Application of MesoCASiMiR: Assessment of *Baetis rhodani* Habitat Suitability

A. M. Mouton[1], M. Schneider[2], I. Kopecki[2], P.L.M. Goethals[1],
N. De Pauw[1]
[1] Laboratory of Environmental Toxicology and Aquatic Ecology, Ghent University, Ans.Mouton@UGent.be
[2] Schneider & Jorde Ecological Engineering sje, Stuttgart

Abstract: The assessment of the ecological status of running waters at a mesohabitat scale is commonly based on fish. Nevertheless, due to their strong dependence on a good physical and chemical habitat, macroinvertebrate species can also be used for this assessment. A new approach is presented by applying the MesoCASiMiR module of the CASiMiR modeling system. In this way, the habitat suitability for the mayfly *Baetis rhodani* was modeled at a mesohabitat scale in the river Zwalm (Flanders, Belgium). The model inference system was based on fuzzy logic. Fuzzy variable sets and rules were derived from expert knowledge and from a database of biological samples. The suitability of the different mesohabitats for *Baetis rhodani* in the river Zwalm could be reliably described by three hydromorphological variables (depth, velocity and dominating substrate) and by the oxygen concentration. As a result, the habitat suitability was calculated and a habitat map of the studied reach was generated. Validation of this map was performed by biological samples at different sites along the reach, indicating that predicted habitat suitability was closely correlated to the observed abundances in most of the sampling sites. Due to the universality of the MesoCASiMiR module, the presented approach is applicable on other rivers and can be used for quick assessment of the ecological river status. This allows identification of the bottlenecks in the river basin and definition of restoration options. By adjusting the input parameters, the model can predict the impact of these restoration actions at a mesohabitat scale. Due to its transparent design and graphical user interface, the model has proved to be a useful tool for river management.

Keywords: Ecological modelling; Meso-scale habitat suitability, fuzzy logic, habitat mapping.

1. INTRODUCTION

Since the eighties, river management in Flanders is mostly conducted at the basin level, using instruments as wastewater treatment plants and enforced effluent standards. Although these measures resulted in a significant improvement of the chemical and the ecological river quality [VMM, 2003], lots of small-scale efforts as remeandering, flood plane restoration and fish passages are still needed to meet the aim set by the Water Framework Directive [EU, 2000]. To allocate these efforts in an efficient way, good river management should be based on a reliable assessment of the ecological bottlenecks in the river basin, at the micro- or mesoscale level.

Several microhabitat models only describe a small reach of the river [Alfredsen, 1997] and extrapolation of this microhabitat model to the river basin scale introduces a high level of uncertainty [Maddock, 1999]. Furthermore, ecological assessment at the micro scale level is very time consuming and thus less suitable for river basin management. Therefore, intermediary methods between the micro- and the macroscale level were developed [Borsányi, 2002], assuming that the river consists of hydromorphological units (HMUs) [Thickner, 2000]. The assessment of the ecological river status at this mesoscale level avoids the problems of time efficiency and upscaling [Maddock, 1999] and is therefore a suitable approach for good river basin management [Parasiewicz, 2003].

The ecological status of rivers is assessed by several mesohabitat models based on fish: MesoHABSIM [Parasiewicz, 2001], Habitat Mapping [Maddock & Bird, 1996], MesoCASiMiR and Habitat [Alfredsen, 1997]. Unfortunately, due to severe disturbance of the aquatic ecosystem by human activities, fish communities are severely reduced in the Zwalm River basin and in the rest of Flanders.

Furthermore, monitoring of ecological river quality in Flanders is done by the Flemish Environmental Agency [VMM] based on macroinvertebrates while the Water Framework Directive aims at a good quality of this community.

Therefore, this paper attempts to apply MesoCASiMiR, one of the present mesohabitat models, on macroinvertebrates, in order to create a first step towards the assessment of the ecological status of Flemish running waters at the mesoscale. MesoCASiMiR is a module of the CASiMiR modelling system [Jorde, 1996; Schneider, 2001], based on fuzzy logic [Zadeh, 1965]. Besides this modelling approach, a practical method to represent the results is proposed.

2. MATERIAL AND METHODS

2.1 Study site

The Zwalm river basin is part of the Scheldt river basin [Carchon & De Pauw, 1997]. The Zwalm River has a length of 21.75 km and its river basin has a total surface of 11.650 ha. (Fig. 1). The average water flow (at Nederzwalm, very near the River Scheldt) is about 1 m³/s.

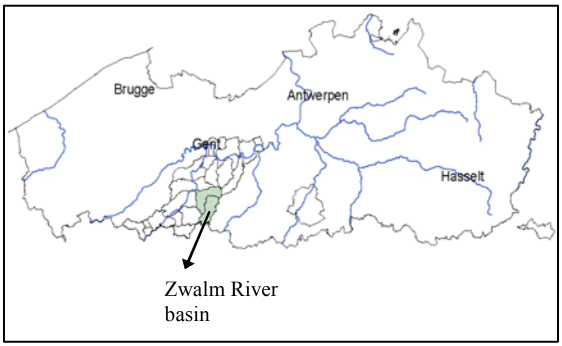

Figure 1. Location of the Zwalm River basin in Flanders

Water quality in the Zwalm river basin improved in the last years, due to investments in sewer systems and wastewater treatment plants [VMM, 2003]. Nevertheless, most parts of the river are still polluted by untreated urban wastewater discharges and by diffuse pollution originating from agricultural activities. Also structural and morphological disturbances are numerous [Carchon & De Pauw, 1997]. Weirs for water quantity control obstruct fish migration and are one of the most important ecological problems within the river basin. Therefore an in-depth study has been made on the development of fish migration channels and also natural overflow systems to reach an ecologically friendly water quantity management in the near future [Soresma, 2000]. Some upper parts of the watercourses in the Zwalm river basin are colonized by very rare fish species and several vulnerable macro-invertebrates.

2.2 Data Collection and Processing

Fuzzy rules describing the habitat suitability for *Baetis rhodani* were derived from biological data collected in the Zwalm River. All data were gathered during 5 consecutive years between August and September (2000 – 2004). At each of the 323 studied sites, 10 m of the present mesohabitat was sampled by means of 5 minutes kick sampling, using a standard handnet with mesh size 500 µm [NBN, 1984] and by in situ exposure of artificial substrates [De Pauw et al., 1983]. The number of present *Baetis rhodani* was expressed as absolute presence. In order to correct for different river width and thus for different sampling areas, the presence of *Baetis rhodani* was expressed as weighted presence using:

$$\textit{weighted presence} = \frac{\textit{absolute presence}}{10 \times W_{avg}} \qquad (1)$$

in which W_{avg} is the average width of the sampled river stretch. This weighted presence was used to express the habitat suitability for *Baetis rhodani*.

Structural and physical variables were measured to describe the different mesohabitats (Table 1). Flow velocity was determined using a propeller flow velocity meter (Höntzsch ZS25GFE). For each 10 m stretch, flow measurements were performed at 40 % of depth on 15 points, divided over 5 transects. Each transect consisted of 3 equidistant points, forming a uniform grid. The dominating substrate was visually assessed and expressed in 4 classes. Field measurements were performed for dissolved oxygen (OXI 330/SET).

Table 1. Measured variables at each sampled river stretch.

Variable	Unit
Flow velocity (m/s)	m/s
Dissolved oxygen	mg/l
Dominating substrate	4 classes (from 1 = pebble to 4 = loam/clay)
Depth	m

2.3 Software

The constructed fuzzy rules were implemented in the MesoCASiMiR module of the CASiMiR modelling system [Jorde, 1996; Schneider, 2001; Schneider et al., 2001]. This module in its current version was developped as an extension of ArcView GIS 3.3 (ESRI) and is not restricted to commonly used habitat parameters as flow velocity, depth and substratum but is designed to be used more universally. It can handle any habitat parameter, which can be defined as a property of a GIS polygon, for the assessement of habitat suitability. The model is adaptable in the way that habitat parameters itself and also their classfication in terms of fuzzy sets (see also 3.1) can be defined specific to the site, the investigation goals or the data availability. While e.g. for the investigation of benthic habitats flow velocity substratum and water quality parameters can be chosen, for fish habitat investigations morphological properties can be of higher importance for habitat assessment.

Any rule combining the classified habitat parameters with a habitat suitability (also classified by the use of fuzzy sets) can be used for the description of habitat preferences. An example of a rule is "IF flow velocity is 'high' AND concentration of dissolved oxygen is 'low' AND … is 'very high' AND … is 'very low' … THEN habitat suitability is 'medium' ". The number of five fuzzy sets, available for definition in the model (e.g. 'very low', 'low', 'medium', 'high', 'very high'), is assumed to be sufficient to describe most habitat preferences adequately.

2.4 Model evaluation

Evaluation of the results was based on two criteria, the percentage of Correctly Classified Instances (CCI) and the weighted Kappa (κ) [Cohen, 1960; Fleiss & Cohen, 1973], which were derived from the confusion matrix [Fielding & Bell, 1997]. The CCI is defined as the number of sites where the modelled habitat suitability class was the same as the monitored one, divided by the total number of sites. The weighted Kappa is a simply derived statistic that measures the proportion of all possible habitat suitability classes that are predicted correctly by a model after accounting for chance.

3. RESULTS

3.1 Fuzzy sets and rulebase

Fuzzy sets and rules were derived from expert knowledge [Adriaenssens et al., in prep.] and from the collected data based on three fold cross validation, using two third of the dataset, randomly chosen. Each variable was divided in a number of fuzzy sets (Table 2), while 192 IF…THEN rules were constructed based on these sets.

Table 2. Number of fuzzy sets in which each variable was divided.

Variable	Number of fuzzy sets
Flow velocity (m/s)	4
Dissolved oxygen	3
Dominating substrate	4
Depth	4
Habitat suitability	4

The rules and sets were implemented in the MesoCASiMiR module of the CASiMiR modelling system. The output of this module, the weigthed presence of *Baetis rhodani*, was converted to the habitat suitability class to which it belonged the most. The fuzzy rules were evaluated by comparing the modelled and the actual habitat suitability class for the remaining one third of the dataset, containing 107 records (Table 3). Based on this confusion matrix, a percentage of CCI of 64.5 % and a weighted Kappa of 0.410 were obtained.

Table 3. Confusion matrix based on the constructed rule base (HS= Habitat suitability).

Predicted HS class	1	2	3	4
Monitored HS class				
1	62	20	1	0
2	6	4	0	0
3	3	3	2	1
4	3	0	1	1

3.2 Habitat suitability mapping

The created fuzzy rules were applied to predict the habitat suitability for *Baetis rhodani* in a 5 km reach of the Zwalm River. This reach was divided in different stretches, each stretch being defined by one HMU. No distinction was made between different mesohabitats within one stretch, as in most stretches, variation of the mesohabitat was not significant. Along the studied reach, 3 hydromorphological variables (depth, flow velocity, dominant substrate) and 1 physical chemical variable (dissolved oxygen) were assessed in each stretch. No significant flow change occurred during the measuring period. The habitat suitability for *Baetis rhodani* was modelled for each mesohabitat based on the created fuzzy rules and these measurements. Visualisation of this habitat suitability was done by means of a habitat suitability map consisting of different polygons, each describing one mesohabitat (Fig. 2).

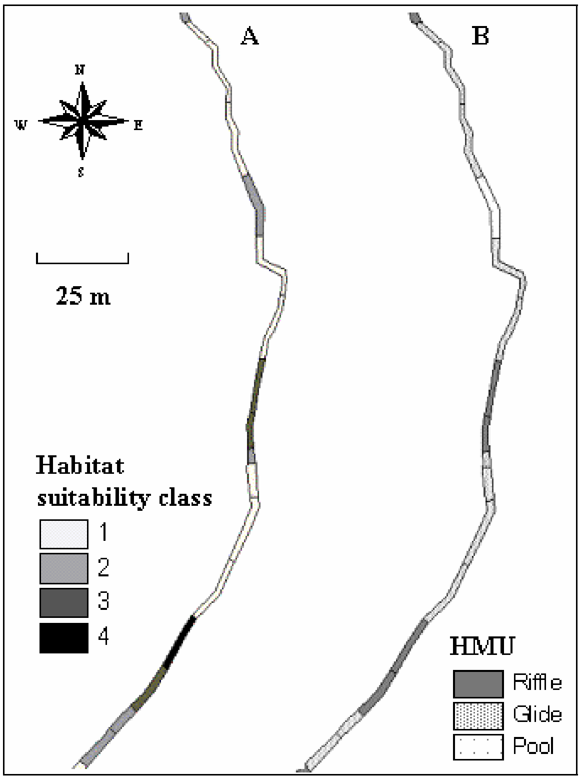

Figure 2. Habitat suitability map for *Baetis rhodani* (A) of a 250 m reach of the studied stretch and HMUs in this reach (B)

Validation was performed by comparison of the modelled results with biological samples of 24 mesohabitats. These samples were not included in the fuzzy rule development process. The resulting percentage of CCI and the weighted Kappa were respectively 62.5 % and 0.554.

Table 4. Confusion matrix based on the validation of the Habitat suitability (HS) map

Predicted HS class	1	2	3	4
Monitored HS class				
1	12	3	1	0
2	2	1	1	0
3	0	2	1	0
4	0	0	0	1

4. DISCUSSION

Most prediction errors occurred due to the prediction of a higher habitat suitability for *Baetis rhodani* than the monitored one. Absence of *Baetis rhodani* can be determined by other variables that were not included in the fuzzy rules, but as well by the limited monitoring efficiency and (re)colonization by this species. Therefore, overestimation of the habitat suitability indicates that the concerned habitat is suitable for *Baetis rhodani* concerning the 4 studied variables, and this is not necessarily a prediction error. For instance, the dissolved oxygen was included in the variable set to take into account the trophic status of the river. Previous research states that conductivity could also have an important effect on macroinvertebrate presence [D'heygere, 2003; Adriaenssens et al., in prep.]. Furthermore, biotic interactions were not included in the fuzzy model although these can also play an important role. As a result, the percentage of correctly classified instances would rise up to 85.1% when the predictions overestimating the habitat suitability would also be considered correct.

The data were unequally divided over the four suitability classes, which is also reflected in the confusion matrix resulting from rule validation. This disproportion is due to the fact that a significant part of the study site is severely impacted by human activities. Only a small part of the river basin contains reference sites, situated in some of the least disturbed brooks in Flanders. This results in more prediction errors in the lower HS classes (1&2) than in the higher ones.

Comparison of the derived rules with expert knowledge derived from literature [Adriaenssens et al., in prep.] indicated that the used sets and rules are not necessarily transferable between different rivers. It is clear that each river consists of specific conditions as geomorphology, (micro)climate, typology,… In that way, the set of key variables determining habitat suitability for macroinvertebrates can be different for each river. Furthermore, the range of the concerned variables changes for different rivers, resulting in other sets and rules.

Generalisation of rules might only give an indication of the impact of some variables on macroinvertebrate habitat suitability. In that way data collection is inevitable in order to establish fuzzy rules, which can be used as a river "blueprint". Once this template is constructed, maintenance can easily be performed by regular rule and set validation. This theory is emphasized by the fact that there is a small difference between the percentage of CCI resulting from rule validation and from map validation, although the sites used for map validation were not included in the rule development process.

In the modelling process, data collection and development of rules and sets require the most efforts. In order to increase efficiency, rules could be derived from the data in a faster model driven way, for instance using Artificial Neural Networks to derive the most important parameters [Dedecker et al., 2004] and Hillclimbing to set the optimal rule base [Van Broekhoven et al., 2004].

In a next step, the presented approach can be used for prediction of the ecological impact of different river restoration options. Restoration decisions are nowadays often based on intuition rather than on rigorous science [Muotka &

Laasonen, 2002]. If an objective of a river restoration option is not obtained, efforts are lost because the measures are already taken. This justifies the need for approaches that can give a reliable indication of the effect on river biology. Ecological models are a powerful tool and can be used for this purpose. Furthermore, modelling will also allow comparing the effects of alternative mitigation options. This will aid river managers in selecting an optimal set of restoration options to obtain a desired ecological quality in a river system. Moreover, implementing such ecological models for macroinvertebrates in a Decision Support System will allow river managers to weigh conflicting demands of different stakeholders such as households, farmers, nature developers, water quantity managers,…

5. CONCLUSIONS

Due to the universality of the MesoCASiMiR module, the presented approach is applicable on other rivers and can be used for prediction of the impact of restoration options at a mesohabitat scale. To achieve this, biological sampling and expert knowledge have to provide fuzzy rules, which can act as a blueprint for the studied river stretch. Furthermore, macroinvertebrates are particularly interesting for assessment of the ecological river status in Flanders, as river degradation has severely reduced fish populations. Other important variables describing this ecological status should be revealed by further research on other macroinvertebrate indicator species. This could result in better and more reliable decisions in integrated river management.

ACKNOWLEDGEMENTS

Collaborative work between Ghent University and sje Engineering was enabled by COST 626 action fundings. The first author is a recipient of a Ph. D grant financed by the Special Research Fund (BOF) of Ghent University.

REFERENCES

Adriaenssens, V. , Verdonschot, P.F.M., De Baets, B. and De Pauw, N., Knowledge base development for macroinvertebrate habitat suitability modelling, *Freshwater Biology*, in prep.

Alfredsen, K., A modelling system for estimation of impacts on fish habitat, Proceedings of the XXVIIth IAHR congress, Environmental and coastal hydraulics: protecting the aquatic habitat, San Francisco, USA, 1997.

Borsányi, P., A meso-scale habitat classification method for production modelling of atlantic salmon in Norway, paper presented at IFIM Users workshop, Fort Collins, C.O., USA, 2002.

Carchon, P. and De Pauw, N., Detailstudie en uitwerking van een methodiek inzake het kwaliteitsbeeld van oppervlaktewateren, Study conducted under authority of the Flemish Environmental Agency (VMM), 1997.

Cohen, J., A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*, 37-46, 1960.

Dedecker, A.P., Goethals, P.L.M., Gabriels, W. and De Pauw, N., Optimization of artificial neural network (ANN) model design for prediction of macroinvertebrates in the Zwalm river basin (Flanders, Belgium), *Ecological Modelling*, *174*, 161-173, 2004.

De Pauw, N. and Vanhooren, G., Method for biological quality assessment of watercourses in Belgium, *Hydrobiologia, 100*, 153-183, 1983.

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

EU, Directive of the European Parliament and of the council 2000/60/EC establishing a framework for community action in the field of water policy, European Union, The European Parliament, The Council, PE-CONS 3639/1/00 REV 1 EN, 2000.

Fielding, A.H. and Bell, J.F., A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservations*, *24*, 38-49, 1997.

Fleiss, J.L. and Cohen, J., The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability, *Educational and Psychological Measurement*, *33*, 613-619, 1973.

Jorde, K., Ökologisch begründete, dynamische Mindestwasserregelungen bei Ausleitungskraftwerken, Ph.D. thesis, Stuttgart University, Stuttgart, 1996.

IBN, Biological water quality: determination of the biotic index based on aquatic macroinvertebrates, NBN T92-402, Institut Belge de Normalisation (IBN), 1984.

Maddock, I., The importance of physical habitat assessment for evaluating river health, *Freshwater Biology*, *41*, 373-391,1999.

Maddock, I. and Bird, D., The application of habitat mapping to identify representative PHABSIM sites on the River Tavy, Devon, UK. *Proceedings of the 2nd International Symposium on Habitats and Hydraulics, Vol. B*, 203-214, Quebec, Canada, 1996.

Parasiewicz, P., MesoHABSIM – a concept for application of instream flow models in river restoration planning, *Fisheries, 29*(9), 6-13, 2001.

Parasiewicz, P., Upscaling: Integrating Habitat Model into River Management, *Canadian Water Resources Journal, 28*(2), 1-17, 2003.

Schneider, M., Habitat- und Abflussmodellierung für Fliessgewässer mit unscharfen Berechnungsansätzen – Weiterentwicklung des Simulationsmodells CASIMIR, Ph.D. thesis, Stuttgart University, Stuttgart, 2001.

Schneider, M., Jorde K., Zoellner F. and Kerle F., Development of a user-friendly software for ecological investigations on river systems, integration of a fuzzy rule-based approach, *Proceedings Environmental informatics 2001, 15th International Symposium, Informatics for Environmental Protection*, ETH Zurich, Switzerland, 2001.

Soresma, Environmental impact assessment report on the development of fish migration channels and natural overflow systems in the Zwalm River basin, Soresma advies- en ingenieursbureau, Antwerp, 2000.

Thickner, D., Armitage, P.D., Bickerton, M.A. and Hall, K.A., Assessing stream quality using information on mesohabitat distribution and character, *Aquatic Conservation: Marine and Freshwater Ecosystems 10*, 179-196, 2000.

Van Broekhoven, E., Interpretability-preserving genetic optimisation of linguistic terms in fuzzy models for fuzzy ordened classification: an ecological case study, *Fuzzy Sets and Systems*, in prep.

VMM, Algemeen Waterkwaliteitsplan 2-6, Bovenschelde, Vlaamse Milieumaatschappij (VMM), 70 pp., Aalst, 2003.

Zadeh, L.A., Fuzzy sets, *Information and Control*, *8*, 338-353, 1965.