EGU23-13470, updated on 17 Jul 2023
https://doi.org/10.5194/egusphere-egu23-13470
EGU General Assembly 2023

Random Forests-Based Surrogate Model as a Tool to Facilitate Groundwater Management

Janire Uribe-Asarta, Vanessa A. Godoy, and J. Jaime Gómez-Hernández

The recovery of aquifers under severe stress requires tools to understand their behaviour to define the most sustainable measures and the collaboration of all the stakeholders involved, water authorities, legislators, and groundwater users. The most used tools for decision-making are numerical models. However, numerical models are usually computationally expensive and due to their complexity are limited to people with advanced technical knowledge. This work proposes to replace the numerical model with a quick and simple-to-use tool accessible to all aquifer's users. For that, we develop a surrogate model based on artificial intelligence methods. This groundwater flow surrogate model allows evaluating the impact of possible changes in pumping extractions or rainfall on piezometric heads in the near future. It has been applied to the Requena-Utiel and Cabrillas-Malacara aquifers, in Spain, considered overexploited by the Júcar River Basin Authority. The surrogate modelling methodology requires a numerical model to create the training data set for the machine learning algorithms. The numerical model is implemented and calibrated using MODFLOW 2005 and FloPy using all available information from 1980 to 2016, with a monthly discretization. The training dataset is obtained by generating 100 MODFLOW realizations with different scenarios of recharge and pumping rate and 145 selected monitoring points from 2016 to 2052. Recharge rate is allowed to decrease up to 60% or increase up to 25% of the average of the last ten years, while, pumping rates are allowed to decrease up to 70% and increase up to 30%. Then, the data is shuffled and 90% is used for training and the remaining 10% is used for testing. Three different algorithms were tried to compare the results: Random Forests, AdaBoost and XGBoost, and random forests were selected as the final surrogate model for its best performance. The surrogate models produce very similar and accurate approximations of the piezometric heads with respect to the data they were trained with and the reduction in computational time is remarkable. The predictions of the surrogate model are interpolated over the study area to obtain piezometric head values maps.

This research was developed under the scope of the InTheMED project, which is part of the PRIMA Program supported by the European Union's Horizon 2020 Research and Innovation Program under Grant Agreement No 1923.