Ecological Modelling 222 (2011) 2202–2212

Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel

Using classification trees to analyze the impact of exotic species on the ecological assessment of polder lakes in Flanders, Belgium

Gert Everaert[a,*], Pieter Boets[a], Koen Lock[a], Sašo Džeroski[b], Peter L.M. Goethals[a]

[a] *Ghent University, Laboratory of Environmental Toxicology and Aquatic Ecology, J. Plateaustraat 22, B-9000 Ghent, Belgium*
[b] *Jozef Stefan Institute, Department Knowledge Technologies, Jamova cesta 39, SI-1000 Ljubljana, Slovenia*

A R T I C L E I N F O

Article history:
Available online 16 September 2010

Keywords:
Brackish water
Classification tree
Ecological assessment
Gammarus tigrinus
Exotic species
Potamopyrgus antipodarum

A B S T R A C T

Polder lakes in Flanders are stagnant waters that were flooded by the sea in the past. Several of these systems are colonized by exotic species, but have hardly been studied until present. The aim of the present study was: (1) to assess the influence of exotic macrobenthic species on the outcome of the Multimetric Macroinvertebrate Index Flanders (MMIF) and (2) to use classification trees for evaluating to what extent physical–chemical characteristics affect the presence of exotic species.

In total, 27 mollusc and 10 macro-crustacean species were present in the monitored lakes of which respectively five and four were exotic. The exclusion of the exotic species from the MMIF resulted in a significant decline of this ecological index (-0.03 ± 0.04; $p = 0.00$). This elimination often resulted into a lower ecological water quality class and more samples were classified into the bad and poor ecological water quality classes.

Single-target classification trees for *Gammarus tigrinus* and *Potamopyrgus antipodarum* were constructed, relating environmental parameters and ecological status (MMIF) to the occurrence of both exotic invasive species. The major advantages of using single-target classification trees are the transparency of the rule sets and the possibility to use relatively small datasets. However, this classification technique only predicts a single-target attribute and the trees of the different species are often hard to integrate and use for water managers. As a solution, a multi-target approach was used in the present study. Exotic molluscs and crustaceans communities were modelled based on environmental parameters and the ecological status (MMIF) using multi-target classification trees. Multi-target classification trees can be used in management planning and investment decisions as they can lead to integrated decisions for the whole set of exotic species and avoid the construction of many models for each individual species. These trees provide general insights concerning the occurrence patterns of individual crustaceans and molluscs in an integrated way.

1. Introduction

The European Water Framework Directive (WFD) forces EU member states to determine the ecological status of water bodies (EU, 2000). The goal of this directive is to ensure that the quality of surface water and groundwater in Europe reaches a good ecological status by the year 2015. The Flemish Environment Agency (VMM) used the Belgian Biotic Index (BBI) for more than two decades to monitor the ecological water quality of Flemish rivers (De Pauw and Vanhooren, 1983). Despite the reliability and robustness of this index, a number of technical shortcomings arose about the potential application of the BBI for WFD implementation, in particular the index was not useful for stagnant waters (Gabriels et al., 2010). Therefore, a new type-specific multimetric index, combining the robustness of the BBI with the versatility of multimetric indices, was developed. This new index, calculated from the taxonomic composition and abundance of the macroinvertebrates, is called the Multimetric Macroinvertebrate Index Flanders (MMIF) and is currently used to assess the ecological water quality in Flanders (Gabriels et al., 2010). The metrics comprised in the MMIF are taxa richness, number of Ephemeroptera, Plecoptera and Trichoptera (EPT), number of other (i.e. non-EPT) sensitive taxa, the Shannon-Wiener diversity index and the mean tolerance score. For each type of river and lake, a set of reference values for all five metrics was determined (Gabriels et al., 2010). Based on the references, a scoring system was developed for each metric consisting of threshold values needed for assigning a score ranging from zero to four (four being nearest to the reference conditions). To obtain the final index, ranging from zero for a very poor ecological quality to one for a very good ecological quality, the five metric scores are summed and subsequently divided by 20. The range of the MMIF index can be considered as an ecological quality ratio (EQR) because the max-

* Corresponding author. Tel.: +32 092643776; fax: +32 092644199.
E-mail address: gert.everaert@ugent.be (G. Everaert).

0304-3800/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.ecolmodel.2010.08.013

imum MMIF value of 1 can only be obtained when all metric values are near the type-specific reference value for that metric (Gabriels et al., 2010). In order to meet the target of the WFD in 2015, aquatic systems should have a MMIF-score of 0.6 or 0.7, depending on the water type (Gabriels et al., 2010).

Macroinvertebrates are identified up to genus or family level for the calculation of the MMIF. Because of this coarse taxonomic identification level, shifts in species composition between native and exotic species often remain hidden (Gabriels et al., 2005). The introduction of exotic species might decrease the alpha diversity, which can be masked due the identification level (Gabriels et al., 2005). For example, the invader *Dikerogammarus villosus* might outcompete native gammarids (Bij de Vaate et al., 2002; Boets et al., 2010), but this will not influence the results of the index calculation at family level of a given sample, since Gammaridae are still present and tolerance classes are defined at family or genus level. Additionally, the inclusion of the exotic invasive species, such as *Corbicula*, can lead to an increase of the ecological water quality index depending on the tolerance class assigned to the invader. Previous examples suggest that the use of a standard list of taxa, where tolerance classes are assigned at specific taxonomic levels (e.g. genus or family level), can result in altered assessment scores if exotic species are present (Gabriels et al., 2005). Therefore, it is necessary to examine the influence of exotic species on the ecological assessment of aquatic ecosystems.

Freshwater exotic invasive species are an issue of growing management concern (Vander Zanden and Olden, 2008). Invasive species have one of the most harmful and least reversible impacts on natural ecosystems as they may change the local fauna and flora all around the world (Vitousek et al., 1996; Ricciardi and MacIsaac, 2000). Exotic invasive species may decrease the ecological quality through changes in biological, chemical and physical properties of aquatic ecosystems (Olenin et al., 2007). These changes include: elimination of sensitive or rare species; alteration of native communities; algal blooms; modification of substrate conditions and the shore zones; alterations of oxygen and nutrient content, pH and transparency of the water; accumulation of synthetic pollutants, etc. (Olenin et al., 2007). For instance, Boets et al. (2009) indicated that the exotic macro-crustacean *Procambarus clarkii* predates on native benthic macroinvertebrates, spreads diseases and affects the physical habitat via burrowing activities. Our research focussed on exotic molluscs and macro-crustaceans, because these have probably the highest impact among all aquatic freshwater invaders in Europe (Orendt et al., 2010).

Stimulated by the expansion of the global transport of goods and people, the numbers and costs of exotic species are rising at an alarming rate (Lovell and Stone, 2006). Exotic species may be unintentionally imported by ships discharging their ballast water (Mills et al., 1993; Lovell and Stone, 2006; Colautti et al., 2006). Leung et al. (2006) found that recreational boaters between lakes are an important pathway of overland dispersal of exotic species. Pathogens and parasites have been introduced unintentionally into the USA via infected stock for aquaculture farms (Naylor et al., 2001). Policy makers spend a lot of money trying to control or remove invaders from our environment (Pimentel et al., 2000; Pimentel et al., 2005). Many USA states have recently created exotic species advisory councils that bring together regulators, researchers and other stakeholders to address research, policy and management needs (Lodge et al., 2006). However, managers lack predictive tools to help them prioritise invasion threats and to help them decide where they should allocate the limited resources for prevention and mitigation most effectively (Ricciardi, 2003).

One of the methods applied by managers in the USA is the national Gap Analysis Program (GAP). This method identifies 'gaps' in the network of conservation land and water areas (Scott et al., 1993). The framework documents biogeographic information and organizational cooperation in ways meaningful to their management and can therefore be useful in the context of exotic species (Jennings, 2000).

Other methods, such as data mining techniques, can be helpful because they allow accurate predictions of species preferences and impacts. Classification trees can give insight in complex, unbalanced, non-linear ecological data where commonly used exploratory and statistical modelling techniques often fail to find meaningful ecological patterns (De'ath and Fabricius, 2000). Classification trees have been applied in numerous ecological studies (Dakou et al., 2007; Boets et al., 2010) and have proven to have a high potential in macroinvertebrate habitat suitability analysis as they combine reliable classifications with a transparent set of rules (Hoang et al., 2009).

Classification trees are decision trees that predict the value of a discrete-valued (nominal) target variable (Breiman et al., 1984). Decision trees are hierarchical structures, where the internal nodes contain tests on the input attributes. Each branch of an internal test corresponds to an outcome of the test and the prediction for the value of the target attribute is stored in a leaf. Each leaf of a decision tree contains a prediction for the target variable. A single-target approach learns a model for each target attribute separately, whereas a multi-target approach builds one model for all target attributes simultaneously (Kocev et al., 2009). Therefore, a single-target approach can be used to predict the possible occurrence of individual exotic species based on physical–chemical parameters, while possible changes in species composition can be highlighted using a multi-target approach.

Polder lakes, situated in the north of Flanders, are brackish, stagnant waters situated on the inland side of the dikes (Delaunois, 1982). They find their origin in history when, due to flooding by the sea, dikes gave way and land was washed out. The salinity of these lakes, determining the fauna and flora, depends on their age and the possible influence of seepage water (van Puijenbroek et al., 2004). The salinity of polder lakes in Flanders decreases from north to south and from west to east (Delaunois, 1982). Many of these shallow lakes are hypertrophic and dominated by algal blooms. The eutrophication process became problematic in the 1950s due to run-off from agriculture and discharges from industry and untreated household waste (van Puijenbroek et al., 2004). In the mid-1980s, projects on lake restoration were started in the Netherlands (Van Huet, 1992). In Flanders, apart from a study by Dumont and Gysels (1971), little research has been carried out on these aquatic systems.

The aims of the present paper were (1) to assess the influence of exotic species on the MMIF and (2) to construct single- and multi-target classification trees to predict the presence of exotic species based on physical–chemical parameters and occurrence of other species.

2. Materials and methods

2.1. Study area and data collection

The dataset contained 108 samples comprising biological and physical–chemical information of 45 polder lakes (Fig. 1). The polder lakes, all located in the Northern part of Flanders, can be divided in three clusters. The first, most westerly oriented cluster of lakes, is situated close to the city of Ostend, whereas the second, most easterly oriented cluster, is located close to the river Scheldt nearby the city of Antwerp. The remaining polder lakes, situated between the first two clusters, are distributed along the Dutch border. Most polder lakes are exploited for recreational purposes: the smaller polder lakes are frequently used for fishing, whereas the bigger lakes are suitable for sailing and windsurfing.