Hoang Thi Thu Huong

Monitoring and assessment of macroinvertebrate communities in support of river management in northern Vietnam

UNIVERSITEIT GENT

FACULTY OF BIOSCIENCE ENGINEERING

Promoters: Prof. Dr. ir. Peter Goethals
Em. Prof. Dr. Niels De Pauw
Laboratory of Environmental Toxicology and Aquatic Ecology
Department of Applied Ecology and Environmental Biology
Ghent University

Dean: Prof. Dr. ir. Guido Van Huylenbroeck

Rector: Prof. Dr. Paul Van Cauwenberge

Hoang Thi Thu Huong

Monitoring and assessment of macroinvertebrate communities in support of river management in northern Vietnam

Thesis submitted in fulfilment of the requirements for the degree of Doctor (PhD) in Applied Biological Sciences

Dutch translation of the title:

Monitoring en evaluatie van macro-invertebratengemeenschappen ter ondersteuning van rivierbeheer in Noord-Vietnam

Please refer to this work as follows:

Hoang, T.H., 2009. Monitoring and assessment of macroinvertebrate communities in support of river management in northern Vietnam. PhD thesis, Ghent University, Gent, Belgium. 235pp

This research was performed at:

Laboratory for Environmental Toxicology and Aquatic Ecology, Department Applied Ecology and Environmental Biology, Faculty of Bio-engineering Sciences, Ghent University
J. Plateaustraat 22, B-9000 Gent (Belgium)
Tel. 0032 (0)9 264 37 65, Fax. 0032 (0)9 264 41 99

Institute for Environmental Science and Technology, Hanoi University of Technology,
Dai Co Viet 1, Hanoi (Vietnam)
Tel. 0084 (0)4 38 681 686, Fax. 0084 (0)4 38 693 551

ISBN-number: 978-90-5989-321-4

Aknowledgements

It is the most difficult part of the thesis, which I wrote first and revised last. But I still have a feeling that I could not express my gratitude to everyone to help me in making this thesis realised.

In the first place I would like to record my deep gratitude to my local promoter Prof. Dr.ir. Dang Kim Chi for her advice from the very early stage of this research as well as giving me extraordinary experiences and support through out the work. Above all and the most needed, she provided me unflinching encouragement and support in various ways. Her truly scientific intuition has made her as a constant oasis of ideas and passions in science, which exceptionally have been inspiring and enriching my growth as a student, researcher, teacher and scientist. Her passion and kindness had helped me in many ways to keep balance between the career and my family so that I never have to make a choice between them. It is so great to have an example in life to look up to.

I gratefully acknowledge my principal promoter Prof. Dr. ir. Peter Goethals for his advice, supervision, and crucial contribution, which made him a backbone of this research and so to this thesis. His involvement with his originality and creativeness has also affected my intellectual maturity that I will benefit from, for a long time to come. I really appreciate his care and support of my study for the last many years. Peter, I am grateful in every possible way and hope to keep up our collaboration as well as friendship in the future.

I would like to express my sincere gratitude to my promoter Em. Prof. Dr. Niels De Pauw. He provided me with helpful suggestions, important advice and constant encouragement during the course of this work. I would also like to thank his spouse Mrs. Françine Moonen for the hospitality and the many interesting discussions about life we had over delicious meals and in the warm Flemish environment of your houses. I feel so privileged to know such wonderful people.

I am deeply indebted to Dr. Koen Lock for his advices, constructive guidance and critical views on my study. His criticism, patience and valuable suggestions helped to improve the quality of this study and this manuscript in many ways.

Collective and individual acknowledgments are also owed to Prof. Colin Jassen and all colleagues at the Laboratory for Environmental Toxicology and Aquatic Ecology. The group offered a refreshing, helpful and memorable atmosphere. Special thanks to Guido, Mark, Marianne, Martine and Sigrid for all their supports to make the lab be my best working place. Thank you Lien for sympathy and encouragement all the time.

I am so proud to be a member of the Aquatic Ecology research group. Thank all past and present members Andy, Ans, Argaw, Diederik, Eleni, Koen, Luis, Marjolein, Minh, Pieter, Rahmat, Tom, Veronique, Wim and Wouter for giving me such a pleasant time when working together. Many thanks go particular to Ans and Argaw for their firm support both professional and mental during the rush to finish this manuscript on time. Argaw, you are a true friend I have found for life.

Thanks are given to my colleagues in the Institute for Environmental Science and Technology. I would like to thank Associate Prof. Nguyen Ngoc Lan, Dr. Nghiem Trung Dung and colleagues in the Department of Environmental Enginering, especially Minh, for all their supports and taking care of my teaching during the hard time working in the field and lab, abroad study and thesis writing.

This work would never be possible without valuable support of the colleagues in the R&D Lab at the INEST and the Centre for Industrial Environmental Monitoring and Polution Control. My gratitude goes to Associate Prof. Huynh Trung Hai and Hưng for all arrangements to provide me an excellent working condition. Thanks also go to Hien and Le Thu for their caring of all chemicals and facilities I had required, even in the funniest time of the days, even if these had not been in the store. Thank to Thu Phương for her courage to join me the first field trips, which were really adventures without a prescribed plan. I am deeply indebted to all staffs in the Lab, particularly to Dao, Du, Dzung, Huong, Lam, Linh, Thuy, as they were always ready to support, ready to share equipment, prepared chemicals and facilities whenever I needed.

My heartiest thanks are due to Prof. Tran Van Nhan for his continuous support, encouragement and sympathy during the hardest stages of my study. His strong belief in me always made me "Stand up and fight" and I really appreciate for that. I have learned a lot from his experiences in managing bunch of works within a limited time available.

And I would like to express my exceptional thank to the our chief driver Anh Ngoc, for his enthusiasm, his constant readiness in supporting me in over 10 hard field trips during the last 4 years. His enthusiasm and songs were always nice encouragement during our long and exhausted field trips.

Very special thanks are given to my students Tuan Anh, Anne, Chung, Ha, Hau, Hoa, Hong, Hung, Khanh, Lien, Le, Mai, Nga, Ninh, Phan, Son, Thao, Thang, Truong, Thuy, Tien, Tu, Van, Viet, especially Lien and Viet from the batch K48-Environmental Engineering for their enthusiasm and help during our hard but productive and cheerful monitoring campaigns disregarding how terrible weather was. We had learned a lot from each other during these field trips. Some of them started working with me while they were students and continued their support as my colleagues. I indeed appreciated their sharing time, health and, the most important, their enthusiasm to ecological research works.

My cordial appreciations go to Prof. Mai Dinh Yen from the Hanoi National University and Assoc. Prof. Ho Thanh Hai from the Institute of Ecology and Biological Resources, Vietnam Academy of Science and Technology. I would like to thank them for their valuable suggestions, and especially for their supportive discussions and openness to a young scientist. They helped me a lot to be self-confident working with a challenging topic, which was from beginning far from my background but was so exciting to get insight.

Many thanks to my official jury committees Prof. Patrick Van Oostveldt (UG), Prof. Dr. Ludwig Triest (VUB), Prof. Dr. ir. Peter Bossier (UG), Dr. Koen Lock (UG) and Dr. ir. Ans Mouton (INBO) for their valuable remarks. This work has dramatically improved thanks to your co-operation. Thank is also given to Dr. N.D. Trung from the Institute of Geoscience and Mineral Resources in Vietnam for his support in providing me a GIS map and helping me with some visualisations.

The financial support from the Belgian Technical Cooperation (BTC) for 4-year PhD sandwich program is highly acknowledged. Thanks to Daphney Windey, Joachim De Frayer and Liesbet Vastenavondt in the BTC headquarter Brussels and all staffs of the BTC Hanoi office for their perfect works to take care of my stays in Belgium as well as support my works in Vietnam. I also acknowledge the Hanoi University of Technology for financial support with 3 university projects and all support during my PhD sandwich program.

Thanks are given to Father Charles de Hemptinne and staffs of the OBSG (Marleen, Isabelle and Annemie) for their hospitality to provide me the best place to live during my stay in Gent. Discussions with Father Charles and participations in social activities of the Club Van Eyek were great and enjoyable opportunities for me to discover culture and people of Belgium, particularly in the Flemish society, that I consider another great outcome during my study in Gent.

Many thanks go to all my friends in Hanoi, H1-group, VinaSAG, VGAC, in Facebook, Blog and from all continents in the world for helping me with documentation, database and especially for their support in keeping me in the world. My keen appreciations go to Hendrik, Hoa, Bình, Linh, Anh Trung, Phuong, Phuc, Anh Tuan, Ha, Tu, especially Thuy and Duc for all their cares and efforts to relieve me of stressfulness during my stay in Belgium. Life is never easy but it is much easier when we are surrounded by friends.

This work is dedicated to my little daughter Xiren (Huong Lien) and my mother, who sacrificed a lot during my PhD study. Their continuous encouragement and absolute trust in me are the most precious things I've ever had and will have. Thank to my dearest Viet Dung for his understanding, encouragement, enthusiasm and love to support all my achievements by taking the entire load off my shoulders. Every woman in the world would be happy to share her life with a man like you. I appreciate the fate to have you in my life. I would never finish this work without supports from my extended family behind.

I would like to appreciate other people, relatives, friends, colleagues, who was also important to the successful realisation of this thesis, as well as express my apology that I could not mention personally one by one.

Thank you all.

Hoang Thu Huong

Gent, 2nd September 2009.

Table of contents

List of Abbreviations

AQEM	The integrated Assessment system for the ecological Quality of streams and rivers throughout Europe using benthic Macroinvertebrates
ASPT	Average Score Per Taxon
AUSRIVAS	Australian River Assessment Scheme
BASTAF	BAffled Septic Tank with up-flow Anaerobic Filter
BBI	Belgian Biotic Index
BEAST	Benthic Assessment of Sediment
BMWP	Biological Monitoring Working Party
BMWP Thai	Biological Monitoring Working Party of Thailand
BMWP Viet	Biological Monitoring Working Party of Vietnam
BOD_5	Biochemical Oxygen Demand (5 days)
CCA	Canonical Correspondence Analysis
CCI	Correctly Classified Instances
COD	Chemical Oxygen Demand
CT	Classification Tree
DCA	Detrended Correspondence Analysis
DO (%)	Dissolved Oxygen (percentage of dissolved oxygen in water)
DO (mg/l)	Dissolved Oxygen (mg oxygen per liter)
DONRE	Department Of Natural Resources and Environment
EPA	Environmental Protection Agency
EPT	Ephemeroptera, Plecoptera, Trichoptera richness index
EQI	Ecological Quality Index
EQR	Ecological Quality Ratio
FAO	Food and Agriculture Organisation
GA	Genetic Algorithm
HKHbios	Hindu Kush-Himalayan biotic index
ICP-MS	Inductively Coupled Plasma Mass Spectrometry
ISRIC	International Soil Reference and Information Centre
K	Cohen's Kappa
MDS	Multi Dimensional Scaling
MRC	Mekong River Commission
NISF	National Institute for Soil and Fertiliser
NEPBIOS	Nepalese Biotic Score
P/R class	Pool/Riffle class

PCA	Principal Component Analysis
PCoA	Principal Coordinates Analysis
PCF	Pruning Confidence Factor
QA/QC	Quality Assurance/ Quality Control
RDA	Redundancy Analysis
RIVPACS	River Invertebrate Prediction and Classification System
SASS	South African Score System
SWEPACS	Sweden Prediction and Classification System
SI	Saprobic Index
SIGNAL	Stream Invertebrate Grade Number Average Level
SMO	Sequential Minimal Optimisation
STOWA	Stichting Toegepast Onderzoek WAterbeheer (Foundation for Applied Water Research)
SVM	Support Vector Machine
TBI	Trend Biotic Index
TCVN	Tieu Chuan Viet Nam
	(Vietnamese environmental standard)
Total N	Total Nitrogen (mg Nitrogen per liter)
TN-sediment	Total Nitrogen in sediment (mg Nitrogen per g Dry Matter)
Total P	Total Phosphorus (mg Phosphorus per liter)
TWINSPAN	Two-Way INdicator SPecies ANalysis
UNESCAP	United Nations Economic and Social Commission for Asia and the Pacific
UN/ECE	United Nation Economic Commission for Europe
VEA	Vietnam Environment Administration
VEM	Vietnam Environment Monitor
VEPA	Vietnam Environmental Protection Agency
VLAREM II	VLAams REglement betreffende de Milieuvergunning
	Flemish regulation concerning environmental licenses
WEKA	Waikato Environment for Knowledge Analysis
WFD	Water Framework Directive
WQI	Water Quality Index
WWTP	Waste Water Treatment Plant

"The flow of rivers is part of a greater flow, the planet's water cycle, which sustains not only the flow of water but the entire web of life. Ultimately, the condition, or health, of the aquatic biota is the best means of understanding and controlling humans' impact on the Earth's watercourses and on the whole water cycle."

James Karr and Ellen Chu (2000)

General introduction

1. Problem definition

Running water is the most important freshwater resource in Vietnam, being used for a variety of life purposes. In recent years, however, the rapid socio-economic development has significantly affected the environment of river basins. Intensive agriculture, industrialisation and rapid urbanisation are the most important threats to deterioration of surface water quality (Fig. 1). Maintaining a high quality of running water has become an increasingly crucial issue in recent years, as greater demand has been placed on water resources (VEPA, 2006).

Figure 1. Waste (water) discharged to the Cau river at Phong Khe, Bac Ninh (Dang et al., 2005).

In running water management in Vietnam nowadays, the assessment of river water quality relies merely on physical chemical characteristics. Until now, biological assessment methods are not yet applied to determine the environmental status of water bodies. The Master plan on *'Development of a national environmental and resource monitoring network in Vietnam'* approved in 2007 defined parameters required in the national monitoring program. Chlorophyll-a and coliforms are the only included parameters related to biological properties of surface water monitoring (VEPA, 2007). However, interest was shown for biological assessment as several sites that were meeting all environmental standards were nevertheless characterised by a decreased biodiversity (Dang et al., 2002; Wilby et al., 2006).

The ecological quality of rivers and streams depends on their physical, chemical and biological properties. The latter are reflected by the types and abundance of living organisms present in the water. Biological monitoring is generally considered to provide a more integrated appraisal of the water and overall environmental quality (Hynes, 1960). Therefore, there is now a widespread recognition that not only chemical but biological analyses are required for an appropriate assessment of the river quality (Wright, 1995). Biological monitoring of the fresh water quality has only recently been initiated in Vietnam with some independent researches, training and public policy development regarding these techniques (Le, 2007; Morse et al., 2007). Consequently, there is an urgent need to develop comprehensive and cost-effective methods for biological monitoring and assessment of running waters in Vietnam (Thorne and Williams, 1997).

Recently, also ecological models are more and more applied to support river management. Ecological models can provide a better elucidation of river status in order to detect the cause of this status, as well as to optimise assessment methods and consequently help to setup cost effective monitoring programs (Vanrolleghem et al., 1999). Potential applications of models for information and decision support in river management are illustrated in Fig. 2.

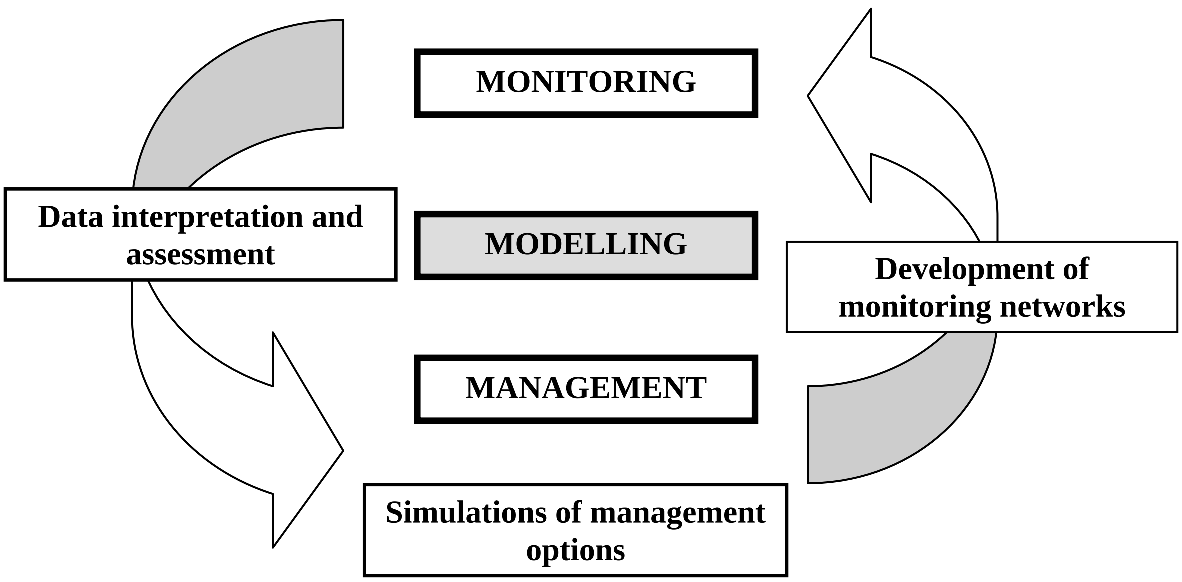


Figure 2. Potential applications of models in river management (Goethals and De Pauw, 2001).

Habitat suitability models, for example, which are one set of ecological models, have proved to be able to predict the impact of water quality on river ecology and can therefore be valuable tools for integrated river management (Poff et al., 2003). For development of these decision support tools, several model development techniques can be applied, such as classification trees (Dzeroski et al., 1997; Dakou et al., 2007), artificial neural networks (Dedecker et al., 2004, Goethals et al., 2007), fuzzy logic (Adriaenssens et al., 2004a), Bayesian Belief Networks (Adriaenssens et al., 2004b) and recently, support vector machines (Sanchez-Hernandez et al., 2007; Ribeiro and Torgo, 2008).

In the past few years, there has been a lot of interest in support-vector machines (SVMs) (Vapnik, 1995; Burges, 1998; Keerthi et al., 2001) because they yield an excellent generalisation performance for a wide range of problems (Keerthi et al., 2001). They also combine a good performance with an efficient training process (Decoste and Scholkopf, 2002; Guo et al., 2005). In sprite of this, only few applications of SVMs in freshwater ecology have so far been published (Shan et al., 2006; Sanchez-Hernandez et al., 2007; Ribeiro and Torgo, 2008).

Making a selection of restoration measures is one of the key tasks of river water managers, often requiring a lot of information concerning financial, social and technical issues, as well as determining the effectiveness of measures to improve the ecological status of the river. Presently, most of these restoration measures are however conducted without any quantitative prediction of their impact on the river ecology (Mouton et al., 2009). Hence, communicative tools are needed to bridge this gap between the results of bioassessment studies on water systems and the information currently used in water management (Roux et al., 2006).

The above challenges draw the core research goals for this PhD study. Obviously it was too ambitious to develop a comprehensive river management program including the application of biological monitoring. However, the approaches presented in this work can be a basis to support the development of biological monitoring methods aiming at an improved river management in Vietnam.

2. Scope and objectives

The present PhD work aimed at developing biological monitoring and assessment techniques for decision-support in river management in northern Vietnam.

The research focused on macroinvertebrates in streams and small rivers. The selected monitoring sites were characterised by a high diversity in disturbance as well as intensities, ranging from near pristine conditions to severely impacted systems. Different ecological indices were tested for their relevance to assess the environmental status of rivers in northern Vietnam. Relations between macroinvertebrate taxa/indices and river characteristics were analysed and based on different multivariate statistical and numerical methods. Moreover, a decision support system, the Water Framework Directive Explorer, was implemented to link human activities with the ecological river conditions and checked for the relevance of several restoration options.

In following paragraphs, the individual chapters of this thesis and the specific goals of the research are briefly described. The manuscript consists of 3 main parts as shown in Fig. 3, which are subdivided in 6 main chapters.

Chapter 1 describes the different approaches to depict relations between river characteristics and biological communities. The chapter also reviews the development and application of biological assessment methods based on macroinvertebrates in the East Asian countries.

Chapter 2 deals with materials and methods applied in this thesis. The chapter contains a description of the study sites in relation to the main anthropogenic impacts. The monitoring methods and the data collection techniques are described in detail. This chapter also provides also the background and specific settings of the applied statistical and numerical methods.

Chapter 3 explores the structure of the macroinvertebrate communities inhabiting in the Du river basin. Clustering and ordination techniques are applied to detect spatial and temporal patterns displayed by these biological communities. In addition, environmental variables responsible for striking shifts in the community composition are searched for on the basis of multivariate analyses.

Chapter 4 evaluates the performance of BMWP-Viet and other biotic indices applied in similar ecoregions for environmental assessment in the River Du basin. This chapter aims (1) to evaluate the performance of different biotic indices in relation to abiotic assessments, (2) to identify the limitations of the BMWP approach and the need for further improvement, and (3) to come up with practical recommendations for river management.

Chapter 5 focuses on the application of two model development techniques to analyse the relation between macroinvertebrates and river conditions, namely classification trees (CTs) and support vector machines (SVMs). These model development and habitat suitability studies focus on (1) prediction of the BMWP-Viet index and (2) habitat suitability models for some common taxa in the study area. The outcomes of the models can provide better insights to improve ecological indices in the future.

Chapter 6 aims to investigate an adaptation of the WFD-Explorer as a communicative tool for water quality management in Vietnam.

The thesis ends with a general discussion about the scientific and practical meaning of the research conducted and the future research prospects.

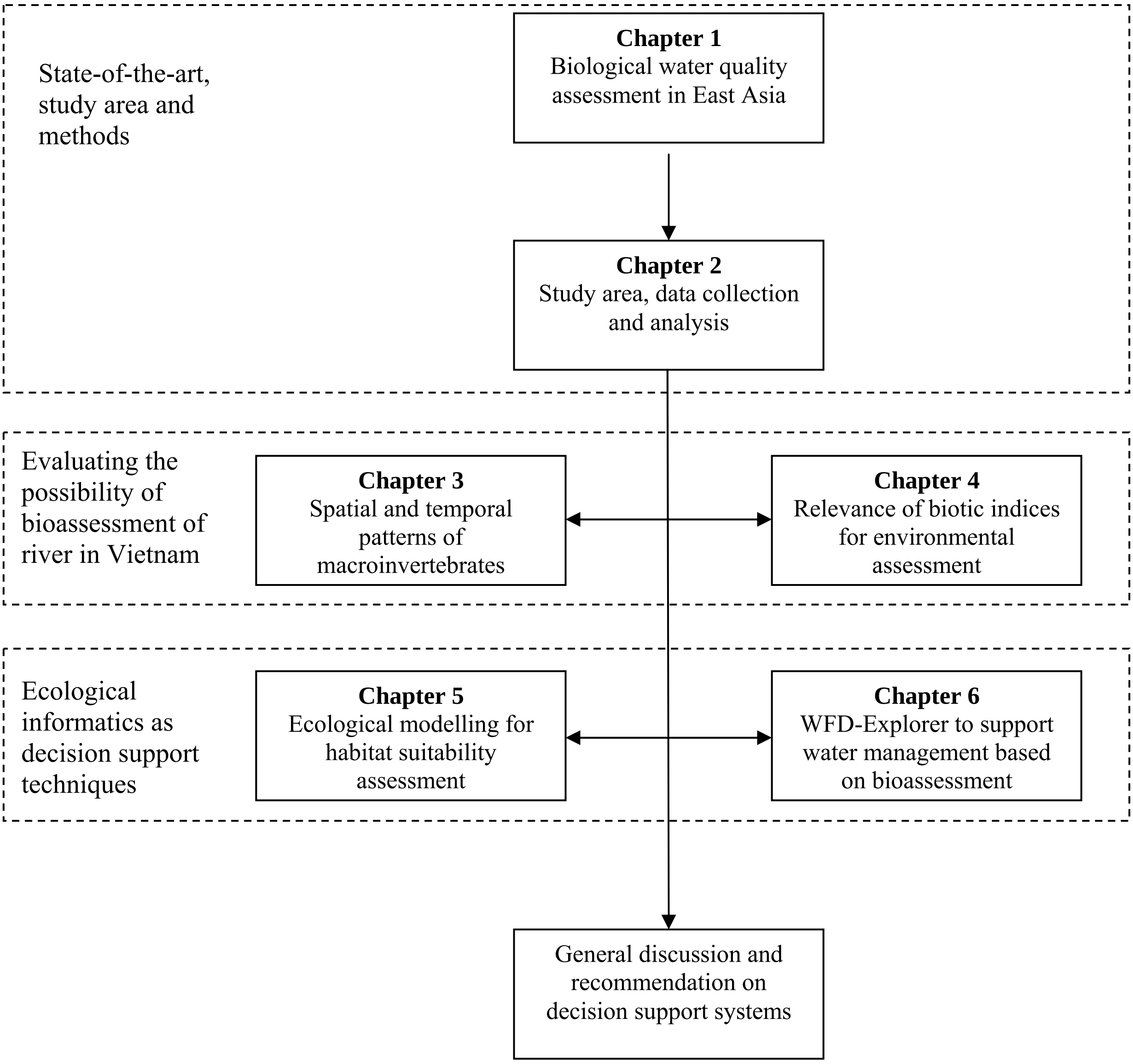


Figure 3. A road map through the dissertation.

Chapter 1

Biological water quality assessment in East Asia based on macroinvertebrate communities

1.1 Introduction

Biological assessment of rivers has proven to be a suitable alternative for physical and chemical methods (Hynes, 1970; Hellawell, 1986; Rosenberg and Resh, 1993; Loeb and Spacie, 1994). Because of this it is now applied in many countries. Biological monitoring can provide an estimate of the ecological water quality, but it may be particularly useful in developing countries as it is relatively cheap and easy to perform (Thorne and Williams, 1997). Several biological assessment methods have been standardised and included in national as well as regional monitoring programs (Barbour et al., 1999; Hering et al., 2003; De Pauw et al., 2006), serving as a basis for policy decisions concerning surface water management.

In recent years, numerous studies have been done to develop and apply the concept of biological assessment in temperate (Robinson et al., 2000; Usseglio-Polatera et al., 2000; Goethals and De Pauw, 2001; Butcher et al., 2003; Skoulikidis et al., 2004; Bij de Vaate and Pavluk, 2004; Lautenschlager and Kiel, 2005; Furse et al., 2006) as well as tropical areas (Capitulo et al., 2001; Hart et al., 2001; Dicken and Graham, 2002; Mustow, 2002, Baptista et al., 2007; Moya et al., 2007). In Asia, especially in developing countries, physical and chemical methods are predominantly used to assess stream and river water quality. Until present, there have only been a few studies testing bioassessment methods in Southeast Asia (Morse et al., 2007).

A survey of international journals from 1992–2001 dealing with freshwater ecology and limnology revealed that the representation of scientists from tropical Asia is extremely low (Dudgeon, 2003). Of a total of 4579 papers included in the survey, only 75 were authored or co-authored by scientists based in tropical Asia; i.e. 1.6 % of the total. Among the articles originating from tropical Asia, none was from Vietnam. However, the interest is growing and some studies were recently carried out about the diversity of aquatic insects in tropical rivers in Vietnam (Nguyen et al., 2004; Hoang and Bae, 2006; Jung et al., 2008).

Several studies encourage the use of biological methods for tropical streams and rivers (e.g. Hart et al., 2001; Fenoglio et al., 2002; Mustow, 2002; Weigel et al., 2002), pointing out that well-established assessment methods developed for temperate streams and rivers (mainly biotic indices and scores) have been successfully applied in tropical areas. Nevertheless, some adaptations are necessary in order to improve their reliability for use in East Asia. Differences between temperate and tropical running waters could be considered responsible for the need of such adjustments (Dominguez-Granda, 2007).

The main topic of this chapter is a review about the development and current efforts to implement macroinvertebrate bioassessment for surface waters in East Asia. Challenges and future developments of this topic in Vietnam are discussed. The development of decision support techniques based on biological indicators for monitoring, assessment and management of rivers in Vietnam is reviewed to set up the scene.

1.2 Biological river assessment based on macroinvertebrates

1.2.1 Advantages and disadvantages of using macroinvertebrates in East Asia

Streams and rivers have a complex nature, which can be explained by the three-dimensional geometry of channels with a long profile, a cross-section and mutual adjustment over a time scale (Allan, 1995). In river ecosystems, the physics, chemistry and biology of the water body are interrelated. Any substances introduced into a river are transported and transformed by physical, biological and biochemical processes. Interactions among chemical and physical processes in running water create conditions at a range of scales that strongly influence biological processes, especially benthic communities (Fig. 1.1).

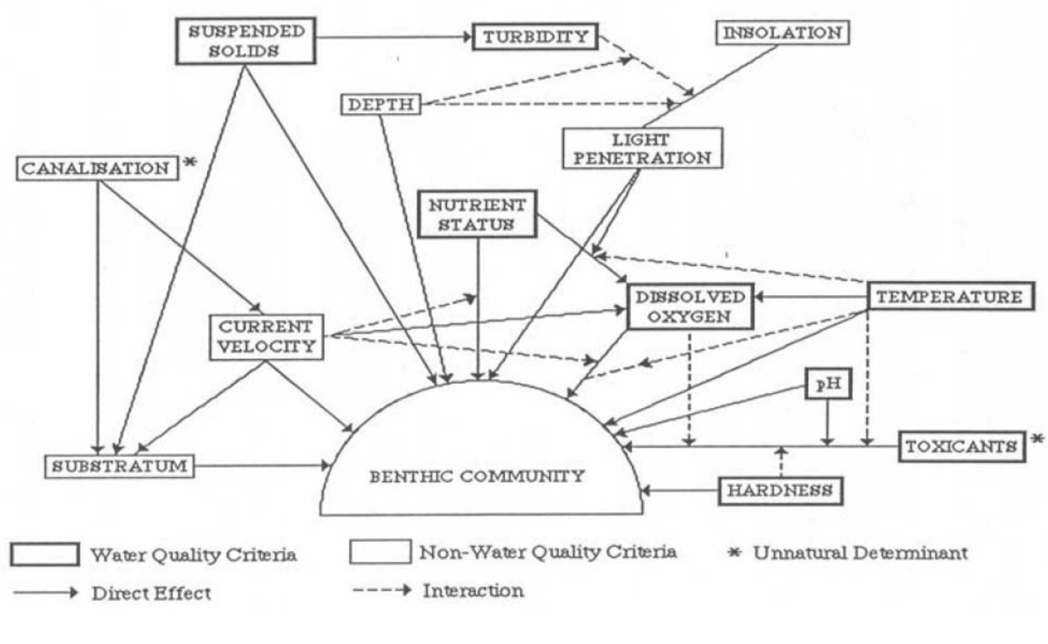


Figure 1.1. Water quality and non water quality determinants of benthic communities in rivers (after De Pauw and Hawkes, 1993).

Aquatic organisms integrate the effects of a variety of pollutants and even reflect short-term, critical fluctuations in water quality. Therefore, biological monitoring is essential to evaluate the environmental health of aquatic ecosystems (Cairns et al., 1993; Loeb and Spacie, 1994). US EPA (1987) recommended a focus on biological assessment of water quality. This recommendation was based on evidence that measurement of the biotic component of aquatic ecosystem provides information about environmental stress that is missed by periodic or continuous monitoring of physical and chemical factors.

Amongst aquatic organisms that can be used for bioassessment, macroinvertebrates have proved to be excellent indicators for the quality of freshwater stream habitats (Hellawell, 1986; Rosenberg and Resh, 1993; Davis and Simon, 1995; Hawkes, 1997). Advantages of macroinvertebrate as bioindicators were described in many previous studies (Ward and Stanford, 1983; Allan, 1995; Hauer and Resh, 1996; Giller and Malmqvist, 1998; De Pauw et al., 2006):

- they are ubiquitous and can be affected by environmental perturbations in different types of aquatic systems.
- the large numbers of species involved offer a spectrum of responses to environmental stresses.
- their sedentary nature allows effective spatial analyses of pollutants or disturbances.

Cranston et al. (1996) assessed all possible biological indicators by 11 criteria and compared them with an ideal rating (Table 1.1). The results showed that macroinvertebrates offer the rating closest to the ideal and they have been widely accepted as key indicators to assess water quality, especially in developing countries.

Table 1.1. Assessment of biological indicators of water quality (Cranston et al., 1996).

Indicator	Criterion										
	1	2	3	4	5	6	7	8	9	10	11
Ideal rating	H	L	H	H	H	H	H	H	H	-	H
Mammals	M	M	H	L	H	M	H	L	M	G	M
Reptiles	M	M	L	L	M	M	H	M	M	G	M
Amphibians	H	L	H	L	M	M	H	M	M	G	M
Water birds	M	M	H	M	L	L	L	L	L	G	M
Fish	M	M	M	M	M	M	M	M	M	GD	L
Plants	H	L	M	M	M	H	M	H	H	G	M
Macroinvertebrates	M	M	H	H	H	H	H	M	H	GD	M
Biomarkers deformity	H	M	H	M	H	M	H	H	M	G	L
Biomarkers asymmetry	H	M	H	M	H	M	H	H	M	G	L
Bioassays	H	H	H	M	M	H	M	M	M	D	L

Selection criteria:

1. Ease to capture (High, Medium, Low)
2. Total cost/ha (H,M,L)
3. Standard methods available (H,M,L)
4. Interpretation criteria available (H,M,L)
5. Significant at catchment scale (H,M,L)
6. Low error associated with measurement (H,M,L)
7. Response to disturbance (H,M,L)
8. Stable over period of measurement (H,M,L)
9. Mappable (H,M,L)
10. Generic (G) / Diagnostic (D) application
11. Context data available (H,M,L)

However, the use of macroinvertebrates as monitors of water quality has also its limitations (Giller and Malmqvist, 1998; Rosenberg and Resh, 1993; Linke et al., 1999; Sandin et al., 2000; D'heygere et al., 2002; De Pauw et al., 2006):

- possibility of wrong identification, particularly in early life stage of insect larvae;
- quantitative sampling is also difficult because of their non-random distribution in the river bed;
- factors other than water quality, such as current velocity and nature of the substratum, are also important determinants of benthic communities;
- restricted geographic distribution of macroinvertebrates, the incidence and frequency of occurrence of some species being different in rivers throughout the region; it is therefore not possible to have a universal system of biological assessment based on the response of the same taxa.

An appreciation of these considerations should be taken into account during adaptation of bioassessment methods for macroinvertebrates used in temperate regions for their application in tropical countries including East Asia.

1.2.2 Assessment approaches based on macroinvertebrates

Several assessment methods based on macroinvertebrates have been developed for stream assessment in temperate areas, mainly in Western Europe and North America and adapted for use in other regions including Australia, some countries in South America, Africa and Asia. Overviews can be found in Metcalfe (1989), De Pauw and Hawkes (1993), Rosenberg and Resh (1993), Sandin et al. (2000), AQEM (2002) and De Pauw et al. (2006). Metcalfe (1989), reviewing river assessment methods based on macroinvertebrates developed in Europe, categorised them on the basis of the approach employed to assess the response of the macroinvertebrate community to pollution. Hence, three approaches were clearly defined: named *saprobic*, *diversity* and *biotic approach*. Since then, new assessment methodologies have been developed, resulting in three new approaches referred to as *multimetric*, *multivariate* and *ecological quality ratio* approaches (De Pauw et al., 2006). These assessment approaches could also be applied for use in East Asia. In the following paragraphs, a brief overview of the six major approaches is presented.

1.2.2.1 Saprobic approach

The original system was as a 'class' system, in which each indicator organism was classified into different 'saprobia'. Such classifications are done in order to express their dependence on decomposing organic nutrients (Metcalfe, 1989).

The saprobic system has been developed to provide a numerical index – the saprobic index. This has resulted in different saprobic indices of different complexities (Sladecek, 1973). The first stage in producing a numerical presentation was to allocate saprobic index (SI) values 1- 4 to express the saprobic classifications oligosaprobic - β, α mespsaprobic - polysaprobic. At the same time, the relative abundance of species was taken into account as a weighting factor for deriving the saprobic index of the site.

The main advantage of the saprobic system is that it includes a wide range of taxa and communities and is thus applicable to all types of rives. However, specific critics on the system are the demanding identifications required - to species level - which makes it cost and time consuming (Carter and Resh, 2001; Bonada et al., 2006). The requirement of the assessment of abundances for index calculation is another constraint. In addition, it is not applicable across large geographic areas and different impact types and is often not consistent in indicating specific forms of organic pollution (Bonada et al., 2006).

1.2.2.2 Diversity approach

Diversity indices use three components of community structure to describe the response of a community to the quality of its environment: namely, richness (number of species present), evenness (uniformity in the distribution of individuals among the species) and abundance (total number of organisms present) (Metcalfe, 1989). The principle is that disturbance of the water ecosystem or communities under stress leads to a reduction in diversity.

Three diversity indices are frequently applied (on an individual basis) in stream studies: the species richness, the total diversity and the evenness index (Margalef, 1958; Shannon and Weaver, 1963; Hill, 1973). These indices, used in the past in the United States for biological assessment of water quality (Weber, 1973), have also been applied in several studies as a comparative ecological criterion for tropical stream and river assessment (Matagi, 1996; Ometo et al., 2000; Buss et al., 2002; Moyo and Phiri, 2002).

Advantages of diversity indices in biomonitoring include their easiness to use and calculate, applicability to all kinds of watercourse with no geographical limitations. Diversity indices are strictly quantitative, dimensionless and rely on statistic analysis and are best suited for comparative purposes (Cook, 1976). They do not rely heavily on the sample size (Pinder et al., 1987a) and are applied equally to measures of biomass (Mason et al., 1985). However, diversity index values are unable to indicate if the community consists of pollution-tolerant or pollution-intolerant species. Furthermore, diversity index values vary greatly since they depend on the sampling method and the nature of the study site (Metcalfe, 1989).

1.2.2.3 Biotic approach

The biotic approach is defined by Tolkamp (1985) as a combination of the diversity with the pollution indication of individual species or higher taxa or groups into a single index or score. The biotic approach incorporates desirable features of the saprobic and diversity approaches combining a quantitative measure of species diversity with qualitative information on the ecological sensitivities of individual taxa into a single numerical expression (Table 1.2).

Table 1.2. Biocoenotic responses of indicator values induced by pollution (De Pauw and Hawkes, 1993).

Response class	Species vs. community response	Response description
A	species	appearance or disappearance of individual species
B	community	reduction in number of species/taxa present i.e. a reduction in diversity
C	community	change in population of individual species
D	community	change in proportional species composition of community

Biotic indices are based on the principle that macroinvertebrate groups disappear as pollution increases, from the most sensitive groups to the more tolerant ones (Mackenthun, 1969). The most sensitive taxon present in a surveyed site, together with the number of relevant taxonomic groups, is translated into a single numerical value. This value can be directly obtained from a table, which combines the information given by the taxa richness and the most sensitive taxon present (i.e. biotic indices) (e.g. the Trent Biotic Index, Woodiwiss, 1964) or by summing individual scores allocated to each taxon (i.e. biotic scores). Although in general biotic indices do not consider abundance per taxon for their calculation (except for the fact that threshold values are sometimes considered for inclusion of taxa), some biotic scores require an abundance measure of the organisms (Chandler, 1970).

The biotic approach has a long history of development and use in Europe (Persoone and De Pauw, 1979; Armitage et al., 1983; De Pauw et al., 1992; De Pauw and Hawkes, 1993, Hering et al., 2003) and North America (Hilsenhoff, 1977, 1988; Lenat, 1993) and has also been successfully adopted in many regions such as South Africa (Chutter, 1994; Thiriron et al., 1995; Dicken and Graham, 2002), neotropical countries (Astorga et al., 1997; Jacobsen, 1998; Fenoglio et al., 2002; Roldan, 2003) and Asia (Sinha et al., 1993; De Zwart and Trivedi, 1994; Mustow, 2002).

The most important advantage of the biotic approach is that only qualitative sampling is required without the need to count abundances per taxon. Most identifications are at family or genus level, which also helps to make the biotic approach widely accepted in East Asia where the lack of taxonomic knowledge remains the biggest constraint in applying bioassessments. However, determining representative reference communities to which the investigated stations can be compared remains a constraint, as well as optimising biological assessment through regional adaptations.

1.2.2.4 Multimetric approach

The first multimetric system was developed for fish (Karr, 1981). Later, this system was also applied on macroinvertebrates (Plafkin et al., 1989; Barbour et al., 1992; Kerans and Karr, 1994; Fore et al., 1996; Karr and Chu, 1999). In multimetric systems, several metrics representing different characteristics of the macroinvertebrate community are summed up into one index value or score (Barbour and Yoder, 2000). In this way, it is assumed that several aspects of ecosystem functioning or different measures of ecological integrity, are combined into a more holistic evaluation. Also, combining several metrics is generally assumed to enhance reliability and robustness of an index, because accidental outliers of one metric can be smoothed by the other metrics. Buffagni et al. (2005) argue that multimetric systems are more suited than single metrics to assess ecological quality and to describe biological communities. Metrics can be classified into several categories, each based on different principles of ecological quality assessment (Resh and Jackson, 1993; Thorne and Williams, 1997; Verdonschot, 2000; De Pauw et al., 2006): richness or diversity, sensitivity metrics, similarity metrics, metrics based on functions, such as feeding groups; and metrics that combine two or more of these categories, such as biotic indices.

An important advantage of multimetric approaches is their flexibility (Gabriels et al., 2006), that makes them potential for use in East Asia. Multimetric indices can be easily adapted to a regional situation by taking the most appropriate metrics into account and by evaluating each metric to an appropriate target.

1.2.2.5 Ecological Quality Ratio approach

In this approach, the assessment is based on comparison of the community composition to a pre-defined target community. This target, usually called the reference, can be based on actual samplings, expert knowledge, historical data or predictive models or a combination of these. This is called the Ecological Quality Ratio (EQR) according to the Water Framework Directive (EU, 2000; Wallin et al., 2003). This ratio is expressed as a numerical value between zero and one, where zero represents a very bad ecological status and one a very good status.

$$\text{EQR} = \frac{\text{Index value of observed community}}{\text{Index value of reference community}}$$

Predictive models nowadays constitute a central part of the EQR approach in the assessment of running water. An example of an EQR is the Environmental Quality Index (EQI) based on the 'River Invertebrate Prediction and Classification System' (RIVPACS) developed in the UK (Armitage et al., 1983; Wright et al., 1993; De Pauw, 2000; Wright et al, 2000). The principle of RIVPACS is that, on the basis of the physical-chemical features of the river, it is possible to predict which macroinvertebrate taxa should be present under these conditions. The predicted reference conditions can then be compared with the observed macroinvertebrate communities. The RIVPACS EQI can be calculated with different metrics or indices, for example the BMWP, the ASPT or the number of taxa (Sweeting et al., 1992). Based on RIVPACS, other similar models have been developed in Australia (AUSRIVAS: 'Australian River Assessment Scheme', Schofield and Davies, 1996; Parsons and Norris, 1996; Smith et al., 1999; Davies, 2000; Coysh et al., 2000; Marchant and Hehir, 2002), Canada (BEAST: 'Benthic Assessment of Sediment', Reynoldson et al., 2000) and Sweden (SWEPACS, Sandin, 2001).

Although nowadays methods based on an EQR approach are commonly developed in temperate areas, their application in other areas seems presently limited because of scarcity of data on reference sites. After basic research on reference sites, the application of this approach in East Asia should be considered.

1.2.2.6 Multivariate approach

Several multivariate techniques have been applied in water quality assessment using macroinvertebrates (Norris and Georges, 1993). The basis for the multivariate approach is the similarity index (Sandin et al., 2000). Classification, ordination and discriminant analysis are some of the most widely used multivariate methods in water quality assessment based on macroinvertebrates (Norris and Georges, 1993). Some of these methods can be computed either by using different similarity measures or the similarity index is hidden in the computations. The most commonly used similarity index is the Jaccard's index (Jaccard, 1908; Washington, 1984). This index expresses the percentage of species shared between two sites. Other examples are the percentage similarity index (Whittaker, 1952), Bray-Curtis dissimilarity index (Bray and Curtis, 1957), Sorensen index (Sorensen, 1948) and Euclidean of ecological distance (Williams, 1971). All these indices give an indication how much a biological community at each sampled site is similar to the median of all reference communities and are not resulting in an assessment class as such. Within the ordination methods used in macroinvertebrate studies, TWINSPAN (Two Way Indicator Species Analysis) (Hill, 1979), PCA (Principal Component Analysis), CA (Correspondence Analysis), RDA (Redundancy Analysis) and CCA (Canonical Correspondence Analysis) are most commonly used.

Multivariate techniques are since the nineties also commonly applied for the development of multimetric systems. The selection of the metrics is based on how complementary or explanatory these are. The complementarity of score systems is necessary to guarantee that correlated metrics do not dominate the overall assessment, while the explanatory aspects are interesting to get insight in the causes of deterioration. Since the new millennium, a shift in the use from multivariate statistical (classification, ordination, regression, clustering etc. based on data distribution functions) to soft computing techniques (based on heuristic search methods such as artificial neural networks and inductive logic programming) has started. Examples of indices of the different assessment approaches based on macroinvertebrates are given in Table 1.3. The use of multivariate techniques is, however, still incipient in tropical areas (Dominguez-Granda, 2007). Multivariate methods may constitute the basis for the development of multimetric systems in tropical areas as they also did in temperate areas. Multivariate methods will allow the correct selection of metrics for multimetric systems, avoiding the use of redundant variables. Therefore, its adequate application is strongly encouraged by scientists working on tropical riverine ecosystems.

1.3 Bioassessment application in East Asian countries

Freshwater biodiversity is under threat worldwide, but the intensity of threat in the east of tropical Asia is exceptional (Fig. 1.2). Asia is the most densely populated region on Earth. Many rivers in that region are strongly polluted and large portions of their drainage basins and floodplains have been deforested or otherwise degraded.

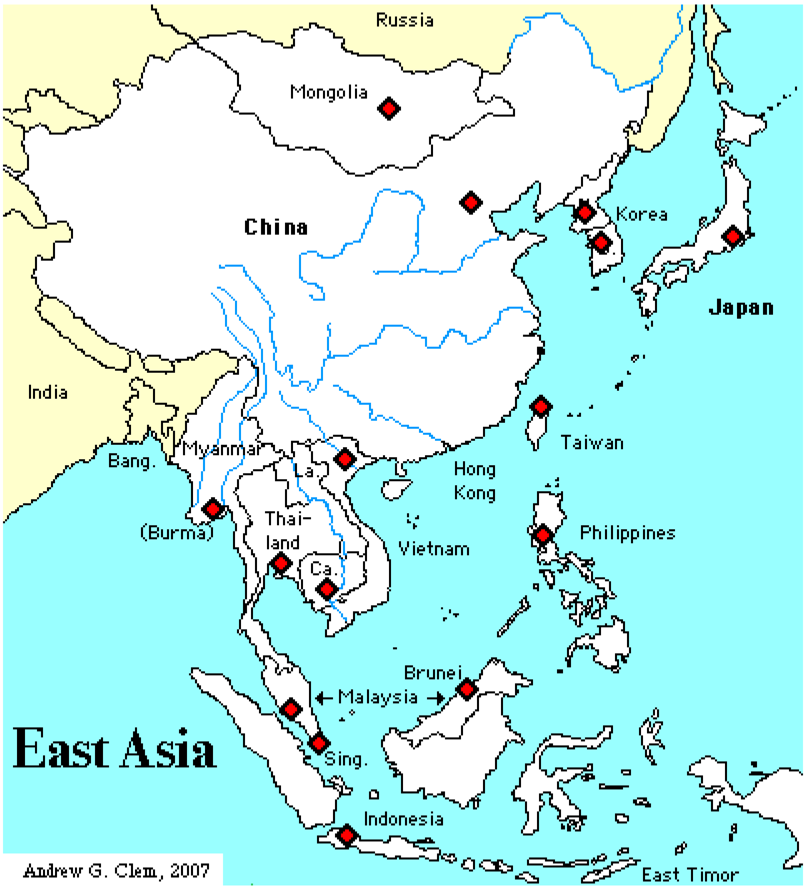


Figure 1.2. Map of East Asia with the major rivers (Clem, 2007).

Protocols for using macroinvertebrates to monitor water quality have been published and implemented in many countries, including the US (Barbour et al., 1999), the European Union (EU, 2000), New Zealand (Stark et al., 2001), Australia (AUSRIVAS, 2005) and Canada (Rosenberg et al., 2005). Until present, Asian biologists have modified these guidelines to meet their own particular needs, primarily to accommodate different habitats and taxa in their country. Direct transfers of approaches used from developed to developing countries are often appropriate, however, techniques dependent on pollution-tolerance values are sometimes region specific and not transferable (Resh, 2007). Countries in East Asia have seen varying levels of progress in implementing appropriate protocols. Below, the status of freshwater macroinvertebrate monitoring in east Asian countries is reviewed.

1.3.1 China

A national hydrobiological survey was performed in the late 1950s and benthic macroinvertebrates and environmental variables from major water bodies were investigated, including the Three Gorges area of the Yangtze River, the Heilongjiang watershed and Lake Taihu. The first Chinese study on the effects of point source pollution on benthic macroinvertebrates was conducted in 1963, in the Songhuajiang River in the north-eastern part of the country (Wang, 2002).

From the late 1970s to the 1980s, government agencies, institutes and universities introduced quantitative methods and different biotic indices to monitor water quality along the Yangtze River, the Zhujiang River, the Jiyun River and others (Wang, 2002). A modified Shannon–Wiener's Species Diversity Index, using benthic macroinvertebrates, was officially applied for biomonitoring by environmental agencies since 1982 (Hwang, 1982).

Later, Qi (1991) used macroinvertebrates as indicators for water quality assessment in the Zhujiang river basin. The results of this study showed that dominant groups in the study area were Polychaeta, Oligochaeta and *Corbicula fluminea*. In that research, BPI (Biological Pollution Index) has been used for analysing the benthic community and this index was strongly correlated with the dissolved oxygen content and BOD of the river water. FMI (Family Monitoring Index) was also utilized to systemise the water quality assessment in Jia-hua river in China (Liangfang et al., 1991).

Semi-quantitative methods that counted EPT taxa and calculated a Hilsenhoff FBI were first used for monitoring stream water quality in 1992 (Yang et al., 1992). After the publication of 'Aquatic insects of China useful for monitoring water quality' (Morse et al., 1994), there have been a number of well-developed studies on biomonitoring (Wang and Yang, 2004; Wang et al., 2005; Ma et al., 2008).

Studies on metrics for water quality bioassessment of 7 small water bodies in Zijinshan Mountain, Nanjing, China (Wang and Yang, 2004), contained biocriteria for the bioassessment of Qinhuai River by using a river biological index, determined tolerance values for benthic macroinvertebrate taxa in eastern China. The studies then evaluated stream ecosystem health and the effects of land use and land cover on stream benthic macroinvertebrates by using the established Benthic Index of Biotic Integrity (B-IBI). Additionally, Wang et al. (2005) determined tolerance values for benthic macroinvertebrates in the Lushan Nature Reserve of China. In addition, benthic macroinvertebrate community dynamics in lakes, such as Taihu Lake, are now often used to highlight the ecological effects of eutrophication (Ma et al., 2008). These studies laid the foundation for the further bioassessment of water quality and aquatic ecosystem health of various freshwater bodies in China.

However, Morse et al. (2007) also pointed out that the Chinese Government is increasingly concerned with the degradation of the environment and pays great attention to ecological restoration, the development of environmental monitoring systems and the conservation of freshwater species. However, financial support, laws and policy related to biomonitoring using benthos is lagging behind in comparison with chemical monitoring because regulatory authorities involved in environmental management rarely collect biological data. Neither multimetric methods nor multivariate approaches have been used so far by governmental environmental monitoring centres. It has also been recommended to overcome current technical obstacles for the application of benthic macroinvertebrate biomonitoring in China by (1) inventorying and describing local faunas, (2) determination of tolerance values for macroinvertebrates of different ecoregions, (3) training more research teams for using benthic macroinvertebrates to assess water quality in universities and institutes and (4) revision of protocols for biomonitoring.

1.3.2 India and adjacent countries

India and adjacent countries, such as Nepal and Bangladesh, can be considered as belonging to East Asia, according to different geographical classifications. Bioassessment approaches have been developed in these countries and therefore also provided important information for the development of methods in other countries in this area.

In a survey of the pollution status of the Khan river in India, Rao et al. (1978) found out that the ratio of Polychaeta and the total number of benthic invertebrates can be used for river water quality assessment. His results pointed out that the water is heavily polluted by nutrients or industrial pollution when the number of Polychaeta accounts for over 80% of the total macrobenthos in the community. Later on, De Zwart and Trivedi (1994) reported the modification and use of the original British BMWP index for water quality monitoring in Indian rivers. The index was accepted in the whole country and recently applied in the national water quality monitoring network comprising 1029 stations of which 592 are rivers (Akolkar et al., 2008).

The NEPBIOS (Nepalese Biotic Score) is an adaptation of the English BMWP for the assessment of highland streams and rivers in Nepal (Sharma and Moog, 1996). The method requires mostly family level identifications and tolerance values are assigned to indicator taxa. Tolerance values are allocated between 1 and 10, with higher values assigned to more sensitive taxa. The NEPBIOS index is calculated by summing the tolerance values of the families present in the sample. Analogous to the ASPT calculation, NEPBIOS can be divided by the number of scoring taxa to obtain the NEPBIOS/ASPT index. Once this last index is calculated, this value can be transformed to water quality classes on the basis of a table. Since developed, several modifications and applications of the NEPBIOS followed and more recently, a new score Hindu Kush-Himalayan biotic index (HKHbios), which is similar to NEPBIOS/ASPT index was developed for rivers in the Hindu Kush-Himalayan region, which is applicable in a wider geographical range including 5 countries: Bangladesh, Bhutan, India, Nepal and Pakistan (Hartman, 2008).

1.3.3 Indonesia

In Indonesia, the Hilsenhoff's biotic index, HBI (Hilsenhoff, 1977) was used to assess the quality of a water body (Rajendra Sarovar) in Dhanbad Town, Bihar (Sinha et al., 1993). The environmental stress on this watershed, exposed to a considerable amount of sewage and organic waste, was reflected successfully by the biotic index. The HBI suggested that the pollution status of the habitat ranged from a state of some organic enrichment to a strongly disturbed level. The authors pointed out that the HBI was a reliable biological method for water quality assessment in the area.

Later on, Trihadiningrum et al. (1996) evaluated the performance of several biotic indices commonly used in temperate rivers for the assessment of the Blawi river (East Java) including the *Trent Biotic Index* - TBI (Woodiwiss, 1964), the *Extended TBI*, EBI (Woodiwiss, 1978), the *modified EBI*, EBIG (Ghetti, 1986), the *Chandler's Score* - CS (Chandler, 1970), the *modified Chandler's score* - MCS (Cook, 1976), the *Indice Biologique Global* - IBG (AFNOR, 1985), the *BMWP* and *ASPT* (Armitage et al., 1983), the *Belgian Biotic Index* - BBI (De Pauw and Vanhooren, 1983) and some traditional diversity indices such as: taxa richness, Shannon-Wiener, Simpson and Margalef (Washington, 1984). Most biotic indices tested in the study were highly correlated with a set of chemical indices of water quality from the lowland rivers, where water was always available and less environmental pressures were observed. Therefore, the need to include biological evaluation together with physical-chemical assessment was advised. Additionally, the authors pointed out that the biotic indices should be geographically specific to guarantee a convenient sensitivity. Consequently, local taxa were included in two new indices that were developed based on the IBG and the BBI (IBGa and BBIa respectively).

An AUSRIVAS-type method was developed by Sudaryanti et al. (2001) for the assessment of the Brantas river in Indonesia and the applicability of such a method in a monitoring programme was discussed by Hart et al. (2001). From their experience, three important aspects were pointed out for the applicability of bioassessment methods in Indonesia and more generally in other tropical countries. First, the confidence that the bioassessment method applied work in the region, second, that the resources needed are available (i.e. technical, scientific and financial support) and third, that the method is accepted by the water resources management agencies. The conclusion from this study also revealed that, although methods based on an EQR approach are commonly developed in temperate areas, their application in tropical areas, especially in developing countries, seems presently limited because of the scarcity of data on reference sites.

1.3.4 Japan

The first comprehensive key to aquatic insects was compiled during the late 1950s (Tsuda, 1962). Together with the key, Kolkwitz and Marsson's (1908) Saprobien System and Beck's (1954) Biotic Index was also introduced (Rosenberg and Resh, 1993). Recently, more comprehensive identification guides for aquatic insects in Japan were developed (Kawai, 1985; Kawai and Tanida, 2005) and widely used in Japan and adjacent countries.

The first bioassessment protocol, namely the Beck-Tsuda Biotic Index (Tsuda, 1964), has been applied for over 20 years in Japan. After the introduction of the Beck system by Tsuda, a number of indices, including the Zelinka–Marvan Saprobic Index (Gose, 1978), the Shannon Diversity Index and the B-IBI (modified from the IBI of Karr (1981)) were used for measuring organic pollution and ecosystem health of Japanese rivers.

In the last few years, several attempts were made to apply bioassessment methods in freshwater rivers and streams in Japan. The relation between benthic invertebrate assemblages and ecological characteristics has been studied in different rivers and streams (Kobayashi and Kagaya, 2005; Yoshimura, 2007). The results of these studies indicated that the macroinvertebrate community structure is likely to vary according to channel characteristics as well as environmental factors. Watanabe et al. (2000) used macroinvertebrate communities to assess the long term effects of recovery of the water quality of rivers receiving mine drainage. Family richness and abundance of selected indicative macroinvertebrate taxa were used to discover the change in water quality.

Recently, also BMWP Scores based on groups of benthic macroinvertebrates found in Japan were introduced. However, there are currently no formally accepted or predominantly used systems other than the Beck system in this country (Morse et al., 2007).

In 1984, the Ministry of the Environment, Japan (MEJ) started a nationwide biomonitoring programme for organic pollution, the Nationwide Survey of Aquatic Organisms (NSAO), which is based on macroinvertebrates (Morse et al., 2007). Thirty macroinvertebrate taxa were designated to four grades of water quality, which correspond to the following pollution levels: oligosaprobic (very slight), beta-mesosaprobic (moderate), alpha-mesosaprobic (high) and polysaprobic (critical). National Censuses on River Environments (NCRE) conducted a similar programme to collect basic information on riverine habitats and organisms. Inventories were made for macroinvertebrates, riparian plants and fish in 109 rivers, streams and major reservoirs. The use of census data for biological indicators is still under development, but these data can be very useful for further study and development of a comprehensive bioassessment protocol in Japan.

1.3.5 Korea

Since the 1970s, biological water quality assessments using aquatic insect or benthic macroinvertebrate communities have been applied in South Korean streams (Yoon and Bae, 1993), but most of them are four-graded saprobic systems, which are unweighted community indices, based on a species diversity index (Pielou, 1966, Wilhm and Dorris, 1968). A 4-tired system including categories of excellent (A class), good (B class), fair (C class) and poor (D class) was proposed. Indicator species including fishes and macroinvertebrates were also designated to each class

Other community indices, i.e. weighted community indices, such as the Beck-Tsuda Biotic Index (Tsuda, 1964), the Hilsenhoff FBI (Hilsenhoff, 1977), the Trent Biotic Index (Woodiwiss, 1978), Biotic score (Chandler, 1970) and BMWP score (Hellawell, 1986) were empirically introduced for Korean streams. Based on these studies a Korean biotic index or Yoon-Kong's Total Biotic Score and a modified Zelinka-Marvan's (1961) method was developed (Yoon et al., 1992a, b, c). This Total Biotic Score method (TBS method) has been modified into a simplified biotic index (Group Pollution Index), which is widely used for South Korean streams these days (Bae et al., 2005). A new biomonitoring method, using an artificial neural network, was also introduced (Chon et al., 1996). In addition, a rapid bioassessment method, the Dominant Species Index, which uses four indicator groups of benthic macroinvertebrates correlated to water quality, was recently developed (Bae et al., 2005).

Biomonitoring programmes using benthic macroinvertebrates have been variously applied to the main watercourse and tributaries of the Han River. Based on results of these biomonitoring projects (Bae and Lee, 2001), some benthic macroinvertebrates were selected as representative indicators in Korean streams in terms of the degree of water quality and habitat degradation (Table 1.3).

Table 1.3. Representative benthic macroinvertebrate taxa in Korean streams according to saprobic value based on Shannon - Wiener's species diversity index (H') (Bae and Lee, 2001)

Saprobic categories	H'	Representative benthic macroinvertebrate taxa
Oligosaprobic	3.0-4.5	*Drunella,* Plecoptera*, Rhyacophila*
β-mesosaprobic	2.0-2.9	*Hydropsyche kozhantschikovi, Uracanthella rufa, Epeorus latifolium*
α-mesosaprobic	1.0-1.9	*Chironomus yoshimatsui* - group
Polysaprobic	0.0-0.9	Tubificidae

Biological assessments in Korea were formalised in the national 10 year master plan (2006-2015), and named as 'The basic plan for water environment management' with the adoption of narrative biological criteria in the law of water quality and aquatic ecosystem conservation in 2006. Related with the biocriteria programme, the government is trying to integrate an ecological concept into the ecosystem restoration in amendment of the related regulations.

One of the major problems facing those carrying out biomonitoring studies in South Korea is insufficient taxonomic knowledge. Molecular taxonomy has recently been introduced to overcome taxonomic problems with larvae of those aquatic insect groups that are hard to identify. As the public demand for clean water grows, more educational programs and materials regarding biomonitoring will be required in South Korea (Morse et al., 2007).

1.3.6 Malaysia

Malaysia is one of the world's 12 'megadiverse' countries. Even though one of the first-ever studies of the macroinvertebrate fauna of a tropical river was undertaken in Malaysia (Bishop 1973), freshwater macroinvertebrates of the region are still poorly known and relatively few species have been described. A guide to the aquatic macroinvertebrate fauna of the region has only recently been published (Yule and Yong, 2004). Consequently, studies of freshwater macroinvertebrates are still mostly in the form of biodiversity surveys (Che-Salmah and Abu-Hassan, 2005). As a result, until the fauna is better known, biomonitoring will remain uncommon, except for some independent short-term surveys.

A preliminary study was conducted to evaluate the potential of benthic macroinvertebrate communities (identified at family level) in predicting the water quality at the Linggi River in Negeri Sembilan (Ahmad et al., 2002). The Shannon diversity index, Margalef richness index and Pielou eveness index were used to compare the water quality status based on community abundance and diversity. The BMWP score was calculated to determine the organic load in the ecosystem. The results of the Linggi River stretch of the study area showed that the diversity index provided higher correlation with the Water Quality Index (WQI) value than with the biotic indices.

Water quality parameters and aquatic insect abundance were studied in the Telipok River in Sabah, Malaysia (Budin et al., 2007). Pearson correlations showed that EPT abundance was negatively correlated with turbidity and total suspended solids, but a positive correlation was observed between the EPT and dissolved oxygen.

More recently, Yap (2005) compared the macroinvertebrate fauna of an urban stream (Langat river) with a pristine river. Another study about the impacts of anthropogenic activities on the distribution and biodiversity of benthic macroinvertebrates and water quality was also conducted in the Langat River (Peninsular Malaysia) (Azrina et al., 2006). Several biotic indices were applied including taxa richness, diversity, evenness indices and BMWP score. This study assessed the impacts of anthropogenic land-based activities such as urban runoff on the distribution and species diversity of benthic macroinvertebrates in the downstream part of the Langat River. The studies contributed to a baseline study for the area and highly supported the development and use of bioassessment methods for Malaysian rivers.

1.3.7 Thailand

In Thailand, bioassessments have only been initiated since the 1990s. Vongprasert (1990) utilized the BBI to assess the water quality of several rivers in Thailand and Indonesia. The final results demonstrated the applicability of the BBI in tropical countries.

Thorne and Williams (1997) tested a variety of rapid assessment methods using macroinvertebrates to study ecological health of the River Ping and Kha Canal (Thailand) as well as from rivers in two other developing countries (Brazil and Ghana). These tests used the original macroinvertebrate data at family level and associated physical and chemical data from pollution gradients. The methods performed well for the assessment of the pollution status of the study sites and the description of the pollution gradient. There were twenty metrics that were tested in this study belonging to five major types as identified by Resh and Jackson (1993): richness indices, enumerations, diversity and similarity measures, biotic indices and functional measures. Seven metrics, including both the BMWP and ASPT, adequately described the pollution gradients. After eliminating methods that duplicate information, the authors recommended a combination of four methods to be used in a multimetric system of bioassessment in order to derive an 'index of biotic integrity (IBI)'. These were two of the richness indices (total number of families and number of EPT families), the BMWP, the ASPT and the Community Loss Index (Courtemanch and Davies, 1987). Although the indices selected and their derived IBI performed well, authors also acknowledged that in order to make the method more sensitive, indices would benefit from modification to suit local conditions, including the addition of local taxa and adjustment of sensitivities of some others.

Mustow (2002) evaluated the performance of the British BMWP at 23 sites on the Mae Ping river. The results proved that the index was capable of distinguishing between sites that were heavily impacted by organic pollution and relatively unpolluted sites. Nevertheless, because of differences between the taxonomic groups of the UK and Thailand, local taxa were included in a new index called BMWP-Thai, considering some of the changes suggested by De Zwart and Trivedi (1994) for rivers in India. In this study, tolerance scores for 10 additional taxa, which do not occur in the UK but were found in the study sites, were incorporated and also some modifications were made for existing taxa. These modifications did not significantly alter the site classification, but the method was easier in use. Although further testing and improvements are advised for the index, Mustow (2002) suggested the application of similar indices in other tropical regions.

In Thailand, the official method for monitoring water quality of inland surface waters is so far mainly based on physical-chemical analysis and an official bioassessment protocol has not yet been established (Morse et al., 2007). More research on the taxonomy and biology of Thai benthic macroinvertebrates is still required as well as a calibration of bioassessment procedures for Thailand's freshwater ecosystems.

Recently, multimetric approach and multivariate analysis were applied for development of a rapid bioassessment approach based on benthic macroinvertebrates (Boonsoong et al., 2009). The method was adapted from the Rapid Biomonitoring Protocols of the USEPA (Barbour et al. 1999) using the multi-habitat approach. Nine core metrics were identified for calculating biological index score, including number of total taxa, Diptera taxa, Ephemeroptera, Plecoptera, Trichoptera, and Coleoptera taxa, (%) Plecoptera, (%) Tolerant organisms, Beck's Biotic Index, (%) intolerant organisms, Shredders taxa and Clingers taxa. This study showed that a multimetric approach appeared to be a good technique for rapid bioassessment protocols of Thai streams.

1.3.8 Vietnam

Since the 1960s, there have been many studies about the impact of pollution on the composition and abundance of macroinvertebrate communities in Vietnam (Nguyen, 1995). However, most studies only focused on water bodies receiving wastewater in the main towns and cities. Dang et al. (2002) mentioned that only few zoobenthos organism were found in the West lake in Hanoi while zooplankton was dominant in the outlet of sewer. Nguyen (1985) also studied the invertebrate community of the To Lich drainage river in Hanoi. Only 14 benthic species were collected due to the impact of pollution by effluents from surrounding leather and paper plants as well as food processing enterprises. Based on physical-chemical factors and presence/absence of invertebrates, Nguyen (1995) also recommended a classification system for the assessment of the water quality in Hanoi. Rotatoria, Copepoda, Cladocera, Chironomidae and Oligochaeta were used as biological indicators.

Sponsored by the UK's Government Darwin initiative (1998-2000), the project 'Conserving Vietnam's Biodiversity through Improved Water Quality Assessment and Monitoring' was carried out based on biological surveillance of freshwaters in Vietnam (Nguyen et al., 2004). The project partners were the Field Studies Council and Institute of Freshwater Ecology in the UK, the Hanoi National University, and the National Environmental Agency (now known as Vietnam Environmental Protection Administration - VEPA) in Vietnam. The project aimed to support training and education for the use of biological indicator groups for water quality monitoring in Vietnam. The project outcomes included 35 BSc dissertations produced, over 600 children involved in school projects, 38 teachers receiving training at three sites, training courses given in five different locations in Vietnam, 500 copies of manuals published and distributed, two identification keys, a database and a species collection established, more than 100 individuals from higher education, participating national parks, the National Environmental Agency, schools and education authorities involved, books and equipment donated.

Within the framework of the project, two studies were about using macroinvertebrates as bioindicators for monitoring and assessing water quality in Vietnam (Hoang and Mai, 2001; Le et al., 2002; Nguyen et al., 2004). The project also resulted in the establishment of a procedure for monitoring and assessment of freshwater quality based on macroinvertebrates. Furthermore, a preliminary BMWP-Viet score system has been modified based on the original BMWP and the BMWP-Thai to meet the Vietnamese natural conditions (Nguyen et al., 2004). Since these studies were merely performed in two small streams in near pristine condition and sites with minor organic pollution, the impact of pollution on the community structure was not evaluated. It was also mentioned that the BMWP-Viet needs further investigation in the other parts of Vietnam to improve its reliability, confidence and applicability.

After this successful project, other studies were following this monitoring approach to test the applicability of the method for several kinds of disturbance during short-term surveys by several universities and institute research groups in many areas in Vietnam (Le, 2003; Nguyen and Ta, 2005; Truong and Ngo, 2007; Ngo, 2007). It was indicated that the $BMWP^{VIET}$ needed further improvement for better applicability in Vietnam. However, these were only independent studies and the approach has so far not been approved for use in a national monitoring programme (VEPA, 2007).

Currently, preliminary studies about biomonitoring are also performed in the framework of a management strategy designed by the Mekong River Commission (MRC) (Campbell et al., 2005; Davidson et al., 2006). The 2004 biomonitoring survey was part of a five-year MRC project, which aimed to develop a biomonitoring method designed specifically for the environmental conditions of the Mekong River and its tributaries. Accordingly, the principal objective of this survey was to test methodologies rather than to provide definitive information about the ecological health of the river and its tributaries. The survey sampled only 20 sites along the length of the Mekong system from northern Thailand, through Laos PDR and Cambodia, to southern Vietnam, four of which are located in Vietnam. Data were collected on four groups of organisms: benthic diatoms, zooplankton, littoral macroinvertebrates and benthic macroinvertebrates, which are believed to be best suited for biomonitoring purposes. Physical and chemical data were also collected to assist in the interpretation of the biological data. Cluster analysis and ordination by CCA have been performed to study spatial patterns. The Shannon–Wiener diversity index and the Berger–Parker dominance index (Stiling, 1996) were used to study the ecological health of the study sites (Davidson et al., 2006).

The first results of the survey demonstrated that biomonitoring is potentially a valuable tool to assess the ecological health of the Lower Mekong river-system. Future surveys will include additional sites and will provide more comprehensive and representative coverage of the Lower Mekong Basin. In the future, the project will also develop and test biological indices that are able to distinguish deleterious human impacts from the effects of natural variation in environmental variables.

Table 1.4. Summary of biological assessment studies in east Asian streams and rivers based on macroinvertebrates, with indication of the assessment methodologies locally developed or used.

Area of study	Assessment approach	Reference
Japan	Beck-Tsuda Biotic Index	Tsuda, 1964
	Saprobic valency	Gose, 1978
	Multimetric	Watanabe et al., 2000
	Multivariate analysis	Kobayashi and Kagaya, 2005 Yoshimura, 2007; Takao et al., 2008
Korea	TBI	IDRC, 1987
	Korean Saprobic Index	Yoon et al., 1992a; Won et al., 2006
	Korean biotic index or Yoon-Kong’s Total Biotic Score	Yoon et al., 1992a, 1992b, 1992c
	Dominant Species Index	Bae, 2005
Indonesia	BBI	Krystiano and Kusjantono, 1991
	AUSRIVAS	Sudaryanti et al, 2001; Hart et al., 2001
	TBI, EBI, EBIG, IBG, CS, MCS, ASPT, BBI, IBGa, BBIa, diversity indices	Trihadiningrum et al., 1996
	Multivariate analysis	Dudgeon, 2006
China	Shannon–Wiener’s Index	Hwang, 1982
	FMI	Liangfang, 1991
	BPI	Qi, 1991
	Hilsenhoff FBI, EPT	Yang et al., 1992
	BMWP, TBI, Goodnight–Whitley Index, Shannon–Wiener’s Index	Wang, 2002
	B-IBI	Wang and Yang, 2004; Wang et al., 2005
	Multimetric	Ma et al., 2008
	Multivariate analysis	Neumann and Dudgeon, 2002
India	RBP III, BMWP, ASPT, BBI, Shannon-Wiener index	Sivaramakrishnan et al., 1996
	HBI	Sinha et al., 1993
	BMWP, ASTP	De Zwart and Trivedi, 1994
	Multivariate analysis	Sharma and Rawat, 2009

Area of study	Assessment approach	Reference
Nepal	IBGN, FBI, BMWP, ASTP, TBI, EBI, IB, BBI, LQI, NEPBIOS	Sharma and Moog, 1996
Hindu-Kush Himalayan region	HKHbios	Hartman et al., 2008
Thailand	Multimetric, multivariate	Boonsoong, 2009
	BMWP, ASPT, BMWP-Thai, ASTP-Thai	Mustow, 2002
	BMWP, ASTP, FBI, Multimetric index	Thorne and Williams, 1997
Malaysia	Multimetric index (BMWP, ASPT, LQI, Pielou evenness index , Shannon diversity index)	Ahmad, 2002
	BMWP	Azrina et al., 2006
Vietnam	BMWP-Viet, ASPT-Viet	Nguyen et al., 2004
	Multimetric and multivariate analysis	Campbell et al., 2005; Davidson et al., 2006

1.4 Discussion

Protection of the environment can only be effective if there are efficient monitoring tools in place to inform decision-makers about the condition of the environment and its trends. Studies performed in many countries in the region (Table 1.4) constitute the first step towards the development and implementation of such monitoring strategies for aquatic ecosystems.

Typically, expressed concerns about the application of biomonitoring in East Asia include poor coordination among agencies, the lack of legislation, identification keys and trained personnel and incomplete information about the ecology of rivers and streams in this area. The developments of some methods for every country are an important achievement for river management in this area. In this sense, as exemplified by standardised methods from temperate areas (e.g. ISO 5667-3, ISO 7828, ISO 8689-1, ISO 9391, NBN T92-402), aspects like sampling device (i.e. type and characteristics), sampling technique, sorting out of organisms (i.e. alive or preserved), processing of organisms (i.e. qualitative or quantitative sorting), taxonomic identification (i.e. level of identification and taxonomic references) and index calculation should be clearly defined and standardised (De Pauw et al., 2006).

Seasonality is known to play an important role in structuring the macroinvertebrate community in temperate watercourses and which can influence the outcomes of bioassessment methods (Linke et al., 1999; Clarke et al., 2002). Although several studies have already been performed in order to increase the knowledge of seasonal patterns of invertebrates in tropical rivers (Arunachalam et al., 1991; Flecker and Feifarek, 1994; Baptista et al., 2001; Ramirez and Pringle, 2001), further research exploring the influence of such a variation on existing and forthcoming bioassessment methods is required. Similarly, spatial variation also plays a remarkable role in bioassessment. The regions should be defined in which the method can be confidently applied to similar aquatic macroinvertebrate communities. These geographical areas, called *ecoregions* (Omernik, 1995), provide a basis for the implementation of biological methods with scoring systems related to reference conditions (Gerritsen et al., 2000).

Taxonomy of running water fauna in temperate regions has been a research topic since the beginning of the past century, which resulted in a vast amount of scientific literature widely distributed compiling the description of their aquatic fauna in publications on a regional basis. This is not the case for tropical areas. Although an increasing effort on taxonomic work describing the fauna from tropical streams can be observed in recent years, identification keys and descriptions are still lacking in many regions. Since different values of a given biotic index could be obtained if different identification keys are used, further efforts regarding the development of taxonomic literature used for bioassessment methodologies are required.

Fundamental to the assessment of river health and biotic integrity is an understanding of the links between the habitat in which organisms live and factors shaping it (Norris and Thoms, 1999). Ecological modelling based on stream monitoring data is a next step in studying and predicting the characteristics of stream habitats. Successful modelling based on biological factors of streams can improve bioassessment results. The suitability of predictive models for assessment programmes is dependent upon their ease of application and practicality in providing management information at minimal cost and effort (Parsons and Norris, 1996). Models developed with the aim of describing the habitat of a certain species in a predictive way are referred to as habitat suitability models. In the habitat suitability modelling approach, one wants to either express the suitability of a habitat for a specific species or use information concerning the physical habitat of a certain species in order to predict its absence/presence or abundance. The analysis of species-environment relationships between macroinvertebrates and their environment has always been a main issue in ecological studies worldwide.

Even though expert knowledge is the basis for the development of biological assessment methods, multivariate data analysis and modelling techniques should also play an important role in the development, testing and optimization of actual and future indices. In this sense, predictive modelling can be considered as a useful tool for decision support in river management (Adriaenssens et al., 2004a). Since it is important to assure the comparability and uncertainty of the data, handling methods (e.g. biotic and abiotic data collection) employed should be harmonized and data quality control should be performed (Cao et al., 1997).

Bioassessment in Vietnam is still in a very early stage. Although the interest for bioassessment has been shown together with increasing concern about the loss of biodiversity in rivers and streams due to the recent increase of anthropogenic influences, the concept of bioassessment has not yet been incorporated in the environmental management system in Vietnam.

The 'Master plan on the national environmental and resource monitoring network towards 2020' has been approved in 2007 with a major aim of the establishment of a uniform network for environmental and resource monitoring in order to develop a comprehensive database for management purposes in Vietnam (VEPA, 2007). Basic parameters required in the monitoring program applied for surface waters include temperature, conductivity, colour, turbidity, pH, TSS, DO, BOD_5, COD, chlorophyll-a, NO_2^--N, NO_3^--N, NH_4^+-N, PO_4^{3-}-P, Kjeldahl N, total phosphorous, SiO_3^{2-}, dissolved minerals (Ca^{2+}, K^+, Mg^{2+}, Na^+, SO_4^{2-}, total Fe, Cl^-), alkalinity, total coliform, fecal coli, CN^- concentration of heavy metals (As, Cd, Cr, Pb, Hg, Zn, Cu), phenol and pesticide residues.

It is the first time the monitoring programme includes biodiversity monitoring including the ecology of standing water as well as running water, benthic and lotic biota, chlorophyll-a, fish and diversity indices. The monitoring frequency is once a year. However, specific guidelines are not available for implementation of this monitoring programme. A lack of reliable and appropriate bioassessment protocols is the main reason for excluding them from the national monitoring programme (VEPA, 2007).

Scarce taxonomic and ecological knowledge is the most important impediment for the development of such methods in Vietnam. Fortunately, during the last years, attempts have been made to fill the gap in the taxonomy of aquatic insects in Vietnam. Many systematic studies were recently carried out about important groups such as Ephemeroptera (Nguyen, 2003; Nguyen and Bae, 2004), Plecoptera (Cao et al., 2007, 2008; Cao, 2008), Trichoptera

(Hoang, 2005; Hoang and Bae, 2006). These studies aimed at reviewing macroinvertebrate groups and describing new species, presenting taxonomic keys for taxa of specific groups in Vietnam and also discussing aspects of their biodiversity and distribution. These are essential for further research on the development and application of bioassessment methods in Vietnam.

1.5 Conclusions

Streams and rivers are complex ecosystems where many interrelated processes take place. River health assessment is a way of examining watercourses using comprehensive tools such as water quality assessment, habitat description, biological monitoring and flow characterisation to create an overall picture of the ecological health of that watercourse. It is obvious that bioassessment approaches developed for temperate region can be used for tropical running waters when appropriately adapted, tested and standardized. The scarcity of taxonomic keys and experts in East Asia has lead to the development and application of assessment methods based on family level identifications. A major conclusion for the water management in East Asia and in particular Vietnam is that bioassessment of freshwater stream and river habitats is an effective tool, which should play an important role in water management. The success in applying bioassessment in freshwater management can also be enhanced by using ecological models, which allow a better allocation of the contribution of all stakeholders to restore deteriorated aquatic systems, as a result of pollution, flooding and ecosystem destruction.

Chapter 2

Study area, data collection and analysis

2.1 Introduction

In this chapter, the applied procedures for data collection are described, both for biotic and abiotic surveys in the study area, the Du river basin, an upstream part of the economically important Cau river basin in Vietnam. Firstly, a brief description of the Cau and Du river basin and available information on environmental status are provided. Next, the data gathering to be used in this study is described. The data can be classified in: (1) a biological dataset containing information on the macroinvertebrate communities sampled in the present study, and (2) a physical and chemical dataset with environmental variables measured in concurrence with the biological data. All the information gathered has been screened in the context of the fulfilment of the objectives of the present thesis. The applied data analysis methods are covered in the last part of this chapter.

2.2 Study area

2.2.1 The Cau river basin, hotspot in river management in Vietnam

Vietnam has a dense river network. There are over 2300 rivers belonging to 13 large river systems with a total watershed covering 80% of the country's territory. Located in the tropical monsoon area, average annual rainfall in Vietnam is between 1,400 and 2,400mm. Three fourths of the total land area is covered by mountains, therefore rainfall is distributed unevenly from one area to another and varies with time. The temporal variation of rainfall greatly affects the river flowing regime and is the main reason of droughts in the dry season and floods in the rainy season.

In Vietnam, irrigation places the largest burden on water resources, followed by industry and the service sector. The total irrigation demand in 2001 was 79.6 billion m^3, representing 80% of the total demand (Fig. 2.1). The government expects the irrigation demand to increase to 88.8 billion m^3 by 2010 serving a total irrigated area of 12 million hectares. The industrial needs currently comprise about 6.5% of the abstracted amount, but the demand is ever increasing with the rapid economic development. The domestic use of water by comparison is very small, accounting only for 2% of total demand. The consumption is expected to increase to 3.1 billion m^3 in 2010 in relation to the population growth (Fig. 2.1).

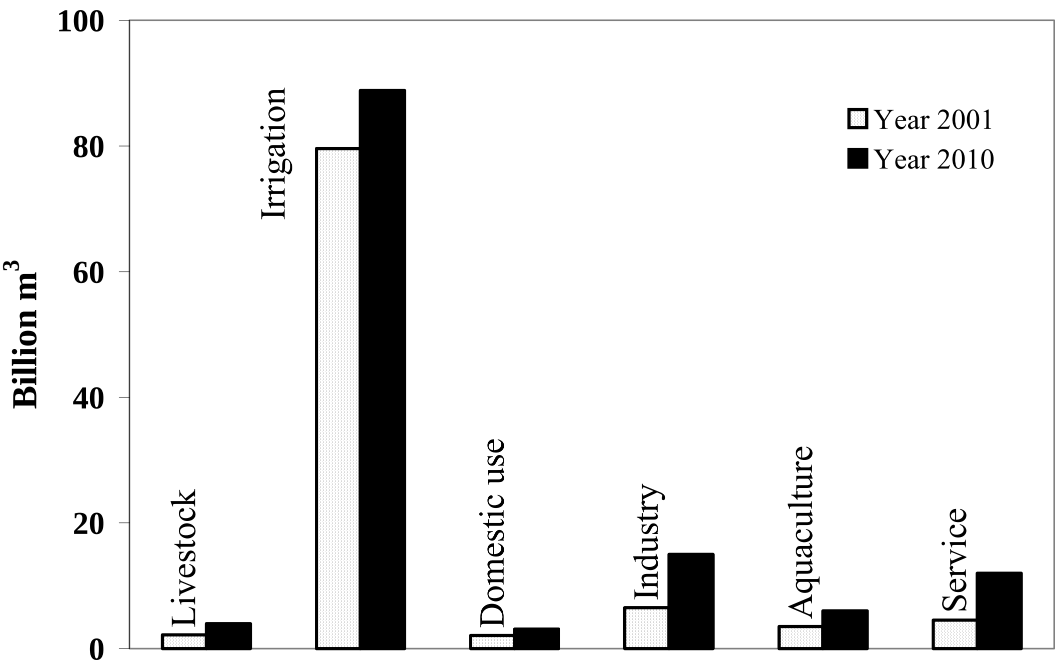


Figure 2.1. Annual water demands according to statistics 2001 and predicted for 2010 (VEM, 2003).

The inland waters are, however, threatened by domestic and industrial wastewater, dam and road construction, dredging and intensive aquaculture. The annual environment report of Vietnam 2006 specially focused on the health of river systems. Three river basins, namely the Cau, the Nhue-Day and the Dong Nai, have been recognised for their important role in the economic development and, at the same time, as the most severely polluted river basins in the country. They are located in two focal economic regions, Northern and Southern, which play the major role in the general development of the whole country and which also are the most densely populated areas. Recently, harmonisation between economic and environmental concerns towards sustainable development is a hot issue in these three river basins (VEPA, 2006).

The Cau river originates from Phia Deng (1527 m altitude) at the southeast of the Pia-Bi-Oc mountain range in Bac Kan. The Cau river is 288 km long flowing through six provinces Bac Kan, Thai Nguyen, Bac Giang, Bac Ninh, Vinh Phuc, Hai Duong and a small part of Hanoi before merging with the Thai Binh river (Fig. 2.2). The Cau river basin has a diverse stream system with six main tributaries, the Chu, Nghinh Tuong, Du, Cong, Ca Lo and Ngu Huyen Khe. The major hydrological characteristics of the Cau river are described in Table 2.2.

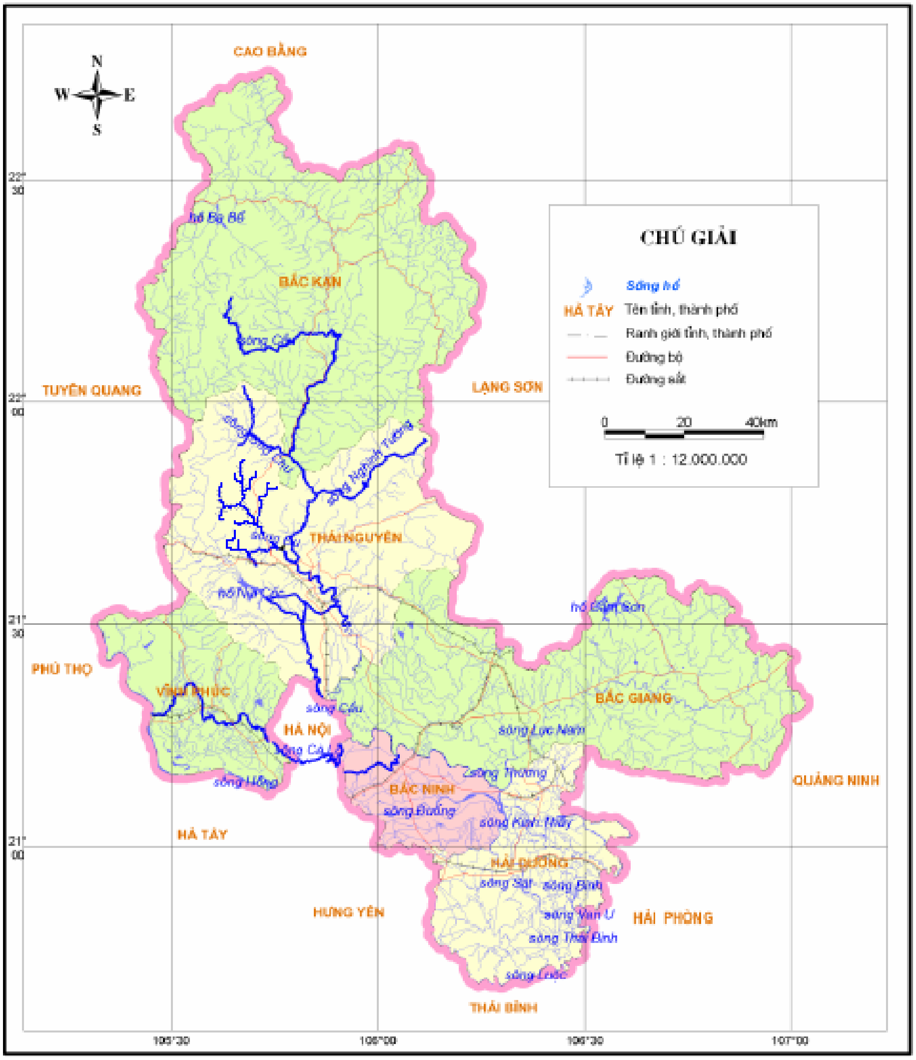


Figure 2.2. The Cau river basin comprising six provinces in the North of Vietnam (VEPA, 2006).

2.2.2 Importance of the Cau river basin for the economic development of the area

The basin covers about 47% of the total natural area of 6 provinces with a population of 6.9 million people, of which 5.9 million live in rural areas (Vietnam Statistical Year book, 2006). The average population density is 427 persons per km^2, double the average population density in the whole country. The Cau river is a main domestic water supplying source for urban and rural inhabitants. Water supply can only be accessed in the towns. In remote and rural areas, ground and river water is directly used as drinking water.

The Cau river basin is rich in natural resources. Several mineral mines are located in the basin including coal, tin, iron, ilmenite, zinc and gold. Forestry, industry and agriculture significantly contribute to the economic development of all provinces along the Cau river basin. The industrial growth rate in the area is higher than the country's average with about 800 industrial entities registered, over 1000 small industries unregistered and about 200 craft villages located in downstream provinces. Mining and ore extraction industries concentrate in the two upstream provinces Bac Kan and Thai Nguyen. The Cau river plays a crucial role by satisfying the water consumption demand for domestic and production activities. The major water consumption requirement for 2005 in the basin was estimated as 400 million m^3 irrigation in three provinces Thai Nguyen, Bac Giang and Bac Ninh; 40 million m^3 for domestic consumption only in Thai Nguyen city, 30 million m^3 only for Thai Nguyen metallurgy and Song Cong industrial zones (Thai Nguyen DONRE, 2006).

2.2.3 Anthropogenic impact and river water quality in the Cau river basin

Deforestation, mining, industrial, intensive agricultural activities and craft works have created high pressure on the natural environment in the basin. The average forest cover in the whole basin was about 45% in 2000. However, in the last decade, unplanned deforestation significantly decreased the forest area in the upstream sites causing land degradation, more severe flooding in the rainy season, and longer lasting droughts in the dry season (VEPA, 2006).

Wastewater from industries, mining areas, craft villages and urban areas is directly discharged into rivers without proper treatment. In agriculture, fertilisers and pesticides are excessively used causing serious nutrient contamination in the intensively cultivated areas in the downstream Bac Ninh and Vinh Phuc provinces.

Upstream the Cau and the Chu tributaries run across remote areas in the Bac Kan province. Domestic and agricultural water consumption is the main stress on this part of river. Organic pollution is detected locally in some places during the dry season (Fig. 2.3a). In addition, sand and gravel exploitation cause bank erosion, changes in flow velocity and increased turbidity.

Before entering Thai Nguyen city, the river shows the first signs of pollution due to industrial, mining and agricultural activities along the river and its two branches. The Nghinh Tuong river is severely impacted by gold mining activities. The downstream part of the Du river receives runoff from the Phan Me coal mine, the Cay Tram ilmenite mine and wastewater from tin extraction workshops along the river.

The Cau river section flowing through Thai Nguyen is seriously polluted (Fig 2.3c) due to wastewater from the industrial enterprises in the city. Annually, the river receives about 35 million m^3 of wastewater without or with only preliminary treatment from over 1,000 business, social services and industrial entities. Main pollution sources are the Thai

Nguyen metallurgy industrial zone, the Cao Ngan thermal power plant, the Hoang Van Thu paper mill and the Song Cong mechanical industrial zone (Thai Nguyen DONRE, 2006).

Downstream sections receive water from the Ca Lo and Ngu Huyen Khe river. The Ca Lo river flows through several industrial zones and urban areas. The organic and nutrient contaminants exceed the allowable standards for surface water. The Ngu Huyen Khe river is the most severely polluted river in the Cau river basin. Pollution of the river water is due to sewage from many craft villages working on food processing, animal husbandry, paper, plastic and scrap recycling, metal refinery, copper and lead casting, carving etc. Wastewater from these villages is directly discharged into the river without any treatment. Waterway transportation in the downstream river also poses threats of pollution by oil spills, wastewater discharges and noise from vessels.

Pollution in the Cau river is at an alarming level that requires comprehensive management and restoration activities.

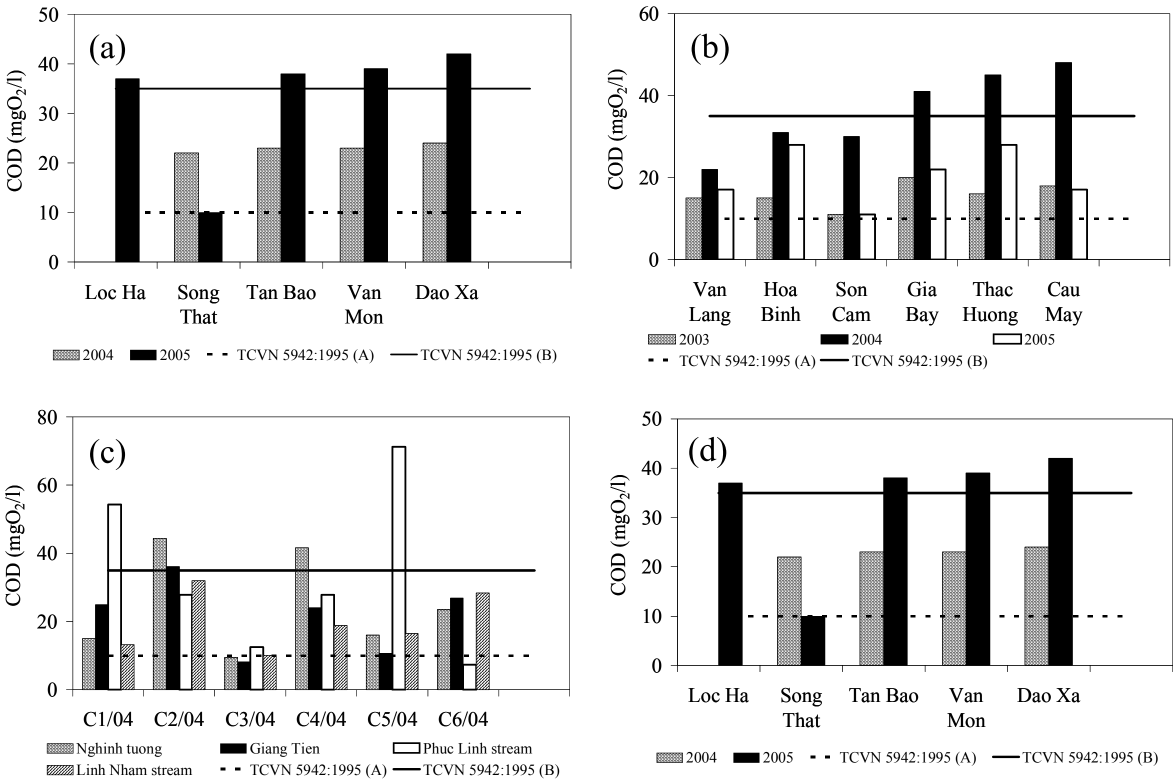


Figure 2.3. COD content in river water along the Cau river basin: (a) upstream sites of the Cau river during the dry season; (b) monitoring sites in Thai Nguyen city; (c) sites before discharging in the main Cau river; (d) downstream sites (Thai Nguyen DONRE, 2006) in comparision with Vietnamese environmental standards for domestic usage (TCVN 5942A) and other purposes (TCVN 5942B).

2.2.4 Ongoing monitoring programme in the Cau river basin

Assessment of the water quality in Vietnam is based on a set of standards. These quality standards vary according to the purposes of water source utilisation. The Vietnamese standard TCVN 5942-1995 and now updated to TCVN 5942-2005 are used for assessment of river water quality. TCVN 5942A and TCVN 5942B are applicable for surface water sources, which can be used for domestic (A) (after being properly treated in accordance with regulations) and others (B).

Recently, the river basin water monitoring activities are implemented by different units, of which the most important one is the National Environmental Monitoring Authority being managed by the Vietnam Environment Administration (VEA). The monitoring programme has been designed with monitoring sites, parameters and frequencies. The monitoring activities in the river basin however have not been fully implemented due to financial limitations and, most important, the lack of mutual co-operation between stakeholders.

Table 2.1. Required parameters in the Cau river basin according to National monitoring programme, and their Vietnamese environmental standards for domestic usage (TCVN 5942A) and other purposes (TCVN 5942B). Pesticide residues and heavy metal concentrations are monitored by special requirements (Thai Nguyen DONRE, 2006).

Criteria	TCVN 5942A	TCVN 5942B	Criteria	TCVN 5942A	TCVN 5942B
pH	6.0-8.5	5.5-9.0	NH_4^+-N , mg/l	0,05	1
SS, mg/l	20	80	NO_3^--N , mg/l	10	15
Turbidity	n/a	n/a	NO_2^--N , mg/l	0,01	0,05
Conductivity	n/a	n/a	PO_4^{3-}-P , mg/l	n/a	n/a
DO, mg/l	> 6	> 2	Oil and grease, mg/l	0	3
BOD_5, mg/l	< 4	< 25	Total Fe, mg/l	1	2
COD , mg/l	>10	>35	Coliform, MPN/100ml	5000	10000

n/a – not applicable.

Apart from the National monitoring program, other activities are performed in the river basin for different purposes by the Ministry of Fisheries, the Ministry of Agriculture and Rural Development, the Ministry of Health and also provincial environmental management authorities. The local activities of river water monitoring have been conducted at a different number of monitoring sites, for different parameters and with different frequencies. Limitations in the monitoring activities included:

- unsystematic monitoring program in terms of monitoring sites, frequency and parameters;
- low reliability of monitoring data especially for special parameters such as heavy metals or pesticide residues due to lack of properly trained man power;
- poor quality analysis and control causing impact on the quality and standardisation of the data;
- biological characteristics of the river water quality are neglected, which cause unilateral arguments in the assessment of river health.

2.2.5 Study area: the Du river basin

The Du river is one of the main tributaries located in the upstream part of the Cau river. The Du river is a medium-sized river with a length of 44 km and a catchment area of 360 km^2 and it is situated between 21°38'-21°49'N and 105°39'-105°47'E. The Du river basin is located in an area with an altitude ranging from 100 to 1050 m. The river originates in the highland of northern Vietnam and discharges into the Cau river basin in Son Cam. The Du river itself has four main tributaries: the river Na Lau and Du, and the streams Khe Coc and Cat (Fig. 2.5).

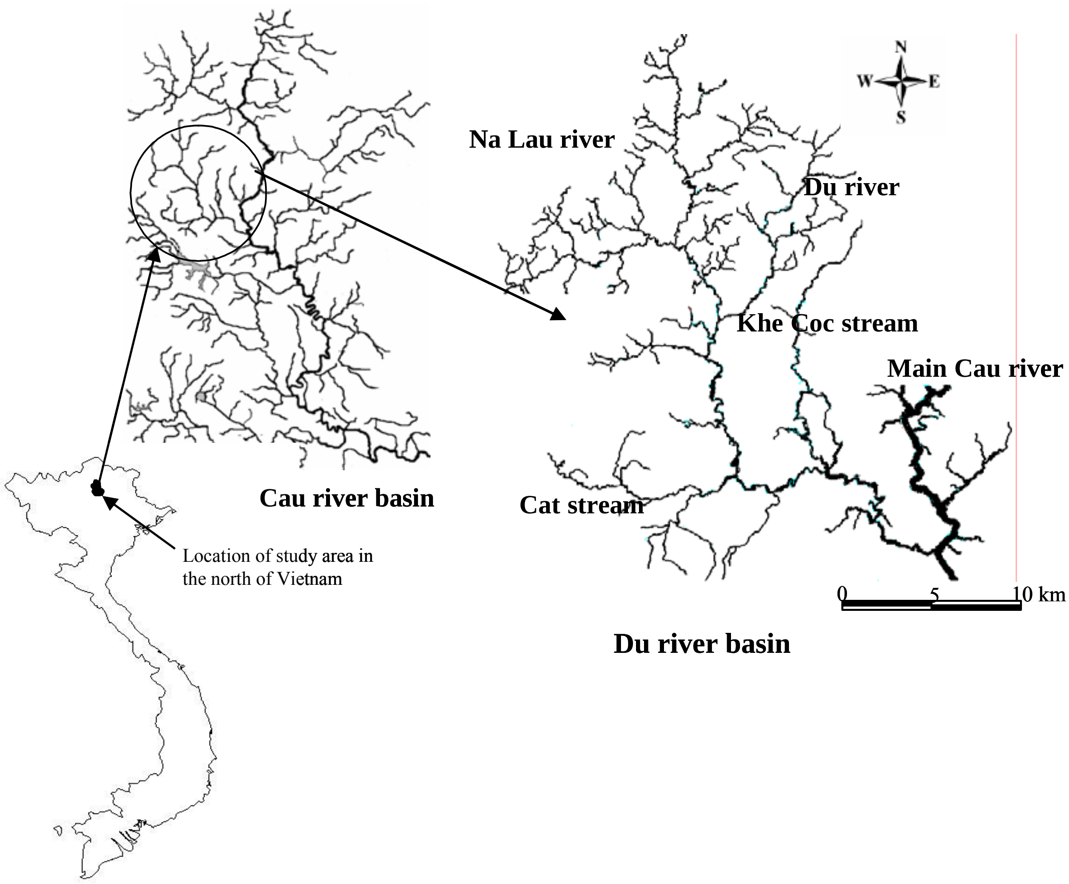


Figure 2.5. The Du river and its location in the Cau river basin.

The Du river basin is located in the subtropical climate zone. The temperatures vary between 10°C and 39°C with the highest and lowest values in August and January, respectively. The annual average temperature is about 22.7°C. The hydrographical regime of the Du river is irregular and strongly seasonally dependent. It can be divided into rainy and dry seasons based on the monthly rainfall. The rainy season usually starts in June and lasts till the end of September. The highest monthly rainfall recorded in the last 50 years is 4344mm. The flow charge in the flooding season occupies 70-80% of the total annual flow. The dry season lasts for seven or eight months, with a monthly rainfall recorded below 200mm. January, February and March are the driest months, which account for only about 5-8 % of the annual flow charge (VEPA, 2006). The average monthly flow discharges within a year can differ by a factor 10, the difference between the highest and lowest water level can be up to 5-6 m. In addition, northern Vietnam, located in the climate transition zone, is characterised by instability in temperature and precipitation as well as in the starting and ending time of the seasons.

Table 2.2. Characteristics of the Cau river and its branch, the Du river.

Characteristics	Cau river	Du river
Total catchment area (km^2)	6030	360
Average elevation (m)	150	129
Average slope degree (in percentage scale, %)	16.1	13.3
Stream network density (km/km^2)	0.7 – 1.2	0.94
The annual water flow (km^3)	4.5	0.264
Average flow out (l/s per km^2)	20-40	19.5-23.2
Average water discharge (m^3/s)	153	8.73
Maximum level of changing water (m)	5-6	5.41
Average maximum frequency of flood (cm/h)	-	58

The lack of information about the Du river was the main constraint in the present study. Social and environmental surveys were therefore conducted in 2006 for the 17 communes belonging to the Du river basin. Surveys investigated information on social (population, land use, infrastructure conditions), environmental (water supply and drainage, solid waste disposal), and economic aspects (agricultural, including irrigation, fertiliser and pesticide use, industrial and service activities), mostly focusing on water consumption and drainage. In addition, separate surveys on production and environmental status of mining companies were also conducted.

The Du river basin encompasses 14 communes with a total population of 112,400 and an average population density of 324 inhabitants per km^2. The most populated places are the Du and Giang Tien town with population densities of 2250 and 1050 inhabitants per km^2, respectively. Over 90% of the inhabitants involved in agricultural activities.

The Du river is a source for irrigation as well as domestic activities. Agriculture occupies about 35% of the total land use. Rice and tea are the main cultivated crops with intensive use of fertilisers and pesticides. Rainwater leaching through cultivated areas brings fertiliser and pesticide residues into the river. Commonly used organochlorine pesticide residues were detected in samples taken at different places in all tributaries of the Du river (Table 2.3) and exceeded the quality standards in many places.

Table 2.3. Organochlorine pesticides residues in the water detected in the Du river basin (2006).

Site Code	**Stream name**	**a-HCH (ng/l)**	**b-HCH (ng/l)**	**Lindane (ng/l)**	**Aldrin (ng/l)**	**Dieldrin (ng/l)**	**DDE (ng/l)**	**DDD (ng/l)**	**DDT (ng/l)**
2	Na Lau	8.33				16.89		12.97	5.24
3	Na Lau	9.26						4.66	6.27
4	Main Du			4.19		3.68			
5	Du		2.66	7.21				1.61	
6	Du	1.07	4.9						
7	Du	0.18							
8	Du		1.96						
9	Cat	3.26			1.68		12.83	5.06	
11	Cat	17.89						5.25	
12	Main Du	22.26	18.23	3.53		22.46		2.35	
13	Main Du					6.73		5.82	
14	Khe Coc	12.83		9.74				22.15	9.44
15	Khe Coc	10.83		8.73				15.44	7.05
TCVN 5942 Standard		0.15 mg/l for total pesticide residues except for DDT							0.15 mg/l
VLAREM II Standard		for organochlorine pesticides M t. < 20 ng/l; in. ≤ 10 ng/l							

VLAREM II : Basic environmental quality standards for surface waters in Flanders, Belgium according to Flemish regulation concerning environmental licenses
M- median, t - total, in. - individual

a-HCH, b-HCH, Lindane : different isomers of 1,2,3,4,5,6-hexachlorocyclohexane ($C_6H_6Cl_6$)
Aldrin : 1,2,3,4,10,10-hexachloro-1,4,4a,5,8,8a-hexahydro-1,4:5,8-dimethanonaphthalene ($C_{12}H_8Cl_6$)
Dieldrin : hexachlorocyclopentadiene ($C_{12}H_8Cl_6O$)
DDE : dichloro-diphenyl-dichloro-ethylene
DDD : dichloro-diphenyl-dichloro-ethane ($C_{14}H_9Cl_5$)
DDT : dichlor-diphenyl-trichloro-ethane ($C_{14}H_9Cl_5$)

The Du river supplies water for mining as well as metal extraction activities in the Na Lau river and the Cat stream. Water consumption and wastewater discharge from mines located in the Du river basin are summarised in Table 2.4. Wastewater from mining activities causes serious impacts on water quality and river health (Thai Nguyen DONRE, 2006).

Table 2.4. Water consumption in mines along the Du river basin (Thai Nguyen DONRE, 2006).

Mines	Annual production	Used water source	Amount of water discharged, m^3
Phan Me open cast coal	100.000 tons	Main Du river	940,000
Cay Tram ilmenite mine	> 100,000 tons	Na Lau river	590,000
Phuc Linh tin mine	200 tons tin product	Cat stream	630,000

The Cay Tram titanium mine is located in the Na Lau river. Titanium, in nature, is always in the form of compounds with other elements such as Fe, Mg, Zn and Cu. The titanium ore in the Cay Tram mine belongs to the most common titanium minerals called *ilmenite* type in the form of placer deposits and contains major $TiFeO_3$. The primary titanium extraction is performed on the premises. Ilmenite ores are sent to multistage crushing and grinding and complex flowsheets involving both gravity and magnetic extraction techniques to obtain enriched titanium products. Water is required during crushing and grinding as well as in gravity extraction. Wastewater from extraction processes contains other minerals and residuals, in which the concentration of Fe is extremely high. Wastewater is collected in a stabilisation pond, part of which is reused. Until recently, except for sedimentation, no treatment measures are being applied. Wastewater is stored in the pond and discharged into the river during rains.

The Phuc Linh tin mine is located in the Ha Thuong commune, which belongs to the Cat stream watershed. The tin extraction used to be processed in the factory, but during the last few years, due to a decrease in both reserve and quality of ore in the mine, extraction now is mainly performed at small scale in private workshops. According to the survey, there are 39 tin extraction workshops along the Cat stream with a total production of 800-1000 tons of raw ore per month. Low grade concentrates need to be employed to chemically upgrade the tin values before smelting. Private workshops along the Cat stream produce low grade concentrates, before they are further enriched in the factory.

Cassiterite is the main tin bearing mineral and is the mineral present in the Phuc Linh mine in a geological occurrence type of placer deposits. The placer deposit contains *cassiterite* in a liberated state and is amenable for beneficiation by gravity methods. *Cassiterite* ore is crushed and grinded before being extracted by gravity method. No extra chemicala are used in the extraction process. However, together with *cassiterite,* other minerals, which are often found as secondary ore, are silicates (quartz and muscovite), sulphides, sulphosalts and oxides. Arsenopyrite (FeAsS) is the most abundant sulphide mineral, followed by pyrite (FeS_2), chalcopyrite ($CuFeS_2$) and stannite $Cu_{1.9}(Fe_{1.0}Zn_{0.1})SnS_4$. Secondary sulphate salts, which are in close contact with Sn-bearing quartz veins as well as in some mineralized veins; these generally consist of mixtures of Ca, Al, Fe and Mg hydrated sulphates such as pickeringite, copiapite, halotrichite, alunogen and gypsum (Gomes and Fava, 2005).

The hydrochemical features of the wastewater from the extraction process appear to be dominated by the oxidation of crushed Fe-bearing sulphide minerals (pyrite, arsenopyrite, chalcopyrite and stannite) producing H^+, SO^{2-}_4 and metals (Me) in solution (aq), as in the following reaction proposed by Cidu et al. (1997):

$$(Fe, Me)S_2 + 3.5O_2 + H_2O \rightarrow Fe_{aq}^{2+}, Me_{aq}^{2+}\ 2SO_4^{2-} + 2H^+$$

$$2Fe_{aq}^{2+} + 0.5O_2 + 2H^+ \rightarrow 2Fe_{aq}^{3+} + H_2O$$

About 840 m^3 wastewater per day from private tin ore extraction workshops is directly discharged into the Cat stream without any treatment. It has an extremely low pH and contains high concentrations of heavy metals (Table 2.5). Wastewater from the ore extraction workshops is the main reason causing acidification and enrichment of heavy metals in the Cat stream.

Table 2.5. Characteristics of wastewater from Sn-ore extraction workshops. Sample taken on 20^{th} Oct 2007 at the discharge point in the Cat stream.

Criteria	**Unit**	**Value Analysed by ICP-MS**	**TCVN 5945-2005 (B)**
pH		3.0	5.5-9
Pb	mg/l	0.05	0.5
Al	mg/l	8.9	-
As	mg/l	9.4	0.1
Cu	mg/l	9.2	2
Zn	mg/l	0.7	3
Mn	mg/l	8.3	1
Fe	mg/l	25.5	5
Sn	mg/l	0.03	1
Hg	mg/l	0.005	0.01
Ni	mg/l	0.04	0.5

Note: TCVN 5945-2005: Industrial waste water - Discharge standards B (wastewater is allowed to be discharged to watersheds used for transportation, irrigation, aquaculture and cultivation).

In addition, all domestic wastewater in the watershed is discharged directly into the streams and rivers. Also a huge amount of rainwater running through the coal open cast is discharged into the main Du river. Structural and morphological disturbances are also present in the basin. Construction of artificial embankments to create suitable water levels for irrigation, removal of river bed sediments for the extraction of sand and gravel was observed downstream of the basin (Fig. 2.6). Major sources of stress due to human activities in the Du river basin are summarised in Table 2.6.

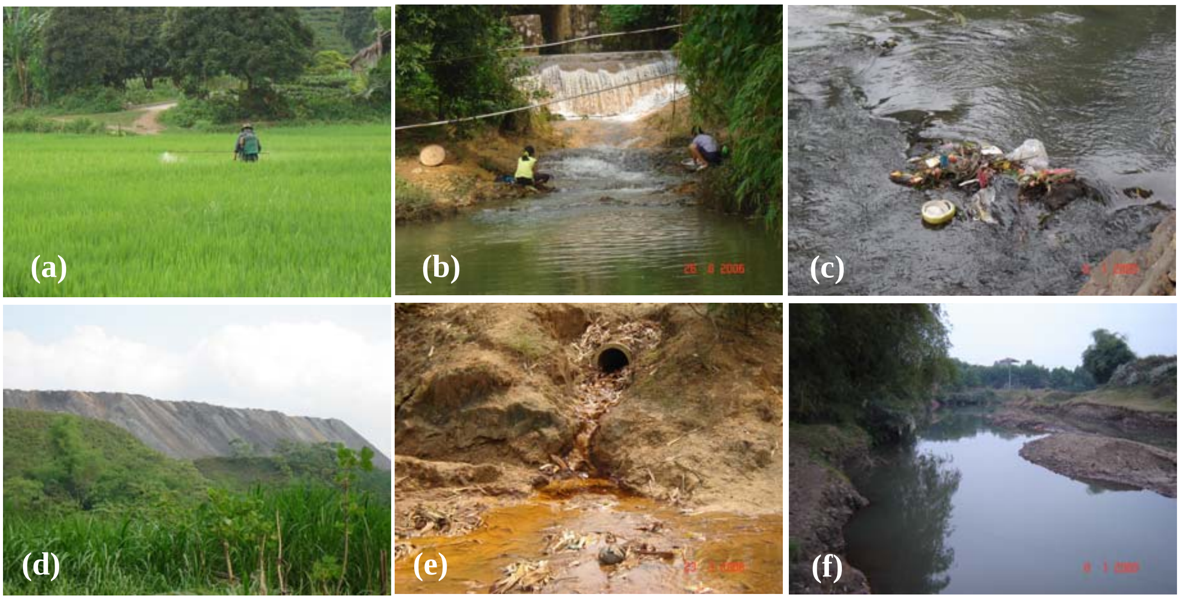


Figure 2.6. Major disturbances observed in the Du river basin: (a) agriculture activities, (b) domestic water use and discharge, (c) solid waste disposal, (d) run-off from coal mining, (e) discharge from metal extraction workshops, (f) extraction of bed sediment.

Table 2.6. Description of major anthropogenic disturbances in the Du river basin.

Physico-chemical disturbances	Structural and morphological disturbances
Point sources	• Extraction of bed sediment
• Effluents from exploitation mines	• River embankment
• Discharge of canals draining the crops	
• Wastewater discharges from extraction workshops	
Diffuse sources	
• Agriculture	
• Scattered housings	
• Waste disposal	

Fig 2.7 presents the average values of BOD_5 and COD measured in 2005 according to National monitoring program. These values of sample taken at the Giang Tien station (belonging to the Du river) exceeded allowable standard. The results also showed that the quality of the river was degraded after receiving water from the Du river. Although a point source, a pulp enterprise, was observed before Gia Bay station, the water quality of the Du river also contributes to the degradation of river water in the main Cau river.

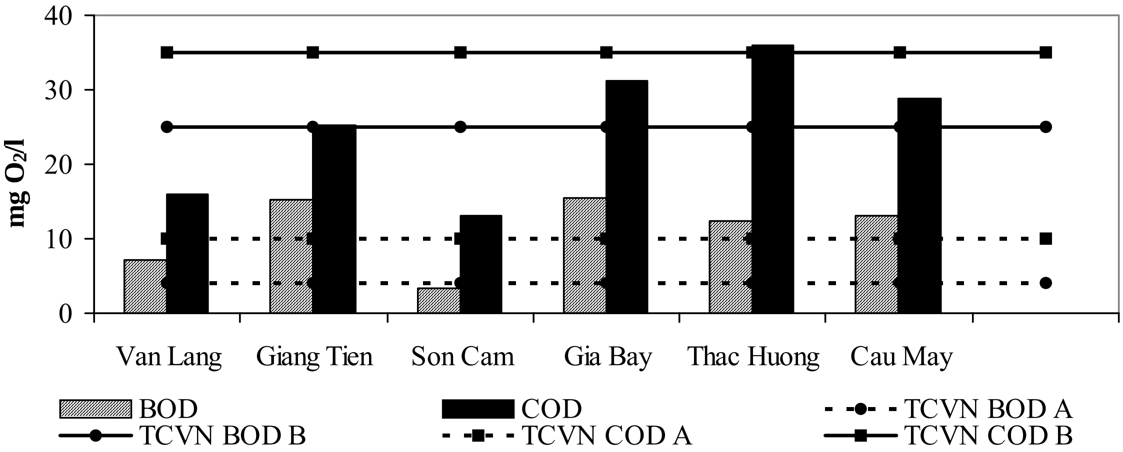


Figure 2.7. BOD and COD concentration of the river water at the National monitoring stations along the Cau river. The Du river joins the main Cau between the Son Cam and the Gia Bay station (Data from the National monitoring program, Thai nguyen DONRE, 2006).

2.3 Data collection in the Du river basin

2.3.1 Monitoring strategy and sampling sites

Several aspects were taken into account during the development of a monitoring strategy including (1) applying appropriate sample sizes, (2) applying random sampling and (3) including environmental information in the design of the sampling strategy (Goethals and De Pauw, 2001; Hirzel and Guisan, 2002).

Additionally, the sampling strategy needed to be based on those gradients that were considered to control the distribution of macroinvertebrates. These gradients were considered before sampling, because otherwise, lack of information could limit assessment accuracy. Random sampling could lead to truncated response curves for some species if the extremities of the main environmental gradients are under-sampled. Stratifying along these gradients and sampling the extremities can assure an efficient sampling of these outer limits (Hirzel and Guisan, 2002). Therefore, it is important not only to sample sites which are degraded to identify point and diffuse pollution sources based on physical-chemical characteristics, but also the pristine sites in the upper reaches.

The Du river covers a high diversity in disturbances consisting of pristine sites upstream and degraded sites caused by different types of impacts. In the Du river basin, 15 sampling sites were selected for monitoring (Fig. 2.8). The following selection criteria were applied:

- including different levels of water quality and different impacts;
- covering different types of land uses;
- including streams from different orders;
- distributing sites over the entire basin and
- being relatively easily accessible.

Seven monitoring campaigns were performed between 2006 and 2008 representing the four seasons of the year: spring, summer, autumn and winter. Summer and autumn often belong to the wet season, while winter and spring are considered as the dry season. Summer and spring are transition periods, which are important in the study of the temporal characteristics of the river.

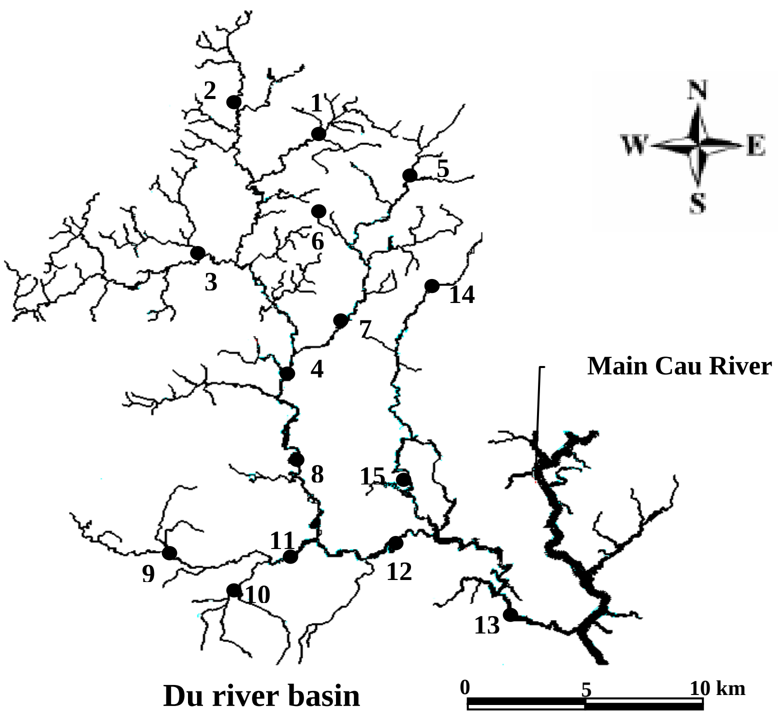


Figure 2.8. Sampling stations selected in the Du river basin for monitoring (●).

Physical, chemical and biological analyses were performed during the monitoring campaigns. Observations regarding structural characteristics were also made. The sampling site number 15 has been added after the first campaign during summer 2006. The database consists thus of 104 instances.

2.3.2 Monitoring of abiotic river characteristics

At each site, 26 environmental variables were recorded according to the following protocols.

Latitude (S), longitude (E), altitude (m): geographical information about site location obtained by using GPS (Global Positioning System) and confirmed readings on 1:50,000 topographic map.

Substrate description visual estimates of the dominance of river substrata (to a depth of 10 cm) and classified into the following substrate particle sizes (Minshall, 1985).

- Boulders (Bo) : >256mm
- Cobbles (Co) : 64 - 256 mm
- Gravel (Gr) : 2 - 64 mm
- Sand (Sa) : 0.062 - 2 mm
- Silt/Clay (Si) : < 62 μm

Dominant substrates were assigned according to the above classification. In many sites, however more than 1 type of substrate dominated the river bed. In these cases, only 2 of the most dominant substrates were recorded.

The Pool/Riffle pattern (P/R) class describes the deep/shallow variation of a site, classified in categories from 1 (well developed) to 6 (absent) (Goethals, 2005).

Water width was measured with a measuring tape according to the EPA guideline for Volunteer Stream Monitoring Partnership (USEPA, 1997). Three cross sections were selected at distances of 20m. Dimensions of 3 cross sections were measured as described in Fig. 2.9. The average water width is the mean of the width (d) measured at 3 cross sections.

Water level: the yardstick was used to measure channel depth at regular intervals across the channel width. The average depth was calculated by adding together all the values and dividing them by the number of depth measurements.

Current velocity: surface velocity was measured by releasing the float into the water. The process was repeated at least three times at each cross section to calculate the average value of the surface velocity. The average velocity was calculated by multiplying average surface velocity by 0.8 for rocky bottom streams and 0.9 for muddy bottom streams (AQEM, 2002).

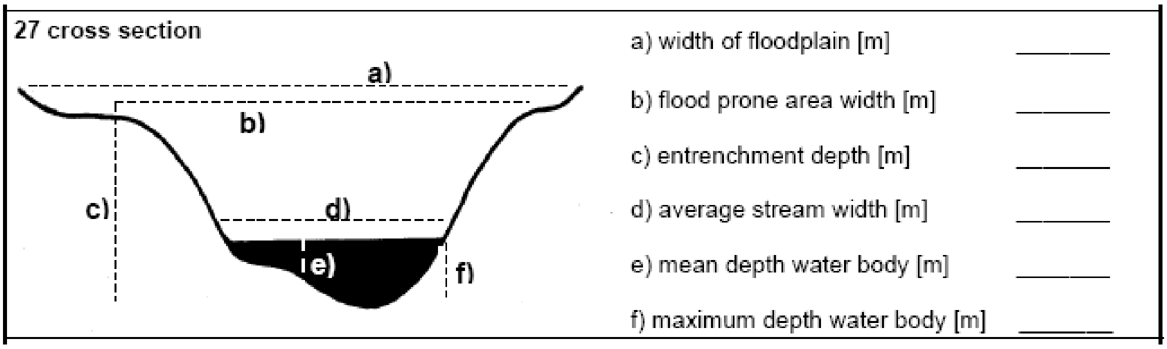


Figure 2.9. Guideline and protocol for stream morphological measurement (AQEM, 2002).

Water temperature (°C), dissolved oxygen (mg/l, %), pH, conductivity (μS/cm), turbidity (NTU) were measured at the site before sampling the macroinvertebrates, which disturbs the streambed.

Season is categorical variable containing 4 values: Sp, S, A, W, meaning the sample was taken in spring, summer, autumn and winter, respectively.

Nitrite (mg NO_2^--N/l), nitrate (mg NO_3^--N/l), phosphate (mg PO_4^{3-}-P/l), ammonium (mg NH_4^+-N/l), BOD_5 (mg O_2/l), COD (mg O_2/l), total phosphorus (mg P/l) and Kjeldahl Nitrogen (mg N/l) and the concentration of seven metals (Fe, Mn, Mg, Cu, Zn, Pb, Al, As) were measured in the laboratory according to the techniques specified in Table 2.7.

In addition, sediment was also taken manually and later dried up in a vacuum drier. Total Nitrogen (mgN/g DM) and Total phosphorus (mgP/g DM) were determined.

Table 2.7. Environmental variables and techniques of measuring.

Environmental variables	Unity and techniques
Habitat conditions	
Substrate	Visually
Altitude, longitude, latitude	GPS, Topographical map 1/50.000
Water velocity	[m/s] EPA guideline
River width	[m] EPA guideline
Water level	[m] EPA guideline
P/R class	Pool/riffle patterns, Categories from 1 (well developed) to 6 (absent) (Goethals, 2005)
Water temperature	[°C], thermometer
Chemical analysis, water column	
pH	pH meter OAKTON 35632
Dissolved oxygen (DO)	[%, mg/l] WTW Oxi 330
Conductivity	[μS/cm] WTW 249 electrode
COD	[mgO_2/l] TCVN 6491:1999, ISO 6060:1989
BOD_5, 20°C	[mgO_2/l] Closed bottle test, TCVN 6828:2001,

Environmental variables	Unity and techniques
	ISO 10707:1994
Kjeldahl Nitrogen	[mgN/l] TCVN 5987:1995, ISO 5663:1984
Total P, PO_4^{3-}-P, NH_4^+-N , NO_3^-N, NO_2^-N,	[mgP/l] or [mgN/l] Liquid-chromatography of ions TCVN 6494-2:2000 ISO 10304-2:1995
Metals (Fe, Mn, Cu, Zn, Pb, Al, As)	[mg/l] ICP-MS, TCVN6193:1996 ISO8288:1986
Chemical analysis, sediment	
Total Nitrogen	[mgN/g DM] TCVN 5987:1995, ISO 5663:1984
Total Phosphorous	[mgP/g DM] TCVN 6494-2:2000 ISO 10304-2:1995

2.3.3 Macroinvertebrate community monitoring and analysis

Macroinvertebrates were collected by means of standard kick-sampling with a handnet consisting of a metal frame holding a conical net (mesh-size 350 μm) (Fig. 2.10). The duration was 10 minutes active sampling (i.e. time used for substrate disturbance to collect organisms). Riffle habitats were sampled by holding the net downstream while the operator disturbed the substratum by kicking the feet directly in front of the net opening. Stream edge habitats were sampled by vigorously sweeping the net along the stream margins disturbing bottom and bank substratum. Additionally, hand sampling on the vegetation, stones and other substrates was adopted to collect the attached species. The sampling aimed at determining the most representative diversity of the macroinvertebrates at the site examined (De Pauw and Vanhooren, 1983).

The large objects were removed from the samples and subsequently the remaining material was sorted on sieves with mesh sizes of 5mm, 1 mm and 350 μm, respectively. The fractions were put in white trays of 60 x 40 cm. The macroinvertebrate specimens were sorted temporarily into main groups such as worms, molluscs, crustaceans and insects. Very large sized macroinvertebrates were also separated in order to avoid damaging the small sized ones. Organisms were fixed on site and preserved in small plastic vessels containing 70% alcohol for further identification in the laboratory.

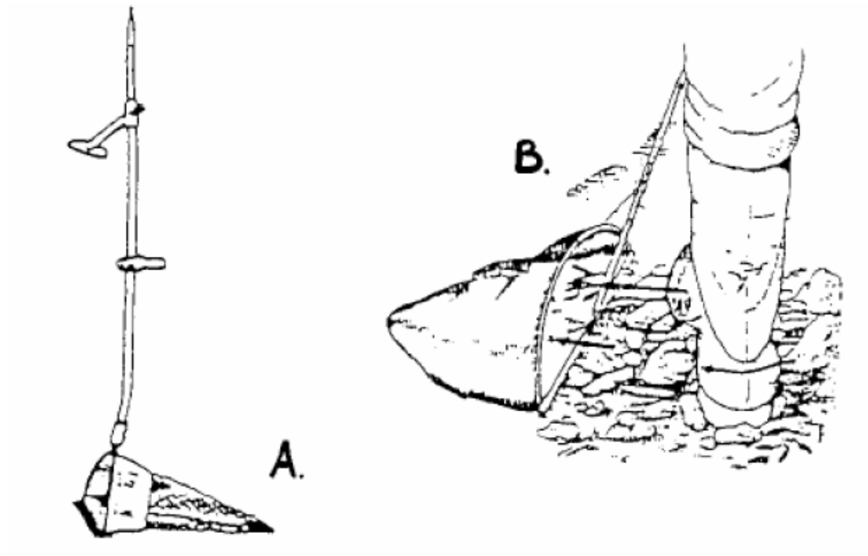


Figure 2.10. Kicksampling technique used to collect macroinvertebrates. (A) handnet with handles; (B) kick method (refer after De Pauw and Vanhooren, 1983)

In the laboratory, the macroinvertebrates were identified up to family level, except for Oligochaeta and Hydracarina at class and order level, respectively, by means of specific keys. Due to the limited taxonomic knowledge of the stream fauna in Vietnam, identification has been done based on the available literature containing identification keys and descriptions of the riverine fauna of the region (e.g. McCafferty and Provonsha, 1981; Dudgeon, 1999; Bouchard, 2004; Nguyen et al., 2004; Nguyen, 2003; Hoang, 2005).

2.4 Data analysis techniques

2.4.1 Correlations

Correlation analysis is a powerful statistical tool for assessing the relationship between two variables (Gotelli and Ellison, 2004). This statistical technique is frequently applied in bioassessment studies to determine significant relationships between or within biological and environmental measures. In this way, for instance a given metric (e.g. biotic index) is considered to have a good performance for biological assessment if it possesses a relatively high correlation with the environmental variable of interest. Correlations between two different environmental variables can also be tested to avoid ‘redundant information’ when performing a multivariate analysis based on these environmental data.

In studies exploring aquatic ecosystems, a coefficient of correlation widely applied for the analysis of relationships between biological and environmental data is the (parametric) Pearson product moment correlation (r) (Legendre and Legendre, 1998). The Pearson correlation coefficient is a way to quantify linear relationships between two random variables. Additionally to the *r* coefficient, a measure of the *significance* of the relationship (the *p*-value) should be simultaneously computed.

The most frequently applied rank based method for correlation analysis is the calculation of the Spearman rank correlation coefficient (ρ). This method assumes that the variables under consideration were measured on at least an ordinal scale, what means that the individual observations (cases) can be ranked into ordered series. This ability of working with ordinal scales makes it a suitable method for the analysis of environmental and ecological scores (e.g. biotic indices, habitat quality indices). The major advantages of using non-parametric correlation methods over parametric ones are that the ρ value is more flexible than *r*: since it is based on ranks, it is insensitive to outliers, requires no assumptions about normality and can be used to assess the strength of monotonic (linear and nonlinear) associations (Sutherland, 2001).

2.4.2 Ordination

In ecological studies, the identification of samples possessing similar features (e.g. species composition, physical-chemical characteristics) is frequently required. Here, ordination methods are applied for this purpose (Clarke and Warwick, 2001). An ordination is a map of the samples, usually in two or three dimensions, in which the placement of samples, rather than representing their simple geographical location, reflects the similarity of their biological communities (Clarke and Warwick, 2001). In this spatial representation, nearby points possess similar communities, while samples which are far apart have few species in common (when presence-absence data are used). Several ordination methods have been applied in bioassessment studies, all of which these techniques aiming to approximate high dimensional information in low dimensional plots. Frequently applied methods in ecological studies are Principal Component Analysis (PCA), Principal Coordinates Analysis (PCoA), Canonical Correspondence Analysis (CCA), Detrended Correspondence Analysis (DCA) and more recently Non-metric Multidimensional Scaling (NMDS).

The step after the identification of groups of samples possessing similar biological characteristics is to determine the environmental factors responsible of this pattern. For this purpose, biological data are frequently accompanied by a set of environmental variables measured at the same set of sites, assessing their natural characteristics and/or suspected contaminants. These latter measurements are in fact of major concern for bioassessment purposes. Some of the methods generally applied for this analysis are PCA and CCA (McCune, 1997).

In cases where the groups of samples have already been defined (e.g. by means of ordination methods), a univariate or a multivariate analysis can be carried out. A simple method frequently observed in ecological studies is the use of a correlation analysis between the measures that best summarize the biological information (e.g. taxa presence/absence) with the axe of a two-dimension plot of the samples (e.g. CCA) that best represents the gradient of interest (i.e. contamination). In this way, environmental factors which present stronger relationships (slopes) with the axe representing this gradient are considered of influence in the spatial data distribution.

In the present study, CCA and hierarchical cluster analysis using Bray-Curtis distance were applied to determine populations of similar characteristics and environmental variables of interest in the study area.

2.4.3 Two-way indicator species analysis (TWINSPAN)

TWINSPAN (TWo-way INdicator SPecies ANalysis) is a computer programme in FORTRAN designed primarily for ecologists and phytosociologists who have collected data on the occurrence of a set of species in a set of samples. TWINSPAN is a development of a method previously published under the name ‘indicator species analysis’ (Hill et al., 1975).

The programme first constructs a classification of the samples and then uses this classification to obtain a classification of the species according to their ecological preferences. The two classifications are then used together to obtain an ordered two-way table that expresses the species' synecological relations as concisely as possible (Hill, 1994). TWINSPAN was applied in the present study for site clustering as well as taxa classification for further analysis on responses of taxa towards ecological stress.

2.4.4 Data mining

Stream and river modelling based on ecological knowledge and monitoring data has proven to considerably facilitate and improve assessment of stream and river habitats, especially the relationship between environmental variables and the occurrence of certain taxa. The availability of reliable data and suitable modelling techniques, which are able to handle the non-linear and complex nature of ecosystems, allow the development of models with a high reliability (Recknagel, 2002).

Habitat suitability models aim to link abiotic variables, describing the general habitat in a river, to the presence or abundance of a species at a site, based on the ecological preferences of this species. River management can benefit from such predictive models as decision support tools to improve the efficiency of monitoring and assessment, for example by choosing the most optimal restoration measure from a set of given river restoration scenarios (Guisan and Zimmerman, 2000; Marshall et al., 2003).

Currently, the assessment of running water bodies relies on a set of physical, chemical and biological features. The latter are based on a set of characteristics of the biological communities inhabiting the watercourse. For management purposes, it is very important to know the impact of particular changes in any environmental aspect on aquatic ecosystems and more specifically on biological communities (Fig. 2.11).

Undoubtedly, river managers can benefit from these predictive models as decision support tools to improve their efficiency in monitoring and assessment and active management for example, by choosing the most optimal solution for a given river restoration scenario.

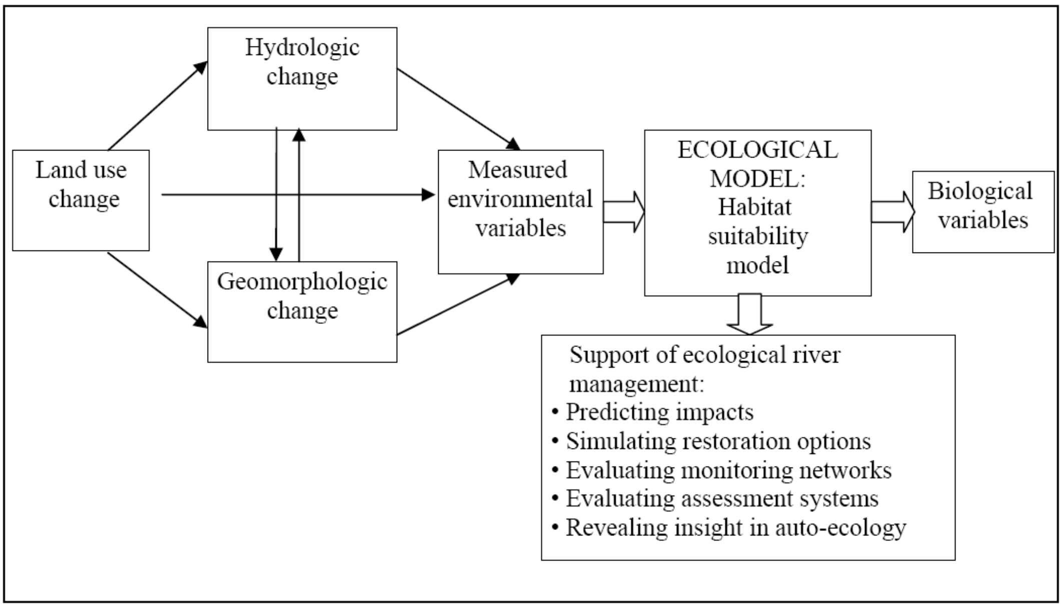


Figure 2.11. Components and objectives of the predictive habitat suitability models for use in ecological river management within the broader environmental modelling framework (Adriaenssens, 2004).

Recently, different modelling techniques have been applied for the assessment of freshwater streams based on the distribution of macroinvertebrates. Artificial neural networks (Hoang et al., 2001; Dedecker et al., 2007; Goethals et al., 2007), fuzzy logic (Adriaenssens et al., 2004a), classification trees (Dzeroski et al., 2000; Goethals et al., 2002; Dakou et al., 2007) and Bayesian belief networks (Adriaenssens et al., 2004b) have proven to have a high potential in macroinvertebrate habitat suitability analysis, as they combine reliable classifications with a convenient interpretation.

In this study, classification trees (CTs) and support vector machines (SVMs), which belong to a generation of inductive modelling techniques, were applied in the present study to develop habitat suitability models to predict the biotic index as well as the presence of selected macroinvertebrate taxa in the Du river. Based on the developed models, variables controlling the macroinvertebrate community structure have been selected for river management purposes.

2.4.5 Water Framework Directive Explorer

Assessing the ecological situation of water bodies and the effectiveness of measures to improve the quality is a complex task for which water managers need an instrument supporting them to analyse the costs and effects of measures on water quality and ecology (Van de Most et al., 2006). Therefore, a user-friendly tool, called the Water Framework Directive Explorer (WFD-Explorer), was developed, which can analyse various options to improve the ecological quality.

The WFD-Explorer has been developed to close the gap between managers and stakeholders and consequently lead to a more integrated water management in the Netherlands (Van Der Most et al., 2006). The toolbox incorporates ecological knowledge as rules which link abiotic variables such as water velocity and nutrient concentrations to the ecological quality of the different water bodies in the river basin of interest. The current river status and the effects of different restoration scenarios are analysed at the river basin level, thus supporting river managers in the development of the river basin management plans needed for WFD implementation.

Although there still remain problems in the application of the WFD-Explorer for estimating the results of physical habitat restoration, the WFD-Explorer proved to be an appropriate tool for modelling changes in chemical water quality and assessing the impact of these changes on aquatic macroinvertebrates (Mouton et al., 2009). Due to its user friendly interface and its clear set-up, the WFD-Explorer can enhance the dialogue on WFD implementation between scientists and water managers. Furthermore, it enables stakeholders to gain insight on how the objectives, the available measures and the impacts of these measures are interrelated, which turns it into a suitable tool for environmental decision making. In this study, the WFD-Explorer is applied in combination with the predictive habitat suitability models to evaluate the possibility to apply this tool as a decision support technique to serve river management purposes in Vietnam.

Chapter 3

Spatial and temporal patterns of macroinvertebrate communities in the Du river basin

3.1 Introduction

During the development of a methodology for river quality evaluation based on bioassessment, the natural and anthropogenic factors influencing the indicator fauna as well as the interactions between and within the physical and chemical components should be determined. Seasonality proved to play a crucial role in the distribution of macroinvertebrates (Dudgeon and Bretschko, 1996; Sporka et al., 2006). Water temperature and water velocity, which are mainly determined by the seasons, are important driving variables for structuring macroinvertebrate communities by both direct effects on habitat conditions as well as indirect effects on variables such as the dissolved oxygen concentration and pollutant loads.

The northern part of Vietnam is situated between the torrid zone and the temperate zone. According to the Köppen system (McKnight and Hess, 2000), the climate of this area is classified as humid subtropical and has four seasons. The humid subtropical climate zone is characterized by hot, humid summers and chilly to mild winters. Humid subtropical climates in Asia differ from those in other continents by having a very pronounced dry winter. In most of this region, there is very little precipitation during winter due to the powerful anticyclonic winds from Siberia and very wet summers due to monsoonal influence. The specific climate conditions in the area may lead to different patterns of macroinvertebrate distributions compared to that in other tropical regions.

The aim of this chapter is to analyse the spatial and temporal patterns in macroinvertebrate community distribution observed in the Du river basin and to assess impacts of environmental stresses on aquatic life. Additionally, the role of the recorded environmental variables in structuring the aquatic fauna is studied.

3.2 Material and methods

3.2.1 Environmental variables

The patterns displayed by biological communities in rivers are determined by both natural and anthropogenic changes in environmental conditions. It is essential for application of biological assessment methods to understand the environmental variables responsible for structuring the macroinvertebrate communities. Physical and chemical variables measured during the whole sampling period with 104 observations were explored to find out potential relationships between these variables. Data reliability was assessed based on expert knowledge during monitoring and analytical work to remove unreliable variables.

Correlations were assessed based on the Pearson correlation coefficient (r). Any significant correlation with an absolute value of at least 0.2 was considered a strong correlation.

3.2.2 Macroinvertebrate communities

A biological database including 7 surveys with 104 samples was explored by means of graphical and statistical methods. A general description of the community structure was provided including dominance and taxa composition. Non-parametric correlations were performed between the biological and environmental variables to determine the major environmental factors accompanying the observed patterns of taxa richness and taxa within insect orders. The Ephemeroptera-Plecoptera-Trichoptera (EPT) taxa richness, an index widely used in bioassessment studies (Marchant, 2007), was also evaluated.

3.2.3 Data analysis

Extensive evidence suggests that seasonality of ecological communities can influence the performance of stream assessment methodologies (Jacobsen and Encalada, 1998; Baptista et al., 2001; Melo and Froehlich, 2001; Ahmad et al., 2002). In order to determine seasonal patterns in macrobenthic communities and environmental variables in the study area, samples collected during the wet and dry periods were compared site by site. Differences between seasons were determined for (a) the total number of taxa, (b) the total number of insect taxa, (c) the EPT index and (d) the environmental variables.

3.2.4 Site clustering

The Student t-test for dependent samples was used to identify differences of environmental characteristics and macroinvertebrate communities between seasons. Correlation analyses were performed with each pair of variables collected during the monitoring campaigns. Any correlation with an absolute value of 0.2 or higher can be considered as a noise source for the analysis performance. The removal of input variables has not only been based on the correlations, but also on data reliability and ecological expert knowledge. The stations were classified into clusters according to the composition of the macrobenthic communities using the classification programme TWINSPAN (Two-Way INdicator SPecies ANalysis) (Hill, 1979). TWINSPAN also yields indicator species characterising the various assemblages. To get insight into the environmental parameters affecting the macroinvertebrate community, the Canonical Correspondence Analysis (CCA) option from the program package PCord4 (McCune and Mettford, 1999) was applied. Prior to CCA analysis, all data were log transformed, except pH, which is already on a log scale. In addition, hierarchical cluster analysis using Bray-Curtis distance was also applied to interpret the main reasons for clustering. It was checked to which conditions of the habitat and anthropogenic impacts classes the clusters belonged.

3.3 Results

3.3.1 Environmental variables

A basic set of environmental variables collected during 2006-2008 were registered in conjunction with the biological data (Appendix 1). Based on the Pearson correlation coefficients and judgment of the reliability of the data collected during monitoring and analysis, redundant variables were removed for further analysis. Based on strong correlations between the iron concentration and the concentration of other metals (r=0.55-0.73) and because of the reliability of the iron measurements, this variable was kept in the dataset as a representative of all metal concentrations in the water column. NO_2-N and Kjeldahl-N were chosen as representatives of nitrogen levels in the water column, while NO_3-N and NH_4-N were removed from the dataset. Since dissolved oxygen saturation and oxygen concentration values were strongly correlated, only the former was retained for analysis. Total phosphorous concentrations in the sediment were not available for all samples, hence orthophosphate and total phosphorous in the water column as well as total nitrogen in the sediment were selected. After removal of the redundant variables, 19 physical-chemical and structural measurements were retained for further analysis. Maximum, minimum, average values and standard deviation of the selected variables observed during the whole study are reported in Table 3.1.

Table 3.1. Observed river characteristics in the Du river basin during the sampling period 2006-2008, (n=104).

Variable	Unit	Min	Max	Mean	Std. Dev.
Altitude	m, a.s.l.	106	380	217	81
Pool/Riffle	category	1	5	-	-
Substrate	category				
Water velocity	m/s	0	0.95	0.27	0.21
Width	m	0.63	28.9	7.51	6.54
Depth	m	0.05	0.95	0.29	0.16
Dissolved oxygen	mg/l	1.0	9.2	6.7	1.4
pH	-	3.10	8.36	6.93	1.15
Conductivity	µS/cm	0	660	250	120

Variable	Unit	Min	Max	Mean	Std. Dev.
Turbidity	NTU	0	441	40	93
Water temperature	^{o}C	14	36	25	5
BOD_5	mg/l	2	20	8	3.7
COD	mg/l	2.7	82.1	23.7	15.2
Kjeldahl N	mg N/l	0.1	14.8	3.2	2.2
NO_2^--N	mg N/l	0.01	0.78	0.08	0.13
Total P	mg/l	0.002	5.65	0.64	0.85
PO_4^{3-}-P	mg P/l	0.001	0.37	0.08	0.09
Cl^-	mg/l	3.9	42.6	17.1	6.9
Total Fe	mg/l	0.014	27.00	2.30	4.04
Total N sediment	mg/kg	0.11	4.98	1.20	0.94

A correlation matrix between the river characteristics is presented in an Appendix 1. Table 3.2 illustrates only correlations, which are considered significant. Results show relatively strong correlations among hydromorphological variables (P/R class, stream width, depth and water velocity). The strongest correlations were observed between the water velocity and the stream depth (r= 0.61; p<0.001), between the altitude and the P/R class and the stream width (r=-0.70 and r=-0.63 respectively; p<0.001) and between the pH and the total Fe concentration (r= -0.72; p<0.001). Analysis also indicated significant negative correlation between the DO and the total Fe concentration (r= -0.64; p<0.001). Significant correlations between BOD_5 and pH (r= 0.4; p<0.001), between BOD_5 and total Fe (r= -0.41; p<0.001) were explained by a low BOD_5 level (BOD_5=2) registered in the Cat stream together with an extremely low pH (down to 3.1) and high concentration of total Fe (up to 27 mg/l) during all monitoring campaigns.

Table 3.2. Pearson correlation coefficients (*r*) calculated between the environmental variables. The presented correlation values are significant at $p<0.05$ (N=104).

	Altitude	**Water velocity**	**Width**	**Depth**	**DO**	**pH**	**Conductivity**	**Temp**	**BOD_5**	**COD**	**Kjd N**	**Total P**
Water velocity	-0.48	1										
Width	-0.63	0.52	1									
Depth	-0.48	0.6	0.58	1								
P/R Class	-0.7	0.33	0.51	0.47								
pH		0.32		0.41	0.37	1						
Temp						-0.25		1				
BOD_5						0.4			1			
COD		0.35		0.25	0.29	0.27				1		
Kjd N								0.53			1	
NO_2^--N				0.26				-0.34		0.23		
Total P											0.31	
PO_4^{3-}-P					0.26		-0.3					1
Cl^-	-0.28	0.32		0.27								-0.24
Total Fe					-0.64	-0.72			-0.41	-0.23		
TN Sed	0.46											

3.3.2 Macroinvertebrate communities

Seventy macroinvertebrate taxa were sampled during seven campaigns carried out between summer 2006 and spring 2008 in the Du river basin. Insects were the dominant group with 48 out the 70 identified taxa. The Hemiptera were the most diverse order with 12 families, followed by Diptera with 8 families, Coleoptera, Odonata and Ephemeroptera with 7 families and Trichoptera with 5 families (Table 3.3). Plecoptera were only represented by the Perlidae family. Non insect taxa include 9 families of Gastropoda, 3 families of Bivalvia, 4 families of Decapoda, 1 family of Isopoda, 3 families of Hirudinea and 2 taxa Oligochaeta and Hydracarina identified at class and order level, respectively.

Table 3.3. List of taxa collected in the Du river basin.

Group	Taxon	Group	Taxon
Oligochaeta	Oligochaeta	Plecoptera	Perlidae
Hirudinea	Erpobdellidae	Hemiptera	Aphelocheiridae
	Glossiphoniidae		Corixidae
	Hirudinidae		Gerridae
Gastropoda	Bithyniidae		Hebridae
	Littorinidae		Hydrometridae
	Pachychilidae		Mesoveliidae
	Pilidae		Naucoridae
	Stenothyridae		Nepidae
	Thiaridae		Notonectidae
	Viviparidae		Pleidae
	Lymnaeidae		Veliidae
	Planorbidae	Megaloptera	Corydalidae
Bivalvia	Corbiculidae	Trichoptera	Brachycentridae
	Pisidiidae		Ecnomidae
	Unionidae		Hydropsychidae
Decapoda	Atyidae		Leptoceridae
	Palaemonidae		Philopotamidae
	Parathelphusidae	Coleoptera	Curculionidae
	Potamidae		Dytiscidae
Isopoda	Corallanidae		Elminthidae
Acarina	Hydracarina		Helodidae
Ephemeroptera	Baetidae		Hydraenidae
	Caenidae		Hydrophilidae
	Ephemerellidae		Psephenidae
	Ephemeridae	Lepidoptera	Pyralidae
	Heptageniidae	Diptera	Athericidae
	Leptophlebiidae		Ceratopogonidae
	Prosopistomatidae		Chironomidae
Odonata	Aeshnidae		Culicidae
	Amphipterygidae		Simuliidae
	Calopterygidae		Stratiomyidae
	Coenagrionidae		Tabanidae
	Corduliidae		Tipulidae
	Gomphidae		
	Libellulidae		

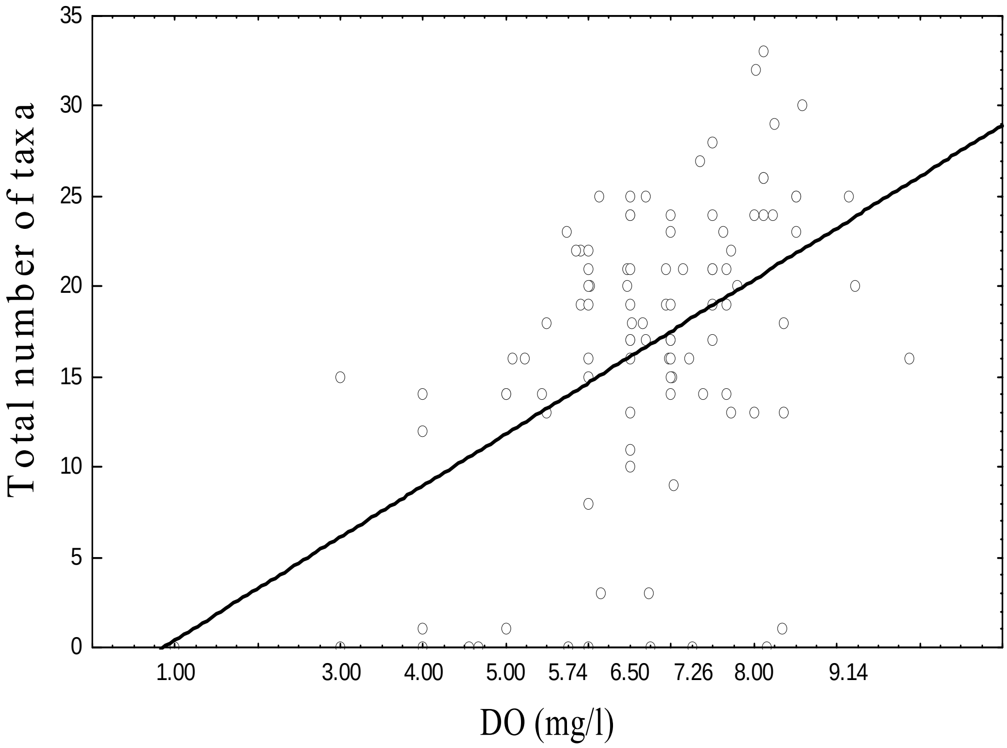
Correlation analysis between the invertebrate taxa richness and the environmental variables registered during the whole study indicated that the total number of taxa, the number of insect families and the number of EPT taxa were significantly related to the P/R Class, altitude and chemical parameters including dissolved oxygen, BOD_5 and total Fe in water column (Fig. 3.1). Significant but weaker correlations were detected between the taxa richness measures and pH, Cl^- and total nitrogen in the sediment. A similar analysis was performed between the environmental variables and the number of families within each insect order, indicating a higher correlation between the number of taxa of Ephemeroptera as well as Odonata and the environmental variables. Total insect taxa and number of Ephemeroptera taxa also showed significant correlations with the stream hydromorphological characteristics represented by altitude and P/R Class (Table 3.4).

Table 3.4. Spearman correlations (ρ) between the taxa richness measures (number of taxa) and the environmental variables registered during the period 2006-2008. The presented correlation values are significant at p<0.05.

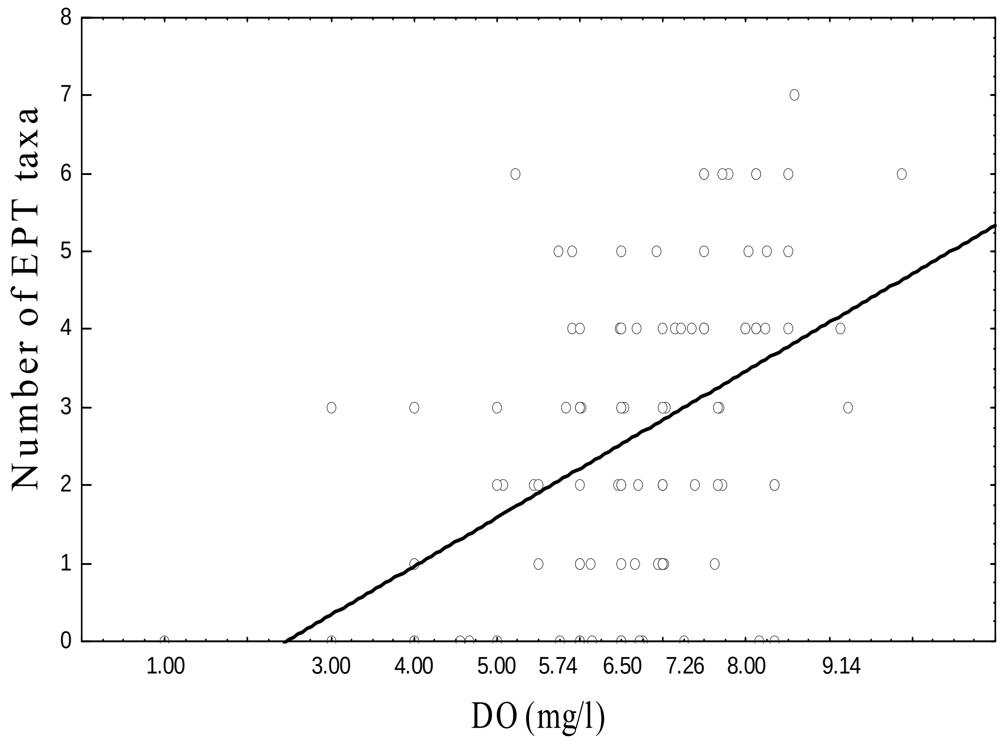
Number of taxa	**Altitude**	**Width**	**P/R Class**	**DO**	**pH**	**BOD_5**	**Cl**	**Total Fe**	**TN Sed**
Total	0.46*	-0.29	-0..34*	0.44*	0.20	0.37*	-0.32	-0.49*	0.31
EPT	0.29		-0.43*	0.45*		0.32*	-0.37*	-0.45*	
Non Insect	0.35*	-0.21	-0.37*	0.41*	0.20	0.36*	-0.39*	-0.45*	0.26
Insect	0.43*	-0.31	-0.21	0.36*	0.24	0.29		-0.47*	0.31
Ephemeroptera	0.34*		-0.44*	0.42*	0.23	0.32	-0.37*	-0.45*	
Odonata	0.27		-0.22	0.35*	0.24*	0.35*	-0.38*	-0.44*	0.23
Hemiptera	0.23	-0.31					-0.23	-0.34*	0.42*
Trichoptera			-0.30	0.33*		0.25	-0.28	-0.32	
Coleoptera			-0.20	0.23	0.3	0.31		-0.24*	
Diptera	0.32*	-0.29	-0.24	0.25	0.35*	0.22		-0.30	0.29
Mollusca	0.38*	-0.23		0.34*	0.22	0.30		-0.46*	0.26

* indicates correlation values are significant at p<0.001.

Taxa richness in rainy and dry season are presented in Fig. 3.2 and Fig. 3.3, however, the extreme sites 10 and 11 were excluded. Although insects were the dominant taxa, a high number of mollusc taxa was found in all sites. During the rainy season, the highest number of taxa was found in upstream sites 1, 5, 6 and 14. The highest number of EPT taxa was also found in these sites. A higher numbers of Odonata taxa occurred in the upstream sites, especially in sites 1 and 5 during the rainy season. During the dry season, the highest number of taxa was found in the downstream sites 4 and 13, with a dominance of mollusc, Coleoptera and Odonata taxa.



(A)



(B)

Figure 3.1. Relationships between taxonomic richness measures (in terms of total number of taxa (A) and number of EPT taxa (B) and dissolved oxygen in the Du river basin during the period 2006-2008.

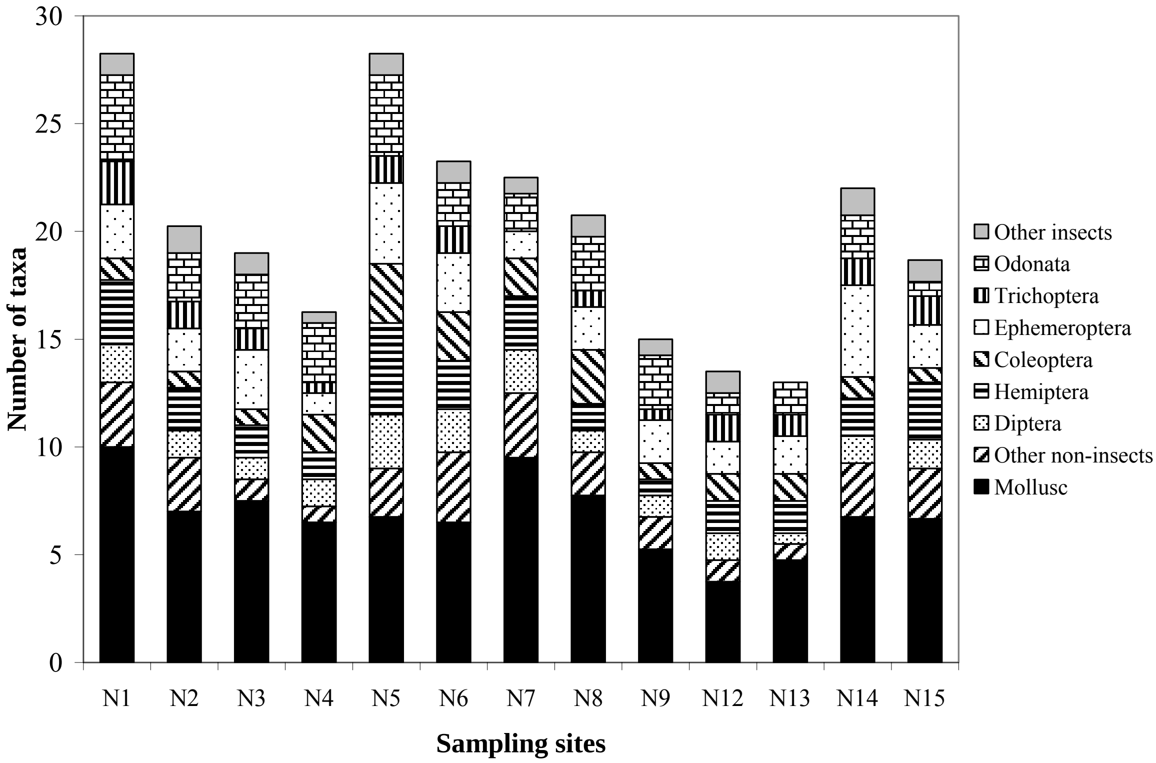


Figure 3.2. Taxa richness composition for sites surveyed during the rainy season.

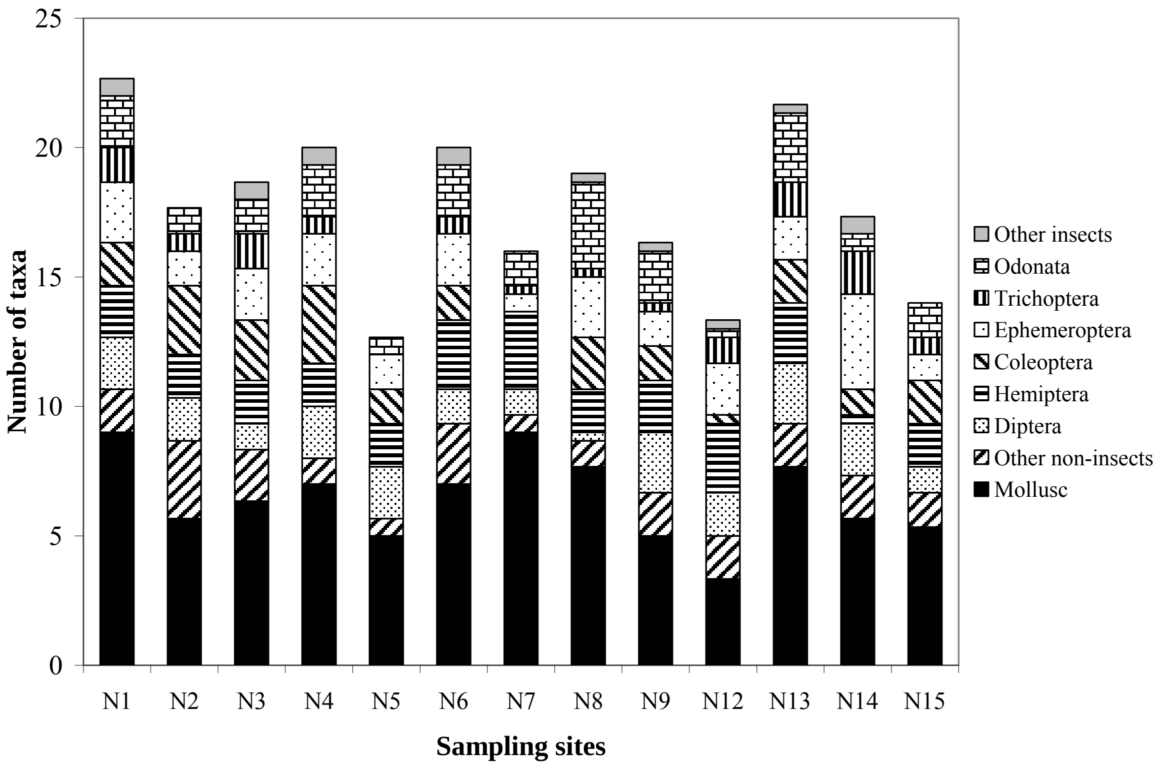


Figure 3.3. Taxa richness composition for sites surveyed during the dry season.

3.3.3 Spatial and temporal patterns

Monitoring took place from summer 2006 until spring 2008 in order to cover 4 campaigns in the wet season (S06, A06, S07 and A07) and 3 in the dry season (Sp07, W07 and Sp08). However, due to instability of the seasonal patterns in the climate transition area and also the fact that a much higher precipitation was observed in the north of Vietnam in 2008 as a consequence of the 'La Niña' phenomenon, the wet season 2008 had already started during the campaign Sp08, which was carried out in mid spring 2008. The dataset therefore contained 74 instances describing river conditions in the wet season and 30 in the dry season.

The results of Student t-tests performed with environmental variables in wet and dry seasons showed that significantly higher values were observed during the wet season for water depth (p=0.001) and water velocity (p<0.001), while mean river width did not change significantly (p=0.16) due to the U-shaped valley of most rivers (Table 3.5). Significantly lower water temperatures during the wet season (p<0.001) reflected the fact that the dry season takes place during winter and spring. Although no significant difference was observed between dissolved oxygen concentration in the water, higher water temperatures lead to increased oxygen saturation during the wet season (p<0.001). The BOD_5 (p=0.009), the Kjeldahl Nitrogen (p=0.005) and the total phosphorous content in the sediment (p=0.001) were higher during the wet season, which could be due to the eroded alluvium. No significant differences between wet and dry season were detected for other habitat characteristics.

Macroinvertebrate communities differed between upstream narrow rivers and wider downstream rivers with an average width of more then eight meters. There were four sites belonging to wide rivers: 4, 8, 12 and 13. During the dry season in winter, spring and early summer, these sites always contained running water and were rich in habitat types while during the wet season, the high water velocity due to the runoff from the tributaries washed away the habitats as well as the invertebrates living at these sites. In the wide river sites, water velocity is therefore a steering variable defining diversity of macroinvertebrate taxa at sampling sites.

The macroinvertebrate communities in wide rivers were therefore poorer during the wet season compared to the dry season (p=0.04). These results were also reflected by the negative correlation between the total number of taxa and the water velocity (r=-0.64). In contrast to the wide rivers, a higher number of taxa was observed during the wet season in the narrow rivers (p=0.02). There were no significant differences in the number of taxa collected between narrow and wide river sites during the dry season, while the diversity in the narrow river sites was relatively higher during the wet season (p=0.0015).

Sensitive taxa seem to favour small tributaries in the warm wet season. A significantly higher number of EPT taxa was observed in narrow sampling sites during the wet seasons compared to the dry seasons. Seasonal dissimilarity in the number of EPT was observed in narrow river sites (p<0.001), but much less in wide rivers (p=0.24).

Table 3.5. Mean (± SD) characteristics of the wet and dry season communities sampled in the Du river basin (excluding site 10, 11) and the significance level of the Student t-test distinguishing differences between them.

	Wet season	Dry season	*p*-value
Taxa richness information			
Number of taxa at wide river sites	16±5	20±4	0.04
Number of taxa at narrow river sites	21±5	18±5	0.02
Number of EPT taxa at wide river sites	2.5±1.8	2.9±1.5	0.24
Number of EPT taxa at narrow river sites	3.8±1.6	2.2±1.0	<0.001
Number of mollusc taxa per site	6.8±2.3	6.7±2.2	0.45
Environmental variables			
Water velocity, m/s	0.34±0.22	0.18±0.15	<0.001
Mean water depth, m	0.34±0.17	0.22±0.1	<0.001
Water temperature (^{o}C)	27±3	19±3	<0.001
DO saturation, %	86±10	71±22	<0.001
BOD_5 (mg/l)	7.2±3.6	9.2±3.3	0.009
Kjeldahl N (mg/l)	2.52±0.9	3.5±2.3	0.005
Cl^-, mg/l	14.4±3	18.5±8.5	0.005
Total P in sediment, mg/g DM	0.095±0.1	0.77±1.1	<0.001

3.3.4 Site clustering

TWINSPAN and CCA were applied for site clustering. Sites 10 and 11 are extremely polluted sites with a pH of 3.14 to 5.52. No taxa were found at site 11, while only Gerridae, Tabanidae and Planorbidae were found at site 10. Samples with no macroinvertebrate taxa presented were excluded from further analysis. Based on the results of TWINSPAN and CCA, the sites were clustered into three groups: a group containing sites in good condition, a group containing sites in poor condition and an intermediate group containing sites with a moderate condition (Fig. 3.4).

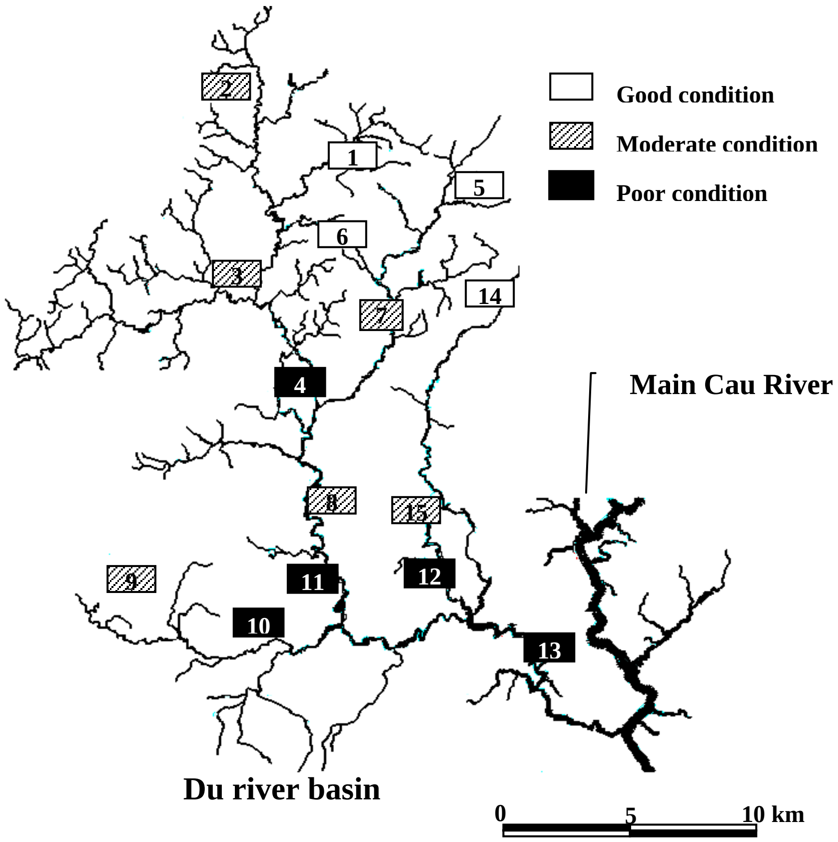


Figure 3.4. Location of the sampling sites in the Du river basin, with indication of the three identified clusters.

CCA performed on the whole dataset showed that the site clustering was based on physical variables (r^2=0.15) including altitude, water velocity, depth, width and P/R class and chemical variables TNSed, DO, PO_4 - P and Total Fe (Fig. 3.5). The Eigenvalues of the first and the second axis were 0.094 and 0.077, respectively. The cumulative percentage variance of species-environment relation explained by the first two axes is 37.2%. Further analysis with CCA was also performed for the wet and the dry season separately. During the wet season, the driving variables were DO, NO_2-N, P/R class, iron content and pH in the water column (Fig. 3.6A). The Eigenvalues for the first and the second axis were 0.108 and 0.102, respectively, whereas percent explained for both axes is 35.3%. In the dry season, the driving variables were the PO_4-P content in the water column, the total nitrogen content in the sediment and the river width (Fig. 3.6B). The Eigenvalues of the first and the second axis were 0.210 and 0.192, respectively, whereas percent explained for both axes is 34.2%.

Hierarchical cluster analysis received an agglomerative coefficient of 0.85. The dendrogram based on Bray-Curtis distance (Fig. 3.7) showed that samples could be clustered into several groups. In the first division, the samples from the extremely polluted site 10 in the Cat stream were separated from the remaining the ones. In the second division, samples from sites downstream with high impacts were split off. In the next division, samples were split off based on altitude and then by impacts. The hierarchical cluster also confirmed clustering by CCA and TWINSPAN. No seasonal variation was observed in the clustering.

TWINSPAN carried out on the whole dataset indicated Hydropsychidae, Caenidae and Corydalidae as indicator taxa of good quality sites. TWINSPAN was also performed separately for the wet and the dry seasons. During the wet seasons, indicator taxa of good quality sites were Hydropsychidae, Caenidae, Leptophlebiidae, Gomphidae and Corydalidae. During the dry season, Hydropsychidae and Caenidae were indicators for sites with a good quality. Two-way ordered tables for the whole dataset and separately for wet and dry seasons are presented in Appendix 2.

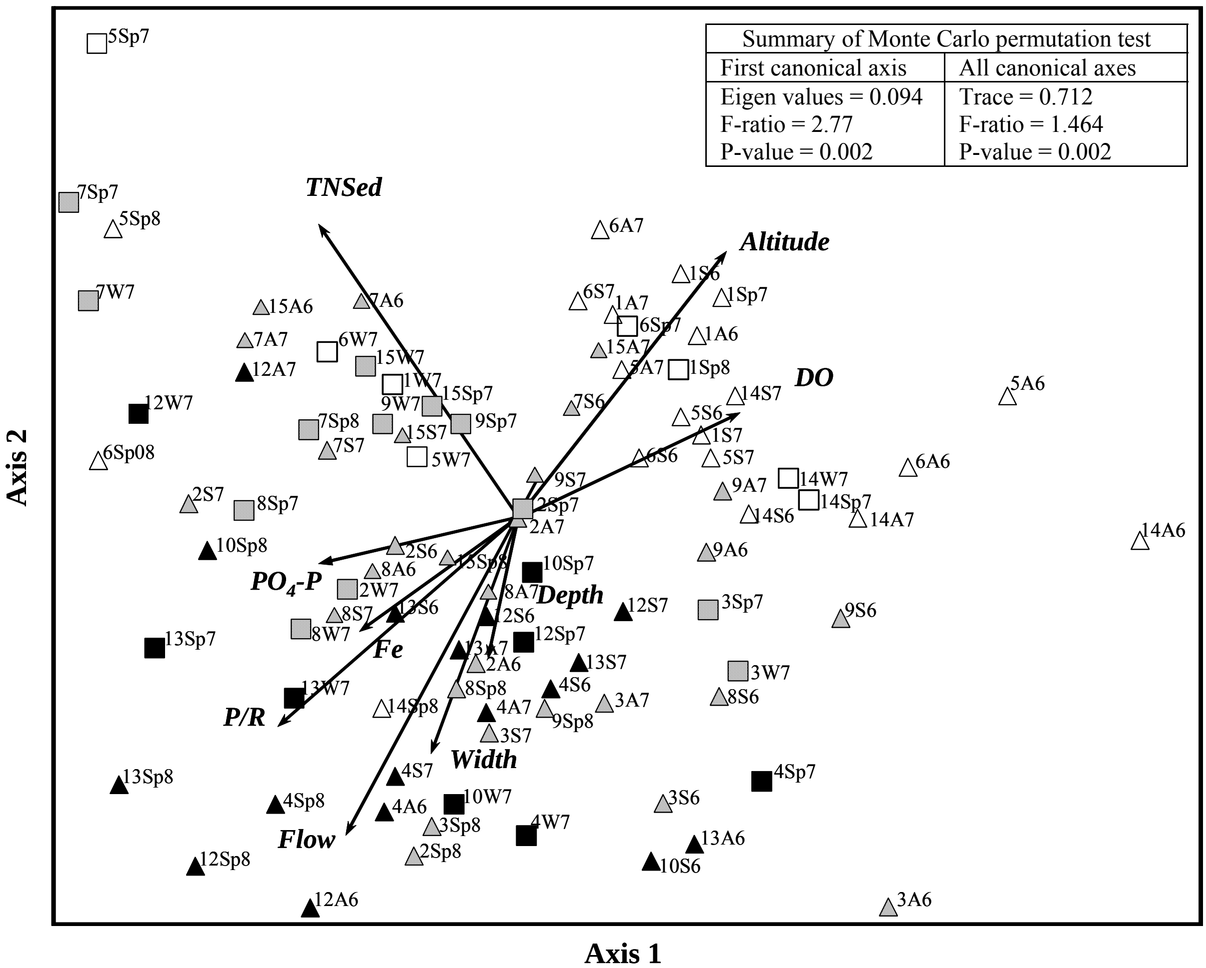
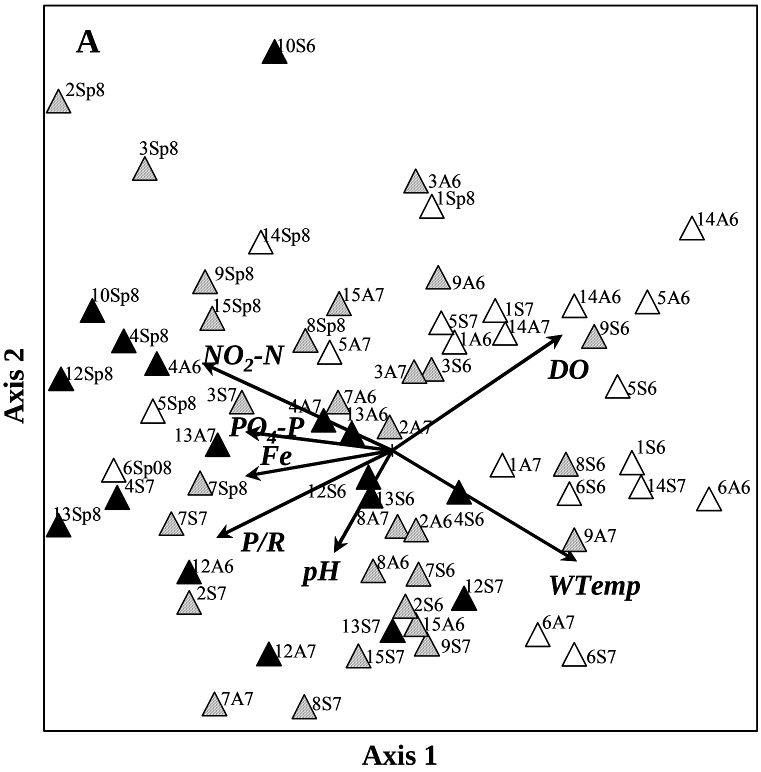
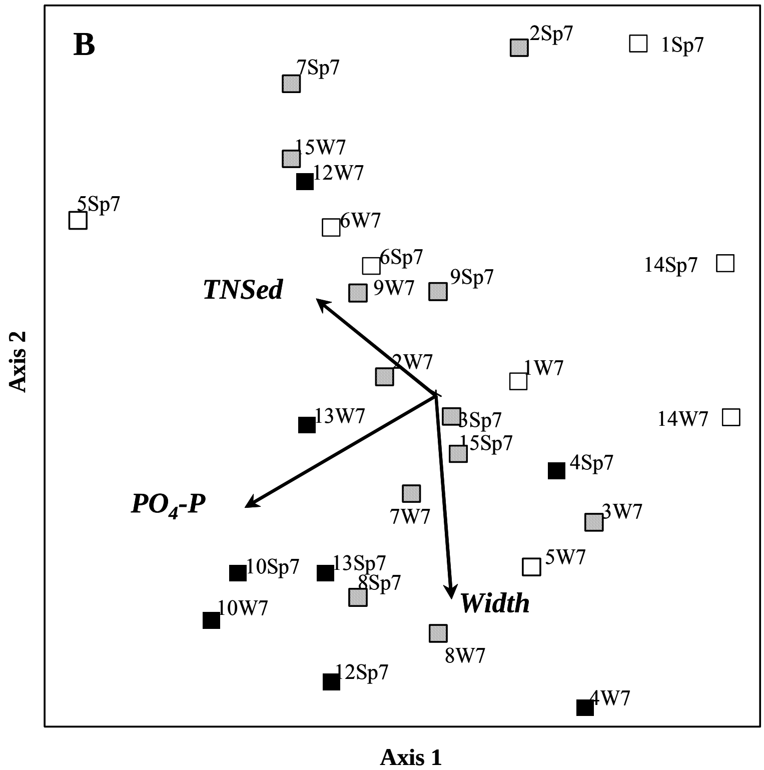


Figure 3.5. Biplot of the sample scores and the environmental variables with indication of the three identified clusters (black: poor condition; dashed: moderate condition; white: good condition) and the sampling season (dry season: squares; wet season: triangles). Labels indicate sampling site (1-15), season (Sp: spring, S: summer, A: autumn, W: winter) and year (200**6**, 200**7**, 200**8**). Cut-off r^2 value: 0.15.



Summary of Monte Carlo permutation test	
First canonical axis	All canonical axes
Eigen values = 0.105	Trace = 0.859
F-ratio = 1.99	F-ratio = 1.248
P-value = 0.05	P-value = 0.002



Summary of Monte Carlo permutation test	
First canonical axis	All canonical axes
Eigen values = 0.208	Trace = 1.736
F-ratio = 0.86	F-ratio = 1.275
P-value = 0.02	P-value = 0.004

Figure 3.6. Biplot of the sample scores and the environmental variables in the wet (A) and the dry (B) season with indication of the three identified clusters (black: poor condition; dashed: moderate condition; white: good condition) and the sampling season (dry season: squares; wet season: triangles). Labels indicate sampling site (1-15), season (Sp: spring, S: summer, A: autumn, W: winter) and year (200**6**, 200**7**, 200**8**). Cut-off r^2 value: 0.15.

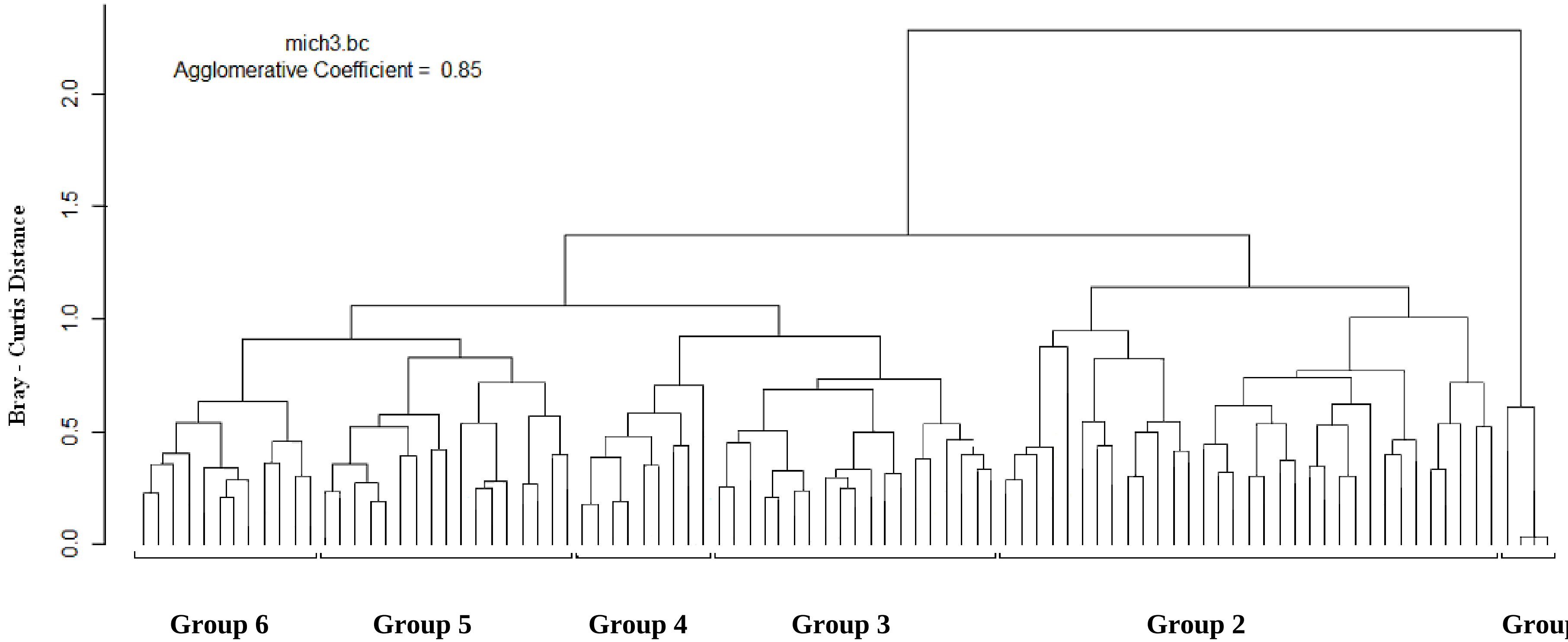


Figure 3.7. Dendrogram of hierarchical cluster analysis based on Bray-Curtis Distance using square root transformed macroinvertebrate data from the Du river basin. Group 1: extremely polluted; Group 2: down stream sites with high impacts; Group 3: sites with low impact at low altitude; Group 4: sites with moderate impacts at low altitude; Group 5: sites with moderate impact at high altitude; Group 6: sites with low impact at high altitude.

3.4 Discussion

The macroinvertebrate diversity in the studied sites (70 taxa including 48 insect families belonging to 9 orders) is relatively poor compared to the diversity in reference sites in Vietnam reported in other studies. For example, in nine sites along a mountain river of Tam Dao National Park (21°30' N, 105°23' E), which also belongs to the Cau river basin in the northern subtropical region of Vietnam, found a higher diversity (63 insect families belonging to 9 orders) (Nguyen et al., 2001). In the Dak Pri river (12°15' N, 107°55' E), a typical tropical mountain river in the south of Vietnam, 91 insect families were detected in eight sites (Hoang and Bae 2006). The high diversity was largely attributed to Odonata, Hemiptera and Coleoptera, which are often considered the most important tropical representatives of aquatic insects (Bishop, 1973; Covich, 1988). A similar community structure was found in the present study with 25 families of Odonata, Hemiptera and Coleoptera.

Although the water quality characteristics of the studied sites did not exceed the norms for fresh watercourses (Surface water quality standard), the diversity in the Du river basin was seriously affected by human activities. The basin is located in a highly populated area with intensive agriculture. The use of fertilizers led to high levels of nitrogen, while wastewater from metal mining and metal extraction activities in some tributaries caused a low pH and high metal concentrations in some sites. Coal mining from the Phan Me open pits located downstream sampling site 8 was also an important source of pollutants. Impacts of anthropogenic activities in the Du river basin on the distribution of the macroinvertebrates were the main topic of this study, which is discussed in following chapters.

The influence of altitude on the invertebrate community composition of running waters has been studied by several authors. Different macroinvertebrate communities could be distinguished along an altitudinal gradient into a lowland (< 400 m), midland (2000 m) and highland (3800 m) community (Suren, 1994; Jacobsen, 2003). Although in the present study, the altitudinal range of the surveyed sites was much narrower in comparison with the previous studies, interesting trends could also be observed. Altitude was significantly correlated with taxa richness of both insects and non insect macroinvertebrate communities (Table 3.4). Additionally, clustering by CCA also discriminated sites from high and lower reaches. All sites located at an altitude above 250m (1, 5, 6 and 14) were clustered separately (Fig. 3.5). Several sensitive taxa occurred strictly in these upper basin sites including Amphipterygidae, Ephemeridae, Ephemerellidae, Heptageniidae.

Although the relationship between the diversity of invertebrate communities and the altitude of sampling sites has been studied by several authors, no definitive trend has been reported at the moment. These studies report both negative relationships (Suren, 1994) and positive (Stout and Vandermeer, 1975; Lang and Reymond, 1993; Pringle and Ramirez, 1998; Jacobsen, 2003) depending on the range of altitudes. However, although such correlations are frequently attributed to natural changes based on the river continuum concept (Vannote et al., 1980), the role of human impacts could be obscuring these natural patterns as is the case in the present study. An inverse relationship between the intensity of human activities and its related environmental degradation (e.g. organic, chemical and morphological) with increasing altitude is commonly observed. The same situation occurred in the Du river basin, where major human activities including agriculture, domestic and industrial activities are carried out in the lowlands. The higher reach of the basin is only slightly affected by human settlements, leaving the running waters of the area in a nearly pristine statue, which is characterised by representatives of the mentioned sensitive taxa.

In riverine ecosystems, the physical environment exerts substantial control over the community composition of macroinvertebrates. Some of these physical factors that are of particular importance to aquatic insects include DO concentrations, water temperature and hydrodynamics. All these factors are characterised by seasonal fluctuations (Wetzel, 2001). Although some dissimilarities between seasons were detected for DO and stream current, those differences were small. Upper reaches were generally characterised by higher DO concentrations and water velocities were higher in wide rivers at the basin outlet. These characteristics lead to a different spatial distribution of macroinvertebrates in the upper narrow streams and wide lowland river sites.

In contrast to temperate areas, tropical regions are not characterized by strong thermal changes over time (Dominguez-Granda, 2007). Many studies on temporal patterns of macroinvertebrates in tropical as well as neotropical areas revealed no significant change in diversity between seasons (Jacobsen and Encalada, 1998; Baptista et al., 2001; Melo and Froehlich, 2001; Ahmad et al., 2002; Hoang and Bae, 2006). In this study, seasonal patterns in macroinvertebrate community composition were detected in the Du river basin, which could be explained by the specific climatic condition in the humid subtropical zone in the north of Vietnam causing temperature variations between seasons. Overall, there was a higher invertebrate diversity during the warm wet season in the upstream sites, while wide river downstream sites showed a higher diversity during the dry season.

Although some differences in macroinvertebrate communities were observed in the Du river between seasons, temporal dissimilarities in the Du river were considered insignificant. The differences between the number of taxa observed in the dry and the wet seasons were usually less than 20%. Low numbers of taxa in downstream sites 4 and 13 were mainly the result of spates, which flushed away habitat and macroinvertebrates at the sites during the wet season. In the dry season, upper sites dried up or became pools causing a reduction of the number of taxa at the sites (e.g. site 5 and 14). However, this happened only in the first order streams during the middle of the dry season, which is not characteristic for the whole river basin.

Minor temporal variations do not require modifications in the bioassessment methodology for watercourses in Vietnam compared to other tropical regions. However, based on the results of the present study, it is advised to avoid sampling large tributaries shortly after spates, while during the dry season, sampling of small tributaries that stop flowing should also be avoided.

3.5 Conclusions

Macroinvertebrate communities in the Du river basin dominated by aquatic insects were considered poorer than reference conditions in other tropical river in Vietnam. Invertebrate taxa richness was higher during the wet season in most of the sampled sites in the upper part of the basin, while in those sites located in the wide river sites in the lowland, an opposite trend was observed. Higher stream currents during the wet season are probably improving the water and habitat quality of the sites located in the upstream reaches. In the sites of the lower basin, where the riverine ecosystems seem to be already maximally colonized during the dry season, organisms are probably flushed away by the effect of higher stream velocities (natural physical stress in combination with higher drift) during the rainy season. Spatial and temporal patterns of the macroinvertebrate communities in the study area need to be considered in the development and evaluation the effectiveness of biological assessment methods applied in watercourses in Vietnam.

Chapter 4

Relevance of biotic indices based on macroinvertebrates for assessment of watercourses in northern Vietnam

4.1 Introduction

During the history of running water quality assessment numerous biological indicator systems have been developed and accepted for general use in many countries. At least 100 indices have been developed over the past tens of years, of which about 60 % are biotic ones based on macroinvertebrate analysis (De Pauw and Hawkes, 1993). The Biological Monitoring Working Party (BMWP) scoring system from the UK (Armitage et al., 1983) for example has been applied in many countries, especially in developing countries because they can be employed with qualitative, family level data. However, the application of biological indices for environmental conditions or pollution types other than the ones they were developed for requires adaptation (Alba-Tercedor and Pujante, 1997).

In Vietnam, a preliminary BMWP Viet scoring system was proposed as an outcome of a research project supported by the UK Government Darwin Initiative (Nguyen et al., 2004). The BMWP score was selected for use mainly because of its simplicity. Initial modifications were proposed for Vietnam based on the original BMWP and the revised BMWP-Thai. However, these studies were only performed in two small streams with near pristine conditions and sites with minor organic pollution, while impacted sites were not assessed. Also stream morphological and hydrological characteristics, which are crucial for macroinvertebrate communities in watercourses, were neglected in these studies. Nguyen et al. (2004) therefore mentioned that the BMWP-Viet needed further improvement.

The aim of this chapter is to evaluate the performance of the present BMWP-Viet in comparison with abiotic as well as similar biotic indices to assess impacts of the environmental stresses on macroinvertebrates in tropical rivers in the northern of Vietnam. Recommendations for practical application and further improvements of the method are suggested at the end of the chapter.

4.2 Material and methods

4.2.1 Database

4.2.1.1 Physical - chemical characteristics and water quality

A brief description of physical and chemical characteristics of the sampling sites is given in the Table 3.1 (see chapter 3). Low levels of pH were measured at site 10 and 11, with average values of 4.0 and 4.1, respectively, which are caused by discharges from small private tin extraction workshops along the Cat stream. In the main Du river, this water is diluted and the pH in these downstream sites therefore returned to its normal range. High levels of BOD_5, COD (up to 20 mg/l and 82 mg/l, respectively) and low levels of dissolved oxygen (0.24 mg/l) in the downstream site (15) of the Khe Coc and the main Du river (12, 13) during several campaigns in the dry seasons (Sp07, W07) suggest that these sites may be impacted by organic pollution due to intensive agricultural activities and domestic wastewater discharges.

Metals are expected to be important pollutants in the Du river due to mining and ore extraction activities in the river basin. ICP-MS was applied to discover potential metal contamination in the Du river. However, due to technical constraints and budget limitations, only the six most easily detectable metals were analysed in only 71 selected monitoring samples including all downstream sites, where metals were expected to be present and some reference sites upstream (Table 4.1). Results are compared with the Vietnamese standards for surface water quality (TCVN 5942b), which are to be used for purposes other than domestic water.

Table 4.1. Metal concentrations in the Du river basin.

Variable	Unit	Max	Min	Mean	Std.Dev.	TCVN 5942	VLAREM II
Al	mg/l	7.63	0.01	0.31	1.12	-	-
As	mg/l	0.74	<0.001	0.02	0.10	0.1	0.03
Cu	mg/l	9.49	<0.001	0.38	1.47	1	0.05
Mn	mg/l	11.48	<0.001	0.69	1.31	0.8	0.2
Pb	mg/l	0.06	<0.001	0.01	0.01	0.1	0.05
Zn	mg/l	1.06	<0.001	0.07	0.13	2	0.2

TCVN 5942. Vietnamese standards for surface water quality.
VLAREM II. Flemish regulation concerning environmental licenses (Basic environmental quality standards for surface waters in Flanders).

4.2.1.2 Macroinvertebrate communities

Seventy macroinvertebrate taxa were found during seven campaigns between summer 2006 and spring 2008 in the Du river basin. Insects were the dominant group, with 48 out of 70 identified taxa (see also chapter 3). During each sampling effort, the lowest number of taxa was found in sites 10 and 11 (0 - 3 taxa), the maximum number of taxa was observed in the headwater sites 1, 5 and 14 (up to 33 taxa). At these locations also the most sensitive taxa were found. The stonefly family Perlidae, the caddisfly families Brachycentridae and Leptoceridae, the mayfly families Heptageniidae, Ephemerellidae and Ephemeridae and the damselfly family Amphipterygidae were found only in these sites, which are located in remote areas with minor human impact. The most common taxa, which are present in more than 50% of the sampling sites, include Mollusca families (Corbiculidae 57%, Pachychilidae 65%, Thiaridae 51%, Viviparidae 60%), Decapoda families (Palaemonidae 65%, Atyidae 62%, Parathelphusidae 65%), Diptera family Chironomidae (72%), Hemiptera family Gerridae (58%), Ephemeroptera families Baetidae (74%), Caenidae (57%) and the Trichoptera family Hydropsychidae (55%).

4.2.2 Biotic indices tested

4.2.2.1 SIGNAL

SIGNAL (Stream Invertebrate Grade Number – Average Level) is a scoring system for macroinvertebrate samples from Australian rivers. A SIGNAL score gives an indication of the water quality in the river where the sample is collected. Rivers with high SIGNAL scores are likely to have low levels of turbidity and nutrients such as nitrogen and phosphorus. They are also likely to contain high concentrations of dissolved oxygen (Chessman, 1995).

Each macroinvertebrate taxon has a 'grade number' between 1 and 10. A low grade number indicates that the macroinvertebrate is tolerant for a range of environmental conditions, including common forms of water pollution. A sensitive macroinvertebrate receives a high grade. The grades are appointed to phylum, class and family level of identification (Chessman, 2003b).

SIGNAL2 scores can be calculated with or without abundance weighing. If no weighing is used, the SIGNAL score is the average of the grade numbers of the collected macroinvertebrate taxa (Chessman, 2003a). Since the data were collected disregarding abundances of macroinvertebrates, the latter calculation was applied for the SIGNAL2 scoring. The status of the sampled sites is classified according to the SIGNAL2 values for clean water up to severe pollution (Chessman, 1995). In the currently updated system for index interpretation, a comparison with reference conditions obtained with the AUSRIVAS software is suggested (Chessman, 2003b). This method of index interpretation has, however, not been applied in this study.

4.2.2.2 SASS

The South African Scoring System Version 5 (SASS5) is being applied for rapid bioassessment of rivers in South Africa. The method is a refinement of the SASS method developed by Chutter (1994). With several improvements in the latest version, the SASS5 is nowadays adopted for the national river monitoring programme in South Africa (Dickens and Graham, 2002). A tolerance score is given to each taxon identified at family level. The values range between 1 and 15, with the higher values assigned to the more sensitive taxa. For Baetidae and Hydropsychidae, their scores are assigned according to the encountered number of species of each family (for both taxa: scores of 4, 6 and 12 are given if only 1, 2 or >2 species are identified at the site). In this study, Baetidae and Hydropsychidae were not identified at lower taxonomical level. Therefore, the lowest tolerance scores were assigned to these taxa (4 in both cases). For this study, the SASS5 results are expressed as the SASS score by summing the tolerance scores of all representative taxa and as SASS-ASPT by dividing the SASS score by the number of taxa involved in the calculation.

4.2.2.3 HKHbios

In the nineties, a first region-specific score based method for Nepalese rivers was developed and called the NEPBIOS (Nepalese Biotic Score), which is an adaptation of the BMWP/ASPT system (Sharma, 1996; Sharma and Moog, 1996). Since then, several modifications and applications of the NEPBIOS followed and more recently, a new score (HKHbios) based on the biological assessment method was developed for rivers in the Hindu Kush-Himalayan region, which is applicable in a wider geographical range including 5 countries: Bangladesh, Bhutan, India, Nepal and Pakistan (Hartman et al., 2008).

Within the HKHbios assessment system, macroinvertebrates are ranked using a ten point scoring system. The scores reflect the sensitivity of taxa to organic pollution (oxygen depletion), chemical pollution as well as to hydromorphological changes. Tolerance values are assigned to indicator taxa, mostly at family level identification and for many families, scores for taxa identified at lower level are also provided. The procedure for assigning scores follows Sharma (1996), where in a first step, a numerical procedure is applied comprising the calculation of a so called ‘Guide Score’, obtained from presence/absence of the taxon determining the number of sites represented by the taxon in each pollution class according to the pre-classification. In a second step, the final value is assigned by expert judgement.

By applying an iterative process, the best combination of scores is evaluated aiming to achieve the highest accuracy in discriminating different river quality classes for each ecoregion. The resulting scoring list includes 199 taxa in total. It comprises 2 taxa at class level, 139 at family level, 4 at subfamily level, 51 at genus level and 3 at species level. For lowlands, 155 taxa are scored (Hartman et al., 2008).

During the development of the HKH bios, the effect of including abundance values and weighting of selected taxa was tested. Weighted average scores per taxon including abundance classes ($ASPT_{WA}$) were calculated as

$$ASPT_{WA} = \frac{\sum_{1}^{i} Score_i * Weight_i * abundance_i}{\sum_{1}^{i} Weight_i * abundance_i}.$$

The HKHbios can be calculated without abundance and weighting (ASPT), including abundance ($ASPT_A$), including weighting of taxa ($ASPT_W$) or with both abundance and weighting ($ASPT_{WA}$).

The derivation of threshold values to discriminate between different stages of stress is based on the index ranges depending on the estimated quality class (pre-classification). It is defined for each ecoregion separately including: IM0120- Lower Gangetic Plains Moist Deciduous Forest; IM0166 - Upper Gangetic Plains Moist Deciduous Forest, IM0301 - Himalayan Subtropical Pine Forests, IM0401- Easter Himalayan Broadleaf Forest, IM0403 - Western Himalayan Broadleaf Forest. The Himalayan Subtropical Pine Forests region is the closest to northern Vietnam. The class system defined for this region (Hartman et al., 2008) has been used for site classification in the present study.

Although Nepal, India and Vietnam belong to the same biogeographic region of Indo-Malaya according to the WWF ecozones classification (Udvardy, 1975), which is characterised by the same evolutionary history of plants and animals, watercourses in Nepal are located at remarkably higher altitude compared to rivers in northern Vietnam. Although the geographical range may cause different sensitivities of macroinvertebrate taxa to stressors, the scoring system of HKHbios can be adapted for application in other areas.

4.2.2.4 BMWP

The Biological Monitoring Working Party (BMWP) index originated from the Trent Biotic Index (TBI), the first biotic index developed for river water assessment (Armitage et al., 1983) and has been improved through many years of research. The index is currently the standardized method for the biological assessment in many countries. The BMWP index has been adapted for use in other European countries such as Poland (Czeniawska-Kusza, 2005) and Spain (Alba-Tercedor and Sanchez-Ortega, 1988), in South America including

Brazil (Baptista et al., 2007), Colombia (Roldan, 2003), Costa Rica (Astorga et al., 1997), Ecuador (Jacobsen, 1998) and in Asian countries like India (De Zwart and Trivedi, 1994), Nepal (Sharma and Moog, 1996), Indonesia (Trihadiningrum et al., 1996) and Thailand (Mustow, 2002).

For the BMWP index calculation, each macroinvertebrate family is scored according to their sensitivity to the organic pollution. Tolerance values are assigned from 1 (the most tolerant taxa) to 10 (the most sensitive taxa). A standard list is used (National Water Council, 1981), which rates 85 common families (including the class Oligochaeta). The value of the index is the total of the scores of all taxa present in the list, where each family in the sample is counted once, regardless the number of species and individuals.

$$BMWP = \sum_{i=1}^{n} ToleranceScore_i$$

The BMWP index can be divided by the number of scoring families present in the taxa list, the result of which is known as the Average Score Per Taxon (ASPT) index.

Adaptations for the use of the BMWP in different countries are made by including different taxa and adapting tolerance scores.

4.2.2.5 BMWP-Thai

The BMWP approach was applied in some rivers in Thailand in association with physical and chemical data from pollution gradient (Mustow, 1997; Thorn and Williams, 1997; Mustow, 1999). The studies showed that the effects of organic pollution on lotic macroinvertebrate assemblages were broadly the same as those that are known to occur in temperate regions. However in response to several perceived problems with using the BMWP score in Thailand, the system was adapted and referred to as the BMWP-Thai score (Mustow, 2002). Major adaptations include:

(1) eleven additional families were included in the BMWP-Thai score, based on frequency of occurrence, distribution and ease of identification;

(2) fifteen redundant families were omitted from the BMWP-Thai score;

(3) families, that were found to be difficult to differentiate during the identification, were combined for index calculation;

(4) all Odonata families allocated a BMWP score of 8 received a lower BMWP-Thai score of 6 due to their frequent occurrence at highly organically polluted sites.

However, there is no specific quality class attribution system available for the interpretation of the index values. Index interpretation of the original BMWP is generally used for river quality classification.

4.2.2.6 BMWP- Viet

The BMWP-Viet scoring system, including scores for 90 taxa, was developed based on a similar system used in the UK with some modifications to take into account some differences between the two regions concerning the tolerance of some families.

The adaptation was based in part on the work done by Mustow (1997) and others in Thailand but also on the distribution of some families in relation to environmental characteristics in two catchments, one in the north and one in the south of Vietnam, both sampled in April and August 1999 (Nguyen et al., 2004). In their study, similar results were obtained as in temperate regions on which the used method of biological monitoring was based. In general, comparing lists of tolerance scores from Thailand, the scores of taxa for Vietnam remained mostly unchanged in the BMWP-Viet scores. However, some adaptations were recommended based on the outcome of this study and expert knowledge:

(1) Odonata seem in general to be more tolerant than their temperate counterparts. A similar observation led Mustow (1997) to downgrade families of Odonata in the BMWP-Thai, a refinement that appears to be justified also by the results of this study. In the BMWP-Viet, the families Coenagrionidae, Corduliidae and Libellulidae were even further downgraded to a score of 4 in order to reflect their tolerance to pollution. A notable exception is the family Amphipterygidae, which was only found in clean, well oxygenated sites, scoring 10;

(2) two Mollusca families, Unionidae and Viviparidae, were also given a lower score than indicated by their position in BMWP-Thai, because they showed evident tolerance to pollution at the sites they inhabited;

(3) additional scores were assigned to a number of families recorded in this study and also to other families that are known to occur or are likely to occur in Vietnam. Several of these were included in the BMWP-Thai. However, Mustow (2002) regarded the two families of prawn Atyidae and Palaemonidae to be pollution intolerant (score of 8), while these were also found in water of very poor quality, just like the crab family Parathelphusidae. These two prawn families were therefore given scores of 3.

A frequency distribution of the BWWP-Viet scoring system shows that 72.2 % of the taxa score from 3 to 6, 18% score 10, no taxa score 9, only 5 taxa score 7 or 8, score 1 is given to Oligochaeta and 2 to Chironomidae, the latter two taxa being given the lowest values in the scoring system (Fig. 4.1).

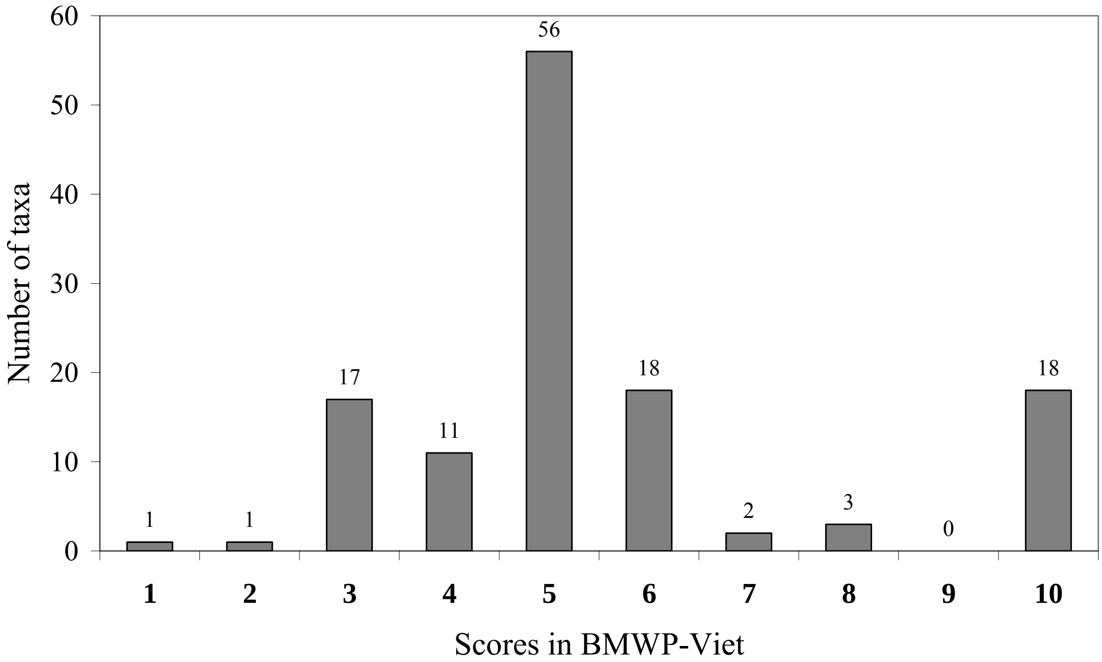


Figure 4.1. Frequency distribution of tolerance scores in the BMWP-Viet.

Nguyen et al. (2004) emphasised that the produced BMWP-Viet is intended as a working solution that requires to be expanded and further improved as the fauna of Vietnam and the distribution and tolerance of macroinvertebrate families in Vietnam rivers become better known as a results of more extensive surveys.

An example is given for taxa collected and identified at site 8 in the spring campaign 2008 (coded 8Sp8). A similar procedure as for the original BMWP has been applied for the calculation of the BMWP-Viet and ASPT-Viet. In the sample 8Sp8, 15 taxa were collected and identified, and scores could be assigned to 12 taxa according to the BMWP-Viet scoring system (Table 4.2).

Table 4.2. List of taxa identified in sample 8Sp8 with tolerance scores according to BMWP-Viet.

Taxa	Score BMWP-Viet	Taxa	Score BMWP-Viet
Pachychilidae	No score	Baetidae	4
Viviparidae	4	Coenagrionidae	4
Lymnaeidae	3	Mesoveliidae	5
Atyidae	3	Nepidae	5
Palaemonidae	3	Dytiscidae	5
Parathelphusidae	3	Elminthidae	5
Hydracarina	No score	Hydraenidae	5
		Pyralidae	No score

The BMWP-Viet is calculated as follows:

$$BMWP = \sum_{i=1}^{n} ToleranceScore_i$$

In which: Tolerance $Score_i$: Tolerance Score of taxon *i* present in the sample according to the BMWP-Viet Scoring system

n: total number of taxa used for the calculation. In the above example n=12

BMWP-Viet = 4+3+3+3+3+4+4 +5+5+5+5+5= 49

$$ASPT - Viet = \frac{BMWP - Viet}{n} = \frac{49}{12} = 4.25$$

The only taxon belonging to the EPT taxa group was Baetidae. The EPT index calculated for this sample was thus one.

4.4.2.7 Ephemeroptera, Plecoptera, Trichoptera Richness Index (EPT)

Among the aquatic insects, Ephemeroptera, Plecoptera and Trichoptera comprise rich assemblages. These organisms occur in well oxygenated waters and are also sensitive to environmental changes. Therefore, EPT assemblages are frequently considered to be good indicators of water quality (Rosenberg and Resh, 1993).

The EPT evaluates the environmental quality by measuring the richness of these sensitive taxa at a site. This index is calculated by summing the number of distinct taxa within these three orders. The result indicates water quality based on the general principle of associating low EPT values to degraded water quality. However, since the use of EPT in water quality classifications remains equivocal, EPT often accompanies other indices and shows comparisons between impacted and reference sites in many studies (Soldner et al, 2004; Sharma et al., 2005; Bispo et al., 2006; Dinakaran and Anbalagan, 2007).

4.4.2.8 Interpretation of biotic indices

Different quality class attribution systems are applied for the interpretation of the tested indices 5 quality varying classes from very good to bad (Table 4.3). The original BMWP and ASPT quality class attribution is applied for all BMWP scoring systems (*http://www.bournestreampartnership.org.uk/bmwp_scoring.htm*). The class system of HKHbios developed for the Himalayan Subtropical Pine Forest region is used for the delineation of the quality classes (Hartmann, 2008). National specific class systems are applied for SASS5 and SIGNAL2 as derived by Dallas and Day (2007) and Chessman, (1995), respectively.

Table 4.3. Quality class attribution systems for the interpretation of the tested biotic indices.

Class	Description	SIGNAL2	SASS5 Score	SASS5 ASPT	HKHbios	BMWP	ASPT
Very good (VG)	Richer than references potential biodiversity 'hot spot'	≥6.0	≥166	≥9.0	≥6.8	>100	≥8.0
Good (G)	Reference	5.0-6.0	137-165	8.2-9.0	5.7-6.8	71-100	6.0-8.0
Moderate (M)	impairment of water quality and/or habitat with loss of pollution-sensitive taxa	4.0-5.0	108-136	7.4-8.2	4.7-5.7	41-70	5.0-6.0
Poor (P)	substantial impairment of water quality and/or habitat; major loss of pollution-sensitive taxa	3.0-4.0	79-107	6.6-7.4	3.4-4.7	11-40	3.0-5.0
Bad (B)	severe impairment; remaining taxa pollution-tolerant	<3.0	<79	<6.6	<3.4	0-10	< 3.0

In the present study, the performance of several biotic indices SIGNAL2 (Chessman, 2003a), SASS5 Score (Dickens and Graham, 2002), BMWP Original (National Water Council, 1981), BMWP Thai (Mustow, 1999), SASS-ASPT and HKHbios (Hartman, 2008) (Table 4.4) have been compared with the BMWP-Viet (Nguyen et al., 2004), ASPT-Viet and EPT for the assessment of the Du river basin in section 4.3.

Table 4.4. List of tolerance scores provided in different biotic indices. The list is based on the taxa available from the BMWP-Viet together with the list of taxa collected during the study in the Du river (marked as *). Blank cells indicate that no tolerance scores have yet been assigned.

Class/Order	No.	Taxa	SIGNAL 2 (1-10)	SASS5 (1-15)	HKHbios (1-10)	BMWP original (1-10)	BMWP Thai (1-10)	BMWP Viet (1-10)
Oligochaeta	1.	Oligochaeta*	2	1	2	1	1	1
Hirudinea	2.	Erpobdellidae*	1	3		3	3	3
	3.	Glossiphoniidae*	1	3	4	3	3	3
	4.	Hirudinidae*		3		3	3	3
Gastropoda	5.	Bithyniidae*	3		5			3
	6.	Littoridinidae						3
	7.	Littorinidae*						3
	8.	Pachychilidae*						
	9.	Pilidae*						4
	10.	Stenothyridae*			8			
	11.	Thiaridae*	4	3	4		3	3
	12.	Viviparidae*	4	5	6	6	6	4
	13.	Lymnaeidae*	1	3	4	3	3	3
	14.	Planorbidae*	2	3	6	3	3	3
Bivalvia	15.	Amblemidae			7			4
	16.	Corbiculidae*	4	5	4		3	3
	17.	Mytilidae						5
	18.	Pisidiidae*			7			
	19.	Unionidae*		6	6	6	6	4
Decapoda	20.	Atyidae*	3	8	6		8	3
	21.	Palaemonidae*	4	10	6		8	3
	22.	Parathelphusidae*			6		3	3
	23.	Potamidae*			7			8
Isopoda	24.	Corallanidae*						
Acarina	25.	Hydracarina*	6	8				
Ephemeroptera	26.	Baetidae*	5	4	5	4	4	4
	27.	Caenidae*	4	6	7	7	7	7
	28.	Ephemerellidae *	9		10	10	10	10
	29.	Ephemeridae*		15		10	10	10
	30.	Heptageniidae*		13	10	10	10	10
	31.	Leptophlebiidae*	8	9	7	10	10	10
	32.	Oligoneuridae		15				10
	33.	Potamanthidae			10	10	10	10
	34.	Prosopistomatidae *	4	15	10			
	35.	Siphlonuridae	10			10		4
Odonata	36.	Amphipterygidae*	6		7			10
	37.	Aeshnidae*	4	8	7	8	6	6
	38.	Calopterygidae*		10	8	8	6	6
	39.	Chlorocyphidae		10	6		6	6
	40.	Coenagrionidae*	2	4	5	6	6	4
	41.	Cordulegastridae			10	8	6	6
	42.	Corduliidae*	5	8	5	8	6	4
	43.	Gomphidae*	5	6		8	6	6
	44.	Lestidae	1	8	7	8		6
	45.	Libellulidae*	4	4	6	8	6	4
	46.	Macromiidae	8		7		6	6
	47.	Platycnemididae		10	7	6		6
	48.	Protoneuridae	4	8	5		3	3
Plecoptera	49.	Leuctridae			10	10		10
	50.	Nemouridae	6	14	9	7	7	7
	51.	Perlidae *	10	12	8	10	10	10
	52.	Perlodidae			9	10		10
Hemiptera	53.	Aphelocheiridae *			7	10	10	10
	54.	Belostomatidae	1	3	6			5
	55.	Corixidae*	2	3	6	5	5	5

Class/Order	**No.**	**Taxa**	**SIGNAL 2 (1-10)**	**SASS5 (1-15)**	**HKHbios (1-10)**	**BMWP original (1-10)**	**BMWP Thai (1-10)**	**BMWP Viet (1-10)**
	56.	Gerridae*	4	5		5	5	5
	57.	Hydrometridae *	3	6		5	5	5
	58.	Hebridae *	3		7			5
	59.	Mesoveliidae *	2	5	6	5	5	5
	60.	Naucoridae *	2	7	7	5	5	5
	61.	Nepidae *	3	3	6	5	5	5
	62.	Notonectidae*	1	3	3	5	5	5
	63.	Pleidae *	2	4	4	5	5	5
	64.	Veliidae*	3	5	7			5
Megaloptera	65.	Corydalidae *	7	8	7		4	4
Neuroptera	66.	Sialidae	5	6	7	4	4	4
Trichoptera	67.	Brachycentridae*			8	10	10	10
	68.	Ecnomidae *	4	8	6			
	69.	Goeridae			9	10	10	10
	70.	Hydropsychidae*	6	4	9	5	5	5
	71.	Hydroptilidae	4	6	7	6	6	6
	72.	Lepidostomatidae		10	8	10	10	10
	73.	Leptoceridae*	6	6		10	10	10
	74.	Molannidae				10	10	10
	75.	Odontoceridae	7		7	10	10	10
	76.	Philopotamidae*	8	10	7	8	8	8
	77.	Phryganeidae			8	10	10	10
	78.	Psychomyiidae			7	8	8	8
Coleoptera	79.	Chrysomelidae	2		8			5
	80.	Curculionidae *	2		5	5	5	5
	81.	Dryopidae			5	5	5	5
	82.	Dytiscidae*	2	5	5	5	5	5
	83.	Elminthidae*	7	8	8	5	5	5
	84.	Gyrinidae	4	5	6	5	5	5
	85.	Haliplidae	2	5	6	5	5	5
	86.	Helodidae *		12		5	5	5
	87.	Hydraenidae *	3	8	8			5
	88.	Hydrophilidae *	2	5	6	5	5	5
	89.	Hygrobiidae	1			5	5	5
	90.	Psephenidae *	6	10	8		5	5
	91.	Ptilodactylidae	10					5
Lepidoptera	92.	Pyralidae *	3	12	7			
Diptera	93.	Athericidae *	8	10	9			
	94.	Blepharoceridae	10	15	10			3
	95.	Ceratopogonidae *	4	5	2			
	96.	Chironomidae*	3	2	1	2	2	2
	97.	Culicidae *	1	1	2			
	98.	Ephydridae	2	3	1			3
	99.	Simuliidae*	5	5	7	5	5	5
	100.	Stratiomyidae *	2					3
	101.	Tabanidae*	3	5	7			
	102.	Tipulidae*	5	5	8	5	5	5

4.2.2 Quality indices based on abiotic water characterisation

4.2.2.1 Water quality index (WQI)

The water quality index (WQI) is a popular quality index based on the use of standard abiotic variables to obtain criteria for surface water classification (Stambuk-Giljanovic, 1999; Pesce and Wunderlin, 2000; Kannel et al., 2007; Sanchez et al., 2007). This index is a mathematical instrument used to transform large quantities of water characteristics into a single number, which represents the water quality level. The use of the WQI is simple, and allows easy classification of the water quality.

Determination of the WQI requires a normalisation step in which each parameter is transformed into a 0–100 scale, where 100 represents the maximum quality. A weighting factor is also applied for each variable in accordance with the importance of the variable as an indicator of water quality (Sanchez et al., 2007). In the present study, eight coupled state variables: DO (mg/l), pH, turbidity (NTU), water temperature (^oC), BOD_5 (mg/l), COD (mg/l), NO_2 -N (mg/l), PO_4-P(mg/l) were selected for index investigation. Variable selection was based on the data obtained during the present study and the available normalisations given in literature.

The subjective water quality index, WQI_{sub}, was calculated on the basis of the WQI as follows:

$$\mathrm{WQI_{sub}} = k\frac{\sum_i C_i x P_i}{\sum_i P_i}$$

in which *k* is a subjective constant. It represents the visual impression of river contamination (as could be evaluated by a person without training in environmental issues). It takes one of the following values according to the condition of the river:

- 1.00 for water without apparent contamination (clear or with natural suspended solids);
- 0.75 for light contaminated water (apparently), indicated by light non-natural colour, foam, light turbidity due to non-natural reasons;
- 0.50 for contaminated water (apparently), indicated by non-natural colour, light to moderate odour, high turbidity (non-natural), suspended organic solids, etc.
- 0.25 for highly contaminated water (apparently), indicated by blackish colour, strong odour, visible fermentation, etc.

C_i is the value assigned to each variable after normalisation (Table 4.4).

P_i is the relative weight assigned to each variable. P_i values range from 1 to 4, with 4 representing the variable that is most important for aquatic life preservation (e.g. dissolved oxygen), while a value of 1 indicates that such a parameter has a small impact (e.g. chloride) (Table 4.5). Relative weights are adapted from Kannel (2007). The subjective constant was evaluated from notes on the river appearance that were taken while sampling.

The objective water quality index (WQI_{obj}) was calculated using the same equation but with k =1 in all the cases to account only for variations due to measured variables.

Table 4.5. Variables considered for WQI calculation (after Kannel, 2007).

Parameter	**Relative weight (P_i)**	**Normalization factor (C_i)**										
		100	90	80	70	60	50	40	30	20	10	0
		Analytical value										
DO, ng/l	4	≤7.5	>7.0	>6.5	>6.0	>5.0	>4.0	>3.5	>3.0	>2.0	≥1.0	<1.0
pH	1	7	7-8	7-8.5	7-9	6.5-9	6-9.5	5-10	4-11	3-12	2-13	1-14
Turbidity, NTU	2	<5	<10	<15	<20	<25	<30	<40	<60	<80	≤100	>100
Water temp., ^oC	1	21/16	22/15	24/14	26/12	28/10	30/5	32/0	36/-2	40/-4	45/-6	>45;<-6
BOD_5, mg/l	3	<0.5	<2	<3	<4	<5	<6	<8	<10	<12	≤15	>15
COD, mg/l	3	<5	<10	<20	<30	<40	<50	<60	<80	<100	≤150	>150
NO_2^--N, mg/l	2	<0.005	<0.01	<0.03	<0.05	<0.1	<0.15	<0.2	<0.25	<0.5	≤1.0	>1.0
PO_4^{3-}-P, mg/l	1	<0.02	<0.05	<0.1	<0.2	<0.3	<0.5	<0.75	<1	<1.5	≤2	>2

Both WQIs were calculated for each sampling site. The list of input variables shows that the WQIs reflect mainly organic pollution. The WQI was therefore used to evaluate the ability of the biotic indices to detect organic pollution. The water quality classification system adopted here was proposed by Jonnalagadda and Mhere (2001). According to these authors, a WQI in the range of 0–25 is very bad, 26–50 is bad, 51–70 is medium, 71–90 is good and 91–100 is excellent.

4.2.2.2 Heavy metal contamination index

In order to assess the relation between biotic indices and inorganic pollution in the Du river, the contamination index related to heavy metals was applied (Bai et al., 2009). Assessment of water contamination by heavy metals was performed by the contamination index (P_i) and the integrated contamination index (P).

A contamination index (Pi) to describe the contamination of a given toxic substance in a basin is calculated using the following formula:

$$Pi = Ci/Cf$$

In which C*i* and C*f* are the mean content and the reference value for the substance. The Vietnamese standards for surface water were used as the reference baselines in this study.

The integrated contamination index (*P*) can be calculated according to the modified Nemerow index (Nemerow, 1985):

$$P = \sqrt{\max(P_i) \times \left(\frac{1}{n}\sum_{i=1}^{n} \bar{P}_i\right)}$$

For the description of the integrated contamination index, the same terminology was used, that is: $P\leq1$ no contamination; $1<P\leq2$ low contamination; $2<P\leq3$ moderate contamination; $P>3$ high contamination (Bai et al., 2009).

Advantages of biotic indices in water quality assessment are discussed based on a comparison with the assessment performance of the abiotic indices described above in this section.

4.2.3 Data analysis

For each site of the Du river basin, all biotic indices described in the previous section were calculated (SASS5 Score, BMWP, BMWP-Thai, BMWP-Viet, SIGNAL2, SASS-ASPT, HKHbios, ASPT-Viet and EPT) according to tolerance values presented in Table 4.3. Water quality indices and contamination index were also calculated where available.

Correlations between biological indices and chemical variables were computed using the non-parametric Spearman rank correlation coefficient (ρ). This correlation was also used to test correlations between all indices. All statistical analyses were performed using the STATISTICA package version 7 (Statsoft Inc, 2004).

4.3 Results

4.3.1 Correlations between biotic indices and environmental variables

Six out of 19 recorded environmental variables showed significant correlations with biotic indices. Among them, significant correlations were found between the BMWP group of indices and the EPT index, and the concentration of Fe in the water column, DO, BOD_5 and three physical variables: altitude, mean water width and P/R class (Table 4.6).

The group of biotic indices based on average scores showed relatively weaker correlations with environmental variables. Correlative relations of DO were encountered with ASPT-Viet and HKHbios.

Table 4.6. Spearman rank correlations between biotic indices and environmental variables for the Du river (n=104).

	P/R Class	**Altitude**	**Width**	**DO**	**BOD5**	**Total Fe**
SASS5 Score	-0.30**	0.31*	n/c	0.41**	0.43**	-0.46**
BMWP original	-0.32**	0.33**	n/c	0.45**	0.38**	-0.50**
BMWP Thai	-0.30**	0.38**	-0.22**	0.48**	0.35**	-0.55**
BMWP Viet	-0.39**	0.43**	-0.23*	0.51**	0.34**	-0.54**
SIGNAL 2	-0.24*	n/c	n/c	n/c	0.22**	n/c
SASS5 - ASPT	n/c	-0.23*	0.44	n/c	0.29**	n/c
HKHbios	n/c	n/c	0.26*	0.25*	n/c	n/c
ASPT Viet	n/c	n/c	n/c	0.33*	n/c	-0.27**
EPT	-0.43**	0.29**	n/c	0.48**	0.32**	-0.45**

*p<0.05; ** p< 0.001; n/c: no significant correlation.

The BMWP-Viet and EPT showed strong correlations with environmental variables for the Du river, while with the ASPT-Viet, significant correlations were only found with DO and the concentration of Fe.

4.3.2 Correlations between biotic and abiotic indices

Spearman rank correlations were also calculated among biotic indices and WQI (Table 4.7). Correlations between biotic indices and the heavy metal contamination index were also calculated for the sites where metal concentrations were measured (Table 4.8).

The indices calculated by summing tolerance scores were strongly correlated (ρ=0.90-0.97, p<0.001), which can be explained by the similar method applied for calculating these indices. The ASPT indices showed lower correlations (ρ=0.39-0.73), indicating dissimilarity of the scoring systems applied in different regions.

Correlations between ASPT and the total-score group of indices were still significant but lower (ρ=0.25-0.57). Among them, ASPT-Viet presented the strongest correlations with the BMWP Viet and BMWP original index.

The EPT provided significant correlations with all other indices applied in this study. The strongest correlation was obtained with the BMWP-Viet. The EPT also showed significant correlations with both subjective and objective WQIs and the heavy metal contamination index.

Table 4.7. Spearman rank correlations between the indices used for water quality assessment.

	SASS5 Scores	BMWP original	BMWP Thai	BMWP Viet	SIGNAL 2	SASS5 - ASPT	HKHbios	ASPT Viet	EPT	WQI sub	WQI obj
SASS5 Scores	1.00	0.90**	0.93**	0.92**	0.34**	0.42**	0.34**	0.40**	0.80**	0.46**	0.63**
BMWP original		1.00	0.94**	0.96**	0.38**	0.25*	0.39**	0.57**	0.84**	0.51**	0.66**
BMWP Thai			1.00	0.97**	0.33*	0.25*	0.33*	0.46**	0.83**	0.54**	0.68**
BMWP Viet				1.00	0.36**	0.23*	0.36**	0.52**	0.84**	0.58**	0.72**
SIGNAL 2					1.00	0.54**	0.63**	0.39**	0.57**	n/c	n/c
SASS5 - ASPT						1.00	0.73**	0.39**	0.35**	n/c	n/c
HKHbios							1.00	0.66**	0.51**	0.30*	n/c
ASPT Viet								1.00	0.56**	0.40**	0.31*
EPT									1.00**	0.45**	0.65**
WQI sub										1.00	0.72**
WQI obj											1.00

*$p<0.05$; **$p< 0.001$; n/c: no significant correlation.

Both subjective and objective water quality indices showed strong correlations with BMWPs and SASS5 Scores (ρ=0.46-0.72, $p<0.001$), while they showed no or weak correlations with the ASPT group of indices. The results revealed that BMWP scores perfectly reflected organic pollution in the Du river.

The heavy metal contamination index also showed significant correlations with the BMWP, EPT and SASS5 scores and no significant correlations with the ASPT group of scores.

Table 4.8. Spearman rank correlations between the heavy metal contamination index and the biotic indices (n=71).

Index	**Spearman ρ**	***p-level***	**Index**	**Spearman ρ**	***p-level***
SASS5 Score	-0.42	<0.001	SIGNAL2	n/c	0.32
BMWP original	-0.40	<0.001	SASS5	n/c	0.62
BMWP Thai	-0.46	<0.001	HKHbios	n/c	0.92
BMWP Viet	-0.43	<0.001	ASPT Viet	n/c	0.41
EPT	-0.38	<0.001			

The BMWP-Viet scoring system included most taxa found in the Du river (58 out of 70 taxa). The SIGNAL 2, SASS5, HKHbios and BMWP-Thai scoring system comprised scores for 54, 54, 55 and 49 taxa present at the studied sites, respectively. The original

BMWP scoring system from the UK only included 42 of the observed taxa. The lack of tolerance scores for many taxa in the original BMWP resulted in lower indicator values for the original BMWP compared to the values for the BMWP-Viet.

The BMWP scoring system for the UK, Thailand and Vietnam are generally similar. The major modifications were the addition of scores given to regionally occurring taxa. The fact that the original BMWP and the BMWP-Thai resulted in lower index values was mainly due to the lack of tolerance scores for the occurring taxa.

The HKHbios and the BMWP-Viet scoring system have 77 taxa in common, 49 of which were assigned with higher scores in the HKHbios than in the BMWP-Viet scoring system (Table 4.3). The HKHbios index therefore resulted in higher values than the ASPT-Viet.

The SIGNAL2 system provides scores for many taxa present in the Du river (54 taxa). However, the SIGNAL2 resulted in the lowest values because the tolerance scores for taxa in SIGNAL2 are generally lower than these in the BMWP-Viet scoring system.

4.3.3 Site classification based on biotic indices

According to the quality class attribution system (Tabl 4.3), the environmental quality of the study sites assessed according to the BMWP-Viet varied from very good to very poor (Table 4.9). The HKHbios, the original BMWP and the BMWP-Thai classified sites from good to very poor quality, while the SASS5 and the SIGNAL2 from moderate to very poor. The SASS5-ASPT classified all sites as poor and very poor. The ASPT-Viet failed in site classification. All sites were moderate, except sites 10 and 11, which were classified as very poor.

Although the WQI showed strong correlations with the BMWP systems, it failed in properly classifying the watercourses in the study area. The objective WQI, which mostly relies on organic pollution, classified all sites as good or very good. The subjective WQI, on the other hand, provided a classification close to the BMWP-Viet, however, it rather reflected a subjective judgement on the environmental condition of the sites.

The heavy metal contamination index classified sites 10 and 11 as highly contaminated. Although average values of the heavy metal contamination index showed no contamination in other sites, metal pollution was detected in sites 4, 12 and 13 during several monitoring campaigns (A6, S7, Sp8), where poor ecological conditions were revealed by low values of BMWP-Viet for these samples.

Table 4.9. Water quality class of the sites in the Du river based on the biotic indices. Sites were classified as very good (VG), good (G), moderate (M), poor (P) and very poor (VP).

Sites	SASS5 Score	SASS5 ASPT	HKHbios	SIGNAL2	BMWP UK	BMWP Thai	BMWP Viet	ASPT Viet
1	M	VP	G	M	G	G	VG	M
2	P	VP	M	P	M	G	M	M
3	P	P	G	M	M	G	G	M
4	P	P	G	M	M	M	M	M
5	M	VP	G	P	G	G	G	M
6	M	VP	P	M	M	G	G	M
7	P	VP	P	P	M	G	M	M
8	M	P	G	M	M	M	G	M
9	M	VP	M	P	M	M	M	M
10	VP	VP	VP	VP	VP	VP	VP	VP
11	VP	VP	VP	VP	VP	VP	VP	VP
12	P	P	G	M	M	M	M	M
13	P	P	G	P	M	P	M	M
14	M	P	G	M	G	M	G	M
15	P	VP	M	P	M	P	M	M

The study sites were classified on the basis of the mean values of the BMWP-Viet index (Fig. 4.2). Extremely low values of the BMWP-Viet in sites 12 and 13 during the campaign August 2006 (minimal values of BMWP-Viet) were due to a severe spate, which occurred about two weeks before the sampling campaign. Dominant macroinvertebrate taxa, which were usually present at these sites, were flushed away during the spate and the communities did not yet recover. The variability of the classification of the sites in the Du river basin is illustrated in Fig. 4.3 according to the assessments based on the BMWP-Viet.

The results of the site classification showed that the BMWP-Viet reflected both organic and inorganic pollution and could differentiate disturbance intensities among the sampling sites along the Du river. The BMWP-Viet values indicated pristine conditions in the head water sites and also detected different pollution levels in the downstream sites.

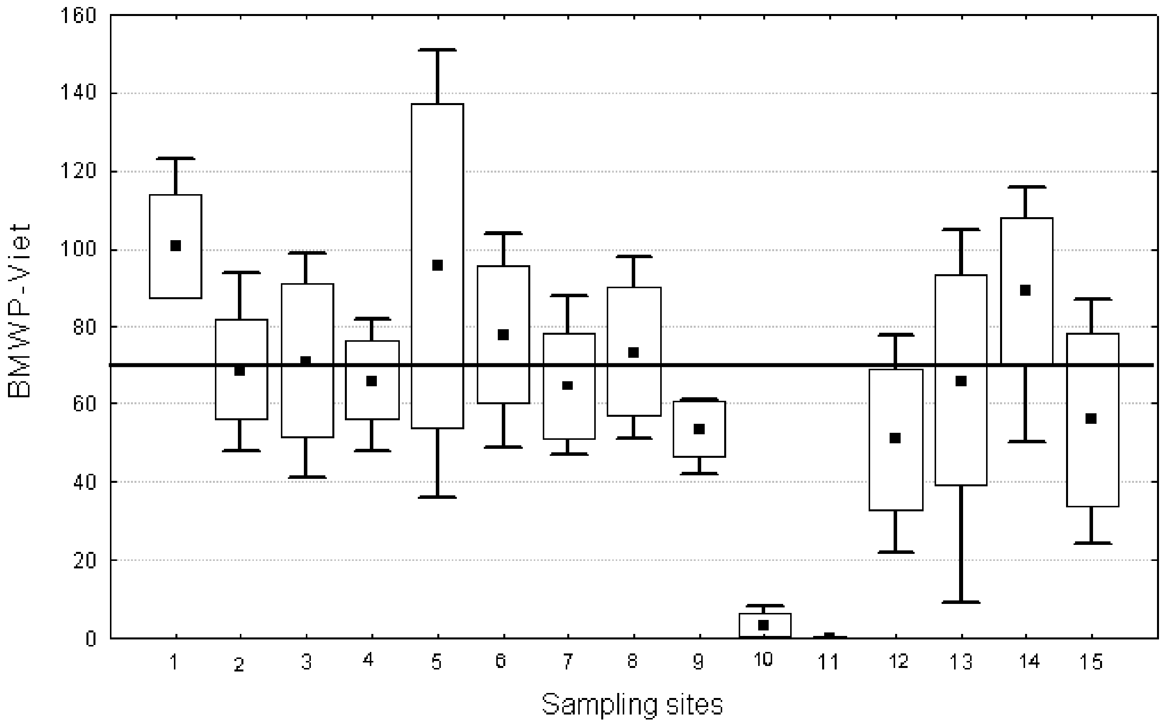


Figure 4.2. Box (95% confidence interval) and whisker (Max-Min values) plot of the BMWP-Viet values for the Du river. The horizontal line indicates the threshold between good and moderate water conditions.

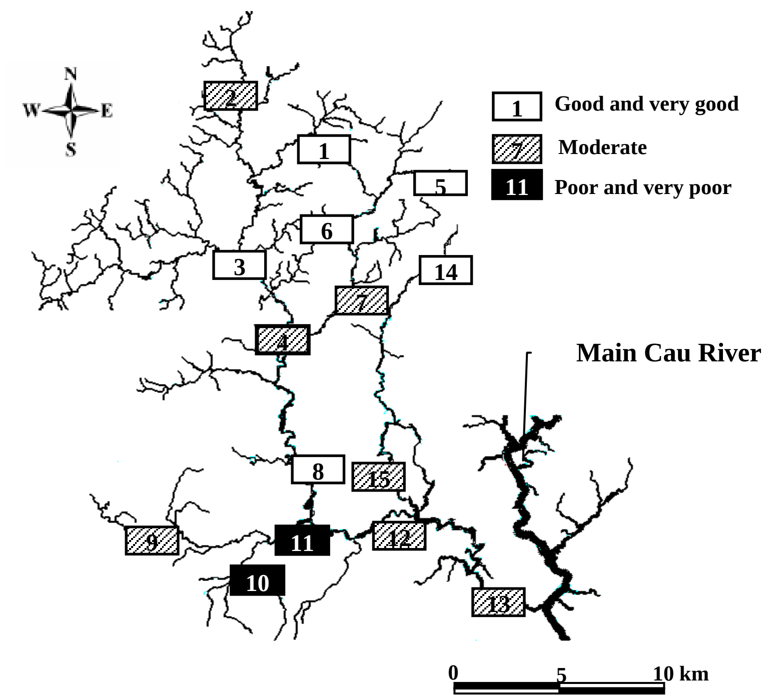


Figure 4.3. Site classifications according to the water quality assessed by the BMWP-Viet in the Du river basin.

4.4 Discussion

4.4.1 Relevance of using biotic indices for rapid assessment

The BMWP system proved to be an appropriate tool for rapid bioassessment, especially in developing countries thanks to the simplicity of the assessment procedure, the index calculation and the interpretation.

One of the most important criteria by mean of which performance should be judged is the accuracy of a biotic index in indicating impact (Resh, 1994). Biotic indices should show a decline in response to anthropogenic stressors. Such stressors are usually expected to eliminate sensitive taxa or reduce their probability of collection and should therefore cause declines in the total number of families, the number of sensitive families and the biotic index. Considering this criterion, ASPT failed in indicating differences in water quality in the study area. The most important reason is probably that over 80% of the taxa scored from 3 to 6 in the BMWP-Viet system. Most samples therefore received values of the ASPT-Viet index from 3.5 to 5 and as a consequence, average ASPT-Viet scores for all sites (except 10 and 11) fell into the same range.

On the other hand, the BMWP considers the number of taxa present at the studied sites and thus performed better in reflecting the presence of pollution along the Du river. Highly significant correlations were found between indices and environmental variables such as BOD_5, DO and total Fe (Table 4.4), which proved the sensitivity of these indices to stress. The BMWP was also significantly correlated with the WQI, which reflects organic pollution. A strong correlation between the BMWP and the heavy metal contamination index proved that the index was also successful in detecting inorganic pollutants, which play an important role in the Du river.

The EPT index assesses the water quality based on the presence of Ephemeroptera, Plecoptera and Trichoptera, which are groups sensitive to organic pollution. In this study, the highest number of EPT taxa was found at sites 1 and 5, headwaters of the Du river located in remote sites with the lowest population density in the study area. The EPT index showed a decrease of sensitive taxa along the river. The high correlation of the EPT index with environmental variables indicated the potential of using this index as a bioindicator in the Du river.

The BMWP-Viet was better fit to classify the 15 sites into different levels of ecological health (from very good to very poor) compared to other biotic indices of the BMWP group.

The present study showed the relevance of the BMWP group of indices in the sub-tropical rivers in northern Vietnam, especially as rapid assessment method. The relationship between the BMWP-Viet values and the environmental variables is investigated in more detail in Chapter 5.

4.4.2 Proposed modifications of the BMWP-Viet scoring system

The present BMWP-Viet scoring system offered some advantages over the other evaluated indices as far as accuracy was concerned. The study proved that approaches used in developed as well as developing countries are appropriate. However, pollution-tolerance values are often region specific and need to be adapted (Resh, 2007). By including more families that occur in Vietnam and adapting tolerance scores to the geographical area, the system could still be improved. Tolerance scoring systems developed in similar ecological regions including HKHbios, SIGNAL2 and SASS5 can be consulted as references for the adaptations including (1) addition of scores for taxa, which are useful water quality indicators, (2) modification of scores of taxa, which evidently behave differently compared to other regions. Some suggestions are illustrated below.

(1) Twelve families were present in the study area but not included in the preliminary BMWP-Viet score. Several are useful water quality indicators and should be given scores in the scoring system

- Three mollusc families, Pisidiidae, Pachychilidae and Stenothyridae and four Diptera families Tabanidae, Ceratopogonidae, Culicidae, Athericidae were present in more than 65% of the sites, ranging from pristine upstream to impacted downstream sites.
- The Trichoptera family Ecnomidae and the Ephemeroptera family Prosopistomatidae are important indicator taxa belonging to the EPT group (Dean, 1996). Different scores were given to these taxa in the SASS5, SIGNAL2 and HKHbios scoring system.
- Isopods are important bioindicators in temperate areas, however, the only isopod family found in the study area (Corallanidae) has not been scored in the tolerance scoring systems.
- The Lepidoptera family Pyralidae and Hydracarina were found during different campaigns at all types of sites except 10 and 11. Scores for these taxa were

assigned in other scoring systems such as SIGNAL2, which revealed their importance as bioindicators to environmental stress.

(2) Scores given to some taxa showed inappropriate judgement during the application and need to be modified.

- The Hemiptera family Aphelocheiridae that receives the highest score of 10 occurred in various sites, mostly in impacted sites 8, 12 and 13.
- Odonata are conspicuous elements of the zoobenthos in many tropical Asian streams (Dudgeon, 2000). Gomphidae are the most common dragonflies in study sites with a presence in 41% of the sampling sites and have been found in all sites except 10 and 11. They show even more tolerance to stress than other taxa in this order including Corduliidae and Libellulidae. Although the Gomphidae taxon was already downgraded in the BMWP-Thai, further downgrading should be considered, by recommending scores given for this taxon in the SASS5 and SIGNAL2 scoring system.
- The Ephemeroptera family Caenidae is one of the most common taxa in the study area. It was also abundantly present in all types of sites. The tolerance score of this taxon had also been downgraded in many adapted BMWP systems such as the IBMWP (Alba-Tercedor and Pujante, 1997), BMWP-CR (Astorga et al., 1997) and also improvements of the BMWP index as done by Walley and Hawkes (1996 and 1997. The sensitivity of this taxon was also lower in the SIGNAL2 and SASS5 scoring systems.
- In the present study, different genera of the Ephemeroptera family Leptophlebiidae showed a different response to environmental conditions. While *Thraulus* showed a high sensitivity and occurred only in pristine sites, other genera were found in sites classified as moderately polluted. In the HKHbios scoring list, *Thraulus* is given a score of 10 while other genera only a score of 7 for their sensitivity towards pollution. In this study, only the family level identification has been considered for assessment. Downgrading of the family Leptophlebiidae is done in SASS5, SIGNAL2 and BMWP systems adapted to tropical countries like Columbia and Costa Rica. Downgrading should be considered for this family and further study on lower levels of identification is recommended for a better interpretation of this important indicator taxon.

- The Trichoptera family Hydropsychidae and the Megaloptera family Corydalidae preferred fast flowing upstream sites and occurred in good and very good sites during all campaigns. These taxa also scored higher in all reference scoring systems and upgrading should be considered for modifying the score of this taxon.
- All Hemiptera families (except Aphelocheiridae) and Coleoptera families were scored 5 in the BMWP-Viet scoring system, while they showed different habitat preferences during this study. The SIGNAL scoring system also provides a wider range of scores (from 1 to 7) for families of these orders.

With the limited database of this study, reasonable judgements can not be made for all taxa and thus no modification was proposed. However, further studies are recommended to develop more reliable scores for all taxa.

TWINSPAN (Two-way indicator species analysis) is usually used for identifying stream groups based on macroinvertebrate community structure. However, this process can not only classify sites, but also taxa simultaneously, based on dividing reciprocal averaging ordination space (Hill et al., 1975). Taxa classification based on characteristics of inhabited sites can provide useful information for developing tolerance scores for indicator taxa. When appropriate databases are available, this type of classification can be useful to detect the taxa, for which scores require adaptation.

4.4.3 Application of bioassessment of river in Vietnam

4.4.3.1 Monitoring programme

Biomonitoring is especially appropriate for use in developing countries because neither sophisticated nor expensive equipment is required (Resh, 2007). This study revealed that there was no significant change in the diversity between seasons in the Du river, which was also found by many studies on temporal patterns of the macroinvertebrate diversity in tropical as well as neotropical areas (Jacobsen and Encalada, 1998; Baptista et al., 2001; Melo and Froehlich, 2001; Ahmad et al., 2002; Hoang and Bae, 2006). However, in upstream sites, especially at high altitude (site 5), streams can dry up during the dry season, causing loss of habitat for macroinvertebrates, leading to a loss of taxa and a decrease of BMWP values at these sites. Flooding during the rainy seasons is also an important reason for the loss of taxa downstream (sites 12, 13) because spates can flush away macroinvertebrates as well as their habitat. These events should be taken into account during the design and implementation of a monitoring programme.

Samples should not be taken during the driest and coldest time of the year in upstream sites. At this time, many upstream sites can dried up or become standing water, not allowing a proper colonisation which characterises the environmental condition of the watercourse. Samples should also not be taken in the middle of the flooding season, especially in downstream sites, to avoid extreme effects of flow on habitat conditions of the macroinvertebrates. However, at this time, reference conditions can be present at upstream sites, because these become suitable for macroinvertebrates. Monitoring thus needs to be carried out during these periods to collect information about reference conditions. Samples should be taken at least 4 weeks after spates, which is an appropriate time for community recovery.

The most crucial aspect for bioassessment is sampling. The method of sampling and the used equipment are strongly influenced by the size of the rivers being sampled but often consist of two parts: a main sample taken with a net and a manual search. Kick-sampling of the river bed, including vegetation should be carried out where possible and a dredge or artificial substrates can be used in large rivers. A manual search of habitats that are not likely to be sampled efficiently by the used method, such as under rocks or submerged wood is essential to collect all representative taxa at the sites. This is also the reason why only qualitative data are used for further data analysis, since quantitative data (abundances) can not be correctly collected with this sampling methodology.

4.4.3.2 Index interpretation

The ranges of the BMWP values that are used to assign quality classes should also be region specific. For the Iberian Peninsula, the IBMWP is an adaptation of the British BMWP scoring system. The modifications included not only the addition of extra families and changing some scores, but also a different division of the scores into five classes, representing varying degrees of pollution (Alba-Tercedor and Sanchez-Ortega, 1988). In Southeast Asia, although the BMWP approach was adopted, specific index interpretations were not considered. There should be further studies on the application of the BMWP approach in different watercourses in order to develop appropriate class boundaries for index interpretation in Vietnam.

An important issue, which needs to be considered, is that the BMWP was originally developed for the purpose of comparing ecological conditions of the same site at different times (Hawkes, 1982). Essential factors determining macroinvertebrate community structure are often physical habitat conditions, which do not reflect imposed stresses such as sewage pollution, toxic wastes and pesticides (Hellawell, 1986). Therefore, it often difficult to compare the biological quality between different watercourses based on a single BMWP index. Classification of the biological quality based on the difference between observed and expected biotic indices at sites with reference conditions can provide a more reliable assessment (Dines and Murray-Bligh, 2000). This is the concept of RIVPACS and AUSRIVAS, which are used in the UK and Australia. An equivalent system could also be developed for Vietnam. Moreover, developing criteria to differentiate sites can reduce the subjectivity in assessing the performance of the BMWP system (Mustow, 2002).

4.4.3.3 Lower level of identification

Essential for the successful application of assessment methods based on macroinvertebrates is the availability of suited keys of identification up to the required level and this in the local language (De Pauw and Hawkes, 1993; De Pauw et al., 2006). The identification level is often the result of a practical trade-off between taxonomic precision and time constraints and financial resources (Guerold, 2000; Gabriels et al., 2005). Various authors recommend identification to a lower than family level to ascertain a detailed insight in the community composition, avoiding information loss due to lumping of taxa and showing a strong assemblage–environment relationship (e.g. Resh and McElravy, 1993; Stubauer and Moog, 2000; Lenat and Resh, 2001; Adriaenssens et al., 2004c). Although species identification is time-consuming and expensive, lower levels of identification, for example to genera level, can be practically possible and provide opportunities for improving robustness of biotic indices such as the BBI and multimetric indices (De Pauw and Vanhooren, 1983; De Pauw et al., 2006; Gabriels et al., 2006), .

The BMWP approach was chosen in Vietnam for rapid bioassessment mainly because it requires only qualitative sampling and family level identification. However, when better identification keys would become available, other methods based on lower level identifications could be applied, which could increase the reliability and robustness of the biotic index.

4.4.3.4 Different bioassessment approaches

Other biological monitoring techniques recently developed, which are most frequently used in the temperate regions, are more informative such as multimetric assessment systems (Gabriels et al., 2006) and may also be useful in the Southeast Asia. The multimetric approach for bioassessment provides a means of integrating ecological information relating elements and processes of naturally functioning aquatic assemblages and current knowledge about those relationships. A multimetric index describes the state of an ecosystem by means of several individual variables (metrics). These metrics each represent a different component of the ecosystem quality and are combined into one index value. The most efficient metrics are combined on the basis of their capacity to respond to multiple environmental pressures (Hering et al., 2006). Multimetric indices were developed or are under development in many recent studies in temperate and tropical areas (Thorne and Willams, 1997; Davis et al., 2003; Hering, 2004). The flexibility of multimetric indices is that they can easily be adjusted by adding or removing metrics or by fine-tuning the metric scoring criteria. It can also be used as a calibration approach with the BMWP index (Munne and Prat, 2009). This advantage makes the multimetric approach potentially suitable for application in developing countries including Vietnam.

4.5 Conclusions

Bioassessment based on the macroinvertebrate communities is a suitable and promising method for the assessment of running water in the Du river. Different levels of river water quality could be detected, from very good to very poor conditions by mean of the BMWP-Viet values, which reflected gradient of human impacts on the river ecology. The BMWP-Viet index proved to be appropriate for use in watercourses in Vietnam as well as in other countries in the region. The BMWP-Viet could be applied to rapidly detect sites suffering from pollution. This would certainly be helpful to the authorities responsible for water quality assessment. However, further research is needed to improve the BMWP-Viet scoring system by inclusion of new taxa as well as adjustments of scores for some other taxa. The scoring systems developed for other geographical areas could serve as a good reference for further improvements. In addition, adjustments in their threshold values for the determination of quality classes are required. It is also recommended that authorities responsible for water quality monitoring in Vietnam utilise this technique routinely. Training should be provided on sample collection, processing and macroinvertebrate identification. The current BMWP-Viet approach can be useful at an early stage of bioassessment application in Vietnam. However, this method should be improved by the development of other assessment approaches such as multimetric indices.

Chapter 5

Application of habitat suitability models for the prediction of macroinvertebrates in the Du river

5.1 Introduction

Different modelling techniques have been applied to freshwater streams, mainly focussing on the environmental response of biological communities to specific disturbances. The distribution of taxa as a function of the abiotic environment, named habitat suitability modelling, has been recognised as a significant component of conservation planning (Austin, 1998, 2002; Guisan and Zimmerman, 2000). Artificial intelligence has played a crucial role in studying these relationships (Recknagel, 2002; Marshall et al., 2003), including classifications trees and support vector machines.

Classification trees (CTs) (Breiman et al. 1984), often referred to as decision trees (Quinlan 1986), predict the value of a discrete dependent variable with a finite set of values (called class) from the values of a set of independent variables (called attributes), which may be either continuous or discrete. Data describing a real system, represented in the form of a table, can be used to learn or automatically construct a decision tree. Due to their transparency and flexibility, the technique has been applied for studying impacts of different anthropogenic disturbances on the distribution of macroinvertebrates (D'heygere et al., 2003; D'heygere et al., 2006; Dakou et al., 2007; Goethals et al., 2007).

Support-vector machines (SVMs) (Vapnik, 1995) belong to a generation of inductive modelling techniques that has been inspired by aspects of biological information processing. Multi-class problems are solved using pairwise classification with this technique (Witten and Frank, 2000). They have the ability to extract temporal or spatial patterns and to analyse highly non-linear and complex data. They produce very competitive results with the best accessible classification methods and they need only the smallest amount of model tuning (Decoste and Scholkopf, 2002; Guo et al., 2005). They also yielded excellent generalisation performance on a wide range of problems and have been successfully employed to solve many nonlinear regression and time serie problems (Shan et al., 2006; Sanchez-Hernandez et al., 2007; Ribeiro and Torgo, 2008).

The aim of this chapter is to develop models predicting the habitat suitability of macroinvertebrates in the Du river in Vietnam using CTs and SVMs and to compare the performance of both modelling techniques. The models developed allow selection of important variables for river restoration and thus may support the implementation of integrated river management in Vietnam. The following results are presented in this chapter:

- Development of habitat suitability models based on CTs and SVMs to predict the BMWP-Viet index in the Du river basin;

- Development of habitat suitability models based on CTs and SVMs to predict the selected indicator taxa in the Du river basin;
- Comparative discussion of the obtained results.

5.2 Material and methods

5.2.1 Database setup

Input variables were selected from environmental variables monitored in the Du river basin. The correlation (*r*) between each pair of variables was calculated to identify 'noise' variables (Appendix 1). If two variables are strongly correlated, then one of these two variables may be removed without adversely affecting the model performance. According to Walczak and Cerpa (1999), any *r* with an absolute value of 0.20 or higher indicates a probable noise source to the performance of models and they advice to consider the removal of one of these variables.

However, there might be practical reasons to leave these correlated variables in, such as ecologically irrelevant correlations. Variables should be included if the effect of these variables on the river biology is interesting for management purposes. The removal of input variables therefore should be overruled by use of ecological expert knowledge to avoid losing important information.

To build a reliable model, a verification method has been applied to exclude or correct data points that do not correspond to the possible real values, known by general knowledge (Witten and Frank, 2000). However, the elimination of the data can lead to decrease of the reliability of the model and the prediction of the water quality (Goethals, 2005). Analysis of the data related to the minima, maxima, averages and standard deviations is an important aspect to discover possible outliers in the dataset. Analysis of the Du river dataset showed several outliers in the variables: turbidity, dissolved oxygen, COD and all the metal concentrations. However, observations indicated that the concerned sites were highly contaminated. Therefore, despite the presence of missing data and high standard deviations, no data (instances) were removed. All 21 physical-chemical and structural measurements (Table 3.1) were retained as inputs for the model development. Most variables were continuous, whereas the P/R class, substrates and sampling seasons were categorical variables.

Extreme values can wrongly increase the bandwidth, causing a decrease in accuracy of the model. To deal with this problem, data can undergo a log-transformation or can be normalised but a trade-off between bandwidth and accuracy is always required (Goethals, 2005). However, the outcomes of classification tools such as CTs and SVMs are generally not influenced by wide ranges of data values. Therefore, no transformation of the data had been applied.

In the models predicting the ecological quality of the river sites, the biotic index BMWP-Viet was the dependent variable, while environmental variables were the independent ones. The BMWP-Viet values were divided into five classes, each representing the ecological quality of the river water. The classification of BMWP-Viet is given in chapter 4.

The applied habitat suitability models focused on presence or absence of macroinvertebrate taxa in the Du river. When presence/absence of macroinvertebrates is predicted, the percentage of Correctly Classified Instances (CCIs) is mostly applied to assess model performance (Goethals, 2005). There is however clear evidence that the CCI is affected by the frequency of occurrence of the taxa being modelled (Fielding and Bell, 1997; Manel et al., 1999). Therefore, rare taxa were excluded from the model development and only the 30 most abundant taxa were used as output variables of the models.

A visual relation analysis between the input and output variables could have several advantages regarding getting insight into outliers, data clusters and missing or scarce variable combinations in certain ranges (Goethals, 2005). Therefore, these methods can be very important to realize the difficulty of the development of well performing models or finding out why some models do perform well and why some data can be classified as outliers (Witten and Frank, 2000). These analyses have gained a lot of popularity during the last years and became standard tools in most data mining and analysis software packages. These analyses were performed in the WEKA toolbox (Witten and Frank, 2000). Visualisation graphs of the variables applied in predictive models of BMWP-Viet and 10 taxa represented different macroinvertebrate groups are illustrated in an Appendix 3.

5.2.2 Classification tree construction

The 'Top-Down Induction of Decision Trees' (also called C4.5 algorithm) is one of the most widely used classification tree induction methods. Tree construction proceeds recursively starting with the entire set of training examples. At each step, the most informative attribute is selected as the root of the tree and the current training set is split into subsets according to the values of the selected attributes. For discrete attributes, a branch of the tree is typically created for each possible attribute value. For continuous ones, a threshold is selected and two branches are created based on that threshold. For the subsets of training examples in each branch, the tree construction algorithm is called recursively. Tree construction stops when all examples in a node are of the same class or if some other stopping criterion is satisfied. Such nodes are called leaves and are labelled with the corresponding values of the class (Quinlan, 1993).

CTs were constructed on the basis of the WEKA software, version 3.4.11 (Witten and Frank, 2000). WEKA is a collection of machine learning algorithms for data mining tasks. The J48 algorithm is a Java re-implementation of C4.5 and is a part of the machine learning package WEKA. In the current study, the J48 algorithm (*weka.classifiers.trees.J48*) with binary splits was applied to induce classification trees. Standard settings were used except for pruning confidence factors (PCF), which had an important effect on the selected variables that were used for the classification (Dakou et al., 2007). Tree pruning was applied to optimise model performance by changing confidence factor values. Models with different intensities of pruning were induced by changing the confidence factor into 0.01, 0.1, 0.25 and 0.5 (default=0.25). The model training and validation was based on ten-fold cross validation (Kohavi, 1995; Witten and Frank, 2005; Dakou et al., 2007). The model with the best performing PCF was run five times after randomisation to check its robustness and reproducibility. The average predictive performance was calculated for further analysis. The following settings were applied to develop the classification trees:

- ‘binary splits’ (whether to use binary splits on nominal attributes when building the trees): false;
- ‘confidence factor’ (the confidence factor is used for pruning; smaller values incur more pruning): 0.5, 0.25, 0.1 and 0.01;
- ‘minimum number of instances per leaf’: 2;
- ‘number of folds’ (determines the amount of data used for reduced-error pruning: one fold is used for pruning, the rest for inducing the tree): 3;
- ‘reduced error pruning’ (whether reduced-error pruning is used instead of C.4.5 pruning): false;
- ‘seed’ (the seed used for randomizing the data when reduced-error pruning is used): 1;
- ‘subtree raising’ (whether to consider the subtree raising operation when pruning): true;
- ‘unpruned’ (whether pruning is performed): false;
- ‘use Laplace’ (whether counts at leaves are smoothed based on Laplace): false.

5.2.3 Support vector machines

SVMs implement Platt's sequential minimal optimisation algorithm for training a support-vector classifier (Keerthi et al., 2001). This implementation replaces all missing values and transforms nominal attributes into binary ones. Multi-class problems are solved using pairwise classification (Witten and Frank, 2000). They consist of input and output layers connected with weight vectors. The arrangements of neurons in the input layer control the input vectors while the output layer comprises a two-dimensional network of neurons arranged on a hexagonal net. SVMs have been designed for two-class problems including positive and negative objects (Guo et al., 2005). SVM can be used for classification and regression. A complete analysis of SVMs entails three steps: model selection, fitting and validation.

The basic idea of SVMs is to transform the data in such a way they become more or even completely linearly separable and subsequently to perform a linear separation. The maximum margin hyperplane is the one that gives the greatest separation between the classes (Fig. 5.1).The instances that are closest to the maximum margin hyperplane are called support-vectors. There is always at least one support-vector for each class and often there are more. The set of support-vectors uniquely defines the maximum margin hyperplane for the learning problem.

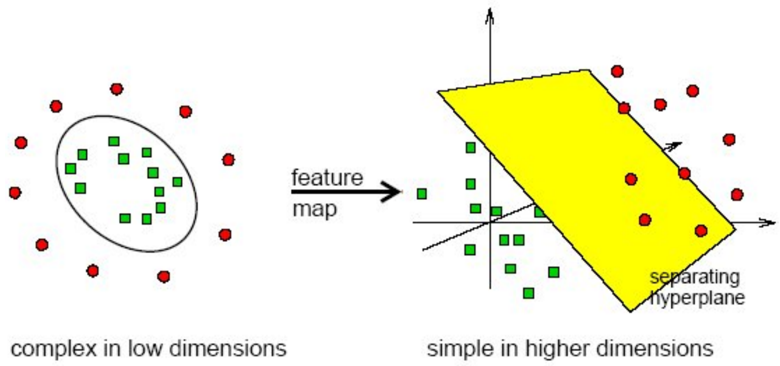


Figure 5.1. Hyperplane to make separation easier in higher dimensions of the data points (Van Looy et al., 2007).

A good separation is achieved by the hyperplane that has the largest distance to the neighbouring datapoints of both classes, since in general, the larger the margin is, the better the generalisation error of the classifier is. The instances that are closest to the maximum margin hyperplane are called support-vectors (Keerthi et al., 2001; Smola, 2004; Milenova et al., 2005).

The Platt's sequential minimal optimisation (SMO) (*weka.classifiers.functions.SMO*) algorithm, which is also a part of the machine learning package WEKA, was applied for the model development. Four standard kernels are usually used in classification problems and also used in regression cases: 1) linear kernel; 2) polynomial kernel; 3) radial basis function (RBF) network and 4) sigmoid kernel. A SVM with a RBF network and one with a sigmoid kernel implement another type of neural network, a multilayer perceptron with one hidden layer. The last three kernel functions are used to implement different non-linear mappings (Table 5.1).

Table 5.1. Kernel equations for non linear mappings (Cortes and Vapnik, 1995).

Kernel Function	Type of Kernel function
$K(x,y) = \exp(-\frac{\|\|x-y\|\|^2}{2\sigma^2})$	RBF
$K(x,y) = (x.y+1)^{degree}$	Polynomial
$K(x,y) = \tanh(x.y-\theta)$	Sigmoid

A difficulty in the application was stated as the choice of an ideal Kernel function (Cortes and Vapnik, 1995; Burges 1998; Witten and Frank, 2000; Meyer et al., 2003). However, the tested run showed that RBF and normalised Kernel did not present a good prediction compared to polynomial Kernel, which was therefore selected for the model development. In comparison with other methods, such as decision tree learners, even the fastest training algorithms for SVMs are slow when applied in non-linear settings.

The parameter settings in the WEKA toolbox were default values except for the exponents of the polynomial Kernel. To obtain the optimal results regarding predictive performances, the SVMs were optimised based on the application of different exponents from one to five. The model with the best performing exponent was run five times after randomisation to check robustness and reproducibility.

The John Platt's SMO algorithm was developed for training a support vector classifier by the WEKA toolbox with following settings (Witten and Frank, 2000):

- ‘build logistic models’ (whether to fit logistic models to the outputs): false;
- ‘c’ (the complexity parameter C): 1.0;
- ‘cache size’ (the size of the kernel cache; this should be a prime number. zero is used for full cache): 250007;
- ‘debug’ (if set to true, the classifier may output additional information to the console): false;
- ‘epsilon’ (the epsilon for round-off error. This should not be changed): 1.0E-12;
- ‘exponent’ (the exponent for the polynomial kernel): from 1 to 5
- ‘feature space normalisation’ (whether feature-space normalisation is performed (only available for non-linear polynomial kernels): false;
- ‘filter type’ (determines how/if the data will be transformed): normalise training data;

- ‘gamma’ (the value of the gamma parameter for RBF kernels): 0.01;
- ‘lower order terms’ (whether lower order polynomials are also used; this is only available for non-linear polynomial kernels): false;
- ‘number folds’ (the number of folds for cross-validation used to generate training data for ‘logistic models’ (-1 means use training data): 3.
- ‘random seed’ (random number seed for the cross-validation): 1;
- ‘tolerance parameter’ (the tolerance parameter; this should not be changed): 0.0010;
- ‘use RBF’ (whether to use a radial basis function (RBF) kernel instead of a polynomial one): false.

5.2.4 Genetic algorithms as an optimisation technique

Genetic algorithms (GAs) are general search algorithms inspired by Charles Darwin’s principle ‘survival of the fittest’ to solve complex optimisation problems (Holland, 1975; Goldberg, 1989; Vose, 1999). A popular approach is the use of GAs as an optimisation tool for resetting the parameters in other classifiers such as SVMs and CTs, especially in the field of river ecology (D’heygere et al., 2006; Goethals et al., 2007).

Using a very large dataset, it might be impractical to construct a decision tree using all of the points. As an alternative, subsets of the original dataset can be extracted, a tree can be constructed for each subset and then parts of individual trees can be combined in a smart way to produce an improved final set of feasible trees or a final tree. Trees can be generated by a decision tree package using GAs to allow them to crossover and mutate for a number of generations. This approach produced uniformly high-quality classification trees and the impact of scaling could be used effectively on very large datasets (Fu et al., 2003).

Through fusions of GAs and SVMs, the ‘optimal detection model’ for a SVM classifier can be determined (Kim et al., 2005). To cope with the problem of highly discriminative features and the set of parameters for the Kernel functions, genetic algorithms were adopted to allow the selection of the optimal set of features and set of Kernel parameters. In this method, the decision models generated by GAs are evaluated by SVMs to obtain the optimal decision model which reduces the modelling time and improves the classification rate of SVMs (Ohn et al., 2004).

The 21 input variables were presented to the GAs as a chromosome to determine the important variables for model development. In a next phase, decision trees and support vector machines were constructed with variables, which were selected as driving variables for most abundant taxa, in order to improve the comprehensiveness of the CTs and performance of the SVMs (D'heygere et al., 2003).

GAs were applied by selection of the attributes for CTs using GeneticSearch (*weka.attributeSelection.GeneticSearch*) as a search method and J48 as a classifier. The attribute evaluator was based on the Wrapper Subset Evaluator function (*weka.attributeSelection.WrapperSubsetEval.J48*) with the full training set. The Classifier settings were similar to those of the CTs. The search was also run five times after randomisation to find variables mostly selected in each run.

The same procedure was applied for the selection of the attributes for SVMs using GeneticSearch. The attribute evaluator was based on the Wrapper Subset Evaluator function (*weka.attributeSelection.WrapperSubsetEval.SMO*) with the full training set. The default settings were developed for the GAs (Witten and Frank, 2000). For the classifier, the settings were similar to those of the SVMs.

After the new variables were selected by GAs, model evaluation, size of the trees and model performances were compared for CTs and SVMs with and without GAs.

5.2.5 Performance criteria

There are a number of criteria for evaluation of model performance including the percentage of correctly classified instances (%CCI), area under the curve (AUC), sensitivity (Sn), specificity (Sp), normalised mutual information statistic (NMI) and odds ratio (Mouton, 2008). When presence/absence of macroinvertebrates is predicted, the percentage of CCI is usually applied to assess model performance (Goethals, 2005). However, CCI is affected by the frequency of occurrence of the taxon being modelled (Fielding and Bell, 1997; Manel et al., 2001). The Cohen's Kappa statistic (K) is therefore a more reliable performance measure compared to other evaluation parameters (Mouton, 2008). K is a derived statistic that measures the proportion of all possible cases of presence or absence that are predicted correctly by a model after accounting for chance predictions. In this study, K and %CCI were used to evaluate the performance of each method as also suggested in other studies (Witten and Frank, 2005; D'heygere et al., 2006; Dakou et al., 2007; Goethals et al., 2007). These studies also proposed that models with a CCI higher than 70% and K higher than 0.4 can be considered reliable.

Further, a visualisation of the predicted and observed classes of BMWP-Viet was presented by plotting the confusion matrix, to analyse the distribution of the incorrectly classified instances.

5. 3 Results

5.3.1 Ecological models predicting BMWP-Viet

Successively, results are given for the application of CT and SVM eventually in combination with GA for predicting BMWP-Viet.

Classification trees were the first habitat suitability models studied. Table 5.2 illustrates the effect of the pruning algorithm on the tree size and the relation with the tree performance based on %CCI and K. The CT developed for the prediction of BMWP-Viet class gave the best predictive performance for PCF = 0.5, with the highest CCI (57.9 %) and Kappa (0.40). However, this high PCF value resulted in a more complex tree compared to lower pruning factors of 0.01 or 0.1.

Table 5.2. The effect of PCF value on CT size, %CCI and Kappa.

	PCF 0.01	**PCF 0.1**	**PCF 0.25**	**PCF 0.5**
Tree size	7	18	41	41
CCI (%)	56.2	57.0	57.3	57.9
Cohen's Kappa	0.36	0.37	0.38	0.40

The simplicity of the tree is also an important aspect of the model results. Simpler trees were obtained with a PCF of 0.1 with slightly lower CCI (57%) and Kappa (0.37) than with a PCF of 0.5. Therefore, a simpler classification tree constructed with a PCF of 0.1 and consisting of 12 leaves (Fig. 5.2) was selected. The confusion matrix (Fig. 5.3) presents the performance of the tree by visualising the distribution of the classified instances around the diagonal of the matrix, which represents the correctly classified instances.

```
pH <= 5.82: VeryPoor (14.0)
pH > 5.82
|  DO <= 7.26: Moderate (57.0/20.0)
|  DO > 7.26
|  |  Substrates = Bo_Co
|  |  |  Conductivity >= 23.5: Good (3.0)
|  |  |  Conductivity < 23.5: VeryGood (3.0)
|  |  Substrates = Gr
|  |  |  Fe >= 932: Good (8.0/1.0)
|  |  |  Fe < 932: VeryGood (2.0/1.0)
|  |  Substrates = Co_Gr: Good (3.0/1.0)
|  |  Substrates = Co_Si: VeryGood (4.0)
|  |  Substrates = Bo_Gr: VeryGood (2.0/1.0)
|  |  Substrates = Gr_Sa
|  |  |  DO <= 8.58: Moderate (6.0)
|  |  |  DO > 8.58: VeryGood (2.0/1.0)
|  |  Substrates = Co_Sa: Good (0.0)
```

Figure 5.2. Classification tree for BMWP-Viet classes (PCF = 0.1).

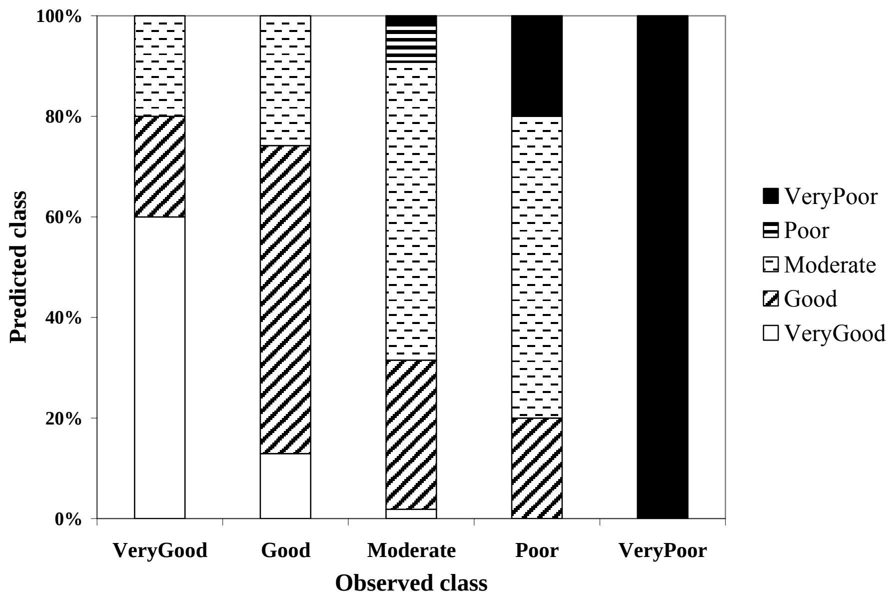


Figure 5.3. Confusion matrix of CT for BMWP-Viet classes (PCF = 0.1).

The application of the genetic algorithm for CTs resulted in the elimination of 17 input variables. The retained variables included substrates, seasons, DO, pH and turbidity. A classification tree was constructed again with the selected variables and the same settings were selected as for CTs without application of GAs. Performance significantly improved (CCI = 65.5% and K = 0.48).

With the application of GAs, a simpler tree was constructed to classify quality classes based on the BMWP-Viet (Fig. 5.4). The corresponding confusion matrix is presented in Fig. 5.5.

```
pH <= 5.82: VeryPoor (14.0)
pH > 5.82
|  DO <= 7.26: Moderate (57.0/20.0)
|  DO > 7.26
|  |  Substrates = Bo_Co
|  |  |  Seasons = D: Good (2.0)
|  |  |  Seasons = W: VeryGood (4.0/1.0)
|  |  Substrates = Gr: Good (10.0/3.0)
|  |  Substrates = Co_Gr: Good (3.0/1.0)
|  |  Substrates = Co_Si: VeryGood (4.0)
|  |  Substrates = Bo_Gr: VeryGood (2.0/1.0)
|  |  Substrates = Gr_Sa
|  |  |  DO <= 8.58: Moderate (6.0)
|  |  |  DO > 8.58: VeryGood (2.0/1.0)
|  |  Substrates = Co_Sa: Good (0.0)
```

Figure 5.4. CT for BMWP-Viet classes after variable selection by GA (PCF = 0.1).

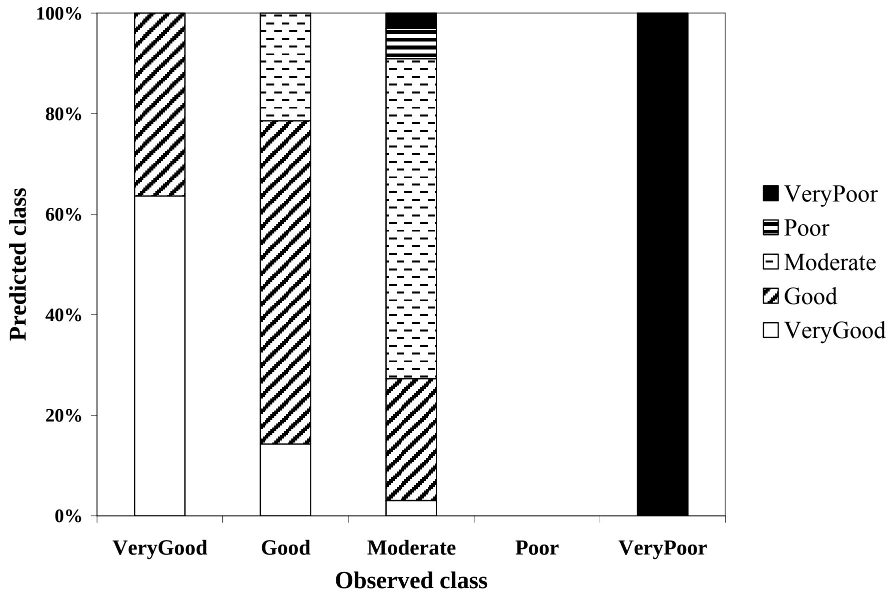


Figure 5.5. Confusion matrix of CT with genetic search algorithm (PCF = 0.1).

To obtain optimal results regarding predictive performance, the SVMs were optimised based on the application of different exponents. The outcomes obtained with different values revealed that the best performance was obtained using an exponent of 2 (Fig. 5.6).

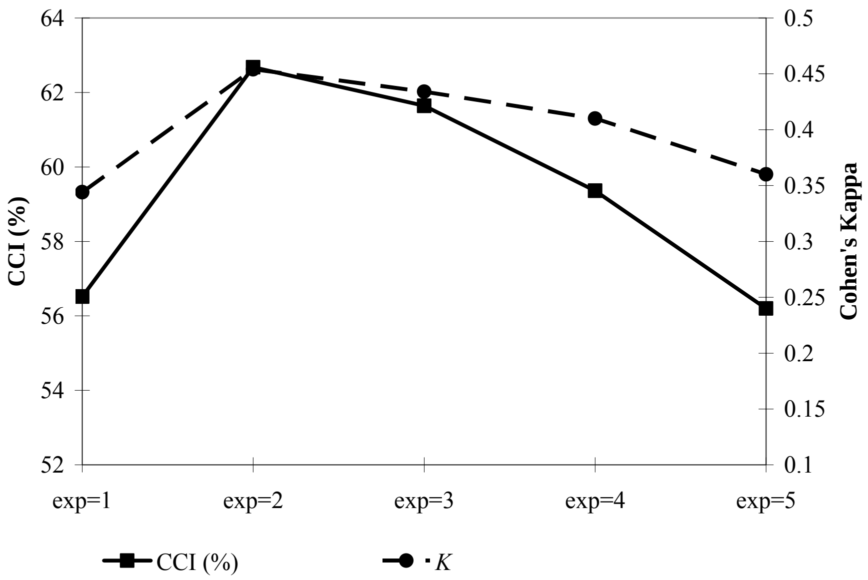


Figure 5.6. Effect of the exponent in the SVMs on %CCI and Kappa.

The overall average CCI and Kappa reached values of 62.9% and 0.46, respectively, which is a fair prediction of the BMWP-Viet classes. The confusion matrix is given in Fig. 5.7.

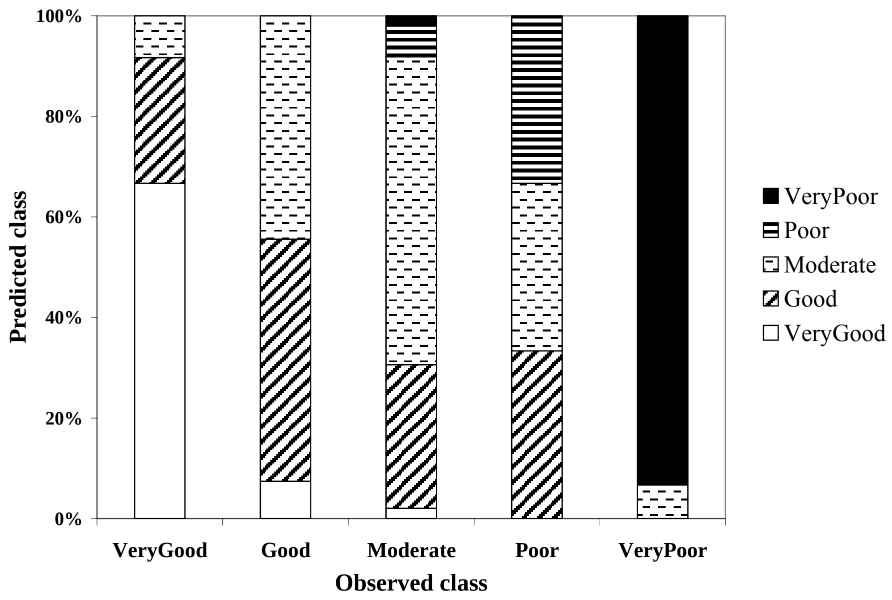


Figure 5.7. Confusion matrix for the polynomial Kernel SVM (exponent = 2).

A combination of GAs and SVMs resulted in the elimination of 15 input variables. Three of the seven selected attributes were associated with physical river habitat variables including altitude, substrate and season. Four other selected were physical-chemical ones (pH, DO, NO_2-N and turbidity). For GAs, the parameter settings were left as default values except for the exponent of the polynomial Kernel, where a value of 2 was set to allow comparison of the outcomes of SVMs before and after the variable selection procedure and to ease the interpretation of the obtained results. After application of GAs, the performance measures were CCI = 64.7 % and Kappa = 0.48, indicating a slight increase in the overall average prediction. Fig 5.8 represents the confusion matrix obtained.

Support vector machines performed significantly better than classification trees without optimisation (p<0.01 both for %CCI and K). The CTs were significantly improved after application of the GAs (p<0.001, both for %CCI and K). Optimisation with GAs did not provide significant improvement for SVMs in terms of %CCI and K. Based on these performance criteria, SVMs and CTs had a similar performance. However, the confusion matrix (Fig. 5.8) showed that SVMs with GAs application provided a better classification than CTs with and without GAs and SVMs without GAs. In addition, smaller standard errors for both %CCI and K (Fig. 5.9) indicate that SVMs delivered more stable performances.

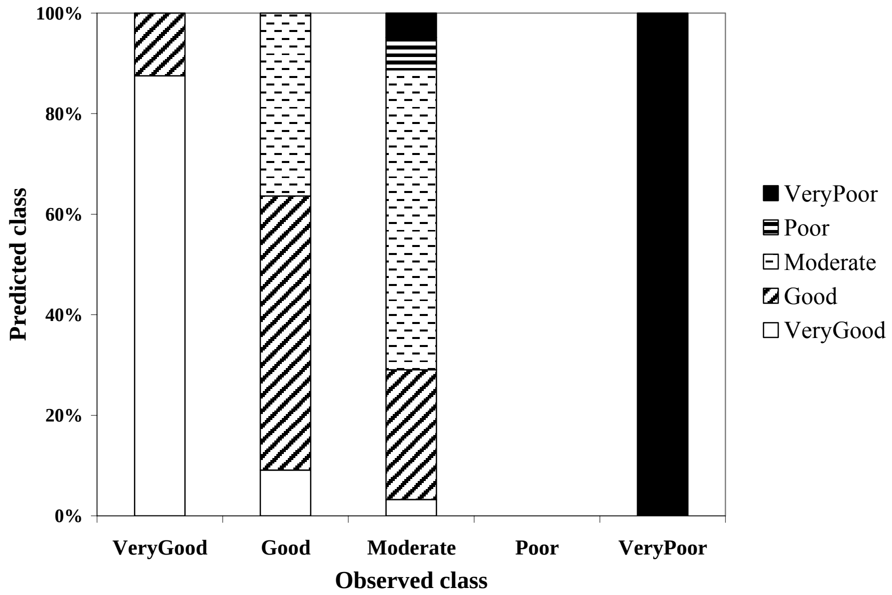


Figure 5.8. Confusion matrix for polynomial Kernel SVMs with a GA (exponent = 2).

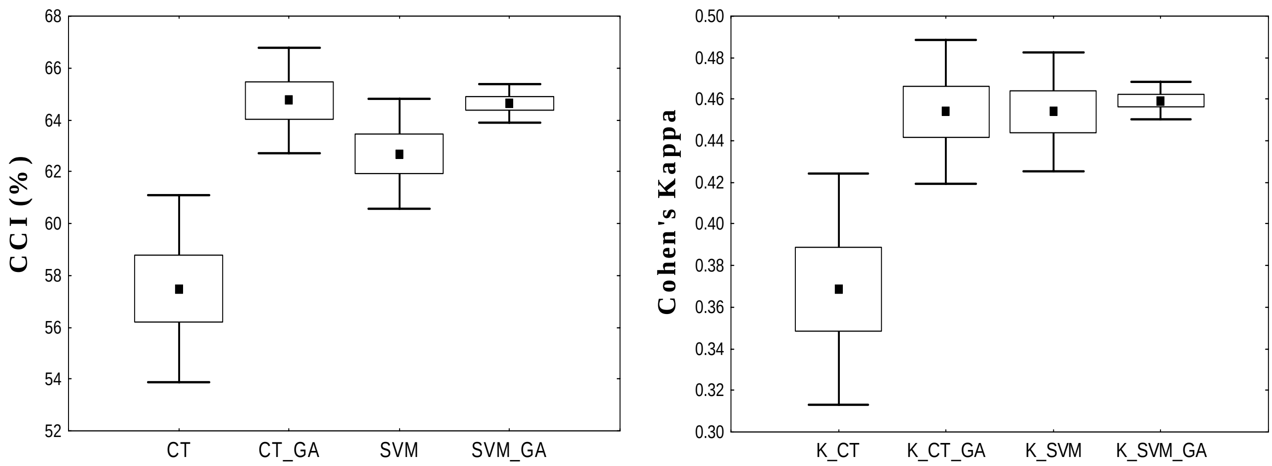


Figure 5.9. Box (standard error) and whisker (95% confidence interval) plot of the performance indicators %CCI (left) and *K* (right) for different applications: classification trees (CT), classification trees with genetic algorithm (CT_GA), support vector machines (SVM) and support vector machines with genetic algorithm (SVM_GA).

5.3.2 Ecological modelling to predict the habitat suitability for taxa

Habitat suitability models were developed to predict presence/absence of the 30 most abundant macroinvertebrate taxa in the Du river. The optimal confidence level of pruning was different for models predicting the habitat suitability of different taxa. The value of PCF = 0.1 was the optimal confidence level for 24 out of 30 models constructed, exceptions were Hirudinidae, Lymnaeidae, Aphelocheiridae, Psephenidae, Veliidae and Libellulidae. This value was therefore selected for further development of all CTs.

The SVMs were optimised based on the application of different exponents between one and five. Unlike the model developed for BMWP-Viet prediction, the best performance was obtained using an exponent of one. This default value was applied for the development of SVM models for the 30 taxa selected in this study.

In general, support vector machines performed significantly better ($p<0.001$ both for %CCI and K) compared to the classification trees (Table 5.3). According to the aforementioned CCI and K thresholds, only 10 taxa could be predicted reliably by the CTs. The SVMs, however, provided good predictions for 20 taxa in terms of both criteria CCI and K. The results illustrate the performance of models obtained by CTs and SVMs. The average value of CCI was higher both for CTs (73±8.1%) and SVMs (78±5.8%) than the considered criterion for good performance. However, the mean value of K obtained by the CTs (0.32±0.16) was lower than 0.4, while the mean value obtained by the SVMs (0.42±0.12) is higher than this threshold. For 25 taxa, better predictions were obtained with SVMs, while only for Viviparidae, Mesoveliidae, Gerridae, Leptophlebiidae and Corydalidae, K was slightly higher for CTs. Based on the CCI, the results seemed satisfying in almost all cases. However, K values indicated that high CCI were achieved just based on chance when predicting taxa having a low prevalence (D'heygere et al., 2006). Except for Aphelocheiridae and Amphipterygidae, low values of K were only obtained from the prediction of relatively rare taxa with a prevalence of less than 30%.

The small size of a dataset has been claimed to be an important cause for the poor performance of most induced trees (D'heygere et al., 2002; Goethals et al., 2002) and therefore, SVMs were proposed as an alternative method to deal with small datasets. The results of the present study showed that even with a small dataset, SVMs can provide a good predictive performance.

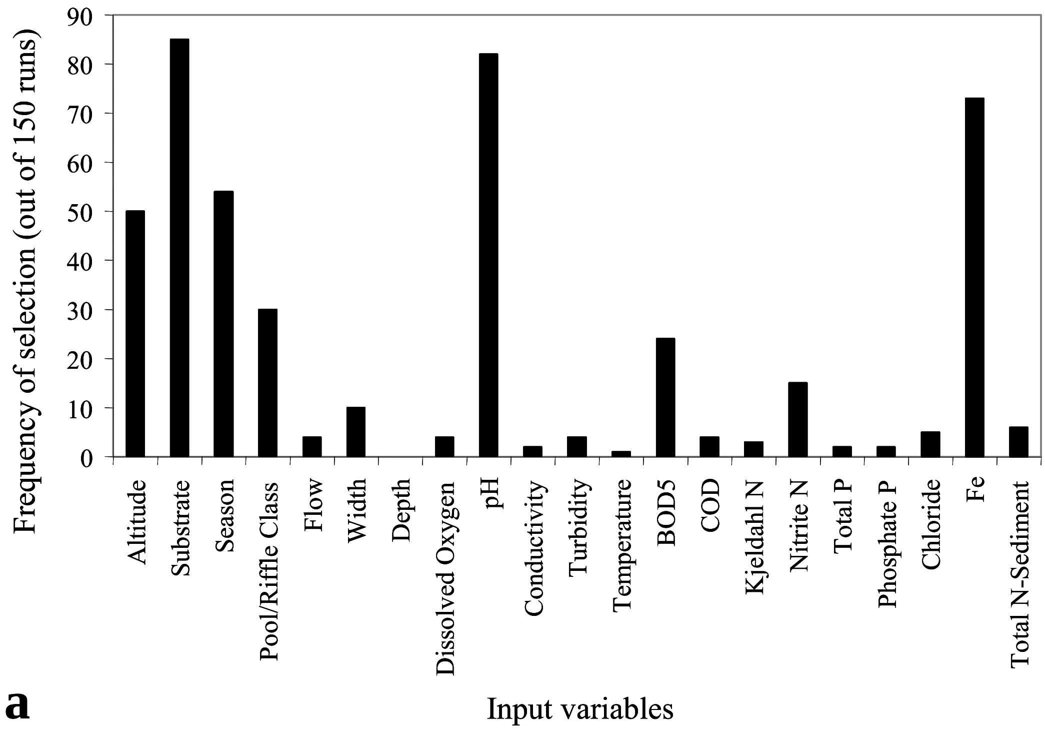
5.3.3 Genetic search algorithms for attributes selection

Different inputs have been selected by GA for different taxa. Results of input selection by the 10-fold cross-validation for the 30 selected taxa are shown in Fig. 5.10a. The numbers indicate how many times of the five runs of the 30 predictive models (150 runs in total) a particular variable was considered relevant for the prediction of the occurrence of the macroinvertebrate taxa. Fig. 5.10b reveals that for the prediction of these selected macroinvertebrate taxa in the Du river, elevation, substrate, season, P/R class, water width, pH, BOD_5, NO_2-N and total Fe in the water column were the most explanatory variables.

In the next step, these nine variables were subsequently used for the development of both CTs and SVMs. Paired Student's t-tests were conducted for the comparison of the predictive performance of models based on the J48 algorithm and SVMs before and after selection of inputs by genetic algorithms. Performances of both CTs and SVMs before and after input selection are shown in the Fig. 5.11 and Fig 5.12, respectively.

The results indicated that input selection by GAs improved the performance of the CTs according to both evaluation criteria CCI (mean CCI=75.4±7.9%) and K (mean K=0.35±0.17). Better performances were achieved with 24 models based on CCI, but only with 20 models in terms of K. This may be explained by the fact that elimination of input variables can result in losing meaningful information on the impact of environmental variables on macroinvertebrate presence.

The performance of SVMs with these 9 selected input variables was better than that of CTs (mean CCI =76±6.4% and mean K=0.35±0.15). However, the performance was worse than that of the SVMs constructed for the whole set of input variables.



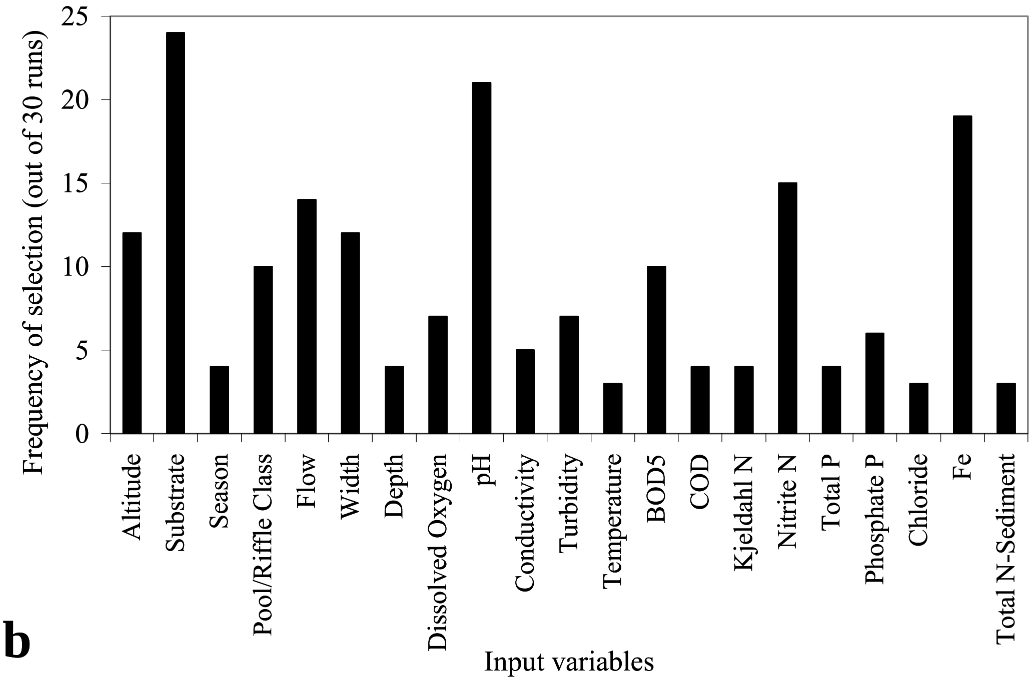


Figure 5.10. Number of runs in which input variables were selected by classification trees with genetic algorithms (a) and number of times that input variables were selected as important in SVM models (b).

Table 5.3. Macroinvertebrate taxa used for modelling based on classification trees (CTs) and support vector machines (SVMs) with and without Genetic Algorithms (GA), prevalence (%) and model performance quantified by percentage correctly classified instances (CCI) and Cohen's Kappa (*K*).

		Classification trees				Support vector machines			
Taxa		No GA		With GA		No GA		With GA	
	Prevalence,%	CCI,%	*K*	CCI,%	*K*	CCI,%	*K*	CCI,%	*K*
Oligochaeta	44	61	0.22	64	0.26	71	0.42	70	0.38
Glossiphoniidae	31	64	0.12	69	0.24	83	0.59	79	0.49
Hirudinidae	25	73	0.04	74	0.01	81	0.36	80	0.31
Pachychilidae	65	77	0.48	79	0.51	82	0.58	80	0.55
Stenothyridae	39	67	0.28	67	0.29	73	0.40	64	0.24
Thiaridae	51	64	0.28	69	0.38	70	0.40	66	0.33
Viviparidae	61	76	0.48	73	0.35	74	0.43	76	0.47
Lymnaeidae	18	80	0.04	82	0.00	81	0.19	80	0.03
Planorbidae	37	75	0.46	73	0.43	71	0.35	69	0.30
Corbiculidae	57	74	0.44	79	0.55	77	0.53	73	0.45
Atyidae	62	74	0.44	79	0.54	82	0.59	81	0.56
Palaemonidae	64	75	0.38	77	0.40	75	0.38	74	0.32
Parathelphusidae	65	73	0.34	78	0.42	76	0.41	77	0.42
Corallanidae	19	80	0.24	82	0.25	84	0.26	82	0.13
Baetidae	69	79	0.43	82	0.49	81	0.49	80	0.42
Caenidae	54	70	0.40	69	0.37	72	0.44	74	0.46
Leptophlebiidae	25	76	0.34	80	0.45	75	0.25	78	0.32
Amphipterygidae	12	94	0.67	95	0.71	91	0.58	91	0.55
Coenagrionidae	46	66	0.32	64	0.27	67	0.34	65	0.31
Gomphidae	40	67	0.32	78	0.53	80	0.57	78	0.54
Libellulidae	34	63	0.14	64	0.02	67	0.20	70	0.30
Aphelocheiridae	15	86	0.34	88	0.39	85	0.50	85	0.00
Gerridae	57	76	0.51	73	0.45	74	0.46	71	0.40
Mesoveliidae	21	81	0.34	85	0.49	82	0.33	82	0.33
Veliidae	21	76	0.13	78	0.04	81	0.31	80	0.29
Corydalidae	39	77	0.53	78	0.54	76	0.49	74	0.46
Hydropsychidae	51	59	0.17	61	0.22	72	0.44	64	0.28
Elminthidae	49	60	0.21	67	0.35	73	0.46	70	0.41
Psephenidae	26	72	0.02	74	0.00	75	0.17	73	0.11
Chironomidae	72	83	0.51	84	0.54	84	0.53	82	0.43

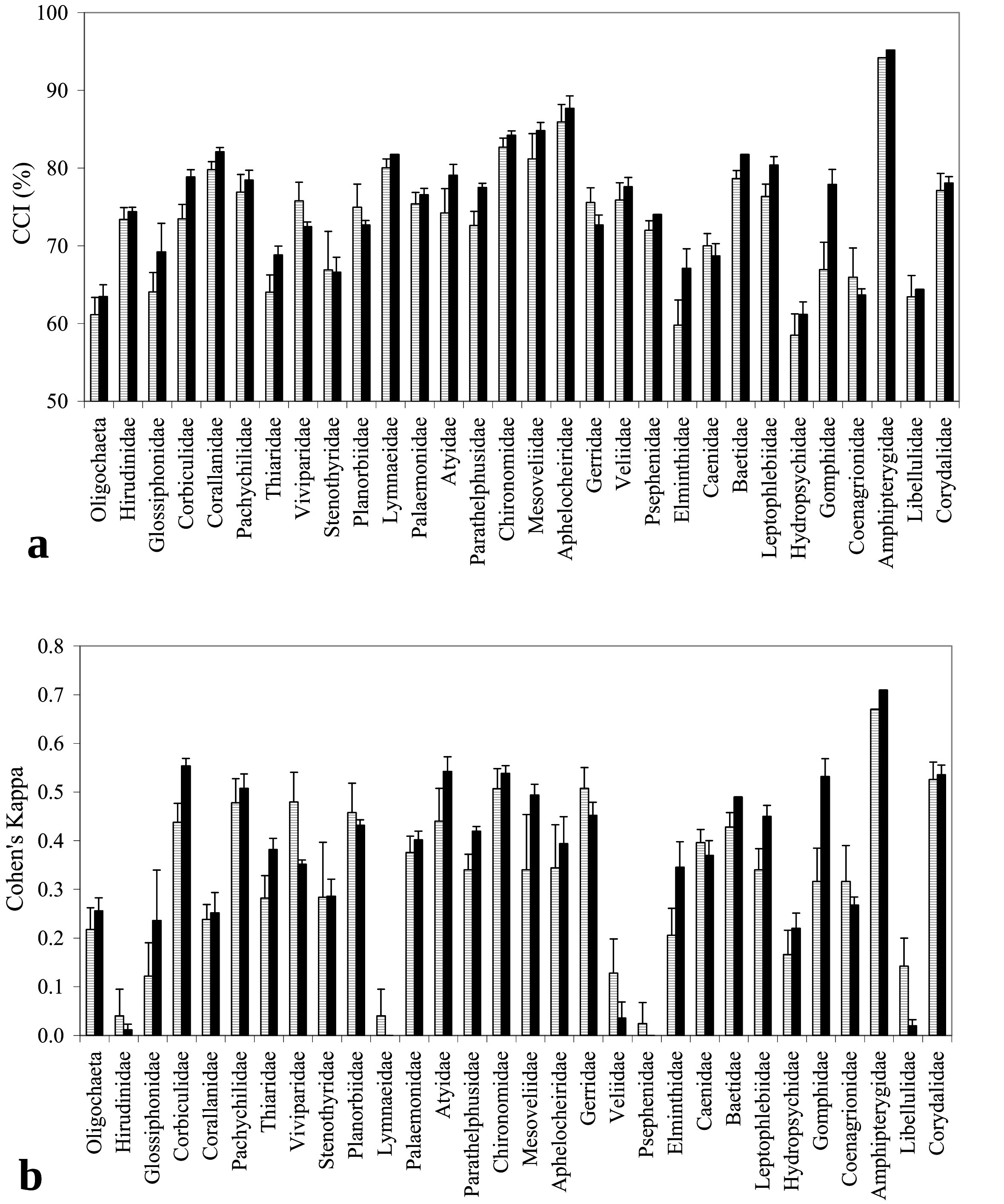
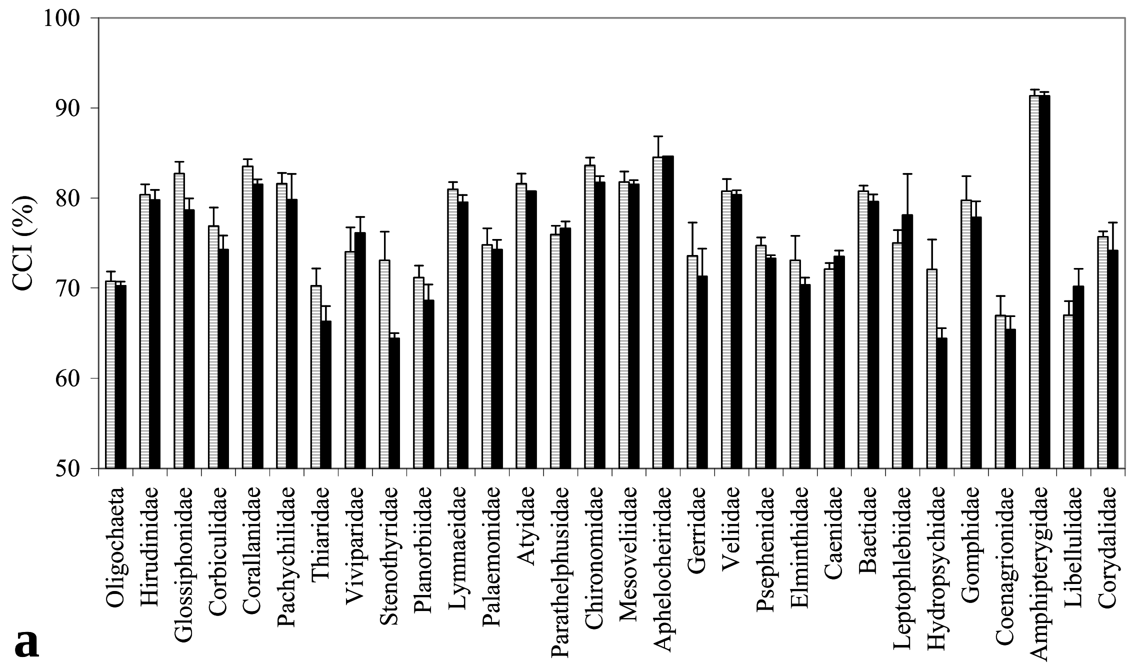


Figure 5.11. Correctly Classified Instances (CCI) (a) Cohen's kappa (b) of classification trees for 30 modelled benthic macroinvertebrate taxa before (dashed bars) and after (black bars) variable selection.



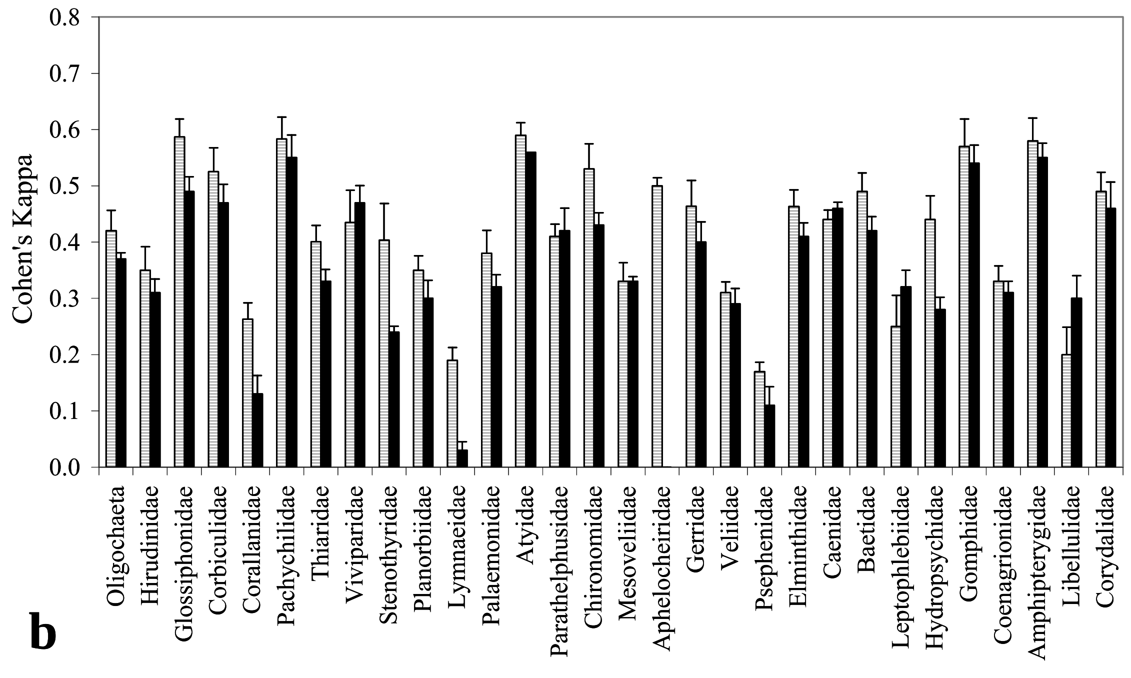


Figure 5.12. Correctly Classified Instances (CCI) (a) and Cohen's kappa (b) of support vector machines for 30 modelled benthic macroinvertebrate taxa before (dashed bars) and after (black bars) variable selection.

5.4 Discussion

Gaining insight in the complex relation between habitat conditions and the presence of macroinvertebrate taxa and consequently the BMWP-Viet index is an important task to support water management, but at the same time very hard, because selection of the key variables determining water quality is often difficult. Classification trees and support vector machines are appropriate techniques for both exploring and modelling such complex ecological data (Breiman et al, 1984).

In this study, both CTs and SVMs provided reliable models applied for the prediction of both the BMWP-Viet index and the presence/absence of macroinvertebrate taxa. The SVMs, however, have showed their robustness over CTs because they were less affected by missing data and multiple collinearity (Witten and Frank, 2005). The small size of the dataset could be an important reason for causing poor performances of the constructed trees (D'heygere et al., 2006, Goethals et al., 2002). Moreover, noise in the ecological data and missing values can negatively affect model reliability (Dakou et al., 2007).

The CTs explain variation of single response variables, by one or more input variables. The tree did merely take a small set of the input variables into account, and trees with a high pruning confidence factor seemed to perform better. However, a lower PCF value decreases tree complexity and due to the ease of interpretation, simple CTs can make more sense than complex ones (Witten and Frank, 2000). In this study, the best performing tree according to %CCI and *K* was obtained with a PCF = 0.5 consisting of 41 leaves and 26 nodes. However, the tree constructed with a PCF = 0.1 was less complex and considered preferable.

The impact of exponent values on the SVMs performance depends on the complexity of the database. The presence/absence models performed better with an exponent of one, while classification models for the BMWP-Viet having five-classes required an exponent of two. However, determining all the parameter settings in SVMs is still a heuristic process that needs further research.

Application of GAs for attribute selection significantly improved the performance of the CTs. Without the input selection by GAs, the CTs were complex and difficult to interpret because redundant attributes seemed to confuse classification systems in practice (D’heygere et al., 2003). In addition, the reduction of the dimensionality of a tree contributed to an easier interpretation of the revealed trends in the data, focusing on the more important variables. This clearly illustrated the strength of the applied machine learning technique as shown by Dzeroski et al. (1997).

In the present study the predictive performance evaluated based on %CCI and *K* did not significantly increase when a variable selection scheme by GAs was integrated with SVMs. During the prediction of the BMWP-Viet, confusion matrices, however, showed a significant improvement in the predictive accuracy of the SVMs after application of the GAs. Input selection is, however, case specific and a group of variables selected may not provide performance improvement for all presence/absence models developed.

Another advantage of SVMs is that these models provide attribute weights for all input variables. The attribute weights were different for each taxon. Fig. 5.10b indicates in how many cases each variable was the most important one in the model. A variable was considered important if it obtained an absolute weight value greater than 0.5. The variables considered most important include altitude, substrate, P/R class, water velocity, water width, pH, total Fe, BOD_5 and NO_2^--N. Compared to the inputs selected by GA, SVMs provided similar result, with water velocity as an exception. The sampling season, which proved to be not important for tropical rivers, was excluded. Consequently, these weights may provide useful information to relate the presence of macroinvertebrates to the water quality, because driving variables can be derived and ecological meaningful information can be extracted for practical application in river management.

CTs have been applied previously for practical applications in water management (D'heygere et al., 2002; Goethals et al., 2002; Dakou et al., 2007). These studies also illustrated that CTs with GA application can be a useful tool for water management. SVMs even have more the advantage of getting insight in the effect of environmental variables. They are able to assess the effect of combinations of driving variables to describe favourable conditions for macroinvertebrates without removing any variables.

The results of the developed models supported measures selected for improving ecological quality in the Du river. The driving variables selected by the models (pH, BOD_5, NO_2^--N and total Fe) revealed that reduction of industrial pollution loads and of nutrient and organic pollution inputs into the river in upstream and middle stream reaches could be effective restoration measures. In the downstream reach, restoration of the river bed affected by sand exploitation was also suggested based on the P/R class, which was selected as relevant by the developed models. Driving variables revealed by the optimised models thus provide meaningful information for river management. This advantage makes SVMs and CTs also potentially useful to support river restoration and conservation strategies. More detailed studies on specific macroinvertebrate taxa may also provide better insight in environmental impacts.

5.5 Conclusions

The present study developed models to explore the habitat suitability of 30 macroinvertebrate taxa in the Du river in Northern Vietnam as well as the BMWP-Viet index for water quality classification. Performances of the SVMs and CTs were compared both with and without genetic algorithms. The SVMs have recently gained popularity in many domains since they are considered as very successful, in particular for classification problems. The results of the present study revealed that the developed models using SVMs outperformed CTs and the provided attribute weights of SVMs may be a reliable alternative to GAs to select input variables. These weights also provide useful information regarding the importance of environmental variables influencing the presence of macroinvertebrates, which makes SVMs a promising tool for river management.

Monitoring strategies should focus on these selected input variables, while keeping physical variable as key factors determining favourable habitat conditions for macroinvertebrates. Comparison between monitoring results of sites with similar physical habitat conditions can indicate the pollution status of these sites.

Chapter 6

Application of the EU Water Framework Directive Explorer to support water management in Vietnam based on biological assessment

6.1 Introduction

Integrated river management is based on monitoring, modelling and assessment of the water cycle. Effective and efficient management is only achieved when all three activities are well interconnected and concerted. The choice of appropriate, high quality monitoring and assessment techniques is probably one of the most crucial factors in the assessment of river systems (Goethals and De Pauw, 2001). Therefore, the EU has adopted a Water Framework Directive (WFD), which came into force on December 22nd 2000. The WFD focuses on the ecology and chemistry to realise an improvement and a sustainable use of the environment.

The Directive aims at maintaining and improving the quality of the aquatic environment in the Community (EU, 2000). The EU WFD prescribes the achievement of a 'good status' by 2015 as a target for all water bodies and stipulates that each member country is responsible for ecologically healthy water systems. However, the evaluation of the ecological situation of water bodies and of the effectiveness of restoration measures is often difficult. Available ecological knowledge must therefore be assessed and compiled into water management and policy.

Due to this information need, a project was started in the Netherlands to develop a Planning Kit, named the WFD-Explorer, which can assist water managers with the development of river basin management plans. The WFD-Explorer comprises a relatively simple tool to develop and analyse various strategies to realise ecologically sound water bodies. The WFD-Explorer performs a number of functions related to the implementation process of the WFD. It creates a platform to integrate and disclose ecological information on water bodies, including a common knowledge base on the effectiveness of measures. It also allows stakeholders to deepen their understanding of how the pursued objectives can be achieved, which measures might be taken and what the impacts are of such measures.

This chapter aims to investigate an application of the EU WFD-Explorer as a communicative tool for water quality management in Vietnam. The application was tested in the study area and based on the outcomes, possibilities of adaptations as well as shortcomings during application for water management in Vietnam are discussed.

6.1.1 The EU Water Framework Directive (WFD) Explorer

The WFD-Explorer has been developed by a consortium of representatives of different Water Boards, the STOWA (Foundation for Applied Water Research), the RIZA (Institute for Inland Water Management and Waste Water Treatment) and a number of other institutes and consultancy firms. Detailed river basin studies are conducted on a fine scale while river managers have to make decisions on a coarse scale for river bodies. Therefore, a communicative tool, integrating the findings of previous research, is needed to bridge this gap between the results of detailed studies and the information used in management (Gottardo et al., 2007).

The goal of the development and application of the WFD-Explorer is to support water managers in the development of management plans for river basins. This support is focused on the discussion and communication about restoration measures implemented at the water body or the river basin scale level. The WFD-Explorer does not aim to optimise the scope of measures at an individual water body level, however, it can give an impression for a particular water body of what a restricted or extended application of a measure could mean.

Application of the WFD-Explorer facilitates a uniform approach and contributes to the harmonisation of local and regional water management processes. The WFD-Explorer can also be used to evaluate the Maximum Ecological Potential (MEP) for heavily modified water bodies. Application of the WFD-Explorer reveals the water body status that can be achieved with the selected measures. The WFD-Explorer also offers a platform for the compilation of information from process and system studies, so that the information stream remains manageable and new insights can be efficiently integrated.

The WFD-Explorer contains a simplified description of the hydromorphology and water quality of the water bodies within the river basin (based on a water and substance balance of the water bodies within the river basin). It aims to provide a rough estimation of the abiotic water body characteristics and to link these characteristics to ecological quality. Consequently, the impacts of river restoration on the abiotic water body conditions and thus on the ecological quality can be analysed.

The most important part of the method developed within the WFD-Explorer is the knowledge rules. The knowledge rules quantify relationships between measures and steering variables and between steering variables and ecological quality. In the WFD-Explorer the effects of measures on the ecological situation of the water bodies are quantified mainly based on the steering variables. The steering variables describe the abiotic environment as a necessity for a certain ecological quality. There are also a number of top-down measures that are directly connected to ecology, for example, active biological management. The method developed for the WFD-Explorer is shown in Fig. 6.1.

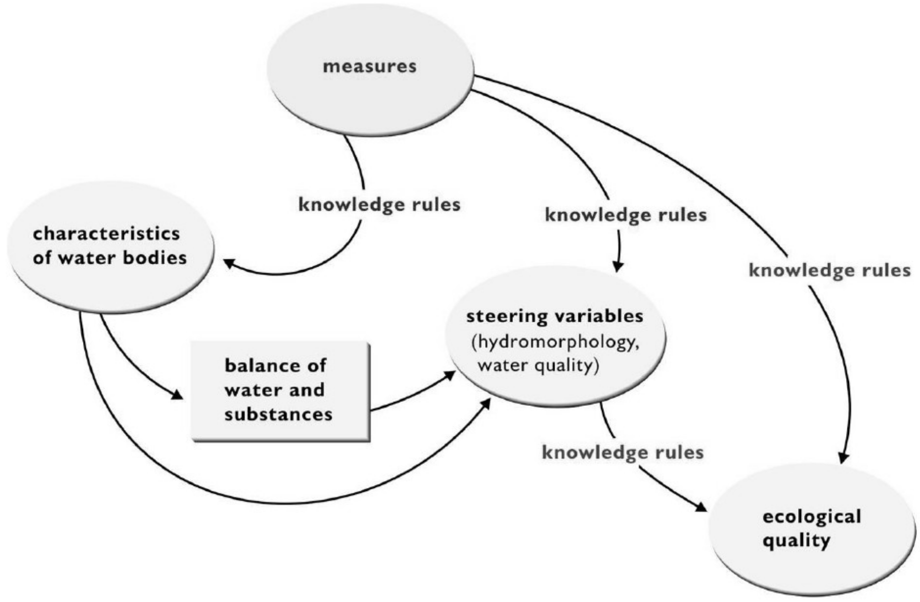


Figure 6.1. Outline of the method developed within the WFD-Explorer (Van der Most et al., 2006).

The development of the knowledge rules for the various ecological quality elements and the various water types is very different. This depends on the availability of suitable datasets and also on the experience and preference of the ecologists involved in the development of knowledge rules.

The first version of the WFD-Explorer has been developed and successfully tested on a pilot river basin in the Netherlands (Van der Most et al., 2006). The prototype was further developed into a base version which can easily be applied elsewhere in the Netherlands and abroad. A number of water boards and river basin districts in the Netherlands are currently in the process of developing applications of the WFD-Explorer for their areas. This further development was carried out in conjunction with a number of ongoing planning studies in the Netherlands (Van der Most et al., 2006).

A pilot study in Flanders (Belgium) was applied to test the usability of the WFD-Explorer in an international context (Maes et al., 2007). The basin of the Burgravenstroom, which is heavily subject to diverse activities (agriculture, residential, drinking water production), was used as a study site. Based on a limited number of data, it was possible to predict the chemical and ecological quality of the water bodies using the WFD-Explorer. Hence, the bottlenecks (impact from agriculture, water velocity, etc.) became apparent and effects of restoration options on the water quality could be predicted. In another study, data from the Zwalm river (Belgium) were implemented in the WFD-Explorer to predict the effect on the river quality of restoration actions proposed by the Flemish government (Mouton et al., 2009). The WFD-Explorer was successful in detecting the most effective restoration measures. However, the study also indicated that the spatial scale on which the WFD-Explorer modelled the impacts of physical habitat restoration might significantly affect the reliability of the results.

Currently, two Dutch enterprises are assisting Romania with the implementation of the Water Framework Directive, in particular with the rehabilitation of 'Heavily Modified Water Bodies' (Madar et al., 2008). The use of the WFD-Explorer becomes common in European countries (Van der Most et al., 2006), which stimulates potential application in other regions. This communicative tool could help to close the gap between managers and stakeholders and consequently lead to a more integrated water management in other countries, including Vietnam.

6.1.2 The WFD-Explorer as communicative tool for water management

In Vietnam, some recent changes have provided a legal and institutional basis that will allow dealing more effectively with water quality and environmental management. However, the lack of comprehensiveness and efficiency in managing water resources has become a serious problem facing developing countries such as Vietnam (Al-Shemmeri et al., 2001; Ito et al., 2001) and no effective mechanisms exist to mobilise participation of private, non-government or industrial organisations.

Integrated water management is costly and time consuming but investments in the development of a powerful predictive tool could provide a good support system (Ito et al., 2001). Recently, new advances in computer technology have enabled widespread improvement in the understanding of key factors and processes involved in the degradation of surface water quality (Assaf and Saahed, 2008). These are useful in water resources planning and efficient management, in combination with an incremental cost-effectiveness analysis (Goethals, 2005). These decision support systems are helpful tools to reduce the subjectivity and to increase insight in cause-effect relationships. These often allow graphical representation of any complex water-resource system, which makes interpretation of the results more easy (Goethals et al., 2001; Ito et al., 2001). Despite these advantages, involving river managers in the model development process is not an easy task and a serious gap exists between the state of the art in water-resource system analysis and the usage by practitioners at the real world level. For integrated management on a large scale, the use of strong directives in combination with decision support systems plays a crucial role, stressing the need for the development of an integrated water management system (Andreu et al., 1995; Goethals, 2005).

As discussed in previous chapters, the Du river basin is facing serious impacts of anthropogenic activities. Various stresses include agricultural activities, domestic water use and discharge, solid waste disposal, run-off from coal and metal ore mining, discharge from metal extraction workshops and extraction of bed sediments. A description of the major anthropogenic disturbances in the Du river basin has been given in chapter 2. These impacts have already drastically reduced the water quality of the Du river, which was reflected in both abiotic and abiotic river characteristics (Chapter 3, 4). Integrated water management systems can provide a platform for the design and the implementation of necessary measures for the improvement of the ecological health of the Du river. The WFD-Explorer has been applied to explore the influence of specific measures on the ecological and chemical quality of the area.

For the WFD-Explorer, the integral predictions of hydrology, water quality and ecology in a water system are based on a good balance between the required data collection, the usability and the fast response to key questions of water managers. Moreover, the simulation time is relatively short and the visualisation allows for a better interpretation (Maes et al., 2007). The WFD-Explorer proposes a list of restoration options that can be implemented to evaluate their effects on water quality (Van der Most et al., 2006). However, a common focus on the major sciences determining water quality (hydrology, geomorphology and ecology) is of the utmost importance when a choice between restoration measures has to be made (Newson et al., 2006; Devesa, 2009).

6.2 Material and methods

An analysis was performed with the WFD-Explorer version 1.04.0025 available on the website www.krwverkenner.nl. Application of the WFD-Explorer for a watercourse should be designed according to the following phases (WFD-Explorer Users' Manual, 2007):

- the definition phase: examines the area and the measures, and defines the problem substances and the parties involved. This phase largely dictates the following phases;
- the construction phase: here the necessary data is gathered and entered in the WFD-Explorer database and the measures are configured. An initial WFD-Explorer application is ready at the end of this phase;
- the acceptance phase: decides whether the application construction phase has been satisfactory. This should be carried out in close consultation with the water management experts. Part of this phase is the calibration and the further refinement of

the measures after testing by the users. At the end of this phase the definitive WFD-Explorer application, that fulfils the required purposes, is set up and ready for use;

- the application phase: here analyses will be carried out by the WFD-Explorer, the tool will be used in discussions and reports will be made.

In the whole process, the construction phase is essential for a successful application of WFD-Explorer.

6.2.1 Definition of the water bodies and basins

The WFD-Explorer operates at the sub-basin scale level, the smallest unit being a water body. Therefore, the Du river has been considered as a system of water bodies. At each intersection point of the stream a new part of the total water body has been defined, resulting in five water bodies: Na Lau river, Du river, Cat stream, Main Du A, Main Du B (Fig.6.2). To simplify an application of the WFD-Explorer, the Khe Coc stream has not been included in the study area. Relatively small contributions of its flow into the main river, especially during the dry season as well as insignificant pollution levels make the Khe Coc stream insignificant in the total Du river system. For each water body, the WFD-Explorer calculated the flow and nutrient concentrations based on monitored and estimated inflows of nutrients and water. The sample values of variables were based on the average measured values in the Du river.

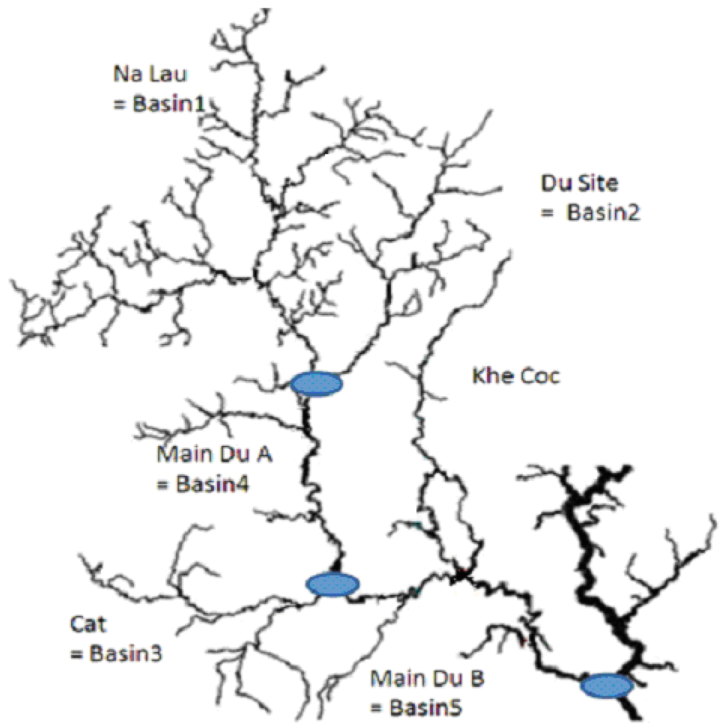


Figure 6.2. Study area used for the WFD-Explorer including five defined water bodies (basins).

All input characteristics of each water body were incorporated in the WFD-Explorer using a MS Access Database. Heterogeneous water body characteristics such as river depth, width and the river bank profile were monitored together with general characteristics such as slope, length and sinuosity for each stream. From these observed water body characteristics and the calculated flow, a one-dimensional hydraulic model, which is implemented in the WFD-Explorer, derived the values of flow velocity and the dominant substrate class for each stream. Based on these values, the WFD-Explorer roughly estimated the average flow velocity and depth as well as the distribution of these parameters. Finally, the current river status and the effects of different restoration options were analysed at the river basin level.

The studied river was visualised using Mapper (Demis bv – Delft Hydraulics WL, the Netherlands), which is a GIS tool that allows the implementation of interactive maps into applications (Mouton et al., 2009). Several water body characteristics can be visualised for each water body on a map of the river basin, such as nutrient concentrations and Ecological Quality Ratios (EQRs), which quantify the ecological quality of a water body for fish or macroinvertebrates.

6.2.2 Knowledge rules base of the WFD-Explorer

The WFD-Explorer presents a simplified description of the transport of substances through an area. All items of the water and substance balance have to be entered in the WFD-Explorer in advance. However, a proper analysis of all available data sources of the project area is necessary.

The WFD-Explorer uses a cascade like water and substance balance to show the flow of water and substances throughout the study area. Discharge and the waste loads from the upper segments are presented to the segment located downstream as an incoming source of water and substances. In the present version of the toolbox, 14 substances have been included into the knowledge base of the WFD-Explorer. These are general substances (BOD_5, Cl^-, total N, total P), heavy metals (Cd, Cu, Ni, Pb, Zn) and organic micro-pollutants (Benzo(a)pyrene, benzo(k)fluoranthene, Diuron, Fluoranthene, HCB).

Nutrient retention in various water systems was analysed by deducing a meta-model for retention of nutrients based on a deterministic water quality model for a lowland stream in The Netherlands (De Klein, 2008).

$$C = \frac{C_{in} + k_0 T}{1 + k_1 T} \text{ and } R = 1 - \frac{C}{C_{in}} = 1 - \frac{1 + \frac{k_o}{C_{in}} T}{1 + k_1 T}$$

in which:	C	concentration (mg/l)
	C_{in}	upstream concentration (mg/l)
	k_0	zero order decay-rate (1/d)
	k_1	first order decay-rate (1/d)
	T	residence time (day)
	R	retention fraction (-)

The knowledge rules were developed based on empirical data of lowland streams which characterize the Netherlands and Flanders, the northern part of Belgium. Consequently, transferability of these rules to other ecoregions in the world may be limited (Randin et al., 2006; Fitzpatrick et al., 2007; Strauss and Biedermann, 2007). To ensure consistency among the results of the different applications of the WFD-Explorer, the user cannot modify the knowledge rules in the current toolbox version. However, experts and model users are continuously being involved in the reviewing process of the expert rules, which may lead to adaptations of these rules in future versions of the WFD-Explorer (WFD-Explorer Users' manual, 2007).

6.2.3 Input variables

For each water body, the WFD-Explorer requires 13 input variables (Table 6.1). For each water body in the region, the main sources of pollution were agricultural activities and domestic households, characterised by the following chemical variables: total phosphorus (TP), total nitrogen (TN) and biochemical oxygen demand (BOD_5). Water body catchment area, length and sinuosity were derived from aerial photos and maps of the river basin using GIS. The river bank profile was monitored visually into four classes, while the water body slope was derived from the geographical maps of the water body by determining the difference in river bed elevation at the start and the end of each water body. For each stream, the polluting villages and the amount of inhabitants was obtained by surveys.

Table 6.1. Input variables required for implementation of the toolbox.

Variables	**Unit**
Input variables	
Profile width	m
Profile depth	m
Profile shape	four classes
Water body slope	m/m
Water body sinuosity	five classes
Water flow directly to the water body	m^3/day
Water flow from catchment to water body	m^3/day
Total N flow directly to the water body	kg N/day
Total N flow from catchment to water body	kg N/day
Total P flow directly to the water body	kg P/day
Total P flow from catchment to water body	kg P/day
BOD_5 flow directly to the water body	kg BOD_5/day
BOD_5 flow from catchment to the water body	kg BOD_5/day
Output variables	
Flow velocity distribution	-
Dominating substrate distribution	-
Depth distribution	-
Average water velocity	m/s
BOD_5 concentration	mg O_2/l
Total N concentration	mg N/l
Total P concentration	mg P/l

The major water inflows into each water body were quantified considering households' wastewaters and surface runoff. The water flow from households was determined using an equivalent volume of 120 l/day/person (UNESCAP, 2009), the volume of wastewater produced per head, in combination with the number of inhabitants in its respective catchment according to the surveys. For the calculation of agricultural runoff, assumptions were made due to lack of reliable data. Therefore, a maximal value was set as 1000 m^3/km^2/day which is the average daily value for the river (Alan, 2006).

The pollution due to agriculture was considered as flowing from the catchment and was derived from the average nutrient runoff (P and N) of agricultural soils and the area of agricultural land use in each catchment. The amount of fertilizer consumption was obtained from the National Institute for Soil and Fertiliser (NISF). The annual chemical fertilizer requirement for agricultural production in Vietnam is estimated as 1,200,000 ton N, 728,600 ton P_2O_5 and 600,000 ton K_2O. Additionally, manure is used to fertilize the field and this amount is estimated equivalent to 210,000 ton N, 105,000 ton P_2O_5 and 230,000 ton of K_2O (NISF, 2008). With a total surface of 5,630,000 ha agricultural land in Vietnam, a rough estimation of fertilizer use per hectare could be calculated. The amount of nutrients finally lost to the river water has been estimated considering the uptake efficiency and soil removal fraction.

The pollution load of domestic households on each water body was derived from the number of inhabitants in its respective catchment. There is not any connection from houses to a wastewater treatment plant. The amount of N, P and BOD5 (kg/day) from domestic sources entering into each water body was calculated from survey data and benchmarks given in the literature.

6.2.4 Model calibration and validation

The calculated values of water velocity, BOD_5, total phosphorus and total nitrogen concentrations were calibrated based on field observations. These field data indicated realistic ranges of the physical and chemical variables assessed, hence allowing calibration of water and nutrient flows in the WFD-Explorer. The flow in each water body was calibrated by adjusting the error term which consisted of in- and outflows due to ground water inflows of local water extractions or discharges. Nutrient inflows were calibrated by adding a nutrient inflow error term for the two chemical variables total phosphorus and BOD_5 concentration. This error term consisted of nutrient inputs due to atmospheric deposition, groundwater inflow or other discharges (Mouton et al., 2009). However, groundwater and atmospheric related processes have not been studied in detail, thus no data were available for calculation. Therefore, assumptions had to be made, resulting in only rough estimates of both water and nutrient flows, only indicating the magnitude of the different sources. Water and nutrient flows to and from each water body *i* could be described by:

$$Q_{out,i} = \sum_{k=1}^{N} Q_{in,i,k} + \varepsilon_i$$

in which: $Q_{out,i}$: the nutrient or water outflow of water body *i* going to its receiving water body *j*,

$Q_{out,i,k}$: the k^{th} inflow of a set of *N* known inflows to water body *i*, a negative inflow being a nutrient or water outflow from water body *i* which is not flowing to water body *j*,

ε: an error term describing all inflow errors or unknown inflows such as ground water inflow or local water extraction or discharge.

Adjustment of these error terms allows further calibration of the WFD-Explorer to field data.

6.2.5 Simulation of restoration options

The basic version of the WFD-Explorer contains 46 different restoration measures. These measures vary from the cleaning of point sources to active organic management. Some developed measures were appropriate for improving water quality in the Du river basin and could be proposed. The working mechanisms of all these measures are implemented in the basic version of the WFD-Explorer. There can be some other measures apart from those included in the WFD-Explorer that can contribute to an improvement of the ecological health of the Du river. However, only the measures already developed in the basic version of the WFD-Explorer, which were considered most important, were selected for this investigation.

Choosing the most appropriate restoration technology is not an easy task but it could reduce the risk of future problems and failures. The preferable method is a technology that is economically affordable, environmentally sustainable and socially acceptable (Massoud et al., 2009). Specifically, the most suitable restoration options could be suggested for each water body based on the impact of these options on the variables affecting water quality. Consequently, the following measures were proposed for restoration of the Du river basin.

6.2.5.1 Cleaning of point sources from mining and metal extraction workshops

Previous chapters 3, 4 and 5 described pH and total Fe as the strong driving variables affecting the quality of the water in the Du river basin. However, the input variable analysis by data mining indicated that extremely low pH values and high metal concentrations locally presented in the downstream part of the Cat stream. Due to dilution and precipitation processes along the stream, the pH returned to a normal range in the main Du river. The concentration of Fe increased in the main B river section compared to the upstream part. However, the average concentration was much lower than the threshold given by the classification tree analysis.

Cleaning these point sources needs to be the prioritised measure for restoration of the Du river basin. The pH of the water can be increased by acid neutralising reactions involving bicarbonate alkalinity. Metals in the wastewater can also be removed by this simple cation exchange. Investigation showed that 95% of Fe can be removed after this application (Nguyen and Dang, 2008). Consequently, the Fe concentration in the stream water could return to the level, which had no direct effect on presence/absence of the macroinvertebrates.

Information about treatment of acid drainage from tin extraction workshops is case specific and thus not provided by the WFD-Explorer. This measure therefore was performed outside the range of the WFD-Explorer. Assumption was made that after application of wastewater treatment, the water body has the same condition of the upstream site 9.

6.2.5.2 Reducing the impact of agricultural runoff

The dominant land use in the study area is agriculture, indirectly affecting environmental and human health. It may result in contaminants leaching or seeping into surface water and can cause changes in the quantity and quality of runoff which may degrade the environment, especially the aquatic environment. These non-point sources are difficult to identify and they cannot readily be monitored. Moreover, taking samples is a complex and costly task. Some options for minimising agricultural impacts involve (Volkman, 2003) (1) selection of the most appropriate land use for the site and circumstances, (2) increase in efficiencies of the application of farming inputs, (3) increase in the resistance of farming systems to losses of nutrients and chemicals and (4) a greater use of field and landscape buffer zones (Fig.6.3), the last belongs to measures proposed by the WFD-Explorer.

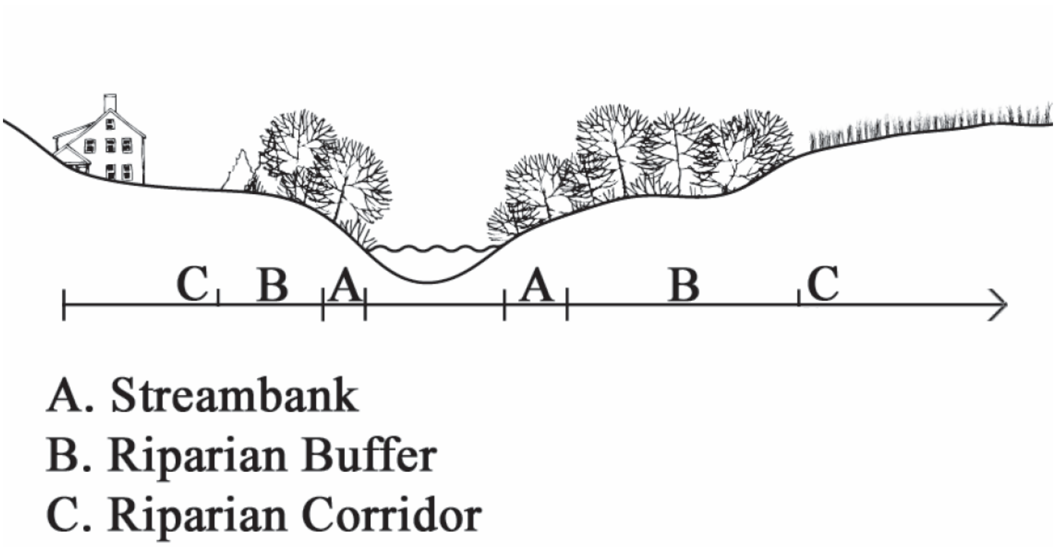


Figure 6.3. Buffer zone structures (MRWP, 2008).

A well maintained buffer serves as a natural filter, reducing the amount of nutrients such as P and N, sediments and chemicals that enter a stream or river (Barling and Moore, 1994). With the increased canopy that trees and shrubs provide, direct sunlight is reduced, lowering the water temperature and thus resulting in higher oxygen concentrations. In addition, riparian buffers are essential to feed, shelter and provide travel paths for a lot of species (Mander et al., 2005).

These filter systems were assumed to remove phosphorus from agricultural runoff with a removal efficiency of at least 50% (Volkman, 2003). However, the WFD-Explorer showed that phosphorous was not an important steering variable for the Du river basin. Consequently, application of this measure is expected to affect the water quality of the Du river insignificantly.

6.2.5.3 Decentralised wastewater collection and treatment

The Du river is located mostly in rural and remote areas. Providing reliable and affordable wastewater treatment in rural areas is a challenge, especially in developing countries (Massoud et al., 2009). A large scale treatment of the wastewater requires a centralised wastewater collection, which is costly to build and to operate, especially in areas with low population densities and dispersed households, like the river basins of the upstream branches of the river Du and the river Na Lau. Therefore, a solution focusing on a decentralised approach for wastewater treatment that consists of a combination of (clustered) onsite systems is suggested (Anh et al., 2002). As these are designed to operate at smaller scales, decentralized wastewater management options require efficient, cost effective and compact wastewater treatment processes (Bakir, 2001; Anh et al., 2007; Massoud et al., 2009). When used effectively, decentralized systems promote the return of treated wastewater within the watershed of origin. Nowadays, these systems can be designed for a specific site, thus overcoming the problems associated with site conditions such as high groundwater tables, impervious soils, shallow bedrock and limestone formations (Massoud et al., 2009).

Because of BOD_5 concentration was the limiting variable for the water quality, a proposed restoration option was the construction of small scale water treatment devices, such as helophyte filters. This restoration measure was included in the WFD-Explorer intervention package.

6.2.6 Combining classification trees with WFD-Explorer in evaluating the effectiveness of restoration measures

In original WFD-Explorer application, the impact of different restoration options on the ecological health of the river basin can be assessed immediately by selecting measures. The measure can be applied locally to one selected water body or generally to all water bodies in the river basin. After selection of a measure it is shown by the WFD-Explorer how much this measure contributes to the realization of ecological objectives by mean of the EQR. Due to this setup, the WFD-Explorer offers to water managers and stakeholders the opportunity to get insight in the relationships between the objectives pursued, the measures that might be taken and the impacts of such measures (including costs) (Van Der Most et al., 2006).

The impact of restoration measures on the EQR were calculated based on the effect of these measures on the physical and chemical water body characteristics. However, the knowledge rules developed for calculation of EQRs in the Netherlands may not be appropriate for conditions in Vietnam. In addition, validation of the EQR requires specific data on the abundance of macroinvertebrates and identification at genus level. The collected macroinvertebrates in this study were only identified at family level and only presence/absence of macroinvertebrates were monitored. Therefore, the EQR concept could not be applied in the present study.

In order to overcome this limitation and take advantages of the WFD-Explorer to support water management in Vietnam based on biological assessment, the WFD-Explorer was combined with datamining techniques. The restoration measures relevant to the study site were configured in the presented WFD-Explorer application to analyse the change of chemical characteristics of water bodies after restoration measures. Results were applied for further calculation of the BMWP-Viet, which can be predicted by the CTs and SVMs as discussed in the chapter 5, to evaluate the effectiveness of these restoration measures.

Although SVMs proved to have better prediction performance than CTs in general, the trees created by the CTs provided clearer elucidation, which is very useful for management purposes. In addition, application of GAs in selecting the most important driving variables also significantly improved the predictive performance of the CTs (Fig. 5.9.).

The WFD-Explorer calculated values of variables characterised the water quality in each river basins according to each restoration scenario. These values were used as inputs for CTs developed in chapter 5 in order to provide ecological classification based on BMWP-Viet for each river segment. In these predictive models, physical habitat conditions were kept as driving variables while chemical ones were changed according to results of WFD-Explorer calculation.

The input variables were selected based on the GA application. These included physical characteristics such as altitude, substrate, P/R class and chemical variables. Although the DO was selected by the GAs, this variable was highly correlated with the BOD (Appendix 1), which was an input of the WFD-Explorer. Therefore, BOD was selected to replace DO as input for the BMWP-Viet calculation. Chemical variables, which were also selected as inputs, include total P, total N and pH.

The tree was constructed with a pruning confidence factor of 0.1 (Table 5.2). The model training was based on the whole set of 104 instances. The trained model was then applied for prediction. The predictive performance provided classification of quality classes of each river segments based on BMWP-Viet. Based on the results predicted by the classification trees, the effectiveness of restoration measures was evaluated. Results were illustrated by visualisation graphs.

6.3 Results

6.3.1 WFD-Explorer model calibration

Calibration of the error terms in the mass balances of the model led to accurate predictions of the chemical water quality in all water bodies. Based on the average values of the measured water velocity, the calculated water velocity was calibrated by adjusting the error term ε_i and this calibration reduced the difference between observed and calculated average water velocity. The toolbox also indicates which variable is most affecting the water quality in each water body of the studied river basin.

The calibration of the WFD-Explorer focused on the BOD_5 and total P concentrations, which were detected to be the most important variables determining water quality. Sinuosity was not limited in the presented study and the values of velocity showed a high variation in time and were therefore not considered as suitable for calibration. TP_{calc} was not correlated with TP_{obs} (R^2=0.1608). This was caused by one outlier value: an extremely high phosphorus concentration was observed during one of the sampling campaigns in the dry season. At site 4 for W07, a value of 5.65 mgP/l was obtained, while all of the other values were less than 1.5 mg/l. Outliers were therefore removed from further analysis.

Table 6.2. Regression coefficients, significance (*p*) values and R^2 of the regression between the observed variable values and the values calculated by the WFD-Explorer.

Dependent variable	Independent variable			Intercept		R^2
	Variable	Coefficient	*p*	Coefficient	*p*	Value
TP_{calc}	TP_{obs}	0.1608	0.068	0.3641	0.0054	0.722
$BOD_{5,calc}$	$BOD_{5,obs}$	0.982	<0.001	0.0589	0.763	0.997
ν_{calc}	ν_{obs}	0.0991	0.8184	0.392	0.035	0.0204

The subscripted indices indicate calculated (calc) or observed (obs) variable values.

6.3.2 Implementation of different restoration options

In terms of physical habitat condition, the Du river was considered as a natural river, without significant human impacts, physical habitat limitations such as sinuosity were thus not implemented in the WFD-Explorer interventions table.

Cleaning the point sources from extraction workshops in the Cat stream can improve quality of the section Cat stream. Wastewater discharged received a pH of 7.1 (Nguyen and Dang, 2008). River water had characteristics of upstream site 9.

The buffer zone could only reduce nutrient loads, which related to phosphorous and nitrogen concentration in water (TP and TN), and thus not provided a significant water quality improvement in the Du river (Table 6.3).

The results indicated that the helophyte filters applied for the tributaries, Main Du A and Main Du B were sufficient to meet good condition for all the water bodies due to a reduction of the BOD_5 concentration (Fig. 6.3). The model thus suggested that the installation of helophyte filters could eliminate chemical macroinvertebrate habitat problems in the studied river basin. These filters also treated agricultural phosphorus and nitrogen (Fig. 6.4, Fig. 6.5).

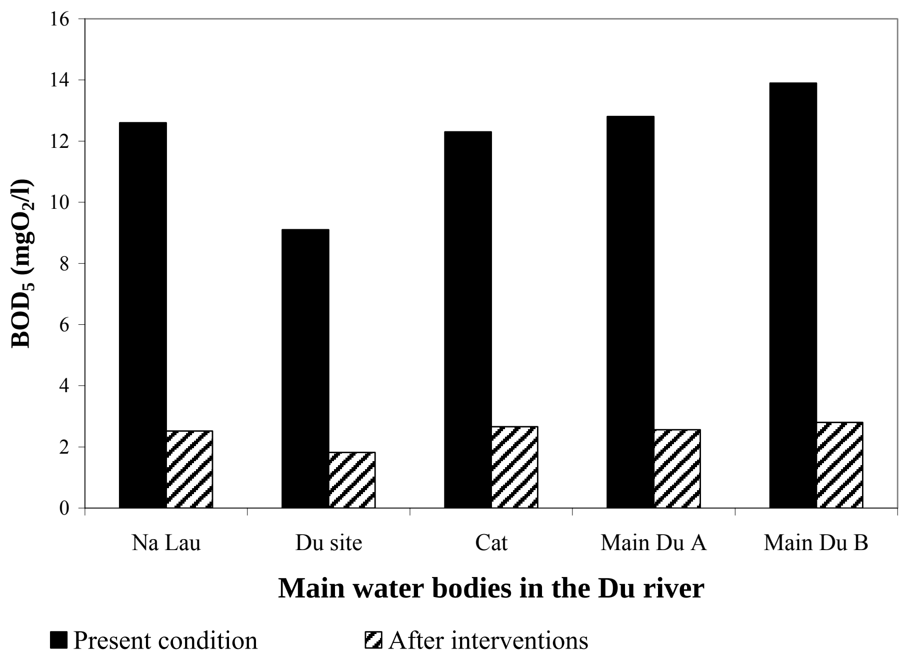


Figure 6.3. Effect of helophyte filters on BOD_5 in the Du river.

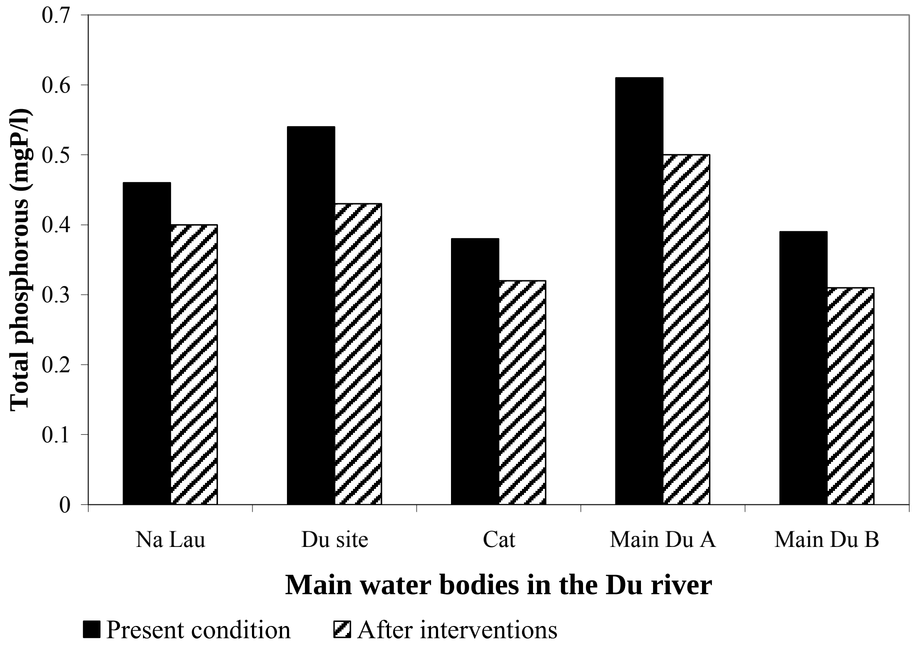


Figure 6.4. Effect of helophyte filters on total phosphorous concentrations in the Du river.

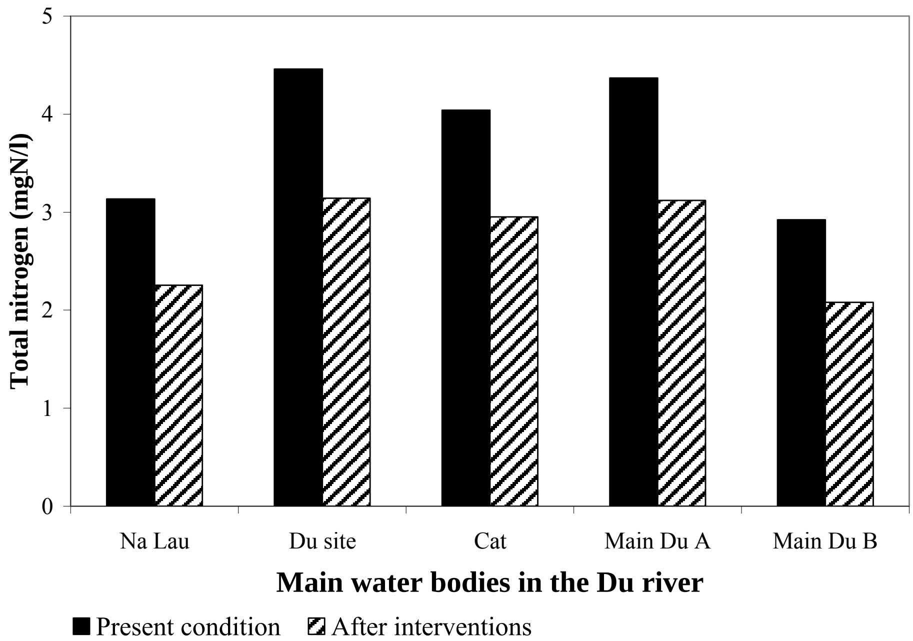


Figure 6.5. Effect of helophyte filters on total nitrogen concentrations in the Du river.

River water characteristics predicted by the WFD-Explorer were applied in the classification trees. A tress constructed with the selected variables was tested with 10 cross validation and provided good predictive performance (CCI=60% and *K*=0.41). The tree developed (Fig. 6.6) was then applied for predict quality classes of each river segment after each restoration step.

```
pH <= 5.82: VeryPoor (14.0)
pH > 5.82
|   BOD5 <= 12
|   |   BOD5 <= 8
|   |   |   Substrates = Bo_Co: Good (5.0/2.0)
|   |   |   Substrates = Gr
|   |   |   |   KjdN <= 1.98: VeryGood (2.0)
|   |   |   |   KjdN > 1.98: Good (6.0)
|   |   |   Substrates = Co_Gr: Good (3.0/1.0)
|   |   |   Substrates = Co_Si: VeryGood (5.0/1.0)
|   |   |   Substrates = Bo_Gr: VeryGood (2.0/1.0)
|   |   |   Substrates = Gr_Sa
|   |   |   |   pH <= 7.56: Moderate (4.0/1.0)
|   |   |   |   pH > 7.56: Good (3.0)
|   |   |   Substrates = Co_Sa: Good (0.0)
|   |   BOD5 > 8
|   |   |   TotalP >= 0.366315: Moderate (19.0/4.0)
|   |   |   TotalP < 0.366315: Good (17.0/6.0)
|   BOD5 > 12: Moderate (24.0/4.0)
```

Figure 6.6. Classification tree for BMWP-Viet classes developed with selected variables (PCF=0.1).

After implementation of all proposed interventions, the Du tributary presented a very good ecological quality while all other basins had good conditions (Fig. 6.9).

Table 6.3. Effect of successive restoration measures of the WFD-Explorer on BMWP-Viet values for all water bodies.

	BOD_5, mg/l	**Total P, mg/l**	**Total N, mg/l**	**BMWP-Viet**
Original EQR values before interventions				
Na Lau	12.6	0.46	3.14	Moderate
Du site	9.1	0.54	4.46	Good
Cat	3.1	0.34	3.22	Very Poor
Main Du A	12.8	0.61	4.37	Moderate
Main Du B	13.9	0.39	2.92	Moderate
Stage 1: Wastewater treatment for point sources from metal extraction workshops				
Na Lau	12.6	0.46	3.14	Moderate
Du site	9.1	0.54	4.46	Good
Cat	12.3	0.38	3.66	Moderate
Main Du A	12.8	0.61	4.37	Moderate
Main Du B	13.9	0.39	2.92	Moderate
Stage 2: Buffer strips applied in the whole Du river basin				
Na Lau	12.6	0.43	2.98	Moderate
Du site	9.1	0.48	4.14	Good
Cat	13.3	0.36	3.15	Moderate
Main Du A	12.8	0.59	4.02	Moderate
Main Du B	13.9	0.39	2.82	Moderate
Stage 3: Helophyte filters applied in the Na Lau, the Du and the Cat tributaries				
Na Lau	2.52	0.40	2.26	Good
Du site	1.82	0.43	3.14	Very Good
Cat	2.66	0.32	2.95	Good
Main Du A	7.74	0.56	3.75	Good
Main Du B	9.82	0.36	2.45	Moderate
Stage 4: Helophyte filters applied in the tributaries and the Main Du A				
Na Lau	2.52	0.4	2.26	Good
Du site	1.82	0.43	3.14	Very Good
Cat	2.66	0.32	2.95	Good
Main Du A	2.56	0.50	3.12	Good
Main Du B	7.34	0.33	2.15	Good
Stage 5: Helophyte filters applied in whole river basin				
Na Lau	2.52	0.4	2.26	Good
Du site	1.82	0.43	3.14	Very Good
Cat	2.66	0.32	2.95	Good
Main Du A	2.56	0.50	3.12	Good
Main Du B	2.80	0.31	2.08	Good

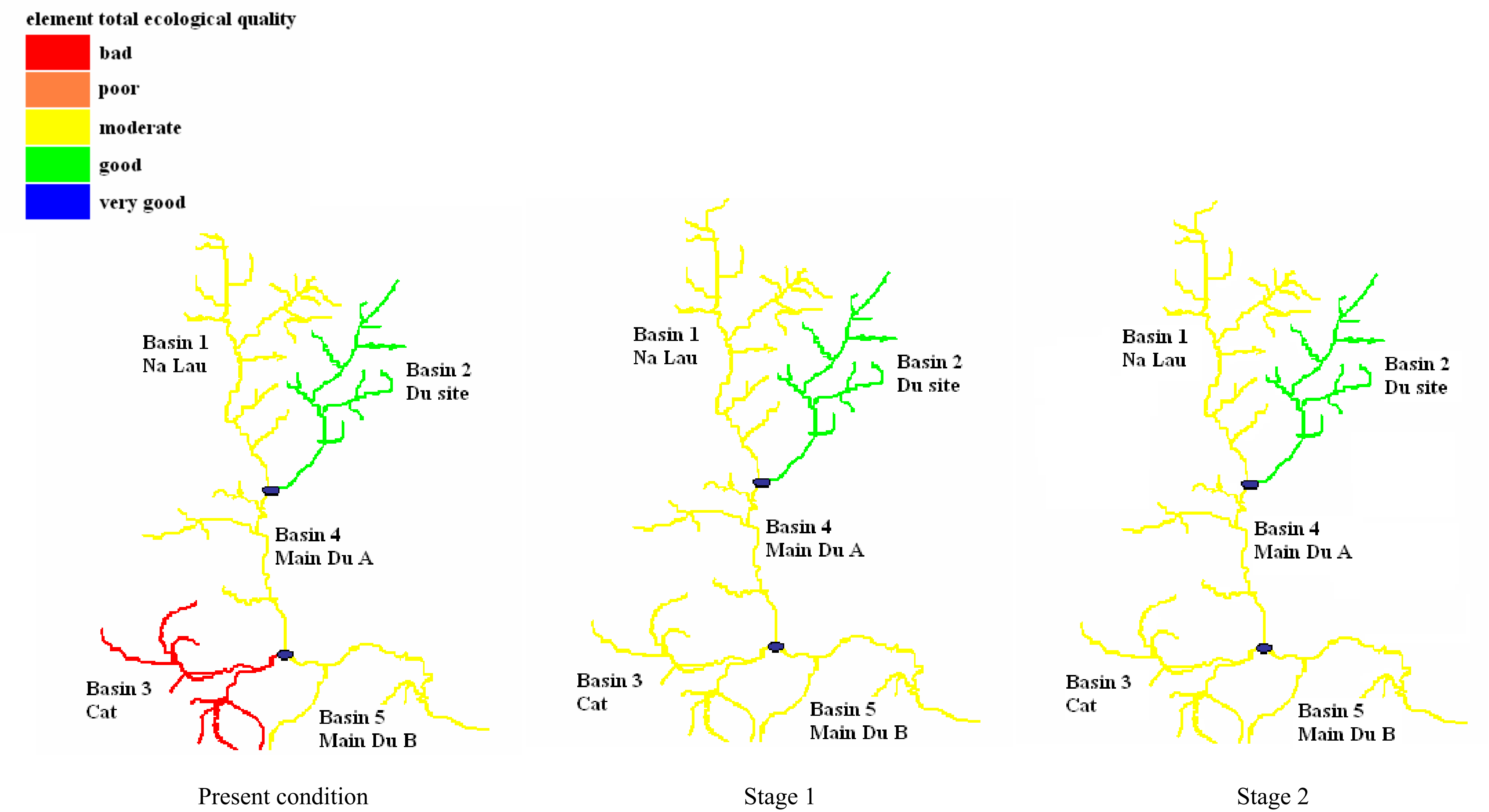


Figure 6.7. Ecological quality of the Du river at different restoration stages.

Stage 1 - Cleaning point sources in the Cat stream. Stage 2 - Buffer strips applied in the whole river basin.

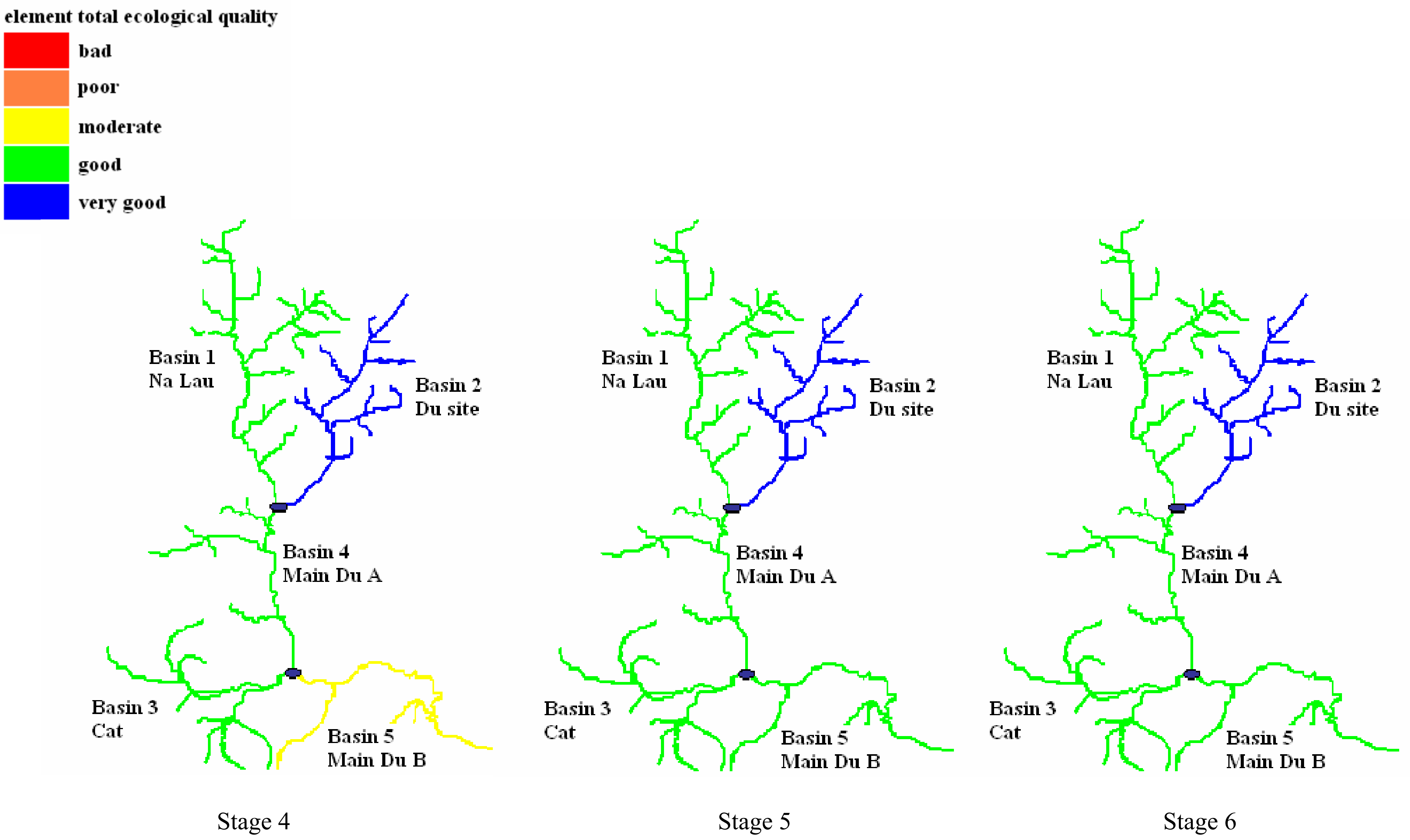


Figure 6.7. Ecological quality of the Du river at different restoration stages (continued)

Stage 3 - Helophyte filters applied in tributaries. Stage 4 - Helophyte filters applied in Main Du A. Stage 5 - Helophyte filters applied in Main Du B.

6.4 Discussion

Implementation of different restoration options was evaluated based on the results of the WFD-Explorer. Classification trees developed with selected variables showed that the BOD_5 concentration was the key factor currently determining the BMWP-Viet values in the Du river basin (Fig. 6.6). This result reflected the current situation in the river basin since it contains many household pollution sources, all of which discharge into the water bodies without any treatment. Therefore, river restoration options affecting the BOD_5 concentration in the water body were the most effective. A construction of helophyte filters as decentralised wastewater treatment systems was proposed. The WFD-Explorer revealed that this intervention significantly decreased BOD_5 values and consequently improved BMWP-Viet values. This restoration measure, applied to all water bodies, was sufficient for the achievement of a good water quality in the whole river basin.

In practical, however, helophyte filters can only be an appropriate intervention option in remote areas such as in the Na Lau, the Du tributaries and the Cat stream. Interventions on these three water bodies already significantly improved BMWP-Viet in both the main Du A and the main Du B. In the main Du water bodies, a larger population is present with residential clusters established along the main Du river A (Du town) and along the main Du river B (Giang Tien town). An appropriate restoration measure would be the collection of wastewater in a sewerage system and the construction of a medium side wastewater treatment plant. The WFD-Explorer can provide an analysis of the cost effectiveness of different restoration measures to enable the selection of an effective measure. However, the lack of input data in the study area does not allow deeper analysis in this study.

The WFD-Explorer showed that buffer strips were not an efficient restoration measure. According to the WFD-Explorer, buffer strips reduced total nitrogen and phosphorous from agriculture runoff (Table 6.3). However, these variables did not significantly affect the BMWP-Viet classification (Fig. 6.6). In practical, buffer strips not only reduce nutrient loads into the river but also create appropriate habitats for macroinvertebrate communities. Buffer trips can also reduce the loads of other potential pollutants such as pesticide residues. This information, however, was not included in the currnt version of the WFD-Explorer.

The physical habitat in the Du river basin was not disturbed significantly by impacts such as weirs for flood control, straightened river channels or artificial embankments. For all streams, a sinuosity value of at least 1.29 was calculated, indicating a high meandering pattern. However, river bed deterioration due to sand extraction observed in the main Du river can be an important reason for deteriorated macroinvertebrate habitat conditions in the river. Analysis of habitat suitability models (Chapter 5) indicated that substrate composition was one of the most important driving variables in determining the distribution of the macroinvertebrates. However, information about habitat conditions was not included in the knowledge rules of the WFD-Explorer, which can be an important limitation of the toolbox when dealing with assessment of river health based on aquatic fauna habitat conditions.

In the presented WFD-Explorer application, the two confluent water bodies A and B, always formed a third water body C if they joined. The quality of each water body affects the quality of its receiving water body (Mouton et al., 2009). However, the lack of available information in the study region hampered the data collection, resulting in a lack of important data for the implementation of the WFD-Explorer. Therefore, large water bodies have to be defined while the impacts on each water body could not be characterized into detail. Van der Most et al. (2006) mentioned the accuracy of the WFD-Explorer being strongly dependent on the level of detail and availability of the data. Moreover, different components of the error term ε_i indicating processes such as evaporation and leaching could not be quantified into detail while their implementation could play a crucial role in several transport processes. Consequently, the simulations in this study may need further calibration and the identification of missing data is an important aspect for reliable simulations and interpretation of WFD implementation impacts in the river basin of interest.

Although the WFD-Explorer needs sufficient data for more detailed simulations, this toolbox aims only to provide a general approach to ease the communication between stakeholders and water managers during WFD implementation discussions (Mouton et al., 2009). Therefore, the WFD-Explorer could still be used as a communicative tool for water managers despite the presence of roughly estimated values or processes such as evaporation, leaching, etc. During the application of the WFD-Explorer, small-scale microhabitat heterogeneity was neglected. As the impact of physical habitat changes on river biology occurs at a smaller scale (Giller and Malmqvist, 1998; Madar et al., 2008), models simulating the impact of physical habitat changes on river biology generally operate at smaller scales such as mesoscale or microscale (Mouton et al., 2009). The Du river indeed consists of several physical habitat patches, which are characterized by features such as water velocity and substrate particle size. These variables were also selected as driving variables by the classification trees and support vector machines in the presented study (Chapter 5).

Restoration options for the studied basin should not only focus on a limited number of variables (BOD_5, total nitrogen and total phosphorous), but also on the other potential pollutants. Results of the habitat suitability models presented in chapter 5 for example showed that substrate composition and other variables play a crucial role for the water quality of the Du river. However, these values were not taken into account in the knowledge rules currently available in the WFD-Explorer.

Combining the results of the WFD-Explorer simulations with habitat suitability models could improve the applicability of the toolbox in tropical conditions by avoiding the use knowledge rules for EQR developed for temperate conditions. The combination also solved the problem dealing validation with EQR, which required semi-quantitative measurement and more detailed identification of macroinvertebrate taxa.

The knowledge rule for organic pollution was developed for temperate conditions in The Netherlands and therefore needs to be adapted for further application under tropical conditions. The additional development of more knowledge rules is important to improve the reliability of the water quality assessment and the subsequent evaluation of restoration measures by the WFD Explorer. However, until today, knowledge rules of the WFD-Explorer cannot be modified by the users (Van der Most et al., 2006). Further studies therefore should be conducted to calibrate the empirical coefficients in the knowledge rules for tropical conditions.

6.5 Conclusions

This chapter aimed to evaluate the effectiveness of different restoration measure on water quality of the Du river basin by a recently developed user friendly and communicative tool, the WFD-Explorer. This tool provides an assessment of river water quality based on knowledge rules developed in The Netherlands and an evaluation of restoration options improving the water quality. The results proved the possibility to apply the tool in water management in Vietnam although the lack of reliable data in the study area significantly limited the use of the WFD-Explorer in simulating restoration measures in the river basin.

A combination of WFD-Explorer simulations with habitat suitability models may provide more reliable simulations for integrated water management. This made the WFD-Explorer more flexible in application in developing countries, where a lack of data remains the biggest constraint in integrated river management.

The WFD-Explorer is mainly an appropriate tool for modelling changes in chemical water quality and assessing the impact of these changes on aquatic macroinvertebrates (Mouton et al., 2009). Due to its user friendly interface and its clear set-up, the WFD-Explorer can enhance the dialogue on WFD implementation between scientists and water managers. Furthermore, it enables stakeholders to gain insight in how the objectives, the available measures and the impacts of these measures are interrelated, which turns it into a suitable tool for environmental decision making. However, it appears that in its current form, the WFD-Explorer is not able to deal with inorganic chemical pollutants. Consequently, more detailed models may be needed if oxygen demand and nutrient load are not the key factors determining macroinvertebrate presence. These models should also enable the integration of other factors such as physical habitat characteristics (Mouton et al., 2007), interspecific interactions (Schoener, 1974) or bio-energetic aspects (Heggenes and Saltveit, 1990; Booker et al., 2004).

Despite these limitations, the WFD-Explorer provides a general approach to ease the communication between stakeholders and water managers during WFD implementation discussions, not only in Europe but also in Vietnam.

General discussion, conclusions and future research

1 Introduction

The results of this study demonstrate that the structure of a macroinvertebrate community can reflect the state of the Du river basin in northern Vietnam. The distribution of macroinvertebrates at family level is determined by a number of environmental factors including physical, structural and chemical variables. The use of freshwater macroinvertebrates as indicator organisms to assess the environmental condition requires considerable understanding of the factors involved in determining these conditions. In contrast, research on the response of macroinvertebrate assemblages to habitat conditions can improve the understanding of environmental stress.

The aim of this last part is to link the results and discussion in the previous chapter and present some general and practical issues with regard to the development and application of techniques based on biomonitoring for decision support in water management.

2 Restoration measures in the Du river basin

2.1 Driving variables and pollution sources

Integrated application of both classical statistical analysis and modern ecological modelling can help to extract important information about pollution sources that is vital for selecting and prioritising appropriate restoration measures for rivers. The results of the present study indicated that driving variables for the distribution of macroinvertebrates in the Du river basin are (1) pH, (2) BOD_5, (3) NO_2-N, (4) total Fe and (5) DO. The condition of the river bed and turbidity were also important. These variables revealed following major pollution sources and physical degradations in the Du river basin:

- acid mine wastewater caused low pH values and high metal concentrations in the water column;
- untreated domestic wastewater discharged into the river caused excessive nutrient and BOD levels;
- runoff from agricultural land caused excessive nutrient loads;
- the river bed was destroyed by uncontrolled sand extraction at the downstream sites.

2.2 Restoration measures

For the Du river basin, the following measures are suggested to restore the degraded parts and to prevent the deterioration of the pristine upstream parts according to the findings of the study:

(1) Acid mine wastewater treatment

Prevention measures should be prioritised. Cleaner production at different stages should be introduced to mining and extraction industries to reduce the amounts of wastes discharged. Water recycling using stabilisation ponds can be efficient for mining and ore processing enterprises in midland areas.

Rainwater collection systems in the mining areas also require major attention. The rainwater running across the surfaces of mining areas flush a great amount of pollutants into the river. Rain water should also be collected into stabilisation ponds. Rainwater collection systems together with wastewater treatment by physical chemical methods, especially absorption by natural minerals, proved to be efficient for removing metals in wastewater from mining activities (Do et al., 2005; Sajidu et al., 2008).

However, technical solutions are not the most important issue. Legal sanctions are not heavy enough to force industries to apply prevention and treatment measures. Moreover, metal extraction in small workshops scattered along the river with primitive technology causes a lot of difficulties in collecting the wastewater and applying appropriate technologies for its treatment. Metal extraction should not be performed at a small scale. In the future, metal extraction workshops should be banned. Before these actions can be performed, it is, however, essential to call for the participation of workshop owners and workers and inform them about the pollution threats of wastewater and simple remediation methods such as neutralising reactions involving bicarbonate alkalinity before discharging into the environment.

(2) Domestic wastewater collection and treatment

Both on-site and concentrated treatments should be considered in dealing with domestic wastewater in the area. In upstream sites, the low population level resulted in scattered households along the river. A centralised WWTP is not suitable in these areas. The combination of on-site primary treatment in baffled septic tanks with up-flow anaerobic filters (BASTAF) and secondary treatment by simple vertical subsurface flowing constructed wetlands can significantly reduce pollution loads to the river. The BASTAF has been successfully applied in rural areas in Vietnam with average BOD removal efficiencies of 75% for small plants (Anh et al., 2002). After treatment in a BASTAF, the treated black wastewater is discharged through combined sewers with grey wastewater into the decentralised community treatment plants, which consist of imhoff tank/waste stabilisation ponds /constructed wetlands and/or other low-cost options. In the two densely populated areas in the Du town and the Giang Tien town (population over 1000 inhabitants per km^2), sewer systems together with medium-sized centralised WWTPs should be constructed for an appropriate treatment of the domestic wastewater.

(3) Riparian buffer zones

The creation of riparian buffer zones can provide effective control of erosion and nutrient run-off from intensive agriculture areas. A well maintained buffer serves as a natural filter, reducing the amount of nutrients such as N, P, sediments and chemicals that enter a stream or river and protect the river banks from erosion, especially along hills planted with tea trees located in upstream sites (Barling and Moore, 1994; Fahlen, 2002; Walter et al., 2009). Theoretically 30 m of buffer zone can reduce nitrate and phosphate content from diffused agricultural sources by about 80% (Gilliam et al., 1997). Practically, however, the appropriate buffer zone width should be carefully studied to ensure effective land use and to avoid social problems.

(4) Improvement of agricultural management practices

The most important land uses in the area are paddy fields and tea tree hills. Intensive cultivation and harvesting of tea buds and also rice around the years require large amount of nutrients. Large amounts of agrochemicals are applied in the area each year including different fertilisers and pesticides. Farmers usually apply agrochemicals on their crop lands spontaneously without proper guidelines. This often results in excessive use of fertiliser and pesticides on farming land, which is an important source of nutrient enrichment of the river water. Appropriate agricultural management practices not only reduce the loads of chemicals into the river but also increase the economical benefits of agricultural activities in the area (Rao et al., 2009). Agriculture management practices should be strongly supported by governmental authorities.

2.3 Management measures

Management issues are decisive in the effective application of restoration measures in the river. Recently, the Cau river committee, involving participations of authorities from six provinces along the Cau river, has proposed programs for restoring the Cau river basin. These programs belong to five groups of measures: (1) pollution prevention and remediation, (2) improvement of habitat quality including flow and river bed, (3) strengthening the policy framework and management capacity, (4) improving the environmental inventory and monitoring activities and (5) education and awareness raising on environmental protection.

Specific actions should focus on capacity building and establishment of a legal platform for environmental management.

Environmental planning for the river basin must be integrated with socio-economic planning. Special attention should be addressed to deforestation in upstream sites, appropriate planning for resource exploitation activities including mining and sand extraction along the river. A balance between economic benefit and environmental degradation should be carefully considered.

Serious impacts were observed from mining and extraction activities. In the Cau river basin, a restriction and prohibition of seriously polluting industrial activities has been proposed. However, implementation faces constraints dealing with both social and economic issues. Economic tools such as fees for wastewater discharge and tax incentives for enterprises applying environmental protection measures can be appropriate management tools.

Participation of the general public is essential for enhancing the effectiveness of river management and restoration programmes. Awareness raising activities to disseminate the existing environmental status and restoration plans should be carried out at different levels. The public should be the part of the policy development as well as the management activities in the rivers. The community must understand the importance of the river water quality as well as their role and therefore, they have to be adequately informed. Environmental education of the public at different levels should be an immediate action to improve the river water quality.

3 Application of bioassessment based on macroinvertebrate communities for water quality monitoring in Vietnam

Protection of the environment can only be effective if there are efficient monitoring tools in place to inform decision-makers about the condition of the environment and its trends. Bioassessment based on macroinvertebrates can be a step forward to the development and implementation of such monitoring strategies for aquatic ecosystems in Vietnam.

3.1 Application of macroinvertebrates as bioindicators

The possibility of using macroinvertebrates as bioindicators to assess the river water quality in Vietnam was clearly demonstrated in this study. It has been discussed in the previous chapters that methods based on macroinvertebrates can indeed be effective tool for biomonitoring streams and rivers in Vietnam. Their sensitivity towards natural and anthropogenic impacts can be expressed through numerous metrics, which can be used for river quality classification. In this study, the qualitative metrics BMWP, ASPT and EPT were applied to discriminate sites according to their quality. Characterisation of macroinvertebrate communities in the impacted and reference sites was essential to assess the degree of human impacts on the river and to propose restoration measures.

Although semi-quantitative sampling was recommended for use throughout the world (Barbour et al., 1999; De Pauw et al., 2006), the river bed covering mostly boulders and cobbles caused difficulties in application of kick sampling technique. Manual search by hand picking proved to be the most convenient method to collect representative taxa in the upstream sites. Since it is impossible to examine all bedrocks, abundance data therefore became unreliable for further data analysis. However, qualitative sampling loses a lot of important information to detect impairment of the sites. Decision support techniques like WFD-Explorer also require abundance data for calibration. In situ exposure of artificial substrates (De Pauw et al., 1994) can be a potential alternative for semi-quantitative sampling.

Lower than family level identification is assumed to deliver the greatest amount of information. However, identification at species level is difficult because of the small size of the organisms and a lack of adequate species-level keys and descriptions. The use of family level identification can be justified depending on the purpose of study, the level of sensitivity required and the type of index or analysis being used (Resh and McEltravy, 1993). In this study, macroinvertebrate assemblages at family level proved to be appropriate to detect pollution that may have effects on the fauna and to provide ‘early warning’ of potential problems or changes in communities. Considering a cost effective bioassessment program in Vietnam in this early stage, family level identification proved to be adequate. However, improved keys are now becoming available in Vietnam for some groups that allow identifications at lower levels.

3.2 Metric selection

In the present study, qualitative biotic indices were evaluated for the biological assessment of the Du river, including number of taxa, BMWP-Viet, ASPT-Viet and EPT taxa richness. BMWP-Viet revealed its robustness as a rapid assessment measure. The BMWP-Viet can be selected for use at an early stage. However, further studies should emphasise on improving its efficiency, accuracy, precision, predictive ability and, most important, its scoring system.

Several other biotic and diversity metrics exist, reflecting other attributes of the benthic community (e.g. structure, community balance, pollution tolerance, feeding preferences), which have been already evaluated satisfactorily in the assessment of tropical running waters (Moya et al., 2007). Indices of satisfactory performance in the assessment of temperate streams seem to be not necessarily responsive in tropical systems (Thorne and Williams, 1997; Moya et al., 2007). Besides, metrics based on functional feeding groups can also provide additional information about disturbance (Ambelu, 2009).

3.3 Technical and human resources improvements

The most important issue in biomonitoring is that the data collection should be performed according to a standardised protocol. A practical manual with detailed guidelines for biomonitoring and assessment should be developed. Guidelines should be given for (1) site selection; (2) site protocol completion; (3) sampling; (4) identification; (5) metric calculation and (6) data interpretation. Many comprehensive manuals have been developed worldwide for consultation (De Pauw and Vanhooren, 1983; Barbour et al., 1999; Coysh et al., 2000; AQEM, 2002). However, adaptation should be considered for appropriate implementation in Vietnam.

Appropriate human resources are also a key to the success of biological monitoring and assessment. Preparing academics and technicians for performance of biomonitoring should be done at national as well as provincial level. There should be training and also assessing performance of practitioners in sample collection, sorting and taxa identification.

Higher education and research institutions should work closely in both research and training activities for improvement of monitoring techniques, data interpretation methods and education for implementation of the biomonitoring.

3.4 Management issues

Fig. 7.1 provides a scheme describing the process of biological monitoring and assessment of rivers in a number of steps.

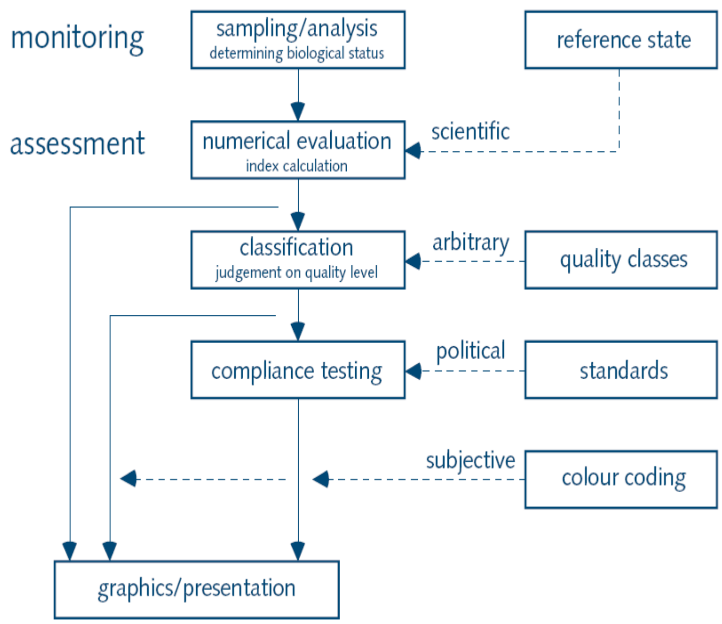


Figure 1. Elements of biological assessment methods (UN/ECE, 1995).

The scheme indicates that there are great contributions of biomonitoring and assessment in river water management, in which a close co-operation is required between the management authorities and the scientific research institutions.

Bioassessment should be enclosed in the national monitoring programme together with the conventional on-going monitoring programme. The biomonitoring programmes based on macroinvertebrates should be further developed and implemented for all river basins.

4 Application of ecological models for decision support in river management

The results of the study have shown that ecological informatics can successfully and meaningfully be applied to analyse causal relationships including identification and assessment of complex impact factors and prediction of system behaviour. Particularly, ecological informatics have advantages if relationships are unknown, very complex or nonlinear, which is typical for river and stream ecosystems.

4.1 Habitat suitability models

In recent years, modern techniques involving artificial intelligence techniques have surged for application in environmental management. Classification trees and support vector machines are popular techniques now applied in ecological modelling. The potential of predictive models in river management for the Du river basin case were clearly illustrated in this study. Based on a classification tree constructed for the identification of sites under high impact of human activities, it was possible to determine the major stressors requiring attention of river managers. Support vector machines proved to be even more powerful in dealing with large databases containing biological and environmental information of the basin. Both techniques also proved their robustness in predicting the presence/absence of indicator taxa in the river. In this manner, the use of these techniques can be included in decision support systems, helping river managers in the selection of restoration/protection options.

4.2 WFD-Explorer for selection of restoration measures

To improve the reliability and efficiency of management actions in the future, decision support systems enabling cost-benefit analysis should play an important role. However, when constructing such systems, careful attention needs to be paid to the possibilities and limitations of the available water system models and how they can be integrated in a user friendly simulation shell (Goethals, 2005). For the latter aspect, the participation of manager and stakeholder is of crucial importance.

The WFD-Explorer is a user friendly, communicative tool that provides an evaluation of restoration options improving the water quality. The WFD-Explorer is an appropriate tool for modelling changes in physical chemical water quality and assessing the impact of these changes on aquatic macroinvertebrates (Mouton et al., 2009). Due to its user friendly interface and its clear set-up, the WFD-Explorer can enhance the dialogue on water management implementations between scientists and water managers.

Although some adaptations should be considered for use in tropical conditions, the general approach of the WFD-Explorer is a useful platform for communication between stakeholders and water managers. Combining the results of the WFD-Explorer simulations with those of more detailed physical habitat models, such as classification trees and support-vector machines, may lead to more reliable simulations useful for integrated water management.

5 Future research

Rarefaction analysis of macroinvertebrate data can be a useful tool for designing a monitoring programme (Ambelu, 2009). This kind of analysis provides information on the appropriate number of sampling sites and the number of sampled individuals required to obtain reliable data.

The scoring system is the core for the calculating biotic indices. This study showed that the current scoring systems adapted from the original BMWP from the UK proved to be inadequate for the conditions of Vietnam. In the near future, a revision of scores for several taxa as well as adding scores for other taxa, which are not included in the BMWP-Viet, should be done to improve the application of the BMWP-Viet scoring system in Vietnam.

Further studies evaluating the performance of additional metrics for the assessment of the river quality in Vietnam could provide better information for the development of a multimetric index. The major advantage of index based on several metrics would be the robust and sensitive insights into how the assemblages respond to natural and anthropogenic stressors.

Another important question for river management is how to determine the threshold level for both effluents discharged into a surface water body as well as the concentration of elements in the aquatic ecosystem in order to maintain and improve its ecological health. Therefore, beside the surface water quality standards (TCVN 5942-1995), which are being applied for quality control of the surface water by specifying parameter limits and maximum allowable concentrations of pollutants in surface water, it would be very useful to have a water quality guideline for rivers and streams to protect aquatic community.

Modern techniques of data mining proved to be useful in the analysis of environmental data. Improving performance and application of classification trees and support vector machines could assist decision makers during the evaluation of the impacts and during the development of integrated and sustainable river management plans.

The application of the WFD-Explorer requires further research for application in river management in Vietnam. Knowledge rules adapted from the Netherlands with many empirical coefficients may not be appropriate for the tropical conditions in Vietnam. However, the methodology for knowledge rule development could be useful to discover region specific empirical coefficients. Further research about the mass balance would also be useful to improve the reliability and could be extended with other substances to allow a more detailed characterization of the chemical water quality in the studied river basin. Additional knowledge rules should be included in the WFD-Explorer toolbox to link the Ecological Quality Ratio with the relevant variables, which can be found by the results of ecological modelling techniques such as classification trees and support vector machines.

6 General conclusions

The aim of this thesis was to develop biological monitoring and assessment techniques for decision support in river management in northern Vietnam. Several conclusions as well as contributions of this thesis can be summarised as follows:

- Information on riverine macroinvertebrate communities in the rivers in northern Vietnam was collected to improve a source of data available in this region. BMWP-Viet biological index based on collected data proved to be available for evaluation of deterioration in the rivers in northern Vietnam, although further research is required to improve its scoring system.
- Conventional statistical analyses as well as different data mining techniques were performed and compared. The study revealed that the developed models using SVMs provided more stable results and better performance than CTs. Data mining techniques indicated the availability to get insights in the relations between river characteristics and the inhabiting biology.
- Combination of data-driven and knowledge-based models was developed and applied to improve the robustness of the both techniques. Together with this combination, the WFD-Explorer could be a useful tool to compare effectiveness of restoration measures in order to select appropriate measures.
- Bioassessment together with ecological informatics can effectively supports river management in Vietnam.

Development of appropriate and effective biological monitoring and assessment is a on-going process. It requires not only contributions and active participations of scientists and river managers, but also involvement of general public. The present research may constitute a step towards the development and implementation of biological monitoring and assessment based on macroinvertebrates for aquatic ecosystems in Vietnam.

References

Adriaenssens, V. (2004). *Knowledge-based macroinvertebrate habitat suitability models for use in ecological river management*. PhD thesis, Ghent University, Gent, Belgium, 296pp.

Adriaenssens, V., De Baets, B., Goethals, P.L.M. and De Pauw, N. (2004a). Fuzzy rule-based models for decision support in ecosystem management. *The Science of the Total Environment,* 319: 1-12.

Adriaenssens, V., Goethals, P.L.M., Charles, J. and De Pauw, N. (2004b). Application of Bayesian Belief Networks for the prediction of macroinvertebrate taxa in rivers. *Annales De Limnologie-International Journal of Limnology,* 40(3): 181-191.

Adriaenssens, V., Simons, F., Nguyen, L.T.H., Goddeeris, B., Goethals, P. and De Pauw, N. (2004c). Potential of bioindication of chironomid communities for assessment of running water quality in Flanders (Belgium). *Belgian Journal of Zoology,* 134: 15-24.

Ahmad, A., Maimon, A., Othman, M. S. and Mohd Pauzi, A. (2002). *The potential of local benthic macroinvertebrates as a biological monitoring tool for river water quality assessment* Regional Symposium on Environment and Natural Resources, Kuala Lumpur, Malaysia: 464-471.

Akolkar, P., Agrawal, S. and Trivedi, R.C. (2008). *Biological monitoring of water quality in India - needs and constraints*. Scientific Conference 'Rivers in the Hindu Kush Himalaya. Ecology and Environmental Assessment', ASSESS-HKH project, Kathmandu and Dhulikhel, Nepal: 125-130.

Alan, D.Z., Tran, T.L., Giambelluca, T.W., Sidle R.C., Sutherland, R. A., Nullet, M.A. and Tran, D.V. (2006). Effective slope lengths for buffering hillslope surface runoff in fragmented landscapes in northern Vietnam. *Forest Ecology and Management*, 224: 104–118.

Alba-Tercedor, J. and Sanchez-Ortega, A. (1988). Un metodo rapido y simple para evaluar la calidad biologica de las aguas corrientes basado en de Hellawell, 1978. *Limnetica*, 4: 51-56.

Alba-Tercedor, J. and Pujante, A.M. (2000). Running-water biomonitoring in Spain: opportunities for predictive approach. 207-216. In: Wright, J. F., Sutcliffe, D. W. and Furse, M. (Eds.), *Assessing the biological quality of freshwater: RIVPACS and other techniques*. Freshwater Biological Association, Ambleside, Cumbria, UK.

Allan, J.D. (1995). *Stream ecology: Structure and Function of Running Water*. Chapman and Hall, London, UK, 388pp.

Al-Shemmeri, T., Al-Kloub, B. and Pearman, A. (2001). Computer aided decision support system for water strategic planning in Jordan. *Forest Ecology and Management*, 143: 65-76.

Ambelu, A. (2009). *Biological monitoring based on macroinvertebrates for decision support of water management in Ethiopia*. PhD thesis, Ghent University, Gent, Belgium, 185pp.

Andreu, J., Capilla, J. and Sanchis, E. (1995). AQUATOOL, a generalized decision-support system for water-resources planning and operational management. *Journal of Hydrology*, 177: 269-291.

Anh, N.V., Ha, T.D., Nhue, T.H., Heinss, U., Morel, A., Moura, M. and Schertenleib, R. (2002). *Decentralized wastewater treatment - new concept and technologies for Vietnamese conditions*. 5th Specialised Conference on Small Water and Wastewater Treatment Systems, Istanbul, Turkey.

Anh, N. V., Pham, N. T., Nguyen, T. H., Morel, A. and Tonderski, K. (2007). *Improved septic tank with constructed wetland, a promising decentralized wastewater treatment alternative in Vietnam*. NOWRA 16th Annual Technical Education Conference & Exposition, Norway: Paper XI-RCS-07-30.

AQEM (2002). *A comprehensive method to assess European streams using benthic macroinvertebrates, developed for the purpose of the Water Framework Directive. Version 1.0*. www.aqem.de, 198pp.

Armitage, P.D., Moss, D., Wright, J.F. and Furse, M.T. (1983). The performance of a new biological water-quality score system based on macroinvertebrates over a wide-range of unpolluted running water sites *Water Research*, 17: 333-347.

Arunachalam, M., Madhusoodanan, K.C.M., Vijverberg, J., Kortmulder, K. and Suriyanarayanan, H. (1991). Substrate selection and seasonal variation in densities of invertebrates in stream pools of a tropical river. *Hydrobiologia*, 213: 141-148.

Asefa, T., Kemblowski, M., Urroz, G. and McKee, M. (2005). Support vector machines (SVMs) for monitoring network design. *Ground Water*, 43: 413-422.

Assaf, H. and Saadeh, M. (2008). Assessing water quality management options in the Upper Litani Basin, Lebanon, using an integrated GIS-based decision support system. *Environmental Modelling & Software*, 23: 1327-1337.

Astorga, Y., De Pauw, N. and Persoone, G. (1997). *Development and application of cost effective methods for biological monitoring of rivers in Costa Rica. Final report, Joint research European Union Project No NCI1* CT-92-0094.* European Community.

AUSRIVAS. (2005). *AUSRIVAS bioassessment: macroinvertebrates.* . Accessed April 2009, from http://ausrivas.canberra.edu.au/Bioassessment/Macroinvertebrates/

Austin, M.P. (1998). An ecological perspective on biodiversity investigations: examples from Australian eucalypt forests. *Annals of the Missouri Botanical Garden*, 85: 2-17.

Austin, M.P. (2002). Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling*, 157: 101-118.

Azrina, M.Z., Yap, C.K., Ismail, A.R., Ismail, A. and Tan, S.G. (2006). Anthropogenic impacts on the distribution and biodiversity of benthic macroinvertebrates and water quality of the Langat River, Peninsular Malaysia. *Ecotoxicology and Environmental Safety*, 64: 337-347.

Bae, Y.J. and Lee, B.H. (2001). Human impacts on stream ecosystems and freshwater arthropods in Korea. *Korean Journal of Entomology*, 31: 63-76.

Bae, Y.J., Kil, H.K. and Bae, K.S. (2005). Benthic macroinvertebrates for uses in stream biomonitoring and restoration. *KSCE Journal Civil Engineering*, 9: 55-63.

Bai, J.H., Cui, B.S., Xu, X.F., Ding, Q.Y. and Gao, H.F. (2009). Heavy Metal Contamination in Riverine Soils Upstream and Downstream of a Hydroelectric Dam on the Lancang River, China. *Environmental Engineering Science*, 26: 941-946.

Bakir, H. (2001). *Wastewater management services in water stressed countries: guiding principles and options for sustainable development*. The 2nd Asian conference of water and wastewater management WHO Regional Centre for Environmental Health Activities (CEHA), Tehran, Iran.

Baptista, D.F., Dorville, L.F.M., Buss, D.F. and Nessiamian, J. L. (2001). Spatial and temporal organization of aquatic insects assemblages in the longitudinal gradient of a tropical river. *Revista Brasileira de Biologia*, 61: 295-304.

Baptista, D.F., Buss, D.F., Egler, M., Giovanelli, A., Silveira, M.P. and Nessimian, J.L. (2007). A multimetric index based on benthic macroinvertebrates for evaluation of Atlantic Forest streams at Rio de Janeiro State, Brazil. *Hydrobiologia*, 575: 83-94.

Barbour, M.T., Plafkin, J.L., Bradley, B.P., Graves, C.G. and Wisseman, R.W. (1992). Evaluation of EPA's Rapid Bioassessment benthic metrics: metric redundancy and variability among reference stream sites. *Environmental Toxicology and Chemistry*, 11: 437-449.

Barbour, M.T., Gerritsen, J., Snyder, B.D. and Stribling, J.B. (1999). *Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish.* . U.S. Environmental Protection Agency; Office of Water, Washington, D.C., USA, 339pp.

Barbour, M.T. and Yoder, C.O. (2000). The multimetric approach to bioassessment, as used in the United States of America. 281-292. In: Wright, J. F., Sutcliffe, D. W. and Furse, M. T. (Eds.), *Assessing the biological quality of fresh waters. RIVPACS and other techniques.* Freshwater Biological Association, Cumbria, UK.

Barling, R.W. and Moore, I.D. (1994). Role of buffer strips in management of waterway pollution: a review. *Environmental Management*, 18(53-558).

Beck, W.M.J. (1954). Studies in stream pollution biology, I: a simplified ecological classification of organisms. *Quarterly Journal of the Florida Academy of Sciences*, 17: 211–227.

Bishop, J.E. (1973). *Limnology of a small Malayan river, Sungai Gombak*. Junk Publishers, Hague, The Netherlands, 485pp.

Bispo, P.C., Oliveira, L.G., Bini, L.M. and Sousa, K.G. (2006). Ephemeroptera, Plecoptera and Trichoptera assemblages from riffles in mountain streams of Central Brazil: environmental factors influencing the distribution and abundance of immatures. *Brazilian Journal of Biology*, 66: 611-622.

Bij de Vaate, A.B. and Pavluk, T.I. (2004). Practicability of the Index of Trophic Completeness for running waters. *Hydrobiologia*, 519: 49-60.

Bonada, N., Prat, N., Resh, V. H. and Statzner, B. (2006). Developments in aquatic insect biomonitoring: A comparative analysis of recent approaches. *Annual Review of Entomology*, 51: 495-523.

Booker, D.J., Dunbar, M.J. and Ibbotson, A. (2004). Predicting juvenile salmonid drift-feeding habitat quality using a three-dimensional hydraulic bioenergetic model. *Ecological Modelling*, 177: 157–177.

Boonsoong, B., Sangpradub, N. and Barbour, M.T. (2009). Development of rapid bioassessment approaches using benthic macroinvertebrates for Thai streams. *Environmental Monitoring and Assessment*, Online publication date: 17-Aug-2008: DOI 10.1007/s10661-008-0423-2.

Bouchard, R.W.J. (2004). *Guide to aquatic macroinvertebrates of the Upper Midwest*. Water Resources Center, University of Minnesota, St. Paul, USA, 208pp.

Bray, J.R. and Curtis, J.T. (1957). An ordination of the upland forest communities of Southern Wisconsin. *Ecological Monographs*, 27: 325-349.

Breiman, L., Friedman, J.H., Olshen, A. and Stone, C.G. (1984). *Classification and regression trees*. Wadsworth International group, Belmon, USA.

Budin, K., Ahmed, A., Abdullah, N. and Dawalih, M. (2007). *Correlation analysis on water quality parameter with aquatic insects abundance in Telipok River, Sabah, Malaysia*. 12th WSEAS International Conference on Applied Mathematics, Cairo, Egypt: 324-327.

Buffagni, A., Erba, S., Birk, S., Cazzola, M., Feld, C., Ofenböck, T., Murray-Bligh, J., Furse, M.T., Clarke, R., Hering, D., Soszka, H. and van de Bund, W. (2005). *Towards European intercalibration for the water framework directive: procedures and examples for different river types from the E.C.* project STAR. IRSA, Rome, Italy, 468pp.

Burges, C.J.C. (1998). A tutorial on Support Vector Machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2: 121-167.

Buss, D.F., Baptista, D.F., Silveiral, M.P., Nessimian, J.L. and Dorvill'e, L.F.M. (2002). Influence of water chemistry and environmental degradation on macroinvertebrate assemblages in a river basin in south-east Brazil. *Hydrobiologia*, 481: 125–136.

Butcher, J.T., Stewart, P.M. and Simon, T.P. (2003). Effects of two classification strategies on a Benthic Community Index for streams in the Northern Lakes and Forests Ecoregion. *Ecological Indicators* 3: 195-202.

Cairns, J.J., McCormick, P.V. and Niederlehner, B.R. (1993). A proposed framework for developing indicators of ecosystem health. *Hydrobiologia*, 263: 1-44.

Campbell, I.C., Barlow, C.G. and Pham, G.H. (2005). Managing the ecological health of the Mekong River: evaluating threats and formulating responses. *Verhandlungen der Internationale Vereinigung fu¨r Theoretische und Angewandte Limnologie*, 29: 497–500.

Cao, Y., Williams, W.P. and Bark, A.W. (1997). Effects of sample size (number of replicates) on similarity measures in river Aufwuchs community analysis. *Water Environment Research*, 69: 107-114.

Cao, T.K.T., Ham, S.A. and Bae, Y.J. (2007). Description of three new species of Neoperla (Plecoptera: Perlidae) and a historical review of tropical Southeast Asian Perlidae. *Zootaxa* 1453: 41-54.

Cao, T.K.T. (2008). *Systematics of the Vietnamese Perlidae (Insecta: Plecoptera)*. PhD thesis, Seoul Women's University, Seoul, Korea, 297pp.

Cao, T.K.T. and Bae, Y.J. (2008). Neoperlops vietnamellus, a rare stonefly from Vietnam (Plecoptera: Perlidae). *Zootaxa*, 1968: 33-38.

Carter, J.L. and Resh, V.H. (2001). After site selection and before data analysis: sampling, sorting, and laboratory procedures used in stream benthic macroinvertebrate monitoring programs by USA state agencies. *Journal of the North American Benthological Society*, 20: 658-682.

Chandler, J.R. (1970). A biological approach to water quality management. *Water Pollution Control*, 69: 415-422.

Che-Salmah, M.R. and Abu-Hassan, A. (2005). Aquatic insects of a peat swamp river, Sungei Bebar, PahangIn: Latiff, A., Hamzah, K. A. and Ahmad, N. (Eds.), *Biodiversity expedition Sungai Bebar, Pekan, Pahang*. Forest Research Institute of Malaysia, Selangor, Malaysia.

Chessman, B.C. (1995). Rapid assessment of rivers using macroinvertebrates. A procedure based on habitat-specific sampling, family level identification and a biotic index *Australian Journal of Ecology*, 20: 122-129.

Chessman, B.C. (2003a). *SIGNAL 2 - A scoring system for macroinvertebrate in Australian rivers*. Commonwealth of Australia, Canberra, Australia, 32pp.

Chessman, B.C. (2003b). New sensitivity grades for Australian river macroinvertebrates. *Marine and Freshwater Research*, 54: 95-103.

Chon, T.-S., Park, Y.S., Moon, K.H. and Cha, E.Y. (1996). Patternizing communities by using artificial neural network. *Ecological Modelling*, 90: 69-78.

Chutter, F.M. (1994). The rapid biological assessment of streams and river water quality by means of macroinvertebrate communities in South Africa. 217-234. In: Uys, M. C. (Eds.), *Classification of rivers and Environmental Health Indicators.* . Water research commission Report No TT 63/94, South Africa.

Cidu, R., Caboi, R., Fanfani, L. and Frau, F. (1997). Acid drainage from sulfides hosting gold mineralization (Furtei, Sardinia). *Environmental Geology*, 30: 231-237.

Clarke, K.R. and Warwick, R.M. (2001) *Change in Marine Communities: An Approach to Statistical Analysis and Interpretation. 2nd edition*. PRIMER-E, Plymouth, UK, 172pp.

Clarke, R.T., Furse, M.T., Gunn, R.J.M., Winder, J.M. and Wright, J.F. (2002). Sampling variation in macroinvertebrate data and implications for river quality indices. *Freshwater Biology*, 47: 1735-1751.

Clem, A. (2007). *Map of East Asia with the most important rivers*. http://www.faculty.sbc.edu/aclem/EastAsianPolitics.html, Accessed 10th May 2009.

Cook, S.E.K. (1976). Quest for an index of community structure sensitive to water pollution. *Environmental Pollution*, 11: 269-287.

Cortes, C. and Vapnik, V. (1995). Support-Vector Networks. *Machine Learning*, 20(3): 273-297.

Courtemanch, D.L. and Davies, S.P. (1987). A coefficient of community loss to assess detrimental change in aquatic communities. *Water Research*, 21: 217–222.

Covich, A.P. (1988). Geographic and historical comparisons of neotropical streams: biotic diversity and detrital processing in highly variable habitats. *Journal of the North American Benthological Society*, 7: 361–386.

Coysh, J., Nichols, S., Simpson, S., Norris, R., Barmuta, L., Chessman, B. and Blackman, P. (2000). *Australian River Assessment System - National River Health Program Predictive Model Manual.* CRC Freshwater Ecology, University of Canberra, Canberra, Australia. Available http://ausrivas.canberra.edu.au, 20pp.

Cranston, P.S., Fairweather, P. and Clarke, G. (1996). Biological indicators of water quality. 143-154. In: Walker, J. and Reuter, D. J. (Eds.), *Indicators of Catchment Health: A Technical Perspective.* CSIRO Publishing, Melbourne, Australia.

Czerniawska-Kusza, I. (2005). Comparing modified biological monitoring working party score system and several biological indices based on macroinvertebrates for water-quality assessment. *Limnologica*, 35(3): 169-176.

Dakou, E., D'heygere, T., Dedecker, A.P., Goethals, P., Lazaridou-Dimitriadou, M. and De Pauw, N. (2007). Decision tree models for prediction of macroinvertebrate taxa in the river Axios (Northern Greece). *Aquatic Ecology*, 41: 399–411.

Dallas, H.F. and Day, J.A. (2007). Natural variation in macroinvertebrate assemblages and the development of a biological banding system for interpreting bioassessment data - a preliminary evaluation using data from upland sites in the south-western Cape, South Africa.

Hydrobiologia, 575: 231-244.

Dang, N.T., Ho, T.H., Duong, D.T. and Mai, D.Y. (2002). *Aquatic biology of the inland freshwater ecosystem in Vietnam (In Vietnamese)*. Science and Technical Pulblisher, Hanoi, Vietnam, 399pp.

Dang, K.C., Nguyen, N.L. and Tran, L.M. (2005). *Craft villages in Vietnam and the environment*. Scientific and Technical Publisher, Hanoi, Vietnam, 392pp.

Davidson, S.P., Kunpradid, T., Peerapornisal, Y., Nguyen, T.M.L., Pathoumthong, B., Vongsambath, C. and Pham, A.D. (2006). *Biomonitoring of the Lower Mekong and selected tributaries. MRC Technical Paper 13*. Mekong River Commission, 106 pp Vientiane.

Davies, P.E. (2000). Development of a national river bioassessment system (AUSRIVAS) in Australia. 113-124. In: Wright, J. F., Sutcliffe, D. W. and Furse, M. T. (Eds.), *Assessing the biological quality of fresh waters: RIVPACS and other techniques.* . Freshwater Biological Association, Ambleside, Cumbria, UK.

Davis, W.S. and Simon, T.P., Eds. **(1995)**. *Biological Assessment Criteria: Tools for Water Resource Planning and Decision Making*. Lewis Publisher., Boca Raton, USA.

Davis, N.M., Weaver, V., Parks, K. and Lydy, M.J. (2003). An assessment of water quality, physical habitat, and biological integrity of an urban stream in Wichita, Kansas, prior to restoration improvements (phase I). *Archives of Environmental Contamination and Toxicology*, 44: 351-359.

De Klein, J.J.M. (2008). *From Ditch to Delta, Nutrient retention in running waters.* PhD thesis, Wageningen University, Wageningen, The Netherlands, 194pp.

De Pauw, N. and Vanhooren, G. (1983). Method for biological quality assessment of watercourses in Belgium. *Hydrobiologia*, 100: 153-168.

De Pauw, N., Ghetti, P.F., Manzini, P. and Spaggiari, P. (1992). Biological assessment methods for running water. 217-248. In: Newman, P., Piavaux, A. and Sweeting, R. (Eds.), *River water quality – assessment and control. Commission of the European Communities*. EUR 14606 EN-FR, 1992-III, Brussels, Belgium.

De Pauw, N. and Hawkes, H.A. (1993). Biological monitoring of river water quality. 161-172. In: Walley, W. J. and Judd, S. (Eds.), *Proceedings of the Freshwater Europe Symposium on River Water Quality Monitoring and Control*, Birmingham, UK.

De Pauw, N., Lambert, V., Van Kenhove, A. and bij de Vaate, A. (1994). Performance of two artificial substrate samplers for macroinvertebrates in biological monitoring of large and deep rivers and canals in Belgium and The Netherlands. *Environmental Monitoring and Assessment*, 30: 25-47.

De Pauw, N. (2000). Using RIVPACS as a modeling tool to predict the impacts of environmental changes. 311-314. In: Wright, J. F., Sutcliffe, D. W. and Furse, M. (Eds.), *Assessing the biological quality of freshwater: RIVPACS and other techniques*. Freshwater Biological Association, Ambleside, Cumbria, UK.

De Pauw, N., Gabriels, W. and Goethals, P. (2006). River monitoring and assessment methods based on macroinvertebrates. 113-134 In: Ziglio, G., Siligardi, M. and Flaim, G. (Eds.), *Biological monitoring of rivers. Application and perspectives*. John Wiley & Sons, UK.

Bij de Vaate, A.B. and Pavluk, T.I. (2004). Practicability of the Index of Trophic Completeness for running waters. *Hydrobiologia*, 519: 49-60.

De Zwart, D. and Trivedi, R.C. (1994). *Manual on integrated water quality evaluation.* National Institute of Public Health and Environmental Protection (RIVM), Bilthoven, The Netherlands.

Dean, J.C. (1996). Prosopistomatidae. 76. In: Dean, J. C. and Suter, P. J. (Eds.), *Mayfly Nymphs of Australia. A guide to Genera.* CRC for Freshwater Ecology, Identification Guide No. 7, Albury, NSW, Australia.

Decoste, D. and Scholkopf, B. (2002). Training invariant support vector machines. *Machine Learning*, 46: 161-190.

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

Dedecker, A., Van Melckebeke, K., Goethals, P. and De Pauw, N. (2007). Development of migration models for macroinvertebrates in the Zwalm river basin (Flanders, Belgium) as tools for restoration management. *Ecological Modelling*, 203: 72-86.

Devesa, F., Comas, J., Turon, C., Freixo, A., Carrasco, F. and Poch, M. (2009). Scenario analysis for the role of sanitation infrastructures in integrated urban wastewater management. *Environmental Modelling & Software*, 24(3): 371-380.

D'heygere, T., Goethals, P. and De Pauw, N. (2002). Optimisation of the monitoring strategy of macroinvertebrate communities in the river Dender, in relation to the EU Water Framework Directive. *The Scientific World Journal*, 2: 607-617.

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

D'heygere, T., Goethals, P.L.M. and De Pauw, N. (2006). Genetic algorithms for optimisation of predictive ecosystem models based on decision trees and neural networks. *Ecological Modelling*, 195: 20-29.

Dickens, C.W.S. and Graham, P.M. (2002). The South African Scoring System (SASS) Version 5 Rapid Bioassessment Method for Rivers. *African Journal of Aquatic Science*, 27: 1-10.

Dinakaran, S. and Anbalagan, S. (2007). Anthropogenic impacts on aquatic insects in six streams of south Western Ghats. *Journal of Insect Science*, 7: No.37.

Dines, R. A. and Murray-Bligh, J. (2000). Quality assurance and RIVPACS. 71-78. In: Wright, J. F., Sutcliffe, D. W. and Furse, M. (Eds.), *Assessing the biological quality of freshwater: RIVPACS and other techniques*. Freshwater Biological Association, Ambleside, Cumbria, UK.

Do, T.A. and Dang, K.C. (2005). Exploring application of Chitosan originated from aquatic product processing waste for removing heavy metals from industrial wastewater. *Interuniversity Journal of Science and Technology*, 35: 112-116.

Dominguez-Granda, L.E. (2008). *Indices based on macroinvertebrate communities for assessment of the quality of the Chaguana river in Ecuador*. PhD thesis, Ghent University, Gent, Belgium, 233pp.

Dudgeon, D. and Bretschko, G. (1996). Allochthonous inputs and land-water interactions in seasonal streams: tropical Asia and temperate Europe. 161-179. In: Schiemer, F. and Boland, K. T. (Eds.), *Perspectives in tropical limnology*. SPB Academic Publishing, Amsterdam, The Netherlands.

Dudgeon, D. (1999). *Tropical Asian streams. Zoobenthos, ecology and conservation*. Hongkong University press, Hongkong, 830pp.

Dudgeon, D. (2000). Riverine biodiversity in Asia: a challenge for conservation biology. *Hydrobiologia*, 418: 1-13.

Dudgeon, D. (2003). The contribution of scientific information to the conservation and management of freshwater biodiversity in tropical Asia. *Hydrobiologia*, 500: 295-314.

Dudgeon, D. (2006). The impacts of human disturbance on stream benthic invertebrates and their drift in North Sulawesi, Indonesia. *Freshwater Biology*, 51: 1710-1729.

Dzeroski, S., Grbovic, J. and Walley, W.J. (1997). Machine learning applications in biological classification of river water quality. 429-448. In: Michalski, R. S., Bratko, I. and Kubat, M. (Eds.), *Machine learning and data mining: methods and applications*. John Wiley and Sons Ltd., New York, USA.

Dzeroski, S., Demsar, D. and Grbovic, J. (2000). Predicting chemical parameters of river water quality from bioindicator data. *Applied Intelligence*, 13: 7-17.

EU (2000). *Directive of the European Parliament and of the Council 2000/60/EC establishing a framework for community action in the field of water policy. Rep. No. PECONS 3639/1/00 REV 1*. Union, E., Luxembourg.

Fahlen, A. (2002). Mixed tree-vegetative barrier designs: Experiences from project works in northern Vietnam. *Land Degradation & Development*, 13: 307-329.

Fenoglio, S., Badino, G. and Bona, F. (2002). Benthic macroinvertebrate communities as indicators of river environment quality: an experience in Nicaragua. *Revista De Biologia*

Tropical, 50: 1125-1131.

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

Fitzpatrick, M., Weltzin, J., Sanders, N. and Dunn, R. (2007). The biogeography of prediction error: why does the introduced range of the fire ant over-predict its native range. *Global Ecology and Biogeography* 16: 24–33.

Flecker, A.S. and Feifarek, B.P. (1994). Disturbance and the temporal variability of invertebrate assemblages in two Andean streams. *Freshwater Biology*, 31: 131-142.

Fore, L.S., Karr, J.R. and Wisseman, R.W. (1996). Assessing invertebrate responses to human activities: Evaluating alternative approaches *Journal of the North American Benthological Society*, 15: 212-231.

Fu, Z., Golden, B.L., Lele, S., Raghavan, S. and Wasel, E.A. (2003). A Genetic Algorithm-Based Approach for Building Accurate Decision Trees. *Informs Journal on Computing*, 15: 3-22.

Furse, M.T., Hering, D., Brabec, K., Buffagni, A., Sandin, L. and Verdonschot, P.F.M. (2006). The ecological status of European rivers: evaluation and intercalibration of assessment methods. *Hydrobiologia*, 566: 1-2.

Gabriels, W., Goethals, P.L.M. and De Pauw, N. (2005). Implications of taxonomic modifications and alien species on biological water quality assessment as exemplified by the Belgian Biotic Index method. *Hydrobiologia*, 542: 137-150.

Gabriels, W., Goethals, P.L.M. and De Pauw, N. (2006). Development of a multimetric assessment system based on macroinvertebrates for rivers in Flanders (Belgium) according to the European Water Framework Directive. *Verhandlungen der Internationale Vereinigung für Theoretische und Angewandte Limnologie*, 29: 2279-2282.

Gerritsen, J., Barbour, M.T. and King, K. (2000). Apples, oranges, and ecoregions: on determining pattern in aquatic assemblages. *Journal of the North American Benthological Society*, 19: 487-496.

Giller, P.S. and Malmqvist, B. (1998). *The Biology of Streams and Rivers*. Oxford University Press, Oxford, NewYork, USA, 296pp.

Gilliam, J.W., Osmond, D.L. and Evans, R.O. (1997). Selected agricultural best management practices to control nitrogen in the Neuse rivr basin. *North Carolina Agricultural Research Service Technical Bulletin*: 311.

Goethals, P. and De Pauw, N. (2001). Development of a concept for integrated ecological river assessment in Flanders, Belgium. *Journal of Limnology* 60: 7-16.

Goethals, P., Dedecker, A.P., Gabriels, W. and De Pauw, N. (2002). Development and application of predictive river ecosystem models based on classification trees and artificial neural networks. 91-107. In: Recknagel, F. (Eds.), *Ecological informatics: understanding* *ecology by biologically-inspired computation*. Springer-Verlag, Berlin, Germany.

Goethals, P. (2005). *Data driven development of predictive ecological models for benthic macroinvertebrates in rivers*. PhD thesis, Ghent University, Gent, Belgium, 377pp.

Goethals, P.L.M., Dedecker, A.P., Gabriels, W., Lek, S. and De Pauw, N. (2007). Applications of artificial neural networks predicting macroinvertebrates in freshwaters. *Aquatic Ecology*, 41: 491-508.

Gomes, M.E.P. and Favas, P.J.C. (2005). *Mineralogical controls on mine drainage of the abandoned Ervedosa tin mine in north-eastern Portugal*. Symposium on Mineralogy and Geochemistry of Acid Mine Drainage and Metalliferous Minewastes held at the 15th Annual VM Goldschmidt Conference, Pergamon-Elsevier Science Ltd, Moscow, ID: 1322-1334.

Gose, K. (1978). Trial to application of saprobic values and indicator value on macroscopic animals by Zelinka-Marvan. *Japanese Journal of Water Treatment Biology*, 14: 9-17.

Gotelli, N.J. and Ellison, A.M. (2004). *A primer of ecological statistics*. Sinauer Associates Inc. , Sunderland, Massachusetts, USA, 510pp.

Gottardo, S., Semenzin, E., Agostini, P., Critto, A. and Marcomini, A. (2007). *Supporting tools for decisional process within Water Framework Directive: From EU context to MODELKEY perspective*. Workshop on Sustainable Use and Development of Watersheds for Human Security and Peace, Springer, Istanbul, TURKEY: 215-225.

Guerold, F. (2000). Influence of taxonomic determination level on several community indices. *Water Research*, 34(2): 487-492.

Guisan, A. and Zimmerman, N.E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, 135: 147-168.

Guo, Q.H., Kelly, M. and Graham, C.H. (2005). Support vector machines for predicting distribution of sudden oak death in California. *Ecological Modelling*, 182: 75-90.

Hart, B.T., Davies, P.E., Humphrey, C.L., Norris, R.N., Sudaryanti, S. and Trihadiningrum, Y. (2001). Application of the Australian river bioassessment system (AUSRIVAS) in the Brantas River, East Java, Indonesia. *Journal of Environmental Management*, 62: 93-100.

Hartmann, A., Moog, O., Ofenbock, T., Korte, T., Sharma, S., Hering, D. and Baki, A. B.M. (2008). *ASSESS-HKH Methodology Manual describing fundamentals & application of three approaches to evaluate the river quality based on benthic invertebrates: HKH screening, HKH score bioassessment & HKH multimetric bioassessment*. University of Natural Resources and Applied Life Sciences (BOKU), Vienna, Austria, 81pp.

Hauer, F.R. and Resh, V.H. (1996). Benthic macroinvertebrates. 339-365. In: Hauer, F. R. and Lamberti, G. A. (Eds.), *Methods in Stream Ecology*. Academic Press, NewYork, USA.

Hawkes, H.A. (1982). Biological surveillance of rivers. *Water Pollution Control*, 81: 329-

342.

Hawkes, H.A. (1997). Origin and development of the biological monitoring working party score system. *Water Resource*, 32: 964-968.

Heggenes, J. and Saltveit, S. (1990). Seasonal and spatial microhabitat selection and segregation in young Atlantic salmon, Salmo salar L., and brown trout, Salmo trutta L., in a Norwegian river. *Journal of Fish Biology*, 36: 707–720.

Hellawell (1986). *Biological indicators of freshwater pollution and environmental management*. Elsevier Applied Science Publishers, London, UK, 546pp.

Hering, D., Buffagni, A., Moog, O., Sandin, L., Sommerhauser, M., Stubauer, I., Feld, C., Johnson, R., Pinto, P., Skoulikidis, N., Verdonschot, P. and Zahradkova, S. (2003). *The development of a system to assess the ecological quality of streams based on macroinvertebrates - Design of the sampling programme within the AQEM project*. 5th International Meeting on River Botton, Lunz, Austria: 345-361.

Hering, D., Moog, O., Sandin, L. and Verdonschot, P.F.M. (2004). Overview and application of the AQEM assessment system. *Hydrobiologia*, 516: 1-20.

Hering, D., Feld, C.K., Moog, O. and Ofenbock, T. (2006). Cook book for the development of a Multimetric Index for biological condition of aquatic ecosystems: Experiences from the European AQEM and STAR projects and related initiatives. *Hydrobiologia*, 566: 311-324.

Hill, M.O. (1973). Diversity and evenness: a unifying notation and its consequences. *Ecology*, 54: 427-432.

Hill, M.O., Bunce, R.G.H. and Shaw, M.W. (1975). Indicator species analysis, a divisive polythetic method of classification, and its application a survey of native pinewoods in Scotland. *Journal of Ecology*, 63: 597-613.

Hill, M.O. (1979). *TWINSPAN - A FORTRAN program for arranging multivariate data in an ordered two-way table by classification of the individuals and attributes*. Cornell University, Ithaca, 90pp.

Hill, M.O. (1994). *DECORANA and TWINSPAN, for ordination and classification of multivariate species data: a new edition, together with supporting programs, in FORTRAN*. Ecology and Systematics, Cornell University, Ithaca, New York 14850, 77pp.

Hilsenhoff, W.L. (1977). *Use of arthropods to evaluate water quality of streams*. Madison: Department of natural resources, Wisconsin, USA, 55pp.

Hilsenhoff, W.L. (1988). Rapid field assessment of organic pollution with a family-level biotic index. . *Journal of the North American Benthological Society*, 7: 65-68.

Hirzel, A. and Guisan, A. (2002). Which is the optimal sampling strategy for habitat suitability modelling. *Ecological Modelling*, 157: 331-341.

Hoang, H., Recknagel, F., Marshall, J. and Choy, J. (2001). Predictive modelling of macroinvertebrate assemblages for stream habitat assessments in Queensland (Australia). *Ecological Modelling*, 195: 195-206.

Hoang, T.H. and Mai, D.Y. (2001). Assessment of water quality of Da Lat city, Dac Ta Dun stream and Da Nhim river using macroinvertebrates (In Vietnamese). *Journal of Biology*, 3:69-75.

Hoang, H., Recknagel, F., Marshall, J. and Choy, J. (2002). Elucidation of hypothetical relationships between habitat conditions and macroinvertebrate assemblages in freshwater streams by Artificial Neural Networks. 179-192. In: Recknagel, F. (Eds.), *Ecological informatics: understanding ecology by biologically-inspired computation*. Springer-Verlag, Berlin.

Hoang, D.H. (2005). *Systematics of the Trichoptera (Insecta)*. PhD thesis, Seoul Women's University, Seoul, Korea, 357pp.

Hoang, D.H. and Bae, Y.J. (2006). Aquatic insect diversity in a tropical Vietnamese stream in comparison with that in a temperate Korean stream. *Limnology*, 7: 45-55.

Holland, J.H. (1975). *Adaptation in Natural and Artificial Systems*. The University of Michigan Press, Ann Arbor, USA.

Hwang, Y.Y., Teng, D.X. and Zhao, Z.X. (1982). Monitoring Jiyunhe estuary pollution by use of macroinvertebrate community and diversity index. *Sinozoologia*, 2: 133–146.

Hynes, H.B.N. (1960). *The Biology of Polluted Water*. Liverpool University Press, Liverpool, UK, 202pp.

Hynes, H.B.N. (1970). *The Ecology of Running Water*. Liverpool University Press, Liverpool, UK, 555pp.

IDRC (1987). *Macroinvertebrate survey of rivers in Korea*. Yonsei University. Project 84011.

Ito, K., Xu, Z.X., Jinno, K., Kojiri, T. and Kawamura, A. (2001). Decision support system for surface water planning in river basins. *Journal of Water Resources Planning and Management-Asce*, 127: 272-276.

Jaccard, P. (1908). Nouvelles recherches sur la distribution florale *Bulletin de la Société Vaudoise des Sciences Naturelles*, 44: 223–270.

Jacobsen, D. (1998). The effect of organic pollution on the macroinvertebrate fauna of Ecuadorian highland streams. *Archiv für Hydrobiologie*, 143: 179-195.

Jacobsen, D. (2003). Altitudinal changes in diversity of macroinvertebrates from small streams in the Ecuadorian Andes. *Archiv für Hydrobiologie*, 158: 145-167.

Jacobsen, D. and Encalada, A. (1998). The macroinvertebrate fauna of Ecuadorian highland streams in the wet and dry season. *Archiv für Hydrobiologie*, 142: 53-70.

Jonnalagadda, S.B. and Mhere, G. (2001). Water quality of the Odzi River in the Eastern Highlands of Zimbabwe. *Water Research*, 35: 2371-2376.

Jung, S.W., Nguyen, V.V., Nguyen, Q.H. and Bae, Y.J. (2008). Aquatic insect faunas and communities of a mountain stream in Sapa Highland, northern Vietnam. *Limnology*, 9: 219-229.

Kannel, P.R., Lee, S., Lee, Y.S., Kanel, S.R. and Khan, S.P. (2007). Application of water quality indices and dissolved oxygen as indicators for river water classification and urban impact assessment. *Environmental Monitoring and Assessment*, 132: 93-110.

Karr, J.R. (1981). Assessment of biotic integrity using fish communities *Fisheries*, 6(6): 21-27.

Karr, J.R. and Chu, E.W. (1999). *Restoring life in running waters. Better biological monitoring*. Island Press, Washington DC., USA, 106pp.

Karr, J.R. and Chu, E.W. (2000). Sustaining living rivers. *Hydrobiologia*, 422: 1-14.

Kawai, T., Ed. **(1985)**. *An illustrated book of aquatic insects of Japan (In Japanese)*. Tokai University Press, Tokyo, Japan, 408pp.

Kawai, T. and Tanida, K., Eds. **(2005)**. *Aquatic insects in Japan: manual with keys and illustrations (in Japanese)*. Japan: Tokai University Press, Hadano, Kanagawa, 1342pp.

Keerthi, S.S., Shevade, S.K., Bhattacharyya, C. and Murthy, K.R.K. (2001). Improvements to Platt's SMO algorithm for SVM classifier design. *Neural Computation*, 13: 637-649.

Kerans, B.L. and Karr, J.R. (1994). A benthic index of biotic integrity (B-IBI) for rivers of the Tennessee Valley. *Ecological Applications*, 4: 768-785.

Kim, D.S., Nguyen, H.N. and Park, J.S. (2005). *Genetic algorithm to improve SVM based network intrusion detection system*. 19th International Conference on Advanced Information Networking and Applications, Taipei, Taiwan: 155-158.

Kobayashi, S. and Kagaya, T. (2005). Hot spots of leaf breakdown within a headwater stream reach: comparing breakdown rates among litter patch types with different macroinvertebrate assemblages. *Freshwater Biology*, 50(6): 921-929.

Kohavi, R. (1995). *The power of decision tables*. 8th European Conference on Machine Learning, Iraklion, Greece: 174-189.

Kolkwitz, R. and Marsson, M. (1908). Ökologie der pflanzlichen Saprobien. . *Berichte der Deutschen Botanischen Gesellschaft* 26: 505-519.

Krystiano, A.I.A. and Kusjantono, H. (1991). *A preliminary study of the use of the BBI (Belgian Biotic Index) to assess the water quality of the river Ledok, Indonesia*. Conservation and management of tropical inland waters: Problems, solutions and perspectives., Hong Kong.

Lang, C. and Reymond, O. (1993). Empirical relationships between diversity of invertebrate communities and altitude in rivers: Application to biomonitoring. *Aquatic Sciences*, 55: 188-196.

Lautenschlager, M. and Kiel, E. (2005). Assessing morphological degradation in running waters using Blackfly communities (Diptera, Simuliidae): Can habitat quality be predicted from land use? *Limnologica*, 35(4): 262-273.

Le, T.H., Nguyen, X.Q. and Mai, D.Y. (2002). Apply BMWP score system to assess water quality of some running water. *Vietnam National University Journal of Science*, 1: 22-28.

Le, T.H. (2003). *Macroinvertebrate community composition in Tam Dao stream and their application as biological indicators*. PhD thesis, Hanoi National Unversity, Hanoi, Vietnam, 199pp.

Le, V.K. (2007). *Environmental biological indicators (In Vietnamese)*. Education Publishing house, Hanoi, Vietnam, 280pp.

Legendre, P. and Legendre, L. (1998). *Numerical Ecology. 2nd English Edition*. Elsevier Science, Amsterdam, The Netherlands, 870pp.

Lenat, D.R. (1993). A biotic index for the south-eastern United-States- Derivation and list of tolerance values with criteria for assigning water-quality ratings. *Journal of the North American Benthological Society*, 12(3): 279-290.

Lenat, D.R. and Resh, V.H. (2001). Taxonomy and stream ecology - The benefits of genus- and species-level identifications. *Journal of the North American Benthological Society*, 20: 287-298.

Liangfang, Y., Youwen, L., Lixin, T. and Changhai, S. (1991). *Community structure of aquatic insects and its use in biomonitoring water quality of the JiuHua River, P.R. China*. Conservation and Management of Tropical Inland Waters: Problems, Solutions and Prospects, Hong Kong.

Linke, S., Bailey, R.C. and Schwindt, J. (1999). Temporal variability of stream bioassessment using benthic macroinvertebrates. *Freshwater Biology*, 42: 575-584.

Loeb, S.L. and Spacie, A. (1994). *Biological Monitoring of Aquatic System*. Lewis Publisher, Boca Raton, Florida, USA, 381pp.

Ma, T., Huang, Q., Wang, H., Wang, Z., Wang, C. and Huang, S. (2008). Selection of benthic macroinvertebrate-based multimetric and preliminary establishment of biocriteria for the bioassessment of the water quality of Taihu Lake, China. *Acta Ecologica Sinica*, 28: 1192−1200.

Mackenthun, K. M. (1969). *The practice of water pollution biology* FWPCA, Washington, USA, 281pp.

Madar, M. and Cretu, G. (2008). *First approach of ecohydrological researches at a regional scale in Banat hydrographical area*. Banat Water Directorate, Timisoara, Romania.

Maes, A., Mouton, A.M., De Pauw, N. and Goethals, P. (2007). KRW (Kaderrichtlijn Water)-Verkenner als instrument voor het ecologische evalueren van rivierherstelopties (In Dutch). *Water*, 27: 6-11.

Mander, U., Kuusemets, V. and Hayakawa, Y. (2005). Purification processes, ecological functions, planning and design of riparian buffer zones in agricultural watersheds. *Ecological Engineering*, 24: 421-432.

Manel, S., Dias, J.M., Buckton, S.T. and Ormerod, S.J. (1999). Alternative methods for predicting species distribution: an illustration with Himalayan river birds. *Journal of Applied Ecology*, 36: 734-747.

Manel, S., Williams, H.C. and Ormerod, S.J. (2001). Evaluating presence-absence models in ecology: the need to account for prevalence. *Journal of Applied Ecology*, 38: 921-931.

Marchant, R. and Hehir, G. (2002). The use of AUSRIVAS predictive models to assess the response of lotic macroinvertebrates to dams in south-east Australia. *Freshwater Biology*, 47(5): 1033-1050.

Marchant, R. (2007). The use of taxonomic distinctness to assess environmental disturbance of insect communities from running water. *Freshwater Biology*, 52(8): 1634-1645.

Margalef, R. (1958). Information theory in ecology. *General systems*, 3: 36-71.

Marshall, J., Hoang, H., Choy, S. and Recknagel, F. (2003). Relationships between habitat properties and the occurrence of macroinvertebrates in Queensland streams (Australia) discovered by a sensitivity analysis with artificial neural networks. *Verhandlungen der Internationale Vereinigung für Theoretische und Angewandte Limnologie*, 28: 1415-1419.

Mason, W.T., Lewis, P.A. and Weber, C.I. (1985). An evaluation of benthic macroinvertebrate biomass methodology. 2. Field assessment and data evaluation. *Environmental Monitoring and Assessment*, 5: 399-422.

Massoud, M.A., Tarhini, A. and Nasr, J.A. (2009). Decentralized approaches to wastewater treatment and management: Applicability in developing countries. *Journal of Environmental Management*, 90: 652-659.

Matagi, S.V. (1996). The effect of pollution on benthic macroinvertebrates in a Ugandan stream. *Archiv für Hydrobiologie*, 137(4): 537-549.

McCafferty, W.P. and Provonsha, A.V. (1983). *Aquatic Entomology: The fishermen's and ecologists' illustrated guide to insects and their relatives*. Jones and Bartlett Publishers, Boston, USA, 448pp.

McCune, B. (1997). Influence of noisy environmental data on Canonical Correspondence Analysis. *Ecology*, 78: 2617-2623.

McCune, B. and Mefford, M.J. (1999). *Multivariate analysis for ecological data Version 4.0*. MjM Software, Oregon.

McKnight, T.L. and Hess, D. (2000). *Climate Zones and Types: The Köppen System,* *Physical Geography: A Landscape Appreciation. Upper Saddle River.* NJ: Prentice Hall, USA, 200pp.

Melo, A.S. and Froehlich, C.G. (2001). Macroinvertebrates in neotropical streams: richness patterns along a catchment and assemblage structure between 2 seasons. *Journal of the North American Benthological Society*, 20: 1-16.

Metcalfe, J.L. (1989). Biological water-quality assessment of running water based on macroinvertebrate communities - History and present status in Europe. *Environmental Pollution*, 60: 101-139.

Meyer, D., Leisch, F. and Hornik, K. (2003). The support vector machine under test. *Neurocomputing*, 55(1-2): 169-186.

Milenova, B.L., Yarmus, J.S. and Campos, M.M. (2005). *SVM in Oracle Database 10g: Removing the Barriers to Widespread Adoption of Support Vector Machines*. 31st VLDB Conference, Trondheim, Norway.

Minshall, G.W., Cummins, K.W., Petersen, R.C., Cushing, C.E., Bruns, D.A., Sedell, J. R. and Vannote R.L. (1985). Developments in stream ecosystem theory. *Canadian Journal of Fisheries and Aquatic Science*, 42: 1045-1055.

Morse, J.C., Bae, Y.J., Munkhjargal, G., Sangpradub, N., Tanida, K., Vshivkova, T.S., Wang, B.X., Yang, L.F. and Yule, C.M. (2007). Freshwater biomonitoring with macroinvertebrates in East Asia. *Frontiers in Ecology and the Environment*, 5: 33-42.

Mouton, A.M., Schneider, M., Depestele, J., Goethals, P. L. M. and De Pauw, N. (2007). Fish habitat modelling as a tool for river management. *Ecological Engineering*, 29: 305-315.

Mouton, A.M. (2008). *A critical analysis of performance criteria for the evaluation and optimisation of fuzzy models for species distribution*. PhD thesis, Ghent University, Gent, Belgium, 181pp.

Mouton, A.M., Van Der Most, H., Jeuken, A., Goethals, P.L.M. and De Pauw, N. (2009). Evaluation of river basin restoration options by the application of the Water Framework Directive Explorer in the Zwalm River basin (Flanders, Belgium). *River Research and Applications*, 25: 82-97.

Moya, N., Tomanova, S. and Oberdorff, T. (2007). Initial development of a multi-metric index based on aquatic macroinvertebrates to assess streams condition in the Upper Isiboro-Secure Basin, Bolivian Amazon. *Hydrobiologia*, 589: 107-116.

Moyo, N.A.G. and Phiri, C. (2002). The degradation of an urban stream in Harare, Zimbabwe. *African Journal of Ecology*, 40: 401-406.

MRWP (2008). *Riparian Buffer Zones - Vital to Healthy Rivers & Streams.* Education brochure. Middlebury River Watershed Partnership, Middlebury, USA. http://207.136.225.66/Downloads/Other/Riparian.pdf. Assessed June 10th 2009.

Munne, A. and Prat, N. (2009). Use of macroinvertebrate-based multimetric indices for water quality evaluation in Spanish Mediterranean rivers: an intercalibration approach with the IBMWP index. *Hydrobiologia*, 628: 203-225.

Mustow, S.E. (1997). *Aquatic Macroinvertebrates and Environmental Quality of Rivers in Northern Thailand*. PhD thesis, London University, London, UK, 391pp.

Mustow, S.E. (1999). Lotic macroinvertebrate assemblages in northern Thailand: altitudinal and longitudinal distribution and the effects of pollution. *Nat. Hist. Bull. Siam Soc*, 47: 225–252.

Mustow, S.E. (2002). Biological monitoring of rivers in Thailand: use and adaptation of the BMWP score. *Hydrobiologia*, 479: 191-229.

National Water Council (1981). *River Quality: The 1980 Survey and Future Outlook.* National Water Council, London, 39pp.

Nemerow (1985). *Stream, Lake, Estuary and Ocean Pollution*. Van Nostrand Reinhold, New York, USA, 215pp.

Neumann, M. and Dudgeon, D. (2002). The impact of agricultural runoff on stream benthos in Hong Kong, China. *Water Research*, 36: 3103-3109.

Newson, M.D. and Large, A.R.G. (2006). 'Natural' rivers, 'hydromorphological quality' and river restoration: a challenging new agenda for applied fluvial geomorphology. *Earth surface processes and landforms*, 31: 1606-1624.

Ngo, X.Q. (2007). *Applied the ASPT index for water quality assessment in the main streams of Nui Chua national park*. Institute of Tropical Biology: Sciense and Technology Symposium, Ho Chi Minh city, Vietnam: 623-631.

Nguyen, T.L., Dang, K.C (2008). *Study on wastewater treatment method in small scale Phuc Linh tin extraction enterprise (Thai Nguyen province)*. Conference Proceeding: Promoting Development of Environmental Science and Industry in Vietnam, Hanoi, Vietnam: 264-267.

Nguyen, X.Q. (1985). Database on macroinvertebrate community in the To Lich river, Hanoi (In Vietnamese). *Vietnam National University Journal of Science*, 3: 51-57.

Nguyen, X.Q. (1995). *Reseach on invertebrates in water bodies in Hanoi.* PhD thesis, Hanoi National Unversity, Hanoi, Vietnam, 130pp.

Nguyen, V.V., Hoang, D.H., Cao, T.K.T., Nguyen, X.Q. and Bae, Y.J. (2001). Altitudinal distribution of aquatic insects from Tam Dao National Park in northern Vietnam. 123-133. In: Bae, Y. J. (Eds.), *The 21st century and aquatic entomology in East Asia*. Jeonghaengsa, Seoul, Korea.

Nguyen, V.V. (2003). *Systematics of the Ephemeroptera (Insecta) of Vietnam*. PhD thesis, Seoul Women's University, Seoul, Korea, 322pp.

Nguyen, V.V. and Bae, Y.J. (2004). Descriptions of Rhoenanthus sapa, new species, and larval stage of R-magnificus Ulmer (Ephemeroptera: Potamanthidae) from Vietnam. *Aquatic Insects*, 26: 9-17.

Nguyen, X.Q., Mai, D.Y., Pinder, C. and Tilling, S. (2004). *Biological surveillance of freshwaters: A practical manual and identification key for use in Vietnam*. Vietnam National University Publishers, Hanoi, Vietnam, 109pp.

Nguyen, V.T. and Ta, H.T. (2005). *Application of ASPT biotic index for rapid assessment of water quality in the Cau river*. The 1st National conference of Ecology and Biological Resources, Hanoi, Vietnam.

NISF (2008). *Ferliliser use and production in Vietnam*. Fertilizer Advisory, Development and Information Network for Asia and the Pacific, http://www.fadinap.org/vietnam/fertilizer.html. Assessed November 2008.

Norris, R. and Georges, A. (1993). Analysis and interpretation of benthic macroinvertebrate surveys. 234–287. In: Rosenberg, D. M. and Resh, V. H. (Eds.), *Freshwater Biomonitoring and Benthic Macroinvertebrates*. Chapman & Hall, NewYork.

Ohn, S.Y., Nguyen, H.N., Kim, D.S. and Park, J. S. (2004). *Determining optimal decision model for support vector machine by genetic algorithm*. 1st International Symposium oncomputational and Information Science, Shanghai, P.R.China: 895-902.

Omernik, J.M. (1995). Ecoregions: a spatial framework for environmental management. 49-62. In: Davis, W. S. and Simon, T. P. (Eds.), *Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making*. Lewis Publishers, Boca Raton, USA.

Ometo, J., Martinelli, L.A., Ballester, M.V., Gessner, A., Krusche, A.V., Victoria, R.L. and Williams, M. (2000). Effects of land use on water chemistry and macroinvertebrates rates in two streams of the Piracicaba river basin, south-east Brazil. *Freshwater Biology*, 44: 327-337.

Parsons, M. and Norris, R. (1996). The effect of habitat - specific sampling on biological assessment of water quality using a predictive model. *Freshwater Biology*, 36: 419-434.

Persoone, G. and De Pauw, N. (1979). Systems of biological indicators for water quality assessment 39–75. In: Ravera, O. (Eds.), *Biological aspects of freshwater polution. Pergamon Press*, Oxford, UK.

Pesce, S.F. and Wunderlin, D.A. (2000). Use of water quality indices to verify the impact of Cordoba city (Argentina) on Suquýa river. *Water Research*, 34(11): 2915–2926.

Pielou, E.C. (1966). Measurement of diversity in different types of biological collections. *Journal of Theoretical Biology*, 13: 131-144.

Pinder, L.C.V., Ladle, M., Gledhill, T., Bass, J. and Mathews, A.M. (1987). Biological surveillance of water quality -1. A comparison of macroinvertebrate surveillance methods in relation to assessment of water quality, in a chalk stream. *Archiv für Hydrobiologie*, 109: 207-226.

Plafkin, J.L., Barbour, M.T., Porter, K.D., Gross, S.K. and Hughes, R.M. (1989). *Rapid bioassessment protocols for use in streams and rivers: Benthic macroinvertebrates and fish*. U.S. Environmental Protection Agency, Office of Water Regulations and Standards, Washington, D.C., USA, 674pp.

Poff, N.L., Allan, J.D., Palmer, M.A., Hart, D.D., Richter, B.D., Arthington, A.H., Rogers, K.H., Meyers, J.L. and Stanford, J.A. (2003). River flows and water wars: emerging science for environmental decision making. *Frontiers in Ecology and the Environment*, 1: 298-306.

Pringle, C.M. and Ramirez, A. (1998). Use of both benthic and drift sampling techniques to assess tropical stream invertebrate communities along an altitudinal gradient, Costa Rica. *Freshwater Biology*, 39: 359-373.

Qi, S. (1991). *Water-quality assessment y using benthic macroinvertebrates in the Zhujiang (Pearl River) Delta, P.R. China.* Conservation and Management of Tropical Inland Waters: Problems, Solutions and Prospects, Hong Kong.

Quinlan, J.R. (1986). Induction of decision trees. *Machine Learning*, 1: 81-106.

Quinlan, J.R. (1993). *C4.5: program for machine learning*. Morgan Kaufmann Publishers, San Francisco, USA, 302pp.

Ramirez, A. and Pringle, C.M. (2001). Spatial and temporal patterns of invertebrate drift in streams draining a Neotropical landscape. *Freshwater Biology*, 46: 47-62.

Randin, C.F., Dirnbo¨ck, T., Dullinger, S., Zimmerman, N.E., Zappa, M. and Guisan, A. (2006). Are niche-based species distribution models transferable in space. *Journal of Biogeography*, 33: 1689–1703.

Rao, S.V.R., Singh, V.P. and Mall, L.P. (1978). Pollution studies of River Khan (Indore) India. 1. Biological assessment of pollution. *Water Research*, 12: 555-559.

Rao, N.S., Easton, Z.M., Schneiderman, E.M., Zion, M.S., Lee, D.R. and Steenhuis, T.S. (2009). Modeling watershed-scale effectiveness of agricultural best management practices to reduce phosphorus loading. *Journal of Environmental Management*, 90: 1385-1395.

Recknagel, F. (2002). *Ecological informatics: understanding ecology by biologically-inspired computation*. Springer-Verlag, Berlin, Germany, 432pp.

Resh, V.H. and Jackson, J.K. (1993). Rapid assessment approaches to biomonitoring using benthic macroinvertebrates. 195-233. In: Rosenberg, D.M. and Resh, V.H. (Eds.), *Freshwater Biomonitoring and Benthic Macroinvertebrates*. Chapman & Hall, New York, USA.

Resh, V.H. and McElravy (1993). Contemporary quantitative approaches to biomonitoring using benthic macroinvertebrates. 159-194. In: Rosenberg, D.M. and Resh, V.H. (Eds.), *Freshwater Biomonitoring and Benthic Macroinvertebrates*. Chapman & Hall, New York, USA.

Resh, V.H. (2007). Multinational, freshwater biomonitoring programs in the developing world: Lessons learned from African and Southeast Asian river surveys. *Environmental Management*, 39: 737-748.

Reynoldson, T.B., Day, K.E. and Pascoe, T. (2000). The development of the BEAST: a predictive approach for assessing sediment quality in the North American Great Lakes. 165-180. In: Wright, J.F., Sutcliffe, D.W. and Furse, M.T. (Eds.), *Assessing the biological quality of fresh waters: RIVPACS and other techniques.* . Freshwater Biological Association, Ambleside, Cumbria, UK.

Ribeiro, R. and Torgo, L. (2008). A comparative study on predicting algae blooms in Douro River, Portugal. *Ecological Modelling*, 212: 86-91.

Robinson, C.T., Minshall, G.W. and Royer, T.V. (2000). Inter-annual patterns in macroinvertebrate communities of wilderness streams in Idaho, USA. *Hydrobiologia*, 421: 187-198.

Roldán, G. (2003). *Bioindicación de la calidad del agua en Colombia. Uso del método BMWP/Col.* Universidad de Antioquia, Medellín, Colombia, 182pp.

Rosenberg, D.M. and Resh, V.H. (1993). *Freshwater biomotoring and benthic macroinvertebrates*. Chapman & Hall, New York-London, 488pp.

Rosenberg, D.M., Davies, I.J., Cobb, D.G. and Wiens, A.P. (2005). *Protocols for measuring biodiversity: Benthic Macroinvertebrates in Fresh Waters*. Department of Fisheries and Oceans, Freshwater Institute, 501 University Crescent, Winnipeg, Manitoba**:** 46.

Roux, D.J., Rogers, K.H., Biggs, H.C., Ashton, P.J. and Sergeant, A. (2006). Bridging the science-management divide: moving from unidirectional knowledge transfer to knowledge interfacing and sharing. *Ecology and Society*, 11: Article number 4.

Royal Haskoning, Witteveen-Bos and Taken Landschapsplanning (2005). *Validatie en Verdere Optimalisatie Van de Concept KRW-Maatlatten Voor de Natuurlijke Rivier-en Meertypen*. RIZA, Rotterdam, The Netherlands (in Dutch).

Sajidu, S.M.I., Persson, I., Masamba, W.R.L. and Henry, E.M.T. (2008). Mechanisms of heavy metal sorption on alkaline clays from Tundulu in Malawi as determined by EXAFS. *Journal of Hazardous Materials*, 158: 401-409.

Sanchez, E., Colmenarejo, M.F., Vicente, J., Rubio, A., Garcia, M. G., Travieso, L. and Borja, R. (2007). Use of the water quality index and dissolved oxygen deficit as simple indicators of watersheds pollution. *Ecological Indicators*, 7: 315-328.

Sanchez-Hernandez, C., Boyd, D.S. and Foody, G.M. (2007). Mapping specific habitats from remotely sensed imagery: Support vector machine and support vector data description based classification of coastal saltmarsh habitats. *Ecological Informatics*, 2:83-88.

Sandin, L., Sommerhäuser, M., Stubauer, I., Hering, D. and Johnson, R. (2000). *Stream* *assessment methods, stream typology approaches and outlines of a European stream typology*. AQEM-EU project EVK1-CT1999-00027**:** 43.

Sandin, L. (2001). SWEPACS: a Swedish running water prediction and classification system using benthic macroinvertebrates. 44-46. In: Bäck, S. and Karttunen, K. (Eds.), *Classification of ecological status of lakes and rivers.* TemaNord Environment 2001.

Schoener, T.W. (1974). Resource partitioning in ecological communities. *Science*, 185:27-39.

Schofield, N.J. and Davies, P.E. (1996). Measuring the health of our rivers. *Water*, 23:39-43.

Shan, Y., Paull, D. and McKay, R.I. (2006). Machine learning of poorly predictable ecological data. *Ecological Modelling*, 195: 129-138.

Shannon, C.E. and Weaver, W. (1949). *The mathematical theory of communication.* . University of Illinois Press, Urbana, USA, 117pp.

Sharma, S. (1996). *Biological Assessment of Water Quality in the Rivers of Nepal*. PhD thesis, University of Agriculture, Forestry and Renewable Natural Resources (BOKU), Vienna, Austria, 398pp.

Sharma, S. and Moog, O. (1996). *The applicability of biotic indices and scores in water quality assessment of Nepalese rivers* International Conference on Ecohydrology of High Mountain Areas, ICIMOD, UNESCO, Kathmandu: 641–657.

Sharma, C.M., Sharma, S., Borgstrom, R. and Bryceson, I. (2005). Impacts of a small dam on macroinvertebrates: A case study in the Tinau River, Nepal. *Aquatic Ecosystem Health & Management*, 8: 267–275.

Sharma, R.C. and Rawat, J.S. (2009). Monitoring of aquatic macroinvertebrates as bioindicator for assessing the health of wetlands: A case study in the Central Himalayas, India. *Ecological Indicators*, 9: 118-128.

Sinha, M., Sinha, R., Khan, E. and Mehrotra, P. (1993). Biotic index as a means of biomonitoring of aquatic habitat. . *Journal of Ecotoxicology and Environmental Monitoring*, 3: 125-128.

Sivaramakrishnan, K.G., Hannaford, M.J. and Resh, V.H. (1996). Biological assessment of the Kaveri River catchment, South India, using benthic macroinvertebrates: Applicability of water quality monitoring approaches developed in other countries. *International Journal of Ecology and Environmental Sciences*, 22: 113-132.

Skoulikidis, N.T., Gritzalis, K.C., Kouvarda, T. and Buffagni, A. (2004). The development of an ecological quality assessment and classification system for Greek running waters based on benthic macroinvertebrates. *Hydrobiologia*, 516: 149-160.

Sladecek, V. (1973). System of water quality from the biological point of view. *Ergebnisse der Limnologie*, 7: 1-128.

Smith, M.J., Kay, W.R., Edward, D.H.D., Papas, P.J., Richardson, K.S., Simpson, J.C., Pinder, A.M., Cale, D.J., Horwitz, P.H.J., Davis, J.A., Yung, F.H., Norris, R.H. and

Halse, S.A. (1999). AusRivAS: using macroinvertebrates to assess ecological condition of rivers in Western Australia. *Freshwater Biology*, 41: 269-282.

Smola, A.J. and Scholkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing*, 14: 199-222.

Soldner, M., Stephen, I., Ramos, L., Angus, R., Wells, N.C., Grosso, A. and Crane, M. (2004). Relationship between macroinvertebrate fauna and environmental variables in small streams of the Dominican Republic. *Water Research*, 38: 863-874.

Sorensen, T. (1948). A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analysis of the vegetation on Danish commons. *Biologiske Skifter. Det Konglige Danske Videnskabernes Selskab*, 5: 1-34.

Sporka, F., Vlek, H.E., Bulankova, E. and Krno, I. (2006). Influence of seasonal variation on bioassessment of streams using macroinvertebrates. *Hydrobiologia*, 566: 543-555.

Stambuk-Giljanovic, N. (1999). Water quality evaluation by index in Dalmatia. *Water Research*, 33(16): 3423-3440.

Stark, J.D., Boothroyd, I.K.G., Harding, J.S., Maxted, J.R. and Scarsbrook, M.R. (2001). *Protocols for sampling macroinvertebrates in wadeable streams. Report No. 1. Prepared for the Ministry for the Environment. Sustainable Management Fund Project No. 5103*. New Zealand Macroinvertebrate Working Group Auckland, New Zealand**:** 57.

Stiling, P.D. (1996). *Ecology: Theories and Applications*. Prentice Hall International, New Jersey, USA, 539pp.

STOWA (2004). *Referenties en Concept-Maatlatten Voor Rivieren Voor de Kaderrichtlijn Water. (in Dutch)*. Hageman Fulfilment, Zwijndrecht, The Netherlands.

Strauss, B. and Biedermann, R. (2007). Evaluating temporal and spatial generality: how valid are species-habitat relationship models. *Ecological Modelling*, 204: 104–114.

Stubauer, I. and Moog, O. (2000). Taxonomic sufficiency versus need for information – comments based on Austrian experience in biological water quality monitoring. *Verhandlungen der Internationale Vereinigung für Theoretische und Angewandte Limnologie*, 27: 2562–2566.

Sudaryanti, S., Trihadiningrum, Y., Hart, B.T., Davies, P.E., Humphrey, C.L., Norris, R., Simpson, J.C. and Thurtell, L. (2001). Assessment of the biological health of the Brantas River, East Java, Indonesia using the Australian River Assessment System (AUSRIVAS) methodology. *Aquatic Ecology*, 35: 135-146.

Suren, A.M. (1994). Macroinvertebrate communities of streams in western Nepal: effects of altitude and land use. *Freshwater Biology*, 32: 323-336.

Sutherland, R.A. (2001). Analysis and commentary on "Statistical methods and pitfalls in environmental data analysis" by Yue Rong. *Environmental Forensics*, 2: 265-274.

Sweeting, R., Lowson, D., Hale, P. and Wright, J.F. (1992). Biological assessment of rivers in the UK. 319-326. In: Newman, P., Piavaux, A. and Sweeting, R. (Eds.), *River water quality – assessment and control*. Commission of the European Communities. EUR 14606 EN-FR, 1992-III., Brussels, Belgium.

Takao, A., Kawaguchi, Y., Minagawa, T., Kayaba, Y. and Morimoto, Y. (2008). The relationships between benthic macroinvertebrates and biotic and abiotic environmental characteristics downstream of the Yahagi dam, central Japan, and the state change caused by inflow from a tributary. *River Research and Applications*, 24: 580-597.

Thai Nguyen DONRE (2006). *Environmental status in Thai Nguyen provinces 2004- 2005*. Thai Nguyen DONRE, Thai Nguyen, Vietnam 138pp.

Thiriron, C., Macke, A. and Woest, R. (1995). *Biological Monitoring of Streams and Rivers Using SASS 4. A user manual.* . Department of Water Affairs and Forestry, Pretoria, South Africa, 46pp.

Thorne, R.S. and Williams, W.P. (1997). The response of benthic macroinvertebrates to pollution in developing countries: A multimetric system of bioassessment. *Freshwater Biology*, 37: 671-686.

Tolkamp, H.H. (1985). Biological assessment of water quality in running water using macroinvertebrates: a case study for Limburrg, the Netherlands. *Water Science and Technology*, 17: 867-878.

Trihadiningrum, Y., De Pauw, N., Tjondronegoro, I. and Verheyen, R.F. (1996). Use of benthic macroinvertebrates for water quality assessment of the Blawi river (East Java, Indonesia). 199-221. In: Schiemer, F. and Boland, K.T. (Eds.), *Perspectives in tropical limnology*. SPB Academic Publishing, Amsterdam, The Netherlands.

Truong, T.C. and Ngo, T.T.A. (2007). Application of macroinvertebrates for assessment of water quality in four main canal systems in Ho Chi Minh city (In Vietnamese). *Journal of Science and Technology Development*, 10: 25-31.

Tsuda, M. (1962). *Aquatic entomology (In Japanese)*. Japan: Hokuryukan Press, Tokyo, Japan.

Tsuda, M. (1964). *Biology of polluted waters (In Japanese)*. Hokuryukan Press, Tokyo, Japan, 258pp.

Udvardy, M.D.F. (1975). *A classification of the biogeographycal provinces in the worlds*. IUCN, Morges, Switzerland**:** 50.

UN/ECE (1995). *UN/ECE Task Force on Monitoring & Assessment: Vol 3. Biological Assessment Methods for Watercourses*. Institute for Inland Water Management and Waste Water Treatment (RIZA), Lelystad, The Netherlands, 86pp.

UNESCAP (2009). *Statistical Yearbook for Asia and the Pacific 2008.* United Nations Publication, Bangkok, Thailand, 224pp.

US EPA (1997). *Surface water monitoring: A framework for change*. U.S. Environmental Protection Agency, Office of Water, Office of Policy Planning and Evaluation, Washington, D.C, USA.

Usseglio-Polatera, P., Bournaud, M., Richoux, P. and Tachet, H. (2000). Biological and ecological traits of benthic freshwater macroinvertebrates: relationships and definition of groups with similar traits (vol 43, pg 175, 2000). *Freshwater Biology*, 43(1): 175-205.

Van Der Most, H., Jeuken, A. and Van Schijndel, S. (2006). *WFD-Explorer Q4: a planning kit to support the implementation of the Water Framework Directive*. The 7th International Conference in Hydroinformatics, Nice Research Publishing Services, Chennai, India.

Van Looy, S., Verplancke, T., Benoit, D., Hoste, E., Van Maele, G., De Turck, F. and Decruyenaere, J. (2007). A novel approach for prediction of tacrolimus blood concentration in liver transplantation patients in the intensive care unit through support vector regression. *Critical Care*, 11(4): R83.

Vannote, R.L., Minshall, G.W., Cummins, K.W., Sedell, J.R. and Cushing, C.E. (1980). The River Continuum Concepts. *Canadian Journal of Fisheries and Aquatic Science*, 37: 130-137.

Vanrolleghem, P.A., Schilling, W., Rauch, W., Krebs, P. and Aalderink, H. (1999). Setting up measuring campaigns for integrated wastewater modelling. *Water Science & Technology*, 39: 257-268.

Vapnik, V. (1995). *The nature of statistical learning theory*. Springer-Verlag, New York, USA, 187pp.

VEM (2003). *Vietnam Emvironmental Monitor 2003: Water*. MONRE, World Bank and DANIDA, Hanoi, Vietnam**:** 74pp.

VEPA (2006). *Environmental report of Vietnam 2006: The state of water environment in three river basins of Cau, Nhue-Day and Dong Nai river system*. VEPA, Hanoi, Vietnam, 93pp.

VEPA (2007). *Master plan on national environmental and resource monitoring network towards 2020*. Office of Prime Minister of Vietnam, Hanoi, Vietnam, 85pp.

Verdonschot, P.F.M. (2000). Integrated ecological assessment methods as a basis for sustainable catchment management. . *Hydrobiologia* 422/423: 389-412.

Vietnam General Statistical Office (2006). *Statistical Yearbook 2005*. Statistical Publishing House, Hanoi, Vietnam.

Volkman, S. (2003). *Sustainable Wastewater Treatment and Reuse in Urban Areas of the Developing World*. Department of Civil and Environmental Engineering. Michigan Technological University, Michigan, USA.

Vongprasert, R. (1990). *Application of biotic index for water quality assessment of rivers* *in tropical countries: case studies in Thailand and Indonesia.* Master thesis, Ghent University, Gent, Belgium, 78pp.

Vose, M.D. (1999). Random heuristic search. *Theoretical Computer Science*, 229: 103-142.

Walczak, S. and Cerpa, N. (1999). Heuristic principles for the design of artificial neural networks. *Information and Software Technology*, 41: 107-117.

Walley, W.J. and Hawkes, H.A. (1996). A computer-based reappraisal of the Biological Monitoring Working Party scores using data from the 1990 river quality survey of England and Wales. *Water Research*, 30: 2086-2094.

Walley, W.J. and Hawkes, H.A. (1997). A computer-based development of the biological monitoring working party score system incorporating abundance rating, site type and indicator value. *Water Research*, 31: 201-210.

Wallin, M., Wiederholm, T. and Johnson, R.K. (2003). *Guidance on establishing reference conditions and ecological status class boundaries for inland surface waters. Final version 7.0*. Produced by CIS Working Group 2.3 - REFCOND: 73 + appendices.

Walter, M.T., Archibald, J.A., Buchanan, B., Dahlke, H., Easton, Z.M., Marjerison, R.D., Sharma, A.N. and Shaw, S.B. (2009). New Paradigm for Sizing Riparian Buffers to Reduce Risks of Polluted Storm Water: Practical Synthesis. *Journal of Irrigation and Drainage Engineering-Asce*, 135: 200-209.

Wang, B.X. (2002). *Water quality biomonitoring using benthic macroinvertebrates in China.* PhD thesis, Nanjing Agricultural University, Nanjing, China, 187pp.

Wang, B.X. and Yang, L.F. (2004). A study on tolerance values of benthic macroinvertebrate taxa in eastern China. *Acta Ecologica Sinica*, 24: 2768-2775.

Wang, B.X., Yang, L.F., Hu, B.J. and Shan, L.N. (2005). A preliminary study on the assessment of stream ecosystem health in south of Anhui Province using benthic-index of biotic integrity. *Acta Ecologica Sinica*, 25: 1481-1490.

Washington, H.G. (1984). Diversity, Biotic and similarity indices. A review with special relevance to aquatic ecosystems. *Water Research*, 6: 653-694.

Watanabe, N.C., Harada, S. and Komai, Y. (2000). Long-term recovery from mine drainage disturbance of a macroinvertebrate community in the Ichi-kawa River, Japan. *Hydrobiologia*, 429: 171-180.

Weber, C.I. (1973). *Biological field and laboratory methods for measuring the quality of surface water and effluents*. US EPA, Office of Research and Development, Cincinnati, Ohio, USA, 194pp.

Weigel, B.M., Henne, L.J. and Martinez-Rivera, L.M. (2002). Macroinvertebrate-based index of biotic integrity for protection of streams in west-central Mexico. *Journal of the North American Benthological Society*, 21: 686-700.

Wetzel, R.G. (2001). *Limnology, Lake and River Ecosystems. 3rd Edition*. Academic Press, London, UK, 1006pp.

Whittaker, R.H. (1952). A study of summer foliage insect communities in the Great Smoky Mountains. *Ecological Monographs*, 22: 1-44.

Wilby, A., Lan, L.P., Heong, K.L., Huyen, N.P.D., Quang, N.H., Minh, N.V. and Thomas, M.B. (2006). Arthropod diversity and community structure in relation to land use in the Mekong delta, Vietnam. *Ecosystems*, 9: 538-549.

Wilhm, J.L. and Dorris, T. C. (1968). Biological parameters of water quality criteria. *Bioscience*, 18: 477-481.

Williams, W.T. (1971). Principles of clustering. *Annual Review of Ecology and Systematics*, 2: 303-326.

Witten, I.H. and Frank, E. (2000). *Data mining: practical machine learning tools and techniques with Java implementations*. Morgan Kaufmann Publishers, San Francisco, USA, 371pp.

Won, D.H., Jun, Y.C., Kwon, S.J., Hwang, S.J., Ahn, K.G. and Lee, J.K. (2006). Development of Korean Saprobic Index using Benthic Macroinvertebrates and Its Application to Biological Stream Environment Assessment. *Journal of Korean Society on Water Quality*, 22: 768-783.

Woodiwiss, F.S. (1964). The biological system of stream classification used by the Trent-river-board. *Chemistry & Industry* 11: 443-447.

Woodiwiss, F.S. (1978). *Comparative study of biological - ecological water quality assessment methods*. Environment and Consumer Protection Service, Commission of the European Communities, UK.

Wright, J.F., Furse, M.T. and Armitage, P.D. (1993). RIVPACS - a technique for evaluating the biological quality of rivers in the UK. *European Water Quality Control*, 3: 15-25.

Wright, J.F. (1995). Development and use of a system for predicting the macroinvertebrate fauna in flowing waters. *Freshwater Biology*, 20: 181-197.

Wright, J.F., Sutcliffe, D.W. and Furse, M.T., Eds. **(2000)**. *Assessing the biological quality of fresh waters: RIVPACS and other techniques*. Freshwater Biological Association, Ambleside, Cumbria, UK, 373pp.

Wright, J.F. (2000). An introduction to RIVPACS. 1-24. In: Wright, J. F., Sutcliffe, D. W. and Furse, M. T. (Eds.), *Assessing the biological quality of fresh waters: RIVPACS and other techniques*. Freshwater Biological Association, Ambleside, Cumbria, UK.

Yang, L.F., Li, Y.W. and Qi, D.G. (1992). Community structure of aquatic insects and biomonitoring of water quality in Jiuhuahe River. *Acta Ecologica Sinica*, 12: 9-15.

Yap, S. (2005). A comparison of the macroinvertebrate communities in two Malaysian streams. *Journal of Aquatic Science*, 20: 13-26.

Yoon, I.B., Kong, D.S. and Ryu, J.K. (1992a). Studies on the biological evaluation of water quality by benthic macroinvertebrates. (I) - Saprobic valency and indicative value. *Korean Journal of Environmental Biology*, 10: 24-39.

Yoon, I.B., Kong, D.S. and Ryu, J.K. (1992b). Studies on the biological evaluation of water quality by benthic macroinvertebrates. (II) - Effects of environmental factors to community. *Korean Journal of Environmental Biology*, 10: 40-55.

Yoon, I.B., Kong, D.S. and Ryu, J.K. (1992c). Studies on the biological evaluation of water quality by benthic macroinvertebrates. (III) - Macroscopic simple water quality evaluation. *Korean Journal of Environmental Biology*, 10: 77-84.

Yoon, I.B. and Bae, Y.J. (1993). Indicator organisms and biological indices in freshwater ecosystems. *Nature Conservation*, 82: 23-27.

Yoshimura, M. (2007). Comparison of stream benthic invertebrate assemblages among forest types in the temperate region of Japan. *Biodiversity and Conservation*, 16(7): 2137-2148.

Yule, C.M. and Yong, H.S., Eds. **(2004)**. *Freshwater invertebrates of the Malaysian region*. Academy of Sciences, Kuala Lumpur, Malaysia, 861pp.

Zelinka, M. and Marvan, P. (1961). Zur Präzisierung der biologischen klassifikation der Reinheid fliessender Gewässer. *Archiv für Hydrobiologie*, 57: 389-407.

Appendices

Appendix 1. Correlation matrix of the 21 environmental variables of the Du river basin dataset (n=104). Correlation coefficients with an absolute value of at least 0.20 are marked in bold.

	Altitude	Water velocity	Width	Depth	P/R Class	DO	pH	Conductivity	Turbidity	Water Temp	BOD5	COD	Kiejdahl N	Nitrite	Total Phosphorous	Orthophosphate	Chloride	Total Fe	TN Sediment
Altitude	1.00	**-0.48**	**-0.63**	**-0.48**	**-0.69**	**0.23**	0.11	0.11	0.07	-0.13	0.04	-0.11	-0.06	-0.02	-0.02	0.18	**-0.28**	-0.18	**0.46**
Water velocity		1.00	**0.51**	**0.60**	**0.33**	0.05	**0.32**	-0.13	-0.01	0.12	0.17	**0.35**	0.01	0.09	0.04	-0.07	**0.32**	-0.17	-0.14
Width			1.00	**0.58**	**0.51**	0.05	0.06	-0.11	-0.05	0.13	0.05	0.01	0.08	-0.06	-0.08	-0.04	0.15	0.05	**-0.22**
Depth				1.00	**0.47**	0.02	**0.41**	0.01	0.03	0.01	0.22	**0.25**	-0.02	**0.26**	-0.14	-0.04	**0.27**	-0.16	-0.10
P/R Class					1.00	**-0.20**	-0.05	-0.01	0.07	0.14	-0.12	0.04	0.16	-0.10	-0.09	-0.09	0.17	0.08	-0.15
DO						1.00	**0.37**	-0.19	**0.23**	0.06	**0.21**	**0.29**	-0.13	-0.17	-0.13	**0.26**	-0.08	**-0.64**	0.06
pH							1.00	0.06	0.15	**-0.25**	**0.40**	**0.27**	-0.18	0.11	-0.03	-0.06	0.06	**-0.72**	0.18
Conductivity								1.00	0.03	-0.08	0.00	-0.08	0.15	-0.02	0.09	-0.30	-0.06	0.03	**0.22**
Turbidity									1.00	-0.06	-0.03	0.00	0.09	0.17	-0.11	0.19	0.06	-0.10	0.02
Water Temperature										1.00	0.16	0.10	**0.53**	**-0.34**	0.13	0.08	-0.02	0.02	-0.19
BOD5											1.00	**0.36**	**0.22**	0.13	0.12	0.00	-0.14	**-0.41**	-0.05
COD												1.00	-0.01	0.23	-0.02	**0.21**	0.08	**-0.23**	-0.09
Kiejdahl N													1.00	-0.10	**0.31**	0.12	-0.14	-0.07	-0.07
Nitrite														1.00	-0.10	0.06	0.17	0.10	0.07
Total Phosphorous															1.00	0.00	**-0.24**	-0.14	-0.11
Orthophosphate																1.00	-0.10	-0.11	0.08
Chloride																	1.00	0.12	-0.01
Total Fe																		1.00	-0.16
TN Sediment																			1.00

Appendix 2. TWINSPAN Two-way ordered table with the indicator species in bold

2.1 With the whole dataset

```
 Sample ID      121355571227345534115      22461462234569166556468583888228867991278889 446677377 13334471379
                9625056709589268932382812356190841767953531690443111350253448371470367927790225718428144840660

 12  Bithyni    -------1--111---111----------1------1------1---11----------------1--------------------11------  0000
 24  Simulid    -1--111--111----1---1----------11--1----------------------------1-----------1--1-----------  0000
  6  Pisidiid   ----1---------1---1--------1----11--11------1-------1------------1-----------------------  00010
 59  Branchy    ------------11--------------1---1---1---------------------------------------------1--------  00010
 51  Caenid     1-11111-11111111111111-11111-11111111111111-1111111111-----1---------------1-1-11--1111-1-----  000110
 56  Perlid     ----------1--------------1-----1-------1-----------------------------------1-------------  000110
 57  Hydropsy   11111-11111111111111111-111111111111-1-11111111111111-----1----1-------1--------1-1-11--1-1----  000110
 68  Corydali   11-1-1-111--1-1-1-11111111111---1-111111--11111-11-----------1----------1-11-1----------1-----  000110
 53  Leptoph    11----1--11-1111---111-1-111-1--------1111----1-1111---------------------1----11-1----------  000111
 49  Ephmerel   -------------------1--------------------------------------------------------------------------  001000
 50  Ephmerid   ------------------1-----------------1---------------------------------------------------------  001000
 58  Philopot   -------------1-----1--------------------------------------------------------------------------  001000
 64  Amphipte   --1-------1-1--1---111-1---------111-111-1-----1---1----------1-------------------------------  001000
 54  Heptagen   ----------1-1------1111--11----------1--1-------------------------------------------1-------  001001
 44  Helodid    -------------------------------------11-1-------11----1---------------------------------------  00101
  2  Hirudin    ------------------1----11111111-11--1-----11-11-111---------------1--1-111-1--------1---------  0011
 60  Ecnomid    ---------------------1111--------1-111-1----111----1--1------11------------------1-----------  0011
 65  Corduli    -----1-------1--------11-11111--------1------111-1-----------------------1--1-----1-1-1--1----  0011
  1  Oligocha   ---1111--111111-1--1-1111-111-1111-111-11--11-11---------------1111-1----11-111--1---1--------  0100
 39  Veliid     -------------1-1-11----------11111111--1-------1--1--------1--1----11-------11---------------  0100
 61  Gomphid    -11111111-----1---111-11111111111111-----1---1-111111-1-1----------1-1------1------1-1-1-11----  0100
 41  Herbid     -1------------------1---------------------------------1----------------------------------------  010100
 42  Psephen    ----1--1-111--111-----1--111111------1---------------11--1-1----------------11-----1-1-1-1-----  010100
 46  Elminthi   ---1111-11-1-1111-111------11-11------1-111111-111111--1---111--111-1-11-------1111111-111----  010100
 67  Libellu    111-1-111------11-1-1-11-11111--111-------1----------------------1--1-11-1-1-11-11-1--1-------  010100
  7  Corbicul   -1111111-1-111--11-1-11111111111-11111-111--1-11111--1111---------11111-1-111-1--11---1111----  010101
  9  Pachychi   11111111111111111-1111111111111111-11--11-11-111111111-1--11-----11111111-1111--11-----11----  010101
 10  Thiarid    11-11----111-11-11111-111111111111-111-------11-111--1--------1--11111--111--11--1111111------  010101
 11  Vivipari   1--1-111-111---1-1111111-1111---111111111111-111-111--111--111111-11111111-1----11-111--1-----  010101
 14  Stenothy   -1-1--1----1-11-1111-1-111---1-1111-1----1-11-1-1---1111--------1-1-11----1---1-----111111----  010101
 16  Planorb    ---------1--11-1---1111---1--11-111-111111--1-1-111--------1------1-111--1------11---11-111---  010101
 18  Palaemon   111111111111---111111-11111--1-1-11111111-111-1111111----1111111111111-11111-1-111--1-1--1----  010101
 20  Parathel   111--11111----111-1111111111-111111111111-11111111--11--1-1111---111-11111111-1111111-1-1-----  010101
 22  Hydraca    -----1---1-1--111--11111--11-111-1111-1111--111111-111-1-1-11-11--1-1-11-1111111-------------  010101
 25  Chironom   1-11111-11111111111-11-1111111111111111111111111111--1--11111-1-1111111-11111-111111111---11----  010101
 30  Atherix    --------------------11-----1---------1----------------------------------------1----------1-----  010101
 31  Mesoveli   ------1------111--1-1--1--11----1--1--1-1----1-1------1-----1--------1---1---1-----1-1--------  010101
 36  Naucorid   -----------------1-1---1--1------------------------------1-----------------------1-----------  010101
 52  Baetid     1-111111111111111-1111-11111111111111111111-111111111-11--1-11111-1111-111111111111--1-1111----  010101
 13  Littorin   -------------------1----------111-----1----------------------------1----------1-1----11-------  01011
 19  Atyid      1-111--11--11----11---11111--11111111111-1111111111111-1111--11111-11111111----11-1-11-1----  01011
 23  Tabanid    1-------------------1-1------------1-------------------1-----------------------------------1--1---  01011
 26  Tipulid    ----------------1---1-----1----1--1--1------------1-----1--------11---11-------1-------------  01011
  3  Glossiph   ---1-1-1-1111111111-11--1--1-1-1--------------1--1------------1---------11--1-11------11111----  011
  4  Erpobde    1-1-------111--1-------------------1--------------------1---1-------------1-------------------  011
 33  Pleidae    -------------------------------------1-------------------------------------1-----------------  1000
 37  Nepid      ---1---------------11--11---1----------1--------1----1111---1-1----------------------1---------  1000
 48  Hydrophi   -----------------------1--111----11-1-1--------------------1-------------111111----------1------  1000
 21  Potamid    ----1------------------------------------------------------------1---------------------------  1001
 34  Gerrid     ----111111-1-111-11--1-11111-111-11-11111-11111-1---11-----1-111-1-1-1-111-1--11-11111-1-11111  1001
 35  Hydromet   -----------------1----------------------------------------------------------------1--------  1001
 40  Notonect   -----------1--11----------11----------1------------------------------11-11--------------1-----  1001
  5  Unionid    -----------------------1----------------------------1------------------------------1--1---------  101
 32  Apheloch   --------------------------1------1-----------111-1-1----------1-----------1-----1-1--111-1-----  101
 43  Dytiscid   --------------------11-------1-1-11---1---1---11-1--11111--11-------1----1-1----1--11-1-1-----  101
 62  Calopter   --------------------1---11-1------------------------------------1-----1-------------1111----  101
 15  Pilidae    -------11-------------1-----------------1-----1----------------------1-1---1----1-1-----1-----  1100
 66  Aeshnid    --------1-----------------------------1-------------------------------------1--11----------  1100
  8  Corallan   ------------------1---1-----------1----------11-11--1------1-----11111--11-1--------1-1--1----  11010
 17  Lymnaei    ----1--------111-1--------1------1---------------1-1-------11--1--1-1--1-----1---1---111------  11010
 38  Corixid    -------------1111-11------1------1---11---1---------11-------11-------1-111111---------11----  11010
 63  Coenagri   11111----11--11--1--11-11-1---1111-1--11-------------11-1---11--111111111111111--1111-1--1----  11010
 45  Hydraeni   --------------------------------------1-----1--------------11----1----1---------------------  11011
 69  Pyralid    ------1---1-------1------------------------------1----1-11----1----1--11--1---1-----------  11011
 27  Ceratop    -------------------------------------1-1-------------1-----1-11--1---------1--------1---------  1110
 28  Stratiom   ---------------------------------------------------1-------------------1---------1--------  1110
 29  Culicid    ---------------1-----------------------------------------1--1------------------11--1------------  1110
 47  Curculio   -------------------------------------------------1--1---1------------------------------------  1110
 55  Prosopis   ------------------------------------------------1-----1--------------------1-1----1-11----  1111

                0000000000000000000000000000000000000000000000000000111111111111111111111111111111111111111111
                0000000000000000000011111111111111111111111111111111000000000000000000000000000000000000001111
                000000000111111111000000000000011111111111110000000000000000000000000000000000001111111111
                000111111000011111100000000000111111110000000111111100000111111111111111111111111100000001111
                   001111    0000110001111111000000011100001110000000101111000000000011111111111111110000001
                                  0111111000111          000111      0000111100000000111111111000111
                                   ********** TWINSPAN completed **********
```

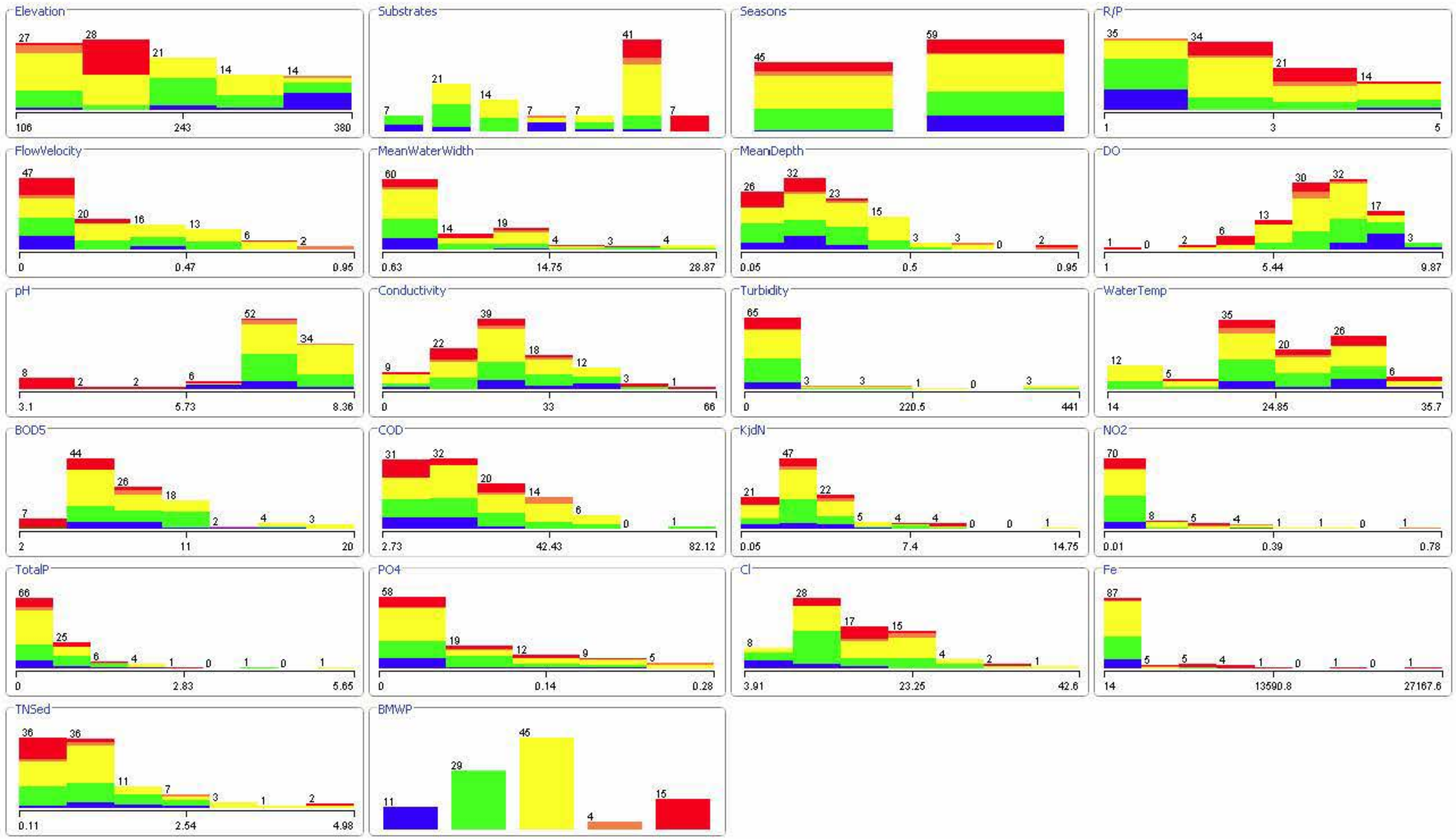
2.2 Samples taken in dry season

```
 Sample ID    111  1112   111 22  2 2222212
              18913573724724682656091358904

  5  Unionid   ------------1----1-----------  00000
 35  Naucorid  -----------1-----------------  00000
 50  Prosopis  1--------------111-----------  00000
 56  Calopter  ---------------11------------  00000
 13  Littorin  1-----------111---1-1--------  000010
 17  Lymnaei   ----------11---1-1111--------  000010
 34  Hydromet  -----------1------1----------  000010
 63  Pyralid   -----------------1--1--------  000010
 31  Apheloch  1-----------1--1-11----------  000011
 59  Corduli   ----1-------1--11--1---------  000011
  2  Hirudin   ---11-------11---------------  0001
 43  Elminthi  111---1--11-111-1111---------  0001
 15  Pilidae   -11--------------------------  001000
 51  Perlid    1----------------------------  001000
 60  Aeshnid   11---------------------------  001000
  4  Erpobde   --------11-------------------  001001
 21  Potamid   ----------1------------------  001001
 32  Pleidae   ---1-------------------------  001001
 39  Notonect  ---------1-------------------  001001
 42  Hydraeni  ------1----------------------  001001
 45  Ephmerid  -------1---------------------  001001
 58  Amphipte  ---1-1-11--------------------  001001
 62  Corydali  -1-1-1111-----------1--------  001001
 12  Bithyni   ---11---11-1---1-------------  00101
 44  Hydrophi  ---1-1--------------1--------  00101
 48  Leptoph   1-1-1---1----------1---------  00101
 49  Heptagen  -------11------1-------------  00101
  1  Oligocha  ---1--11111--11--1111--1-----  0011
 16  Planorb   --1111111----1-11111---------  0011
 29  Atherix   -------1------------1--------  0011
 53  Branchy   ----1--11---1------1---------  0011
 46  Caenid    1111111111111111--1111-------  0100
 52  Hydropsy  11-1111-1111111-11-1----1----  0100
 18  Palaemon  11111111-1111-1111--1-1-1----  0101
 55  Gomphid   -1--111---1-11111----11------  0101
 61  Libellu   -11-11----1----1-1--1--1-----  0101
 14  Stenothy  ---1111--1-11-111-11-1-1-1---  01100
 19  Atyid     -1-111111111111111---11--1---  01100
 38  Veliid    ---111-----1-11----11-1-1----  01100
  3  Glossiph  ----1---11-1--111-111--11----  01101
  8  Corallan  -----1------1--11-----1------  01101
 30  Mesoveli  ------------------111----1---  01101
  6  Pisidiid  ---1---1--1----------1-------  0111
 23  Simulid   ---1-----11------1-----1-----  0111
 25  Tipulid   1----1-1------1-----------1--  0111
 47  Baetid    11111111111-111111-1111111---  100
  7  Corbicul  ---111111111-11111-1-111-11--  101000
 20  Parathel  11111111----1111-1---111--1--  101000
 33  Gerrid    11-11111-111111-1111-1111--11  101000
 40  Psephen   ----1--1-11-11-1-----1---1---  101000
  9  Pachychi  -1-111111111-11-11-11111-11--  101001
 22  Hydraca   1--1111--1---11-----11-11----  10101
 24  Chironom  111111111111111-11-11111111--  10101
 11  Vivipari  --11-1-1-1-11----11---1-111--  1011
 54  Ecnomid   ---1--11---------1------11---  1011
 10  Thiarid   ---11--1-1111111-11111111----  110
 57  Coenagri  ----------11111111-11111-11--  110
 26  Ceratop   -------1----1--------1--1----  1110
 41  Dytiscid  --1-11------1-11-----1---11--  1110
 27  Stratiom  ------------------1-------1--  1111
 28  Culicid   -----------------1--1--1--1--  1111
 36  Nepid     ------------1--------1--111--  1111
 37  Corixid   ----------------1--111-111---  1111

               00000000000000000000011111111
               00000000000111111111100000011
               0001111111100000001110000011
                  000000110000111
                  000011

       ********** TWINSPAN completed **********
```

2.3 Samples taken in wet season

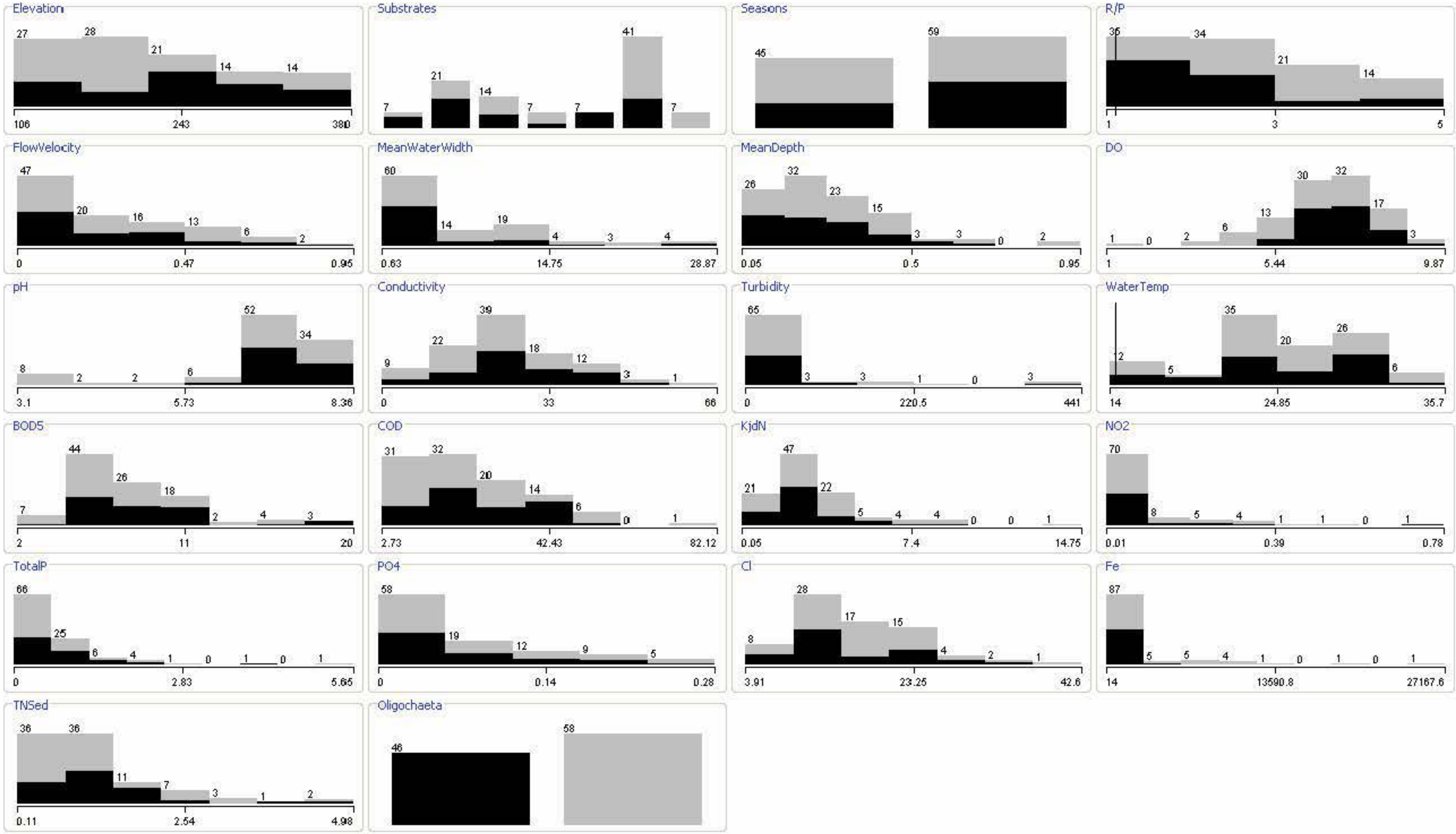
```
  Sample ID      122111345612413335 2343 3444223553    2 41122 4434556126665556516
                942635652159558189357244870178960201236143123476938893701464560702

  24  Simulid   -1----1-1--1111-----1------1---------------------------1-------1--  0000
   5  Unionid   -------------------------1---------1------------------------------  00010
  58  Branchy   --------------------1---------------------------------------------  00010
  65  Aeshnid   ---------------1------------------------1--1----------------------  00010
  30  Atherix   --------------1-1------1--------------1---------------------------  000110
  56  Hydropsy  111-111-111111111111111--11-111111111111-1111-----------1---------  000110
  48  Ephmerel  --------------1---------------------------------------------------  000111
  49  Ephmerid  ----1-------------------------------------------------------------  000111
  52  Leptoph   1---1-1111111-111111-11--11-1-1----1-11-1-------------------------  000111
  53  Heptagen  ----1----1--1-1-1--1----1------------1----------------------------  000111
  55  Perlid    -1-------1--1------1----------------------------------------------  000111
  63  Amphipte  -11-1----1--1-111-1-1-1---1--------1------------1-----------------  0010
  12  Bithyni   ------------11---------------11-111------1-------------1----------  0011
  50  Caenid    111111--1111111111111111-11111111--1111---11---1------------11----  0011
  60  Gomphid   -11-111111----1---1111-11111111-1--1111111------------1----1------  0011
  67  Corydali  11-11111---1-11111-1-1-11-111111-1-11111-11---111-----------------  0011
   6  Pisidiid  -1------------------11-------1-1------1----------------1----------  010
  32  Apheloch  -------------------11--1-11-1--11---------------11----------------  010
  59  Ecnomid   ---1------------11--1---111----1---11-----------1-----------------  010
  64  Corduli   ---------------1---1----1-111--11--1-1111--------11---------------  010
   3  Glossiph  -----1-----1111-1----111--1111----1-1-1--1----11---------1--------  01100
  10  Thiarid   11-11111---1111---1111--1-1-1111--1111111--1--1--1-11-11----------  01100
  31  Mesoveli  -1-1----11----11---1-11------1-11--1--1----1-----1--1---------1---  01100
  43  Helodid   ----------1----1-11---------1------------------------------1------  01100
  23  Tabanid   1-------------1-----1---1---------1-------------------------1---1-  01101
  35  Naucorid  ----1--------------1---------------1----1--------------------1----  01101
  36  Nepid     ----11--------1--1------------1----11--1-------------------1--11--  01101
   2  Hirudin   -1----------------111-----1-1111-1-11111--1---11-1--1--1-1--------  0111
  61  Calopter  --------------1--------1----------1-111---------------1--1--------  0111
  13  Littorin  -1--1----------1---------------------------------------1----------  10000
  37  Corixid   ----1--1-----1-1-1-11111-----1---1-------------1-11--1---1--------  10000
  38  Veliid    -1-1-------------11-1-1------1--1-----------------11---------1----  10000
   1  Oligocha  -1-111-111-111--11-111--1-11----11-11-11---------111-111--1-------  100010
  16  Planorb   -1--1----111--1111111-111-1-1-1------------------1-11--11----1--1-  100010
  41  Psephen   -----------111-----1-1111------------11111--1--1--1----------1-1--  100010
  66  Libellu   111---1-1----11----11-1-1----1-----1-111-111---1-1-1--1-11--------  100010
   7  Corbicul  --1111111111-1--11111--11-111-1-1-11111111----11--111--11--1---1--  100011
   9  Pachychi  111-111-1111111-1-11111111111-1-11-1111111----11-1111--111--1111--  100011
  11  Vivipari  11-111--11111-1111111-11111111-111-1-111-111--1111-111-11111-111--  100011
  14  Stenothy  -1--11111-1--1--1---11-1--1--11--11111-------------11--1--111-----  100011
  20  Parathel  11111-1111-1-11111-111111-11111111-1111-1111-111111-1-1111--1111--  100011
  45  Elminthi  ----11--1111-111--1--111-1-11111111---111-11----1--1--111111111---  100011
  51  Baetid    111111-11111111111111111-111111111111111-1--1-111111-11111---11---  100011
  19  Atyid     1111-1-1-1-----1-1111---111-1111111111--111-1111-11-11111111111---  100100
  25  Chironom  1111-1--1111111111111111--11-1-111-111111-111-11111-11111111-1-1--  100100
  34  Gerrid    --------11-1---111-1111---11-111-111111-1111--1111--111--1--11--11  100100
  47  Hydrophi  -------1-------1---11-------------11--11------11-11----------1----  100100
  42  Dytiscid  --------------111---1--1--1-1---1---------11---1-1-1-------1111---  100101
  18  Palaemon  1-11111111-11111-1111-1-1111111-11-111---11-111111-11111-11-111---  10011
  22  Hydraca   ---11--1-111-11111-1111-11111-111--1--1-----1--1-111-1-111-1111---  10100
  62  Coenagri  1111-111---11-1111-111-------------11---1--1--1111111-11111---1---  10100
   4  Erpobde   1-11--------1---------1-----------------------1------------1-1----  10101
  40  Herbid    ----1-1-----------------------------------------------------1-----  10101
  57  Philopot  --------------1------1-----------------------------------------1--  10101
  15  Pilidae   -----------------1-----11-1-------------11-----1----1----1--------  1011
  54  Prosopis  -----------------------1-1------------------1---------------------  1011
  17  Lymnaei   -------------------1111--1--1-----1----------------1-1-1-1---11---  1100
  29  Culicid   ----------------------1--------------------------------------1----  1100
  39  Notonect  ---------------1---1-111--------------1-------1--1--1---1---------  1100
   8  Corallan  ------------------------1-1-1111--------------11-1-11-11----11----  1101
  26  Tipulid   -------------11---11----------------------------------1111--------  1101
  68  Pyralid   -------11---1----------------1--------------1----1----1-1---111---  1101
  27  Ceratop   -----------------1-----------------------------11-----1------1-1--  1110
  28  Stratiom  -----------------------------------------------1---------------1--  1110
  44  Hydraeni  ---------------1--------------------------------------1-1----11---  1110
  46  Curculio  -----------------------------------------------------------111----  1110
  21  Potamid   ------------------------------------------------------1-----------  1111
  33  Pleidae   -------------------------------------------------1----------------  1111
                000000000000000000000000000000000000000000001111111111111111111111
                00000000000000011111111111111111111111111111000000000000000000011
                0000111111111110000000011111111111111111111100000000000000011111
                    0000111111100001111000000000000111111111001111111111111000001
                        0000001        000000111111000000011  0001111111111
                        000111         0111110000010000011       0000011111
                                  ********** TWINSPAN completed **********
```

Appendix 3. Data relation visualisation graphs for BMWP classification (Very Good (blue), Good (green), Moderate (yellow), Poor (Orange) and Very Poor (red) in the Du river basin (in total 104 instances) in relation to the 21 environmental variables.



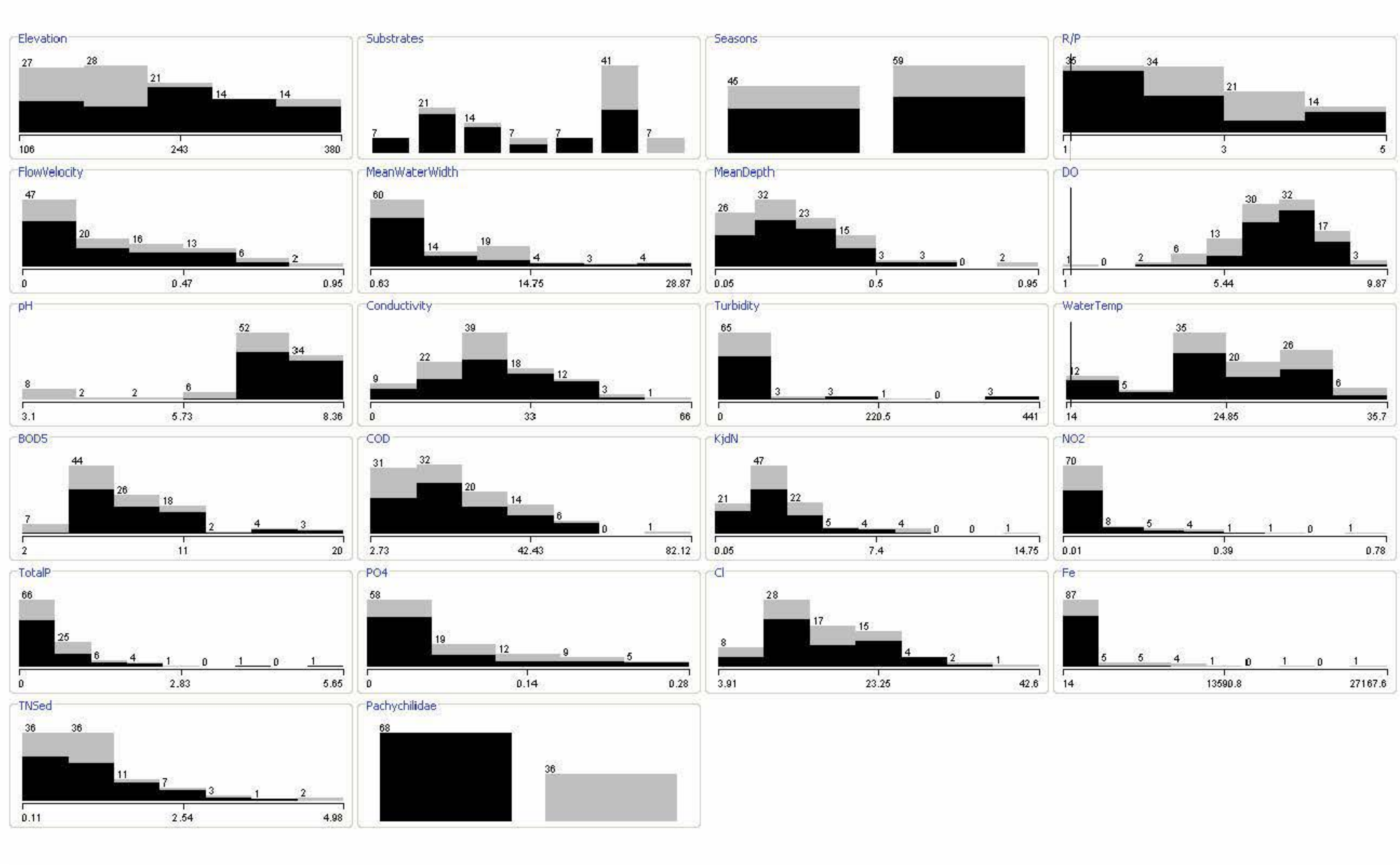
Appendix 3 (continued)

Data relation visualisation graphs for taxa presence/absence in the Du river basin (in total 104 instances) in relation to the 21 environmental variables. Taxon Oligochaeta (Class Oligochatea) present in 46 instances (black), absent in 58 instances (grey).



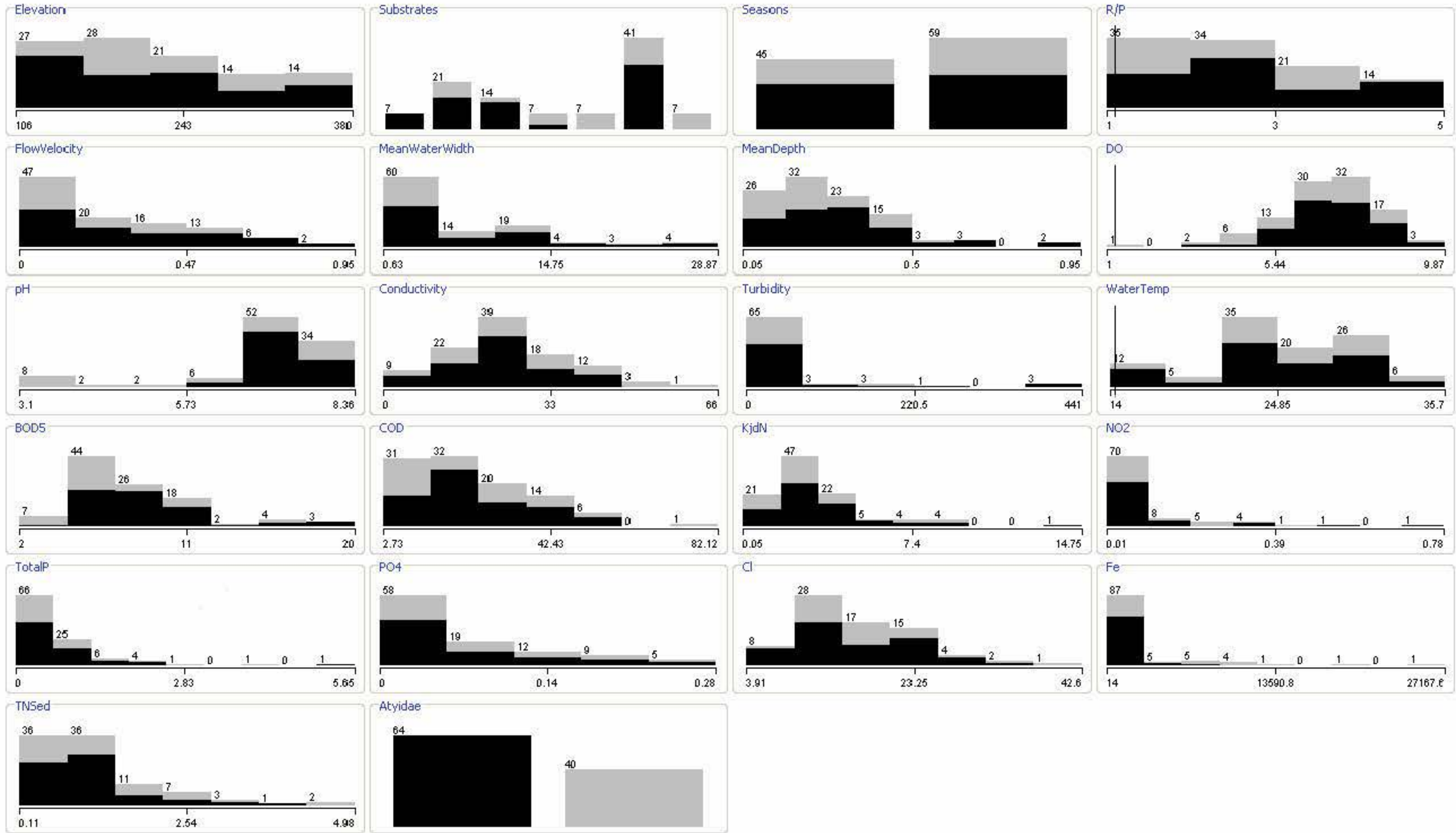
Appendix 3 (continued)

Taxon Pachychilidae (Class Gastropoda, Order Sorbeoconcha) present in 68 instances (black), absent in 36 instances (grey).



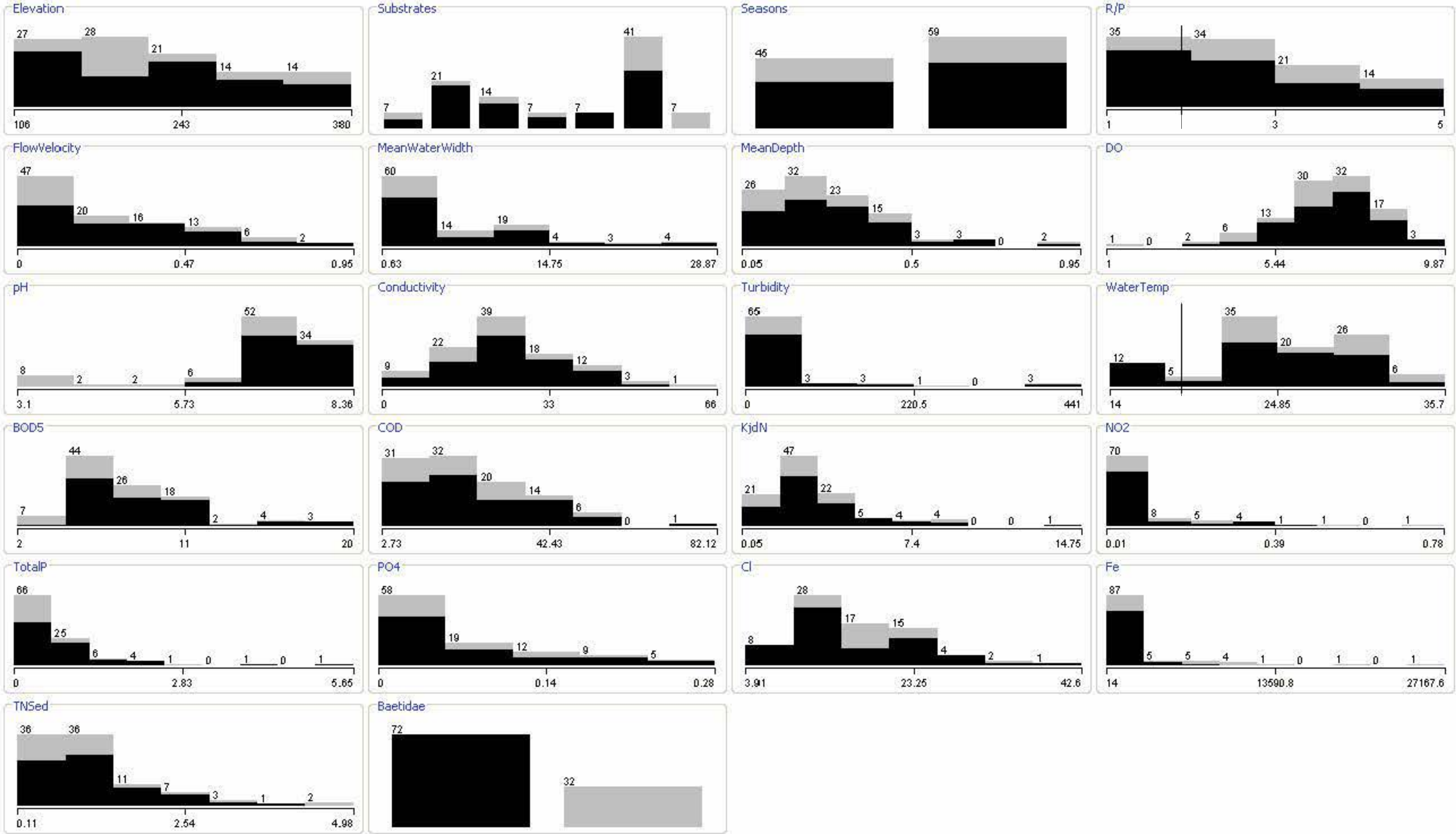
Appendix 3 (continued)

Taxon Atyidae (Class Malacostraca, Order Decapoda) present in 64 instances (black), absent in 40 instances (grey).



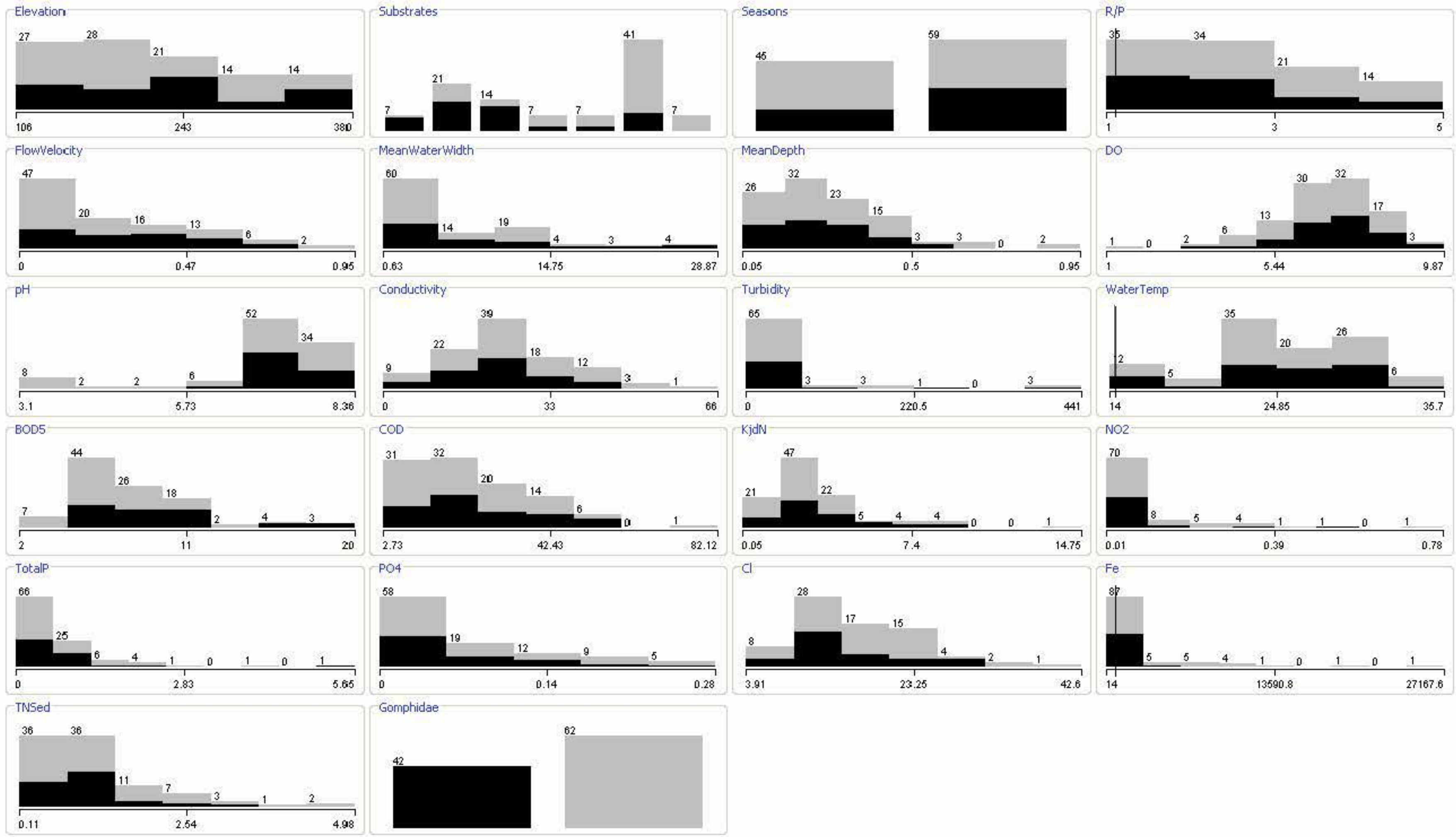
Appendix 3 (continued)

Taxon Baetidae (Class Insecta, Order Ephemeroptera) present in 72 instances (black), absent in 32 instances (grey).



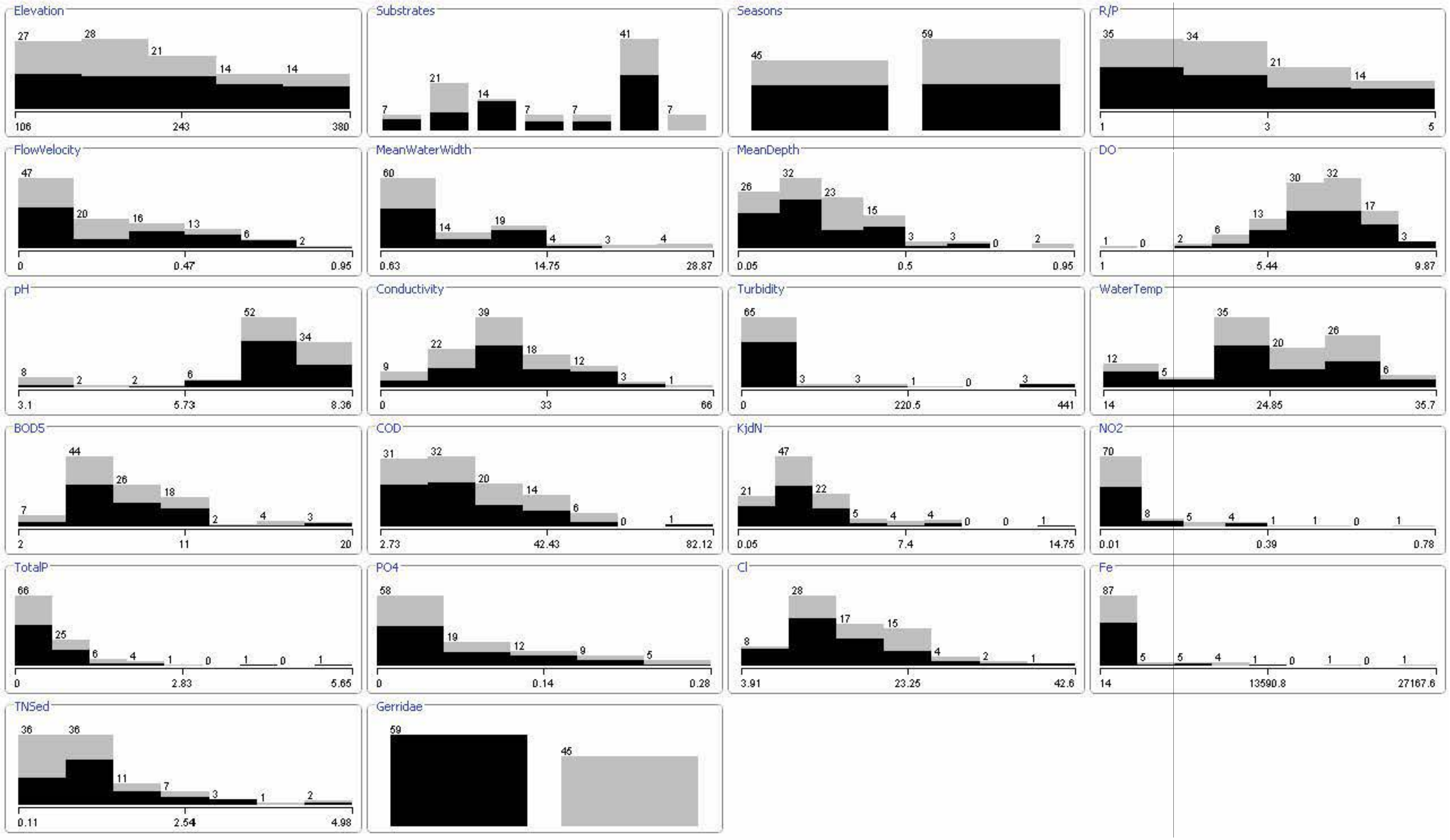
Appendix 3 (continued)

Taxon Gomphidae (Class Insecta, Order Odonata) present in 42 instances (black), absent in 62 instances (grey).



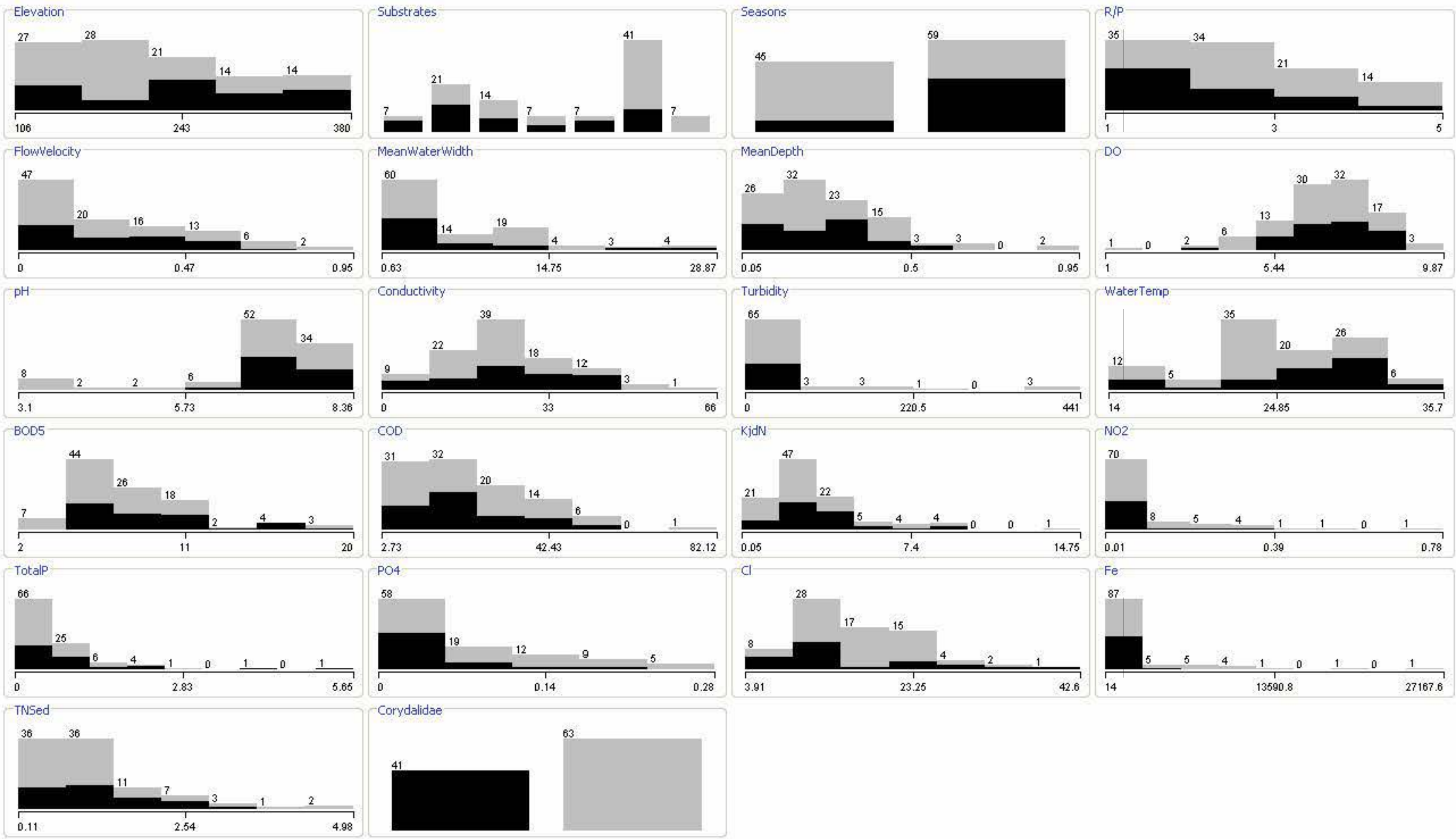
Appendix 3 (continued)

Taxon Gerridae (Class Insecta, Order Hemiptera) present in 59 instances (black), absent in 45 instances (grey).



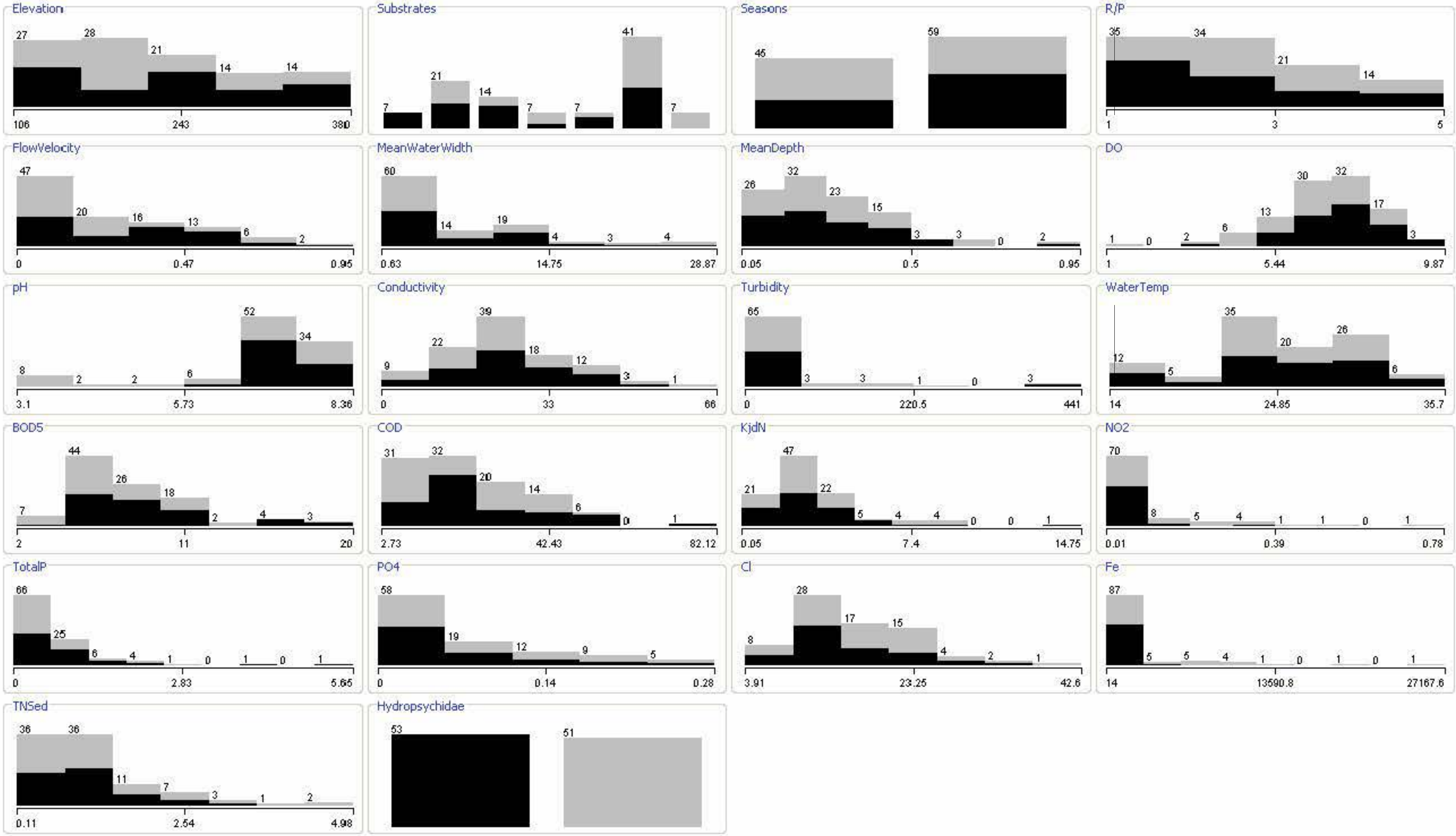
Appendix 3 (continued)

Taxon Corydalidae (Class Insecta, Order Megaloptera) present in 41 instances (black), absent in 63 instances (grey).



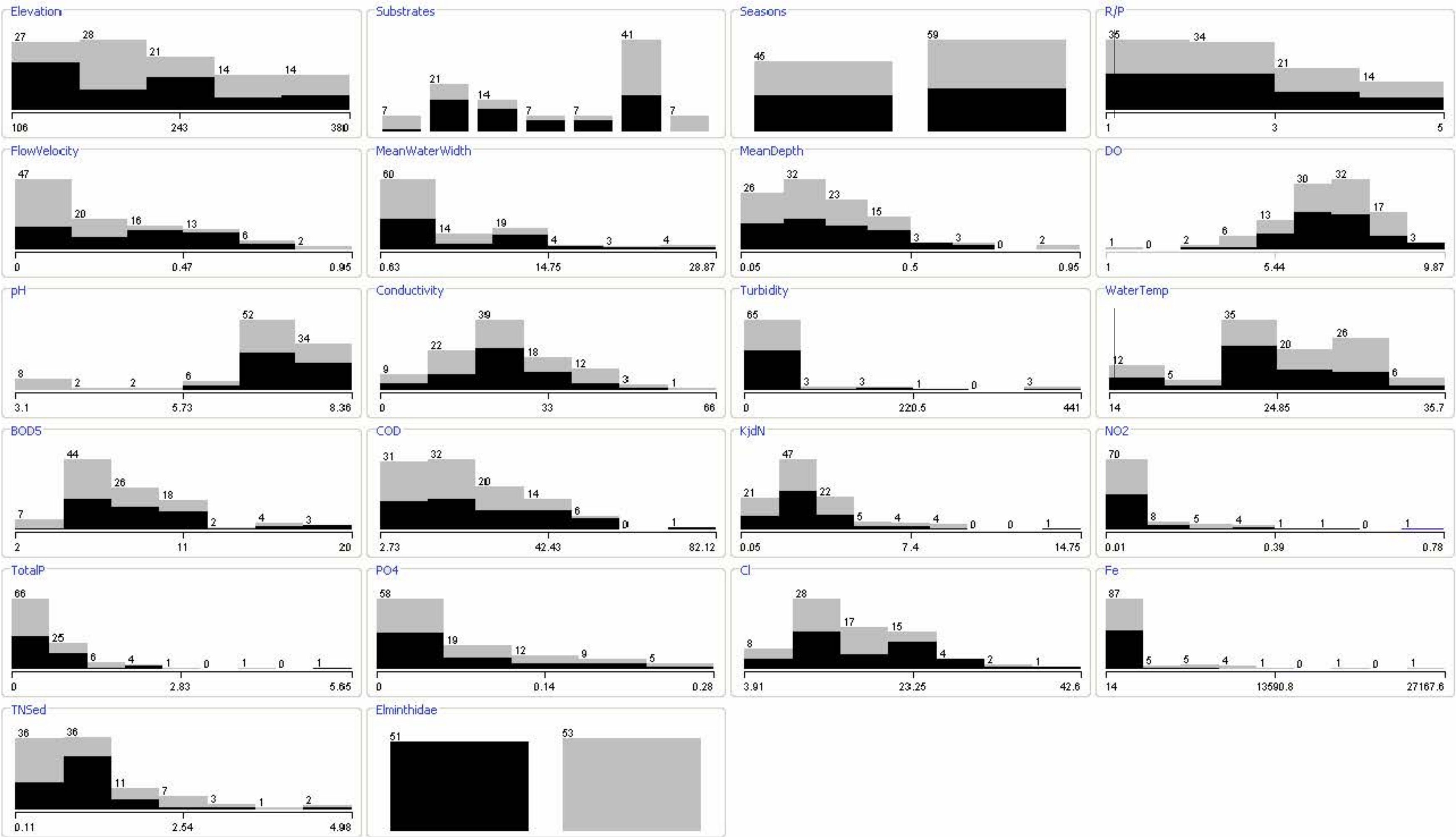
Appendix 3 (continued)

Taxon Hydropsychidae (Class Insecta, Order Trichoptera) present in 53 instances (black), absent in 51 instances (grey).



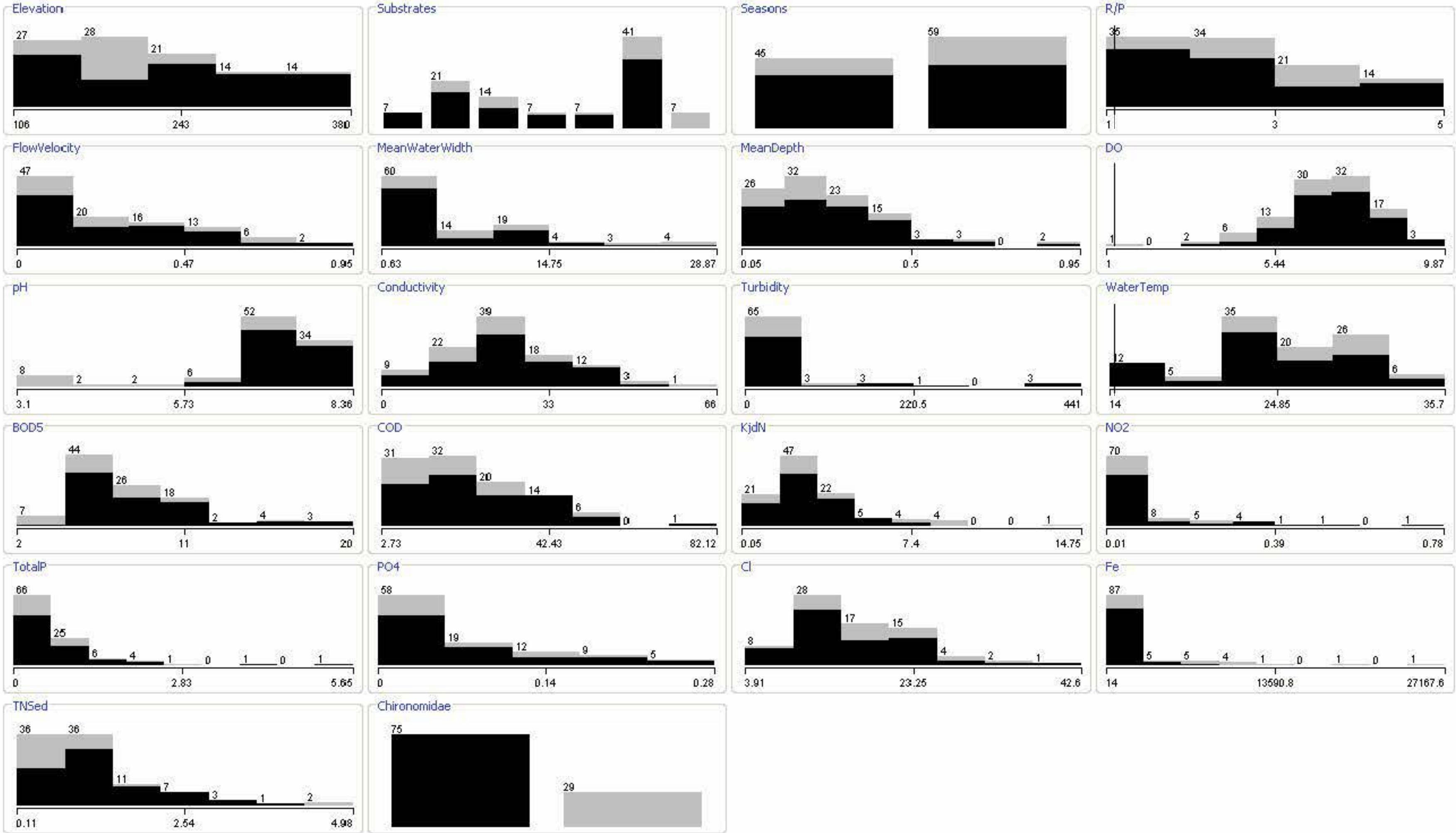
Appendix 3 (continued)

Taxon Elminthidae (Class Insecta, Order Coleoptera) present in 51 instances (black), absent in 53 instances (grey).



Appendix 3 (continued)

Taxon Chironomidae (Class Insecta, Order Diptera) present in 75 instances (black), absent in 29 instances (grey).



Summary

In Vietnam, most rivers are threatened by the consequences of the rapid socio-economic developments during the last decades. Many rivers are suffering from urbanisation, industrialisation and intensive agricultural practices. The protection of high quality freshwater resources has therefore become an important social issue in recent years. In order to attain sustainable use of surface water resources, effective water quality monitoring and assessment programme to support river management is crucial. Bioassessment based on macroinvertebrates can be applied to analyse the status of watercourses and to select sustainable restoration measures.

The Du river basin located in northern Vietnam is experiencing negative impacts of different human activities. Major disturbances include deterioration of river morphology by sand extraction and water pollution caused by mining and ore extraction practises, runoff from intensive agriculture practices and domestic waste(water) disposal (and discharge). However, there are still some pristine river segments in the upstream remote areas. Environmental variables and macroinvertebrates were measured and collected at 15 monitoring sites during the period from 2006 till 2008, based on seven monitoring campaigns. This resulted in a total of 104 samples. Environmental variables reflect the morphological river characteristics as well as physical and chemical water conditions. The data were analysed using statistical and multivariate methods and data mining techniques.

A general data analysis showed that macroinvertebrate communities in the Du river basin were poorer than reference conditions in Vietnam. Spatial and temporal analysis showed that variations in macroinvertebrate communities in the Du river were not only driven by morphological characteristics but also by water quality issues. This reflects the different anthropogenic impacts in the river basin. A relatively small temporal variation was detected that requires no remarkable modifications in the development of a bioassessment methodology for watercourses in the specific river. Multivariate analyses using CCA and Bray-Curtis cluster analysis provided a similar discrimination between pristine and impacted sites in the Du river basin.

Qualitative biotic indices based on the community composition of macroinvertebrates were calculated to analyse their relevance for biological assessment. The BMWP-Viet proved to be appropriate for use in the studied watercourses in Vietnam. The BMWP-Viet could differentiate study sites into classes ranging from very good to very poor ecological conditions. The current BMWP-Viet approach can be useful at an early stage of bioassessment application in Vietnam. However, this method should be improved by

optimising the scoring system for common taxa as well as development of more robust assessment approaches such as multimetric indices.

Data mining techniques including classification trees and support vector machines were applied to develop predictive models for BMWP-Viet as well as presence/absence of macroinvertebrate taxa (ecological indicators). The results of the presented study revealed that the developed models using SVM provided more stable results and better performance than CT. Optimised models indicate the major environmental variables influencing the presence/absence of macroinvertebrates, which in the mean time also reflect the river characteristics that river managers have to consider in their policy plans.

A decision support system, the WFD-Explorer in combination with data mining techniques, was implemented to link human activities with the ecological river conditions and analyse the relevance of several restoration options. The river condition was evaluated by the biological index BMWP-Viet predicted by classification trees based on chemical characteristics provided by the WFD-Explorer. Elimination of point sources from ore extraction practises and decentralised domestic wastewater collection and treatment proved to be the most effective measure to improve the ecological condition of the Du river.

However, restoration actions should integrate technical and management issues, especially awareness raising for involvement of general public to the conservation of the river ecosystem. Biological monitoring and assessment based on macroinvertebrates can constitute a step towards the development and implementation of effective monitoring strategies for aquatic ecosystems in Vietnam.

Samenvatting

In Vietnam worden de meeste rivieren bedreigd als gevolg van de snelle sociaal-economische ontwikkelingen gedurende de voorbije decaden. Veel rivieren zijn verstoord door verstedelijking, industrialisatie en intensieve landbouw. De bescherming van zoetwatervoorraden van goede kwaliteit is daarom een belangrijk maatschappelijk thema geworden. Voor het bereiken van een duurzaam watersysteemgebruik, zijn goede waterkwaliteits-monitoringsprogramma's onontbeerlijk. Biologische beoordelingssystemen gebaseerd op macro-invertebraten kunnen gebruikt worden om de status van waterlopen te bepalen en duurzame herstelplannen op te stellen.

De Du-rivier ligt in het noordelijk deel van Vietnam en wordt gekenmerkt door allerlei negatieve humane impacten. De voornaamste impacten zijn gerelateerd aan morfologische degeneratie als gevolg van zandextractie, waterverontreiniging door ontginningsactiviteiten, uitstroming vanuit de landbouw en het storten (lozen) van afval(water). Desalniettemin zijn er tevens een aantal nagenoeg natuurlijke riviersegmenten in bovenstroomse afgelegen gebieden. Milieuvariabelen en macro-invertebraten werden opgemeten en verzameld op vijftien locaties gedurende de periode 2006-2008. Deze metingen werden tijdens zeven achtereenvolgende staalnamecampagnes uitgevoerd, wat resulteerde in de gegevensbank met 104 stalen. De betreffende milieuvariabelen weerspiegelen zowel de morfologie als waterkwaliteit van de onderzochte rivieren. De gegevens werden geanalyseerd aan de hand van multivariate statistische methoden en data mining technieken.

Een algemene gegevensanalyse gaf aan dat macroinvertebratengemeenschappen in het Du-rivierbekken een lagere diversiteit vertoonden in vergelijking met Vietnamese referentierivieren. Spatio-temporele analyse toonde ook aan dat de variatie in de gemeenschappen niet alleen gerelateerd was aan morfologische aspecten, maar dat tevens de waterkwaliteit hierin een rol speelt. Dit weerspiegelt het brede gamma aan verstroringsbronnen binnen het betreffende rivierbekken. De analyse toonde ook aan dat de seizoenale variatie relatief gezien onbelangrijk is voor de uitgeteste biologische beoordelingsmethoden in deze specifieke rivier. Multivariate analyses met CCA en Bray-Curtis clustering analyse gaven daarnaast een vergelijkbaar onderscheid weer tussen natuurlijke en verstoorde locaties van het Du-rivierbekken.

Kwalitatieve biotische indexen gebaseerd op de gemeenschapssamenstelling van macro-invertebraten werden berekend met het oog op de analyse van hun relevantie voor biologische beoordeling. De BMWP-Viet bewees relevante resultaten op te leveren voor het analyseren van Vietnamese waterlopen, gezien deze index een goed onderscheid kan maken tussen rivierkwaliteitsklassen gaande van zeer verstoord tot natuurlijk. Daarom kan op korte termijn reeds de huidige BMWP-Viet nuttig zijn voor biologische beoordeling in Vietnam. Deze methode dient in de toekomst evenwel geoptimaliseerd te worden door betere scores te ontwikkelen voor de algemeen voorkomende lokale taxa en tevens door het uitwerken van een meer robuuste multimetrische index.

Data mining methoden zoals klassificatiebomen en support vector machines (SVMs) werden gebruikt om voorspellingsmodellen te ontwikkelen van BMWP-Viet en voor de aan- en afwezigheid van macro-invertebratentaxa (ecologische indicatortaxa). De resultaten van voorliggende studie wezen uit dat SVMs stabielere en meer betrouwbare resultaten opleverden dan klassificatiebomen. Geoptimaliseerde modelen gaven weer welke de belangrijkste rivierkarakteristieken zijn voor de aan- en afwezigheid van de macro-invertebraten, wat op zijn beurt een indicatie is voor het type van maatregelen waar waterbeheerders bijzondere aandacht dienen voor te hebben.

Daarnaast werd een beslissingsondersteunend systeem, de Kaderrichtlijn Water Verkenner, geïmplementeerd om menselijke activiteiten te reletateren aan de ecologische riviercondities en het mogelijk effect van rivierherstelmaatregelen te analyseren. De ecologische riviercondities werden geëvalueerd aan de hand van de Ecologische Kwaliteitscoëfficiënt (zoals deze in de Europese lidstaten worden aangewend binnen de Europese Kaderrichtlijn Water). De bekomen resultaten werden daarnaast ook getoetst en vergeleken met de BMWP-Viet. Gedecentraliseerde afvalwaterverzameling en –zuivering bleken op basis van de analyse de meest effectieve maatregelen te zijn om de ecologische conditie van de Du rivier te verbeteren.

Desalniettemin dient beklemtoond te worden dat technische maatregelen dienen gekoppeld te worden met een algemene bewustmakingsstrategie van de lokale bevolking om riviersystemen succesvol te verbeteren in het betreffende gebied. Biologische monitoring en beoordeling aan de hand van macro-invertebraten kan in deze context een eerste belangrijke stap zijn naar de ontwikkeling en toepassing van effectieve monitoringsstrategieën voor aquatische ecosystemen in Vietnam.

Curriculum vitae

Curriculum vitae

Family name	: Hoang
Given name	: Thi Thu Huong
Date and place of birth	: 15/March/1973, Hanoi (Vietnam)
Marital status	: Married, 1 child
Address (private)	: 21 Loduc, Haibatrung, Hanoi, Vietnam
Address (work)	: Institute for Environmental Science and Technology (INEST) Hanoi University of Technology (HUT) 1 Dai Co Viet, Haibatrung, Hanoi, Vietnam E-mail: huonghtt@yahoo.com huonght-inest@mail.hut.edu.vn Tel: 0084(4) 38681686/ Ext.12 Fax: 0084(4) 38693551

Education

1990-1995: Chemical Engineer, option Environmental Protection and Rational Resources Utilisation, Odessa State Academy of Refrigeration – Ukraine (5 year full time academic training).

Thesis 'Optimisation of thermal and humid transferring processes in ejector scrubber by mathematical models'. Promoter: Prof. Dr. ir. V.P. Onishenco

2000-2001: Master of Applied Science in Soil and Water, Adelaide University, Australia (one and half year full time academic training)

Thesis 'Predicting freshwater habitat conditions by the distribution of macroinvertebrates using artificial neural network'. Promoter: Prof. Dr. F. Recknagel

2005-2009: PhD student

Laboratory of Environmental Toxicology and Aquatic Ecology,

Department of Applied Ecology and Environmental Biology,

Ghent University, Belgium.

Certificate

2002. Training course "Environmental Management Accounting". Vietnam Cleaner Production Center, Hanoi.

2004. Training course "Management of Hazardous Substances and Goods" University of Applied Sciences Basel (FHBB), Institute of Environmental Technology, Switzerland.

2006. Training course "Project Cycle Management", Hanoi University of Technology VLIR-HUT Institutional University cooperation program, Hanoi.

2006. Training of Trainers on "Management of Urban Environmental Infrastructure and Services", Asian Institute of Technology, SEA-UEMA Project and AIT Extension, Bangkok, Thailand

2007. Training course "Biomaker and ecotoxicological test", VEPA, Hanoi Vietnam

Work experience

1996 –12/1999: Environmental Consultant

Centre for Environment and Technology CET/FPT

Ministry of Science, Technology and Environment, Vietnam

2001 - 2003: Researcher.

Institute for Environmental Science and Technology (INEST)

Hanoi University of Technology (HUT)

2003-to present: Lecturer.

Department of Environmental Enginering

Institute for Environmental Science and Technology (INEST)

Hanoi University of Technology (HUT)

Publications

Hoang, T.H, Lock, K., Mouton, A., Goethals, P. Application of decision trees and support vector machines to model the presence of macroinvertebrates in rivers in Vietnam. Ecological Informatics (Submitted)

Hoang, T.H, Lock, K., Dang, K.C., De Pauw, N., Goethals, P. Spatial and temporal patterns of macroinvertebrate communities in the Du river basin in northern Vietnam. Limnology (in preparation)

Hoang, T.H, Mouton, A., Lock, K., Goethals, P. Combination of data-driven and knowledge-based models to support river management in the Du River basin (Vietnam). Ecological Informatics (in preparation)

Hoang, T.H, Lock, K., De Pauw, N., Goethals, P., Dang, K.C. (2009). Application of multivariate statistical analysis to study the human impact on the water quality of the Du tributary in the Cau river basin. *Journal of Science and Technology – technical universities (Vietnam)* 72A: 88-95.

Marshall, J., Hoang, H., Choy, S. and Recknagel, F. (2003). Relationships between habitat properties and the occurrence of macroinvertebrates in Queensland streams (Australia) discovered by a sensitivity analysis with artificial neural networks.

Hoang, H., Recknagel, F., Marshall, J. and Choy, J. (2002). Elucidation of hypothetical relationships between habitat conditions and macroinvertebrate assemblages in freshwater streams by Artificial Neural Networks. p. 179-192. In: F. Recknagel

(Eds.) *Ecological informatics: understanding ecology by biologically-inspired computation*. Springer-Verlag, Berlin.

Hoang, H., Recknagel, F., Marshall, J. and Choy, J. (2001). Predictive modelling of macroinvertebrate assemblages for stream habitat assessments in Queensland (Australia). *Ecological Modelling* 195: 195-206.

Attended symposia, conferences and workshops (presenter is underlined)

Hoang, T.H, Lock, K., Goethals, P. Application of decision trees and support vector machines to model the presence of macroinvertebrates in rivers in Vietnam. Session: Combining data driven and knowledge-based modelling. International Conference on Ecological Informatics ISEI 6, Cancun, Mexico, December 2008.

Hoang, T.H., Dang, K.C., De Pauw, N. Goethals, P. Application of biological monitoring in assessing pollution impacts on river water quality in Vietnam. Promoting Development of Environmental Science and Industry in Vietnam, Hanoi, Vietnam, September 2008.

Hoang, T.H., Goethals, P., Dang, K.C., Mai, D.Y, De Pauw, N. Impact assessment of environmental conditions on macroinvertebrate communities in northern streams in Vietnam based on classification tree techniques. Oral Session 16. Ecological Monitoring and Assessment. EcoSummit 2007, Beijing, China, May 2007.

Hoang, T.H., Dang, K.C., Goethals, P. Assessing the impact of mining activities on river water quality and the biota in the Du river. International Workshop on GeoEcology and Environmental Technology, Hanoi, Vietnam, November 2007.

Educational activities during candidature

Supervision of Master scripts

Nguyen Thi Hoa (2006-2007). Script “Ecologcial impact assessment based on macroinvertebrate in the Du river (Vietnam)” Master in Environmental Sanitation, Ghent University, Gent, Belgium. Promoters: Prof. Dr. N. De Pauw and Dr. ir. P. Goethals. Tutor. Hoang T.T. Huong.

Anne Temmerman (2008-2009). Script “Ecological evaluation of river in northern Vietnam”. Master of Bioscience Engineering: Environmental Technology, Ghent Univrsity, Gent, Belgium. Promoters: Prof. Dr. ir. P. Goethals and Dr. ir. Ans Mouton. Tutor. Hoang T.T. Huong.

Lecturing in international training courses

Training course: Dynamic modeling and simulation of water quality - The Integrated Urban Water Management System. 17-19 June, 2008, Sofia, Bulgaria. Course on ‘Monitoring, modelling and assessment of freshwater ecosystems’ (6h). Program organiser: WECOVAL project.