G Model
ENVSCI 1886 No. of Pages 11

ARTICLE IN PRESS

Environmental Science & Policy xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Environmental Science & Policy

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

Fuzzy modelling to identify key drivers of ecological water quality to support decision and policy making

Marie Anne Eurie Forio[a,*], Ans Mouton[a], Koen Lock[a], Pieter Boets[a,b], Nguyen Thi Hanh Tien[a,c], Minar Naomi Damanik Ambarita[a], Peace Liz Sasha Musonge[a], Luis Dominguez-Granda[d], Peter L.M. Goethals[a]

[a] *Laboratory of Environmental Toxicology and Aquatic Ecology, Ghent University, Coupure Links 653, B-9000 Ghent, Belgium*
[b] *Provincial Centre of Environmental Research, Godshuizenlaan 95, B-9000 Ghent, Belgium*
[c] *Research Institute for Aquaculture No1, Dinh Bang, Tu Son, Bac Ninh, Vietnam*
[d] *Facultad de Ciencias Naturales y Matemáticas, Centro del Agua y Desarrollo Sustentable, Campus Gustavo Galindo, Escuela Superior Politécnica del Litoral (ESPOL) Km. 30.5 Vía Perimetral, P.O. Box 09-01-5863, Guayaquil, Ecuador*

A R T I C L E I N F O

Article history:
Received 14 December 2015
Received in revised form 5 December 2016
Accepted 6 December 2016
Available online xxx

Keywords:
Fuzzy logic
Environmental variables
Decision support systems
River basin management

A B S T R A C T

Water quality modelling is an effective tool to investigate, describe and predict the ecological state of an aquatic ecosystem. Various environmental variables may simultaneously affect water quality. Appropriate selection of a limited number of key-variables facilitates cost-effective management of water resources. This paper aims to determine (and analyse the effect of) the major environmental variables predicting ecological water quality through the application of fuzzy models. In this study, a fuzzy logic methodology, previously applied to predict species distributions, was extended to model environmental effects on a whole community. In a second step, the developed models were applied in a more general water management context to support decision and policy making. A hill-climbing optimisation algorithm was applied to relate ecological water quality and environmental variables to the community indicator. The optimal model was selected based on the predictive performance (Cohen's Kappa), ecological relevance and model's interpretability. Moreover, a sensitivity analysis was performed as an extra element to analyse and evaluate the optimal model. The optimal model included the variables land use, chlorophyll and flow velocity. The variable selection method and sensitivity analysis indicated that land use influences ecological water quality the most and that it affects the effect of other variables on water quality to a high extent. The model outcome can support spatial planning related to land use in river basins and policy making related to flows and water quality standards. Fuzzy models are transparent to a wide range of users and therefore may stimulate communication between modellers, river managers, policy makers and stakeholders.

1. Introduction

Deterioration of water quality has become an issue of global concern (UN, 2014). The expansion of industries, agriculture, and tourism often causes water quality degradation (Goethals and Volk, 2016; Teixeira et al., 2014). Impacts on surface and groundwater resources limit their use for drinking, bathing, industrial, or agricultural purposes (UN, 2014). Furthermore, polluted water threatens both quality of life and public health (Barnhoorn et al., 2015; Hauck et al., 2015). Thus, conservation and restoration of good water quality are important. Monitoring, assessment, modelling and implementation of appropriate management actions are necessary to avoid further deterioration of water quality (He et al., 2014).

Water quality modelling is an effective tool to investigate, describe and predict the ecological state of the aquatic ecosystem. In the last decade, several methods were applied to model environmental systems. One of the most promising modelling approaches is fuzzy models (Zadeh, 1965), which are based on fuzzy set theory (Akerkar and Sajja, 2010). An advantage of fuzzy models is its flexibility, transparency and user friendliness (Adriaenssens et al., 2006; Akerkar and Sajja, 2010). These models take into account the inherent uncertainty of ecological variables and allow the expression of non-linear relations between

* Corresponding author.
E-mail addresses: marie.forio@ugent.be, anne_forio@yahoo.com (M.A.E. Forio).

http://dx.doi.org/10.1016/j.envsci.2016.12.004

ecological variables. A fuzzy rule-based system was constructed that connects the input variables to the output by means of if-then rules. Fuzzy models use linguistic descriptions such as 'low', 'moderate' or 'high' for the quantification of variables and transform these descriptions into a mathematical framework in which data processing can be performed (Kampichler et al., 2000). As such, fuzzy models have the ability to convey a logical and reliable stream of information (Adriaenssens et al., 2004). Furthermore, fuzzy models enable the incorporation of ecological aspects such as the ecological gradient theory (Cadenasso et al., 2003; Yarrow and Salthe, 2008).

Most of the developed fuzzy models have been used for assessment of ecosystem sustainability (Canavese et al., 2014; Liu et al., 2012), environmental impact assessment (de Siqueira et al., 2006; Peche and Rodríguez, 2011), development of water quality index (Gharibi et al., 2012; Icaga, 2007), development of environmental quality index (Peche and Rodríguez, 2012), water quality assessment (Liu and Zou, 2012; Scannapieco et al., 2012) and habitat suitability and prediction of organisms' occurrence (Adriaenssens et al., 2006; Mouton et al., 2011). These studies illustrate the potential of fuzzy models in ecological modelling, particularly in analysing and predicting ecological water quality.

The physicochemical and hydromorphological aspects of an aquatic ecosystem influence water quality and the occurrence of certain organisms (Alvarez-Mieles et al., 2013; Mereta et al., 2012). Various environmental variables may simultaneously cause water quality deterioration; however, selection of a limited number of relevant variables is essential to delineate management priorities and facilitates water resource restoration and conservation in a cost-effective way. Thus, input variable selection plays an important role in the development of data-driven models in environmental modelling (Li et al., 2015).

This paper aims to diagnose the major environmental variables predicting ecological water quality through the application of fuzzy models. Furthermore, an insight on the interactions of variables and the significance of different variables determining the water quality was examined. Model reliability was evaluated based on model's predictive performance and ecological relevance. A sensitivity analysis was performed as an extra element to evaluate the reliability and performance of the optimal model. Although the methodology was previously applied to predict species distributions in other river systems, in this study, the methodology was extended to model a whole community indicator and was applied in water management and policy making. In this paper, a case study is presented in modelling the ecological water quality in a single tropical river basin. However, the approach could be applied to other river basins as decision support in water management due to its simplicity, universality and transparency.

2. Materials and method

2.1. Study area

The Guayas River basin is located in central-western Ecuador and is one of the largest river systems in South America (Fig. 1). It occupies a land surface area of 34,000 km^2. The Daule Peripa dam is located in the Guayas River basin and covers an area of 18,000 ha (Arriaga, 1989). The Guayas River, with a total length of 60 km,

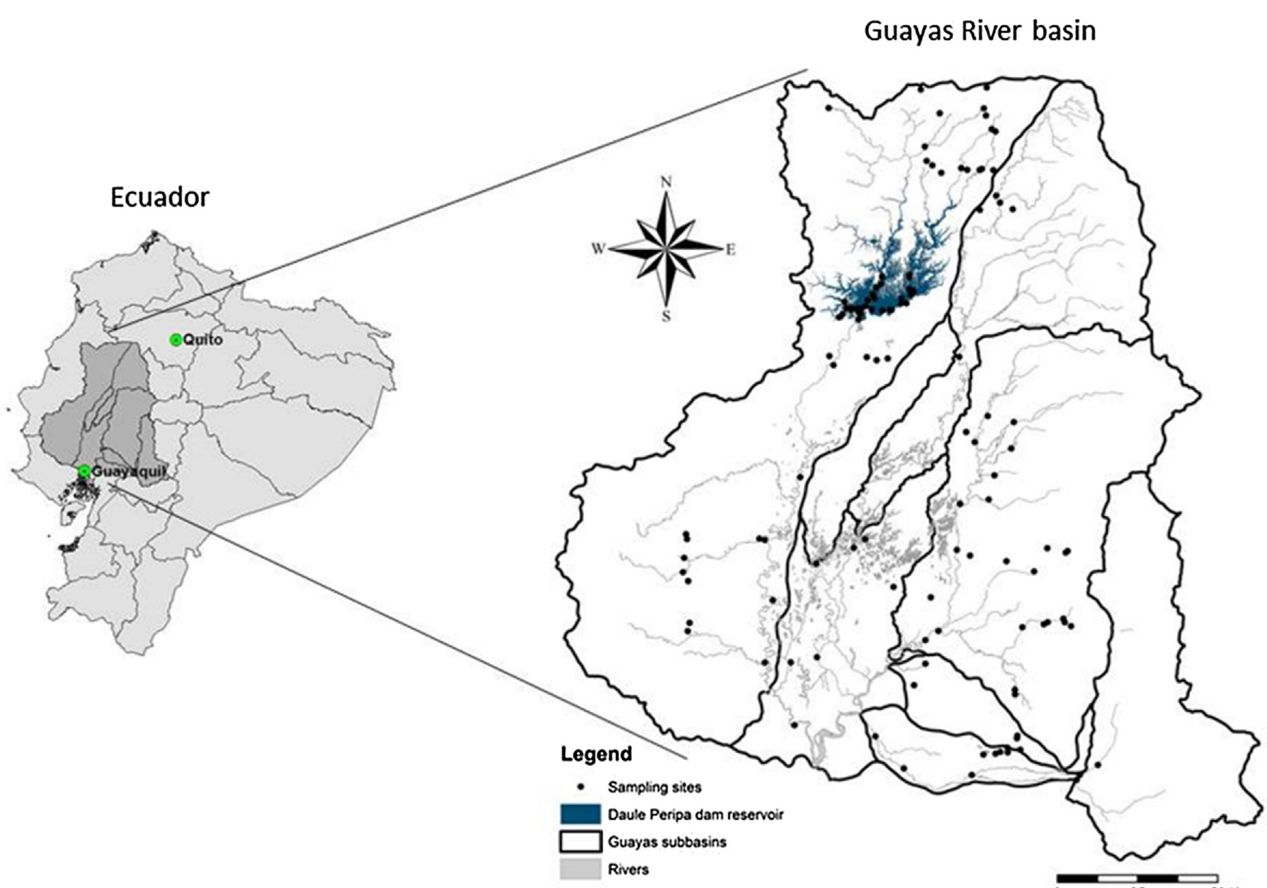

Fig. 1. The sampling sites at Guayas River Basin. The main rivers and the Daule Peripa dam within the basin are represented.

discharges in the Pacific Ocean. The dry season falls between May and November (Arias-Hidalgo, 2013).

Within the Guayas River basin, various human activities such as urban-industrial development, agriculture and aquaculture threaten the quality of the water and resulted in biodiversity loss and reduced water availability (Alvarez-Mieles et al., 2013; Arias-Hidalgo, 2013). Currently, water quality issues receive limited attention as projects have been focusing on flood control and irrigation, which were centralized on water supply requirements (Andres, 2009).

2.2. Data collection

Benthic macroinvertebrates are commonly used in water quality monitoring and assessment (Forio et al., 2016; Lock et al., 2011). Macroinvertebrates function as bioindicators for water quality as they (1) are sensitive to disturbance, (2) have varying degrees of tolerances towards pollution and (3) have a limited mobility (Hilsenhoff, 1988; Molozzi et al., 2012). Furthermore, the life cycles of most taxa are long enough to integrate the environmental stresses that have occurred over an extended period (Rosenberg and Resh, 1996). In this study, macroinvertebrates were used to monitor and assess the ecological water quality.

Macroinvertebrates were sampled in the Guayas River basin at 120 different locations from October to November, 2013. The sampling campaign was conducted during the dry season for safety reasons and to ensure accessibility of all sampling locations. Macroinvertebrates were monitored through kick sampling with a standard handnet (conical net with a frame size of 20×30 cm and a mesh size of 500 μm, attached to a stick) as described by Gabriels et al. (2010). For each sampling site, a 10–20 m stretch was sampled during 5 min. Sampling effort was proportionally distributed over all aquatic habitats present at the sampling site, including bed substrates (stones, sand, or mud), macrophytes (floating, submerged, emerging) and other floating or submerged natural and artificial substrates. All the collected materials were transferred to buckets with covers. Afterwards, samples were sieved and organisms were sorted alive in the laboratory. Macroinvertebrates were identified to family level. To assess the ecological water quality, the Biological Monitoring Working Party/Colombia (BMWP/Col, Alvarez, 2006) index was calculated for each sampling point based on the presence or absence of macroinvertebrate taxa and classified into five discrete water quality classes. For more details on the calculation of the water quality index, we refer to Alvarez (2006).

For each sampling site, physicochemical water quality characteristics were measured (Table 1). Additionally, elevation, average stream depth and width were measured. Hydromorphological characteristics, such as river morphology, water flow, the presence or absence of macrophytes, substrate characteristics and main land use were determined via field inspection.

2.3. Pre-selection of input variables

Input variables predicting water quality were selected based on the correlation matrix (Appendix A), ecological relevance and input variables incorporated in the Bayesian Belief Network (BBN) model obtained from a previous study (Forio et al., 2015). This resulted in a pre-selection of 7 input variables for fuzzy rule-base modelling (Table 2). Sampling sites with missing values were omitted prior to analysis. As a result, there were 100 sampling sites left for data analysis.

2.4. Fuzzy set parameters

Linguistic values such as 'very bad', 'bad', 'poor', 'moderate' and 'good' were assigned to the output variable (BMWP/Col). Values such as 'low' and 'high' were assigned to the six input variables (chlorophyll, dissolved oxygen, flow velocity, dominant bed substrate, elevation and chemical oxygen demand) while values such as 'forest', 'arable', 'residential' and 'orchard' were assigned to the input variable 'main land use'. These linguistic values were defined by fuzzy sets (Zadeh, 1965). A membership function of a particular fuzzy set indicates the degree to which an element belongs to these fuzzy sets, with membership ranging from zero to one. Consequently, the membership functions of the fuzzy sets have overlapping boundaries. A flow velocity which has a membership degree of 0.4 to the fuzzy set 'low' and of 0.6 to the fuzzy set 'high' can be translated into the linguistic statement 'the flow velocity is quite low but is tending to be high'. In this study, all membership functions had trapezoidal shapes and were defined by four parameters (a,b,c,d). The membership degree linearly increases between a and b from 0 to 1 and is equal to 1 between b and c and linearly decreases from 1 to 0 between c and

Table 1
All measured chemical, physical and hydromorphological variables considered in the study.

Chemical variables (units)	Measuring Device	Physical variables (units)	Measuring device	Hydromorphological variables
Specific conductivity (μS/cm)	Multiprobe[a]	Water temperature (°C)	Multiprobe[a]	Valley form
Water pH	Multiprobe[a]	Average flow velocity (m/s)	Current meter[c]	Channel form
Dissolved oxygen (mg/L)	Multiprobe[a]	Turbidity (NTU)	Multiprobe[a]	Variation in width
Chemical oxygen demand (mg/L)	Spectrophotometer[b]			Extent of erosion
Chlorophyll a concentration (μg/L)	Multiprobe[a]			Bank profile
Chloride concentration (mg/L)	Multiprobe[a]			Variation of flow
Total phosphorus (mg/L)	Spectrophotometer[b]			Depth of sludge layer
Nitrate-N (mg/L)	Spectrophotometer[b]			Pool/Riffle class
Nitrite-N (mg/L)	Spectrophotometer[b]			Bank shape
Ammonium-N (mg/L)	Spectrophotometer[b]			Bank slope
Total Nitrogen (mg/L)	Spectrophotometer[b]			Bed compaction
				Sediment matrix
				Sediment angularity
				Main sediment type

[a] model YSI 6600 V2 and YSI 6600 V1, YSI manufacturer.
[b] Hach Lange GmbH spectrophotometric method.
[c] model höntzsch HFA, Höntzsch GmbH manufacturer.

Please cite this article in press as: M.A.E. Forio, et al., Fuzzy modelling to identify key drivers of ecological water quality to support decision and policy making, Environ. Sci. Policy (2016), http://dx.doi.org/10.1016/j.envsci.2016.12.004

Table 2
Measured variables at each sampled river, the linguistic values assigned to the input and output variables of the water quality models and the fuzzy sets describing these linguistic values.

Variable	Unit	Linguistic value	Thresholds	Fuzzy set parameters
Chlorophyll	µg/L	Low	0–3	(0.00, 0.00, 2.70, 3.30)
		High	3–67	(2.70, 3.30, 67.00,69.00)
Dissolved oxygen	mg/L	Low	0–7	(0.00, 0.00, 6.30, 7.70)
		High	7–14	(6.30, 7.70, 14.00, 15.00)
Flow velocity	m/s	Low	0–0.4	(0.00, 0.00, 0.36, 0.44)
		High	0.4–1.6	(0.36, 0.44, 1.60, 1.80)
Dominant bed substrate	2 classes			
	1 = coarse	Low		(0.00, 0.00, 1.10, 1.20)
	2 = fines	High		(1.90, 2.00, 2.00, 2.10)
Elevation	m.a.s.l.	Low	0–200	(0.00, 0.00, 180.00, 220.00)
		High	200–1076	(180.00, 220.00, 1076.00, 1078.00)
Chemical oxygen demand	mg/L	Low	0–10	(0.00, 0.00, 9.00, 11.00)
		High	10–20	(9.00, 11.00, 118.00, 120.00)
Main land use	4 classes			
	1 = forest	very low		(0.00, 0.00, 1.10, 1.20)
	2 = arable	Low		(1.90, 2.00, 2.00, 2.10)
	3 = residential	Medium		(2.90, 3.00, 3.00, 3.10)
	4 = orchard	High		(3.90, 4.00, 4.00, 4.10)
Water quality	5 classes			
	1 = very bad	very bad		(0.00, 0.00, 1.10, 1.20)
	2 = bad	Bad		(1.90, 2.00, 2.00, 2.10)
	3 = poor	Poor		(2.90, 3.00, 3.00, 3.10)
	4 = moderate	Moderate		(3.90, 4.00, 4.00, 4.10)
	5 = good	Good		(4.90, 5.00, 5.00, 5.10)

d. Table 2 presents the fuzzy set parameters of the output and input variables.

The fuzzy set parameters of the continuous input variables were derived from the thresholds in the BBN water quality model in a previous study. For more details on the thresholds of the continuous input variables, we refer to Forio et al. (2015). The thresholds were assigned with a 0.5 membership function. 10% of the thresholds were carried out as an uncertainty which created the trapezoidal shapes.

2.5. Fuzzy rule-based modelling and rule base optimisation

Fuzzy rule-based models relate the input variables to the water quality classes and consist of if-then rules, such as 'IF chlorophyll is low AND dissolved oxygen is high THEN the ecological water quality is moderate'. The if-part of the rule, called the antecedent, describes in which situation this rule applies, while the then-part, called the consequent, indicates the predicted ecological water quality of the river. Both observed and modelled values of the output variable were assigned to the fuzzy set with the highest fulfilment degree, which allowed comparison of the modelled output with the observed output and calculation of performance measures. Specifically, if a water quality value has 0.4 membership in the fuzzy set 'good' and 0.6 membership in the fuzzy set 'moderate', then the value is assigned to 'moderate'. If both 'moderate' and 'good' have a membership of 0.5, then the value is assigned as good.

To generate a reliable ecological water quality model, the consequents of the fuzzy rules were optimized using a nearest ascent hill climbing algorithm (Michalewicz and Fogel, 2000). Starting from fixed sets (Table 2) and a randomly selected rule base, the consequent of one rule is changed into the linguistic value of one of its neighbouring fuzzy sets (e.g. the replacement of 'high' by 'low') and the impact on model performance is calculated. If model performance increases, the algorithm continues with the adjusted rule base, if not, it continues with the original one. As three-fold cross validation revealed to be the most stable estimator for the model's predictive performance, three-fold cross validation was applied to indicate the robustness of the optimization results. The original data were divided into three groups, stratified based on the target variable's classes (good, moderate, poor, bad, very bad). Two of these groups were used to train the model, while the third subset was used to test the model. Each training iteration was stopped when no further increase of the performance measure on the test fold was observed. Each iteration was repeated starting from a random point in the search space and the obtained rule base was compared to each rule base obtained in previous iteration steps. The number of iterations was increased until more than 70% of the rule base was similar among iterations.

Models were optimized until the highest Cohen's *Kappa* (Cohen, 1960) was attained as Cohen's Kappa appeared to be an appropriate performance measure for the rule base optimization (Mouton et al., 2008). Cohen's *Kappa* ranges from −1 to 1 and is derived from the confusion matrix that cross-tabulates observations and model predictions. In contrast to other performance measures that only quantify predictive performance, *Kappa* measures the proportion of all possible presences and absences that are predicted correctly by a model after accounting for chance. For N data points, Kappa is given by:

$$Kappa = \frac{(a+b) - \left(\frac{(a+c)(a+b)+(b+d)(c+d)}{N}\right)}{N - \left(\frac{(a+c)(a+b)+(b+d)(c+d)}{N}\right)}$$

where a is the number of true positive predictions, b is the number of false positive predictions, c is the number of false negative predictions and d is the number of true negative predictions.

2.6. Variable selection

The variable selection method is a combination of a step-forward and step-backward procedure. A step-forward variable selection starts from a model which contains one variable and then this model expands by adding the other variables one by one. A step-backward selection method starts from a model which includes all variables and then reduces this model by removing variables one by one. In both approaches, the contribution of each variable is the difference between the performances of the models with and without this variable. If the input variables are not independent of each other, the step-forward variable contribution to the total model performance differs from the step-backward contribution. Specifically, the step-forward procedure considers different correlations between variables than the correlations which are included in the step-backward approach. Therefore, the variable selection method applied is a combination of both step-forward and step-backward procedures. Although this approach is computationally more expensive, it considers all possible correlations between variables and thus generates consistent and reliable results (Mouton et al., 2009).

This selection was based on model performance, which is quantified by Cohen's *Kappa* (Cohen, 1960) and the Aikaike information criterion (Akaike, 1974). AIC quantifies the balance between model performance and complexity and is calculated as:

$$AIC = 2K + N\ln\left(\frac{RSS}{N}\right)$$

with K is the number of parameters in the model, N is the total number of data points and RSS is the residual sum of squares of the model. Although the AIC assigns an arbitrary weight to performance and complexity, it provides an indication of how likely a model will overfit the data.

2.7. Assessment of ecological relevance

The ecological relevance of the different models was analysed by plotting the cumulative predictive ecological classes of each model for each input variable. Comparison of these plots with the cumulative observed ecological water quality classes reveals the extent to which a model is over- or underpredicting the observations. A model is assumed to be ecologically more relevant if the shape of its resulting cumulative prediction curves resembles the shape of the cumulative observed ecological water quality and if it showed no underprediction of this ecological water quality. An underpredicting model is ecologically irrelevant, whereas an overprediction may not necessarily imply a model error. For more details, we refer to Mouton et al. (2009).

2.8. Sensitivity analysis

To evaluate the optimal model, a sensitivity analysis was performed. A sensitivity analysis provides qualitative and/or quantitative information regarding the effects of different input variables on the output variable of the model. In this paper, a sensitivity analysis was performed to determine the influence of the different variables on the ecological water quality in the model. Based on the optimal model, the predicted ecological water quality is plotted as a function of an input variable while the remaining variables are kept constant.

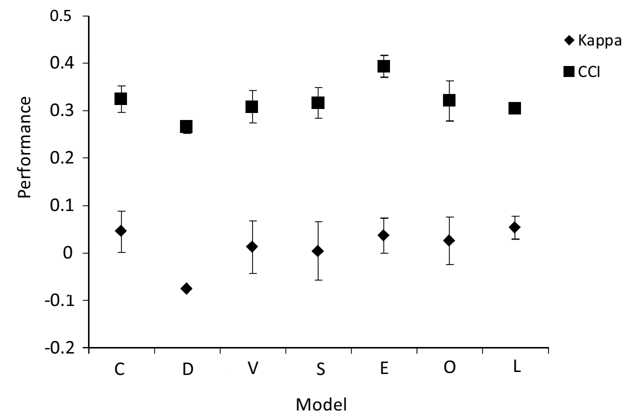

Fig. 2. Performance of seven fuzzy rule optimization models that incorporate one variable: chlorophyll (C), dissolved oxygen (D), velocity (V), sediment type (S), elevation (E), chemical oxygen demand (O) and land use (L). Model performance was quantified by the percentage of correctly classified instances (CCI) and Kappa. Values of these performance criteria were averaged over the three folds.

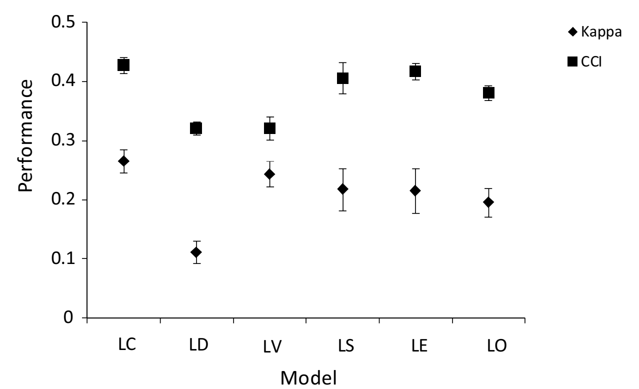

Fig. 3. Performance of six fuzzy rule optimization models that incorporated land use (L) and another variable: chlorophyll (C), dissolved oxygen (D), velocity (V), sediment type (S), elevation (E) and chemical oxygen demand (O). Model performance was quantified by the percentage of correctly classified instances (CCI) and Kappa. Values of these performance criteria were averaged over the three folds.

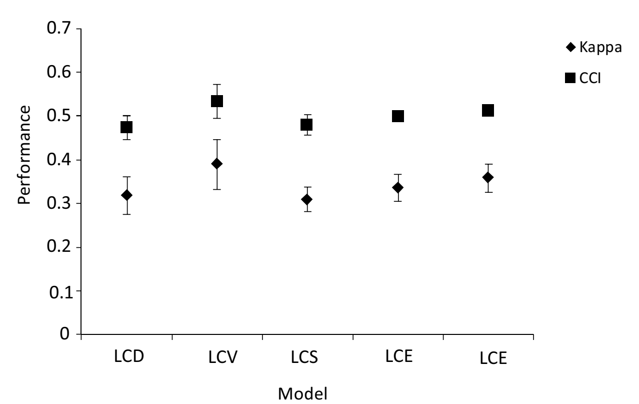

Fig. 4. Performance of five fuzzy rule optimization models that incorporated land use (L), chlorophyll (C) and another variable: dissolved oxygen (D), velocity (V), sediment type (S), elevation (E) and chemical oxygen demand (O). Model performance was quantified by the percentage of correctly classified instances (CCI) and Kappa. Values of these performance criteria were averaged over the three folds.

Please cite this article in press as: M.A.E. Forio, et al., Fuzzy modelling to identify key drivers of ecological water quality to support decision and policy making, Environ. Sci. Policy (2016), http://dx.doi.org/10.1016/j.envsci.2016.12.004

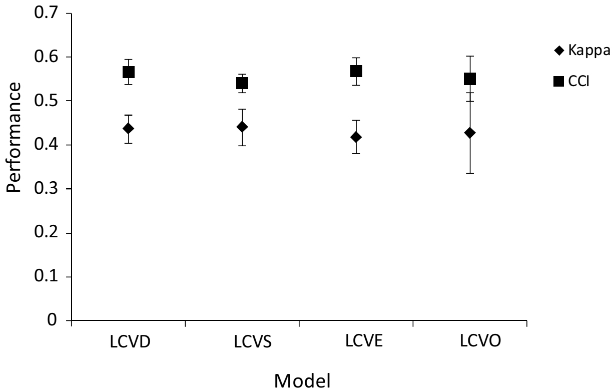

Fig. 5. Performance of five fuzzy rule optimization models that incorporated land use (L), chlorophyll (C), velocity (V) and another variable: dissolved oxygen (D), sediment type (S), elevation (E) and chemical oxygen demand (O). Model performance was quantified by the percentage of correctly classified instances (CCI) and Kappa. Values of these performance criteria were averaged over the three folds.

3. Results

The variable selection procedure indicates that the optimal fuzzy rule base model included three variables: main land use, chlorophyll concentration and flow velocity (LCV model). Land use appears to be the most important variable explaining the ecological water quality (Fig. 2). The inclusion of chlorophyll and flow velocity contributed substantially to an increase in model performance (Figs. 3 and 4). The inclusion of a fourth variable to the LCV model shows only a marginal contribution to the model's predictive performance (Fig. 5). Furthermore, the similar performances among these four variable models may demonstrate that these four input variables only provide limited additional information on the ecological water quality (Fig. 5).

The cumulative prediction curves of all models resemble the shape of the cumulative observed ecological water quality (Figs. 6 and 7). The model incorporating only main land use (L) reveals an underprediction of the ecological water quality, while the model incorporating both land use and chlorophyll (LC) overpredicted the observations the most. Furthermore, the model incorporating main land use, chlorophyll concentration and velocity (LCV) is also overpredicting the observations. Incorporation of the sediment type to the LCV model led to a marginal model improvement and showed overprediction of observations.

The model incorporating all the pre-selected input variables has the highest AIC value (Table 3), which implies model overfitting. LC model has the lowest AIC value and the AIC of the LCV model was comparably low. Both the L and LCVS model have similar intermediate AIC values.

The optimised fuzzy rules are consistent and proved to be coherent. They reveal that rivers surrounded by forest land use predicted water quality to range from poor to good, while rivers situated in the residential area predicted water quality ranging from very bad to good (Table 4). Arable and orchard land use were related to very bad to poor and very bad to moderate water quality, respectively. Among the different land uses, the combination of a low chlorophyll concentration and a high flow velocity showed the best water quality, while the combination of a high chlorophyll concentration and a low flow velocity resulted in the worst water quality.

The sensitivity graph (Fig. 8) illustrates the influence of chlorophyll, velocity and land use on the ecological water quality classes. The results reveal that for each land use class, the ecological water quality is best at higher flow velocities (0.5 and 1 m/s). Furthermore, the degree of shift in water quality class as a result of a change in chlorophyll concentration is most dramatic for residential land use.

4. Discussion

4.1. Quantification of model performance

The most common measure of model performance in ecological modelling has been predictive accuracy (Fielding and Bell, 1997). In the last decade, aspects such as ecological relevance (Manel et al., 2001; Mouton et al., 2011) and model's applicability (Boets et al., 2013; Goethals et al., 2007) are also important in assessing model performance. However, the choice of performance aspects is usually driven by the goal of the study.

In this study, besides the predictive accuracy of the model, the distinction between over- and underprediction of observations was considered as an important issue for the selection of a reliable ecological model. In a previous study, it has been shown that models overpredicting the observations may not necessarily be ecological irrelevant, while underprediction of the observation

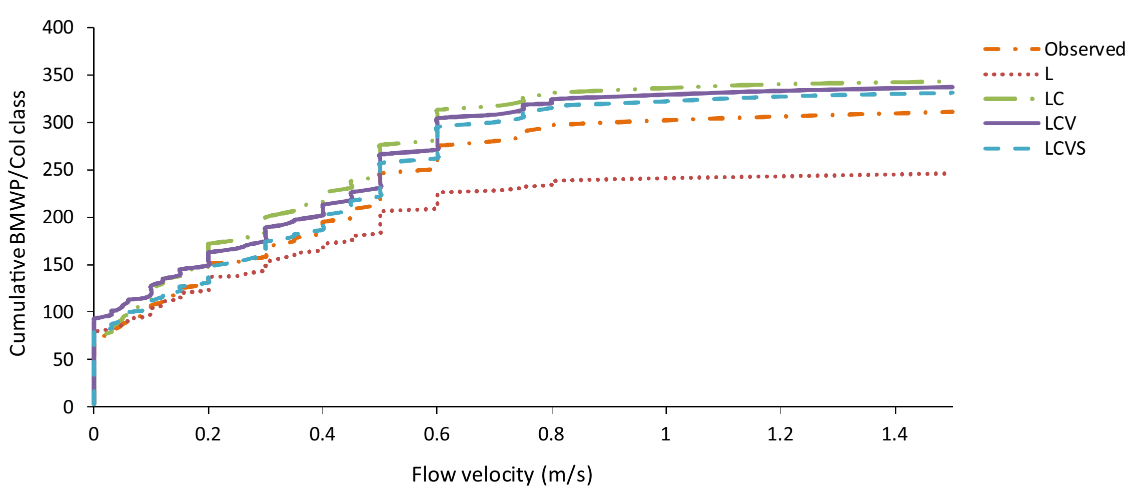

Fig. 6. Cumulative predicted ecological water quality classes of the models selected in the model development process compared to the cumulative observed ecological water quality classes as a function of the flow velocity. The four models incorporate land use only (model L), land use and chlorophyll (model LC), land use, chlorophyll and velocity (model LCV) and land use, chlorophyll, velocity and sediment type (model LCVS).

Please cite this article in press as: M.A.E. Forio, et al., Fuzzy modelling to identify key drivers of ecological water quality to support decision and policy making, Environ. Sci. Policy (2016), http://dx.doi.org/10.1016/j.envsci.2016.12.004

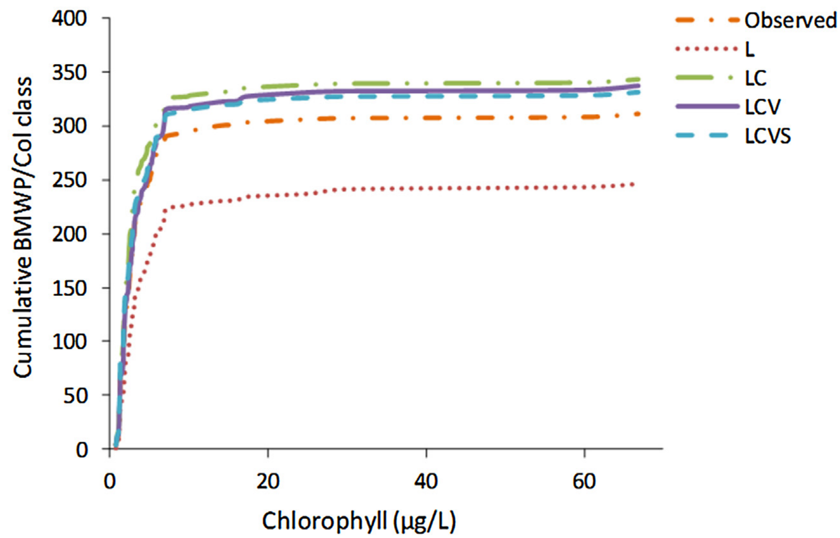

Fig. 7. Cumulative predicted ecological water quality classes of the models selected in the model development process compared to the cumulative observed ecological water quality classes as a function of the chlorophyll concentration. The four models incorporate land use only (model L), land use and chlorophyll (model LC), land use, chlorophyll and velocity (model LCV) and land use, chlorophyll, velocity and sediment type (model LCVS).

Table 3
Performance of the four optimal models obtained after addition of variables to the initial one-variable model and the models with all the 7 pre-selected input variables.

Model	Mean AIC ± sd	Mean Kappa ± sd	Mean CCI ± sd
L	61 ± 9	0.05 ± 0.04	0.30 ± 0.02
LC	42 ± 12	0.26 ± 0.03	0.42 ± 0.02
LCV	45 ± 7	0.39 ± 0.10	0.53 ± 0.07
LCVS	64 ± 13	0.44 ± 0.07	0.56 ± 0.05
LCVSDEO	495 ± 11	0.51 ± 0.03	0.62 ± 0.02

always implies a model error (Mouton et al., 2009). As ecological water quality is based on a weighted score on the presence of a certain macroinvertebrate taxon, the model's under- or overprediction can be explained similarly with the occurrence of certain taxa. A false negative prediction, which implies a prediction of certain taxon/taxa to be absent despite its presence, is always false and implies a model error. In contrast, false positive predictions are inherent to the classification of ecological data, which could be a result of monitoring inefficiency, migration barriers and temporal population variation (Barry and Elith, 2006; Lütolf et al., 2006; MacKenzie et al., 2003). Monitoring inefficiency could result in the absence of a certain macroinvertebrate taxon, despite its presence. This may result in a lower observed water quality, despite the site's better water quality.

4.2. Variable and model selection

The selection of a limited number of relevant environmental variables is a common problem in ecological modelling. Thus, variable selection is necessary as it decreases model complexity, increases processing speed and decreases the amount of data required to estimate model parameters efficiently (D'heygere et al., 2003; Walczak and Cerpa, 1999). In literature, various methods were implemented to determine the optimal set of input variables (Goethals et al., 2007; Zuur et al., 2009). These include the (1) acquisition of standard knowledge and elimination of highly correlated variables, (2) analytical techniques (i.e. assessment of model performance after addition of remaining input variables) and (3) input variable selection through genetic algorithms. In this study, both first two methods were implemented. The third

Table 4
The fuzzy rule base of the optimal fuzzy model (model LCV). The number of habitats that covered each environmental condition is shown in the last column.

Main land use	Chlorophyll	Velocity	Ecological water quality	Rule coverage
Forest	L	L	Moderate	9
Forest	H	L	Poor	27
Forest	L	H	Good	14
Forest	H	H	Moderate	2
Arable	L	L	Bad	12
Arable	H	L	Very bad	7
Arable	L	H	Poor	8
Arable	H	H	Bad	2
Residential	L	L	Moderate	5
Residential	H	L	Very bad	5
Residential	L	H	Good	3
Residential	H	H	Poor	1
Orchard	L	L	Poor	1
Orchard	H	L	Very bad	1
Orchard	L	H	Moderate	2
Orchard	H	H	Bad	1

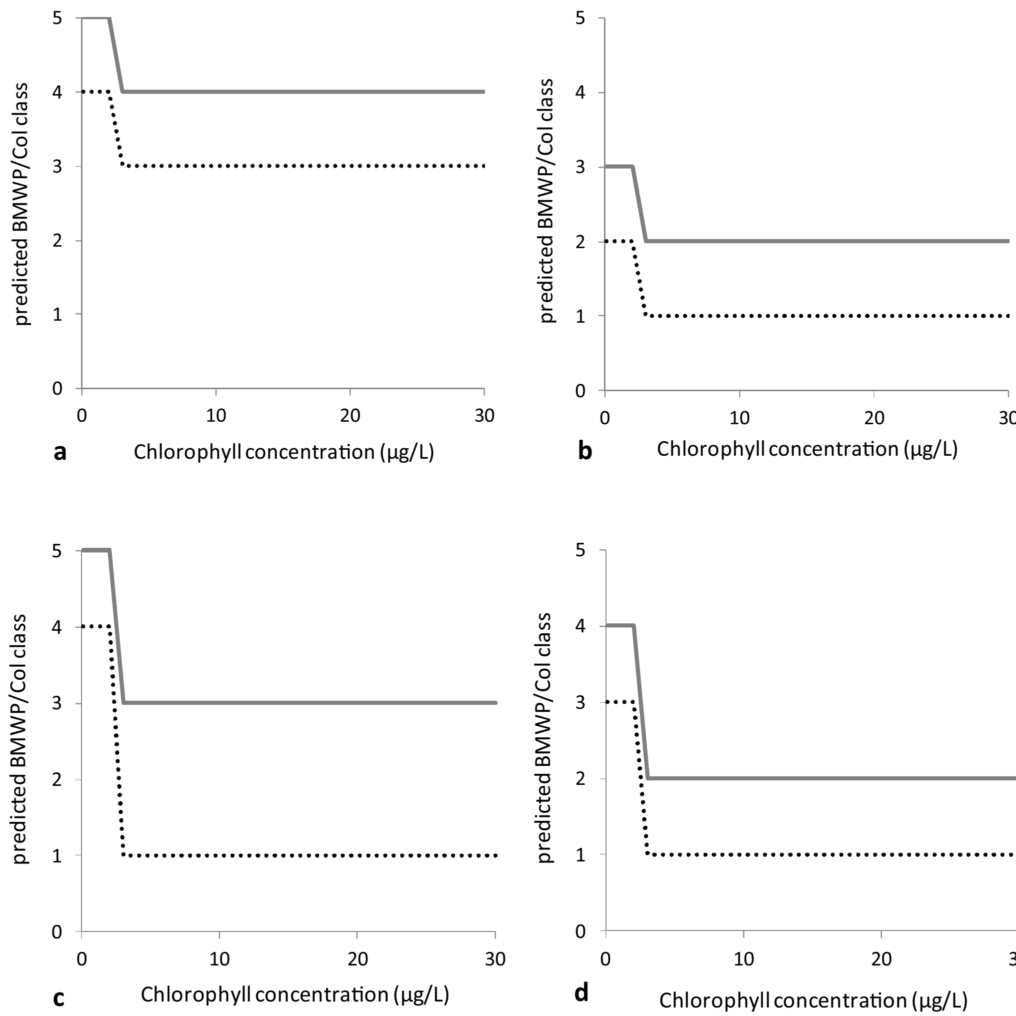

Fig. 8. Predicted ecological water quality classes of model LCV as a function of the chlorophyll concentration at 0.1 m/s flow velocity (broken line) and 0.5 and 1 m/s flow velocity (solid line; both flow velocity gave the same result) at forest land use (a), arable land use (b), residential land use (c) and orchard land use (d).

method was not applied as the dataset was too small. During the pre-selection of variables, not all correlated variables were eliminated. For instance, chlorophyll and velocity were highly correlated, but one of the variables was not excluded during model development. From an ecological point of view, correlated variables do not necessarily involve the same ecological processes, thus relationships between environmental variables should be considered with caution.

In this paper, the optimal model contained three variables, while the addition of a fourth variable only led to a marginal increase in model performance. Both the backward and forward selection steps showed the same results, however, caution should be taken in selecting an appropriate model. The optimal models with two to three variables reveal to be ecologically relevant as the curves resemble the shape of the cumulative observed ecological water quality and there was no underprediction of this ecological water quality. As the model including only land use and chlorophyll (LC model) showed the most overprediction, this model was not considered as the most appropriate model. Although the LCVS model had a high predictive performance, the model was not selected as the optimal model because the model slightly underpredicted the observations. Furthermore, the inclusion of extra variables in the model decreases model readability and interpretability as rule sets become very complex.

The inclusion of all pre-selected variables (LCVSDEO model) showed a substantial increase in model performance. However, the considerably high AIC of the model indicates significant model overfitting. The inclusion of all variables may be interesting to determine the evolution of ecological water quality upon the combination of these variables. However, the development of reliable, but more complex data-driven models often requires more data and incorporation of a wider environmental range in the training data, which results in a substantial increase in costs and efforts. The optimal model was selected based on its predictive performance, ecological relevance and model's readability and interpretability.

Land use appeared to be the most important variable predicting water quality based on variable and model selection procedure. This result is further confirmed by the sensitivity analysis. The degree of shift in water quality class as a result of a change in chlorophyll concentration depends on the land use class. Land use affects water quality as illustrated in various studies (Bu et al., 2014; Erol and Randhir, 2013; Gyawali et al., 2013). Furthermore, land use may indirectly reflect unmeasured environmental variables such as pesticide or pharmaceutical concentrations in the water. These unmeasured variables may have significant effects on the water quality. Consequently, the inclusion of land use in the fuzzy model could be considered ecologically relevant and demonstrates that a data-driven fuzzy approach may provide a wider scope of interpretability.

Chlorophyll concentration appeared to be an important variable predicting the ecological water quality. Chlorophyll concentration represents algal biomass in river systems (Boyer et al., 2009). The occurrence of algal blooms is caused by a complex interaction of various variables, such as nutrient availability (Veraart et al., 2008), canopy cover (Mosisch et al., 2001), turbidity (Munn et al., 1989), water temperature (Morin et al., 1999) and hydrologic disturbances (Riseng et al., 2004). Although algae can become abundant in water with open canopies, their growth is limited by nutrient availability. Under nutrient enrichment and heavy modification of hydrological conditions, algal biomass can abundantly increase and become a nuisance in rivers (Frankforter et al., 2010; Sabater et al., 2011). In the optimal model, the variable chlorophyll reflects that nutrient inputs in the rivers are important variables predicting the ecological water quality. Various areas of the river basin are characterized by agricultural land which may contribute to nutrient runoff into the rivers as a consequence of the extensive use of either nitrate- or phosphate-rich fertilizers.

The final important variable explaining water quality in the fuzzy model was flow velocity. Alteration of flow velocity can be caused by water abstraction and existing dams within the river basin. Altering flow velocity can cause shifts in macroinvertebrate communities and may consequently result in a poor ecological water quality (Forio et al., 2015). Furthermore, low flow velocity may cause accumulation or slow movement of undesirable chemical pollutants (e.g. pesticides, micropollutants, nutrients, organic pollutants). As seen in the sensitivity analysis, impacts of nutrients were apparent at low flow velocities. Additionally, flow velocity facilitates the dissolution of oxygen in the water. The higher the flow velocity, the better is the dissolution of oxygen. As dissolved oxygen is essential for the occurrence of most macroinvertebrates, a higher flow velocity aids the presence of many taxa. The chain effects caused by a low flow velocity can reduce macroinvertebrate diversity and abundance and an insight on variable interaction is successfully demonstrated in fuzzy models.

4.3. Model applicability in decision and policy making

Fuzzy logic models allow to determine the major variables that determine the ecological water quality. This approach can assist river managers in prioritizing restoration activities to improve water quality. As land use is the major variable that determines the ecological water quality, the model outcome can support spatial planning and policy making. For instance, incorporating a forested area in between arable land use may alleviate the water quality. However, the acceptability of the models by water managers is mainly defined by their perception of the practical value of the model. Although water managers are interested in models that describe water systems in a very reliable and detailed manner, also the model's functionality and user-convenience are of paramount importance. Fuzzy logic models are characterized by their high transparency and simplicity, and these aspects help to convince stakeholders to use this modelling technique to guide their decisions. Although fuzzy logic models have several advantages over other techniques, the main challenge is the development time and practical knowledge to integrate both knowledge and data from different sources in these models. Stable predictive performance and model transparency are moreover inherent to BBNs. However, extra information (data or knowledge) is necessary to build the network structures (cause-effect relationship) and conditional probabilities in BBNs. In comparison to BBNs, fuzzy models are easier to implement, particularly in situations where the cause-effect relations are not well understood.

Other techniques such as classification trees, artificial neural networks (ANNs) and random forests can be used to determine the major variables that affect ecological water quality. Classification trees may be transparent, but the predictive performance and in particular the selection of the key variables can be unstable (Goethals, 2005). The latter also counts for ANNs and random forests, that are less transparent in comparison to classification trees and in particular to fuzzy and BBN models (Mouton et al., 2011; Goethals et al., 2007). Thus, fuzzy models may be appropriate for situations where field data are limited and expert knowledge is available, and in particular for situations where stable predictive performance is preferred, combined with model transparency and simplicity.

The national regulatory authority for environmental management in Ecuador recognized the necessity of applying a holistic and integrated approach to set up a sustainable management of aquatic resources (Nolivos et al., 2015). The Ecuadorian government translated this need into a national policy aiming for more water resources oriented research. Consequently, this policy facilitates the communication between science and policy to solve problems in order to achieve sustainable natural resources management. Thus, the fuzzy model developed in this study can contribute to increasing the efficiency of investments in environmental protection and restoration. Furthermore, this study aids and facilitates the current call of the Ecuadorian government not only in the establishment of monitoring and assessment programs at river basin scale, but also to provide cost-effective solutions to the problems in the water sector.

Although our study discusses an application of a fuzzy logic modelling approach on a single river basin, the optimal model in this study and the modelling approach itself can be applied on any other river basin. However, extrapolation of the optimal model should be done with care. As environmental conditions in other regions may be slightly different, the model should be tested and validated before application. Application of the modelling approach on other systems is relatively flexible since other variable categories (i.e. main land use) and suitable ecological water quality indices in a particular region can be easily implemented in the model and are in general available. As the modelling approach makes use of empirical data, an improvement on data collection can contribute to a better model performance. For instance, implementation of a balanced collection of variable classes (e.g. the same number of cases with forest, orchard, arable, residential land use) is recommended.

Acknowledgements

This research was performed in the context of the VLIR Ecuador Biodiversity Network project. This project was funded by the Vlaamse Interuniversitaire Raad – Universitaire Ontwikkelingssamenwerking (VLIR-UOS), which supports partnerships between universities and university colleges in Flanders and the South. The authors would also like to thank the assistance of the Department of Chemical and Environmental Sciences (CADS-ESPOL) staff during the sampling campaign in Ecuador and Dr. Christine Vander Heyden (HoGent) for sharing her expertise on chemical analysis of water samples. Marie Anne Eurie Forio and Minar Naomi Damanik Ambarita are supported by the Special Research Fund of Ghent University (BOF). Pieter Boets was supported by a post-doctoral fellowship from the Fund for Scientific Research Flanders (FWO-Vlaanderen).

Appendix A.

Correlation matrix of continuous variables: temperature (Temp), conductivity (Cond), pH, chlorophyll concentration (Chloro), chloride concentration (Cl), dissolved oxygen (DO), turbidity (Turbid), total nitrogen (TotN), total phosphorus (TotP), nitrate-N concentration (NO3), nitrite-N concentration (NO2), ammonia-N (NH4), average stream width (Width), average water depth (Depth), average velocity (Vel), sampling site elevation (Ele).

	Temp	Cond	pH	Chloro	Cl	DO	Turbid	COD	TotN	TotP	NO3	NO2	NH4	Width	Depth	Vel	Ele
Temp	1.00	0.53	0.04	0.36	0.48	−0.11	0.31	0.34	0.10	0.24	0.04	0.25	0.16	0.26	−0.09	−0.45	−0.70
Cond		1.00	0.04	0.36	0.53	−0.21	0.29	−0.24	−0.02	0.25	0.18	0.19	−0.08	0.03	−0.09	−0.30	−0.49
pH			1.00	−0.06	0.11	0.83	−0.12	0.32	−0.14	0.13	−0.16	0.01	−0.37	−0.02	0.00	0.37	0.34
Chloro				1.00	0.42	−0.17	0.50	0.48	0.24	0.20	0.04	0.28	0.53	0.05	0.11	−0.48	−0.35
Cl					1.00	0.00	−0.25	−0.24	−0.09	0.18	0.10	0.18	0.31	0.29	0.16	−0.29	−0.40
DO						1.00	−0.25	−0.24	−0.09	0.05	−0.12	0.04	−0.28	0.12	0.07	0.33	0.39
Turbid							1.00	0.32	0.10	0.20	0.04	0.28	0.32	0.09	0.12	−0.26	0.19
COD								1.00	0.08	0.11	0.32	0.30	0.31	−0.03	0.15	0.54	−0.41
TotN									1.00	0.21	0.33	0.19	0.29	−0.21	−0.13	−0.33	−0.02
TotP										1.00	0.16	0.26	0.15	−0.17	−0.21	−0.10	−0.10
NO3											1.00	0.27	−0.11	−0.13	0.07	−0.03	−0.03
NO2												1.00	0.15	0.11	0.09	−0.17	−0.27
NH4													1.00	0.07	0.22	−0.41	0.32
Width														1.00	0.28	0.08	−0.33
Depth															1.00	−0.10	−0.12
Vel																1.00	0.55
Ele																	1.00

References

Adriaenssens, V., De Baets, B., Goethals, P.L.M., De Pauw, N., 2004. Fuzzy rule-based models for decision support in ecosystem management. Sci. Total Environ. 319 (1–3), 1–12.

Adriaenssens, V., Goethals, P.L.M., De Pauw, N., 2006. Fuzzy knowledge-based models for prediction of Asselus and Gammarus in watercourses in Flanders (Belgium). Ecol. Model. 195, 3–10.

Akaike, H., 1974. A new look at statistical-model identification. IEEE Trans. Autom. Control 19 (6), 716–723.

Akerkar, R., Sajja, P., 2010. Knowledge-based Systems. Jones and Battlett Publishers, MA, USA pp. 129.

Alvarez, L.F., 2006. Metodología Para La Utilización De Los Macroinvertebrados Acuáticos Como Indicadores De La Calidad Del Agua. Instituto Alexander von Humboldt, Colombia.

Alvarez-Mieles, G., Irvine, K., Griensven Arias-Hidalgo, M., Torres, A., Mynett, A.E., 2013. Relationship between aquatic biotic communities and water quality in a tropical river-wetland system (Ecuador). Environ. Sci. Policy 34, 115–127.

Andres, C., 2009. Sedimentation Processes at the Confluence of the Daule and Babahoyo Rivers, Guayaquil, Ecuador. El Palmar Island. Manejo Ambiental De Ríos. MS Dissertation. University of Birmingham, UK http://repositorio.educacionsuperior.gob.ec/bitstream/28000/1039/1/T-SENESCYT-0415.pdf (Accessed 23 December 2014).

Arias-Hidalgo, M.E., 2013. A decision framework for wetland management in a river basin context: the Abras de Mantiquilla case study in the Guayas River Basin, Ecuador. Environ. Sci. Policy 34, 103–114.

Arriaga, L., 1989. The Daule-peripa dam project, urban development of Guayaquil and their impact on shrimp mariculture. In: Olsen, S., Arriaga, L. (Eds.), A Sustainable Shrimp Mariculture Industry for Ecuador. Narragansett RI: Coastal Resources Center, University of Rhode Island, USA, pp. 148–152.

Barnhoorn, I.E.J., van Dyk, J.C., Genthe, B., Harding, W.R., Wagenaar, G.M., Bornman, M.S., 2015. Organochlorinepesticide levels in Clarias gariepinus from polluted freshwater impoundments in South Africa and associated human health risks. Chemosphere 120, 391–397.

Barry, S., Elith, J., 2006. Error and uncertainty in habitat models. J. Appl. Ecol. 43 (3), 413–423.

Boets, P., Lock, K., Goethals, P., 2013. Modelling habitat preference, abundance and species richness of alien macrocrustacens in surface waters in Flanders (Belgium) using decision trees. Ecol. Inf. 17, 73–81.

Boyer, J.N., Kelble, C.R., Ortner, P.B., Rudnick, D.T., 2009. Phytoplankton bloom status: chlorophyll a biomass as an indicator of water quality condition in the southern estuaries of Florida, USA. Ecol. Indic. 9 (6), S56–S67.

Bu, H., Meng, W., Zhang, Y., Wan, J., 2014. Relationship between land use patterns and water quality in the Taizi river basin, China. Ecol. Indic. 41, 187–197.

Cadenasso, M.L., Pickett, S.T.A., Weathers, K.C., Jones, C.G., 2003. A framework for a theory of ecological boundaries. Bioscience 53 (3), 751–758.

Canavese, D., Ortega, N.R.S., Queirós, M., 2014. The assessment of local sustainability using fuzzy logic: an expert opinion system to evaluate environmental sanitation in the Algarve region, Portugal. Ecol. Indic. 36, 711–718.

Cohen, J., 1960. A coefficient of agreement for nominal scales. Educ. Psychol. Meas. 20 (1), 37–46.

de Siqueira, A., Boclin, C., de Mello, R., 2006. A decision support method for environmental impact assessment using a fuzzy logic approach. Ecol. Econ. 58 (1), 170–181.

D'heygere, T., Goethals, P.L.M., De Pauw, N., 2003. Use of genetic algorithms to select input variables in decision tree models for the prediction of benthic macroinvertebrates. Ecol. Model. 160, 291–300.

Erol, A., Randhir, T., 2013. Watershed ecosystem modelling of land-use impacts on water quality. Ecol. Model. 270, 54–63.

Fielding, A.H., Bell, J.F., 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ. Conserv. 24 (1), 38–49.

Forio, M.A.E., Landuyt, D., Bennetsen, E., Lock, K., Nguyen, T.H.T., Ambarita, M.N.D., Musonge, P., Boets, P., Everaert, G., Dominguez, L., Goethals, P.L.M., 2015. Bayesian belief network models to analyse and predict ecological water quality in rivers. Ecol. Model. 312, 222–238.

Forio, M.A.E., Van Echelpoel, W., Dominguez-Granda, L., Mereta, S.T., Argaw, A., Hoang, T.H., Boets, P., Goethals, P.L.M., 2016. Analysing the effects of water quality on the occurrence of freshwater macroinvertebrate taxa among tropical river basins from different continents. AI Commun. 29 (6), 665–685 http://content.iospress.com/articles/ai-communications/aic712.

Frankforter, J., Weyers, H., Bales, J., Moran, P., Calhoun, D., 2010. The relative influence of nutrients and habitat on stream metabolism in agricultural streams. Environ. Monit. Assess. 168, 461–479.

Gabriels, W., Lock, K., De Pauw, N., Goethals, P.L.M., 2010. Multimetric Macroinvertebrate Index Flanders (MMIF) for biological assessment of rivers and lakes in Flanders (Belgium). Limnologica 40 (3), 199–207.

Gharibi, H., Mahvi, A.H., Babizadeh, R., Arabalibeik, H., Yunesian, M., Swolat, M.H., 2012. A novel approach in water quality assessment based on fuzzy logic. J. Environ. Manag. 112, 87–95.

Goethals, P.L.M., Volk, M., 2016. Implementing sustainability in water management: are we still dancing in the dark? Sustain. Water Qual. Ecol. 7, 1–4.

Goethals, P.L.M., Dedecker, A.P., Gabriels, W., Lek, S., De Pauw, N., 2007. Application of artificial neural network predicting macroinvertebrates in freshwaters. Aquat. Ecol. 41, 491–508.

Goethals, P.L.M., 2005. Data Driven Development of Predictive Ecological Models for Benthic Macroinvertebrates in River. PhD Dissertation. Ghent University, Belgium.

Gyawali, S., Techato, K., Yuangyai, C., Musikavong, C., 2013. Assessment of relationship betweeen land uses of riparian zone and water quality of river for sustainable development of river basin, a case study of U-Tapao river basin, Thailand. Proced. Environ. Sci. 17, 291–297.

Hauck, J., Winkler, K.J., Priess, J.A., 2015. Reviewing drivers of ecosystem change as input for environmental and ecosystem services modelling. Sustain. Water Qual. Ecol. 5, 9–30.

He, H., Cheng, Y., Zhou, J., Keo, S., Goethals, P.L.M., 2014. Monitoring and assessment of water quality and ecology. Sustain. Water Qual. Ecol. 3–4, 1–2.

Hilsenhoff, W.L., 1988. Rapid field assessment of organic pollution with a family-level biotic index. J. N. Am. Benthol. Soc. 7, 65–68.

Please cite this article in press as: M.A.E. Forio, et al., Fuzzy modelling to identify key drivers of ecological water quality to support decision and policy making, Environ. Sci. Policy (2016), http://dx.doi.org/10.1016/j.envsci.2016.12.004

Icaga, Y., 2007. Fuzzy evaluation of water quality classification. Ecol. Indic. 7 (3), 710–718.
Kampichler, C., Barthel, J., Wieland, R., 2000. Species density of foliage-dwelling spiders in field margins: a simple, fuzzy rule-based model. Ecol. Model. 129 (1), 87–99.
Lütolf, M., Kienast, F., Guisan, A., 2006. The ghost of past species occurence: improving species distribution models for presence-only data. J. Appl. Ecol. 43 (4), 802–815.
Li, X., Maier, H.R., Zecchin, A.C., 2015. Improved PMI-based input variable selection approach for artificial neural networks and other data driven environmental and water resource models. Environ. Model. Software 65, 15–29.
Liu, D., Zou, Z., 2012. Water quality evaluation based on improved fuzzy matter-element method. J. Environ. Sci. 24 (7), 1210–1216.
Liu, G., Baniyounes, A., Rasul, M.G., Amanullah, M.T.O., Khan, M.M.K., 2012. Fuzzy logic based environmental indicator for sustainability assessment of renewable energy system using life cycle assessment. Proced. Eng. 49, 35–41.
Lock, K., Asenova, M., Goethals, P.L.M., 2011. Benthic macroinvertebrates as indicators of the water quality in Bulgaria: a case-study in the Iskar river basin. Limnologica 41 (4), 334–338.
MacKenzie, D.I., Nichols, J.D., Hines, J.E., Knutson, M.G., Franklin, A.B., 2003. Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. Ecology 84 (8), 2200–2207.
Manel, S., Williams, H.C., Ormerod, S.J., 2001. Evaluating presence-absence models in ecology: the need to account for prevalance. J. Appl. Ecol. 38 (5), 921–931.
Mereta, S.T., Boets, P., Ambelu, A., Malu, A., Ephrem, Z., Sisay, A., Endale, H., Yitbarek, M., Jemal, A., De Meester, Luc, Goethals, P.L.M., 2012. Analysis of environmental factors determining the abundance and diversity of macroinvertebrate taxa in natural wetlands of Southwest Ethiopia. Ecol. Inf. 7 (1), 52–61.
Michalewicz, Z., Fogel, D.B., 2000. How to Solve It: Modern Heuristics. Springer-Verlag, Berlin, Heidelberg pp. 467.
Molozzi, J., Feio, M.J., Salas, F., Marques, J.C., Callisto, M., 2012. Development and test of a statistical model for the ecological assessment of tropical reservoirs based on benthic macroinvertebrates. Ecol. Indic. 23, 155–165.
Morin, A., Lamoureux, W., Busnarda, J., 1999. Empirical models predicting primary productivity from chlorophyll a and water temperature for stream periphyton and lake and ocean phytoplankton. J. N. Am. Benthol. Soc. 18, 299–307.
Mosisch, T., Bunn, S., Davies, P., 2001. The relative importance of shading and nutrients on algal production in subtropical streams. Freshw. Biol. 46, 1269–1278.
Mouton, A.M., Schneider, M., Peter, A., Holzer, G., Müller, R., Goethals, P.L.M., De Pauw, N., 2008. Optimization of a fuzzy habitat model for spawning European grayling (*Thymallus thymallus* L.) in the Aare river (Thun, Switzerland). Ecol. Model. 215 (1–3), 122–132.
Mouton, A.M., De Baets, B., Goethals, P.L.M., 2009. Knowledge-based versus data-driven fuzzy habitat suitability models for river management. Environ. Model. Software 24 (8), 982–993.
Mouton, A.M., Alcaraz-Hernández, J.D., De Baets, B., Goethals, P.L.M., Martínez-Capel, F., 2011. Data-driven fuzzy habitat suitability models for brown trout in Spanish Mediterranean rivers. Environ. Model. Software 26 (5), 615–622.
Munn, M., Osborne, L.L., Wiley, M.J., 1989. Factors influencing periphyton growth inagricultural streams of central Illinois. Hydrobiologia 174, 89–97.
Nolivos, I., Villacis, M., Vázquez, R., Mora, D.E., Domínguez-Granda, L., Hampel, H., Velarde, E., 2015. Challenges for a sustainable management of Ecuadorian water. Sustain. Water Qual. Ecol. 6, 101–106.
Peche, R., Rodríguez, E., 2011. Environmental impact assessment by means of a procedure based on fuzzy logic: a practical application. Environ. Impact Assess. Rev. 31 (2), 87–96.
Peche, R., Rodríguez, E., 2012. Development of environmental quality indexes based on fuzzy logic a case study. Ecol. Indic. 23, 555–565.
Riseng, C.M., Wiley, M.J., Stevenson, R.J., 2004. Hydrologic disturbance and nutrient effects on benthic community structure in Mid-Western US streams: a coveriance structure analysis. J. N. Am. Benthol. Soc. 23, 309–326.
Rosenberg, D.M., Resh, V.H., 1996. Use of aquatic insects in biomonitoring. In: Merritt, R.W., Cummins, K.W. (Eds.), An Introduction to the Aquatic Insects of North America. Kendall/Hunt, Dubuque, IA, USA, pp. 87–96.
Sabater, S., Artigas, J., Gaudes, A., Muñoz, I., Urrea, G., Romaní, A.M., 2011. Long-term moderate nutrient inputs enhance autotrophy in a forested Mediterranean stream. Freshw. Biol. 56, 1266–1280.
Scannapieco, D., Naddeo, V., Zarra, T., Belgiorno, V., 2012. River water quality assessment: a comparison of binary- and fuzzy logic-based approaches. Ecol. Eng. 47, 132–140.
Teixeira, Z., Teixeira, H., Marques, J.C., 2014. Systematic processes of land use/land cover change to identify relevant driving forces: implications on water quality. Sci. Total Environ. 470–471, 1320–1335.
UN, 2014. International Decade for Action 'water for Life' 2005–2015. http://www.un.org/waterforlifedecade/quality.shtml (Accessed 1 December 2014).
Veraart, A.J., Romaní, A.M., Tornés, E., Sabater, S., 2008. Algal response to nutrient enrichment in forested oligotrophic stream. J. Phycol. 44, 564–572.
Walczak, S., Cerpa, N., 1999. Heuristic principles for the design of artificial neural networks. Information and Software Technology 41 (2), 107–117.
Yarrow, M.M., Salthe, S., 2008. Ecological boundaries in the context of hierarchy theory. Biosystems 92, 233–244.
Zadeh, L.A., 1965. Fuzzy sets. Inf. Control 8 (3), 338–353.
Zuur, A.F., Ieno, E.I., Walker, N.J., Saveliev, A.A., Smith, G., 2009. Mixed Effects Models and Extensions in Ecology with R. Springer Science + Business Media. LLC, 233 Spring Street, New York, NY 10013 USA.