

DOI : <https://doi.org/10.54302/mausam.v75i3.5606>Homepage: <https://mausamjournal.imd.gov.in/index.php/MAUSAM>

UDC No. 633.11 : 004.383.8

Weather based wheat yield prediction using machine learning

SHREYA GUPTA, ANANTA VASHISTH, P.KRISHNAN, ACHAL LAMA*,
SHIVPRASAD and ARAVIND K.S.

ICAR-Indian Agricultural Research Institute, Pusa, New Delhi – 110 012

*ICAR-Indian Agricultural Statistics Research Institute, Pusa, New Delhi – 110 012

(Received 21 March 2022, Accepted 11 May 2023)

e mail :ananta.iari@gmail.com

सार – मौसम के प्राचलों के प्रभाव से गेहूं की फसलें अत्यधिक प्रभावित होती हैं। इसके विश्वसनीय पूर्वानुमान के लिए मशीन लर्निंग का उपयोग करके मौसम-आधारित मॉडल विकसित और मान्य करने की आवश्यकता है। फसल उगाने की अवधि के दौरान गेहूं की उपज और मौसम के आंकड़े आईएआरआई, नई दिल्ली, हिसार, अमृतसर, लुधियाना और पटियाला से एकत्र किए गए। उपज अनुमान मॉडल को चरणबद्ध बहुरेखिक समाश्रयण (SMLR), सपोर्ट वेक्टर समाश्रयण (SVR), लीस्ट एब्सोल्यूट श्रिकेज & सलेक्शन ऑपरेटर (LASSO) और हाइब्रिड मशीन लर्निंग मॉडल LASSO-SVR तथा आर सॉफ्टवेयर में SMLR-SVR का उपयोग करके विकसित किया गया। अंशांकन के लिए 70% डेटा और सत्यापन के लिए शेष 30% डेटा तय करके विश्लेषण किया गया। अध्ययन क्षेत्रों के लिए गेहूं की उपज अनुमान मॉडल 46^{वें} से 15^{वें} मानक मौसम सप्ताहों के दीर्घकालिक दैनिक मौसम डेटा के साथ-साथ दीर्घकालिक फसल उपज डेटा का उपयोग करके विकसित किए गए। विभिन्न स्थानों के लिए गेहूं की उपज के अनुमान के लिए इन मॉडलों की जांच करने पर, LASSO ने सबसे अच्छा प्रदर्शन किया, जिसमें nRMSE मान पटियाला के लिए 0.6% से लुधियाना के लिए 4.8% के बीच था। यदि LASSO और SMLR के संयोजन में एक हाइब्रिड मॉडल लागू किया जाता है तो SVR का मॉडल प्रदर्शन बढ़ जाएगा। हाइब्रिड मॉडल LASSO-SVR ने SMLR-SVR की तुलना में SVR मॉडल में अधिक सुधार दिखाया है।

ABSTRACT. Wheat crops are highly affected by the influence of weather parameters. Thus, there is a need to develop and validate weather-based models using machine learning for its reliable prediction. Wheat yield and weather data during the crop growing period were collected from IARI, New Delhi, Hisar, Amritsar, Ludhiana and Patiala. The yield prediction model was developed using stepwise multi linear regression (SMLR), support vector regression (SVR), least absolute shrinkage and selection operator (LASSO) and hybrid machine learning model LASSO-SVR and SMLR-SVR in R software. Analysis was done by fixing 70% of the data for calibration and the remaining 30% data for validation. Wheat yield prediction models for study areas were developed using long term crop yield data along with long period daily weather data from the 46th to 15th standard meteorological weeks. On examining these models for wheat yield prediction for different locations, LASSO performed best having nRMSE value ranged between 0.6 % for Patiala to 4.8% for Ludhiana. The model performance of SVR is increased if a hybrid model in combination with LASSO and SMLR is applied. The hybrid model LASSO-SVR has shown more improvement in SVR model compared with SMLR-SVR.

Key words– Weather variable, Machine learning model, Support vector regression, Least absolute shrinkage and selection operator, Stepwise multi linear regression, Yield prediction.