Random forests in water resources

Hristos Tyralis\textsuperscript{1,2}, Georgia Papacharalampous\textsuperscript{3,4}, and Andreas Langousis\textsuperscript{5}

\textsuperscript{1}Hellenic Air Force General Staff, Hellenic Air Force, Athens, Greece (montchrister@gmail.com)
\textsuperscript{2}School of Civil Engineering, National Technical University of Athens, Zografou, Greece (hristos@itia.ntua.gr)
\textsuperscript{3}School of Engineering, University of Patras, Patras, Greece (geopap@upatras.gr)
\textsuperscript{4}Department of Engineering, Roma Tre University, Rome, Italy (papacharalampous.georgia@gmail.com)
\textsuperscript{5}School of Engineering, University of Patras, Patras, Greece (andlag@alum.mit.edu)

Random forests is a supervised machine learning algorithm which has witnessed recently an exponential increase in its implementation in water resources. However, the existing implementations have been restricted in applications of Breiman's (2001) original algorithm to regression and classification models, while numerous developments could be also useful for solving diverse practical problems. Here we popularize random forests for the practicing hydrologist and present alternative random forests based algorithms and related concepts and techniques, which are underappreciated in hydrology. We review random forests applications in water resources and provide guidelines for the full exploitation of the potential of the algorithm and its variants. Relevant implementations of random forests related software in the R programming language are also presented.