



## **Predictive pike (*Esox lucius*) and tench (*Tinca tinca*) population models based on classification trees**

R. Zarkami, P.L.M. Goethals and N. De Pauw

Ghent University, Dept. Applied Ecology, J. Plateaustraat 22, B-9000 Gent, Belgium  
(peter.goethals@UGent.be)

Many activities in Flanders (Belgium) are threatening various fish species in particular pike "*Esox lucius*" and tench "*Tinca tinca*" population. In this study, 307 sites (for tench) and 90 sites (for pike) were monitored over a period of 5 years. Dataset consisted of physical-habitats variables (width, depth, gradient slope and distance from source), physical-chemical (pH, DO, water temperature and conductivity) and biological parameters (individuals' fish and individuals' biomass). The physical-habitats and physical-chemical variables were used as input while the biological ones as output variables to predict the absence and presence of pike and tench populations by means of classification tree models induced via the J48 algorithm. To achieve the best model performance, three-fold cross validations were applied. In order to evaluate the model stability, the dataset was shuffled 5 times, obtaining in total thus 15 different model training and evaluation events. The effect of pruning on the reliability and model complexity was tested in each subset. The performance evaluation was based on a combination of the number of Correctly Classified Instances (CCI) and kappa. The predictive performance for tench models was better than pike. The highest overall average of CCI and kappa for prediction of tench was respectively 74.4 and 0.48 while these values for pike prediction was respectively computed 65.5 and 0.3.