Aquatic Botany 120 (2015) 142–149

Contents lists available at ScienceDirect

Aquatic Botany

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

Distribution and predictive occurrence model of charophytes in Estonian waters

Kaire Torn[a,*], Anastasiia Kovtun-Kante[a], Kristjan Herkül[a], Georg Martin[a], Helle Mäemets[b]

[a] Estonian Marine Institute, University of Tartu, Mäealuse 14, Tallinn 12618, Estonia
[b] Centre for Limnology, Rannu, Tartumaa 61117, Estonia

ARTICLE INFO

Article history:
Received 1 August 2013
Received in revised form 21 April 2014
Accepted 1 May 2014
Available online 15 May 2014

Keywords:
Charophytes
Distribution
Baltic Sea
Lakes
Habitat modelling

ABSTRACT

Material collected during the years 1995–2011 was used to describe the distribution and environmental preferences of charophyte species in Estonian lakes and its coastal Baltic Sea. Altogether 22 species of charophytes were found in Estonian waters. Five taxa occurred in less than 10 localities and were classified as rare. *Chara aspera* and *Tolypella nidifica* were the most frequent and widespread species. The majority of species preferred shallow water less than 1 m in Estonian lakes and the coastal sea. Mud was the prevailing substrate on locations where charophytes were found, sandy substrate was characteristic for species which tolerate more exposed localities. Most of freshwater species preferred water alkalinity over 80 mg HCO_3^- l^{-1}. A model was developed to predict the probability of the occurrence of *Chara* spp. in the extent of the whole Estonian marine waters based on several environmental variables. Boosted regression trees (BRT) was chosen as the modelling technique. Based on the model prediction, the vast majority of charophyte habitats are situated in the sea areas of the West Estonian Archipelago. That sea area is characterized by favourable conditions for charophytes: high proportion of shallow areas protected from wave exposure.

1. Introduction

Charophytes occupy several ecological niches in aquatic ecosystems. They may inhabit the deepest areas of clear-water lakes but also form shallow-water pioneer vegetation in recently formed ponds and wetlands (Chambers and Kalff, 1985; Casanova and Brock, 1999). Charophyte communities are an important habitat for a number of invertebrate species and provide feeding and nursery areas for several species of fish and birds (e.g. Schubert and Blindow, 2003; Torn, 2008).

Human impact and consequent environmental changes has caused a progressive decrease in the abundance, occurrence and diversity of charophyte species in past decades (Romanov, 2009). Some became rare and several species of charophytes are Red Listed in Europe (Blindow et al., 2003). Charophytes are among the species listed in Annex I of the EU Habitat Directive as characteristic species of the habitat type no. 1150 "Coastal lagoons" and are used as indicators in procedures of assessment of coastal water quality in many countries (e.g. Germany, Sweden) (European Commission, 2007; Steinhardt et al., 2009). Among inland waters charophyte lakes are distinguished as an EU Habitat Directive Annex I habitat type no. 3140 "Hard oligo-mesotrophic waters with benthic vegetation of *Chara* spp.".

Studies on the distribution and ecological demands of charophytes in several countries display large disproportions in time and space. The species richness is commonly directly related to the field sampling effort and the activity of aquatic botanists. Despite the fact that the Estonian coastal sea is well-studied and data on charophytes in this area are constantly being updated (Torn et al., 2004; Kovtun et al., 2011), published information about charophyte distribution in the inland waters is old (Pork, 1954). An important shortcoming is the absence of a common charophyte database for both coastal sea and inland waters. The lack of a common database has (1) hindered development of a holistic understanding of the distribution and ecology of charophytes as several species are present in both inland and marine waters, and (2) caused misinformation: e.g. in some publications only data on brackish water species have been used or new data have been combined with 60-year old records (Urbaniak, 2007; Romanov, 2009). Therefore one of the aims of this paper is to give a review of the distribution and

* Corresponding author. Tel.: +372 671 8940; fax: +372 6718900.
E-mail address: kaire.torn@ut.ee (K. Torn).

http://dx.doi.org/10.1016/j.aquabot.2014.05.005
0304-3770/

environmental preferences of charophyte species in Estonian brackish and fresh waters.

Greater sampling effort can certainly improve our knowledge of the distribution of charophytes and identify threatened species. However, traditional sampling-point field work is not suitable for covering large areas in high detail as it yields data only from visited sampling sites and leaves most of the study area unsampled. Moreover, extensive in situ field work is very time-consuming and expensive. Predictive modelling enables a general assessment of the distribution of species in large spatial extents that cannot be fully covered with in situ sampling (Zimmermann et al., 2010). A seamless map of the probability of occurrence gives a significantly more relevant view of the distribution of a species than simple plotting of field localities on a map. This is especially so, when considering that sites of field sampling are commonly spatially unequally distributed over extensive areas. Additionally, predictive modelling provides an opportunity to examine the effects of environmental variables on the distribution of a species at various spatial scales and help to determine appropriate management actions (Kumar et al., 2009). Accordingly, the second aim of this paper was to predict the potential distribution for charophytes in coastal waters based on available georeferenced environmental data (depth, wave exposure etc.).

2. Material and methods

2.1. Data collection

The material for the present study was collected during 1995–2011 and is based on databases of the Estonian Marine Institute (University of Tartu) and Centre of Limnology (Estonian Agricultural University) (Fig. 1). Sampling in brackish water (salinity over 0.5 psu) has been predominantly performed by SCUBA diving from a boat or directly from the shore. For each locality, GPS position, depth, sediment type and abiotic water column properties (e.g. salinity, oxygen content, Secchi depth) were recorded. Sampling in fresh water (salinity below 0.5 psu) was performed by dredging with a hook from a boat or directly from a shore. The type of water body (lake, pond, ditch), GPS position, depth and sediment type were fixed for each site. Six types of sediment (mud, sand, clay, gravel, peaty mud and clayey mud) was distinguished based on content, consistency, grain size and/or colour of the soil. Mud was defined as the remains of biota and inorganic particles, peaty mud mainly consists detritus of *Sphagnum* spp. Water alkalinity (HCO_3^- mg l^{-1}) and dichromate oxygen consumption (COD_{Cr} mg O l^{-1}) which reflects the organic content were used for the characterization of freshwater locations. Samples for chemical analyses were collected from the surface layer of water column in midsummer. Alkalinity was titrated with HCl, dichromate oxygen consumption determined by the oxidation of organic matter by a solution of $K_2Cr_2O_7$ in H_2SO_4. Collected charophyte samples were packed, labelled and frozen or preserved in formaldehyde solution until determination in the laboratory.

For species identification, the determination keys of Krause (1997), Schubert and Blindow (2003) and Langangen (2007) were used. Sterile specimens of *Nitella flexilis* (Linnaeus) C. Agardh could not be distinguished from *Nitella opaca* (Bruzelius) C. Agardh, therefore these species were treated as a group of *N. opaca/flexilis.*

2.2. Distribution modelling

We aimed to build a model that best predicts the spatial distribution of genus *Chara* in the Estonian coastal waters. Boosted regression trees (BRT) was chosen as the modelling technique as its predictive performance has been shown to be superior to most other modelling methods (Elith et al., 2006; Revermann et al., 2012). BRT is an ensemble method that combines the strength of two algorithms: regression trees and boosting (Elith et al., 2008). Regression trees are good at selecting relevant predictor variables and can model interactions. Boosting enables a building of a large number of trees in a way that each successive tree adds small modifications to parts of the model space to fit the data better (Friedman et al., 2000). BRT has no need for prior data transformation or elimination of outliers, can fit complex nonlinear relationships, can handle different types of predictor variables, and can model interaction effects among predictors (Elith et al., 2006). Important parameters in building BRT models are learning rate and tree complexity. Learning rate determines the contribution of each tree to the growing model and tree complexity defines the depth of interactions allowed in a model. The BRT modelling was performed in the statistical software R version 2.15.1 (R Development Core Team, 2012) using packages 'gbm' (Ridgeway, 2012) and 'dismo' (Hijmans et al., 2012).

The predictor variables included different bathymetrical (depth, slope of seabed), hydrodynamic (wave exposure, current speed), geological (seabed substrate), and physico-chemical (temperature, salinity, oxygen content) variables. Altogether 26 abiotic predictor variables were used (Table 1) that were all available as georeferenced raster layers. Input data for the dependent variable, i.e. the sampling point-wise presence-absence data of *Chara* spp., were compiled from the benthos database of the Estonian Marine Institute. The input dataset on charophytes included 11 149 sampling sites distributed over the Estonian marine area from the period 1995–2011 (Fig. 1). *Chara* spp. were present in 1146 sites corresponding to 10.3% of the total number of sampling sites. *Tolypella nidifica* (O.F. Müller) Leonhardi was excluded because of somewhat different environmental preferences (e.g. wider depth distribution, salinity tolerance) compared to genus *Chara* species. Due to the lack of good environmental data from freshwater, the spatial prediction of the occurrence of charophytes was made only for the coastal sea.

Two groups of BRT models were built that had tree complexity of 1 and 5, respectively. Tree complexity of 1 fits an additive model without interactions between predictors while tree complexity of 5 fits a model with up to five-way interactions. In both groups, models with learning rates of 0.005, 0.01, 0.05 and 0.1 were built and their predictive performance was estimated by calculating predictive deviance and Area Under the Receiver Operating Curve (AUROC, generally abbreviated to AUC) (Fielding and Bell, 1997) using 10-fold cross validation. An AUC value of 0.5 indicates that the model prediction is not better than random while the value of 1 shows a perfect match between the model prediction and real value (Fielding and Bell, 1997). The model with the highest cross-validation AUC value was chosen and it was further subjected to simplification as implemented in the package 'dismo': the routine performs a backwards elimination of variables to drop those that give no evidence of improving predictive performance (Hijmans et al., 2012). After simplification, the model was used for making the spatial prediction of the probability of occurrence of *Chara* spp. in the Estonian sea area. The prediction was modelled over a 200×200 m grid covering water depths of 0 to 15 m.

3. Results

3.1. Distribution of charophytes

Charophytes were found from 1365 locations in coastal area and from 176 lakes or ponds. Altogether 22 species of charophytes were found in Estonian waters (Fig. 2). In brackish waters, seven species of stoneworts were found, representing the genera *Chara*