Original Research Article

Key management rules for agricultural alpine newt breeding ponds based on habitat suitability models

Niels De Troyer [a], *, Marie Anne Eurie Forio [a], Karen Roels [a], Luc De Meester [b, c, d], Pieter Lemmens [b], Steven A.J. Declerck [b, e], Koen Martens [f], Peter Goethals [a]

[a] Department of Animal Science and Aquatic Ecology, Ghent University, Coupure Links 653, B-9000, Ghent, Belgium
[b] Laboratory of Aquatic Ecology, Evolution and Conservation, KU Leuven, Charles Deberiotstraat 32, B-3000, Leuven, Belgium
[c] Leibniz Institut für Gewässerökologie und Binnenfischerei (IGB), Müggelseedamm 310, 12587, Berlin, Germany
[d] Institute of Biology, Freie Universität Berlin, Königin-Luise-Straße 1-3, 14195, Berlin, Germany
[e] Department of Aquatic Ecology, Netherlands Institute of Ecology (NIOO-KNAW), P.O. Box 50, 6700 12 AB, Wageningen, the Netherlands
[f] Royal Belgian Institute of Natural Sciences, OD Nature, Freshwater Biology, Vautierstraat 29, B-1000 Brussels, Belgium and Ghent University, Department of Biology, K.L. Ledeganckstraat 35, B-9000, Ghent, Belgium

A R T I C L E   I N F O

A B S T R A C T

The alpine newt, *Ichthyosaura alpestris*, is very sensitive to habitat destruction and alteration which has led to declining populations across Europe. As this species is protected through the Bern Convention, it is essential to have a comprehensive understanding of its habitat requirements to ensure proper conservation measures. We trained, validated and optimized classification tree models based on data on local aquatic habitat conditions from 125 farmland ponds scattered over Belgium and Luxembourg where the alpine newt commonly occurs. To obtain user-friendly and representative models, data was pre-processed and stratified after which different degrees of pruning were applied for model optimization. In order to check the model's applicability for management, we predicted alpine newt occurrence with an independent dataset. The most robust and reliable model revealed that fish absence was the major driving factor followed by the thickness of the sludge layer. We found that fish presence established alpine newt absence and that fishless ponds with a sludge layer of 15 cm or more were predicted to host no alpine newts. The latter provides quantitative information for decision makers. Moreover, our results indicated that the amount of sludge could be associated with eutrophication and erosion. Regarding management practices, it is advised to assure the absence of fish and reduce sludge accumulation in ponds designated for the conservation of alpine newts, for example by designing temporary ponds not fed through fish-containing surface waters. Furthermore, we recommend to install buffer strips around a pond to reduce nutrient and soil run-off from the terrestrial environment.

* Corresponding author.
E-mail address: Niels.DeTroyer@UGent.be (N. De Troyer).

1. Introduction

Across Europe, salamander or urodelan populations are facing dramatic declines and extinctions. Almost 30% of all European urodelan species are categorized as vulnerable, endangered or critically endangered which is the second highest of all European vertebrates (cf. 37% for freshwater fishes, 19.4% for reptiles, 16.5% for mammals, 12.6% for birds and 7.5% for marine fishes) (BirdLife International, 2015; Cox and Temple, 2009; Freyhof and Brooks, 2011; Nieto et al., 2015; Temple and Cox, 2009; Temple and Terry, 2007). Urodeles are exposed to several threats. Human induced habitat loss and degradation are the most prominent stressors at present (Hof et al., 2011). Many of the European salamanders require stagnant water bodies, such as ponds, for their reproduction (Temple and Cox, 2009). Agricultural intensification is a major cause for pond loss and degradation due to eutrophication or pesticide pollution (Curado et al., 2011; Temple and Cox, 2009). Furthermore, the destruction of ponds and other small landscape elements, such as shrubs, hedges and rough vegetation enhances spatial homogenisation and fragmentation of a landscape thereby reducing connectivity amongst habitats (Hehl-Lange, 2001). Next to habitat alteration, urodelans are threatened by the fungal disease caused by the species *Batrachochytrium dendrobatidis* and *Batrachochytrium salamandrivorans*. The severity of the disease depends on the affected species. For example, fire salamanders, *Salamandra salamandra*, are lethally susceptible to *B. salamandrivorans* independent of the dose while alpine newts, *Ichthyosaura alpestris*, show a dose-dependent *B. salamandrivorans* disease course making them moderately susceptible (Martel et al., 2013). Due to their dispersal abilities and their frequent co-occurrence with the highly susceptible fire salamander, alpine newts are seen as possible vectors of chytridiomycosis (Stegen et al., 2017). They are considered to be the most consistently infected species in the wild for *B. dendrobatidis*, thus further enhancing their disease carrying and transmission potential (Spitzen-Van Der Sluijs et al., 2014).

The alpine newt is currently listed as a "least concern" species in many European countries, including Belgium, the Netherlands and Luxembourg (the Benelux union), owing to its wide distribution, low habitat requirements and presumed large populations (Arntzen et al., 2009; Jooris et al., 2012). The species and its habitat are lawfully protected by the Bern Convention, a European treaty aiming at conserving Europe's natural heritage. Nonetheless, a continuously decreasing trend is observed for European populations (Arntzen et al., 2009). Habitat destruction and fish introduction are the most prominent factors responsible for the population declines (Arntzen et al., 2009). The alpine newt requires both suitable terrestrial and suitable aquatic habitats. Generally, they are absent from large water bodies and rivers, which often contain fish. Most other water bodies, such as (garden) ponds, temporary pools, ditches, fens, (concrete) cattle drinking basins and even ruts can be occupied for reproduction. On land, they are present in a wide variety of habitats ranging from forests over pastures and gardens to heavily disturbed lands where they use rocks, wood and trash as hiding places during hibernation. They seem to avoid large cultivated agricultural areas and prefer deciduous forested zones owing to the presence of multiple hiding places (Arntzen et al., 2009; van Delft, 2009). Even though knowledge is available about the habitat this species inhabits, quantitative information, useful for development of effective management strategies of natural environments, is lacking.

Decision tree models have widely been used for quantitative habitat suitability analyses owing to their ease of interpretation and reliability (Boets et al., 2010; Everaert et al., 2011; Hoang et al., 2010). The classification tree is a purely data-driven ecological model describing the non-linear statistical relationship between a categorical response variable and numerical predictors which are split according to certain threshold values (Van Echelpoel et al., 2015). In this way, the occurrence of a certain species can be linked to environmental conditions. This modelling technique thus allows to non-parametrically extract the most essential information in the ecological multidimensional data cloud that determine a species' occurrence. Here, we aim to: (1) develop a classification tree based habitat suitability model relating local pond conditions (predictor variables) to alpine newt presence/absence (response variable); (2) evaluate the obtained model using a combination of technical criteria, expert knowledge and validation with independent field data; and (3) provide information for decision makers on the practical implications of our results for management.

2. Materials and methods

2.1. Study area

A habitat suitability decision tree model was trained, optimized and validated using data on a set of 125 farmland ponds in Belgium and Luxembourg which were sampled between 1 July and August 7, 2008 (Lemmens et al., 2018). The ponds were distributed over five biogeographical regions (i.e. Gutland, Chalk region, Sand region, Polders and Loam region). Each region had five clusters (within 38 km^2) of five randomly selected ponds. In this way, ponds within each cluster were expected to share a similar regional species pool.

The alpine newt is generally omnipresent across Belgium. Therefore, we wanted to independently test the model's applicability for a similar area outside the regions from 2008. Additional field data were collected from 18 ponds in the province of East Flanders between 8 and May 13, 2017 (Fig. 1). The area was comparable to the ones used for the model development in terms of land use, aquatic characteristics and the occurrence of alpine newts. Four of the selected ponds were located in rural urban areas and nine ponds were situated in nature reserves. These natural areas were characterised by wet meadows, brushwood and swampy forests. Furthermore, we selected five ponds in a more agricultural environment (i.e. arable land and pasture). Information on newt presence was obtained from local stakeholders (landowners and nature
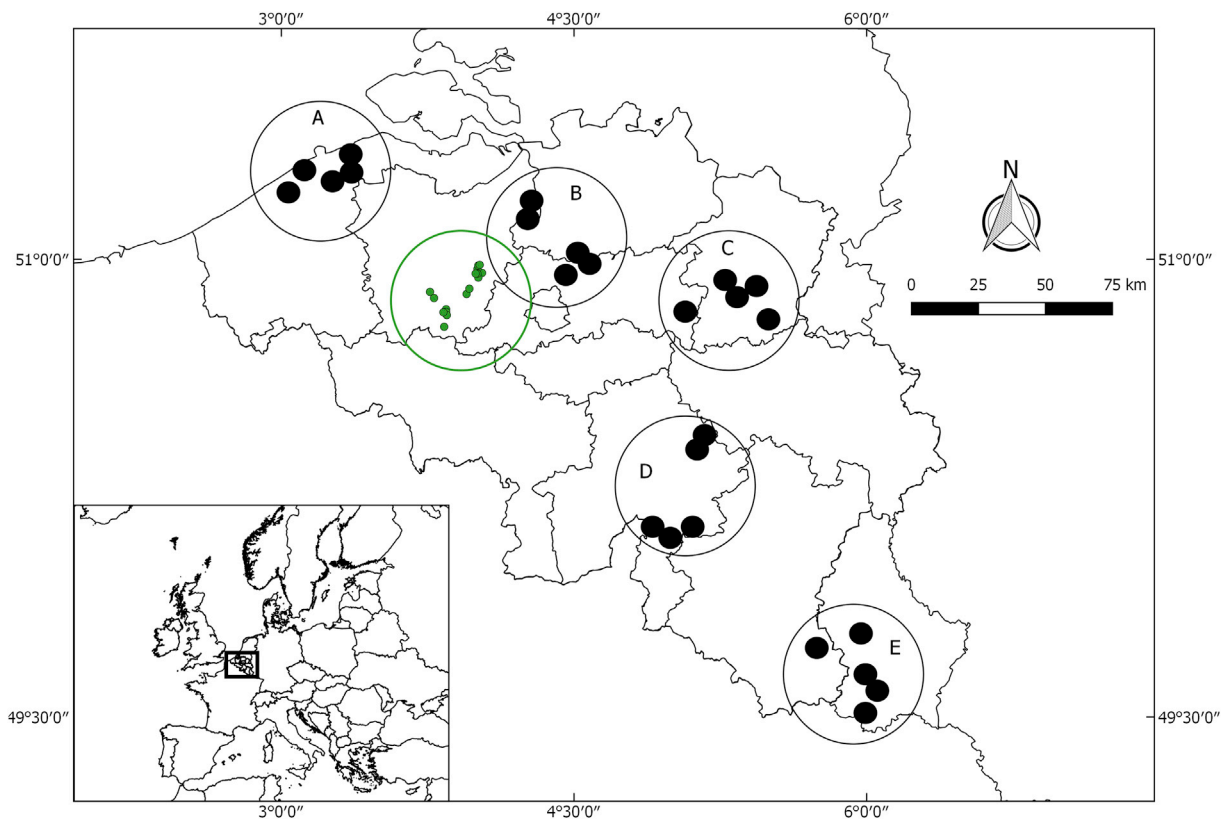
Fig. 1. Map of the study area with the sampling regions Polders (A), Sand region (B), Loam region (C), Chalk region (D) and Gutland (E) each containing five clusters represented by the black dots. In every cluster five ponds were sampled. The sites which were additionally sampled in 2017 are indicated in green. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

conservation organisations). We aimed to include ponds both with and without alpine newts in order to test the model's prediction power for both situations. Finally, we measured the same aquatic variables as during 2008 to verify the similarity.

2.2. Field data collection and sample analyses

Decision tree model training data consisted of morphometric, physical-chemical, and biological data. These were collected once for each of the 125 farmland ponds during the summer of 2008 based on the methodology of Declerck et al. (2006). A detailed description can be found in Appendix A. Pond sampling started with visually gathering information about pond characteristics. The model excluded data from land use and structural pond connectivity (i.e. number of nearby ponds) thus focusing solely on the suitability of the individual aquatic habitat for alpine newt. From each pond, a water sample was taken to quantify pH, water transparency, electrical conductivity (EC), dissolved oxygen (DO), temperature (T), suspended solids (SS), chlorophyll *a* (Chl_a), total nitrogen (TN), total phosphorous (TP), alkalinity, hardness and ions of calcium (Ca^{2+}), sulphate (SO_4^{2-}) and chloride (Cl^-). For assessing the fish community, the Point Abundance Sampling by electrofishing (PASE) was applied as this method is suitable for catching small juvenile fish (body size > 6−7 mm) in standing water bodies (Garner, 1996; Perrow et al., 1996). The amount of sampling locations varied according to the pond size. Per are (1 dam^2), 6 to 8 anode immersions were done randomly. Per are (1 dam^2), 6 to 8 anode immersions were done randomly. Fish were collected with a hand net, identified, counted and returned to the pond. The presence of amphibian species was determined through direct visual observation and capture during sweep net sampling (25 cm × 25 cm; mesh size: 250 μm) in the open water areas and vegetated zones. The total sweep net sampling time varied with pond size and the time spent in each zone was adapted according to their respective percentage. Additionally, amphibians might have been accidentally caught whilst electrofishing. All caught amphibians were identified, counted and released back to their natural environment.

For the model testing (the additional 18 samples) we measured the same morphometric and physical-chemical variables as in 2008 according to the procedure described in Appendix A. The presence of fish and amphibians was determined using floating fikes (50 cm × 30 cm x 25 cm) which were placed in the ponds during one night. The number of fikes depended on the pond size and varied from two to eight aiming at a catching effort of 3−4 fikes per are. They were equidistantly placed close to the shores (Figure A1). All species were identified and released back to the pond. This sampling procedure differed

from the one from 2008 as the purposes of both studies were not the same. For the present research, solely amphibian and fish occurrences were required for which both methods provided accurate information, despite their different sampling strategies. Indeed, the sampling of newts in 2008 took place in summer thereby aiming at catching larvae instead of adults for which sweep net sampling is most effective (Verbelen and Jooris, 2009). On the other hand, newt sampling in May 2017 targeted adults given their considerable occurrence in ponds during that time. Fikes are more efficient for catching adult newts (Verbelen and Jooris, 2009).

2.3. Data pre-processing

In order to quantitatively investigate the relationship between the occurrence of alpine newt in pond ecosystems and local pond conditions, classification tree models were built. The modelling procedure consisted of three parts (Fig. 2). As this type of black-box model is purely data-driven, its quality is highly dependent on the data. We therefore checked all data for missing values, any type of error and skewed distributions prior to model development and selection (Fig. 2 − Step 1).

The raw data was converted into an optimized modelling dataset in three steps based on Van Echelpoel et al. (2015). We conducted a *Pearson correlation analysis* for numerical data aiming at finding strongly correlated variables characterised by a correlation coefficient ρ of at least 0.7 (Dancey and Reidy, 2004). This analysis was performed with the software R, version 3.3.1, using the packages Hmisc and corrplot (Harrell, 2017; R Core Team, 2016; Wei and Simko, 2017). Oxygen and temperature were excluded from the dataset as the measurements can strongly fluctuate during the day and were therefore considered not representative (Andersen et al., 2017; Whitney, 1942). Similarly, we omitted data regarding presence of water birds and cattle due to the strong time dependence of these variables. In case of collinearity between two predictor variables, we decided to retain only one for model development. This decrease in dataset dimension is favourable as it results in a lower computation time and reduces model complexity. Associations among numerical environmental variables were also investigated via a Principal Component Analysis (PCA) in R with the FactoMineR package (Le et al., 2008). *Missing values* were beforehand statistically imputed via the expectation-maximization algorithm (EM-PCA) after which a biplot was constructed from the completed dataset using the R package missMDA (Josse and Husson, 2016). Furthermore, we applied multiple regression analyses with the MASS package in R to check the contribution of several metric predictors to a certain metric dependent variable (Venables and Ripley, 2002). A backwards step-wise elimination procedure, assessed via the Akaike's Information Criterion (AIC), was used to retain the most significantly contributing variables to the regression model. In case the model's residuals were not normally distributed, the normalizing Yeo-Johnson power transformation was performed and the model was recalibrated. *Outliers* were considered to be an inherent and valuable part of the data, as argued by Orr et al. (1991). Decision trees in general are quite robust against skewed distributions due to outliers, especially models built with the C4.5 algorithm, which was used for model development in the present research (John, 1995).

The classification tree model has to be able to predict both presence and absence of alpine newts with an equal accuracy. Our data however contained less presence records than absences. To achieve a balanced design, we *stratified the dataset* by removing absence records thereby making the urodelan species distribution uniform. The elimination procedure was done geographically so that every region (Fig. 1: A − E) occurred in the subset. Per region, a same fraction of ponds was randomly selected so that the relative record abundances amongst regions between the original dataset and the subset were
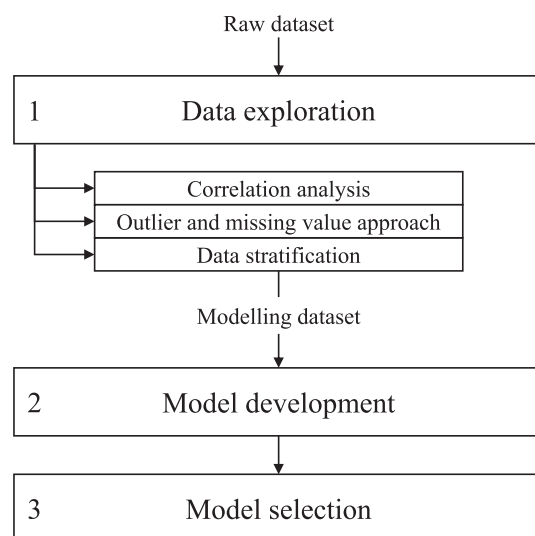
Fig. 2. Overall modelling procedure for classification tree construction. Data pre-processing via correlation, outlier, missing value and stratification analyses (step 1) preceded model development (step 2) and model selection (step 3).

approximately the same. This fraction was equal to the proportion of the total amount of presence records to the total absences. A major drawback of this procedure is the loss of data. The manipulation was repeated ten times and optimal classification tree models were developed for each data subset to verify the model robustness against data manipulation.

2.4. Model development and selection

Decision trees are suitable models when it comes to small datasets, which is the case in the present research (Everaert et al., 2011). The widely used C4.5 machine learning algorithm for tree induction selects the most appropriate attributes for splitting the data based on information gain criteria (Witten and Eibe, 2005). In the first step of the procedure, the algorithm splits the data for each attribute, calculates the information gain and chooses the one that gains the most information to split on. The algorithm then continues recursively until the data can't be split any further. After data pre-processing, model building (Fig. 2 — Step 2) was done with the software Waikato Environment for Knowledge Analysis (Weka, version 3.8.3) for all ten stratified datasets using the J48 algorithm, which is the Java implementation of the C4.5 algorithm (Witten and Eibe, 2005). The software accounts for missing values by classifying them with weights proportional to frequencies of the observed non-missing values. As model parameterization can considerably influence the model outcome and thus the applicability towards end-users, we tested different settings related to the size and growth of the classification tree (Everaert et al., 2016). Data overfitting during model training often leads to large and complex trees at the expense of predictive power. Pruning is a means to prevent overfitting by removing knowledge rules which contribute little and thus limiting tree size (Fürnkranz, 1997; Witten and Eibe, 2005). The degree of pruning was verified via the pruning confidence factor PCF ('confidenceFactor' in Weka) and the minimum number of instances per leaf ('MinNumObj' in Weka). PCF is a post-pruning parameter as it compares model reliability of a classification tree with its subtrees after replacing branches (subtree) by leaves (terminal nodes), thereby affecting the error estimates in each node. More specifically, increasing PCF values decreases the difference between the error estimate of a parent node and its splits. We therefore tested PCF values of 0.05, 0.10, 0.15, 0.25 and 0.35 for each of the ten datasets during model training. Pre-pruning was also applied by varying the minimum amount of instances a leaf should contain before it is split (test settings of MinNumObj: 2, 5, 7, 10, 12). Higher values for MinNumObj reduce tree expansion as this criterion hinders the splitting procedure. These pruning settings covered the overall range of model outcome and associated performances as further increasing values of PCF and MinNumObj did not alter the output anymore.

Regarding model validation, we applied a ten-fold cross-validation, as recommended by Han et al. (2011). Parameter testing led to the construction of 25 classification tree models per stratified set, i.e. the sum of all PCF combinations per MinNumObj element without repetition. For each array of models, we identified the most optimal parameter set by assessing the classification trees mainly on the models' technical performance. The percentage of correctly classified instances (%CCI) and Cohen's kappa (κ) served as criteria for testing the models' fit. The %CCI corresponds to the percentage of ponds which were classified correctly as a pond with or without alpine newts, i.e. true positive and true negative predictions, respectively. κ accounts for all correctly predicted presences and absences adjusted to the amount of agreement expected by chance (Cohen, 1960; Hoang et al., 2010; Manel et al., 2001). Models with %CCI higher than or equal to 70% and κ exceeding 0.4 were considered to be sufficiently reliable (Goethals et al., 2007). In addition to technical criteria, model selection (Fig. 2 — Step 3) also depended on: (1) the complexity of the model, determined via the number of nodes (less complex trees are more user-friendly); (2) model robustness (the frequency of recurrence of a specific model over all models and stratified subsets); and (3) ecological relevance. For the latter, expert-knowledge was used to check if the observed patterns were plausible from an ecological point of view. This selection assessment eventually led to the proposal of one model which was tested with the independent data as collected in 2017. The if/then rules were implemented in Microsoft Excel, after which %CCI and κ were calculated and evaluated via a confusion matrix which visualizes the model predictions against the observations for both alpine newt presence and absence.

3. Results

3.1. Data optimization

The initial dataset consisted of 36 predictor variables, of which 7 were categorical (nominal/ordinal) and 29 were numerical (Table A1). The dimension of the dataset was reduced by conducting multiple one on one correlation analyses and by selecting one of each pair of variables which showed significant correlations (Figure A2 & Table A2). The percentages of shade, overhanging trees and trees on the margin of the pond were found to be highly correlated. We therefore decided to only include the percentage of shade as a proxy for the light entering the pond in subsequent analyses. Similarly, we decided to retain surface area of the pond above other dimension variables (i.e. length, width, volume) due to their high mutual correlation. The average bank angle was chosen over minimum and maximum bank angle. After correlation analysis the dataset contained 8 categorical and 21 numerical variables. Five records had missing values: three with unavailable data for the variables hardness, sulphate and chloride, and two other instances did not have data about fish presence. The total dataset contained information on 125 ponds of which 41 had only alpine newts with no fish, while 42 were solely inhabited by fish and contained no alpine newts (Fig. 3). The most commonly encountered fish were *Pungitius pungitius* (ninespine
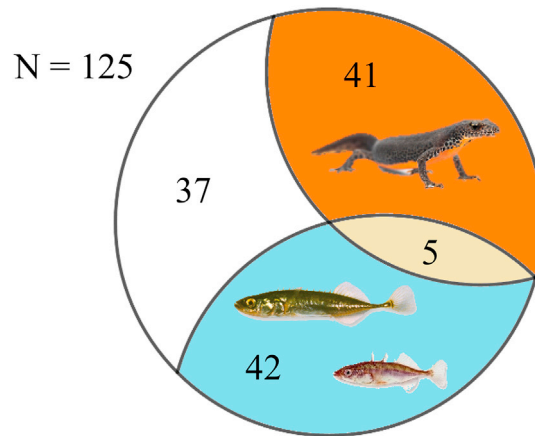
Fig. 3. Venn diagram of the prevalence of fish and alpine newt in the optimized dataset containing 125 ponds (N).

stickleback), *Gasterosteus aculeatus* (three-spined stickleback) and *Carassius gibelio* (Prussian carp), occurring in 71.1%, 26.1% and 21.7% of the ponds with fish, respectively. There were only 5 ponds in which alpine newts and fish co-occur (Fig. 3).

3.2. Model development, selection and testing

All constructed models were first assessed on their technical performance, calculated via %CCI and κ. An extensive overview of the results of all parameterizations and stratifications is provided in Table A3. There were 220 models of 250 possible combinations (88.0%) which fulfilled the requirements of a %CCI and κ of at least 70% and 0.40 respectively. A total of 42 different classification trees could be distinguished. The model which returned most frequently, in 32.3% of all cases, was also amongst the least complex models (maximum of two nodes) and robust as it recurred in 6 out of 10 stratified datasets. The key factors in this model were the dichotomous variable fish occurrence (absence/presence) followed by sludge thickness split at 15 cm (Fig. 4). If fish are present in a pond ecosystem, alpine newts are predicted absent by the model, and in fishless ponds the species tend to be absent when the sludge layer is > 15 cm thick. The PCA biplot shows that sludge thickness was associated with variables related to eutrophication (Figure A3). Positive associations were observed between sludge thickness and concentrations of TN, TP, Chl$_a$, SS while water transparency, measured via sneller depth, negatively associated with this variable. The relationship between eutrophication and the sludge amount was quantitatively investigated via a multiple regression model for which solely the following eutrophication related predictors were retained during model development: TN, TP, Chl$_a$, SS, Sneller. The model had an adjusted R^2 of 16.9% and revealed that only sneller depth had a significant negative relation with sludge on the 5% significance level (Equation A1, Table A4 & Table A5). Total nitrogen showed a positive correlation with sludge on the 10% significance level. All other predictor variables linked to eutrophication (Chl$_a$, SS and TP) were removed during the backwards elimination procedure and therefore did not significantly contribute to the model's predictions.

The second most common model, in 14.1% of all reliable models, was the same as the one described in Fig. 4 apart from the threshold value for sludge thickness, which was deemed 22 cm. Furthermore, in 9.6% of all reliable models, sneller depth (root node) and fish (internal node) determined habitat suitability. A pond characterised by a sneller depth smaller than or equal to 10 cm, is predicted to contain no alpine newts. In less turbid ponds (i.e. sneller >10 cm), alpine newts are likely to be absent in case the pond is inhabited by fish. Both models were present in 3 and 4 out of 10 stratifications, respectively. Due to this lower model robustness, we deemed the model in Fig. 4 to be more representative of alpine newt occurrence in ponds in our data. All other reliable models occurred rather sporadically (<4.6%), and were less robust as they were all the result from a specific
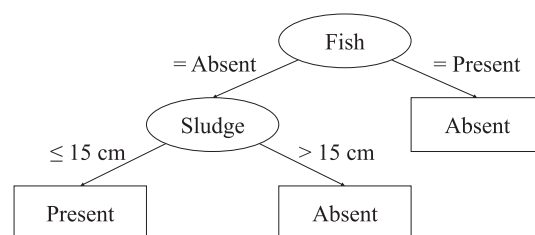

Fig. 4. Most consistently returning model of all reliable classification trees (%CCI ≥ 70% and κ ≥ 0.40). Fish presence results in absence of alpine newts, while in fishless ponds sludge depth is considered to be the most decisive factor.

subset, and were more complex consisting of 3–6 nodes. Note that the most simple models occurred at higher levels of pre-pruning, while more complex and lower performing models on average could be linked to less pre-pruning.

The obtained model described in Fig. 4 was tested with independent newly collected field data. The test ponds were similar to the ponds from 2008 regarding physical-chemical, morphometric and biological characteristics (Table A6). Alpine newts were present in 8 out of 18 sampled ponds and we found two fish species in 5 of the ponds, namely *Pungitius pungitius* and *Gasterosteus aculeatus*. Sludge depth varied from 1 cm to 42 cm with an average of 10.9 cm and a standard deviation of 12.2 cm (Table A7). The model was able to correctly classify 13 out of 18 cases and combined with κ equal to 0.46 this indicated a satisfactory model performance in an ecological context (Goethals et al., 2007).

4. Discussion

4.1. Model relevance

The alpine newt is a urodelan species which is generally believed to have few specific habitat requirements (van Delft, 2009). This idea is supported by the relatively simple structure of the best performing model from the present study (Fig. 4), which only contains fish presence/absence and sludge thickness as major steering variables determining the presence/absence of alpine newts. This classification tree is user-friendly, due to its simplicity and reliability (κ > 40% for test data). We thus propose that the model can be used for management purposes in Belgium and Luxembourg, where alpine newts are generally ubiquitous.

The proposed model indicates a lack of co-occurrence between alpine newts and fish. Indeed, fish compete for the same resources and/or directly predate on the amphibian community in the same habitat (Winandy et al., 2015). For example, Pearson and Goater (2009) observed an almost complete extinction of *Ambystoma macrodactylum* larvae (long-toed salamander) and a 39% reduction in salamander survival in the presence of the predatory fish *Oncorhynchus mykiss* (rainbow trout). The authors also found that the non-piscivorous *Pimephales promelas* (fathead minnows) outcompeted the urodele for zooplankton as a food resource. Several other authors have reported similar outcomes on the effect of fish presence on newt communities (Cabrera-Guzmán et al., 2017; Monello and Wright, 2001; Pagnucco et al., 2011). In ponds, fish either lead to complete extirpation of newts (here: 42 of 125 ponds) or, in case of co-existence (here: 5 of 125 ponds), decrease their population sizes substantially (Knapp, 2005). The most commonly encountered fishes in the assessed ponds of the present study were sticklebacks which can predate on newt egg masses and larvae (Jarvis, 2010). They are also strong competitors as they feed on similar resources, such as zooplankton and macroinvertebrates (Jakubavičiūtė et al., 2017).

Fish presence can also induce behavioural changes in newts and have non-consumptive effects. For example, Winandy and Denoël (2013) reported that alpine newts use shelters (micro-habitats) more frequently and significantly reduce their sexual activity in the presence of *Carassius auratus* (goldfish). Multiple non-predatory fish can have detrimental impacts on aquatic vegetation and can increase water turbidity thereby impeding foraging and signalling efficiency during mating (Richardson et al., 1995; Secondi et al., 2007). Furthermore, fish have a substantial negative effect on newt paedomorphs, i.e. a non-metamorphosized alternative adult phenotype with larval traits such as gills (Denoël et al., 2005). This uncompleted form of transformation is common amongst alpine newts, smooth newts (*Lissotriton vulgaris*) and palmate newts (*Lissotriton helveticus*), making them even more susceptible to fish co-occurrence (Denoël et al., 2009).

In fishless ponds, our data suggest that sludge thickness determines alpine newt occurrence. The amount of sludge is not necessarily a causal factor, as it can be linked to several processes. Sludge thickening occurs during the natural succession of pond ecosystems. Unmanaged shallow ponds will gradually fill up, thereby reducing pond depth and allowing vegetation to gradually colonize the entire ecosystem (Chauchan, 2008). The alpine newt, however, commonly occurs in vegetated pools where they fold eggs in leaves to protect them from predation (van Delft, 2009). This makes it rather unlikely that solely natural succession is responsible for making ponds unsuitable for this urodelan species. The multiple regression model for sludge and our PCA results demonstrate that sludge thickness is mainly negatively related to water transparency and is, to a lesser extent, positively associated with the level of total nitrogen. This result indicates that eutrophication and erosion also potentially influence the sludge accumulation in ponds. Eutrophication leads to algae-dominated systems with sparse submerged vegetation and oxygen depletion during the night. The lack of submerged vegetation reduces the amount of suitable egg depositing habitats, while high turbidity affects the feeding and mating behaviour of newts (Secondi et al., 2007).

4.2. Implications for management

Regarding management practices for alpine newt conservation in an agricultural landscape, decision makers are advised to avoid fish presence and reduce sludge accumulation in ponds. Recurrent dry-stands of such ecosystems might provide suitable habitats as this hampers the persistency of fish populations and limits build-up of sludge due to organic oxidation during the dry phase (Collinson et al., 1995). Besides, this management strategy reduces the mechanisms induced by eutrophication and can stabilize the pond ecosystem in the clear water macrophyte dominated state (Lemmens et al., 2013). Furthermore, hydroperiod, i.e. the duration and frequency of inundation, generally has a major effect on the composition of the biological pond communities (De Meester et al., 2005). Temporary ponds, i.e. ponds with an annual dry phase between 3 and 8 months, shelter very specific biota which are adapted to bridge the dry periods. Such ecosystems commonly provides habitat for very active, rapidly developing organisms and contain fewer predators, such as fish and invertebrate predators,

compared to more permanent waterbodies (Collinson et al., 1995; Wellborn et al., 1996). The unique communities of temporary ponds can therefore contribute significantly to regional biodiversity resulting in a high pond conservation value (Biggs et al., 1994; Lemmens et al., 2013). However, if these waterbodies are located nearby lotic systems or permanent fish-containing ponds, fish can be introduced during flooding or rainfall events reducing habitat suitability for the alpine newt. Furthermore, it is advisable that the hydroperiod is not too short and that ponds dry out earliest in late summer, by which time offspring of alpine newts has left the pond (Bauwens and Claus, 1996). Larval growth rate can, however, vary depending on a pond's hydroperiod due to developmental plasticity (Griffiths, 1997).

In addition to regular dry-stands, buffer strips consisting of e.g. reed, soft rush and water plantain should be encouraged as a protective measure against nutrient input from surrounding land, especially in case of permanent (cattle drinking) ponds (Puglis and Boone, 2012; Van Damme et al., 1997). This can also reduce the negative impact of agricultural contaminants, such as pesticide residues and heavy metals (Dorioz et al., 2006; Schmitt et al., 1999). Patches of dense, high grass can be added to the buffer zone so that adult newts can forage and find shelter. The herb layer then preferably gradually evolves to shrubs and forest as these elements provide excellent hibernation and shelter places, such as mouldered branches and trunks (Van Damme et al., 1997; van Delft, 2009).

Sludge accumulation can also be countered by dredging and vegetation removal, mainly to control natural succession and to preserve the aquatic ecosystem. However, it is inadvisable that all ponds in a landscape are dredged simultaneously as this could result in a significant decrease in aquatic gamma diversity through habitat homogenisation (Hassall, 2014; Biggs et al., 1994). Teurlincx et al. (2018), for example, promote successional stage heterogeneity via the Cyclic Rejuvenation through Management (CRM) approach, where aquatic water bodies are periodically and asynchronously reset.

5. Conclusions

The present research provides a reliable and user-friendly classification tree model as a management tool to assess the habitat suitability of farmland ponds for alpine newts. The major steering variables within local habitats are the presence of fish and the thickness of the sludge layer. Ponds without fish provide suitable habitat for alpine newts, while the species also prefers a sludge thickness less or equal to 15 cm. We advise decision makers to create temporary ponds to eliminate fish and reduce sludge accumulation. These temporary ponds should preferably not receive water from adjacent fish-containing water bodies. In addition, construction of a buffer zone around ponds limits the nutrient and soil input from land to the waterbody so build-up of sludge due to eutrophication and erosion is hindered, especially in agricultural areas where fertilization is widely applied.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.gecco.2020.e01086.

Data availability statement

Data available from GBIF.org https://doi.org/10.15468/zs6sky (Lemmens et al., 2018).

References

Andersen, M.R., Kragh, T., Sand-Jensen, K., 2017. Extreme diel dissolved oxygen and carbon cycles in shallow vegetated lakes. Proc. R. Soc. B 284 (1862). https://doi.org/10.1098/rspb.2017.1427.

Arntzen, J.W., Denoël, M., Kuzmin, Sergius, Ischchenko, V., Beja, P., Andreone, F., et al., 2009. Ichthyosaura Alpestris. *The IUCN Red List of Threatened Species 2009.* IUCN Red List. https://doi.org/10.2305/iucn.uk.2009.rlts.t59472a11946568.en. https://search.crossref.org/?q=Arntzen%2C+J.+W.%2C+Deno%C3%ABl%2C+M.%2C+Kuzmin+Sergius%2C+Ischchenko%2C+V.%2C+Beja%2C+P.%2C+Andreone%2C+F.%2C+et+al.+%282009%29.+Ichthyosaura+Alpestris.+The+IUCN+Red+List+of+Threatened+Species+2009.+IUCN+Red+List.

Bauwens, D., Claus, K., 1996. Verspreiding van amfibieën en reptielen in Vlaanderen. De Wielewaal, Turnhout.

Biggs, J., Corfield, A., Walker, D., Whitfield, M., Williams, P., 1994. New approaches to the management of ponds. Br. Wildl. 5 (5), 273–287.

BirdLife International, 2015. European Red List of Birds. Office for Official Publications of the European Communities, Luxembourg. https://doi.org/10.2779/975810.

Boets, P., Lock, K., Messiaen, M., Goethals, P.L.M., 2010. Combining data-driven methods and lab studies to analyse the ecology of *Dikerogammarus villosus*. Ecol. Inf. 5 (2), 133–139. https://doi.org/10.1016/J.ECOINF.2009.12.005.

Cabrera-Guzmán, E., Díaz-Paniagua, C., Gomez-Mestre, I., 2017. Competitive and predatory interactions between invasive mosquitofish and native larval newts. Biol. Invasions 19 (5), 1449–1460. https://doi.org/10.1007/s10530-017-1369-5.

Chauchan, B.S., 2008. Environmental Studies, second ed. Laxmi Publications.

Cohen, J., 1960. A coefficient of agreement for nominal scales. Educ. Psychol. Meas. 20 (1), 37–46. https://doi.org/10.1177/001316446002000104.

Collinson, N.H., Biggs, J., Corfield, A., Hodson, M.J., Walker, D., Whitfield, M., Williams, P.J., 1995. Temporary and permanent ponds: an assessment of the effects of drying out on the conservation value of aquatic macroinvertebrate communities. Biol. Conserv. 74 (2), 125–133. https://doi.org/10.1016/0006-3207(95)00021-U.

Cox, N.A., Temple, H.J., 2009. European Red List of Reptiles. Office for Official Publications of the European Communities, Luxembourg.

Curado, N., Hartel, T., Arntzen, J.W., 2011. Amphibian pond loss as a function of landscape change - a case study over three decades in an agricultural area of northern France. Biol. Conserv. 144 (5), 1610–1618. https://doi.org/10.1016/j.biocon.2011.02.011.

Dancey, C.P., Reidy, J., 2004. Statistics without Maths for Psychology: Using SPSS for Windows. Pearson Education, Harlow, United Kingdom.

De Meester, L., Declerck, S., Stoks, R., Louette, G., Van De Meutter, F., De Bie, T., et al., 2005. Ponds and pools as model systems in conservation biology, ecology and evolutionary biology. Aquat. Conserv. Mar. Freshw. Ecosyst. 15 (6), 715–725. https://doi.org/10.1002/aqc.748.

Declerck, S., De Bie, T., Ercken, D., Hampel, H., Schrijvers, S., Van Wichelen, J., et al., 2006. Ecological characteristics of small farmland ponds: associations with land use practices at multiple spatial scales. Biol. Conserv. 131 (4), 523–532. https://doi.org/10.1016/J.BIOCON.2006.02.024.

Denoël, M., Džukić, G., Kalezić, M.L., 2005. Effects of widespread fish introductions on paedomorphic newts in Europe. Conserv. Biol. 19 (1), 162–170. https://doi.org/10.1111/j.1523-1739.2005.00001.x.

Denoël, M., Ficetola, G.F., Ćirović, R., Radović, D., Džukić, G., Kalezić, M.L., Vukov, T.D., 2009. A multi-scale approach to facultative paedomorphosis of European newts (Salamandridae) in the Montenegrin karst: distribution pattern, environmental variables, and conservation. Biol. Conserv. 142 (3), 509–517. https://doi.org/10.1016/J.BIOCON.2008.11.008.

Dorioz, J.M., Wang, D., Poulenard, J., Trévisan, D., 2006. The effect of grass buffer strips on phosphorus dynamics—a critical review and synthesis as a basis for application in agricultural landscapes in France. Agric. Ecosyst. Environ. 117 (1), 4–21. https://doi.org/10.1016/J.AGEE.2006.03.029.

Everaert, G., Boets, P., Lock, K., Džeroski, S., Goethals, P.L.M., 2011. Using classification trees to analyze the impact of exotic species on the ecological assessment of polder lakes in Flanders, Belgium. Ecol. Model. 222 (14), 2202–2212. https://doi.org/10.1016/J.ECOLMODEL.2010.08.013.

Everaert, G., Pauwels, I., Bennetsen, E., Goethals, P.L.M., 2016. Development and selection of decision trees for water management: impact of data pre-processing, algorithms and settings. AI Commun. 29 (6), 711–723. https://doi.org/10.3233/AIC-160711.

Freyhof, J., Brooks, E., 2011. European Red List of Freshwater Fishes. Publications Office of the European Union, Luxembourg. https://doi.org/10.2779/85903.

Fürnkranz, J., 1997. Pruning algorithms for rule learning. Mach. Learn. 27 (2), 139–172. https://doi.org/10.1023/A:1007329424533.

Garner, P., 1996. Microhabitat use and diet of 0+ cyprinid fishes in a lentic, regulated reach of the River Great Ouse, England. J. Fish. Biol. 48 (3), 367–382. https://doi.org/10.1111/j.1095-8649.1996.tb01434.x.

Goethals, P.L.M., Dedecker, A.P., Gabriels, W., Lek, S., De Pauw, N., 2007. Applications of artificial neural networks predicting macroinvertebrates in freshwaters. Aquat. Ecol. 41 (3), 491–508. https://doi.org/10.1007/s10452-007-9093-3.

Griffiths, R.A., 1997. Temporary ponds as amphibian habitats. Aquat. Conserv. Mar. Freshw. Ecosyst. 7 (2), 119–126. https://doi.org/10.1002/(sici)1099-0755(199706)7:2<119::aid-aqc223>3.0.co;2-4. https://search.crossref.org/?q=Griffiths%2C+R.+A.+%281997%29.+Temporary+ponds+as+amphibian+habitats.+Aquat.+Conserv.+Mar.+Freshw.+Ecosyst.%2C+7%282%29%2C+119%E2%80%93126.

Han, J., Kamber, M., Pei, J., 2011. Data Mining: Concepts and Techniques, third ed. Morgan Kaufmann.

Harrell Jr, F.E., 2017. Hmisc: Harrell Miscellaneous. R package version 4.0-3. https://cran.r-project.org/package=Hmisc.

Hassall, C., 2014. The ecology and biodiversity of urban ponds. Wiley Interdiscipl. Rev.: Water 1 (2), 187–206. https://doi.org/10.1002/wat2.1014.

Hehl-Lange, S., 2001. Structural elements of the visual landscape and their ecological functions. Landsc. Urban Plann. 54 (1–4), 107–115. https://doi.org/10.1016/S0169-2046(01)00129-3.

Hoang, T.H., Lock, K., Mouton, A., Goethals, P.L.M., 2010. Application of classification trees and support vector machines to model the presence of macroinvertebrates in rivers in Vietnam. Ecol. Inf. 5 (2), 140–146. https://doi.org/10.1016/J.ECOINF.2009.12.001.

Hof, C., Araújo, M.B., Jetz, W., Rahbek, C., 2011. Additive threats from pathogens, climate and land-use change for global amphibian diversity. Nature 480 (7378), 516–519. https://doi.org/10.1038/nature10650.

Jakubavičiūtė, E., Bergström, U., Eklöf, J.S., Haenel, Q., Bourlat, S.J., 2017. DNA metabarcoding reveals diverse diet of the three-spined stickleback in a coastal ecosystem. PLoS One 12 (10). https://doi.org/10.1371/journal.pone.0186929 e0186929.

Jarvis, L.E., 2010. Non-consumptive effects of predatory three-spined sticklebacks (*Gasterosteus aculeatus*) on great crested newt (*Triturus cristatus*) embryos. Herpetol. J. 20, 271–275.

John, G.H., 1995. Robust decision trees: removing outliers from databases. In: Proceedings of the First International Conference on Know Ledge Discovery and Data Mining. AAAI Press, Montreal, Canada, pp. 174–179.

Jooris, R., Engelen, P., Speybroeck, J., Lewylle, I., Louette, G., Bauwens, D., Maes, D., 2012. De IUCN Rode Lijst van de amfibieën en reptielen in Vlaanderen. In: *Rapporten van Het Instituut Voor Natuur- En Bosonderzoek*. Brussels: Instituut voor Natuur- en Bosonderzoek.

Josse, J., Husson, F., 2016. missMDA: a package for handling missing values in multivariate data analysis. J. Stat. Software 70 (1), 1–31. https://doi.org/10.18637/jss.v070.i01.

Knapp, R.A., 2005. Effects of nonnative fish and habitat characteristics on lentic herpetofauna in Yosemite National Park, USA. Biol. Conserv. 121 (2), 265–279. https://doi.org/10.1016/J.BIOCON.2004.05.003.

Le, S., Josse, J., Husson, F., 2008. FactoMineR: an R package for multivariate analysis. J. Stat. Software 25 (1), 1–18. https://doi.org/10.18637/jss.v025.i01.

Lemmens, P., De Wever, A., Bonjean, N., Castiaux, A., Colson, L., De Bie, T., et al., 2018. Database of the PONDSCAPE project (Towards a sustainable management of pond diversity at the landscape level). Freshwater Metadata J. 31, 1–10. https://doi.org/10.15504/FMJ.2018.31.

Lemmens, P., Mergeay, J., De Bie, T., Van Wichelen, J., De Meester, L., Declerck, S.A.J., 2013. How to maximally support local and regional biodiversity in applied conservation? Insights from pond management. PLoS One 8 (8). https://doi.org/10.1371/journal.pone.0072538 e72538.

Manel, S., Ceri Williams, H., Ormerod, S.J., 2001. Evaluating presence-absence models in ecology: the need to account for prevalence. J. Appl. Ecol. 38, 921–931.

Martel, A., Spitzen-Van Der Sluijs, A., Blooi, M., Bert, W., Ducatelle, R., Fisher, M.C., et al., 2013. Batrachochytrium salamandrivorans sp. nov. causes lethal chytridiomycosis in amphibians. Proc. Natl. Acad. Sci. Unit. States Am. 110 (38), 15325–15329. https://doi.org/10.1073/pnas.1307356110.

Monello, R.J., Wright, R.G., 2001. Predation by goldfish (Carassius auratus) on eggs and larvae of the eastern long-toed salamander (Ambystoma macrodactylum columbianum). J. Herpetol. 35 (2), 350. https://doi.org/10.2307/1566132.

Nieto, A., Ralph, G.M., Comeros-Raynal, M.T., Kemp, J., Criado, M.G., Allen, D.J., et al., 2015. European Red List of Marine Fishes. Publications Office of the European Union, Luxembourg. https://doi.org/10.2779/082723.

Orr, J.M., Sackett, P.R., Dubois, C.L.Z., 1991. Outlier detection and treatment in I/O psychology: a survey of researcher beliefs and an empirical illustration. Person. Psychol. 44 (3), 473–486. https://doi.org/10.1111/j.1744-6570.1991.tb02401.x.

Pagnucco, K.S., Paszkowski, C.A., Scrimgeour, G.J., 2011. Wolf in sheep's clothing: effects of predation by small-bodied fish on survival and behaviour of salamander larvae. Ecoscience 18 (1), 70–78. https://doi.org/10.2980/18-1-3395.

Pearson, K.J., Goater, C.P., 2009. Effects of predaceous and nonpredaceous introduced fish on the survival, growth, and antipredation behaviours of long-toed salamanders. Can. J. Zool. 87 (10), 948–955. https://doi.org/10.1139/Z09-084.

Perrow, M.R., Jowitt, A.J.D., Zambrano González, L., 1996. Sampling fish communities in shallow lowland lakes: point-sample electric fishing vs electric fishing within stop-nets. Fish. Manag. Ecol. 3 (4), 303–313. https://doi.org/10.1046/j.1365-2400.1996.d01-152.x.

Puglis, H.J., Boone, M.D., 2012. Effects of terrestrial buffer zones on Amphibians on golf courses. PloS One 7 (6). https://doi.org/10.1371/journal.pone.0039590 e39590.

R Core Team, 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org/.

Richardson, M.J., Whoriskey, F.G., Roy, L.H., 1995. Turbidity generation and biological impacts of an exotic fish Carassius auratus, introduced into shallow seasonally anoxic ponds. J. Fish. Biol. 47 (4), 576–585. https://doi.org/10.1111/j.1095-8649.1995.tb01924.x.

Schmitt, T.J., Dosskey, M.G., Hoagland, K.D., 1999. Filter strip performance and processes for different vegetation, widths, and contaminants. J. Environ. Qual. 28 (5), 1479–1489. https://doi.org/10.2134/jeq1999.00472425002800050013x.

Secondi, J., Aumjaud, A., Pays, O., Boyer, S., Montembault, D., Violleau, D., 2007. Water turbidity affects the development of sexual morphology in the palmate newt. Ethology 113 (7), 711–720. https://doi.org/10.1111/j.1439-0310.2007.01375.x.

Spitzen-Van Der Sluijs, A., Martel, A., Hallmann, C.A., Bosman, W., Garner, T.W.J., Van Rooij, P., et al., 2014. Environmental determinants of recent endemism of Batrachochytrium dendrobatidis infections in Amphibian assemblages in the absence of disease outbreaks. Conserv. Biol. 28 (5), 1302–1311. https://doi.org/10.1111/cobi.12281.

Stegen, G., Pasmans, F., Schmidt, B.R., Rouffaer, L.O., Van Praet, S., Schaub, M., et al., 2017. Drivers of salamander extirpation mediated by Batrachochytrium salamandrivorans. Nature 544 (7650). https://doi.org/10.1038/nature22059.

Temple, H.J., Cox, N.A., 2009. European Red List of Amphibians. Office for Official Publications of the European Communities, Luxembourg.

Temple, H.J., Terry, A., 2007. The Status and Distribution of European Mammals. Office for Official Publications of the European Communities, Luxembourg.

Teurlincx, S., Verhofstad, M.J.J.M., Bakker, E.S., Declerck, S.A.J., 2018. Managing successional stage heterogeneity to maximize landscape-wide biodiversity of aquatic vegetation in ditch networks. Front. Plant Sci. 9, 1013. https://doi.org/10.3389/fpls.2018.01013.

Van Damme, R., Bervoets, L., De Losse, L., 1997. Poelen, spiegels van het landschap. In: Hermy, M., De Blust, G. (Eds.), Punten en lijnen in het landschap, first ed. Mark Van de Wiele, Brugge, p. 336.

van Delft, J., 2009. Alpenwatersalamander. In: Nederlandse Fauna: Amfibieën en reptielen van Nederland. KNNV Uitgeverij, pp. 96–104.

Van Echelpoel, W., Boets, P., Landuyt, D., Gobeyn, S., Everaert, G., Bennetsen, E., et al., 2015. Species distribution models for sustainable ecosystem management. In: Developments in Environmental Modelling. Elsevier B.V, pp. 115–134. https://doi.org/10.1016/B978-0-444-63536-5.00008-9.

Venables, W.N., Ripley, B.D., 2002. Modern Applied Statistics with S, fourth ed. Springer, New York http://www.stats.ox.ac.uk/pub/MASS4.

Verbelen, D., Jooris, R., 2009. Monitoringhandleiding natuurpunt module F7 - poelenonderzoek amfibieën. Natuurpunt - Hyla werkgroep 1–10.

Wei, T., Simko, V., 2017. R Package "Corrplot": Visualization of a Correlation Matrix (Version 0.84). https://github.com/taiyun/corrplot.

Wellborn, G.A., Skelly, D.K., Werner, E.E., 1996. Mechanisms creating community structure across a freshwater habitat gradient. Annu. Rev. Ecol. Systemat. 27 (1), 337–363. https://doi.org/10.1146/annurev.ecolsys.27.1.337.

Whitney, R.J., 1942. Diurnal fluctuations of oxygen and pH in two small ponds and A stream. J. Exp. Biol. 19 (1), 92–99.

Winandy, L., Darnet, E., Denoël, M., 2015. Amphibians forgo aquatic life in response to alien fish introduction. Anim. Behav. 109, 209–216. https://doi.org/10.1016/j.anbehav.2015.08.018.

Winandy, L., Denoël, M., 2013. Introduced goldfish affect amphibians through inhibition of sexual behaviour in risky habitats: an experimental approach. PLoS One 8 (11), 1–11. https://doi.org/10.1371/journal.pone.0082736.

Witten, I.H., Eibe, F., 2005. Data Mining: Practical Machine Learning Tools and Techniques, second ed. Morgan Kaufmann Publishers, San Francisco.