EGU21-5269
https://doi.org/10.5194/egusphere-egu21-5269
EGU General Assembly 2021
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Gradient Boosting Machine for Phosphorus Removal Prediction in Multi-Soil-Layering (MSL) system operated in a rural area

Sofyan Sbahi[1], Naaila Ouazzani[2], Abderrahmane Lahrouni[3], Abdessamed Hejjaj[4], and Laila Mandi[5]
[1]Cadi Ayyad University, Faculty of Sciences Semlalia, Morocco (sofyansbahi@gmail.com)
[2]Cadi Ayyad University, Faculty of Sciences Semlalia, Morocco (ouazzani@uca.ma)
[3]Cadi Ayyad University, Faculty of Sciences Semlalia, Morocco (lahrouni@uca.ma)
[4]Cadi Ayyad University, Faculty of Sciences Semlalia, Morocco (dashaouz@yahoo.fr)
[5]Cadi Ayyad University, Faculty of Sciences Semlalia, Morocco (mandi@uca.ma)

The quality of effluents from wastewater treatment plants still challenging especially in underprivileged rural areas where water resources are mostly affected by pollution, depletion and excessive exploitation. Thus, the prediction of phosphorus removal is one of the most important tasks in the management of wastewater effluent. Predictive model accuracy is crucial for safe reuse of treated water for public health and the environment. However, linear models that use a high dimensional dataset may be unable to build accurate and interpretable models. To address this complexity, the current study evaluates the effect of hydraulic retention time (HRT) on the removal of orthophosphates (PO_4–P) and total phosphorus (TP) by the multi-soil-layering (MSL) eco-friendly technology. In addition, it attempts to predict this removal from domestic wastewater using a combined approach based on feature selection technique and gradient boosting machine algorithm (GBM). Sixteen physicochemical and bacterial indicators were monitored for a one-year period. The results show that the HRT impact significantly ($p < 0.01$) the removal of phosphorus content by the MSL system. The HRT, pH, PO_4–P and TP were suggested relevant for predicting the removal of TP, while HRT and PO_4–P were sufficient for predicting the removal rate of PO_4–P. The analysis of accuracy using the validation dataset demonstrates that GBM models have high credibility as they achieve an $R^2 > 0.92$, while the analysis of sensitivity reveals that the HRT was the most important factor affecting phosphorus removal in the MSL system. In addition, the modeling results show that the GBM model has proven to be useful for predicting pollutant removal in the MSL technology and investigating its behavior.