

## Prediction of Dissolved Oxygen Concentration in the Yeşilirmak River Using Machine Learning Models

### Eda Göz

Department of Chemical Engineering, Ankara University, Türkiye

### Göktuğ Yüceer

Department of Computer Engineering, Sakarya University, Türkiye

### Erdal Karadurmuş

Department of Chemical Engineering, Hitit University, Türkiye

### Mehmet Yüceer

Department of Chemical Engineering, İnönü University, Türkiye

### Abstract

Dissolved oxygen (DO) concentration is one of the most critical indicators of water quality, directly affecting aquatic ecosystem health and biological processes. Accurate prediction of DO remains challenging due to the nonlinear, dynamic, and non-stationary nature of aquatic environments. In this study, machine learning-based models were developed to predict dissolved oxygen concentration in the Yeşilirmak River (Türkiye) using real-time sensor data obtained from an online monitoring station. A total of 33,374 observations were analyzed, including water temperature, pH, electrical conductivity, total organic carbon, and nitrate as input variables. Three machine learning algorithms—Support Vector Regression (SVR), Random Forest (RF), and eXtreme Gradient Boosting (XGBoost)—were implemented and evaluated using a hold-out strategy combined with k-fold cross-validation. Model performance was assessed using RMSE, MAPE, and R<sup>2</sup> metrics. Among the tested models, XGBoost demonstrated superior predictive performance, achieving high accuracy and strong generalization capability. To enhance model transparency and interpretability, explainable artificial intelligence (XAI) was integrated using the SHAP method, which revealed water temperature as the dominant factor influencing dissolved oxygen levels, followed by total organic carbon and pH. The results indicate that the proposed XGBoost-based framework provides an accurate and interpretable approach for dissolved oxygen prediction and can support real-time water quality monitoring and management applications.

### Keywords

Dissolved oxygen, Machine learning, XGBoost, Water quality monitoring, Explainable AI.