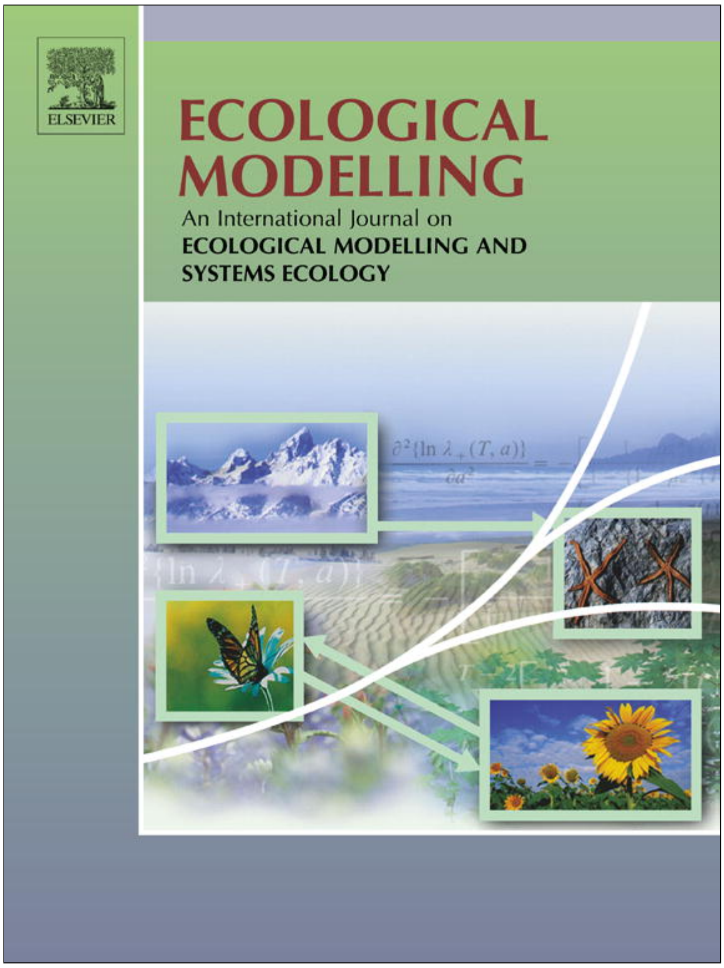

ELSEVIER
ECOLOGICAL MODELLING
An International Journal on
ECOLOGICAL MODELLING AND SYSTEMS ECOLOGY

Ecological Modelling 251 (2013) 187–198

Contents lists available at SciVerse ScienceDirect

Ecological Modelling

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

Consistency of fuzzy rules in an ecological context

Juan C. Gutiérrez-Estrada[a,*], Inmaculada Pulido-Calvo[a], David T. Bilton[b]

[a] *Dpto. Ciencias Agroforestales, Escuela Técnica Superior de Ingeniería, Campus de La Rábida, Universidad de Huelva, 21819 Palos de la Frontera, Huelva, Spain*
[b] *Marine Biology and Ecology Research Centre, School of Marine Science and Engineering, University of Plymouth, Drake Circus, Plymouth, PL4 8AA, UK*

ARTICLE INFO

Article history:
Received 16 July 2012
Received in revised form
12 December 2012
Accepted 19 December 2012

Keywords:
Evolutionary algorithm
Water beetle
Shannon diversity index
Contingency
Generalized additive models

ABSTRACT

In this paper, we assess the performance of fuzzy inference systems (FISs) and the consistency of fuzzy rules generated from a meta-analysis exploring diversity–environment relationships, in a system of temporary and fluctuating ponds located in two regions of southern England. The analyses focus on aquatic coleopteran assemblages, which act as excellent surrogates of wider freshwater macroinvertebrate diversity. Evaluated FISs were calibrated using evolutionary algorithms and the consistency of the rules examined using a consistency index specifically developed in this work. The best fit accounted for 76% of observed variability in the Shannon diversity index across ponds in the validation phase, which was 56 points better than the benchmark value established by a generalized additive model (GAM). The analysis of fuzzy rules indicated that the basic dynamics of this system are controlled by 8 rules. Another 10 complementary rules were detected, suggesting that more than a single dimension controlled the dynamics of the system. Therefore, water beetle diversity appears to be driven by a relatively short set of rules which relate diversity and environmental factors in a non-linear manner. These rules can be grouped according to their consistency levels, which reflect differences in coleopteran community composition.

1. Introduction

Ever since Lotfi Zadeh developed his theory of fuzzy sets, there are many fields of science and engineering in which this theory has been implemented efficiently to solve complex problems (Zadeh, 1973; Pappis et al., 1977; Yu et al., 2004). Basically, Zadeh formalized mathematically the way in which humans interact with the environment. The human mind tends to filter out fuzziness so that decisions and actions are more easily made and, therefore, tends to perceive discrete objects and events, distinct boundaries and definite classes of things (Bosserman and Ragade, 1982). These features of fuzzy sets theory have clearly been of interest to ecologists, which explains the significant increase of applications of this heuristic to ecological problems (Meesters et al., 1998; Kampichler et al., 2000; Addriaenssens et al., 2004; Cheung et al., 2005; Mouton et al., 2009).

There are several specific properties that have popularized the use of fuzzy sets theory among ecologists. The first is that it is an intelligent computational method which does not require a very detailed mathematical description of the process to be analyzed. Generally, these models are easily integrated-implemented in fuzzy logic controllers (FLCs) or fuzzy inference systems (FISs) and utilize a form of many-valued logic. This means that a fuzzy set can be divided in different regions by mean geometric partitions associated with linguistic concepts which allow us to describe a discrete point as a function of its membership to different sets (Zeldis and Prescott, 2000). Other very interesting property of fuzzy sets theory and FISs is that the rules that control the system dynamic can be grouped and modelled into a defined rule-base as a set of identifiable and comprehensible linguistic labels. In this rule-base (also known as fuzzy associative memory or FAM), a set of antecedents or premises which also could be interpreted as independent variables (x_i), are related with consequents or dependent variables (y) in the following form: R_i = IF x_1 is $A_{i1}(x_1)$ and x_2 is $A_{i2}(x_2)$ and. . .and x_n is $A_{in}(x_n)$, THEN y is $B_i(y)$, where $A_{i1}(x_1)$, $A_{i2}(x_2)$, $A_{in}(x_n)$ and $B_i(y)$ are linguistic concepts, and R_i is the ith rule of the FAM. Therefore, from an ecological point of view, an important advantage of analyzing an ecosystem with fuzzy sets theory is that the information or knowledge is structured in the form of rules.

However, the knowledge extraction via rules can be an inconvenient in some circumstances. For example, let us suppose that, using an FIS, we are trying to explain the effects of five environmental variables (chemical or physical factors) on the species richness of a faunistic group. Also, let us suppose that each environmental variable, that is, each independent variable is made up of three fuzzy partitions. This implies that, theoretically, the FAM could have a total of 3^5 rules (243 rules). In this FAM we could include rules that apparently provide contradictory information, which would reduce our ability to find an ecological meaning of the variables implied in the system. This is particularly common when the FISs are calibrated from real environmental data. Therefore, when we

---

[*] Corresponding author. Tel.: +34 959217528; fax: +34 959217528.
E-mail address: juanc@uhu.es (J.C. Gutiérrez-Estrada).

0304-3800/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ecolmodel.2012.12.013

meta-analyze an ecosystem with FISs, it is not enough to provide an FAM. It is also necessary to analyze the consistency of rules, that is, the absence of contradictory rules in the FAM.

In the literature there are very few works that analyze the consistency of FAMs and all are focused to engineering applications (Alonso et al., 2008). For example, Jin et al. (1999) proposed a methodology based on evolutionary algorithms for generating consistent and compact FAMs for the automatic control of safety distance in cars. They provided a consistency index calculated as a function of the similarities of rule premises and rule consequents. On the other hand, Cheong and Lai (2000) used genetic algorithms (GAs) to design FISs for controlling three different industrial plans. These authors used a genetic strategy to obtain an FIS with restrictions in the FAM. In this way, the number of rules fired at the same time was minimized to improving the consistency of the FAM. In both these works it is assumed that the consistency of the FAM results in the absence of contradictory rules, in the sense that rules with similar premises or antecedents should have similar consequent parts (Gacto et al., 2011). In other words, two inconsistent rules will have similar premise parts, but will have rather different consequents.

This initial assumption on the consistency of rules could generate erroneous conclusions when we carry out a meta-analysis in an ecological context because unlike engineering problems, the environmental data base could be made up of contingent records. Contingency arises when the nature and composition of ecosystems in different locations are different realizations of the same underlying processes (Schmitz, 2010). Therefore, any consistency analysis of rules generated from environmental data must explicitly confront the issue of contingency. To cope with the above mentioned problem, a new approach of consistency based on the contingent nature of environmental data is proposed in this paper.

In summary, we use an FIS optimized with evolutionary algorithms to extract the set of rules that controls the composition of a local assemblage. We selected a system which has been previously analyzed, and therefore relatively well studied, providing a set of reference results. The system selected was composed of a series of temporary and fluctuating ponds in two regions of southern England (Lizard and New Forest). In both regions, we focused our attention on the different diversity patterns of the community of aquatic Coleopera and their relationships with environmental factors. This macroinvertebrate assemblage was selected because this group is relatively diverse, ecologically well understood and occurs across a wide spectrum of pond types (Bilton et al., 2006; Sánchez-Fernández et al., 2006). Evaluation of the accuracy and validity of each FIS model was carried out via the comparison of error levels with those obtained in a reference model, which in our case was a generalized additive model (GAM). Finally, the consistency of the extracted rules from the best FISs were analyzed with our new approach and compared with existing methodologies.

2. Methods

2.1. Study area and sampling

The dataset used in this study was obtained from 76 temporary ponds located in the New Forest (Hampshire) and the Lizard Peninsula (Cornwall), both in southern England. These regions contain a high density of temporary and fluctuating ponds, differing widely in biological, physical and chemical characteristics. A detailed description of these ecosystems can be found in Bilton et al. (2001), Rundle et al. (2002), Bilton et al. (2006) and Bilton et al. (2009).

Each pond was sampled during February and March 2000, a time when the spatial extent and the presence of ponds were at their maximum (Bilton et al., 2006). Ponds were sampled using a hand net (1 mm mesh, dimensions 20 cm × 25 cm), taking semi-quantitative 1 m sweeps amongst aquatic vegetation. Each 1 m sweep involved approximately 10 s of back and forth netting over the same area of habitat. This sampling protocol has been favourably evaluated in several investigations on invertebrate assemblages (Rundle et al., 2002; Foggo et al., 2003; Bilton et al., 2006). Two or three such samples were taken from the largest sites according to their area. Sweeps were pooled and samples preserved in 95% ethanol in the field.

In addition, a wide range of environmental variables was recorded. Before Coleoptera were sampled pH, temperature compensated conductivity and turbidity readings were taken on-site using a Solomat 520C probe (Zellweger Analytics, Poole, UK). Water depth in the area sampled was estimated using a 1 m rule (mean of five measurements). For analysis of metal cations (calcium, magnesium, aluminium, nickel, chromium, cobalt, iron, zinc and copper) and nutrient concentrations (organic nitrate and soluble reactive phosphorus), two water samples from each pond were also collected. However, Gutiérrez-Estrada and Bilton (2010) have previously demonstrated that only four environmental variables (conductivity, turbidity, magnesium concentration and depth) were enough to explain more than 82% of the variation of water beetles diversity in New Forest and Lizard regions. Therefore, in this study only these four variables were considered as inputs to the FIS.

Later, in the laboratory beetles were counted and determined to species level. Shannon's index (H') was calculated for Coleoptera from each pond following Brower et al. (1998). This diversity measure (the output of the FIS) was selected because it is easy to calculate and reflects both species richness and the relative abundance of species within assemblages. H' normally varies between 1.5 and 3.5, with values higher than 3 being seen as representing diverse communities whilst those below 2 are relatively uniform (Cowell et al., 2004).

2.2. Fuzzy inference system (FIS)

A fuzzy inference system is a model in which the theory of fuzzy sets has been implemented. In these models, three types of parameters must be optimized in function of input and outputs variables: (1) the shape of the fuzzy sets or geometrical partitions; (2) the degree of overlap between fuzzy sets; and (3) the definition of the IF-THEN rules. In our case, both the independent variable or input variable (the diversity of aquatic Coleoptera community measure as Shannon's index) and dependent variables (environmental factors: conductivity, turbidity, magnesium concentration and depth) were divided in three geometrical partitions labelled as 'Low Level', 'Normal Level' and 'High Level' for which four types of normal and convexes membership functions were tested: (a) symmetric and non-symmetric triangular forms; (b) symmetric and non-symmetric trapezoidal forms; (c) symmetric and non-symmetric non-continuous multipoint forms (with seven points); and (d) symmetric and non-symmetric continuous PI forms (Fig. 1).

In this work, an evolutionary algorithm was used to find the optimal values of the model parameters. Evolutionary algorithms are non-linear search and optimization methods inspired by the biological processes of natural selection and survival of the fittest (Holland, 1975; Goldberg, 1989). Generally, in an evolutionary algorithm the basic unit is the gene. Various genes contain the information required to define a chromosome whose decoding is interpreted as an individual. Thus, the parameters of the model were coded as genes in the chromosome. Once the initial information has been coded, four classical types of operators (reproduction, crossover, mutation and abort) were used in order to evolve towards a suboptimal fuzzy configuration. The steps in the

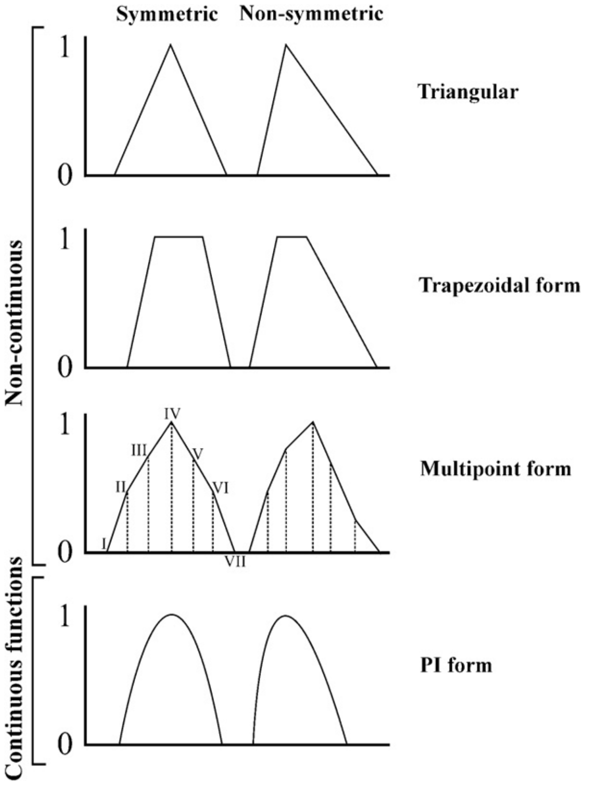


Fig. 1. Non-continuous and continuous membership functions tested. In symmetric multipoint form, the number of points codified is shown.

optimization of the FIS parameters by means an evolutionary algorithm are summarized in Fig. 2.

In order to facilitate the application of this methodology, the Nubecilla 1.0 (Beta version) computer program, written in Microsoft Visual Basic®, was designed (Fig. 3). The Nubecilla program integrates an importation data module, a module to design the architecture of the FIS and a parameter's optimizer based on evolutionary methodologies. Additionally, the Nubecilla program can carry out a sensitivity analysis based on the fuzzy input presence measure index (FIPM) proposed by Baraldi et al. (2009).

2.3. Rules consistency

In this work, the evaluation of rule consistency is based on a similarity concept. A fuzzy similarity measure provides the degree to which two fuzzy sets are equal. Following Jin et al. (1999), the fuzzy similarity can be calculated as:

$$S(C,D)=\frac{M(C\cap D)}{M(C\cup D)}=\frac{M(C\cap D)}{M(C)+M(D)-M(C\cap D)} \tag{1}$$

where C and D are two fuzzy sets and $M(C)$ and $M(D)$ are the areas of C and D, respectively. From the similarity concept it is possible to define two new concepts: (a) the global similarity of rule premises (GSRP) and; (b) the similarity of rule consequents (SRC). GSRP(i,k) and SRC(i,k) are defined as:

$$\mathrm{GSRP}(i,k)=\frac{\sum_{j=1}^{n}S(A_{ij},A_{kj})\cdot W_j}{\sum_{j=1}^{n}W_j} \tag{2}$$

$$\mathrm{SRC}(i,k)=S(B_i,B_k) \tag{3}$$

where $S(A_{ij},A_{kj})$ is the similarity of the jth premises of the ith and kth rules, n is the number of premises, W_j is the weight of the premise in the model estimated from the sensitivity analysis and $S(B_i,B_k)$ is the similarity between the fuzzy consequents of the ith and kth rules. From these definitions, GSRP(i,k) = 0 if no premise is overlapped and GSRP(i,k) = 1 if the all premises of the rules i and k are the same. On the other hand, SRC(i,k) = 0 if B_i and B_k are not overlapped and SRC(i,k) = 1 if $B_i = B_k$. From GSRP and SRC is possible to formulate a consistency index (CC) that incorporates the contingency concept. This index is a continuous function for any value of GSRP and SRC with five decision regions, that is, two regions with values close or equal to one, two regions with values close or equal to zero and a region with values between zero and one and close to 0.5. The first two regions will be associated to rules which are clearly compatible, with either very similar or very different premises and consequents. Instead, the second two regions will be associated to contingent rules, with very similar premises but very different consequents, or very different premises but very similar consequents. The fifth region will group compatible rules in greater or lesser extent as a function of their CC values. Therefore CC is defined as:

$$CC=1-\mathrm{SRC}-\mathrm{GSRP}+2(\mathrm{SRC}\cdot\mathrm{GSRP}) \tag{4}$$

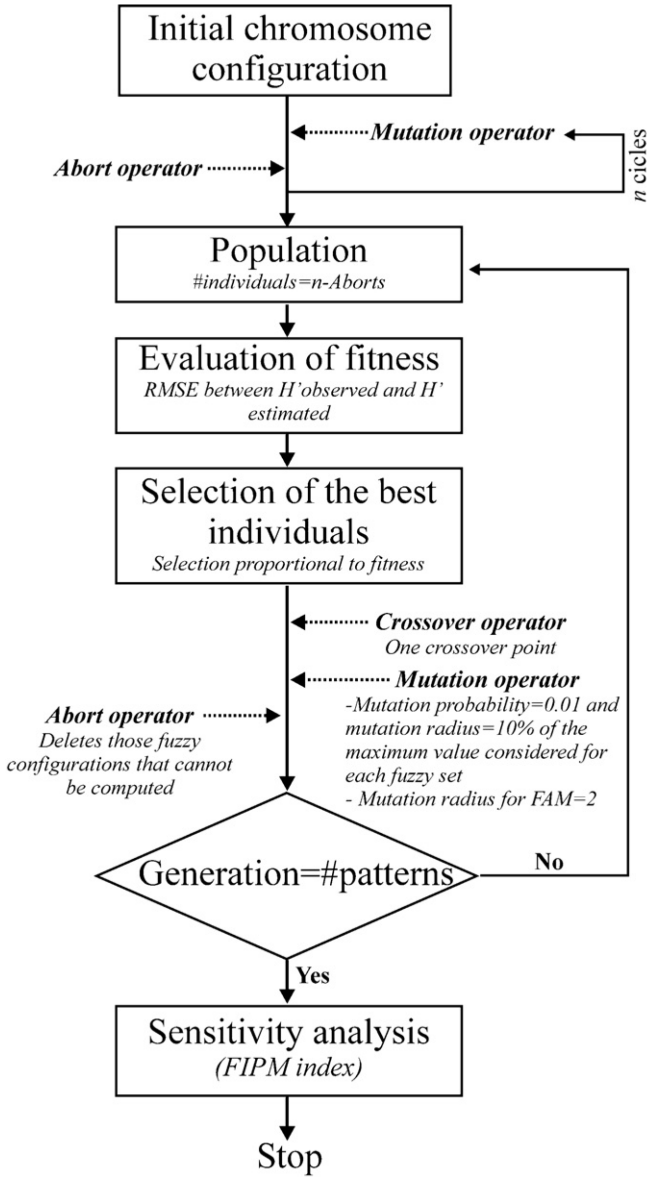
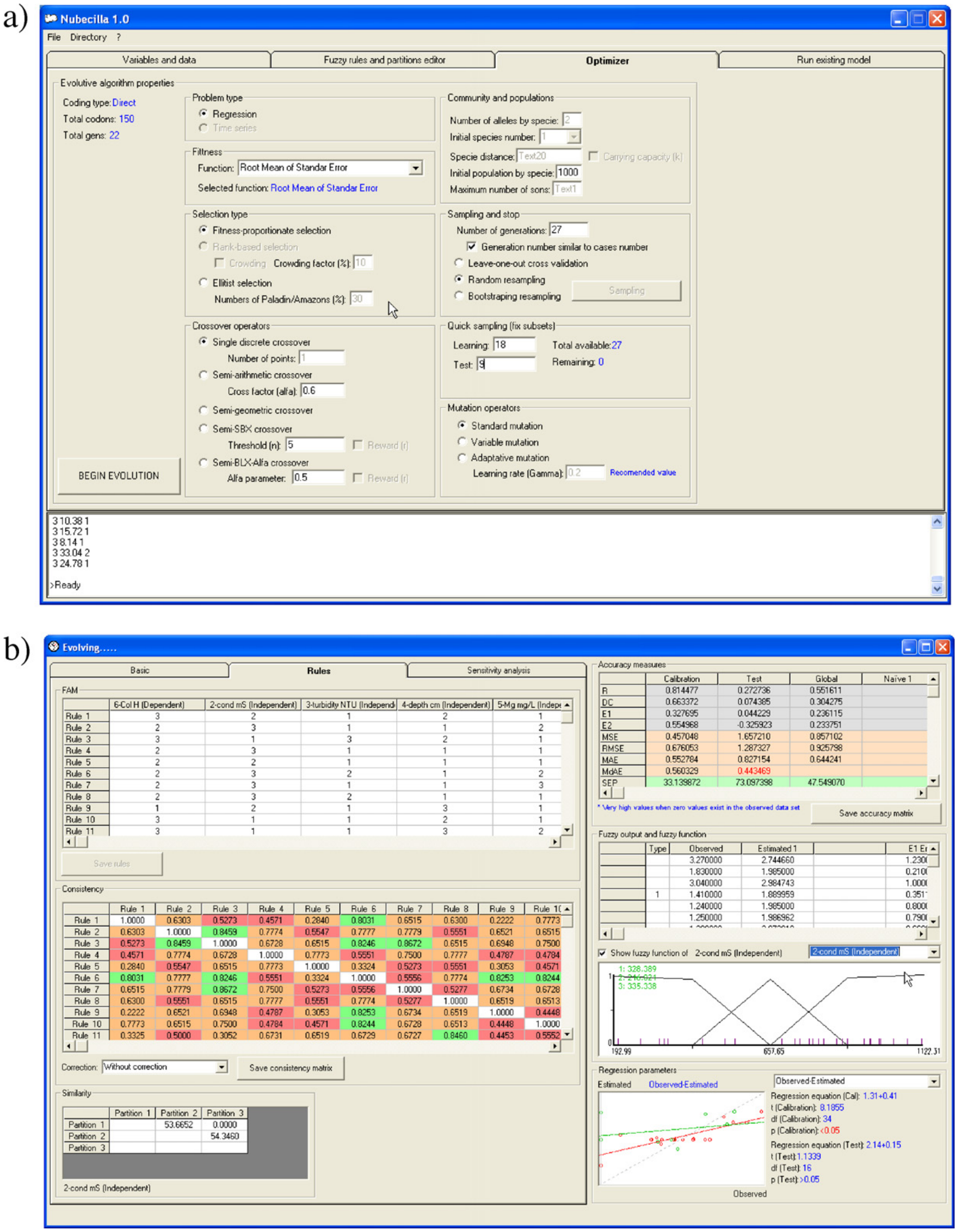


Fig. 2. Flow chart of the FIS model optimization with evolutionary algorithms.



Evolutive algorithm properties

Coding type: Direct
Total codons: 150
Total gens: 22

Problem type
● Regression
○ Time series

Fittness
Function: Root Mean of Standar Error
Selected function: Root Mean of Standar Error

Selection type
● Fitness-proportionate selection
○ Rank-based selection
☐ Crowding Crowding factor (%): 10
○ Ellitist selection
Numbers of Paladin/Amazons (%): 30

Crossover operators
● Single discrete crossover
Number of points: 1
○ Semi-arithmetic crossover
Cross factor (alfa): 0.6
○ Semi-geometric crossover
○ Semi-SBX crossover
Threshold (n): 5 ☐ Reward (r)
○ Semi-BLX-Alfa crossover
Alfa parameter: 0.5 ☐ Reward (r)

Community and populations
Number of alleles by specie: 2
Initial species number: 1
Specie distance: Text20 ☐ Carrying capacity (k)
Initial population by specie: 1000
Maximum number of sons: Text1

Sampling and stop
Number of generations: 27
☑ Generation number similar to cases number
○ Leave-one-out cross validation
● Random resampling
○ Bootstraping resampling

Quick sampling (fix subsets)
Learning: 18 Total available: 27
Test: 9 Remaining: 0

Mutation operators
● Standard mutation
○ Variable mutation
○ Adaptative mutation
Learning rate (Gamma): 0.2 Recomended value

```
3 10.38 1
3 15.72 1
3 8.14 1
3 33.04 2
3 24.78 1

>Ready
```

FAM

	6-Col H (Dependent)	2-cond mS (Independent)	3-turbidity NTU (Independ	4-depth cm (Independent)	5-Mg mg/L (Indepe
Rule 1	3	2	1	2	1
Rule 2	2	3	1	1	2
Rule 3	3	1	3	2	1
Rule 4	2	3	1	1	1
Rule 5	2	2	1	1	1
Rule 6	2	3	2	1	2
Rule 7	2	3	1	1	3
Rule 8	2	3	2	1	1
Rule 9	1	2	1	3	1
Rule 10	3	1	1	2	1
Rule 11	3	1	1	3	2

Consistency

	Rule 1	Rule 2	Rule 3	Rule 4	Rule 5	Rule 6	Rule 7	Rule 8	Rule 9	Rule 1
Rule 1	1.0000	0.6303	0.5273	0.4571	0.2840	0.8031	0.6515	0.6300	0.2222	0.7773
Rule 2	0.6303	1.0000	0.8459	0.7774	0.5547	0.7777	0.7779	0.5551	0.6521	0.6515
Rule 3	0.5273	0.8459	1.0000	0.6728	0.6515	0.8246	0.8672	0.6515	0.6948	0.7500
Rule 4	0.4571	0.7774	0.6728	1.0000	0.7773	0.5551	0.7500	0.7777	0.4787	0.4784
Rule 5	0.2840	0.5547	0.6515	0.7773	1.0000	0.3324	0.5273	0.5551	0.3053	0.4571
Rule 6	0.8031	0.7777	0.8246	0.5551	0.3324	1.0000	0.5556	0.7774	0.8253	0.8244
Rule 7	0.6515	0.7779	0.8672	0.7500	0.5273	0.5556	1.0000	0.5277	0.6734	0.6728
Rule 8	0.6300	0.5551	0.6515	0.7777	0.5551	0.7774	0.5277	1.0000	0.6519	0.6513
Rule 9	0.2222	0.6521	0.6948	0.4787	0.3053	0.8253	0.6734	0.6519	1.0000	0.4448
Rule 10	0.7773	0.6515	0.7500	0.4784	0.4571	0.8244	0.6728	0.6513	0.4448	1.0000
Rule 11	0.3325	0.5000	0.3052	0.6731	0.6519	0.6729	0.6727	0.8460	0.4453	0.5552

Correction: Without correction

Similarity

	Partition 1	Partition 2	Partition 3
Partition 1		53.6652	0.0000
Partition 2			54.3460
Partition 3			

2-cond mS (Independent)

Accuracy measures

	Calibration	Test	Global	Naïve 1
R	0.814477	0.272736	0.551611	
DC	0.663372	0.074385	0.304275	
E1	0.327695	0.044229	0.236115	
E2	0.554968	-0.325923	0.233751	
MSE	0.457048	1.657210	0.857102	
RMSE	0.676053	1.287327	0.925798	
MAE	0.552784	0.827154	0.644241	
MdAE	0.560329	0.443469		
SEP	33.139872	73.097398	47.549070	

* Very high values when zero values exist in the observed data set

Fuzzy output and fuzzy function

	Type	Observed	Estimated 1		E1 Er
		3.270000	2.744660		1.230
		1.830000	1.985000		0.210
		3.040000	2.984743		1.000
	1	1.410000	1.889959		0.351
		1.240000	1.985000		0.800
		1.250000	1.986962		0.790

☑ Show fuzzy function of 2-cond mS (Independent) 2-cond mS (Independent)

Regression parameters
Observed-Estimated
Regression equation (Cal): 1.31+0.41
t (Calibration): 8.1855
df (Calibration): 34
p (Calibration): <0.05
Regression equation (Test): 2.14+0.15
t (Test): 1.1339
df (Test): 16
p (Test): >0.05

Fig. 3. Two windows of the Nubecilla program: (a) different options of the evolutionary algorithm and sampling; (b) evolving window where the FAM, consistency and similarity matrixes, accuracy measures and fuzzy functions are shown.

2.4. General procedure

The general procedure employed for the calibration of FIS models is outlined in Fig. 4. Before the calibration of any FIS, ponds were grouped following the cluster analysis carried out by Gutiérrez-Estrada and Bilton (2010), which was based on contribution profiles of conductivity, turbidity, magnesium concentration and depth on the H' index. We selected the cluster that grouped ponds of both regions in an equilibrated way and had the most highly weighted contribution profile variable. This cluster (cluster 1 of conductivity profile contribution) grouped 23 Lizard and New Forest ponds (see Gutiérrez-Estrada and Bilton, 2010). The objective of this selection was to have a set of ponds which behaved as uniformly as possible in relation to H' and, at the same time, spanned both regions. In this

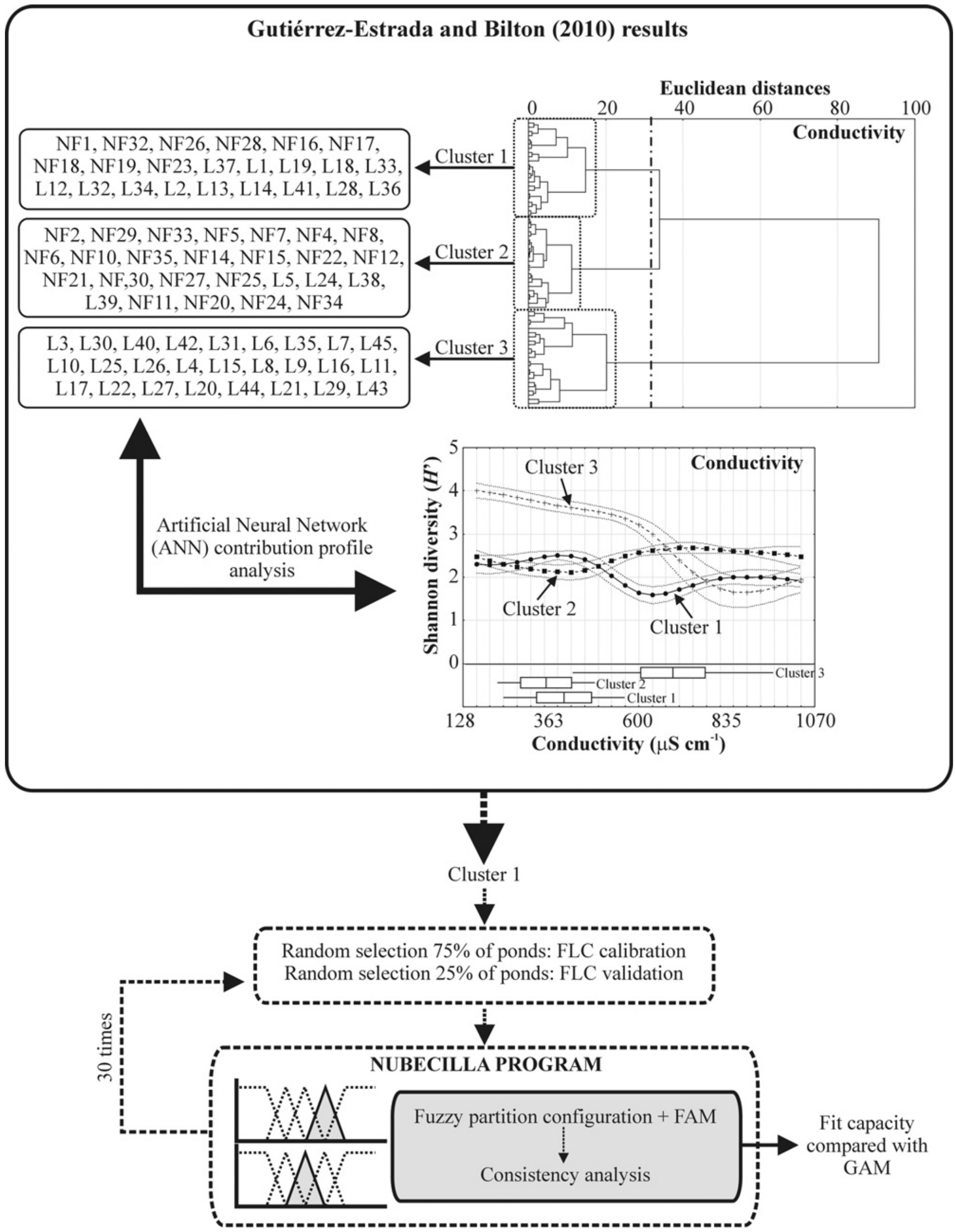


Fig. 4. Schematic representation of the general procedure followed for the calibration of FIS models.

way, the ability of *CC* to discriminate amongst contingent rules can be evaluated.

Later, this group of ponds was randomly divided into two subsets: the first one (75% of data) was used for the FIS calibration; the second one (25% remaining, not employed to calculate the FIS parameters) was used to validate the model. During the calibration, the FIS configurations that provided the best accuracy measures for the validation set were retained. To mitigate the effect of the random selection of data in the calibration and validation sets, each FIS configuration was calibrated and validated 30 times. Therefore, a similar procedure to that followed to calibrate other heuristics was adopted in this work.

In order to test the coherency of the results obtained with the FISs, the fit capacity of each FIS was compared with a generalized additive model (GAM) (Hastie and Tibshirani, 1990). In this sense, six accuracy measures were calculated in the calibration and validation phases of each FIS and GAM models examined: the Pearson's correlation coefficient (R), the determination coefficient (r^2), the square root of the mean square error (RMSE), the mean absolute error (MAE), the standard error of prediction (SEP), and the efficiency coefficient (E_2) (Nash and Sutcliffe, 1970; Kitanidis and Bras, 1980; Ventura et al., 1995; Legates and McCabe, 1999). The values of accuracy measures obtained with the GAM model were used as benchmark levels. In this way, the performance of a particular FIS was considered coherent when all accuracy measures associated with this FIS were better than the benchmark level given by the GAM model. Later, an ensemble FIS model was made using all calibrated models and those FISs that obtained better values for each accuracy measures during testing than the benchmark levels establish by GAM model.

3. Results

A total of 68 Coleoptera species were recorded across the two regions, and 35% of the species were of high conservation

Table 1
Average values of accuracy measures of FISs[a] in the calibration and test phases. Calibration + test = all data.

Accuracy measure	All FLC models			Best FLC models[b]			Generalized additive model[c]		
	Calibration	Test	Calibration + test	Calibration	Test	Calibration + test	Calibration	Test	Calibration + test
R	0.8169	0.7981	0.8188	0.8564	0.8721	0.8543	0.7939	0.4538	0.7475
r^2	0.6697	0.6618	0.6745	0.7335	0.7605	0.7300	0.6303	0.2059	0.5588
RMSE	0.3575	0.3360	0.3605	0.3474	0.2721	0.3325	0.3946	0.4866	0.4163
MAE	0.2870	0.2530	0.2796	0.2709	0.2491	0.2661	0.3514	0.3823	0.3581
SEP (%)	16.4464	15.0944	16.5228	15.7156	13.1068	15.2407	17.9194	22.9809	19.0659
E_2	0.6442	0.6041	0.6580	0.7059	0.7409	0.7143	0.6209	0.2042	0.5525

[a] Non-symmetric multipoint membership functions.

[b] Only models with all accuracy measures in the test phase better than the benchmark values.

[c] Gamma distribution of H'; link function = log; all degrees of freedom set to 3; final deviance = 0.9012.

importance (Bilton et al., 2009). In relation to Shannon's diversity index, the maximum values recorded were 3.968 on the Lizard and 3.471 in New Forest. As expected, some ponds lacked beetles (or contained a single species) resulting in $H' = 0$. In this way, although the maximum values in both regions were relatively high, the overall mean H' index was not higher than 2. On the other hand, none of the four environmental variables were significantly and linearly correlated with the H' index (H' index versus conductivity: $R = -0.21$, $p = 0.073$; H' index versus turbidity: $R = -0.02$; $p = 0.858$; H' index versus depth: $R = 0.16$, $p = 0.157$; H' index versus magnesium concentration: $R = 0.06$, $p = 0.630$). However, a clear statistical correlation was found between conductivity and magnesium concentration ($R = 0.78$, $p < 0.05$).

In relation to the ponds grouped in cluster 1, the maximum value of Shannon's diversity index recorded was 3.471, while the minimum value was 1.198. Among the environmental variables, a significant linear correlation was found between conductivity and magnesium concentration ($R = 0.42$, $p < 0.05$). This value was lower than that found for the whole data set. In contrast the correlations calculated between environmental variables and the H' index were clearly higher than those for the whole data set and significant in the case of conductivity (H' index versus conductivity: $R = -0.54$, $p < 0.05$; H' index versus turbidity: $R = -0.14$; $p = 0.533$; H' index versus depth: $R = 0.21$, $p = 0.344$; H' index versus magnesium concentration: $R = 0.19$, $p = 0.384$).

For all configurations of FISs calibrated, the capacity of generalization was good although it was significantly high when non-symmetric multipoint functions were used (Table 1). Using this type of membership functions the explained variance and the efficiency in the test phase were slightly higher than 66% and 0.60, respectively. Also, the RMSE provided statistically significant values, which implied an average standard error of prediction of 15%. These results were significantly improved when those FISs with a low convergence level (correlation values lower than 0.6 in the test phase) were rejected. In this way, the determination coefficient in the test phase increased 10 points ($r^2 = 0.7605$) and the standard error was improved 2% points (SEP = 13.1068%). Likewise, remaining accuracy measures provided better results. The RMSE and MAE decreased to 0.2721 and 0.2491 respectively, and the efficiency coefficient increased to 0.7059. These results were similar to those found by Gutiérrez-Estrada and Bilton (2010) using ensemble artificial neural networks and clearly better than the benchmark levels determined by a generalized additive model (Table 1).

In particularly, the best individual FIS explained in the test phase more than 76% of data variance ($R = 0.8769$; $r^2 = 0.7689$) and provided a standard error of prediction lower than 13% (SEP = 12.6445%) which was the lowest RMSE and MAE for all FISs calibrated (RMSE = 0.2676; MAE = 0.2474). In this model, the dependence relationships between H' and the three independent variables (conductivity, turbidity and magnesium concentration) was different, although the specific weights of each independent variable were around 25% (Table 2).

In this FIS, the level of variation of H' as a function of conductivity, turbidity, depth and magnesium concentration was only regulated by 18 rules from a maximum of 81 (3 premises for conductivity × 3 premises for turbidity × 3 premises for depth × 3 premises for magnesium concentration) (Table 3). Only two (rule 1 and 11) of these 18 rules were activated in almost 25% of cases, and 50% of data were associated to 5 rules and 75% of data fired ten rules. In addition the five most activated rules covered the complete range of variation seen in H' (two rules when H' = 'Low'; two rules when H' = 'Normal'; and one rule when H' = 'High'). On the other hand, the rule structure was incomplete because 'Normal level' was not found for turbidity, and here more than 90% of cases were grouped into the premise 'Low level'. In relation to the remaining independent variables, depth (ranked as the most important variable by the sensitivity analyses) showed the most equilibrated structure (37% of cases were grouped into the premise 'Low level', 45% of cases into the premise 'Normal level' and the remaining, 18%, into the premise 'High level').

The consistencies associated with the rule base of the best FIS (see Table 3) are shown in Tables 4 and 5. In these tables those rules with the same consequents and rules with very similar premises can be identified. Rules were considered to have similar premises when they were different in only one of them (for example, see premises of rules 1 and 11 in Table 3). In contrast, rules were considered to have very different premises when the sum of their absolute differences was at least equal to 5 (for example see premises of rules 8 and 18). Table 4 shows *CC* values without considering the effects of the sensitivity analysis. On the other hand, Table 5 shows the values of consistency taking into account the FIMP index.

Also contingent rules (rules with equal consequents and very different premises and rules with similar premises and very different consequents) can be found. In Tables 4 and 5, these rules are

Table 2
Sensitivity analysis of the FIS[a] with the best accuracy measures in the test phase.

	Conductivity (μS cm^{-1})	Turbidity (NTU)	Depth (cm)	Mg (mg L^{-1})
FIPM (ranking)	6.7776 (2)	6.7775 (3)	7.2667 (1)	6.5648 (4)
Specific weight (%)	24.7479	24.7475	26.5338	23.9708

[a] Non-symmetric multipoint membership functions.

Table 3

Fuzzy associative memory (rule base) of the FIS[a] with the best accuracy measures in the test phase. Dependent and independents variables are composed by three premises: 'Low level' = 1, 'Normal level' = 2, 'High level' = 3. The way to read this rule base is the following: Rule 1 = IF conductivity is 'Low' and turbidity is 'Low' and depth is 'Normal' and magnesium concentration is 'Low' THEN H' is 'Normal'. Rules sorted in function of number of times that each rule is activated (expressed in percentage). The last column shows the accumulated percentage.

Rule	H'	Conductivity	Turbidity	Depth	Magnesium concentration	Fired (%)	Accumulated fired (%)
Rule 1	2	1	1	2	1	11.7	11.7
Rule 11	2	1	1	2	2	11.7	23.4
Rule 7	3	1	1	1	1	8.8	32.2
Rule 16	1	1	1	1	2	7.3	39.5
Rule 15	1	2	1	2	1	7.3	46.8
Rule 13	2	1	1	3	2	5.8	52.6
Rule 14	2	2	1	1	1	5.8	58.4
Rule 10	2	2	1	1	3	5.8	64.2
Rule 2	3	1	1	3	1	5.8	70.0
Rule 8	3	2	1	2	2	5.8	75.8
Rule 9	1	3	1	2	3	4.4	80.2
Rule 17	2	3	1	1	2	3.5	83.7
Rule 6	1	1	3	1	1	2.9	86.6
Rule 5	1	1	3	2	1	2.9	89.5
Rule 18	1	2	1	1	2	2.9	92.4
Rule 3	1	2	1	3	1	2.9	95.3
Rule 4	1	3	1	3	1	2.9	98.2
Rule 12	2	1	3	2	2	1.8	100.0

[a] Non-symmetric multipoint membership functions.

marked with 'e' and 'f' superscripts, respectively. In relation to rules with very different consequents and very similar premises, the values of *CC* were lower than 0.25. In this case it was possible to find 8 contingent rules (rules 7 and 16; 8 and 15–18; 2 and 3). On the other hand, those rules with the same consequents and very different premises (superscript 'f') had values of *CC* lower than 0.28. In this case the rules implied were: rules 3 and 6–16; 4 and 5–6–16; 5 and 9–16–18; 6 and 9, 10 and 12; 12 and 14–17. Between these, the rules 5–18, 12–17, 10–12 and 6–9 had *CC* values significantly low.

Rules 6 and 9 were associated with a *CC* value of 0.0133. Therefore, they were the most contingent rules of the system. Rule 6 established that H' is 'Low' if conductivity is 'Low' and turbidity is 'High' and depth is 'Low' and magnesium concentration is 'Low'. Instead rule 9 predicted that H' is 'Low' if conductivity is 'High' and turbidity is 'Low' and depth is 'Normal' and magnesium concentration is 'High'. These two rules were fired in 2.9% (rule 6) and 4.4% (rule 9) of cases. On the other hand, rules 3–12, 4–12, 5–10 and 5–17 showed significantly high values of *CC* which is associated with highly consistent rules. For example rule 5 established that H' is 'Low' if conductivity is 'Low' and turbidity is 'High' and depth is 'Normal' and magnesium concentration is 'Low', and instead rule 17 indicated that H' is 'Normal' if conductivity is 'High' and turbidity is 'Low' and depth is 'Low' and magnesium concentration is 'Normal'. In this case the premises were rather different and the consequents were 'Low' and 'Normal', respectively. *A priori*, this is concordant with the assumption that two rules are consistent if they have different premises and consequents, which is in good agreement with the philosophy of fuzzy set theory.

Table 6 shows the consistencies calculated using the index proposed by Jin et al. (1999). In this matrix it is clear that some consistencies cannot be calculated because for these rules $\mathrm{SRC}(i,k) = 0$ or $\mathrm{SRP}(i,k) = 0$ (SRP = similarity of rule premises defined by Jin et al., 1999). When $\mathrm{SRC}(i,k) > 0$ and $\mathrm{SRP}(i,k) > 0$ this index provided values between 0.7369 and 1. The minimum consistency value (0.7369) was only calculated for the rules 1–2. Rule 1 predicted that H' is 'Normal' if conductivity is 'Low' and turbidity is 'Low' and depth is 'Normal' and magnesium concentration is 'Low'. Instead rule 2 indicated that higher H' would be reached if conductivity is 'Low' and turbidity is 'Low' and depth is 'High' and magnesium concentration is 'Low'. For the remaining rules, the consistency values oscillated between 0.8835 and 1, indicating a high level of coherency between them.

4. Discussion

In this work the coherence of the rules that control coleopteran diversity in a series of temporary and fluctuating ponds was analyzed using a continuous consistency index based on the similarity of consequents and premises, and the weight of environmental parameters. The rule set was obtained using a regressive fit between coleopteran diversity and environmental parameters, which was carried out with FISs optimized with evolutionary algorithms.

As Gutiérrez-Estrada and Bilton (2010) previously concluded, results indicate that water beetle diversities are driven by physical and chemical parameters and that the relationships between these variables and coleopteran diversity are clearly non-linear. Our study also highlights the fact that, in spite of in fact that analyses have been carried out with a homogeneous data set, contingent rules can be found, and that these may represent the different and subtle ways that water beetles respond to changes in environmental gradients.

The FIS approach applied in this study resulted in models which, in the test phase, explained >76% of the variance in diversity observed across ponds. This level of explained variance is in line with previous applications of other heuristic techniques like artificial neural networks (ANNs) in other systems. For example Lek-Ang et al. (1999) applied ANNs to relate the diversity of wetland Collembola (expressed as Shannon index) to environmental factors. In this work, the authors explained 79% of the variation of springtail diversity, clearly demonstrating the fit power of ANNs versus more classical approaches such as multiple linear regression. Likewise remaining accuracy measures obtained with FISs in the test phase were clearly better than the benchmark values established by a GAM model, demonstrating the power of this approach in predicting and understanding community–environment relationships in these systems.

An important part of the high efficiency of FIS was a consequence of the optimization method used. Genetic algorithms working with

Table 4

Consistency matrix without considering the effects of sensitivity analysis (in equation 2, ∀j $W_j = 1$). Green cells or 'a' superscripts are rules with consequents 'Low level'; oranges cells or 'b' superscripts are rules with consequents 'Normal level'; blue cells or 'c' superscripts are rules with consequents 'High level'. Black frames or 'd' superscripts are rules with very similar premises. Red frames or 'e' superscripts are rules with very different consequents and very similar premises. Bold numbers or 'f' superscripts are rules with the same consequent and very different premises. FIS calibrated with non-symmetric multipoint membership functions.

	Rule 1	Rule 2	Rule 3	Rule 4	Rule 5	Rule 6	Rule 7	Rule 8	Rule 9	Rule 10	Rule 11	Rule 12	Rule 13	Rule 14	Rule 15	Rule 16	Rule 17	Rule 18
Rule 1	1.0000	0.2596^{d}	0.4580	0.4800	0.2642	0.4875	0.2666^{d}	0.4734	0.5000	0.2866^{b}	0.7567^{b,d}	0.5067^{b}	0.5279^{b}	0.5366^{b}	0.2422^{d}	0.4811	0.2700^{b}	0.6950
Rule 2		1.0000	0.2269^{e}	0.2502	0.4788	0.5000	0.7500^{c}	0.3012^{c}	0.7286	0.7009	0.4752	0.6968	0.2724^{d}	0.4793	0.4555	0.4933	0.7156	0.7198
Rule 3			1.0000	0.7644^{a,d}	0.2945^{a}	**0.2733**^{a,f}	0.4767	0.4721	0.2856^{a}	0.5000	0.6875	0.9233	0.4717	0.2642	0.7712^{a,d}	**0.2800**^{a,f}	0.7159	0.5067^{a}
Rule 4				1.0000	**0.2712**^{a,f}	**0.2500**^{a,f}	0.5000	0.7075	0.5212^{a}	0.7223	0.7095	0.9453	0.4937	0.4864	0.5356^{a}	**0.2567**^{a,f}	0.4937	0.2711^{a}
Rule 5					1.0000	0.7633^{a,d}	0.4867	0.7198	**0.2500**^{a,f}	0.9372	0.4937	0.2578^{d}	0.7095	0.7013	0.5233^{a}	**0.2700**^{a,f}	0.9528	**0.0433**^{a,f}
Rule 6						1.0000	0.2502	0.9564	**0.0133**^{a,f}	0.7138	0.7170	0.4811	0.7295	0.4780	0.2866^{a}	0.5067^{a}	0.7295	0.2800^{a}
Rule 7							1.0000	0.2933^{c}	0.7366	0.4793	0.4823	0.7039	0.4940	0.2577^{d}	0.4635	0.2435^{e}	0.4940	0.4700
Rule 8								1.0000	0.4732	0.4773	0.2577^{d}	0.4793	0.4605	0.4823	0.2435^{e}	0.4635	0.4755	0.2369^{e}
Rule 9									1.0000	0.4739	0.4883	0.7241	0.7041	0.7098	0.5144^{a}	0.2757^{a}	0.4758	0.2900^{a}
Rule 10										1.0000	0.2990^{b}	**0.0490**^{b,f}	0.2857^{b}	0.7500^{b}	0.4875	0.4663	0.5268^{b}	0.2525^{d}
Rule 11											1.0000	0.7500^{b}	0.7712^{b,d}	0.2933^{b}	0.4717	0.2517^{d}	0.5133^{b}	0.4655
Rule 12												1.0000	0.5212^{b}	**0.0433**^{b,f}	0.7075	0.4875	**0.2633**^{b,f}	0.7013
Rule 13													1.0000	0.2800	0.6875	0.2642	0.5000^{b}	0.4780
Rule 14														1.0000	0.2517^{d}	0.4717	0.5211^{b}	0.2578^{d}
Rule 15															1.0000	0.2933^{a}	0.7034	0.5200^{a}
Rule 16																1.0000	0.2642	0.7733^{a,d}
Rule 17																	1.0000	0.2506^{d}
Rule 18																		1.0000

Table 5

Consistency matrix considering the effects of FIPM index. Green cells or ‘a’ superscripts are rules with consequents ‘Low level’; oranges cells or ‘b’ superscripts are rules with consequents ‘Normal level’; blue cells or ‘c’ superscripts are rules with consequents ‘High level’. Black frames or ‘d’ superscripts are rules with very similar premises. Red frames or ‘e’ superscripts are rules with very different consequents and very similar premises. Bold numbers or ‘f’ superscripts are rules with the same consequent and very different premises. FIS calibrated with non-symmetric multipoint membership functions.

	Rule 1	Rule 2	Rule 3	Rule 4	Rule 5	Rule 6	Rule 7	Rule 8	Rule 9	Rule 10	Rule 11	Rule 12	Rule 13	Rule 14	Rule 15	Rule 16	Rule 17	Rule 18
Rule 1	1.0000	0.2720^{d}	0.4691	0.4909	0.2618	0.4988	0.2795	0.4625	0.4879	0.2846^{b}	0.7667^{b,d}	0.5193^{b}	0.5239^{b}	0.5243^{b}	0.2400^{d}	0.4854	0.2680^{b}	0.6971
Rule 2		1.0000	0.2246^{e}	0.2477	0.4903	0.5128	0.7347^{c}	0.2995^{c}	0.7299	0.7034	0.4788	0.6982	0.2636^{d}	0.4909	0.4673	0.4986	0.7181	0.7228
Rule 3			1.0000	0.7668^{a,d}	0.2853^{a}	**0.2628**^{a,f}	0.4898	0.4761	0.2842^{a}	0.5048	0.6891	0.9226	0.4601	0.2786	0.7572^{a,d}	**0.2770**^{a,f}	0.7187	0.5014^{a}
Rule 4				1.0000	**0.2622**^{a,f}	**0.2397**^{a,f}	0.5128	0.7092	0.5175^{a}	0.7248	0.7109	0.9444	0.4818	0.4987	0.5239^{a}	**0.2539**^{a,f}	0.4987	0.2682^{a}
Rule 5					1.0000	0.7487^{a,d}	0.4987	0.7050	**0.2653**^{a,f}	0.9366	0.4818	0.2484^{d}	0.7109	0.7105	0.5281^{a}	**0.2680**^{a,f}	0.9523	**0.0436**^{a,f}
Rule 6						1.0000	0.2477	0.9561	**0.0141**^{a,f}	0.6996	0.7189	0.4854	0.7321	0.4735	0.2769^{a}	0.5193^{a}	0.7153	0.2949^{a}
Rule 7							1.0000	0.2911^{c}	0.7383	0.4682	0.4863	0.7056	0.4988	0.2557^{d}	0.4757	0.2335^{e}	0.4829	0.4577
Rule 8								1.0000	0.4611	0.4815	0.2557^{d}	0.4751	0.4710	0.4863	0.2335^{e}	0.4757	0.4863	0.2514^{e}
Rule 9									1.0000	0.4854	0.4767	0.7102	0.7058	0.7115	0.5271^{a}	0.2734^{a}	0.4803	0.2877^{a}
Rule 10										1.0000	0.2965^{b}	**0.0490**^{b,f}	0.2824^{b}	0.7603^{b}	0.4915	0.4549	0.5389^{b}	0.2432^{d}
Rule 11											1.0000	0.7525^{b}	0.7572^{b,d}	0.2911^{b}	0.4601	0.2654^{d}	0.5013^{b}	0.4770
Rule 12												1.0000	0.5097^{b}	**0.0436**^{b,f}	0.6935	0.4988	**0.2538**^{b,f}	0.7105
Rule 13													1.0000	0.2770	0.6891	0.2786	0.4872^{b}	0.4903
Rule 14														1.0000	0.2654^{d}	0.4601	0.5335^{b}	0.2484^{d}
Rule 15																0.2911^{a}	0.7054	0.5155^{a}
Rule 16																1.0000	0.2618	0.7756^{a,d}
Rule 17																	1.0000	0.2484^{d}
Rule 18																		1.0000

Table 6

Consistency matrix considering the consistency index proposed by Jin et al. (1999).[a] In these cases '—' the consistency cannot be calculated. Green cells or 'a' superscripts are rules with consequents 'Low level'; oranges cells or 'b' superscripts are rules with consequents 'Normal level'; blue cells or 'c' superscripts are rules with consequents 'High level'. Black frames or 'd' superscripts are rules with very similar premises. Red frames or 'e' superscripts are rules with very different consequents and very similar premises. Bold numbers or 'f' superscripts are rules with the same consequent and very different premises.

	Rule 1	Rule 2	Rule 3	Rule 4	Rule 5	Rule 6	Rule 7	Rule 8	Rule 9	Rule 10	Rule 11	Rule 12	Rule 13	Rule 14	Rule 15	Rule 16	Rule 17	Rule 18
Rule 1	1.0000	0.7369^{d}	1.0000	0.8465	---	1.0000	0.9998	0.9936	---	0.9977^{b}	0.9946^{b,d}	0.9946^{b}	0.9946^{b}	0.9977^{b}	0.9900^{d}	1.0000	0.9977^{b}	1.0000
Rule 2		1.0000	---^{e}	---	---	---	---^{c}	0.9946^{c}	---	0.9778	0.9936	0.9936	0.9936^{d}	0.9778	---	---	0.9936	---
Rule 3			1.0000	0.9977^{a,d}	0.9946^{a}	**0.9919**^{a,f}	---	---	0.9977^{a}	---	1.0000	1.0000	1.0000	---	0.9803^{a,d}	**0.9946**^{a,f}	1.0000	0.9946^{a}
Rule 4				1.0000	**0.9803**^{a,f}	**---**^{a,f}	---	---	0.9977^{a}	1.0000	0.9977	0.9977	0.9977	1.0000	0.9977^{a}	**0.9946**^{a,f}	0.9977	0.9977^{a}
Rule 5					1.0000	0.9946^{a,d}	---	---	**---**^{a,f}	1.0000	0.9977	0.9977^{d}	0.9977	1.0000	0.9919^{a}	**0.9977**^{a,f}	1.0000	**0.9977**^{a,f}
Rule 6						1.0000	---	---	**0.9977**^{a,f}	1.0000	1.0000	1.0000	1.0000	1.0000	0.9977^{a}	0.9946^{a}	0.9977	0.9946^{a}
Rule 7							1.0000	0.9975^{c}	---	0.9778	0.9998	0.9998	0.9936	0.9778^{d}	---	---^{e}	0.9936	---
Rule 8								1.0000	---	0.9998	0.9776^{d}	0.9776	0.9778	0.9998	---^{e}	---	0.9998	---^{e}
Rule 9									1.0000	1.0000	0.8835	0.8835	0.8835	1.0000	0.9977^{a}	0.9977^{a}	1.0000	0.9977^{a}
Rule 10										1.0000	0.9977^{b}	**0.9977**^{b,f}	0.9919^{b}	---^{b}	1.0000	1.0000	0.9977^{b}	0.8835^{d}
Rule 11											1.0000	---^{b}	0.9803^{b,d}	0.9977^{b}	1.0000	1.0000^{d}	0.9977^{b}	1.0000
Rule 12												1.0000	0.9803^{b}	**0.9977**^{b,f}	1.0000	1.0000	**0.9977**^{b,f}	1.0000
Rule 13													1.0000	0.9946	1.0000	---	---	0.9900
Rule 14														1.0000	1.0000^{d}	1.0000	0.9977^{b}	0.9977^{d}
Rule 15															1.0000	0.9977^{a}	1.0000	0.9977^{a}
Rule 16																1.0000	---	0.9999^{a,d}
Rule 17																	1.0000	1.0000^{d}
Rule 18																		1.0000

[a] Consistency $(R_i, R_k) = \exp\left(-\frac{(((\mathrm{SRP}(i,k))/(\mathrm{SRC}(i,k))-1)^2}{(1/\mathrm{SRP}(i,k))^2}\right)$; $\mathrm{SRP}(i, k) = \min_{j=1,2\ldots n} S(A_{ij}, A_{kj})$; $\mathrm{SRC}(i, k) = S(B_i, B_k)$.

chromosomes of variable length and coding dependent on the context (all of which is possible using the Nubecilla program), allow rapid convergence towards a compact base of fuzzy rules. In this way the data structure analyzed in this work indicated that coleopteran diversity in this system was controlled by only 18 rules. From these, only four rules (rules 1, 2, 11 and 7) were associated with more than 38% of data set, establishing the basic response of this community in relation to the environmental variables analyzed. For these four rules, the model predicted that the coleopteran diversity was 'Normal' or 'High' if conductivity was 'Low', turbidity was 'Low', depth was 'Normal', 'Low' or 'High' and magnesium concentration was 'Low' or 'Normal', which broadly corresponds to commonly observed responses of lentic macroinvertebrate communities to these environmental variables.

The influence of conductivity on macroinvertebrate composition is broadly documented (Kapoor, 1978; Lemly, 1982; Williams et al., 1997; Williams and Williams, 1998; Blasius and Merritt, 2002; Biggs et al., 2005; De Jonge et al., 2008). In the majority of these studies a negative relationship between conductivity and richness was documented, which is driven by the osmotic challenges most freshwater invertebrates face in waters with high ion concentrations (Macan, 1974; Blasius and Merritt, 2002). Several studies have reported that densities and diversity of macroinvertebrates are negatively related with turbidity (Hentges and Stewart, 2010; Hopkins et al., 2011), and likewise depth has been identified as an important direct factor influencing macroinvertebrate community responses (Muehlbauer et al., 2011), although here responses could also be related with other co-varying parameters such as the duration of the dry phase (Nicolet et al., 2004). In relation to magnesium concentration, it is difficult to provide a direct functional explanation, particularly bearing in mind the fact that magnesium concentration was significantly correlated with conductivity. However, as concluded Gutiérrez-Estrada and Bilton (2010), proxy effects related with total organic nitrogen (TON) may be important for this parameter.

Rules 3, 4, 9 and 18 complemented the basic response. These rules predicted in an average way that 'Low' levels of diversities were found when conductivity was 'High', turbidity was 'Low', depth was 'Normal-High' and magnesium concentration was 'Normal', which was coherent with rules 1, 2, 7 and 11. On the other hand, the remaining rules seem to be consequences of alternatives strategies that reflect the variety of possible response patterns to environmental variations. Several studies have reported that the animal community found in an individual pond can be significantly different from that in an adjacent pond, despite similar environmental conditions (Jeffries, 1989; Jenkins and Buikema, 1998), such differences perhaps reflecting the contingency of colonization and 'monopolization' by early colonists (De Meester et al., 2002). Jeffries (2003) used logistic regressions to analyze the relationships between pond invertebrates and several environmental factors, and reported that all variables were significant predictors of presence–absence of the different species studied for all models tested, but that no one environmental factor occurred in all models. Jeffries concluded that different species showed alternatively positive or negative relationships to the same environmental factor, which should be reflected in any model containing different predictors of diversity. Likewise, in our study the different diversity patterns provided several groups of coherent rules associated with different response patterns of H' to the environmental factors analyzed.

Rules 5–18, 6–9, 10–12 and 12–14 showed very low values of consistency, denoting high contingency levels. The interactions of these rules in a group of ponds with a relatively homogeneous response versus coleopteran diversity suggest the existence of different environmental scenarios, determined by interactions among many factors, rather than that one single dimension controls the dynamic of this system. This is coherent with one of the principal thesis of alternative state models theory which establishes that a system can shift abruptly between two or more states as a consequence of minor perturbations in environmental conditions (Scheffer et al., 2001). Bilton et al. (2009) reported that ponds in Lizard and New Forest regions varied substantially in several biotic and environmental factors like area, depth and vegetation composition, but all were relatively small water bodies, something which favours low levels resiliency. Therefore, these rules could be associated with different hysteresis effects in the collapse or recovery phases as a consequence of the non-linear behaviour of coleopteran diversity in temporary and fluctuating ponds.

The consistency index proposed in this work has shown its ability to highlight contingent fuzzy rules, even when the rule base has been obtained from an apparently homogeneous data base. Highly contingent rules take values close to zero whilst rules with a good level of coherence take values close to one. This information combined with the percentage of rules fired can help detect different behaviours of the estimated variable in relation to the explicatory variables. Likewise, the *CC*'s formulation allows the weight of the variables to be incorporated into the system, which facilitates the interpretation of the rule in the context analyzed. Although some works have dealt the problem of the consistencies of a set of fuzzy rules (Jin et al., 1999; Pedrycz, 2003; Alonso et al., 2008), to our knowledge no approach in an ecological context exists in the literature. In addition, previous works have not considered the contingent concept, which implies that for some rules the consistency cannot be calculated, or is barely evaluable. For example, the consistency index proposed by Jin et al. (1999) cannot be calculated when the similarity between consequents is zero, or it always provides high consistency levels when the similarity of premises is very low, independently of the similarity of consequents.

5. Conclusions

In summary, this study has highlighted the non-linear nature of the relationship between water beetle diversity and environmental variables in a set of ponds from two regions in the South of England, and demonstrates the strong performance of fuzzy inference systems in modelling diversity–environment interactions. FIS's generalization capacity (explaining more than 76% of the variance in diversity) suggest that, as with other heuristic techniques, it can be used to simulate the response of organismal diversity under different environmental scenarios. Also FISs have shown that a compact and interpretable rule set can be extracted from meta-analysis of a data set. Finally the consistency index proposed in this work, combined with the percentage of rules fired, allows the identification of contingent responses which could illuminate adaptation mechanisms of pond communities to environmental changes.

Acknowledgements

Much of the fieldwork which generated the data base of this work was conducted under a PhD studentship undertaken by Louise McAbendroth. We are grateful to Natural England and the University of Plymouth for financial support for this project.

References

Addriaenssens, V., De Baets, B., Goethals, P.L.M., De Pauw, N., 2004. Fuzzy rule-based models for decision support in ecosystem management. The Science of the Total Environment 319, 1–12.

Alonso, J.M., Magdalena, L., Guillaume, S., 2008. HILK: a new methodology of designing highly interpretable linguistic knowledge bases using the fuzzy logic formalism. International Journal of Intelligent Systems 23, 761–794.

Baraldi, P., Librizzi, M., Zio, E., Podofillini, L., Dang, V.N., 2009. Two techniques of sensitivity and uncertainty analysis of fuzzy expert systems. Expert Systems with Applications 36, 12461–12471.

Biggs, J., Williams, P., Whitfield, P., Nicolet, P., Weatherby, A., 2005. 15 years of pond assessment in Britain: results and lessons learned from the work of pond conservation. Aquatic Conservation: Marine and Freshwater Ecosystems 15, 693–714.

Bilton, D.T., Foggo, A., Rundle, S.D., 2001. Size, permanence and the proportion of predators in ponds. Archiv für Hydrobiologie 151, 451–458.

Bilton, D.T., McAbendroth, L.M., Bedford, A., Ramsay, P.M., 2006. How wide to cast the net? Cross-taxon congruence of species richness, community similarity and indicator taxa in ponds. Freshwater Biology 51, 578–590.

Bilton, D.T., McAbendroth, L.C., Nicolet, P., Bedford, A., Rundle, S.D., Foggo, A., Ramsay, P.M., 2009. Ecology and conservation status of temporary and fluctuating ponds in two areas of southern England. Aquatic Conservation: Marine and Freshwater Ecosystems 19, 134–146.

Blasius, B.J., Merritt, R.W., 2002. Field and laboratory investigations on the effects of road salt (NaCl) on streams macroinvertebrate communities. Environmental Pollution 120, 219–231.

Bosserman, R.W., Ragade, R.K., 1982. Ecosystem analysis using fuzzy set theory. Ecological Modelling 16, 191–208.

Brower, J.E., Zar, J.H., von Ende, C.N., 1998. Field and Laboratory Methods for General Ecology. McGraw and Hill, Dubuque, IA.

Cheong, F., Lai, R., 2000. Constraining the optimization of a fuzzy logic controller using an enhanced genetic algorithm. IEEE Transactions on System, Man and Cybernetics Part B: Cybernetics 30, 31–46.

Cheung, W.W.L., Pitcher, T.J., Pauly, D., 2005. A fuzzy logic expert system to estimate intrinsic extinction vulnerabilities of marine fishes to fishing. Biological Conservation 124, 97–111.

Cowell, B.C., Remley, A.H., Lynch, D.M., 2004. Seasonal changes in the distribution and abundance of benthic invertebrates in six headwater streams in central Florida. Hydrobiologia 522, 99–115.

De Jonge, M., Van de Vijver, B., Blust, R., Bervoets, L., 2008. Responses of aquatic organisms to metal pollution in a lowland river in Flandes: a comparison of diatoms and macroinvertebrates. Science of the Total Environment 407, 615–629.

De Meester, L., Gomez, A., Okamura, B., Schwenk, K., 2002. The monopolization hypothesis and the dispersal-gene flow paradox in aquatic organisms. Acta Oecologica 23, 121–135.

Foggo, A., Rundle, S.D., Bilton, D.T., 2003. The net result: evaluating species richness extrapolations techniques for pond invertebrates. Freshwater Biology 48, 1–9.

Gacto, M.J., Alcalá, R., Herrera, F., 2011. Interpretability of linguistic fuzzy rule-based systems: an overview of terpretability measures. Information Sciences 181, 4340–4360.

Goldberg, D., 1989. Genetic Algorithms in Search, Optimization and Machine Learning. Addison Wesley, Reading, MA, USA.

Gutiérrez-Estrada, J.C., Bilton, D.T., 2010. A heuristic approach to predicting water beetle diversity in temporary and fluctuating water. Ecological Modelling 221, 1451–1462.

Hastie, T.J., Tibshirani, R.J., 1990. Generalized Additive Models. Chapman & Hall/CRC, Boca Raton.

Hentges, A.V., Stewart, T.W., 2010. Macroinvertebrate assemblages in Iowa prairie pothole wetlands and relation to environmental features. Wetlands 30, 501–511.

Holland, J., 1975. Adaptation in Natural and Artificial Systems. MIT Press, Cambridge, MA, USA.

Hopkins, J.M., Marcarelli, A.M., Bechtold, H.A., 2011. Ecosystem structure and function are complementary measures of water quality in a polluted, spring-influenced river. Water, Air and Soil Pollution 214, 409–421.

Jeffries, M.J., 1989. Measuring Talling's element of chance in pond populations. Freshwater Biology 21, 383–393.

Jeffries, M.J., 2003. Idiosyncratic relationships between pond invertebrates and environmental, temporal and patch-specific predictors of incidence. Ecography 26, 311–324.

Jenkins, D.G., Buikema, A.L., 1998. Do similar communities develop in similar sites? A test with zooplankton structure and function. Ecological Monograph 68, 421–443.

Jin, Y., Von Seelen, W., Sendhoff, B., 1999. On generating FC3 fuzzy rule systems from data using evolution strategies. IEEE Transactions of Systems, Man, and Cybernetics 29, 829–845.

Kampichler, C., Barthel, J., Wieland, R., 2000. Species density of foliage-dwelling spiders in field margins: a simple, fuzzy rule-based model. Ecological Modelling 129, 87–99.

Kapoor, N.N., 1978. Effect of salinity on the osmoregulatory cells in the tracheal gills of the stonefly nymph, *Paragnetia media* (Plecoptera: Perlidae). Canadian Journal of Zoology 56, 2608–2613.

Kitanidis, P.K., Bras, R.L., 1980. Real time forecasting with a conceptual hydrological model. 2. Applications and results. Water Resources Research 16 (6), 1034–1044.

Legates, D.R., McCabe Jr., G.J., 1999. Evaluating the use of 'goodness-of-fit' measures in hydrologic and hydroclimatic model validation. Water Resources Research 35 (1), 233–241.

Lek-Ang, S., Deharverng, L., Lek, S., 1999. Predictive models of collembolan diversity and abundance in a riparian habitat. Ecological Modelling 120, 247–260.

Lemly, D.A., 1982. Modification of benthic insect communities in polluted streams: combined effects of sedimentation and nutrient enrichment. Hydrobiologia 87, 229–245.

Macan, T.T., 1974. Freshwater Ecology, 2nd ed. Longman Group, London.

Meesters, E.H., Bak, R.P.M., Westmacott, S., 1998. A Fuzzy logic model to predict coral reef development under nutrient and sediment stress. Conservation Biology 12 (5), 957–965.

Mouton, A.M., De baets, B., Van Broekhoven, E., Goethals, P.L.M., 2009. Prevalence-adjusted optimisation of fuzzy models for species distribution. Ecological Modelling 15, 1776–1786.

Muehlbauer, J.D., Doyle, M.W., Bernhardt, E.S., 2011. Macroinvertebrate community responses to a dewatering disturbance gradient in a restores stream. Hydrology and Earth System Sciences 15, 1771–1783.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models. I. A discussion of principles. Journal of Hydrology 10, 282–290.

Nicolet, P., Biggs, J., Fox, G., Hodson, M.J., Reynolds, C., Whitfield, M., Williams, P., 2004. The wetland plant and macroinvertebrate assemblages of temporary ponds in England and Wales. Biological Conservation 120, 261–278.

Pappis, C.P., Ebrahim, H., Mamdani, E.H., 1977. Fuzzy logic-controller for a traffic junction. IEEE Transactions of Systems, Man, and Cybernetics 7 (10), 707–717.

Pedrycz, W., 2003. Expressing relevance interpretability and accuracy of rule-based systems. In: Casillas, J., Cordón, O., Herrera, F., Magdalena, L. (Eds.), Interpretability Issues in Fuzzy Modeling. Springer-Verlag, Berlin, Heidelberg/New York.

Rundle, S.D., Foggo, A., Choisel, V., Bilton, D.T., 2002. Are distribution patterns linked to dispersal mechanism? An investigation using pond invertebrate assemblages. Freshwater Biology 4, 1571–1581.

Sánchez-Fernández, D., Abellán, P., Mellado, A., Millán, A., Velasco, J., 2006. Are water beetles good indicators of biodiversity in Mediterranean aquatic ecosystems? The case of the Segura river basin (SE Spain). Biodiversity and Conservation 15, 4507–4520.

Scheffer, M., Carpenter, S., Foley, J.A., Folke, C., Walker, B., 2001. Catastrophic shifts in ecosystems. Nature 413, 591–596.

Schmitz, O.J., 2010. Resolving Ecosystem Complexity. Monographs in Population Biology, vol. 47. Princeton University Press, Oxfordshire, USA.

Ventura, S., Silva, M., Pérez-Bendito, D., Hervás, C., 1995. Artificial neural networks for estimation of kinetic analytical parameters. Analytical Chemistry 67 (9), 1521–1525.

Williams, D.D., Williams, N.E., Cao, Y., 1997. Spatial differences in macroinvertebrates community structure in springs in southeastern Ontario in relation to their chemical and physical environments. Canadian Journal of Zoology 75, 1404–1414.

Williams, D.D., Williams, N.E., 1998. Aquatic insects in an estuarine environment: densities, distribution and salinity tolerance. Freshwater Biology 39, 411–421.

Yu, J.Z., Tan, M., Wang, S., 2004. Development of a biomimetic robotic fish and its control algorithm. IEEE Transactions of Systems, Man, and Cybernetics Part B: Cybernetics 34 (4), 1798–1810.

Zadeh, L.A., 1973. Outline of a new approach to the analysis of complex systems and decision processes. IEEE Transactions of Systems, Man, and Cybernetics 2, 28–44.

Zeldis, D., Prescott, S., 2000. Fish disease diagnosis program-problems and some solutions. Aquacultural Engineering 23, 3–11.