RESEARCH ARTICLE Bhartiya Krishi Anusandhan Patrika, Volume 37 Issue 1: 18-22 (March 2022)

डिकंपोजिसन आधारित हाइब्रिड मॉडल का उपयोग करके कृषि मूल्य पूर्वानुमान

कपिल चौधरी[1], गिरीश कुमार झा[2], राजीव रंजन कुमार[1], रोनित जैसवाल[1]

10.18805/BKAP435

सारांश

वैश्वीकरण और बाजार एकीकरण के कारण सभी स्तरों पर निर्णय लेने के लिए कृषि मूल्य सूचना की जरूरतें बढ़ रही हैं। कृषि की जैविक प्रक्रियाओं पर अत्यधिक निर्भरता के कारण कृषि मूल्य पूर्वानुमान समय श्रृंखला विश्लेषण के सबसे कठिन क्षेत्रों में से एक है। इस अध्यन में, आलू की कीमतों का अनुमान लगाने के लिए इंपीरिकल मोड डिकंपोजिसन (ईएमडी) पर आधारित एक न्यूरल नेटवर्क मॉडल का उपयोग किया गया है। चेन्नई बाजार के आलू की मासिक थोक मूल्य श्रृंखला को पांच स्वतंत्र इनट्रिनसिक मोड फंक्शन (आईएमएफ) और विभिन्न आवृत्तियों के साथ एक अवशेष में विघटित किया गया है फिर, इन आईएमएफ एवं अवशेष घटकों को स्वतंत्र रूप से पूर्वानुमान लगाने के लिए, एक हिडन लेयर के साथ आर्टिफिशियल न्यूरल नेटवर्क बनाया गया है। अंत में, सभी आईएमएफ के पूर्वानुमान परिणामों को एकत्रित करके मूल मूल्य श्रृंखला के लिए समेकित आउटपुट बनाया गया है।

कुंजी शब्दः आर्टिफिशियल न्यूरल नेटवर्क, इंपीरिकल मोड डिकंपोजिसन (ईएमडी), इनट्रिनसिक मोड फंक्शन, मूल्य पूर्वानुमान।

Agricultural Price Forecasting using Decomposition-based Hybrid Model

Kapil Choudhary[1], Girish Kumar Jha[2], Rajeev Ranjan Kumar[1], Ronit Jaiswal[1]

ABSTRACT Agricultural price information needs for decision-making at all levels are increasing due to globalization and market integration. Due to its great reliance on biological processes, agricultural price forecasting is one of the most difficult fields of time series analysis. In this paper, a neural network model based on empirical mode decomposition is used to forecast potato prices. The monthly wholesale price series of potato from Chennai market was decomposed into five independent intrinsic modes (IMFs) and one residual with various frequencies. Then, to forecast these IMFs and residual components independently, an artificial neural network with a single hidden layer was built. Finally, the ensemble output for the original price series is formed by aggregating the forecast outcomes of all IMFs, including residuals. In terms of root mean square error and directional prediction statistics, empirical data show that the suggested ensemble model outperforms a single model.

Key words: Artificial neural network, Empirical mode decomposition, Intrinsic mode function, Price forecasting.

प्रस्तावना

समय पर और विश्वसनीय कृषि मूल्य पूर्वानुमान किसानों, नीति नियोजकों और कृषि आधारित उद्योगों के लिए उपयोगी होते हैं। कृषि की जैविक घटनाओ पर निर्भरता के कारण कृषि वस्तुओं की मूल्यों का पूर्वानुमान एक चुनौतीपूर्ण विषय है। बनर्जी और अन्य (2021) के अनुसार कृषि वस्तुओं की कीमतें अक्सर यादृच्छिक होती हैं क्योंकि वे बड़े पैमाने पर मौसम चर घटनाओं से प्रभावित होती हैं। इससे कृषि वस्तुओं के मूल्य मॉडलिंग और पूर्वानुमान की प्रक्रिया में काफी जोखिम और अनिश्चितता पैदा होती है। बॉक्स और अन्य (1970) के अनुसार

[1]ICAR-Indian Agricultural Statistics Research Institute, New Delhi-110 012, India.
[2]ICAR-Indian Agricultural Research Institute, New Delhi-110 012, India.

Corresponding Author: Rajeev Ranjan Kumar, ICAR-Indian Agricultural Statistics Research Institute, New Delhi-110 012, India. Email: rrk.uasd@gmail.com

How to cite this article: Choudhary, K., Jha, G.K., Kumar, R.R. and Jaiswal, R. (2022). Agricultural Price Forecasting using Decomposition-based Hybrid Model. Bhartiya Krishi Anusandhan Patrika. 37(1): 18-22. DOI: 10.18805/BKAP435.

Submitted: 18-01-2022 **Accepted:** 11-05-2022 **Online:** 16-05-2022

समय श्रृंखला पूर्वानुमान का मुख्य लक्ष्य अपने स्वयं के ऐतिहासिक डेटा के साथ–साथ संबंधित चर के आधार पर उपयोगी चर के भविष्य के मूल्यों का पूर्वानुमान करना है। कुमार और अन्य (2019) ने पाया कि पिछले कई दशकों में, शोधकर्ताओं ने प्रभावी और कुशल समय श्रृंखला पूर्वानुमान मॉडल के डिजाइन और विकास में गहरी रुचि ली है।

मौजूदा साहित्य के अनुसार, प्रचुर मात्रा में डेटा विश्लेषण रूपों का प्रस्ताव दिया गया है और समय श्रृंखला डेटा पर उपयोग किया गया है। डेटा विश्लेषण के पारंपरिक रूप जैसे फूरियर स्पेक्ट्रल विश्लेषण ने समय श्रृंखला डेटा का विश्लेषण करने के लिए एक सामान्य विधि प्रदान की है। हालांकि फूरियर ट्रांसफॉर्म बेहद सामान्य परिस्थितियों में मान्य है। इस अध्ययन में नान–स्टेसनरी और नान–लिनियर समय श्रृंखला से विशिष्ट जानकारी निकालने के लिए इंपीरिकल मोड डिकंपोजिसन (ईएमडी) का प्रयोग किया गया है। ईएमडी का मुख्य उद्देश्य एक नान–स्टेसनरी एवं नान–लिनियर समय श्रृंखला डेटा को कई स्वतंत्र इनट्रिनसिक मोड फंक्शंस (आईएमएफ) और एक अवशेष में विभिन्न आयामों और आवृत्तियों के साथ विघटित करना है। इस अध्यन में, चेन्नई बाजार की एक जटिल मासिक आलू थोक मूल्य श्रृंखला के लिए ईएमडी–एएनएन मॉडल का उपयोग किया गया है।

सामग्री एवं परीक्षण विधि

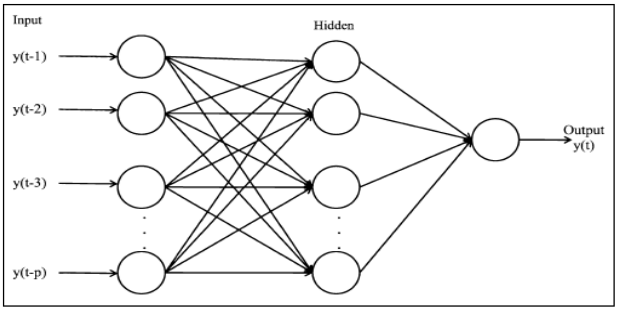
हुंग और अन्य (1998) के अनुरूप इंपीरिकल मोड डिकंपोजिसन (ईएमडी), नान–स्टेसनरी और नान–लिनियर समय श्रृंखला डेटा के लिए अनुकूल अपघटन तकनीक का एक रूप है। ईएमडी का मुख्य उद्देश्य एक नान–स्टेसनरी और नान–लिनियर समय श्रृंखला डेटा को कई सरल मोड (आईएमएफ और अवशेष) में विघटित करना है। इसका अपघटन मूल्य श्रृंखला के स्थानीय विशिष्ट समय पैमाने पर आधारित है। चौधरी और अन्य (2019) ने बताया कि मूल श्रृंखला की तुलना में, विघटित श्रृंखला का पूर्वानुमान आसानी से बेहतर सटीकता के साथ किया जा सकता है। मशीन लर्निंग तकनीकों में, विशेष रूप से आर्टिफिशियल न्यूरल नेटवर्क (एएनएन) का बड़े पैमाने पर अध्ययन किया गया है और समय श्रृंखला पूर्वानुमान के लिए उपयोग किया जाता है। पारंपरिक पूर्वानुमान दृष्टिकोणों के विपरीत, एएनएन व्यापक डेटा या ज्ञान के बिना नान–लिनियर और अनुमानित जटिल संबंधों को अनुकूलित करने में सक्षम हैं। टेम्पोरल डेटा को दो तरह से न्यूरल नेटवर्क का उपयोग करके मॉडल किया जा सकता है, या तो आवर्तक न्यूरल नेटवर्क की मदद से या नोड्स के आउटपुट पर बफर प्रदान करके। इस अध्ययन में, एकल हिडन लेयर के साथ फीड–फॉरवर्ड आर्टिफिशियल न्यूरल नेटवर्क (एएनएन) को विघटित आईएमएफ और अवशेष घटक के मॉडलिंग के लिए बहु–स्तरीय शिक्षण उपकरण के रूप में नियोजित किया गया है। झा और अन्य (2013) के द्वारा प्रस्तावित एकल हिडन लेयर के साथ फीड–फॉरवर्ड आर्टिफिशियल न्यूरल नेटवर्क (एएनएन) को विघटित आईएमएफ और अवशेष घटक के मॉडलिंग के लिए उपयोग किया गया है।

चौधरी और अन्य (2019) ईएमडी और एएनएन क्रमशः एक समय श्रृंखला की नान–स्टेसनरी और नान–लिनियर विशेषताओं के लिए फायदेमंद हैं। ईएमडी और एएनएन को मिलाकर कृषि मूल्य श्रृंखला के नान–स्टेसनरी और नान–लिनियर व्यवहार को ध्यान में रखने का प्रयास किया गया है। चित्र 1 में एक आर्टिफिशियल न्यूरल नेटवर्क जिसे वर्तमान अध्ययन में इस्तेमाल किया गया है, को दिखाया गया है।

यह शोध कार्य भा.कृ.अनु.प.–भारतीय कृषि सांख्यिकी अनुसंधान संस्थान, नई दिल्ली में वर्ष २०२२ में संपन्न किया गया है।

परिणाम एवं विवेचना

वर्तमान अध्ययन के लिए, जनवरी 2010 से जुलाई 2020 की अवधि तक राष्ट्रीय बागवानी अनुसंधान और विकास फाउंडेशन (एनएचआरडीएफ) से एकत्र किए गए चेन्नई बाजार के मासिक आलू थोक मूल्य (रुपये प्रति क्विंटल) का प्रयोग किया गया है। मूल मूल्य श्रृंखला को दो सेटों में विभाजित किया गया है, अर्थात् प्रशिक्षण सेट (जनवरी 2010 से जुलाई 2019) एवं परीक्षण सेट (अगस्त 2019 से जुलाई 2020)। यहां, डार्बेले और अन्य (2000) के अनुसार प्रशिक्षण सेट का उपयोग पैरामीटर के आकलन के साथ–साथ मॉडल की सामान्यीकरण क्षमता को मापने के लिए किया गया है। नेल्सन और अन्य (1999) के अनुसार परीक्षण सेट का उपयोग मॉडल के नमूना प्रदर्शन से बाहर का आकलन करने के लिए किया गया है। चित्र 2 में



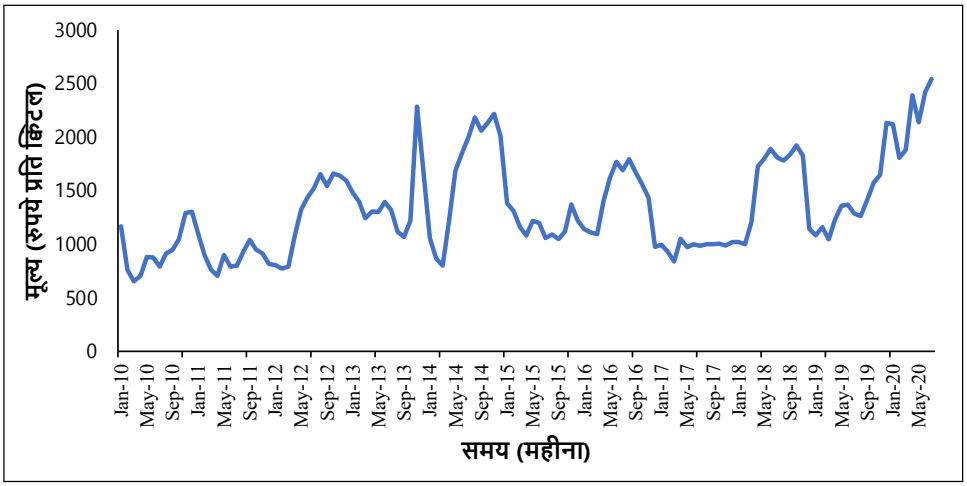
चित्र 1: एक हिडन लेयर के साथ एएनएन।

समय श्रृंखला का प्लॉट दिया गया है, जिससे स्पष्ट रूप से समय श्रृंखला की नान–स्टेसनरी और नान–लिनियर व्यवहार का पता चलता है। इस अध्ययन में प्रयुक्त मूल्य श्रृंखला के मूल वर्णनात्मक आंकड़े तालिका 1 में प्रस्तुत किए गए हैं।

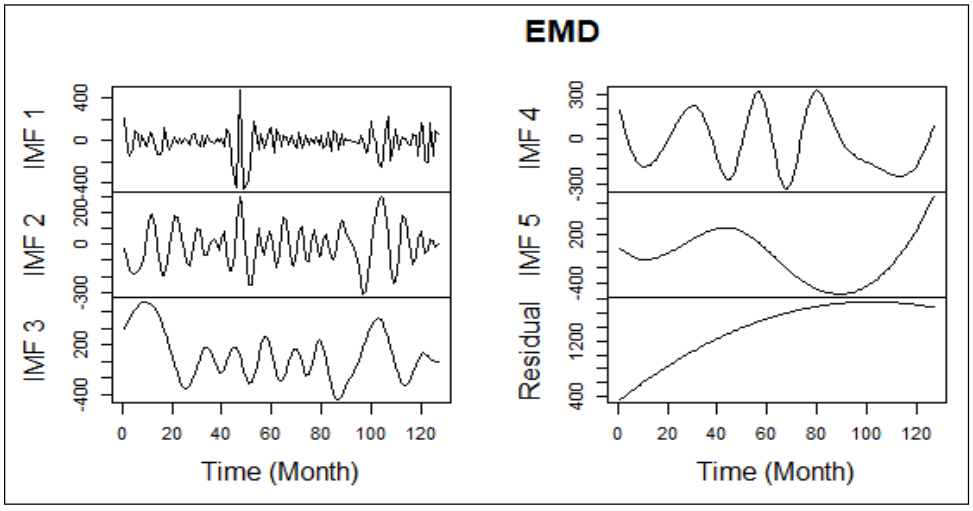
समय श्रृंखला आकड़ों को विघटित करने के लिए ईएमडी का उपयोग करने से पहले, इसकी स्टेसनरी और लिनियर प्रकृति का परीक्षण किया जाता है। डेटा श्रृंखला की स्टेसनरीटी का परीक्षण करने के लिए, हम लोगो ने ऑगमेंटेड–डिकी–फुलर (ADF) टेस्ट का उपयोग किए हैं। एडीएफ परीक्षण के परिणाम से पता चलता है कि मूल्य श्रृंखला नान–स्टेसनरी हैं। मूल्य श्रृंखला की रैखिकता विशेषताओं का परीक्षण करने के लिए, हमने Brock-Decher-Scheikman (BDS) परीक्षण का उपयोग किया है। इस परीक्षण में आयाम 2 और 3 को एम्बेड करने के लिए, संभाव्यता मान 0.001 से कम थे, जो दर्शाता है कि मूल्य श्रृंखला 1 प्रतिशत स्तर पर भी गैर–रैखिक है।

इसके बाद, ईएमडी के माध्यम से मूल श्रृंखला को अलग–अलग स्वतंत्र पाँच इनट्रिनसिक मोड फंक्शंस (आईएमएफ) तथा एक अवशेष में विघटित किया गया है। जिसे चित्र 3 में दिखाया गया है, जो स्पष्ट रूप से प्रदर्शित करता है कि आईएमएफ 1 से आईएमएफ 5 तक आयाम बढ़ रहा है जबकि आवृत्ति घट रही है।

ईएमडी–एएनएन मॉडल के पूर्वानुमान परिणामों की तुलना आर्टिफिशियल न्यूरल नेटवर्क (एएनएन) मॉडल से की गयी है। झांग और अन्य (2008) के अनुसार ईएमडी कृषि मूल्य पूर्वानुमान के लिए उपयुक्त है क्योंकि यह एक जटिल डेटा को सरल स्वतंत्र इनट्रिनसिक मोड फंक्शंस में विघटित कर सकता है। चूंकि अपघटन डेटा के स्थानीय विशिष्ट समय पैमाने पर अ धारित होता है, इसलिए विघटित मोड प्रकृति में स्टेसनरी और नान–लिनियर हो जाते हैं। ईएमडी समय श्रृंखला के छिपे हुए पैटर्न और प्रवृत्तियों को भी प्रकट कर सकता है,



चित्र 2: चेन्नई बाजार की मासिक आलू थोक मूल्य श्रृंखला का टाइम प्लॉट।



चित्र 3: चेन्नई बाजार के मासिक आलू थोक मूल्य के लिए आईएमएफ और अवशेष।

तालिका 1: मूल्य (रुपये प्रति क्विंटल) श्रृंखला के वर्णनात्मक आंकड़े।

सांख्यिकी	वैल्यू
औसत	1329.20
अधिकतम	2543.50
न्यूनतम	652.90
मानक विचलन	434.84
तिरछापन	0.72
कुकुदता	2.71
जार्क–बेरा	11.54

तालिका 2: चेन्नई बाजार के मासिक आलू थोक मूल्य के लिए आरएमएसई और क्त्रेजंज से प्रत्येक पूर्वानुमान मॉडल की तुलना।

मॉडल	आरएमएसई	क्त्रेजंज
एएनएन	36.91	54.54
ईएमडी–एएनएन	19.75	72.73

जो एएनएन की पूर्वानुमान क्षमता को बढ़ाता है। किसी समय श्रृंखला पूर्वानुमान के लिए एएनएन आर्किटेक्चर में प्रत्येक लेयर में नोड्स की कुल संख्या का निर्धारण किया जाता है। झांग और अन्य (2015) के अनुसार न्यूरल नेटवर्क एक सार्वभौमिक सन्निकटन होने के कारण किसी भी नान–लिनियर फंक्शन को एक हिडन लेयर के साथ मैप कर सकता है जिसमें हिडन लेयर पर पर्याप्त संख्या में नोड्स और प्रशिक्षण के लिए पर्याप्त डेटा बिंदु दिए गए हैं। तदनुसार, झू (2012) के अनुसार इस अध्ययन में, एक हिडन लेयर के साथ एएनएन का उपयोग किया गया है। इसके अलावा हिडन लेयर पर नोड्स की संख्या नमूना पूर्वानुमान क्षमता में सुधार करती है और ओवरफिटिंग समस्या से भी बचाती है। छिपे हुए नोड्स की संख्या प्रयोग की मदद से निर्धारित की जाती है। हम बुनियादी क्रॉस सत्यापन पद्धति के साथ छिपे हुए नोड्स की संख्या को 2 से 20 तक भिन्न करते हैं। यह भी देखा जा सकता है कि 12 छिपे हुए नोड्स वाला एक न्यूरल नेटवर्क मॉडल नमूना पूर्वानुमान सटीकता उपायों के संबंध में अन्य प्रतिस्पर्धी मॉडल की तुलना में बेहतर प्रदर्शन करता है। लॉजिस्टिक सिग्मॉइड और पहचान कार्यों को क्रमशः छिपे हुए और आउटपुट नोड्स के लिए एक सक्रियण फंक्शन के रूप में नियोजित किया गया है।

इस अध्यन में, हमारा ध्यान अल्पकालिक पूर्वानुमान पर है, और इसलिए, हम बारह महीनो तक के पूर्वानुमान क्षितिज पर विचार किए हैं। क्लेमेंट्स और अन्य (1977) ने ईएमडी आधारित एएनएन की पूर्वानुमान क्षमता की तुलना आरएमएसई और D_{stat} के संदर्भ में एएनएन से की है। तालिका 2, स्पष्ट रूप से प्रदर्शित करती है कि ईएमडी–एएनएन की पूर्वानुमान सटीकता एएनएन मॉडल की तुलना में बेहतर है क्योंकि इसके कम आरएमएसई और उच्च D_{stat} है। विशेष रूप से, एएनएन के आरएमएसई (36.91) की तुलना में, ईएमडी–एएनएन का आरएमएसई केवल 19.75 है। इसी तरह एएनएन का D_{stat} 54.54 है जबकि ईएमडी–एएनएन मॉडल के मामले में D_{stat} 72.73 है।

निष्कर्ष

इस अध्ययन में आलू के चेन्नई बाजार के लिए मासिक थोक मूल्य पूर्वानुमान करने के लिए एक हाइब्रिड (ईएमडी–एएनएन) मॉडल प्रस्तावित किया गया है। जैसा कि हम जानते हैं कि एएनएन मॉडल समय श्रृंखला की नान–लिनियर प्रकृति को संभालते हैं, लेकिन जब श्रृंखला में नान–स्टेसनरी प्रकृति भी होता है तो यह ऐसी स्थिति को संभालने में असमर्थ होता है। ऐसी श्रृंखला के लिए हमने नान–स्टेसनरी से निपटने के लिए ईएमडी को प्रीप्रोसेसर के रूप में इस्तेमाल किया है। प्रस्तावित मॉडल से बुवाई एवं फसल की कटाई से पहले सटीक मूल्य पूर्वानुमान किया जा सकता है, जिससे किसानों को सही निर्णय लेने में मदद मिलेगी। इससे सरकार के लिए आयात और निर्यात नीति की अग्रिम योजना बनाने में भी आसानी होगी।

REFERENCES

Banerjee, R., Das, P., Ahmad, T. and Kumar, M. (2021). Modeling and forecasting of agricultural commodity production under changing climatic condition: A review. Bhartiya Krishi Anusandhan Patrika. 36(4): 273-279.

Box, G.E.P. and Jenkins, G.M. (1970). Time Series Analysis: Forecasting and Control. Holden-Day, San Francisco, CA.

Clements, M.P. and Smith, J. (1997). The performance of alternative forecasting methods for SETAR models. International Journal of Forecasting. 13(4): 463-475.

Choudhary, K., Jha, G.K. and Kumar, R.R. (2019). Delhi Potato price analysis using ensemble empirical mode decomposition. Bharatiya Krishi Anusandhan Patrika. 34(1): 33-37.

Choudhury, K., Jha, G.K., Kumar, R.R. and Mishra, D.C. (2019). Agricultural commodity price analysis using ensemble empirical mode decomposition: A case study of daily potato price series. Indian Journal of Agricultural Sciences. 89(5): 882-886.

Darbellay, G.A. and Slama, M. (2000). Forecasting the short-term demand for electricity: Do neural networks stand a better chance? International Journal of Forecasting. 16(1): 71-83.

Huang, N.E., Shen, Z., Long, S.R., Wu, M. C., Shih, H.H., Zheng, Q. and Liu, H.H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. In Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences. 454: 903-995.

Jha, Girish, K. and Sinha, K. (2013). Agricultural price forecasting using neural network model: An innovative information delivery system. Agricultural Economics Research Review. 26(2): 229-239.

Nelson, M., Hill, T., Remus, W. and O'Connor, M. (1999). Time series forecasting using neural networks: Should the data be deseasonalized first?. Journal of Forecasting. 18(5): 359-367.

Kumar, R.R., Jha, G.K., Choudhary, K. and Budhlakoti, N. (2019). Examining integration between Agra and Delhi potato markets. Bhartiya Krishi Anusandhan Patrika. 34(1): 62-64.

Zhang, X., Lai, K.K. and Wang, S.Y. (2008). A new approach for crude oil price analysis based on empirical mode decomposition. Energy Economics. 30(3): 905-918.

Zhang, J.L., Zhang, Y.J. and Zhang, L. (2015). A novel hybrid method for crude oil price forecasting. Energy Economics. 49: 649-659.

Zhu, B. (2012). A novel multiscale ensemble carbon price prediction model integrating empirical mode decomposition, genetic algorithm and artificial neural network. Energies. 5(2): 355-370.