

## **ABSTRACT**

Concerns over prospects for groundwater monitoring were raised by the rising groundwater demand, particularly on the western India. Thus, this study attempted to address the major determinants that influence groundwater and to highlight the anticipated groundwater zones in the Pune district. Well locations for the years 2001 to 2021 were acquired from Groundwater Survey and Development Agency (GSDA) and comprised of about 15,898 records which were pertaining well locations for 12 months (January to December). Additionally, non-well locations were randomly generated using the ArcGIS Pro. Reference data was divided into two segments, (i) both well and non-well locations (4000) points were merged to get the training (80%) and validation (20%) datasets and, (ii) Well (3000) locations were divided into training (80%) and validation (20%) datasets. The first set of reference data was utilized to run the Random Forest (RF) algorithm along with ten groundwater effective factors to analyse the dominant factors among them. Actual Map showing both absence and present points were prepared with the help of both RF and SVM algorithms. In addition, Frequency Ratio (FR) model was implemented to calculate the Prediction Rate (PR) which took the second set of reference data and all the groundwater influential factors as input to produce the Groundwater Potential Zone (GWPZ) map. The percentage of area under the five zones (Very Low, Low, Moderate, High and Very High) was calculated. Finally, the area under the ROC curve (AUC) was implemented to validate the maps.

**Keywords:** Machine learning algorithm, dominant factors, groundwater potential areas, ROC-AUC curve.