भारत में खाद्य मुद्रास्फीति का पूर्वानुमान: मशीन लर्निंग के माध्यम से मूल्य और गैर-मूल्य संकेतों का लाभ उठाना

निशांत सिंह और अभिरुचि राठी द्वारा ^

भारत के उपभोक्ता मूल्य सूचकांक (सीपीआई) में खाद्य पदार्थों का महत्व उनके भार और मूल्य अस्थिरता के संदर्भ में खाद्य मुद्रास्फीति के सटीक पूर्वानुमान को हेडलाइन मुद्रास्फीति अनुमानों के लिए महत्वपूर्ण बनाता है। नाउकास्ट, जो वर्तमान अवधि की मुद्रास्फीति का अनुमान है, लघु और मध्यम-क्षितिज पूर्वानुमानों का मार्गदर्शन करता है। डेटा उपलब्धता में प्रौद्योगिकी-संचालित उछाल को देखते हुए, यह अध्ययन भारत में खाद्य मुद्रास्फीति का नाउकास्टिंग करने के लिए उच्च आवृत्ति मूल्य और गैर-मूल्य संकेतकों की भविष्य कहनेवाला शक्ति की जांच करता है। इसके अलावा, मशीन लर्निंग (एमएल) तकनीकों का उपयोग करते हुए, यह अध्ययन पारंपरिक बेंचमार्क मॉडल पर उनकी उपयोगिता का पता लगाता है। अनुभवजन्य निष्कर्ष संकेत देते हैं कि सूचना सेट का विस्तार करने से नाउकास्ट सटीकता में सुधार होता है, जिसे नियमितीकरण और एमएल विधियों को नियोजित करके और बढ़ाया जाता है।

परिचय

भारतीय रिजर्व बैंक अधिनियम ने 2016 में लचीले मुद्रास्फीति लक्ष्यीकरण (एफआईटी) 1 के लिए एक वैधानिक आधार प्रदान किया और केंद्रीय बैंक को विकास के उद्देश्य को ध्यान में रखते हुए मूल्य स्थिरता बनाए रखने के प्राथमिक उद्देश्य के साथ मौद्रिक नीति संचालित करने का काम सौंपा। विशिष्ट मौद्रिक संचरण अंतराल को देखते हुए, मुद्रास्फीति पूर्वानुमान एक एफआईटी ढांचे में केंद्रीय बैंक के लिए एक मध्यवर्ती लक्ष्य के रूप में कार्य करते हैं, जो मौद्रिक नीति कार्यों और रुख का मार्गदर्शन करते हैं। इस प्रकार, मुद्रास्फीति (साथ ही आर्थिक गतिविधि) के सटीक पूर्वानुमान एफआईटी की सफल खोज और मुद्रास्फीति को टिकाऊ आधार पर लक्ष्य के अनुरूप बनाने में महत्वपूर्ण भूमिका निभाते हैं। भारत में, खाद्य और पेय पदार्थ समूह की उपभोक्ता मूल्य सूचकांक-संयुक्त (सीपीआई-सी) टोकरी में उच्च हिस्सेदारी (45.86 प्रतिशत) है साथ ही, खाद्य कीमतों में महत्वपूर्ण व्यापक आर्थिक निहितार्थ होते हैं क्योंकि उनके उतार-चढ़ाव गरीबों के कल्याण को दृढ़ता से प्रभावित करते हैं (शेखर, एट अल., 2018) और खाद्य सुरक्षा के आसपास के विकास की निगरानी में मदद करते हैं (कैचिया, 2014)। आपूर्ति श्रृंखला गतिशीलता (भोई एट अल., 2019) में सुधार के कारण 2014 से नरम पड़ने के बाद, सीपीआई खाद्य मुद्रास्फीति 2 में मौसम संबंधी गड़बड़ी, कोविड-19 महामारी और भू-राजनीतिक संघर्षों के कारण आपूर्ति में व्यवधानों के फिर से उभरने के कारण 2019-20 से औसत और अस्थिरता दोनों में वृद्धि देखी गई।

खाद्य और पेय पदार्थ समूह (जिसे आगे 'खाद्य' कहा जाएगा) के सीपीआई बास्केट में इसके उच्च वजन और संबंधित बड़ी अस्थिरता के कारण हेडलाइन मुद्रास्फीति पर महत्वपूर्ण प्रभाव को देखते हुए, खाद्य मुद्रास्फीति के विश्वसनीय पूर्वानुमान बनाने के लिए सिस्टम का होना अत्यंत महत्वपूर्ण है। क्रॉस-कंट्री हेडलाइन मुद्रास्फीति पूर्वानुमान प्रदर्शन की तुलना से पता चलता है कि जिन देशों के सीपीआई बास्केट में खाद्य पदार्थों का बड़ा हिस्सा होता है, वे उच्च पूर्वानुमान त्रुटियों का अनुभव करते हैं (आरबीआई, 2020), भारत में मुद्रास्फीति पूर्वानुमान की चुनौतीपूर्ण प्रकृति को उजागर करता है और सटीक खाद्य मुद्रास्फीति पूर्वानुमानों के महत्व को दोहराता है।

पूर्वानुमान लगाने की प्रक्रिया का एक महत्वपूर्ण घटक नाउकास्टिंग है, जिसका उद्देश्य आधिकारिक डेटा प्रकाशित होने से पहले वर्तमान अवधि (महीना या तिमाही) में मुद्रास्फीति का पूर्वानुमान लगाना है (क्लार्क, एट अल., 2022)। ये नाउकास्ट न केवल अपने आप में दिलचस्प हैं, बल्कि अल्पावधि और मध्यम अवधि में पूर्वानुमानों के लिए महत्वपूर्ण इनपुट के रूप में भी काम करते हैं (क्रुगर एट अल., 2017; फॉस्ट एंड राइट, 2013)। हालाँकि, खाद्य मुद्रास्फीति के नाउकास्ट को बेहतर बनाने पर ध्यान केंद्रित करने वाले अध्ययन अपेक्षाकृत कम हैं (मैकियास, एट अल., 2023), जिसके कारण इस क्षेत्र में अधिक ध्यान देने की आवश्यकता है।

^ निशांत सिंह और अभिरुचि राठी भारतीय रिजर्व बैंक (आरबीआई), मुंबई के आर्थिक और नीति अनुसंधान विभाग (डीईपीआर) में प्रबंधक हैं। लेखक बिनोद बिहारी भोई के बहुमूल्य सुझावों के लिए हम उनके आभारी हैं। लेख में व्यक्त विचार लेखकों के हैं और आरबीआई के विचारों का प्रतिनिधित्व नहीं करते हैं।

[1] एफआईटी ढांचे के तहत मुद्रास्फीति का लक्ष्य 4 प्रतिशत निर्धारित किया गया है, जिसमें शीर्ष मुद्रास्फीति के संदर्भ में +/-2 प्रतिशत की सहनशीलता बैंड है, जिसे आधार वर्ष 2012=100 के साथ अखिल भारतीय उपभोक्ता मूल्य सूचकांक-संयुक्त (सीपीआई-सी) श्रृंखला में वर्ष-दर-वर्ष (वाईओवाई) प्रतिशत परिवर्तन द्वारा मापा जाता है।

[2] खाद्य मुद्रास्फीति को खाद्य एवं पेय पदार्थ समूह की अखिल भारतीय उपभोक्ता मूल्य सूचकांक-संयुक्त (सीपीआई-सी) श्रृंखला (आधार वर्ष 2012=100) में वर्ष-दर-वर्ष प्रतिशत परिवर्तन के आधार पर मापा जाता है।

घरेलू और अंतर्राष्ट्रीय कारकों की बहुलता के प्रति खाद्य कीमतों की संवेदनशीलता और इसके बदलते सांख्यिकीय गुणों के कारण, पारंपरिक एकतरफा मॉडलिंग विधियों का उपयोग करके खाद्य कीमतों में बड़े और अप्रत्याशित उतार-चढ़ाव को पकड़ने का कार्य चुनौतीपूर्ण हो गया है। यह वैकल्पिक, फिर भी प्रासंगिक, जानकारी की खोज करने का मामला बनाता है जो पारंपरिक रूप से उपयोग किए जाने वाले स्रोतों की तुलना में उच्च आवृत्तियों पर उपलब्ध हो सकती है, और मूल्य उतार -चढ़ाव की दिशा और परिमाण के वास्तविक समय के संकेत निकालने के लिए ऐसी जानकारी का लाभ उठाने के गैर-पारंपरिक तरीकों का मूल्यांकन करती है।

हाल ही में तकनीकी प्रगति, डिजिटलीकरण का उदय, और विशेष रूप से, अर्थव्यवस्था के विभिन्न क्षेत्रों से उच्च आवृत्ति वाली जानकारी के उद्भव ने भारतीय संदर्भ में खाद्य मूल्य गतिशीलता की बेहतर समझ को सुगम बनाया है। जबकि कुछ हालिया अध्ययनों ने अल्पकालिक खाद्य मुद्रास्फीति (यादव और दास, 2023; प्रताप एट अल., 2022) की भविष्यवाणी करने के लिए उच्च आवृत्ति वाली मूल्य जानकारी का उपयोग करने वाले मॉडल को नियोजित किया है, वहीं बहुत सी अन्य मूल्य और गैर-मूल्य जानकारी उपलब्ध है, जिसे मूल्य जानकारी के साथ समेकित और उपयोग किए जाने पर, खाद्य कीमतों और सीपीआई खाद्य मुद्रास्फीति को प्रभावित करने वाली आपूर्ति-पक्ष गतिशीलता को बेहतर ढंग से दर्शा सकता है। ऐसी मूल्य और गैर-मूल्य जानकारी के संभावित सेट में उच्च आवृत्ति खुदरा और थोक/मंडी खाद्य कीमतें, घरेलू और अंतर्राष्ट्रीय कमोडिटी कीमतें, मौसम संबंधी जानकारी, जलाशय स्तर, फसल बुवाई, उत्पादन और बाजार/मंडी आवक, मजदूरी दरें, विनिमय दर में उतार-चढ़ाव (आईएनआर/यूएसडी), खाद्य कीमतों के संबंध में जनता के बीच इंटरनेट-खोज रुझान और सरकारी नीतियां और हस्तक्षेप शामिल हैं। इसलिए, खाद्य मुद्रास्फीति का पूर्वानुमान लगाने के लिए नियमित आधार पर ऐसी उभरती हुई सूचनाओं पर नज़र रखने की आवश्यकता होती है।

टर्निंग पॉइंट या उच्च मुद्रास्फीति प्रकरणों के आसपास पूर्वानुमान त्रुटियाँ बड़ी हो सकती हैं क्योंकि ऐसे चरणों के दौरान मुद्रास्फीति की समय श्रृंखला गुणधर्म और प्रमुख मैक्रोइकॉनोमिक चर के साथ इसका संबंध अस्थिर हो सकता है (एंड्रियास जोसेफ एट अल., 2024)। ऐसी परिस्थितियों में, ऑटोरिग्रैसिव इंटीग्रेटेड मूविंग एवरेज (एआरआईएमए), सीजनल एआरआईएमए (एसएआरआईएमए) और लीनियर रिग्रेशन सहित व्यापक रूप से उपयोग किए जाने वाले पारंपरिक रैखिक अर्थमितीय मॉडल, जो रैखिक और समय-अपरिवर्तनीय संबंधों को मानते हैं, संभावित गैर-रैखिकता और बदलते संबंधों को छोड़ सकते हैं (बिनर एट अल., 2005)। इसलिए, न केवल विभिन्न क्षेत्रों में वैकल्पिक जानकारी का लाभ उठाना मूल्य उतार -चढ़ाव और टर्निंग पॉइंट्स का पता लगाने के लिए उपयोगी हो सकता है, बल्कि डेटा में संबंधों को पकड़ने के लिए वैकल्पिक तकनीकों को नियोजित करने से बेहतर नाउकास्टिंग में और मदद मिल सकती है। इसलिए, एमएल मॉडल का लाभ उठाने से नाउकास्टिंग अभ्यास में मूल्य जोड़ा जा सकता है, क्योंकि वे बड़े विषम डेटा से निपटने और बदलते संबंधों को पकड़ने में उपयुक्त हैं (चक्रवर्ती और जोसेफ, 2017)। उच्च-आयामी डेटा पर, नियमितीकरण या संकुचन विधियों को नियोजित करने से मॉडलिंग प्रदर्शन को बढ़ाने में भी मदद मिल सकती है (जोसेफ एट अल., 2024; रिचर्डसन एट अल., 2021)। इसलिए, पारंपरिक रैखिक तकनीकों के अलावा, यह अध्ययन नियमितीकरण (संकुचन) तकनीकों जैसे रिज रिग्रेशन (होर्ल और केनार्ड, 1970), डीप लर्निंग (डीएल) सहित मशीन लर्निंग (एमएल) तकनीकों की भी खोज करता है जो संभावित गैर-रैखिकता को पकड़ते हैं (सिंह और भोई, 2022; चक्रवर्ती और जोसेफ, 2017; लेकुन एट अल., 2015) और सपोर्ट वेक्टर रिग्रेशन (एसवीआर) जो ओवरफिटिंग को कम करता है और उच्च-आयामी डेटा पर अच्छा प्रदर्शन करने के लिए जाना जाता है (नोबल, 2006; ड्रकर एट अल., 1996)।

भारत में खाद्य मुद्रास्फीति के बारे में अब तक के अध्ययन दुर्लभ हैं और खुदरा और थोक मूल्यों के अलावा अन्य जानकारी का लाभ उठाने वाले अध्ययन और भी दुर्लभ हैं। इसलिए, यह अध्ययन खाद्य मुद्रास्फीति के अब तक के पूर्वानुमान की सटीकता को बढ़ाने के लिए एमएल और नियमितीकरण सहित वैकल्पिक अब तक के पूर्वानुमान तकनीकों और विधियों (डेटा विज्ञान) का उपयोग करके बड़ी वैकल्पिक जानकारी (बड़ा डेटा) की भविष्य कहनेवाला शक्ति की जांच करके मौजूदा साहित्य में योगदान करने का प्रयास करता है। अनुभवजन्य अभ्यास के लिए, अध्ययन सभी तकनीकों के सेट को तीन व्यापक श्रेणियों में विभाजित करता है, यानी, यूनीवेरिएट लीनियर, मल्टीवेरिएट लीनियर और

[3] इसे समझने के लिए, अध्ययन ने अनुभवजन्य अभ्यास में इनपुट के रूप में गूगल ट्रेंड्स डेटा से जानकारी पर विचार किया। गूगल ट्रेंड्स से प्राप्त जानकारी उपभोक्ता भावना को दर्शा सकती है और मूल्य उतार -चढ़ाव की भविष्यवाणी करने में उपयोगी है (सीबोल्ड और कोप्पला, 2015)।

मल्टीवेरिएट एमएल-आधारित यह जांचने के लिए कि क्या डेटा कवरेज और मॉडल जटिलता दोनों के परिष्कार का बढ़ता स्तर अब तक के प्रदर्शन को बढ़ाता है। यह जांचने के लिए कि क्या वे व्यक्तिगत मॉडलों में सुधार करते हैं, साहित्य से आकर्षित होकर, व्यक्तिगत मॉडलों की तुलना में सटीकता में संभावित सुधार का सुझाव देते हुए, अलग-अलग अब तक के संयोजन भी तैयार किए गए हैं (जॉन एट अल।, 2020; स्टॉक और वॉटसन, 2004)।

अध्ययन के बाकी हिस्से को पाँच खंडों में व्यवस्थित किया गया है। खंड II भारत में खाद्य मुद्रास्फीति पर शैलीगत तथ्य प्रदान करता है। खंड III प्रासंगिक साहित्य की समीक्षा करता है। खंड IV कार्यप्रणाली और अनुभवजन्य रणनीति का वर्णन करता है, उसके बाद खंड V में परिणाम दिए गए हैं । खंड VI पेपर का समापन करता है।

II. भारत में खाद्य मुद्रास्फीति पर शैलीगत तथ्य

भारत के सीपीआई में खाद्य समूह का पर्याप्त भार है (सारणी 1)। भारत में खाद्य मुद्रास्फीति ने 2014 से 2019 की पहली छमाही तक गिरावट का रुख दर्ज किया। जबकि मांग के सापेक्ष बेहतर आपूर्ति की स्थिति को इस नरमी के पीछे प्राथमिक प्रेरक शक्ति के रूप में रेखांकित किया गया है, आपूर्ति श्रृंखला की गतिशीलता में वृद्धि भी इस प्रवृत्ति में योगदान देने वाला कारक थी (भोई एट अल., 2019)। हालांकि, 2019-21 के दौरान, खाद्य मुद्रास्फीति और इसका योगदान (2017-19 के दौरान लगभग

सारणी 1: भारत में सीपीआई मुद्रास्फीति

अवधि: जनवरी 2014 – दिसंबर 2023

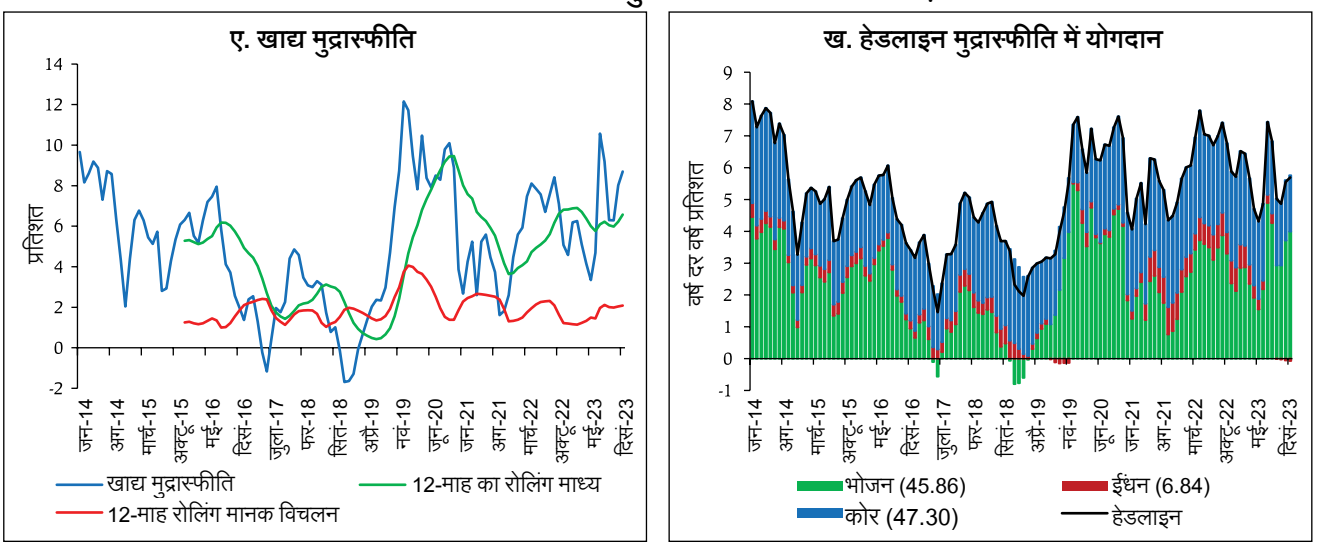
सीपीआई ग्रुप	सीपीआई वजन	अर्थ	मानक विचलन	तिरछापन	कुकुदता
खाना	45.86	5.1	3.0	-0.1	-0.6
ईंधन	6.84	5.4	3.6	0.3	-0.1
मुख्य	47.30	5.2	0.8	0.1	-0.8
शीर्षक	100	5.1	1.5	-0.1	-0.7

नोट: कोर ग्रुप खाद्य और ईंधन समूहों को छोड़कर सीपीआई हेडलाइन को संदर्भित करता है।
स्रोत: राष्ट्रीय सांख्यिकी कार्यालय (एनएसओ), सांख्यिकी और कार्यक्रम कार्यान्वयन मंत्रालय (एमओएसपीआई), भारत सरकार (जीओआई), और आरबीआई स्टाफ अनुमान।

13 प्रतिशत और 2019-21 के दौरान 53 प्रतिशत) हेडलाइन मुद्रास्फीति में तेजी से बढ़ा (चार्ट 1ए और 1बी) बारिश से प्रेरित खाद्य मूल्य दबाव और महामारी से प्रेरित आपूर्ति व्यवधानों के कारण। वैश्विक आपूर्ति श्रृंखला की गड़बड़ी कम होने के साथ, 2021-22 में खाद्य मुद्रास्फीति कम हुई। हालाँकि, यह नरमी अल्पकालिक थी क्योंकि यूरोप में संघर्ष के कारण वैश्विक खाद्य और ऊर्जा की कीमतों में नए सिरे से उछाल आया (आईएमएफ, 2022), जिसने मौसम संबंधी गड़बड़ियों के साथ मिलकर 2022 से खाद्य मुद्रास्फीति को बड़े पैमाने पर ऊंचा रखा।

सीपीआई खाद्य गति 4 में भी कुछ सांख्यिकीय परिवर्तन हुए हैं जैसे समय के साथ उसी सीज़न के अंतराल (12वें अंतराल) के साथ सहसंबंध में गिरावट (चार्ट 2ए), नाउकास्टिंग के लिए

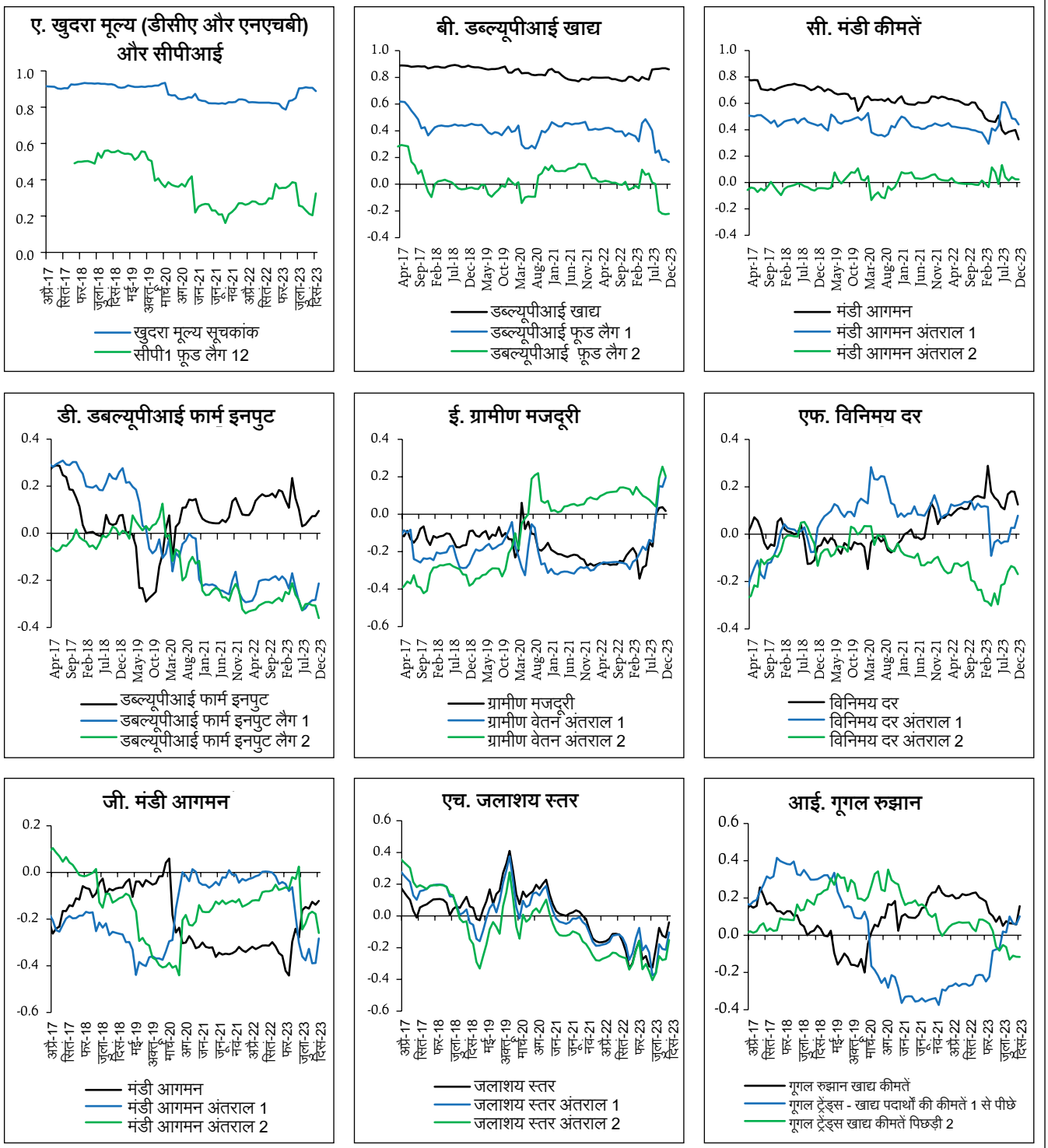
चार्ट 1: भारत में खाद्य मुद्रास्फीति के सारांश आंकड़े



नोट: कोष्ठक में दिए गए आंकड़े सीपीआई में भार दर्शाते हैं। कोर ग्रुप का तात्पर्य खाद्य और ईंधन समूहों को छोड़कर सीपीआई हेडलाइन से है।
स्रोत: एनएसओ, एमओएसपीआई; तथा आरबीआई स्टाफ अनुमान।

[4] मासिक गति से तात्पर्य माह-दर-माह (मासिक) परिवर्तन (प्रतिशत में) से है।

चार्ट 2: कुछ वैकल्पिक चरों के साथ सीपीआई खाद्य का रोलिंग सहसंबंध



नोट: 1. चार्ट में चर, जलाशय स्तर (दीर्घावधि औसत (एलपीए) से प्रतिशत विचलन) को छोड़कर गति के रूप में हैं।
2. वाई -अक्ष सहसंबंध गुणांक को दर्शाता है। रोलिंग सहसंबंध 36 महीने की अवधि पर आधारित होते हैं।
3. खुदरा मूल्य सूचकांक (आरपीआई) तैयार करने के लिए, उपभोक्ता मामले विभाग (डीसीए) से 22 आवश्यक वस्तुओं और राष्ट्रीय बागवानी बोर्ड (एनएचबी) से 11 वस्तुओं के खुदरा मूल्यों को 2012 (सीपीआई आधार वर्ष) में औसत मूल्यों को आधार मानकर सूचकांक में परिवर्तित किया गया और फिर संबंधित सीपीआई भार का उपयोग करके समेकित किया गया। अन्य चरों के बारे में जानकारी के लिए, कृपया सारणी ए2 देखें।
4. विनिमय दर को आईएनआर प्रति यूएसडी के रूप में व्यक्त किया जाता है। गूगल ट्रेंड्स इंडेक्स भारतीय क्षेत्र के लिए "खाद्य मूल्य" शब्द से मेल खाता है।

स्रोत: आरबीआई स्टाफ का अनुमान।

यूनीवेरिएट मॉडलिंग विधियों को कमजोर करना, जो खाद्य मुद्रास्फीति की गतिशीलता को बेहतर ढंग से पकड़ने के लिए अतिरिक्त उच्च आवृत्ति जानकारी का पता लगाने के लिए कहता है। खुदरा और थोक मूल्यों के अलावा अन्य जानकारी संभावित रूप से आपूर्ति पक्ष की गतिशीलता के बारे में अतिरिक्त वास्तविक समय की जानकारी प्रदान कर सकती है, जो खुदरा मूल्य परिवर्तनों में तब्दील हो सकती है, जो शुरुआती/अग्रणी संकेतक के रूप में कार्य करती है, विशेष रूप से, उन खाद्य वस्तुओं के लिए जिनके लिए मूल्य-संबंधी जानकारी उपलब्ध नहीं है। खाद्य मुद्रास्फीति और कुछ संबंधित चर के बीच संबंधों में भी समय के साथ बदलाव आया है, जैसा कि चर और उनके अंतराल (चार्ट 2) के साथ सहसंबंध में परिलक्षित होता है

III. साहित्य समीक्षा

नाउकास्टिंग अभ्यास विशेष रूप से मैक्रोइकॉनोमिक चर के लिए प्रासंगिक हैं जो कम आवृत्तियों पर उपलब्ध होते हैं, अक्सर तिमाही आधार पर, जैसा कि सकल घरेलू उत्पाद (जीडीपी) के मामले में होता है। इन चरों के लिए डेटा आम तौर पर महत्वपूर्ण अंतराल के साथ जारी किया जाता है, जिससे नाउकास्ट की आवश्यकता होती है - अधिक आसानी से उपलब्ध और अक्सर अपडेट किए गए डेटा स्रोतों का लाभ उठाकर प्राप्त किए गए शुरुआती अनुमान (बनबुरा एट अल., 2010)। मुद्रास्फीति के डेटा आमतौर पर मासिक आधार पर उपलब्ध होने के कारण, नाउकास्टिंग, विशेष रूप से खाद्य मुद्रास्फीति के लिए, वैश्विक स्तर पर अपेक्षाकृत कम खोज की गई है, खासकर प्रमुख उन्नत अर्थव्यवस्थाओं में सीपीआई बास्केट में इसके अपेक्षाकृत कम वजन को देखते हुए। सीपीआई बास्केट में खाद्य की बड़ी हिस्सेदारी वाली उभरती और विकासशील अर्थव्यवस्थाओं में, खाद्य मुद्रास्फीति की नाउकास्टिंग मौद्रिक नीति के संचालन और निर्माण को काफी हद तक मजबूत कर सकती है।

कई अध्ययनों से पता चलता है कि मुद्रास्फीति की भविष्यवाणी करने के लिए उच्च आवृत्ति की जानकारी का महत्व है। सिल्वा एट अल. (2024) के अनुसार, खाद्य मूल्य मुद्रास्फीति की भविष्यवाणी करने में दैनिक खाद्य कीमतों का महत्व है। बेक एट अल. (2022) जर्मन मुद्रास्फीति की भविष्यवाणी करने के लिए उत्पाद-स्तर की कीमतों और मात्राओं पर घरेलू स्कैनर डेटा का लाभ उठाते हैं, यह प्रदर्शित करते हुए कि यह बारीक जानकारी प्रत्येक महीने की शुरुआत में मुद्रास्फीति की गतिशीलता में समय पर अंतर्दृष्टि प्रदान करती है। यादव और दास (2023) के अनुसार, दैनिक भीड़-स्रोत वाले खाद्य मूल्यों पर गतिशील कारकों और मिश्रित आवृत्ति मॉडल का उपयोग करने वाला दृष्टिकोण मुद्रास्फीति की भविष्यवाणी करने में पारंपरिक दृष्टिकोणों से बेहतर प्रदर्शन करता है। मैकियास एट अल. (2023) ने पाया कि प्रमुख ऑनलाइन खुदरा विक्रेताओं के वेबपेजों से स्क्रैप किए गए खाद्य और गैर-अल्कोहल पेय पदार्थों की कीमतों के व्यापक डेटासेट का उपयोग करने से खाद्य मुद्रास्फीति की भविष्यवाणी की सटीकता बढ़ जाती है। गूगल ट्रेंड्स डेटाबेस का लाभ उठाते हुए, सीबोल्ड और कोपोला (2015) ने पाया कि इंटरनेट सर्च इंडेक्स को एकीकृत करने से मध्य अमेरिका में कीमतों की भविष्यवाणी में सुधार होता है। मोडुग्नो (2013) ने दिखाया कि ऊर्जा और कच्चे माल की कीमतों पर उच्च आवृत्ति डेटा को शामिल करने से मुद्रास्फीति के नाउकास्ट का प्रदर्शन बेहतर होता है। नॉटेक और ज़मान (2017) ने भी अमेरिका में वास्तविक समय के डेटा का उपयोग करके मुद्रास्फीति के नाउकास्टिंग का पता लगाया।

हाल के साहित्य ने असंरचित डेटा को संसाधित करने और मजबूत गैर-रैखिकता (देसाई, 2023; गौलेट कूलोम्बे एट अल., 2022) को पकड़ने की उनकी क्षमता के लिए एमएल के भीतर वैकल्पिक नाउकास्टिंग तकनीकों के उपयोग पर भी प्रकाश डाला है, साथ ही संकुचन और नियमितीकरण तकनीकें जो ओवरफिटिंग को कम करके और सामान्यीकरण में सुधार करके उच्च-आयामी डेटा पर अच्छा प्रदर्शन करती हैं। जोसेफ एट अल. (2024) यूके मुद्रास्फीति की भविष्यवाणी करने में प्रमुख घटक विश्लेषण (पीसीए), रिज रिग्रेशन जैसी संकुचन विधियों और सपोर्ट वेक्टर मशीन (एसवीएम) और न्यूरल नेटवर्क जैसे एमएल मॉडल जैसी आयाम कम करने वाली तकनीकों की प्रभावशीलता का पता लगाते हैं और पाते हैं कि रिज रिग्रेशन और अन्य संकुचन विधियां उच्च-आयामी डेटा का उपयोग करते समय सबसे अच्छा प्रदर्शन करती हैं, और प्रभावी दंड के साथ बड़ी और प्रासंगिक सूचना सेट को संयोजित करने से मॉडल का प्रदर्शन बेहतर होता है। यूरो क्षेत्र मुद्रास्फीति डेटा का उपयोग करते हुए, अलियाज एट अल. (2023) पाते हैं कि लैसो रिग्रेशन, एक अन्य लोकप्रिय नियमितीकरण तकनीक, यूरो क्षेत्र मुद्रास्फीति की नाउकास्टिंग में मानक विधियों से बेहतर प्रदर्शन करती है।

IV. कार्यप्रणाली, डेटा और अनुभवजन्य रणनीति

भारत में खाद्य मुद्रास्फीति (वर्ष-दर-वर्ष) प्रकृति में गैर-स्थिर है। हालांकि, सीपीआई खाद्य गति स्थिर पाई गई है (सारणी ए1), और इसलिए इसे इस अध्ययन में अनुभवजन्य कार्य के लिए लक्ष्य चर के रूप में उपयोग किया गया है (सारणी 2)। इस पत्र में सभी अनुभवजन्य कार्य मासिक डेटा पर आधारित हैं। नमूना अवधि पूर्वाग्रह को नियंत्रित करने के लिए विस्तारित नमूना आधार पर नाउकास्ट तैयार किए गए हैं। प्रत्येक तकनीक के लिए मॉडल विनिर्देश दिसंबर 2022 तक के डेटा का उपयोग करके तय किया गया है, जिसके बाद चयनित मॉडल 12 मासिक नाउकास्ट उत्पन्न करने के लिए एक समय में एक क्रमिक महीने के प्राप्त डेटा को जोड़कर विस्तारित नमूना आधार पर प्रशिक्षित किए जाते हैं। एमएल मॉडल को आम तौर पर प्रशिक्षित मॉडल की सटीकता का आकलन करने और परीक्षण डेटा सेट पर प्राप्त न्यूनतम त्रुटि 6 के आधार पर इष्टतम मॉडल चुनने के लिए परीक्षण डेटा सेट 5 की आवश्यकता होती है।

इस शोधपत्र में, वैकल्पिक नाउकास्टिंग तकनीकों को उनके नाउकास्ट प्रदर्शन की व्यापक तुलना के लिए नियोजित किया गया है। सभी तकनीकों के सेट को तीन व्यापक श्रेणियों में विभाजित किया गया है, अर्थात, यूनीवेरिएट लीनियर, मल्टीवेरिएट लीनियर और मल्टीवेरिएट एमएल-आधारित (सारणी 3) यह आकलन करने के लिए कि क्या डेटा कवरेज और मॉडल जटिलता के संदर्भ में परिष्कार का बढ़ता स्तर नाउकास्ट प्रदर्शन में सुधार करता है।

जबकि डीप लर्निंग (डीएल) को डेटा में संभावित गैर-रैखिकता को पकड़ने के लिए नियोजित किया गया है, रिज रिग्रेशन 7 और एसवीआर 8 का उपयोग यह जांचने के लिए किया गया है कि क्या उच्च-आयामी डेटा के लिए उनकी उपयुक्तता नाउकास्ट सटीकता को बढ़ाती है।

अध्ययन में खाद्य मुद्रास्फीति में भिन्नता को पकड़ने के लिए मूल्य और गैर-मूल्य दोनों संकेतकों पर विचार किया गया है (सारणी ए 2)। खुदरा मूल्य, थोक/मंडी मूल्य, घरेलू और अंतर्राष्ट्रीय कमोडिटी मूल्य, थोक मूल्य सूचकांक (डबल्यूपीआई) आधारित इनपुट मूल्य, ग्रामीण मजदूरी, वर्षा विचलन, जलाशय स्तर, बाजार आगमन, विनिमय दर और गूगल रुझान डेटा पर विचार किया गया है। अनुभवजन्य कार्य सभी व्याख्यात्मक चरों को समसामयिक और विलंबित दोनों रूपों में मानता है ताकि उनके डेटा रिलीज में देरी को ध्यान में रखते हुए डबल्यूपीआई और ग्रामीण मजदूरी को छोड़कर, नाउकास्ट पीढ़ी के लिए उनके तत्काल और विलंबित प्रभाव को पकड़ा जा सके, जैसा कि परिशिष्ट की सारणी ए 3 में विस्तृत है ।

चरों के सेट और उनकी उपयुक्त अंतराल संरचना (एआईसी पर आधारित) को अंतिम रूप देने के बाद, अधिकतम प्रमुख घटकों

सारणी 2: अध्ययन में प्रयुक्त कार्यप्रणाली का विवरण

वस्तु	पारंपरिक तकनीक	एमएल तकनीक
अध्ययन अवधि	जनवरी 2014 – दिसंबर 2023	जनवरी 2014 – दिसंबर 2023
मॉडल पहचान अवधि	जनवरी 2014 – दिसंबर 2022	जनवरी 2014 – दिसंबर 2022
लक्ष्य चर	सीपीआई खाद्य गति	सीपीआई खाद्य गति
परीक्षण डेटा का आकार	-	12 महीने
मॉडल निर्माण अवधि	प्रशिक्षण अवधि	प्रशिक्षण + परीक्षण अवधि
प्रथम नमूना प्रशिक्षण अवधि	जनवरी 2014 – दिसंबर 2022	जनवरी 2014 – दिसंबर 2021
प्रथम नमूना परीक्षण अवधि	-	जनवरी 2022 – दिसंबर 2022
नाउकास्ट अवधि	जनवरी 2023 – दिसंबर 2023	जनवरी 2023 – दिसंबर 2023

5 परीक्षण डेटा एक डेटासेट है, जो प्रशिक्षण डेटासेट से अलग है, जिस पर प्रशिक्षित मॉडलों का परीक्षण करके उनकी भविष्यवाणी की सटीकता का आकलन किया जाता है।

6 अध्ययन में नाउकास्ट प्रदर्शन तुलना के लिए त्रुटि के माप के रूप में रूट मीन स्क्वेयर्ड एरर (आरएमएसई) और मीन एब्सोल्यूट एरर (एमएई) का उपयोग किया गया है।

7 रिज रिग्रेशन - एक लोकप्रिय नियमितीकरण तकनीक, बड़ी संख्या में चरों की उपस्थिति में मॉडल की जटिलता को कम करती है, नियमितीकरण (L2) के रूप में उनके आकार पर जुर्माना लगाकर प्रत्येक के गुणांक को छोटा करके, जहाँ दंड गुणांक के वर्ग परिमाण पर लागू होता है (रिचर्डसन एट अल., 2021)। यह दृष्टिकोण न केवल आउटलेयर के प्रति इसकी संवेदनशीलता को कम करके मॉडल को स्थिर करने में मदद करता है, बल्कि इसके सामान्यीकरण प्रदर्शन को भी बेहतर बनाता है, जिससे यह उच्च-आयामी डेटा और मल्टीकोलिनियरिटी से निपटने के दौरान उपयोगी हो जाता है।

8 सपोर्ट वेक्टर रिग्रेशन (एसवीआर) - एक लोकप्रिय एमएल तकनीक, रैखिक रिग्रेशन की तुलना में आउटलेयर के प्रति कम संवेदनशील है, क्योंकि यह एक एप्सिलॉन-असंवेदनशील क्षेत्र, रिग्रेशन लाइन के चारों ओर एक ट्यूब प्रस्तुत करता है जहां अवशेषों की उपेक्षा की जाती है, जो इसे छोटे विचलनों को अनदेखा करके अधिक मजबूत रेखा को फिट करने की अनुमति देता है, ओवरफिटिंग को कम करता है क्योंकि इसका नुकसान फ़ंक्शन इस असंवेदनशील क्षेत्र के बाहर केवल अवशेषों को दंडित करता है (ड्रकर एट अल., 1996)।

सारणी 3: अध्ययन में प्रयुक्त वैकल्पिक तकनीकें

तकनीक	प्रकार
यादृच्छिक चाल	एकतरफा रैखिक
अरिमा और सारिमा	एकतरफा रैखिक
रेखीय प्रतिगमन	बहुभिन्नरूपी रैखिक
एरिमैक्स और सारीमैक्स	बहुभिन्नरूपी रैखिक
रिज रिग्रेशन	बहुभिन्नरूपी रैखिक
कृत्रिम तंत्रिका नेटवर्क (एएनएन)	बहुभिन्नरूपी एमएल-आधारित
सपोर्ट वेक्टर रिग्रेशन (एसवीआर)	बहुभिन्नरूपी एमएल-आधारित

नोट: रिज रिग्रेशन और एसवीआर के अलावा अन्य मॉडलों के विवरण के लिए कृपया सिंह और भोई, 2022 देखें।

(पीसी) को उत्पन्न करने हेतु आयामी कमी हेतु खुदरा कीमतों [9] को छोड़कर सभी व्याख्यात्मक चरों पर प्रमुख घटक विश्लेषण (पीसीए) तकनीक का उपयोग किया गया है (सारणी ए3)। इसके बाद, थ्रेशोल्ड के रूप में 10 प्रतिशत के महत्व स्तर का उपयोग करके प्रशिक्षण डेटा के आधार पर सबसे अधिक प्रासंगिक पीसी को शॉर्टलिस्ट करने के लिए फॉरवर्ड सिलेक्शन [10] (एफएस) तकनीक का उपयोग किया गया है। पीसीए आयोजित करने से पहले और बाद में इनपुट चरों की सूची अनुलग्नक (सारणी ए4) में दी गई है। जबकि रिज प्रतिगमन को इसके नियमितीकरण (संकुचन) क्षमता के कारण सभी 35 पूर्वानुमानों के कच्चे डेटा का लाभ उठाने की अनुमति है, अन्य मॉडल शॉर्टलिस्ट किए गए प्रमुख घटकों को व्याख्यात्मक चर के रूप में मानते हैं (सारणी ए5)।

डेटा पर मौसमी समायोजन नहीं किया गया है क्योंकि (1) पेपर केवल नाउकास्ट सटीकता पर ध्यान केंद्रित करता है और प्रभाव मूल्यांकन पर नहीं और (2) मौसमी समायोजन के परिणामस्वरूप कुछ जानकारी का नुकसान होता है, भले ही ठीक से आयोजित किया गया हो (आईएमएफ, 2017)।

V. परिणाम

चूंकि मुद्रास्फीति के पूर्वानुमान महीने-दर-महीने (एमओएम) गति के रूप में हैं, इसलिए उन्हें वास्तविक सीपीआई खाद्य मुद्रास्फीति दरों (वाईओवाई) के साथ समान तुलना के लिए वर्ष-दर-वर्ष (वाईओवाई) संख्याओं में परिवर्तित किया जाता है। मॉडलों में प्रदर्शन तुलना के लिए नाउकास्ट के मूल माध्य वर्ग त्रुटि (आरएमएसई) और माध्य निरपेक्ष त्रुटि (एमएई) दोनों की गणना की गई है। रैंडम वॉक को आधार मॉडल मानकर गणना की गई सापेक्ष आरएमएसई की भी तुलना की गई है (चार्ट 3)। चूंकि 2023 ति. 3 में खाद्य कीमतों में भारी उतार-चढ़ाव देखा गया, इसलिए अधिक मजबूत तुलना के लिए ति. 3 को छोड़कर नाउकास्ट की भी गणना की गई है।

परिणाम यह संकेत देते हैं कि सूचना सेट का विस्तार करने के साथ-साथ मॉडल जटिलता बढ़ाने से 2023 के डेटा के आधार पर नाउकास्ट की सटीकता में सुधार होता है (सारणी 4 , चार्ट 3)। सटीकता तब उल्लेखनीय रूप से बढ़ जाती है जब खुदरा और थोक खाद्य मूल्य की जानकारी को उन परिदृश्यों की तुलना में शामिल किया जाता है जहां केवल यूनीवेरिएट मॉडल (कोई वैकल्पिक जानकारी नहीं) नियोजित होते हैं, जो ऑटोरिग्रैसिव मॉडल (चार्ट ए 1) में महत्वपूर्ण तत्काल अनुक्रमिक लैग की उपस्थिति के कारण समयबद्ध रूप से पिछड़े (प्रतीत होता है कि दाईं ओर शिफ्ट) नाउकास्ट का उत्पादन करते हैं। हालांकि, बहुभिन्नरूपी रैखिक मॉडल (रैखिक प्रतिगमन और एसएआरआईएमएएक्स) के भीतर, खुदरा और थोक मूल्यों के अलावा अन्य सूचनाओं को शामिल करने से प्रदर्शन लाभ नहीं मिलता है क्योंकि पारंपरिक रैखिक मॉडल उच्च आयामों पर अच्छा प्रदर्शन नहीं कर सकते हैं और डेटा में संभावित गैर-रैखिकता को पकड़ नहीं सकते हैं।

एमएल मॉडल के भीतर, जबकि एसवीआर उच्च-आयामी डेटा (प्रेडिक्टर वेरिएबल्स के बीच नोइस और विविधता द्वारा विशेषता) के लिए मजबूत अनुमान उत्पन्न करके सटीकता को बढ़ाता है, एएनएन गैर-रैखिकता को पकड़ने की अपनी क्षमता के कारण नाउकास्ट प्रदर्शन को और बढ़ाता है। हालांकि, रिज रिग्रेशन, एक रैखिक मॉडल, अधिकतम प्रदर्शन लाभ प्रदान करता है क्योंकि इसका नियमितीकरण (संकोचन) सुविधा उच्च-आयामी डेटा के साथ आम तौर पर जुड़े ओवरफिटिंग और मल्टीकोलिनियरिटी को कम करके गुणांक को स्थिर करती है। ये परिणाम संकेत देते हैं कि अतिरिक्त जानकारी को शामिल करने

[9] समकालीन रूप में सीपीआई खाद्य के साथ खुदरा मूल्य की जानकारी के मजबूत संबंध को देखते हुए, इसका स्वतंत्र रूप से उपयोग किया गया है और इसे पीसीए के अंतर्गत शामिल नहीं किया गया है।

[10] फॉरवर्ड सिलेक्शन (एफ़एस) एक वैरिएबल सिलेक्शन तकनीक है जो मॉडल में किसी भी वैरिएबल के बिना शुरू होती है और एक सीमा तक एक-एक करके वैरिएबल जोड़ती है (मान लीजिए, 10 प्रतिशत का महत्व स्तर) जिसके बाद कोई और सुधार नहीं होता है। चयन का क्रम सांख्यिकीय महत्व और स्पष्टीकरण शक्ति के क्रमिक जोड़ पर आधारित है।

सारणी 4: वैकल्पिक मॉडलों का प्रदर्शन तुलना

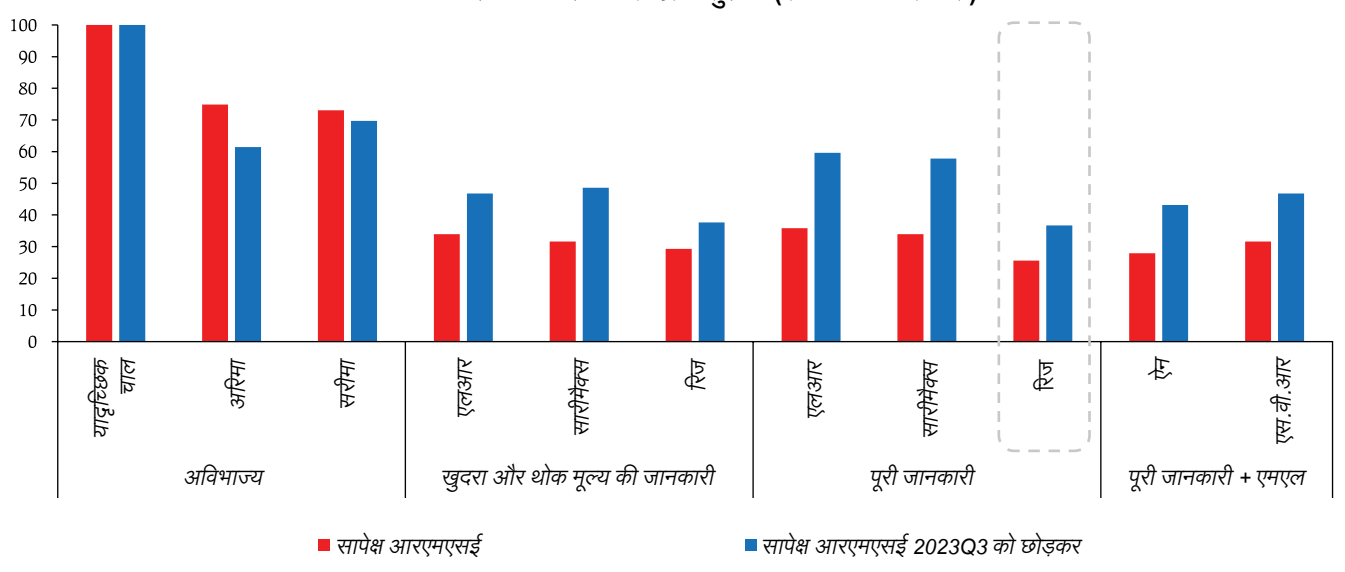
वास्तविक खाद्य मुद्रास्फीति की तुलना में (वर्ष-दर-वर्ष) *नाउकास्ट अवधि: जनवरी 2023 – दिसंबर 2023*

मॉडल प्रकार	नमूना	आरएमएसई		मॅई		जटिलता
		पूर्ण नमूना	2023Q3 को छोड़कर	पूर्ण नमूना	2023Q3 को छोड़कर	
एकतरफा रैखिक	यादृच्छिक चाल	2.15	1.09	1.54	0.93	कोई वैकल्पिक जानकारी नहीं
	अरिमा	1.61	0.67	1.07	0.52	
	सारिमा	1.57	0.76	1.07	0.64	
बहुभिन्नरूपी रैखिक	एलआर	0.73	0.51	0.54	0.39	खुदरा और थोक मूल्य की जानकारी
	सारीमैक्स	0.68	0.53	0.55	0.43	
	रिज	0.63	0.41	0.49	0.35	
	एलआर	0.77	0.65	0.62	0.54	पूरी जानकारी
	सारीमैक्स	0.73	0.63	0.60	0.51	
	रिज	0.55	0.40	0.45	0.33	
बहुभिन्नरूपी एमएल-आधारित	ऐन	0.60	0.47	0.51	0.38	पूरी जानकारी + एमएल
	एसवीआर	0.68	0.51	0.52	0.41	
युग्म	रिज(पीआई)+एएनएन	0.59^	0.41	0.47^	0.33^	
	रिज(एफआई)+एएनएन	0.55	0.40	0.44^	0.31^	

टिप्पणियाँ: 1. एमएई : माध्य निरपेक्ष त्रुटि; एलआर : रैखिक प्रतिगमन; पीआई : मूल्य सूचना (खुदरा और थोक); एफ़आई : पूर्ण सूचना।
2. सभी मॉडलों में व्याख्यात्मक चर के रूप में लक्ष्य चर का मौसमी अंतराल (12वाँ अंतराल) शामिल है।
3. संयोजन अब पूर्वानुमान के आरएमएसई -भारित औसत पर आधारित हैं। ^ यदि आरएमएसई (संयोजन) < न्यूनतम (व्यक्तिगत मॉडलों के आरएमएसई) है तो अब पूर्वानुमान संयोजन के साथ सुधार दर्शाता है।
स्रोत: आरबीआई स्टाफ अनुमान।

से खाद्य मुद्रास्फीति में भिन्नताओं को समझाने में मदद मिलती है, जो पारंपरिक रैखिक बेंचमार्क पर नाउकास्टिंग प्रदर्शन को बढ़ाता है। सर्वश्रेष्ठ प्रदर्शन करने वाले मॉडल का उपयोग करके, आरएमएसई -भारित संयोजन नाउकास्ट की भी गणना की जाती है जो व्यक्तिगत मॉडल से प्राप्त किए गए लोगों की तुलना में अतिरिक्त प्रदर्शन लाभ प्रदान करते हैं, रिज और एएनएन

चार्ट 3: वैकल्पिक मॉडलों का प्रदर्शन तुलना (रैंडम वॉक के सापेक्ष)



स्रोत: आरबीआई स्टाफ का अनुमान।

मॉडल के संयोजन का उपयोग करके अधिकतम प्रदर्शन का सुझाव देते हैं, उपयोग किए गए डेटा में गैर-रैखिकता को पकड़ने के साथ नियमितीकरण की प्रभावशीलता को रेखांकित करते हैं। रिज (एफआई) और एएनएन के संयोजन का आरएमएसई - 0.55 पर बेसलाइन रैंडम वॉक मॉडल का लगभग एक-चौथाई और पारंपरिक यूनीवेरिएट समय-श्रृंखला मॉडल का लगभग एक-तिहाई है, जो इस पेपर में अपनाए गए संवर्धित मॉडलिंग दृष्टिकोण से बड़े पैमाने पर नाउकास्टिंग लाभ का सुझाव देता है।

VI. निष्कर्ष

यह अध्ययन भारत में खाद्य मुद्रास्फीति की अब-प्रत्याशा में सुधार करने के लिए वैकल्पिक सूचना और मॉडलिंग तकनीकों का लाभ उठाने की प्रभावशीलता की जांच करता है। अनुभवजन्य निष्कर्ष दर्शाते हैं कि इनपुट सूचना सेट का विस्तार करना और पारंपरिक एकतरफा मॉडलिंग से आगे बढ़कर अतिरिक्त उच्च आवृत्ति मूल्य और गैर-मूल्य संकेतक शामिल करना भारत में खाद्य कीमतों को प्रभावित करने वाली जटिल आपूर्ति-पक्ष गतिशीलता को कैप्चर करके अब-प्रत्याशा सटीकता में काफी सुधार करता है। इसके अलावा, अध्ययन पारंपरिक रैखिक मॉडलों पर नियमितीकरण (संकोचन) और एमएल-आधारित तकनीकों सहित वैकल्पिक मॉडलिंग दृष्टिकोणों को नियोजित करने के लाभों पर प्रकाश डालता है, जो उच्च-आयामी डेटा को संसाधित करने, ओवरफिटिंग को कम करने, गैर-रैखिकता के लिए लेखांकन और इसलिए, पूर्वानुमान प्रदर्शन को बढ़ाने में उत्कृष्टता के लिए जाने जाते हैं। परिणाम यह भी संकेत देते हैं कि विविध मॉडलों को मिलाने से अब-प्रत्याशा की सटीकता और बढ़ जाती है, जो पूर्वानुमान मॉडलिंग अभ्यासों में एक समूह दृष्टिकोण को अपनाने की वकालत करता है।

इसलिए, इस अध्ययन से प्राप्त अंतर्दृष्टि भारत में आवर्ती और बड़े पैमाने पर अस्थिर खाद्य मूल्य गतिशीलता के माहौल में सूचित नीति निर्माण के लिए मूल्यवान समर्थन प्रदान कर सकती है। भविष्य के शोध में अतिरिक्त असंरचित डेटा स्रोतों के एकीकरण, अधिक उन्नत डीप लर्निंग आर्किटेक्चर को नियोजित करने और मुद्रास्फीति के अन्य घटकों के लिए इस ढांचे को विस्तारित करने की व्यवहार्यता की जांच की जा सकती है।

References

Aliaj, T., Ciganovic, M., and Tancioni, M. (2023). Nowcasting Inflation with Lasso-Regularized Vector Autoregressions and Mixed Frequency Data. *Journal of Forecasting*, 42(3), 464-480.

Banbura, M., Giannone, D., and Reichlin, L. (2010). Nowcasting. Working Paper Series, No. 1275, European Central Bank.

Beck, G. W., Carstensen, K., Menz, J. O., Schnorrenberger, R., and Wieland, E. (2023). Nowcasting Consumer Price Inflation Using High-Frequency Scanner Data: Evidence from Germany. Deutsche Bundesbank Discussion Paper No. 34/2023.

Bhoi, B. B., Kundu, S., Kishore, V., and Suganthi, D. (2019). Supply Chain Dynamics and Food Inflation in India. *RBI Bulletin*, Vol. 73(10), 95-111.

Binner, J. M., Bissoondeeal, R. K, Elger, T., Gazely, A. M., and Mullineux, A. W. (2005). A Comparison of Linear Forecasting Models and Neural Networks: An Application to Euro Inflation and Euro Divisia. *Applied Economics*, Vol. 37(6), 665-680.

Cachia, F. (2014). Nowcasting Regional Consumer Food Inflation. Food and Agriculture Organisation of the United States Working Paper Series, ESS/14-07.

Chakraborty, C., and Joseph, A. (2017). Machine Learning at Central Banks. Bank of England Staff Working Papers, No. 674.

Clark, T. E., Leonard, S., Marcellino, M., and Wegmüller, P. (2022). Weekly Nowcasting US Inflation with Enhanced Random Forests. Federal Reserve Bank of Cleveland.

Desai, A. (2023). Machine Learning for Economics Research: When What and How? Staff Analytical Note 2023-16, Bank of Canada.

Drucker, H., Burges, C. J., Kaufman, L., Smola, A., and Vapnik, V. (1996). Support Vector Regression Machines. *Advances in Neural Information Processing Systems*, 9.

Faust, J., and Wright, J. H. (2013). Forecasting Inflation. *Handbook of Economic Forecasting*, Vol. 2 (Elsevier), 2–56.

Goulet Coulombe, P., Leroux, M., Stevanovic, D., and Surprenant, S. (2022). How is Machine Learning Useful for Macroeconomic Forecasting? *Journal of Applied Econometrics*, 37(5), 920-964.

Hoerl, A. E., and Kennard, R. W. (1970). Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics*, 12(1), 55-67.

IMF. (2022). Countering The Cost-Of-Living Crisis. World Economic Outlook, October 2022, International Monetary Fund (IMF).

John, J., Singh, S., and Kapur, M. (2020). Inflation Forecast Combinations – The Indian Experience. Reserve Bank of India Working Paper Series WPS (DEPR): 11/2020.

Joseph, A., Potjagailo, G., Chakraborty, C., and Kapetanios, G. (2024). Forecasting UK Inflation Bottom Up. *International Journal of Forecasting*.

Knotek, E. S., and Zaman, S. (2017). Nowcasting US Headline and Core Inflation. *Journal of Money, Credit and Banking*, 49(5), 931-968.

Krüger, F., Clark, T. E., and Ravazzolo, F. (2017). Using Entropic Tilting to Combine BVAR Forecasts with External Nowcasts. *Journal of Business and Economic Statistics*, 35(3), 470–485.

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep Learning. *Nature*, 521(7553), 436-444.

Macias, P., Stelmasiak, D., and Szafranek, K. (2023). Nowcasting Food Inflation with a Massive Amount of Online Prices. *International Journal of Forecasting*, Vol. 39, Issue 2, April–June 2023, 809-826.

Modugno, M. (2013). Now-casting Inflation Using High Frequency Data, *International Journal of Forecasting*, Vol. 29, Issue 4, 2013, 664-675.

Noble, W. S. (2006). What is a Support Vector Machine? *Nature Biotechnology*, 24(12), 1565-1567.

Pratap, B., Ranjan, A., Kishore, V., and Bhoi, B. B. (2022). Forecasting Food Inflation Using News-based Sentiment Indicators. *Reserve Bank of India Occasional Papers*, 42(2).

RBI. (2020). Monetary Policy Report, April.

Richardson, A., Van Florenstein Mulder, T., and Vehbi, T. (2021). Nowcasting GDP Using Machine-Learning Algorithms: A Real-Time Assessment. *International Journal of Forecasting*, 37(2), 941-948.

Seabold, S. and Coppola, A. (2015). Nowcasting Prices Using Google Trends: An Application to Central America, Policy Research Working Paper Series 7398, The World Bank.

Sekhar, C. S. C., Roy, D., and Bhatt, Y. (2018). Food Inflation and Volatility in India: Trends and Determinants. *Indian Economic Review* (2018) 53:65–91.

Silva e Silva, L., Mongeau Ospina, C. A., and Fabi, C. (2024). Food Price Inflation Nowcasting and Monitoring. *Statistical Journal of the IAOS*, Vol. 40, No. 2, 325-339.

Singh, N., and Bhoi, B. B. (2022). Inflation Forecasting in India: Are Machine Learning Techniques Useful? *Reserve Bank of India Occasional Papers*, Vol. 43, No. 2: 2022, 46-88.

Stock, J. H., and Watson, M. W. (2004). Combination Forecasts of Output Growth in a Seven-Country Data Set. *Journal of Forecasting*, *23*(6), 405-430.

अनुबंध

सारणी ए1: स्थिरता परीक्षण

चर	एडीएफ परीक्षण सांख्यिकी	पी-मान	परिणाम
खाद्य मुद्रास्फीति (वर्ष दर वर्ष)	-2.19	0.21	गैर स्थिर
सीपीआई खाद्य गति (एम-ओ-एम)	-6.20***	0.00	अचल

नोट: 1. अनुमान जनवरी 2014-दिसंबर 2022 के आंकड़ों पर आधारित हैं।
2. *: पी < 0.10; **: पी < 0.05; ***: पी < 0.01

स्रोत: आरबीआई स्टाफ अनुमान।

सारणी ए 2: इनपुट जानकारी का विवरण

जानकारी	सूत्रों का कहना है
खुदरा मूल्य	डीसीए, एमसीई, भारत सरकार; एनएचबी, एमएएफडब्ल्यू, भारत सरकार
थोक और मंडी कीमतें	डबल्यूपीआई; एगमार्कनेट, एमएएफ़डबल्यू , भारत सरकार
अंतर्राष्ट्रीय मूल्य	अंतर्राष्ट्रीय मुद्रा कोष
डबल्यूपीआई इनपुट कीमतें	डबल्यूपीआई, एमसीआई, भारत सरकार
ग्रामीण मजदूरी	श्रम ब्यूरो, एमएलई, भारत सरकार
वर्षा	आईएमडी, भारत सरकार
जलाशय स्तर	सीएमआईई
बाजार/मंडी आगमन	एगमार्कनेट, एमएएफडब्लू, भारत सरकार
तेल/ऊर्जा की कीमतें	पेट्रोलियम एवं प्राकृतिक गैस मंत्रालय, भारत सरकार; आईएमएफ
विनिमय दर	एफबीआईएल
गूगल ट्रेंड्स	गूगल

नोट: जीओआई: भारत सरकार; डीसीए: उपभोक्ता मामले विभाग; एमसीई: उपभोक्ता मामले मंत्रालय; एनएचबी: राष्ट्रीय बागवानी बोर्ड; एमएएफडब्ल्यू: कृषि और किसान कल्याण मंत्रालय; डब्ल्यूपीआई: थोक मूल्य सूचकांक; एमसीआई: वाणिज्य और उद्योग मंत्रालय; आईएमएफ: अंतर्राष्ट्रीय मुद्रा कोष; एमएलई: श्रम और रोजगार मंत्रालय; आईएमडी: भारतीय मौसम विभाग; सीएमआईई: भारतीय अर्थव्यवस्था की निगरानी के लिए केंद्र; एमओपीएनजी: पेट्रोलियम और प्राकृतिक गैस मंत्रालय; एफबीआईएल: फाइनेंशियल बेंचमार्क इंडिया प्राइवेट लिमिटेड।

सारणी ए 3: चयनित प्रमुख घटक और चर महत्व

थोक/मंडी मूल्य चर	पीसी 1	पीसी 2
मंडी भाव एमओएम$_{t}$	*0.28*	*0.01*
मंडी भाव एमओएम$_{t-1}$	*0.15*	*0.22*
मंडी भाव एमओएम$_{t-2}$	*0.25*	*0.28*
डबल्यूपीआई खाद्य एमओएम$_{t-1}$	*0.10*	*0.31*
डबल्यूपीआई खाद्य एमओएम$_{t-2}$	*0.22*	*0.18*

अन्य चर	पीसी 1	पीसी 2	पीसी 3	पीसी 4	पीसी 5	पीसी 6
विनिमय दर एमओएम$_{t}$	*0.02*	*0.00*	*0.00*	*0.13*	*0.01*	*0.02*
विनिमय दर एमओएम$_{t-1}$	*0.00*	*0.03*	*0.07*	*0.01*	*0.00*	*0.00*
विनिमय दर एमओएम$_{t-2}$	*0.06*	*0.02*	*0.08*	*0.03*	*0.00*	*0.11*
गूगल रुझान एमओएम$_{t}$	*0.03*	*0.16*	*0.01*	*0.09*	*0.19*	*0.08*
गूगल ट्रेंड्स एमओएम$_{t-1}$	*0.04*	*0.24*	*0.00*	*0.02*	*0.00*	*0.04*
गूगल ट्रेंड्स एमओएम$_{t-2}$	*0.08*	*0.04*	*0.04*	*0.02*	*0.08*	*0.01*
आईएमएफ कृषि कच्चे माल एमओएम$_{t}$	*0.02*	*0.01*	*0.00*	*0.01*	*0.01*	*0.04*
आईएमएफ कृषि कच्चे माल एमओएम$_{t-1}$	*0.00*	*0.02*	*0.01*	*0.03*	*0.00*	*0.00*
आईएमएफ कृषि कच्चे माल एमओएम$_{t-2}$	*0.00*	*0.00*	*0.01*	*0.00*	*0.06*	*0.01*
आईएमएफ खाद्य और पेय पदार्थ एमओएम$_{t}$	*0.04*	*0.05*	*0.01*	*0.03*	*0.01*	*0.07*
आईएमएफ खाद्य और पेय पदार्थ एमओएम$_{t-1}$	*0.00*	*0.02*	*0.00*	*0.00*	*0.02*	*0.08*
आईएमएफ खाद्य और पेय पदार्थ एमओएम$_{t-2}$	*0.01*	*0.00*	*0.01*	*0.00*	*0.07*	*0.01*
मंडी आगमन एमओएम$_{t}$	*0.12*	*0.09*	*0.00*	*0.10*	*0.01*	*0.01*
मंडी आगमन एमओएम$_{t-1}$	*0.01*	*0.12*	*0.12*	*0.00*	*0.01*	*0.00*
मंडी आगमन एमओएम$_{t-2}$	*0.17*	*0.00*	*0.12*	*0.07*	*0.01*	*0.06*
तेल की कीमत एमओएम$_{t}$	*0.01*	*0.04*	*0.02*	*0.05*	*0.15*	*0.09*
तेल की कीमत एमओएम$_{t-1}$	*0.03*	*0.00*	*0.08*	*0.00*	*0.02*	*0.15*
तेल की कीमत एमओएम$_{t-2}$	*0.02*	*0.00*	*0.01*	*0.00*	*0.01*	*0.00*
वर्षा विचलन$_{t}$	*0.00*	*0.00*	*0.25*	*0.00*	*0.01*	*0.00*
वर्षा विचलन$_{t-1}$	*0.03*	*0.03*	*0.01*	*0.25*	*0.08*	*0.10*
वर्षा विचलन$_{t-2}$	*0.01*	*0.03*	*0.05*	*0.02*	*0.05*	*0.05*
जलाशय विचलन$_{t}$	*0.01*	*0.00*	*0.00*	*0.00*	*0.01*	*0.01*
जलाशय विचलन$_{t-1}$	*0.03*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
जलाशय विचलन$_{t-2}$	*0.04*	*0.00*	*0.01*	*0.01*	*0.01*	*0.02*
ग्रामीण मजदूरी एमओएम$_{t}$	*0.07*	*0.00*	*0.09*	*0.02*	*0.05*	*0.00*
डबल्यूपीआई फार्म इनपुट एमओएम$_{t-1}$	*0.00*	*0.03*	*0.00*	*0.03*	*0.07*	*0.00*
डबल्यूपीआई फार्म इनपुट एमओएम$_{t-2}$	*0.09*	*0.03*	*0.00*	*0.03*	*0.04*	*0.04*
डबल्यूपीआई औद्योगिक इनपुट एमओएम$_{t-1}$	*0.01*	*0.01*	*0.00*	*0.01*	*0.01*	*0.00*
डबल्यूपीआई औद्योगिक इनपुट एमओएम$_{t-2}$	*0.05*	*0.03*	*0.00*	*0.05*	*0.03*	*0.00*

नोट: प्रत्येक सेल प्रत्येक पीसी में चर के महत्व को मापता है। प्रत्येक पीसी के भीतर उच्च महत्व को गहरे हरे रंग की छाया से दर्शाया गया है।
स्रोत: आरबीआई स्टाफ अनुमान।

सारणी ए4: अंतिम इनपुट चर की सूची

	पीसीए से पहले		पीसीए और एफएस के बाद
क्र.सं.	**इनपुट चर (एमओएम, प्रतिशत)**	**समकालीन + विलम्बित**	**अंतिम चर**
1	खुदरा मूल्य सूचकांक (आरपीआई)	1	खुदरा मूल्य (डीसीए और एनएचबी) सूचकांक गति
2	डब्ल्यूपीआई खाद्य	2	2 प्रमुख घटक
3	मंडी मूल्य सूचकांक	3	
4	डबल्यूपीआई कृषि इनपुट सूचकांक	2	6 प्रमुख घटक
5	डबल्यूपीआई औद्योगिक इनपुट सूचकांक		
6	ग्रामीण मजदूरी	1	
7	एलपीए से पूर्ण वर्षा विचलन (प्रतिशत में)	3	
8	एलपीए से जलाशय स्तर विचलन (प्रतिशत में)		
9	बाजार आगमन सूचकांक		
10	आईएमएफ खाद्य एवं पेय सूचकांक		
11	आईएमएफ कृषि कच्चा माल सूचकांक		
12	तेल की कीमत		
13	विनिमय दर		
14	गूगल ट्रेंड्स सूचकांक		
	कुल	35	9

नोट: 1. सभी चर महीने-दर-महीने (एमओएम) प्रतिशत परिवर्तन के रूप में हैं, एलपीए (प्रतिशत) से पूर्ण वर्षा विचलन और एलपीए (प्रतिशत) से जलाशय स्तर विचलन को छोड़कर।

2. एफएस: फॉरवर्ड सिलेक्शन; एलपीए: लॉन्ग पीरियड एवरेज।

3. गूगल ट्रेंड्स इंडेक्स भारतीय क्षेत्र के लिए “खाद्य मूल्य” शब्द से मेल खाता है।

सारणी ए 5: वैकल्पिक नाउकास्टिंग मॉडल पर विस्तार से जानकारी

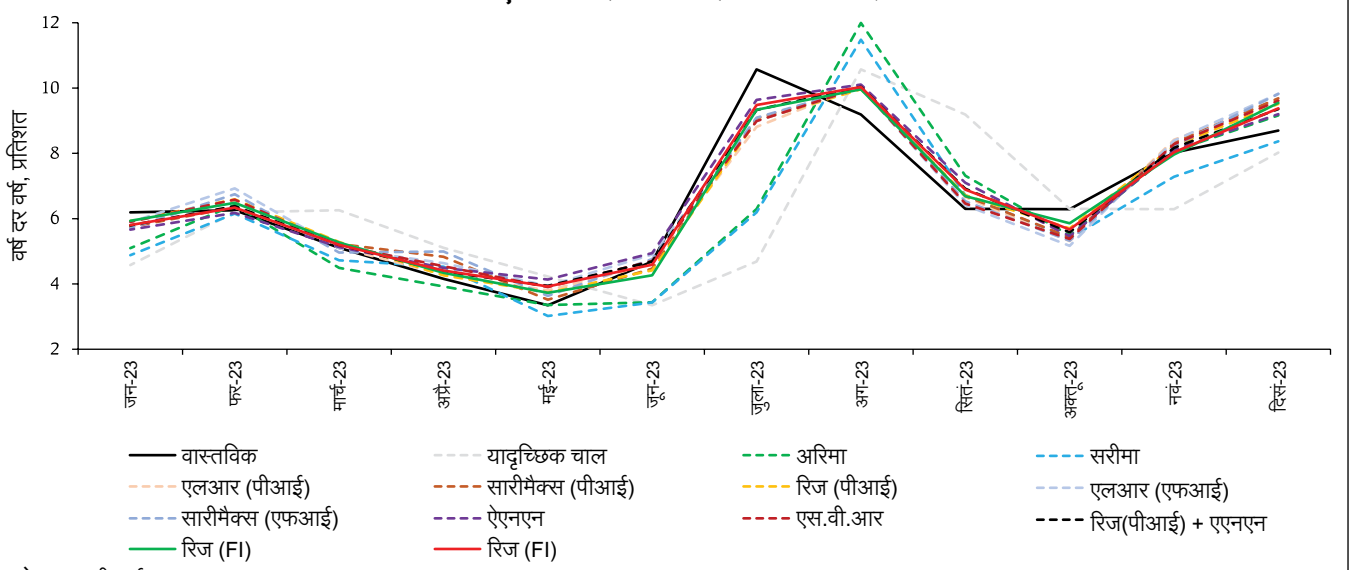
आश्रित चर: वाई = सीपीआई खाद्य गति					
मॉडल प्रकार	**जटिलता**	**नमूना**	**व्याख्यात्मक चर और मॉडल संरचना**	**आर²**	**एआईसी**
एकतरफा रैखिक	*कोई वैकल्पिक जानकारी नहीं*	*अरिमा*	*सी, (पी,डी,क्यू) = (4,0,2)*	*0.43*	*2.79*
		सारिमा	*सी, (पी,डी,क्यू)(पी,डी,क्यू) = (1,0,2)(1,0,1)*	*0.50*	*2.77*
बहुभिन्नरूपी रैखिक	*खुदरा और थोक मूल्य की जानकारी*	*एलआर*	*सी, सीपीआई खाद्य गति टी -12 , आरपीआई गति टी , डब्ल्यूपीआई और मंडी खाद्य कीमतों को कैप्चर करने वाले पीसी*	*0.79*	*1.63*
		सारीमैक्स	*सी, (पी,डी,क्यू)(पी,डी,क्यू) = (1,0,1)(0,0,0), आरपीआई मोमेंटम टी , पीसी डब्ल्यूपीआई और मंडी खाद्य कीमतों को कैप्चर कर रहे हैं*	*0.81*	*1.65*
		रिज	*सी, आरपीआई मोमेंटम टी , डब्ल्यूपीआई और मंडी खाद्य मूल्य चर विधि = के-फोल्ड*	*0.81*	*-*
	पूरी जानकारी	*एलआर*	*सी, सीपीआई खाद्य गति$_{टी-12}$, आरपीआई गति$_{टी}$, डब्ल्यूपीआई और मंडी खाद्य कीमतों और बाकी जानकारी कैप्चर करने वाले पीसी*	*0.85*	*1.45*
		सारीमैक्स	*सी, (पी,डी,क्यू)(पी,डी,क्यू) = (1,0,1)(0,0,0), आरपीआई मोमेंटम टी , डब्ल्यूपीआई और मंडी खाद्य कीमतों और बाकी जानकारी को कैप्चर करने वाले पीसी*	*0.88*	*1.47*
		रिज	*सी, सभी चर खाद्य कीमतों और शेष जानकारी को कैप्चर करते हैं विधि = के-फोल्ड*	*0.88*	*-*
बहुभिन्नरूपी एमएल-आधारित	*पूरी जानकारी + एमएल*	*ऐन*	*सी, सीपीआई खाद्य गति टी -12 , आरपीआई गति टी , डब्ल्यूपीआई और मंडी खाद्य कीमतों और बाकी जानकारी कैप्चर करने वाले पीसी हाइपरपैरामीटर: 3 नोड्स के साथ 1 छिपी हुई परत, सक्रियण फ़ंक्शन = सिग्मॉइड, सीखने की दर = 0.005, रन = 1000*	*-*	*-*
		एसवीआर	*सी, सीपीआई खाद्य गति$_{टी-12}$, आरपीआई गति$_{टी}$, डब्ल्यूपीआई और मंडी खाद्य कीमतों और शेष जानकारी कैप्चर करने वाले पीसी हाइपरपैरामीटर: कर्नेल = लीनियरसी (नियमितीकरण पैरामीटर) = 0.1एप्सिलॉन = 0.01*	*-*	*-*

टिप्पणियाँ: 1. 'p', 'd' और 'q' ऑटोरिग्रैसिव, डिफरेंसिंग और मूविंग एवरेज ऑर्डर्स को संदर्भित करते हैं, जबकि 'P', 'D' और 'Q' मौसमी ऑटोरिग्रैसिव, मौसमी डिफरेंसिंग और मूविंग एवरेज ऑर्डर्स हैं; 'X' एक्सोजेनस चरों के सेट को संदर्भित करता है।

2. C: स्थिर; एआईसी: अकाइक सूचना मानदंड; 'X tn ' चर X के n वें अंतराल को संदर्भित करता है।

स्रोत: RBI स्टाफ अनुमान।

चार्ट ए 1: वैकल्पिक मॉडलों की नाउकास्ट



स्रोत: आरबीआई स्टाफ का अनुमान।