



## Prediction of rainfall and groundwater using machine learning Algorithms for Nagpur division

TULSHIDAS M. JIBHAKATE and YASHWANT B. KATPATAL\*

*Dept. of Civil Engineering, Visvesvaraya National Institute of Technology, Nagpur, Maharashtra – 440 010, India*

*\*Dept. of Civil Engineering, Visvesvaraya National Institute of Technology, Nagpur, Maharashtra – 440 010, India*

*(Received 30 May 2023, Accepted 8 April 2024)*

**e mails : [tulshimj@gmail.com](mailto:tulshimj@gmail.com); [ybkatpatal@rediffmail.com](mailto:ybkatpatal@rediffmail.com)**

**सार** – वर्षा आर्द्र भूजल का पूर्वानुमान जल संसाधन योजना के लिए आर्द्र सूखे आर्द्र बाढ़ जैसी आपदाओं के परिणामों को कम करने के लिए भी महत्वपूर्ण है। वर्तमान अध्ययन में, सकारात्मक आर्द्र नकारात्मक विसंगतियों की गणना के लिए 20 साल की अवधि (2001-2020) के लिए वर्षा विसंगति सूचकांक (RAI) का आकलन किया गया। भूजल स्तर में उतार-चढ़ाव पर वर्षा के प्रभाव की तुलना करने के लिए अनुमानित न्यूनतम आर्द्र उच्चतम RAI वर्षों का उपयोग किया गया। मशीन लर्निंग एल्गोरिदम का उपयोग करके वर्षा आर्द्र भूजल का अनुमान लगाया गया। नाइव (N), एक्सपोनेंशियल स्मूथिंग (ES), डिसीजन ट्री रिग्रेसर (DT), रैंडम फॉरेस्ट रिग्रेसर (RF), ऑटोएरिमा (AA), के - नेबर रेजिस्टर (KN), आर्द्र ग्रेडिएंट बूस्टर रेजिस्टर (GB) जैसे मशीन लर्निंग एल्गोरिदम का उपयोग करके अध्ययन क्षेत्र में वर्षा आर्द्र भूजल स्तर का अनुमान लगाने के लिए Sktime आर्द्र scikit - लर्न लाइब्रेरी का उपयोग किया गया। नागपुर भूखंड के लिए 2001 से 2015 तक प्रेक्षित मौसमी वर्षा आर्द्र भूजल डेटा के आधार पर, वर्तमान अध्ययन 2016-2020 की अवधि के लिए मूल्यों का अनुमान लगाता है। फिर, 2016-2020 के लिए प्रेक्षित आर्द्र अनुमानित मूल्यों का उपयोग करते हुए, सहसंबंध गुणांक (r), माध्य निरपेक्ष त्रुटि (MAE), नैश-सटक्लिफ दक्षता (NSE), मूल माध्य वर्ग त्रुटि (RMSE), आर्द्र टेलर आरेख जैसे सटीकता मूल्यांकन प्राचलों का सत्यापन आर्द्र सबसे अच्छे आर्द्र सबसे खराब मॉडल पूर्वानुमानकर्ताओं की जांच करने के लिए मूल्यांकन किया गया। वर्तमान अध्ययन में पाया गया है कि वर्षा के मामले में, Auto ARIMA पूर्वानुमान सबसे उपयुक्त मॉडल है, आर्द्र भूजल के मामले में, नैवे पूर्वानुमान मॉडल सबसे उपयुक्त है। डिसीजन ट्री पूर्वानुमान मॉडल वर्षा आर्द्र भूजल मानों दोनों में सबसे खराब-फिट है। फिर, वर्ष 2021-2025 के लिए क्रमशः वर्षा आर्द्र भूजल मूल्यों का पूर्वानुमान करने के लिए Auto ARIMA आर्द्र नैवे मॉडल का उपयोग किया गया। ENSO घटना 2001-2010 के दौरान अधिक प्रमुख थी, आर्द्र 2011-2020 के दौरान, दोनों ISMR को प्रभावित करने वाले प्रेरक कारक हो सकते हैं।

**ABSTRACT.** Rainfall and groundwater predictions are important for water resource planning and also to reduce the consequences of catastrophes like drought and floods. In the present study, Rainfall Anomaly Index (RAI) was estimated for 20 years period (2001-2020) to calculate the positive and negative anomalies. The estimated lowest and highest RAI years were used to compare the effect of rainfall on groundwater level fluctuations. Predictions of rainfall and groundwater were performed using machine learning algorithms. Sktime and scikit-learn libraries were used to predict the rainfall and groundwater levels in the study area using machine learning algorithms such as Naive (N), Exponential Smoothing (ES), Decision Tree Regressor (DT), Random Forest Regressor (RF), Auto ARIMA (AA), K-Neighbour regressor (KN), and Gradient Booster Regressor (GB). Based on the observed seasonal rainfall and groundwater data from 2001 to 2015 for Nagpur division, present study predicts values for the 2016-2020 period. Then, using observed and predicted values for 2016-2020, accuracy assessment parameters like correlation coefficient (r), mean absolute error (MAE), Nash-Sutcliffe efficiency (NSE), root mean square error (RMSE), and Taylor diagram were assessed for validation and to investigate the best and worst model forecasters. The present study observes that in the case of rainfall, the Auto ARIMA forecaster is the best-fitted model, and in the case of groundwater, the naïve forecaster is the best-fitted model. The decision tree forecaster is the worst-fitted model in both rainfall and groundwater data. Then, the Auto ARIMA and Naive models were used to predict rainfall and groundwater values, respectively, for the years 2021-2025. Impact of ENSO and IOD on ISMR has been assessed. The ENSO phenomenon was more prominent during 2001-2010, and during 2011-2020, both may be the driving factors impacting ISMR.

**Key words** – Rainfall, Rainfall anomaly index (RAI), Groundwater, Machine learning algorithms, Accuracy assessment parameters, Taylor diagram, ENSO, IOD and ISMR.