह है, N - प्रेक्षणों बनाम सभी (अनिर्देशित) निकटवर्ती वृक्षों का $n \times v$ अभिकल्पना आव्यूह है, δ - सभी तंत्र (निकटवर्ती)

प्रभावों का सदिश है, $X - n \times (2v+1)$ अभिकल्पना आव्यूह है, θ —प्राचलों का $(2v+1) \times 1$ सदिश है और ε —अपेक्षा $E(\varepsilon) = \mathbf{0}$ एवं प्रकीर्णन $D(\varepsilon) = \sigma^2 I_n$ सहित $n \times 1$ त्रुटियों का सदिश है।

7. कुछ परिभाषाएँ

मान लीजिये कि m प्लॉट g कतार—स्तम्भ सारणी में इस प्रकार हैं कि प्रत्येक सारणी में n प्लॉट अंदर के हैं और $m - n$ प्लॉट सीमा पर हैं। कतार में अंदर के प्लॉट की संख्या कतार की लंबाई है, स्तम्भ में अंदर के प्लॉट की संख्या स्तम्भ की लंबाई है और मान लीजिये कतार लंबाई $= \alpha$ व स्तम्भ लंबाई $= \beta$ है। $n (= \alpha\beta)$ प्लॉटों में विभिन्न वृक्ष प्रजातियाँ v लगी हुई हैं। कतार (b) और स्तम्भ (k) में i वे प्लॉट के लिए लगाने वाली वृक्ष प्रजातियों को हम $i(b,k)$ द्वारा निर्दिष्ट करते हैं, जहाँ पर $b = 0, 1, 2, \dots, \beta, \beta + 1$; $k = 0, 1, 2, \dots, \alpha, \alpha + 1$ हैं। अवयव 0 एवं $\alpha + 1$ क्रमशः बायें एवं दायें छोर तथा 0 एवं $\beta + 1$ अवयव क्रमशः ऊपर एवं नीचे के छोर वाले प्लॉटों को निर्दिष्ट करते हैं। इस सबको स्पष्ट समझने के लिए निम्नलिखित परिभाषाएँ दी जा रही हैं।

निकटवर्ती आव्यूह: निकटवर्ती आव्यूह A को $n \times n$ क्रम के समरूप द्विपदीय आव्यूह के रूप में परिभाषित किया जाता है जिसमें i वे एवं k वे प्लॉटों के मध्य संबंध $a_{ij} = 1$ द्वारा दिया गया है यदि i एवं k प्लॉट वृक्ष तंत्र के पदों में संयोजित हैं अन्यथा $(i, k = 1, 2, \dots, n)$ । A बहुत ही महत्वपूर्ण है क्योंकि यह परीक्षण एकाईयों के मध्य संयोजन दर्शाता है जो कि रैखिक तंत्र प्रभाव मॉडल में मूल पूर्व धारणा है। लगभग सभी तंत्र A विश्लेषण पर आधारित हैं क्योंकि यह मापे जा रहे सम्बन्धों द्वारा मानचित्रित परीक्षणात्मक इकाईयों के “सामाजिक स्थान” में किसके आगे कौन है, अथवा किसके निकटवर्ती कौन है, प्रदर्शित करता है।

तंत्र: तंत्र परीक्षणात्मक इकाईयों (प्लॉटों) के समूहों को उद्धृत करता है जो प्रकृति में हैं और एक दूसरे के निकटवर्ती होने से संयोजित है।

वृक्षतंत्र प्रभाव: यह प्लॉटों के निकटवर्ती होने के परिणामस्वरूप आस-पास के प्लॉटों पर प्रजातियों $i(b,k)$ द्वारा पड़ने वाले प्रभाव को उद्धृत करता है।

निष्कर्ष: कृषि—वानिकी बहुत पुरानी कृषि प्रणाली है और लघु एवं सीमांत कृषकों के कार्यों का अभिन्न अंग है। कृषि—वानिकी प्रणाली में उन्नत उपादेयता सुनिश्चित करने के लिए परीक्षणात्मक परीक्षणों का नियोजन सावधानीपूर्वक करना आवश्यक है। कृषि परीक्षणों के विपरीत, कृषि—वानिकी परीक्षण, वृक्षों एवं फसलों जैसे बहुघटक निहित होने के कारण अधिक जटिल हैं। विशेष रूप से वृक्षों की प्रवृत्ति न केवल अपने प्लॉट में उगने वाली फसलों बल्कि निकटवर्ती प्लॉटों की फसलों को प्रभावित करने वाली होती हैं। इस अध्ययन में निकटवर्ती प्लॉटों के वृक्षों को अ—दिशित माना गया है और उन्हें वृक्ष तंत्र प्रभाव कहा गया है। तंत्र अभिकल्पना, जो की लेखाचित्र सिद्धान्त से ली गयी है, उस पर आधारित एक रैखिक तंत्र प्रभाव मॉडल से परिचित करवाया गया है।

संदर्भ

- अमनोर, के. एस. (1996). *द रोल ऑफ ट्रीज इन फार्मिंग: द पर्सपेक्टिव ऑफ फार्मर्स, समरी रिपोर्ट*। इंस्टीट्यूट ऑफ अफ्रिकन स्टडीज, यूनिवर्सिटी ऑफ घाना।
- चार्टरेंड, जी. एवं झांग, पी. (2012). *ए फर्स्ट कोर्स इन ग्राफ थ्योरी*। डोवर पब्लिकेशन इंक. माइन्सोला, न्यूयार्क।
- दे बेट्स, एन. गैरियेपी, एस. एवं वैजिना, ए. (2007). *पोर्टरेट ऑफ एग्रोफोरेस्ट्री इन क्यूबेक*। एग्रीक्यूल्चर समरी। कनाडा। पीपी. 16।
- ध्यानी, एस. के. एवं हांडा, ए. के. (2014). *एग्रोफोरेस्ट्री इन इण्डिया एंड इट्स पोटेन्शियल फॉर इकोसिस्टम सर्विसेज*। इन एग्रोफोरेस्ट्री सिस्टम इन इण्डिया: लाईवलीहुड सिक्योरिटी एंड इकोसिस्टम सर्विसेज।

डागर, जे. सी., सिंह, ए. के. एवं अरुणाचलम, ए. (ईडीएस) एडवांसेज इन एग्रोफोरेस्ट्री 10, 345–365 / सिंजरा।

जग्गी, एस. शर्मा, वी. के., गिल, ए. एस. एवं वर्गीस, सी. (2007). सम इन्वेस्टिगेशन ऑन डिज़ाइन एंड एनालिसिस ऑफ एग्रोफोरेस्ट्री एक्सपेरीमेंट्स। प्रोजेक्ट रिपोर्ट, भा.कृ.अनु.प.–भा.कृ.सां.अनु.सं.।

कहन, डी. (2008). फार्म मैनेजमेंट एक्सटेंशन गाइड सिरीज़: मैनेजिंग रिस्क इन फार्मिंग। एफएओ, रोम। पीपी. 107।

लंगरेन, बी. ओ. एवं रेनट्री, जे. बी. (1982). सस्टेंड एग्रोफोरेस्ट्री। इन: नेस्टेल, बी. (इंडी.)।

एग्रीकल्चरल रिसर्च फॉर डेवलेपमेंट: पोटेन्शीयल एंड चेलेंजेज़ इन एशिया, पीपी. 37–49। आइएसएनएआर, द हेग, द नीदरलैंड्स।

नायर, पी. के. आर. (1993). एन इंट्रोडक्शन टु एग्रोफोरेस्ट्री। लूवर अकेडेमिक पब्लिशर्स इन कोऑपरेशन विद आइसीआरएएफ, बॉस्टन/लंदन।

पार्कर, बी. एम., गिलमोर, एस. जी. एवं श्चोरमन्स, जे. (2017). ओप्टिमल डिज़ाइन ऑफ एक्सपेरीमेंट्स ऑन कनेक्टेड यूनिट्स विद एप्लिकेशन टु सोशल नेटवर्क्स। अप्लाइड स्टेटिस्टिक्स, सिरीज़ सी, 66(3) 455–480। डीओआई: 10.1111/ आरएसएससी. 12170

“

माना की अंधेरा घना है,
लेकिन दिया जलाना कहाँ मना है।

”

-नरेन्द्र मोदी

विशेषता निष्कर्षण तकनीकों और समानता आव्यूह की समीक्षा: सामग्री-आधारित छवि पुनर्प्राप्ति

मधु, सपना निगम, संचिता नाहा एवं चंदन कुमार देब
भा.कृ.अनु.प.—भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली—110012

सार

कृषि हजारों वर्षों से मनुष्य की आजीविका का सबसे प्राथमिक स्रोत रही है। आज भी यह दुनिया की लगभग 50% आबादी को निर्वाह प्रदान करती है। दुनिया भर में हर साल, फसल उत्पादन के बड़े नुकसान का गंभीर कारण पौधों की बीमारियाँ हैं। पौधों की बीमारियों से होने वाले वित्तीय नुकसान से निपटने के लिए पौधों को उनकी वृद्धि/विकास के विभिन्न चरणों में स्वस्थ रखना आवश्यक है। संक्रमण के लक्षण मुख्य रूप से पौधों की पत्तियों पर दिखाई देते हैं। इस प्रकार पत्तियों का उपयोग आमतौर पर रोगों का पता लगाने और उनकी पहचान करने के लिए किया जाता है। दृश्य अवलोकन के माध्यम से रोग का पता लगाना अपने आप में एक चुनौतीपूर्ण कार्य है और इसके लिए बहुत अधिक मानवीय विशेषज्ञता की आवश्यकता होती है। यह लेख छवि पुनर्प्राप्ति के लिए विभिन्न चरणों का वर्णन करता है। फिर निम्न स्तर की विशेषताओं जैसे रंग, बनावट और आकार की विशेषताओं को निकालने के लिए उपयोग की जाने वाली विभिन्न पद्धतियों का वर्णन किया गया है। साथ ही यह लेख छवियों के बीच समानता का पता लगाने के लिए विभिन्न दूरी के उपायों की विस्तृत समीक्षा प्रदान करता है।

1. प्रस्तावना

छवियां दृश्य सामग्री (जैसे रंग, बनावट और आकार) में बहुत समृद्ध हैं, जिनका उपयोग पाठ आधारित छवि

पुनर्प्राप्ति प्रणाली की कमियों को दूर करने के लिए किया जा सकता है, यदि कंप्यूटर द्वारा सुलभ गणितीय वैक्टर (फीचर वैक्टर) के रूप में प्रतिनिधित्व किया जा सके। एक सामग्री आधारित छवि पुनर्प्राप्ति (CBIR) असामान्य राज्य ढांचे (मानव मन) और निम्न स्तर के ढांचे (एक पीसी (Personal Computer)) के बीच एक इंटरफेस है। मानव मस्तिष्क जटिल दृश्य विवेक करने के लिए उपयुक्त है, हालांकि गति में विवश है जबकि एक पीसी प्रतिबंधित दृश्य क्षमताओं के लिए काफी अधिक दरों पर उपयुक्त है। एक CBIR में, दृश्य चित्र सामग्री को चित्र हाइलाइट्स के रूप में बोला जाता है, जो स्वाभाविक रूप से निकाले जाते हैं और कोई मैनुअल हिमायत नहीं होती है, इसलिए घटक निष्कर्षण व्यवस्था में लोगों पर निर्भरता को हटा दिया जाता है। ये कम्प्यूटरीकृत निष्कर्षण दृष्टिकोण कम्प्यूटेशनल रूप से महंगे, परेशानी वाले हैं और क्षेत्र विशेष होने की प्रवृत्ति रखते हैं। इसलिए कम्प्यूटेशनल पेचीदगियों को सीमित करने के लिए विस्तार का पार्सल, पुनर्व्यवस्था और गैर-अनन्य प्रयास CBIR में अन्वेषण के लिए उपलब्ध हैं। CBIR या सामग्री आधारित छवि पुनर्प्राप्ति रंग, बनावट और आकार जैसी दृश्य विशेषताओं के आधार पर छवियों की पुनर्प्राप्ति है। CBIR में, डेटाबेस में संग्रहीत प्रत्येक छवि में इसकी विशेषताओं को निकाला जाता है और क्वेरी छवि (Query Image) की विशेषताओं की तुलना में किया जाता है। इसमें निम्नलिखित चरण शामिल हैं:

(i) छवि अधिग्रहण

यह छवि से एक डिजिटल छवि निकालने की प्रक्रिया का पहला चरण है, जो डेटाबेस में संग्रहीत है। डेटाबेस में बड़ी संख्या में छवियां होती हैं जो उपयोगकर्ता की पसंद और विकल्पों पर निर्भर करती हैं।

(ii) प्री-प्रोसेसिंग

सुविधाओं को निकालने के लिए छवि को पहले संसाधित किया जाता है, जो इसकी सामग्री का वर्णन करता है। प्रसंस्करण में फिल्टरिंग, सामान्यीकरण, विभाजन और ऑब्जेक्ट (Object) पहचान शामिल है। इस चरण का उत्पादन महत्वपूर्ण क्षेत्रों और वस्तुओं का एक समूह है।

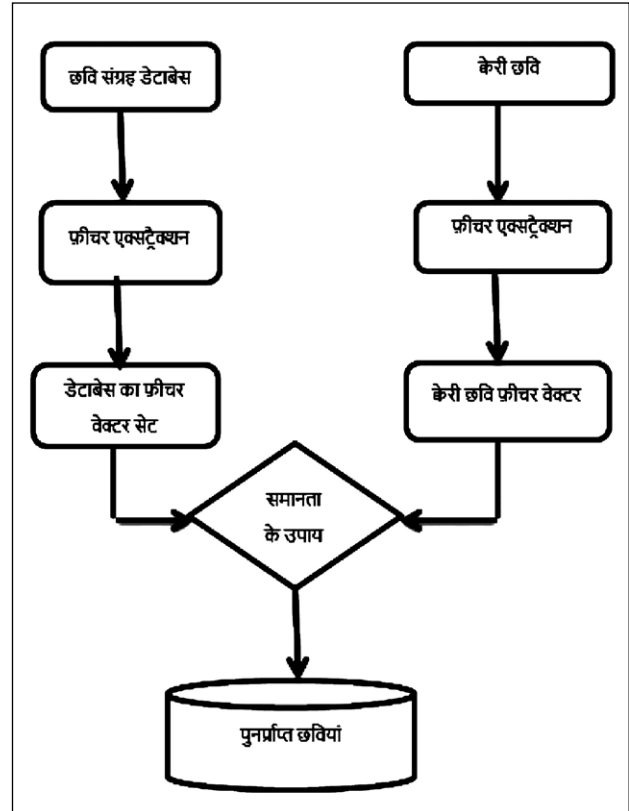
(iii) फीचर एक्सट्रैक्शन

प्रक्रिया में पहला कदम छवि सुविधाओं को एक अलग सीमा तक निकालना है। छवि की सामग्री का वर्णन करने के लिए आकृति, बनावट, रंग आदि जैसी विशेषताओं का उपयोग किया जाता है। हम विभिन्न स्तरों पर सुविधाओं को निकाल सकते हैं।

(iv) समानता के उपाय

यह एक ऐसी प्रक्रिया है जिसमें गणना प्रक्रिया के लिए प्रत्येक छवि के बारे में जानकारी को उसके फीचर वैक्टर में संग्रहीत किया जाता है और इन फीचर वैक्टर का मिलान क्वेरी इमेज के फीचर वैक्टर से किया जाता है जो समानता को मापने में मदद करता है। इस चरण में उपरोक्त वर्णित विशेषताओं का मिलान सम्मिलित है ताकि ऐसे परिणाम प्राप्त किये जा सके जो समानता माप विधि के उपयोग के साथ दृष्टि से समान हैं, इसे दूरी विधि कहा जाता है। यूक्लिडियन दूरी, सिटी ब्लॉक दूरी, और केनबरा दूरी, जैसी विभिन्न दूरी विधियां उपलब्ध हैं। चित्र 1 में छवियों की पुनर्प्राप्ति के सभी चरण दिखाए गए हैं।

खंड I में CBIR और छवि पुनर्प्राप्ति चरणों, खंड II में, विभिन्न रंग फीचर निष्कर्षण विधियों, खंड III में, विभिन्न बनावट सुविधा निष्कर्षण विधियों, खंड IV में,



चित्र 1 सामग्री आधारित छवि पुनर्प्राप्ति में फीचर निष्कर्षण

विभिन्न आकृति सुविधा निष्कर्षण पर और तदुपरान्त खंड V में, विभिन्न समानता मेट्रिक्स पर चर्चा की गई है।

2. रंग सुविधा निष्कर्षण

(i) रंग हिस्टोग्राम

CBIR की पद्धति में, चित्रों के रंगों से संबंधित विवरणों के प्रसारण को इस आरेख द्वारा दर्शाया जाता है। इसमें सभी बार (bar) में इस्तेमाल होने वाले विभिन्न प्रकार के रंगों में से एक विशेष रंग होता है। आंतरिक रूप से आंकड़ों की श्रेणी को समान आकार के डिब्बों में विभाजित किया जाता है। उसके बाद प्रत्येक डिब्बे से रंगीन पिक्सेल की गणना की जाती है और कुल बिंदुओं की ओर से व्यवस्थित किया जाता है। यह उस बिन (bin) में पिक्सल के गिरने की संभावना प्रदान करता है। यह उन स्थितियों में बहुत उपयोगी है जहां चित्र रंग का वितरण विश्व स्तर पर किया जाता है। संचयी

रंग हिस्टोग्राम, रंग हिस्टोग्राम का एक उन्नत रूप है।

(ii) कलर कोरेलोग्राम

कलर कोरेलोग्राम तस्वीर के रंग से संबंधित जानकारी को मापने के लिए उपयोग किया जाता है। यह एक 3-डी चार्ट है जो पिक्सेल के बीच छाया और स्थान द्वारा सूचीबद्ध है। यह उस तरीके का प्रतिनिधित्व करता है जिसमें रंग का त्रि-आयामी लिंक सहेजे गए चित्र के अंदर स्थान के साथ बदलता रहता है।

(iii) प्रमुख रंग वर्णनकर्ता

DCD किसी छवि या छवि क्षेत्र में विशिष्ट रंगों का वर्णन करता है। एकल या विभिन्न रंग मानों के आधार पर इसका उपयोग डेटाबेस से समान छवियों को पुनर्प्राप्त करने और छवि डेटाबेस की ब्राउज़िंग के लिए किया जाता है। DCD को $F = \{ \{c_i, p_i, v_i\}, s \}$ के रूप में परिभाषित किया गया है

$$(i=1, 23, N)$$

जहाँ N – प्रमुख रंगों की संख्या ,

c_i – संगत रंग स्थान घटक मानों का सदिश।

p_i – रंग सीआई के अनुरूप छवि या छवि क्षेत्र में पिक्सेल का अंश, और

v_i – विचरण जो रंग मूल्यों की भिन्नता को दर्शाता है।

(iv) रंग क्षण

रंग हिस्टोग्राम में परिमाणीकरण प्रभाव की कमी थी। उस पर काबू पाने के लिए, चित्र निष्कर्षण के समय, रंग के मूल्य का उपयोग रूप सुविधा वाहक में किया जाता है। किसी भी रंग वितरण को चिह्नित करने के लिए उपयोग किए जाने वाले क्षण मुख्य रूप से फीचर वाहक के रूप में औसत, अंतर और तिरछापन पर केंद्रित होते हैं। इसका मजबूत बिंदु यह है कि, वितरण में अनियमितता के स्तर को निर्धारित करने वाली क्रियाओं के रूप में इस के तिरछेपन को लागू करना संभव है।

(v) रंग जुटना वेक्टर (Colour Assembly Vector)

रंग जुटना वेक्टर एक विभाजित हिस्टोग्राम है जहां पिक्सल को उनकी त्रि-आयामी स्थिरता के आधार पर विभाजित किया जा सकता है। सभी चित्र तत्वों को दो खंडों के बीच में विभाजित किया गया है, अर्थात्, एक ही रंग के बड़े क्षेत्र का एक हिस्सा सुसंगत या असंगत है या नहीं। उसके बाद, तार्किक और अतार्किक चित्र तत्वों के समर्थन में अकेला (individual) बार चार्ट बनाना संभव है, जिसमें फीचर सेट में स्थान की जानकारी शामिल है।

(vi) औसत RGB

ए की मदद से, लाल, नीले और हरे रंग के रंग स्थान में रंग का औसत परिभाषित करना संभव हो जाता है:

$E = (\text{आर (औसत)}, \text{जी (औसत)}, \text{बी (औसत)})$ टी जहां आर (औसत) = लाल छवि औसत मूल्य,

जी (औसत) = हरे रंग की छवि का औसत मूल्य और

बी (औसत) = नीली तस्वीर के साथ औसत मूल्य कनेक्शन।

इस सुविधा के उपयोग के पीछे मूल उद्देश्य पहले चरण में बड़ी दूरी की कंपनी में तस्वीर को परिष्कृत करना है जब कई फीचर प्रश्न सम्मिलित होते हैं।

3. बनावट सुविधा निष्कर्षण

(i) तमुरा बनावट सुविधा

यह एक विशेष प्रकार की विशेषता है जो मनुष्य के देखने के दृष्टिकोण पर निर्भर करती है। यह विभिन्न कोणों से संरचनात्मक विशेषताओं का निरूपण करता है। छह दृश्य बनावट गुण इस प्रकार हैं:

- | | |
|---------------------|-----------------|
| 1. खुरदरापन | 4) विचरण |
| 2. दिशात्मक विशेषता | 5) रेखा सादृश्य |
| 3. संगति | 6) असमानता |

(ii) स्टीयरेबल पिरामिड

यह विधि के रूप में मौजूद है जो संरचनात्मक विशेषताओं की पहचान करता है। इस पद्धति में, एक चित्र दो भागों के बीच अलग हो जाता है। उसके बाद, निर्देशित फिल्टर के परिणामों से संरचनात्मक विशेषताएं निर्धारित की जाती हैं।

(iii) वेवलेट ट्रांसफॉर्म

यह फोटो एडिटिंग सॉफ्टवेयर के विभिन्न हिस्सों में उपयोगी हो जाता है। यह विधि संरचनात्मक विशेषताओं के परीक्षण और पहचान में सहायक होती है। इस विधि का उपयोग चित्रों के निष्कर्षण के लिए किया जाता है। उच्च स्तर की निष्कर्षण प्रामाणिकता तरंगिका गुणों द्वारा प्राप्त नहीं की जाती है। ये गुण जो असतत तरंगिका के कारकों के माध्यम से निर्धारित होते हैं, CBIR में उपयोगिता में सुधार के लिए विभिन्न मापदंडों को निर्दिष्ट करते हैं।

(iv) गैबर वेवलेट ट्रांसफॉर्म

गैबर फिल्टर में तरंगों का एक सेट मौजूद होता है। सभी सेट एक विशेष समाधान और दिशा में तीव्रता को पकड़ लेते हैं। इसका अर्थ है, पूर्ण संकेत या चित्र के आसपास मौजूद ऊर्जा को गैबर फिल्टर द्वारा जब्त कर लिया जाता है। इस प्रकार के परिवर्तन का उपयोग मुख्य रूप से उन विशेषताओं को हटाने के लिए किया जाता है जो किसी चित्र की सतह से संबंधित होती हैं।

(v) फूरियर ट्रांसफॉर्म

इस पद्धति में, एक संकेत को विभिन्न ज्या वक्रों में विभाजित किया जाता है जिनकी आवर्तता समान नहीं होती है। सिग्नल के फ्रीक्वेंसी घटकों को मापने के लिए, सिग्नल को टाइम डोमेन से फ्रीक्वेंसी डोमेन में परिवर्तित किया जाता है। यह एन्क्लेव और ऑर्थोगोनल भाग के बीच की एक तस्वीर को खराब कर देता है। यह विधियों के एक परिवार से संबंधित है जो असतत फूरियर ट्रांसफॉर्म (डीएफटी) की गणना बहुत प्रभावी तरीके से करता है। इसका अर्थ है कि फास्ट फूरियर रूपांतरण एफटी का दृष्टिकोण नहीं है। यह डिजिटल

सिग्नल प्रोसेसिंग में उपयोगी है।

(vi) ग्रे लेवल सह-घटना मैट्रिक्स (GLCM)

GLCM में पिक्सेल के बीच में रिक्त स्थान और दिशाओं ने एक आव्यूह का गठन किया। तदुपरांत, सतह के गुणों के रूप में आव्यूह से महत्वपूर्ण आंकड़े निकालता है। आमतौर पर शोधकर्ता चार मुख्य विशेषताओं पर काम करते हैं जिनमें शक्ति, असमानता, संबंध, निरंतरता ऐसे गुण हैं जिन पर शोधकर्ता काफी शोध कार्य करते हैं।

(vii) कंटूरलेट ट्रांसफॉर्म

मूल रूप से यह डायरेक्शनल फिल्टर बैंक (DFB) और एक लैप्लासियन पिरामिड (LP) का मिश्रण है। इन दोनों में से प्राथमिक विभिन्न दिशाओं में ब्रेकडाउन देता है और बाद वाला अलग-अलग पैमाने पर ब्रेकडाउन देता है। यह डबल फिल्टर बैंक के रूप में मौजूद है। यह द्वि-आयामी क्रमिक सरल संकेतों के समर्थन में सीमित डिज़ाइन की दिशा में योगदान प्रदान करता है, जो चित्रों के समान है।

4. आकार सुविधा निष्कर्षण

(i) फूरियर डिस्क्रीप्टर

फूरियर रूपांतरण को आकार सीमा पर लागू किया जाता है और फूरियर रूपांतरित गुणांक प्राप्त किए जाते हैं। ये रूपांतरित कारक, आकार के फूरियर डिस्क्रीप्टर (FD) के रूप में प्रसिद्ध हो जाते हैं। फास्ट फूरियर ट्रांसफॉर्म से कम्प्यूटरीकृत गणना उपयोगी हो जाती है।

(ii) ज्यामितीय क्षण (GM)

जीएम को कुछ विशेष भारित औसत माना जाता है, छवि पिक्सल की तीव्रता के एक फलन के रूप में, या इन क्षणों में, चित्र संपादन जैसे विभिन्न क्षेत्रों में। मूल रूप से, विभाजन के बाद इसका उपयोग वस्तुओं के विवरण के लिए किया जा सकता है। समग्र शक्ति, इसकी दिशा से संबंधित विवरण और गुरुत्वाकर्षण का केंद्र चित्र गुणों के रूप में प्रदर्शित होता है जो इस क्षण के माध्यम से निर्धारित होते हैं।

(iii) बीजगणितीय क्षण अपरिवर्तनीय (AMI)

एएमआई प्रारंभिक एम प्राथमिक तत्काल के आधार पर निर्धारित किया जाता है। यह उन आव्यूहों के ईजिन मूल्यों के रूप में प्रसिद्ध हो जाता है जिन्हें पहले से $M [j, k]$ के रूप में परिभाषित किया जाता है। इसके घटक बड़े पैमाने पर एम प्राथमिक तत्काल के प्राचल बन जाते हैं।

(iv) वेवलेट ट्रांसफॉर्म

वेवलेट ट्रांसफॉर्म का उपयोग एक पदानुक्रमित प्लानर कर्व डिस्क्रिप्टर विकसित करने के लिए किया जाता है। इस डिस्क्रिप्टर द्वारा विभिन्न तत्वों के बीच में एक वक्र टूट जाता है, जिसका स्तर समान नहीं होता है। वेवलेट डिस्क्रिप्टर की तुलना में इसकी मिलान प्रक्रिया बहुत ही किफायती है।

5. समानता दूरी मेट्रिक्स

विभिन्न छवियों के समानता मिलान के लिए विभिन्न दूरी आव्यूह का उपयोग किया जाता है। इन दूरी के उपायों या आव्यूह को इस प्रकार दर्शाया गया है:

(i) L1 दूरी

इसे मैनहट्टन दूरी के नाम से भी जाना जाता है। यदि $x=(a1, b1)$ और $y=(a2, b2)$ दो बिंदु हैं, तो x और y के बीच L1 द्वारा, इस प्रकार दिया जाता है:-

$$L1(x, y) = |a1-a2|+|b1-b2|$$

(ii) L2 दूरी

इसे यूक्लिडियन दूरी के रूप में भी जाना जाता है। यदि $x=(a1, b1)$ और $y=(a2, b2)$ दो बिंदु हैं, तो x और y के बीच L2 द्वारा, इस प्रकार दिया जाता है:-

$$L2(x, y) = \sqrt{(a1-a2)^2 + (b1-b2)^2}$$

(iii) मानकीकृत L2

मानकीकृत L2 दूरी का अर्थ है L2 दूरी की गणना मानकीकृत डेटा पर की जाती है। मानकीकृत मान = (मूल मान - माध्य)/मानक विचलन:-

$$d = \sqrt{\sum (1/s_i^2) (x_i - y_i)^2}$$

(iv) सिटी ब्लॉक दूरी

यूक्लिडियन दूरी की तुलना में सिटी ब्लॉक दूरी छवियों को तेज दर से पुनर्प्राप्त करती है। यह बनावट विशेषता समानता पर एचएसवी हिस्टोग्राम तुलना के साथ अच्छी तरह से काम करता है। यह दूरी नीचे समीकरण में दर्शाई गई है:-

$$D(x,y) = \sum_{i=1}^n |x_i - y_i|$$

(v) कोसाइन समानता माप (Cos)

कोसाइन दूरी मीट्रिक इकाई लंबाई के लिए सभी फीचर वेक्टर को सामान्य करता है और छवि सामग्री के सापेक्ष इन-प्लेन स्केलिंग परिवर्तन के सम्मुख इसे अपरिवर्तनीय बनाता है। यह समीकरण में दिखाया गया है:-

$$d(x,y) = \sum_i x_i y_i / \sqrt{\sum_i x_i^2 \sum_i y_i^2}$$

(vi) मिकोव्स्की दूरी

दो बिंदुओं X और Y के बीच की मिकोव्स्की दूरी या क्रम p फीचर स्पेस R , $X=(x1, x2, x3, x4, \dots, x_i, \dots, x_n)$ और $Y=(y1, y2, y3, y4) \dots, y_i, \dots, y_n) \in R$ नीचे समीकरण में दिया गया है:

$$d(x,y) = \sum_{i=1}^n ((x_i - y_i)^p)^{1/p}$$

p = दूरी मीट्रिक का क्रम

(vii) सामान्यीकृत L2

महालनोबिस दूरी में, जब सहप्रसरण आव्यूह समरूपता आव्यूह होता है, तो यह यूक्लिडियन दूरी तक कम हो जाता है और जब सहसंयोजक आव्यूह विकर्ण होता है, तो यह सामान्यीकृत L2 तक कम हो जाता है:

$$d(x,y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2 / s_i^2}$$

निष्कर्ष

CBIR कृषि, इमेज प्रोसेसिंग, पैटर्न रिकग्निशन और कंप्यूटर विज्ञान में एक सक्रिय शोध विषय है। CBIR का उद्देश्य क्वेरी छवि और डेटाबेस छवियों के बीच समानता को मापकर, एक बड़े छवि डेटाबेस से

प्रासंगिक छवियों को पुनः प्राप्त करना है। यह मोटे तौर पर तीन बुनियादी निम्न स्तर की विशेषताओं पर आधारित है, अर्थात् रंग, बनावट और आकार, और उनमें से रंग सबसे महत्वपूर्ण है। इस लेख में रंग, बनावट, आकार विशेषता निष्कर्षण के लिए उपयोग की जाने वाली विभिन्न विधियों का वर्णन किया गया है। इन विधियों का उपयोग और चयन उस अनुप्रयोग के अनुसार किया जाता है जिसके लिए उन्हें नियोजित किया जाता है। छवि की स्थानिक जानकारी पर विचार करने वाली विधियां अधिक लोकप्रियता प्राप्त कर रही हैं, क्योंकि वे विभिन्न अन्य दृष्टिकोणों में भी लाभप्रद हैं। इसके अतिरिक्त विधि शोर, स्केलिंग, रोटेशन, पिलपिंग आदि के सम्मुख मजबूत होनी चाहिए, और भंडारण स्थान और समय जटिलता मापदंडों में कुशल होनी चाहिए। इन तकनीकों की तुलना से यह दिखाया जा सकता है कि बनावट सुविधा के कुशल भेदभाव की आवश्यकता होने पर गैबर वेवलेट ट्रांसफॉर्म का उपयोग किया जाता है। CBIR में रंग और बनावट विशेषता निष्कर्षण के लिए अन्य तकनीकों का भी अच्छी तरह से उपयोग किया जा सकता है। साथ ही यह लेख छवियों के बीच समानता खोजने के लिए विभिन्न दूरी उपायों की विस्तृत समीक्षा प्रदान करता है।

संदर्भ

- दानिश, एम., रावत, आर., एवं शर्मा, आर. (2013). "एक सर्वेक्षण: रंग, बनावट, आकार और न्यूरो फ़ज़ी के आधार पर सामग्री आधारित छवि पुनर्प्राप्ति". इंटर. जर्नल ऑफ़ इंजीनियरिंग रिसर्च एंड एप्लीकेशन, वॉल्यूम 3- नंबर 5, पीपी.839-844।
- चड्ढा, ए., मलिक, एस. एवं जौहर, आर. (2012). "कंटेंट बेस्ड इमेज रिट्रीवल के लिए फीचर-एक्सट्रैक्शन तकनीकों का तुलनात्मक अध्ययन और अनुकूलन", कंप्यूटर अनुप्रयोगों के अंतर्राष्ट्रीय जर्नल, वॉल्यूम 52- नंबर 20, पीपी. 35-42, डीओआई: 10.5120/8320-1959।
- बोर्डे, स्वप्ना एवं भोसले, उद्धव (2012). "फीचर वेक्टर बेस्ड सीबीआईआर इन स्पेसियल एंड ट्रांसफॉर्म डोमेन," इंटरनेशनल जर्नल ऑफ़ कंप्यूटर एप्लीकेशन (0975 - 8887), वॉल्यूम 60- नंबर 19, पीपी. 34-42, डीओआई: 10.5120/9810-4399।
- कचौरी, आर., जेमल, के., मारेफ, एच., सेलामी, मसमुदिल डी., एवं डरबेल, एन. (2008). "विषम छवि डेटाबेस मान्यता के लिए फीचर निष्कर्षण और प्रासंगिकता मूल्यांकन: पर" छवि प्रसंस्करण सिद्धांत, उपकरण और अनुप्रयोग, आईईईई, पीपी.1-6, डीआई: 10.1109/इप्टा. 2008.4743738।

“

धर्म के बिना विज्ञान लंगड़ा है,
विज्ञान के बिना धर्म अंधा है।

”

-अल्बर्ट आइंस्टाइन

फसल कटाई परीक्षणों की कम संख्या का उपयोग करके जिला स्तरीय प्रमुख फसल उपज अनुमान

कौस्तव आदित्य¹, पंकज दास¹, भारती¹, अंकुर विश्वास¹, श्रीला दास² एवं मंजू गौतम¹

¹भा.कृ.अनु.प.—भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली—110012

²भा.कृ.अनु.प.—भारतीय कृषि अनुसंधान संस्थान, नई दिल्ली—110012

सार

भारतीय अर्थव्यवस्था में कृषि एक महत्वपूर्ण भूमिका निभाती है और इसलिए कृषि सांख्यिकी का संग्रह और रखरखाव बहुत महत्व रखता है। पिछले कुछ कृषि वर्षों के दौरान यह देखा गया है कि सामान्य फसल अनुमान सर्वेक्षण (GES) के अंतर्गत कई प्रमुख और लघु फसलों के फसल उपज अनुमानों का पता लगाने के लिए हर साल भारत में लगभग 1300000 फसल कटाई प्रयोग (CCE) आयोजित किए जाते हैं। जनशक्ति की कमी और दिन-प्रतिदिन भारी मात्रा में कार्य के कारण आंकड़ों की गुणवत्ता संदिग्ध होती जा रही है। इस समस्या से निपटने के लिए, भा.कृ.अनु.प.—भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली द्वारा अर्थशास्त्र एवं सांख्यिकी निदेशालय (DES), कृषि एवं किसान कल्याण मंत्रालय (MOA&FW), सरकार द्वारा प्रायोजित एक मूल अध्ययन किया गया। यह भारत के राज्यों से चुने गए गांवों के कम प्रतिदर्शों के आकार से प्रमुख फसल उपज के जिला स्तर के अनुमान उत्पन्न करने के लिए किया गया। गांवों की संख्या में कमी के साथ, अध्ययन के दौरान कुछ जिलों में प्रतिदर्श आकार नहीं होने की समस्या का सामना करना पड़ा, जहां फसल उपज के सामान्य अभिकल्पना आधारित अनुमान उत्पन्न नहीं किए जा सकते थे। इस समस्या से निपटने के लिए एग्रीगेट लेवल स्मॉल एरिया एस्टीमेशन (SAE) का उपयोग किया गया। इस परीक्षणात्मक अध्ययन में उत्तर प्रदेश राज्य में दो प्रमुख फसलों धान और गेहूं के लिए दो मौसमों अर्थात् कृषि वर्ष 2015–16 की खरीफ और रबी और असम में धान के लिए कृषि

वर्ष 2015–16 के खरीफ के लिए प्राप्त परिणाम पर चर्चा की गयी। उपज अनुमानों की तुलना निर्धारण वर्ष 2015–16 के लिए GCES के अंतर्गत जारी अनुमानों से की गई थी। यह पाया गया कि CCE की संख्या के कम प्रतिदर्श आकार से प्राप्त अनुमान के सापेक्ष GCES ने स्वीकार्य स्तर की सटीकता के साथ समान अनुमान प्रस्तुत किए।

मुख्य शब्द: प्रतिदर्श आकार; कृषि सांख्यिकी; फसल कटाई परीक्षण; लघु क्षेत्र अनुमान, प्रायोगिक सर्वेक्षण।

परिचय

विकासशील देशों में नीतियाँ बनाने में सर्वेक्षण ज्यादातर महत्वपूर्ण होते हैं और अधिकांश विकासशील देशों में कृषि एक महत्वपूर्ण भूमिका निभाती है, जैसे भारत में 50 प्रतिशत से अधिक ग्रामीण परिवार अपनी आजीविका के प्रमुख साधन के रूप में कृषि पर निर्भर हैं। कृषि क्षेत्र की प्रमुख स्थिति को देखते हुए, कृषि सांख्यिकी का संग्रह और रखरखाव बहुत महत्व रखता है। यादृच्छिक प्रतिदर्श (Random Sampling) की वैज्ञानिक तकनीकों का पालन करते हुए राज्य एजेंसियों द्वारा आयोजित GCES के माध्यम से प्रमुख फसलों की उपज दरों का अनुमान लगाया जाता है। GCES योजना के तहत, एक या एक से अधिक राज्य एजेंसियों द्वारा समय पर रिपोर्टिंग योजना (TRS) से प्रतिदर्श गांवों में CCE किया जाता है (कृषि सांख्यिकी पर विशेषज्ञ समिति की रिपोर्ट, 2011)। आपसी सहमति से, सभी राज्य प्रतिदर्श गांवों और भूखंडों के चयन के

लिए समान अभिकल्पना और प्रक्रियाओं के साथ-साथ विशेषज्ञों द्वारा तय CCE के संचालन और पर्यवेक्षण के लिए क्षेत्र प्रक्रियाओं का पालन करते हैं। लक्ष्य यह था कि प्रतिदर्श त्रुटि की एक उचित, निर्दिष्ट सीमा (Margin) के भीतर विभिन्न फसलों की प्रति हेक्टेयर उपज का अनुमान प्राप्त किया जा सके। पिछले कुछ कृषि वर्षों के दौरान, भारत के विभिन्न राज्यों/केंद्र शासित प्रदेशों में हर साल 52 खाद्य और 16 गैर-खाद्य फसलों को समाविष्ट (cover) करते हुए लगभग 1300000 फसल कटाई परीक्षण किए गए। फसल कटाई परीक्षणों की संख्या बढ़ रही है और इस प्रकार, विभिन्न प्रकार की गैर-प्रतिदर्श त्रुटियों आदि ने आंकड़ों की गुणवत्ता को प्रभावित किया है। इसके अतिरिक्त, विभिन्न सरकारी विभागों में श्रमिकों की कमी व इनके संगठनों ने इन सर्वेक्षणों के संचालन को पहले की तुलना में अधिक कठिन बना दिया है। इस समस्या के समाधान के रूप में, भा.कृ.अनु.प.-भा.कृ.सां.अनु.सं., नई दिल्ली द्वारा अर्थशास्त्र एवं सांख्यिकी निदेशालय, कृषि एवं किसान कल्याण मंत्रालय, भारत सरकार द्वारा प्रायोजित एक प्रारंभिक मार्ग-दर्शन (pilot) अध्ययन किया गया था। इस अध्ययन का उद्देश्य सर्वेक्षणों के संचालन में सम्मिलित कर्मचारियों पर काम के बोझ को कम करने और आंकड़ों की गुणवत्ता में सुधार के लिए गांवों के कम प्रतिदर्श के आकार, अर्थात् फसल कटाई परीक्षणों की संख्या में कमी से प्रमुख फसल उपज के जिला स्तर के अनुमान प्राप्त करने का था।

अध्ययन का मुख्य उद्देश्य GCES के तहत हर साल आयोजित CCE की कम संख्या से प्रमुख फसल उपज का विश्वसनीय अनुमान उत्पन्न करने की संभावना का मूल्यांकन करना था। अर्थशास्त्र एवं सांख्यिकी निदेशालय, गुवाहाटी, द्वारा किए गए इस अध्ययन के परिणाम कृषि वर्ष 2015-16 की खरीफ की फसल के लिए असम राज्य में स्वतंत्र रूप से एकत्र किए गए आंकड़ों पर आधारित थे। राज्य से आंकड़े प्राप्त करने के बाद, आंकड़े विश्लेषण भा. कृ. अनु. प.-भा. कृ. सां. अनु. सं., नई दिल्ली में किया गया और परिणाम प्राप्त किए गए। उपज अनुमानों की तुलना संबंधित राज्य सरकार द्वारा वर्ष 2015-16 के लिए जारी किए

गए अनुमानों से की गई थी। निम्नलिखित अनुभागों में हम कार्यप्रणाली की चर्चा के साथ-साथ निष्कर्ष के बाद आंकड़ों से प्राप्त परिणामों पर चर्चा करेंगे।

सामग्री एवं विधि

इस अध्ययन के तहत भारत के 5 राज्यों असम, उत्तर प्रदेश, ओडिशा, कर्नाटक और गुजरात में एक सर्वेक्षण किया गया। भारत में कृषि वर्ष 2015-16 के लिए खरीफ और रबी दोनों मौसमों के लिए सर्वेक्षण किया गया था। इस लेख में हम असम पर ध्यान केंद्रित करेंगे। किए गए अध्ययन के तहत गांवों की कुल संख्या और प्रस्तावित फसल कटाई प्रयोगों की संख्या को तालिका 1 में दर्शाया गया है। इस तालिका से यह स्पष्ट है कि हम TRS योजना के तहत चयनित प्रतिदर्श आकार के संबंध में असम राज्य के लगभग 6% गांवों पर विचार कर रहे हैं। तदनुसार, CCE की कुल संख्या भी उसी अनुपात में कम हो जाएगी।

फसल उपज के प्रत्यक्ष अनुमानों को उत्पन्न करने के लिए प्रस्तावित प्रतिदर्श आकार को कुशलतापूर्वक कार्यान्वित करने के लिए, स्तरीकृत बहु-चरण यादृच्छिक प्रतिचयन का प्रयोग किया गया है जिसका विवरण इस प्रकार है:

- प्रत्येक चयनित राज्य में जिले को एक स्तर के रूप में माना गया है।
- एक जिले में 50% तहसीलों/तालुकों को बिना प्रतिस्थापन (SRSWOR) सरल यादृच्छिक प्रतिचयन द्वारा प्रथम चरण प्रतिदर्श इकाईयों (FSU) के रूप में चुना गया है। एक FSU के भीतर के गांवों को दूसरे चरण की प्रतिदर्श इकाईयों (SSU) के रूप में लिया जाता है, और प्रत्येक SSU के भीतर, सर्वेक्षण/उप-सर्वेक्षण संख्या को तीसरे चरण की प्रतिदर्श इकाईयों (TSU) के रूप में लिया गया है। चयनित सर्वेक्षण/उप-संख्या में निर्दिष्ट आकार के परीक्षात्मक प्लॉट को उपज अनुमान के लिए अंतिम चरण इकाई माना जाता है। SRSWOR का उपयोग करके SSU और TSU का चयन किया गया है। CCE भूखंडों के चयन के लिए फ्रेम के निर्माण के लिए चयनित गांव के भीतर 5 सर्वेक्षण

- संख्या के 20 समूहों के रूप में 100 सर्वेक्षण संख्या का चयन किया गया है।
- iii. निर्धारण वर्ष 2015–16 के खरीफ और रबी मौसम के दौरान फसल की गणना के लिए प्लॉट से प्लॉट तक के लिए प्राथमिक कार्यकर्ता द्वारा प्रत्येक चयनित सर्वेक्षण संख्या का दौरा किया गया है।
 - iv. CCE के लिए सर्वेक्षण संख्या के चयन के लिए फसल—वार फ्रेम यादृच्छिक रूप से चयनित 100 सर्वेक्षण संख्याओं के आधार पर तैयार किया गया है। प्रत्येक प्रमुख फसल के लिए, SRSWOR का उपयोग करते हुए CCE आयोजित करने के लिए दो सर्वेक्षण/उप—सर्वेक्षण संख्याओं का चयन किया गया है।
 - v. द्वितीय चरण इकाइयों के लिए प्रतिदर्श इकाइयों का चयन भा.कृ.अनु.प.—भा.कृ.सां.अनु.सं., नई दिल्ली द्वारा किया गया है, जबकि तीसरे चरण इकाइयों का चयन, अर्थात् सर्वेक्षण / उप—सर्वेक्षण संख्या दोनों क्षेत्र गणना और CCE के लिए, संबंधित राज्य क्षेत्र के अधिकारियों द्वारा किया गया है। क्षेत्र आंकड़े संग्रह के लिए, फ्रेम तैयार करने के लिए एक प्रश्नावली (Schedule) और फसल कटाई के परीक्षण के लिए दो प्रश्नावली तैयार की गई हैं। अनुमानों को विकसित करने के लिए असमान प्रथम चरण इकाइयों के साथ स्तरीकृत द्विचरणीय प्रतिचयन (सुखात्मे इत्यादि, 1984) का उपयोग किया और छोटे क्षेत्र के आकलन के लिए फे और हैरियट (1979) द्वारा प्रस्तावित समग्र स्तर SAE पद्धति (EBLUP मॉडल) का उपयोग किया गया था।

परिणाम और चर्चा

असम राज्य में, प्रस्तावित गांवों की संख्या 300 थी और कुल 1800 CCE का संचालन किया जाना था, जबकि खरीफ में केवल 288 गांवों और रबी सीजन के सभी 300 गांवों का सर्वेक्षण किया गया था और केवल 1282 CCE आयोजित किए गए थे। असम राज्य में खरीफ मौसम के दौरान प्रमुख फसल शीतकालीन धान पाई गई, जबकि रबी में प्रमुख फसल श्वेतसरसों (Rapeseed) और सरसों (Mustard) पाई गई। हम इस

अध्ययन के लिए केवल शीतकालीन धान पर विचार कर रहे हैं क्योंकि धान/चावल भारत में प्रमुख फसलों में से एक है। प्रस्तावित से कम CCE के संचालन के कारण, प्रतिदर्श आकार कम हो गया और कुछ जिलों में प्रस्तावित अनुमानों के प्रतिशतता गुणांक (% CV) में वृद्धि हुई। इसके अतिरिक्त, CCE की संख्या में कमी के कारण, कुछ जिलों में फसल उपज के जिला स्तर के प्रत्यक्ष अनुमानों के आकलन के लिए पर्याप्त प्रतिदर्श आकार नहीं पाया गया था। उन जिलों को गैर— प्रतिदर्श जिलों के रूप में माना जाता है और अनुमानों का आकलन क्षेत्र स्तर EWLUP अनुमानक के आधार पर SAE का उपयोग करके अन्य सभी जिलों के अनुमानों में सुधार के साथ लगाया जाता है। SAE मॉडल को फिट करने के लिए, उर्वरक सांख्यिकी (2015–16), The Fertilizer Association of India, नई दिल्ली द्वारा जारी उत्तर प्रदेश और असम राज्य के लिए वर्ष 2015–16 के जिला स्तरीय उर्वरक खपत आंकड़ों को सहायक जानकारी के रूप में उपयोग किया गया था। SAE अनुमानक के निर्माण के लिए नाइट्रोजन (N), फास्फोरस (P) और पोटेशियम (K) उर्वरकों की कुल उर्वरक खुराक आंकड़ों का उपयोग सहायक जानकारी के रूप में किया गया था। यह पाया गया कि असम राज्य में चावल की उपज और उर्वरक की मात्रा में सहसंबंध 0.5–0.6 के मध्य था।

जिला स्तरीय उपज अनुमान प्रस्तावित अभिकल्पना आधारित पद्धति का उपयोग करके तैयार किए गए हैं और प्रतिदर्श जिलों के लिए संबंधित राज्यों में स्वीकार्य % CV के साथ GCES योजना के तहत जारी अनुमानों के साथ लगभग समान पाया गया है। चावल की उपज और धान की उपज के बीच का अंतर एक गुणक है जिसका उपयोग धान से चावल के उत्पादन के अनुमान के दौरान भूसी की दर को निरूपित करने के लिए किया जाता है। असम राज्य में आयोजित CCE की औसत संख्या इस अध्ययन के तहत शीतकालीन धान के लिए आयोजित CCE की औसत संख्या की केवल 10% है। अब GCES के तहत धान/चावल और गेहूं जैसी प्रमुख फसलों के लिए जिला स्तर पर आयोजित CCE की औसत संख्या लगभग 100–120

तालिका 1. असम राज्य में पायलट सर्वेक्षण के तहत प्रस्तावित गांवों और फसल कटाई परीक्षणों की कुल संख्या

समय पर रिपोर्टिंग योजना (TRS) के तहत प्रस्तावित गांवों की संख्या	प्रायोगिक सर्वेक्षण में चयनित गांवों की संख्या	फसल कटाई परीक्षणों की संख्या	
		खरीफ़	रबी
5327	300	1200	600

तालिका 2. कृषि वर्ष 2015-16 के लिए असम राज्य में % CV के साथ उपज का जिलावार अनुमान

क्र.सं.	जिला	उपज (कि./हे.)	%CV	EBLUP अनुमान	hCV	GCES अनुमान
1	बक्सा	2975.30	4.00	3009.97	3.88	3267
2	बारपेटा	2193.20	13.57	2392.52	10.05	2272
3	बोंगोईगांव	2595.12	10.56	2803.90	8.97	2510
4	कछार	3169.66	5.76	3256.39	5.39	3345
5	चिरांग	3087.56	4.29	3046.86	4.20	2679
6	दरांग	3236.45	11.81	3278.41	9.94	3469
7	धेमाजी	2465.29	3.10	2241.58	3.07	2293
8	धुबरी	3989.76	5.61	3943.58	5.41	2658
9	डिब्रूगढ़	4350.06	3.32	4282.31	3.31	3042
10	दीमा हसाओ			3620.83	15.84	3766
11	गोलपाडा	4061.14	16.29	3749.91	10.97	3295
12	गोलाघाट	3271.85	7.52	3353.06	6.84	3438
13	हैलाकांडी	3514.34	8.17	3474.35	7.55	3725
14	जोरहाट	3547.44	10.03	3112.68	8.62	3404
15	के. आंगलॉग			2807.98	14.46	2575
16	कामरूप (एम)			3366.51	16.63	3395
17	करीमगंज	3594.01	16.51	3681.08	10.77	2635
18	कोकराझार	2458.64	8.40	2610.83	7.53	2595
19	लखीमपुर	3789.12	5.06	3754.29	4.86	3138
20	मोरीगांव	2718.85	16.02	2877.89	10.88	2960
21	नागांव	3015.46	4.70	3050.81	4.62	3377
22	नलबाडी	2957.99	5.04	2955.38	4.89	3192
23	शिवसागार	3200.00	1.83	3185.68	1.83	3155
24	सोनितपुर	3539.60	9.05	3304.97	7.99	3294
25	तिनसुकिया	2989.91	8.55	3152.84	7.43	2890
26	उदलगुरी	3343.35	6.47	3360.67	6.09	3298

श्रोत: वैद्यनाथन समिति की रिपोर्ट (2019-20)

है। इसलिए असम में यह CCE के राष्ट्रीय औसत की केवल 10% है।

असम राज्य के लिए तालिका 2 से यह स्पष्ट है कि अभिकल्पना आधारित अनुमानक का % CV औसतन 8.07 है जिसकी सीमाएं 1.83 से 16.51 तक है जबकि SAE अनुमानों का औसत % CV सभी प्रतिदर्श जिलों के लिए 6.74 है जिसकी सीमाएं 1.83 से 10.97 तक है। गैर-प्रतिदर्श जिलों के लिए, SAE अनुमानक का % CV 15.84 से 16.63 की सीमा में है। इस मामले में हम यह भी देख सकते हैं कि SAE ने गैर-प्रतिदर्श जिलों के अनुमानों के साथ प्रस्तावित अभिकल्पना आधारित अनुमानों की तुलना में अधिक सटीक अनुमान प्रदान किए। GCES के अनुमानों के साथ परिणामों की तुलना करने के बाद, यह पाया गया कि धुबरी (जिला) को छोड़कर अन्य सभी जिलों के लिए धान की उपज का अनुमान बहुत कम विचलन के साथ लगभग समान है। गैर-प्रतिदर्श वाले जिलों में SAE का उपयोग करके सृजित अनुमान जिला स्तर पर GCES योजना के स्वीकार्य % CV के तहत प्रकाशित अनुमानों के समान पाए गए हैं। असम के लगभग 96% जिलों के इस अध्ययन से हम प्रत्येक जिले में GCES के तहत औसतन लगभग 10% CCE से गेहूं धान की फसल की उपज का सटीक अनुमान लगाने में सक्षम हैं।

निष्कर्ष

प्रारंभिक मार्ग-दर्शन (pilot) अध्ययन का उद्देश्य, कृषि आंकड़े अर्थात् खेती की गई फसलों के उत्पादन अनुमानों के लिए राज्य और राष्ट्रीय एजेंसी के कार्य भार को कम करना था। अब प्रत्येक फसल के उत्पादन अनुमान के लिए उपज अनुमान के साथ क्षेत्र के अनुमानों को गुणा करके व्युत्पन्न किया जाता है। क्षेत्र के अनुमानों की संरचना के मामले में, स्मार्ट प्रतिदर्श योजना के साथ दूर संवेदी (Remote Sensing) और एरियल फोटोग्राफी जैसी आधुनिक तकनीकें बहुत

उपयोगी, कम समय लेने वाली और काफी सटीक पाई गई और प्रत्येक के तहत क्षेत्र अनुमान तैयार करते समय एजेंसी के लिए राहत के रूप में आई। लेकिन फसल उपज अनुमानों के उत्पादन के मामले में निहित बड़ी समस्या यह थी कि आज तक ऐसी कोई मौजूदा तकनीक नहीं थी जो फसल की उपज का उद्देश्यपूर्ण तरीके से अनुमान लगाने में सक्षम हो। कई तकनीकों जैसे दूर संवेदी, GIS, मशीन लर्निंग आदि को खाद्य एवं कृषि संगठन (FAO), रोम जैसे राष्ट्रीय और अंतर्राष्ट्रीय संगठनों द्वारा फसल उपज के वस्तुनिष्ठ अनुमानों की उत्पत्ति के लिए आजमाया गया लेकिन असफल रहे। ये सभी तकनीकें निश्चित स्तर की सटीकता के साथ भविष्यवाणी आधारित अनुमान उत्पन्न करती हैं जो CCE द्वारा उत्पन्न वस्तुनिष्ठ अनुमानों के रूप में स्वीकार्य नहीं पाया जाता है। इसलिए CCE को फसल उपज के आकलन के लिए एक स्वर्ण मानक विधि कहा जाता है (सूद इत्यादि, 2017)। कार्यभार को कम करने और उपलब्ध मानव शक्ति के साथ आंकड़ों की गुणवत्ता में सुधार करने के लिए, GCES योजना के तहत किए गए फसल कटाई परीक्षण की कुल संख्या को कम करने का यही एकमात्र तरीका बचा था। इस लक्ष्य को प्राप्त करने के लिए, हमारे अध्ययन को सफलता मिली, जहां हम अपने देश की दो प्रमुख फसलों के लिए उत्तर प्रदेश और असम के अधिकांश जिलों (90% से अधिक) के लिए फसल उपज के समान सटीक अनुमान प्राप्त करने में सक्षम हैं। GCES योजना के तहत आयोजित 50% से कम CCE से चावल और गेहूं फसल उपज के सटीक अनुमान प्राप्त हुए। इसके अतिरिक्त, SAE तकनीक का उपयोग करते हुए, गैर-प्रतिदर्श जिलों के अनुमानों के साथ-साथ अनुमानों की शुद्धता में सुधार किया गया। इसलिए हम यह निष्कर्ष निकाल सकते हैं कि प्रमुख खाद्यान्न फसलों के लिए GSES योजना के तहत आयोजित 50% से कम CCE से अधिक और स्वीकार्य सटीकता के साथ फसल उपज का अनुमान लगाया जा सकता है।

संदर्भ

फे, आर. ई. एवं हेरियट, आर. ए. (1979). एस्टिमेशन ऑफ़ इनकम फ्रॉम स्माल प्लेसिज: एन एप्लीकेशन ऑफ़ जेम्स-स्टेन प्रोसिजरस टू सेन्सस डेटा। *जर्नल ऑफ़ द अमेरिकन स्टैटिस्टिकल एसोसिएशन*, 74, 269–277।

फर्टिलाइज़र स्टैटिस्टिक्स (2015–16). द फर्टिलाइज़र एसोसिएशन ऑफ़ इंडिया, नई दिल्ली।

प्रसाद, एन. जी. एन. एवं राव, जे. एन. के. (1990). द एस्टिमेशन ऑफ़ द मीन स्कोरड एरर ऑफ़ द स्माल एरिया एस्टिमेशन। *जर्नल ऑफ़ द अमेरिकन स्टैटिस्टिकल एसोसिएशन* 85, 163–171

राव, जे. एन. के. (2003). *स्माल एरिया एस्टिमेशन*। विली, न्यूयॉर्क।

रिपोर्ट ऑन एक्सपर्ट कमिटी ऑन एग्रीकल्चरल स्टैटिस्टिक्स (2011). डायरेक्टरेट ऑफ़ इकोनॉमिक्स

एंड स्टैटिस्टिक्स, मिनिस्ट्री ऑफ़ एग्रीकल्चर एंड फार्मर वेलफेयर, गवर्नमेंट ऑफ़ इंडिया।

सिंह, आर. एवं मंगत, एन. पी. एस. (1996). *एलिमेंट्स ऑफ़ सर्वे सैप्लिंग*। कलुवेर ऐकडेमिक पब्लिशर्स।

सूद, यू. सी., अहमद, टी., गुप्ता, वी. के., चंद्रा, एच., साहू, पी. एम., आदित्य, के., सिंह, एम. एवं बिस्वास, ए. (2017). क्रॉप 1 एंड 2: मेंअज़रिंग क्रॉप एरिया एंड यील्ड अंडर प्योर स्टैंड, मिक्स्ड एंड कंटीन्यूअस क्रॉपिंग। वर्किंग पेपर नंबर 16 ग्लोबल स्ट्रेटेजी, FAO रोम पब्लिकेशन।

सुखात्मे पी. वी., सुखात्मे बी. वी., सुखात्मे एस., एवं अशोक, सी. (1984). *सैप्लिंग थ्योरी ऑफ़ सर्वेज विद एप्लीकेशन्स*, थर्ड एडिशन, आईओवा स्टेट यूनिवर्सिटी प्रेस, युएसए एंड इंडियन सोसाइटी ऑफ़ एग्रीकल्चरल स्टैटिस्टिक्स, नई दिल्ली।

“

विज्ञान मानवता के लिए सुंदर उपहार है,
हमें इसे विकृत नहीं करना चाहिए।

”

बहिर्जात चर के साथ गॉम्पर्ट्ज़ प्रसंभाव्य विभिन्नात्मक समीकरण पद्धति

हिमाद्रि घोष एवं सविता वधवा

भा.कृ.अनु.प.—भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली—110012

सारांश

गुतिरेज़ इत्यादि (2005) ने बहिर्जात चर के साथ गॉम्पर्ट्ज़ समरूप व्यापन (diffusion) प्रक्रिया का अध्ययन किया जो प्रवृत्ति को प्रभावित करती है। इस लेख का उद्देश्य गुतिरेज़ इत्यादि के कार्य में दो संशोधन करना है। पहला संशोधन गुतिरेज़ प्रतिमान में व्यापन अवधि को काल-निर्भर बनाना है, जो अधिक यथार्थवादी है। दूसरा संशोधन वेवलेट विश्लेषण की प्रभावशाली तकनीक को ड्रिफ्ट अवधि का अनुमान लगाने के लिए नियोजित करना है। उपरोक्त दो पहलुओं को शामिल कर पद्धति विकसित की गई। प्रयोग के लिए उपयुक्त कंप्यूटर प्रोग्राम एसएएस (SAS) और आर (R) सॉफ्टवेयर पैकेज में लिखे गए। अंत में, उदाहरण के रूप में, भारत के कुल खाद्यान्न उत्पादन के काल-श्रृंखला आँकड़ों का अध्ययन किया गया जो बाहरी चर के रूप में वर्षा, उर्वरक और कीटनाशक काल-श्रृंखला आँकड़ों पर निर्भर हैं। गुतिरेज़ इत्यादि (2005) द्वारा प्रस्तावित प्रतिमान की तुलना में इस लेख में प्रस्तावित प्रतिमान की श्रेष्ठता दर्शाई गई।

प्रस्तावना

अरैखिक गॉम्पर्ट्ज़ विकसित प्रतिमान को विभिन्न विषयों जैसे कृषि, दवा, और उद्योग में व्यापक रूप से नियोजित किया जाता है। यह प्रतिमान क्रियाविधिक है अतः अंतर्निहित प्राचलो की विशिष्ट जैविक व्याख्या होती है। प्रसंभाव्य विभिन्नात्मक समीकरण (एसडीई) प्रतिमान के कई लेख (घोष एवं प्रज्ञेष् 2017) जो कि गॉम्पर्ट्ज़ प्रसंभाव्य विभिन्नात्मक समीकरण को दर्शाते हैं, प्रकाशित किये गए हैं। इन प्रतिमानों की कमी

यह है कि उनकी प्रक्रिया के विकास का प्रतिमान काल-निर्भर नहीं है। सेमिनल शोधपत्र में गुतिरेज़ इत्यादि (2005) ने काल के फलन के रूप में बहिर्जात चर का उपयोग किया जो प्रवृत्ति को प्रभावित करता है तथा प्रसंभाव्य गॉम्पर्ट्ज़ प्रतिमान के प्रयोग के लिए पद्धति विकसित की। इस लेख का उद्देश्य गुतिरेज़ इत्यादि (2005) की पद्धति में दो संशोधन करना है। पहला संशोधन गुतिरेज़ प्रतिमान में व्यापन अवधि को काल-निर्भर बनाना है। गुतिरेज़ प्रतिमान में टुकड़ों द्वारा काल-आधारित रैखिक फलन का उपयोग बहिर्जात कारक द्वारा ड्रिफ्ट अवधि का अनुमान लगाने के लिए किया गया जो कि यथार्थवादी नहीं है। अतः दूसरे संशोधन में ड्रिफ्ट अवधि का अनुमान लगाने के लिए वेवलेट विश्लेषण (घोष इत्यादि, 2010) की प्रभावशाली तकनीक का प्रयोग किया गया। परिणाम स्वरूप प्राप्त हुए प्रतिमान का नाम “संशोधित गुतिरेज़ प्रतिमान” दिया गया।

विधि

गुतिरेज़ इत्यादि (2005) ने अरैखिक गॉम्पर्ट्ज़ समरूप व्यापन प्रक्रिया का अध्ययन काल-फलन के माध्यम से किया जो इसकी प्रवृत्ति को प्रभावित करता है। हालाँकि उपरोक्त पद्धति की एक सीमा यह है कि यदि काल-अंतराल में घातीय प्रवृत्ति उपस्थित हो तो यह आँकड़ों का वर्णन करने में सक्षम नहीं है। इस पद्धति में लघुगणक प्रक्रिया के गॉशियन माध्य फलन $X_G^{(ln)}(t) = \log X_G(t)$ के लिए घातीय प्रवृत्ति के कारण $X_G(t)$ का प्रतिबंधित माध्य समरूप नहीं होता। इसलिए इस लेख में प्रक्रिया $X_G(t)$ का प्रक्रिया

$X_{G,d}(t)$ के लिए सामान्यीकरण का प्रस्ताव दिया गया है। काल-प्रवृत्ति के कारण $b(x, t) = g^2(t)x^2$ है जहाँ $g(t) = \exp(K + \gamma t)$ है।

इसके इलावा अरैखिक गॉम्पर्टज़ एसडीई प्रतिमान को काल-निर्भर व्यापन $b(x, t) = \exp\{2(K + \gamma t)\}x^2$ के साथ समान्य करने के लिए गुतिरेज़ इत्यादि (2005) के अनुसार यह पाया गया कि $f(y, \tau|x, t)$ प्रक्रिया $X_{G,d}(t)$ के माध्य के साथ लॉग नार्मल है जो कि निम्न में दर्शाया गया है:

$$E[X_{G,d}^{(ln)}(\tau)|\mathcal{F}_t] = \left\{ \exp(-\beta(\tau - t)) \log x - \frac{\exp(2K)}{(\beta + 2\gamma)} (\exp(2\gamma\tau) - \exp(-\beta(\tau - t) + 2\gamma t)) + \int_t^\tau h(s) \exp(-\beta(\tau - s)) ds \right\}$$

$$V[X_{G,d}^{(ln)}(\tau)|\mathcal{F}_t] = \frac{(1 - \exp(-2(\beta + \gamma)(\tau - t))) \exp(2(\gamma\tau + K))}{(\beta + \gamma)} \quad (1)$$

$X_{G,d}^{(ln)}(\tau)$ का प्रतिबंधित प्रसरण समरूप नहीं है। अतः यह भिन्न काल में ड्रिफ्ट एवं व्यापन गुणांक को निर्दिष्ट करते हुए घातीय प्रवृत्ति का वर्णन करने में सक्षम है। निम्न विकसित किये गए:

$$c_1(t) = (2/g(t))(h(t) - g^2(t) - \beta \log z), \quad c_2(t) = -2(\beta + (g'_t/g_t))$$

प्रस्तावित प्रतिमान का अध्ययन करने के लिए x_1, x_2, \dots, x_n प्रेक्षित मान हैं। $v_1 = x_1$ और $v_{i,\beta} = \log x_i - \exp(-\beta) \log x_{i-1}$ मान लिए गए। इसका लाईक्लीहुड फलन है:

$$L_{v_{2,\beta}, v_{3,\beta}, \dots, v_{n,\beta}} = \left(\frac{\beta}{2\pi\sigma^2(1 - \exp(-2\beta))} \right)^{\binom{n-1}{2}} \exp\left(-\frac{\beta(v_\beta - \gamma_\beta U'_\beta a)'(v_\beta - \gamma_\beta U'_\beta a)}{2\sigma^2(1 - \exp(-2\beta))} \right), \quad (2)$$

$$\text{जहाँ } \mathbf{u}_{i,\beta} = \left(1, \frac{1}{\gamma_\beta} \int_{t_{i-1}}^{t_i} g_1(\tau) \exp(-\beta(t_i - \tau)) d\tau, \dots, \frac{1}{\gamma_\beta} \int_{t_{i-1}}^{t_i} g_1(\tau) \exp(-\beta(t_i - \tau)) d\tau \right)'$$

$$\text{साथ } \mathbf{v}_\beta = (v_{2,\beta}, v_{3,\beta}, \dots, v_{n,\beta})'$$

$$\mathbf{a} = (\alpha_0 - \sigma^2, \alpha_1, \dots, \alpha_q)'$$

$$\gamma_\beta = ((1 - \exp(-\beta))/\exp(-\beta)),$$

और $U_\beta = (\mathbf{u}_{2,\beta}, \mathbf{u}_{3,\beta}, \dots, \mathbf{u}_{n,\beta}) \quad (q + 1) \times (n - 1)$ मैट्रिक्स है।

इसी प्रकार प्रक्रिया $X_{G,d}(t)$ के संभावित (Likelihood) फलन का परिकलन किया गया। गुतिरेज़ इत्यादि (2005) के समान प्रस्तावित लेख में भी काल-अंतराल $[t_{i-1}, t_i]$ की लम्बाई "एक" अनुमानित की गई। अधिकतर काल-श्रृंखला आँकड़े समान अंतराल पर उपलब्ध होते हैं। अतः यह प्रतिबंधन हमेशा लागू नहीं किया जा सकता। गुतिरेज़ इत्यादि (2005) में $(q-1)$ टुकड़ों के रैखिक फलन $g_{ij}(t); i = 2, 3, \dots, n; j = 1, 2, \dots, q$ के बहुभुज फलन का प्रयोग किया गया जो निर्देशांक (t_i, y_i) के माध्यम से गुजरते हैं। जहाँ y_i वृद्धि को दर्शाता है। इसलिए वेवलेट विश्लेषण (घोष इत्यादि, 2010) तकनीक का उपयोग किया गया। फूरियर विश्लेषण के विपरीत वेवलेट विश्लेषण का उपयोग प्रेक्षित आँकड़ों को विघटित करने के लिए किया गया। पिरामिड परिकलन आधारित बहुक्रिया विश्लेषण का उपयोग वेवलेट विश्लेषण के भिन्न स्केल पर बहिर्जात चर अनुमानित करने के लिए वेवलेट रूपांतरण किया गया। हार एवं डाबेचिस फिल्टर का उपयोग विभिन्न स्केल पर नामिनल पास-बैंड में किया गया, जो उच्च पास फिल्टर का वर्णन करने में सक्षम है। (smooth) भाग को काल में माध्य फलन के रूप में माना गया। इसे $g_j(t)$ फलन की रचना के लिए उपयोग किया गया। तदानुसार बहुभुज फलन के आकलन के लिए एक नई पद्धति वेवलेट रूपांतरण का प्रस्ताव बहिर्जात चर के प्रेक्षित मान के स्थान पर दिया गया। $\{X_G(t); t \in [t_0, T]\}$ प्रक्रिया में ड्रिफ्ट अवधि के प्रभाव को स्मूथ करने के लिए वेवलेट रूपांतरण \mathcal{W} की उप-मैट्रिक्स को बहिर्जात चरों के

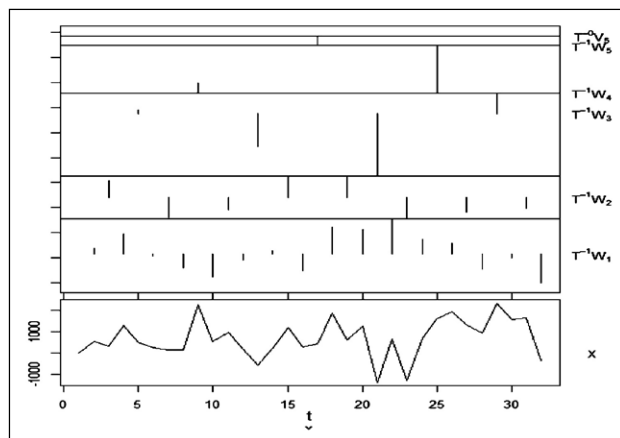
वेवलेट रूपांतरण वेक्टर से गुणा किया गया। संशोधित गॉम्पर्टज़ प्रतिमान का आकलन करने के लिए उपयुक्त कोड आर और एसएसएस पैकेज में लिखे गये।

उदाहरण

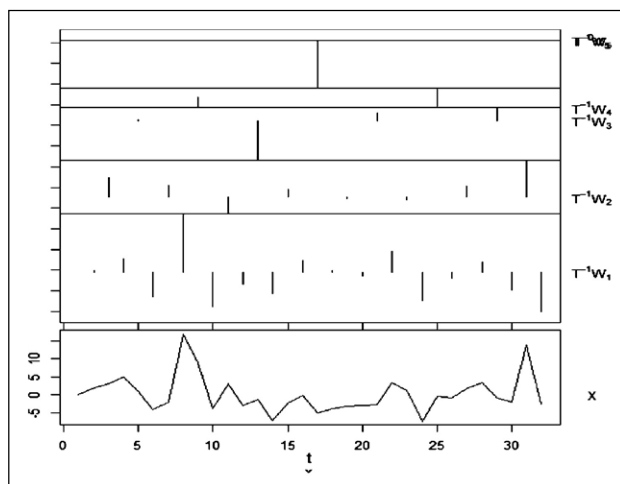
उदाहरण के रूप में भारत के कुल खाद्यान्न उत्पादन 1982–83 से 2015–16 की अवधि के लिए काल-श्रृंखला आँकड़े “कृषि सांख्यिकी एक झलक 2015” से लिए गए। 2013–14 तक की अवधि के आँकड़ों का उपयोग प्रतिमान के आकलन के लिए एवं शेष 2 वर्षों के आँकड़ों का उपयोग सत्यापन के लिए किया गया। कुल खाद्यान्न उत्पादन मुख्य रूप से वर्षा, उर्वरक की खपत और कीटनाशक की खपत पर निर्भर करता है। इन तीनों चरों को बहिर्जात चर के रूप में लिया गया। वर्षा पर काल-श्रृंखला आँकड़े भारतीय उष्णकटिबंधीय मौसम विज्ञान संस्थान की वेबसाइट www.tropmet.res.in पर उपलब्ध हैं। उर्वरक और कीटनाशक की खपत के आँकड़े क्रमशः उर्वरक सांख्यिकी और वेबसाइट www.Indiastat.com से लिए गए।

इन तीन बहिर्जात चरों पर आँकड़ों को स्मूथ किया गया। हार एवं डाबेचिस फिल्टर का उपयोग किया गया और यह पाया कि डाबेचिस (db4) फिल्टर ने सर्वोत्तम परिणाम प्रदान किए। जिन्हें रेखाचित्र-1 में प्रदर्शित किया गया। गुतिरेज़ प्रतिमान को आँकड़ों के आकलन के लिए प्रयोग किया गया और परिणाम तालिका-1 के दूसरे सतंभ में दर्शाये गए। विकसित एसएसएस कोड का प्रयोग करते हुए संशोधित गुतिरेज़ प्रतिमान को आँकड़ों के आकलन के लिए प्रयोग किया गया। तालिका-1 के तीसरे सतंभ में इसके परिणाम दर्शाये गए। तालिका-1 में दिये गए आकलित मान का उपयोग करते हुए भारत के खाद्यान्न उत्पादन; मिलियन टन पर दोनों प्रतिमानों की गणना की गई एवं परिणाम तालिका-2 में दर्शाये गए। औसत वर्ग माध्य त्रुटि (एमएसई) की गणना की गई। संशोधित गुतिरेज़ प्रतिमान की एमएसई का मान 209.61 है और गुतिरेज़ प्रतिमान की एमएसई का मान 224.82 है। अतः न्यूनतम मान, संशोधित गुतिरेज़ प्रतिमान की श्रेष्ठता

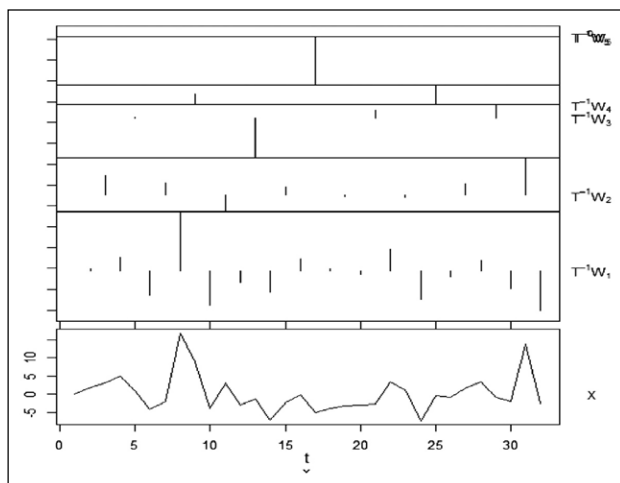
(i) वर्षा:



(ii) उर्वरक:



(iii) कीटनाशी:



रेखाचित्र 1: डाबेचिस (db4) फिल्टर द्वारा बहिर्जात चरों की स्मूथिंग

को दर्शाता है। सत्यापन के उद्देश्य से दोनों प्रतिमानों के प्रदर्शन की तुलना दो वर्षों के पूर्वानुमान के लिए की गई। परिणाम तालिका-3 में दर्शाये गए। तालिका-3 से पता चलता है कि संशोधित गुतिरेज प्रतिमान के पूर्वानुमान गुतिरेज प्रतिमान की तुलना में वास्तविक मूल्यों के बहुत समीप हैं। अतः विचाराधीन आँकड़ों पर पूर्वानुमान के लिए भी संशोधित गुतिरेज प्रतिमान का प्रदर्शन उत्तम सिद्ध हुआ। आकलित संशोधित गुतिरेज प्रतिमान का रेखाचित्र प्रेक्षित आँकड़ों के साथ रेखाचित्र-2 में दर्शाया गया।

तालिका 1: गुतिरेज प्रतिमान और संशोधित गुतिरेज प्रतिमान के प्राचलों का आकलन

प्राचल	आकलन	
	गुतिरेज प्रतिमान	संशोधित गुतिरेज प्रतिमान
α_0	4.700×10^{-2}	1.800×10^{-2}
α_1	3.000×10^{-5}	-1.607×10^{-5}
α_2	1.412×10^{-5}	-1.266×10^{-10}
α_3	-8.249×10^{-3}	-1.007×10^{-2}
B	4.172×10^{-3}	8.021×10^{-9}
σ_0	3.213×10^{-3}	-
K	-	2.749×10^{-3}
γ	-	7.401×10^{-3}

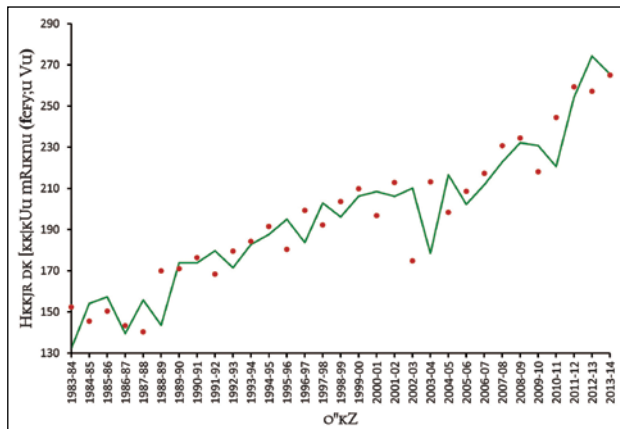
तालिका-2: भिन्न प्रतिमानों के लिए भारत के खाद्यान्न उत्पादन (मिलियन टन) का वार्षिक आकलन

वर्ष	वास्तविक मान	गुतिरेज प्रतिमान	संशोधित गुतिरेज प्रतिमान
1983-84	152.37	133.31	132.32
1984-85	145.54	160.97	154.11
1985-86	150.44	151.86	157.27
1986-87	143.42	152.46	139.47
1987-88	140.35	146.56	155.79
1988-89	169.92	143.07	143.50
1989-90	171.04	174.58	173.87
1990-91	176.39	175.35	173.79

वर्ष	वास्तविक मान	गुतिरेज प्रतिमान	संशोधित गुतिरेज प्रतिमान
1991-92	168.38	180.95	179.67
1992-93	179.48	172.73	171.43
1993-94	184.26	184.07	182.74
1994-95	191.50	188.96	187.61
1995-96	180.42	196.35	194.98
1996-97	199.34	185.00	183.70
1997-98	192.26	204.53	202.87
1998-99	203.61	196.80	196.16
1999-00	209.80	204.69	206.17
2000-01	196.81	211.78	208.45
2001-02	212.85	198.12	206.09
2002-03	174.77	214.32	210.13
2003-04	213.19	179.42	178.42
2004-05	198.36	218.41	216.58
2005-06	208.60	203.08	202.24
2006-07	217.28	214.53	211.75
2007-08	230.78	222.69	222.83
2008-09	234.47	230.07	232.15
2009-10	218.11	234.28	230.80
2010-11	244.49	225.70	220.60
2011-12	259.29	261.27	254.30
2012-13	257.13	275.20	274.23
2013-14	265.04	266.69	265.50
औसत एमएसई		224.82	209.61

तालिका 3: भिन्न प्रतिमानों के लिए भारत के खाद्यान्न उत्पादन (मिलियन टन) का वार्षिक पूर्वानुमान

वर्ष	वास्तविक मान	गुतिरेज प्रतिमान	संशोधित गुतिरेज प्रतिमान
2014-15	252.02	274.43	263.18
2015-16	251.57	280.02	272.37



रेखाचित्र 2: प्रेक्षित मान के साथ संशोधित गुतिरेज प्रतिमान के आकलित मान

निष्कर्ष

बहिर्जात चर और काल-निर्भर व्यापन के साथ अरैखिक गॉम्पर्टज़ व्यापन प्रतिमान के लिए पद्धति विकसित की गई। हालाँकि यह पद्धति उन काल-श्रृंखला आँकड़ों पर लागू होती है जो सभी चरों के साथ समान अंतराल पर उपलब्ध हों। काल-श्रृंखला आँकड़ों का संग्रह करने के लिए समय, कर्मियों और बजट आदि की बाधाएँ रहती हैं। डेनिस और पॉसियानो (2014) ने महत्व दिया कि असमान काल-अंतराल के आँकड़े या लुप्त आँकड़े अवलोकन के लिए संभावित रूप से सूचनात्मक हैं और इस तरह के आँकड़ों को छोड़कर किये गए विश्लेषण से निष्कर्षों पर प्रतिकूल प्रभाव पड़ता है अतः इस लेख में विकसित पद्धति का विस्तार करने की आवश्यकता है ताकि असमान अंतराल के

आँकड़ों का प्रयोग किया जा सके। इस दिशा में कार्य प्रगति पर है। इस लेख में विकसित पद्धति का उपयोग रिचर्ड्स और लॉजिस्टिक प्रतिमान के आकलन के लिए किया जा सकता है।

संदर्भ

कोहेन, एस. एवं इलियट, आर. जे. (2015). *स्टोकेस्टिक केलकुलस और ऐप्लीकेशन्स*, 2 एडिशन बिकरहॉसर, स्विट्ज़रलैंड।

डेनिस, बी. एवं पॉसियानो, जे. एम. (2014). डेनिसिटी डिपेंडेंट स्टेट स्पेस मॉडल फॉर पापुलेशन ऐबनडेंस डेटा विद अनइक्युअल टाईम इन्टरवल। *इकोलोजी*, **95**, 2069–2076।

घोष, एच., पॉल, आर. के. एवं प्रज्ञेषु (2010). वेवलेट फ्रिक्यूवेंसी डोमेन ऐपरोच फॉर स्टेटिस्टिकल मॉडलिंग फॉर इंडियन मानसून रेनफॉल टाईम-सीरीज़ डेटा। *जर्नल ऑफ स्टेटिस्टिकल थ्योरी एण्ड प्रेक्टिस*, **4**, 813–825।

घोष, एच. एवं प्रज्ञेषु (2010). गॉम्पर्टज़ ग्रोथ मॉडल इन रेंडम एन्वायरमेंट विद टाईम-डिपेंडेंट डिफ्यूजन। *जर्नल ऑफ स्टेटिस्टिकल थ्योरी एण्ड प्रेक्टिस*, **11**, 746–758।

गुतिरेज, आर., गुतिरेज-सांचेज, आर., नफीदी, ए., रोमन, पी. एवं टॉरेस, एफ. (2005). इनफरेंस इन गॉम्पर्टज़-टाईप नॉनहोमोजीनिस स्टोकेस्टिक सिस्टम बाए मीन्स ऑफ डिसक्रीट सैम्पलिंग। *साइब्रन सिस्टम*, **36**, 203–216।

“

फुटबाल की तरह जिंदगी में तुम तब तक आगे नहीं बढ़ सकते,
जब तक तुम्हें यह नहीं पता होता की गोल-पोस्ट कहाँ है।

”

—आर्नोल्ड ग्लासो

कुमारस्वामी बंटन: प्रसामान्य बंटन की एक नई श्रेणी

पंकज दास, कौस्तव आदित्य एवं भारती

भा.कृ.अनु.प.—भारतीय कृषि सांख्यिकी अनुसन्धान संस्थान, नई दिल्ली—110012

सार

कुमारस्वामी ने जल विज्ञान संबंधी अनुप्रयोगों के साथ द्विक परिबंध यादृच्छिक प्रक्रियाओं के लिए बंटन को प्रस्तावित किया। किसी भी सतत बेसलाइन जी बंटन के लिए, जी.एम. कॉर्डेइरो और एम. डी कास्ट्रो कई प्रसिद्ध बंटनों में सामान्य, वीबुल, गामा बंटन का विस्तार करने के लिए एक व्यापकीकृत बंटन की नई श्रेणी का वर्णन करते हैं। नए श्रेणी में कुछ विशेष बंटनों जैसे Kw-सामान्य, Kw-वीबुल, और Kw-गामा बंटन पर चर्चा की गयी है। हम किसी भी Kw व्यापकीकृत बंटन के सामान्य आघूर्णों की चर्चा मूल बंटन के प्रायिकता भारित आघूर्णों के रैखिक फलनों के रूप में करते हैं। हमने आधारभूत बंटन के संभाव्यता भारित आघूर्णों के फलन के रूप में क्रम-प्रतिदर्शज के सामान्य आघूर्ण भी प्राप्त किये हैं। नए वर्ग में बंटन को फिट करने के लिए अधिकतम संभावना प्राकलन विधि का उपयोग किया है और वास्तविक डेटा (data) के लिए दो अनुप्रयोगों के साथ नए मॉडल की क्षमता का वर्णन किया है।

मुख्य शब्द: गामा बंटन, कुमारस्वामी बंटन, आघूर्ण, प्रसामान्य बंटन, क्रम-प्रतिदर्शज, वीबुल बंटन

1. परिचय

बीटा बंटन एक बहुमुखी बंटन है और विभिन्न प्रकार की अनिश्चितताओं को इसके द्वारा प्रतिरूपण किया जा सकता है। वास्तविक स्थिति में, कई परिमित श्रेणी बंटनों को आसानी से मानक बीटा बंटन में परिवर्तित किया जा सकता है। अर्थमिति में, कई बार डेटा को परिमित श्रेणी बंटनों द्वारा प्रतिरूपित किया जाता है। व्यापकीकृत बीटा बंटन का व्यापक रूप से आँकड़ों में अध्ययन किया गया है और कई लेखकों ने इन बंटनों की विभिन्न श्रेणियों को विकसित किया है।

यूजीन एट अल (2002) ने दो मापदंडों को नियोजित करके बीटा यादृच्छिक चर के लॉगिट से परिभाषित एक यादृच्छिक चर के लिए बंटन के एक व्यापक वर्ग का प्रस्ताव रखा। यूजीन एट अल (2002) के बीटा प्रसामान्य बंटन को परिभाषित करने के बाद, नादराजाह और कोट्ज़ (2004) ने बीटा गंबेल बंटन को प्रस्तावित किया, नादराजा और गुप्ता (2004) ने बीटा फ्रीचेट बंटन का प्रस्ताव रखा और नादराजा और कोट्ज़ (2004) ने बीटा घातीय बंटन पर काम किया। हालाँकि, इसमें कुछ गणितीय जटिलता हैं क्योंकि बीटा बंटन ट्रैक्टेबल नहीं है और इसके संचयी बंटन फलन (cdf) में अपूर्ण बीटा फलन अनुपात शामिल है। कुमारस्वामी ने निम्न और ऊपरी परिबंध वाले चरों के लिए एक नया प्रायिकता बंटन प्रस्तावित किया। प्रायिकता और सांख्यिकी में, कुमारस्वामी द्विक परिबंध बंटन वितरण अंतराल (0, 1) पर परिभाषित सतत प्रायिकता बंटन की एक श्रेणी है, जो दो गैर-नकारात्मक आकार मानकों, a और b के मान में भिन्न है। यूजीन एट अल (2004) और जोन्स (2004) ने अंतराल (0, 1) पर कुमारस्वामी व्यापकीकृत बंटन (Kw-G बंटन) की एक नई श्रेणी का निर्माण किया। दो आकार मापदंडों $a > 0$ और $b > 0$ के साथ प्रायिकता घनत्व (पीडीएफ) और संचयी बंटन फलन को इस तरह परिभाषित किया जा सकता है:

$$f(x) = abx^{a-1}(1-x^a)^{b-1} \text{ and } F(x) = 1 - (1-x^a)^b$$

(1) जहां, $-\infty < x < \infty$, $(x) = dF(x)/dx$ और $a, b > 0$ बंटन F के अतिरिक्त आकार प्राचल है। संबंधित हैजर्ड रेट फंक्शन (hrf) इस प्रकार है:

$$h(x) = \frac{abg(x)G^{a-1}(x)}{1-G^a(x)}$$

2. बंटन का Kw-G बंटन में रूपांतरण

माना मूल सतत बंटन का cdf G(x) और pdf g(x)

है। फिर अंतराल $(0, 1)$ पर क्वांटाइल फंक्शन लागू करके Kw-G बंटन (कॉर्डेइरो और डी कास्ट्रो, 2009) बना सकते हैं। Kw-G के cdf $F(x)$ को इस प्रकार परिभाषित किया गया है

$$F(x) = 1 - \{1 - G(x)^a\}^b \quad (2)$$

जहां $a > 0$ और $b > 0$ दो अतिरिक्त पैरामीटर हैं जिनकी भूमिका विषमता को प्रस्तुत करना और टेल वेइट्स बदलना है। इसी तरह बंटन की इस श्रेणी के प्रायिकता घनत्व का सरल रूप इस प्रकार है:

$$f(x) = abg(x)G(x)^{a-1} \{1 - G(x)\}^{b-1} \quad (3)$$

2.1. कुछ विशेष Kw व्यापकीकृत बंटन

2.2. Kw- प्रसामान्य

$G(\cdot)$ और $g(\cdot)$ को प्रसामान्य बंटन का cdf और pdf मान कर समीकरण (3) से KN प्रायिकता घनत्व प्राप्त किया जाता है, ताकि

$$f(x) = \frac{ab}{\sigma} \phi\left(\frac{x-\mu}{\sigma}\right) \left\{ \Phi\left(\frac{x-\mu}{\sigma}\right) \right\}^{a-1} \left\{ 1 - \Phi\left(\frac{x-\mu}{\sigma}\right) \right\}^{b-1} \quad (4)$$

जहां $x \in \mathbb{R}$, $\mu \in \mathbb{R}$ लोकेशन पैरामीटर है, $\sigma > 0$ एक स्केल पैरामीटर है, $a, b > 0$ आकार पैरामीटर हैं, $\phi(\cdot)$ और $\Phi(\cdot)$ क्रमशः मानक प्रसामान्य बंटन के pdf और cdf हैं। एक यादृच्छिक चर ऊपरोक्त घनत्व $f(x)$ के साथ इस तरह दर्शाया जा सकता है $X \sim Kw-N(a, b, \mu, \sigma^2)$ । $\mu = 0$ और $\sigma = 1$ के लिए मानक Kw-N बंटन प्राप्त कर सकते हैं। इसके अलावा, Kw-N बंटन ($a = 2$ और $b = 1$), विषम प्रसामान्य बंटन (आकार पैरामीटर = 1) के साथ मेल खाता है।

2.3. Kw-वीबुल

वीबुल बंटन का cdf पैरामीटर $\beta > 0$ और $c > 0$ के साथ इस प्रकार है $G(x) = 1 - \exp\{-(\beta x)^c\}$; $x > 0$. इसके अनुरूप, Kw- वीबुल बंटन का घनत्व, $Kw-W(a, b, c, \beta)$, इस प्रकार बदल जाता है

$$f(x) = abc\beta^c x^{c-1} \exp\{-(\beta x)^c\} [1 - \exp\{-(\beta x)^c\}]^{a-1} \times \{1 - [1 - \exp\{-(\beta x)^c\}]^a\}^{b-1} \quad (5)$$

$x, a, b, c, \beta > 0$

यदि $c = 1$ तो Kw-घातांक बंटन प्राप्त होता है। $Kw-W(1, b, 1, \beta)$ बंटन β^b पैरामीटर के साथ घातीय बंटन से मेल खाता है।

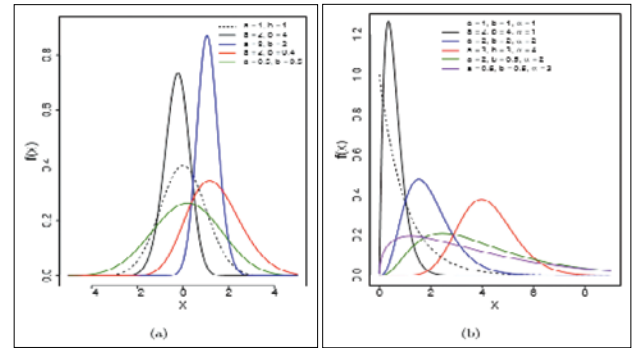
3.3. Kw- गामा

माना Y एक एक गामा यादृच्छिक चर है जिसका cdf $G(y) = \Gamma_{\beta y}(\alpha) / \Gamma(\alpha)$; $y, \alpha, \beta > 0$, है। जहां $\Gamma(\cdot)$ गामा फंक्शन है और $\Gamma_z(\alpha) = \int_z^\infty t^{\alpha-1} e^{-t} dt$ अपूर्ण गामा फंक्शन है। Kw-Ga बंटन वाले एक यादृच्छिक चर $X \sim Kw-Ga(a, b, \beta, \alpha)$ के घनत्व को इस रूप में व्यक्त किया जा सकता है

$$f(x) = \frac{ab\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)^{ab}} \Gamma_{\beta x}(\alpha)^{a-1} \{\Gamma(\alpha)^a - \Gamma_{\beta x}(\alpha)\}^{b-1} \quad (6)$$

$x, a, \beta, \alpha, b > 0$

जब $\alpha = 1$ है, तो Kw-घातांक बंटन प्राप्त होता है।



चित्र 1: Kw-G बंटन के घनत्व फलन के कुछ संभावित आकार (a) Kw-प्रसामान्य ($a, b, 0, 1$) और (b) Kw-गामा ($a, b, 1, \alpha$) घनत्व फलन (धराशायी रेखाएं मूल बंटन को दर्शाती हैं)

4. घनत्व फलन का एक व्यापक विस्तार

कॉर्डेइरो और डी कास्ट्रो (2009) ने इस बंटन के व्यापक विस्तार को विस्तृत किया।

For $b > 0$ वास्तविक गैर-पूर्णांक के लिए, बंटन का रूप

$$\{1 - G(x)^a\}^{b-1} = \sum_{i=0}^{\infty} (-1)^i \binom{b-1}{i} G(x)^{ai} \quad (7)$$

उपरोक्त विस्तार और सूत्र (3) से, हम Kw-G घनत्व को इस प्रकार लिख सकते हैं

$$f(x) = g(x) \sum_{i=0}^{\infty} w_i G(x)^{a(i+1)-1} \quad (8)$$

जहाँ $w_i = w_i(a, b) = (-1)^i ab \binom{b-1}{i}$ and $\sum_{i=0}^{\infty} w_i = 0$

5. आघूर्ण के लिए सामान्य सूत्र

पूर्णाक और वास्तविक गैर-पूर्णाक के लिए, Kw-G बंटन के s -वें आघूर्ण को समीकरण (8) से मूल बंटन G के क्रम (s, r) के PWM के अनंत भारित योग के रूप में व्यक्त किया जा सकता है। हम मानते हैं कि Y और X क्रमशः बेसलाइन G और Kw-G वितरण का अनुसरण करते हैं।

माना X का s -th आघूर्ण μ'_s है जिसको को (s, r) -वें PWM के रूप में व्यक्त किया जा सकता है Y के लिए $\tau_{sr} = E\{Y^s G(Y)^r\}$, $r = 0, 1, \dots$ के लिए, जैसा कि ग्रीनवुड एट अल (1979) द्वारा परिभाषित किया गया है। a पूर्णाक के लिए,

$$\mu'_s = \sum_{r=0}^{\infty} w_r \tau_{s, a(r+1)-1} \quad (9)$$

जबकि a वास्तविक गैर पूर्णाक के लिए सूत्र

$$\mu'_s = \sum_{i,j=0}^{\infty} \sum_{r=0}^{\infty} w_{i,j,r} \tau_{s,r} \quad (10)$$

हम G बंटन के PWM के अनंत भारित योग के रूप में Kw-G बंटन के आघूर्ण की गणना कर सकते हैं।

6. प्रायिकता भारित आघूर्ण

PWM के सामान्य सिद्धांत में सैद्धांतिक प्रायिकता बंटन का सारांश और विवरण, अवलोकन किए गए डेटा नमूनों का सारांश और विवरण, अवलोकित नमूने के अंतर्निहित बंटन का नॉन-पैरामीट्रिक अनुमान, प्रायिकता बंटन के मापदंडों अनुमान और प्रायिकता बंटन के लिए परिकल्पना परीक्षण शामिल हैं। (बराकत और अब्देलकादर, 2004)

Kw-G बंटन वाले X का (s, r) -th PWM को इस तरह परिभाषित किया जा सकता है

$$\tau_{s,r}^{Kw} = E\{X^s F(X)^r\} = \int_{-\infty}^{\infty} x^s F(x)^r f(x) dx \quad (11)$$

इस सूत्र को निम्न रूप में भी लिखा जा सकता है:

$$\tau_{s,r}^{Kw} = \sum_{m,u,v=0}^{\infty} \sum_{l=0}^v p_{r,m}(a,b) w_{u,v,l} \tau_{s,m+l} \quad (12)$$

जहाँ $\tau_{s,m+l}$ G बंटन का $(s, m+l)$ -th PMW और

$$p_{r,m}(a,b) = \sum_{k=0}^u \binom{u}{k} (-1)^k \sum_{m=0}^{\infty} \sum_{l=r}^{\infty} (-1)^{m+l} \binom{kb}{m} \binom{ma}{l} \binom{l}{r}$$

7. क्रम-प्रतिदर्शज

i -वें क्रम-प्रतिदर्शज का घनत्व $f_{in}(x)$ इस प्रकार है; $i = 1, \dots, n$ । यादृच्छिक X_1, \dots, X_n i.i.d वितरित है और Kw-G बंटन का अनुसरण करते हैं।

$$\begin{aligned} f_{in}(x) &= \frac{f(x)}{B(i, n-i+1)} F(x)^{i-1} \{1-F(x)\}^{n-i} \\ &= \frac{ab}{B(i, n-i+1)} g(x) G(x)^{i-1} [1 - \{1-G(x)^a\}^b] \\ &\quad \{1-G(x)^a\}^{b(n-i)-1} \end{aligned} \quad (13)$$

$$f_{in}(x) = \frac{f(x)}{B(i, n-i+1)} \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} F(x)^{i+j-1} \quad (14)$$

सभी पदों का विस्तार करने पर हमें निम्नलिखित दो रूप प्राप्त होते हैं:

$$\begin{aligned} \text{जब } a = \text{ गैर पूर्णाक} \\ f_{in}(x) &= \frac{g(x)}{B(i, n-i+1)} \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} \\ &\quad \sum_{r,u,v=0}^{\infty} \sum_{t=0}^v w_{u,v,t} p_{r,i+j-1}(a,b) G(x)^{r+t} \end{aligned} \quad (15)$$

$$\begin{aligned} \text{और जब } a = \text{ पूर्णाक} \\ f_{in}(x) &= \frac{g(x)}{B(i, n-i+1)} \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} \\ &\quad \sum_{r,u=0}^{\infty} w_u p_{r,i+j-1} ab G(x)^{a(u+1)+r-1} \end{aligned} \quad (16)$$

सूत्र (15) और (16) Kw-G बंटन के क्रम-प्रतिदर्शज के घनत्व को आधारभूत बंटन के घनत्व फलन और G(x) की घात के अनंत भारित योग के गुणांक के रूप में प्रदान करते हैं।

8. L आघूर्ण

L-आघूर्ण सामान्य आघूर्ण के अनुरूप होते हैं जिसका क्रम प्रतिदर्शज के रैखिक संयोजनों द्वारा अनुमान लगाया जा सकता है। L-आघूर्ण अपेक्षित क्रम प्रतिदर्शज के रैखिक फलन हैं जिन्हें इस तरह परिभाषित किया गया है

$$\lambda_{r+1} = (r+1)^{-1} \sum_{k=0}^r (-1)^k \binom{r}{k} E(X_{r+1-k:r+1}) \quad (17)$$

पहले चार L-आघूर्ण हैं:

$$\lambda_1 = E(X_{1:1}), \quad \lambda_2 = \frac{1}{2} E(X_{2:2} - X_{1:2}),$$

$$\lambda_3 = \frac{1}{3} E(X_{3:3} - 2X_{2:3} + X_{1:3})$$

$$\lambda_4 = \frac{1}{4} E(X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4})$$

L-आघूर्ण की गणना (12) में दिए गए सूत्र के अनुसार पीडब्लूएम के रूप में भी की जा सकती है:

$$\lambda_{r+1} = \sum_{k=0}^{\infty} (-1)^{r-k} \binom{r}{k} \binom{r+k}{k} \tau_{1,k}^{Kw} \quad (18)$$

विशेष रूप से,

$$\lambda_1 = \tau_{1,0}^{Kw}, \lambda_2 = 2\tau_{1,1}^{Kw} - \tau_{1,0}^{Kw}, \lambda_3 = 6\tau_{1,2}^{Kw} - 6\tau_{1,1}^{Kw} + \tau_{1,0}^{Kw}$$

9. माध्य विचलन

माना $X \sim Kw-G(a, b)$, माध्य से माध्य विचलन ($\delta_1(X)$) और माध्यिका से माध्य विचलन ($\delta_2(X)$) को इस रूप में व्यक्त किया जा सकता है

$$\delta_1(X) = E(|X - \mu'_1|) = 2\mu'_1 F(\mu'_1) - 2T(\mu'_1)$$

और

$$\delta_2(X) = E(|X - M|) = \mu'_1 - 2T(M)$$

जहाँ $\mu'_1 = E(X)$, $M =$ माध्यिका, $F(\mu'_1)$ pdf से निकाला जा सकता है और $T(z) = \int_{-\infty}^z xf(x)dx$

10. प्राचल का आकलन

मान लीजिए समीकरणों (2) और (3) में γ आधारभूत बंटन का पी-आयामी पैरामीटर वेक्टर है। माना स्वतंत्र यादृच्छिक चर X_1, \dots, X_n में प्रत्येक X_i पैरामीटर वेक्टर $\theta = (a, b, \gamma)$ के साथ Kw-G बंटन का अनुसरण करता है। (3) से प्राप्त मॉडल मापदंडों के लिए लॉग-

लाक्लिहुड $\ell = \ell(\theta)$ फंक्शन इस प्रकार है

$$\ell(\theta) = n\{\log(a) + \log(b)\} + \sum_{i=1}^n \log\{g(x_i; \gamma)\} + (a-1)$$

$$\sum_{i=1}^n \log\{G(x_i; \gamma)\} + (b-1) \sum_{i=1}^n \log\{1 - G(x_i; \gamma)^a\}$$

स्कोर वेक्टर के तत्व इस प्रकार हैं

$$\frac{d\ell(\theta)}{da} = \frac{n}{a} + \sum_{i=1}^n \log\{G(x_i; \gamma)\} \left\{1 - \frac{(b-1)G(x_i; \gamma)^a}{1 - G(x_i; \gamma)^a}\right\}$$

$$\frac{d\ell(\theta)}{db} = \frac{n}{b} + \sum_{i=1}^n \log\{1 - G(x_i; \gamma)^a\}$$

$$\text{और } \frac{d\ell(\theta)}{d\gamma_j} = \sum_{i=1}^n \left[\frac{1}{g(x_i; \gamma)} \frac{dg(x_i; \gamma)}{d\gamma} + \frac{1}{G(x_i; \gamma)} \right]$$

$$\frac{dG(x_i; \gamma)}{d\gamma} \left\{1 - \frac{a(b-1)}{G(x_i; \gamma)^a - 1}\right\}$$

उपरोक्त लॉग- लाक्लिहुड का संख्यात्मक अधिकतमकरण आर में gamlss पैकेज में उपलब्ध आरएस विधि (रिग्बी और स्टैसिनोपोलोस, 2005) का उपयोग करके किया गया है। चूंकि सामान्यीकृत बीटा बंटन के प्राचल के आकलन की तुलना में Kw-G बंटन के मापदंडों की अधिकतम संभावित आकलन अधिक सरल है इसलिए बंटन के दूसरे श्रेणी के स्थान पर Kw-G बंटन का उपयोग करने की सलाह दे सकते हैं।

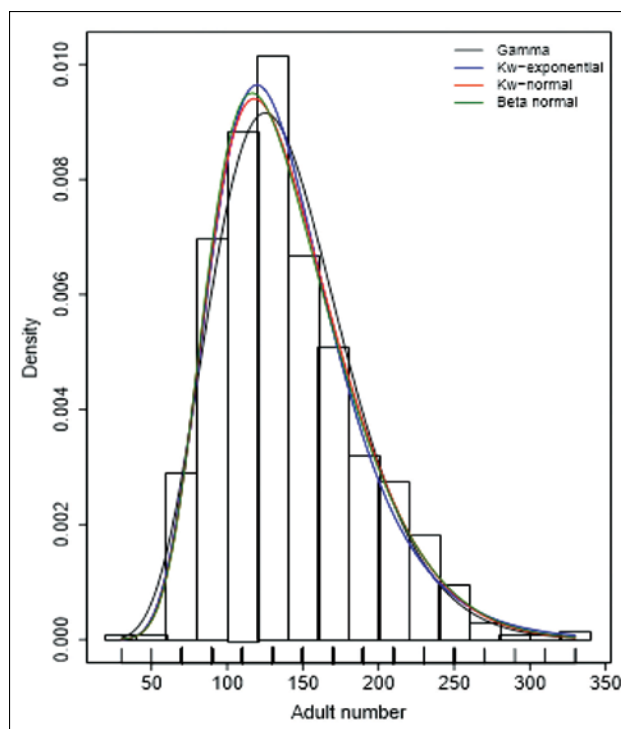
11. सेंसर्ड डेटा में अनुप्रयोग

इस अनुभाग में कॉर्डेइरो और डी कास्ट्रो (2009) द्वारा प्रस्तुत 29 डिग्री सेल्सियस पर कल्चर्ड फलौरी बीटल (टी कन्फ्यूसम) की वयस्क संख्या के डेटा के साथ एक उदाहरण प्रस्तुत किया गया है।

AIC के अनुसार, बीटा प्रसामान्य और Kw-प्रसामान्य बंटन बाकी चुने हुए बंटन से बेहतर प्रदर्शन कर रहे हैं। गामा बंटन के लिए चित्र 3 में फीटेड बंटन को आयतचित्र के साथ आरोपित किया गया है जो तालिका 1 में दिए गए परिणाम को सुदृढ़ करते हैं। बीटा प्रसामान्य और Kw-प्रसामान्य बंटन लगभग अप्रभेद्य हैं। तालिका 2 में प्रेक्षित और अपेक्षित आवृत्तियों के बीच तुलना इस दावे को और मजबूत बनती है। अपेक्षित और प्रेक्षित आवृत्तियों के बीच औसत

तालिका 1:

चर का वितरण	AIC	पैरामीटर			
		a	b		
Beta normal	7176.91	18.20 (12.70)	0.25 (0.07)	$\mu=12.69$ (23.40)	$\sigma=42.76$ (8.89)
Kw-normal	7177.41	14.86 (1.74)	0.27 (0.05)	$\mu=25.52$ (0.59)	$\sigma=42.24$ (3.89)
Kw-exponential	7181.31	15.54 (3.61)	1.34 (0.33)	$B=46.22$ (6.98)	0.18
Kw-gamma	7180.92	1.85 (1.36)	0.67 (0.48)	$\alpha= 7.37$ (6.31)	$\beta= 14.21$ (11.70)
Gamma	7183.96			$\alpha= 8.99$ (0.47)	$\beta= 15.56$ (0.84)



चित्र 3. वयस्क संख्या का आयतचित्र और फिटिड प्रायिकता घनत्व फलन।

निरपेक्ष विचलन Kw-प्रसामान्य बंटन के लिए न्यूनतम पाया गया है।

LR प्रतिदर्शज (अनुभाग 8) के मान के आधार पर, Kw-गामा और Kw-घातीय बंटन (LR = 1.542 (1 d.f., p = 0.214)) ज्यादा भिन्न नहीं हैं। यदि Kw-गामा और गामा बंटन की तुलना की जाये तो एक महत्वपूर्ण अंतर देखा जा सकता है (LR = 6.681, 2 d.f., p = 0.035)।

12. सार

कुमारस्वामी बंटन और बीटा सामान्यीकृत बंटन की श्रेणियों पर विचार करने के बाद, हमने प्रसामान्य, वीबुल, गामा और गंबेल बंटन जैसे कई व्यापक रूप से ज्ञात बंटन का विस्तार करने के लिए Kw सामान्यीकृत (Kw-G) बंटन की एक नए श्रेणी को परिभाषित किया है। प्राचल आकलन के लिए अधिकतम संभावित आकलक पर चर्चा की है। Kw-G बंटन में अधिकतम संभावित आकलक बीटा सामान्यीकृत बंटन में अनुमान

तालिका 2: वयस्क संख्याओं की प्रेक्षित और अपेक्षित आवृत्तियां और आवृत्तियों के बीच माध्य निरपेक्ष विचलन (MAD)

व्यस्क संख्या	असल	अपेक्षित			
		Gamma	Kw-exponential	Kw-normal	Beta normal
30	1	0.75	0.19	0.21	0.22
50	1	9.85	6.32	5.77	5.67
70	40	39.72	37.82	37.43	37.69
90	96	83.59	90.75	91.85	93.70
110	122	117.15	127.14	125.53	127.54
130	140	124.56	127.81	123.95	123.73

व्यस्क संख्या	अपेक्षित				
	असल	Gamma	Kw-exponential	Kw-normal	Beta normal
150	92	108.51	104.22	102.41	100.77
170	70	81.35	74.35	75.95	74.30
190	44	54.24	48.68	52.00	50.97
210	38	32.93	30.16	33.18	32.80
230	25	18.51	18.03	19.80	19.86
250	13	9.76	10.53	11.06	11.32
270	4	4.87	6.06	5.79	6.08
290	1	2.32	3.46	2.84	3.08
310	1	1.06	1.96	1.31	1.47
330	2	0.47	1.10	0.56	0.66
Total	690	689.70	688.60	689.61	689.52
MAD		6.17	4.74	4.60	4.39

की तुलना में बहुत सरल है। वास्तविक डेटा पर इस अध्ययन का प्रयोग भी किया गया है।

सन्दर्भ

- बराकत, एच. एम. एवं अब्देलकडेर, वाई. एच. (2004). कंप्यूटिंग द मोमेंट्स ऑफ आर्डर स्टेटिस्टिक्स फ्रॉम नॉन आईडेंटिकल रैंडम वेरिएबल्स, स्टेटिस्टिकल मेथड्स एंड एप्लिकेशन्स 13:15–26
- कोरडीरोआ, गॉस एम. एवं कास्ट्रोब, मारिओ, डी. (2009). ए न्यू फॅमिली ऑफ जेनेरलाइज्ड डिस्ट्रीब्यूशनस, जर्नल ऑफ स्टेटिस्टिकल कम्प्यूटेशन एंड सिमुलेशन, 1–17
- यूजीने, ऍन., ली., सी. एवं फामोये, (2002). बीटा नार्मल डिस्ट्रीब्यूशन एंड इट्स ऍप्लिकेशन्स, कम्प्युनिकेशन्स इन स्टेटिस्टिक्स: थेओरी एंड मेथड्स 31:497–512.

- जॉस, एम. सी. (2004). फॅमिलीज ऑफ डिस्ट्रीब्यूशनस अराइजिंग फ्रॉम डिस्ट्रीब्यूशन ऑफ आर्डर स्टेटिस्टिक्स (विथ डीसक्शन) टेस्ट 13:1–43
- कुमारस्वामी, पी. (1980). जेनेरलाइज्ड प्रोबेबिलिटी डेंसिटी फंक्शन फॉर डबल बॉण्डेड रैंडम प्रोसेसेज, जर्नल ऑफ हाइड्रोलॉजी 462:79–88.
- नादराज, एस. एवं गुप्ता, ए. के. (2004). द बीटा फ्रेंचेट डिस्ट्रीब्यूशन, फार ईस्ट जर्नल ऑफ थ्योरेटिकल स्टेटिस्टिक्स 14:15–24
- नादराज, एस. एवं कोट्ज, एस. (2006). द बीटा एक्सपोनेंशियल डिस्ट्रीब्यूशन, रिलायबिलिटी इंजीनियरिंग एंड सिस्टम सेफ्टी 91: 689–697
- रिग्बी आर. ए. एवं स्टासिनोपोइओस डी. एम. (2005). जेनेरलाइज्ड ऐडिटीव मॉडल्स फॉर लोकेशन, स्केल एंड शेप, एप्लाइड स्टेटिस्टिक्स 54:507–554.

“ अच्छा स्वभाव सुंदरता के अभाव को पूरा कर देता है,
लेकिन सुंदरता अच्छे स्वभाव के अभाव की पूर्ति नहीं कर सकती। ”

दीर्घकाल अनुस्मरण समय श्रृंखला मॉडल द्वारा भारत में चावल की कीमतों का पूर्वानुमान

रंजीत कुमार पॉल, अमृत कुमार पॉल, दीपांकर मित्रा एवं सत्य पाल सिंह
भा.कृ.अनु.प.—भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली—110012

1 परिचय

वर्तमान प्रतिस्पर्धी और अस्थिर बाजार में मूल्य पूर्वानुमान एक आवश्यक विश्लेषणात्मक तकनीक है। ऐसे कई क्षेत्र हैं, जहां मूल्य पूर्वानुमान एक महत्वपूर्ण भूमिका निभाता है जैसे कि इष्टतमीकरण उद्देश्यों के लिए खुदरा और थोक मूल्य पूर्वानुमान, दूरसंचार ध्वनि-मार्ग मूल्य पूर्वानुमान, वर्तमान एवं आगे की कीमतों की पूर्वानुमान के लिए बिजली की कीमत का पूर्वानुमान, आवश्यक कृषि वस्तुओं का मूल्य पूर्वानुमान और कई अन्य। चूंकि मूल्य पूर्वानुमान पर कई व्यावसायिक निर्णय और सरकारी नीतियां बनाई जाती हैं, इसलिए इसे सटीक रूप से प्राप्त करना महत्वपूर्ण है। पिछले कुछ दशकों में, बॉक्स जीईपी, जेनकिंस का ऑटोरेग्रेसिव मूविंग एवरेज मेथडोलॉजी (अरिमा) का पूर्वानुमान प्राप्त करने के लिए कुशलतापूर्वक उपयोग किया जाता है। हालांकि पारंपरिक स्थिर ऑटोरेग्रेसिव मूविंग एवरेज (एआरएमए) प्रक्रियाएं अक्सर वित्तीय समय-श्रृंखला डेटा में उच्च कोटि की दृढ़ता पर कब्जा नहीं कर सकती हैं, गैर-स्थिर इकाई रूट या $I(1)$ प्रक्रियाओं के वर्ग में वित्तीय डेटा के लिए कुछ अप्राप्य गुण हैं। पिछले तीस वर्षों में, दीर्घकाल अनुस्मरण प्रक्रियाओं का उपयोग करके अधिक शोध कार्य किए गए हैं, जो पारंपरिक स्थिर $I(0)$ प्रक्रियाओं और गैर-स्थिर $I(1)$ प्रक्रियाओं के बीच आधे रास्ते में हैं। समग्र श्रृंखला में अविभाज्य दीर्घकाल अनुस्मरण मॉडल हो सकते हैं और एकीकृत, या अनंत लंबाई स्थानांतरण फंक्शन संबंध का पालन कर सकते हैं। लंबी दूरी की दृढ़ता या लंबे समय तक निर्भरता दिखाने वाले समय-श्रृंखला डेटा

को मॉडल करने के लिए आमतौर पर ऑटोरेग्रेसिव फ्रैक्शनल इंटीग्रेटेड मूविंग एवरेज (अरफिमा या फरिमा) मॉडल का उपयोग किया जाता है। अरफिमा मॉडल की मदद से दीर्घकाल अनुस्मरण संरचना को प्रदर्शित करने वाली एक समय-श्रृंखला प्रक्रिया को अच्छी तरह से तैयार किया जा सकता है। कृषि सामग्री कीमतों के पूर्वानुमान के लिए अरफिमा मॉडल लागू किया गया है। मॉडल के अच्छे प्रदर्शन ने परिवर्तनशीलता की व्याख्या और पूर्वानुमान के प्रदर्शन की शर्तों में प्रदर्शन किया है।

भारत सफेद चावल और भूरे चावल के लिए दुनिया के सबसे बड़े उत्पादकों में से एक है, जो विश्व के चावल उत्पादन का 20% हिस्सा है। भारत में चावल एक प्रमुख फसल है और इसकी खेती के तहत सबसे बड़ा क्षेत्र है। भारत के सभी भागों में चावल की खेती की जाती है। भारत के प्रमुख चावल उत्पादक राज्य हैं— पश्चिम बंगाल, ओडिशा, उत्तर प्रदेश, आंध्र प्रदेश, बिहार, असम, तमिलनाडु। यह भारत की कुल आबादी का लगभग 65% (मुख्य रूप से देश के पूर्वी और दक्षिणी भागों में) का मुख्य भोजन है। इसलिए, उपयुक्त सांख्यिकीय मॉडलों का उपयोग करके चावल की कीमत का पूर्वानुमान लगाने की आवश्यकता है। वर्तमान अध्ययन में चावल के अखिल भारतीय दैनिक अधिकतम, न्यूनतम और मोडल मूल्य के पूर्वानुमान मूल्य के लिए अरफिमा मॉडल की जांच करने का प्रयास किया गया है। फिटेड अरफिमा मॉडल की पूर्वानुमान क्षमता की तुलना अरिमा मॉडल से भी की जाती है।

2. कार्यप्रणाली

2.1 लघु अनुस्मरण प्रक्रिया

एक समय-श्रृंखला $\{(Y_t)_t\}$ को सख्ती से स्थिर कहा जाता है, यदि किसी $t_1, t_2, \dots, t_n \in \mathbb{Z}$, किसी $k \in \mathbb{Z}$ और $n = 1, 2, \dots$ $F_{Y_{t_1}, Y_{t_2}, \dots, Y_{t_n}}(y_1, y_2, \dots, y_n) = F_{Y_{t_1+k}, Y_{t_2+k}, \dots, Y_{t_n+k}}(y_1, y_2, \dots, y_n)$ के लिए (1)

जहाँ, $F_X(x)$ वितरण फलन को बिंदु x पर यादृच्छिक चर X दर्शाता है। एक सख्ती से स्थिर समय-श्रृंखला प्रक्रिया $\{(Y_t)_t\}$ को शॉर्ट मेमोरी (या शॉर्ट-रेंज निर्भरता) प्रक्रिया कहा जाता है यदि ρ_k संतुष्ट करता है $\sum_{k=0}^{\infty} |\rho_k| < \infty$, जहाँ $\rho_k = cov(y_t, y_{t-1})/var(y_t)$ k के समय अंतराल के साथ समय-श्रृंखला का स्वतःसहसंबंध कार्य है, जैसा कि दिया गया है

$$\rho_k = cov(y_t, y_{t-1})/var(y_t) \quad (2)$$

एक स्थिर लघु स्मृति प्रक्रिया के लिए ऑटोसहसंबंध फलन ρ_k भी निम्नलिखित शर्त को पूरा करता है

$$\lim_{k \rightarrow \infty} \rho_k = 0 \quad (3)$$

अधिकांश स्थिर और उलटा (एआरएमए) समय-श्रृंखला प्रक्रिया के स्वतः सहसंबंध कार्य एक घातीय दर पर बहुत तेजी से क्षय हो जाते हैं, ताकि $\rho_k \approx |m|^k$, where $|m| < 1$

2.2 अरिमा मॉडल

अरिमा मॉडल का उपयोग ज्यादातर इनवर्टिबल और स्थिर शॉर्ट मेमोरी प्रक्रिया के रूप में किया जाता है। अरफिमा मॉडल में, यह माना जाता है कि एक चर का भविष्य मूल्य चर के पिछले मूल्यों का एक रैखिक फलन है और यादृच्छिक त्रुटियाँ भी हैं। यह एक रेखीय अविभाज्य समय-श्रृंखला मॉडल है जो एक समय-श्रृंखला प्रक्रिया को व्यक्त करता है, मान लीजिए, $\{y_t\}, t = 1, 2, \dots, n$ मापदंडों के तीन सेटों के रूप में

$$\varphi(B)(1-B)^d y_t = \theta(B)u_t \quad (4)$$

जहाँ y_t और u_t क्रमशः t समय पर वास्तविक मान और यादृच्छिक त्रुटि हैं; $\varphi(B)$ और $\theta(B)$ क्रमशः p और q के लैग ऑपरेटर B के ऑटोरेग्रेसिव (एआर) और

मूविंग एवरेज (एमए) बहुपद हैं, यूनिट सर्कल के बाहर रूट के साथ; यादृच्छिक त्रुटियाँ, u_t को स्वतंत्र माना जाता है और समान रूप से माध्य शून्य और प्रसरण σ^2 के साथ वितरित किया जाता है। अरिमा मॉडलिंग की बॉक्स-जेनकिस पद्धति में तीन पुनरावृत्त चरण शामिल हैं, अर्थात् मॉडल पहचान, पैरामीटर आकलन और निदानार्थ जांच।

2.3 दीर्घकाल अनुस्मरण प्रक्रिया

समय-श्रृंखला विश्लेषण में अधिकांश शोध कार्य यह मानते हैं कि दीर्घकाल समय के अंतराल से अलग किए गए अवलोकन एक दूसरे से स्वतंत्र हैं या लगभग ऐसे ही हैं। लेकिन कई प्रायोगिक स्थितियों में यह देखा जाता है कई प्रयोगसिद्ध आर्थिक श्रृंखला का पता चलता है कि पृथक अवलोकन, निर्भर कर रहे हैं, हालांकि सहसंबंध छोटा है लेकिन नगण्य नहीं है। दीर्घकाल अनुस्मरण प्रक्रियाओं के लिए, स्वतः सहसंबंध कार्यों का क्षय बहुत धीमी दर (हाइपरबोलिक दर) पर होता है जो $\rho_k \approx Ck^{2d-1}$, के अनुरूप होता है, क्योंकि k अनिश्चित काल तक बढ़ता है, जहाँ C स्थिर है और d दीर्घकाल अनुस्मरण पैरामीटर है। एक स्थिर समय-श्रृंखला प्रक्रिया $\{(Y_t)_t\}$ को दीर्घकाल अनुस्मरण (या लंबी दूरी की निर्भरता) प्रक्रिया कहा जाता है यदि $\rho_k, \sum_{k=0}^{\infty} |\rho_k| = \infty$ संतुष्ट करता है।

2.4 अरफिमा मॉडल

ऑटोरेग्रेसिव फ्रैक्शनली इंटीग्रेटेड मूविंग-एवरेज अरफिमा मॉडल का उपयोग दीर्घकाल अनुस्मरण की उपस्थिति में समय-श्रृंखला मॉडलिंग के लिए किया जाता है। भिन्नात्मक एकीकरण पूर्णांक एकीकरण का एक सामान्यीकरण है जो दीर्घकाल अनुस्मरण पैरामीटर को -1 से 1 के बीच कोई भी भिन्नात्मक मान लेने की अनुमति देता है। यहाँ, समय-श्रृंखला को आमतौर पर ऑर्डर शून्य या एक के एकीकृत d माना जाता है। उदाहरण के लिए, एक ऑटोरेग्रेसिव चल-औसत प्रक्रिया आदेश d की एकीकृत (अरफिमा (p, d, q) निरूपित किया जाता है और a से दर्शाया जा सकता है

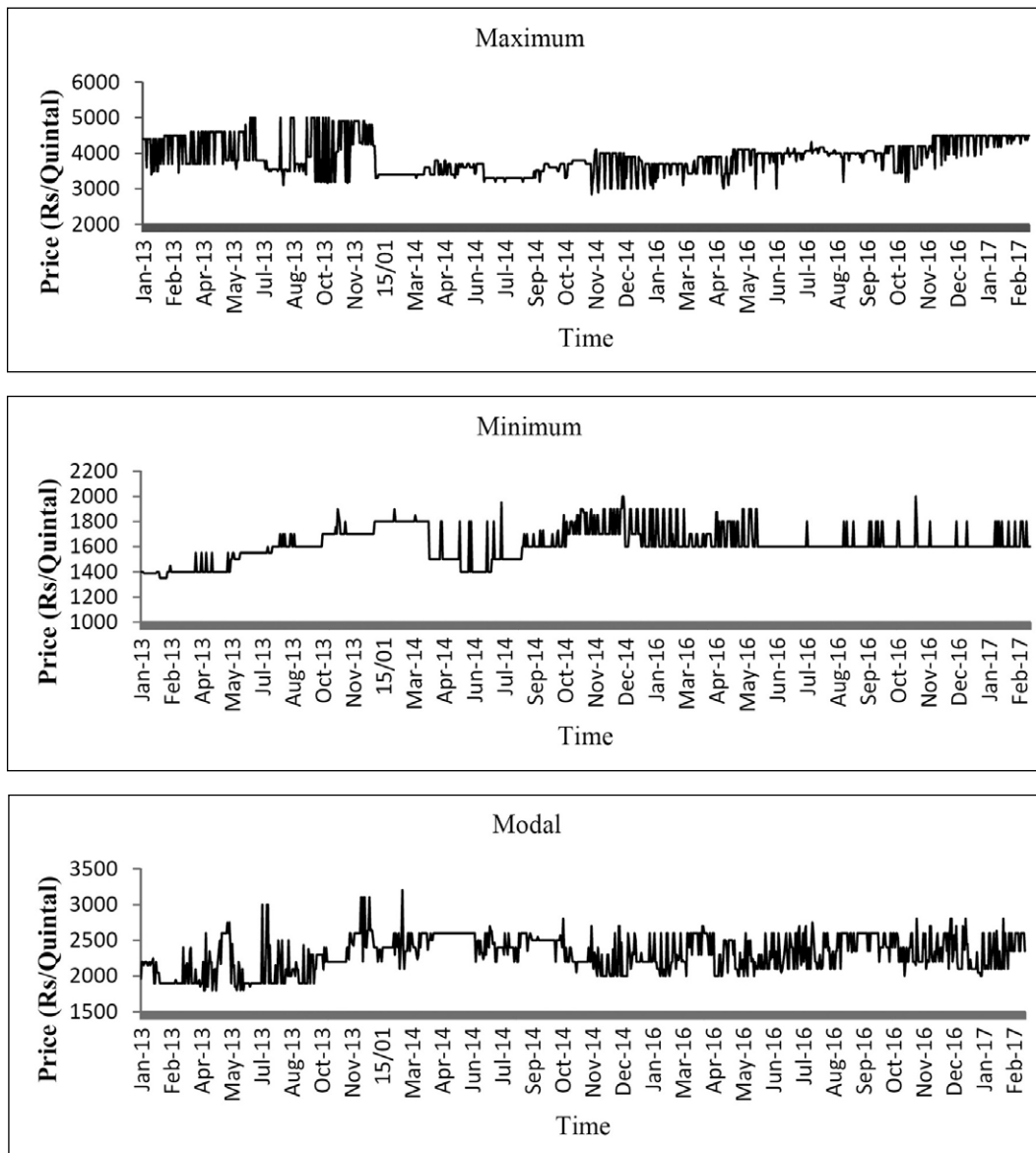
$$\varphi(B)y_t = (1 - B)^{-d}\theta(B)u_t$$

जहाँ u_t , $\phi(B)$ और $\theta(B)$ ऊपर वर्णित हैं। $d = 0$, के लिए, प्रक्रिया स्थिर है, $-0.5 < d < 0.5$ के लिए प्रक्रिया y_t स्थिर और उलटा है $d \in (0, 1/2)$ के लिए प्रक्रिया को दीर्घकाल अनुस्मरण कहा जाता है।

2.4.1 दीर्घकाल अनुस्मरण पैरामीटर का परीक्षण

दीर्घकाल अवधि की दृढ़ता या लंबी दूरी की निर्भरता की संभावित उपस्थिति का परीक्षण एक आवश्यक कार्य है क्योंकि दीर्घकाल अनुस्मरण का कोई सबूत

अरफिमा जैसे दीर्घकाल अनुस्मरण मॉडल के उपयोग का समर्थन करेगा। समय-श्रृंखला साहित्य में दीर्घकाल अनुस्मरण पैरामीटर d का अनुमान लगाने के लिए कई विधियां हैं। आमतौर पर इस्तेमाल की जाने वाली कुछ विधियों में अनुमान की अधिकतम संभावना विधि (एमएलई), स्थानीय व्हिटल (एलडब्ल्यू) अनुमान, जीपीएच, तरंगिका पद्धति आदि शामिल हैं। वर्तमान अध्ययन में जीपीएच परीक्षण का उपयोग किया जाता है जो मूल मूल्य डेटा में दीर्घकाल अनुस्मरण पैरामीटर का परीक्षण करता है।



चित्र 1: श्रृंखला का समय प्लॉट

3. परिणाम और चर्चा

3.1 डेटा विवरण

वर्तमान अध्ययन के लिए चावल का अखिल भारतीय अधिकतम, न्यूनतम और मोडल मूल्य जनवरी, 2013 से फरवरी, 2017 की अवधि के लिए उपभोक्ता मामलों के निदेशालय (consumeraffairs-nic-in वेबसाइट) से एकत्र किया जाता है। डेटासेट कुल 941 टिप्पणियों से बना है। पहले 891 प्रेक्षणों का उपयोग मॉडलिंग भवन के उद्देश्य के लिए किया जाता है और अंतिम 50 अवलोकनों का उपयोग मॉडल सत्यापन उद्देश्य के लिए किया जाता है।

समय-श्रृंखला विश्लेषण में पहला और सबसे महत्वपूर्ण कदम डेटासेट को प्लॉट करना और कई समय-श्रृंखला घटकों और मौजूदा पैटर्न की उपस्थिति की कल्पना करना है। विचाराधीन श्रृंखला के समय भूखंड चित्र 1 में दिए गए हैं। आकृति 1 का अवलोकन यह दर्शाता है कि विचाराधीन श्रृंखला के स्थिर होने की संभावना है। स्थिरता की पुष्टि करने के लिए संवर्धित डिकी-फुलर (एडीएफ) परीक्षण और फिलिप्स-पेरॉन (पीपी) परीक्षण किया गया है और यह पाया गया है कि परीक्षण श्रृंखला की स्थिरता को इंगित करने वाले 5% स्तर पर महत्वपूर्ण हैं।

3.2 वर्णनात्मक विशेषताएं

अध्ययनाधीन डेटासेट की विशिष्ट विशेषताएं तालिका 1 में बताई गई हैं। तालिका 1 का अवलोकन यह दर्शाता है कि विचाराधीन सभी श्रृंखलाएं सकारात्मक रूप से तिरछी और प्लेटिकर्टिक हैं। सभी श्रृंखलाओं के लिए भिन्नता के गुणांक (CV) के निम्न मान डेटासेट में अस्थिरता या अस्थिरता की अवधि की अनुपस्थिति को दर्शाते हैं। यह भी स्पष्ट है कि मॉडल मूल्य डेटा लगभग सममित है।

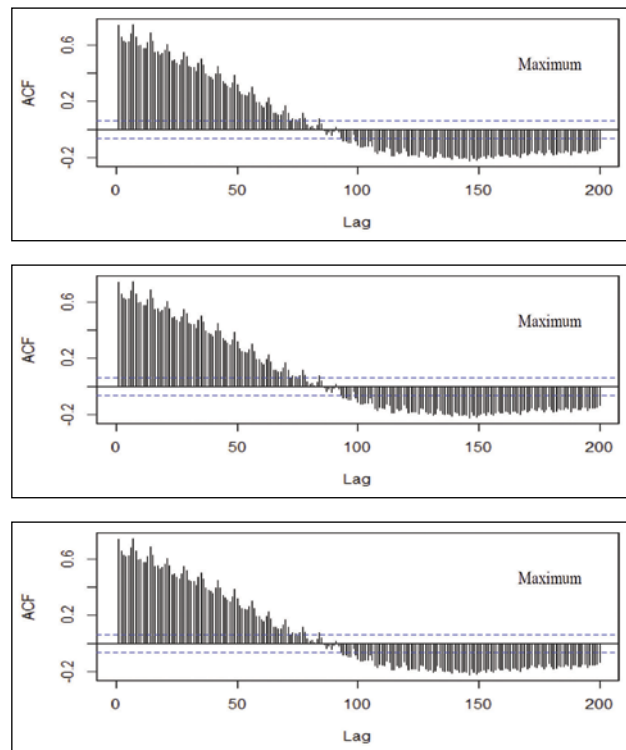
तालिका 1: मूल्य डेटा के वर्णनात्मक आँकड़े

सांख्यिकी	अधिकतम	न्यूनतम	मोडल
प्रेक्षण	941	941	941
माध्य (रु/ क्विंटल)	3886.33	1616.26	2315.72

सांख्यिकी	अधिकतम	न्यूनतम	मोडल
माधिका	3800.00	1600.000	2325.00
न्यूनतम	2840.00	1350.00	1800.00
अधिकतम	5000.00	2000.00	3200.00
रेंज	2160.00	650.00	1400.00
मानक विचलन	407.27	131.49	254.36
भिन्नता का गुणांक (%)	12.10	8.12	10.98
तिरछापन	0.39	0.25	0.03
कुर्टोसिस	-0.53	-0.08	-0.55

3.3 डेटासेट का एसीएफ और पीएसीए प्लॉट

ऑटोसहसंबंध फलन (एसीएफ) की आलेखन अनुमानी दृष्टिकोण एक समय श्रृंखला डेटा में सांख्यिकीय निर्भरता वर्तमान को देखने के लिए में से एक है। डेटासेट के एसीएफ प्लॉट चित्र-2 में दिए गए हैं। चूंकि ऑटोसहसंबंध और आंशिक ऑटोसहसंबंध दूर



चित्र 2: श्रृंखला के एसीएफ आलेखन

के अंतराल (200 अंतराल के बाद भी) पर महत्वपूर्ण हैं, इसलिए मूल्य डेटा में दीर्घकाल अनुस्मरण का स्पष्ट संकेत है। एसीएफ भूखंडों में बिंदीदार रेखाएं परीक्षण आंकड़ों के 95% महत्वपूर्ण मूल्यों का प्रतिनिधित्व करती हैं।

3.4 दीर्घकाल अनुस्मरण के लिए परीक्षण

डेटासेट के एसीएफ और पीएसीएफ आलेखन की कल्पना करने के बाद, दीर्घकाल अनुस्मरण की उपस्थिति का परीक्षण करने के लिए दीर्घकाल अनुस्मरण परीक्षण लागू किया जाता है और परिणाम तालिका 2 में प्रदान किए जाते हैं। तालिका 2 का एक अवलोकन इंगित करता है कि सभी श्रृंखलाओं में महत्वपूर्ण दीर्घकाल अनुस्मरण पैटर्न है। (जैसा कि परिकलित Z आंकड़ा महत्व के 5% स्तर अर्थात् 1.96 पर महत्वपूर्ण मान से अधिक है)।

तालिका 2: अरहर की कीमत के आंकड़ों का दीर्घकाल अनुस्मरण पैरामीटर आकलन

	अधिकतम	न्यूनतम	मॉडल
d	0.476	0.485	0.456
$S.E.$	0.140	0.056	0.057
Z	3.400	8.670	8.000

तालिका 3: सर्वश्रेष्ठ फिटेड अरफिमा मॉडल के पैरामीटर आकलन

श्रृंखला	पैरामीटर आकलन				एआईसी मान
	Intercept	AR(1)	d	MA(1)	
अधिकतम	3859.350 (138.017)	0.976 (0.012)	0.155 (0.0569)	0.892 (0.038)	10485.100
न्यूनतम	1607.260 (95.417)	0.989 (0.006)	0.242 (0.045)	0.924 (0.019)	7825.380
मॉडल	6558.840 (312.136)	0.973 (0.014)	0.174 (0.079)	0.127 (0.073)	8761.290

तालिका 4: सर्वश्रेष्ठ फिटेड अरिमा मॉडल के पैरामीटर आकलन

श्रृंखला	पैरामीटर आकलन			एआईसी मान
	Intercept	AR(1)	MA(1)	
अधिकतम	3881.108 (102.410)	0.977 (0.009)	-0.794 (0.028)	13000.240
न्यूनतम	1602.277 (53.422)	0.991 (0.005)	-0.808 (0.023)	10345.590
मॉडल	6415.620 (324.146)	0.991 (0.004)	-0.075 (0.058)	11146.580

3.5 अरिमा और अरफिमा मॉडल की फिटिंग

अरफिमा मॉडल दीर्घकाल अनुस्मरण पैटर्न को केचर करने के लिए श्रृंखला में फिट किया जाता है और न्यूनतम एआईसी मान के आधार पर सर्वश्रेष्ठ मॉडल का चयन किया जाता है। तदनुसार, अरिमा मॉडल को डेटासेट में भी फिट किया जाता है और न्यूनतम एआईसी मान के आधार पर सर्वश्रेष्ठ मॉडल का चयन किया जाता है। चावल की अधिकतम, न्यूनतम और मॉडल कीमत के लिए सर्वोत्तम फिटेड अरफिमा और अरिमा मॉडल के पैरामीटर आकलन क्रमशः इसकी मानक त्रुटि के साथ तालिका 4 और तालिका 5 में दिए गए हैं। तालिकाओं का अवलोकन इंगित करता है कि पैरामीटर महत्व के कम से कम 5% स्तर पर महत्वपूर्ण हैं।

3.6 पूर्वानुमान मूल्यांकन और मॉडल पर्याप्तता जांच

पिछले 50 अवलोकनों (अर्थात् $h = 50$ के लिए) के मूल्य के एक-चरण आगे के पूर्वानुमान की गणना चयनित अरफिमा और अरिमा मॉडल का उपयोग करके की जाती है। फिट किए गए मॉडलों की सटीकता को निम्नलिखित सूत्रों का उपयोग करते हुए सापेक्ष माध्य निरपेक्ष प्रतिशत त्रुटि (आरएमएपीई), मूल माध्य वर्ग त्रुटि (आरएमएसई) और माध्य निरपेक्ष त्रुटि (एमएई) के रूप में मापा जाता है।

$$\text{RMAPE} = \frac{1}{h} \sum_{t=1}^h |y_t - \hat{y}_t| / y_t \times 100 \quad (6)$$

$$\text{RMSE} = \sqrt{\frac{1}{h} \sum_{t=1}^h (y_t - \hat{y}_t)^2} \quad (7)$$

$$\text{MAE} = \frac{1}{h} \sum_{t=1}^h |y_t - \hat{y}_t| \quad (8)$$

पांच मूविंग विंडोज (10— स्टेप, 20— स्टेप, 30— स्टेप, 40— स्टेप और 50— स्टेप आगे) के लिए पूर्वानुमान मूल्यांकन किया जाता है। प्रत्येक पूर्वानुमान क्षितिज के लिए चावल के अखिल भारतीय अधिकतम, न्यूनतम और मोडल मूल्य के लिए परिकलित आरएमएपीई, आरएमएसई और एमएई मान क्रमशः तालिका 5, तालिका 6 और तालिका 7 में सूचीबद्ध हैं। प्रत्येक

श्रृंखला के लिए, "औसत" के रूप में लेबल की गई तालिका 5, तालिका 6 और तालिका 7 के अंतिम कॉलम क्रमशः सभी श्रृंखलाओं के लिए सभी पूर्वानुमान क्षितिजों में औसत आरएमएपीई, आरएमएसई और एमएई दिखाते हैं। फिट किए गए मॉडलों की पर्याप्तता की जांच फिट किए गए मॉडलों से प्राप्त अवशेषों के एसीएफ और पीएसीएफ की आलेखन की जाती है ताकि यह देखा जा सके कि क्या अवशेषों में कोई व्यवस्थित पैटर्न मौजूद है जिसे आगे समझाया जा सकता है। अवशिष्टों के स्वतःसहसंबंधों की गणना की जाती है और यह पाया जाता है कि लगभग सभी स्वसहसंबंध 95% महत्वपूर्ण मूल्यों के भीतर हैं।

तालिका 5: अधिकतम मूल्य के लिए पूर्वानुमान मूल्यांकन

मूल्यांकन पद्धतियां	पूर्वानुमान क्षितिज (h)					
	10	20	30	40	50	औसत
अरफिमा						
आरएमएपीई	2.613	2.770	2.718	2.367	2.342	2.562
आरएमएसई	133.638	137.829	146.956	129.632	111.909	131.993
एमएई	115.973	122.867	121.856	117.793	104.278	116.553
अरिमा						
आरएमएपीई	4.110	3.860	2.705	3.383	3.042	3.420
आरएमएसई	190.976	196.535	173.203	172.485	154.523	177.544
एमएई	181.340	169.132	164.108	150.409	133.936	159.785

तालिका 6: न्यूनतम मूल्य के लिए पूर्वानुमान मूल्यांकन

मूल्यांकन पद्धतियां	मूल्यांकन के पद्धतियां पूर्वानुमान क्षितिज (एच) (h)					
	10	20	30	40	50	औसत
अरफिमा						
आरएमएपीई	0.240	2.960	3.472	2.647	3.042	2.472
आरएमएसई	4.718	84.301	60.348	65.310	59.960	54.927
एमएई	7.000	51.718	56.820	44.914	51.460	42.382
अरिमा						
आरएमएपीई	0.700	3.353	5.327	3.663	4.158	3.440
आरएमएसई	15.462	109.610	89.937	79.389	82.434	75.366
एमएई	11.202	60.317	87.266	61.991	70.615	58.278

तालिका 7: मॉडल मूल्य के लिए पूर्वानुमान मूल्यांकन

मूल्यांकन पद्यतियां	मूल्यांकन के पद्यतियां पूर्वानुमान क्षितिज (एच) (h)					
	10	20	30	40	50	औसत
अरफिमा						
आरएमएपीई	8.023	7.028	8.325	4.644	4.695	6.543
आरएमएसई	216.00	188.596	229.644	155.298	130.136	183.935
एमएई	181.313	167.693	195.733	116.173	120.225	156.227
अरिमा						
आरएमएपीई	7.216	8.123	9.402	5.826	5.442	7.202
आरएमएसई	218.000	250.375	248.886	187.500	144.266	209.805
एमएई	165.884	198.742	218.633	144.559	138.874	173.338

4 निष्कर्ष

दीर्घकाल अनुस्मरण परीक्षण के महत्वपूर्ण परिणाम मूल्य श्रृंखला में दीर्घकाल अनुस्मरण के अस्तित्व का संकेत देते हैं। अरिमा और अरफिमा मॉडल प्रत्येक मूल्य श्रृंखला में फिट होते हैं और न्यूनतम एआईसी मूल्य के आधार पर सर्वश्रेष्ठ मॉडल का चयन किया जाता है। सभी विंडो लंबाई के लिए कम आरएमएपीई, आरएमएसई और एमएई मान फिट किए गए अरफिमा मॉडल के बेहतर प्रदर्शन का संकेत देते हैं। अरिमा और अरफिमा मॉडल के पूर्वानुमान प्रदर्शन के बीच एक तुलनात्मक अध्ययन किया गया है और यह पाया गया है कि सबसे अच्छा फिट अरफिमा मॉडल आरएमएपीई, आरएमएसई और एमएई मूल्यों के संदर्भ में सबसे अच्छा फिटेड अरिमा मॉडल से बेहतर प्रदर्शन करता है। इसलिए यह स्पष्ट है कि चावल के थोक अधिकतम, न्यूनतम और मॉडल मूल्य का वर्णन और मॉडलिंग करने में दीर्घकाल अनुस्मरण एक महत्वपूर्ण और प्रमुख भूमिका निभाती है।

संदर्भ

- बॉक्स, जी. ई. पी., जेनकिंस, जी. एम. एवं रीन्सेल, जी. सी. (2007). टाइम-सीरीज एनालिसिस: फोरकास्टिंग एंड कंट्रोल थर्ड एडीशन पियर्सन एजुकेशन, इंडिया
- पॉल, आर. के. (2014). फोकास्टिंग होल प्राइस आफ पीजन पी लॉन्ग मेमोरी टाइम-सीरीज मॉडल। एजइकॉन सर्च 27(2): 167-1766.
- पॉल, आर. के., गुरुंग, बी. एवं पॉल, ए. के. (2014). मॉडलिंग एन्ड फोरकास्टिंग आफ रिटेल प्राइस आफ अरहर दाल इन करनाल, हरियाणा। इंडियन एग्रिक साइंस 85(1):69-72
- बेरन, जे. (1995). मैक्सजीमम लाइकलीहूड एस्टीमेशन आफ डिफरेंसिंग पैरामीटर फोर इनवर्टिबल शॉर्ट एंड लॉन्ग मेमोरी ऑटोरेग्रेसिव इंटीग्रेटेड मूविंग एवरेज मॉडल्स। जे आर स्टेट सोक सीरीज बी स्टेट मेथोडोल 57(4):659-672
- डिकी, डी. एवं फुलर, डब्ल्यू. (1979). डिस्ट्रीब्युशन आफ एस्टीमेशन फार ऑटोरेग्रेसिव टाइम सीरीज विद ए यूनिट रूट। जे एम स्टेट एसोक 74: 427-431
- फिलिप्स, पी. सी. बी. एवं पेरोन, पी. (1988). टेस्टिंग फार यूनिट रूट्स इन टाइम सीरीज रिग्रेसन बायोमैट्रिका। 75: 335-346

कृषि परीक्षण हेतु प्रचलन प्रतिरोधी लागत प्रभावी द्विस्तर घटकीय रन क्रम

बिजोय चन्द¹, अर्पण भौमिक¹, सीमा जग्गी², एल्दो वर्गीस³, सिनी वर्गीस¹,
अनिंदिता दत्ता¹ एवं देवेन्द्र कुमार¹

¹भा.कृ.अनु.प.—भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली—110012

²कृषि शिक्षा विभाग, भा.कृ.अनु.प. नई दिल्ली—110012

³भा.कृ.अनु.प.—केन्द्रीय समुद्री मत्स्यिकी अनुसंधान संस्थान, कोच्चि—682018

सार

बहुत वर्षों से कृषि परीक्षण निरंतर विकास की रीढ़ की हड्डी जैसे है और कृषि परीक्षणों का महत्व अभी भी दिन प्रतिदिन बढ़ रहा है क्योंकि प्रेक्षणीय घटनाएं बहुत सारे कारकों के संयोजन से प्रभावित है। इसमें घटकीय परीक्षण निर्णायक भूमिका अदा करता है। प्रस्तुत लेख में, तीन घटकों वाले द्विस्तर घटकीय रन क्रम जहाँ पर परीक्षण में महंगे घटकों की उपस्थिति के कारण परीक्षण का यादृच्छिकीकरण बहुत कठिन है, के कुछ पहलुओं पर परिचर्चा की गई है। जब परीक्षणों में महंगे घटक निहित होते हैं तो परीक्षणों की लागत कम करने में ये घटकीय रन क्रम प्रभावी हो सकते हैं। इसके अतिरिक्त, ये घटकीय रन क्रम प्रभावों के लिए प्रतिरोधी हैं जो कि एक महत्वपूर्ण प्रभाव है क्योंकि ये न्यूनतम लागत रन क्रम, यादृच्छिकीकरण की कमी के कारण प्रचलित प्रभावों से प्रभावित होने के अधीन हैं।

कुंजी शब्द: घटकीय परीक्षण, रनक्रम, न्यूनतम स्तर परिवर्तन, न्यूनतम लागत, परिवर्तन के लिए कठिन घटक, प्रचलन प्रतिरोधी

1. परिचय

सांख्यिकीय शब्दावली में, घटकीय परीक्षण, बहुघटक परीक्षणों का ही एक विशेष प्रकार है जिनका कृषि अनुसंधान के बहुत सारे क्षेत्रों में अत्याधिक अनुप्रयोग है। बहुघटक कृषि परीक्षणों से प्राप्त परिणामों पर आधारित निष्कर्ष निरंतर विकास के लक्ष्यों को प्राप्त करने में तेज गति पा रहे हैं क्योंकि घटकीय

परीक्षण हमें केवल किसी विशेष कारक के प्रभाव का अध्ययन ही नहीं करवाते हैं अपितु ये हमें यह स्पष्ट जानने में सहायता करते हैं कि दो अथवा अधिक घटक एक दूसरे से किस प्रकार अन्तःक्रिया करते हैं। यह सदैव अपेक्षित है कि प्रेक्षणों को स्वतंत्र करने के लिए किसी बहुकारक परीक्षण के घटकीय रन क्रमों को यादृच्छिक रूप से संचालित किया जाए ताकि पक्षपात को न्यूनतम कर एक मान्य अनुमान प्राप्त किया जाए जो अन्ततः परीक्षण की परिशुद्धता बढ़ा देगा। यद्यपि यदि बहुकारक परीक्षण के नियोजन स्तर के दौरान समुचित यादृच्छिकीकरण कर लिया जाए तो प्रयोक्ता कारक स्तर परिवर्तनों को बड़ी संख्या में देख सकता है, जो कि अन्ततः परीक्षण को महंगा कर देगा, विशेष रूप से तब यदि परीक्षण में परिवर्तन में कठिन घटक निहित हो (जहां घटकों का स्तर, इस प्रकार के परीक्षणों में निहित घटकों के मूल्य विन्यास के कारण अथवा संचालन पद्धति के कारण कठिन है)। मूल्य प्रभावी न्यूनतम परिवर्तित रन क्रमों का उपयोग अधिकतर परिवर्तन के लिए कठिन घटकों से निहित परीक्षणों हेतु एक अनुकूल विकल्प हो सकता है क्योंकि मूल्य, स्तर परिवर्तनों की संख्या के समानुपाती है। (संदर्भ के लिए कॉक्स (1951), ड्रेपर एवं स्टोनमैन (1968), डिकिन्सन (1974), दे लियोन इत्यादि (2005), कॉरिया इत्यादि (2009), कॉरिया इत्यादि (2012), हाइलो (2013), भौमिक इत्यादि (2015), वर्गीस इत्यादि (2017), भौमिक इत्यादि (2017), ऑप्रइम इत्यादि (2017), वर्गीस इत्यादि (2019), भौमिक इत्यादि (2020) प्योरजा इत्यादि (2020)

आदि देखिए)। यद्यपि किसी विशिष्ट घटकीय संयोजन हेतु बहुत से न्यूनतम परिवर्तित रन क्रम हो सकते हैं।

सीमित यादृच्छिकीकरण के कारण उनका सुनियोजित प्रचलन का प्रभाव देखना अपेक्षित है। प्रस्तुत लेख का सरोकार तीन घटकों वाले प्रचलन प्रतिरोधी न्यूनतम मूल्य द्विस्तर घटकीय रन क्रमों के कुछ पहलुओं से है।

2. प्रचलन घटक सहित घटकीय मॉडल

मान लीजिए X_1, X_2, \dots, X_k k घटक हैं, Y अनुक्रिया चर का $n \times 1$ सदिश है। तब प्रचलन घटक की उपस्थिति में घटकीय रन क्रमों हेतु मॉडल को इस प्रकार परिभाषित किया जा सकता है।

$$Y = F\alpha + G\beta + \varepsilon \quad (1)$$

जहां F ; $n \times p$ क्रम के अभिकल्पना आव्यूह को निर्दिष्ट करता है जहां p मॉडल के घटकीय हिस्से से अनुमानित किये जाने वाले प्राचालों की संख्या है जिसमें सामान्य माध्य, सभी मुख्य एवं अंतःक्रिया प्रभाव निहित हैं। इस प्रकार F आव्यूह के प्रथम स्तम्भ में सभी अवयव 1 होंगे। यह ध्यान देने की बात है कि प्रस्तुत अनुसंधान हेतु 2^3 घटकीय एवं 2^k आंशिक घटकीय की अर्ध प्रतिकृति ली गई है और F आव्यूह, माध्य एवं मुख्य प्रभावों से संरचित माना गया है अर्थात् $p = k + 1$ । यहां α , रुचि के प्राचालों का $p \times 1$ सदिश है। यहां $n \times q$ क्रम का G , प्रचलन प्रभाव मापने के लिए लांबिक बहुपदीय गुणांक को निर्दिष्ट करता है एवं β , प्रचलन प्रभावों का $q \times 1$ सदिश है। यहां, लांबिक बहुपदों के सापेक्ष के रैखिक गुणांक लिये गए हैं अर्थात् $q = 1$ ।

3. परिभाषाएं एवं शब्दावली

समीकरण (1) में दिये गए मॉडल के आधार पर कुछ महत्वपूर्ण शब्दावली नीचे दी जा रही है जो प्रस्तुत अनुसंधान के अध्ययन में ली गई है।

परिभाषा 3.1: उपरोक्त परीक्षण विन्यास को लेकर सामान्य प्रसरण अथवा समकक्ष को न्यूनतम सूचना आव्यूह के निर्धारक को अधिकतम कर D -इष्टतम अभिकल्पना इस प्रकार खोजी गई

$$D = |F'F| \quad (2)$$

अतः अभिकल्पना $d \in \Omega$ (जहां पर Ω समान गुणों वाली अभिकल्पनाओं का एक अ-विकृत वर्ग है) को D -इष्टतम अभिकल्पना कहा जाता है यदि $D = |F'F|$, किसी अन्य अभिकल्पना $d_1 \in \Omega$ की तुलना में d के लिए अधिकतम है।

परिभाषा 3.2: समय प्रचलन की उपस्थिति में, टैक एवं वैन्डेब्रोक (2001) ने इष्टतमता मानदंड में बदलाव किया और एक नया मानदंड जिसे D_t -इष्टतमता मानदंड कहा जाता है, प्रारम्भ किया। टैक एवं वैन्डेब्रोक (2001) के अनुसार, D_t -इष्टतमता मानदंड सामान्य प्रसरण को न्यूनतम कर अथवा समीकरण (3)

$$D_t = |F'F - F'G(G'G)^{-1}G'F| \quad (3)$$

में दिए अनुसार प्रचलन की उपस्थिति में सूचना को समकक्ष अधिकतम कर खोजा गया।

अतः $d \in \Omega$ अभिकल्पना (जहां पर Ω समान गुणों वाली अभिकल्पनाओं का एक अ-विकृत वर्ग है) को D -इष्टतम अभिकल्पना कहा जाता है यदि $D_t = |F'F - F'G(G'G)^{-1}G'F|$ किसी अन्य अभिकल्पना $d_1 \in \Omega$ की तुलना में d के लिए अधिकतम है। यहां पर यह ध्यान देने योग्य है कि यदि अभिकल्पना पूर्णतया प्रचलन मुक्त है तो $F'G = 0$, अतः D_t मूल्य तब D मूल्य के समान होगा। तदुपरान्त, टैक एवं वैन्डेब्रोक (2021) ने एक नया मापन प्रारम्भ किया जिसे विभिन्न रन क्रमों पर प्रचलन घटक के प्रभावों का अध्ययन करने के लिए प्रचलन गुणांक के नाम से जाना जाता है। टैक एवं वैन्डेब्रोक (2001) ने प्रचलन गुणांक को निम्नांकित अनुसार परिभाषित किया।

परिभाषा 3.3 : घटकीय रन क्रम पर प्रचलन प्रभाव देखने के लिए टैक एवं वैन्डेब्रोक (2001) प्रचलन गुणांक परिभाषित किया जो कि किसी विशेष रन क्रम के लिए D_t मूल्य का D मूल्य से अनुपात है अर्थात् प्रचलन गुणांक

$$\text{Trend Factor(TF)} = \left[\frac{D_t}{D} \right]^{\frac{1}{p}} \quad (4)$$

रन क्रम जिसके लिए रूचि को घटकीय प्रभाव के लिए प्रचलन घटक सदैव लांबिक है अर्थात वह रन क्रम जिनके लिए, $F'G = 0$, उन्हें पूर्णतया प्रचलन मुक्त कहा जाता है। पूर्णतया प्रचलन मुक्त अभिकल्पनाओं के लिए प्रचलन गुणांक मूल्य 1 होगा। यदि प्रचलन गुणांक मूल्य 1 के पास हो तो अभिकल्पना को लगभग प्रचलन मुक्त अथवा प्रचलन प्रतिरोधी कहा जाएगा। यद्यपि, यदि प्रचलन गुणांक 0 है तो अभिकल्पना समय प्रचलन से पूर्णतया प्रभावित है। अतः $0 \leq TF \leq 1$ रन क्रम का प्रचलन गुणांक मूल्य इस बात का संकेत है कि समय प्रचलन के विरुद्ध रन क्रम कितने सुदृढ़ हैं।

4. घटकीय परीक्षणों में न्यूनतम परिवर्तित रन अनुक्रमों हेतु शर्तें

घटक स्तरों में परिवर्तनों की संख्या का लागत पर महत्वपूर्ण प्रभाव होता है और परीक्षणों को संचालित करने हेतु प्रयास विशेष रूप से जब किसी घटक के स्तर के परिवर्तन में बहुत अधिक प्रयास की आवश्यकता हो अथवा कुछ अन्य कारणों से कुछ निश्चित समय के लिए प्रतीक्षा करनी आवश्यक हो।

घटक अनुसार स्तर परिवर्तनों की न्यूनतम संख्या तभी प्राप्त की जा सकती है जब कोई एक घटक स्तर दो लगातार परीक्षणात्मक परीक्षणों पर परिवर्तित हो अथवा अभिकल्पना आव्यूह में किसी एक कतार से

सारणी 5.1: 23 घटकीय परीक्षण के कुछ न्यूनतम परिवर्तित रन क्रम

क्रम सं.	घटक अनुसार स्तर परिवर्तन	परिवर्तनों की कुल संख्या	रन क्रम	D- मूल्य	D _t - मूल्य	प्रचलन गुणांक
1	(1, 2, 4)	7	ab abc ac a (1) c bc b	4096	975.2381	0.6985342
2	(1, 4, 2)	7	(1) b bc c ac abc ab a	4096	975.2381	0.6985342
3	(2, 1, 4)	7	(1) c ac a ab abc bc b	4096	975.2381	0.6985342
4	(2, 4, 1)	7	(1) b ab a ac abc bc c	4096	975.2381	0.6985342
5	(4, 1, 2)	7	(1) a ac c bc abc ab b	4096	975.2381	0.6985342
6	(4, 2, 1)	7	(1) a ab b bc abc ac c	4096	975.2381	0.6985342
7	(1, 3, 3)	7	ac abc ab a (1) c bc b	4096	585.14286	0.6147882
8	(3, 1, 3)	7	(1) c ac a ab b bc abc	4096	585.14286	0.6147882
9	(3, 3, 1)	7	(1) a ab b bc c ac abc	4096	585.14286	0.6147882
10	(2, 2, 3)	7	(1) b bc abc ab a ac c	4096	1755.4286	0.8091067
11	(2, 3, 2)	7	ab b (1) c bc abc ac a	4096	1755.4286	0.8091067
12	(3, 2, 2)	7	b (1) a ac c bc abc ab	4096	1755.4286	0.8091067

अगली कतार में जाते हुए केवल एक चिन्ह ही परिवर्तित हो। भौमिक इत्यादि (2017, 2020) ने यह दर्शाया है कि यदि i वें घटक के साथ k घटक हों और $i = 1, 2, \dots, k$, के लिए s_i स्तर हों तो $\prod_{i=1}^k s_i$ घटकीय अभिकल्पना हेतु स्तर परिवर्तनों की कुल संख्या $\left(\prod_{i=1}^k s_i \right) - 1$ होगी अर्थात रनों की संख्या से 1 कम।

5. प्रचलन प्रतिरोधी न्यूनतम लागत 2^3 घटकीय रन क्रम

रन अनुक्रमों में न्यूनतम परिवर्तनों वाले 2^3 घटकीय परीक्षणों हेतु स्तर परिवर्तनों की कुल संख्या 7 होगी (अर्थात $2^3 - 1 = 7$) और न्यूनतम परिवर्तित रन अनुक्रमों की कुल संभावित संख्या 144 हैं। घटक अनुसार स्तर परिवर्तनों के 12 संयोग हैं जहां परिवर्तन की संख्या 7 है अर्थात (1, 2, 4), (1, 4, 2), (2, 1, 4), (2, 4, 1), (4, 1, 2), (4, 2, 1), (1, 3, 3), (3, 1, 3), (3, 3, 1), (2, 2, 3), (2, 3, 2), (3, 2, 2)। इनमें से प्रथम 6 संयोग 8 बार आ रहे हैं और शेष 6 संयोग 16 बार आ रहे हैं। प्रत्येक विशेष संयोग हेतु हमने इष्टतम मानदंड मूल्यों के साथ एक विशेष अभिकल्पना अथवा रन क्रम का वर्णन किया है।

क्योंकि परीक्षण की लागत, परिवर्तनों की संख्या के समानुपाती है इसलिए 2^3 घटकीय परीक्षण लागत हेतु सभी न्यूनतम परिवर्तन के लिए रन अनुक्रमों की संख्या

समान है क्योंकि परिवर्तनों की संख्या नियत (अर्थात् 7) है। यद्यपि प्रचलन का प्रभाव भिन्न होगा क्योंकि घटक अनुसार स्तर परिवर्तन के सभी संयोगों हेतु प्रचलन गुणांक मूल्य समान नहीं हैं। लागत एवं प्रचलन प्रभाव के सापेक्ष (2, 2, 3) के विभिन्न घटक अनुसार स्तर परिवर्तन संयोग प्रचलन प्रभावों के लिए अधिक प्रतिरोधी है क्योंकि प्रचलन गुणांक मूल्य दो संयोगों के अलावा अधिक है इस प्रकार परिवर्तन में कठिन घटकों से निहित परीक्षणों हेतु घटक अनुसार स्तर वाला कोई भी रन क्रम परिवर्तित होता है क्योंकि (2, 2, 3) में विभिन्न क्रमचयों को प्राथमिकता दी जाती है और उनके द्वारा अच्छे प्रचलन प्रतिरोध उपलब्ध करवाने के अलावा वे न्यूनतम लागत रन क्रम हैं।

संदर्भ

- भौमिक, ए., वर्गीस, ई., जग्गी, एस. एवं वर्गीस, सी. (2015). फैक्टोरियल एक्सपेरीमेंट्स विद मिनिमम चेंलेल इन रन सिक्वेंसेज। *जर्नल ऑफ इण्डियन सोसायटी ऑफ एग्रीकल्चरल स्टैटिस्टिक्स*, **69(3)**, 243-255।
- भौमिक, ए., वर्गीस, ई., जग्गी, एस. एवं वर्गीस, सी. (2017). मिनिमली चेन्ज्ड रन सिक्वेंसेज इन फैक्टोरियल एक्सपेरीमेंट्स। *कम्यूनिकेशन इन स्टैटिस्टिक्स -थ्योरी एण्ड मैथड्स*, **46(15)**, 7444-7459।
- भौमिक, ए., वर्गीस, ई., जग्गी, एस. एवं वर्गीस, सी. (2020). ऑन द जैनेरेशन ऑफ फैक्टोरियल डिजाइन्स विद मिनिमम लेवल चेन्जेज। *कम्यूनिकेशन इन स्टैटिस्टिक्स -सिमुलेशन एण्ड कम्यूटेशन*। डी. ओ.आई. 10.1080/0361.918.2020.17202441
- कोरिया, ए. ए., गिरिमा, पी. एवं टार्ट- मार्टोरल, एक्स. (2009). ऐक्पैरीमेंटेशन आर्डर विद गुड प्रापर्टीज फॉर 2^k मफैक्टोरियल डिजाइन्स। *जर्नल ऑफ एप्लाइड स्टैटिस्टिक्स*, **36(7)**, 743-754
- कोरिया, ए. ए., गिरिमा, पी. एवं टार्ट- मार्टोरल, एक्स. (2012). ऐक्पैरीमेंटेशन आर्डर इन फैक्टोरियल डिजाइन्स: न्यू फाइंडिंग्स। *जर्नल ऑफ एप्लाइड स्टैटिस्टिक्स*, **39(7)**, 1577-1591
- कॉक्स, डी. आर. (1951). सम सिस्टेमेटिक एक्सपैरीमेंटल डिजाइन्स। *बायोमैट्रिका*, **38**, 312-323।
- दे लियोन, जी., गिरिमा, पी. एवं टार्ट- मार्टोरल, एक्स. (2005). ऐक्पैरीमेंटेशन आर्डर इन फैक्टोरियल डिजाइन्स विद 8 और 16 रन्स। *जर्नल ऑफ एप्लाइड स्टैटिस्टिक्स*, **32(3)**, 1297-313।
- डकिनसन, ए. डब्ल्यू. (1974). सम आर्डर्स रिक्वायरिंग ए मिनिमम नम्बर ऑफ फैक्टर लेवल चेन्जेज फॉर 2^4 एण्ड 2^5 मेन इफैक्ट्स प्लान्स। *टैक्नोमैट्रिक्स*, **16(1)**, 31-37।
- ड्रेपर, एन. एवं स्टोनमैन, डी. (1968). फैक्टर चेन्जेज एण्ड लीनियर ट्रैंड्स इन ऐट-रन टू-लेवल फैक्टोरियल डिजाइन्स। *टैक्नोमैट्रिक्स*, **10**, 301-311।
- हाइलो, एच. (2013). कम्पैरिजन अमंग रन आर्डर एल्गोरिद्मस फॉर सिक्वेंशियल फैक्टोरियल एक्सपैरीमेंट्स। *कंप्यूटेशनल स्टैटिस्टिक्स एण्ड डेटा एनालिसिस*, **58**, 397-406।
- ओप्राइम, पी. सी., पुरेजा, वी. एम. एम. एवं ऑलीवियरा, एस. सी. डी. (2017). सिस्टेमेटिक सिक्वेंसिंग ऑफ फैक्टोरियल एक्सपैरीमेंट्स एज एन आल्टरनेटिव टू रेण्डम एरर। *जैस्टाओ एण्ड प्रोड्काओ*, **24**, 2018-122।
- पुरेजा, वी. एम. एस., ओप्राइम, पी. सी., कोस्ता, ए. एफ. एवं मोरेल्स, डी. (2020). बार्ड-ऑब्जेक्टिव मैथेमेटिकल मॉडल फॉर ऑपटीमल सिक्वेंसिंग ऑफ टू लेवल फैक्टोरियल डिजाइन्स। *ब्रीजीलियन जर्नल ऑफ प्रोबेबिलिटी एण्ड स्टैटिस्टिक्स*। **34(4)**, 712-727।
- वर्गीस, ई., भौमिक, ए., जग्गी, एस., वर्गीस, सी. एवं कौर, सी. (2017). ऑन जैनेरेशन ऑफ कास्ट इफैक्टिव रेस्पॉन्स सरफेस डिजाइन्स। *कम्यूटर्स एण्ड इलेक्ट्रानिक्स इन एग्रीकल्चर*, **133**, 37-45।
- वर्गीस, ई., भौमिक, ए., जग्गी, एस., वर्गीस, सी. एवं लाल, एस. (2019). ऑन द कंस्ट्रक्शन ऑफ रैस्पॉन्स डिजाइन्स विद निनिमम लेवल चेन्जेज। *यूटिलिटास मैथेमैटिका*, **110**, 293-303।

पशुओं में गैर-रैखिक मिश्रित प्रभाव मॉडल का उपयोग करके वृद्धि मापदंडों का अनुमान और स्थिर प्रभाव मॉडल के साथ तुलना

पंकज दास, भारती, कौस्तव आदित्य, एवं अमृत कुमार पॉल
भा.कृ.अनु.प.—भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली—110012

सार

जानवरों में वृद्धि मॉडलिंग एक जटिल प्रक्रिया है, क्योंकि इसमें जैविक व्याख्या के साथ कुछ मापदंडों के साथ अनुदैर्घ्य माप का वर्णन करने की आवश्यकता होती है। अनुदैर्घ्य आंकड़ों में, समय (आयु) के साथ प्रेक्षणों का प्रसरण बढ़ सकता है, और किसी व्यक्ति पर बार-बार माप में सहसंबद्ध पाया जाता है। आंकड़ों की गैर-स्वतंत्रता कई सांख्यिकीय प्रक्रियाओं में अंतर्निहित एक महत्वपूर्ण धारणा का उल्लंघन करती है और अधिकांश पारंपरिक गैर-रैखिक स्थिर प्रभाव मॉडल में इसे अनदेखा कर दिया जाता है। गैर-रैखिक मिश्रित प्रभाव मॉडल (NLMM) का उपयोग करके इस समस्या का समाधान किया जा सकता है। इसकी परिवर्तनीय सहप्रसरण संरचना और असंतुलित आंकड़ों को समायोजित करने की क्षमता के कारण, अनुदैर्घ्य आंकड़ों, जैसे कि वृद्धि के लिए मिश्रित प्रभाव मॉडल ने लोकप्रियता अर्जित की है। इस अध्ययन में बकरी के वजन के आंकड़ों का उपयोग किया गया है और वॉन-बर्टलान्फी मिश्रित मॉडल को विकसित करने का प्रयास किया गया है। इन आंकड़ों के लिए लॉजिस्टिक, गोम्पटर्ज़ और वॉन-बर्टलान्फी फिक्स्ड और मिश्रित मॉडल को भी उपयुक्त किया गया है। मॉडलों की तुलना के लिए वर्ग माध्य त्रुटि (MSE) और वर्ग मूल माध्य त्रुटि (RMSE) की गणना की गई है। सबसे अच्छे उपयुक्त मॉडल के मापदंडों के साथ-साथ उनकी संबंधित मानक त्रुटि का अनुमान लगाया गया है। इस अध्ययन में यह पाया गया है कि बकरी के वजन आंकड़ों के लिए लॉजिस्टिक मिश्रित प्रभाव मॉडल सबसे अच्छा है और इस मॉडल द्वारा की गयी भविष्यवाणी भी बेहतर है।

मुख्य शब्द: NLMM, अनुदैर्घ्य आंकड़ों, यादृच्छिक सहसंयोजक, MSE, RMSE

परिचय

पशु परीक्षण में, उत्पादक और उपभोक्ता दोनों के लिए निवेश लागत और लाभ के लिए एक जानवर की वृद्धि दर में भिन्नता महत्वपूर्ण है। वृद्धि वंशानुगत और पर्यावरणीय प्रभावों का संयोजन है जिसे शरीर के आकार में समय के साथ बढ़ती के रूप में परिभाषित किया जाता है। ब्रॉडी (1995) के अनुसार वृद्धि, एक अपेक्षाकृत अपरिवर्तनीय प्रक्रिया है। शरीर के किसी पक्ष में जो परिवर्तन होता है, उसे वृद्धि कहते हैं। इस प्रकार से वृद्धि, शरीर के विभिन्न अंगों के आकार में परिवर्तन के साथ-साथ शरीर की लम्बाई, चौड़ाई, ऊंचाई व भार को व्यक्त करता है। जानवरों में वृद्धि एक महत्वपूर्ण अध्ययन है, क्योंकि यह उत्पादन के विभिन्न रूपों, जैसे दूध, मांस आदि को प्रभावित करता है। शरीर के वजन और उम्र के बीच संबंध का अध्ययन, विशेष रूप से मांस पैदा करने वाले जानवरों जैसे बकरी, सुअर, भेड़ आदि के लिए महत्वपूर्ण है। पशु वृद्धि आंकड़ों में अधिकतर स्थिर प्रभाव मॉडल का उपयोग किया जाता है, लेकिन इन मॉडलों में कुछ त्रुटियाँ हैं, जैसे कि विभिन्न जानवरों के शरीर के भार को मापने की तुलना में एक ही जानवर के शरीर से वजन को बार-बार मापने में अधिक संबंध होने की संभावना है, और एक जानवर के भार को अधिक अंतराल में मापने की तुलना में कम अंतराल में मापने पर अधिक सहसंबद्ध होने की संभावना होती है (लिटल इत्यादि, 2000)। अनुदैर्घ्य वृद्धि आंकड़ों में, किसी व्यक्ति में भीतरी और अंतर्वर्ती भिन्नता पायी जाती है। व्यक्ति के बीच परिवर्तनशीलता

को स्थिर प्रभाव मॉडल सम्मिलित नहीं करता है (क्रेग और शिकेल, 2001)। अनुदैर्घ्य आंकड़ों में, समय में वृद्धि (आयु) के साथ प्रेक्षणों के विचरण के बीच संबंध होता है। यह कई सांख्यिकीय तकनीकों को प्रभावित करता है। NLMM मॉडल पारंपरिक गैर-रेखीय स्थिर प्रभाव मॉडल की इन सभी कमियों का समाधान करता है, जिसमें एक यादृच्छिक उपादान शामिल है। वांग और जुइडहोफ़ (2004) ने पक्षियों के बीच भिन्नता और विषम विचरण के लिए मिश्रित गोम्पटर्ज़ वृद्धि मॉडल का उपयोग किया। शिकेल इत्यादि (2005) ने बतख वृद्धि के अध्ययन के लिए वैकल्पिक गैर-रैखिक मिश्रित प्रभाव मॉडल का उपयोग किया। ईडुरन इत्यादि (2008) ने गैर-रेखीय वृद्धि मॉडल निर्धारित करने के लिए निरपेक्ष न्यूनन अनुपात (ARR) जैसे नए दृष्टिकोण प्रस्तुत किए। सिंह इत्यादि (2009) ने बकरी पर वृद्धि वक्र के आनुवंशिक प्राचलों के सांख्यिकीय गुणों का अध्ययन किया। फू इत्यादि (2013) ने दक्षिण-मध्य चीन में चीनी देवदार (कनिंघमिया लेंसियोलेटा) के पेड़ों के लिए गैर-रैखिक मिश्रित-प्रभाव वाले क्राउन विड्थ मॉडल का उपयोग किया। इस शोध पत्र का उद्देश्य बकरी की वृद्धि प्रवृत्ति का वर्णन करने के लिए गैर-रेखीय मिश्रित विकास मॉडल उपयोग कर स्थिर प्रभाव मॉडल के साथ गैर-रेखीय मिश्रित प्रभाव विकास मॉडल की तुलना करना है।

पद्धति

वर्ष 2005 में भा.कृ.अनु.प.- केंद्रीय बकरी अनुसंधान संस्थान, मखदूम, उत्तर प्रदेश के बकरी फार्म से 142 बकरियों के शरीर के वजन में वृद्धि (किलोग्राम) के आंकड़ों को एकत्र किया गया। बकरी के शरीर के वजन के आंकड़े निम्नलिखित रूप में लिए गए थे:

W1: पहले महीने में शरीर का वजन, W2: दूसरे महीने में शरीर का वजन, W3: तीसरे महीने में शरीर का वजन, W4: चौथे महीने में शरीर का वजन, W5: 5वें महीने में शरीर का वजन, W6: 6वें महीने में शरीर का वजन, W7: 7वें महीने में शरीर का वजन, W8: 8वें महीने में शरीर का वजन, W9: 9वें महीने में शरीर का वजन, W10: 10वें महीने में शरीर का वजन, W11:

11वें महीने में शरीर का वजन और W12: 12वें महीने में शरीर का वजन।

इस अध्ययन के लिए 3 गैर-रेखीय वृद्धि मॉडल (गोम्पटर्ज़, लॉजिस्टिक मॉडल और वॉन-बर्टलान्फी) लिए गए। बकरी के शरीर के वजन में वृद्धि मापदंडों का अनुमान लगाने के लिए गैर-रेखीय स्थिर प्रभाव मॉडल को लिया गया और फिर एक यादृच्छिक फलन अर्थात् औसत माध्य से यादृच्छिक विचलन को स्थिर प्रभाव मॉडल को मिश्रित प्रभाव मॉडल में बदलने के लिए सम्मिलित किया गया। फिर नवगठित मिश्रित प्रभाव मॉडल से मापदंडों का अनुमान लगाया गया।

उदाहरण के लिए गोम्पटर्ज़ मॉडल में, कुटनर इत्यादि (1996) द्वारा प्रस्तावित एक स्थिर प्रभाव गोम्पटर्ज़ वृद्धि फलन इस प्रकार है

$$W_{it} = W_m \exp(-\exp(b(t-t^*))) + e_{it}$$

जहां W_m एक ही समूह के सभी व्यष्टियों का औसत परिपक्व (mature) BW है। b परिपक्व दर है। t^* दिनों की संख्या, जिस पर वृद्धि दर अधिकतम थी। e_{it} व्यष्टि i का t आयु पर अवशिष्ट BW है। मजज को, जिसे समान रूप से वितरित माना जाता है और जिसका औसत शून्य और प्रसरण स्थिर है। W_{it} व्यष्टि i का t आयु पर अपेक्षित BW है।

मॉडल में एक यादृच्छिक फलन नप सम्मिलित किया गया है। यादृच्छिक फलन अपने जीनोटाइप W_m के औसत परिपक्व BW से अवशिष्ट के परिपक्व BW का यादृच्छिक विचलन है। अब मिश्रित प्रभाव मॉडल को इस प्रकार भी लिखा जा सकता है

$$W_{it} = f(t; \theta) + u_i g(t; \theta) + e_{it}$$

जहां $\theta = \{W_m, b\}$, जैसे कि यादृच्छिक प्रभाव u_i को $g(t; \theta) = \exp(-\exp(b(t-t^*)))$ से गुणा किया जाता है और मिश्रित प्रभाव वृद्धि मॉडल का माध्य, विचरण और सहप्रसरण इस प्रकार है: $E(W_{it}) = f(t; \theta)$, $V(W_{it}) = \sigma_u^2 g^2(t; \theta) + \sigma_e^2$ & $Cov(W_{it}, W_{it+j}) = \sigma_u^2 g(t; \theta)g(t+j; \theta)$.

इसी तरह मिश्रित प्रभाव लॉजिस्टिक मॉडल इस प्रकार है

$$W_{it} = \frac{(W_m + u)}{1 + e^{-\frac{(t-t^*)}{b}}} + e_{it} \quad \dots 2$$

मॉडल के चिन्ह उपरोक्त मॉडल में परिभाषित हैं। मिश्रित प्रभाव वॉन-बर्टलान्फी मॉडल का संशोधित रूप इस प्रकार है

$$W_{it} = (W_m + u)(1 - Be^{-b(t-t^*)})^3 + e_{it} \quad \dots 3$$

B समाकलन फलन है

वर्तमान अध्ययन में, SAS सॉफ्टवेयर पैकेज 9.2 का उपयोग करके लेवेनबर्ग-माक्वार्ड विधि द्वारा मापदंडों का आकलन किया गया है।

सर्वोत्तम उपयुक्त मॉडल के चयन के लिए निम्नलिखित मानदंड का उपयोग किया गया है:

$$i) \text{ RMSE} = \sqrt{\frac{\sum_{j=1}^N \sum_{i=1}^{t_j} \{W(t) - \hat{W}(t)\}^2}{\sum_{j=1}^N (t_j - p)}} \quad \dots 4$$

जहां p : उपयोग किए गए मापदंडों की संख्या, छ: जानवरों की संख्या है, t_j : j th जानवर के वजन की संख्या।

ii. प्रतिशत पूर्व-सूचना त्रुटि

$$(PPE): \left| \frac{W(t) - \hat{W}(t)}{W(t)} \right| \times 100 \quad \dots 5$$

iii. औसत निरपेक्ष त्रुटि

$$(MAE): \sum_{j=1}^N \sum_{i=1}^{t_j} |W(t) - \hat{W}(t)| / \sum_{j=1}^N (t_j - p) \quad \dots 6$$

मॉडल फिटिंग मानदंड: अध्ययन के परिणामों की तुलना निम्नलिखित मानदंडों से की गयी:

a. AIC (एकाइक सूचना मानदंड)

$$AIC = -\ln L + p \quad \dots 7$$

जहां LP प्राचल वाले मॉडल के लिए संभाव्यता (likelihood) फलन है।

b. BIC (बेयेसियन सूचना मानदंड):

$$BIC = -2L_p + p \ln n \quad \dots 8$$

जहाँ n : प्रतिदर्श आकार, L_p : मॉडल की अधिकतम

Long-likelihood, और p : मॉडल में मापदंडों की संख्या है।

स्थिर और मिश्रित प्रभाव मॉडल की सटीकता

वास्तविक स्थिति में, पारंपरिक सांख्यिकीय मूल्यांकन के आँकड़े जैसे कि वर्गमाध्य त्रुटि (MSE), वर्गमूल माध्य त्रुटि (RMSE) आदि। इनका उपयोग पूर्वानुमान परिणामों का मूल्यांकन और तुलना करने के लिए किया जाता है। इस अध्ययन में डाइबोल्ड मारियानो परीक्षण का प्रयोग किया गया है। डाइबोल्ड मारियानो परीक्षण (डाइबॉल्ड और मारियानो, 1995) निम्नलिखित आंकड़ों का उपयोग करता है:

$$DM = \frac{\bar{d}}{\sqrt{2\pi \hat{f}_d(0)}} \quad \dots 9$$

$d = L(e1t) - L(e2t)$, \bar{d} = हानि फलन (Loss Function) के बीच औसत दूरी, $2\pi \hat{f}_d(0)$: यह $\sqrt{T\bar{d}}$ के अनंतस्पर्शी विचरण का एक सुसंगत आकलक है।

मॉडल उपयुक्तता की जाँच के लिए अवशिष्टों का विवरण विश्लेषण महत्वपूर्ण है। अवशिष्टों की प्रसामान्यता का परीक्षण शापिरो-विल्क टेस्ट, एंडरसन-डार्लिंग टेस्ट और कोलमोगोरोव-स्मिरनोव आदि टेस्ट की मदद से किया जाता है।

परिणाम और चर्चा: इस विश्लेषण के लिए 142 बकरियों (71 नर और 71 मादाओं) से कुल 1704 शरीर के वजन का अवलोकन किया गया। गोम्पर्ट्ज, लॉजिस्टिक और वॉन-बर्टलान्फी मॉडल बकरियों (नर और मादा दोनों) के शरीर के वजन के डेटासेट में उपयोग किए गए। SAS सॉफ्टवेयर (SAS, Institute, 1999) की NLMixed तकनीक का उपयोग करते हुए एक ही डेटासेट को मिश्रित और स्थिर प्रभाव मॉडल दोनों के लिए उपयोग किया गया है।

प्रारंभिक अनुमान: प्रारंभिक अनुमानों का पता लगाने के लिए रैखिककरण तकनीकों का उपयोग किया जाता है। वृद्धि वक्र मापदंडों के प्रारंभिक अनुमान मान तालिका संख्या 1 में दिए गए हैं।

तालिका 1: मापदंडों का प्रारंभिक मान

बकरी के प्रारंभिक मान	मादा	नर
W_m	1	1
B	17	17
σ_e^2	22.74	35.23
σ_u^2	1.5	1.5

मॉडल फिटिंग मानदंड: MSE और RMSE जैसे मॉडल उपयोग मानदंड की गणना तीनों मॉडलों के लिए की गयी। उनके मान तालिका 2 और 3 में दिए गए हैं।

तालिकाओं के आधार पर यह पाया गया है कि मिश्रित प्रभाव मॉडल प्रत्येक वृद्धि मॉडल के लिए सर्वोत्तम हैं। यह भी पाया गया कि अन्य दो मॉडलों (गोम्पटर्ज और वॉन-बर्टलान्फी) की तुलना में लॉजिस्टिक मॉडल में MSE और RMSE कम है। मिश्रित प्रभाव वाले लॉजिस्टिक मॉडल में उच्चतम परिशुद्धता पायी गयी। उपयुक्त मॉडल के प्रदर्शन के मूल्यांकन के लिए गोम्पटर्ज, लॉजिस्टिक और वॉन-बर्टलान्फी मॉडल का उपयोग कर आंकड़ों की गणना की गयी है। 2लॉग संभाव्यता, एआईसी (एकाइक सूचना मानदंड), बीआईसी (बेयेसियन सूचना मानदंड) और बायेसियन सूचना मानदंड (बीआईसी) को तालिका 4 और तालिका 5 में संक्षेपित किया गया है।

तालिका 2: मादा बकरियों के लिए मॉडल फिटिंग मानदंड

मादा	गोम्पटर्ज मॉडल		लॉजिस्टिक मॉडल		वॉन-बर्टलान्फी मॉडल	
	स्थिर प्रभाव	मिश्रित प्रभाव	स्थिर प्रभाव	मिश्रित प्रभाव	स्थिर प्रभाव	मिश्रित प्रभाव
एमएसई (MSE)	5.344	4.270	6.270	1.190	29.637	26.745
आरएमएसई (RMSE)	2.312	2.066	2.500	1.091	5.444	5.171

तालिका 3: नर बकरियों के लिए मॉडल फिटिंग मानदंड

नर	गोम्पटर्ज मॉडल		लॉजिस्टिक मॉडल		वॉन-बर्टलान्फी मॉडल	
	स्थिर प्रभाव	मिश्रित प्रभाव	स्थिर प्रभाव	मिश्रित प्रभाव	स्थिर प्रभाव	मिश्रित प्रभाव
एमएसई (MSE)	6.647	5.595	7.676	1.756	22.132	21.679
आरएमएसई (RMSE)	2.578	2.365	2.772	1.313	4.704	4.656

तालिका 4: मादा बकरियों के लिए मॉडल फिटिंग मानदंड

	गोम्पटर्ज मॉडल		लॉजिस्टिक मॉडल		वॉन-बर्टलान्फी मॉडल	
	स्थिर प्रभाव	मिश्रित प्रभाव	स्थिर प्रभाव	मिश्रित प्रभाव	स्थिर प्रभाव	मिश्रित प्रभाव
-2लॉग संभाव्यता (-2 Log Likelihood)	5475.6	3496.9	3925.3	2873.9	6596.0	4648.7
एआईसी (AIC)	5483.6	3504.9	3933.3	2881.9	6604.0	4656.7
एआईसीसी AICC	5483.7	3505.0	3933.3	2882.0	6604.0	4656.8
बीआईसी (BIC)	5502.6	3514.0	3952.2	2891.0	6622.9	4665.8

तालिका 5: नर बकरियों के लिए मॉडल फिटिंग मानदंड

	गोम्पटर्ज मॉडल		लॉजिस्टिक मॉडल		वॉन-बर्टलान्फी मॉडल	
	स्थिर प्रभाव	मिश्रित प्रभाव	स्थिर प्रभाव	मिश्रित प्रभाव	स्थिर प्रभाव	मिश्रित प्रभाव
-2लॉग संभाव्यता (-2 Log Likelihood)	4371.0	3965.5	4094.6	3158.7	4985.3	4984.8
एआईसी (AIC)	4381.0	3973.5	4102.6	3166.7	4993.3	4992.8
एआईसीसी AICC	4381.1	3973.6	4102.7	3166.7	4993.4	4992.8
बीआईसी (BIC)	4404.7	3982.6	4121.5	3175.7	5012.3	5001.8

तालिका 6. लॉजिस्टिक मॉडल में मानक त्रुटि के साथ प्राचल आकल

	मादा				नर			
	स्थिर प्रभाव		मिश्रित प्रभाव		स्थिर प्रभाव		मिश्रित प्रभाव	
	आकल	मानक त्रुटि	आकल	मानक त्रुटि	आकल	मानक त्रुटि	आकल	मानक त्रुटि
<i>b</i>	7.857	0.067	7.815	0.148	7.028	0.056	7.015	0.115
<i>W_m</i>	41.793	0.353	41.633	0.776	49.925	0.397	49.763	0.865
σ_e^2	6.265	0.305	1.297	0.066	7.665	0.374	1.855	0.094
σ_u^2	—	—	40.681	7.007	—	—	50.062	8.684

मिश्रित प्रभाव मॉडल के उपयुक्त आंकड़े मान संबंधित स्थिर प्रभाव मॉडल की तुलना में कम थे। इन के आधार पर यह निष्कर्ष निकाला गया कि मिश्रित प्रभाव मॉडल, स्थिर प्रभाव मॉडल की तुलना में आंकड़ों में बेहतर तरीके से उपयुक्त होते हैं। गोम्पटर्ज और वॉन-बर्टलान्फी मॉडल की तुलना में लॉजिस्टिक मॉडल सभी आंकड़ों के लिए सबसे उपयुक्त पाया गया।

प्राचल आकल

मॉडल उपयोग मानदंड और उपयुक्त आंकड़ों के आधार पर यह निष्कर्ष निकाला जा सकता है कि दिए गए आंकड़ों के उपयोग के लिए लॉजिस्टिक मॉडल सबसे उपयुक्त था। इसलिए लॉजिस्टिक मॉडल का उपयोग करके स्थिर और मिश्रित प्रभाव मॉडल के प्राप्त वृद्धि मापदंडों के आकल को तालिका 6 में संक्षेपित किया गया है।

यह पाया गया है कि दोनों मॉडलों के लिए

परिपक्व शरीर के वजन (डब्ल्यूएम) और परिपक्वता दर (बी) के अनुमान समान हैं क्योंकि अपेक्षित औसत समान हैं। लॉजिस्टिक ग्रोथ मॉडल के मिश्रित प्रभाव और स्थिर प्रभाव मॉडल में मादा बकरियों का औसत अनुमानित एसिम्टोटिक शरीर का वजन (W_m) 41.63 किलोग्राम और 41.79 किलोग्राम क्रमशः पाया गया। नर आंकड़ों के मामले में, मिश्रित प्रभाव मॉडल में शरीर का वजन 49.76 किलोग्राम और स्थिर प्रभाव मॉडल में 49.92 किलोग्राम पाया गया। इस से यह निष्कर्ष निकाला जा सकता है कि नर का परिपक्व शरीर का वजन (W_m) मादा की तुलना में अधिक है। मादा के लिए औसत अनुमानित अधिकतम वृद्धि दर (बी) 7.8579 ± 0.0676 महीने (स्थिर प्रभाव मॉडल) और 7.8152 ± 0.1484 महीने (मिश्रित प्रभाव मॉडल) है, जबकि नर बकरियों में यह 7.0282 ± 0.0564 महीने (स्थिर प्रभाव मॉडल) और 7.0155 ± 0.1152 महीने (मिश्रित प्रभाव मॉडल) हैं। इससे पता चलता है कि मादाओं की

परिपक्वता दर उच्च होती है। उपयोग किए गए मॉडल (स्थिर और मिश्रित प्रभाव मॉडल) की सटीकता का मूल्यांकन नर और मादा दोनों के लिए सबसे अच्छा मॉडल खोजने के लिए किया जाता है। इस अध्ययन में (आर सॉफ्टवेयर के पूर्वानुमान पैकेज) डाइबोल्ड मारियानो परीक्षण का उपयोग स्थिर और मिश्रित प्रभाव मॉडल दोनों की सटीकता का पता लगाने के लिए किया गया है। परीक्षण के विस्तृत परिणाम तालिका 7 में दिए गए हैं।

तालिका 7: बकरी डेटा के लिए डाइबोल्ड मारियानो टेस्ट के परिणाम

मादा		नर	
DM का मान	p का मान	DM का मान	p का मान
3.160	0.004	3.812	0.001

यह पता लगाया गया है कि मिश्रित प्रभाव मॉडल स्थिर प्रभाव मॉडल से बेहतर प्रदर्शन कर रहा है। मिश्रित प्रभाव मॉडल द्वारा किये गए शरीर के वजन के आकलन वास्तविक शरीर के वजन के करीब पाए गए। उपरोक्त परिणामों के आधार पर यह निष्कर्ष निकाला जा सकता है कि मिश्रित प्रभाव लॉजिस्टिक मॉडल इस प्रकार के शरीर के वजन आंकड़ों के लिए बेहतर है। वर्तमान अध्ययन में, शापिरो-विल्क परीक्षण (शापिरो और विल्क, 1965) का उपयोग मादा और नर

दोनों के शरीर के वजन आंकड़ों में स्थिर और मिश्रित लॉजिस्टिक मॉडल की सामान्यता की जांच के लिए किया गया है। परीक्षण के सारांशित परिणाम तालिका 8 में दिए गए हैं।

यह पाया गया कि नर और मादा में स्थिर और मिश्रित प्रभाव मॉडल के लिए p का मान .05 से अधिक है। परिणाम स्पष्ट रूप से दर्शाता है कि दोनों मॉडल प्रसामान्यता की धारणा को पूरा करते हैं। इसलिए आंकड़ों के लिए उपयुक्त स्थिर और मिश्रित प्रभाव लॉजिस्टिक मॉडल मान्य है।

निष्कर्ष:

अब हम यह निष्कर्ष निकाल सकते हैं कि मिश्रित प्रभाव वाले मॉडल जानवर, जैसे कि बकरी के लिए बेहतर हैं। बकरी आंकड़ों में लॉजिस्टिक मॉडल उपयुक्त पाया गया। मिश्रित लॉजिस्टिक मॉडल की सटीकता पशु में स्थिर प्रभाव मॉडल की तुलना में अधिक होती है। सामान्य तौर पर, यह पाया गया है कि जहां तक शरीर के वजन की मॉडलिंग का संबंध है, नरों की तुलना में मादाओं में सटीकता अधिक है। मॉडल उपयुक्तता के लिए शापिरो विल्क परीक्षण का उपयोग करके उपयोग किए गए मॉडलों के अवशिष्टों की जांच की गई। यह पाया गया है कि सभी मामलों में अवशिष्ट अवशेष सामान्य रूप से वितरित हैं।

तालिका 8. बकरी डेटा के लिए शापिरो टेस्ट का परिणाम।

मादा				नर			
स्थिर प्रभाव		मिश्रित प्रभाव		स्थिर प्रभाव		मिश्रित प्रभाव	
Test value (W)	p value	Test value (W)	p value	Test value (W)	p value	Test value (W)	p value
0.935	0.446	0.963	0.825	0.915	0.247	0.880	0.089

संदर्भ

- ब्रॉडी, एस. (1995). बायोइनरजेटिक्स एंड ग्रोथ, रेनहोल्ड न्यू यॉर्क.
- क्रैग, बी. ए. एवं सचिनकेल, ए. पी. (2001). नॉन लीनियर मिक्सड इफेक्ट्स मॉडल फॉर स्वाइन ग्रोथ, प्रोफेशनल एनिमल साइंस, 17: 99–103
- डीबॉल्ड, एफ. एक्स. (2007). एलिमेंट ऑफ फोरकार्स्टिंग, 4, थॉमसन साउथ वेस्टर्न: सिनसिनाटी, यू एस ए
- डीबॉल्ड, एफ. एक्स. एवं मरियनो, आर. (1995). कम्पेरिंग प्रेडिक्टिव एकुरेसी, जर्नल ऑफ बिजनेस एंड इकनोमिक स्टेटिस्टिक्स, 13: 253–265
- आईदुरान, ई., कुसुक, एम., कारकुस, के. एवं ओज़डेमिर, टी. (2008). न्यू अप्रोचेस टू डेटर्मिनेशन ऑफ द बेस्ट नॉन लीनियर फंक्शन डेस्क्रिबिंग ग्रोथ एट अर्ली फेजेज ऑफ किवीरिक एंड मोरकारमन ब्रीड्स, जर्नल ऑफ एनिमल एंड वेटरनरी एडवांसेज 7(7): 799–804.
- फु, लियांग, सुन, एच., शर्मा, पी. आर. एवं ले, वाई. (2013). नॉन लीनियर मिक्सड इफेक्ट्स क्राउन विड्थ मॉडल्स फॉर इंडिविजुअल ट्रीज ऑफ चार्निज़ फर (कन्निघमा लेंसीयोलेटा) इन साउथ सेंट्रल चाइना, फारेस्ट इकोलॉजी एंड मैनेजमेंट 302: 210–220
- कुटनेर, एम. एच., नाचतसेम एवं वास्सरमैन (1996). एप्लाइड लीनियर स्टेटिस्टिकल मॉडल्स, 4 एडिशन, रिचर्ड डी इरविन इंक शिकागो
- लिटल, आर. सी., पेंडरगसत, जे. एवं नटराजन, आर. (2000). टुटोरिअल इन बायो स्टेटिस्टिक्स: मॉडलिंग कोवैरिअन्स स्ट्रक्चर इन द एनालिसिस ऑफ रिपेटेड मेजरड डेटा, स्टेटिस्टिक्स इन मेडिसिन, 19: 1793–1819
- एसएसएस इंस्टिट्यूट (1999). एसएसएस/ एसटीएटी यूजरस गाइड (8.2), एसएसएस इंस्टिट्यूट इंक केरी एंनसी
- सचिनकेल, ऐ. पी., अडोला, ओ. एवं आइंस्टीन, एम. ई. (2005). इवैल्यूएशन ऑफ अल्टरनेटिव नॉन लीनियर मिक्सड इफेक्ट्स मॉडल्स ऑफ डक ग्रोथ, पोल्ट्री साइंस, 84 (2): 256–264
- शेपिरो, एस. एस. एवं विल्क्स, एम. बी. (1965). एन एनालिसिस ऑफ वैरिअन्स टेस्ट फॉर नॉर्मलिटी (कम्प्लीट सैम्पल्स). बायोमेट्रिका 52(3): 591–611
- सिंह, एन. एम. (2009). सम ऐमिपरिकल इन्वेस्टीगेशंस ऑन स्टेटिस्टिकल प्रॉपर्टीज ऑफ ग्रोथ कुवेस, अनपब्लिशड थीसिस आईएआरआई
- वांग, जेड. एवं जुइडोफ, एम. जे. (2004). एस्टिमेशन ऑफ ग्रोथ पैरामीटर्स यूजिंग ए नॉन लीनियर मिक्सड गोम्पेटर्ज मॉडल, पोल्ट्री साइंस, 83(6):847–852.

“

विज्ञान में इतनी विभूति (शक्ति) है,
कि वह काल के चिन्हों को भी मिटा दे।

”

—प्रेमचंद

स्मार्ट खेती में इन्टरनेट ऑफ थिंग्स का उपयोग

सपना निगम¹, अक्षय धीरज², संचिता नाहा¹, मो. अशराफुल हक¹, सौमेन पाल¹ एवं मधु¹

¹भा.कृ.अनु.प.—भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली—110012

²भा.कृ.अनु.प.—भारतीय मृदा एवं जल संरक्षण संस्थान, देहरादून—248195

परिचय

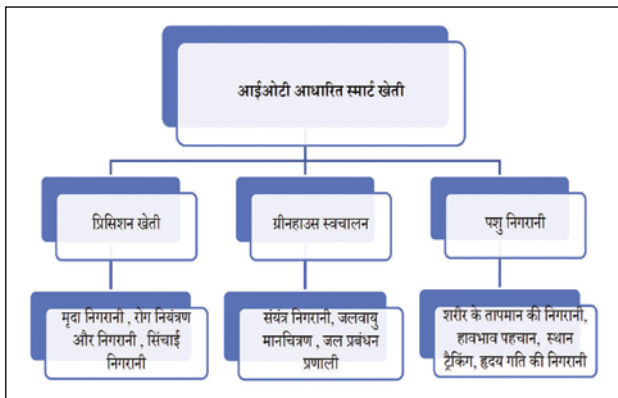
खेती ने पिछले दशकों में कई तकनीकी परिवर्तन देखे हैं, जो अधिक औद्योगिक और प्रौद्योगिकी-संचालित हो गए हैं। विभिन्न स्मार्ट कृषि उपकरणों का उपयोग करके, किसानों ने पशुधन और फसल उगाने की प्रक्रिया पर बेहतर नियंत्रण प्राप्त किया है, जिससे यह अधिक अनुमानित हो गया है और इसकी दक्षता में सुधार हुआ है। दूसरी ओर, स्मार्ट खेती का उपयोग ज्यादातर कृषि में आईओटी (Internet of Things) समाधानों के अनुप्रयोग को दर्शाने के लिए किया जाता है। पर्यावरण और मशीन एकत्र करने के लिए आईओटी संवेदक (Sensor) का उपयोग करके, किसान सूचित निर्णय ले सकते हैं, और अपने काम के हर पहलू में सुधार कर सकते हैं। पशुधन से लेकर फसल की खेती तक, गुणवत्ता और मात्रा के मामले में भोजन की बढ़ती मांग ने कृषि क्षेत्र में औद्योगिकरण और गहनता की आवश्यकता को बढ़ा दिया है। आईओटी एक अत्यधिक आशाजनक तकनीक है जो कृषि क्षेत्र के आधुनिकीकरण के लिए कई नवीन समाधान प्रस्तुत कर रही है। कृषि के विभिन्न क्षेत्रों पर ध्यान देने के लिए आईओटी का उपयोग करके समाधान और उत्पाद देने के लिए अनुसंधान संस्थान और वैज्ञानिक समूह लगातार काम कर रहे हैं।

स्मार्ट खेती को अपनाने से खेती में बेहतर तकनीकों का उपयोग, उत्पादकता में सुधार करता है क्योंकि इसमें निर्गत (Outputs) को अधिकतम करने और अपव्यय को कम करने पर ध्यान दिया जाता है। स्मार्ट खेती एक कृषि प्रबंधन अवधारणा को संदर्भित करती है जो कृषि उत्पादों की गुणवत्ता और मात्रा

बढ़ाने के उद्देश्य से आधुनिक तकनीक का उपयोग करती है। इस दृष्टिकोण में इन्टरनेट ऑफ थिंग्स, आंकड़े प्रबंधन, मृदा स्केनिंग, साथ ही अन्य स्मार्ट प्रौद्योगिकियों के बीच Geo Positioning System (GPS) का उपयोग जैसे पहलू सम्मिलित हैं। आईओटी का पूरा नाम इन्टरनेट ऑफ थिंग्स है। इन्टरनेट ऑफ थिंग्स के जरिए कई प्रकार की प्रौद्योगिकियों और उपकरणों को एक साथ जोड़ा जा सकता है। इन्टरनेट ऑफ थिंग्स नेटवर्किंग का विकास, बड़ी सफलता है। एक पूर्ण आईओटी सिस्टम चार अलग-अलग घटकों को एकीकृत कर कार्य करता है। ये घटक सेंसर या उपकरण, कनेक्टिविटी, डेटा प्रोसेसिंग और यूजर इंटरफेस हैं। सबसे पहले, सेंसर या डिवाइस अपने वातावरण से आंकड़े एकत्रित करता है। जिसके बाद उन आंकड़ों को क्लाउड (इन्टरनेट-आधारित प्रक्रिया) पर भेजा जाता है। वहीं सेंसर सेलुलर, सैटेलाइट, वाईफाई, ब्लूटूथ, जैसी वस्तुओं के माध्यम से अपना डेटा क्लाउड को भेजता है। एक बार जब आंकड़े क्लाउड पर जाते हैं, तो सॉफ्टवेयर उस पर कुछ प्रकार की प्रोसेसिंग करता है। जिसके बाद आपको जानकारी भेजी जाती है।

1. कृषि में आईओटी का अनुप्रयोग

आईओटी कृषि समाधानों में कई निगरानी, नियंत्रण और ट्रैकिंग अनुप्रयोग शामिल हैं जो हवा की निगरानी, तापमान की निगरानी, नमी की निगरानी, मिट्टी की निगरानी, पानी की निगरानी, निषेचन, कीट नियंत्रण, रोशनी नियंत्रण और स्थान ट्रैकिंग जैसे कई प्रकार के चर को मापता है।



- i. **जलवायु निगरानी:** संभवतः सबसे लोकप्रिय स्मार्ट कृषि गैजेट (Gazette) मौसम स्टेशन हैं, जो विभिन्न स्मार्ट खेती सेंसरों को मिलाते हैं। पूरे क्षेत्र में स्थित, वे पर्यावरण से विभिन्न आंकड़े एकत्र करते हैं और इसे क्लाउड पर भेजते हैं। प्रदान किए गए माप का उपयोग जलवायु परिस्थितियों को मैप (Map) करने, उपयुक्त फसलों का चयन करने और उनकी क्षमता में सुधार के लिए आवश्यक उपाय करने के लिए किया जा सकता है। इस उप-डोमेन (Sub-domain) का उद्देश्य हवा की स्थिति का मूल्यांकन और निर्धारण करना है ताकि इसे हानिकारक प्रभावों से रोका जा सके। प्रणाली (System) में एक तापमान और आर्द्रता सेंसर होता है जो ज़िगबी नामक संचार तकनीक द्वारा समर्थित होता है और सौर पैनलों द्वारा संचालित होता है।
- ii. **ग्रीनहाउस (Green House) स्वचालन:** आमतौर पर, किसान ग्रीनहाउस पर्यावरण को नियंत्रित करने के लिए हस्तचालित हस्तक्षेप का उपयोग करते हैं। आईओटी सेंसर का उपयोग उन्हें ग्रीनहाउस स्थितियों जैसे प्रकाश, तापमान, मिट्टी की स्थिति और आर्द्रता पर सटीक वास्तविक समय की जानकारी प्राप्त करने में सक्षम बनाता है। पर्यावरणीय डेटा सोर्सिंग (Environmental Data Sourcing) के अलावा, मौसम स्टेशन (Station) द्वारा दिए गए मापदंडों से मेल खाने के लिए स्वचालित रूप से स्थितियों को समायोजित कर सकते हैं। विशेष रूप से, ग्रीनहाउस ऑटोमेशन सिस्टम (Green House Automation System) एक समान सिद्धांत

का उपयोग करते हैं।

- iii. **मवेशी निगरानी और प्रबंधन:** फसल की निगरानी की तरह ही, आईओटी कृषि सेंसर हैं जो जानवरों के स्वास्थ्य और लॉग प्रदर्शन की निगरानी के लिए एक खेत में संलग्न किए जा सकते हैं। पशुधन ट्रैकिंग और निगरानी स्टॉक स्वास्थ्य, भलाई और भौतिक स्थान पर डेटा एकत्र करने में मदद करती है। फसल को चोरी और जंगली हमलों से बचाने के लिए विभिन्न निगरानी उपकरणों और सेंसर को खेत में तैनात किया गया है। इसके अलावा, खेत में चरने वाले मवेशियों के झुंड की निगरानी आरएफआईडी और डब्ल्यूएसएन का उपयोग करके की जा सकती है। इस तरह, पशुपालक मवेशियों की वास्तविक समय पर निगरानी कर सकते हैं। उदाहरण के लिए, ऐसे सेंसर बीमार जानवरों की पहचान कर सकते हैं ताकि किसान उन्हें झुंड से अलग कर सकें और प्रदूषण से बचा सकें। वास्तविक समय में मवेशियों की ट्रैकिंग के लिए ड्रोन का उपयोग करने से किसानों को स्टाफिंग खर्च कम करने में भी मदद मिलती है।
- iv. **मिट्टी और पानी की निगरानी:** इन में खेतों में मिट्टी की नमी और तापमान की निगरानी के लिए समाधान प्रस्तावित किए जाते हैं। पीएच सेंसर, तापमान और रसायनों को महसूस करके पानी की गुणवत्ता या जल प्रदूषण की निगरानी करने में मदद करते हैं, जो पानी की सामान्य स्थिति को बदल सकते हैं। मिट्टी का तापमान फसल उत्पादकता में महत्वपूर्ण भूमिका निभाता है। सतह और भूजल के बीच पोषक तत्वों की मात्रा की निगरानी के लिए एक प्रणाली प्रस्तावित की गई है। मिट्टी में पोषक तत्वों की मात्रा को मापने के लिए विद्युत रासायनिक प्रतिबाधा लागू की गई थी। मृदा परीक्षण के परिणामों की निगरानी एक अधिष्ठापन (एल), केपेसिटेंस (सी), और प्रतिरोध (आर) (एलसीआर) मीटर के माध्यम से की जाती है, और परिणामों की गणना मानक पुस्तकालय माप के माध्यम से की जाती है।

- v. **फसल और पौधों की वृद्धि की निगरानी:** यह किसी भी बीमारी या संक्रमण को प्रभावी ढंग से रोकने के लिए फसल की वृद्धि और किसी भी विसंगति की निगरानी करता है जो आपकी उपज को नुकसान पहुंचा सकता है। इस प्रस्तावित प्रणाली का मूल उद्देश्य प्रबंधन लागत को कम करना, फसलों की गुणवत्ता में सुधार करना और कीटों के हमलों से बचाव करना है।
- vi. **कृषि व्यवसाय प्रबंधन:** आईओटी आधारित एकीकृत सूचना वातावरण, खेतों, परिवहन प्रणालियों, प्रसंस्करण इकाइयों, बाजारों और एक ही मंच पर प्राप्त निर्यात घर, गुणवत्ता में सुधार, लागत में कमी, बेहतर लाभ और प्रतिस्पर्धी बिक्री मूल्य आदि से जानकारी के साथ।

2. कृषि में आईओटी- आधारित उपकरण/सेंसर

स्मार्ट खेती के लिए तापमान, आर्द्रता और मिट्टी

की नमी को सबसे महत्वपूर्ण माना जाता है। नीचे दी गई तालिका में कृषि के क्षेत्र में आईओटी आधारित विभिन्न सेंसर और उपकरणों के संचालन का निरूपण है।

3. कृषि में आईओटी के लाभ

- आईओटी कृषि प्रक्रिया के दौरान चुनौतियों और मुद्दों को दूर करना संभव बनाता है।
- न्यूनतम अपव्यय के साथ आईओटी का उपयोग करके जल प्रबंधन कुशलतापूर्वक किया जा सकता है।
- यह उत्पादकता बढ़ाने के लिए भूमि की लगातार निगरानी करने में मदद करता है। हस्तचालित काम, समय और संसाधन को कुशलता से कम करता है।
- फसल की वृद्धि का निरीक्षण करने के लिए फसल की निगरानी आसानी से की जा सकती है।

क्र.सं.	सेंसर	संचालन
1	पीएच सेंसर	मिट्टी में पोषक तत्वों की सही मात्रा की निगरानी के लिए, पीएच सेंसर का उपयोग किया जाता है, जो पौधों और फसलों के स्वस्थ विकास के लिए कुशल है।
2	गैस सेंसर	इन्फ्रारेड (Infrared) विकिरणों के अवलोकन के माध्यम से यह सेंसर पशुधन और ग्रीनहाउस में जहरीली गैसों की सटीक मात्रा को मापता है।
3	गति डिटेक्टर (Detector) सेंसर	सेंसर का उपयोग जानवरों और खेत के स्थान को ट्रैक (Track) / ट्रेस (Trace) करने के लिए किया जाता है, इसके अलावा यह खेत या खेत में किसी अवांछित वस्तु की गति का भी पता लगाता है और किसान को समय पर कार्रवाई करने और फसल के नुकसान को रोकने के लिए चेतावनी (alert) उत्पन्न करता है।
4	अल्ट्रा वायलेट सेंसर और निष्क्रिय इन्फ्रारेड (PIR) सेंसर	एक अल्ट्रा वायलेट (Ultraviolet) सेंसर फसलों के प्रभावी विकास के लिए यूवी (Ultraviolet) किरणों की निगरानी करता है। पीर (Peer) सेंसर में, एक मोशन (Motion) डिटेक्टर लगाया जाता है जो क्षेत्र में किसी व्यक्ति की गति की सीमा का पता लगाता है। सेंसर में प्रकाश का पता लगाने वाला गुण भी होता है। किसी वस्तु को ट्रैक करते समय, यह फसल के विकास का विश्लेषण करने के लिए बढ़ते तापमान को वोल्टेज (Voltage) में बदल देता है।
5	मिट्टी की नमी सेंसर	मृदा संवेदक पूरे खेत में पानी की मात्रा और नमी के स्तर को मापता है।
6	तापमान सेंसर	मिट्टी के तापमान में परिवर्तन मिट्टी के पोषक तत्वों और नमी के अवशोषण को प्रभावित करता है। मिट्टी और पानी की सतह में पोषक तत्वों की सही मात्रा का नक्शा बनाने के लिए एक नवीन संवेदन दृष्टिकोण प्रस्तुत किया गया है।

- मृदा प्रबंधन जैसे पीएच स्तर और नमी की मात्रा को आसानी से पहचाना जा सकता है।

4. कृषि में आईओटी की चुनौतियां

- आईओटी स्मार्ट कृषि को इंटरनेट कनेक्टिविटी (Internet Connectivity) की उपलब्धता की आवश्यकता है। हमारे देश का ग्रामीण क्षेत्र अभी भी निरंतर इंटरनेट कनेक्टिविटी से जूझ रहा है।
- आईओटी आधारित उपकरण अपनाना एक मुख्य चुनौती है क्योंकि इसके लिए किसानों को उचित प्रशिक्षण और प्रौद्योगिकी की समझ की आवश्यकता होती है।
- सिस्टम पर कई तरह के साइबर (Cyber) हमले हो सकते हैं।
- आईओटी प्रणाली के लिए बड़ी तकनीक की डिजाइनिंग, विकास और रखरखाव की प्रणाली काफी जटिल है।

निष्कर्ष

आईओटी आधारित स्मार्ट खेती सेंसर (मिट्टी की नमी, प्रकाश, तापमान, आदि) की मदद से कृषि भूमि की निगरानी और सिंचाई प्रथाओं को स्वचालित करने के लिए बनाई गई प्रणाली है। छोटे और बड़े पैमाने पर स्मार्ट खेती सभी किसानों के लिए उपयोगी हो गई है, इसमें किसानों को प्रौद्योगिकियों और उपकरणों तक पहुंच मिलती है, जो खेती की लागत को कम करते हुए उत्पादों की गुणवत्ता और मात्रा को अधिकतम करने में मदद करते हैं। सटीक खेती में आईओटी का सबसे आम अनुप्रयोग है। यह वास्तविक समय पर फसल और मिट्टी की स्थिति की निगरानी, पौधे के स्वास्थ्य पर

नज़र रखने और मौसम की भविष्यवाणी जैसी प्रक्रियाओं को सम्मिलित करके कृषि प्रथाओं को अधिक सटीक और सुनियोजित बनाता है। किसान इस प्रणाली द्वारा प्राप्त अंतर्दृष्टि के आधार पर अपने खेतों का प्रबंधन कर सकते हैं। इसके अलावा, मोबाइल उपकरणों, हाई-स्पीड इंटरनेट, कम लागत वाले उपग्रहों (इमेजरी और पोजिशनिंग के लिए) का उपयोग करके, आईओटी-एकीकृत आर्टिफिशियल इंटेलिजेंस सिस्टम (IOT-Integrated Artificial Intelligence System) फसल विकास प्रक्रियाओं को अनुकूलित करने के लिए नवीनतम (up-to-date) चेतावनी और आंकड़े दे सकता है। यह कृषि पद्धति संसाधनों के सटीक उपयोग का आश्वासन देती है और प्रक्रिया की स्थिरता को बनाए रखते हुए क्षेत्र की उत्पादकता को बढ़ाती है।

संदर्भ

- अयाज़, एम., अम्माद-उद्दीन, एम., शरीफ, जेड., मंसूर, ए., एवं अगगौने, ई. एच. एम. (2019). इंटरनेट ऑफ थिंग्स बेस्ड स्मार्ट एग्रीकल्चर: टुवर्ड मेकिंग द फील्ड्स टॉक। आईईईई एक्सेस, 7, 129551–129583.
- फारूक, एम. एस., रियाज़, एस., आबिद, ए., उमर, टी., एवं ज़िक्रिया, वाई.बी. (2020). कृषि में आईओटी प्रौद्योगिकी की भूमिका: एक व्यवस्थित साहित्य समीक्षा। इलेक्ट्रॉनिक्स, 9(2), 319.
- लखवानी, के., गियानी, एच., अग्रवाल, एन., एवं गुप्ता, एस. (2019). स्मार्ट कृषि के लिए आईओटी का विकास एक समीक्षा-विशेषज्ञ अनुप्रयोगों और सुरक्षा में उभरते रुझानों में। सिंग्रार, 425–432.

“

विज्ञान मानवता के लिए सुंदर उपहार है,
हमें इसे विकृत नहीं करना चाहिए।

”

स्वसंयोजित लक्षण प्रतिचित्र का उपयोग करते हुए फसल जीन प्रारूप का वर्गीकरण

सुकान्त दाश, अनिल कुमार, बैद्यनाथ मंडल, सुशील कुमार सरकार एवं देवेन्द्र कुमार
भा.कृ.अनु.प.—भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली—110012

1. परिचय

कोहेनेन तंत्र अथवा स्वसंयोजित प्रतिचित्र कृत्रिम तंत्रिकीय नेटवर्क (ANNS) के लिए संरचना के अन्य प्रकार की अपेक्षा एक समुचित भिन्न दृष्टिकोण प्रदान करता है। उनकी अभिकल्पना प्राथमिक तौर पर्यवेक्षित समस्याओं से अधिक अपर्यवेक्षित अधिगम हेतु की गई है। गुच्छण में, प्रशिक्षण वस्तुओं की सामूहिक सदस्यता अथवा गुच्छों का पहले से ही ज्ञान होना आवश्यक नहीं है। अतः यहाँ पर प्रशिक्षण आँकड़े सेटों में केवल आगत (इनपुट) चर होते हैं। कोहेनेन नेटवर्क एक स्व-संयोजित प्रणाली है जो अपने आप को इस प्रकार अपना लेता है कि समान आगत वस्तुएं कृत्रिम तंत्रिकीय नेटवर्क (कृ.त.न.) में सांस्थितिक निकट न्यूरोन्स से संबद्ध हों। अर्थात् वे न्यूरोन जो कृ.त.न. में भौतिक रूप से एक-दूसरे के निकट हों, समान आगत के लिए समान प्रतिक्रिया करेंगे जबकि वे न्यूरोन जो कृ.त.न. के विन्यास में अधिक दूर हैं, वे समान आगत के लिए भिन्न प्रकार से प्रतिक्रिया करेंगे। प्रमुख लक्ष्य एच्छक विमा के आने वाले आगत स्वरूप का एक द्विविमीय असतत मानचित्र में परिवर्तित करना और इसे क्रमीय आकार में परिवर्तित करना है। नेटवर्क में न्यूरोन एक द्वि-विमीय ग्रिड में व्यवस्थित हैं और इन न्यूरोन में आगत स्वरूप निर्दिष्ट करने की प्रतियोगिता होती है। 'जीतने वाले' न्यूरोन एवं समान स्वरूप वाले न्यूरोन अर्थात् निकटवासी न्यूरोन निर्गत स्थान में संलग्न स्थानों पर व्यवस्थित हैं।

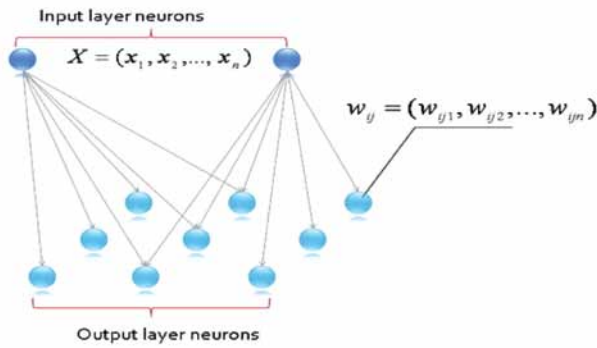
2. आँकड़ों के वर्गीकरण में स्वसंयोजित प्रतिचित्र (SOM) के अवयव एवं उपयोग

मान लीजिए कि हमारे पास कुछ वस्तुओं (उदाहरणार्थ, कारों) का वर्णन करने वाले n - विमीय सदिशों का एक समुच्चय है। प्रत्येक सदिश अवयव वस्तु का (कार के मामले में, जैसे कि चौड़ाई, ऊँचाई, भार, इंजिन का प्रकार, शक्ति, गैसोलीन टैंक, आयतन, आदि-आदि) एक प्राचल है। इस प्रकार का प्रत्येक प्राचल, भिन्न वस्तु के लिए भिन्न है। यदि आपको केवल इन सदिशों को देखकर कार के प्रकार के बारे में तय करना आवश्यक है तो यह आसानी से किया जा सकता है।

2.1 स्वसंयोजित लक्षण प्रतिचित्र (OFM) की संरचना

स्वसंयोजित लक्षण प्रतिचित्र तंत्रिकीय नेटवर्क (SOFMNN) न्यूरोन की दो परतों में बना होता है, (चित्र 1)। प्रथम परत वास्तव में न्यूरोन परत नहीं होती है, केवल आगत आँकड़े प्राप्त करती है और दूसरी परत को स्थानांतरित करती है। एक साधारण प्रकरण (मामला) लेते हैं जब दूसरी परत के न्यूरोन एक द्वि-विमीय ग्रिड से संयुक्त हो जाते हैं। अन्य संरचना जैसे कि त्रिविमीय गोला, बेलन इत्यादि इस लेख के अंतर्गत नहीं आते हैं। तीसरी परत का प्रत्येक न्यूरोन, दूसरी परत के प्रत्येक न्यूरोन से जुड़ जाता है। दूसरी परत में न्यूरोन की संख्या इच्छानुसार चुनी जा सकती

है और अलग कार्य के लिए अलग होती है। दूसरी परत के प्रत्येक न्यूरॉन का अपना भार सदिश होता है। जिसकी विमा आगत परत की विमा के समान होती है। न्यूरॉन अपने साथ वाले न्यूरॉन से निकटवर्ती संबंध द्वारा जुड़े होते हैं जो प्रतिचित्र की सांस्थिति अथवा संरचना का आदेश देता है। विशिष्ट फलन द्वारा निर्दिष्ट इस प्रकार के निकटवर्ती संबंध को सांस्थितिकीय निकटवर्ती कहा जाता है (चित्र 1) देखिए।



चित्र 1: स्वसंयोजित लक्षण प्रतिचित्र (SOFMNN) की संरचना

2.2 अध्ययन नियम

कार्य के प्रारम्भ में, दूसरी परत के न्यूरॉन के सभी भार सदिश को यादृच्छिक मूल्य दिया जाता है। तदुपरान्त, अध्ययन सदिश के समुच्चय से कुछ आगत-सदिश का चयन किया जाता है और तंत्रिकीय नेटवर्क का आगत किया जाता है। इस चरण पर, आगत सदिश एवं सभी न्यूरॉन सदिश के मध्य अन्तर की गणना इस प्रकार की जाती है:

$$D_{ij} = |X^t - W_{ij}| = \sqrt{(x_1 - w_{ij1})^2 + \dots + (x_n - w_{ijn})^2}$$

जहाँ पर i एवं j निर्गत परत में न्यूरॉन के सूचकांक हैं। तदुपरान्त, तंत्रिकीय नेटवर्क 'जीतने वाला- न्यूरॉन' चुनता है अर्थात् वह न्यूरॉन जिसका भार सदिश, आगत सदिश के सर्वाधिक समान हो :

$$D(k_1, k_2) = \min_{i,j} D_{i,j}$$

यहाँ पर, k_1 व k_2 , 'जीतने वाले न्यूरॉन' के सूचकांक हैं। अब, हमें जीतने वाले एवं उस के न्यूरॉनों के भार

सदिशों को सही करने की आवश्यकता है। न्यूरॉन की निकटवर्तीता, सांस्थितिकीय निकटवर्ती फलन द्वारा निर्धारित की जाती है, जो इस प्रकार है :

$$h(\rho, t) = \exp\left(\frac{\rho^2}{2\sigma^2(t)}\right)$$

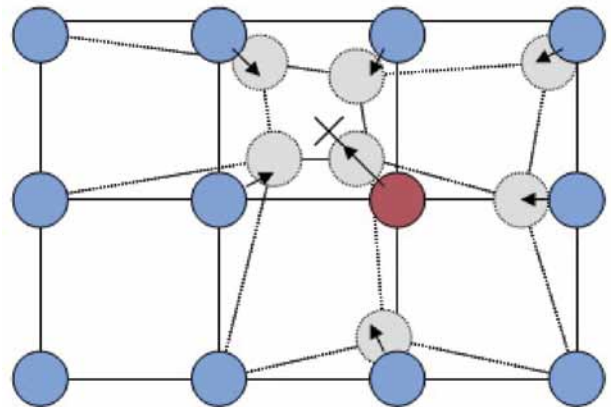
यहाँ पर, ρ जीतने वाले-न्यूरॉन की दूरी है।

अगले चरण में, प्रत्येक न्यूरॉन के लिए सांस्थितिकीय निकटवर्ती फलन की गणना करने के उपरान्त सभी न्यूरॉनों के भारों का अद्यतन निम्नानुसार किया जाता है :

$$\rho = \sqrt{(k_1 - i)^2 + (k_2 - j)^2}$$

यहाँ पर, $a(t)$ अध्ययन दर फलन है जो समय के साथ घटता भी रहता है। यदि कोई न्यूरॉन जीतने वाला है अथवा जीतने के पास वाला है, तो इसका भार सदिश अद्यतन होता है, अन्यथा अपरिवर्तित रहता है। प्रत्येक चरण पर तंत्रिकीय नेटवर्क उस न्यूरॉन को निर्धारित करता है जिसका भार सदिश, आगत सदिश के सर्वाधिक समान है, इसको और इस के निकटवर्ती भार सदिश को आगत सदिश के पास लाने के लिए सही करता है (चित्र 2)।

यहाँ चित्र 2 में ठोस एवं बिन्दुदार रेखाएं क्रमशः अद्यतन से पहली व बाद वाली परिस्थितियों के सदृश हैं। अक्सर प्रशिक्षण समुच्चय से प्रत्येक आगत सदिश



चित्र 2: x द्वारा चिन्हित आगत सदिश की ओर 'जीतने वाला- न्यूरॉन' एवं इसके निकटवर्तियों का अद्यतन

तंत्रिकीय नेटवर्क को प्रस्तुत किया जाता है और अध्ययन या तो किसी चक्र की नियत संख्या तक चलता रहता है या तब तक जब तक कि आगत एवं भार संदिश का अन्तर किसी एपसाइलन मूल्य तक पहुँचे। आस-पास वाले न्यूरॉनों के मध्य अन्तर समय के साथ घटता रहता है और इसीलिए वे समूहों (प्रतिचित्रों) में संयोजित हो जाते हैं जो अध्ययन समुच्चय से किसी एक श्रेणी के सदृश है।

3. कोहोनेन लघुगणक

1. प्रत्येक न्यूरॉन संदिश $w_i = [w_{i1}, w_{i2}, \dots, w_{im}]^T$ एवं $1 \leq i \leq n$, जहाँ पर $n = f \times f$ को यादृच्छिकीय ढंग से प्रारम्भ करें।
2. आगत प्रतिमान $x = [x_1, x_2, \dots, x_m]^T$ प्रस्तुत करें और आगत प्रतिमान x व न्यूरॉन भार संदिश w_i के मध्य यूक्लीडियन दूरी की गणना करें। इस दूरी द्वारा जीतने वाला न्यूरॉन अथवा सर्वाधिक मिलने वाला न्यूरॉन b का पता इस प्रकार लगाया जाता है:

$$d(x - w_b) = \min \{d(x - w_i)\}$$

3. न्यूरॉन i के $w_i(n)$ न्यूरॉन का सूत्रयुग्मक भार संदिश समय n पर किया गया अद्यतन, भार संदिश $w_i(n+1)$ समय $n+1$ पर के लिए इस प्रकार परिभाषित किया गया है :

$$w_i(n+1) = \eta(n) h_{bi}(n) (x(n) - w_i(n)).$$

यहाँ पर $h_{bi}(n)$ समय का फलन है और जीतने वाले न्यूरॉन b एवं मॉडलिंग न्यूरॉन i के बीच की दूरी इस प्रकार परिभाषित की गई है:

$$h_{bi}(n) = \exp\left(\frac{|S_b - S_i|}{2\sigma^2(n)}\right)$$

S दोविमीय प्रतिचित्र गिडों में न्यूरॉन की स्थिति है और $\sigma(n)$ निकटवर्ती फलन परिभाषित करता है। अध्ययन दर प्राचल $\eta(n)$ बढ़ते समय n के साथ धीरे-धीरे घटता है, जो कि η_0 से प्रारम्भ होकर

निम्नलिखित समीकरण द्वारा दिया गया है

$$\eta(n) = \eta_0 \exp\left(\frac{-n}{\tau}\right)$$

यहाँ पर η समय – नियतांक है।

4. जब तक अभिसारिता मानदंड प्राप्त नहीं होता, तब तक चरण 2 व 3 पर वापस जाना है। अभिसारिता मानदंड यहाँ पर विशिष्ट काल द्वारा परिभाषित किया जाता है अर्थात् सभी आगत संदिश कितनी बार प्रशिक्षण के लिए स्व-संयोजित प्रतिचित्र में दिए जाते हैं।

प्रशिक्षण के उपरान्त अद्यतन किये गए भारों को आगत स्वरूप के नमूना बतौर माना जा सकता है और निकटस्थ वर्गीकरण कर्ता को तब अज्ञात स्वरूप के वर्गीकरण करने के लिए प्रयुक्त किया जा सकता है। स्वरूप x हेतु यदि $d(x - w_b) = \min \{d(x - w_i)\}$, तो x को नमूना w_b के लिए नियत किया जाता है। एक वर्ग के नमूने बाह्य स्थान में एक-दूसरे के करीब हैं।

4. उदाहरण

एक परीक्षण में 75 चावल जीनोटाइप एवं नौ गुण सूत्र जैसे कि प्रति पौधा कृषक संख्या (TN), पौधे की ऊँचाई (PH) सेमी में, पुष्प-गुच्छ लम्बाई (PL) सेमी में, पुष्प-गुच्छ भार (PW) ग्राम में, परीक्षण भार (TW) ग्राम में, जैव द्रव्यमान (BM) प्रति पौधा ग्राम में, कटाई सूचकांक (HI) प्रतिशत में, भूसा उपज (SY) प्रति पौधा ग्राम में एवं अन्न उपज (GY) प्रति पौधा ग्राम में, का उपयोग किया गया। इन 75 जीनोटाइप में लम्बे, मध्यम एवं बौने चावल जीनोटाइप हैं और इन्हें भा.कृ. अनु.प.-भा.कृ.अनु.सं., नई दिल्ली के अनुवांशिकी प्रभाग से प्राप्त किया गया है।

अब हमारा उद्देश्य इन 75 चावल जीनोटाइप, जिन के लिए हमने गुच्छण पद्धति का उपयोग किया है, उन पर आधारित समूहों को बनाना है। यद्यपि स्व-संयोजित प्रतिचित्र अपर्यवेक्षित समस्याएं (गुच्छन जैसी) हल करने में सक्षम है। जहाँ पर पहले से ही यह जानना आवश्यक हो जाता है कि गुच्छों के सदस्य

जिसमें प्रशिक्षण गुच्छा अथवा समूह संबंधित है तब हम स्व-संयोजित प्रतिचित्र बनाते हैं। यहाँ पर हमने विश्लेषण Matlab में किया है।

5. कृत्रिम तंत्रिकीय नेटवर्क (ANN) में वर्गीकरण की पद्धति की तरह स्व-संयोजित प्रतिचित्र हेतु Matlab कोड का उपयोग कर प्राप्त परिणाम

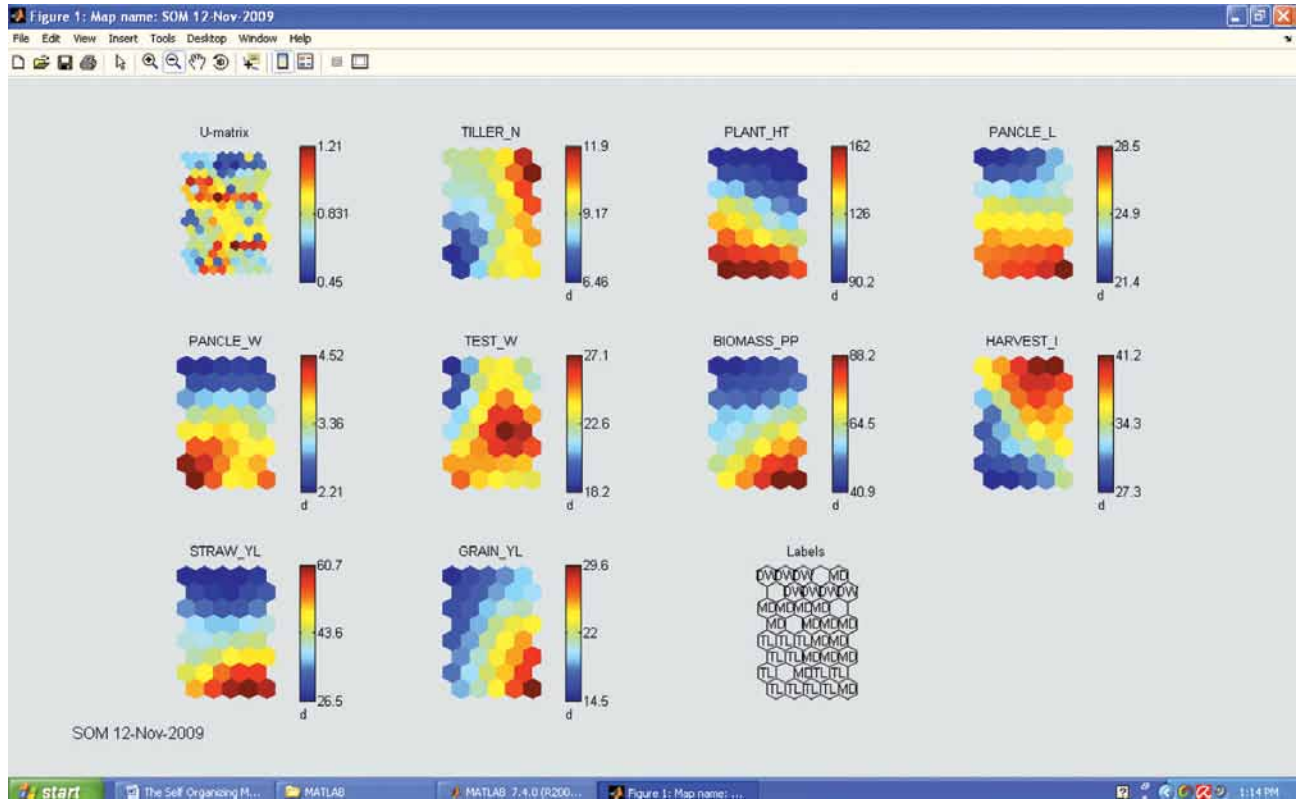
som.show का उपयोग कर स्वसंयोजित प्रतिचित्र (SOM) कल्पित किया गया है। सभी नौ अवयव सतहों के साथ U- आव्यूह दर्शाया गया है। som_show_add का उपयोग रिक्त ग्रिड में कर प्रत्येक प्रतिचित्र इकाई के उपनाम भी दर्शाये गए हैं। अवयवों के मूल्य असामान्य बनाए हैं ताकि रंगीन छड़ पर दर्शाये गए मूल्य मौलिक मूल्य परास में हों। कल्पना चित्र 3 में दिखायी गई है।

U- आव्यूह से यह देखना आसान है कि SOM की ऊपर की तीन कतारें एक बहुत स्पष्ट गुच्छा बनाती है।

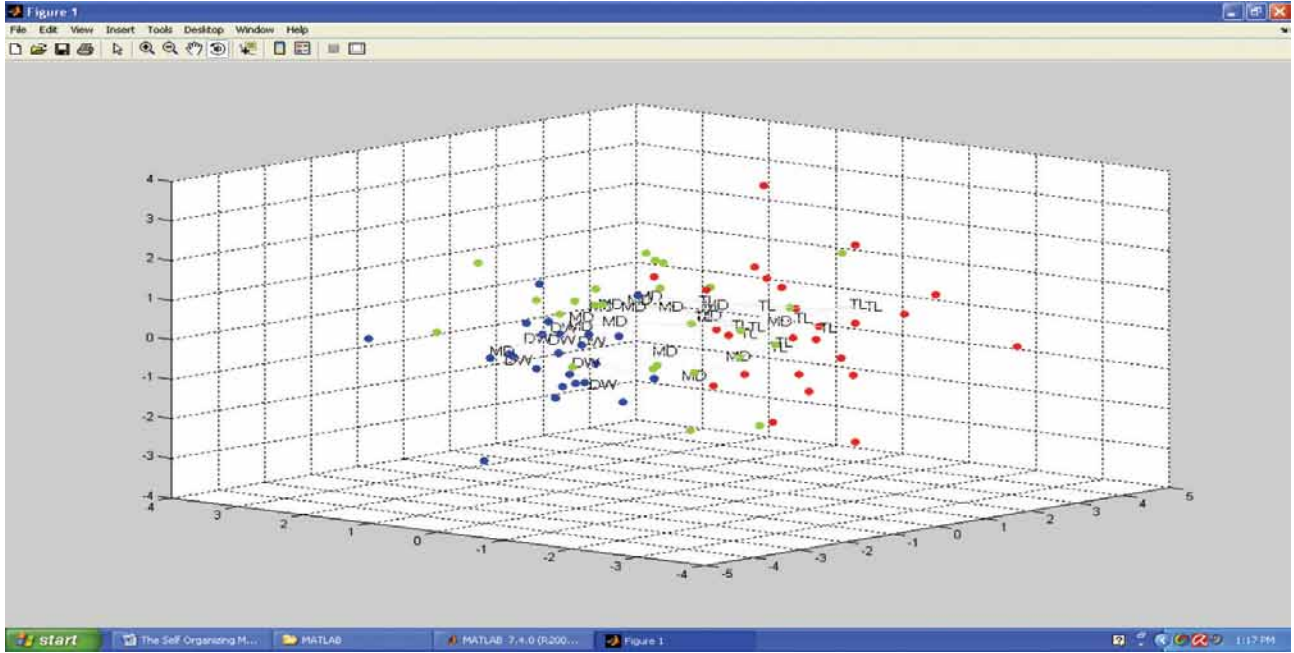
उपनामों को देखने पर यह तुरंत देखा है कि ये ऊंची जीनोटाइप से संबद्ध है। दो अन्य प्रकार के जीनोटाइप अन्य गुच्छे बनाते हैं। U-आव्यूह उन दोनों के मध्य स्पष्ट अन्तर नहीं दर्शाता है, लेकिन उपनामों से ऐसा प्रतीत होता है कि वे गुच्छों के दो विभिन्न हिस्सों से संबद्ध हैं। अवयव सतहों से यह देखा जा सकता है कि वे एक दूसरे से बहुत निकट से संबंधित हैं।

अवयव सतहें तब बहुत ही सुविधाजनक हैं जब किसी को बहुत सारी सूचनाएं तुरंत कल्पित करनी हों। यद्यपि जब थोड़े से चरों में रुचि हो तो प्रकीर्ण रेखाचित्र अधिक दक्ष होते हैं। चित्र 4 आंकड़ों एवं रेखाचित्र ग्रिड दोनों का मुख्य अवयव विश्लेषण (PCA)-प्रक्षेपण दर्शाता है।

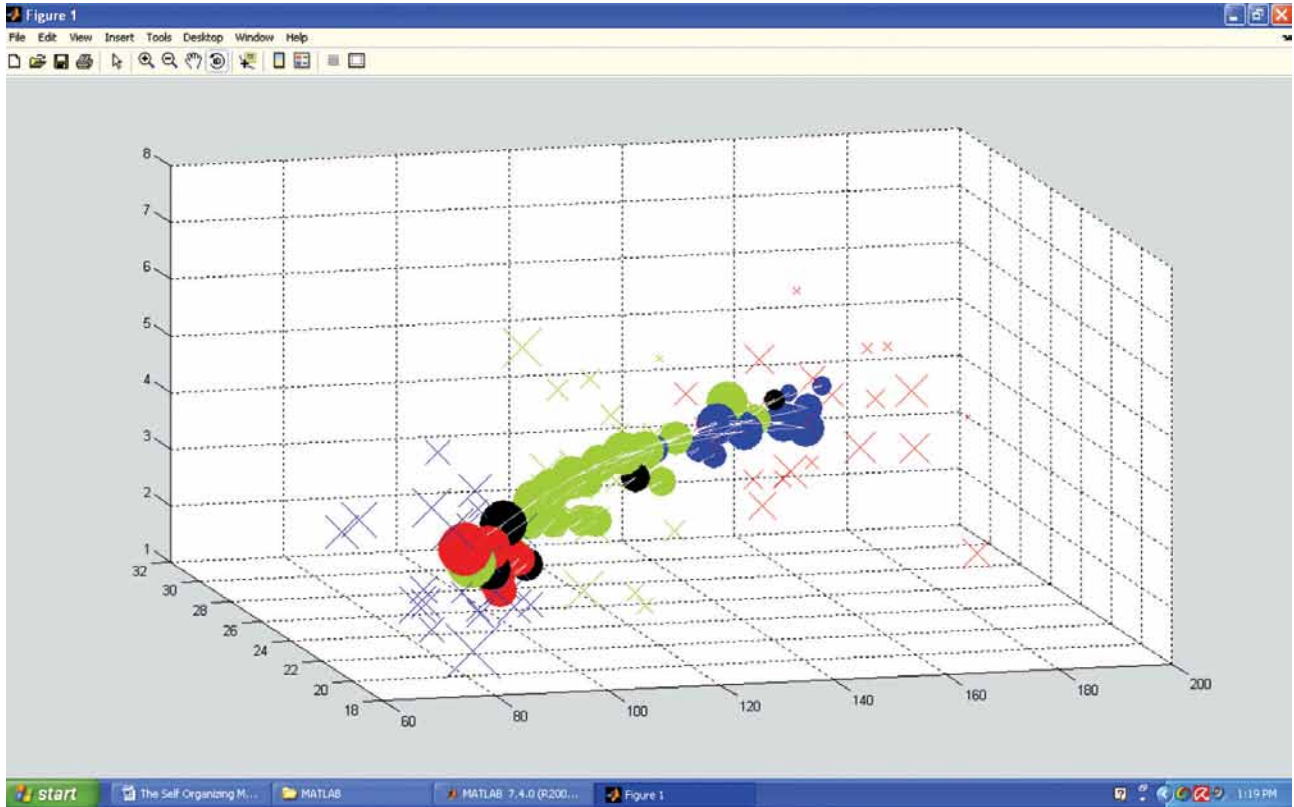
चित्र 5 एवं 6, SOM के सभी नौ चर एवं रेखाचित्र (ग्राफ) में तीन निर्देशांक, चिह्नक आकार एवं रंगों का उपयोग कर जीनोटाइप सूचना कल्पित करता है।



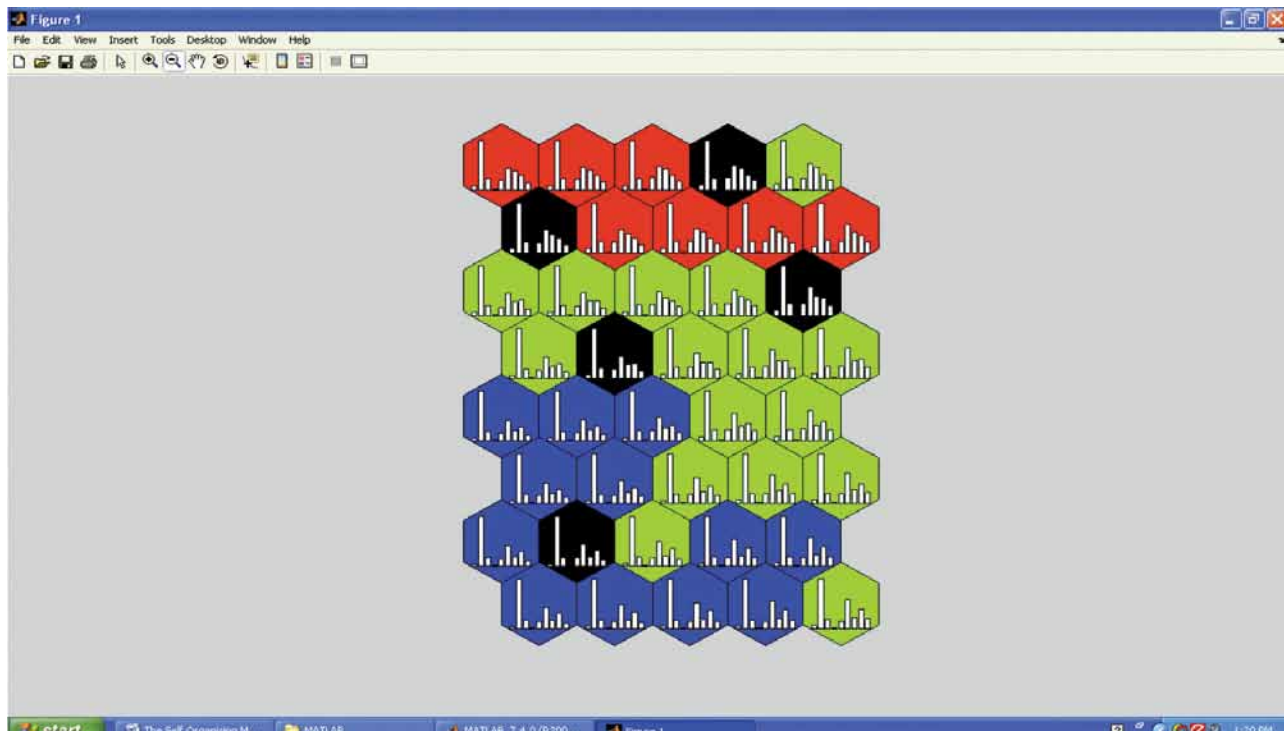
चित्र 3: सभी नौ अवयव सतहों के साथ U- आव्यूह प्रदर्शन



चित्र 4: आँकड़ों एवं रेखाचित्र (ग्राफ) ग्रिड दोनों का मुख्य अवयव विश्लेषण (PCA)- प्रक्षेपण प्रदर्शन



चित्र 5: एस ओ एम के सभी नौ चर एवं रेखाचित्र (ग्राफ) में तीन निर्देशांक, चिह्नक आकार एवं रंगों का उपयोग कर जीनोटाइप सूचना का प्रदर्शन



चित्र 6: रेखाचित्र उपयोग कर उपरोक्त सूचना का प्रदर्शन

5. परिचर्चा

कोहोनेन नेटवर्क एक स्व-संयोजित प्रणाली है, जो अपने आप को इस प्रकार बना लेती है कि एक जैसे आगत वस्तुएं ANN में सांस्थितिकी रूप से पास वाले न्यूरॉन संयुक्त रहें। इस लिए तांत्रिकीय नेटवर्क की यह पद्धति वर्गीकरण उद्देश्य जैसे कि गुच्छण में अत्यधिक उपयोगी है और आँकड़े खनन में भी उपयोगी है। साहित्य में वर्गीकरण (गुच्छण) की विभिन्न पद्धतियां उपलब्ध हैं किन्तु जब आँकड़े अधिक हो तो ये पद्धतियां उपयोग नहीं की जा सकती हैं। इन परिस्थितियों में सामान्यतः स्व-संयोजित पद्धति का अनुप्रयोग किया जाता है। तांत्रिकीय भार आगत स्थान के लिए संकेतक की तरह काम करता है। ये प्रशिक्षण प्रतिदर्शों के वितरण का असतत लगभग सही अनुमान निकटीकरण करते हैं। उच्च प्रशिक्षण प्रतिदर्श सांद्रता वाले क्षेत्रों के लिए अधिक न्यूरॉन बिन्दु एवं कम बिन्दु वहाँ जहाँ प्रतिदर्श की कमी हो। SOM को मुख्य अवयव विश्लेषण (PCA) का अरैखिक सामान्यीकरण माना जा

सकता है। इसे कृत्रिम एवं वास्तविक दोनों प्रकार के आँकड़ों का उपयोग कर के इस प्रकार दर्शाया गया है कि SOM पारम्परिक लक्षण निष्कर्ष पद्धति जैसे कि प्रयोगसिद्ध लाम्बिक फलन (EOF) अथवा PCA की अपेक्षा अधिक लाभप्रद है।

संदर्भ:

टुडु, बी., भट्टाचार्य, एन., जन, ए., घोष, डी. एवं बंदोपाध्याय, आर. (2007). सैल्फ-ऑर्गेनाइजिंग मैप बेस्ड क्लासिफिकेशन ऑफ स्मैल स्टेजेज ऑफ टी फर्मेन्टेशन प्रोसेस यूजिंग इलेक्ट्रानिक नोज, थर्ड इण्डियन इंटरनेशनल कॉन्फ्रेंस ऑन आर्टिफिशियल इंटेलिजेंस, 1626–1635।

दाश, एस., वाही, एस. डी. एवं राव, ए. आर. (2012). क्लासिफिकेशन ऑफ मेज जीनोटाइप्स बाई आर्टिफिशियल न्यूरल नेटवर्क-बेस्ड मैथड: सैल्फ. आर्गेनाइजिंग फीचर मैप। इण्डियन जर्नल ऑफ एग्रीकल्चरल साइंस, 82(2), 161–163।

रिपीटेड मेज़रमेंट डेटा में प्रेडिक्शन दृष्टिकोण

राहुल बनर्जी¹, संघमित्रा पाल², तौकीर अहमद¹ एवं भारती¹

¹भा.कृ.अनु.प-भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली-110012

²सांख्यिकी विभाग, पश्चिम बंगाल राज्य विश्वविद्यालय, बारासात, पश्चिम बंगाल-700126

सार

कृषि, चिकित्सा जैसे जैविक विज्ञान एवं सामाजिक विज्ञान के क्षेत्र में हमें अक्सर रिपीटेड मेज़रमेंट वाले डेटा देखने को मिलते हैं। रिपीटेड मेज़रमेंट का अर्थ यह है कि समय के साथ एक ही व्यक्ति पर बार-बार माप लिया गया है। यदि समष्टि में रिपीटेड मेज़रमेंट मौजूद है, तो यह बाध्य है कि समष्टि में सहसम्बन्ध विद्यमान होगा। रिपीटेड मेज़रमेंट वाले डेटा के सांख्यिकीय विश्लेषण के लिए सहसंबंध संरचना के कारण विशेष ध्यान की आवश्यकता होती है, जो मानक विश्लेषण तकनीक को अमान्य प्रस्तुत कर सकता है। समष्टि में अंतर वर्ग सहसंबंध की उपस्थिति में प्रेडिक्शन दृष्टिकोण पर पहले शोध नहीं किया गया है। इस अध्ययन में हमने एक ऐसे समष्टि का चयन किया है जिसमें रिपीटेड मेज़रमेंट वाले डेटा मौजूद है, फिर उस समष्टि से पीपीएसडब्ल्यूआर (PPSWR) सैंपलिंग (sampling) तकनीक का इस्तेमाल करके प्रतिदर्श का चयन किया गया है एवं ब्रेवर (1963), रॉयल (1970) द्वारा दिए गए प्रेडिक्शन दृष्टिकोण का उपयोग करते हुए समष्टि कुल का आकलन किया गया एवं एक तरफ़ा यादृच्छिक मॉडल का इस्तेमाल करके एनोवा (ANOVA) के जरिये σ एवं ρ का आकलन किया गया है।

मुख्य शब्द: रिपीटेड मेज़रमेंट; प्रेडिक्शन दृष्टिकोण; पीपीएसडब्ल्यूआर (PPSWR); एनोवा; यादृच्छिक प्रभाव मॉडल

प्रस्तावना

कृषि अनुसंधान में रिपीटेड मेज़रमेंट वाले डेटा बहुत अधिक पाए जाते हैं। रिपीटेड मेज़रमेंट का अर्थ यह है

कि समय के साथ एक ही व्यक्ति इकाई पर बार-बार माप लिया गया है। सामान्य अर्थ में, शब्द "इकाई" एक डिज़ाइन किए गए प्रयोग में एक प्रयोगात्मक इकाई, एक प्रतिदर्श सर्वेक्षण में एक प्रतिदर्श इकाई, या एक पूर्वव्यापी अध्ययन में एक विषय को संदर्भित कर सकता है। ज़्यादातर स्थितियों में, रिपीटेड मेज़रमेंट समय के साथ होते हैं, लेकिन वे स्थान के साथ भी हो सकते हैं। कृषि अनुसंधान के लगभग सभी क्षेत्रों में रिपीटेड मेज़रमेंट वाले डेटा देखने को मिलते हैं, इनमें सबसे अधिक समय के साथ पौधों एवं जानवरों में होने वाली वृद्धि के अध्ययन में देखा जाता है। समय के साथ दोहराए गए मेज़रमेंट के अन्य उदाहरणों में प्रायोगिक भूखंडों पर कई फसल कटाई से फसल की पैदावार, व्यक्ति गायों से दैनिक दूध की पैदावार एवं व्यक्ति नीलामी बाज़ारों में साप्ताहिक पशुधन की कीमतें शामिल हैं। स्थान के साथ दोहराए गए मेज़रमेंट के अन्य उदाहरणों में, मिट्टी के कोर नमूनों में कई गहराई पर नमी का निर्धारण, एक लाइन ट्रांज़ेक्ट पर कई बिंदुओं पर प्रदूषक का मापन, साइट्रस पेड़ों के भीतर कई स्थलों पर कीटनाशक स्प्रे जमाव की मात्रा इत्यादि शामिल है।

रिपीटेड मेज़रमेंट वाले डेटा के सांख्यिकीय विश्लेषण के लिए सहसंबंध संरचना के कारण विशेष ध्यान की आवश्यकता होती है, जो मानक विश्लेषण तकनीक को अमान्य प्रस्तुत कर सकता है। संतुलित डेटा के लिए, विचरण विधियों के बहुभिन्नरूपी विश्लेषण को नियोजित किया जा सकता है और समायोजन को अविभाज्य विधियों पर लागू किया जा सकता है, सहसंबंध संरचना के लिए लेखांकन के साधन के रूप

में। रिपीटेड मेज़रमेंट वाले डेटा को "संतुलित" कहा जाएगा यदि प्रत्येक इकाई में एक ही समय बिंदुओं पर पूरा डेटा मौजूद हो। संतुलित डेटा के लिए, पारंपरिक विधि विचरण के समय विश्लेषण में स्प्लिट प्लॉट एनोवा है। इसे यूनीवेरिएट (Univariate) रिपीटेड वाले डेटा का एनोवा भी कहा जाता है।

डेटा विवरण

इस अध्ययन के लिए वृद्धावस्था के नागरिकों की शारीरिक ताकत पर पोषण एवं व्यायाम के प्रभावों के एक बड़े अध्ययन से डेटा लिया गया है। जैसा कि प्रत्येक नागरिक के वजन जो नागरिक की शारीरिक ताकत को दर्शाता है, का माप हर दूसरे दिन दो सप्ताह तक मापा गया था। पहला कार्यक्रम एक नियंत्रण (CONTD) था, जिसमें कोई प्रशिक्षण नियोजित नहीं था, दूसरे कार्यक्रम में एक भार प्रशिक्षण प्रणाली का उपयोग किया गया जिसमें समय के साथ अभ्यास के दोहराव की संख्या में वृद्धि हुई (R1) एवं तीसरे कार्यक्रम में, समय के साथ वजन की मात्रा में वृद्धि हुई (W1)। नागरिकों के सिस्टोलिक रक्तचाप को भी बार-बार अंतराल में मापा गया, नागरिकों के सिस्टोलिक रक्तचाप को सहायक चर के रूप में माना जा सकता है क्योंकि रक्तचाप और अध्ययन चर यानी नागरिकों के वजन के बीच एक महत्वपूर्ण संबंध है। लोगों की ऊंचाई आकार माप चर बनाती है जिसका उपयोग क्रमशः CONTD, R1 और W1 की समष्टि से PPSWR का उपयोग करके प्रतिदर्श लेने के लिए किया जा सकता है।

पद्धति:

व्यायाम चिकित्सा अध्ययन कई दोहराए गए डेटा के डिजाइनों के लिए विशिष्ट है जिसमें विषयों को यादृच्छिक रूप से "उपचार" समूहों को सौंपा जाता है, एवं एक प्रतिक्रिया चर समय के साथ बार-बार मापा जाता है। इस प्रकार के अध्ययन के लिए निम्नलिखित मॉडल उपयोग कर सकते हैं:

$$y_{ijk} = \mu_{ik} + \varepsilon_{ijk} \quad \dots(1)$$

जहां, y_{ijk} , k^{th} समय पर j^{th} उपचार में j^{th} व्यक्ति की प्रतिक्रिया है, और μ_{ik} समय k पर उपचार के लिए समष्टि माध्य है। त्रुटियों ε_{ijk} को सामान्य रूप से माध्य शून्य के साथ वितरित माना जाता है; $V(\varepsilon_{ijk}) = \mathbf{V}$ एवं $\varepsilon_{ij} = [\varepsilon_{ij1}, \varepsilon_{ij2}, \dots, \varepsilon_{ijt}]$

रिपीटेड मेज़रमेंट वाले डेटा के साथ एक प्रमुख मुद्दा सहसंयोजक मैट्रिक्स (Matrix) V की संरचना है। यहाँ, V के 'पंक्ति k , स्तंभ k' ' के तत्व, $\sigma_{kk'}$ पर गणितीय स्थितियों के संदर्भ में पाँच विशेष संरचनाएँ दी गई हैं।

मान लीजिए कि तीन समष्टि (CONTD, R1, W1) निम्नलिखित समीकरण द्वारा व्यक्त की गई हैं:

$$y_i = \beta x_i + \varepsilon_i, i = 1, 2, \dots, N \quad \dots(2)$$

जहां, x_i के मान >0 हैं और गैर-स्टोकेस्टिक (non-stochastic) हैं एवं ε_i ऐसे यादृच्छिक चर हैं जो कि: $E_m(\varepsilon_i) = 0; Var_m(\varepsilon_i) = \sigma_i^2, Cov_m(\varepsilon_i, \varepsilon_j)$

$$= \rho_{ij} \sigma_i \sigma_j, \forall i = 1(1)N, \text{ and } i \neq j = 1(1)N$$

यहाँ, E_m, Var_m और Cov_m क्रमशः मॉडल पर निर्भर

क्रम संख्या	संरचना	प्रतिबंध	गणितीय स्थिति
1.	असंरचित	कोई प्रतिबंध नहीं	$\mathbf{V} = (\sigma_{kk'})$
2.	वृत्ताकार	समान विचरण एवं शून्य सहप्रसरण	$\sigma_{kk} = \sigma^2; \sigma_{kk'} = 0$
3.	कंपाउंड सिमेट्रिक	समान विचरण एवं समान सहप्रसरण	$\sigma_{kk} = \sigma^2;$ $\sigma_{kk'} = \delta \sigma^2$
4.	हुई-फ़ैल्ड	अप्रतिबंधित विचरण; प्रतिबंधित सहप्रसरण	$\sigma_{kk} = 2\tau_k + \phi$ $\sigma_{kk'} = \tau_k + \tau_{k'}$
5.	ऑटोरिग्रेसिव	सहप्रसरण समय अंतराल का फंक्शन है	$\sigma_{kk'} = \theta^{ k-k' }$

एक्सपेक्टेडेशन, (Expectation) विचरण एवं सहप्रसरण हैं।

समष्टि को कुल $Y = \sum_{i=1}^N y_i$ को दो घटकों में विभाजित किया जा सकता है:

$$Y = \sum_{i=1}^N y_i = \sum_s y_i + \sum_r y_i \quad \dots(3)$$

जहाँ, $\sum_s y_i$ प्रतिदर्श में शामिल हिस्सा हैं एवं $\sum_r y_i$ वो हिस्सा है जो प्रतिदर्श में मौजूद नहीं है; इस हिस्से को हम एक यादृच्छिक चर मान सकते हैं एवं इसकी प्रेडिक्शन करना आवश्यक है, चूंकि, y_i भी एक यादृच्छिक चर है, समीकरण 1 को इस प्रकार लिखा जा सकता है:

$$\sum_r y_i = \beta \sum_r x_i + \sum_r \varepsilon_i$$

समीकरण 4 के दोनों पक्षों पर E_m लेने से:

Y का प्रेडिक्टर बन जाता है:

$$t = \sum_s y_i + \hat{\beta} \sum_r x_i \quad \dots(5)$$

जहाँ, $\hat{\beta} = \sum_s B_i y_i; Y$ का एक रैखिक फंक्शन (function) है।

यदि, $B_{0i}, i \in s$, न्यूनतम गुणांक है फिर प्रेडिक्टर का प्रारूप ऐसा लिखा जा सकता है:

$$t_0 = \sum_{i \in s} y_i + \sum_{i \in r} x_i \sum_{i \in s} B_{0i} y_i \quad \dots(6)$$

इस प्रेडिक्टर को Y के लिए श्रेष्ठतम रैखिक अनभिन्नत आकलक (BLUP) कहा जाता है।

एक रैखिक m - अनभिन्नत आकलक के लिए त्रुटि को इस प्रकार लिखा जा सकता है :

$$\begin{aligned} Var_m &= E_m[(t - Y) - E_m(t - Y)]^2 \\ &= E_m[\hat{\beta} \sum_{i \in r} x_i - \sum_{i \in r} y_i]^2 \quad \dots(7) \\ &= E_m[(\sum_{i \in r} x_i)(\hat{\beta} - \beta) - \sum_{i \in r} (y_i - \beta x_i)]^2 = M \end{aligned}$$

$$(मान लीजिए) \quad \dots(8)$$

B_i का एक विकल्प जो M को कम से कम करता है; और इन शर्तों: $\sum_{i \in s} B_i x_i = 1$ को संतुष्ट करता है; $B_i = \frac{1}{n x_i}; i \in s$ मान लीजिये कि $n; s$ के आकार का है M का परिणामी न्यूनतम मान, M_0 है:

$$\begin{aligned} M_0 &= \sigma^2(1 - \rho) \left[\sum_r x_i^2 + \frac{(\sum_r x_i)^2}{n} \right] \\ &= Var_m(t_0 - Y) = E_m(t_0 - Y)^2 \quad \dots(9) \end{aligned}$$

उपरोक्त B_i के साथ रैखिक m - अनभिन्नत आकलक के लिए t_0 लिखना BLUP कहा जाता है, अर्थात:

$$t_0 = \sum_s y_i + \frac{1}{n} \left[\sum_s \frac{y_i}{x_i} \right] \left[\sum_r x_i \right] = \sum_s y_i + \hat{\beta} \sum_s x_i \quad \dots(10)$$

और: $\hat{\beta} = \frac{1}{n} \sum_s \frac{y_i}{x_i}$; जहाँ t_0, β का BLU प्राकलक हैं।

अब अगर हम मान लें:

$$\begin{aligned} \rho_{ij} &= 0; \forall i \neq j, E_m(y_i) = \beta x_i, Var_m(y_i) \\ &= \sigma_i^2; Cov_m(y_i, y_j) = 0; i \neq j; \end{aligned}$$

तो हमें मिलता है $M_1 = \mu_i = \beta x_i$; फिर Y के BLUP को हम इस प्रकार लिख सकते हैं:

$$t_{BLU} = \sum_s y_i + \frac{1}{n} \left[\sum_s \frac{y_i}{x_i} \right] \left[\sum_r x_i \right] = X \frac{[\sum_s y_i]}{[\sum_s x_i]} = X \frac{\bar{y}}{\bar{x}} \quad \dots(11)$$

अगर विशेष रूप से, $\sigma_i^2 = \sigma^2 X_i$, $\bar{y}(x)$ को $y(x)$ के प्रतिदर्श माध्य के रूप में लिखें तो:

$$\begin{aligned} M_0 &= Var_m(t_R - Y) = E_m(t_R - Y)^2 \\ &= E_m \left[\frac{[\sum_s y_i]}{[\sum_s x_i]} \sum_r x_i - \sum_r y_i \right]^2 \\ &= \frac{N^2}{n} (1 - f) \frac{\bar{x} \bar{x}_r}{\bar{x}} \sigma^2 \quad \dots(12) \end{aligned}$$

स्रोत	माध्य वर्ग	अपेक्षित माध्य वर्ग
क्लस्टरों (clusters) के बीच	$B = \frac{\sum_{i \in s} m_i (\bar{y}_{sl} - \bar{y}_s)^2}{n - 1}$	$\sigma^2(1 - \rho) + \frac{\rho \sigma^2}{n - 1} (m - \sum_{i \in s} \frac{m_i^2}{m})$
क्लस्टर (cluster) के भीतर	$W = \frac{1}{n} \sum_{i \in s} \sum_{j \in s_i} \frac{(y_{ij} - \bar{y}_{sl})^2}{m_i - 1}$	$\sigma^2(1 - \rho)$

ANOVA प्राक्कलक

अनुमानकों के आधार पर अनुमानकों का एक सेट विचरण का नमूना विश्लेषण (ANOVA) निम्नलिखित तालिका में दर्शाया गया है। तालिका में समग्र नमूना माध्य को इस प्रकार परिभाषित किया गया है

$$\bar{y}_s = \sum_s \frac{y_i \bar{y}_{sl}}{m} \quad \dots(13)$$

नियमित गणना यह सत्यापित करने के लिए पर्याप्त है कि माध्य के मॉडल के तहत अपेक्षित मान दूसरे कॉलम के वर्ग अंतिम कॉलम के मानों के बराबर हैं। एनोवा तालिका दो अज्ञात प्राचल में दो समीकरण प्रदान करता है, σ^2 और ρ जिसे आसानी से हल किया जा सकता है।

दृष्टांत

नियंत्रण (CNTD) समष्टि के लिए, $N = 140$, नमूना आकार $n = 50$ पीपीएसडब्ल्यूआर द्वारा लिया गया। इस नमूने के आधार पर CONTD समष्टि के लिए t_0 और M_0 के मान इस प्रकार हैं:

$$\hat{\beta} = \frac{1}{n} \sum_s \frac{y_s}{x_s} = 0.76$$

$$t_0 = \sum_s y_s + \hat{\beta} \sum_r x_r = 4004 + (0.76 \times 9661) = 11346.36$$

$$M_0 = Var_m(t_0 - Y) = E_m(t_0 - Y)^2$$

$$= \frac{N^2}{n} (1 - f) \frac{\overline{X x_r}}{\bar{x}} \sigma^2 = 213600.57$$

R1 समष्टि जिसका आकार $N=112$ है, के लिए नमूना आकार $n=35$ पीपीएसडब्ल्यूआर द्वारा लिया गया। इस नमूने के आधार पर R1 समष्टि के लिए t_0 और M_0 के मान इस प्रकार हैं:

$$\hat{\beta} = \frac{1}{n} \sum_s \frac{y_s}{x_s} = 0.75$$

$$t_0 = \sum_s y_s + \hat{\beta} \sum_r x_r = 2810 + (0.75 \times 8376) = 9092$$

$$M_0 = Var_m(t_0 - Y) = E_m(t_0 - Y)^2$$

$$= \frac{N^2}{n} (1 - f) \frac{\overline{X x_r}}{\bar{x}} \sigma^2 = 366381.03$$

W1 समष्टि के लिए, $N = 91$, नमूना आकार $n = 30$ पीपीएसडब्ल्यूआर द्वारा लिया गया। इस नमूने के आधार पर W1 समष्टि के लिए t_0 और M_0 के मान इस प्रकार हैं:

$$\hat{\beta} = \frac{1}{n} \sum_s \frac{y_s}{x_s} = 0.74$$

$$t_0 = \sum_s y_s + \hat{\beta} \sum_r x_r = 2427 + (0.74 \times 6538) = 7265.12$$

$$M_0 = Var_m(t_0 - Y) = E_m(t_0 - Y)^2$$

$$= \frac{N^2}{n} (1 - f) \frac{\overline{X x_r}}{\bar{x}} \sigma^2 = 259863.09$$

विचरण σ^2 एवं इंटरक्लास सहसंबंध गुणांक ρ का अनुमान लगाने के लिए, सभी तीन समष्टि के लिए एनोवा अनुमानक का उपयोग किया गया था जिसके परिणाम नीचे दिए गए हैं।

एनोवा अनुमानकों द्वारा σ एवं ρ का अनुमान

समष्टि नियंत्रण (CNTD) के लिए, $NM=140$, (20 क्लस्टर प्रत्येक का आकार 7) के 10 क्लस्टरों का एक नमूना (प्रत्येक क्लस्टर का आकार 7) पीपीएसडब्ल्यूआर द्वारा लिया गया:

स्रोत	माध्य वर्ग	अपेक्षित माध्य वर्ग
क्लस्टरों के बीच	60.66	60.314
क्लस्टर के भीतर	0.80	0.79

$$\rho = 0.914, \sigma^2 = 9.302$$

समष्टि R1 के लिए, $NM=112$, (16 क्लस्टर प्रत्येक का आकार 7) के हम 8 क्लस्टरों का एक नमूना (प्रत्येक क्लस्टर का आकार 7) पीपीएसडब्ल्यूआर द्वारा लिया गया:

स्रोत	माध्य वर्ग	अपेक्षित माध्य वर्ग
क्लस्टरों के बीच	91.71	91.224
क्लस्टर के भीतर	22.89	22.88

$$\rho = 0.299, \sigma^2 = 32.65$$

समष्टि W1 के लिए, NM=91, (13 क्लस्टर प्रत्येक का आकार 7) के हम 6 क्लस्टरों का एक नमूना (प्रत्येक क्लस्टर का आकार 7) पीपीएसडब्ल्यूआर द्वारा लिया गया:

स्रोत	माध्य वर्ग	अपेक्षित माध्य वर्ग
क्लस्टरों के बीच	181.64	150.61
क्लस्टर के भीतर	1.87	1.87

$$\rho = 0.919, \sigma^2 = 23.12$$

निष्कर्ष:

यहां, इस अध्ययन में हमने देखा है कि रिपीटेड मेज़रमेंट वाले डेटा में समष्टि मानकों की प्रेडिक्शन के लिए प्रेडिक्शन दृष्टिकोण एक व्यवहार्य विधि है, मानक विचलन σ एवं इंटरक्लास सहसंबंध ρ गुणांक का अनुमान एनोवा प्राक्कलक के माध्यम से लगाया जा सकता है। इस प्रकार, सहसम्बन्धित समष्टि के लिए यह विधि व्यवहार्य और बेहतर विकल्प प्रदान करती है।

सन्दर्भ:

- बासु, डी. (1958). ऑन सैपलिंग विद एवं विदाउट रिप्लेसमेंट, *संख्या*, **20**, 287–294.
- बासु, डी. (1969). रोल ऑफ सपफीसिएन्सी एवं लाइकलीहुड प्रिंसिपल्स इन सैपल सर्वे थ्योरी, *संख्या*, सीरीज़ I., **31**, 441–454.
- बॉक्स, जी. ई. पी. (1954). सम थेओरेम्स ऑन क्वाड्रेटिक फॉर्मस एप्लाइड इन द स्टडी ऑफ एनालिसिस ऑफ वरिएंस प्रॉब्लम्स, *एनल्स ऑफ मैथमेटिकल स्टैटिस्टिक्स*, **25**, 290–302.
- ब्रेवर, के. आर. डब्लू. (1963). रेश्यो एस्टिमेशन एवं फाइनाइट पॉपुलेशन: सम रिजल्ट्स देड्यूसिबल

- फ्रॉम द अस्समपशन ऑफ एन अंडरलाइंग स्टोकास्टिक प्रोसेस, *एजेएस*, **5**, 93–105.
- ब्रेवर, के. आर. डब्लू. (1979). रेश्यो एस्टिमेशन एवं फाइनाइट पॉपुलेशन: सम रिजल्ट्स देड्यूसिबल फ्रॉम थी अस्समपशन ऑफ एन अंडरलाइंग स्टोकास्टिक प्रोसेस, *एजेएस*, **5**, 93–105.
- ब्रेवर, के. आर. डब्लू. (1979). ए क्लास ऑफ रोबस्ट सैपलिंग डिज़ाइन फॉर लार्ज स्केल सर्वे, *जर्नल ऑफ अमेरिकन स्टैटिस्टिकल एसोसिएशन*, **74**, 911–915.
- चौधरी, ए. (2010). *एसेंशियल ऑफ सर्वे सैपलिंग*, पी एच लर्निंग प्राइवेट लिमिटेड, नई दिल्ली
- चौधरी, ए. (2014). *मॉडर्न सर्वे सैपलिंग*, सीआरसी प्रेस-टेलर एंड फ्रांसिस ग्रुप, लन्दन
- गोदाम्बे, वी. पी. (1955). ए यूनिफाइड थ्योरी ऑफ सैपलिंग फ्रॉम फाइनाइट पॉपुलेशन, *जर्नल ऑफ द रॉयल स्टैटिस्टिकल सोसाइटी*, **17**, 269–278.
- गोदाम्बे, वी. पी. (1966). ए न्यू एप्रोच टू सैपलिंग फ्रॉम फाइनाइट पॉपुलेशन II डिस्ट्रीब्यूशन फ्री सपफीसिएन्सी, *जर्नल ऑफ थी रॉयल स्टैटिस्टिकल सोसाइटी*, सीरीज़-B, **28(2)**, 310–319.
- जेनरिक, आर. आई. एवं स्च्लुटर, एम. डी. (1986). अनबैलेंस्ड रिपीटेड मेज़र्स मॉडल्स विद स्ट्रक्चर्ड कवरीयन्स मैट्रीसेस, *बायोमैट्रिक्स*, **42**, 805–820.
- सुखात्मे पी. वी., सुखात्मे बी. वी., सुखात्मे एस., एवं अशोक, सी. (1984). *सैपलिंग थ्योरी ऑफ सर्वेज विद ऐप्लिकेशन*, थर्ड एडिशन, आईओवा स्टेट यूनिवर्सिटी प्रेस, युएसए एंड इंडियन सोसाइटी ऑफ एग्रीकल्चरल स्टैटिस्टिक्स, नई दिल्ली.
- वॉलिगंट, आर., डोर्फमैन, ए. एच्. एवं रॉयल एम. एम. (2000). फाइनाइट पापुलेशन सैपलिंग एंड इनफरेंस-ए प्रेडिक्शन, जॉन विलेय, न्यूयॉर्क.

प्रोटीओमिक्स डेटा विश्लेषण का संक्षिप्त विवरण

सुधीर श्रीवास्तव, स्नेहा मुर्मू, मो. समीर फारुकी, नीरज बुढलाकोटी, द्विजेश चंद्र मिश्र,
यू. बी. अंगडि एवं के. के. चतुर्वेदी

भा.कृ.अनु.प.—भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली—110012

भूमिका

प्रोटीन एक महत्वपूर्ण जैविक स्थूल अणु है जो कि विभिन्न प्रकार के कार्य करते हैं। शब्द "प्रोटीओम" को जीव द्वारा उत्पादित या संशोधित प्रोटीन के पूरे समूह के रूप में परिभाषित किया गया है। "प्रोटीओम" शब्द 1994 में मार्क विल्किंस द्वारा सिएना, इटली में आयोजित "2D Electrophoresis% from protein maps to genomes" पर एक संगोष्ठी में गढ़ा गया। प्रोटीओम स्थिर नहीं होता और प्रत्येक कोशिका में भिन्न होता है और समय के साथ बदलता है। प्रोटीओमिक्स आमतौर पर किसी कोशिका प्रारूप के लिए प्रोटीन के बड़े पैमाने पर मात्रात्मक / गुणात्मक अध्ययन को संदर्भित करता है। अब यह विभिन्न क्षेत्रों में एक शक्तिशाली साधन (tool) के रूप में उभरा है जैसे कि बायोमेडिसिन (मुख्य रूप से रोगों के लिए), कृषि और पशु विज्ञान। यह पौधों के कार्यों के विभिन्न पहलुओं के अध्ययन के लिए तेजी से महत्वपूर्ण होता जा रहा है जैसे कि कीटों से पौधों की रक्षात्मक प्रतिक्रिया में प्रत्याशी प्रोटीन की पहचान, फसल उत्पादन पर भू-मंडलीय मौसम परिवर्तन का प्रभाव, इत्यादि। पशु विज्ञान के अध्ययन में प्रोटीओमिक्स महत्वपूर्ण भूमिका निभाते हैं जैसे शरीर क्रिया विज्ञान, प्रतिरक्षा विज्ञान, प्रजनन, लैक्टेशनल बायोलॉजी, इत्यादि। प्रोटीओमिक्स के अनुप्रयोगों में प्रोटीओमिक्स एक्सप्रेसन/एबंडेंस, संरचनात्मक प्रोटीओमिक्स, बायोमार्कर का आविष्कार, इंटरैक्शन प्रोटीओमिक्स, प्रोटीन नेटवर्क, आदि शामिल हैं।

सामान्यतः प्रोटीन एक्सप्रेसन डेटा (आंकड़ा) मास स्पेक्ट्रोमीटर जैसे उच्च-थ्रूपुट तकनीक का उपयोग करके उत्पन्न किया जाता है। जटिल मिश्रण में प्रोटीन और पेप्टाइड्स की पहचान और परिमाणन के लिए प्रोटीओमिक्स में तरल क्रोमैटोग्राफी – मास स्पेक्ट्रोमेट्री

(एम.एस.) का उपयोग एक विधि के रूप में किया जाता है। प्रोटीओमिक्स के दो बुनियादी दृष्टिकोण हैं, अर्थात् बॉटम-अप (नीचे-ऊपर) और टॉप-डाउन (ऊपर-नीचे)। सबसे आम प्रोटीओमिक्स दृष्टिकोण बॉटम-अप है जिसमें एक सैम्पल में प्रोटीन एंजाइमेटिक क्रिया से पेप्टाइड्स में टूट जाता है और उसके पश्चात क्रोमैटोग्राफिक पृथक्करण, आयनीकरण और मास विश्लेषण किया जाता है। बॉटम-अप दृष्टिकोण में, पूर्ण / अखंड प्रोटीन को मास स्पेक्ट्रोमीटर में डाला जाता है जहां उनका खंडीकरण किया जाता है। इसके विपरीत, टॉप-डाउन प्रोटीओमिक्स पूर्ण प्रोटीन के अध्ययन को संबोधित करता है और अक्सर इसका उपयोग शुद्ध या आंशिक रूप से शुद्ध प्रोटीन के लिए किया जाता है।

मात्रात्मक प्रोटीओमिक्स डेटा विश्लेषण

मात्रात्मक प्रोटीओमिक्स डेटा विश्लेषण में शामिल बुनियादी कदम इस प्रकार हैं:

- पेप्टाइड और प्रोटीन की पहचान
- प्रोटीन बहुतायत मात्रा का परिमाणन
- डेटा सफाई
- डेटा सामान्यीकरण
- इम्प्युटेशन (imputation) तकनीकों का उपयोग करके अनुपस्थित मूल्यों को हैंडल (handle) करना
- डेटा विजुअलाइजेशन और व्याख्या
- प्रोटीओमिक्स डेटा का सांख्यिकीय विश्लेषण, आदि रौ (raw) डेटा संचयन, डेटा जमा करने और विश्लेषण के लिए विभिन्न ऑनलाइन संसाधन हैं:
- PRIDE (<http://www.ebi.ac.uk/pride/archive>)

- PeptideAtlas (<http://www.peptideatlas.org/>)
- iProX (<http://iprox.hupo.org.cn/>)
- Tranche (<https://proteomecommons.org/tranche>)
- GPMDB (<http://www.thegpm.org/>)

रौ डेटा के विभिन्न प्रारूप (वित्तउंज) हैं (उदाहरण: cdf, mgf, mzData, mzTab, इत्यादि), जहाँ से एक्सप्रेसन (expression) / एबंडेंस (abundance) डेटा प्राप्त किया जा सकता है। मात्रात्मक प्रोटीओमिक्स डेटा विश्लेषण में शामिल मुख्य कदम बाद के खंड (प्रोटीओमिक्स एक्सप्रेसन / एबंडेंस विश्लेषण) में चर्चा की गई है। इस भाग में पहले दो चरण की चर्चा नीचे की गई है।

पेप्टाइड और प्रोटीन की पहचान

पेप्टाइड्स के अनुक्रम को निर्धारित करने के लिए दो प्रमुख दृष्टिकोण:

- विखंडन स्पेक्ट्रा डेटाबेस के विरुद्ध खोज करना
उदाहरण: मैस्कॉट (Mascot), एंड्रॉमेडा (Andromeda – मैस्कॉट का हिस्सा), सीक्वेस्ट (SEQUEST), X!Tandem / X!!Tandem, इत्यादि
- डी नोवो पेप्टाइड अनुक्रमण
उदाहरण: ihDI(PEAKS), शेरेन्गा (SHERENGA), पेकन (PECAN), इत्यादि

प्रोटीन बहुतायत मात्रा का परिमाणन

विभिन्न स्थितियों के बीच प्रोटीन एबंडेंस में अंतर का पता लगाने के लिए पेप्टाइड्स / प्रोटीन का परिमाणन लेबल-मुक्त या लेबल-युक्त (मेटाबोलिक, एंजाइमेटिक या रासायनिक) हो सकता है। लेबल-मुक्त परिमाणन में, फीचर्स (प्रोटीन अथवा पेप्टाइड्स) की एम.एस. आयन तीव्रता और स्पेक्ट्रल गणना प्रमुख दृष्टिकोण हैं।

लेबल-युक्त मात्रात्मक प्रोटीओमिक्स के टूल्स (ज्वसे) का उदाहरण:

- मैक्सक्वान्ट MaxQuant
- प्रोटीओम डिस्कवरर – थर्मो साइंटिफिक (Proteome Discoverer - Thermo Scientific)

- ऐक्सप्रेस (XPRESS)

लेबल-मुक्त मात्रात्मक प्रोटीओमिक्स के टूल्स का उदाहरण:

- MaxLFQ - मैक्सक्वान्ट में यह मॉड्यूल उपलब्ध है
- emPAI - एक्सपोनेंसीयली मोडिफाइड प्रोटीन बहुतायत सूचकांक (Exponentially modified protein abundance index)
- मैस्कॉट डिस्टिलर – मैट्रिक्स साइंस (Mascot Distiller - Matrix Science)

प्रोटीओमिक्स आंकड़ों में विषमता एवं अनुपस्थित अवलोकन (Missing values)

प्रोटीओमिक्स डेटा विश्लेषण के लिए विभिन्न दृष्टिकोण विद्यमान हैं, जिसमें पहला कदम फीचर्स (Features) की आयन तीव्रता को संक्षेप में प्रस्तुत करना है और इसके बाद कुछ परिवर्तन जैसे कि लॉग परिवर्तन का उपयोग कर उसे नॉर्मल डिस्ट्रिब्युसन (Normal Distribution) के समीप लाते हैं। प्रोटीओमिक्स डेटा विश्लेषण के टूल्स / विधियाँ की उपलब्धता के बावजूद, प्रोटीओमिक्स डेटा का विश्लेषण करने में विभिन्न सांख्यिकीय चुनौतियाँ हैं, जैसे कि डेटा विविधता और अनुपस्थित अवलोकन। प्रत्येक विधि में कई कमियाँ हैं जिनका अध्ययन इन विधियों के सांख्यिकीय गुणों का अध्ययन करके किया जा सकता है।

सैम्पल्स (samples) के बीच जैविक परिवर्तितता (variability) और डेटा उत्पादन (generation) के तकनीकी दृष्टिकोण विषमता को जन्म देती है। जैविक परिवर्तितता आनुवंशिक और पर्यावरणीय कारकों से उत्पन्न होती है। तकनीकी दृष्टिकोण, जैसे कि सैम्पल (Samples) निकालना, संचयन, बफर (buffer) के विभिन्न बैच (batch), मास स्पेक्ट्रोमीटर रन को दोहराना, आदि एबंडेंस डेटा में परिवर्तन लाते हैं।

जब एक डेटा सेट के प्रत्येक समूह में समान संख्या में सैम्पल (विषय) होते हैं, और जब फीचर्स में कोई अनुपस्थित अवलोकन नहीं होता है, तो उस डेटा सेट को संतुलित कहा जाता है। यह स्थिति सर्वदा नहीं

होती। डेटा असंतुलित भी हो सकता है, जिसमें सैम्पल की असमान संख्या, या मिसिंग वेल्यु या दोनों हो सकते हैं। प्रोटिओमिक्स डेटा में मिसिंग वेल्यु जैविक और / या तकनीकी मुद्दों के कारण हो सकते हैं। मिसिंग वेल्यु तीन प्रकार के होते हैं:

- (i) मिसिंग कम्प्लिटली एट रैंडम (MCAR) – इसमें मिसिंग वेल्युस दोनों अप्रत्यक्ष और प्रत्यक्ष डेटा से स्वतंत्र होते हैं। यह Random errors या stochastic fluctuations से उत्पन्न होते हैं।
- (ii) मिसिंग एट रैंडम (MAR) – इसमें प्रत्यक्ष डेटा पर सशर्त, मिसिंग वेल्युस मिसिंग माप से स्वतंत्र हैं। यह मानते हैं कि एक variable के लिए मिसिंग वेल्यु की संभावना अन्य observed variable द्वारा निर्धारित की जाती है।
- (iii) मिसिंग नॉट एट रैंडम (MNAR) – जब डेटा न तो MCAR है और न ही MAR है। ये तब होते हैं जब विशेष स्थितियों में कुछ फीचर्स की मात्रा को निर्धारित नहीं किया जा सकता है क्योंकि वे कुछ सैम्पल्स में पहचान की सीमा से नीचे होते हैं।

मिसिंग वेल्युस वाले डेटा का विश्लेषण या तो मिसिंग वेल्युस वाली फीचर्स को हटा कर, या सांख्यिकीय तरीकों का उपयोग करके जो कि असंतुलित डेटा को हैंडल कर सकता है, या इंप्यूटेशन विधियों का उपयोग करके किया जा सकता है। यदि मिसिंग वेल्युस वाली फीचर्स को हटा दिया जाए, तो जानकारी की हानि होगी। इसलिए, मिसिंग वेल्युस को हैंडल करने वाले तरीकों का उपयोग, जैसे कि इंप्यूटेशन विधियां, आमतौर पर चुनी जाती है।

प्रोटिओमिक्स एक्सप्रेसन/ एबंडेंस विश्लेषण

प्रोटिओमिक्स एबंडेंस विश्लेषण दो या दो से अधिक स्थितियों में महत्वपूर्ण फीचर्स का पता लगाने के लिए किया जाता है, जैसे कि स्वस्थ बनाम विभिन्न रोग स्थितियां। डेटा को विभिन्न परिवर्तन और/ या सामान्यीकरण विधियों (उदाहरण: लॉग परिवर्तन (logarithmic transformation), मात्रात्मक सामान्यीकरण (quantile normalization), विचरण स्थिर सामान्यीकरण (variance stabilizing normalization),

आदि) का उपयोग करके सामान्यीकृत किया जाता है। यदि डेटा में मिसिंग वेल्यु है तो इंप्यूटेशन तकनीकों (उदाहरण: सिंगुलर वेल्यु डिकम्पोसिसन (singular value decomposition), k-निकटतम पड़ोसी (k-nearest neighbour), अधिकतम संभावना प्राक्कलन (maximum likelihood estimation), आदि) का उपयोग किया जा सकता है। महत्वपूर्ण फीचर्स की पहचान के लिए विभिन्न सांख्यिकीय पद्धतियां हैं जैसे टी-टेस्ट (t-test), मॉडरेटेड टी-टेस्ट (moderated t-test), अनोवा (ANOVA), लीनियर मिक्स्ड मॉडल (linear mixed model), इत्यादि।

फीचर्स के अंतर एक्सप्रेसन विश्लेषण के लिए विभिन्न टूल्स और पैकेज उपलब्ध हैं जैसे कि "RepExplore", "MSqRob", "MSstats", "PANDA", "PANDA-view", "DanteR", "DAPAR", "ProStaR", इत्यादि।

वेब सर्वर "RepExplore" का उपयोग कर विश्लेषण

वेब सर्वर "RepExplore" का उपयोग करके प्रोटिओमिक्स एबंडेंस विश्लेषण का एक उदाहरण नीचे वर्णित है। केस-कंट्रोल अध्ययन के लिये एक टेस्ट डेटासेट का भाग उदाहरण के तौर पर नीचे दिया गया है (चित्र 1)। डेटासेट के दोनों समूह (केस और कंट्रोल) में 2 जैविक सैम्पल हैं तथा प्रत्येक जैविक सैम्पल के 2 तकनीकी रेप्लिकेट हैं।

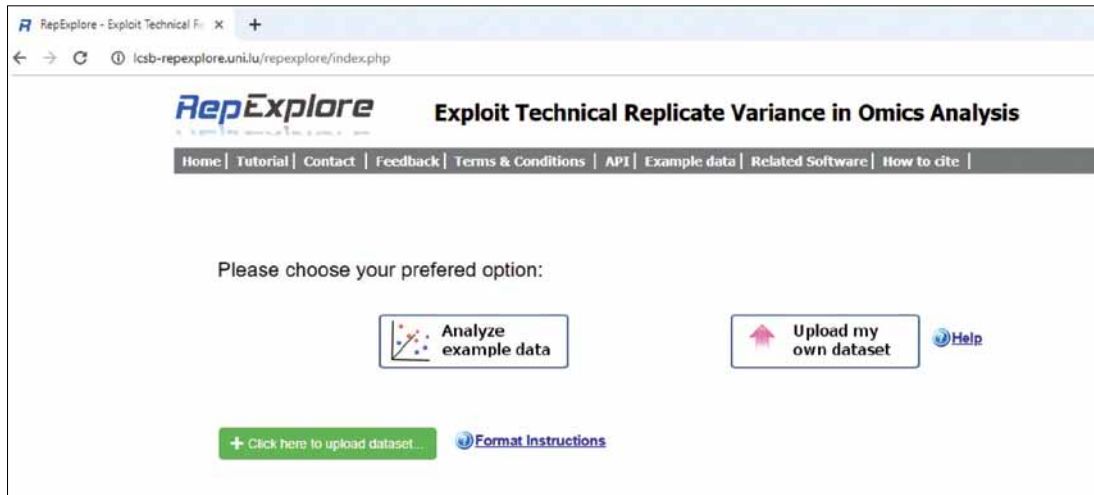
उपयोगकर्ता को डेटा अपलोड करना होता है (चित्र 2)। डेटा अपलोड करने के बाद, उपयोगकर्ता को अन्य विकल्पों का चयन करना होता है (चित्र 3)। इसके बाद उपयोगकर्ता को "Run Analysis!" बटन पर क्लिक करना है (चित्र 3)।

तब उपयोगकर्ता को विभिन्न परिणाम मिलते हैं जो कि नीचे दिए गए हैं (चित्र 4-5)। डिफ्रेंसियली एबंडेंट/एक्सप्रेस्ड फीचर्स की रैंकिंग तालिका नीचे दी गई है (चित्र 4)।

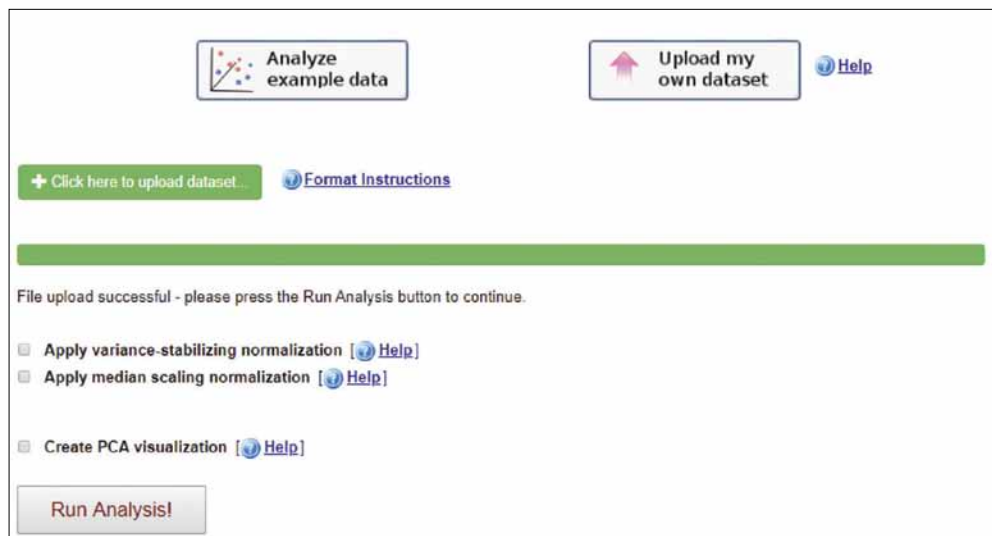
उपयोगकर्ता किसी भी फीचर का बार प्लॉट "generate-bar-plot" बटन पर क्लिक करके प्राप्त कर सकता है जिसका एक उदाहरण नीचे दिया गया है।

	Control					Case					
	control_1_1	control_1_2	control_2_1	control_2_1	control_2_2	case_3_1	case_3_2	case_3_3	case_4_1	case_4_2	case_4_3
biomolecule_1	20.84	19.93	20.78	19.24	20.03	20.87	19.65	20.07	20.7	21.17	20.48
biomolecule_2	19.18	18.79	18.88	18.43	18.97	18.88	18.82	18.64	18.91	19.1	18.87
biomolecule_3	19.5	18.84	20.14	19.06	19.58	19.29	19.1	19.31	18.91	19.3	19.04
biomolecule_4	19.23	18.52	19.67	17.73	19	18.6	16.4	18.44	18.67	19.05	18.88
biomolecule_5	19.64	19.25	19.99	18.78	19.5	19.31	19.16	19.41	18.91	19.19	18.98
biomolecule_6	19.89	19.45	19.93	18.8	19.46	18.76	18.84	18.94	18.96	19	19.04
biomolecule_7	22.07	21.72	23.26	21.35	22.74	21.65	20.97	22.17	22.06	21.81	22.42
biomolecule_8	21.84	21.47	22.81	21.22	22.35	21.58	21.18	22.01	21.58	21.42	22.03
biomolecule_9	17.56	17.41	17.46	17.7	16.76	18.13	18.51	17.3	17.77	17.16	17.76
biomolecule_10	20.34	19.81	21.02	19.23	20.38	19.6	19.06	19.8	19.87	20.16	20.09
biomolecule_11	19.15	18.79	17.98	19.03	17.81	19.55	19.89	18.76	18.89	18.7	18.6
biomolecule_12	24.64	24.12	23.21	24.38	23.31	24.77	25.04	24.21	23.88	23.98	23.68
biomolecule_13	26.51	26.06	26.74	25.23	26.32	25.67	25.15	25.95	25.76	26.19	25.86
biomolecule_14	25	24.42	23.27	24.79	23.48	25.16	25.45	24.58	24.16	24.23	23.91
biomolecule_15	18.05	18.3	18.51	17.98	18.52	17.4	18.36	16.85	18.32	18.23	18.38
biomolecule_16	17.82	17.34	18.24	17.07	17.8	17.66	17.45	17.65	17.77	18.05	18.16
biomolecule_17	17.98	17.31	18.28	17.27	17.77	17.37	17.47	17.31	17.31	17.39	17.66
biomolecule_18	19.32	18.13	19.33	18.04	18.81	18.64	18.66	17.88	18.66	18.93	18.52
biomolecule_19	24.89	24.43	24.82	23.74	24.4	24.24	24.1	24.35	24.27	24.54	24.34
biomolecule_20	17.94	17.25	18.39	17.19	17.19	17.08	16.88	16.82	17.51	18.18	17.42

चित्र 1. केस-कंट्रोल अध्ययन के लिये टेस्ट डेटासेट का भाग



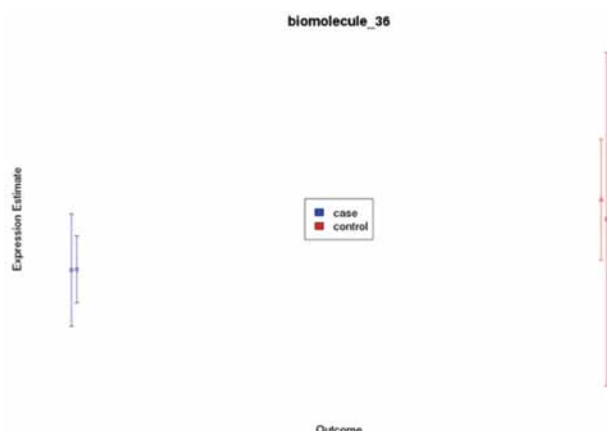
चित्र 2. डेटा अपलोड करना



चित्र 3. विकल्पों का चयन

Biomolecule identifier	Log ₂ fold change	Probability of positive likelihood ratio (PPLR)	P-like significance score (from PPLR, 1-PPLR)	eBayes T-score	eBayes P-value	eBayes adj. P-value
biomolecule_36 generate_bar_plot	-1.02	0.734	0.266	-4.12	0.0068	0.541
biomolecule_87 generate_bar_plot	-0.709	0.773	0.227	-2.96	0.0261	0.541
biomolecule_72 generate_bar_plot	-0.771	0.719	0.281	-2.75	0.0349	0.541
biomolecule_85 generate_bar_plot	-0.696	0.82	0.18	-2.49	0.0489	0.541
biomolecule_5 generate_bar_plot	-0.61	0.824	0.176	-2.47	0.0501	0.541
biomolecule_84 generate_bar_plot	-0.845	0.714	0.286	-2.41	0.0543	0.541
biomolecule_73 generate_bar_plot	-0.567	0.732	0.268	-2.36	0.0579	0.541
biomolecule_50 generate_bar_plot	-0.572	0.721	0.279	-2.34	0.0593	0.541

चित्र 4. डिफरेंसियली एबंडेंट/ एक्सप्रेस्ड फीचर्स की रैंकिंग उपयोगकर्ता किसी भी फीचर का बार प्लॉट



चित्र 5. एक फीचर के बार प्लॉट का उदाहरण

(ProStaR) का उपयोग करके तुलनात्मक विश्लेषण

इस अध्ययन में उपयोग किए गए डेटासेट में दो स्थितियों से प्रोटीन तीव्रता मान शामिल हैं और प्रत्येक स्थिति में तीन रेप्लिकेट हैं (स्रोत: <http://live-prostar-proteomics-org/>; Exp2_R100_prot)। तालिका 1 डेटासेट का मात्रात्मक अवलोकन दिखाती है।

तालिका 1. डेटासेट का अवलोकन

विवरण	वैल्यू (Value)
प्रतिदर्श की संख्या	6
स्थिति की संख्या	2

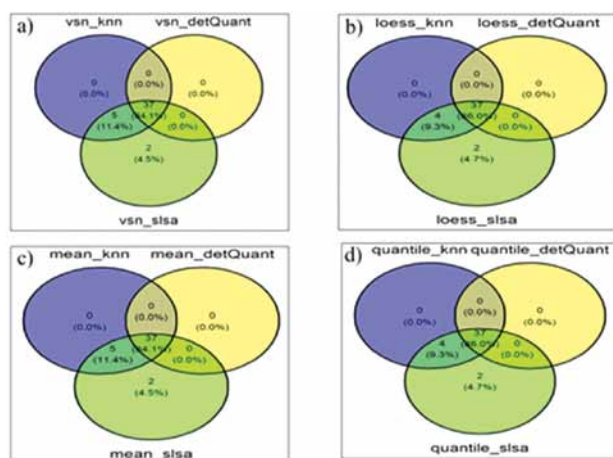
विवरण	वैल्यू (Value)
लाइन की संख्या	923
मिसिंग वैल्यू की संख्या	252
मिसिंग वैल्यू का प्रतिशत (%)	4.55
खाली लाइनों की संख्या	4

किसी एक स्थिति में अगर तीन डेटा missing हैं तो उन लाइन को हटा दिया गया है। इस प्रकार पहली 33 लाइनों को हटा दिया गया है। संदूषक को हटाने के लिए प्लस चिह्न (+) वाले "Contaminant" और "Reverse" कॉलम पर फ़िल्टर लागू किया गया। इस प्रक्रिया में 17 लाइनें हटा दी गई हैं। फ़िल्टर करने के बाद, डेटासेट में 890 लाइन बची।

हमने सामान्यीकरण और इंप्यूटेशन विधियों के विभिन्न संयोजनों का उपयोग किया है। चार सामान्यीकरण विधियों का उपयोग किया जाता है, अर्थात्, क्वांटाइल सेंटरिंग (Quantile Centering), मीन सेंटरिंग (Mean Centering), वेरिएंस स्टेबिलाइजिंग नॉर्मलाइजेशन (VSN) और लोकल वेटेड स्केटरप्लॉट स्मूथिंग (Quantile Centering - LOESS) नॉर्मलाइजेशन। missing मूल्यों को इंप्यूट करने के लिए तीन अलग-अलग इंप्यूटेशन तकनीकों का उपयोग किया गया, अर्थात्, एस.एल.एस.ए. (स्ट्रक्चर्ड लीस्ट स्क्वायर एडेप्टिव - Structured Least Square Adaptive), डेटक्वांटाइल (DetQuantile) और के-एन.एन. (k-NN). सामान्यीकरण और इंप्यूटेशन के बाद टी-टेस्ट (t-test) का उपयोग करके हाइपोथेसिस टेस्टिंग किया गया ताकि यह परीक्षण किया जा सके कि प्रोटीन दोनों स्थितियों के बीच डिफरेंसियली एबंडेंट/ एक्सप्रेस्ड है या नहीं। उन्ही प्रोटीन को दो स्थितियों के बीच डिफरेंसियली एबंडेंट में माना गया जिनका fold change ≥ 2 और adjusted p-value ≤ 0.05 है। सामान्यीकरण और इंप्यूटेशन तकनीकों के विभिन्न संयोजनों का उपयोग करके पहचाने गए डिफरेंसियली एक्सप्रेस्ड प्रोटीन की संख्या को तालिका 2 और चित्र 6 में दर्शाया गया है।

तालिका 2. सामान्यीकरण और इंप्यूटेशन तकनीकों के विभिन्न संयोजनों का उपयोग करके पहचाने गए डिफरेंसियली एक्सप्रेस्ड प्रोटीन की संख्या

क्र. सं.	सामान्यीकरण	इंप्यूटेशन	डिफरेंसियली एक्सप्रेस्ड प्रोटीन की संख्या	अप-रेगुलटेड प्रोटीन की संख्या	डाउन-रेगुलटेड प्रोटीन की संख्या
1	VSN	KNN	42	39	3
2	VSN	Det quantile	37	35	2
3	VSN	SLSA	44	40	4
4	LOESS	KNN	41	38	3
5	LOESS	Det quantile	37	35	2
6	LOESS	SLSA	43	39	4
7	Mean centering	KNN	42	39	3
8	Mean centering	Det quantile	37	35	2
9	Mean centering	SLSA	44	40	4
10	Quantile centering	KNN	41	38	3
11	Quantile centering	Det quantile	37	35	2
12	Quantile centering	SLSA	43	39	4



चित्र 6. डिफरेंसियली एक्सप्रेस्ड प्रोटीन के लिए वेन आरेख

सबसे अधिक डिफरेंसियली एक्सप्रेस्ड प्रोटीन VSN सामान्यीकरण और SLSA इंप्यूटेशन तकनीक का उपयोग करके पहचाने गए हैं। 37 प्रोटीन सभी संयोजन में पहचाने गये हैं।

निष्कर्ष

यह लेख प्रोटीओमिक्स और जैव सूचना विज्ञान के क्षेत्र में काम करने वाले शोधकर्ताओं के लिए उपयोगी

है। इसमें प्रोटीओमिक्स एबंडेंस डेटा का विश्लेषण करने के लिए अलग-अलग तरीकों का उल्लेख किया है। एक ऑनलाइन टूल "RepExplore" का उपयोग करके प्रोटीओमिक्स डेटा विश्लेषण का प्रदर्शन किया है। "ProStaR" टूल के द्वारा सामान्यीकरण और इंप्यूटेशन तकनीकों के विभिन्न संयोजनों का उपयोग करके तुलनात्मक विश्लेषण भी किया गया है। यह निष्कर्ष निकाला गया है कि किसी को प्रोटीओमिक्स डेटा विश्लेषण के लिए एक ही विधि पर निर्भर नहीं रहना चाहिए। भविष्य में, बेंचमार्क डेटासेट का उपयोग करके और अधिक तुलनात्मक विश्लेषण सम्मिलित करेंगे। इसके अतिरिक्त, इन तरीकों का उपयोग समान प्रयोगों के एक्सप्रेसन डेटा (जैसे, माइक्रोएरे और मेटाबॉलोमिक्स डेटा) का विश्लेषण करने के लिए किया जा सकता है।

संदर्भ

एंडरसन, एन. एल. एवं एंडरसन, एन. जी. (1998). प्रोटीन और प्रोटीओमिक्स: न्यू टैक्नीक, न्यू हायपोथेसिस एंड न्यू वर्ड्स, *इलेक्ट्रोफोरेसिस*, **19** (11), 1853–1861.

- विल्किंस, मा. (2009). प्रोटिओमिक्स डेटा माइनिंग”, प्रोटिओमिक्स की विशेषज्ञ समीक्षा, इंग्लैंड, **6(6)**, 599–603.
- करपीविच, वाई. वी., डबनी, ए. आर. एवं स्मिथ, आर. डी. (2012). लेबल-मुक्त LC-MS विश्लेषण के लिए सामान्यीकरण और मिसिंग वेल्यु इंप्यूटेशन, *बी.एम.सी. बायोइन्फोर्मेटिक्स*, **13 सप्ल 16**, S5।
- रुबिन, डी. बी. (1976). इंफेरेंस एवं मिसिंग डेटा, *बायोमेट्रिका*, **63 (3)**, 581–592.
- ग्लैब, ई. एवं शनाइडर, आर. (2015). RepExplore: प्रोटिओमिक्स और मेटाबॉलिकम डेटा विश्लेषण में तकनीकी रेप्लिकेट वैरियेंस को संबोधित करते हुए, *बायोइन्फोर्मेटिक्स*, **31 (13)**, 2235–7.
- गोएमीन, एल. जे. ई., गेवर्ट, के., एवं क्लेमेंट, एल. (2018). लेबल-मुक्त मात्रात्मक LC/MS प्रोटिओमिक्स में प्रयोगात्मक डिजाइन और डेटा-विश्लेषण: MSqRob के साथ एक ट्यूटोरियल, *ज. प्रोटिओमिक्स*, **171**, 23–36.
- चांग, सी., एट अल. (2018). PANDA-view: सांख्यिकीय विश्लेषण और मात्रात्मक प्रोटिओमिक्स डेटा के विजुआलाइजेशन के लिए एक उपयोग करने में आसान वाला टूल, *बायोइन्फोर्मेटिक्स*.
- चोई, एम., एट अल. (2014). MSstats: मात्रात्मक द्रव्यमान स्पेक्ट्रोमेट्री आधारित प्रोटिओमिक प्रयोगों के सांख्यिकीय विश्लेषण के लिए एक आर पैकेज, *बायोइन्फोर्मेटिक्स*, **30 (17)**, 2524–6.
- विएक्जोरेक, स., कॉंबेस, फ., लजार, सी., गियाई गियानेट्टों, क., गत्तो, ल., दोफ़र्, अ., हेससे, अ., कौते, य., फेररो, म., बुलेय, सी., एवं बर्गर, त. (2017). DAPAR और ProstaR: सॉफ्टवेर टु परफ़ोर्म स्टेटिस्कल एनालिसिस इन क्वॉंटिटेटिव डिस्कवरी प्रोटिओमिक्स, *बायोइन्फोर्मेटिक्स*, **33(1)**, 135–136.

“

हर्ष और आनंद से परिपूर्ण जीवन,
केवल ज्ञान और विज्ञान के आधार पर संभव है।

”

—सर्वपल्ली राधाकृष्णन

स्थायित्व विश्लेषण- अप्राचालिक दृष्टिकोण

प्रकाश कुमार, रंजीत कुमार पॉल, अमृत कुमार पॉल, राजू कुमार, राजीव रंजन कुमार,
मृन्मय राय एवं मो. यासीन

भा.कृ.अनु.प.—भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली—110012

1. परिचय

फसल सुधार की गतिविधियों की सफलता बहुत हद तक बड़े पैमाने पर उत्पादन के लिए बेहतर किस्मों की पहचान पर निर्भर करती है। एक जीनोटाइप बेहतर माना जा सकती है यदि यह अनुकूल वातावरण के तहत उच्च उपज क्षमता और एक ही समय में प्ररूपी स्थायित्व है। फेनोटाइप, जीनोटाइप (G), पर्यावरण (E) एवं जी. ई. इंटरैक्शन (जी.ई.आई.) का एक मिश्रण है। स्थिरता की अवधारणा को समझने के लिए शोधकर्ता अनुकूलन, प्ररूपी स्थायित्व और उपज स्थायित्व जैसे पदों का उपयोग करते हैं (बेकर और लीओन, 1988)।

यहाँ पद अनुकूलन के बारे में दावोलकर (1999) ने विस्तृत चर्चा की है। सभी जीवित वस्तुओं में शारीरिक समायोजन होता है जो उन्हें अपने तात्कालिक परिवेश में उतार-चढ़ाव से सामना करने की अनुमति देता है। ये समायोजन स्वयं से अनुकूलन के रूप में जाना जाता है। अनुकूलन एक जीनोटाइप का लक्षण है जो चयन के तहत अपने अस्तित्व के लिए अनुमति देती है। सायमंड्स (1981) ने अनुकूलन के दो पहलू दिए हैं।

1. विशिष्ट जिनोटाइपिक अनुकूलन: यह एक सीमित पर्यावरण के लिए अनुकूलन के करीब है।
2. सामान्य जिनोटाइपिक अनुकूलन: यह एक विभिन्न वातावरण के लिए अनुकूलन है।

ब्रीडर / किसान के अंतिम लक्ष्य और विचाराधीन विशेषता के लिए स्थायित्व की दो अवधारणा हैं, जैविक और कृषिशास्त्रीय अवधारणा, इन्हें क्रमशः स्थिर और गतिशील अवधारणा (लीओन, 198) भी कहते हैं।

जैविक अवधारणा के तहत एक स्थिर जीनोटाइप जिसका फेनोटाइप है, जो कि विशेषता के स्तर से कम

विचलन दिखता है, जब जीनोटाइप का परीक्षण एक संख्या से अधिक वातावरण में करते हैं। अलग-अलग वातावरण में स्थिर प्रदर्शन के कारण, इस अवधारणा को स्थायित्व की अवधारणा भी कहा जाता है।

जैविक अवधारणा में अनाज की उपज को देखते हैं। एक स्थिर जीनोटाइप अनुकूल वातावरण के तहत इतनी अच्छी पैदावार नहीं देता जितना कि प्रतिकूल वातावरण में भी औसतन अच्छी पैदावार करता है।

सारणी 2: स्थायित्व माप और उनके अंतर्निहित स्थायित्व अवधारणा

स्थायित्व माप	अंतर्निहित स्थायित्व अवधारणा
Environmental variance	जैविक
Ecovalence	कृषिशास्त्रीय
Stability variance	कृषिशास्त्रीय
Regression coefficient	जैविक एवं कृषिशास्त्रीय
Deviation mean square	कृषिशास्त्रीय
Coefficient of determination	कृषिशास्त्रीय
Hanson's stability Measure	कृषिशास्त्रीय
Huhn's Measures	जैविक / कृषिशास्त्रीय
Kang's rank-sum	कृषिशास्त्रीय
Ketata's ranking sum	कृषिशास्त्रीय

प्रचलन में उपज स्थायित्व अधिकांश कृषिशास्त्रीय अवधारित है। जैविक अवधारणा के लिए केवल दो माप उपलब्ध हैं और वे पर्यावरणीय विचरण $S_{Y_i}^2$ और विविधता का पर्यावरण गुणांक (CV_i) है।

$$S_{Y_i}^2 = \sum_j (Y_{ij} - \bar{Y}_i.)^2 / (s-1) \quad (1)$$

$$CV_i = (S_{Y_i} / \bar{Y}_i.) \times 100 \quad (2)$$

इनमें से कम मान वाला सर्वथा उच्च स्थायित्व दिखता है। यद्यपि जो मापांक सैद्धांतिक रूप से अधिक अच्छा हो निम्नलिखित कारण से व्यावहारिकता में उपयोगी नहीं होता है (i) जैविक अवधारणा के तहत स्थायित्व आमतौर पर अपेक्षाकृत कम उपज दिखता है (ii) उच्च स्तरीय उपज वाले जीनोटाइप को वृहत वातावरण की सीमा में अमल में लाना कठिन है।

स्थायित्व विश्लेषण को दो तरीकों, प्राचालिक और अप्राचालिक समूहों में वर्गीकृत करते हैं (हुन, 1979)। प्राचालिक दृष्टिकोण जीनोटाइप, पर्यावरण और GEI प्रभाव के वितरण के बारे में सांख्यिकीय मान्यताओं पर आधारित है। लेकिन, अप्राचालिक कोई विशेष सांख्यिकीय मान्यताओं पर आधारित नहीं है।

अप्राचालिक स्थायित्व विधि, प्राचालिक स्थायित्व विधि से कई मामले में अच्छी है। पर्यवेक्षित मान के वितरण के बारे में कोई मान्यताओं की आवश्यकता नहीं है। यह आउटलायर से उत्पन्न पक्षपात को कम करता है एवं उपयोग तथा व्याख्या करने के लिए आसान है और एक या अधिक जीनोटाइप को हटाने एवं जोड़ने

से परिणामों में अधिक भिन्नता नहीं देता है।

यह सच है कि अप्राचालिक विधि, प्राचालिक विधि समकक्ष से कम शक्तिशाली है, लेकिन जब जीनोटाइप की संख्या अधिक हो तो अप्राचालिक विधि की क्षमता प्राचालिक विधि की क्षमता के अधिक समकक्ष होती है (रिगर और प्रभाकरण, 2000)। स्थायित्व मापने के लिए बहुत सारी विधियाँ हैं। प्रजनन कार्यक्रमों / किस्म के चयन में जीनोटाइप परीक्षण के लिए कुछ में जोड़ने से अच्छा और उचित estimators के लिए अप्राचालिक विधि एवं प्राचालिक विधि के बीच सांख्यिकीय संबंधों का अध्ययन करने के लिए आवश्यक है।

2. सामग्री और विधियाँ

इस अध्ययन में उपयोग किए आंकड़े आंध्र प्रदेश के विभिन्न कृषि जलवायु क्षेत्रों में स्थित अनुसंधान केंद्रों पर आयोजित बहु स्थान साल परीक्षणों से एकत्र किए गए थे। प्रायोगिक लेआउट (layout) 3 प्रतिकृतियों के साथ यादृच्छिक खण्ड अभिकल्पना में किया गया था। आंकड़ों की आपूर्ति क्षेत्रीय कृषि अनुसंधान स्टेशन (RARs), Palem, AN-GRAU, आंध्र प्रदेश के द्वारा की गई है। फली पैदावार प्रति हेक्टेयर किलो के रूप में व्यक्त किया गया। 10 जीनोटाइप और 12 वातावरण के मध्य आंकड़े (तालिका 1) में प्रस्तुत है।

तालिका 1: 10 जीनोटाइप एवं 12 वातावरण के मध्य उपज आंकड़े

	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12
G1	1773	880	2841	2020	856	1382	1458	282	1190	1001	2708	1832
G2	1715	861	2497	2020	505	1104	1153	275	1394	882	1956	1907
G3	1241	424	3266	1717	1148	1225	1130	113	701	705	1688	1568
G4	1472	917	3172	2222	1505	1475	1222	632	1308	334	2833	1157
G5	1208	1435	3625	1919	903	1432	921	862	1081	539	2303	1778
G6	1893	1310	2716	2374	1320	1476	1482	680	1498	591	2877	2333
G7	1852	1169	2527	2222	903	1220	1407	455	1637	521	2042	1732
G8	1266	993	2245	1869	292	972	1171	275	1419	767	2184	2037
G9	1736	792	2376	2172	981	1113	1051	364	1579	364	2940	1500
G10	1442	695	2800	2071	1051	1890	1051	605	1684	67	2083	1419

3. नास्सर, एवं हुन द्वारा जीनोटाइप के आधार पर प्ररूपी स्थायित्व के कुछ अप्राचालिक विधियों के विवरण

3.1. प्रत्येक वातावरण में प्रसरण रैंक (Si⁽²⁾)

$$S_i^{(2)} = \frac{\sum_{j=1}^E (r_{ij} - \bar{r}_i)^2}{(E-1)}$$

जहाँ E वातावरण की संख्या है,

\bar{r}_i वातावरण में रैंक (rank) का माध्य है, r_{ij} प्रत्येक वातावरण में इंटरैक्शन रेसिडूअल (Intraction re-schedule) V_{ij} पर आधारित जीनोटाइप का रैंक है, जहाँ रैंक कम से उच्चतम की ओर आवंटित है।

3.2. \bar{r}_i के प्रति इकाई में जीनोटाइप के निरपेक्ष रैंक के अंतर का योग (Si⁽³⁾)

$$S_i^{(3)} = \frac{\sum_{j=1}^E |r_{ij} - \bar{r}_i|}{\bar{r}_i}$$

जहाँ E वातावरण की संख्या है, \bar{r}_i वातावरण में रैंक का माध्य है, r_{ij} प्रत्येक वातावरण में इंटरैक्शन रेसिडूअल V_{ij} पर आधारित i^{th} जीनोटाइप एवं j^{th} वातावरण का रैंक है, जहाँ रैंक कम से उच्चतम की ओर आवंटित है।

3.3. प्रत्येक वातावरण में रैंक में मानक विचलन (Si⁽⁴⁾)

$$S_i^{(4)} = \sqrt{\frac{\sum_{j=1}^E (r_{ij} - \bar{r}_i)^2}{E}}$$

जहाँ E वातावरण की संख्या है, \bar{r}_i वातावरण में रैंक का माध्य है, r_{ij} प्रत्येक वातावरण में इंटरैक्शन रेसिडूअल V_{ij} पर आधारित i^{th} जीनोटाइप एवं j^{th} वातावरण का रैंक है, जहाँ रैंक कम से उच्चतम की ओर आवंटित है।

3.4. प्रत्येक वातावरण में रैंक में माध्य विचलन (Si⁽⁵⁾)

$$S_i^{(5)} = \frac{\sum_{j=1}^E |r_{ij} - \bar{r}_i|}{E}$$

जहाँ E वातावरण की संख्या है, \bar{r}_i वातावरण में रैंक का माध्य है, r_{ij} उपज के माध्य पर आधारित i^{th} जीनोटाइप एवं j^{th} वातावरण का रैंक है।

3.5. \bar{r}_i के प्रति इकाई में प्रसरण रैंक (Si⁽⁶⁾)

$$S_i^{(6)} = \frac{\sum_{j=1}^E (r_{ij} - \bar{r}_i)^2}{\bar{r}_i}$$

जहाँ E वातावरण की संख्या है, \bar{r}_i वातावरण में रैंक का माध्य है, r_{ij} प्रत्येक वातावरण में इंटरैक्शन रेसिडूअल V_{ij} पर आधारित i^{th} जीनोटाइप एवं j^{th} वातावरण का रैंक है, जहाँ रैंक कम से उच्चतम की ओर आवंटित है।

4. रिक्की एकोवालेंस (ecovalence)

इंटरैक्शन वर्गों के योग के लिए जीनोटाइप के योगदान से सरल और गणना करने के लिए आसान, एकोवालेंस मापांक (W_i) मिलता है

$$W_i = \sum_j (X_{ij} - \bar{Y}_i - \bar{Y}_j + \bar{Y}_{..})^2 = \sum_j \hat{g}_{ij}^2 \quad (3.3)$$

कम ecovalence मान वाला जीनोटाइप, उपज-स्थायित्व की दृष्टि से आदर्श माना जाता है।

5. शुक्ला का स्थायित्व प्रसरण (Shukla's stability variance measure)

रेसिडूअल $g_{ij} + \bar{e}_{ij}$ के विचरण का अनुमान $\hat{\sigma}_i^2$, i^{th} जीनोटाइप की स्थायित्व का एक उपयोगी सूचक होता है। $\hat{\sigma}_i^2$ को स्थायित्व विचरण के रूप में परिभाषित किया गया है:

$$\hat{\sigma}_i^2 = \frac{t}{(s-1)(t-2)} W_i - \frac{MS(GE)}{(t-2)} \quad (3.4)$$

जहाँ W_i , (3.3) द्वारा परिभाषित किया गया है और MS(GE) जी×ई इंटरैक्शन का माध्य वर्ग है,

$$[MS(GE) = \sum_{i,j} \hat{\sigma}_{ij}^2 / (s-1)(t-1)]$$

यह सांख्यिकी (Statistic) W_i का रैखिक संयोजन है इसीलिए रैंकिंग के लिए दोनों मापांक W_i और $\hat{\sigma}_i^2$ बराबर है।

6. कांग रैंक-योग (Kang Rank-Sum(RS))

कांग रैंक-योग एक अप्राचालिक विधि है, जहाँ उपज एवं शुक्ला का प्रसरण दोनों को लिया जाता है। यहाँ उपज एवं स्थायित्व दोनों के लिए एक ही भार देते हैं ताकि हमें उच्च उपज के साथ-साथ स्थायित्व जीनोटाइप भी मिल सके। इस विधि में, उच्चतम उपज जीनोटाइप और सबसे कम स्थायित्व प्रसरण वाले जीनोटाइप दोनों को रैंक 1 देते हैं उसके बाद सभी जीनोटाइप के लिए दोनों मानकों के रैंकों पर आधारित योग निकालते हैं। (Akcura & Kaya, 2008) के अनुसार प्रत्येक जीनोटाइप से कम रैंक वाले रैंक-योग (RS) सबसे अधिक वांछनीय जीनोटाइप माना जाता है।

$RS_i =$ माध्य उपज की रैंक + σ_i^2 की रैंक

यहाँ रैंकिंग के लिए हम शुक्ला की स्थायित्व प्रसरण प्राचाल; (σ_i^2) को Wricke ecovalence (W_i) के सामान मानते हैं क्योंकि शुक्ला का स्थायित्व प्रसरण प्राचाल; (σ_i^2), Wricke ecovalence (W_i) के रेखीय संयोजन से बना होता है।

उदाहरण

सबसे पहले हम ऊपर लिए गए आंकड़ों से MS Excel में $V_{ij} = (Y_{ij} - \bar{Y}_i - \bar{Y}_j + \bar{Y}_{..})$ का मान निकालते हैं। तदुपरान्त, Wricke ecovalence (W_i) का मान निकालते हैं, उसके बाद प्रत्येक जीनोटाइप का Wricke ecova-lence (W_i) पर आधारित रैंक निकलते हैं जिस में अधिकतम उपज-माध्य मान वाले जीनोटाइप को अधिकतम रैंक देते हैं।

प्रत्येक जीनोटाइप का उपज के माध्य मान पर आधारित रैंक निकालते हैं जिसमें अधिकतम उपज-माध्य मान वाले जीनोटाइप को न्यूनतम रैंक देते हैं, उसके बाद दोनों मानकों से निकाले गये रैंको का रैंकयोग निकालते हैं।

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
34		E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	Y _i	
35	G1	1773	880	2841	2020	856	1382	1458	282	1190	1001	2708	1832	1518.583333	
36	G2	1715	861	2497	2020	505	1104	1153	275	1394	882	1956	1907	1355.75	
37	G3	1241	424	3266	1717	1148	1225	1130	113	701	705	1688	1568	1243.833333	
38	G4	1472	917	3172	2222	1505	1475	1222	632	1308	334	2833	1157	1520.75	
39	G5	1208	1435	3625	1919	903	1432	921	862	1081	539	2303	1778	1500.5	
40	G6	1893	1310	2716	2374	1320	1476	1482	680	1498	591	2877	2333	1712.5	
41	G7	1852	1169	2527	2222	903	1220	1407	455	1637	521	2042	1732	1473.916667	
42	G8	1266	993	2245	1869	292	972	1171	275	1419	767	2184	2037	1290.833333	
43	G9	1736	792	2376	2172	981	1113	1051	364	1579	364	2940	1500	1414	
44	G10	1442	695	2800	2071	1051	1890	1051	605	1684	67	2083	1419	1404.833333	
45	Y _j	1560	947.6	2807	2060.6	946.4	1328.9	1204.6	454.3	1349.1	577.1	2361.4	1726.3	1443.55	
46															
47	$V_{ij} = (Y_{ij} - \bar{Y}_i - \bar{Y}_j + \bar{Y}_{..})$														
48	G1	=B35-\$N35-\$B\$45+\$N\$45				-165.4	-21.93	178.37	-247	-234.1	348.9	271.57	30.667	426501.2867	2
49	G2	243	1.2	-222	47.2	-353.6	-137.1	36.2	-91.5	132.7	392.7	-317.6	268.5	608725.82	4
50	G3	-119	-323.9	659.2	-143.9	401.3	95.817	125.12	-142	-448.4	327.6	-473.7	41.417	1314757.737	9
51	G4	-165	-107.8	288.3	84.2	481.4	68.9	-59.8	100.5	-118.3	-320	394.4	-646.5	1069322.42	8
52	G5	-409	430.5	761.6	-198.6	-100.4	46.15	-340.6	350.8	-325.1	-95.1	-115.4	-5.25	1350969.27	10
53	G6	64.25	93.45	-359	44.45	104.7	-121.9	8.45	-43.3	-120.1	-255	246.65	337.75	426155.87	1
54	G7	261.8	191	-310	131.03	-73.77	-139.3	172.03	-29.7	257.53	-86.5	-349.8	-24.67	470294.8867	3
55	G8	-141	198.1	-409	-38.88	-501.7	-204.2	119.12	-26.6	222.62	342.6	-24.68	463.42	918351.5367	6
56	G9	205.8	-126.1	-401	140.95	64.15	-186.4	-124.1	-60.8	259.45	-184	608.15	-196.8	806331.87	5
57	G10	-79.1	-213.9	32.22	49.117	143.3	599.82	-114.9	189.4	373.62	-471	-239.7	-268.6	976223.9367	7
58															

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	
1		E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	Mean	Rank of mean W _i	Rank of W _{RS}				
2	G1	1773	880	2841	2020	856	1382	1458	282	1190	1001	2708	1832	1518.58333	=RANK(N2,\$N\$2:\$N\$11,0)			5	2	
3	G2	1715	861	2497	2020	505	1104	1153	275	1394	882	1956	1907	1355.75	RANK(number,ref,[order])		4	12	6	
4	G3	1241	424	3266	1717	1148	1225	1130	113	701	705	1688	1588	1243.83333		10	2.6E+07	9	19	10
5	G4	1472	917	3172	2222	1505	1475	1222	632	1308	334	2833	1157	1520.75		2	3.5E+07	8	10	4
6	G5	1208	1435	3625	1919	903	1432	921	862	1081	539	2303	1778	1500.5		4	3.3E+07	10	14	7
7	G6	1893	1310	2716	2374	1320	1476	1482	680	1498	591	2877	2333	1712.5		1	3.9E+07	1	2	1
8	G7	1852	1169	2527	2222	903	1220	1407	455	1637	521	2042	1732	1473.91667		5	2.8E+07	3	8	3
9	G8	1266	993	2245	1869	292	972	1171	275	1419	767	2184	2037	1290.83333		9	2.4E+07	6	15	9
10	G9	1736	792	2376	2172	981	1113	1051	364	1579	364	2940	1500	1414		6	2.9E+07	5	11	5
11	G10	1442	695	2800	2071	1051	1890	1051	605	1684	67	2083	1419	1404.83333		7	2.9E+07	7	14	7
12																				

निकाले गये रैंक-योग (RS) का उपयोग वांछनीय जीनोटाइप का चयन करने में करते हैं।

निष्कर्ष:

प्रत्येक जीनोटाइप में से सबसे कम रैंक वाले रैंक-योग (RS) को सबसे अधिक वांछनीय जीनोटाइप माना जाता है।

7. केतता की रैंकिंग योग विधि (Ketata's ranking sum methods)

Ketata (1988) ने एक रैंकिंग विधि दी है जिस में जीनोटाइप की रैंक सभी प्रकार के वातावरण में अनाज की उपज पर आधारित होती है। प्रत्येक जीनोटाइप की रैंक उपज के माध्य एवं मानक विचलन पर आधारित होती है। इस विधि में जो जीनोटाइप अधिकतम प्रदर्शन करता है उसका रैंक 1 लेते हैं और यदि एक जीनोटाइप का प्रदर्शन माध्य रैंक 1 के करीब एवं कम रैंक के मानक विचलन के करीब हो उसे सबसे अधिक स्थायी किस्म के रूप में जाना जाता है।

8. संशोधित स्थायित्व मापांक

यहाँ स्थायित्व मापांक निकालने के लिए दो मापांकों का उपयोग करते हैं। पहला एक जीनोटाइप का सभी वातावरण से रैंक-योग जो कि इंटरैक्शन रेसिडूअल

$$\hat{V}_{ij} = (Y_{ij} - \bar{Y}_i - \bar{Y}_j + \bar{Y}) \text{ or } (Y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j)$$

पर आधारित हो एवं दूसरा प्रत्येक वातावरण में उसी जीनोटाइप प्रसरण जो कि इंटरैक्शन रेसिडूअल

$$\hat{V}_{ij} = (Y_{ij} - \bar{Y}_i - \bar{Y}_j + \bar{Y}) \text{ or } (Y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j)$$

पर आधारित हो, दोनों को लेते हैं। एक बार पुनः रैंक-योग का रैंक जिस में अधिकतम मान वाले जीनोटाइप को अधिकतम रैंक देते हैं एवं इसी प्रकार जीनोटाइप प्रसरण के लिए भी रैंकिंग करते हैं जिस में अधिकतम मान वाले जीनोटाइप को अधिकतम रैंक देते हैं। दोनों मापांक का रैखिक संयोजन करके संशोधित स्थायित्व मापांक विकसित करते हैं जिसे रैंक आधारित स्टेबिलिटी इंडेक्स (RSI) से चिन्हित करते हैं।

$$RSI = ((rRS + rSi(2))/2) \text{ का रैंक,}$$

उदाहरण

सबसे पहले हम ऊपर लिए गए आंकड़ों से MS Excel में $V_{ij} = (Y_{ij} - \bar{Y}_i - \bar{Y}_j + \bar{Y})$ का मान निकालते हैं। $V_{ij} = (Y_{ij} - \bar{Y}_i - \bar{Y}_j + \bar{Y})$ उसके बाद का निरपेक्ष मान निकालते हैं। प्रत्येक जीनोटाइप का सभी वातावरण में $V_{ij} = (Y_{ij} - \bar{Y}_i - \bar{Y}_j + \bar{Y})$ का निरपेक्ष मान के आधार पर रैंकिंग करते हैं, उसके बाद $V_{ij} = (Y_{ij} - \bar{Y}_i - \bar{Y}_j + \bar{Y})$ का निरपेक्ष मान के आधार पर प्रत्येक जीनोटाइप का सभी वातावरण में प्रसरण निकाल कर रैंकिंग करते हैं।

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	
1	V_i														
2		1773	880	2841	2020	856	1382	1458	282	1190	1001	2708	1832	1518.583	
3		1715	861	2497	2020	505	1104	1153	275	1394	882	1956	1907	1355.75	
4		1241	424	3266	1717	1148	1225	1130	113	701	705	1688	1568	1243.833	
5		1472	917	3172	2222	1505	1475	1222	632	1308	334	2833	1157	1520.75	
6		1208	1435	3625	1919	903	1432	921	862	1081	539	2303	1778	1500.5	
7		1893	1310	2716	2374	1320	1476	1482	680	1498	591	2877	2333	1712.5	
8		1852	1169	2527	2222	903	1220	1407	455	1637	521	2042	1732	1473.917	
9		1266	993	2245	1869	292	972	1171	275	1419	767	2184	2037	1290.833	
10		1736	792	2376	2172	981	1113	1051	364	1579	364	2940	1500	1414	
11		1442	695	2800	2071	1051	1890	1051	605	1684	67	2083	1419	1404.833	
12		1559.8	947.6	2806.5	2060.6	946.4	1328.9	1204.6	454.3	1349.1	577.1	2361.4	1726.3	1443.59	
13	V_j														
14		=B2-\$N2-B\$12+\$N\$12			-115.633	-165.433	-21.9333	178.3667	-247.333	-234.133	348.8667	271.5667	30.66667		
15		243	1.2	-221.7	47.2	-353.6	-137.1	36.2	-91.5	132.7	392.7	-317.6	268.5		
16		-119.083	-323.883	659.2167	-143.883	401.3167	95.81667	125.1167	-141.583	-448.383	327.6167	-473.683	41.41667		
17		-165	-107.8	288.3	84.2	481.4	68.9	-59.8	100.5	-118.3	-320.3	394.4	-646.5		
18		-408.75	430.45	761.55	-198.55	-100.35	46.15	-340.55	350.75	-325.05	-95.05	-115.35	-5.25		
19		64.25	93.45	-359.45	44.45	104.65	-121.85	8.45	-43.25	-120.05	-255.05	246.65	337.75		
20		261.8333	191.0333	-309.867	131.0333	-73.7667	-139.267	172.0333	-29.6667	257.5333	-86.4667	-349.767	-24.6667		
21		-141.083	198.1167	-408.783	-38.8833	-501.683	-204.183	119.1167	-26.5833	222.6167	342.6167	-24.6833	463.4167		
22		205.75	-126.05	-400.95	140.95	64.15	-186.35	-124.05	-60.75	259.45	-183.55	608.15	-196.75		
23		-79.0833	-213.883	32.21667	49.11667	143.3167	599.8167	-114.883	189.4167	373.6167	-471.383	-239.683	-268.583		
24															

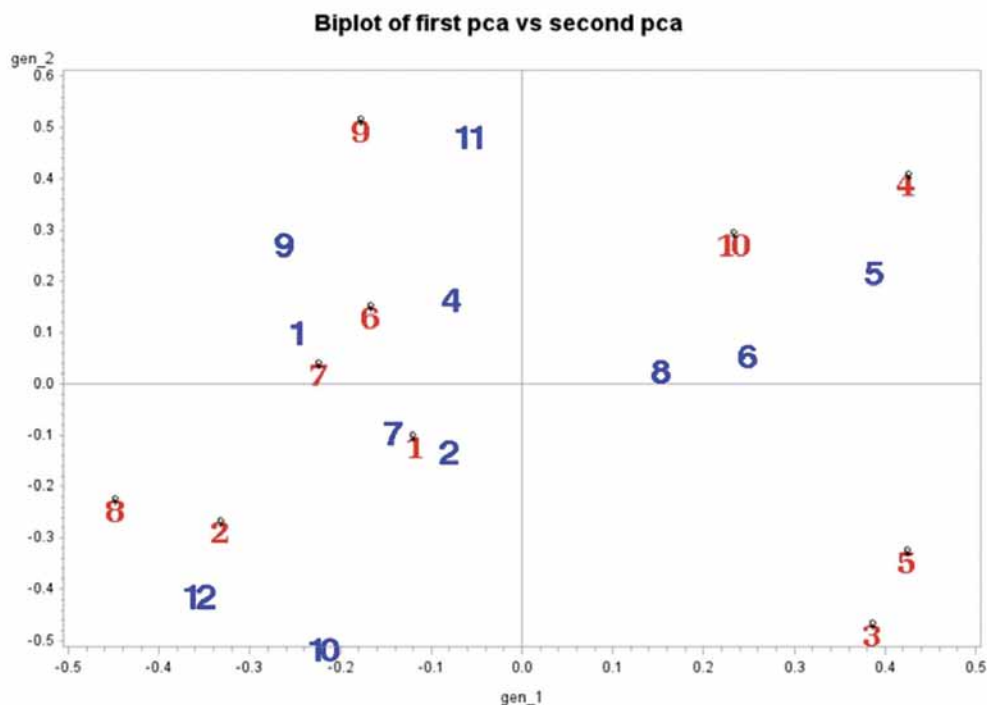
RSI= rank of $((rRS + rSi(2))/2)$ का मान निकालते हैं।

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
23	G10	-79.0833	-213.883	32.21667	49.11667	143.3167	599.8167	-114.883	189.4167	373.6167	-471.383	-239.683	-268.583								
24	$ V_j $																				
25	G1	138.1667	142.6333	40.53333	115.6333	165.4333	21.93333	178.3667	247.3333	234.1333	348.8667	271.5667	30.66667								
26	G2	243	1.2	221.7	47.2	353.6	137.1	36.2	91.5	132.7	392.7	317.6	268.5								
27	G3	119.0833	323.8833	659.2167	143.8833	401.3167	95.81667	125.1167	141.5833	448.3833	327.6167	473.6833	41.41667								
28	G4	165	107.8	288.3	84.2	481.4	68.9	59.8	100.5	118.3	320.3	394.4	646.5								
29	G5	408.75	430.45	761.55	198.55	100.35	46.15	340.55	350.75	325.05	95.05	115.35	5.25								
30	G6	64.25	93.45	359.45	44.45	104.65	121.85	8.45	43.25	120.05	255.05	246.65	337.75								
31	G7	261.8333	191.0333	309.8667	131.0333	73.76667	139.2667	172.0333	29.66667	257.5333	86.46667	349.7667	24.66667								
32	G8	141.0833	198.1167	408.7833	38.88333	501.6833	204.1833	119.1167	26.58333	222.6167	342.6167	24.68333	463.4167								
33	G9	205.75	126.05	400.95	140.95	64.15	186.35	124.05	60.75	259.45	183.55	608.15	196.75								
34	G10	79.08333	213.8833	32.21667	49.11667	143.3167	599.8167	114.8833	189.4167	373.6167	471.3833	239.6833	268.5833								
35	Ranked $ V_j $																				
36	G1	4	5	2	6	6	1	9	9	5	8	5	3		RS	r_{RS}	S_i^2	r_{fact}	$((rRS + rSiRSI)$		
37	G2	8	1	3	3	7	6	2	5	3	9	6	6		63	4	=VAR(B25:M25)	2.5	2		
38	G3	3	9	9	9	8	4	7	7	10	6	9	4		59	2	=VAR(number1,number2,...)	3	4		
39	G4	6	3	4	5	9	3	3	6	1	5	8	10		85	10	36973.42	9	9.5	9	
40	G5	10	10	10	10	3	2	10	10	8	2	2	1		63	4	36305.88	8	6	5	
41	G6	1	2	6	2	4	5	1	3	2	4	4	8		78	9	46312.26	10	9.5	9	
42	G7	9	6	5	7	2	7	8	2	6	1	7	2		42	1	14215.08	3	2	1	
43	G8	5	7	8	1	10	9	5	1	4	7	1	9		62	3	11629.39	2	2.5	2	
44	G9	7	4	7	8	1	8	6	4	7	3	10	5		67	6	28595.54	6	6	5	
45	G10	2	8	1	4	5	10	4	8	9	10	3	7		70	7	23774.58	5	6	5	
46															71	8	30409.56	7	7.5	8	

निष्कर्ष:

प्रत्येक जीनोटाइप में से सबसे कम रैंक आधारित स्टेबिलिटी इंडेक्स (RSI) का मान वांछित परिणाम देता है।

स्थायित्व की अप्राचालिक विधियों से प्राप्त परिणाम की तुलना के लिए हम उसी डेटासेट (dataset) में योगात्मक मुख्य प्रभाव और गुणक अंतःक्रिया (एएमएमआई) मॉडल से बाइप्लॉट विश्लेषण करने पर परिणाम संतोषप्रद मिलता है।



संदर्भ

दावोलकर, ए. आर. (1999). एलिमेंट्स ऑफ बायोमेट्रिकल जेनेटिक्स। कान्सेप्ट प्रकाशन कंपनी, नई दिल्ली, 373–374।

नास्सर, आर., एवं हुन, एम. (1987). स्टडीज ओन एस्टीमेशन ऑफ फिनोटाइपिक स्टेबिलिटी टेस्ट ऑफ सिग्निफिकेंस ऑफ नॉन-पैरामीट्रिक मेजर्स ऑफ फिनोटाइपिक स्टेबिलिटी। बायोमेट्रिक्स, **43**, 45–53।

कांग, एम. एस. (1988). ए रैंक-सम मेथड फॉर सेलेक्टिंग हाई इल्लिडिंग, स्टेबल कॉर्न जीनोटाइप। सिरियल रिसर्च कम्युनिकेशन, **16**, 113–115।

राजू, बी. एम. के. एवं भाटिया, वी. के. (2003). कोम्परिजन ऑफ वेरियस मेजर्स ऑफ स्टेबिलिटी

विथ रेस्पेक्ट टू रैंकिंग एबिलिटी अंडर वर्सिंग सिचुएशन। जर्नल ऑफ इंडियन सोसाइटी ऑफ एग्रीकल्चरल स्टेटिस्टिक्स, **56 (3)**, 276–293।

शुक्ला, जी. के. एच. (1972). सम स्टैटिस्टिकल आस्पेक्ट ऑफ पार्टिशनिंग जीनोटाइप-एनवायरनमेंट कंपोनेंट्स ऑफ वरिअबिलिटी। हेरेडिटी, **29**, 237–245।

राव, आर. (1997). जेनेटिक पैरामीटर्स के बेहतर आकलन के लिए कुछ योगदान। अप्रकाशित पी.चडी थीसिस, पी जी स्कूल, भा.कृ.अ.प., नई दिल्ली।

रिक्की, जी. (1962). क्षेत्र में अनुसंधान में जैविक विविधता को समझने की एक विधि जेड. पफलांजेनजुक्त, **47**, 92–46।

गार्च और एस वी आर के संयोजित मॉडल का उपयोग करके कृषि कमोडिटी की कीमतों में अस्थिरता का पूर्वानुमान

रंजीत कुमार पॉल, तनिमा दास, अमृत कुमार पॉल, प्रकाश कुमार, मोहम्मद यासीन एवं सत्य पाल सिंह
भा.कृ.अनु.प.—भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली

सार

कृषि वस्तुओं (commodities) की कीमतों में उतार-चढ़ाव का पूर्वानुमान लगाना आजकल बहुत महत्वपूर्ण हो गया है। मूल्य श्रृंखला में अस्थिरता को पकड़ने में पारंपरिक पैरामीट्रिक मॉडल का उपयोग अक्षम पाया गया है। इस संदर्भ में, मशीन लर्निंग (एमएल) तकनीक जैसे सपोर्ट वेक्टर रिग्रेशन (एसवीआर) को पूर्वानुमान की सटीकता में सुधार के लिए लागू किया जा सकता है। वर्तमान जांच में, पैरामीट्रिक अरेखीय टाइम सीरीज़ मॉडल अर्थात् सामान्यीकृत ऑटोरेग्रेसिव कंडीशनल हेटेरोसेडैस्टिक (गार्च) मॉडल और पर्यवेक्षित एमएल के संयोजन पर आधारित एक एल्गोरिथ्म है। एस वी आर का उपयोग भारत के दो प्रमुख बाजारों दिल्ली और कोलकाता में प्याज की कीमतों में उतार-चढ़ाव के पूर्वानुमान के लिए किया गया है। गार्च मॉडल की तुलना में प्रस्तावित एल्गोरिथ्म के आउटपरफॉर्मेंस को रूटमीनस्क्वायर एरर (आरएमएसई), मीन एब्सोल्यूटएरर (एमएई) और R^2 लॉग के माध्यम से तुलनात्मक प्रदर्शन किया गया है।

प्रमुख शब्द: गार्च, एसवीआर, प्याज की कीमत, अस्थिरता।

परिचय

इस अध्ययन में भारत के दो प्रमुख बाजारों, दिल्ली और कोलकाता में प्याज की कीमत में उतार-चढ़ाव का पूर्वानुमान लगाने के लिए मशीन लर्निंग तकनीक (एसवीआर) के साथ नॉनलाइनियर वोलैटिलिटी मॉडल (गार्च) के संयोजन पर आधारित एक एल्गोरिदम बनाया गया है। ताकि, सटीक और समय पर पूर्वानुमान से किसानों को आय बढ़ाने में मदद मिलेगी।

विभिन्न महत्वपूर्ण कृषि वस्तुओं में मुख्य रूप से खराब होने वाले उत्पाद अस्थिरता को दर्शाते हैं। इनमें से प्याज एक बहुत ही महत्वपूर्ण ताजी सब्जी है जिसका सेवन दुनिया भर में साल भर किया जाता है। दुनिया भर में भारत, प्याज की फसल के क्षेत्र में पहले और चीन के बाद उत्पादन में दूसरे स्थान पर है। अस्थिर मौसम और उत्पादन में उतार-चढ़ाव बाजार मूल्य में उतार-चढ़ाव के प्रमुख कारक हैं। अस्थिरता को पकड़ने के लिए सबसे व्यापक रूप से इस्तेमाल किए जाने वाले पैरामीट्रिक नॉनलाइनियर टाइम-सीरीज़ मॉडल ऑटोरेग्रेसिव कंडीशनल हेटेरोसेडैस्टिक (गार्च) मॉडल (1) और सामान्यीकृत आर्च (गार्च) मॉडल (2) हैं। किसी एक मॉडल की कमियों को दूर करने के लिए, विभिन्न संकर मॉडलों का उपयोग प्रस्तावित किया गया है (3, 4)। पारंपरिक समय-श्रृंखला मॉडल पर "मशीन लर्निंग तकनीक" की श्रेष्ठता कई शोधपत्रों (5, 6, 7) में साबित हुई है। वित्तीय समय श्रृंखला की भविष्यवाणी के लिए सपोर्ट वेक्टरमशीन (एसवीएम) और तंत्रिका नेटवर्क के प्रदर्शन की तुलना पहले की गयी है (8, 9)। लूएट. अल. (10) वित्तीय समय श्रृंखला की पूर्वानुमान के लिए स्वतंत्र घटक विश्लेषण (आईसीए) और एसवीआर पर आधारित एक एल्गोरिदम विकसित किया था। पॉलएटअल (11) गुजरात, भारत में अरहर में रोग की घटना के पूर्वानुमान के लिए एसवीआर तकनीक लागू किया है। वर्तमान शोध पत्र में, अस्थिरता को अधिक कुशलता से पकड़ने के लिए गार्च और एसवीआर तकनीक को मिला कर हाइब्रिड मॉडल को लागू करने का प्रयास किया गया है। एक अनुप्रयोग के रूप में, दिल्ली और कोलकाता में प्याज की कीमतों में उतार-चढ़ाव के पूर्वानुमान के लिए कार्यप्रणाली लागू

किया गया है।

एसवीआर मॉडल रिग्रेशन सेटअप के समान है और यह निम्नलिखित फंक्शन द्वारा पिछले अवलोकनों से भविष्य के मूल्य का पूर्वानुमान करता है:

$$y_t = f(y_{t-1}, y_{t-2}, \dots, y_{t-p}) + \varepsilon_t$$

यहाँ y_t आउटपुट है, p सम्मिलित किए जाने वाले लैग्स की संख्या है और ε_t जर्नॉइज़डर्म है। एक समय-श्रृंखला (σ_t^2) की अस्थिरता को दो भागों, सशर्त विचरण (h_t) और अवशिष्ट (R_t) के रूप में शामिल किया जा सकता है:

$$\sigma_t^2 = h_t + R_t$$

यहाँ h_t पहले गार्च द्वारा प्राप्त किया जा सकता है, उसके बाद अवशिष्ट (R_t) का अनुमान इस प्रकार होगा:

$$R_t = \sigma_t^2 - \hat{h}_t$$

R_t को एसवीआर द्वारा निम्नानुसार मॉडल किया जा सकता है:

$$R_t = \varphi(R_{t-1}, R_{t-2}, \dots, R_{t-p}) + \varepsilon_t$$

जहाँ p अंतराल है और $\varphi(\cdot)$ एसवीआर फंक्शन है। इस एसवीआर को लागू करके \hat{R}_t प्राप्त किया जा सकता है। इसलिए, संयुक्त पूर्वानुमान है

$$\sigma_t^2 = \hat{h}_t + \hat{R}_t$$

यहाँ, अस्थिरता का पूर्वानुमान करने की प्रक्रिया में निम्नलिखित चरण शामिल हैं:

1. प्रतिशत लॉग-रिटर्न श्रृंखला (r_t) की गणना मौसमी रूप से समायोजित मासिक डेटा पर की जाती है।
2. सर्वश्रेष्ठ उपयुक्त माध्यमॉडल (एरिमा) को r_t श्रृंखला पर फिट किया जाता है।
3. सशर्त विषम लैंगिकता का परीक्षण करने के लिए अवशेषों पर आर्च-एलएम परीक्षण किया जाता है।
4. सशर्त अस्थिरता (\hat{h}_t) को मॉडल करने के लिए एक विशिष्ट गार्च मॉडल फिट किया गया है।
5. अवशिष्टों की गणना फिटिड गार्च से की जाती है: $R_t = (r_t^2 - \hat{h}_t)$, जहाँ वास्तविक अस्थिरता को वर्गलॉगरिटर्न (लूएट. अल, 2013) के रूप में माना जा सकता है, जिसकी गणना इस प्रकार की जा

सकती है:

$$r_t = 100 * (\ln(Y_t) - \ln(Y_{t-1}))$$

6. अवशिष्ट निदान किया जाता है। यदि अवशिष्ट सहसंबद्ध हैं, तो अगले चरण पर आगे बढ़ें। अन्यथा अंतिम परिणाम के लिए जाएं।
7. एसवीआर को अवशिष्ट श्रृंखला ($\hat{R}|t$) के पूर्वानुमान के लिए लागू किया जाता है।
8. अस्थिरता का अंतिम पूर्वानुमान ($\hat{h}_t + \hat{R}_t$) के रूप में प्राप्त किया जा सकता है।

उपरोक्त एल्गोरिथम की पूर्वानुमान सटीकता की प्रयोगाश्रित तुलना के लिए, गार्च मॉडल और एसवीआर, में निम्नलिखित तीन मानदंडों का उपयोग किया जाता है:

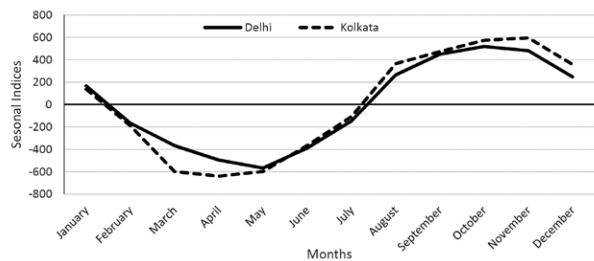
$$MAE = \sum_{t=1}^n |r_t^2 - \hat{r}_t^2| / n$$

$$MSE = \sum_{t=1}^n (r_t^2 - \hat{r}_t^2)^2 / n$$

$$R^2 \log = \sum_{t=1}^n \left[\ln \left(\frac{r_t^2}{\hat{r}_t^2} \right) \right]^2 / n$$

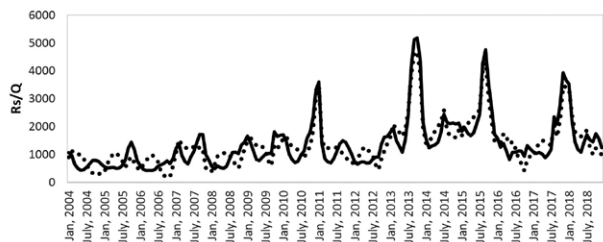
वर्तमान अध्ययन के लिए, दिल्ली और कोलकाता के लिए जनवरी, 2004 से दिसंबर, 2018 के दौरान राष्ट्रीय बागवानी अनुसंधान और विकास फाउंडेशन (एनएचआरडीएफ), भारत से प्याज का थोक मूल्य एकत्र किया गया है। मौसमी सूचकांकों को चित्र 1 में दर्शाया गया है। जो मौसमी सूचकांकों का न्यूनतम मूल्य मई के महीने में दिल्ली में और अप्रैल में कोलकाता के बाजार में मिले जबकि अधिकतम मूल्य अक्टूबर के महीने में दिल्ली में लिए और नवंबर में कोलकाता के बाजार में देखा जा सकता है।

दिल्ली और कोलकाता के बाजारों के मासिक मूल्य के साथ-साथ मौसमी रूप से समायोजित मूल्य क्रमशः चित्र 2 और 3 में दर्शाए गए हैं।

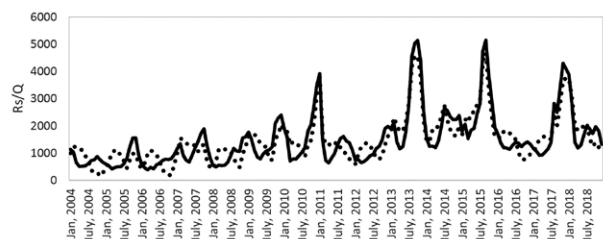


चित्र. 1 अध्ययन किए गए दो बाजारों में प्याज के थोक मूल्य के मौसमी सूचकांक

ऑगमेंटेडडिक्रीफुलर (एडीएफ) परीक्षण, फिलिप्सपेरोन (पीपी) परीक्षण और क्वाटकोस्की-फिलिप्स-स्मिथशिन (केपीएसएस) परीक्षण श्रृंखला की स्थिरता के परीक्षण के लिए किए गए हैं। सभी परीक्षण श्रृंखला की स्थिरता का संकेत देते हैं। श्रृंखला में अरेखीय रेखा की उपस्थिति सुनिश्चित करने के लिए बीडीएस अरेखिकता परीक्षण भी किया गया है। आर्च-एलएम परीक्षण का उपयोग करके श्रृंखला में सशर्त हेटेरोसेडैस्टिसिटी की उपस्थिति का परीक्षण किया गया था।



चित्र. 2 दिल्ली बाजार के लिए वास्तविक प्याज मूल्यश्रृंखला (ठोसरेखा) और मौसमी रूप से समायोजित श्रृंखला (बिंदीदार रेखा)



चित्र. 3 कोलकाता बाजार के लिए वास्तविक प्याज मूल्यश्रृंखला (ठोसरेखा) और मौसमी रूप से समायोजित श्रृंखला (बिंदीदार रेखा)

व्यक्तिगत लॉगरिथ्म श्रृंखला के लिए अविभाजित गार्च मॉडल को फिट करने के बाद, अवशिष्टों की जांच ऑटो सहसंबंध (एसीएफ) और आंशिक ऑटो सहसंबंध (पीएसीएफ) प्लॉट के माध्यम से ऑटो सहसंबंध की संभावित उपस्थिति के लिए की गई थी। दोनों बाजारों के लिए स्वतः सहसंबद्ध अवशेषों को खोजने पर, एसवीआर को फिटेड गार्च मॉडल से प्राप्त अवशिष्ट श्रृंखला पर लागू किया गया है। अवशिष्ट श्रृंखला में स्वतः सहसंबंध की उपस्थिति स्पष्ट रूप से इंगित करती है कि एसवीआर जैसी अन्य उन्नत सांख्यिकीय तकनीक का उपयोग कर के कुछ पैटर्न अभी भी अवशेषों से निकाला जा सकता है।

पारंपरिक गार्च मॉडल और गार्च-एसवीआर के बीच प्रयोगाश्रित तुलना आरएमएसई, एमएई और R^2 लॉगमानों द्वारा मापी गई उनके पूर्वानुमान प्रदर्शन के आधार पर की जाती है। तुलना के उद्देश्य से, 18 डेटा बिंदुओं के सत्यापन से इसका उपयोग किया जाता है। संबंधित परिणाम तालिका 1 में प्रस्तुत किए गए हैं। दोनों बाजारों में गार्च के आरएमएसई और हाइब्रिड तकनीक के बीच एक सटीक अंतर है। इसी तरह, एमएई और R^2 लॉगमानों पर भी विचार करते हुए, गार्च विधि पर गार्च एसवीआर हाइब्रिड मॉडल के बेहतर प्रदर्शन को दृढ़ता से स्थापित किया जा सकता है।

तालिका 1 एक कदम आगे गार्च और गार्च-एसवीआर मॉडल का पूर्वानुमान प्रदर्शन

बाजार	सत्यापन मानदंड	गार्च	गार्च-एसवीआर
दिल्ली	आरएमएसई	865.29	674.50
	एमएई	567.15	463.76
	R^2 लॉग	41.64	6.86
कोलकाता	आरएमएसई	833.07	779.47
	एमएई	517.19	470.81
	R^2 लॉग	66.47	8.73

भविष्य के पूर्वानुमान और सटीकता की विशेषता के आधार पर महत्वपूर्ण निर्णय लेने के लिए एक सांख्यिकीय मॉडल की सिफारिश की गई है। यहाँ,

पर्यवेक्षित एमएल तकनीक अर्थात अस्थिरता से निपटने के लिए पारंपरिक गार्च मॉडल के साथ एसवीआर लागू किया गया है। एसवीआर आधारित हाइब्रिड दृष्टिकोण दिल्ली और कोलकाता दोनों बाजारों में प्याज की कीमतों में उतार-चढ़ाव के पूर्वानुमान और मॉडलिंग में पारंपरिक गार्च मॉडल से बेहतर पाया गया है। आरएमएसई, एमएई और आर स्कायर लॉग वैल्यू को दक्षता मानदंड मानते हुए, गार्च और गार्च-एसवीआर मॉडल के बीच प्रयोगाश्रित तुलना की गई है। वस्तुओं की कीमतों का सटीक और समय पर पूर्वानुमान न केवल किसानों को अपनी उपज बेचने के लिए उचित निर्णय लेने में मदद करेगा बल्कि नीति निर्माताओं को उचित नीति तैयार करने में भी मदद करेगा जो बदले में किसानों की आय बढ़ाने में मदद करेगा।

संदर्भ

- एंगल, आर. एफ. (1982). ऑटोरेग्रेसिव कंडीशनल हेटेरो सेडैस्टिसिटी विद एस्टीमेट आफ द वैरिएन्स आफ यूनाइटेड किंगडम इकोनामेट्रिका। 50, 987-1007
- बोलर्सलेव, टी. (1986). जर्नलाइज्ड ऑटोरेग्रेसिव कंडीशनल हेटेरो सेडैस्टिकटी। इकोनामेट्रिका जर्नल, 31, 307-327.
- पगन, ए. आर. एवं स्वर्ट, जी. डब्लू. (1990). अल्टरनेटिव मॉडल फोर कंडीशनल स्टॉक वोलेटिलिटी. जे. इकोनोमेट्रीस 45: 267-290।
- खशी, एम एवं बिजरी, एम (2012). ए न्यू क्लास आफ हाइब्रिड मॉडल फार टाइम सीरीज फोरकास्टिंग एक्सपर्ट सिस्ट एप्लाइड। 39(4), 4344- 4357।
- वेडिंग, डी. के. एवं सिओस, के. जे. (1996). टाइम सीरीज फोरकास्टिंग वाई काम्बाइनिंग आरबीएफ नेटवर्क, सर्टीनिटी फैक्टर्स एन्ड द बॉक्स-जेनकिंस मॉडल। न्यूरो कंप्यूटिंग, 10(2), 149-168।
- मित्रा, डी. एवं पॉल, आर. के. (2017). हाइब्रिडटाइम-सीरीज मॉडल फार फोरकास्टिंग एग्रीकल्चर कमोडिटी प्राईसेज। मॉडल असिस्ट स्टेटएपल 12(3): 255-264.
- झांग, वाई. जे. एवं झांग, जे. एल. (2018). वोलेटिलिटी फोर कास्टिंग आफ क्रूड आयल मार्केट: ए न्यू हाइब्रिड मैथड। जे फोरकास्ट 37(8): 781-789।
- टे, ए. एफ. ई. एवं काओ, एल. (2001). एप्पलीकेशन आफ सपोर्ट वेक्टर मशीन इन फाइनेशियल टाइम-सीरीज फोरकास्टिंग। ओमेगा 29(4): 309-317.
- किम, के. जे. (2003). फाइनेशियल टाइम-सीरीज फोर कास्टिंग यूजिंग सपोर्ट वेक्टर मशीन। न्यूरोकंप्यूटिंग 55(1-2): 307-319।
- लूसी, जे., ली, टी. एस. एवं चिउ, सी. सी. (2009). फाइनेशियल टाइम-सीरीज फोरकास्टिंग यूजिंग इन्डेपेन्डेंट कम्पोनेन्ट एनालाइसिस एन्ड सपोर्ट वेक्टररिगेशन। डेसिससपोर्टसिस्ट 47(2): 115-125।
- पॉल, आर. के., वेन्निला, एस., सिंह, एन., चंद्र, पी., यादव, एस. के., शर्मा, ओ. पी., शर्मा, वी. के., निसार, एस., भट, एम. एन., राव, एम. एस. एवं प्रभाकर, एम. (2018). सीजनल डायनेमिक्स ऑफ स्टैटिस्टिकल मोजेक ऑफ पिजनपी एंड इट्स प्रेडिक्शन यूजिंग स्टैटिस्टिकल मॉडल फार बनासकाठारीजन आफ गुजरात, इंडिया जेइन्ड सोस एग्रिल स्टेट 72(3): 213-233

“

विज्ञान के चमत्कार कभी-कभी,
ईश्वर के अस्तित्व को भी दर्शाते हैं।

”

फसल उपज के पूर्वानुमान के लिए उन्नत मौसम सूचकांक आधारित बेसिअन समाश्रयण मॉडल

मोहम्मद यासीन¹, के.एन. सिंह¹, अचल लामा¹, बिशाल गुरुंग¹, रंजीत कुमार पॉल¹, प्रकाश कुमार¹,
अमृत कुमार पॉल, हिमाद्रि शेखर राय¹ एवं विशाल दिनकर²

¹भा.कृ.अनु.प.—भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली—110012

²भा.कृ.अनु.प.—केन्द्रीय शीतोष्ण बागवानी संस्थान, श्रीनगर—190007

सार:

कृषि भारतीय अर्थव्यवस्था की रीढ़ है, इसलिए सरकार को नई योजनाएँ बनाने के लिए फसल उपज के एक विश्वसनीय पूर्वानुमान की आवश्यकता होती है। समाश्रयण विश्लेषण, फसल उपज के पूर्वानुमान के लिए व्यापक रूप से उपयोग की जाने वाली तकनीक है। मापदंडों का महत्व समाश्रयण विश्लेषण की प्रमुख समस्याओं में से एक है। गैर-महत्वपूर्ण प्राचल निरर्थक पूर्वानुमान मूल्यों की ओर ले जाते हैं जो कि विश्वसनीय नहीं होते हैं। मॉडल को बेहतर बनाने के लिए हमने बेसिअन तकनीक के माध्यम से पूर्वगामी जानकारी को सम्मिलित किया है और बेसिअन ढांचे के तहत इन मॉडलों की श्रेष्ठता की जांच की है। बेसिअन तकनीक सांख्यिकी के आधुनिक युग में सबसे शक्तिशाली विधियों में से एक है। हमने विभिन्न प्रकार के पूर्वगामी (सूचनात्मक, गैर-सूचनात्मक और संयुग्म पूर्वज) पर चर्चा की है। बेसिअन ढांचे के तहत मापदंडों के आकलन के लिए एम.सी.एम.सी (MCMC) पद्धति पर संक्षेप में चर्चा की गई है। इन मॉडलों को स्पष्ट करने के लिए केला, आम और गेहूँ की फसलों के उत्पादन आंकड़ों को ध्यान में रखा जाता है। हमने पारंपरिक समाश्रयण मॉडल की तुलना बेसिअन समाश्रयण मॉडल से की और निष्कर्ष निकाला कि बेसिअन ढांचे के तहत अनुमानित मॉडल, शास्त्रीय दृष्टिकोण के तहत अनुमानित मॉडल की तुलना में बेहतर परिणाम प्रदान करते हैं।

मुख्य शब्द: बेसिअन तकनीक, मार्कोव चेन मॉंटे कार्लो (एम.सी.एम.सी), पूर्वगामी वितरण, सरल समाश्रयण

मॉडल, मौसम सूचकांक।

1. परिचय

कृषि क्षेत्र, भारतीय अर्थव्यवस्था में सबसे महत्वपूर्ण योगदानकर्ताओं में से एक है। कृषि खाद्य सुरक्षा प्रदान कर अधिक घनी आबादी वाले और स्तरीकृत समाजों के विकास को सक्षम बनाती है। भारत को कृषि प्रधान देश कहा जाता है क्योंकि भारत के लगभग 43% भौगोलिक क्षेत्र पर कृषि की जाती है और भारत की लगभग 70% आबादी प्रत्यक्ष या अप्रत्यक्ष रूप से कृषि क्षेत्र में कार्यरत हैं। भारत जैसे कृषि प्रधान देशों में खाद्य खरीद से संबंधित कई नीतियों के नियोजन व कार्यान्वयन फसल कटने से पहले उपज के पूर्वानुमान पर बहुत अधिक निर्भर हैं। फसल कटने से पहले उपज का अनुमान लगाने के लिए हमें उन कारकों के बारे में जानना होगा जो फसल की उपज को प्रभावित करते हैं। मौसम चर फसलों की वृद्धि और उत्पादन के लिए महत्वपूर्ण कारकों में से एक हैं। कटाई से पहले फसल की पैदावार की मॉडलिंग और भविष्यवाणी करना शोधकर्ता के लिए चिंता का विषय है। बहु-समाश्रयण विश्लेषण (Multiple Regression Analysis) कृषि मौसम विज्ञान-संबंधी (Agro&meteorological) फसल उपज पूर्वानुमान के लिए शक्तिशाली उपकरणों में से एक है (गोम्स, 1998)। समाश्रयण विश्लेषण (Regression analysis) एक बहुभिन्नरूपी तकनीक (Multivariate technique) है जिसका उपयोग व्याख्यात्मक चर के रूप में पर्यावरणीय कारकों का विश्लेषण करने तथा फसल की पैदावार पर उनका प्रभाव ज्ञात करने एवं

निर्णय लेने में किया जाता है (सेलम और पूवम्मल, 2016)। अधिकतर मामलों में पारंपरिक समाश्रयण मॉडल मापदंडों का एक बहुत ही कुशल अनुमान देता है, लेकिन कभी-कभी मॉडल के प्राचल सांख्यिकीय रूप से महत्वपूर्ण नहीं होते हैं (चटर्जी और हादी, 2015)। इसलिए समाश्रयण प्राचल अनुमानों में और सुधार की गुंजाइश है। इसलिए वर्तमान लेख में हमने इस संदर्भ में बेसिअन तकनीक (Bayesian technique) को नियोजित किया है।

बेसिअन विश्लेषण इस मूल धारणा पर आधारित है कि मॉडल के सभी प्राचल यादृच्छिक प्रकृति के हैं। हम बेहतर पूर्वानुमान मान प्राप्त करने के लिए मापदंडों के अनुमानित मूल्य में पूर्वगामी ज्ञान को शामिल कर सकते हैं। बेसिअन सांख्यिकी (Bayesian Statistics) चिरसम्मत आवृत्तिवादी सांख्यिकी (Classical Frequentist Statistics) के विपरीत हैं जहाँ सभी मापदंडों को निश्चित और अज्ञात राशियों के रूप में माना जाता है। बेसिअन विश्लेषण संभाव्यता के सिद्धांतों (Theory of Probability) पर आधारित है। बेसिअन समाश्रयण न केवल एक एल्गोरिथम है, बल्कि विभिन्न सांख्यिकीय अनुमानों के लिए दृष्टिकोण का एक अलग तरीका भी है। बेसिअन ढांचे का उपयोग करके, हम बिंदु अनुमान के बजाय अनुमानात्मक समाधानों की एक श्रृंखला प्राप्त करते हैं (मारिन और अन्य, 2014)। समान्यतः, एक गैर-बेसिअन समाश्रयण मॉडल (Non-Bayesian Regression model) या कहें कि एक रैखिक समाश्रयण मॉडल (Linear Regression Model) डेटा को ओवरफिट करता है। बेसिअन रैखिक समाश्रयण (Bayesian Linear Regression) 'भावी सूचक वितरण' (Predictive distribution) को शामिल करके इस समस्या को दूर करने में मदद करता है (ग्रेगरी, 2005)। बेसिअन तकनीक रैखिक मॉडल के अज्ञात मापदंडों का अनुमान लगाने के लिए एक व्यक्तिपरक प्रक्रिया है और यह मॉडल मापदंडों का अधिक कुशल अनुमान देता है (बन, 1975)। बेसिअन समाश्रयण अनुमान का लाभ इसके उस गुण के कारण है जिससे कोई भी पूर्वगामी वितरण (Prior distribution) को शामिल कर सकता है, या "विश्वासों" की वर्तमान स्थिति

के बारे में अनुमानित ज्ञान का उपयोग कर सकता है और अनुमानित मान को अधिक सटीक व कुशल बना सकता है। समस्याओं के एक विस्तृत वर्ग को हल करने के लिए उपयुक्त पूर्वगामी के साथ बेसिअन पद्धति को लागू करने के लिए विभिन्न विश्लेषणात्मक और संख्यात्मक तकनीकें उपलब्ध हैं (बर्नार्डो और स्मिथ, 2001)। बेसिअन तकनीक को मॉडल के सुधार के लिए सामान्यीकृत रैखिक मॉडल (Generalized Linear Model, जी.एल.एम) पर सफलतापूर्वक लागू किया जा सकता है और यह सुधार पूर्वगामी वितरण के चुनाव पर निर्भर करता है (दास, 2008)।

इस लेख में हमने बेसिअन तकनीक का उपयोग करके फसल की पैदावार के पूर्वानुमान को और बेहतर बनाने का प्रयास किया गया है। हमें साहित्य में समान कार्य नहीं मिला और आशा है कि हमारा दृष्टिकोण मौसम सूचकांक आधारित फसल उपज पूर्वानुमान मॉडल के उपलब्ध साहित्य को समृद्ध करेगा। आगामी अनुभागों में हम भौतिक विधियों, चित्रण, परिणामों और उसके निष्कर्षों पर चर्चा करेंगे।

2. सामग्री और प्रणालियां

2.1. आँकड़ों के स्रोत

मौसम के मापदंडों जैसे अधिकतम और न्यूनतम तापमान, सुबह और शाम की सापेक्षिक आर्द्रता, 23 साल (1984 से 2007 तक) के लिए वर्षा की मात्रा पर दैनिक आंकड़े भा.कृ.अनु.सं, नई दिल्ली में स्थित मौसम केंद्र से एकत्र किया गया है। मॉडल को स्पष्ट करने के लिए तीन फसलों अर्थात गेहूँ, केला और आम के उत्पादन आंकड़े श्रृंखला एकत्र की गई है। गेहूँ की उपज के आंकड़े भा.कृ.अनु.सं, नई दिल्ली से एकत्र किए गए हैं।

2.2. आँकड़ों के समुच्चयों का रूपान्तरण

अगर मान लिया जाए कि w सप्ताह ($w=1, 2, \dots, 7$) को दर्शाता है, जिस पर फसल की पैदावार की कटाई पूर्व अनुमान जारी करने की आवश्यकता होती है। यदि हम p चर में m सप्ताह पर साप्ताहिक डेटा का उपयोग करते हैं, तो नीचे दी गई प्रक्रिया

का उपयोग करके अब प्रत्येक मौसम चर के संबंध में नए मौसम चर व इंटरैक्शन (interaction) घटक उत्पन्न किए जा सकते हैं। समय प्रवृत्ति (T) सहित सभी उत्पन्न चरों का एक साथ उपयोग करके एक पूर्वानुमान मॉडल विकसित किया गया है (अग्रवाल एवं अन्य 2001)।

2.3. पारंपरिक समाश्रयण मॉडल

एक साधारण समाश्रयण मॉडल को y के रूप में व्यक्त किया जा सकता है जो x और β का एक फलन है:

$$y \approx f(x, \beta)$$

जहाँ y एक आश्रित चर (dependent variable) को दर्शाता है, x एक आश्रित चर है और β एक अज्ञात गुणांक है।

एक बहु-समाश्रयण मॉडल (Multiple Regression model) में आश्रित चर एक से अधिक स्वतंत्र चर (independent variables) से संबंधित होता है।

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon$$

जहाँ y एक आश्रित चर को दर्शाता है, x_i स्वतंत्र चर को दर्शाता है और ε त्रुटि घटक है। β_0 अवरोधन

मापदंड है और β_i ढलान मापदंड हैं।

हमारी मुख्य रुचि अज्ञात मापदंडों में निहित है। मॉडल के अज्ञात मापदंडों का निष्पक्ष व कुशलतापूर्वक अनुमान लगाया जाना चाहिए।

2.4. बेसिअन दृष्टिकोण

बेसिअन निष्कर्ष अनुमान लगाने की सबसे शक्तिशाली तकनीकों में से एक है। आकलन की इस तकनीक के बिंदु अनुमान (Point estimation) की तुलना में कई लाभ हैं। आकलन की इस तकनीक का मुख्य लाभ यह है कि यह पूर्वगामी सूचना का उपयोग करती है। पूर्वगामी सूचना को सम्मिलित करने से हमें पश्च वितरण (Posterior Distribution) प्राप्त होता है। पूर्वगामी वितरण के रूप में अतिरिक्त जानकारी को शामिल करने के कारण पश्च वितरण में और अधिक जानकारी प्राप्त होती है। हमारी रुचि संयुग्म पूर्वगामी (Conjugate priors) में निहित है क्योंकि पश्च वितरण का रूप इसके पूर्व वितरण के समान ही रहता है, केवल हाइपर-मापदंड (hyper & parameters) सामयिक किए जाते हैं (चेन और इब्राहिम, 2003)। समाश्रयण आकलन के बेसिअन दृष्टिकोण के लिए उठाए गए कदम इस प्रकार हैं:

तालिका 1: केले के लिए सरल समाश्रयण मॉडल (Simple Regression Model) के मापदंड अनुमान

गुणांक (Coefficients)	अनुमान (Estimate)	मानक त्रुटि (S.E.)	टी-मान (t-value)	पी-मान (p-value)	महत्व (Significance)
B1	334.13	227.48	1.47	0.16	No
B2	68.28	14.98	4.56	0.00	Yes
B3	120.92	352.97	0.34	0.74	No
B4	-85.72	120.31	-0.71	0.49	No
B5	-236.56	417.00	-0.57	0.58	No
B6	-209.81	375.03	-0.56	0.58	No
B7	-7389.48	6851.06	-1.08	0.30	No
B8	3702.16	1487.62	2.49	0.02	Yes
B9	315.71	166.71	1.89	0.08	No
B10	3.55	12.42	0.29	0.78	No
B11	1799.32	936.62	1.92	0.07	No
B12	-630.70	364.02	-1.73	0.10	No

तालिका 2: आम के लिए सरल समाश्रयण मॉडल (Simple Regression Model) के मापदंड अनुमान

गुणांक (Coefficients)	अनुमान (Estimate)	मानक त्रुटि (S.E.)	टी-मान (t-value)	पी-मान (p-value)	महत्व (Significance)
B1	37.84	5.70	6.64	0.00	Yes
B2	39.87	87.59	0.46	0.65	No
B3	0.97	42.08	0.02	0.98	No
B4	-112.78	76.65	-1.47	0.16	No
B5	53.94	23.37	2.31	0.03	Yes
B6	4006.64	1682.53	2.38	0.03	Yes
B7	-2269.65	980.86	-2.31	0.03	Yes
B8	12.95	16.74	0.77	0.45	No
B9	-20.53	9.30	-2.21	0.04	Yes
B10	-663.12	341.90	-1.94	0.07	No
B11	141.50	269.66	0.53	0.60	No
B12	53.04	92.43	0.57	0.58	No
B13	-10.28	51.51	-0.20	0.84	No

- I. मॉडल विनिर्देश
- II. पूर्वगामी वितरण का चयन
- III. प्रायिकता फलन (likelihood function) ज्ञात करना
- IV. बेयस प्रमेय (Bayes theorem) लागू करना और मार्कोव चेन मॉंटे कार्लो (एम.सी.एम.सी) विधि की मदद से पश्च वितरण उत्पन्न करें
- V. पश्च वितरण की प्रत्याशा प्राप्त करना

3. परिणाम और चर्चा

हमने उपरोक्त चर्चा किए गए मॉडलों को केले, आम और गेहूँ के अकड़ों के समुच्चय की मदद से स्पष्ट किया है। पहले इन मॉडलों के प्रदर्शन की जांच करने के लिए पारंपरिक समाश्रयण विधियों का उपयोग किया है। साधारण न्यूनतम वर्ग (Ordinary Least Square, vks-,y-,l) विधियों का उपयोग करके मापदंडों के अनुमानित मान की गणना की है। अनुमानित मान के महत्व की स्थिति को 5% महत्व के स्तर (Level of significance) पर जांचा है। इसकी अनुमानित मान मानक त्रुटियों व टी-मान (t-value) के साथ निम्न तालिका में सूचीबद्ध है।

तालिका 1 से यह पाया कि कुल मापदंडों में से केवल 16.67% महत्वपूर्ण है और 83.33% प्राचल जो कि एक बड़ा भाग है 5% महत्व स्तर पर गैर-महत्वपूर्ण है। यह भी देखा जा सकता है कि मापदंडों की मानक त्रुटि बहुत अधिक है।

तालिका 2 में कुल मापदंडों का 38.46% महत्वपूर्ण है जबकि शेष प्राचल (61.54%) न केवल गैर-महत्वपूर्ण है बल्कि इसकी मानक त्रुटि भी बहुत अधिक है।

तालिका 3 में यह देखा जा सकता है कि मापदंडों की गैर-महत्व दर बहुत अधिक 81.82% है और मापदंडों की मानक त्रुटि अधिक है। केवल 18.18% प्राचल महत्वपूर्ण है।

निम्नलिखित परिणाम से हमने पाया कि अधिकांश प्राचलों का अनुमान महत्व के 5% स्तर पर सार्थक नहीं है। इस प्रकार महत्व दर के स्तर को बढ़ाकर और मापदंडों अनुमानों की मानक त्रुटि को कम करके मापदंड अनुमानों में सुधार की गुंजाइश है। इसलिए मॉडल के सुधार की जांच करने के लिए दिए गए आंकड़ों के समुच्चयों के लिए बेसिअन ढांचे के तहत रैखिक समाश्रयण लागू किया है। बेसिअन तकनीक

तालिका 3: गेहूं के लिए सरल समाश्रयण मॉडल (Simple Regression Model) के मापदंड अनुमान

गुणांक (Coefficients)	अनुमान (Estimate)	मानक त्रुटि (S.E.)	टी-मान (t-value)	पी-मान (p-value)	महत्व (Significance)
B1	62.32	13.83	4.51	0.00	Yes
B2	3.21	4.50	0.71	0.49	No
B3	10.20	36.03	0.28	0.78	No
B4	1.48	7.12	0.21	0.84	No
B5	223.21	94.70	2.36	0.03	Yes
B6	-0.20	3.21	-0.06	0.95	No
B7	10.25	18.34	0.56	0.58	No
B8	-1.04	1.32	-0.79	0.44	No
B9	13.13	17.42	0.75	0.46	No
B10	-1.17	2.45	-0.48	0.64	No
B11	7.52	12.32	0.61	0.55	No

तालिका 4: केले के लिए बेसिअन समाश्रयण मॉडल (Bayesian regression model) के प्राचल अनुमान

गुणांक (Coefficients)	अनुमान (Estimate)	मानक त्रुटि (S.E.)	टी-मान (t-value)	पी-मान (p-value)	महत्व (Significance)
B1	317.04	214.32	1.48	0.16	No
B2	65.75	14.00	4.70	0.00	Yes
B3	400.94	335.20	1.20	0.25	No
B4	-91.16	112.83	-0.81	0.43	No
B5	-481.95	390.82	-1.23	0.24	No
B6	-117.01	351.92	-0.33	0.75	No
B7	-1431.38	6528.48	-0.22	0.83	No
B8	3492.54	1384.81	2.52	0.02	Yes
B9	323.66	157.84	2.05	0.05	Yes
B10	2.31	11.82	0.20	0.84	No
B11	2033.01	891.70	2.28	0.04	Yes
B12	-628.52	344.26	-1.83	0.09	No

मापदंडों के बारे में कुछ पूर्वगामी ज्ञान का उपयोग करती है। पूर्वगामी ज्ञान को प्राचल में शामिल करने के लिए हम पूर्वगामी वितरण के रूप में सामान्य वितरण (एक संयुग्म और सूचनात्मक पूर्वगामी) का चयन करते हैं। बेसिअन लीनियर समाश्रयण मॉडल का विश्लेषण करने के लिए हमने R सॉफ्टवेयर (संस्करण R-3.3.3) का उपयोग किया है (<http://www.R-project->

[org](http://www.R-project-))(R Core Team, 2013)। पश्च के परिणाम निम्नलिखित तालिकाओं में सूचीबद्ध हैं।

तालिका 4 में मापदंडों की गैर-महत्व दर 83.33% से घटकर 66.67% हो गई है और महत्व दर 16.67% से बढ़कर 33.33% हो गई है। मानकों की मानक त्रुटियों को भी कम किया गया है।

तालिका 5: आम के लिए बेसिअन समाश्रयण मॉडल (Bayesian regression model) के प्राचल अनुमान

गुणांक (Coefficients)	अनुमान (Estimate)	मानक त्रुटि (S.E.)	टी-मान (t-value)	पी-मान (p-value)	महत्व (Significance)
B1	37.58	5.73	6.56	0.00	Yes
B2	8.14	92.02	0.09	0.93	No
B3	-0.41	42.14	-0.01	0.99	No
B4	-91.99	78.13	-1.18	0.26	No
B5	54.41	23.20	2.35	0.03	Yes
B6	3366.10	1819.72	2.25	0.04	Yes
B7	-2256.69	987.31	-2.29	0.04	Yes
B8	11.04	16.71	0.66	0.52	No
B9	-19.76	9.39	-2.10	0.05	Yes
B10	-680.15	337.38	-2.02	0.05	Yes
B11	87.23	271.66	0.32	0.75	No
B12	45.83	91.77	0.50	0.62	No
B13	-4.10	51.44	-0.08	0.94	No

तालिका 5 में आम के आंकड़ों के समुच्चय में हमने पाया कि मापदंड की गैर-महत्व दर पारंपरिक समाश्रयण मॉडल से थोड़ी कम हो जाती है। बेसिअन समाश्रयण मॉडल के मापदंडों की गैर-महत्व दर 53.84% है पहले यह समाश्रयण मॉडल के लिए 61.54%

थी। मानकों की मानक त्रुटि को भी कम किया गया है।

तालिका 6 में हम देखते हैं कि मापदंड की गैर-महत्व दर में कोई कमी नहीं हुई है, लेकिन पारंपरिक समाश्रयण मॉडल की तुलना में मापदंडों की मानक त्रुटियां कम हो गई हैं।

तालिका 6: गेहूँ के लिए बेसिअन समाश्रयण मॉडल (Bayesian regression model) के प्राचल अनुमान

गुणांक (Coefficients)	अनुमान (Estimate)	मानक त्रुटि (S.E.)	टी-मान (t-value)	पी-मान (p-value)	महत्व (Significance)
B1	62.89	13.74	4.58	0.00	Yes
B2	3.01	4.56	0.66	0.52	No
B3	10.09	35.53	0.28	0.78	No
B4	1.71	7.18	0.24	0.81	No
B5	224.23	94.41	2.38	0.03	Yes
B6	-0.09	3.34	-0.03	0.98	No
B7	9.75	18.71	0.52	0.61	No
B8	-1.13	1.48	-0.76	0.46	No
B9	12.17	17.88	0.68	0.51	No
B10	-1.42	2.85	-0.50	0.62	No
B11	8.41	12.88	0.65	0.52	No

तालिका 7: केले के लिए रूट मीन स्क्वायर एरर (आर.एम.एस.ई) तुलना

अकड़ों के समुच्चय	सरल समाश्रयण के लिए आर.एम.एस.ई	बेसिअन समाश्रयण के लिए आर.एम.एस.ई
केला	223.14	216.87
आम	75.33	72.99
गेहूँ	266.35	265.85

तालिका 7 से आर.एम.एस.ई की तुलना से हम पाते हैं कि बेसिअन समाश्रयण मॉडल के रूट मीन स्क्वायर एरर (Root Mean Square Error, आर.एम.एस.ई) में थोड़ी कमी आई है। केला, आम और गेहूँ के अकड़ों के समुच्चय के लिए आर.एम.एस.ई में कमी का प्रतिशत क्रमशः 0.03, 0.03 और 0.002 है।

निष्कर्ष

इस अध्ययन का मुख्य उद्देश्य बेसिअन ढांचे के तहत संशोधित समाश्रयण मॉडल के साथ पारंपरिक समाश्रयण मॉडल की तुलना करना था। समाश्रयण विश्लेषण की दक्षता और सटीकता गैर-संदिग्ध हैं क्योंकि यह परिभाषित मान्यताओं के तहत सर्वश्रेष्ठ रैखिक निष्पक्ष अनुमानक (Best Linear Unbiased Estimator, बी.एल.यू.ई) अनुमान देता है। लेकिन कुछ परिस्थितियों में यह मापदंडों का एक कुशल अनुमान देने में विफल रहता है या कभी-कभी मॉडल का प्राचल अनुमान महत्वपूर्ण नहीं होता है। ऐसी परिस्थिति में समाश्रयण मॉडल को बेहतर बनाने और मापदंडों का एक कुशल अनुमान प्राप्त करने के लिए बेसिअन तकनीक को लागू किया जा सकता है। अपने अध्ययन के आधार पर हम आनुभविक रूप से तीन निष्कर्ष निकाल सकते हैं: पहला, बेसिअन दृष्टिकोण के तहत समाश्रयण, समाश्रयण मॉडल के मापदंडों की गैर-महत्व दर को कम करता है; दूसरा, बेसिअन तकनीक भी मापदंडों की मानक त्रुटि को कम करने में मदद करती है; और तीसरा कि यह मॉडल के आर.एम.एस.ई को भी कम करता है जो इंगित करता है कि बेसिअन

समाश्रयण पारंपरिक समाश्रयण मॉडल की तुलना में अधिक सटीक पूर्वानुमान प्रदान करता है। यह अध्ययन मौसम-आधारित समाश्रयण मॉडल के आकलन के लिए एक वैकल्पिक पद्धति देगा जहां मानक समाश्रयण दृष्टिकोण (Standard Regression Approach) ऐसा करने में विफल रहता है। वर्तमान जांच को विभिन्न अन्य आंकड़ों के समुच्चय पर लागू किया जा सकता है और परिणामों की तुलना की जा सकती है जो मौसम आधारित मॉडलिंग कार्यक्षेत्र के साहित्य को और समृद्ध करेगा।

संदर्भ

- अग्रवाल, आर., जैन, आर. सी. एवं मेहता, एस. सी. (2001). "यील्ड फोरकास्ट बेस्ड ऑन वैदर वैरीएबलेस एंड एग्रिकल्चरल इनपुट्स ऑन एग्रो-क्लाएमेटिक जोन बेसिस", इंडियन जर्नल ऑफ एग्रिकल्चरल साइंस, 71, 7, 487-490.
- बर्नार्डो, जे. एम. एवं स्मिथ, ए. एफ. (2001). "बेसिअन थ्योरी", न्यूयॉर्क. विली सीरीज.
- बन, डी. डब्ल्यू. (1975). "ए बाएसियान अप्रोच टू दी लिनियर कॉम्बिनेशन ऑफ फोरकास्टस" जर्नल ऑफ दी ऑपरेशनल रिसर्च सोसाइटी, 26, 2, 325-329.
- चटर्जी, एस. एवं हादी, ए. एस. (2015). "रिग्रेशन एनालिसिस बाय एकजाम्पल", जॉन विले एंड संस.
- चेन, एम. एच. एवं इब्राहिम, जे. जी. (2003), "कोंजुगते प्रायर्स फॉर गेनेरलाइज्ड लिनियर मॉडेल्स' स्टैटिस्टिका सिनिका, 13, 2, 461-476.
- दास, एस. (2008). "जेनेरलाइज्ड लिनियर मॉडेल्स एंड बियांड", डिसेर्टेशन. कनेक्टिकट विश्वविद्यालय.
- गोम्स, आर. (1998). "जग्रीमीटरोलॉजिकल क्रॉप यील्ड फोरकास्टिंग मथड्स", द नीदरलैंड्स: इंटरनेशनल स्टैटिस्टिकल इंस्टीट्यूट.
- ग्रेगरी, पी. (2005). "बेसिअन लॉजिकल डेटा एनालिसिस फॉर द फिजिकल साइंसेज: ए कम्परेटिव अप्रोच विद मैथेमेटिका सपोर्ट", केम्ब्रिज यूनिवर्सिटी प्रेस.

- मारिन, जे. एम., पिल्लै एन. एस., रॉबर्ट, सी. पी. एवं रूसो, जे. (2014). "रेलेवेंट स्टैटिस्टिक्स फॉर बेसिअन मॉडल चॉइस", जर्नल ऑफ दी रॉयल स्टैटिस्टिकल सोसाइटी: सीरीज बी (सांख्यिकीय मेथोडोलॉजी), 76, 5, 833–859.
- आर कोर टीम, (2013) "आर: ए लेंगवेज अँड एनवायरनमेंट स्टैटिस्टिकल कंप्यूटिंग", आर फाउंडेशन फॉर स्टैटिस्टिकल कंप्यूटिंग, वियना, ऑस्ट्रिया (<http://www.R-project.org/>).
- सेलम, वी. एवं पूवम्मल, ई. (2016). "प्रेडिक्सन ऑफ क्रॉप यील्ड युसिंग रिग्रेसन एनालिसिस", इंडियन जर्नल ऑफ साइंस एंड टेक्नोलॉजी, 9, 38,1–5.
- स्मिथ, ए. एफ. एवं रॉबर्ट्स, जी. ओ. (1993). "बेसिअन कम्प्यूटेशन बाइ दी गिब्स सैंपलर एंड रीलेटेड मार्कोव चेन मॉन्टे कार्लो मेथड्स", जर्नल ऑफ दी रॉयल स्टैटिस्टिकल सोसाइटी: सीरीज बी (मेथोडोलॉजिकल), 55, 1, 3–23.

“

विज्ञान को विज्ञान तभी कह सकते हैं,
जब वह शरीर, मन और आत्मा की भूख मिटाने की पूरी ताकत रखता हो।

”

-महात्मा गांधी

मैग्नापोर्थे ओरेजे से एफेक्टर प्रोटीन और चावल से माइटोकॉन्ड्रिया संबद्ध प्रोटीन कॉम्प्लेक्स के बीच परस्पर प्रभाव की जांच

भारती पाण्डेय, स्नेहा मुर्मू, सौम्या शर्मा, ऋत्विका दास, नीरज बुढ़लाकोटी, द्विजेश चंद्र मिश्र, दीपा भट्ट, सुधीर श्रीवास्तव एवं मोहम्मद समीर फारूकी
भा.कृ.अनु.प.—भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली

सारांश

माइटोकॉन्ड्रिया पौधे और पशु प्रतिरक्षा के लिए महत्वपूर्ण है। एफेक्टर प्रोटीन एमओसीडीआईपी4 (MoCDIP4) माइटोकॉन्ड्रिया से जुड़े ओएसडीजेए9 (OsDjA9)—ओएसडीआरपी1ई (OsDRP1E) प्रोटीन कॉम्प्लेक्स को लक्षित करके चावल की प्रतिरोधक क्षमता को कम करता है। वर्तमान अध्ययन में आणविक मॉडलिंग, आणविक गतिशीलता सिमुलेशन, और मुक्त ऊर्जा बाध्यकारी स्कोर का उपयोग करके जांच की गई कि, कैसे मैग्नापोर्थे ओरेजा (चावल में ब्लास्ट रोग उत्पन्न करने वाला कवक), माइटोकॉन्ड्रियाल गतिशीलता को प्रभावित करता है और चावल के पौधों में जन्मजात प्रतिरक्षा प्रतिक्रियाओं को रोकता है। आर्टिफिशियल इंटेलिजेंस आधारित टूल का उपयोग करके एफेक्टर प्रोटीन और टारगेट प्रोटीन संरचना प्राप्त की गई। एफेक्टर प्रोटीन के एमडी सिमुलेशन विश्लेषण ने माध्यमिक संरचना में महत्वपूर्ण परिवर्तन दिखाए। डॉकिंग स्कोर के आधार पर सर्वश्रेष्ठ ओएसडीजेए9—ओएसडीआरपी1ई कॉम्प्लेक्स को एमओसीडीआईपी4—(ओएसडीजेए9—ओएसडीआरपी1ई) कॉम्प्लेक्स के निर्माण के लिए चुना गया। ओएसडीजेए9—ओएसडीआरपी1ई और एमओसीडीआईपी4—(ओएसडीजेए9—ओएसडीआरपी1ई) कॉम्प्लेक्स की गणना आणविक यांत्रिकी के साथ सामान्यीकृत बॉर्न एंड सरफेस एरिया सॉल्वेशन (MM/GBSA) के साथ की गई और उनकी संबंधित बंधन मुक्त ऊर्जा -64.14 kcal/mol और -80 kcal/उवस पाई गई। इंटरेक्शन विश्लेषण ने हाइड्रोजन-बॉन्ड, सॉल्ट ब्रिज और नॉन बॉन्डेड कॉन्टेक्ट्स के माध्यम से

एमओसीडीआईपी4—(ओएसडीजेए9—ओएसडीआरपी1ई) कॉम्प्लेक्स के बीच मजबूत बाध्यकारी इंटरेक्शन का खुलासा किया। इस अध्ययन में एफेक्टर प्रोटीन के साथ-साथ माइटोकॉन्ड्रिया से जुड़े ओएसडीजेए9—ओएसडीआरपी1ई प्रोटीन कॉम्प्लेक्स के प्रमुख रैसिडु का पता चला, जो पादप-रोगजनक अंतःक्रिया का अध्ययन करने का मार्ग प्रशस्त करेगा।

1. परिचय

बढ़ती वैश्विक जनसंख्या हमें लगभग 50% अधिक भोजन का उत्पादन करने के लिए मजबूर कर रही है, जबकि हम 2050 तक 2.2 बिलियन लोगों के भोजन की तैयारी कर रहे हैं (एफएओ 2018)। लगातार बढ़ती खाद्य मांगों को पूरा करने में मुख्य जैविक चुनौतियों में से एक पौधों के फिलामेंटस रोगाणु हैं। यूकेरियोटिक फिलामेंटस रोगजनकों में कवक और ऊमाइसीटीस शामिल हैं, जो पौधों को व्यापक नुकसान पहुंचाते हैं और फसल की उपज और आर्थिक नुकसान का कारण बनते हैं (सावरी एट अल., 2019)। संक्रमण को रोकने के लिए रोगजनक कवक और ऊमाइसीटीस, प्रोटीन का स्राव करते हैं जिन्हें एफेक्टर के रूप में जाना जाता है (स्टरजीओपोलोस एट अल., 2009; डोड्स एट अल., 2009)। एफेक्टर प्रोटीन के कार्य अत्यधिक विविध हैं, जैसे कि मेजबान कोशिका भित्ति का क्षरण, मेजबान-स्रावित लाइटिक एंजाइम का स्राव को रोकना एवं मेजबान रक्षा प्रणाली को संशोधित करना। स्रावी, स्रावित प्रोटीन के पूरे सेट का गठन करता है, जो किसी जीव के 30% प्रोटीओम को दर्शाता करता है (गणेशन, 2016)। छोटे स्रावित प्रोटीन (एसएसपी) कुल कवक

स्राव का 40–60% होता है। माइटोकॉन्ड्रिया, जानवरों और पौधों की प्रतिरक्षा के लिए आवश्यक हैं। क्योंकि माइटोकॉन्ड्रियल गतिकी रोग प्रतिरोधक क्षमता को प्रभावित करती है, रोगजनकों ने प्रतिरक्षा प्रतिक्रियाओं को दबाने के लिए माइटोकॉन्ड्रियल गतिकी में हस्तक्षेप करने के लिए कई रणनीतियां विकसित की हैं (स्पियर एट अल., 2019)। कई रोगजनक एफेक्टर को हाल ही में डीआरपी1 और माइटोफ्यूसिन जो माइटोकॉन्ड्रियल विखंडन और संलयन मशीनरी के प्रमुख घटक हैं को सीधे लक्षित करने के लिए उत्तरदायी पाया गया था, जिससे माइटोकॉन्ड्रियल गतिकी के साथ हस्तक्षेप और स्तनधारी कोशिकाओं में जन्मजात प्रतिरक्षा प्रतिक्रियाओं को अवरुद्ध किया जा सकता है (स्मिरनोवा एट अल., 2001; सुजुकी एट अल., 2014)। जू एट अल., 2020, ने प्रदर्शित किया कि कवक रोगजनक मैग्नापोर्थे ओरेजा के एफेक्टर प्रोटीन एमओसीडीआईपी4, चावल की प्रतिरक्षा को कम करने के लिए माइटोकॉन्ड्रिया से जुड़े ओएसडीजेए9–ओएसडीआरपी1ई प्रोटीन कॉम्प्लेक्स को लक्षित करता है। डीएनएजे प्रोटीन ओएसडीजेए9 डायनामिन से संबंधित प्रोटीन ओएसडीआरपी1ई के साथ इंटरैक्ट करता है और ओएसडीआरपी1ई के क्षरण को बढ़ावा देता है, जो माइटोकॉन्ड्रियल विखंडन में कार्य करता है। इसके विपरीत, एमओसीडीआईपी4 ओएसडीजेए9 को ओएसडीआरपी1ई के साथ प्रतिस्पर्धा करने के लिए बांधता है, जिसके परिणामस्वरूप ओएसडीआरपी1ई संचय होता है।

वर्तमान शोध यह पता करने की कोशिश कर रही है कि चावल के ब्लास्ट रोग का कारक एजेंट कवक एम.ओरीजा, कैसे माइटोकॉन्ड्रियल गतिशीलता में हस्तक्षेप करता है और कैसे चावल के पौधों में जन्मजात प्रतिरक्षा प्रतिक्रियाओं को अवरुद्ध करता है। एम.ओरीजा से एमओसीडीआईपी4 के इंटरैक्शन की और चावल से ओएसडीजेए9–ओएसडीआरपी1ई के इंटरैक्शन की क्रिस्टल संरचना अभी तक अनुमानित नहीं की गयी है। इस अध्ययन में, हमने आणविक मॉडलिंग और आणविक गतिकी सिमुलेशन का उपयोग करके संरचना स्तर पर एमओसीडीआईपी4 और ओएसडीजेए9–ओएसडीआरपी1ई प्रोटीन कॉम्प्लेक्स के

बीच इंटरैक्शन की व्यवस्थित रूप से जांच की। इसके बाद, अंतःक्रियाओं के गतिकी व्यवहार का वर्णन करने के लिए बाध्यकारी ऊर्जा आकलन (एमएम–जीबीएसए) भी किया गया।

2.विधि

2.1 प्रोटीन की पूर्व-सूचना और प्रोटीन-प्रोटीन डॉकिंग

राइस जीनोम एनोटेशन प्रोजेक्ट (<http://rice.uga.edu/>) से एमओसीडीआईपी4 (यूनीप्रोटकेबी आईडी: G4MVX4; 295 अमीनो एसिड रैसिडु), ओएसडीजेए9 (LOC_Os06g02620.1; 422 अमीनो एसिड रैसिडु) और ओएसडीआरपी1ई (LOC_Os09g39960.1; 626 अमीनो एसिड रैसिडु) के सीक्वेंसेस डाउनलोड किये गए।

आर्टिफिशियल इंटेलिजेंस आधारित सर्वर अल्फाफोल्ड 2 (जम्पर एट अल., 2021) का उपयोग करके तीनों प्रणालियों के लिए प्रोटीन संरचना करके प्राप्त की गई। प्रोटीन कॉम्प्लेक्स बनाने के लिए ओएसडीजेए9 और ओएसडीआरपी1ई के बीच प्रोटीन-प्रोटीन डॉकिंग, हैडॉक सर्वर (यान एट अल., 2017) का उपयोग किया गया।

2.2 एफेक्टर एमओसीडीआईपी4 के आणविक गतिशीलता सिमुलेशन

आणविक गतिकी (एमडी) दृष्टिकोण प्रोटीन के गतिकी व्यवहार का सफलतापूर्वक अध्ययन कर सकता है। इसलिए, ग्रोमोस 96 54a7 बल क्षेत्र का उपयोग करके ग्रोमैक्स 5.0 सॉफ्टवेयर पैकेज (अब्राहम एट अल., 2015) में एमडी सिमुलेशन किया गया। सिस्टम में स्टीपेस्ट डिसेंट मिनिमाइजेशन विधि का उपयोग करके ऊर्जा को न्यूनतम किया गया। इसके अलावा, सिस्टम को दो चरणों में तापमान और दबाव को स्थिर करके संतुलित किया गया था: 1000 पीएस के लिए एनवीटी एनसेंबल और एनपीटी एनसेंबल।

अच्छी तरह से संतुलित प्रणालियों को 100 एनएस के लिए 300 K और 1 एटीएम बार दबाव पर 2fs समय चरणों के साथ परीणाम के दो सेटों में वितरित किया गया, निर्देशांक हर 10.0 पीएस पर नोट किये

गए। चित्रों को यूसीएसएफ काइमेरा1.12 (पेटर्सन एट अल., 2004) का उपयोग करके प्रस्तुत किया गया।

2.3 एमएमजीबीएसए का उपयोग करके बंधन मुक्त ऊर्जा विश्लेषण

हॉकडॉक सर्वर (वेंग एट अल., 2019) से आणविक यांत्रिकी सामान्यीकृत बॉर्न सॉल्वेंट एक्सेसिबिलिटी (एमएमजीबीएसए) का उपयोग करके प्रोटीन-प्रोटीन इंटरैक्शन की बाध्यकारी मुक्त ऊर्जा (ΔG बाइंड) की गणना की गई थी।

$$G_{\text{बाइंड}} = G_{\text{कॉम्प्लेक्स}} - (G_{\text{प्रोटीन}} + G_{\text{लिगैंड}}) \quad (1)$$

$$\Delta G = RT \ln k_{-D} \quad (2)$$

जहां R आदर्श गैस स्थिरांक है (किलो कैलोरी K⁻¹ मोल⁻¹ में), T तापमान 298.15 K पर सेट है और ΔG अनुमानित मुक्त ऊर्जा है।

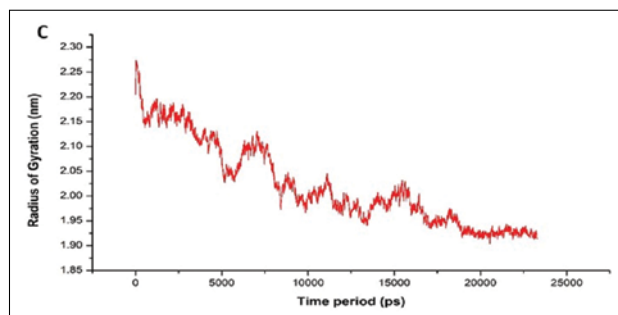
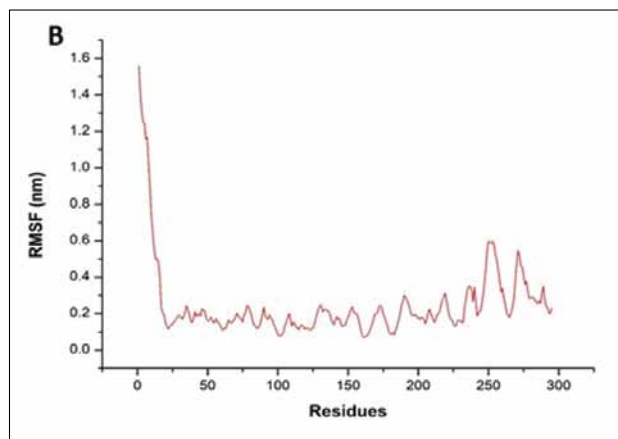
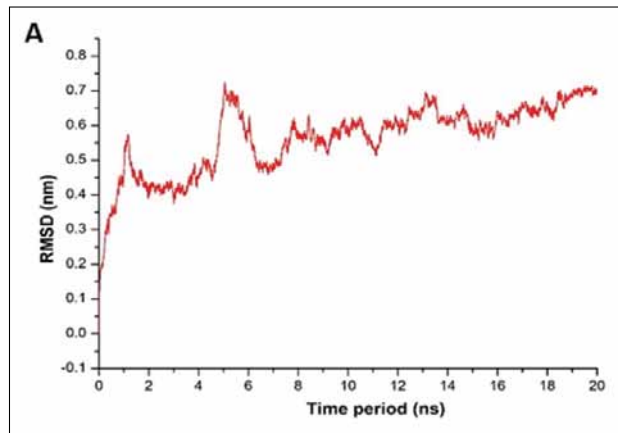
3. परिणाम और चर्चा

3.1 एमडी सिमुलेशन विश्लेषण

एमओसीडीआईपी4 प्रोटीन की अनुमानित संरचना ने एफेक्टर एमओसीडीआईपी4 प्रोटीन में चार डाइसल्फ़ाइड बांड दिखाए; CYS62-CYS182, CYS-128-CYS-139, CYS-268-CYS-284 और CYS-279-CYS-294। संरचना में एन-टर्मिनल पर 18 रैसिडु लंबे अनफोल्डेड क्षेत्र हैं। किसी संदर्भ फ्रेम के सापेक्ष किसी दिए गए फ्रेम के लिए परमाणुओं के एक विशेष सेट के विस्थापन में औसत परिवर्तन की गणना रूट मीन स्क्वायर विचलन (आरएमएसडी) का उपयोग करके की जाती है। इसकी गणना सूत्र के अनुसार की जाती है

$$RMSD = \sqrt{\frac{1}{N} \sum_{i=1}^N \delta_i^2} \quad (3)$$

जहां δ संदर्भ संरचना के एन जोड़े के बीच की दूरी है



चित्र 1: एमडीएस विश्लेषण (ए) रूट मीन स्क्वायर विचलन (बी) रूट मीन स्क्वायर उतार-चढ़ाव (सी) मैग्नापोर्थे ओरेजा से एफेक्टर प्रोटीन एमओसीडीआईपी4 के गार्डरेशन की त्रिज्या।

पूरे सिमुलेशन में प्रोटीन के आरएमएसडी को ट्रैक करने से प्रोटीन संरचनात्मक संरचना के बारे में जानकारी मिल सकती है। एमओसीडीआईपी4 प्रोटीन की बैकबोन के परमाणु की आरएमएसडी संख्या इसकी प्रारंभिक स्थिति में 0.7 एनएम तक बढ़ गई, यह दर्शाता

है कि सिमुलेशन रन टाइम अवधि के दौरान संरचना कुछ गठनात्मक परिवर्तनों से गुजर रही थी चित्र 1 (ए)। प्रोटीन गतिकी और रैसिडु के परमाणु गतिशीलता व्यवहार पर एकल बिंदु प्रक्षेपास्त्र उत्परिवर्तन के परिणामों को निर्धारित करने के लिए, प्रोटीन बैकबोन से प्रत्येक परमाणु स्थिति के मूल माध्य वर्ग उतार-चढ़ाव (त्सैथ) की गणना निम्नलिखित समीकरण का उपयोग करके की गई:

$$RMSF = \sqrt{\frac{1}{T} \sum_{t=1}^T (v_t - \bar{v})^2} \quad (3)$$

जहां T वह समय है जिस पर आरएमएसएफ की गणना की जाती है और अज कण ज की संदर्भ स्थिति है।

प्रत्येक अमीनो एसिड के आरएमएसएफ विश्लेषण से पता चला कि एन- और सी-टर्मिनल से रैसिडु में महत्वपूर्ण गति हुई थी। आरएमएसएफ ग्राफ में महत्वपूर्ण शिखर के उतार-चढ़ाव प्रमुख रूप से प्रोटीन के एन-और सी-टर्मिनल पर देखे गए। हालांकि, एमडी सिमुलेशन रन के दौरान एमओसीडीआईपी4 प्रोटीन के कोर अमीनो एसिड में कोई उतार-चढ़ाव नहीं देखा गया चित्र 1 (बी)। गार्डरेशन की त्रिज्या (आरओजी) प्रोटीन संरचना की सघनता का सूचक है। यह आरएमएसडी के अलावा समष्टि की स्थिरता को भी स्पष्ट करता है। आरओजी मानों में वृद्धि का तात्पर्य प्रोटीन संरचना की कॉम्पैक्टनेस में कमी है, जिससे लचीलेपन में वृद्धि और कम स्थिरता का पता लगता है। चित्र 1 (सी) से यह स्पष्ट है कि एमडी सिमुलेशन समय अवधि के दौरान एमओसीडीआईपी4 प्रोटीन आरओजी मानों में कमी के साथ सबसे कॉम्पैक्ट संरचना को अपनाता है। एमओसीडीआईपी4 प्रोटीन की प्रारंभिक और अंतिम फ्रेम संरचना को डाली सर्वर ([//ekhidna2.biocenter.helsinki.fi/dali/](http://ekhidna2.biocenter.helsinki.fi/dali/)) का उपयोग करके आरोपित किया गया था और उनका विचलन मूल माध्य वर्ग विचलन (आरएमएसडी) में मापा गया जो 2.5Å पाया गया। एमडी के स्थिर समय सीमा से एमओसीडीआईपी4 प्रोटीन के पीडीबी कोऑर्डिनेट्स का

आंकलन किया गया जिन्हें आगे के विश्लेषण के लिए उपयोग में लाया गया (चित्र 2क)।

3.2 ओएसडीजेए9- ओएसडीआरपी1ई के बीच डॉकिंग विश्लेषण

ओएसडीजेए9 और ओएसडीआरपी1ई के बीच प्रोटीन-प्रोटीन डॉकिंग के फलस्वरूप दस कॉम्प्लेक्स उत्पन्न किए गए। क्रमशः उच्चतम डॉकिंग स्कोरवाली समष्टियों को कॉन्फिडेंस स्कोर -296.1 के साथ डॉक किया गया (तालिका 1; चित्र 2बी)। कॉम्प्लेक्स के बीच इंटरैक्शन विश्लेषण से पता चला कि कॉम्प्लेक्स क्रमशः 7, 9 और 208 साल्ट ब्रिज, हाइड्रोजन बॉन्ड और नॉन-बॉन्डेड संपर्कों के साथ अत्यधिक स्थिर था (तालिका 1)।

साल्ट ब्रिज इंटरैक्शन के जोड़े ARG439-ASP137, ARG443-ASP142, HIS579-GLU188, LYS583-GLU188, GLU468-ARG189, GLU584-LYS280 और LYS582-ASP316 थे। आगे की जांच के लिए सबसे अच्छे डॉक हुए माइटोकॉन्ड्रिया से जुड़े ओएसडीजेए9 - ओएसडीआरपी1ई समष्टियों का उपयोग किया गया। ओएसडीजेए9 -ओएसडीआरपी1ई समष्टियों के इलेक्ट्रोस्टैटिक पोटेंशियल सरफेस विश्लेषण ने इलेक्ट्रोस्टैटिक पोटेंशियल सरफेस में बड़ी भिन्नता दिखाई (चित्र 2C)।

3.3 एमओसीडीआईपी4 प्रोटीन और ओएसडीजेए9-ओएसडीआरपी1ई के बीच डॉकिंग विश्लेषण

कवक रोगजनक मैग्नापोर्थ ओरेज़ा का एफेक्टर एमओसीडीआईपी4 चावल की प्रतिरक्षा को कम करने के लिए माइटोकॉन्ड्रिया से जुड़े ओएसडीजेए9 -ओएसडीआरपी1ई प्रोटीन कॉम्प्लेक्स को लक्षित करता है। ओएसडीजेए9 प्रोटीन डायनामिन से संबंधित प्रोटीन ओएसडीआरपी1ई के साथ इंटरैक्ट करता है और ओएसडीआरपी1ई के क्षरण को बढ़ावा देता है, जो माइटोकॉन्ड्रियल विखंडन में कार्य करता है। इसलिए, इंटरैक्शन के मौलिक तंत्र को समझने के लिए हमने एमओसीडीआईपी4 प्रोटीन को ओएसडीजेए9 -ओएसडीआरपी1ई कॉम्प्लेक्स में डॉक किया। सर्वश्रेष्ठ

एमओसीडीआईपी4—(ओएसडीजेए9—ओएसडीआरपी1ई) कॉम्प्लेक्स को डॉकिंग स्कोर और कॉन्फिडेंस स्कोर—259.98 और 0.9002 के आधार पर चुना गया (चित्र 3)। सबसे अच्छे एमओसीडीआईपी4—(ओएसडीजेए9—ओएसडीआरपी1ई) कॉम्प्लेक्स इंटरैक्शन विश्लेषण से पता चला कि ओएसडीजेए9 और ओएसडीआरपी1ई प्रोटीन दोनों के रेसिडु एमओसीडीआईपी 4 प्रोटीन के साथ इंटरैक्ट करते हैं और कॉम्प्लेक्स को स्थिर करते हैं। एमओसीडीआईपी4— ओएसडीजेए9 क्रमशः 1, 3, और 110 साल्ट ब्रिज, हाइड्रोजन बॉन्ड नॉन-बॉडेड संपर्कों के माध्यम से इंटरैक्ट करते हैं। LYS306-ASP135 (3.18Å) सॉल्ट ब्रिज इंटरैक्शन को स्थिर करते हैं।

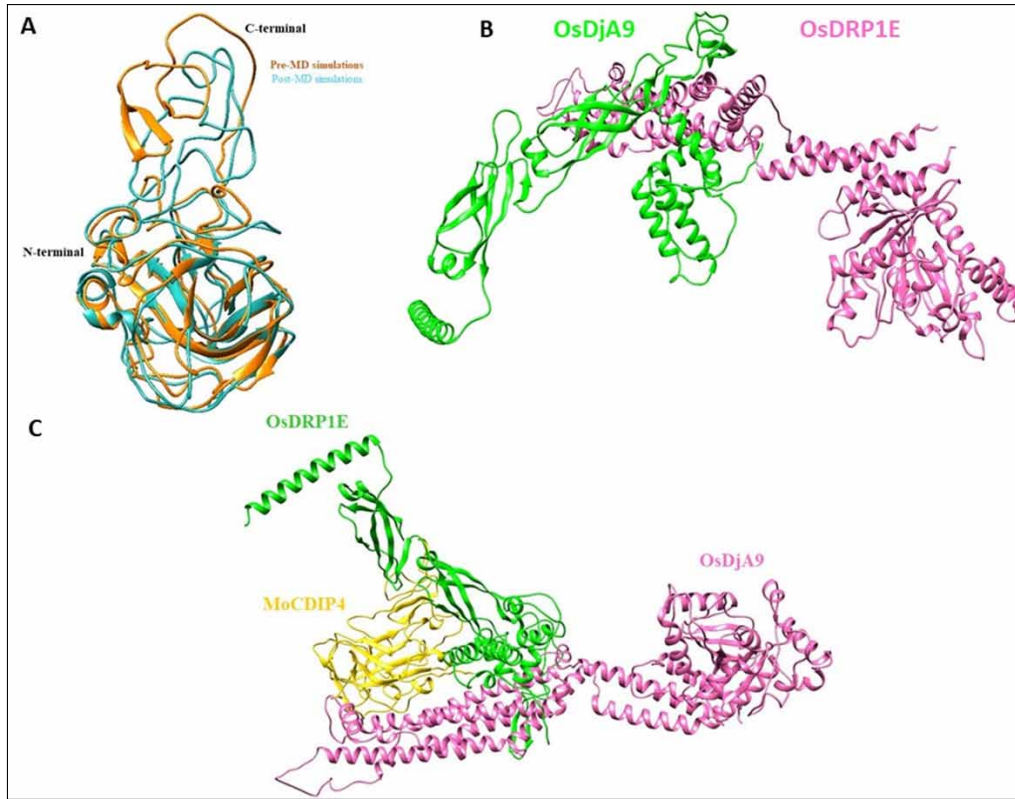
इसके अलावा, एमओसीडीआईपी4—ओएसडीआरपी1ई क्रमशः 1, 8, और 140 साल्ट ब्रिज, हाइड्रोजन बॉन्ड और नॉन-बॉडेड संपर्कों के माध्यम से इंटरैक्ट करते हैं। इंटरैक्शन को स्थिर करने के लिए GLU417-LYS41 (3.28Å) सॉल्ट ब्रिज इंटरैक्शन पाया गया। हाइड्रोजन बांड बनाने वाले रेसिडुस का उल्लेख तालिका 2 और 3 में किया गया है।

3.4 एमएम—जीबीएसए विश्लेषण

एमएम—जीबीएसए बाध्यकारी मुक्त ऊर्जा (ΔG) विश्लेषण ने दिखाया कि ओएसडीजेए9—ओएसडीआरपी1ई और एमओसीडीआईपी4—(ओएसडीजेए9— ओएसडीआरपी1ई) समष्टियों की बाध्यकारी मुक्त ऊर्जा (ΔG) क्रमशः -64.14 kcal/mol, और -80 kcal/mol है, जो स्थिर संरचना को दर्शाता है।

4. निष्कर्ष

कवक रोगजनक मैग्नापोर्थे ओरेज़ा का एफेक्टर एमओसीडीआईपी4 चावल की प्रतिरक्षा को कम करने के लिए माइटोकॉन्ड्रिया से जुड़े ओएसडीजेए9—ओएसडीआरपी1ई प्रोटीन कॉम्प्लेक्स को लक्षित करता है। एम.ओरिजे से एमओसीडीआईपी4 और चावल से ओएसडीजेए9— ओएसडीआरपी1ई की क्रिस्टल संरचना अभी तक हल नहीं हुई है। इस प्रकार, तीनों के लिए प्रोटीन संरचना तैयार करने और प्रोटीन—प्रोटीन डॉकिंग अध्ययन करने का औचित्य एमओसीडीआईपी4 और ओएसडीजेए9— ओएसडीआरपी1ई प्रोटीन कॉम्प्लेक्स के बीच इंटरैक्शन का विवरण प्राप्त करना था। मैग्नापोर्थे ओरिजे और माइटोकॉन्ड्रिया से जुड़े ओएसडीजेए9—ओएसडीआरपी1ई प्रोटीन कॉम्प्लेक्स से एफेक्टर प्रोटीन एमओसीडीआईपी4 के बीच इंटरैक्शन का विश्लेषण संरचना स्तर पर किया गया। इस डॉकिंग अध्ययन ने कॉम्प्लेक्स के निर्माण और चावल की प्रतिरक्षा को कम करने में शामिल महत्वपूर्ण रेसिडुस का खुलासा किया। एमओसीडीआईपी4 और ओएसडीजेए9 और एमओसीडीआईपी4— ओएसडीआरपी1ई के तीन और आठ अमीनो एसिड रेसिडु हाइड्रोजन बॉन्डिंग में शामिल थे। महत्वपूर्ण रूप से, सभी अमीनो एसिड जो हाइड्रोजन बॉन्डिंग में शामिल थे, महत्वपूर्ण पाए गए क्योंकि वे हाइड्रोफोबिक इंटरैक्शन द्वारा भी इंटरैक्ट थे। इस प्रकार, यह अध्ययन माइटोकॉन्ड्रिया से जुड़े ओएसडीजेए9— ओएसडीआरपी1ई प्रोटीन कॉम्प्लेक्स और एमओसीडीआईपी4— ओएसडीजेए9—ओएसडीआरपी1ई प्रोटीन इंटरैक्शन में शामिल प्रमुख रेसिडुस में महत्वपूर्ण अंतर्दृष्टि प्रदान करता है जो चावल की प्रतिरक्षा को कम करते हैं।



चित्र 2: संरचनात्मक विश्लेषण (ए) एमओसीडीआईपी4-की रिबन आकार की आरोपित संरचना दिखाई गई है, नारंगी रंग की संरचना 0 एनएस पर प्रणाली का प्रारंभिक चरण है और आणविक गतिशीलता सिमुलेशन के 20 एनएस के बाद नीले रंग की संरचना प्रणाली का अंतिम चरण है। (B) ओएसडीजेए9 ओएसडीआरपी1ई प्रोटीन कॉम्प्लेक्स का रिबन चित्रण (C) एमओसीडीआईपी4-(ओएसडीजेए9- ओएसडीआरपी1ई) कॉम्प्लेक्स का रिबन चित्रण

तालिका 1: ओएसडीजेए9- ओएसडीआरपी1ई कॉम्प्लेक्स के बीच हाइड्रोजन बॉन्ड

रैसि. परमाणु	रैसि. नाम	रैसि. संख्या	श्रृंखला	रैसि. परमाणु	रैसि. नाम	रैसि. संख्या	श्रृंखला	दूरी
O	ALA	73	A	NH2	ARG	443	B	2-58
O	ASP	137	A	NH2	ARG	439	B	2-64
O	GLN	138	A	ND2	ASN	467	B	1-95
OE1	GLN	143	A	N	ALA	460	B	2-59
NE2	GLN	143	A	OE1	GLN	459	B	2-94
OE2	GLU	188	A	NZ	LYS	583	B	3-02
O	GLN	259	A	ND2	ASN	574	B	2-65
NZ	LYS	280	A	OE2	GLU	584	B	2-67
OD1	ASP	316	A	NZ	LYS	582	B	3-26

तालिका 2: एमओसीडीआईपी4 प्रोटीन और ओएसडीआरपी1ई के बीच हाइड्रोजन बॉन्ड

परमाणु नाम	रैसि. नाम	रैसि. संख्या	श्रृंखला	परमाणु संख्या	परमाणु नाम	रैसि. नाम	रैसि. संख्या	श्रृंखला
OH	TYR	401	A	O	THR	74	C	2-64
OH	TYR	401	A	OG1	THR	76	C	2-64
NE2	GLN	410	A	O	TYR	36	C	2-86
NH1	ARG	413	A	O	SER	70	C	2-19
NH2	ARG	413	A	O	SER	70	C	3-13
NE	ARG	414	A	O	ASP	35	C	2-61
OE2	GLU	417	A	NZ	LYS	41	C	3-28
OD2	ASP	486	A	OH	TYR	46	C	2-73

तालिका 3: एमओसीडीआईपी4 प्रोटीन और ओएसडीजेए9 के बीच हाइड्रोजन बॉन्ड

परमाणु नाम	रैसि. नाम	रैसि. संख्या	श्रृंखला	परमाणु संख्या	परमाणु नाम	रैसि. नाम	रैसि. संख्या	श्रृंखला	दूरी
OH	TYR	77	B	150	N	LEU	10	C	2-94
O	ARG	84	B	150	N	LEU	10	C	2-29
NE	ARG	392	B	1931	O	CYS	128	C	2-2

श्रृंखला ए; ओएसडीआरपी1ई, चैन बी; ओएसडीजेए9, चैन सी; एमओसीडीआईपी4

संदर्भ

सावरी, एस., विलोकेट, एल., पेथीब्रिज, एस.जे., एस्कर, पी., मैकरॉबर्ट्स, एन. एवं नेल्सन, ए. (2019). प्रमुख खाद्य फसलों पर रोगजनकों और कीटों का वैश्विक बोझ। नेचर इकोलॉजी – इवॉल्यूशन 3(3), पीपी.430–439।

स्टर्गिओपॉलोस, आई. एवं दी. विट, पी. जे. (2009). फंगल इफेक्टर प्रोटीन। एनुअल रिव्यू ऑफ़ फाइटोपैथोलॉजी, 47, पीपी.233–263।

डोड्स, पी.एन., रफीकी, एम., गन, पी.एच., हरधम, ए.आर., जोन्स, डी.ए. एवं एलिस, जे.जी. (2009). बायोट्रोफिक फंगी और ओमीसाइकेट्स के प्रभावकारक: रोगजनक कारक और मेजबान प्रतिरोध के ट्रिगर। न्यू फाइटोलॉजिस्ट, 183(4), पीपी.993–1000।

रोवेनिच, एच., बोशोवेन, जे. सी. एवं थॉम्मा, बी.पी. (2014). फिलामेंटस रोगजनक प्रभावकारी कार्य: रोगजनकों, मेजबानों और माइक्रोबायोम के। करंट ओपिनियन इन प्लांट बायोलॉजी, 20, पीपी. 96–103।

गणेशन, बी. (2016). एस्परगिलस सेक्रेटोम: एक अवलोकन. न्यू एंड फ्यूचर डेवलपमेंट्स इन माइक्रोबियल बायोटेक्नोलॉजी एंड बायोइंजीनियरिंग, पीपी.69–77।

स्पायर, ए., स्टावरू, एफ. एवं कोसर्ट, पी. (2019). इंट्रासेल्युलर बैक्टीरियल रोगजनकों और मेजबान सेल माइटोकॉन्ड्रिया के बीच सहभागिता। माइक्रोबायोलॉजी स्पेक्ट्रम, 7(2), पीपी.7–2।

सुजुकी, एम., डेनिलचांका, ओ. और मेकलानोस, जे.जे. (2014). विब्रियो कोलरा टी3 एस. एस. इफेक्टर वोप इ मिरो जी टी पीअसेस को लक्षित करके

- माइटोकॉन्ड्रियल गतिकी और जन्मजात प्रतिरक्षा संकेतन को नियंत्रित करता है। सेल होस्ट – माइक्रोब, 16(5), पीपी.581–591.
- स्मिर्नोवा, ई., ग्रिपरिक, एल., शूरलैंड, डी. एल. एवं वैन डेर ब्लाडक, ए. एम. (2001). स्तनधारी कोशिकाओं में माइटोकॉन्ड्रियल विभाजन के लिए डायनामिन से संबंधित प्रोटीन डीआरपी1 की आवश्यकता होती है। मॉलिक्यूलर बायोलॉजी ऑफ़ द सेल, 12(8), पीपी.2245–2256।
- यान, वाई., झांग, डी., झोउ, पी., ली, बी. एवं हुआंग, एस. वाई. (2017). एचडीओसीके: प्रोटीन-प्रोटीन और प्रोटीन-डीएनए/आरएनए डॉकिंग के लिए एक वेब सर्वर एक संकर रणनीति पर आधारित है। न्यूक्लिक एसिड रिसर्च 45(डब्ल्यू1), पीपी. डब्ल्यू365–डब्ल्यू 373।
- अब्राहम, एम. जे., मुर्तोला, टी., शुल्ज़, आर., पल, एस., स्मिथ, जे. सी., हेस, बी. एवं लिंडाहल, ई. (2015). ग्रोमेक्स: लैपटॉप से सुपर कंप्यूटर तक बहु-स्तरीय समानता के माध्यम से उच्च प्रदर्शन आणविक सिमुलेशन। सॉफ्टवेयरएक्स 1: 19–25।
- पीटरसन, ई. एफ., गोडार्ड, टी. डी., हुआंग, सी. सी., काउच, जी. एस., ग्रीनब्लाट, डी. एम., मेंग, ई. सी. एवं फेरिन, टी.ई. (2004). यूसीएसएफ चिमैरा-अन्वेषी अनुसंधान और विश्लेषण के लिए एक विजुअलाइज़ेशन सिस्टम। जर्नल ऑफ़ कम्प्यूटेशनल केमिस्ट्री 25: 1605–1612।
- वेंग, जी., वांग, ई., वांग, जेड., लियू, एच., झू, एफ., ली, डी. एवं हौ, टी. (2019). हॉकडॉक: कम्प्यूटेशनल डॉकिंग और एमएम/जीबीएसए के आधार पर प्रोटीन-प्रोटीन कॉम्प्लेक्स की भविष्यवाणी और विश्लेषण करने के लिए एक वेब सर्वर। न्यूक्लिक एसिड रिसर्च। 47:W322-W330।
- जम्पर, जे., इवांस, आर., प्रिट्ज़ेल, ए. एट अल. (2021). अल्फाफोल्ड के साथ अत्यधिक सटीक प्रोटीन संरचना की भविष्यवाणी। प्रकृति 596, 583–589

आणविक गतिशीलता सिमुलेशन का उपयोग करके मंकीपॉक्स प्रोटीन के विरुद्ध फाइटोकेमिकल अवरोधकों की पहचान

स्नेहा मुर्मू, सुधीर श्रीवास्तव, दीपा भट्ट, भारती पांडे, सौम्या शर्मा, ऋत्विका दास, के. के. चतुर्वेदी एवं मोहम्मद समीर फारुकी
भा.कृ.अनु.प.—भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली

सारांश

अप्रैल 2022 से पहले, मनुष्यों में मंकीपॉक्स वायरस एम.पी.एक्स.वी. (MPXV) संक्रमण शायद ही कभी अफ्रीकी क्षेत्रों (जहां यह स्थानिक है) के बाहर रिपोर्ट किया गया था। अब दुनियाभर में मामले सामने आ रहे हैं। एम.पी.एक्स.वी. के जीनोमिक्स और महामारी विज्ञान पर व्यापक जानकारी के बावजूद, एम.पी.एक्स.वी. के प्रोटीओम का अध्ययन अच्छी तरह से नहीं किया गया है। कंप्यूटर मॉडल के अलावा एम.पी.एक्स.वी. जीनोम से एन्कोडेड लगभग कोई प्रोटीन संरचना नहीं है, A42R को छोड़कर, जो एक ज्ञात संरचना वाला एकमात्र एम.पी.एक्स.वी. —एन्कोडेड प्रोटीन है। वर्तमान में, मंकीपॉक्स संक्रमण के लिए कोई विशिष्ट उपचार उपलब्ध नहीं हैं। वर्तमान अध्ययन में, हमने एकीकृत आणविक मॉडलिंग पद्धति का उपयोग करके मंकीपॉक्स वायरस के A42R के खिलाफ संभावित अवरोधकों के रूप में नीम—व्युत्पन्न फाइटोकेमिकल्स, निंबोसोन, निंबिन, निंबिओनॉल और निंबियोल की प्रभावकारिता का विश्लेषण किया। आणविक डॉकिंग और आणविक डायनामिक (एम.डी.) सिमुलेशन के संयोजन से, हमने पाया कि निंबोसोन ने एम.डी. सिमुलेशन के 100 एनएस के दौरान एक स्थिर कॉम्प्लेक्स का गठन किया। इसके अलावा, आणविक मैकेनिक पॉइसन—बोल्ट्ज़मान सतह क्षेत्र (एम.एम.पी.बी.एस.ए.) मुक्त ऊर्जा गणना से पता चला है कि इस अणु में स्थिर और अनुकूल ऊर्जा है। यह अध्ययन एक बुनियादी आधार प्रदान करता है और सुझाव देता है कि निंबोसोन, टारगेट के कार्य को विनियमित करने और वायरल प्रतिकृति को नियंत्रित

करने में संभावित अवरोधक के रूप में उपयोगी हो सकता है।

1. परिचय

स्थानिक और गैर—स्थानिक दोनों क्षेत्रों में कई देशों को शामिल करने वाले मंकीपॉक्स के 2022 के प्रकोप ने महत्वपूर्ण अंतर्राष्ट्रीय रुचि पैदा की है (टाइटनजी एट अल., 2022)। मंकीपॉक्स पॉक्सविरिडे परिवार, कॉर्डोपोकसविरिनए सबफैमिली, और जीनस ऑर्थोपॉक्सवायरस से संबंधित है। इंसानों में मंकीपॉक्स बीमारी की शुरुआत सिरदर्द, बुखार और पलू जैसे लक्षणों से होती है। हालांकि अधिकांश रोगी दो से चार सप्ताह में संक्रमण का समाधान कर लेते हैं, लेकिन कुछ मामलों में यह रोग घातक हो सकता है। विश्व स्वास्थ्य संगठन ने मानव मंकीपॉक्स को वैश्विक चिंता का एक सार्वजनिक स्वास्थ्य आपातकाल घोषित किया है (माइंडस्पंज, 2022)।

ऑर्थोपॉक्सवायरस बड़े वायरस होते हैं जिनकी संरचना ईंट जैसी होती है, जिसमें लगभग 200–500kb का एक जीनोम होता है जो 200 से अधिक जीनों को कोड करता है। ओ.पी.एक्स.वी. जीनोम द्वारा एन्कोड किए गए कई जीन सेल कल्चर में वायरस प्रतिकृति के लिए आवश्यक नहीं हैं, लेकिन होस्ट एंटीवायरल प्रतिक्रिया में यह जीन महत्वपूर्ण भूमिका निभा सकते हैं (मैकफैडेन एट अल., 2005)। नेशनल सेंटर फॉर बायोइन्फॉर्मेटिक इंफॉर्मेशन (एनसीबीआई) जेनबैंक में 2022 प्रकोप के स्ट्रेन के 400 से अधिक जीनोम पहले से ही उपलब्ध हैं। ऑर्थोपॉक्सवायरस के जीनोमिक्स पर व्यापक शोध के बावजूद, एम.पी.एक्स.वी. के प्रोटीओम

एवं प्रोटीन संरचना का अच्छी तरह से अध्ययन नहीं किया गया है। A42R एकमात्र ज्ञात संरचना वाला एम.पी.एक्स.वी.-एन्कोडेड प्रोटीन है इसके अलावा एमपीएक्सवी जीनोम से एन्कोडेड लगभग किसी भी प्रोटीन की संरचना ज्ञात नहीं है। एम.पी.एक्स.वी. प्रोटीन। 42R की 1.52Å एक्स-रे संरचना, प्रोटीन डेटाबैंक (पी. डी. बी) में आई डी: 4QWO के साथ उपलब्ध है। यह प्रोटीन इन विट्रो में पॉक्सवायरस की प्रतिकृति के लिए आवश्यक पाया गया है (मिनासोव एट अल., 2022)। वर्तमान में, मंकी पॉक्स संक्रमण के लिए कोई विशिष्ट उपचार उपलब्ध नहीं हैं। चेचक के लिए जो एंटीवायरल दवाएं हैं उनका उपयोग कुछ परिस्थितियों में मंकी पॉक्स के इलाज के लिए भी किया जा सकता है। टेकोविरमट चेचक के टीके, सिडोफोविर, ब्रिन्किडोफोविर और वक्सीनिआ प्रतिरक्षा ग्लोब्युलिन (वी.आई.जी) का उपयोग मंकीपॉक्स के प्रकोप को नियंत्रित करने के लिए किया जा सकता है (झू एट अल., 2022)। हालांकि, मंकीपॉक्स की जटिलताओं के उपचार में सिडोफोविर, ब्रिनसीडोफोविर और वीआईजी की प्रभावशीलता पर कोई डेटा उपलब्ध नहीं है।

फ्लेवोनोइड्स, टेरपीनोइड्स, लिग्निन, अल्कलॉइड्स और क्यूमरिन जैसे फाइटोकेमिकल्स का उपयोग पारंपरिक रूप से कई बीमारियों के लिए किया गया है, और यह वायरल प्रतिकृति/ट्रांसक्रिप्शन को बाधित करने के लिए भी वर्णित है। इनमें से अधिकांश या तो मेजबान कोशिका के अंदर वायरल प्रवेश के दौरान या उनकी प्रतिकृति के दौरान वायरस को रोकते हैं। अजदिरचता इंडिका (नीम), का उपयोग सदियों से कई मानव रोगों के पारंपरिक उपचार के लिए किया जाता रहा है। हालांकि अधिकांश अध्ययनों ने नीम की जीवाणुरोधी और कवकरोधी क्षमता की जांच की है, लेकिन हाल के कुछ कार्यों ने नीम की विषाणुरोधी गतिविधियों का भी पता लगाया है। आज तक, अधिकांश प्रकाशन मानव इम्यूनोडेफिशिएंसी वायरस (एचआईवी), साथ ही साथ हरपीज, डेंगू और इन्फ्लूएंजा वायरस के आसपास केंद्रित रहा है। हाल की कुछ रिपोर्टों में सार्स-कोव-2 भी शामिल है, जो कोविड-19

महामारी के लिए जिम्मेदार है। इसके अलावा, कुछ समूहों ने अन्य वायरस के खिलाफ नीम उत्पादों के उपयोग के साथ सफलतापूर्वक प्रयोग किया है, जैसे कि जापानी एन्सेफलाइटिस वायरस (द्विवेदी एट अल., 2021), हेपेटाइटिस सी (अशफाक एट अल., 2016), और कॉक्ससेकी वायरस (यूनुस एट अल., 2016)। जैसा कि हाल के अध्ययनों से स्पष्ट है की वायरल रोगों के खिलाफ नीम-व्युत्पन्न फाइटोकेमिकल्स का संतुष्टिजनक प्रभाव है, इस कारण ने हमें मंकीपॉक्स वायरस के खिलाफ इसके प्रभाव का पता लगाने के लिए प्रेरित किया।

इस अध्ययन में, हमने उन नीम-व्युत्पन्न फाइटोकेमिकल्स का उपयोग करके A42R प्रोटीन को लक्षित करने की कोशिश की है जिनका एंटीवायरल प्रभाव वर्णित है। इन फाइटोकेमिकल्स में निबोसोन, निबिन और निबिओल (आरा एट अल., 1998) शामिल हैं। फाइटोकेमिकल्स की बाध्यकारी आकर्षण (बाइंडिंग एफिनिटी) खोजने के लिए आणविक डॉकिंग का उपयोग किया गया। लिगैंड्स और वायरल प्रोटीन की अनुमानित बाध्यकारी ऊर्जा का आकलन करने के लिए शीर्ष स्कोरिंग यौगिकों को एम.डी. सिमुलेशन और पॉइसन-बोल्ट्जमैन सतह क्षेत्र सातत्य सॉल्वेशन (एमएम-पीबीएसए) विधि का उपयोग किया गया।

2. विधि

2.1 प्रोटीन-लिगैंड संरचना का रिट्रिवल और आणविक डॉकिंग अध्ययन

आरसीएसबी- पीडीबी डेटाबेस से प्रोटीन A42R (आई डी: 4 क्यू डब्ल्यू ओ) की एक उच्च रिज़ॉल्यूशन (1.52Å) वाली क्रिस्टल संरचना प्राप्त की गई।

आगे बढ़ने से पहले, लिगैंड संरचनाओं को पबकेम (किम एट अल., 2016) से संरचना डेटा फाइलों (एसडीएफ) के रूप में डाउनलोड किया गया और पाईमोल का उपयोग करके पीडीबी में परिवर्तित किया गया था। आणविक डॉकिंग (मॉरिस एट अल., 2009) के लिए ऑटोडॉकटूल (एडीटी) का उपयोग करके प्रोटीन और लिगैंड को तैयार किया गया।

2.2 एम.डी. सिमुलेशन

A42R के साथ चार फाइटोकेमिकल्स, निंबोसोन, निंबिन, निंबियोनॉल और निम्बियोल के डॉक कॉम्प्लेक्सेस की स्थिरता का आकलन करने के लिए ग्रोमैक्स 2021 पैकेज (वैन डेर स्पोएल एट अल., 2005) द्वारा एम.डी. सिमुलेशन किया गया। इन यौगिकों को उनके आणविक डॉकिंग स्कोर के आधार पर जांचा गया और छं+ आयनों को डालकर प्रोटीन-लिगैंड प्रणाली को न्युट्रिलाइज कर दिया गया। न्युट्रिलाइज प्रणाली को दो चरणों में संतुलित किया गया: पहले चरण में, तापमान को एन.वी.टी. एनसैमबल के साथ संतुलित किया गया था जहां 'एन' कणों की स्थिर संख्या है, 'वी' मात्रा है, और 'टी' तापमान है। दूसरे चरण में, एन.पी.टी. एनसैमबल के साथ 300K पर दबाव को संतुलित किया जाता है, जहां 'एन' कणों की स्थिर संख्या है, 'पी' दबाव है, और 'टी' तापमान है। संतुलन के बाद, समष्टि के मोशनल व्यवहार में विस्तृत जानकारी प्राप्त करने के लिए अंततः 100 एनएस का प्रोडक्शन रन स्थापित किया गया।

2.3 प्रक्षेपवक्र विश्लेषण

एम.डी. सिमुलेशन के पूरा होने पर, प्राप्त प्रक्षेपवक्र का विश्लेषण, मूल माध्य वर्ग विचलन (आर.एम.एस.डी.), मूल माध्य वर्ग फ्लक्चुएशन (आर.एम.एस.एफ.), इंटरमॉलिक्युलर हाइड्रोजन बॉन्ड, गार्डरेशन की त्रिज्या (आर.जी.), और विलायक सुलभ सतह क्षेत्र (एस.ए.एस.ए.) गणना के लिए ग्रोमैक्स पैकेज के इन-बिल्ट टूल्स का उपयोग करके किया गया।

2.4 बाध्यकारी मुक्त ऊर्जा (बाइंडिंग फ्री एनर्जी) गणना

आणविक यांत्रिकी/पॉइसन-बोल्ट्ज़मान सतह क्षेत्र (एम. एम.पी. बी. एस. ए.) विधि को जी.एम.एक्स-एम.एम.पी.बी.एस.ए. उपकरण का उपयोग करके नियोजित किया गया ताकि 100 एनएस के ग्रोमैक्स एम.डी. प्रक्षेपवक्र से डॉक किए गए समष्टियों के बीच संबंध की बाध्यकारी मुक्त ऊर्जा का अनुमान लगाया जा सके (कुमारी एट अल., 2014)।

प्रोटीन और लिगैंड के दिए गए समष्टियों के लिए बाध्यकारी मुक्त ऊर्जा की गणना निम्नलिखित समीकरण का उपयोग करके की जा सकती है:

$$\Delta\text{जी-बाइंड} = \text{जी-कॉम्प्लेक्स} - \text{जी-प्रोटीन} - \text{जी-लिगैंड}$$

जहां जी-बाइंड समष्टियों की बाइंडिंग एफिनिटी है, जी-कॉम्प्लेक्स डॉक किए गए कॉम्प्लेक्स की कुल मुक्त ऊर्जा को दर्शाता है, जबकि जी-प्रोटीन और जी-लिगैंड क्रमशः पानी या विलायक में अनबाउंड प्रोटीन और लिगैंड की कुल मुक्त ऊर्जा का प्रतिनिधित्व करता है।

3. परिणाम और चर्चा

3.1 फाइटोकेमिकल्स के साथ ई.बी.वी.-I42R की आणविक डॉकिंग

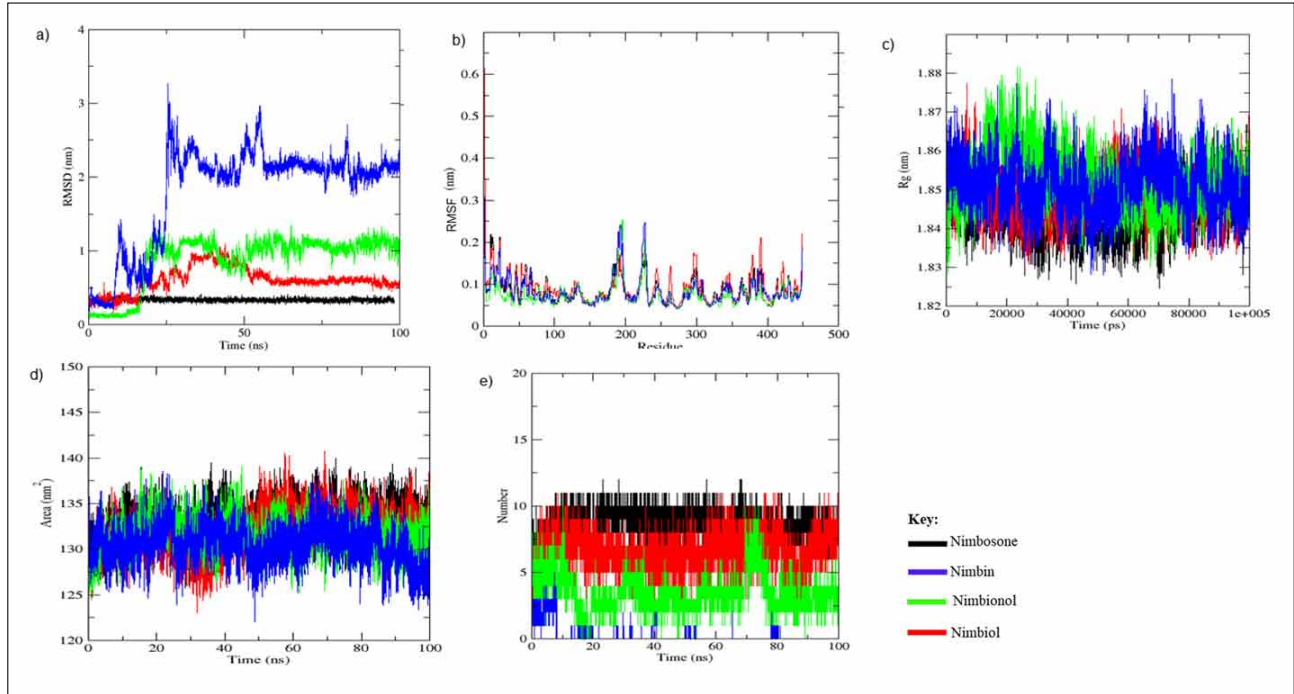
हमने बाइंडिंग के निम्नतम आर.एम.एस.डी मूल्यों के आधार पर फाइटोकेमिकल्स और टारगेट प्रोटीन A42R के सबसे विश्वसनीय बाइंडिंग पोज़ का चयन किया। चार लिगैंड्स, अर्थात, निंबोसोन, निंबिन, निम्बियोनॉल और निम्बियोल की बंधन मुक्त ऊर्जा -7.1 से -7.8 किलो कैलोरी/मोल तक थी। इसके बाद, चार समष्टियों का एम.डी. सिमुलेशन विश्लेषण किया गया।

3.2 एम.डी. सिमुलेशन विश्लेषण

डॉक किए गए कॉम्प्लेक्सेस को उनकी संरचनात्मक स्थिरता का आकलन करने के लिए 100 एनएस के लिए एम.डी. सिमुलेशन के अधीन किया गया था। सिमुलेशन के सफल समापन के बाद प्राप्त प्रक्षेपवक्र से आरएमएसडी, आर.एम.एस.एफ, हाइड्रोजन बांड, आर.जी और एस.ए.एस.ए की गणना करके सिम्युलेटेड सिस्टम की स्थिरता का और अधिक मूल्यांकन किया गया था।

3.2.1 विचलन-वर्ग-माध्य मूल

डॉक किए गए समष्टियों की स्थिरता का आकलन करने के लिए आर.एम.एस.डी आमतौर पर इस्तेमाल किए जाने वाले मात्रात्मक उपायों में से एक है। छोटे



चित्र 1: आणविक गतिकी विश्लेषण: ए) आर.एम.एस.डी. बी) आर.एम.एस.एफ. सी) जाइरेशन की त्रिज्या, डी) सॉल्वेंट सुलभ घुलनशील क्षेत्र और इ) हाइड्रोजन बांड विश्लेषण

विचलन डॉक किए गए समष्टियों की अधिक स्थिरता का संकेत देते हैं। जैसा कि चित्र 1ए से स्पष्ट है, निंबोसोन वाले समष्टि की औसत आर.एम.एस.डी मान लगभग 0.1 एनएम है, जो सभी पांच परिसरों में सबसे कम था। निंबियोनॉल-A42R समष्टि ने 70 एनएस पर ~1 एनएम के नियत आर.एम.एस.डी मान तक पहुंचने से पहले, जहां यह निंबोसोन के साथ ओवरलैप होता है, 20 एनएस के आसपास प्रमुख उतार-चढ़ाव और 50 और 60 एनएस के आसपास दो मामूली उतार-चढ़ाव दिखाए हैं। निंबिन के साथ A42R समष्टि ने पूरे सिमुलेशन में अस्थिरता दिखाई है। इन अवलोकनों के आधार पर, हम मानते हैं कि A42R और निंबिन समष्टि में सबसे अधिक रूपात्मक परिवर्तन होते हैं, इसके बाद निंबियोनॉल, निंबोसोन और सबसे कम परिवर्तन निंबिओल समष्टि द्वारा होता है। नीचे दिए गए आर.एम.एस.एफ. प्लॉट द्वारा इन समष्टियों के अवशेष स्तर पर उनके स्थानीय परिवर्तनों की अस्थिरता की पुष्टि की गई। लिगैंड-बाउंड प्रोटीन के आरएमएसडी विश्लेषण ने सबसे कम आरएमएसडी मान निंबिओल के लिए

और उसके बाद निंबोसोन के लिए दिखाया। चार फाइटोकेमिकल लिगैंड्स में से 100 एनएस सिमुलेशन के दौरान निंबोसोन, टारगेट प्रोटीन A42R से सबसे अधिक मजबूती से बांधता है।

3.2.2 वर्ग माध्य मूल अस्थिरता

आर.एम.एस.एफ. का उपयोग समय के साथ सिस्टम के प्रत्येक रेसिडु की फ्लैक्सिबिलिटी को मापने के लिए किया जाता है। उच्च स्कोर अस्थिर और लचीले बंधनों को दर्शाता है, जबकि निचला स्कोर प्रोटीन-लिगैंड यौगिक में स्थिर क्षेत्रों को इंगित करता है। सभी चार समष्टियों के सी α परमाणुओं का आर.एम.एस.एफ. निर्धारित किया गया।

निंबियोनॉल, निंबोसोन, निंबिन और निंबिओल के साथ गठित नियंत्रण नमूने और समष्टियों का औसत आर.एम.एस.एफ. मान क्रमशः ~0-09, ~0-084, ~0-078, ~0-074, ~0-084, और ~0-075 एनएम था, जैसा कि चित्र 1 बी में दर्शाया गया है। समष्टियों के कुछ उतार-चढ़ाव, फाइटोकेमिकल्स और प्रोटीन की सक्रिय

साइट पर वितरित फाइटोकेमिकल्स और रेसिडु के बीच अभिप्रायपूर्ण बातचीत का संकेत देते हैं। निंबोसोन के लिए सी α परमाणुओं के कम आर.एम.एस.एफ. मान ने A42R के लिए उनके स्थिर बंधन की पुष्टि की। निंबोसोन (0.74 एनएम) के लिए सी α परमाणुओं का आर.एम.एस.एफ. मान निंबियोल (0.75 एनएम) के करीब है।

3.2.3 गायरेशन की त्रिज्या और विलायक सुलभ सतह क्षेत्र विश्लेषण

आरजी प्रोटीन और प्रोटीन-लिगैंड समष्टियों की सघनता और समावेशी आयाम को दर्शाता है जिसमें उनकी उचित परस्पर क्रिया शामिल हो सकती है। स्थिर रूप से चलन प्रोटीन अपेक्षाकृत संतुलित आरजी मान बनाए रख सकता है। जैसा कि चित्र 1सी से स्पष्ट है, सभी चार प्रणालियों ने पूरे सिमुलेशन में 1.842 और 1.852 एनएम के बीच एक बहुत ही संकीर्ण सीमा के साथ एक जैसे आरजी मान दिखाया है। इससे पता चलता है कि फाइटोकेमिकल्स कॉम्प्लेक्स सघन हैं। विलायक सुलभ सतह क्षेत्र (एस.ए.एस.ए) एक और महत्वपूर्ण मापदंड है जो विलायक अणु के सुलभ क्षेत्र की गणना करता है। अनुमानित एस.ए.एस.ए मानों में भी एक समान पैटर्न है, जो 130.5 और 133 एनएम² के बीच में है, जैसा कि चित्र 1 डी में दिखाया गया है, उच्चतम मूल्य निंबिओल का है। ये अवलोकन सभी प्रणालियों की स्थिरता की पुष्टि करते हैं। औसत आरजी और एस.ए.एस.ए मानों में छोटे विचलन भी प्रोटीन-लिगैंड समष्टियों की अधिक सघनता और मजबूत बंधन का संकेत देते हैं।

3.2.4 हाइड्रोजन बांड विश्लेषण

प्रोटीन-लिगैंड समष्टियों की अंतःक्रियाओं की विशिष्टता और बाध्यकारी ताकत का निर्धारण करने में हाइड्रोजन बांड विश्लेषण आवश्यक है। जैसा कि चित्र 1ई में दर्शाया गया है, निंबियोल के साथ गठित समष्टि ने लगातार सबसे उच्च संख्या में इसके बाद निंबियोनॉल समष्टि ने 100 एनएस के अनुकरण के दौरान हाइड्रोजन बांड दिखाए। इन्होंने सिमुलेशन के अलग-अलग समय पर दस से अधिक हाइड्रोजन बांड भी दिखाए। दूसरी ओर, निंबिन ने सिमुलेशन के शुरुआती बिंदु की तुलना में हाइड्रोजन बांड की संख्या में कमी दिखाई, जबकि निंबोसोन ने पूरे सिमुलेशन में स्थिर संख्या में हाइड्रोजन बांड दिखाए।

3.2.5 बंधन मुक्त ऊर्जा अनुमान

प्रत्येक समष्टि के लिए बाध्यकारी मुक्त ऊर्जा का अनुमान लगाने के लिए एम.एम-पी.बी.एस.ए पद्धति का उपयोग किया गया। समष्टियों की अनुमानित बाध्यकारी ऊर्जा जितनी कम होगी, उसकी स्थिरता उतनी ही अधिक होगी। तालिका 1 प्रत्येक समष्टि की अनुमानित बाध्यकारी ऊर्जा और अन्य घटकों को दिखाती है जो आणविक अंतःक्रियाओं में योगदान करते हैं, जैसे वैन डेर वाल्स ऊर्जा, इलेक्ट्रोस्टैटिक ऊर्जा, और सॉल्वेशन ऊर्जा। परिणाम बताते हैं कि सभी प्रकार की ऊर्जा ने अंतःक्रियाओं में महत्वपूर्ण योगदान दिया। उच्चतम नकारात्मक अनुमानित बाध्यकारी ऊर्जा वाले फाइटोकेमिकल निंबोसोन का उपयोग A42R रिसेप्टर के लिए संभावित माड्युलेटर्स के रूप में किया जा सकता है।

तालिका 1. बाध्य परिसरों का एम.एम-पी.बी.एस.ए विश्लेषण

समष्टि	अनुमानित बाध्यकारी ऊर्जा (किलोजूल/मोल)	वैन डेर वाल्स ऊर्जा (किलोजूल/मोल)	इलेक्ट्रोस्टैटिक ऊर्जा (किलोजूल/मोल)	सॉल्वेशन ऊर्जा (किलोजूल/मोल)
निंबियोल	-15.703	-1.828	-10.946	-2.927
निंबोसोन	-16.723	-1.849	-11.649	-3.267
निंबिन	-16.396	-1.685	-11.678	-3.030
निंबियोनॉल	-16.254	-1.769	-11.649	-2.925

4. निष्कर्ष

A42R प्रोटीन मंकीपॉक्स वायरस का एकमात्र वायरस प्रोटीन है जिसे प्रयोगात्मक रूप से मान्य संरचना के लिए जाना जाता है। यह अध्ययन, निष्कर्ष निकालता है कि नीम-व्युत्पन्न चार प्राकृतिक फाइटोकेमिकल समष्टियां निंबोसोन, निबिन, निम्बियोनॉल और निम्बियोल को A42R के खिलाफ संभावित अवरोधक के रूप में उपयोग किया जा सकता है। आणविक डॉकिंग विश्लेषण और 100एनएस के एम.डी. सिमुलेशन से पता चला कि निंबोसोन अणु में अन्य तीन अणुओं की तुलना में कम बाध्यकारी ऊर्जा है। इस वर्तमान अध्ययन से प्राप्त जानकारी का उपयोग भविष्य में मंकीपॉक्स वायरस के खिलाफ अधिक फाइटोकेमिकल आधारित चिकित्सा विज्ञान के विकास के लिए किया जा सकता है।

संदर्भ

टाइटनजी, बी. के., तेगोमोह, बी., नेमातोल्लाही, एस., कोनोमोस, एम., एवं कुलकर्णी, पी. ए. (2022). मंकीपॉक्स: स्वास्थ्य पेशेवरों के लिए एक समकालीन समीक्षा। इन ओपन फोरम इंफेक्शंस डिजीज (वॉल्यूम 9, नंबर 7, पी। ऑफएसी 310)। ऑक्सफोर्ड यूनिवर्सिटी प्रेस।

ब्रेमन, जे. जी., रूटी, के., स्टेनियोवस्की, एम. वी., जानोतो, ई., ग्रोमीको, ए. आई., अरीता, आई., एवं विश्व स्वास्थ्य संगठन (1980). ह्यूमन मंकीपॉक्स 1970-1979/जेजी ब्रेमन, कालिसा-रूटी, एमवी स्टेनियोवस्की, ई. जानोतो, एआई ग्रोमीको और आई. अरीता द्वारा।

मैकफैडेन, जी. (2005). पॉक्सवायरस ट्रॉपिज्म। नेचर रिव्यूज माइक्रोबायोलॉजी, 3(3), 201-213।

माइंडस्पंज, ए. आई. एस. डी. एल. (2022). क्या मंकीपॉक्स के वैश्विक प्रकोप के साथ एक और इंफोडेमिक होगा?

मिनासोव, जी., इनिस, एन. एल., शुवालोवा, एल., एंडरसन, डब्ल्यू. एफ., एवं सैचेल, के. जे. (2022). मंकीपॉक्स प्रोफिलिन जैसे प्रोटीन A42R की संरचना सेलुलर प्रोफाइल से संभावित कार्य अंतर

को प्रकट करती है। बायोरेक्सिव।

झू. एम., जी. जे., शि. डी., लू. एक्स., वांग, बी., वू. एन., ... एवं ली, एल. (2022). मंकीपॉक्स का असामान्य वैश्विक प्रकोप: हमें क्या करना चाहिए?. फ्रंटियर्स ऑफ मेडिसिन, 1-11।

द्विवेदी, वी. डी., सिंह, ए., एल-काफ्रावे, एस. ए., अलंदिजनी, टी. ए., फैज़ो, ए. ए., बजराय, एल. एच., ... एवं अजहर, ई. आई. (2021). अज़ादिराछा इंडिका के बायोफ्लेवोनोइड्स द्वारा जापानी इंसेफेलाइटिस वायरस आरएनए डिपेंडेंट आरएनए पोलीमरेज़ प्रोटीन निषेध में यांत्रिक अंतर्दृष्टि। साइंटिफिक रिपोर्ट, 11(1), 1-13.

अशफाक, यू. ए., जलील, ए., एवं उल कमर, एम. टी. (2016). अज़ादिराछा इंडिका से एंटीवायरल फाइटोकेमिकल्स की पहचान एचसीवी एनएस 3 प्रोटीज के खिलाफ निकलती है: एक सिलिको दृष्टिकोण। नेचुरल प्रोडक्ट्स रिसर्च, 30(16), 1866-1869।

यूनस, आई., सिद्दीक, ए., इशाक, एच., अनवर, एल., बदर, एस., एवं अशरफ, एम. (2016). इन विट्रो में पैर और मुंह रोग वायरस के खिलाफ पौधे के अर्क की एंटीवायरल गतिविधि का मूल्यांकन। क्लीन जे फार्म साइंस, 29(4), 1263-1268।

आरा, आई., सिद्दीकी, बी. एस., फैज़ी, एस., और सिद्दीकी, एस. (1988). अज़ादिराछा इंडिका के तने की छाल से ट्राइसाइक्लिक डाइटरपेनोइड्स। जर्नल ऑफ नेचुरल प्रोडक्ट्स, 51(6), 1054-1061।

किम, एस., थिएसेन, पी. ए., बोल्टन, ई. ई., चैन, जे., फू, जी., गिन्डुलाइट, ए., ... एवं ब्रायंट, एस. एच. (2016). पबकेम पदार्थ और यौगिक डेटाबेस। न्यूक्लिक एसिड रिसर्च, 44(व1), व1202-व1213।

मॉरिस, जी. एम., ह्युई, आर., लिंडस्ट्रॉम, डब्ल्यू., सननर, एम. एफ., बेलेव, आर. के., गुडसेल, डी. एस., एवं ओल्सन, ए. जे. (2009). ऑटोडॉक4 और ऑटोडॉकटूल्स4: चयनात्मक रिसेप्टर लचीलेपन के साथ स्वचालित डॉकिंग। जर्नल ऑफ कम्प्यूटेशनल केमिस्ट्री, 30(16), 2785-2791।



वैन डेर स्पोएल, डी., लिंडाहल, ई., हेस, बी., ग्रोएनहोफ़, जी., मार्क, ए.ई., एवं बेरेन्डसन, एच.जे. (2005). ग्रोमेक्स: तेज़, लचीला और फ्री। जर्नल ऑफ़ कम्प्यूटेशनल केमिस्ट्री, 26(16), 1701–1718।

कुमारी, आर., कुमार, आर., ओपन सोर्स ड्रग डिस्कवरी कंसोर्टियम, एवं लिन, ए. (2014). जी-ऍमऍमपीबीएस ए: उच्च-थ्रूपुट ऍमऍम-पीबीएसए गणनाओं के लिए एक ग्रोमेक्स उपकरण। जर्नल ऑफ़ केमिकल इनफार्मेशन एंड मॉडलिंग, 54(7), 1951–1962।

संख्यकी-विमर्श

2021

अंक
17

राजभाषा
खण्ड

संस्थान की राजभाषा यात्रा: वर्ष 2021

उदय वीर सिंह

भारतीय कृषि सांख्यिकी अनुसंधान संस्थान में वर्ष-दर-वर्ष हिन्दी के प्रगामी प्रयोग में अभिवृद्धि हो रही है। राजभाषा नीति को संस्थान में सुचारु रूप से कार्यान्वित किया जा रहा है। भारत सरकार, गृह मंत्रालय, राजभाषा विभाग द्वारा जारी वार्षिक कार्यक्रम में निहित लक्ष्यों को इस संस्थान में लगभग पूरा कर लिया गया है। संस्थान द्वारा समस्त प्रशासनिक कार्य शत-प्रतिशत हिन्दी में और यथाआवश्यक द्विभाषी हो रहा है।

संस्थान में राजभाषा हिन्दी की प्रगति का जायजा लेने के लिए उपमहानिदेशक (शिक्षा), भारतीय कृषि अनुसंधान परिषद् **मुख्यालय द्वारा 05 अक्टूबर 2021** को संस्थान का राजभाषा सम्बन्धी निरीक्षण किया गया। उपमहानिदेशक (शिक्षा), महोदय द्वारा निरीक्षण रिपोर्ट में संस्थान में हिन्दी में हो रहे कार्यों की प्रगति पर संस्थान की सराहना की गयी है।

संस्थान में प्रशासनिक कार्य के अतिरिक्त वैज्ञानिक प्रकृति के कार्यों में भी हिन्दी के उपयोग को प्रोत्साहित किया जाता है। साथ ही, हिन्दी के प्रयोग में गुणवत्ता की ओर भी ध्यान दिया जा रहा है। संस्थान के विभिन्न प्रभागों के वैज्ञानिकों द्वारा आयोजित प्रशिक्षण कार्यक्रमों की ई-संदर्भ पुस्तिकाओं में आमुख, प्राक्कथन एवं कवर पेज द्विभाषी रूप में प्रस्तुत करने के साथ-साथ वैज्ञानिकों ने अपनी परियोजना रिपोर्टों के कवर पेज, आमुख, प्राक्कथन एवं सारांश द्विभाषी रूप में प्रस्तुत किये तथा कुछ वैज्ञानिकों द्वारा अपनी परियोजना रिपोर्टों में विषय-सूची एवं तालिकाएँ भी द्विभाषी रूप में प्रस्तुत की गयीं। संस्थान के वैज्ञानिकों द्वारा हिन्दी में वैज्ञानिक विषय पर हिन्दी कार्यशालाओं का आयोजन किया गया। इसके अतिरिक्त, संस्थान में एम.एस.सी. तथा पीएच.डी. के विद्यार्थियों द्वारा अपने शोध-प्रबन्धों में द्विभाषी रूप में सार प्रस्तुत किये गये। वैज्ञानिकों एवं तकनीकी कर्मियों द्वारा शोध-पत्र हिन्दी में प्रकाशित

किये गये।

राजभाषा विभाग द्वारा जारी वार्षिक कार्यक्रम में निहित लक्ष्यों को पूरा करते हुए संस्थान के अधिकारियों/कर्मचारियों द्वारा समस्त पत्राचार हिन्दी में अथवा द्विभाषी रूप में किया गया। संस्थान के विभिन्न वैज्ञानिकों, प्रभागों तथा प्रशासनिक अनुभागों द्वारा आयोजित बैठकों की कार्यसूची तथा कार्यवृत्त शत-प्रतिशत हिन्दी में अथवा द्विभाषी रूप में जारी किये गये। गृह मंत्रालय, राजभाषा विभाग द्वारा जारी विभिन्न नकद पुरस्कार योजनाएँ संस्थान में लागू हैं तथा संस्थान के कर्मियों ने इन योजनाओं में भाग लिया।

प्रतिवेदनाधीन अवधि के दौरान संस्थान में राजभाषा कार्यान्वयन समिति की बैठकें (ऑन-लाइन) आयोजित की गयीं। इन बैठकों में राजभाषा अधिनियम 1963 की धारा 3(3) के अनुपालन को सुनिश्चित करने, राजभाषा विभाग द्वारा जारी वार्षिक कार्यक्रम के विभिन्न मदों, राजभाषा विभाग एवं परिषद् मुख्यालय से समय-समय पर प्राप्त निदेशों का अनुपालन सुनिश्चित करने, कार्यशालाओं के नियमित आयोजन, हिन्दी पत्रिका सांख्यिकी विमर्श के प्रकाशन, हिन्दी पखवाड़ा के आयोजन इत्यादि पर विस्तार से चर्चा हुई।

प्रतिवेदनाधीन अवधि के दौरान संस्थान के विभिन्न वर्गों के कर्मियों के लिए पाँच हिन्दी कार्यशालाएँ (ऑन-लाइन) आयोजित की गयीं। पहली कार्यशाला जनवरी-मार्च 2021 तिमाही के दौरान संस्थान के सांख्यिकी आनुवंशिकी प्रभाग के वैज्ञानिकों, डॉ. उपेंद्र कुमार प्रधान एवं डॉ. समरेन्द्र दास द्वारा 18 से 20 मार्च 2021 तक "सांख्यिकीय आनुवंशिकी और कृषि में इसके अनुप्रयोग" विषय पर आयोजित की गयी जिसमें 14 वक्ताओं द्वारा विषय से संबन्धित 15 उप-विषयों पर व्याख्यान दिए गए। इस कार्यशाला में संस्थान के 08

वैज्ञानिकों एवं तकनीकी कर्मियों के अतिरिक्त भारतीय कृषि अनुसंधान परिषद के 14 राज्यों के 22 संस्थानों से 30 वैज्ञानिकों एवं तकनीकी कर्मियों द्वारा सहभागिता की गयी। इस अवसर पर कार्यशाला आयोजकों द्वारा हिन्दी में ई-संदर्भ संहिता भी तैयार की गयी।

दूसरी कार्यशाला अप्रैल-जून 2021 तिमाही के दौरान 24-26 जून 2021 को संस्थान के **पूर्वानुमान एवं कृषि प्रणाली मॉडलिंग** के प्रधान वैज्ञानिक, डॉ. वसी आलम, डॉ. प्रवीण आर्य प्रधान वैज्ञानिक, और डॉ. कंचन सिन्हा, वैज्ञानिक" द्वारा "कृषि में सांख्यिकीय मॉडलिंग एवं पूर्वानुमान" विषय पर आयोजित की गयी जिसमें 11 वक्ताओं द्वारा विषय से संबंधित 12 उप-विषयों पर व्याख्यान दिए गए। इस कार्यशाला में संस्थान के 14 वैज्ञानिकों, 05 तकनीकी कर्मियों के अतिरिक्त भारतीय कृषि अनुसंधान परिषद के 05 राज्यों के 06 संस्थानों से 05 वैज्ञानिकों एवं 05 तकनीकी कर्मियों द्वारा सहभागिता की गयी। इस अवसर पर कार्यशाला आयोजकों द्वारा हिन्दी में संदर्भ संहिता भी तैयार की गयी।

तीसरी कार्यशाला जुलाई-सितम्बर 2021 तिमाही के दौरान 28-30 सितम्बर, 2021 को संस्थान के **परीक्षण अभिकल्पना प्रभाग** के प्रधान वैज्ञानिक, डॉ. अनिल कुमार, डॉ. सुशील कुमार, वैज्ञानिक एवं डॉ. सुकान्त दाश, वैज्ञानिक द्वारा "परीक्षण अभिकल्पना के अनुप्रयोग" विषय पर आयोजित की गयी जिसमें 11 वक्ताओं द्वारा विषय से संबंधित 14 उप-विषयों पर व्याख्यान दिए गए। इस कार्यशाला में संस्थान के 11 वैज्ञानिकों के अतिरिक्त भारतीय कृषि अनुसंधान परिषद के अन्य संस्थानों से 07 वैज्ञानिकों द्वारा सहभागिता की गयी।

चौथी कार्यशाला अक्टूबर-दिसम्बर 2021 तिमाही के दौरान 26 नवम्बर 2021 को राजभाषा संबंधी नियमों एवं उपनियमों की जानकारी देने हेतु संस्थान के मुख्य प्रशासनिक अधिकारी एवं कार्यालय प्रधान, श्री ए.के. श्रीवास्तव, हिन्दी एकक के मुख्य तकनीकी अधिकारी एवं प्रभारी, हिन्दी श्री उदय वीर सिंह द्वारा "राजभाषा नीति: कार्यान्वयन" विषय पर

आयोजित की गयी जिसमें 03 वक्ताओं द्वारा विषय से संबंधित 03 उप-विषयों पर व्याख्यान दिए गए। इस कार्यशाला में संस्थान के 40 आधिकार्यों एवं कर्मचारियों ने सहभागिता की।

पाँचवी कार्यशाला अक्टूबर-दिसम्बर 2021 तिमाही के ही दौरान 30 नवम्बर 2021 से 02 दिसम्बर 2021 को संस्थान के प्रतिदर्श सर्वेक्षण प्रभाग के वैज्ञानिक, डॉ. राहुल बनर्जी, डॉ. पंकज दास एवं डॉ. भारती द्वारा "कृषि में प्रतिदर्श तकनीकों का अनुप्रयोग एवं आंकड़ों का विश्लेषण" विषय पर आयोजित की गयी जिसमें 11 वक्ताओं द्वारा विषय से संबंधित 11 उप-विषयों पर व्याख्यान दिए गए। इस कार्यशाला में संस्थान के 05 वैज्ञानिकों के अतिरिक्त भारतीय कृषि अनुसंधान परिषद के अन्य संस्थानों से 05 वैज्ञानिकों एवं 07 तकनीकी आधिकारियों तथा राज्य कृषि विश्वविद्यालय से 09 वैज्ञानिकों द्वारा सहभागिता की गयी।

संस्थान में कार्यरत सभी हिन्दीतर भाषी अधिकारियों/ कर्मचारियों द्वारा हिन्दी ज्ञान सम्बन्धी प्रशिक्षण पूरा किया जा चुका है। आज तक की स्थिति के अनुसार, संस्थान में अब कोई ऐसा हिन्दीतर अधिकारी/कर्मचारी शेष नहीं रह गया है जिसे हिन्दी ज्ञान सम्बन्धी प्रशिक्षण दिया जाना शेष हो। इसके अतिरिक्त, हिन्दी शिक्षण योजना' के अन्तर्गत हिन्दी आशुलिपि के प्रशिक्षण का लक्ष्य पूरा है, अभी तक हिन्दी टंकण के प्रशिक्षण का लक्ष्य भी पूरा था और आशुलिपिकों का लक्ष्य पूरा है। अभी तक हिन्दी टंकण के प्रशिक्षण का लक्ष्य भी पूरा था परन्तु दिसम्बर, 2018 से संस्थान में नव-नियुक्त 06 अवर लिपिकों में से 05 अवर लिपिक हिन्दी शिक्षण योजना' के अन्तर्गत प्रशिक्षण प्राप्त कर चुके हैं। 01 अवर लिपिक को 'हिन्दी शिक्षण योजना' के अन्तर्गत फरवरी 2021 से आरम्भ सत्र में हिन्दी टंकण के प्रशिक्षण हेतु भेजा गया है तथा हाल ही में वर्ग घ से वर्ग ग में पदोन्नत 02 अवर लिपिकों में से 01 अवर लिपिक जुलाई 2017 में प्रशिक्षण प्राप्त कर चुका है। शेष 01 अवर लिपिक को अगस्त, 2021 सत्र से हिन्दी टंकण के प्रशिक्षण हेतु भेजा जा चुका है। इसके साथ ही हिन्दी टंकण के प्रशिक्षण का लक्ष्य भी पूरा हो जायेगा।

संस्थान द्वारा प्रकाशित हिन्दी पत्रिका: 'सांख्यिकी-विमर्श' के सोलेहवें अंक का प्रकाशन मार्च 2021 में किया गया। इस पत्रिका में संस्थान के कीर्तिस्तम्भ, सम्बन्धित वर्ष में किये गये अनुसंधानों व अन्य कार्यों के संक्षिप्त विवरण, राजभाषा से सम्बन्धित कार्यों आदि की जानकारी के साथ-साथ कृषि सांख्यिकी, संगणक अनुप्रयोग एवं कृषि जैव-सूचना विज्ञान से सम्बन्धित विभिन्न लेखों एवं शोध-पत्रों को भी प्रस्तुत किया गया है। पाठकों के हिन्दी ज्ञानवर्धन के लिए दैनिक स्मरणीय शब्द-शतक हिन्दी व अँग्रेजी में दिया गया है।

संस्थान में 01 से 14 सितम्बर 2021 के दौरान हिन्दी पखवाड़े का आयोजन किया गया। कोविड-19 महामारी के कारण इस वर्ष अधिकांश कार्यक्रम/प्रतियोगितायें ऑन-लाइन आयोजित की गयीं। दिनांक 01 सितम्बर, 2021 को हिन्दी पखवाड़े का उदघाटन संस्थान के निदेशक, डॉ. राजेन्द्र प्रसाद जी द्वारा किया गया। हिन्दी पखवाड़े के उदघाटन के पश्चात काव्य-पाठ का आयोजन किया गया। हिन्दी पखवाड़े के दौरान डॉ. दरोगा सिंह स्मृति व्याख्यान' के साथ-साथ प्रभागों में हिन्दी में सर्वाधिक वैज्ञानिक कार्य करने के लिए प्रभागीय चल-शील्ड के साथ-साथ डिजिटल शोध पोस्टर प्रस्तुति, हिन्दीतर कर्मियों के लिए हिन्दी श्रुतलेख एवं शब्दार्थ लेखन प्रतियोगिता आयोजित की गयी। सभी प्रतियोगिताओं में छात्रों सहित संस्थान के

विभिन्न वर्गों 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 जिसमें किसी सुप्रसिद्ध वैज्ञानिक द्वारा किसी भी वैज्ञानिक विषय पर हिन्दी में व्याख्यान दिया जाता है। इस वर्ष इस कड़ी का तीसवां व्याख्यान सांख्यिकी विज्ञान प्रभाग, टोरंटो विश्वविद्यालय, कनाडा के डॉ. मुरारी सिंह जी द्वारा डॉ. दरोगा सिंह की स्मृति को समर्पित विविधता के गुणांक पर विविध दृष्टिकोण: एक चयनात्मक समीक्षा" विषय पर दिया गया और इस कार्यक्रम की अध्यक्षता आई.सी.एम.आर. के पूर्व अपर महानिदेशक एवं राष्ट्रीय सांख्यिकीय आयोग के पूर्व सदस्य, डॉ. पदम सिंह जी द्वारा की गयी। दिनांक 14 सितम्बर, 2021 को हिन्दी पखवाड़ा के समापन समारोह के अवसर पर इस दौरान आयोजित प्रतियोगिताओं के सफल प्रतियोगियों को पुरस्कृत करने के साथ-साथ वर्ष 2020-21 के दौरान "सरकारी कामकाज मूल रूप से हिन्दी में करने के लिए प्रोत्साहन योजना" के अन्तर्गत भी नकद पुरस्कारों की घोषणा की गयी। इसके अतिरिक्त, जुलाई 2020 से सितम्बर, 2021 तक की अवधि के दौरान संस्थान में आयोजित हिन्दी कार्यशालाओं के वक्ताओं एवं संस्थान द्वारा प्रकाशित हिन्दी पत्रिका: सांख्यिकी विमर्श 2020 के संपादक मण्डल के सदस्यों को प्रशस्ति पत्र प्रदान करने की भी घोषणा की गयी।

हिन्दी पखवाड़ा-2021





दैनिक स्मरणीय शब्द-शतक

1	Aberration	अपेरण-त्रुटि
2	Alternative hypothesis	वैकल्पिक परिकल्पना
3	Absolute magnitude	निरपेक्ष कांतिमान
4	Accuracy	यथार्थता
5	Accumulation ratio	संचय अनुपात
6	Activation rate	सक्रियण दर
7	Analytical Statistics	विश्लेषिक सांख्यिकी
8	Average	औसत, माध्य
9	Beta function	बीटा-फलन
10	Bell shaped curve	घंटाकार वक्र
11	Binomial theorem	द्विपद-प्रमेय
12	Breakdown rate	भंजन-दर
13	Capillary curve	केशिका-वक्र
14	Cardinal points	प्रधान बिन्दु
15	Carrier frequency	वाहक तरंग आवृत्ति
16	Cartesian coordinates	कार्तीय निर्देशांक
17	Central confidence interval	केन्द्रीय विश्वास्यता-अंतराल
18	Centrogram	केंद्र-आरेख
19	Certainty	निश्चितता
20	Characteristic data	अभिलक्षणिव दत्त
21	Correlation coefficient	सहसंबंध गुणांक
22	Confidence band	विश्वास्यता पट्टी
23	Coefficient	गुणांक
24	Covariance	सहप्रसरण विधि
25	Deactivation	निषिक्रयण

26	Dead space	निष्क्रय भाग
27	Debugged model	दोषमार्जित निदर्श
28	Decay curve	क्षय वक्र
29	Declassify	विवर्गीकरण
30	Depreciation	अवमूल्यन
31	Dependent variable	परतंत्र चर
32	Differentiable curve	अवकलनीय वक्र
33	Direct sampling	प्रत्यक्ष प्रतिचयन
34	Descriptive statistics	वर्णात्मक सांख्यिकी
35	Discrete	विविक्त
36	Endoergic process	ऊर्जा-शोषी प्रक्रम
37	Equation of state	अवस्था-समीकरण
38	Equivalent conductivity	तुल्यांकी चालकता
39	Even polynomial	सम बहुपद
40	Fluctuation theory	उच्चावचन वाद
41	Factor analysis	उपादान विश्लेषण
42	Frequency distribution	बारंबारता-बंटन
43	Heat of activation	सक्रियण ऊष्मा
44	High frequency current	उच्च आवृत्ति-धारा
45	Homogeneity	समांगता
46	Horizontal acceleration	क्षैतिज त्वरण
47	Host plant	परपोषी-वेधन
48	Hypothesis testing	परिकल्पना परीक्षण
49	Identical sets	सर्वसम समुच्चय
50	Imbibition	अंतः शोषण

51	Immobile	निश्चल, अचल
52	Independent variable	स्वतंत्र चर
53	Inaccurate	अयथार्थ
54	Inferential analysis	आनुमानिक विश्लेषण
55	Indefinite integral	अनिश्चित समाकल
56	Induction current	प्रेरण धारा
57	Inorganic salt	अकार्बनिक लवण
58	Interval rule	अंतराल नियम
59	Inverse	प्रतिलोम
60	Irradiance	किरणित ऊर्जा मान
61	Isostatic	समस्थितिक
62	Joint set	संयुक्त समुच्चय
63	Kinetic energy	गतिज ऊर्जा
64	Kurtosis	ककुदता
65	Laminar motion	स्तरीय गति
66	Multivariate normal distribution	बहुचर प्रसामान्य बंटन
67	Multiple correlation	बहु सहसंबंध
68	Median	मध्य, मध्यस्थ
69	Measure of variability	विचारणशीलता—माप
70	Measure of skewness	वैषम्य—माप
71	Mode	विधि, रीति, प्रणाली, विधा, बहुलक
72	Marginal error	उपांत त्रुटि
73	Null hypothesis	निराकरणीय परिकल्पना
74	Ordered	क्रमित
75	Orthonormal function	प्रसामान्य लंबिक फलन

76	Parametric method	प्राचल विधि
77	Population	समिष्ट
78	Precision	परिशुद्धता
79	Quota sampling	नियतमात्रात्मक प्रतिचयन
80	Quartile	चतुर्थक
81	Random variable	यादृच्छिक चर
82	Reduced equation	समानीत समीकरण
83	Sampling error	प्रतिचयन त्रुटि
84	Statistical inference	सांख्यिकीय अनुमिति
85	Standard deviation	मानक विचलन
86	Symmetric	सममित
87	Tangential	स्पर्शीय
88	Term series	पद श्रेणी
89	Unbiased test	अनभिनत त्रुटि
90	Uncertainty relation	अनिश्चितता—संबंध
91	Unimodal	एकमापांकी
92	Variance	प्रसरण
93	Virulent strain	उग्र विबेध
94	Weighted mean	भारित माध्य
95	Windward	पवानाभिमुख
96	Zero ideal	शून्य गुणजावली
97	Zero order	शून्य कोटि
98	Zone melting	अंचल पिघलन
99	Zero correlation	शून्य सहसंबंध
100	Zoological garden	प्राणि—उपवन

(संकलनकर्ता : नेहा नारंग)
 श्रोत : ब्रह्म पारिभाषिक शब्द—संग्रह
 विज्ञान: खण्ड 1 एवं 2

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
- एनएसएसएम

प्रोफेसर (कृषि सांख्यिकी)
प्रोफेसर (संगणक अनुप्रयोग)
प्रोफेसर (जैवसूचना विज्ञान) वाईन

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



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