Promoters:

Prof. Dr. ir. Peter Goethals

Aquatic Ecology Research Unit

Department of Animal Sciences and Aquatic Ecology

Ghent University

Dr. ir. Gert Everaert

Flanders Marine Institute

Local project promoter:

Prof. Dr. Luis Dominguez

Centro del Agua y Desarrollo Sustentable

ESPOL (Ecuador)

Dean

Prof. Dr. ir. Marc Van Meirvenne

Rector

Prof. Dr. ir. Rik Van de Walle

This study was performed in the context of the VLIR Ecuador Biodiversity Network Project and was funded by the Vlaamse Interuniversitaire Raad-Universitaire Ontwikkelingssamenwerking (VLIR-UOS). The author was supported by the special research fund of Ghent University (BOF).

MSc. Minar Naomi Damanik Ambarita

Ecological impact assessment of land use on river systems

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

Dutch translation of the title:

Ecologische impact evaluatie van landgebruik op riviersystemen

To cite this work:

Damanik Ambarita, M.N., 2018. Ecological impact assessment of land use on river systems. PhD thesis. Ghent University, Belgium.

ISBN: 9789463570831

Acknowledgements

I would like to express my sincere gratitude to the people and sponsors who have supported me during and in finalizing this PhD studies.

I really thank my promoter Prof. Dr. ir. Peter L.M. Goethals who has given me the opportunity to continue doing a PhD after promoting me during my masters' programme. I thank you for giving me the direction of my PhD, for the freedom and confidence in carrying my work. I also want to thank my second promoter Dr. ir. Gert Everaert who has embarked in the structure of my PhD at a crucial stage. Without your guidance and quick replies, I would not be able to reach this day, not even to finalize my thesis. I thank both my promoters Prof. Goethals and Dr. Everaert for their continuous guidance and support in all stages of my PhD life.

I would like to express my gratitude to my examination committee Prof. Dr. Koen Dewettinck, Prof. Dr. Stefaan De Neve, Prof. Dr. Olivier Thas, Prof. Dr. Martin Volk, Dr. Aymere Awoke Assayie and Prof. Dr. Luis Dominguez for their constructive remarks on my thesis. Their remarks have critically improved the quality and clarity of my thesis. Several external pairs of eyes can indeed shed beneficial lights on my path. Special thanks are due to Prof. Dominguez who has hosted and supported my sampling team members and me in Ecuador.

I also would like to thank other people and sponsors who have helped me finishing my work. First, I would like to thank Dr. Pieter Boets who was a supportive co-author in my papers, giving suggestions on data analyses and writing. I would like to thank and appreciate Sigrid, Marianne and Nancy for their immense and swift support in any matters: administrative, logistics, and talks. Thank you for the support in preparing for the sampling campaign and beyond until the preparation for my defense. You are always so very helpful and friendly anytime I needed, even for things that might not be your tasks. Thank you for that. Marianne, thank you for the jokes, I will surely miss them in the future! Thank you for my sampling team Tien, Eurie, Peace, Koen and Natalija. It was a great time working together with you guys. Unforgettable moments... Thank you Tien for your friendship and support, always be there sharing your experiences and stories with me. Thanks Eurie for your lots of help. Also for the people from CADS-ESPOL Ecuador: Patricia, David and others that I cannot mention one by one. This work could not have been possible without financial support from VLIR Ecuador Biodiversity Network project that has supported the sampling campaign and to the Special Research Fund of Ghent University for providing me with PhD scholarship.

Thanks to my other friends-colleagues in the lab, still or not anymore there: guys, the order doesn't matter here ☺. Wout (I'll miss checking your baking skills and jokes!), mi hermano Niels (never forget OASIS – the band), Sacha (with our last-stage stories), Jana, Arne, the 'my name is Long' with your help and never-ending jokes, Selamawit, Tu Tri, Natalia and your kind soft heart, Stijn (you got me once with your joke), Shewit, Daniel, Ratha (and your prayers), Ruben, Tan and Elina. Guys, I enjoyed various kinds of jokes with you. Nice stories to share with (fun, sad, happy, all kinds). Prof. Colin Janssen, Prof. Karel D.S., Ilias, Emmanuel, Cecilia, Dimitri, Samuel, Josef, Karel V, Jonathan, Jana A. and Abegail. Gisèle (and our toilet schedule), Emmy (thanks for the nuts), Jolien (who will open a new binding service soon) and Marc. Sylvie and Veerle who are always so very friendly since my masters' time, I thank you both for your support and our talks.

I would like to express my gratitude to my ICCG community. I especially indebted to Pastor Leo, my biggest moral support and spiritual advisor whom I considered my father here. Without your continuous support, I would have stopped long time ago. Thank you my ICCG family for the warm and encouraging environment. Thank you Pieter-Mariale and the Dros, you have welcomed me as your additional family member. Thank you Sandy for our talks and times together, you are like a mother to me; and of course thank you Craig. Thank you Kasun-Grace and Yohan, I enjoyed my times and talks with you always. Thank you Jeverton-Juliana, I enjoyed our supporting each other always. My music team, I will surely miss our music practice and service together. Thank you Carlos-Yuliana and Caleb, I will surely miss my times at your house. Thank you Paulo-Thayenny, we can do more games night at your place later. Thank you Geoffrey-Kseniya, it was nice talking with you and spending time together. Thank you Theo for helping me with the thesis and the stories. Thank you Anneke, Janneke, Willem-Inge and Vik, Martine, Tita Donna, Amie, Arllet, Claire-Bem, Joyce, George, Blessings, Lillian, Antonio-Gislene, Karel and family, Erwin, and many more that I cannot mention here.

I want to thank my other supporters here. I really thank Rudo for his continuous support, comfort, listening ears, patience, understanding and prayers especially towards the last stage of my PhD. My friends Christine-Sunny, teh Nuri, teh Okky, mbak Angga and our hang out times and stories, ci Lanny-ko Juliando and family, wa Ann-wa Francis and family, my other Indonesian friends and Indonesian student association PPI Ghent that I cannot name individually. My friends who have helped taking me back and forth to hospital after surgery, doing my grocery shop and even cooking. You won't be forgotten for sure. I sincerely apologize in case I've forgotten to mention some names here.

Thank you my father, sisters and brother kak Itin and her family, kak Ika, kak Mery, bang Uhum and family and Maria, and my whole extended families who have supported me wholeheartedly through talks and prayers and comforts. This is for you Pak… Selamat ya Pak, akhirnya tercapai juga harapan Bapak…

Last but foremost, I thank my Lord God Jesus Christ. Without you Lord, I cannot handle all these struggles, all the pain, all the heartaches I experienced during this PhD. For the many times I wanted to quit, the many times I have lost my motivation, the lack of sleep during the last weeks, You carried me back up and strengthened me over and over again. Thank you for all these experiences, even the difficulties that You used to develop my character. Thank You for being so true and faithful to me. **Soli Deo Gloria.**

Minar Naomi Damanik Ambarita
Ghent, 15 February 2018

Every good and perfect gift is from above, coming down from the Father of the heavenly lights, who does not change like shifting shadows.

James 1:17

Contents

List of abbreviations

A

ABI	Andean Biotic Index
ACF	auto correlation function
AICs	Akaike information criteria
ANN	artifical neural network
ASPT	average score per taxon
AUSRIVAS	Australian River Assessment System

B

BBI	Belgian Biotic Index
BBN	Bayesian belief networks
BIIM	Biotic Integrity Index of Macroinvertebrates
BIP	Biotic Index of Pollution
BMI	Benthic Multimetric Index
BMP	best management practices
BMPS	Biotic Monitoring Patagonian Streams
BMWP	Biological Monitoring Working Party
BMWP-Col	Biological Monitoring Working Party-Colombia

C

CA	correspondence analysis
CART	classification and regression trees
CCA	canonical correspondence analysis
CERA	Calidad Ecológica de Ríos Altoandinos
COD	chemical oxygen demand
CPOM	coarse particulate organic matter
CR	continuous regression
CV	cross validation

D

db-RDA	distance-based redundancy analysis
DFA	discriminant function analysis
DO	dissolved oxygen

E

EC	European Commission
EPT	Ephemeroptera Plecoptera Trichoptera
EQR	ecological quality ratio

F

FAO	Food and Agriculture Organization
FBI	Family Biotic Index
FFG	functional feeding group
FP	field protocol
FPOM	fine particulate organic matter

G

GAM	generalized additive model
GIS	geographical information system
GLM	general(ized) linear model
GMMI	Guapiac, u-Macau Multimetric Index

H

HFBI	Hilsenhoff Family Biotic Index

I

IBE	Indice Biotico Esteso
IBE-IOC	Índice Biótico Estendido - Instituto Oswaldo Cruz
IBG	Indice Biologique Global
IBIAMA	Index of Biotic Integrity for Aquatic Macroinvertebrate Associations
IBIAP	Biotic Integrity Index using aquatic invertebrates
IBIRP	Index of Biotic Integrity for the Río de la Plata
IBMG	Indice Biotique Macroinvertébrés de Guyane
IBPAMP	Biotic Index for Pampean rivers and streams
ICERN-MAE	Índice de Calidad Ecológica con base en macroinvertebrados acuáticos para la cuenca del rio Negro

IMEERA	Índice Multimétrico del Estado Ecológico para Ríos Altoandinos
IMRP	Index for Macroinvertebrates of Pampean Rivers

L

LICOI	Limnological Conditions Index
LOWESS	locally weighted scatterplot smoothing
LS	least squares

M

MAE	Ministerio del Ambiente del Ecuador
MAGAP	Ministerio de Agricultura, Ganadería, Acuacultura y Pesca
MISB	Multimetric Index for Serra da Bocaina
MLC	maximum-likelihood classification
MMIF	Multimetric Macroinvertebrate Index Flanders

N

NEPBIOS	Nepalese Biotic Score
NLSMI	Neotropical Low-land Stream Multimetric Index
NMDS	nonmetric multidimensional scaling

P

(P)CRBI	Proposed Costa Rican Biotic Index
PCA	principal component analysis
PCoA	principal coordinate analysis
PO	proportional odds
PPPMI	Piabanha-Paquequer-Preto Multimetric Index

R

RCC	river continuum concept
RDA	redundancy analysis
RHS	river habitat survey

S

SASS	South African Scoring System
SIGNAL	Stream Invertebrate Grade Number Average Level
SOMI	Multimetric Index for Serra dos Orgaos

SWOT	strengths, weaknesses, opportunities and threats

T

TDS	total dissolved solids
Total N	total nitrogen
Total P	total phosphorus
TSI-BI	Trophic State Index for Benthic Invertebrates

U

UNEP	United Nations Environment Program
USACE	United States Army Corps of Engineers
USEPA	United States Environmental Protection Agency

V

VIF	variance inflation factor
VIS	visible

W

WA	weighted averaging
WFD	Water Framework Directive
WHO	World Health Organization

Chapter 1: General introduction

1.1 Land use

Land use is a description of how people use the land in relation to socio-economic activities and purposes (Fisher *et al.*, 2005), which can be characterized by the actual goods and services obtained as well as the type of management applied on the land (LADA, 2008). The Food and Agriculture Organization (FAO) classified land use system into nine ecosystem types (i.e. forest, grasslands, shrubs, agriculture, urban, wetlands, sparsely vegetated areas, bare areas and open water). Globally, forestry constitutes the largest area, followed by agricultural and bare areas (30%, 18% and 17%, respectively); while open water constitutes the lowest percentage (2%, Fig. 1.1, showing different types and intensities of global land use). Global land use information is generally scarce and not regularly updated (LADA, 2008).

Different types of agricultural activities, livestock grazing, settlement and construction, reserves and protected land, and timber extraction are the most important global land uses spatially and economically. These land uses have been transforming global land cover (Turner *et al.*, 1994). For millennia, humans have modified 75% of terrestrial surface into cropland and pasture to provide food and bioenergy, and only less than a quarter of terrestrial surface remains as wildlands (Erb *et al.*, 2013; Kehoe *et al.*, 2015). The occurring land use intensification can reduce biodiversity and might threaten ecosystem services (Allan *et al.*, 2015; Kehoe *et al.*, 2015). However, spatial patterns of land use intensity are not well understood (Erb *et al.*, 2013; Kehoe *et al.*, 2015; Kuemmerle *et al.*, 2013; Vaclavik *et al.*, 2013). Fortunately, researchers are getting interested in studying this pattern not only globally but locally as well (Li *et al.*, 2017; Robillard and Kerr, 2017; Roder *et al.*, 2015; van der Zanden *et al.*, 2016).

Some land has a single use, but multiple uses are also common especially in large areas such as river catchments (Crétaz and Barten, 2007; Guhathakurta, 2005). Some catchment may have different uses such as urban, forest and agricultural (Ferreira *et al.*, 2016; Tu, 2009). Some land may have multiple uses simultaneously, such as plantation forestry that can be used for recreation, grazing and hunting at the same time. Other land may be used alternately, such as a reservoir that provides flood control in the spring and electricity in the winter. Land use information is required in policy and planning purposes (Fisher *et al.*, 2005).

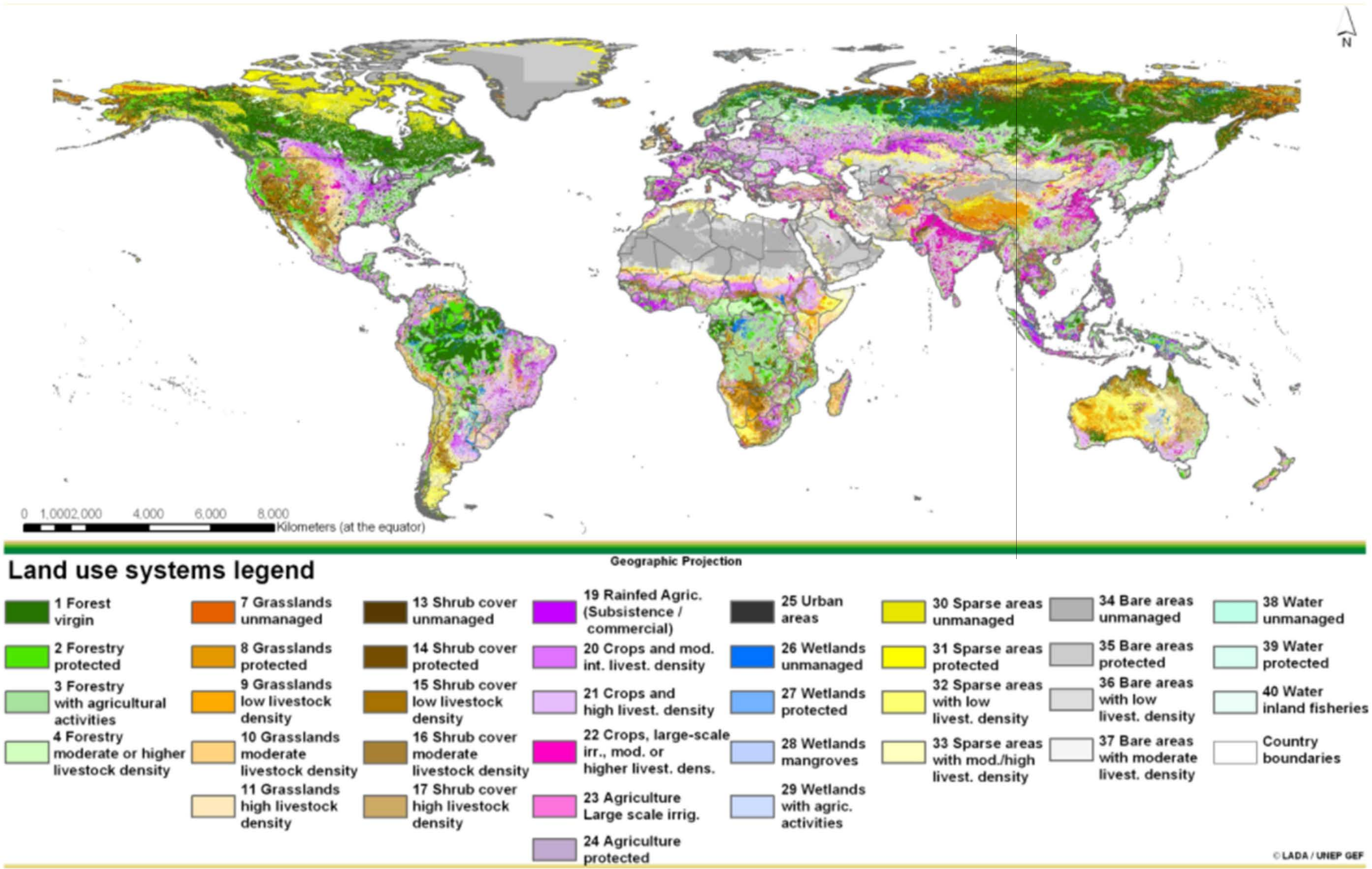


Figure 1.1 Land use and land use intensity distribution

1.2 Water quality

Due to its complexity, it is difficult to give a simple definition of water quality. Here, two profound definitions of water quality according to Bartram and Ballance (1996) and Chapman (1996) are provided.

According to Bartram and Ballance (1996), water quality is the suitability of a water body to sustain various uses or processes. Each use will have certain requirements for the physical, chemical or biological variables of the water; therefore water quality can be defined by a range of variables which limit the use of water. Several countries or regions have developed their water quality criteria to be applied nationally or regionally. For example, United States Environmental Protection Agency (USEPA) has formulated water quality criteria for the use in United States. Besides their use in United States, these water quality criteria have been referred to in various water quality studies ((National Academy of Sciences and National Academy of Engineering, 1972; USEPA, 1986), Table 1.1). Countries such as Ecuador has also defined its water quality criteria based on its own research ((Ministerio del Ambiente del Ecuador - MAE, 2015), Table 1.1). Different countries might set dissimilar environmental standards, and the level of required specifics and details might also be different. Many uses have some common requirements for certain variables; however, quantity and quality demands of different uses are not always compatible. Moreover, the composition of waters is dependent on natural factors (geological, topographical, meteorological, etc) and anthropogenic influences (Bartram and Ballance, 1996).

Table 1.1 Quality criteria for freshwater according to USEPA (National Academy of Sciences and National Academy of Engineering, 1972; USEPA, 1986) and The Ministry of Environment of Ecuador (Ministerio del Ambiente del Ecuador - MAE, 2015).

Variables	**Unit**	**USEPA criteria**	**MAE criteria**
Ammonia	mg/L	0.02 (aquatic life)	0.02 (aquatic life)
Nitrate-N	mg/L	10 (domestic water supply)	13 (aquatic life)
Nitrite-N	mg/L	1 (domestic water supply)	0.2 (aquatic life)
Phosphorus	µg/L	0.1 (estuarine water)	10 (aquatic life)
Chloride	mg/L	250 (domestic water supply)	1000 (max discharged to water body)
Chlorine residue	mg/L	0.003 (aquatic life)	0.01 (aquatic life)
Sulfate	mg/L	250 (domestic water supply)	250 (human consumption)
Hydrogen sulfide	µg/L	2 (aquatic life)	0.2 (aquatic life)
Alkalinity ($CaCO_3$)	mg/L	Min 20 (aquatic life), 400 naturally (human consumption)	500 (human consumption) which require disinfection)
Dissolved oxygen	mg/L	5 (invertebrates), 4 (fish)	5-6 (aquatic life)
pH		5-9 (domestic water supply), 6.5-9 (aquatic life)	6-9 (human consumption), 6.5-9 (aquatic life)
Turbidity	NTU	1 (drinking water), should not change the compensation point more than 10% of its seasonal norm (aquatic life)	Natural condition plus 5% (for natural turbidity 0-50), natural condition plus 10% (for natural turbidity 50-100), natural condition plus 20% (for natural turbidity > 100), 5 (drinking water)
Temperature	°	Max 32 (benthic organisms)	Max 32 (aquatic life)
E. coli	Organisms per 100 mL	126 (freshwater bathing)	-
enterococci	Organisms per 100 mL	33 (freshwater bathing)	-
Fecal coliform	Colony forming unit (CFU) per 100 mL	-	20 (human consumption)
Total coliform	Colony forming unit (CFU) per 100 mL	-	200 (human consumption)

Whereas Chapman (1996) defined water quality following two aspects: quality and pollution of the aquatic environment. The quality of the aquatic environment is first defined as a set of concentrations, speciation and physical partitions of inorganic and organic substances of the water. Second, it is defined as the composition and state of aquatic biota living in the water. Since most aquatic organisms are sensitive toward any changes in their environment (e.g. increasing chemical concentration and modification of water bodies), their responses (death, migration or decreasing abundance) toward environmental changes define their state. Lastly, it describes the temporal and spatial variations of the water body due to internal and external factors. The pollution of the aquatic environment is the substances that are introduced by man either directly or indirectly via point sources (such as domestic wastewaters, industrial wastes and animal husbandry) and non-point sources (such as fertilizer and pesticides). The water pollution may have effects on both the biotic and abiotic variables of the water and can cause harm to living resources, hazards to human health, and impairment to agricultural, industrial, aquatic activities and other economic affairs (Chapman, 1996).

1.3 Water quality assessment

Water quality variables can be measured either through quantitative measurements or through semi-quantitative and qualitative evaluations. The quantitative measurements are done by measuring physico-chemical variables of the water (such as nutrient concentrations and particulate material) and biochemical/biological tests (such as BOD and toxicity tests). The semi-quantitative and qualitative evaluations are done by calculating biotic indices, inventorying taxa that are present in the water, evaluation of visual aspects, odor, etc. The quantitative, semi-quantitative and qualitative assessments can be carried out in the field and in the laboratory (Bartram and Ballance, 1996; Chapman, 1996).

Following the measurements of water quality variables, the water quality can be assessed using metrics and indices. Various indices have been developed to calculate the water quality status of water bodies, which assess abiotic, biotic or a combination of both factors. For example, Prati index calculates water quality based on oxygen concentration of the water and then classifies the water quality status into excellent, acceptable, slightly polluted, polluted and heavily polluted (Prati *et al.*, 1971). Whereas the biotic indices calculate the water quality based on the composition of aquatic biota, where each biotic taxon is given a certain tolerance score according to their sensitivity toward environmental disturbances. Examples of biotic indices are the Multimetric Macroinvertebrate Index Flanders – MMIF (Gabriels *et al.*, 2010), the Biological Monitoring Working Party – BMWP (Armitage *et al.*, 1983) and its adapted versions, the Hilsenhoff Biotic Index (Hilsenhoff, 1987; Hilsenhoff, 1988) and the Neotropical Low-land Stream Multimetric Index – NLSMI (Helson and Williams, 2013). The biotic indices are generally related to environmental conditions of the water to determine the key stressors to the presence of aquatic biotas.

The use of biotic indices has been supporting water quality assessment and management decision on water quality monitoring. For example, the European Community required its member countries to take actions to avoid long-term deterioration of freshwater quality by the year 2000 through its Water Framework Directive (WFD) 2000 (European Commission, 2000). The WFD as a regional effort did not specify the assessment methods, which have been allowing the European countries to use various indices to improve and monitor their freshwater quality nationally (Birk *et al.*, 2012; Hering *et al.*, 2010). Besides improving the freshwater quality of European waters, the WFD has resulted in standardized sampling and analysis procedures across Europe (Hering *et al.*, 2010).

The spatial variation of water quality is largely determined by the hydrodynamic characteristics of the water body. Besides, water quality is influenced by flow direction, discharge and time. Therefore, water quality assessment of a water body need to be done at several sampling sites (Chapman, 1996). Depending on the purpose and resources, water quality assessment can be performed repeatedly as a regular monitoring campaign. Nowadays, various resources are available to support sound water quality assessments such as field protocols and manuals for physico-chemical measurements (Bartram and Ballance, 1996; Chapman, 1996).

1.4 The impacts of land use on water quality

During the 20th century, urbanization rate was high. The urban population has increased from 220 million to 2.9 billion globally, with 3.4 billion extra inhabitants expected by 2050. The world’s largest cities with more than 750,000 inhabitants occupy less than 1% of the earth's surface but utilize 41% of the water resources (McDonald *et al.*, 2016a). Population growth requires extra provision of housing, water and food through agriculture and industry. As a result, land use conversion from natural to agriculture and urban cannot be avoided. For example, the global percentage of agricultural area increased only from 34% in 1961 to 38% in 2014. This increase in global agricultural area can be observed mainly in developing countries, whereas in some developed countries the opposite trend can be observed (Fig. 1.2). As a comparison, countries such as Brazil and Indonesia converted much of their forested area from 1990 to 2014 (Fig. 1.3) to agricultural and urban purposes (FAOSTAT, 2017). Furthermore, many rivers and streams have been modified to support urban and agricultural development (Harding *et al.*, 1998).

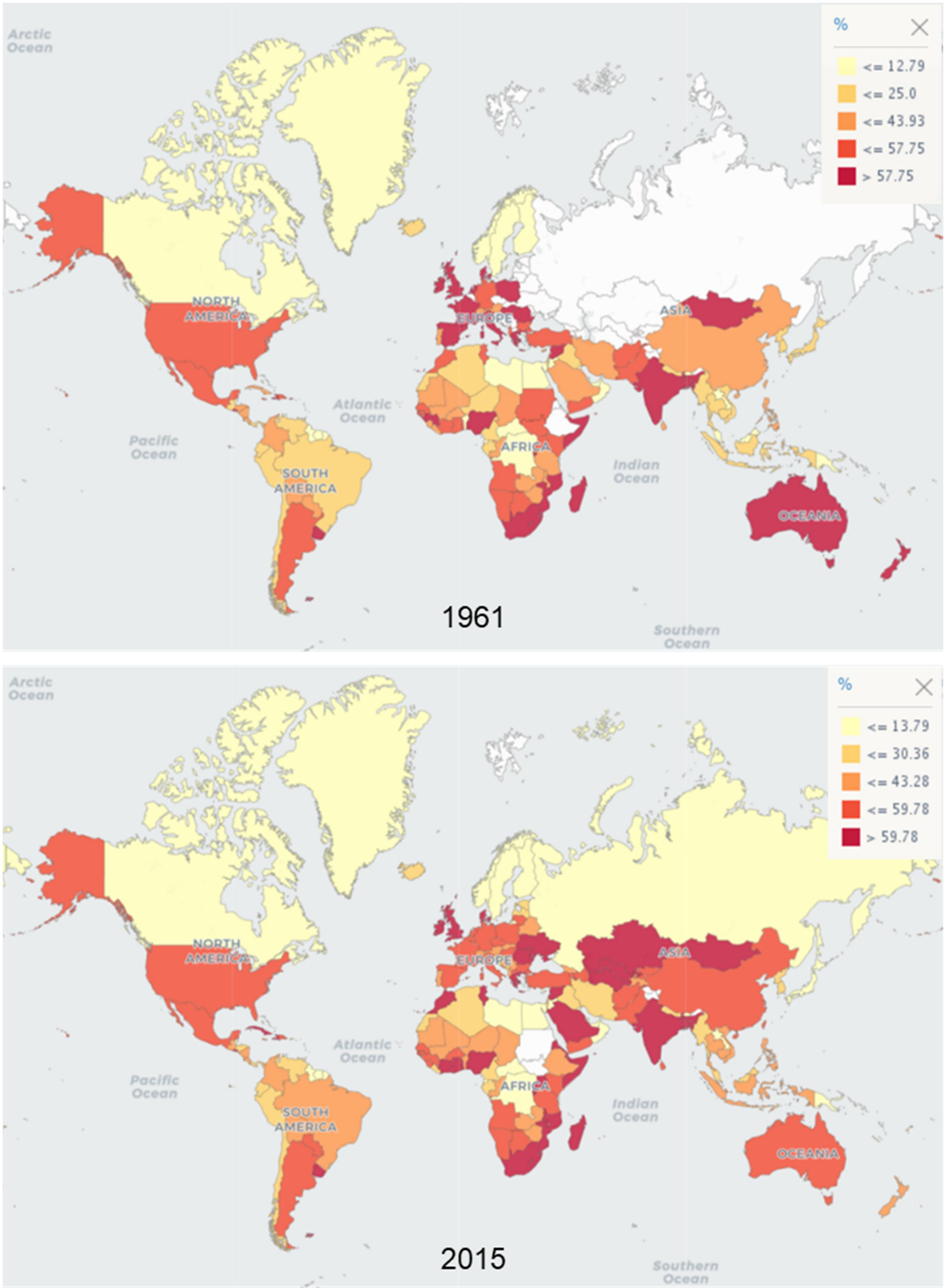


Figure 1.2 Percentage of agriculture area globally for 1961 and 2015, white area means that no data is available (FAOSTAT, 2017).

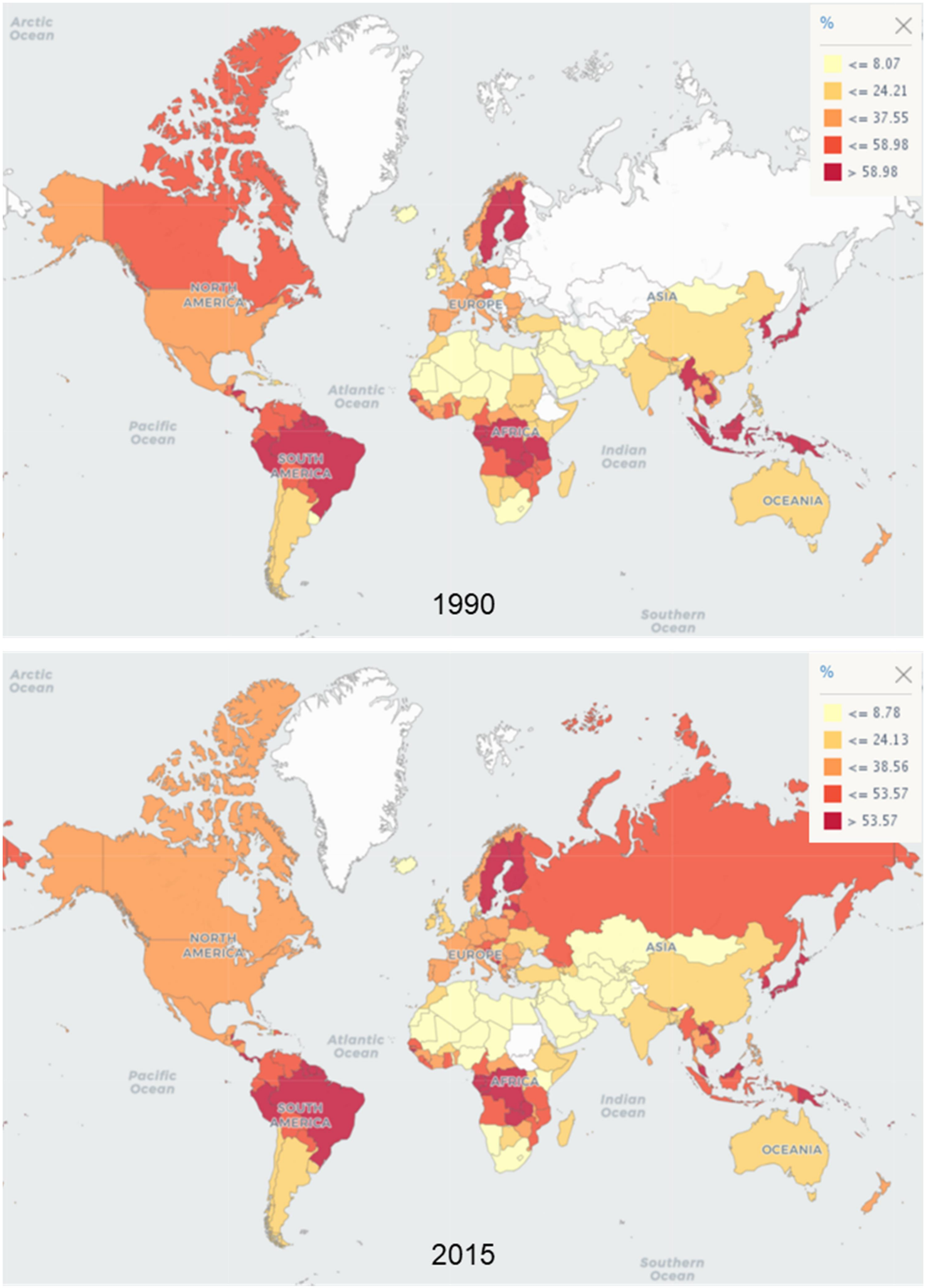


Figure 1.3 Percentage of forested area globally for 1990 and 2015, white area means that no data is available (FAOSTAT, 2017).

Natural land use such as forest at the upstream parts of watersheds provides important ecosystem services for the entire watersheds; however, land use conversion from natural to other uses such as intensive agricultural or urban decreases ecosystem services and increases pollution (McDonald *et al.*, 2016a). It should be noted that managed forests might also decrease water quality and ecosystem services even though generally less degrading than intensive agricultural and urban (Baillie and Neary, 2015; Futter *et al.*, 2016). However, the impacts of agriculture on water quality can be minimized through various agricultural practices such as limited or avoidance of agrochemical use in small-scale farms (Kehoe *et al.*, 2015), application of crop rotations, conservation tillage systems (Roth, 2017; Yates *et al.*, 2006) and ecoagriculture (Scherr and McNeely, 2008). Diffuse pollution originating from land use such as agricultural and urban are more complex than point source pollution originating from industries and sewage treatment plants because diffuse pollution also contains run-off and landscape interactions, and are considered key elements affecting water quality (Brogna *et al.*, 2017). In any case, (waste)water discharge from residential, agricultural and industrial lands entering the rivers and streams can change water quality variables such as nutrient concentrations and sediment composition (da Silva *et al.*, 2015; Ferreira *et al.*, 2016; Goss *et al.*, 2014). Furthermore, non-forested land use also decreases oxygen concentration and increases other physico-chemical variables of the water such as pH, temperature, conductivity and heavy metals (Englert *et al.*, 2015; Ferreira *et al.*, 2016; Robinson *et al.*, 2014; Yun and An, 2016).

Therefore, land use is an important variable in water quality studies. By including land use information in water quality studies, it is possible to associate water quality with the type of land use influencing it. Simply put, it is possible to determine the source of disturbance and therefore possible to plan management actions related to the land use as the main cause (Barbour *et al.*, 1999; Berger *et al.*, 2017; Bolstad and Swank, 1997; Crétaz and Barten, 2007).

1.5 Land use quantification

Land use can be observed in various ways. Field observation, remote sensing and geographical information system (GIS) data are common methods and sources in collecting land use data. Field observation is done through transect walk within a certain distance from the sampling site (Erba *et al.*, 2015), generally with the help of assessment protocols such as river habitat survey (RHS) of the United Kingdom and the Isle of Man (Raven *et al.*, 1998) or the Australian River Assessment System (AUSRIVAS) physical assessment protocol (Parsons *et al.*, 2002). These two assessment protocols have been frequently used worldwide since they cover various aspects of habitat assessment. This type of observation generally has limited area coverage based on the accessibility of the observer. In contrast to field observations, data collection via satellite imagery and GIS can be done remotely or through available data from government or research institution (Baltazar *et al.*, 2016; Carlisle and Meador, 2007; Einheuser *et al.*, 2012). These types of observation do not limit area coverage; therefore it is possible to gather data from local to regional scales.

Land use is classified based on the study purposes or utilized assessment protocols. Common categories include: (1) urban, (2) agricultural and (3) forest (Feio *et al.*, 2007; Guse *et al.*, 2015). Several studies classified land use into more distinguished categories such as (1) residential, (2) industrial, (3) road (Mantyka-Pringle *et al.*, 2014; Van Sickle *et al.*, 2004), (4) orchard, (5) pasture, (6) bare land (Clapcott *et al.*, 2017; Cortes *et al.*, 2013), (7) arable land (Dahm and Hering, 2016), (8) needle-leaved forest and (9) broad-leaved forest (Brogna *et al.*, 2017; Cortes *et al.*, 2013). Generally, the classification is based on the dominant presence of certain land use type at the study site (e.g. Fig. 1.4).

Figure 1.4 Examples of land use: pasture (top left), banana plantation (top right), forest (bottom left) and urban (bottom right).

1.6 Land use and biodiversity

Biodiversity has an essential role in supporting ecosystem services (Teillard *et al.*, 2016). However, human activities have altered the environment and have caused biodiversity loss. Current biodiversity loss is reported at >100 species extinct per million species per year. Biodiversity loss occurs from local to regional level and can have pervasive effects on ecosystem functioning globally (Rockstrom *et al.*, 2009). Worldwide, twenty-five hotspots were registered where exceptional concentrations of endemic species are present while at the same time experiencing exceptional loss of habitat. The loss of habitat is generally due to deforestation or the loss of an area's primary vegetation. Sixteen hotspots are located in developing tropical countries (e.g. Ecuador, Brazil, Indonesia, Madagascar and The Philippines) where biodiversity loss is huge while conservation resources are scarce (Myers *et al.*, 2000).

Taxa extinction happens naturally, but land use changes accelerate the process. Both intensive and extensive agriculture have been causing diversity loss globally (Lanz *et al.*, 2018). Agricultural (including pastoral) lands are considered to be the main driver of biodiversity loss in The Philippines, Colombia, Ecuador, Venezuela, Brazil and South Africa; whereas forestry is the biggest threat in countries such as Indonesia, Papua New Guinea, India, Madagascar, Peru and DR Congo (Chaudhary *et al.*, 2018). Moreover, agriculture is considered a major cause of pollution by changing nitrogen and phosphorus concentrations and cycle in the environment, which at the end destroys the habitat and cause biodiversity loss (Rockstrom *et al.*, 2009). One example is the conversion of some part of the Amazonian rainforest to intensive agricultural lands which at the end destroys the habitat and causes biodiversity loss. Urban expansion is another type of habitat degradation that has been causing biodiversity loss (Chaudhary *et al.*, 2018; Guneralp *et al.*, 2018; Teillard *et al.*, 2016).

1.7 Macroinvertebrate usage in water quality studies

Bioassessments using aquatic biotas are necessary to assess impairments in aquatic life because aquatic biotas reflect an overall ecological integrity of the water (Barbour *et al.*, 1999). The type of organism being affected and the degree of destruction in aquatic lives reflect the type and extend of environmental disturbance occurring in the water. Since aquatic organisms have various life cycles and sensitivities toward environmental disturbance, the information can be used to assess pollution history and current effects on the water (Cairns and Dickson, 1971; De Pauw *et al.*, 2006). Periphyton, benthic macroinvertebrate and fish assemblages are common aquatic biotas assessed during bioassessments, with benthic macroinvertebrates being the most used biotas in European countries (Birk *et al.*, 2012; De Pauw *et al.*, 2006). The use of each organism group as bioindicator has advantages and disadvantages. However, despite their disadvantages, the use of aquatic macroinvertebrates in water quality studies is considered more advantageous than other biotas such as periphyton and fish assemblages, as discussed by Cairns and Dickson (1971), Chapman (1996), Barbour *et al.* (1999), De Pauw *et al.* (2006) and Verissimo *et al.* (2012):

- Macroinvertebrates generally have limited migration pattern (except for several taxa that might drift in moving waters) and are therefore good indicators to study localized conditions, as opposed to fish that are highly mobile.
- Macroinvertebrates have a complex life-cycle of minimum one year and therefore can integrate the effects of short- and long-term environmental changes. Whereas periphyton only has a short life cycle (several days) and highly sensitive to short-term changes in the environment, which makes it a good bioindicator in a snapshot survey. However, due to complex life-cycle of macroinvertebrates, knowledge of their life cycles might be necessary in certain studies assessing the absence of some taxa.
- Some macroinvertebrate taxa might be difficult to identify. However, generally, they can be identified relatively easily to minimum family level and for some taxa can be easily identified to lower taxonomic levels. Furthermore, identification keys and macroinvertebrate experts are relatively easily available. Whereas periphyton’s identification is relatively difficult under a microscope and periphyton experts are scarce.
- Macroinvertebrates consist of taxa from broad range of trophic levels and pollution tolerances that can provide good information regarding cumulative effects based on their presences/absences.
- Macroinvertebrate’s sampling is relatively inexpensive and easy in terms of techniques and equipment, and has minimal detrimental effect on the resident biota, as opposed to some fish sampling methods that might be size and species

selective as well (Han *et al.*, 2016). Nevertheless, quantitative sampling of macroinvertebrate is difficult and sediment type is important during sampling.

- Macroinvertebrates generally present abundantly, and serve as a primary food source for fish; thus determine fish presence as well.

Many biotic indices based on macroinvertebrates have been developed especially in developed countries, such as the Trent Biotic Index (Woodiwiss, 1964), the Chandler Biotic Index (Chandler, 1970), the Belgian Biotic Index (De Pauw and Vanhooren, 1983) and the Multimetric Macroinvertebrate Index Flanders (Gabriels *et al.*, 2010). These indices were developed specifically for the purpose of the country or region they were developed, thus less applicable in other countries or regions.

Generally, indices are country- or region-specific, means they perform well in calculating the ecological water quality status of the water bodies where they were developed and less applicable in other countries or regions (Chapman, 1996; Gabriels *et al.*, 2010). To address this issue, the Biological Monitoring Working Party (BMWP, Armitage *et al.* (1983)) index was developed as a standard international method. The BMWP was also developed to limit time and effort to identify organisms only to family level, instead of species level (Armitage *et al.*, 1983; Chapman, 1996). However, the BMWP's applicability is still debatable in other regions, resulting in development of the BMWP's adaptations such as the BMWP^{THAI} (Mustow, 2002), the Andean Biotic Index (Rios-Touma *et al.*, 2014) and the BMWP-Col (Alvarez, 2005; Roldán Pérez, 2003). Moreover, country- or region-wise indices were developed in Central and South American countries, such as the Neotropical Low-land Stream Multimetric Index (NLSMI) for Panama (Helson and Williams, 2013), the Índice Multimétrico del Estado Ecológico para Ríos Altoandinos (IMEERA) for Peru and Ecuador (Villamarin *et al.*, 2013), Multimetric Index for Serra da Bocaina (MISB) for Brazil (Baptista *et al.*, 2013) and the Biotic Index for Pampean rivers and streams (IBPAMP) for Argentina (Capítulo *et al.*, 2001).

1.8 Relationship between land use and macroinvertebrates

Aquatic macroinvertebrates require a stable and healthy environment to live and reproduce, which includes hydrological, morphological and physico-chemical variables (Chapman, 1996; Hering *et al.*, 2006). Healthy environments provide sufficient food for macroinvertebrates through leaf and wood litter from terrestrial vegetation that enters the water (Townsend *et al.*, 1997). Good water quality will ensure macroinvertebrate's survival for its entire life and will guarantee the stability of the aquatic system. Sensitive macroinvertebrates may not tolerate short-term exposure to environmental disturbance and their disappearance from the surface waters will change community structure and decrease ecosystem services (Cairns and Dickson, 1971).

Since land use can change water quality variables, this also changes habitat conditions of aquatic macroinvertebrates. For example, land use can increase domestic effluents and nutrient concentration of the water; increase erosion, reduce shading and organic matter due to agriculture, urbanization and forestry activities; alter river channel and flow; and increase toxic chemical runoff from point and non-point sources (da Silva *et al.*, 2015; Goss *et al.*, 2014; Karr, 1991). Habitat and water qualities are therefore degraded. Besides immediate disappearance of very sensitive taxa, habitat degradation can also increase predation and competition among macroinvertebrate assemblages in order to survive. Under multiple stressor conditions for a given period of time, they will fasten the disappearance of other sensitive and tolerant taxa from the assemblages (Barbour *et al.*, 1999; Cairns and Dickson, 1971; Hering *et al.*, 2006). Overall, a decrease in water quality due to land use is detrimental to aquatic macroinvertebrate composition, where natural land use (i.e. forest) will support the presence of sensitive taxa, and land use with intensive anthropogenic activities (e.g. agriculture and residential) might only support more tolerant taxa. For example, the shift of taxa presence can be seen at rivers having different land use types in the Gilgel Gibe watershed, Southwest Ethiopia (Fig. 1.5). Sensitive taxa Coenagrionidae and Simuliidae were only present in rivers having forested land use, whereas tolerant taxa could present in both forested and agricultural areas (e.g. Chironomidae and Hydrophilidae were present in rivers having agricultural and residential land uses, while Baetidae could present in both forested and agricultural areas) (Mereta *et al.*, 2012),. In conclusion, land use can determine the ecological water quality status of the waters through its influence in macroinvertebrate composition of the water.

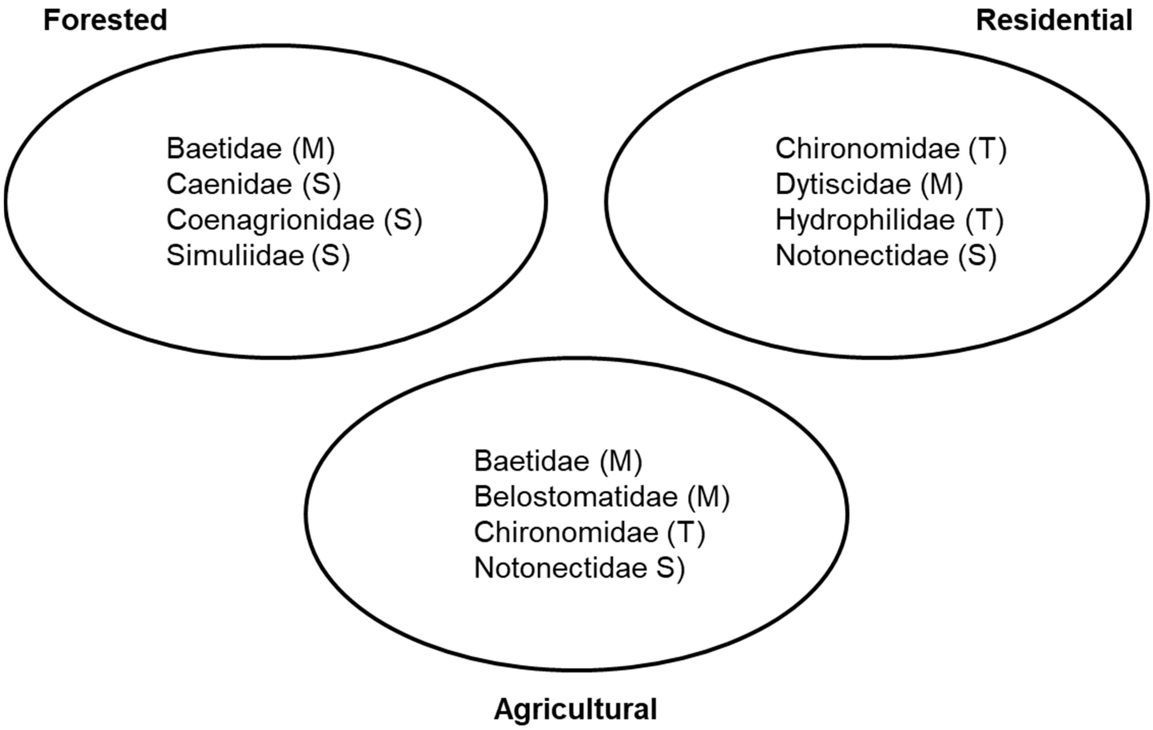


Figure 1.5 Example of land use types and the presence of macroinvertebrate taxa, observed in the Gilgel Gibe watershed, Southwest Ethiopia (Mereta *et al.*, 2012); M: moderate sensitivity, S: sensitive, T: tolerant.

1.9 Scope and objectives

1.9.1 Case studies Guayas river basin, Ecuador

This PhD study is intended to support water quality studies in developing countries, where anthropogenic activities have been increasingly threatening the water quality but lack of regular monitoring campaigns. The focus is on the Guayas river basin, a major watershed in Ecuador as the case study (see chapter 3).

The Guayas river basin has been experiencing intensive agricultural and urbanization activities due to population growth. Around one-third of Ecuador population (5.5 million inhabitants, national census 2010) resides in the Guayas river basin (UNSD, 2017), creating bustling economic activities such as agricultural and industrial. Cultivation of banana, rice, maize, African palm, cacao productions and fisheries are important agriculture and industries here (Alvarez-Mieles *et al.*, 2013; Caceres *et al.*, 2002; Gerebizza, 2009). These population growth and extensive agriculture and industries have been resulting in increasing demand for farm and domestic lands.

According to Myers *et al.* (2000), agricultural including pastoral activities is the main driver of biodiversity loss in Ecuador. Ecuador is one of the most biodiverse tropical ecosystems on earth (having 8% of amphibian, 5% of reptile, 8% of mammal and 16% of bird species of the earth). Among Ecuadorian fauna, invertebrate biodiversity is the least known and identified (The Biodiversity Group, 2016). However, Ecuador has been considered as one of the hotspots having exceptional concentrations of endemic species while at the same time experiencing exceptional loss of habitat (Myers *et al.*, 2000). As suggested by (Iniguez-Armijos *et al.*, 2014), biodiversity loss and ecosystem degradation are associated with deforestation for agricultural and industrial intensification. These threats have been occurring not only in Ecuador as a country, but also in the Guayas river basin as a center of Ecuador's agricultural activities.

The current agricultural and industrial practices can influence water quality of the Guayas river basin. Despite its intensive agricultural and urbanization activities, ecological water quality monitoring in the Guayas river basin is still lacking. For example, this study is the first ecological water quality study in the entire Guayas river basin. The previous sampling campaign was only done in one wetland area under the WETwin project (Alvarez-Mieles *et al.*, 2013; Arias-Hidalgo *et al.*, 2013). This is due to the lack of human and financial resources. The Ecuadorian government and research institutes did not have adequate knowledge and practice to perform the ecological water quality monitoring. The use of macroinvertebrates and biotic indices in ecological water quality studies was relatively new and there was no systematic method to perform the assessment. Moreover, land use conversion for agricultural and domestic purposes is ongoing; however, updated land use information is not available. Therefore, an assessment of land use impacts on ecological water quality is still lacking. Hence, this study is acting as the starting point for ecological water quality assessment in the entire Guayas river basin. The methodologies and findings from this study can be used for future water quality monitoring in the Guayas river basin and for other river basins in the developing countries as well. Restoration and management of the water quality as proposed in this study will consequently protect the biodiversity of the Guayas river basin and Ecuador in general.

1.9.2 Objectives

In order to support developing countries in regular water quality monitoring, four sets of questions were raised in this PhD study:

1. Why is land use information often not included in ecological water quality studies? What is the best way to include land use information in ecological water quality studies? Are ecological models useful to quantify the relationship between land use and ecological water quality? (Chapter 2)

To answer these questions, published scientific papers studying ecological water quality that utilized models in their analyses, where land use was considered a key environmental variable were consulted. There are two hypotheses as to why land use information was often not integrated in ecological water quality studies:

- Land use information is not easily available.
- There is insufficient methodology in quantifying the relationship between land use and the ecological water quality.

2. What is the current ecological water quality of the Guayas river basin, Ecuador? (Chapter 4)

The ecological water quality in the Guayas river basin was expected to be generally bad. It was expected that nutrient concentrations of the water would be relatively high due to intensive agriculture activities in the Guayas river basin. To quantify the ecological water quality, two biotic indices that were developed in neighboring countries were calculated: the BMWP-Col and the NLSMI.

3. How is the relationship between the presence of macroinvertebrate and physico-chemical variables? What are the key environmental variables affecting the ecological water quality? What management actions can be proposed? (Chapters 4 and 5)

Two hypotheses to be tested are:

- Macroinvertebrate composition is highly influenced by physico-chemical variables.
- Agriculture-related nutrients and land use are the key environmental variables influencing the ecological water quality.

To test above hypotheses, multivariate analyses (chapter 4) and regression analyses (chapter 5) were used.

4. Which type of data collection is most suitable to classify land use and to quantify its effect on the ecological water quality? Which spatial scale is most appropriate when quantifying local land use change and its effect on the ecological water quality? Which environmental variables are associated with each type of data collection? (Chapter 6)

Regression analyses were utilized to select the most suitable method or data source in quantifying land use effect and to define the associated variables with observation methods and sources. Three hypotheses to be tested are:

- Field observation is the best method to quantify local land use.
- Land use within direct vicinity to sampling sites is the most influencing land use on the ecological water quality.
- The three methods and sources (field protocol, Google maps and GIS data) are associated with similar environmental variables, since they all define local land use.

1.9.3 Thesis structure

Overall, there are seven chapters within the thesis. To help understanding the sequence among chapters, a scheme that shows the link between chapters is provided here (Fig. 1.6).

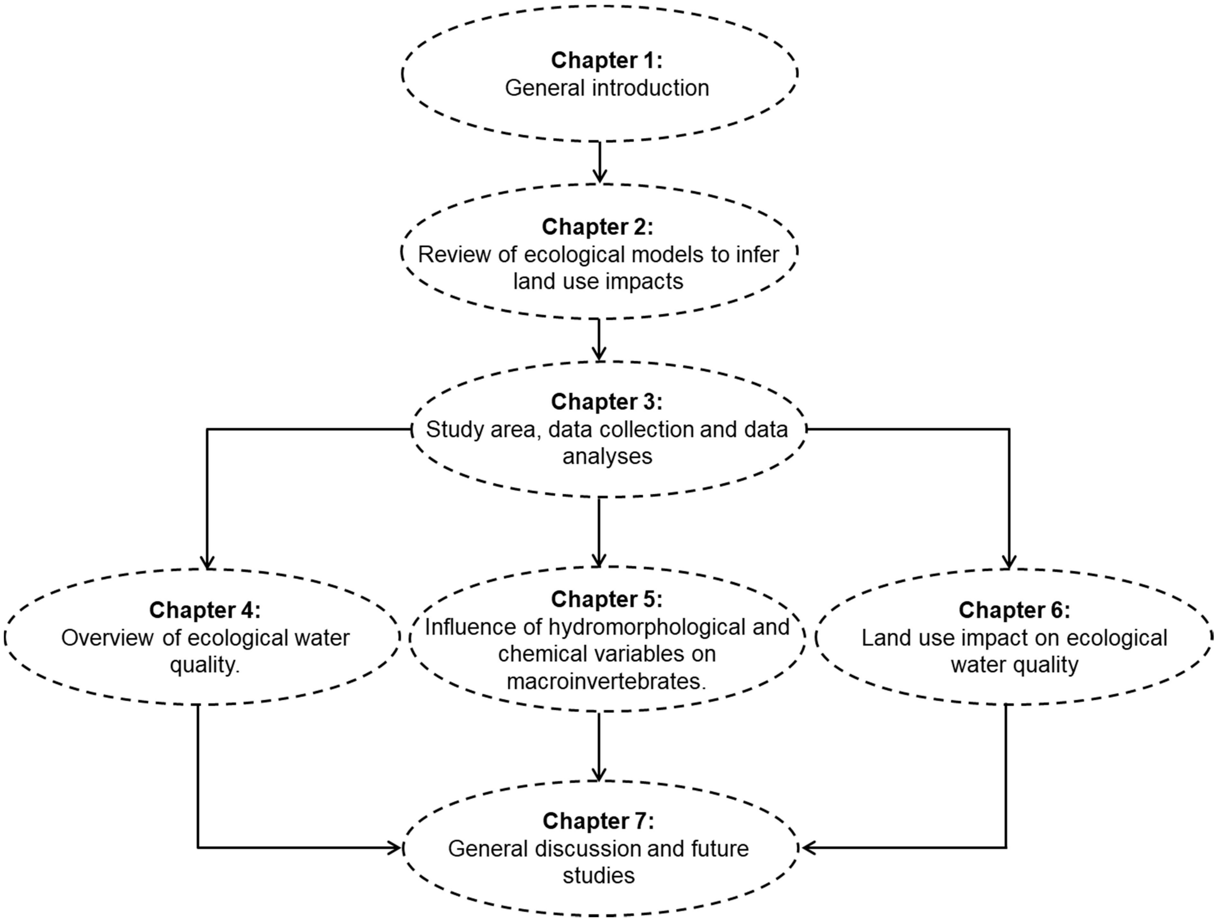


Figure 1.6 Schematic presentation showing the link between chapters within the thesis.

In this first chapter, a general introduction about the importance of land use in water quality studies and how land use is quantified were provided. Furthermore, the relationship between land use and aquatic macroinvertebrates, and why do researchers use aquatic macroinvertebrates in ecological water quality studies are discussed.

The next chapter provides a review regarding the use of models in ecological water quality studies integrating land use information, with a focus on macroinvertebrate communities. Despite the potential negative effects land use pose on the ecological water quality, land use was often not included in the ecological water quality studies. The selected papers published in Web of Science provided insights to the use of land use information in ecological water quality studies and the use of models in analysis.

Chapter 3 provides the description of the Guayas river basin and the methodologies for data collection and data analysis.

In chapter 4, the current ecological water quality status of the Guayas river basin, Ecuador, is evaluated. The BMWP-Col and the NLSMI are used to calculate the ecological water quality index of 120 sites within the Guayas river basin. Both biotic indices are used to define a more suitable index for the Guayas river basin. The potential environmental variables that influenced the presence of macroinvertebrates are also determined using a correspondence analysis.

Chapter 5 highlights the importance of environmental conditions on the ecological water quality and the key environmental variables affecting the ecological water quality. Based on the selected key variables, possible management actions that can be implemented in the Guayas river basin are also proposed. General linear models and sensitivity analyses are utilized to answer the questions.

In chapter 6, the suitability of three observation techniques to quantify land use effect on the ecological water quality with a focus on macroinvertebrate communities is assessed. The three observation techniques varie in terms of area coverage within the local land use. Using general linear models and sensitivity analyses the most suitable observation technique to be applied in the Guayas river basin is selected. Additionally, the key environmental variables associated with each observation technique are defined.

Finally, a summary of each set of research questions is provided in chapter 7. A discussion regarding the methodologies used for data collection and reliability, data analyses and the use of macroinvertebrates as biological indicators is also provided. Besides, recommendations for ecological water quality monitoring and future studies are provided.

Chapter 2: Ecological models to infer the quantitative relationship between land use and the aquatic macroinvertebrate community

Adapted from:

Minar Naomi Damanik Ambarita, Gert Everaert, Peter L.M. Goethals. Ecological models to infer the quantitative relationship between land use and the aquatic macroinvertebrate community. Water 10(184). DOI: 10.3390/w10020184.

Abstract

Land use changes influence the ecological water quality. In spite of this knowledge, land use information is often missing in ecological water quality studies. Therefore, in this chapter, 39 peer-reviewed model-based scientific papers that study the relationship between land use and aquatic macroinvertebrates were consulted. From the selected papers, it was found that certain water bodies responded more to local land use, while other water bodies were more likely to be affected by catchment land use. Hence, combined land use information from both local scale and catchment scale will provide a better understanding of the impact of land use changes on the ecological water quality. To gain this knowledge, efforts need to be taken to acquire land use information from field observations and remote sensing or GIS data source. Furthermore, the benefits of using models to better understand the relationship between the ecological water quality and environmental variables were concluded on. Depending on the aim of the study and the nature of the data, researchers can select the most suitable model to ensure fast analysis.

2.1 Introduction

Anthropogenic activities that are taking place upstream and in the surrounding of surface waters can influence the water quality by altering its physico-chemical and hydromorphological characteristics (Garnier *et al.*, 2013; Pilgrim *et al.*, 2014). There is a clear link between land use and water quality, either positive or negative. For example, urbanization, industries and intensive agriculture activities may increase erosion and sediment accumulation (Beasley and Kneale, 2002; da Silva *et al.*, 2015; Goss *et al.*, 2014; Smucker and Detenbeck, 2014), increase the input of chemicals such as nitrogen and phosphorus (Beasley and Kneale, 2002; Colin *et al.*, 2016; Goss *et al.*, 2014; Raper *et al.*, 2015), and create more homogeneous flow and bed substrate of streams (da Silva *et al.*, 2015; Walsh *et al.*, 2004). The impact of land use changes due to agriculture can be minimized by reducing the use of agrochemicals, i.e. in small-scale farms (Kehoe *et al.*, 2015) or by applying crop rotation and conservation tillage systems (Roth, 2017; Yates *et al.*, 2006). The impact that land use poses on surface waters is not limited to river ecosystems (Cortes *et al.*, 2013; Manfrin *et al.*, 2016), but also affects ponds (Thornhill *et al.*, 2017) and lakes (Alahuhta *et al.*, 2017; Pietron *et al.*, 2017). The change in the physico-chemical and hydromorphological characteristics of the impacted river or catchment will consequently affect the richness and abundance of aquatic organisms such as fish (Hook *et al.*, 2017; Wright *et al.*, 2017), macroinvertebrates (Baillie and Neary, 2015; Gerth *et al.*, 2017) and plants (Liu *et al.*, 2015; Raapysjarvi *et al.*, 2016).

Despite the clear linkage between land use and water quality, land use was not always included in water quality studies. Many studies relating water quality and aquatic organisms only focused on water quality variables, such as physico-chemical characteristics and hydromorphological conditions (Bonada *et al.*, 2008; Brown *et al.*, 2012). Other studies only focused on the potential effects of a certain type of land use on the water quality (e.g. residential (Yang and Toor, 2017), agriculture (Lee *et al.*, 2017), mining (Pietron *et al.*, 2017) and forest (Brogna *et al.*, 2017)) or the occurrence of certain aquatic organisms (Cunha *et al.*, 2015; Epele and Miserendino, 2015; Sueyoshi *et al.*, 2016). Restoration projects focused sometimes only on monitoring water chemistry or the change in hydromorphological conditions, instead of addressing land use as the main cause (Palmer *et al.*, 2014; Palmer *et al.*, 2010). Previous studies performed by Berger *et al.* (2017) at 184 German rivers quantified the benefit of the inclusion of land use in studying water quality to improve ecological quality from diffuse pollution. Shrestha *et al.* (2017) and Bussi *et al.* (2017) also included land use in their studies. Shrestha *et al.* (2017) successfully studied related water yield and nutrient release to it in Onkaparinga catchment (Australia), while Bussi *et al.* (2017) studied the water quality of the River Thames catchment (United Kingdom).

When land use is included in the study, it is important to consider its spatiotemporal aspects, because land use that takes place within different locations, size and time provokes various biogeochemical and hydrological responses (Pilgrim *et al.*, 2014). However, the spatial coverage of studies that assess the impact of land use varies largely. To date, there is no consensus whether the impacts of land use is only present within local or direct vicinity (da Silva *et al.*, 2015; Fierro *et al.*, 2015; Jun *et al.*, 2011), within a certain buffer zone (Park *et al.*, 2011) or as wide as the catchment area (Hughes *et al.*, 2016; Schmalz *et al.*, 2015) of the surface water. Some researchers studied the impacts of land use on the water quality based on a single monitoring campaign (Cortes *et al.*, 2013; Manfrin *et al.*, 2016) or based on a long time data-series, such as within three time periods of 1971, 1985 and 1999 (Tu, 2009), over 75 years (Pilgrim *et al.*, 2014) and over the past century (McDonald *et al.*, 2016b). Unfortunately, Tu (2009), Pilgrim *et al.* (2014) and McDonald *et al.* (2016b) only studied land use impacts on physico-chemical characteristics of the water, thus the impact of land use on macroinvertebrates or other aquatic biotas is unidentified. Studies of land use impact on water quality also vary in the applied methods of acquiring the land use data. Several studies were based on field observations (Bucker *et al.*, 2010; Mwedzi *et al.*, 2016), geographic information system (GIS) data (Hughes *et al.*, 2016; Pilgrim *et al.*, 2014) or combined methods and sources (Cortes *et al.*, 2013; Strehmel *et al.*, 2016). Methods bring highly variable outcomes that are difficult to compare with each other.

Due to the complexity of aquatic ecosystems, water quality studies can be challenging. Aquatic ecosystems are influenced by multiple variables, and it is difficult to decide which variable to focus on in the studies. In this context, using ecological models for studying water quality can be beneficial (Arias-Hidalgo *et al.*, 2013; Everaert *et al.*, 2013; Schuwirth *et al.*, 2016; Tchakonte *et al.*, 2015). Slevers *et al.* (2017) used linear mixed effects models to assess trout response to the change in riparian conditions in North America, Europe, South America and Australia, while Ferreira *et al.* (2017) used partial least squares regression models to assess water quality degradation and biodiversity decline (fish and macroinvertebrates) as the consequence of anthropogenic pressures. Other models such as random forest models for diatom (Larras *et al.*, 2017), multiple regression for macroinvertebrates (Berger *et al.*, 2017) and boosted regression trees for fish and macroinvertebrates (Dahm and Hering, 2016) have also been used in ecological related studies that integrated land use data.

Based on published articles in Web of Science, water quality studies where land use was determined a key stressor influencing the presence of aquatic organisms were reviewed. The selection was based on studies that implemented ecological models to infer and quantify the relation between macroinvertebrate communities and environmental variables in river ecosystems. How to better study the impacts of land use on macroinvertebrates in developing countries where available updated land use information is limited was discussed. An integrated approach of evaluating land use impacts on macroinvertebrates was also recommended. Throughout the chapter, the term ecological water quality is used to define water quality based on aquatic organisms especially macroinvertebrates.

2.2 Materials and methods

The internationally peer-reviewed papers were accessed via the Web of Science for the period between 1955 and 23 May 2017. The search was performed by including key words ‘water quality’, ‘macroinvertebrate*’ and ‘river*’ and excluding key word ‘diatom*’ as topic, then continued with key word(s) ‘land use*’ and ‘model*’ as topic and title, in substitution (Fig. 2.1). During the search, several papers studied ecological water quality based on macroinvertebrates and diatoms. As our primary focus and expertise was on aquatic macroinvertebrates, we excluded papers that solely dealt with diatoms. Using ‘land use’ as key word(s) in the title resulted in 15 papers, while using ‘model*’ as key word in the title resulted in 28 papers (Fig. 2.1). Note that four papers among these 28 papers were also listed in the 15 papers. Hence, in total 39 (= 15 + 28 – 4) papers were retained that covered a wide range of internationally available studies related to the objectives of the thesis.

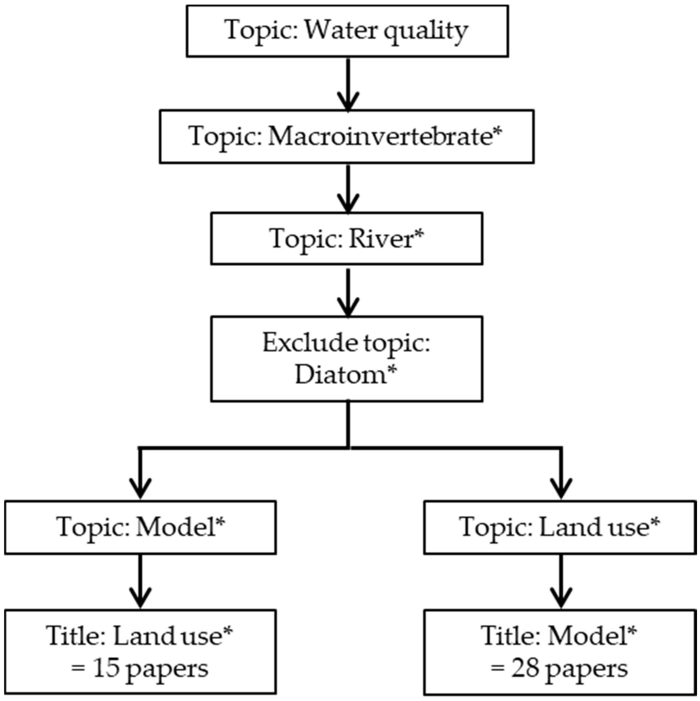


Figure 2.1 Scheme of search category and key word(s) for paper selection in the Web of Science and the number of resulting papers.

The papers were assessed in terms of input variables included in the models, spatial scale of land use information, ecological community that is assessed, biological index used, type of ecological model and country of study. A strengths, weaknesses, opportunities and threats (SWOT) analysis was utilized to evaluate the use of models in ecological water quality studies and the inclusion of land use information in the analysis. Finally, the methods were compiled to provide recommendation for worldwide studies especially in developing countries.

2.3 Ecological water quality studies and land use

2.3.1 Introduction

Most of the 39 papers used macroinvertebrate data identified up to family level (19 papers). Seventeen papers used macroinvertebrate data up to species or genus level for most taxa and up to family level for the remaining taxa, while 2 papers only used order level and 1 paper did not mention the level of identification. Macroinvertebrate data were collected either from national/regional databases (20 papers) or during tailor-designed sampling campaigns (19 papers). Macroinvertebrate sampling was done mainly using kick-net method (13 papers), surber method (4 papers), while 2 papers did not mention the type of sampling they performed. Several papers studied macroinvertebrate data based on taxa richness (16 papers), using various biotic indices (17 papers) or various diversity indices such as Simpson's diversity and Shannon-Wiener index (2 papers), or a combination of biotic and diversity indices (4 papers). In one paper, the authors performed their assessment based on biological, physiological and ecological macroinvertebrate traits (Cortes *et al.*, 2013) while in two other papers, the assessment was based on the functional feeding group (Pearson *et al.*, 2016; Weigel, 2003).

Table 2.1 Macroinvertebrate data used in the selected papers; kick: kick net, surber: surber sampler, slack: slack sampler.

Identification level				**Data source**		**Biotic index**		
Family level	**Mostly species or genus level, some up to family level**	**Order level**	**No information**	**Sampling (kick, surber)**	**National/regional databases**	**No biotic index, only taxa richness**	**Biotic index (e.g. Hilsenhoff, EPT, BMWP, ASPT)**	**Diversity indices (Simpson's diversity, Shannon-Wiener index)**
Abouali *et al.* (2016), Alemneh *et al.* (2017), Alvarez-Cabria *et al.* (2017), Baltazar *et al.* (2016), Cortes *et al.* (2013), (Damanik-Ambarita *et al.*, 2016a), Einheuser *et al.* (2012), Erba *et al.* (2015), Forio *et al.* (2015), Forio *et al.* (2017), Hrodey *et al.* (2009), Hughes *et al.* (2016), Mantyka-Pringle *et al.* (2014),	Barton (1996), Bennetsen *et al.* (2016), Carlisle and Hawkins (2008), Carlisle and Meador (2007), Clapcott *et al.* (2017), Dahm and Hering (2016), Davies and Jackson (2006), Feio *et al.* (2009), Feio *et al.* (2007), Guse *et al.* (2015), Hawkins *et al.* (2000), Hawkins and Yuan (2016), Maloney	Lock and Goethals (2013), Lock and Goethals (2014)	Van Sickle *et al.* (2004)	Alemneh *et al.* (2017) kick, Baltazar *et al.* (2016) kick, Barton (1996) kick, Cortes *et al.* (2013), Damanik-Ambarita *et al.* (2016a) kick, Erba *et al.* (2015) surber, Feio *et al.* (2007) kick, Forio *et al.* (2015) kick, Forio *et al.* (2017) kick, Hawkins *et al.* (2000) surber, Hrodey *et al.* (2009) Ekman dredge+kick+surber, Lock and Goethals (2013)	Abouali *et al.* (2016), Alvarez-Cabria *et al.* (2017) kick, Bennetsen *et al.* (2016), Carlisle and Hawkins (2008) slack, Carlisle and Meador (2007) slack, Clapcott *et al.* (2017) kick+surber, Dahm and Hering (2016), Davies and Jackson (2006), Einheuser *et al.* (2012), Feio *et al.* (2009) kick, Guse *et al.*	Alemneh *et al.* (2017), Bennetsen *et al.* (2016), Carlisle and Hawkins (2008), Carlisle and Meador (2007), Dahm and Hering (2016), Davies and Jackson (2006), Feio *et al.* (2009), Feio *et al.* (2007), Guse *et al.* (2015), Hawkins *et al.* (2000), Hawkins and Yuan (2016), Lock and	Abouali *et al.* (2016), Alvarez-Cabria *et al.* (2017), Baltazar *et al.* (2016), Barton (1996), Clapcott *et al.* (2017), Cortes *et al.* (2013), Damanik-Ambarita *et al.* (2016a), Einheuser *et al.* (2012), Erba *et al.* (2015), Forio *et al.* (2015), Forio *et al.* (2017), Hrodey *et al.* (2009), Hughes *et al.*	Baltazar *et al.* (2016), Erba *et al.* (2015), Moreno *et al.* (2009), Pearson *et al.* (2016), Terrado *et al.* (2016), Weigel (2003)

Identification level				**Data source**		**Biotic index**		
Family level	**Mostly species or genus level, some up to family level**	**Order level**	**No information**	**Sampling (kick, surber)**	**National/regional databases**	**No biotic index, only taxa richness**	**Biotic index (e.g. Hilsenhoff, EPT, BMWP, ASPT)**	**Diversity indices (Simpson's diversity, Shannon-Wiener index)**
Moreno *et al.* (2009), Pearson *et al.* (2016), Sanchez *et al.* (2014), Sheldon *et al.* (2012), Woznicki *et al.* (2016), Zhang *et al.* (2010)	and Weller (2011), Schmalz *et al.* (2015), Sueyoshi *et al.* (2016), Terrado *et al.* (2016), Weigel (2003)			kick, Lock and Goethals (2014) kick, Maloney and Weller (2011) kick, Moreno *et al.* (2009) surber, Pearson *et al.* (2016) kick, Schmalz *et al.* (2015), Sueyoshi *et al.* (2016) surber, Zhang *et al.* (2010) kick	(2015), Hawkins and Yuan (2016), Hughes *et al.* (2016), Mantyka-Pringle *et al.* (2014), Sanchez *et al.* (2014), Sheldon *et al.* (2012), Terrado *et al.* (2016), Van Sickle *et al.* (2004), Weigel (2003), Woznicki *et al.* (2016)	Goethals (2013), Lock and Goethals (2014), Mantyka-Pringle *et al.* (2014), Schmalz *et al.* (2015), Sueyoshi *et al.* (2016),	(2016), Maloney and Weller (2011), Pearson *et al.* (2016), Sanchez *et al.* (2014), Sheldon *et al.* (2012), Van Sickle *et al.* (2004), Weigel (2003), Woznicki *et al.* (2016), Zhang *et al.* (2010)	

The compiled papers suggested that more studies addressed urban and industrial land uses. Moreover, urban and industrial areas pose more negative consequences toward aquatic ecosystems (7 papers), compared to agricultural (5 papers). A combination of agricultural and urban was also considered to negatively influence the aquatic ecosystems (3 papers, Table 2.2). This result corroborated with the report published by the United Nations Environment Program (UNEP, Table A1). The UNEP has published a list of economic activities with their effects on aquatic ecosystems where industries were identified to pose most threats toward aquatic ecosystems (Carr and Neary, 2008). However, many papers only included land use information to support the analysis but did not specifically study land use effect on the aquatic ecosystems. Moreover, several papers did not mention land use classification following typical classification system (e.g. urban, agricultural and forest). Depending on the purpose of the study, land use was sometimes classified into more detailed classes (e.g. heavy and light pastoral (Clapcott *et al.*, 2017)).

Studies on the effect of land use on ecological water quality in developing countries are still limited. From the 39 selected papers, only eight studies were performed in developing countries (Table A2). Four of these studies were performed in South America, three studies were done in Asia, and one study was done in Africa. However, it is possible that most studies in developing countries have been published in local journals that are not accessible via the Web of Science portal.

Most of the 39 studies mainly focused on local or riparian scale, and only 25% of the papers studied land use effects at both local or riparian and catchment scales. Among the 39 papers, only two papers included land use change (temporal aspect, Table A2) and five papers studied effects of land use change by creating a scenario of future conditions (Table A2).

The land use information is collected in different ways. In addition to the conventional way of field observation, other observation methods and data sources for acquiring land use data exist (Table A3). For example, land use data have been collected via remote sensing (Einheuser *et al.*, 2012; Terrado *et al.*, 2016) and GIS sources (Feio *et al.*, 2007; Mantyka-Pringle *et al.*, 2014), available national database; or a combination of the methods and sources (Table A3). National database and GIS can be available in various forms, e.g. shape file and digital map; however, this was not always specified in the selected papers. Hence, both were considered as separate sources in Table A3. By combining different methods and data sources, the area coverage of land use information can be enlarged beyond the dimensions of field observation.

As explained by Kuemmerle *et al.* (2013), the limited availability of comparable land use data is due to varying land use categories between disciplines. Another reason is adequate approaches to quantify land use and integrate various data sources are often missing. The problem is observed more in developing countries, where sometimes countries lack consistent data collection and data sharing frameworks among institutions (Kuemmerle *et al.*, 2013).

Table 2.2 Effects of land use based on the selected published papers.

Used land use information	Land use effects		
	Positive	**Negative**	**Not defined or not studied**
Urban, industrial		Alemneh *et al.* (2017), Baltazar *et al.* (2016), Carlisle and Meador (2007), Cortes *et al.* (2013), Lock and Goethals (2014), Lock and Goethals (2013), Sanchez *et al.* (2014)	
Agricultural (arable, pasture, orchard, etc)		Barton (1996), Hrodey *et al.* (2009), Pearson *et al.* (2016), Sueyoshi *et al.* (2016), Weigel (2003)	
Forest	Sheldon *et al.* (2012)		
Agricultural + urban		Maloney and Weller (2011), Van Sickle *et al.* (2004), Zhang *et al.* (2010)	
Land use is divided into clear classes			Abouali *et al.* (2016), Alvarez-Cabria *et al.* (2017), Clapcott *et al.* (2017), Dahm and Hering (2016), Damanik-Ambarita *et al.* (2016a), Erba *et al.* (2015), Feio *et al.* (2009), Feio *et al.* (2007), Forio *et al.* (2015), Forio *et al.* (2017), Hawkins *et al.* (2000), Mantyka-Pringle *et al.* (2014), Woznicki *et al.* (2016)
Land use classification is not provided			Bennetsen *et al.* (2016), Davies and Jackson (2006), Hawkins and Yuan

Used land use information	**Land use effects**		
	Positive	**Negative**	**Not defined or not studied**
			(2016), Moreno *et al.* (2009)
Scenario best management practices	Einheuser *et al.* (2012), Hughes *et al.* (2016), Schmalz *et al.* (2015), Terrado *et al.* (2016)		
Scenario crop rotations		Guse *et al.* (2015)	
Mixed use (combination of agricultural, residential, forest, etc)		Carlisle and Hawkins (2008)	

2.3.2 Local or riparian land use scale

Most of the selected papers included land use information at local or riparian scale as this information can be relatively easily collected through field observations during a dedicated sampling campaign (Alemneh *et al.*, 2017; Baltazar *et al.*, 2016; Barton, 1996) (Table 3). Here riparian zone (as described by Crétaz and Barten (2007)) is considered to be comprised of stream valley and terrace slope including stream channel, floodplain and parts of adjacent uplands where aquatic and terrestrial ecosystems interact. Riparian zone acts as storage for flood waters, organic material and nutrients that are transported from uplands to streams. However, the function of riparian zone varies according to residence time of pollutants in the buffer, the thickness of the unsaturated zone and the upland land use (Crétaz and Barten, 2007). Having defined the view on riparian zone, note that in scientific literature the term local was sometimes use for riparian land use. Therefore in this chapter local and riparian are combined into one scale. Here examples from selected papers and other land use related studies are provided (Table 3). Several studies have confirmed the importance of local land use on the water quality (Damanik-Ambarita *et al.*, 2016a). For example, Sanchez *et al.* (2014) studied the importance of urban and Hawkins and Yuan (2016) studied the influence of agricultural areas where human interventions are generally expanded until the edge of the streams. However, many studies included the information of local land use but did not specifically assess its potential effects on the water quality (Davies and Jackson, 2006; Hawkins *et al.*, 2000) or did not find its importance on the ecological water quality after analyses (Feio *et al.*, 2009).

Table 3 Various scales in quantifying land use at local or riparian and catchment scales from selected papers and other land use related studies: otherwise mentioned, the local scale is not described as length, width or radius; scale is given as length×width.

Local or riparian scale (m)	Authors	Catchment scale (km^2)	Authors
30	Abouali *et al.* (2016), Hrodey *et al.* (2009)	17	Rios-Touma *et al.* (2015)
1000 radius	Cortes *et al.* (2013), Feio *et al.* (2007)	6378	Waite (2014)
150 radius	Molina *et al.* (2017)	33	Molina *et al.* (2017)
10, 100, 250, 500, 1000, 2000	Usio *et al.* (2017)	447	Lee *et al.* (2012)
50, 100, 250, 500, 1000, 2500	Thornhill *et al.* (2017)	5896	Wen *et al.* (2016)
250 radius	de Morais *et al.* (2017)	181	Raymond and Vondracek (2011)
200×300	Jayawardana *et al.* (2017)	765	Jayawardana *et al.* (2017)
500-, 1000-, 2500-, 5000×100	Dahm and Hering (2016)	173	Merriam *et al.* (2011)
100, 1000	Meyer *et al.* (2015)	35	Carvalho *et al.* (2011)
500 length or radius	Erba *et al.* (2015), Pearson *et al.* (2016), Mantyka-Pringle *et al.* (2014)	2000	Bellucci *et al.* (2011)
30, 120 width	Van Sickle *et al.* (2004)	9162	Park *et al.* (2011)

2.3.3 Catchment or regional land use scale

The effect of land use at catchment scale has not been studied as much as impact of land use at local or riparian scale (only seven out of 39 papers studied it), despite the potential impact that land use at catchment scale poses on the ecological water quality. Since the area coverage of a catchment can be relatively large (i.e. of a large river), it requires relatively much time and human resources to assess the land use through field observation. Remote sensing via satellite images and aerial surveys (Clapcott *et al.*, 2017) and available GIS data (Sueyoshi *et al.*, 2016) are common methods and source in assessing the catchment land use. The scale of catchment land use varies and is not always mentioned (examples in Table 3). Some studies did not classify the catchment land use or did not study specifically its effects on the ecological water quality (Alvarez-Cabria *et al.*, 2017). However, Carlisle and Hawkins (2008) and Carlisle and Meador (2007) successfully defined land use effects at the catchment scale on macroinvertebrates. They found the degree of land use effects following a sequence of land use classes: mixed land use and urban were reported to have the most adverse effects, whereas forests posed a positive effect. Lastly, Woznicki *et al.* (2016) assessed and classified the catchment land use. However, their study did not assign a key importance to land use and therefore they focused on water quality variables instead.

2.3.4 Recommendation for integrated local or riparian and catchment or regional land use scales

Since the effectiveness of local or riparian areas to store flood waters, organic material and nutrients depends on the catchment's characteristics and regional climate (Crétaz and Barten, 2007), studies on the impact of land use changes on aquatic communities should integrate both local or riparian areas and catchment land use information. For example, Lowrance *et al.* (1997) studied the effectiveness of riparian forest buffer at the Chesapeake Bay watershed based on nutrient transport from agricultural watershed into the coastal plain and the Chesapeake Bay. The diverse and complex relation between local or riparian and catchment land use scales was the reason why 11 out of 39 papers studied the impacts of land use at both riparian and catchment scales. The complementary benefit of combining both land use scales can be seen from the studies done by Weigel (2003) and Cortes *et al.* (2013). Weigel (2003) found out the influence of each scale to determine macroinvertebrate distribution was dominant at certain parts of his study area, but not exclusive of each other. However, Van Sickle *et al.* (2004) found out that riparian land use explained the land use impacts better than catchment land use, while Sheldon *et al.* (2012) concluded the opposite.

When field observation and either remote sensing observation or GIS data are combined, land use data become more informative and area coverage can be enlarged more than what is possible through field observation alone. In the future, more land use data will become available for developing countries through the open source data, especially with the improvement of satellite images, aerial surveys and digital data globally (Rocchini *et al.*, 2017b). For example, Baltazar *et al.* (2016) could access the land use data of Niyugan River Sub-watershed, the Philippines, through Google Earth; while Moreno *et al.* (2009) accessed the land use data of the das Velhas River, Brazil, through digital cartography data. Similarly for this PhD study, remote sensing was done using Google Earth and the GIS data were accessed from the Ministerio de Agricultura, Ganadería, Acuacultura y Pesca (MAGAP) of Ecuador to collect land use data of the Guayas river basin, Ecuador. This way, developing countries nowadays have some modest initial access to land use data and thus have the possibility to improve their ecological water quality studies in relation to land use. For future studies, we recommend combining field observations, remote sensing and whenever possible GIS data sources for local or riparian land use. For catchment land use, remote sensing can be utilized and GIS data sources can be accessed. By combining methods and sources, land use can be quantified for both local and catchment land use scales.

2.3.5 Land use change

Only two out of the 39 papers included temporal aspects of land use and both papers had similar conclusions. Maloney and Weller (2011) found that past land use occurring 50 years ago still influence the present day conditions of streams. Similarly, Schmalz *et al.* (2015) also found negative effects of deforestation on the streams and aquatic ecosystems within a 30 years period.

Besides land use change due to anthropogenic activities, water quality variables may also change due to natural processes (Crétaz and Barten, 2007; Harding *et al.*, 1998) and land use change due to extreme events or natural disasters such as climate change, floods, fires and earthquakes (Barber *et al.*, 2017; Milliman *et al.*, 2017; Strauch *et al.*, 2015; Verkaik *et al.*, 2015). For example, an increase in ammonium-N and nitrate-N concentrations of the Swedish' streams and a decrease in aquatic macroinvertebrate richness and abundance were observed after a flashflood event (Lofgren *et al.*, 2014). Another example is wildfire together with post-wildfire rainfall on riparian vegetation. Besides altering microclimatic conditions, increasing runoff and enhancing erosion, wildfire and post-wildfire rainfall may consequently decrease the richness and abundance of aquatic biota (Bixby *et al.*, 2015).

However, data on past land use changes are often not available or not stored compared to the current day situation and in these cases the effect of land use change is difficult to quantify. The poor availability of land use change information is probably the reason why several studies used land use scenarios to study land use impacts using current situation but without information of past land use (Einheuser *et al.*, 2012; Hughes *et al.*, 2016). Indeed, the need of land use change information depends on the purposes of the studies and is not necessarily required when the study purpose is to assess the effect of current land use. We recommend local and regional government in the developing countries to store their land use information. Data from past or current surveys and projects should be added to local or regional databases and the databases need to be updated and completed for other parts of the country. To update their land use data, developing countries can also access global databases that are continuously developing and are freely available (e.g. GRASS GIS (Rocchini *et al.*, 2017a)). To be able to track and study changes (e.g. in the perspective of climate change, agro-economic developments …), it is important to have both historical and recent data available in these databases.

2.4 Use of models in ecological water quality studies

2.4.1 Input variables

When studying the impact of land use on macroinvertebrates, different types of input variables were used in the models of the selected papers (Table 2.4). Geomorphological variables (e.g. elevation, river banks and sediment type) are the most common type of variables being used in ecological water quality studies (37 papers), followed by physico-chemical (e.g. nutrients and pH; 35 papers) and hydrological variables (e.g. annual discharge and flow; 23 papers). Geomorphological and hydrological variables can be gathered via field observation and in situ sampling. Both geomorphological and hydrological variables can provide information on anthropogenic alteration on the water body. Physico-chemical variables are easily changed within a short period of time; therefore the change in water quality can be relatively easily detected based on long-term data originating from regular monitoring campaigns. Such long-term data series are also required to unravel the variability due to land use changes from the natural variability of the aquatic ecosystem. Some authors were interested in studying certain types of variables only; however, most papers combined different types of variables (Table 2.4).

Table 2.4 Type of input variables.

Type of variables	# of studies	References
Geomorphology (e.g. elevation, river banks and sediment type)	1	Barton (1996)
Hydrology (e.g. annual discharge and flow) + physico-chemical (e.g. nutrients and pH)	1	Sanchez *et al.* (2014)
Geomorphology + meteorology (e.g. rainfall and snow fall)	1	Carlisle and Meador (2007)
Meteorology + physico-chemical	1	Sheldon *et al.* (2012)
Geomorphology + physico-chemical	12	Baltazar *et al.* (2016), Bennetsen *et al.* (2016), Cortes *et al.* (2013), Davies and Jackson (2006), Hrodey *et al.* (2009), Lock and Goethals (2014), Lock and Goethals (2013), Moreno *et al.* (2009), Sueyoshi *et al.* (2016), Terrado *et al.* (2016), Weigel (2003), Zhang *et al.* (2010)
Geomorphology + hydrology	1	Dahm and Hering (2016)
Geomorphology + hydrology + meteorology	1	Van Sickle *et al.* (2004)
Geomorphology + hydrology + physico-chemical	9	Alemneh *et al.* (2017), Damanik-Ambarita *et al.* (2016a), Erba *et al.* (2015), Forio *et al.* (2015), Forio *et al.* (2017), Guse *et al.* (2015), Hawkins *et al.* (2000), Hawkins and Yuan (2016), Maloney and Weller (2011)

Type of variables	# of studies	References
Geomorphology + meteorology + physico-chemical	1	Pearson *et al.* (2016)
Geomorphology + hydrology + meteorology + physico-chemical	11	Abouali *et al.* (2016), Alvarez-Cabria *et al.* (2017), Carlisle and Hawkins (2008), Clapcott *et al.* (2017), Einheuser *et al.* (2012), Feio *et al.* (2009), Feio *et al.* (2007), Hughes *et al.* (2016), Mantyka-Pringle *et al.* (2014), Schmalz *et al.* (2015), Woznicki *et al.* (2016)

2.4.2 Ecological models

The selected papers used different mathematical and statistical techniques to identify, assess and quantify the effect of land use changes on the aquatic community (Table 2.5). Both multivariate techniques and decision trees have been often used to predict the presence of macroinvertebrate taxa based on environmental variables. Several papers used more than one model from the same type or a combination of different types of models in their analyses (Table 2.5).

Table 2.5 Types of models used in ecological water quality studies.

Type of models	# of studies	References
Multivariate analyses (e.g. ordination, taxa distribution, community composition)	10	Barton (1996), Bennetsen *et al.* (2016), Davies and Jackson (2006), Feio *et al.* (2009), Feio *et al.* (2007), Hawkins *et al.* (2000), Hawkins and Yuan (2016), Hrodey *et al.* (2009), Moreno *et al.* (2009), Van Sickle *et al.* (2004)
Regression analyses (i.e. linear, multiple, mixed, structural equation)	4	Damanik-Ambarita *et al.* (2016a), Erba *et al.* (2015), Maloney and Weller (2011), Sheldon *et al.* (2012)
Decision trees (i.e. random forest, regression trees, fuzzy, Bayesian belief networks)	4	Alvarez-Cabria *et al.* (2017), Dahm and Hering (2016), Forio *et al.* (2015), Forio *et al.* (2017)
Ordination + regression	6	Alemneh *et al.* (2017), Carlisle and

Type of models	# of studies	References
analyses		Meador (2007), Sanchez *et al.* (2014), Sueyoshi *et al.* (2016), Weigel (2003), Zhang *et al.* (2010)
Ordination + decision trees analyses	2	Carlisle and Hawkins (2008), Mantyka-Pringle *et al.* (2014)
Decision trees + regression analyses	2	Clapcott *et al.* (2017), Einheuser *et al.* (2012)
Ordination + regression + decision trees analyses	3	Cortes *et al.* (2013), Lock and Goethals (2014), Lock and Goethals (2013)
Software programming model (i.e. Stella visual programming and simulation, SWAT eco-hydrological model, InVEST habitat quality module)	3	Baltazar *et al.* (2016), Guse *et al.* (2015), Terrado *et al.* (2016)
Software programming + ordination	2	Schmalz *et al.* (2015), Woznicki *et al.* (2016)
Software programming + regression	1	Hughes *et al.* (2016)
Software programming + decision trees + regression	1	Abouali *et al.* (2016)
Propensity modelling + regression	1	Pearson *et al.* (2016)

Multivariate analyses were most often used to study the relationship between water quality and environmental variables. Multivariate analyses are useful in analyzing the structure or pattern in the data together with the contributions of the variables. These techniques are useful for a dataset that contains a large number of variables (Crawley, 2007; Greenacre and Primicerio, 2013; Zuur *et al.*, 2007). Ordination, a common multivariate technique, integrates regression and permutation methods and provides easy-to-read graphical outputs (Crawley, 2007; Zuur *et al.*, 2007). Due to their relative simplicity they have been often used in ecological water quality studies. For example, Carlisle and Meador (2007) used multiple discriminant analysis, Feio *et al.* (2009) used multi-dimensional scaling and stepwise multiple discriminant function analysis, and Mantyka-Pringle *et al.* (2014) used principal components analysis. Some disadvantages of these techniques are that the outputs can be difficult to interpret and that associations among variables and distribution patterns do not inherently imply causality (Paliy and Shankar, 2016).

The second most frequently applied methods in the selected papers are regression-based techniques, comprising linear, polynomial, multiple and non-linear regression. Regression analysis estimates parameter values and standard errors of a given dataset by analyzing the relationship between the response and the explanatory variables (Crawley, 2007; Dalgaard, 2008; Zuur *et al.*, 2007). From the selected papers, partial least square regression was used to analyze the ecological water quality of Flint River watershed in Michigan, USA, by Abouali *et al.* (2016), while generalized linear model was used to study the water quality of Alto Minho region, Portugal, by Hughes *et al.* (2016). Linear and logistic regression techniques are useful to develop a precise and concise model from a large dataset. However, linear regression cannot handle missing values, while logistic regression will divide variables with missing values into classes (Tuffery, 2011).

Other types of ecological models that are commonly applied in ecological water quality studies are decision tree models based on classification and regression trees (CART). Decision tree models are simple techniques that can provide clear structure of the data having many explanatory variables and the type of interactions between variables. The basic principle of multivariate analyses lays in its binary recursive partitioning, which is splitting the data along coordinate axes of the explanatory variables. Classification trees are applicable when the response variable is nominal, while regression trees are applicable when the response variable is continuous (Berk, 2008; Crawley, 2007; Zuur *et al.*, 2007). Decision trees are also able to deal with relatively small datasets (Van Echelpoel *et al.*, 2015). For example, Dahm and Hering (2016) utilized boosted regression tree to identify recolonization of source sites for fish and macroinvertebrates in Germany, while Lock and Goethals (2014) used classification trees and random forest to predict the occurrence of Plecoptera in Belgium. Despite their simplicity and ability to deal with datasets containing many variables, decision trees are not robust and should be avoided when there are only few observations in the data (Tuffery, 2011).

A combination of different model types, the so-called ensemble methods, was also proven beneficial in the ecological water quality studies. Alemneh *et al.* (2017)

combined multiple regression analysis and canonical correspondence analysis to identify environmental disturbance affecting macroinvertebrate communities in the Upper Blue Nile, Ethiopia. Analysis of covariance, random forest and boosted regression tree were utilized by Clapcott *et al.* (2017) to predict the expected reference condition for macroinvertebrate communities in New Zealand. Stepwise linear regression in combination with adaptive neuro-fuzzy inference systems were used to define the relationship between macroinvertebrates and environmental variables in Saginaw River watershed, USA (Einheuser *et al.*, 2012). Depending on the purpose, the application of ensemble methods can improve the quality of the results.

2.4.3 Recommendation for statistical analysis and model selection

Researchers studying the effect of land use changes on the ecological water quality can rely on a myriad of ecological models or statistical analyses. The selection of the type of analysis to be used depends on the nature of the data (the type of response and explanatory variables) and the aim of the study. Model selection can also depend on the experience of the modeler because no model can be considered as the best option in every situation (Van Echelpoel *et al.*, 2015). In a regression-based model, the selected model should fit best to the data and produces the least unexplained variation, while bearing in mind the parsimony principle and that all model parameters are statistically significant. Several models may explain a given dataset equally well, while in other cases no single best model can explain a dataset (Crawley, 2007; Zuur *et al.*, 2007). The provided guidelines here on data exploration and model selection serve as a recommendation on how analysis can be done in ecological water quality studies.

Zuur *et al.* (2010) have formulated a scheme for various data exploration techniques, which is a very important step before applying a model (Table 2.6). Not every dataset requires each step, because different model requires different assumptions. Without having the ambition to give a full overview on how to perform a data analysis (for that we refer to specific books, e.g. Witten and Frank (2005) and Zuur *et al.* (2007)), process for example a histogram analysis is not required prior to principal component analysis (PCA). Similarly, normality and homogeneity do not need to be checked before developing regression models, since normality and homogeneity can be verified using the residuals produced by the regression models (Zuur *et al.*, 2010).

Table 2.6 Scheme for data exploration techniques. Y: response variable, X: explanatory variable (Zuur *et al.*, 2010).

1.	**Checking for outliers in Y & X**	boxplot & Cleveland dotplot
2.	**Homogeneity Y**	conditional boxplot
3.	**Normality Y**	histogram or QQ-plot
4.	**Zero trouble Y**	frequency plot or corrgram
5.	**Collinearity X**	variance inflation factor (VIF), scatterplots, correlations & principal component analysis (PCA)
6.	**Relationships Y & X**	(multi-panel) scatterplots, conditional boxplots
7.	**Interactions**	coplots
8.	**Independence Y**	auto correlation function (ACF) & variogram

When the aim of the study is only to understand the data, standard inferential statistics can be applied to get the statistics of the data (Witten and Frank, 2005). In many cases, it is also needed to understand the structure and the underlying causal relationship of the data (descriptive methods) or to find association and make predictions for future observations (predictive methods). Prior to modeling, the aim of the study must be specified to optimize the criterion of interest. Since both descriptive and predictive methods have statistical background, a model will possess some level of explanatory and predictive accuracy (Shmueli, 2010; Witten and Frank, 2005). Therefore, both explanatory and predictive qualities of the models need to be retained and reported (Shmueli, 2010). Here the classification (Table 2.7) and comparison (Table 2.8) of various descriptive and predictive modeling based on Tuffery (2011) are provided to help selecting an appropriate model for analysis. Table 2.8 summarizes the advantages and disadvantages of descriptive and predictive modeling in terms of the required assumptions regarding the problem to be solved, the capacity of the model in treating the data exhaustively within a reasonable period for all cases, and the possibility of the model in handling heterogeneous or incomplete data (Tuffery, 2011). For more detailed explanation on a specific method, the readers are refered to Tuffery (2011), Van Echelpoel *et al.* (2015), Berk (2008) and Zuur (2009).

Table 2.7 Classification of descriptive and predictive modeling and purposes/examples of using them; grey background shows methods that integrate statistics and data analysis (Tuffery, 2011); PLS: partial least squares, (M)ANOVA: (multivariate) analysis of variance, (M)ANCOVA: (multivariate) analysis of covariance.

Type	Family	Sub-family	Algorithm	Purposes/examples of use
Descriptive models	Geometrical models	Factor analysis	Principal component analysis (PCA)	Finding predictors for macroinvertebrate composition (Cortes *et al.*, 2013)
			Correspondence analysis (CA), multiple correspondence analysis (MCA)	CA to understand the distribution of macroinvertebrate taxa among sites (Damanik-Ambarita *et al.*, 2016b)
		Cluster analysis	Partitioning methods (moving centres, k-means, dynamic clouds, k-medoids, etc.)	Classifying reference sites (Hawkins *et al.*, 2000)
			Hierarchical methods (agglomerative, divisive)	Macroinvertebrate classification into biologically similar groups (Carlisle and Meador, 2007)
		Cluster analysis + dimension reduction	Neural clustering (Kohonen maps)	Determining macroinvertebrate distribution (Cereghino *et al.*, 2001)
	Combinatorial models		Clustering by aggregation of similarities	
	Logical rule-based models	Link detection	Search for association rules	
			Search for similar sequences	
Predictive models	Logical rule-based models	Decision trees	Decision trees	Classification and regression trees to define trait and tolerance values that distinguished taxa presence (Carlisle and Hawkins, 2008)
	Models based on mathematical functions	Neural networks	Supervised learning networks (perceptron, radial basis function network, etc.)	Predicting macroinvertebrate occurrence based on environmental variables (Goethals *et al.*, 2007)

Type	Family	Sub-family	Algorithm	Purposes/examples of use
		Parametric or semi-parametric models	Continuous dependent variable: linear regression, ANOVA, MANOVA, ANCOVA, MANCOVA, general linear model (GLM), PLS regression	ANOVA to determine differing average values among steams (Carlisle and Hawkins, 2008), PLS to refine selection of predictors after PCA (Cortes *et al.*, 2013)
			Qualitative dependent variable: Fisher's discriminant analysis, logistic regression, PLS logistic regression	Discriminant analysis to select environmental variables estimating probability of a site belongs to a group (Carlisle and Meador, 2007)
			Count dependent variable: log-linear model	
			Continuous, discrete, count or qualitative dependent variable: generalized linear model (GLM), generalized additive model (GAM)	GLM to identify and quantify interactions between drivers and response variables (Hughes *et al.*, 2016)
	Prediction without model	Probabilistic analysis	k nearest neighbours	Predicting macroinvertebrate presence in a river (Yang *et al.*, 2017)

Table 2.8 Comparison of methods based on Tuffery (2011); CHAID: Chi-squared automatic interaction detector.

Method	Assumptions on the problem to be solved	Capacity in exhaustive processing of databases	Possibility of handling heterogeneous or incomplete data
Clustering models			
Moving centers method and its variants	Yes (fixed number of initial clusters and centers)	yes	Numerical variable and variables without missing values
Hierarchical clustering	No (clusters at level *n* are determined by those at level *n*-1)	No (non-linear algorithm), impossible to process more than several thousand observations	Yes (possible to process non-numeric variables with an *ad hoc* distance)
Neural clustering (Kohonen)	Yes (fixed number of clusters)	Yes	Binary variables must be transformed
Clustering by aggregation of similarities	no	In principle yes, but depends on the implementation	Qualitative variables
Classification and prediction models			
Decision trees	Similar to hierarchical clustering	No (but does not reach the limit as soon as hierarchical clustering)	Some trees such as CHAID must discretize continuous variables
Neural networks perceptrons	No (but the number of hidden neurons must be specified)	No (no learning on several hundred variables)	Binary variables must be transformed
Radial basis function networks	No (but the number of hidden neurons must be specified)	yes	Binary variables must be transformed
Discriminant analysis	Yes (assumptions on the conditional distributions between dependent and independent variables)	yes	Numerical variables and variables without missing values
Discriminant analysis on factorial coordinate of MCA (DISQUAL method)	No (assumptions on conditional distributions between dependent and independent variables can be	yes	Yes (missing values are treated as entirely separate values)

Method	Assumptions on the problem to be solved	Capacity in exhaustive processing of databases	Possibility of handling heterogeneous or incomplete data
	dispensed with)		
Linear regression	Yes (linearity + assumptions on the residuals)	yes	Numerical variables and variables without missing values
Logistic regression, generalized linear model	Yes (linearity + non-complete separation	Yes (using a powerful machine if the number of observations is very large)	Yes (continuous variables with missing values are divided into classes)
Association models			
Search for association	no	Depends on the parameter settings	yes
Similar sequences	no	Depends on the parameter settings	yes

Another modeling type is mechanistic modeling, that derives the relationships between significant variables based on theories and principles that govern the studied system. The resulting model is given in mathematical equations. Examples for surface water are modeling of discharges from wastewater treatment plant, industries and storm water; agricultural/urban runoff; and food chain (Nirmalakhandan, 2002). Paillex *et al.* (2017) and Schuwirth *et al.* (2016) showed the use of such mechanistic models in ecological water quality studies. Mechanistic models allowed them to understand the mechanism behind the presence of taxa based on a combination of traits and environmental conditions (Paillex *et al.*, 2017; Schuwirth *et al.*, 2016). Nevertheless, these mechanistic models have disadvantages. Besides the required long process in building such a mathematical model, there is no guarantee that the mechanistic explanation of the model is correct (Nirmalakhandan, 2002). Especially in ecological studies, available trait information that is necessary in a mechanistic model might not be complete, and there is a possibility that an important variable to understand the system is missing (Paillex *et al.*, 2017; Schuwirth *et al.*, 2016). With the complexities and uncertainties of aquatic ecosystems, it is not surprising that this technique is not as popular as descriptive and predictive models. However, it is not our intention to provide a lengthy discussion on mechanistic models. For those interested, we refer to Nirmalakhandan (2002).

For practicality, a list of typical ecological models based on the nature of the response variable (Table 2.9) is provided here, adapted from Guisan and Zimmermann (2000).

Table 2.9 Statistical approaches for three types of response variables: quantitative, semi-quantitative and qualitative (Guisan and Zimmermann, 2000); WA: weighted averaging, LS: least squares, LOWESS: locally weighted scatterplot smoothing, GLM: generalized linear model, GAM: generalized additive model, PO: proportional odds, CR: continuous regression, MLC: maximum-likelihood classification, DFA: discriminant function analysis.

Type of response variable	Probability distribution	Statistical approach	Modelling technique
Quantitative (continuous)	Gaussian	Multiple regression	WA, LS, LOWESS, GLM, GAM, regression tree
		Ordination	CANOCO
	Poisson	Multiple regression	GLM, GAM
	Negative binomial	Multiple regression	GLM, GAM
Semi-quantitative (ordinal)	Discretized continuous	Multiple regression	PO model, CR model
	True ordinal	Multiple regression	Stereotype model
Qualitative (categorical, nominal)	Multinomial	Multiple regression	Polychotomous logit regression
		Classification	Classification tree, MLC, rule-based class
		Discriminant	DFA
		Environmental envelopes	Boxcar, Convex Hull, point-to-point metrics
	Binomial	Multiple regression	GLM, GAM, regression tree
		Classification	Classification tree
		Environmental envelopes	Boxcar, Convex Hull, point-to-point metrics
		Bayes	Bayes formula

2.4.4 Strengths, weaknesses, opportunities and threats (SWOT) analysis

The most important strength of using ecological models is time saving for analysis, despite the possible large number and various types of input variables in the studies (Table 2.10). Second, researchers can use models to test hypotheses, to understand a studied system and to define further research (Waite, 2014). Ecological models can be used to conceptualize the relationships in ecosystems and, despite their limitations; they allow researchers to integrate expert knowledge in the modeling process, which in its turn is beneficial for management purposes. The third strength of using ecological models is that they can be used for any land use scale or for a specific land use type. Fourth, when land use information is included in the models, certain stressor can be related or traced to its source and the degree of its effect on the water bodies can be estimated. Moreover, as a categorical variable, land use information can be easily quantified during a dedicated sampling campaign without specific equipment. Lastly, the ecological models are also widely applicable in terms of the methodology and results, and could facilitate the communication between researchers and public (e.g. studies by Alvarez-Mieles *et al.* (2013) and Van Sickle *et al.* (2004)).

In the present chapter, several weaknesses of the use of models related to land use have been identified. First, due to the complexity of environmental processes, there is no model that can perfectly explain all environmental processes as a whole (Guisan and Zimmermann, 2000) and pre-analysis may be required to select an appropriate model. Second, models can simplify the selection process of model variables, which might result in final model containing variables that are less suitable based on general ecological knowledge. Third, since current ecological models can accommodate more input variables, sampling campaigns might require higher budgets to collect more data. Yet, financial means were not to be discussed in this chapter. Fourth, available land use and land use change information that can be collected via remote sensing and other sources is still lacking especially in developing countries. Fifth, land use data is not regularly updated, thus any possible land use change and its effects are unknown. Besides, not all countries have all their land use registered, and in some cases the land use is recorded only when a specific sampling campaign is taking place. Lastly, there is no consensus regarding land use assessment methods and their scale effects. Hence, studies on land use effect are still lacking.

The first identified threat regarding the use of ecological models in ecological water quality studies is possible over- or under-fitting of the models compared to reality. This goes hand in hand with the nature of ecological models that over simplify the reality (Guisan and Zimmermann, 2000). The second threat is the use of less appropriate models that may provide misleading results. Moreover, due to ongoing development of models, researchers without sufficient knowledge in modelling might use more recent models instead of older ones which might threat proper use and proper selection of the models. Third, there is over- (where researchers accept the results of the models even though not all variables contained in the models are ecologically suitable) or under-reliance of models' results (when the results of the models are not accepted to support decision making). Fourth, due to lack availability of land use data to be accessed via remote sensing and other sources, sometimes researchers had to use outdated data that might not be useful in the analysis or may give misleading result. After some time, a model also might not be applicable anymore on the area where land use data were collected to develop the model, because land use tends to change quickly. Lastly, over simplification of land use classification to be included in the model may shield the real land use effects in the model results.

Despite the abovementioned weaknesses and threats, there are two main opportunities of using ecological models in studying land use impacts (more detail in Table 2.10). The first obvious opportunity is related to model development. Model development to improve its applicability is ongoing, for instance via involving potential users from an early stage of the development process. Moreover, there are various models available for different ecological study purposes. Thus, the qualities and quantities of collected variables are also improved. Continuous model development is also supported by ongoing capacity building in both developed and developing countries. Free software such as R (R-Core-Team, 2013) is also developed to support modelling activities and is accessible worldwide. Second opportunity is related to land use information. Nowadays, researchers are aware that land use change has a potential anthropogenic impact on the aquatic system, and should be included when assessing multiple stressors conditions. Moreover, land use data can be gathered in various ways, such as during sampling campaign and by accessing the global databases (e.g. GRASS GIS (Rocchini *et al.*, 2017a)), thus increase the availability of land use information. New technologies such as the use of drones to record land use data are promising and cost saving compared to common manned-aircraft survey (Hubbart *et al.*, 2017).

Table 2.10 SWOT analysis for the use of models in studying land use impacts on ecological water quality focusing on macroinvertebrates.

STRENGTHS	**WEAKNESSES**
- Time saving during analysis - Can relate land use and aquatic ecosystems' health - Can incorporate land use impacts in general, per land use class or spatially - Can relate certain pollution to certain land use: source of pollution and its degree - Can incorporate many and different types of variables - Can incorporate expert knowledge in variables selection - Can select key variables influencing the ecological water quality - Wide practical applicability of models and model's results - Ease of communication using model's results - Land use is categorical information that is easily collected - Can support management decision regarding land use	- No one-size-fits-all model - No model can explain/assess all environmental process/interaction as a whole - Simplification of variables selection - Increasing sampling cost to collect more data - Requirement of pre-analysis to select appropriate models for an intended purpose - Lack of available land use data in developing countries - Lack of data of land use change - No consensus of land use assessment methods and their scales of effect
OPPORTUNITIES	**THREATS**
- Continuous model development to improve model applicability - Availability of various models for different purposes - Ongoing capacity building in	- Model's over- or under-fitting - Model's over simplification of reality - Use of less appropriate models may provide misleading results - Over- and under-reliance of model's

developed and developing countries - Availability of free software to run the models - Improvement of variables' qualities and quantities - Higher awareness of land use as the source of anthropogenic pollution - Increasing availability of land use data in developed and developing countries - Availability of different land use assessment methods and sources - Possibility to gather land use data via new technologies (e.g. drones) - Access to global databases	results - The newer the model the better - Use of outdated land use data might not be useful or may be misleading - Fast change of land use - Over simplification of land use classification may shield the real land use effects

2.5 Conclusions

Land use can highly influence ecological water quality but its information is often not included in ecological water quality studies. Since land use can influence the ecological water quality and it can change quickly, it is recommended to include land use information in ecological water quality studies on both local and catchment scales. Various methods and sources to collect land use information are available and are continuously developing; therefore efforts need to be taken to collect land use data through field observation, remote sensing and other sources. Moreover, prior to selecting the most appropriate type of ecological models, one should exactly know what the aim of the study is, how the related research hypothesis is formulated and what type of data are available. Despite model's limitation to explain environmental processes as a whole, models can support a fast and quantitative analysis especially when influence of many variables needs to be evaluated. Developing countries can benefit from huge opportunities of using various ecological models to integrate land use information in ecological water quality studies to support their decision making.

Chapter 3: Materials and methods

Abstract

This chapter provides an overview of the study area, the data collection and the methodologies performed throughout the thesis. The Guayas river basin is a major watershed in Ecuador that has been experiencing intensive agriculture and urbanization activities. The methodologies were performed to determine the ecological water quality status of the Guayas river basin and to determine the effects of land use on the ecological water quality. The data analysis procedure is divided into bioassessment based on macroinvertebrates, modelling techniques and sensitivity analysis. The bioassessment part describes the use of the BMWP-Col and NLSMI biotic indices to calculate the ecological water quality, while the modelling part describes the use of ordination and general linear model (GLM) in relating the BMWP-Col and the environmental variables. Lastly, sensitivity analysis describes how to assess the effect of a certain independent variable on the BMWP-Col.

3.1 Study area

The Guayas river basin is located between 1–3°S and 79–81°W, in the central-western part of Ecuador (Caceres *et al.*, 2002) (Fig. 3.1). The Guayas is one of the major watersheds in Ecuador, together with the Esmeraldas and the Amazon, covering an area of 34,000 km^2 (Gerebizza, 2009; United States Army Corps of Engineers - USACE, 1998). The dry season occurs between July and November, while the rainy season occurs between January and May (Alvarez-Mieles *et al.*, 2013). The Guayas river basin receives 1,849 mm average annual precipitation and discharges in average 200 m^3/s during the dry season and 1,600 m^3/s at the peak flow (Frappart *et al.*, 2017). It drains its water towards the Gulf of Guayaquil (Arriaga, 1989; Frappart *et al.*, 2017; United States Army Corps of Engineers - USACE, 1998).

The whole Guayas river basin consists of two main rivers: the Daule river and the Babahoyo river (Arias-Hidalgo *et al.*, 2013). Within the basin, a large amount of the water is diverted towards the Daule-Peripa reservoir. The Daule-Peripa reservoir has a surface area of approximately 30,000 ha, 6 billion m^3 of water storage capacity and 14,350 m^3/s spillway natural maximum discharge. The reservoir was built to generate electricity, to supply water for irrigation, to control floods and to supply drinking water (Arriaga, 1989; CELEC, 2013; Nguyen *et al.*, 2015; United States Army Corps of Engineers - USACE, 1998). Despite its economic benefits, the Daupe-Peripa reservoir has shown negative impacts on the ecological water quality of the rivers within the Guayas river basin. One major impact is the development of aquatic macrophyte water hyacinth in the reservoir which results in water quality degradation in the downstream regions and hindrance in sustainable operation of the hydroelectric schemes (Nguyen, 2017).

One-third of the population of Ecuador (5.5 million inhabitants, national census 2010) resides in the Guayas river basin (UNSD, 2017). Guayaquil (located at the mouth of the river basin) is the largest and most populous city in Ecuador, where many industries are located. Several other cities (e.g. Quevedo and Vinces) and intensive human activities (e.g. agricultural and industries) are located along the main channels of the Daule and Babahoyo rivers. Agricultural land covers 49% of the Guayas river basin, whereas forest and pasture cover 29% and 13%, respectively (Frappart *et al.*, 2017). Forests are located at higher elevations where human activities are either absent or not intensive. Agriculture, and here especially the cultivation of banana, rice, maize, African palm but also cacao production and fisheries are important industries in the Guayas river basin (Alvarez-Mieles *et al.*, 2013; Caceres *et al.*, 2002; Gerebizza, 2009). Aerial spraying is a common technique for pesticide application especially in banana plantation (Deknock, 2017), while intensive and continuous grazing is a general practice for cattle. In many places, both agriculture and cattle farming occupy the land until the edge of the rivers and reservoir. These extensive industries have been resulting in an increasing demand for farm and domestic lands.

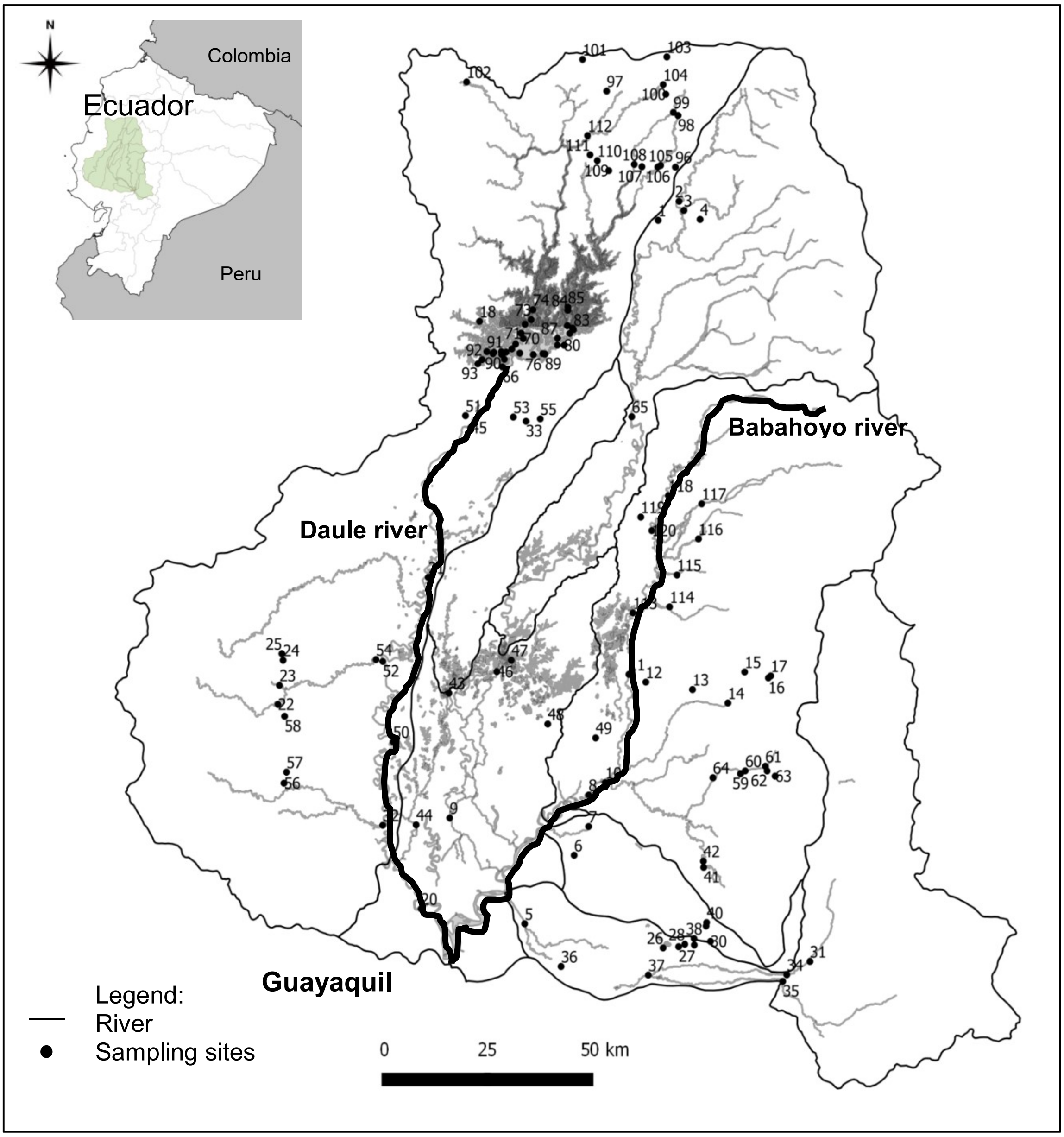


Figure 3.1 Map of the Guayas river basin with indication of the 120 sampling sites; thick lines show the main rivers and gray lines show tributaries.

3.2 Data collection

The sampling campaign was performed from 23 October until 26 November 2013, at the end of the dry season which occurs between July and December (Arriaga, 1989). Data regarding environmental (physico-chemical) and biological (macroinvertebrates) variables were collected at each sampling site. Each site was sampled once. There was no extreme weather, such as heavy rain, during the sampling campaign, and since Ecuador is located in a tropical region, seasonal differences are not as distinct as in temperate regions (Kang and Seager, 2013). This way, extreme environmental conditions affecting the ecological water quality were captured (e.g. conductivity).

The sampling sites (Fig. 3.1) were selected based on an expected gradient of disturbance from relatively pristine (mountainous, clear water, less intensive human activities, and less populated areas) to degraded (low elevation, colored water, intensive human activities, and densely populated areas). Main anthropogenic activities (residential and agricultural) within the Guayas river basin occur along the two main rivers (the Daule and Babahoyo rivers), while forests are located at upstream locations where tributaries are also located (due to the scale of the map, tributaries are not clearly visible). Therefore, a gradient of disturbance can be observed within the selected sampling sites. No exact proportion was allocated for different land uses (i.e. agricultural, forested and residential), however, all sampling sites cover enough representative of each land use type and a quarter of total number of sites was allocated for reservoir. Practical consideration such as accessibility to sampling sites was also considered, because several sites were inaccessible during rainy season. Within the Guayas river basin, 88 sites were sampled along the up- and down-stream locations of the rivers (Fig. 3.1). Since the Daule Peripa reservoir is located at the upstream part of the Guayas river basin and might influence downstream conditions of the rivers, sampling campaign was also performed at 32 sites at the reservoir (Fig. 3.1). By combining rivers and reservoir, the general conditions of the entire Guayas river basin could be assessed.

3.2.1 Physico-chemical variables

Temperature (°C), conductivity (µS/cm), total dissolved solids (TDS, mg/L), pH, chlorophyll a (µg/L), chloride (mg/L), dissolved oxygen (DO, mg/L) and turbidity (NTU) were measured on site using two YSI®6920-V2 multiparameter probes (Yellow Springs, Ohio, United States). To measure the variables, both probes were inserted into a bucket containing a 10 L water sample. When the reading of the probes was stable, the value of each variable was noted.

The measurements of chemical oxygen demand (COD, mg/L), total nitrogen (total N, mg/L), nitrate-N (NO_3^--N, mg/L), nitrite-N (NO_2^--N, mg/L), ammonium-N (NH_4^+-N, mg/L) and total phosphorus (total P, mg/L) were done in the laboratory using Hach-Lange®DR 3900 spectrophotometer kits (Loveland, Colorado, United States). Kits having the lowest detection limits of 5 mg/L, 1 mg/L, 0.5 mg/L, 0.23 mg/L, 0.015 mg/L and 0.015 mg/L for COD, total N, total P, nitrate-N, nitrite-N and ammonium-N, respectively, were used. Water samples from each sampling site were pooled into one sample then stored in cool and dark containers before being transferred into refrigerator to preserve the samples until laboratory analysis. For COD measurement, sulfuric acid H_2SO_4 was added until pH < 2 to preserve the samples. Different treatments were performed to measure different variables using ready-to-use reagents and cuvettes that came together with the Hach-Lange®DR 3900 spectrophotometer kits. The step by step treatment for different measurements was done based on the manual that came together with the kits. The kits also provide necessary liquid for the measurements. The reading of each measurement was done using the kits' visible (VIS) spectrophotometer that has a wavelength range of 320–1100 nm and a wavelength resolution of 1 nm. When water samples were turbid, the samples had to be diluted by adding deionized water that also came with the kits. For quality control, both YSI®6920-V2 multiparameter probes and Hach-Lange®DR 3900 spectrophotometer kits were calibrated following their respected guidelines. Stream width and water depth were quantified using a tape measure, while elevation was measured using a Garmin GPSMap® (Kansas, United States). Flow velocity was measured using the float method as described in the U.S. Environmental Protection Agency (United States Environmental Protection Agency - USEPA, 2012) protocol with a standard length of 5 m.

Using a modified field protocol based on the Australian River Assessment System (AUSRIVAS) physical assessment protocol (Parsons *et al.*, 2002) and the United Kingdom and the Isle of Man River Habitat Survey (RHS) (Raven *et al.*, 1998), the site and its surroundings were also assessed (Table B1). Similar field protocol was used to assess the site and surroundings of rivers and reservoir. The collected information includes land use, macrophytes, riparian vegetation, river banks, channel types, flow types and sediment types. Each variable was divided into different categories. Additionally, aerial photographs from Google and Flash earth maps (from here on is called "Google land use") were used to assess the dominant land use within the direct vicinity of each site in addition to the field protocol (FP, from here on is called "FP land use"). Both Google and Flash earth were consulted in January 2014; however, the resolution and the date when Google collected the data were not recorded. The "FP land use" was assessed within a stretch of 100×10 m (length×width) on the left and right banks of the sampling sites, while the "Google land use" was assessed for a stretch of 100×100 m on the left and right banks of the sampling sites (Table B2 and B3). This way, two types of land use classification within the direct vicinity of sampling sites were collected: the "FP land use" and the "Google land use". Besides the field protocol and Google maps, data regarding dominant land use available in a geographical information system (GIS, from here on is called "GIS land use") were also collected from the Ministerio de Agricultura, Ganadería, Acuacultura y Pesca of Ecuador (Ministerio de Agricultura Ganadería Acuacultura y Pesca - MAGAP, 2015), which was published in 2012. The MAGAP classified the sampling sites into seven categories: residential; agriculture; a mix of agriculture, livestock, forest and urban; a mix of agriculture, livestock and forest; livestock; a mix of livestock and conservation and protection; and conservation and protection. The "GIS land use" was assessed for a stretch of 200×200 m on the left and right banks of each site (Table B2 and B3). All the three land use data were used in chapter 6, while in chapter 4 and 5 only "FP land use" was used.

3.2.2 Biological variable

The macroinvertebrates were sampled using the standardized kick-net method, following the method described by Gabriels *et al*. (2010). A net with a mesh size of 500 µm that was attached to a 0.2×0.3 m metal frame and a 2 m-long handle was used. The sampling was performed for 5 min to cover a stretch of approximately 10–20 m and to cover all different habitats that are present at the site such as macrophytes, bed substrate, litter and parts of terrestrial vegetation that are immersed in the water. Macroinvertebrates were also picked manually from stones and leaves to collect their highest possible richness. For sites located at the reservoir, the macroinvertebrates were sampled at the shorelines. Whereas for sites located away from the shorelines, the macroinvertebrates were sampled from the macrophytes. The macroinvertebrates were then sorted from the samples and identified to family level (Bailey *et al.*, 2001; Barbour *et al.*, 1999; Marshall *et al.*, 2006) in the laboratory. Macroinvertebrate's identification was done using the identification keys of De Pauw *et al.* (1996) and Domínguez and Fernández (2009). The identification keys guided a step by step identification of macroinvertebrate's physical characteristics such as legs and thorax. For certain taxa, the identification had to be done under a microscope for a better visualization of body parts. Each macroinvertebrate family was also identified according to its functional feeding group (FFG) based on Mereta *et al.* (2013), Barbour *et al.* (1999) and Helson and Williams (2013), with relevance to the river continuum concept (Vannote *et al.*, 1980). The FFG was classified into scrapers, shredders, collectors and predators, which distinguish taxa's behavior of food acquisition.

3.3 Data analysis

3.3.1 Bioassessment based on macroinvertebrates

Bioassessment based on macroinvertebrates has been increasingly used in ecological water quality studies together with physico-chemical assessment of the water. Macroinvertebrates are considered useful bioindicators because they are sensitive to organic pollution and environmental change of their habitats, are ubiquitous and present abundantly, have relatively long life cycles, and have varying feeding habits (Fierro *et al.*, 2015; Mwedzi *et al.*, 2016; Rosenberg and Resh, 1993). Ecuador does not have its own biotic index and the goal of the PhD study was not to develop a new biotic index. Therefore, the ecological water quality of the Guayas river basin was calculated using available biotic indices: the Biological Monitoring Working Party (BMWP) adapted for Colombia (BMWP-Col) and the Neotropical Low-land Stream Multimetric Index (NLSMI). The applied indices for the PhD study were selected based on a review over several biotic indices that were locally developed and used in the middle and South America (Damanik-Ambarita *et al.*, 2016b). A brief description of the applied indices is provided here. The calculation of biotic indices was done altogether for 120 sampling sites, except for the NLSMI where the calculation was done first for 120 sampling sites and separately for reservoir and rivers lower and higher than 250 m (see chapter 4).

3.3.1.1 Biological Monitoring Working Party for Colombia (BMWP-Col)

The Biological Monitoring Working Party (BMWP) adapted for Colombia (BMWP-Col) (Roldán Pérez, 2003) was used to calculate the ecological water quality index of the sampling sites. For the present study, the BMWP-Col based on Alvarez (2005) was chosen among all available BMWP-Col versions since it contained most of the encountered taxa. The BMWP-Col was used since Ecuador does not have its own water quality index so far. This index is considered an appropriate index for Ecuador since it was developed in a country having relatively similar environmental conditions and fauna to Ecuador (Dominguez-Granda *et al.*, 2011b). The BMWP-Col was calculated based on macroinvertebrate community composition, where each macroinvertebrate taxon is assigned with a certain tolerance score, ranging from 1 to 10. Low tolerance scores represent tolerant taxa, while high scores represent sensitive taxa. The BMWP-Col score for each site was obtained by adding up the scores of all families that are present at a site. A good ecological water quality has a BMWP-Col score of more than 100, moderate, poor, bad and very bad ecological water qualities have scores of 61–100, 36–60, 16–35 and 0–15, respectively (Alvarez, 2005).

In addition to the BMWP-Col, the average score per taxon (ASPT) index was calculated. The ASPT was calculated to define an ecological water quality index that is independent of taxonomic richness. The ASPT was calculated by dividing the BMWP-Col score with the number of taxa encountered per site, which ranges from 0 to 10. An ASPT score higher than 6 indicates clean water, 5-6 indicates doubtful quality, 4-5 probable moderate pollution, and lower than 4 indicates probable severe pollution (Armitage *et al.*, 1983; Mandaville, 2002). Moreover, high ASPT scores indicate clean upstream sites containing relatively large numbers of high scoring taxa; while opposite is true for low ASPT scores (Armitage *et al.*, 1983).

3.3.1.2 Neotropical Low-land Stream Multimetric Index (NLSMI)

The Neotropical Low-land Stream Multimetric Index (NLSMI) incorporates several individual metrics. The NLSMI was chosen because it was developed specifically for low-land areas of Panama, a country with a relatively similar climate to Ecuador. It was calculated using the formula described by Helson and Williams (2013) based on seven individual metrics. The metrics used in the calculation are: % of scrapers, Margalef's index, ratio of Chironomidae/Diptera individuals, number of Ephemeroptera Plecoptera Trichoptera (EPT) taxa, % of Trichoptera, % of shredders and Shannon-Wiener Evenness index. The sampling sites were divided into reference and impaired sites, which were required to standardize the metrics values to unit-less scores ranging between 0 and 1. The final NLSMI values were calculated by multiplying the sum of the seven metrics values with 1.43. A NLSMI value higher than 8 indicates reference condition, 6-8 indicates good condition, 4-6 moderate condition, 2-4 indicates poor condition, and lower than 2 indicates bad condition (Helson and Williams, 2013). For this PhD study, the selection of reference sites was based on both the dissolved oxygen Prati index (Goethals and De Pauw, 2001; Prati *et al.*, 1971) and the degree of habitat degradation as described by Barbour *et al.* (1999), Hruby (2004), USEPA (2002) and Mereta *et al.* (2013).

3.3.2 Modelling techniques

To find the relationship between the ecological water quality and environmental variables, ordination and general linear model (GLM) were used. The ordination was used to find the key environmental variables influencing the distribution of taxa composition within the sampling sites, while the GLM was used to find key environmental variables influencing the BMWP-Col. Data from 120 sampling sites were used as one dataset to understand the ecological water quality of the Guayas river basin as a whole.

3.3.2.1 Ordination

Ordination is a multivariate technique where sample distribution is arranged based on eigen analysis or the similarity/dissimilarity among the samples. Ordination projects a multidimensional system onto a two- or three-dimensional map (Beals, 1984; Guo *et al.*, 2015). There are two types of ordination: constrained and unconstrained ordinations. Constrained ordination associates two or more datasets in the ordination process at the same time. This technique includes redundancy analysis (RDA), distance-based redundancy analysis (db-RDA), canonical correspondence analysis (CCA) and multiple factor analysis. Unconstrained ordination analyzes only one dataset using a reduced set of orthogonal axes. The major structure of the dataset is presented in a graph for the user to interpret. This technique includes correspondence analysis (CA), principal components analysis (PCA), principal coordinate analysis (PCoA) and nonmetric multidimensional scaling (NMDS) (Guo *et al.*, 2015).

In this PhD study, a correspondence analysis (CA) was executed to find the relationship between environmental variables and the abundance of macroinvertebrates, using the Vegan package (Oksanen *et al.*, 2013) that is available in R software (R-Core-Team, 2013). The CA was selected because an indirect ordination with taxa data will already reflect environmental influence interpreted by taxa distribution, whereas a direct ordination using environmental variable data will focus more on the environmental variables than the taxa composition (Beals, 1984). The CA was applied on taxa count abundance data. As an unconstrained ordination technique, the CA calculates the ecological distance between sites and taxa. To find the influencing environmental variables, the data of environmental variables were then fitted on the CA graph. The fitted environmental variables show their direction in the ordination graph for sites with the environmental values higher than the average (Kindt and Coe, 2005; Oksanen *et al.*, 2013). Taxa count abundance and continuous variables were $\log_{10}(x+1)$ transformed before analysis.

3.3.2.2 General linear model (GLM)

General linear model (GLM) is a method to determine the relationship between dependent (response) and one or multiple independent (explanatory) variables. GLM works with a response variable having a Gaussian distribution (Helsel and Hirsch, 1992; Zuur, 2009). GLM was applied to define the relationship between the BMWP-Col (a continuous variable with Gaussian distribution) and environmental variables and to determine the key environmental variables influencing the BMWP-Col (Weirich *et al.*, 2011; Zuur, 2009). GLM has proven its ability in studying the relationship among variables in ecological-related data (Guisan *et al.*, 2006; Nguyen *et al.*, 2015; Thuiller, 2003).

For the objectives of the PhD study, the continuous variables were not transformed before analysis to avoid complication and difficult interpretation of the models afterwards (Shmueli, 2010) and to avoid changing functional relationship between response and explanatory variables (Austin, 2002). There was no removal of outliers from the analysis since they are real observations (not technical errors) and to avoid reducing the number of observations; whereas the categorical variables were set as factors (Zuur *et al.*, 2007). To measure model's stability, the GLM was developed and validated using three-fold cross validation (CV). The CV was done because there is only one dataset and assigning some part of the data only for validation will reduce the number of observation for the analysis (Witten and Frank, 2005). The three-fold CV was done by splitting the complete dataset randomly into three equal subsets, where the BMWP-Col classes were used to stratify the dataset prior to splitting. The use of the BMWP-Col classes in stratifying the dataset was based on the study by Everaert *et al.* (2013) who used the ecological quality ratio (EQR) status in their analysis. Two subsets were used to develop (train) the model and the remaining subset was used for model validation (testing) (Dedecker *et al.*, 2005; Witten and Frank, 2005). Each subset was used for model validation once. Hence, the dataset produces three final models.

Model fitness was examined using the drop1 command in R that is applicable as a standard command (R-Core-Team, 2013) that removed one variable each time, starting from the variable having the least significant p-value in a model. As a standard procedure, the drop1 command also performed an F-test based on the residual sum of squares and provides Akaike information criterion (AIC) of the model. Variable removal using the drop1 command is continued and the AICs of different model configurations were compared. The model having the lowest AIC was retained as the best model, because a model with a lower AIC better fits the data (Zuur, 2009). R software version 3.0.2 (2013-09-25) was used to perform the GLM analyses, and the drop1 command is available in R without specific packages (R-Core-Team, 2013).

3.3.3 Sensitivity analyses

To assess the effect of a certain independent variable on the dependent variable (i.e. BMWP-Col), a sensitivity analysis was performed by varying variable values one at a time (Jackson *et al.*, 2000). Model sensitivity analysis is useful to get reliable outputs from various model predictions and results, because model predictions and results may not always match the observed data (Guo *et al.*, 2015). The effect was tested under a given situation: the values of a variable that needed to be assessed were ranged between its minimum and maximum values, while the values of the remaining variables were set constant to their median values (Everaert *et al.*, 2010; Goethals *et al.*, 2007; Mouton *et al.*, 2010). In the current PhD study, the sensitivity analysis was performed on the model having the best performance. A similar way of performing the sensitivity analysis was used for the continuous and categorical variables. However, to simplify the analysis, the median category for categorical variables was chosen from the most prevalent class of the categorical variables.

Chapter 4: Ecological water quality analysis of the Guayas river basin (Ecuador) based on macroinvertebrates indices

Adapted from:

Minar Naomi Damanik Ambarita, Koen Lock, Pieter Boets, Gert Everaert, Thi Hanh Tien Nguyen, Marie Anne Eurie Forio, Peace Liz Sasha Musonge, Natalija Suhareva, Elina Bennetsen, Dries Landuyt, Luis Elvin Dominguez Granda, Peter L.M. Goethals. (2016). Ecological water quality analysis of the Guayas river basin (Ecuador) based on macroinvertebrates indices. Limnologica 57, 27-59. DOI: 10.1016/j.limno.2016.01.001.

and

Minar Naomi Damanik Ambarita, Gert Everaert, Marie Anne Eurie Forio, Thi Hanh Tien Nguyen, Koen Lock, Peace Liz Sasha Musonge, Natalija Suhareva, Luis Elvin Dominguez Granda, Elina Bennetsen, Pieter Boets, Peter L.M. Goethals. (2016). Generalized Linear Models to Identify Key Hydromorphological and Chemical Variables Determining the Occurrence of Macroinvertebrates in the Guayas River Basin (Ecuador). Water 8(7). DOI: Artn 297 10.3390/W8070297.

Abstract

In this chapter, the general ecological water quality of the Guayas river basin, Ecuador was assessed. The Guayas river basin is one of the major watersheds in Ecuador, where increasing human activities are affecting water quality and related ecosystem services. The aims of this chapter were (1) to assess the ecological water quality based on macroinvertebrate indices and (2) to determine the major environmental variables affecting the distribution of macroinvertebrate taxa. To do so, two biotic indices were calculated to assess the water quality with an ecological approach: the Biological Monitoring Working Party Colombia (BMWP-Col) and the Neotropical Low-land Stream Multimetric Index (NLSMI). Both the BMWP-Col and NLSMI indicated a good water quality at the (upstream) forested locations, a lower water quality for sites situated at arable lands and a bad water quality at residential areas. Both indices gave relevant assessment outcomes and can be considered valuable for supporting the local water management. Additionally, the average score per taxon (ASPT) was also calculated to assess the calculation of the BMWP-Col that is independent of taxonomic richness. The comparison between the BMWP-Col and NLSMI proved the suitable use of the BMWP-Col to evaluate the ecological water quality of the Guayas river basin, and this conclusion was further confirmed by the ASPT calculation. A correspondence analysis (CA) applied on macroinvertebrate abundance data and subsequently fitted with environmental variables suggested that flow velocity, chlorophyll a concentration, conductivity, land use, sludge layer and sediment type were the major environmental variables determining the ecological water quality. Since actual concentrations of nutrients were not available for all sampling sites, the real influence of nutrients on the ecological water quality could not be evaluated. Therefore, future monitoring needs to be done to investigate the influence of nutrients and other variables such as pesticides in the area where intensive agricultural activities take place. .

4.1 Introduction

Human activities such as agriculture, residential expansion, reservoir development as well as hydrological alterations of the water body can change the environmental conditions of the water and thus affect the presence of aquatic macroinvertebrates. Prevailing water conditions determine the diversity of benthic macroinvertebrates, which make them an ideal indicator to study water quality (Helson and Williams, 2013). The information of benthic macroinvertebrates was used to develop biological indices such as the Biological Monitoring Working Party (BMWP) for Great Britain (Armitage *et al.*, 1983; Hawkes, 1998) and the Multimetric Macroinvertebrate Index Flanders (MMIF) for Flanders, Belgium (Gabriels *et al.*, 2010). Nowadays, many developed countries have used biotic indices together with physico-chemical water quality variables for their routine water-quality monitoring (e.g. Water Framework Directive for member states of European Union (Hering *et al.*, 2010), Clean Water Act for United States (Govenor *et al.*, 2017)). Usually, the condition of water bodies under examination is compared with the condition of water bodies at reference sites, which are less impacted by environmental stressors (Romero *et al.*, 2013; Van den Brink *et al.*, 2011).

Recognizing the need to assess the water quality, several South American countries have performed water quality analyses by applying the methodologies developed in Europe and North America. Examples of studies are the water quality assessment of the Cauca river (Holguin-Gonzalez *et al.*, 2013) and Opia river (Forero-Cespedes *et al.*, 2013) in Colombia and the wetland area of Abras de Mantequilla in the Guayas river basin (Alvarez-Mieles *et al.*, 2013; Arias-Hidalgo *et al.*, 2013) of Ecuador. Moreover, the BMWP index and its adapted versions were used to study water quality in several countries, such as Brazil and Colombia (Ferreira *et al.*, 2011; Forero-Cespedes *et al.*, 2013). However, since biological monitoring methods were mostly developed in Europe or North America, their applicability in developing countries can be debated (Everaert *et al.*, 2014). To solve this problem, several countries have developed their own biological indices, such as the Neotropical Low-land Stream Multimetric Index (NLSMI) to assess rivers in Panama (Helson and Williams, 2013) and the Multimetric Macroinvertebrate Index to assess wetlands in southwest Ethiopia (Mereta *et al.*, 2013).

Previous water quality studies in the Guayas river basin that incorporated macroinvertebrates were only performed in one wetland area (Alvarez-Mieles *et al.*, 2013; Arias-Hidalgo *et al.*, 2013) and consequently did not represent the water quality of the whole river basin. Due to the multiple anthropogenic pressures present in the basin, water quality and quantity can be compromised, so there is a need to study the complete Guayas river basin on a broader scale. As one of the major watersheds in Ecuador, the Guayas river basin plays an important role as a water source in the country (United States Army Corps of Engineers - USACE, 1998). In this chapter, the change in water quality based on macroinvertebrates was evaluated by studying the conditions of the up- and down-stream water bodies situated in the river basin. The Daule-Peripa reservoir and the major rivers were included to get a better understanding of the water quality status of the Guayas river basin. Thus, the objectives of this chapter are (1) to determine the ecological water quality of the Guayas river basin based on macroinvertebrate indices and (2) to identify physico-chemical variables significantly affecting the distribution of macroinvertebrate taxa.

4.2 Data analysis

In this chapter, the BMWP-Col (Roldán Pérez, 2003) based on Alvarez (2005) and the NLSMI (Helson and Williams, 2013) were used to calculate the ecological water quality of the Guayas river basin. These two indices were selected based on a review of several indices that have been locally developed and used in the middle and South America (Table C8). To calculate the NLSMI index, the reference sites should have oxygen Prati index lower than 2 (Table C1) and adapted habitat disturbance score (adapted from Barbour *et al.* (1999), Hruby (2004), USEPA (2002) and Mereta *et al.* (2013)) lower than 18 (calculated based on Table C2). The NLSMI was first calculated for all 120 sites, and then separately, based on the elevation and the types of sampling sites (rivers or reservoir). For each calculation, five sampling sites were chosen as reference sites and the boxplots of seven individual metrics were produced. Since elevation might influence macroinvertebrate community composition, it might influence the BMWP-Col calculation. Therefore, the ASPT (Armitage *et al.*, 1983; Mandaville, 2002) index was also calculated to define an ecological water quality that is independent of taxonomic richness. The ASPT values were then related with the elevation and their coefficient of determination (R^2) was calculated. A strong correlation would confirm the influence of elevation on index calculation and vice versa. The degree of habitat degradation was calculated as well, using an adapted habitat disturbance score (Table C2) as described by Barbour *et al.* (1999), Hruby (2004), USEPA (2002) and Mereta *et al.* (2013). The functional feeding group (FFG) was also checked in relation to elevation.

All analyses including data exploration were done using R software (R-Core-Team, 2013) and following the methods described by Zuur *et al*. (2009). The summary statistics of all measured continuous variables are presented in Table 4.1. Due to a human error, the COD of 30 sites could not be measured, whereas the COD of 6 sites and the missing values of total N, total P, nitrate-N, nitrite-N and ammonium-N were due to their concentrations below the detection limits of the kits. Due to practical limitations, the width and depth of sampling sites located at the reservoir and at big rivers could not be measured either. By taking a summary of the original data, information about the missing values is gained. In case of missing values due to the concentrations below the detection limits, they were assigned the values of the detection limits. The values of the detection limits were chosen to replace the missing values for concentrations below the detection limits in order to accommodate possible highest concentrations the samples could have. A set of data was prepared in which no missing values are left. By doing so, only three variables had missing values in the preprocessed data (i.e. due to human errors and practical limitations): COD, stream width and water depth.

For this chapter, original data (Table 4.1) were used. All variables with missing values (i.e. COD, total N, total P, nitrate-N, nitrite-N, ammonium-N, stream width and water depth) were removed before all analyses. Elevation was also excluded from analyses, since it might influence the distribution of macroinvertebrates in correspondence analysis (CA). FP land use was used for land use information for this chapter. All continuous variables (e.g. DO, conductivity, chlorophyll a, turbidity and velocity) were $\log_{10}(x+1)$-transformed before Pearson correlation analysis (using a cut-off value of 60%) and CA to have more normally distributed data. Another reason is to avoid a strong influence of variables with extreme values from dominating the analysis. Since several variables are correlated with one or more variables, using only one of them as a proxy (e.g. turbidity and total dissolved solids are correlated and using either of them is enough to assess the influence of dissolved solids) is useful to reduce the number of variables to be included in the analysis. Another reason of excluding correlated variables is to avoid the arch effect that might occur when using CA when many variables are used. Based on Pearson correlation analysis, temperature, TDS, pH, chloride, sediment matrix, bed compaction, valley form, and width variation were removed. A CA was performed on $\log_{10}(x+1)$-transformed taxa count abundance data to find the distribution of macroinvertebrate taxa. One site (site 10) was excluded from the CA since no macroinvertebrates was found. The non-correlated environmental variables were fitted on the CA graph to define their relationship with the abundance of macroinvertebrates, using the Vegan package in R software (Oksanen *et al.*, 2013). A detailed explanation of each analysis is given in chapter 3.

Table 4.1 Mean, median, minimum, maximum and standard deviation of continuous variables measured in the Guayas river basin. Lowest detection limits (LDL) by the Hach-Lange kits were 5 mg/L, 1 mg/L, 0.5 mg/L, 0.23 mg/L, 0.015 mg/L and 0.015 mg/L for COD, total N, total P, nitrate-N, nitrite-N and ammonium-N, respectively. Original data show variables with missing values, preprocessed data show variables where missing values due to below detection limits were replaced by kit's LDL values. *Measurements below detection limits are reported as the detection limits.

Variables	Original data						Preprocessed data					
	Mean	Median	Min	Max	Std. dev.	# missing values	Mean	Median	Min	Max	Std. dev	# missing values
Temperature (° C)	26.0	26.0	19.0	34.0	2.5	-	26.0	26.0	19.0	34.0	2.5	-
Conductivity (µS/cm)	200	123	37	1981	238	-	200	123	37	1981	238	-
Total dissolved solids (g/L)	0.13	0.08	0.05	1.27	0.15	-	0.13	0.08	0.05	1.27	0.15	-
pH	7.7	7.6	6.6	8.9	0.5	-	7.7	7.6	6.6	8.9	0.5	-
Chlorophyll a (µg/L)	5.6	3.1	0.7	66.8	8.7	-	5.6	3.1	0.7	66.8	8.7	-
Dissolved oxygen (mg/L)	7.5	7.8	2.0	13.6	1.7	-	7.5	7.8	2.0	13.6	1.7	-
Turbidity (NTU)	9.8	3.4	0.0	355.6	35.1	-	9.8	3.4	0.0	355.6	35.1	-
Chemical oxygen demand (mg/L)	18.0	16.1	5.2	117.6	23.9	36	17.0	13.3	5.0*	117.6	14.9	30
Total nitrogen (mg/L)	1.7	1.0	1.0	7.7	3.8	102	1.1	1.0	1.0*	7.7	0.6	-
Total phosphorus (mg/L)	2.7	2.7	0.8	4.5	0.2	118	0.5	0.5	0.5*	4.5	0.4	-
Nitrate-nitrogen (mg/L)	0.53	0.35	0.24	2.00	11.90	64	0.37	0.23	0.23*	2.00	0.30	-
Nitrite-nitrogen (mg/L)	0.105	0.027	0.015	0.792	0.210	107	0.025	0.015	0.015*	0.792	0.073	-
Ammonium-nitrogen (mg/L)	0.205	0.056	0.016	8.800	0.841	3	0.204	0.055	0.015*	8.800	0.837	-
Chloride (mg/L)	7.3	2.5	0.5	181.7	22.8	-	7.3	2.5	0.5	181.7	22.8	-
Flow velocity (m/s)	0.2	0.2	0.0	1.5	0.3	-	0.2	0.2	0.0	1.5	0.3	-
Elevation (m)	135	82	2	1075	187	-	135	82	2	1075	187	-
Average stream width (m)	22.5	12.0	1.5	230.0	32.1	32	22.5	12.0	1.5	230.0	32.1	32
Average stream depth (m)	0.40	0.36	0.03	1.00	0.22	40	0.40	0.36	0.03	1.00	0.22	40

4.3 Results

The summary statistics of all measured physico-chemical variables are presented in Table 4.1. Temperature ranged from 19° C to 34° C, due to differences in the time of sampling (early morning or midday). The lowest conductivity was observed at the reservoir (36.5 µS/cm), while the highest at a small tributary of the Daule river, which was almost dry (1981 µS/cm). The pH ranged from 6.56 to 8.87. Chlorophyll a ranged from 0.73 µg/L to 66.84 µg/L, with the lowest value (0.73 µg/L) was observed at an upstream location of a small tributary of the Babahoyo river and the highest value (66.84 µg/L) was observed at the location where also the highest conductivity was observed. DO ranged from 1.97 mg/L to 13.63 mg/L, where the highest value (13.63 mg/L) was observed at the location where the highest chlorophyll a and conductivity values were measured. The lowest oxygen concentration (1.97 mg/L) was observed at a tributary of the Daule river. A higher turbidity was observed at downstream locations of both the Daule and Babahoyo rivers (more than 10 times the mean value). Based on the Pearson correlation analysis temperature, TDS and pH were excluded from further analyses since they were highly correlated.

In total, more than 19,000 macroinvertebrates were sorted and identified, which lead to 83 different families. At one location of the Babahoyo river, no macroinvertebrates were found. The highest richness was observed in two locations situated in mountainous areas, each containing 26 families. Insect larvae constituted the highest number of families (61 out of 83 families), with Coleoptera, Diptera, Hemiptera and Trichoptera (12, 11, 11 and 10 families, respectively) as the main orders. Chironomidae was the most frequently encountered taxon, followed by Baetidae and Acari (100, 64 and 56 sites, respectively). Chironomidae was also the most abundant family, succeeded by Thiaridae and Acari (in total 5683, 2357 and 2170 animals, respectively). Table C3 presents the list of encountered taxa, their tolerance scores based on BMWP-Col by Alvarez (2005), the number of presences in the sampling sites and the functional feeding group (FFG) based on Mereta *et al.* (2013), Barbour *et al.* (1999) and Helson and Williams (2013).

4.3.1 Comparison between the BMWP-Col and NLSMI

The water quality for all 120 sampling sites based on the BMWP-Col ranged from 0 to 168 (Fig. 4.1), and from 0 to 9.1 for the NLSMI (Fig. 4.2). Both indices had high values at sites located at higher elevations (Fig. C1) with DO concentrations between 6 and 10 mg/L, a conductivity lower than 300 µS/cm, a chlorophyll a concentration lower than 4 µg/L, a turbidity lower than 20 NTU, a flow velocity higher than or equal to 0.2 m/s and a thin sludge layer (less than 5 cm). High BMWP-Col was also indicated by sites with a water depth lower than 100 cm, while water depth lower than or equal to 50 cm indicated high NLSMI values. A coarse sediment type indicated high BMWP-Col as well, whereas the type of sediment did not influence the NLSMI. The highest BMWP-Col value was noticed at one of the two locations where the number of taxa was also the highest (26 taxa), at an upstream location of the tributary of the Babahoyo river (Fig. 4.3). In addition, the number of taxa plays a bigger role in determining the ecological water quality compared to the highest tolerance score observed at each site (Fig. C2). The highest NLSMI value was observed at an upstream location of the tributary of the Babahoyo river. Since the NLSMI values based on the elevation differentiate the rivers (lower or higher than 250 m, Fig. C3), the NLSMI values for these two types of rivers and reservoir were plotted separately (Fig. C4). This plot indicated that rivers located at an elevation higher than 250 m had higher NLSMI values than rivers located at an elevation lower than 250 m and sites located at the reservoir. For comparison, the plot of the BMWP-Col values for both types of rivers and reservoir is also presented (Fig. C5).

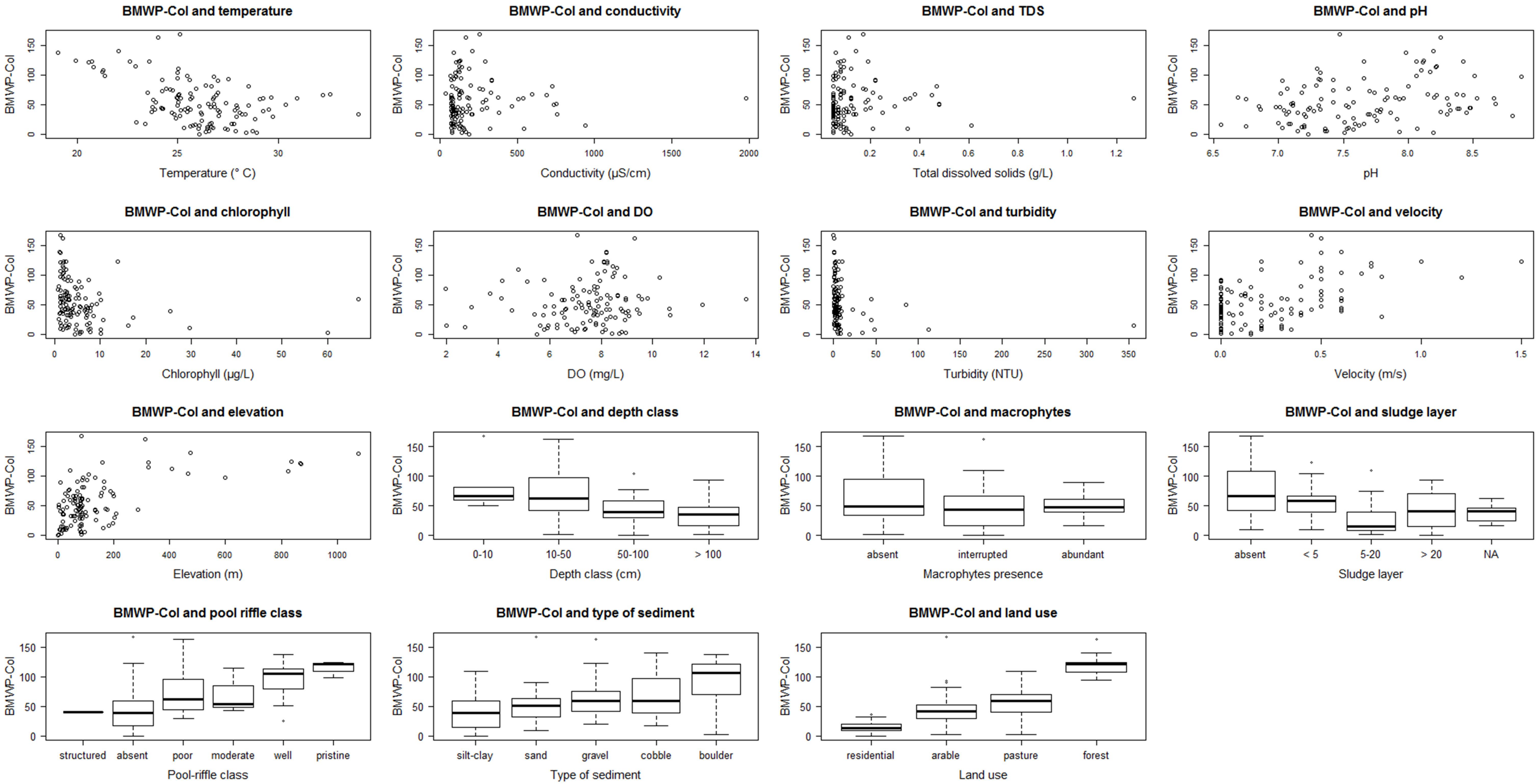


Figure 4.1 Data exploration of the physico-chemical variables plotted against the BMWP-Col for 120 sampling sites. The classification of depth class, presence of macrophytes, sludge layer, pool-riffle class, type of sediment and land use is based on Table B1.

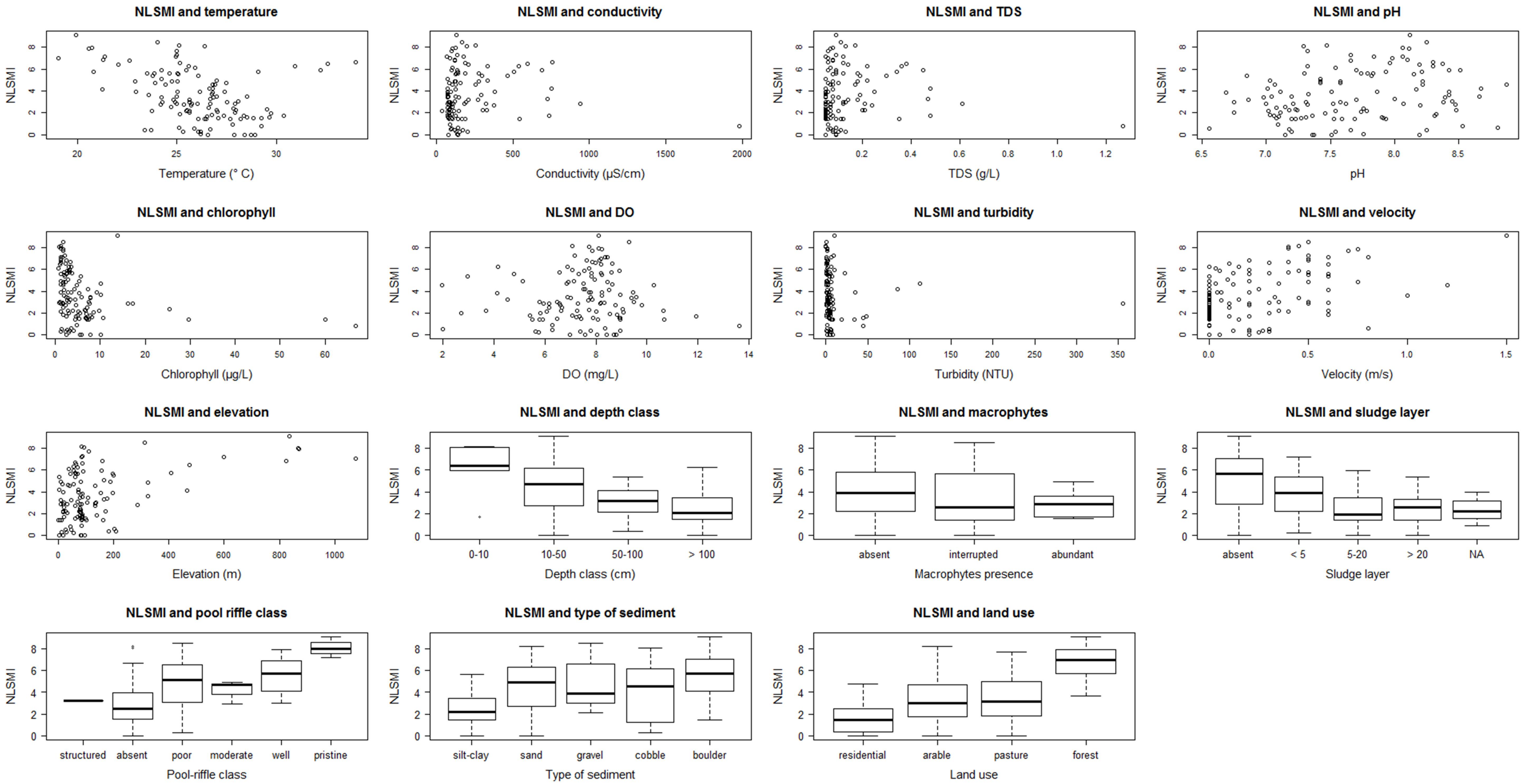


Figure 4.2 Data exploration of the physico-chemical variables plotted against the NLSMI for 120 sampling sites. The classification of depth class, presence of macrophytes, sludge layer, pool-riffle class, type of sediment and land use is based on Table B1.

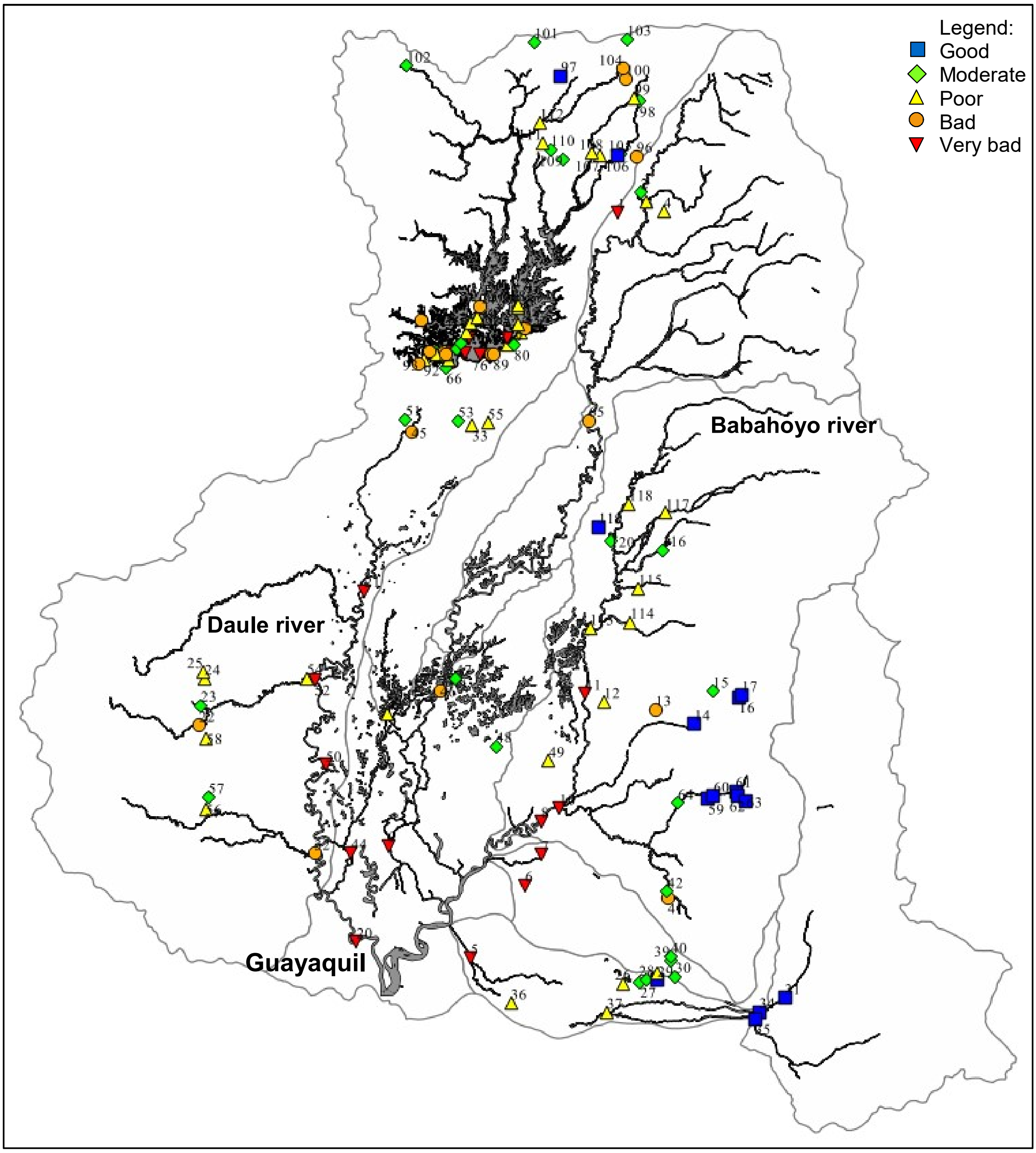


Figure 4.3 Sampling sites in the Guayas river basin with indication of the ecological water quality based on the BMWP-Col ranging from good to bad, as shown in the legend.

There was a positive correlation between the BMWP-Col and NLSMI. The coefficient of determination R^2 was relatively good (0.6) and the p-value approximated zero ($p < 0.001$) for the correlation between both indices (Fig. 4.4 and C6).

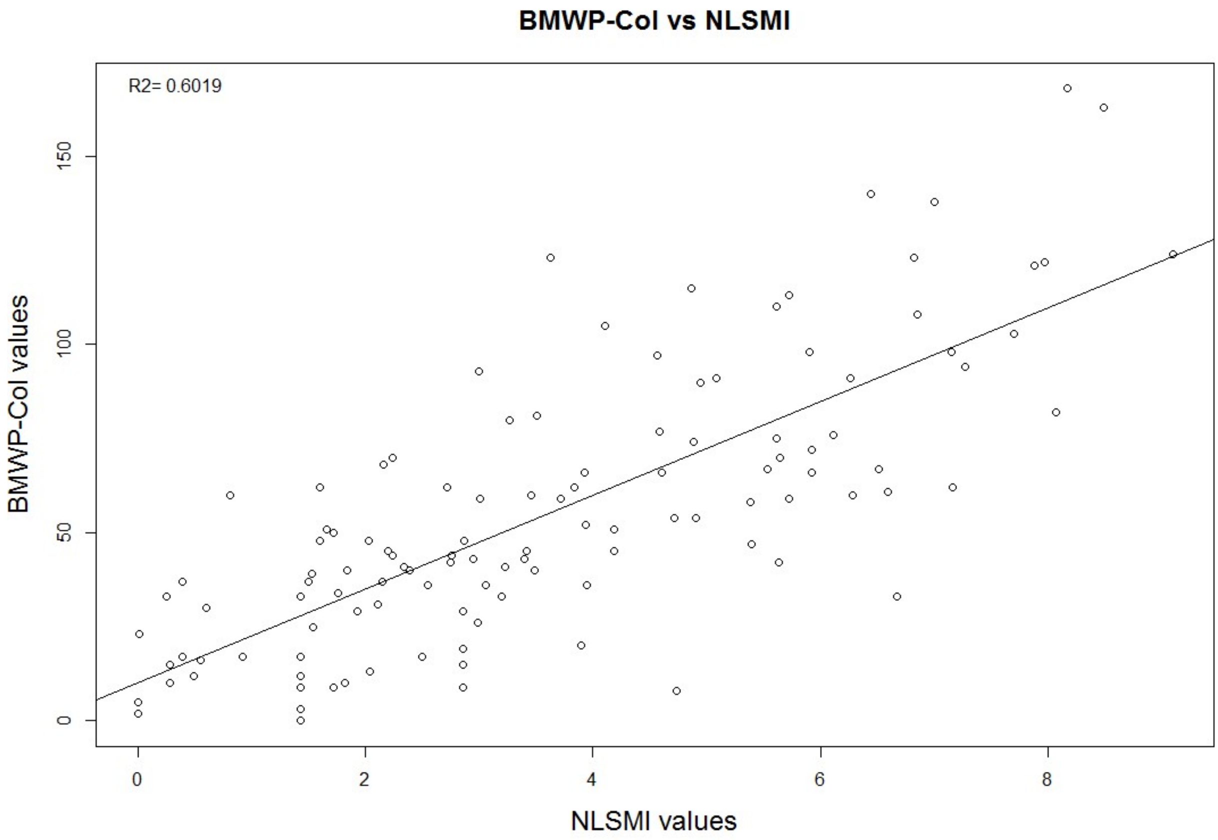


Figure 4.4 Plot correlation between BMWP-Col and NLSMI for 120 sampling sites.

The oxygen Prati index ranged from 0.0 to 7.3, while the habitat disturbance score ranged from 11 to 26. The box plots of the seven individual metrics to compute the NLSMI based on the 120 sites displayed broad ranges of values between impaired and reference sites for the number of EPT taxa, the Margalef index and the % of Trichoptera. This was not the case for the Shannon-Wiener Evenness index, the ratio of Chironomidae/Diptera, the % of scrapers and the % of shredders (Fig. C7-C13). The remaining results for the three separate calculations (rivers lower than 250 m, rivers higher than 250 m and reservoir) are presented in Supporting Information (Fig. C7-C26, Table C4-C6) including the plots of the seven metrics needed to calculate the NLSMI, the relation between the BMWP-Col and NLSMI with each environmental variable, and the correlation between the BMWP-Col and NLSMI.

4.3.2 ASPT calculation, habitat disturbance and functional feeding group

The ecological water quality based on the ASPT ranged from 0 to 7.3. High ASPT values were observed at sites located at higher elevations having forested land use and mountainous areas. High values were also observed at tributaries of the rivers located at lower elevations (Figures C26 and C27). Generally, high ASPT values were observed at sites with a low concentration of chlorophyll a, nitrate-N and nitrite-N. A 90% of shading, a sludge layer of less than 5 cm and the presence of dead wood in the rivers were related to a high ASPT (Figure C28). Hence, ASPT indicated similar environmental conditions to those of the BMWP-Col, as can be seen from their positive correlation (Figure C29). According to the ASPT classification, poor scores indicate the effect of pollution. However, the data showed that poor ASPT scores might have been caused by habitat alteration as well. Since the coefficient of determination (R^2) between ASPT and elevation was low (0.22, Fig. C27), it was concluded that elevation did not influence the calculation of the BMWP-Col. Therefore, this also confirmed that further analyses could be done using the BMWP-Col values.

The habitat disturbance score ranged from 11 to 26 (Figure C30), where low index scores were found in both undisturbed (indicated by low habitat disturbance scores) and disturbed (indicated by high habitat disturbance scores) habitats. Figures C26 and C27 indicated the effect of elevation (i.e. < or > 250 m) on the ecological water quality. Therefore, the FFG was plotted separately for sites located at elevation higher and lower than 250 m, as well as for the reservoir (Figure 4.5). Collectors were dominant at both higher and lower than a 250-m elevation (mean percentage 60.3% and 40.2%, respectively), while predators and collectors dominated the sites at the reservoir (mean percentage 45% and 33%, respectively).

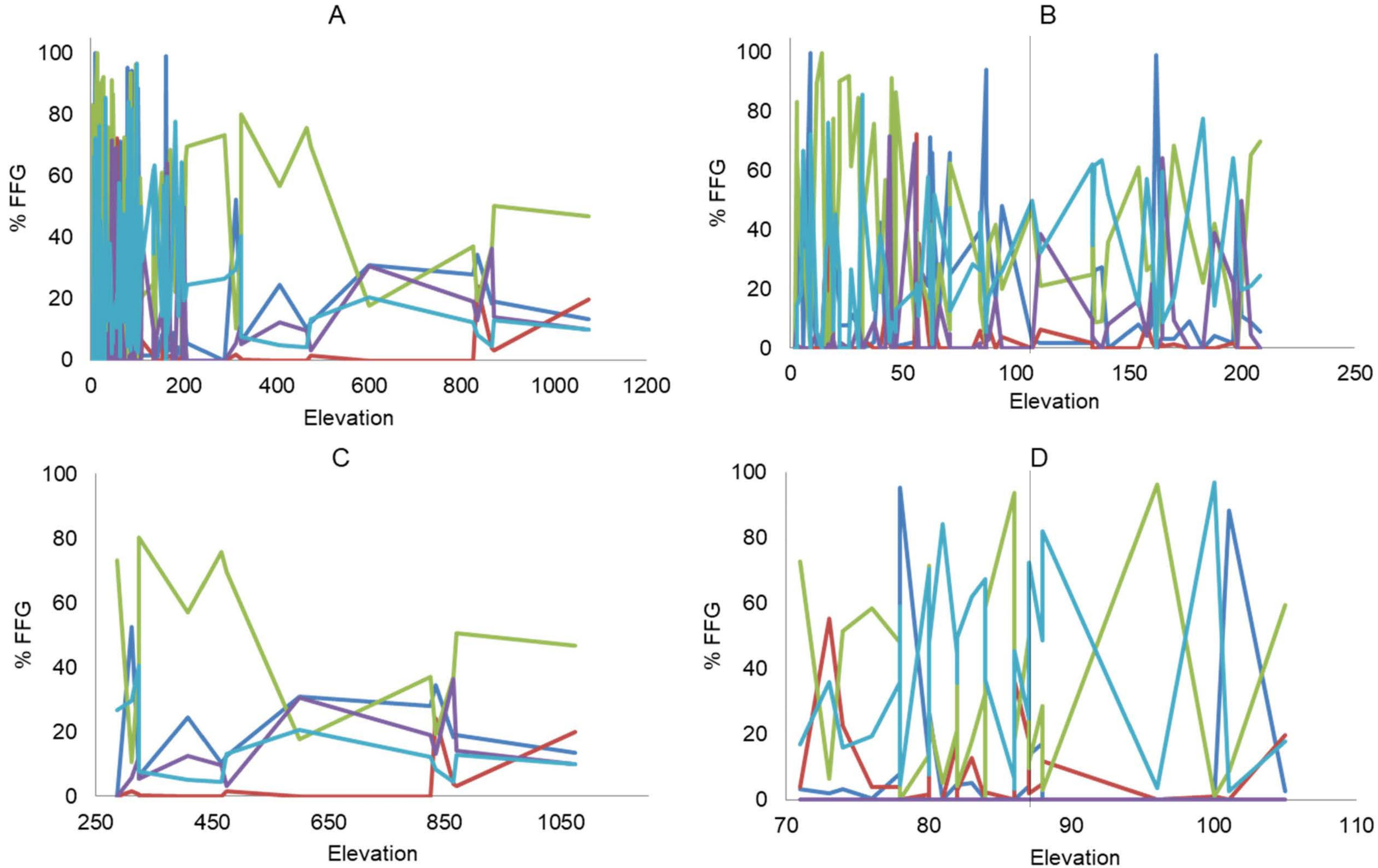


Figure 4.5 Percentage of functional feeding group (FFG) comprises percentage of scrapers ( ), shredders ( ), collector-gatherer ( ), collector-filterer ( ) and predator ( ) encountered at the sampling sites: for 120 sampling sites (A); for sites located at the elevation lower than 250 m (B); for sites located at the elevation higher than 250 m (C) and for sites located at the reservoir (D).

4.3.3 Correspondence analysis

The CA graph (Fig. 4.6) showed that sampling sites with a high flow velocity, a thin sludge layer, a low chlorophyll a concentration, a coarse sediment type and less intensive land use (forest) were separated from other locations along axis one. Most of the sites having a good water quality were located along this axis (on the left). Along axis two, the sampling sites with a high conductivity were separated from other sampling sites. The CA results for both the BMWP-Col and the NLSMI are similar, since the CA plot was based on the composition of macroinvertebrates per sampling site (Fig. 4.7). The difference is on the water quality class of sampling site (Fig. 4.5 for BMWP-Col and Fig. C32 for NLSMI, showing only environmental variables significant at $p < 0.001$ in relation to taxa abundance and distribution in the CA graph).

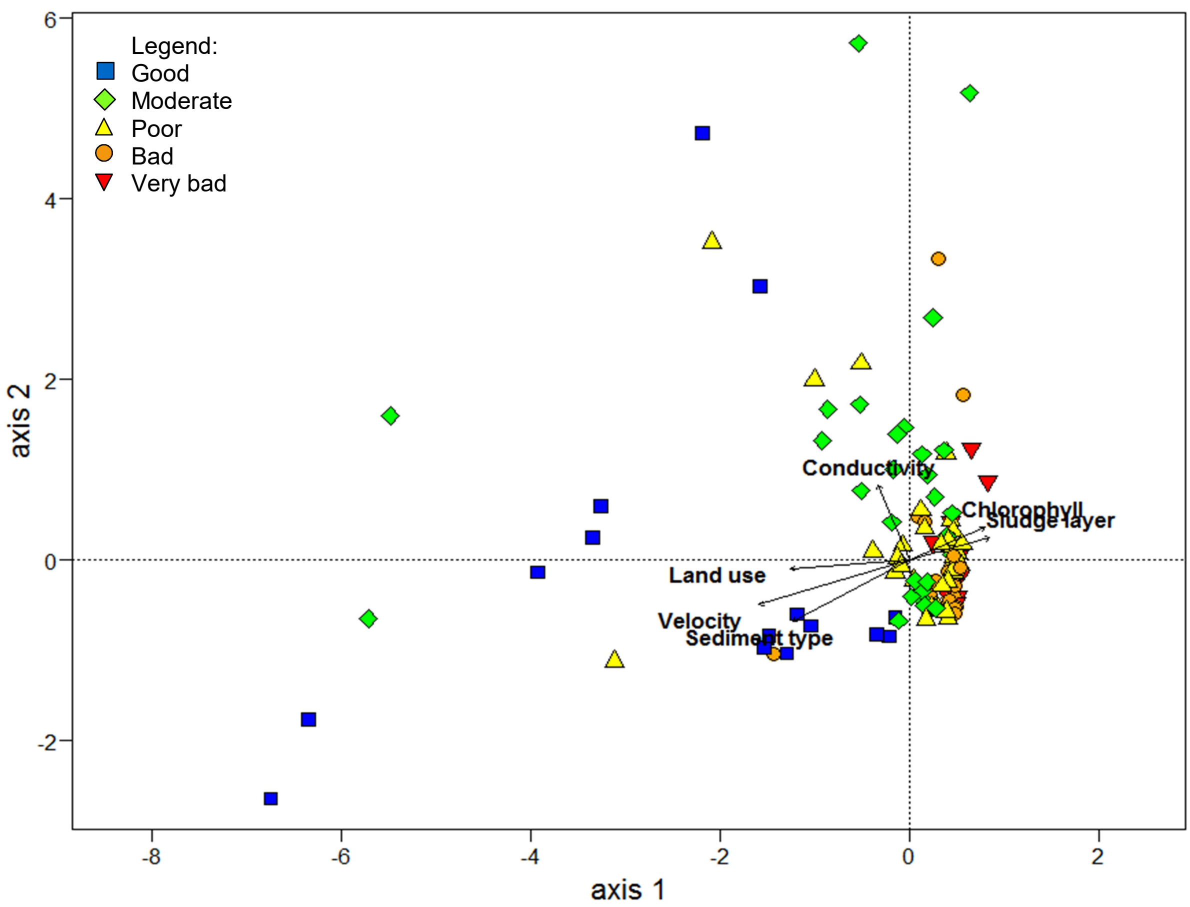


Figure 4.6 Correspondence analysis of taxa abundance (83 taxa) and fitted environmental variables with indication of the ecological water quality of 119 sampling sites expressed as BMWP-Col ranging from good to bad, as shown in the legend.

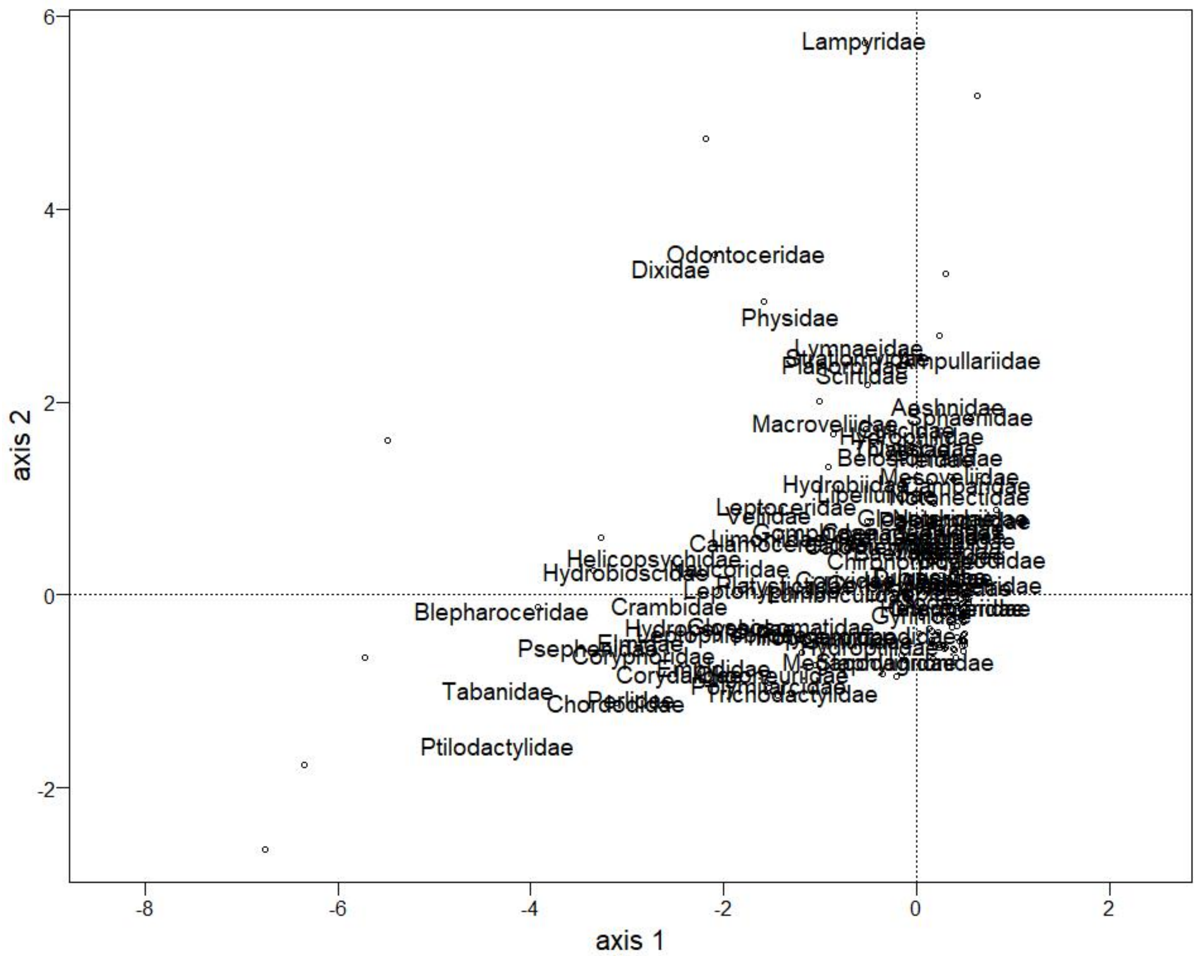


Figure 4.7 Correspondence analysis of taxa count abundance (83 taxa) showing the distribution of macroinvertebrates in 119 sampling sites.

4.4 Discussion

4.4.1 Water quality assessment based on biotic indices

Generally, BMWP-Col values were positively associated with DO concentrations, with flow velocity and with coarse sediment type. A similar situation was observed for the pool-riffle condition: a more pristine pool-riffle pattern was associated with a better water quality. Higher conductivity, chlorophyll a and turbidity were associated with a lower water quality. Deeper water (> 1 m), more abundant macrophytes and a thicker sludge layer were associated with a lower water quality.

Based on the BMWP-Col values, sampling sites were classified ranging from very bad to good, while the NLSMI classified the water quality from bad to reference (Helson and Williams, 2013). Sampling sites located at the upstream sites with less human influence generally had a better water quality (67-168 for BMWP-Col and 6-

9.1 for NLSMI) compared to sampling sites located at the downstream locations where the anthropogenic influence was high due to for example wastewater discharges. As expected, sites located in the forest had a good water quality (> 100 for BMWP-Col and > 8 for NLSMI), while sites located around residential areas had a bad or very bad water quality. In general, the results of both indices at the sites located at the elevation lower than 250 m follow similar behavior as the results of the whole sampling sites (120 sites). The patterns observed at sampling locations above 250 m differed considerably from those at lower elevations. The sites located at the elevation higher than 250 m (Fig. C18-C19 and Table C4-C6) showed forest as the dominant land use and had a higher mean flow velocity when compared to the sites located at the elevation lower than 250 m and at the reservoir (0.7, 0.3 and 0.0 m/s, respectively). The high flow velocity enabled the transport of fine sediments from the upstream to the downstream locations, thus the absent of fine sediments (Beisel *et al.*, 1998; Younes-Baraille *et al.*, 2005). At the reservoir, arable and pasture were the dominant land use, which could explain the presence of silt-clay as the dominant type of sediment. Here, pool-riffle pattern was absent and flow velocity was zero. Moreover, a bad water quality was observed at the main channels of both the Daule and Babahoyo rivers. The observations are in line with an earlier study performed in the Chaguana river basin, situated in the southwest of Ecuador (Dominguez-Granda *et al.*, 2011a).

The high diversity (in total 83 taxa, Shannon-Wiener Evenness index max = 1, mean = 0.52 and median = 0.56, Fig. C9) and the presence of sensitive taxa are an indication of a good water quality. However, the relationship between the BMWP-Col and the number of taxa or the tolerance score suggested that the number of taxa plays a bigger role in determining the water quality compared to the tolerance score (Fig. C2). This could be expected since the samples contained more sensitive than tolerant taxa, so the higher the number of the taxa encountered in the samples, the higher the BMWP-Col value.

When the NLSMI was calculated based on the whole data set, it gave some unexpected results as the NLSMI either over-estimated or under-estimated the water quality of several sites. Usually at the sites located at an upstream location less intensive human activities exist and a good water quality was found (i.e. high diversity and presences of sensitive taxa, low score of oxygen Prati index and less disturbed habitat) (Prati *et al.*, 1971; Wang *et al.*, 2013). In this study, however, a few sites were considered reversely. The over- and under-estimation were mainly occurring for sites located at higher elevations, at the reservoir and at large rivers. Because of these over- and under-estimated results, the NLSMI was recalculated separately based on different groups. However, the same unexpected results were still observed. The unexpected results were also obvious when comparing the NLSMI and BMWP-Col. For example for site 20 (located at the downstream of the main channel of the Daule river, at the city of Guayaquil, few encountered taxa, oxygen Prati score of 1.1 and habitat disturbance score of 24), the BMWP-Col suggested a very bad water quality, while the NLSMI calculated a moderate water quality. For site 34 (located at the tributary of the Babahoyo river, at a mountainous area, a diverse encountered taxa, oxygen Prati score of 0.25 and habitat disturbance score of 11), the BMWP-Col suggested a good water quality, while the NLSMI calculated a poor water quality. These results illustrate that the NLSMI is a river type specific index and performs satisfactory only when applied to assess the water quality of small rivers located at an elevation lower than 250 m above sea level (Helson and Williams, 2013), but does not perform well for other river types. Therefore, the BMWP-Col is considered more suitable to assess the ecological water quality of the Guayas river basin than the NLSMI. The ASPT also indicated similar environmental conditions to those of the BMWP-Col. Moreover, the correlation between the ASPT and elevation showed that elevation did not influence the calculation of the BMWP-Col. These results gave extra confirmation that further analyses in the study could be done using the BMWP-Col.

4.4.2 Observed macroinvertebrates and their relation with environmental variables

When investigating the CA result, a good water quality was associated with a high flow velocity, a coarse sediment type, less intensive land use (forest) and a low conductivity. This is the typical condition found in mountainous areas, and indicated more natural influence on taxa distribution and presence. Opposite conditions (e.g. thicker sludge layer, higher chlorophyll a concentration, higher conductivity and finer sediment) indicated lower water quality, which also indicated anthropogenic influence (e.g. agriculture and residential). Note that the detection limits of nutrients were relatively high. Hence, based on our analyses it cannot be excluded that nutrients in combination with land use exert an effect on macroinvertebrate communities. In the samples, several sensitive taxa were associated with these specific more natural environmental conditions (Fig. 4.7), such as Ptilodactylidae, Blepharoceridae and Perlidae (all with tolerance score 10). Whereas tolerant taxa such as Chironomidae and Ceratopogonidae (tolerance scores 2 and 3, respectively) were present in both good and bad water quality. The tolerant taxa did not show strong association with specific environmental condition, as shown by sensitive taxa. However, the identification of macroinvertebrates in this study was only done to family level and did not take into account the sensitivity differences among for example Chironomidae. These results agree with the comparative study for three tropical countries (Ecuador, Ethiopia and Vietnam) performed by Everaert *et al.* (2014), who found that sensitive taxa such as Leptophlebiidae (tolerance score 9) prefer a high flow velocity and a low conductivity, while Chironomidae occurred at a wide range of physico-chemical conditions (Everaert *et al.*, 2014). Helson and Williams (2013) also concluded in their study in Panama that sites surrounded by a forest and characterized by a coarse sediment type and a low conductivity had a higher ecological water quality.

Related to the FFG (Fig. 4.5), collectors dominated the rivers located at the elevation lower than 250 m, followed by predators and scrapers (mean percentage 40.2 %, 31.6 % and 22.1 %, respectively). Collectors also dominated the rivers located at the elevation higher than 250 m, followed by scrapers and predators (mean percentage 60.3 %, 19.8 % and 15.6 %, respectively). At the reservoir, predators were dominant and followed by collectors, where shredders and scrapers were relatively equal (mean percentage 45 %, 33 %, 12 % and 10 %, respectively). This situation was not totally in accordance with the river continuum concept that describes the dominancy of shredders at the upstream locations, and the dominancy of collectors at the downstream locations together with the scrapers (Vannote *et al.*, 1980). Besides their presence at the reservoir, shredders were present for mean percentage of 4.2 % at the elevation higher than 250 m and 4.6 % at the elevation lower than 250 m. It was expected that the surrounding land use and the type of sediments influenced the presence of certain FFG at the sampling sites (Compin and Cereghino, 2007; Grubaugh *et al.*, 1996; Rios and Bailey, 2006; Strayer *et al.*, 2003).

Since the sampling campaign was performed at the end of dry season, several environmental variables seemed to have reached their extreme values (e.g. conductivity, maximum value was 1981 µS/cm) and created harsh conditions for aquatic lives. As discussed by Blanchette and Pearson (2013), Garcia *et al.* (2015) and Helson and Williams (2013), generally, extreme levels of environmental variables coupled with a habitat shrinkage, an increase in predation and an interruption from upstream assemblages during the dry season resulted in a decline in the abundance and diversity of macroinvertebrates, as compared to the wet season or at the beginning of the dry season. The percentage of the FFG collector-gatherers is generally higher with increasing disturbance, while shredders will decrease with increasing disturbance. Nevertheless, macroinvertebrates' responses towards environmental changes might vary across sites and habitats (Blanchette and Pearson, 2013; Garcia *et al.*, 2015; Helson and Williams, 2013). Dominguez-Granda *et al.* (2011a) observed these different macroinvertebrate responses in Chaguana river basin, where no systematic differences in macroinvertebrates' richness and abundance was found. However, because a similar sampling campaign was not performed during the wet season, the assumption could not be tested.

4.4.3 Importance of nutrient and pesticide measurements for water quality

Much of the measured nutrient concentrations were below the detection limits of the Hach-Lange®DR 3900 spectrophotometer kits. For instance, only in two sampling sites the concentration of total P was above the detection limit. The concentration of total N could only be quantified in 18 sampling sites, nitrate-N in 56 sampling sites and nitrite-N in 13 sampling sites. This was a surprising finding. Since agriculture (including rice, banana and cattle farming (Arias-Hidalgo *et al.*, 2013; Flood, 2000; Seo *et al.*, 2010)) is the main industry in the Guayas river basin, it was expected that nutrient concentrations would be high and thus could be detected by the kit. Nutrient levels in the water can increase due to the use of manure and chemical fertilizers in agricultural areas (Bainbridge *et al.*, 2009; Borbor-Cordova *et al.*, 2006). Moreover, high nutrient concentrations were expected as the sampling activities took place at the end of the dry season when water levels were low and concentrations increased. Borbor-Cordova *et al.* (2006) suggested that some parts of the Guayas river basin have experienced nutrient loss and soil degradation due to their intensive farming activities. They stated that the amount of nutrient that leaves the soil through the exported crops is higher than the original soil content and the applied chemical fertilizers (Borbor-Cordova *et al.*, 2006). Their finding might explain the observed nutrient concentrations in this study. It is recommended that future monitoring campaigns use nutrient measuring kits having lower detection limits than being used in this study (1 mg/L, 0.5 mg/L, 0.23 mg/L, 0.015 mg/L and 0.015 mg/L for total N, total P, nitrate-N, nitrite-N and ammonium-N, respectively), in order to correctly measure nutrient concentrations and determine the influence of nutrients on the ecological water quality of the Guayas river basin. Due to intensive agricultural and residential activities in the Guayas river basin, nutrient concentrations need to be maintained below the guidelines for surface waters (e.g. nitrate-N: 13 mg/L and 1 mg/L for class I to > 11.3 mg/L for class V, nitrite-N: 0.2 mg/L and ≤0.01 mg/L for class I - >0.3 mg/L for class V based on the Ecuador Ministry of Environment (Ministerio del Ambiente del Ecuador - MAE, 2015) and the European Commission (EU, 1998), respectively).

Other studies suggested that there are other variables besides nutrients that affect the water quality. Since the Guayas river basin is one of the major banana and rice producing areas of Ecuador and the fact that farmers use high quantity of different types of agrochemicals such as pesticides (Borbor-Cordova *et al.*, 2006), it is possible that pesticides played an important role in determining the water quality. As stated by FAO (2011), Ecuador is number 14 of the world's largest intensive pesticide users based on the amount of pesticides used per unit of cultivated area. Pesticides can end up in the surface water through runoff and leaching and can be very toxic to aquatic organisms (Kidd and James, 1991). The impacts of pesticides use have been studied for example in Ecuador (Caceres *et al.*, 2002; Horgan *et al.*, 2014) and Costa Rica (Castillo *et al.*, 2000), where the studies concluded that monitoring of pesticides residues is highly important. For the present chapter, however, pesticides were not measured. It is therefore suggested that future studies in this area should incorporate pesticide measurements when assessing the water quality.

Chapter 5: General linear models to identify key hydromorphological and chemical variables determining the occurrence of macroinvertebrates in the Guayas river basin (Ecuador)

Adapted from:

Minar Naomi Damanik Ambarita, Gert Everaert, Marie Anne Eurie Forio, Thi Hanh Tien Nguyen, Koen Lock, Peace Liz Sasha Musonge, Natalija Suhareva, Luis Elvin Dominguez Granda, Elina Bennetsen, Pieter Boets, Peter L.M. Goethals. (2016). Generalized Linear Models to Identify Key Hydromorphological and Chemical Variables Determining the Occurrence of Macroinvertebrates in the Guayas River Basin (Ecuador). Water 8(7). DOI: Artn 297 10.3390/W8070297.

Abstract

The ecological water quality of the Guayas River basin in Ecuador is at risk due to extensive anthropogenic activities. In this chapter, the potential impacts of hydromorphological and chemical variables on ecological water quality using macroinvertebrate-based bioassessments were investigated. The used bioassessment method was the Biological Monitoring Working Party adapted for Colombia (BMWP-Col), via an extensive sampling campaign that was completed throughout the river basin at 120 sampling sites. The BMWP-Col classified the ecological water quality from very bad to good. General linear models (GLMs) and sensitivity analysis were used to relate the ecological water quality to hydromorphological and chemical variables. It was found that elevation, nitrate-N, sediment angularity, logs, presence of macrophytes, flow velocity, turbidity, bank shape, land use and chlorophyll a were the key environmental variables affecting the BMWP-Col. From the analyses, it was observed that the rivers at the upstream higher elevations of the river basin were in better condition compared to lowland systems and that a higher flow velocity was linked to a better BMWP-Col score. Although the results of the models provided insights into the ecosystem, cross fold model development and validation also showed that there was a level of uncertainty in the outcomes. Limitation of nitrate-N measurement might influence model's ability to evaluate thre relationship between the BMWP-Col and environmental variables. However, the results of the models and sensitivity analysis can support water management actions to determine and focus on alterable variables, such as the land use at different elevations, monitoring of nitrate and chlorophyll a concentrations, macrophyte presence, sediment transport and bank stability.

5.1 Introduction

Water quality monitoring involves the measurement of different water quality variables, including physical and chemical conditions, sediment and the biological composition of an aquatic system. Monitoring allows managers to maintain a good water quality by enabling them to make necessary decisions and to take actions prior to ecosystem degradation. As it is more sustainable to keep a clean environment compared to restoring a polluted one (Goethals, 2013), monitoring thus plays a crucial role in water quality management.

Agriculture, urban settlements, irrigation and industries are examples of anthropogenic threats that may change the ecological water quality (Arimoro *et al.*, 2015; Helson and Williams, 2013). Generally, agricultural land use and hydromorphological alteration negatively affect taxa richness and the ecological quality of aquatic communities. Agriculture can alter rivers and riparian integrity, habitat quality and bank stability. Anthropogenic alteration of flow regimes, such as dam constructions, can affect aquatic organisms since they cannot tolerate rapid changes in flow (Bruno *et al.*, 2014). Agricultural areas often result in elevated nutrient concentrations in rivers (Bruno *et al.*, 2014; Frankforter *et al.*, 2010), which can increase the biomass of algae. This condition will consequently cause a decrease of oxygen levels in the water and alter the habitat of aquatic organisms (Frankforter *et al.*, 2010). Moreover, disturbed areas also show higher nutrient transport in the rivers compared to forested watersheds (Silva *et al.*, 2012).

As described by Karr (1991), biotic integrity is the ability of an ecosystem to support and maintain community composition in relation to the environmental conditions of a region. Biomonitoring using benthic macroinvertebrates has been effectively used to assess water quality conditions in rivers, in addition to the hydromorphological condition altered by poor land use practices in watersheds. Thus, bioassessments are good means to define the ecological water quality status of an aquatic ecosystem. Arimoro *et al.* (2015) found that biological oxygen demand and the concentrations of nutrients were important variables to define the macroinvertebrate’s structure of the Ogba River (Nigeria), a river that receives discharges of wastewater from housing and farming. Blanchette and Pearson (2013) reported the influence of riparian vegetation, substratum type, depth and flow velocity on macroinvertebrate’s assemblages in Burdekin catchment (Australia), where mainly agriculture takes place. In the Mediterranean lowland Odelouca River (Portugal), Hughes *et al.* (2009) found that land use and flow velocity had an impact on the structure and functioning of macroinvertebrates due to surrounding agricultural activities. Depending on the region and watershed, studies have found different key variables explaining the structure and functioning of the macroinvertebrate community.

To date, limited information is available on the bioassessments and water quality of river basins in the tropics (Everaert *et al.*, 2014), such as South America, where biodiversity is rich, but threatened by anthropogenic influences (Dudgeon *et al.*, 2006). Previous studies in the Guayas River basin using the BMWP-Col index were only performed in one wetland area, where flow velocity and sediment type influenced taxa distribution, abundance, richness and diversity (Alvarez-Mieles *et al.*, 2013). The study of the Intag cloud forest region in northwestern Ecuador also used the BMWP-Col index; however, no relation between environmental variables and macroinvertebrates was identified (Knee and Encalada, 2014). Other studies used macroinvertebrate richness and composition to define temporal and spatial changes (Blanchette and Pearson, 2013; Rezende *et al.*, 2014). A biological index for the region is still lacking, despite several new indices that have been developed to better study the water quality, such as the Índice Multimétrico del Estado Ecológico para Ríos Altoandinos (IMEERA) (Villamarin *et al.*, 2013), the Andean Biotic Index (ABI) (Rios-Touma *et al.*, 2011) and the Neotropical Low-land Stream Multimetric Index (NLSMI) (Helson and Williams, 2013).

Moreover, water quality studies in the tropics, especially in South America, are still lacking; thus, the relationship between macroinvertebrate communities and habitat disturbance is poorly understood in these regions (Rios-Touma *et al.*, 2011). Consequently, it is difficult for decision makers to determine how to invest limited financial resources to improve the water quality. Fortunately, previous studies have shown the benefits of using ecological models in studying the water quality (Arias-Hidalgo *et al.*, 2013; Everaert *et al.*, 2013; Forio *et al.*, 2015; Hoang *et al.*, 2010), despite the challenge in selecting the variables to be included in the model due to the considerable impacts that multiple variables have on water quality (Everaert *et al.*, 2013). Hence, modelling can be a helpful means to support management actions by identifying the key variables that need to be monitored.

In this chapter, the importance of environmental conditions on the ecological water quality of the Guayas River basin in Ecuador, based on macroinvertebrates was investigated. The Guayas River basin is an important watershed in Ecuador (Arriaga, 1989), and its ecological water quality is at risk due to extensive agriculture and industrial activities in the area (Nguyen *et al.*, 2015). GLMs were used to determine the key environmental variables influencing the ecological water quality. Furthermore, a sensitivity analysis was performed to propose potential restoration or maintenance actions of the tropical river basins' management, as well as for other river basins with similar environmental conditions.

5.2 Data analysis

The BMWP-Col (Roldán Pérez, 2003) based on Alvarez (2005) was used to calculate the ecological water quality of the 120 sampling sites, since it was considered more suitable in determining the ecological water quality of the Guayas river basin based on the results in Chapter 4.

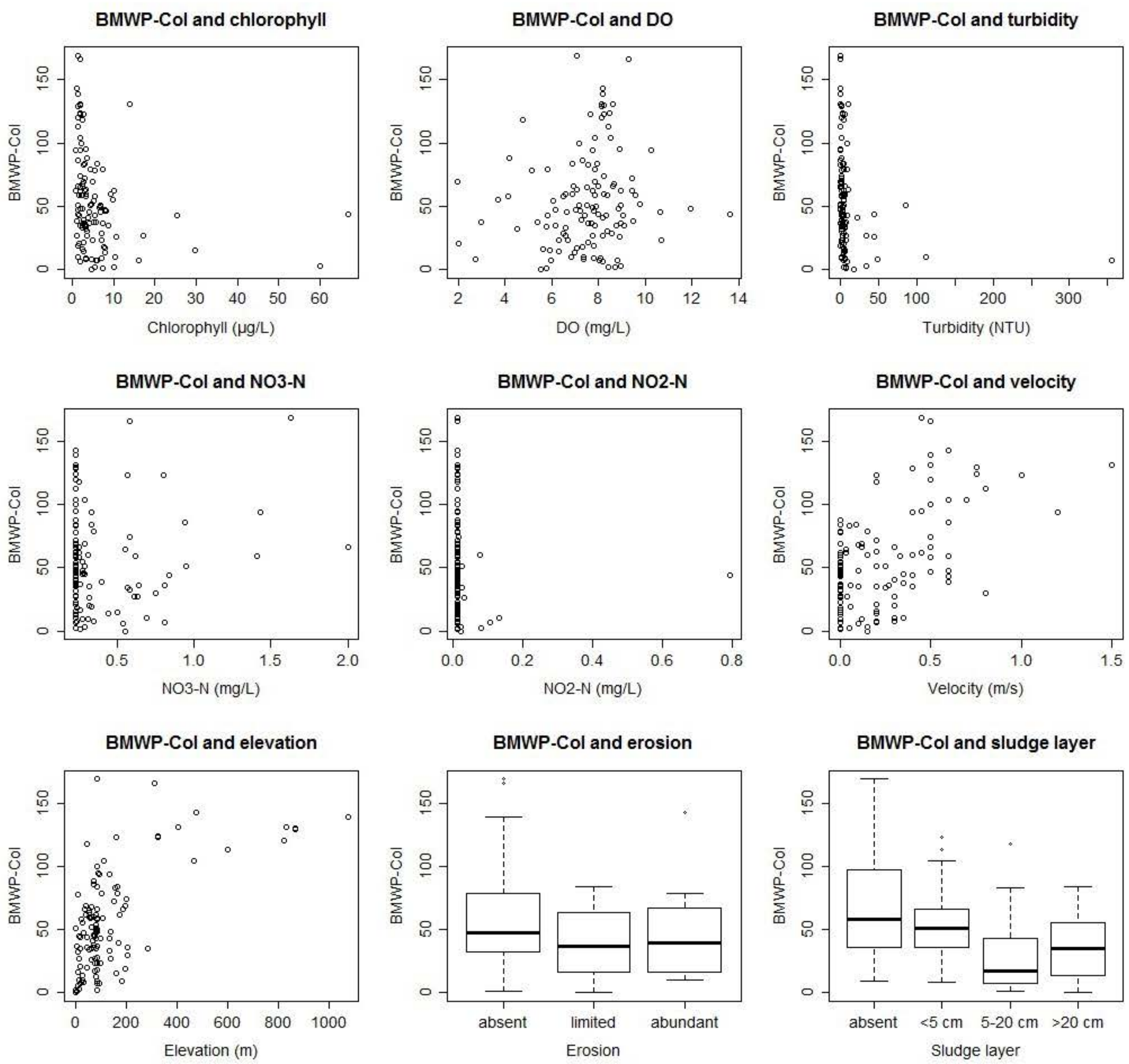
Models were developed to identify key environmental variables influencing the presence of macroinvertebrates in the Guayas river basin, Ecuador, following a scheme shown in Fig. D1. For this chapter, the preprocessed data (Table 4.1) and FP land use (for land use information) were used for analyses. To start the analyses, in total 39 variables were used (Tables 4.1 preprocessed data and B1). However, chemical oxygen demand (COD), stream width and stream depth were removed before the analysis due to missing values. Furthermore, following the procedure described in Zuur (2009) and Zuur *et al.* (2007), 12 collinear variables were removed based on variance inflation factors (VIF), where variables with VIF values higher than three were regarded as collinear. Based on these pre-processing steps, 24 variables were included in the analysis (Tables 4.1 preprocessed data and D1). Next, a general linear model (GLM) was used to determine key environmental variables influencing the ecological water quality (Weirich *et al.*, 2011; Zuur, 2009), expressed as BMWP-Col. Three-fold cross validation (CV) was used to train and validate the GLMs (more detailed explanation on GLM and the use of three-fold CV is given in chapter 3). To assess the robustness of the three-fold CV, the models developed based on 2/3 of the data were compared with a model that was developed based on the complete dataset. Hence, two sets of models were inferred: model developed from and validated on the complete dataset (120 sites) and models developed from and validated on three-fold CV (Fig. D1).

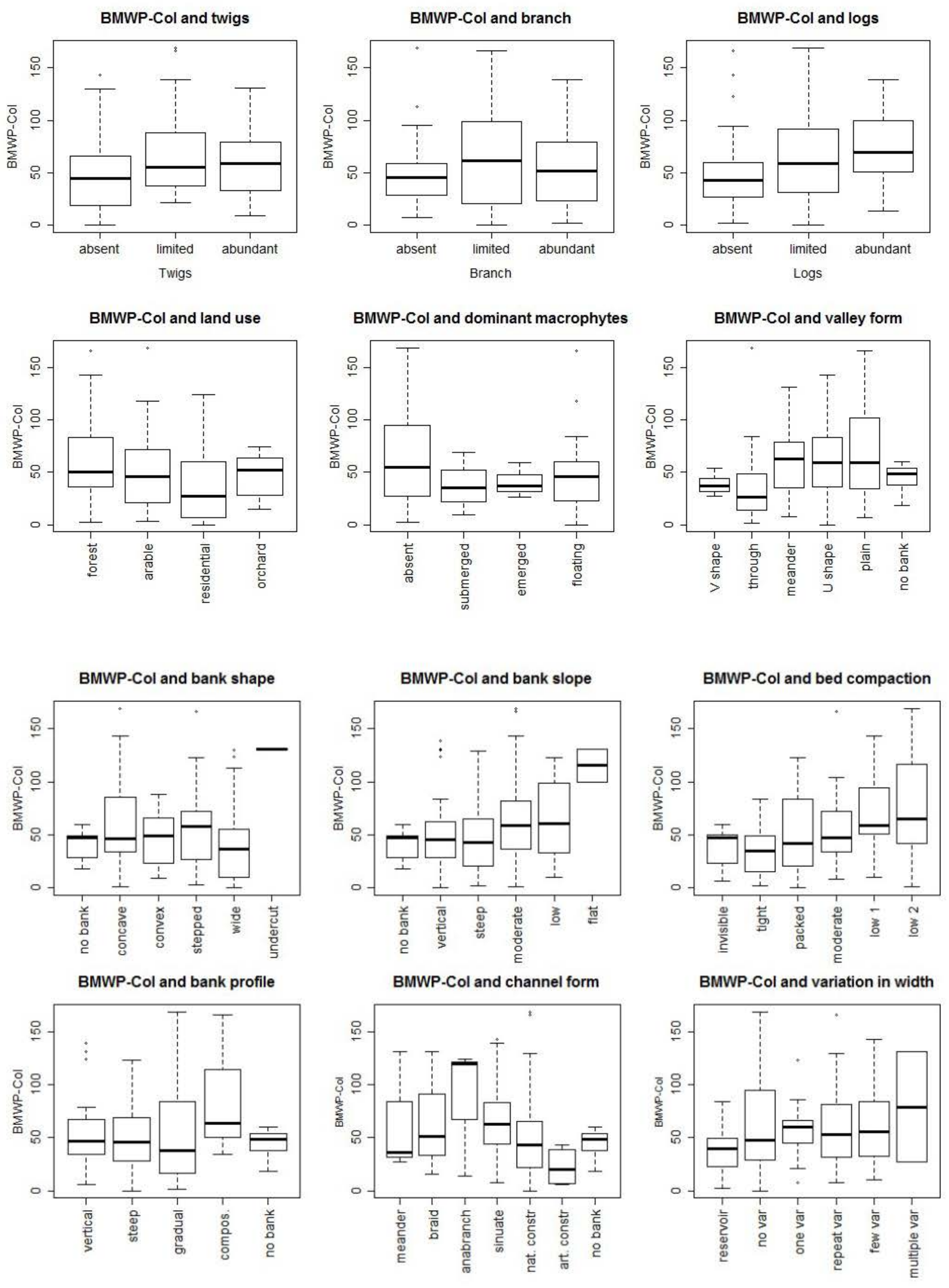
The best model was selected having the lowest AIC. However, models with the lowest AICs did not always contain all variables with *p*-values significant at $p < 0.05$. To address the situation, variable removal using the drop1 command was continued until the models with all variables significant at $p < 0.1$ and $p < 0.05$, respectively, were reached. Two *p*-value criteria were used to see the significant difference among variables contained in models from different partitions. The stability of the results of the models was evaluated by ranking the input variables based on their presence in each model. To do this, each variable was listed according to its significance in the model (based on its *p*-value). The variable lists from all models were then combined to get the final ranks of the variables. Lastly, sensitivity analyses were performed to assess the effect of selected key environmental variables on the BMWP-Col.

5.3 Results

5.3.1 Biotic index and ecological water quality

The ecological water quality based on the BMWP-Col ranged from 0 to 168. High values of the BMWP-Col were observed at sites located at higher elevations having forested land use and mountainous areas. High BMWP-Col values were also observed at tributaries of the rivers located at lower elevations (Figures D2 and D3). Generally, high BMWP-Col values were observed at sites with a low concentration of chlorophyll a, nitrate-N and nitrite-N. High BMWP-Col values were also witnessed at sites where DO concentrations ranged from 6 to 10 mg/L, turbidity was lower than 20 NTU and flow velocity was higher than or equal to 0.2 m/s. A 90% of shading, a sludge layer of less than 5 cm and the presence of dead wood in the rivers were also related to a high BMWP-Col (Figures 5.1).


BMWP-Col and chlorophyll
BMWP-Col
Chlorophyll (µg/L)
BMWP-Col and DO
DO (mg/L)
BMWP-Col and turbidity
Turbidity (NTU)
BMWP-Col and NO3-N
NO3-N (mg/L)
BMWP-Col and NO2-N
NO2-N (mg/L)
BMWP-Col and velocity
Velocity (m/s)
BMWP-Col and elevation
Elevation (m)
BMWP-Col and erosion
absent
limited
abundant
Erosion
BMWP-Col and sludge layer
absent
<5 cm
5-20 cm
>20 cm
Sludge layer


BMWP-Col and twigs
BMWP-Col
absent
limited
abundant
Twigs
BMWP-Col and branch
Branch
BMWP-Col and logs
Logs
BMWP-Col and land use
forest
arable
residential
orchard
BMWP-Col and dominant macrophytes
absent
submerged
emerged
floating
BMWP-Col and valley form
V shape
through
meander
U shape
plain
no bank
BMWP-Col and bank shape
no bank
concave
convex
stepped
wide
undercut
BMWP-Col and bank slope
no bank
vertical
steep
moderate
low
flat
BMWP-Col and bed compaction
invisible
tight
packed
moderate
low 1
low 2
BMWP-Col and bank profile
vertical
steep
gradual
compos.
no bank
BMWP-Col and channel form
meander
braid
anabranch
sinuate
nat. constr
art. constr
no bank
BMWP-Col and variation in width
reservoir
no var
one var
repeat var
few var
multiple var

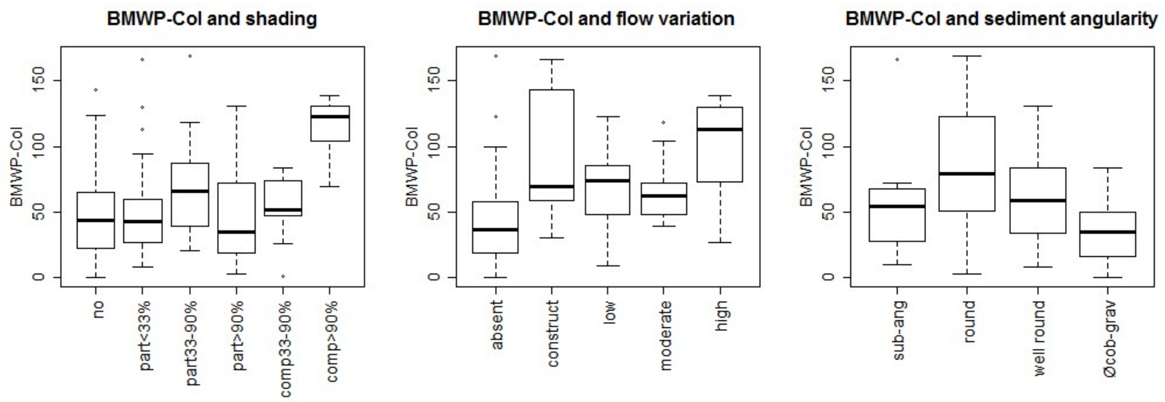


Figure 5.1 Plots showing the distribution of the data for physico-chemical variables in relation to BMWP-Col for 120 sampling sites, classification of categorical variables are based on Table B1; compos: composite, nat: natural, art: artificial, constr: construction, var: variation, part: partly, comp: completely, ang: angular, cob-grav: cobble-pebble-gravel.

5.3.2 Statistical model

Physico-hydromorphological (i.e. elevation, sediment angularity, logs, main macrophytes, flow velocity, turbidity, bank shape and land use) and chemical (i.e. nitrate-N and chlorophyll a concentrations) variables were selected as the key drivers of the ecological water quality expressed as the BMWP-Col, out of the three final models. In total, 16 variables were contained in the three-fold cross-validation models (i.e. nitrate-N, chlorophyll a, turbidity, flow velocity, elevation, sediment angularity, valley form, twigs, branches, logs, land use, bank slope, bank shape, main macrophytes, erosion and variation in flow). However, different data partitions from the three-fold cross-validation resulted in varied selected variables and significant levels.

Elevation was the most significant variable, while nitrate-N was the only nutrient variable that came up in each criterion. For Training Set 1 + 2, 11 variables were selected based on the model with the lowest AIC: elevation, main macrophytes, nitrate-N, sediment angularity, logs, land use, erosion, chlorophyll a, flow variation, velocity and bank slope, with *p*-values of 0.001, 0.013, 0.024, 0.027, 0.044, 0.048, 0.048, 0.064, 0.067, 0.151 and 0.181, respectively. The variables' selection is presented in Table D1, while the final models are shown in Table D2 together with their ranks. Fold 1 (Training Set 1 + 2 and Testing Set 3) had the highest R^2 value for testing set compared to other folds. The R^2 values were 0.57 and 0.49 for Training Set 1 + 2 and Testing Set 3, respectively. Compared to other criteria, the model with the lowest AIC gave the highest R^2 value (Table D3). The results of other data partitions are presented in the Supporting Information (Tables D1–D3). Residual plots and model validation are presented in the Supporting Information (Figures D2–D17). For the model based on the complete dataset, 10 variables were selected that corroborated the results of the three-fold cross-validation (Tables D4 and D5 and Figures D18–D26).

5.3.3 Sensitivity analysis

Besides elevation, all other variables in the models were also investigated in the sensitivity analysis to assess their effects on the BMWP-Col values (Table D6). Here, the impacts of different elevations and nitrate-N concentrations based on the models from all folds (Figures 5.2 and 5.3) are presented. The sensitivity analysis of elevation clearly showed that the BMWP-Col increased from 35 (bad) to 122 (good) if the elevation ranged from 2 to 1080 m. The data showed that a nitrate-N concentration higher than 0.6 mg/L was associated with a poor ecological water quality, whereas the sensitivity analysis suggested an improvement in ecological water quality from 39 (poor) to 118 (good) for nitrate-N concentrations between 0 and 2.1 mg/L. Due to this finding, the relationship between the nitrate-N and other variables that might be related to nitrate-N, i.e., chlorophyll a and dominant macrophytes was further checked. Several sites with nitrate-N concentrations higher than 0.5 mg/L were found where chlorophyll a concentrations were lower than 10 µg/L. Nitrate-N concentrations higher than 0.5 mg/L were also detected at sites where macrophytes were absent or where floating macrophytes were present (Figure D27).

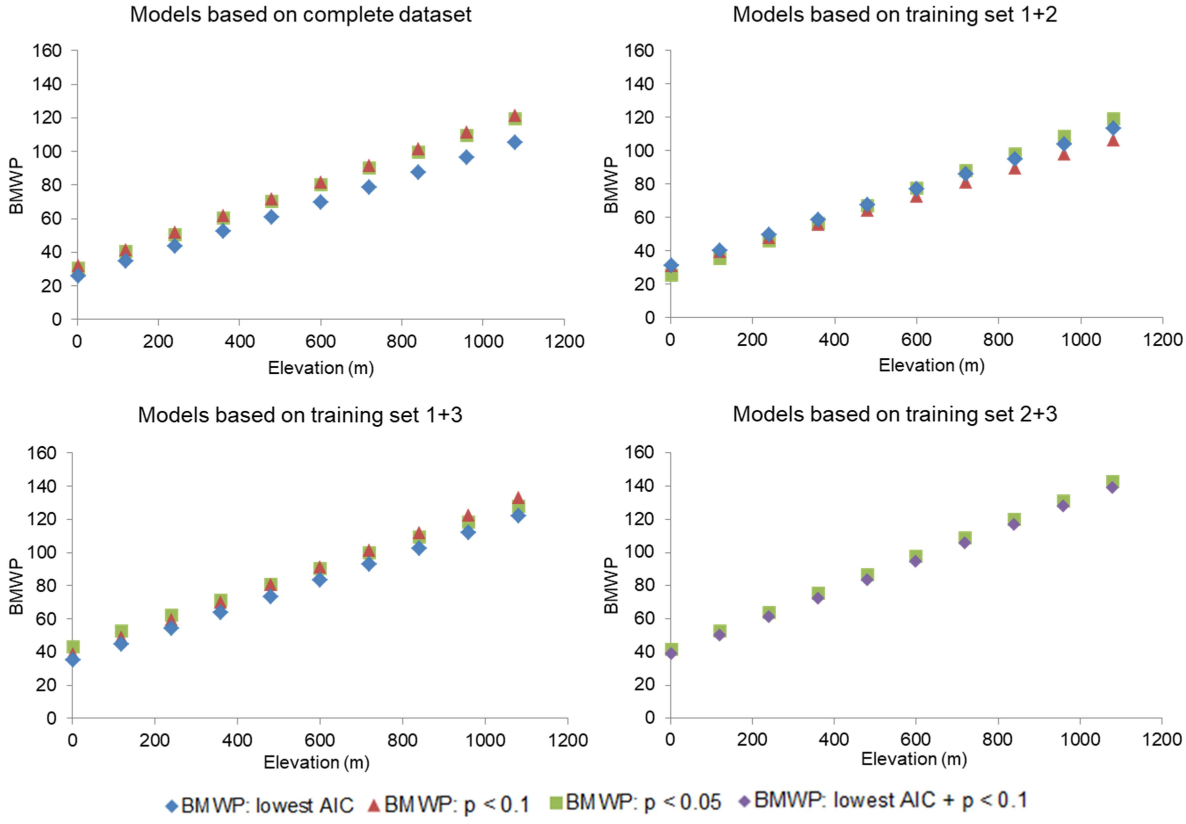


Figure 5.2 The impact of different elevations on the ecological water quality expressed as BMWP-Col for models from different folds. The values used in the analysis were based on Table D6.

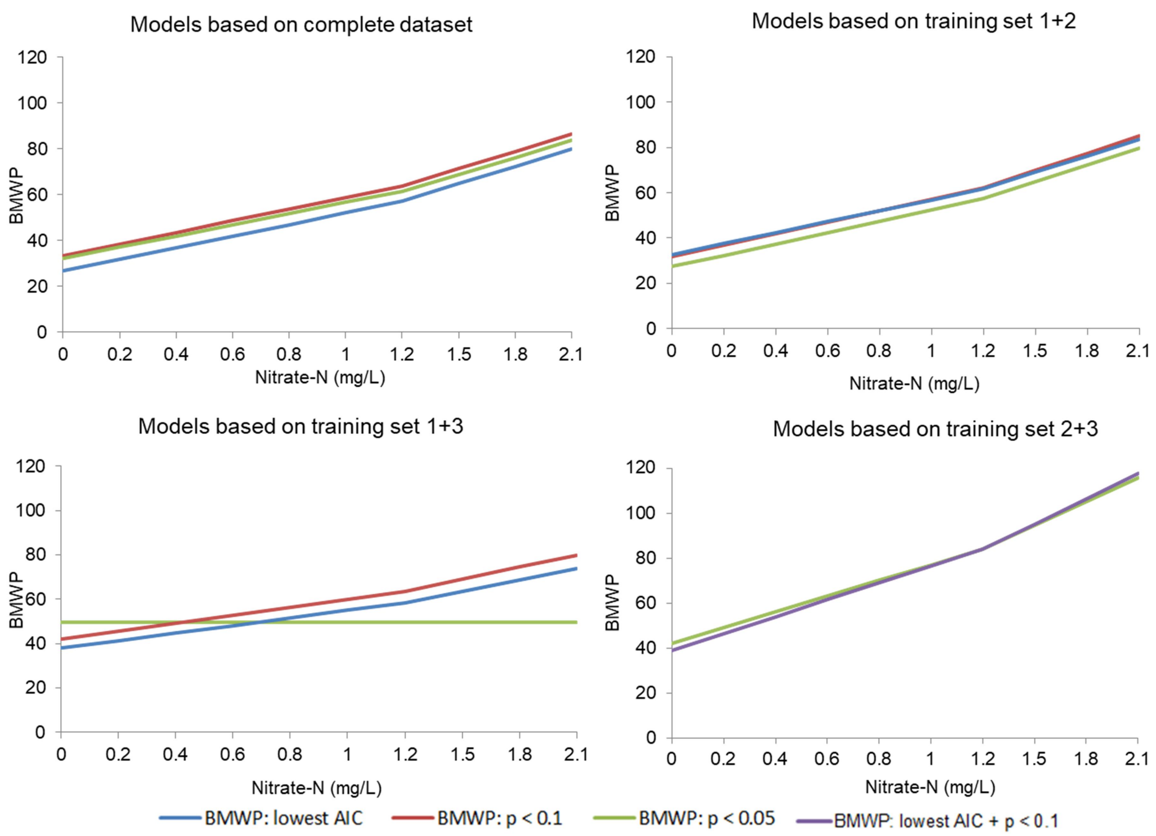


Figure 5.3 The impact of different nitrate-N concentrations on the ecological water quality expressed as BMWP-Col for models from different folds. The values used in the analysis were based on Table D6.

The figures of other variables are given in the Supporting Information (Figures D28-D30), namely flow velocity, sediment angularity and chlorophyll a. The sensitivity analysis of different flow velocity from 0 to 1.5 m/s showed that the BMWP-Col will increase from 34 (bad) to 88 (moderate). More angular sediment (sub-angular and round types) could promote the ecological water quality and a decrease in chlorophyll a concentration will promote the ecological water quality.

5.4 Discussion

5.4.1 Ecological water quality and potential restoration actions

Elevation, nitrate-N concentration, sediment angularity, logs, main macrophytes, flow velocity, turbidity, bank shape, land use and chlorophyll a concentration were the major variables that influenced the ecological water quality expressed as the BMWP-Col in the Guayas River basin. For management purposes, ensuring proper land use at different altitudes and monitoring the concentrations of nutrients that enter the surface waters can address most of the aforementioned variables.

5.4.1.1 Elevation

Elevation was present in all models and, thus, is an important variable explaining the observed ecological water quality of the river basin. The importance of elevation in determining the water quality has often been reported (Malmqvist and Maki, 1994; Rezende *et al.*, 2014; Younes-Baraille *et al.*, 2005). However, its impacts often depend on several physico-chemical variables that are correlated with the altitude, such as temperature and oxygen levels, the type of substrates (coarser sediment is present more at a higher elevation), flow velocity and the level of disturbance related to land use and waste water discharges, due to less intensive human activities at higher elevation. For example, Malmqvist and Maki (1994) related the importance of elevation with temperature, while Rezende *et al.* (2014) linked elevation with the richness and density of macroinvertebrates. Younes-Baraille *et al.* (2005) found a correlation between the elevation and more intensive human activities along the Andorran rivers. Intensification of human settlements at the lower elevation in Andorra increases the organic and nutrients load into the water that consequently decreases the water quality (Younes-Baraille *et al.*, 2005).

The elevation also influences the presence of macroinvertebrates. The river continuum concept (RCC) suggested that upstream rivers are generally characterized by the presence of shredders due to the rich presence of coarse particulate organic matter (CPOM) in the water, while downstream rivers are generally characterized by collectors that take advantage of fine particulate organic matter (FPOM) (Vannote *et al.*, 1980). The data showed the dominance of collectors at higher and lower elevations and low presence of shredders at higher elevations (see chapter 4 and Fig. C31), as opposed to the RCC for higher elevations

(Damanik-Ambarita *et al.*, 2016b). The RCC was observed at the reservoir, where predators and collectors were dominant. The land use surrounding the sites and the type of sediments might influence the presence of FFG (Compin and Cereghino, 2007; Grubaugh *et al.*, 1996; Rios and Bailey, 2006; Strayer *et al.*, 2003), and the dry season might not provide enough CPOM upstream for the shredders to survive. The increased temperature during the dry season might also negatively influence certain taxa (Vannote *et al.*, 1980).

5.4.1.2 Land use

Although one cannot alter elevation, the land use can be managed adequately at different altitudes. Previous studies have shown the impacts of land use on the water quality, in relation to the elevation. Different land uses at different altitudes are present in the study area, which means that different management actions are needed. At higher elevations, preserving forest in mountainous areas is necessary to maintain a low conductivity, low temperature, low turbidity, low TDS and high DO concentration of the water (Kasangaki *et al.*, 2008a). Forest also provides food for macroinvertebrates, through its leaf and wood litter (Townsend *et al.*, 1997), prevents nonpoint-source pollutants from entering the streams, and enhances in-stream processing of pollutants (Sweeney *et al.*, 2004). Revegetation of riparian areas can decrease the TDS concentration of the water, and its canopy cover also reduces water temperature (Ellison *et al.*, 2009).

Since agricultural activities are more likely to occur in flatter landscapes (Hutchens *et al.*, 2009), its proper management is needed to preserve water quality. Ellison *et al.* (2009) argued that reducing animal grazing in riparian zones is a necessary management option, especially during the summer/dry season, because grazing animals might degrade river banks, lower the water table, and increase water turbidity. Moreover, proper regulation and management of agrochemical use are crucial to reduce the impacts on water quality and macroinvertebrates (Hutchens *et al.*, 2009). Other options to improve the water quality are providing more sanitary infrastructures (Von Sperling and Chernicharo, 2002) and installing a wastewater treatment plant (Younes-Baraille *et al.*, 2005) to treat urban wastewater. Nevertheless, the exclusion of elevation from future studies to analyze environmental impacts on the ecological water quality that is independent of elevation is suggested.

5.4.1.3 Nutrients

The next factor that influenced the ecological water quality was the concentration of nitrate-N in the surface water. Generally, a nitrate-N concentration higher than 5 mg/L in surface waters indicates pollution, and concentrations higher than 0.2 mg/L may stimulate algal growth and indicate eutrophic conditions in lakes (Chapman, 1996). The data confirm this principle, while the sensitivity analysis suggested an improvement in ecological water quality with increasing nitrate-N concentration. Since aquatic plants require nitrogen compounds as their nutrient source (Ballance, 1996), perhaps the results explain this relationship. A previous study by Borbor-Cordova *et al.* (2006) suggested that some parts of the Guayas River basin have experienced nutrient loss and soil degradation due to their intensive farming activities. According to their research, the amount of nutrients that leave the soil through harvested crops is higher than the original soil content plus the applied chemical fertilizers (Borbor-Cordova *et al.*, 2006). Their finding suggests that the Guayas river basin might require an extra amount of nitrate-N for its productivity. However, with regard to general conditions, there is the possibility of a turning point in the sensitivity analysis when the nitrate-N concentration has reached a certain tipping point, which was not studied here.

The presence of nutrients, especially nitrate and phosphate, can also promote the concentration of chlorophyll a in surface waters. High concentrations of chlorophyll a can indicate pollution, in particular eutrophication (Chapman, 1996). However, the data did not show a positive relationship between nitrate-N and chlorophyll a. Garcia *et al.* (2015) suggested that the increase in chlorophyll a concentration is highly influenced by long exposure of the surface water to sunlight and rapid uptake of nutrients by primary producers, thus explaining the high chlorophyll a concentration, but low nitrate-N concentration; whereas a positive relationship between nitrate-N and macrophytes was observed at several sites, especially at sites with the presence of floating macrophytes. Chapman (1996) has discussed the role of nutrients in the development of macrophytes, and Arimoro *et al.* (2015) argued the importance of macrophytes presence in the rivers to provide a suitable microhabitat for certain macroinvertebrates, such as dipterans and odonatas, which was the case in the current study. Thus, macrophyte presence can improve the ecological water quality. Moreover, Nguyen *et al.* (2015) has confirmed a positive correlation between water hyacinth (floating macrophytes) and macroinvertebrate's diversity and the water quality. O'Toole *et al.* (2008) suggested an association between mesotrophic waters and most macroinvertebrate taxa, whereas plecopterans are more associated with oligotrophic and chironomids and tubificids are tolerant with eutrophic waters.

Furthermore, the concentrations of nitrate-N observed in the study were generally lower than the detection limits of the Hach-Lange®DR 3900 spectrophotometer kits. With observed maximum concentration of 2 mg/L, all observed concentrations were below the guidelines for surface waters from the Ministerio del Ambiente del Ecuador (MAE, 13 mg/L) and the European Commission (EC, 1 mg/L for class I to >11.3 mg/l for class V) (EU, 1998; Ministerio del Ambiente del Ecuador - MAE, 2015). Nevertheless, the observed maximum concentration of nitrate-N equaled the appropriate maximum level to protect the most sensitive freshwater taxa (2 mg/L) (Camargo *et al.*, 2005; Kincheloe *et al.*, 1979). Therefore, future management of ecological water quality needs to keep the concentration of nitrate-N lower than the guidelines, especially lower than 2 mg/L in order to protect the presence of sensitive freshwater taxa. Moreover, since the concentrations of nitrate-N were not exactly measured due to the limitation of the kit, future monitoring needs to use kits having a lower detection limit than being used in this study. When correct measurement of nitrate-N concentration is taken, the results of the current models and sensitivity analysis can be tested and the influence of nitrate-N on the ecological water quality can be further quantified.

5.4.1.4 Sediment and river banks

Angular sediment was also found to promote the ecological water quality, since more angular sediment allows macroinvertebrates to attach onto the sediment surface and avoid their drifting (Holomuzki and Biggs, 2003). The angularity or roundness of sediment indicates the amount of transport it had, and fine sediment deposition in the water can reduce the angularity of the rock (Flügel, 2004). Regarding bank shape, several studies suggested the importance of stable river banks to improve the water quality and the macroinvertebrate community. Raymond and Vondracek (2011), for example, suggested a positive correlation between a stable river bank and the macroinvertebrate assemblage by converting conventional grazing to rotational grazing in farming. Similar to land use management, Lester and Boulton (2008) also suggested that bank stability can be improved through the exclusion of grazing animals from river banks and revegetation of the river banks.

5.4.1.5 Flow velocity

Another key variable was the flow velocity, which is often highly diverse in a river basin. Flow velocity is generally related to the elevation (Forio *et al.*, 2015), the amount of rainfall and water transport through the basin. Flow velocity is also linked with the substrate, land use and channel slopes in the up-stream locations (Townsend *et al.*, 1997). The importance of velocity in studying water quality was also deduced by Hughes *et al.* (2009) and Arimoro *et al.* (2015). A slow flow velocity allows the deposition of fine sediments (Wyzga *et al.*, 2014; Wyzga *et al.*, 2009), which consequently inhibits water exchange and oxygen transport (Boulton *et al.*, 1997) and supports nutrients and contaminants transfer (Collins and Walling, 2007) within the water, a condition that can be harmful to aquatic animals. A high flow velocity provides more suitable habitat and offers continuous food and oxygen supply for aquatic animals, thus improving the ecological water quality (Dominguez-Granda *et al.*, 2011b; Fornaroli *et al.*, 2015; Kairo *et al.*, 2012). However, altering the flow velocity of the rivers is difficult, especially in low-land areas, where flow increase can only be induced by a lower water use (e.g., irrigation) or the removal of obstructions at the upstream, such as hydropower dams.

5.4.1.6 Seasonal aspect

At the downstream parts of the rivers and at tributaries that were disconnected from their main channels, elevated levels of several environmental variables, such as conductivity, were observed. It is assumed that this is related to the seasonality, where the late dry season is usually characterized by the lower water quality conditions of the surface waters, since environmental variables have reached their extreme levels. Generally, temperature, conductivity, chlorophyll a and turbidity highly increase through the dry season (Blanchette and Pearson, 2013; Everaert *et al.*, 2014; Garcia *et al.*, 2015). The temperature, pH, conductivity, turbidity and DO vary temporally, and more specifically, the temperature follows seasonal trajectories, while DO can vary significantly within a 12-h period at similar depths (Blanchette and Pearson, 2013). Low nutrient levels, but increasing chlorophyll a and primary production through the dry season indicate rapid nutrient uptake by primary producers due to long exposures to sunlight (Garcia *et al.*, 2015).

Increasing disturbance also influences the presence of more tolerant macroinvertebrates, and the interruption from upstream assemblages during the dry season reduces macroinvertebrates abundance and diversity. However, macroinvertebrates' responses towards environmental changes might vary spatially and across habitats (Blanchette and Pearson, 2013; Damanik-Ambarita *et al.*, 2016b; Garcia *et al.*, 2015; Helson and Williams, 2013). The dry season is also characterized by low flow periods, whereas high flooding flows characterize the wet season. During the wet season, wet season floods support ecosystem replenishment, and habitat conditions are getting more stable when floods recede, which then allows the settlement and growth of macroinvertebrate communities (Damanik-Ambarita *et al.*, 2016b; Garcia *et al.*, 2015). However, the conditions during the dry and wet seasons could not be compared, since the sampling campaign was only performed at the end of the dry season. Moreover, Greenwood and Booker (2015) stressed the importance of studying the temporal variations of hydrological and ecological data to capture the full picture of aquatic systems and to define the response of aquatic organisms towards disturbances. Because the sampling campaign was only performed once and within a short period, the degree of hydrological and ecological variability of the rivers over time, as well as the variations in community compositions could not be assessed as well. Thus, continuous monitoring of the aforementioned variables during the dry and wet seasons can provide better understanding of the temporal variations of the ecological water quality of the river basin.

5.4.2 Model development and validation

Dealing with a complex and dynamic system in aquatic ecology, the coefficient of determination R^2 values of all models, both models using the complete dataset and those using three-fold cross-validation, indicated a good model fitness to predict the ecological water quality. The robustness of the outcome of the modelling exercises was tested and it was found that models based on the complete dataset had similar

R^2 values for development and validation. However, when assessing each of the separate folds, the R^2 values of the training datasets ranged from 0.52 to 0.62, and the validation datasets ranged from 0.31 to 0.49.

Besides the differences in the coefficient of determination, the cross validation also showed some differences in the importance of environmental variables in the models. These differences are visualized in the sensitivity analyses. For example, flow velocity was selected as a key variable in training set 1+3 for models with the lowest AIC, $p < 0.1$ and $p < 0.05$, while the same results were not taken from all models with training set 2+3. Another example is chlorophyll a that was selected as a key variable in training sets 1+2 and 1+3 for models with the lowest AIC, $p < 0.1$ and $p < 0.05$, while it was not selected as a key variable in all models with training set 2+3. However, certain variables were always selected as key variables, despite their relative importance in the models based on the *p*-values. As such, the use of cross-validation is helpful to avoid the model overfitting. Cross-validation also allows model validation using an independent dataset without reducing the number of samples that can be used (Zuur, 2009). Thus, this shows the importance of the variables ranking in defining the most influencing variables from all key variables, instead of choosing one best model. This way, more options are available for monitoring and restoration actions. However, the use of the lowest AIC to select the best model in future studies is recommended.

The parameter used to stratify the dataset before splitting was assumed to cause the presence of several ‘outliers’ in the residual plots of the models. Most of them represented the same sites with very high BMWP-Col values within the dataset. The models under-predicted the ecological water quality values as compared to the actual values, while the remaining few other sites were overly predicted. These results suggested that the models can predict the ecological water quality within a certain range of values. To improve model performance, it is recommended that future studies can be done by splitting the dataset based on the BMWP-Col values, instead of its classes. Another recommendation is to analyze the reservoir, up- and down-stream parts of the river basin separately. This idea was not tested in this study, since the composition of macroinvertebrate taxa was relatively similar in the reservoir, up- and down-stream parts of the river basin. However, future monitoring might observe a different taxa composition and the results of future monitoring can then be compared to this study.

The results also proved the ability of GLMs to determine the relative importance of each environmental variable towards the ecological water quality and macroinvertebrate communities in particular, which is an advantage over other techniques, such as artificial neural network – ANN (Mouton *et al.*, 2010; Thuiller, 2003). However, one limitation of using GLMs in R software as compared to other techniques was also experienced. The use of GLM using R software will treat the dataset based on complete cases, thus variables with missing values (i.e. COD, stream width and stream depth) had to be removed before starting the model selection process in order to keep all 120 observations. This might cause the model to miss important variables (e.g. COD), since the model could not include the influence of organic material of the water in studying the relationship between ecological water quality and environmental variables. Besides COD, data regarding oxygen demand of the water (both biological and chemical) were not available which might further limit model's ability to define the exact influence of environmental variables on the ecological water quality. Bayesian belief networks (BBN) can easily deal with missing values (Forio *et al.*, 2015); however, the BBN cannot deal with many variables and continuous variables need to be discretized. Since there are many variables in this study, the BBN was not an option. Other model such as DISQUAL discriminant analysis treats missing values as entirely separate values (Tuffery, 2011). However, DISQUAL method deals with nominal dependent variables (Tuffery, 2011), which was not the case in this study. To complete the data, maximum likelihood, multiple imputation, Bayesian and weighted estimating equations can also be used in handling missing values (Donner, 1982; Ibrahim *et al.*, 2005). When missing values cannot be avoided, any of the aforementioned methods can be applied to complete the data. Nevertheless, recommending the use of one particular model for a given problem is practically impossible, and each study may require a different modelling technique (Solomatine and Ostfeld, 2008).

5.5 Conclusions

The physico-hydromorphological (i.e., elevation, sediment angularity, logs, main macrophytes, flow velocity, turbidity, bank shape and land use) and chemical (i.e., nitrate-N and chlorophyll a concentrations) variables were found as the major variables that influenced the macroinvertebrates of the Guayas River basin in Ecuador. The relevance of the variables analyzed via a sensitivity analysis and cross-fold validation provided insights for the stability of the outcomes. Limitation on nitrate-N measurement and the exclusion of COD from analyses might influence model's ability to evaluate the relationship between the BMWP-Col and environmental variables. To restore and protect river ecosystems and their functions, and in particular, macroinvertebrate communities, policy actions need to focus on alterable variables, such as management of land use at different elevations, management and monitoring of nitrate-N and chlorophyll a concentrations, macrophyte presence, sediment transport and bank stability. Measuring nutrients and oxygen demand of the water and monitoring during the rainy season will provide more information in future ecological water quality studies of the Guayas river basin.

Chapter 6: Impact assessment of local land use on ecological water quality of the Guayas river basin (Ecuador)

Adapted from:

Minar Naomi Damanik Ambarita, Pieter Boets, Thi Hanh Tien Nguyen, Marie Anne Eurie Forio, Gert Everaert, Koen Lock, Peace Liz Sasha Musonge, Natalija Suhareva, Elina Bennetsen, Sacha Gobeyn, Tuan Long Ho, Luis Elvin Dominguez Granda, Peter L.M. Goethals. Impact assessment of local land use on ecological water quality of the Guayas river basin (Ecuador). Ecological Informatics (under review).

Abstract

Extensive anthropogenic activities including land conversion have been taking place in the Guayas river basin (Ecuador) due to increasing population growth. Land use changes are considered one of the key sources affecting the ecological water quality of the Guayas river basin. Therefore, this chapter investigated the effect of land use on the ecological water quality both within direct vicinity and within a distance of 200 m from the sampling sites. This chapter investigated which of three land use assessment methods (i.e. field protocols, Google maps data and GIS data) is most suitable to quantify the impact of local land use on the ecological water quality and which key environmental variables influence the ecological water quality. To do so, the relation between the BMWP-Col, local land use and other environmental variables was investigated using general linear models (GLMs) and sensitivity analyses. Based on multi-model comparison, the ecological water quality was best associated with the land use close to the sampling sites (Google land use, $R^2 = 0.93$, $p < 0.05$). Models involving land use assessed using Google maps were associated mainly with physico-chemical variables, whereas models involving land use originating from field protocols and GIS data were associated mainly with hydromorphological variables.

6.1 Introduction

During the last decades, human population growth has resulted in an increasing demand of human settlements and economic development such as food production. Worldwide, forest and rural areas have been converted into residential and agricultural lands to meet human demands (Schmalz *et al.*, 2015; Smucker and Detenbeck, 2014). Urbanization has been intensified and cities were enlarged to accommodate the worldwide population increase. Rivers have been regulated to create reservoirs to provide electricity, drinking water and water for irrigation (Bertone *et al.*, 2016; Strehmel *et al.*, 2016). In addition, industries have grown exponentially and have consumed a lot of the available water (Tuan *et al.*, 2016). These anthropogenic activities constitute a pressure on the water quantity and quality as well as ecosystem services of water bodies (Courtonne *et al.*, 2016; Smucker and Detenbeck, 2014).

A river basin usually comprises more than one type of land use varying in size and spatial distribution. An area may be fragmented and mixed with different types of land use, such as forest, residential or farm land (Ferreira *et al.*, 2016; Goss *et al.*, 2014). Watercourses are not restricted to one type of land use and might flow through residential or agriculture land before passing through a forest downstream (Wilkins *et al.*, 2015). Therefore, pressure on the water bodies can come from different types of land use at varying spatial scales (Cortes *et al.*, 2013; Palmer *et al.*, 2014).

Water bodies such as rivers and lakes act as the receiving environment of (waste)water discharge from residential, agricultural and industrial lands (Ferreira *et al.*, 2016; Poff *et al.*, 2006). Agricultural run-off often ends up in surface waters, changing water quality variables such as nutrient concentrations and sediment composition (da Silva *et al.*, 2015; Goss *et al.*, 2014). Therefore, water quality determination is required to assess the impact of land use, especially in areas where anthropogenic presence is apparent. Moreover, previous studies have shown the importance of local land use in determining the ecological water quality of the river basin (Damanik-Ambarita *et al.*, 2016a; Damanik-Ambarita *et al.*, 2016b). For example, impact of land use on water quality has been investigated in previous studies within direct vicinity of the sampling sites (Cortes *et al.*, 2013; Epele and Miserendino, 2015; Zhang *et al.*, 2010), within a certain radius of the sampling sites

(Manfrin *et al.*, 2016), at the catchment level (Erba *et al.*, 2015) or based on a combination of spatial scales (Erba *et al.*, 2015; Leps *et al.*, 2015). Cortes *et al.* (2013) and Manfrin *et al.* (2016) agreed with the positive influence of forest on the water quality; whereas residential (Cortes *et al.*, 2013; Manfrin *et al.*, 2016; Zhang *et al.*, 2010) and agriculture related land uses (Leps *et al.*, 2015; Zhang *et al.*, 2010) influence the water quality negatively.

Land use has been assessed through field observation (Erba *et al.*, 2015), using aerial maps (Manfrin *et al.*, 2016; Zhang *et al.*, 2010) or geographic information system (GIS) data (Cortes *et al.*, 2013; Epele and Miserendino, 2015; Leps *et al.*, 2015). Field observations follow a certain protocol such as river habitat survey (Raven *et al.*, 1997) where the observer is required to do transect walks within a certain distance from the sampling sites (Erba *et al.*, 2015). Land use assessment using aerial maps and GIS data do not require field observations, since the assessor can access the data online or through a research institute (Cortes *et al.*, 2013; Manfrin *et al.*, 2016; Zhang *et al.*, 2010). However, studies comparing the differences among different spatial scales or different assessment methods are still limited. Therefore, in the present chapter the integration of different assessment methods in evaluating local land use impact on the ecological water quality was investigated.

Impact of land use on the water quality can be studied by monitoring the physico-chemical variables of the water (Poff *et al.*, 2006). Since physico-chemical variables of the water can fluctuate easily, integrating bioassessment using aquatic organisms such as macroinvertebrates has been regarded beneficial in studying the water quality. Bioassessment can be performed by calculating biotic indices to determine the ecological water quality status of surface waters (Oliveira *et al.*, 2011; Sundermann *et al.*, 2015; Verissimo *et al.*, 2012; Wilkins *et al.*, 2015). Moreover, the relationship between the biotic index and environmental variables can be evaluated to determine key environmental stressors affecting the ecological water quality using statistical models (Everaert *et al.*, 2014; Holguin-Gonzalez *et al.*, 2013; Schuwirth *et al.*, 2016; Tchakonte *et al.*, 2015).

The present chapter aims to evaluate which type of land use assessment method is most suitable to quantify the impact of human stressors on the ecological water quality of the Guayas river basin, Ecuador. To do so, three different types of land use data, originating from three different sources were used. The land use data based on field protocols and Google maps were used to quantify the local land use within direct vicinity of sampling sites, while the land use data based on GIS information were used to quantify the local land use within a 200-m distance from the sampling sites. These data in combination with the conventional physico-chemical variables were used to quantify the ecological water quality changes. Furthermore, the key environmental variables associated with both land use scales were also investigated. General linear models (GLMs) and sensitivity analyses were utilized to meet these objectives.

6.2 Data analysis

In this chapter, three land use assessment methods were used: “FP land use”, “Google land use” and “GIS land use” (Table B.2). Both the “FP land use” and the “Google land use” were classified into four categories: forest, arable, residential and orchard (“FP land use”); forest arable, residential and pasture (“Google land use”). The “GIS land use” were classified into seven categories: residential; agriculture; a mix of agriculture, livestock, forest and urban; a mix of agriculture, livestock and forest; livestock; a mix of livestock and conservation and protection; and conservation and protection. Since several “GIS land use” categories were not supported by sufficient observations, this classification had to be condensed for analyses. Because the categories do not share relatively similar use to be combined into one category in order to have enough observations, the sampling sites were divided into two categories for the analyses: (1) agriculture and (2) all other categories. This condensation was intended to have enough representation of each category for data partition in analysis. Several sampling sites had different categories according to the three land use assessment methods and sources (Tables E1-E3). The BMWP-Col (Roldán Pérez, 2003) based on Alvarez (2005) was used to calculate the ecological water quality index of the 120 sampling sites. The continuous variables listed in Table 4.1 preprocessed data were used for analyses.

To assess which land use assessment methods worked best to quantify the ecological water quality, a five-step approach (Fig. E1) was followed. The summary statistics of continuous variables are presented in Table 4.1 preprocessed data, while the categorical variables are presented in Table B1. In the first step, COD, stream width and stream depth were removed from the dataset due to missing values. In the second step, collinear variables based on variance inflation factors (VIF) were removed; where variables having VIF values higher than three were considered as collinear, following the procedure described in Zuur (2009) and Zuur *et al.* (2007). The VIF analysis was done involving the three land use assessment methods separately: one time involving the “FP land use”, the “Google land use”, and the “GIS land use”, respectively. These pre-processing steps resulted in three sets of environmental variables: 23 variables involving the “FP land use” (Table E4), 24 variables involving the “Google land use” (Table E5), and 24 variables involving the “GIS land use” (Table E6).

In the third step, general linear models (GLMs) involving these three sets of non-collinear variables were developed separately. The GLMs were developed and validated using three-fold cross validation (CV). The three-fold CV was done by splitting the complete dataset randomly into three equal subsets, where the BMWP-Col classes and land use assessment methods have been used to stratify the dataset prior to splitting. Three sets of models were developed: models involving the “FP land use”, the “Google land use”, and the “GIS land use”, respectively. The three-fold CV resulted in three models for each set; therefore there are nine models in total (more detailed explanation on GLM and the use of three-fold CV is given in chapter 3).

In the fourth step, model performance was evaluated by calculating the number of input variables, the coefficient of determination (R^2) between actual and calculated BMWP-Col values, weighted and unweighted Kappa between actual and calculated BMWP-Col classes, the p-values and the average of all folds. Lastly, the key environmental variables were selected based on a significant presence ($p < 0.05$) of minimum two times in the three-fold CV models for each land use set. This was done by ranking the variables according to their significant levels in each model. Only variables having significant level $p < 0.05$ can be considered a key variable. Since there were three final models from the three-fold CV, each variables having $p < 0.05$ had to appear in minimum two final models to be selected as key variables. R software version 3.0.2 was used to perform all statistical analyses (R-Core-Team, 2013). The Kappa value was calculated using the *psych* package in R (Revelle, 2016). A sensitivity analysis was also performed after all the five steps were finalized. Each land use category from the three land use assessment methods was used to define its influence on the ecological water quality together with a key variable that was assessed. The sensitivity analysis was performed on the model configuration that resulted in the best model fitness, using values listed in Table E7.

6.3 Results

6.3.1 Observed ecological water quality

The BMWP-Col index ranged from very bad (0) to good (168) (Fig. E2). The "Google land use" input data was more effective in discriminating the different ecological water quality among sites characterized by different land use types, compared to the "FP land use" and the "GIS land use" input data (Fig. 6.1). For example, the "Google land use" input data indicated that sites characterized by forested land use had a good ecological water quality, followed by pasture, arable and residential. A clear distinction of the ecological water quality was not provided by the "FP land use" input data, whereas the "GIS land use" input data indicated that sites characterized by agriculture had better ecological water qualities compared to sites characterized by all other categories. Nutrient and turbidity concentrations were generally low, whereas the DO concentration mainly ranged between 6 and 10 mg/L. Shaded sites generally had a good ecological water quality, as well as sites with some flow variations, a flat bank slope, and more angular sediment (Fig. E2).

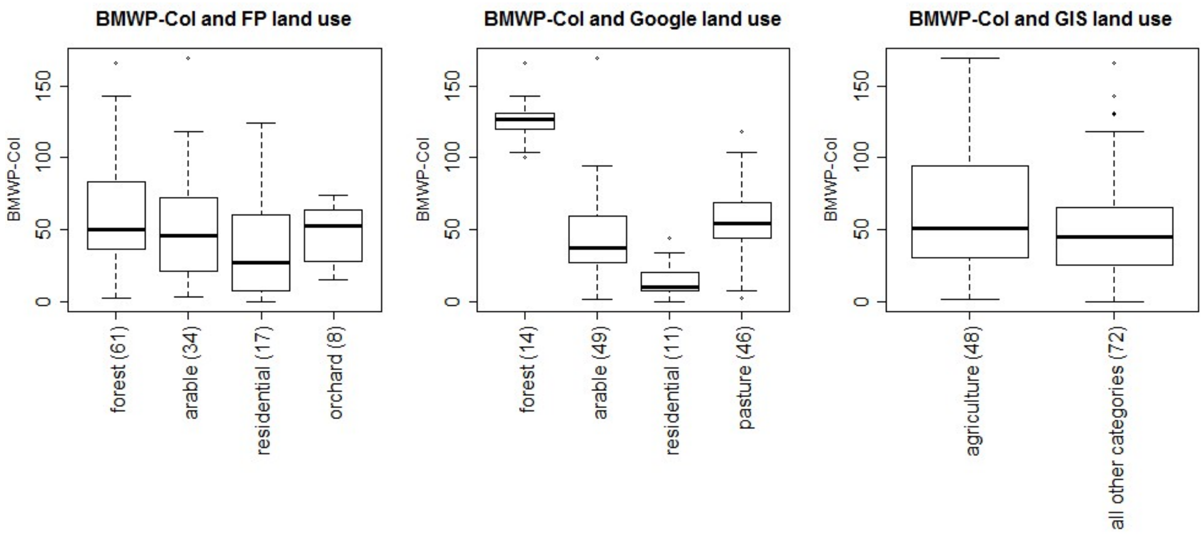


Figure 6.1 Plots showing the distribution of the data for the three land use assessment methods and source in relation to BMWP-Col for 120 sampling sites, number of observations is shown in brackets.

6.3.2 Land use and associated environmental variables

The selected variables resulting from the three-fold CV did not vary much among the models (Fig. E3-E11 and Tables E8-E12). Models involving the “Google land use” covariate had the best R^2 (0.93), weighted Kappa (0.91) and unweighted Kappa (0.73) compared to models involving the “FP land use” and the “GIS land use” covariates (Fig. 6.2 and Table E11). The “FP land use” was selected as a key variable in three folds of the three-fold CV and the “Google land use” was selected as a key variable in two folds of the three-fold CV. However, the “GIS land use” was only selected as a key variable in one fold. The key variables for the “FP land use” and the “GIS land use” models were mainly associated with hydromorphological variables, whereas the “Google land use” models were associated more with physico-chemical variables (Table 6.1).

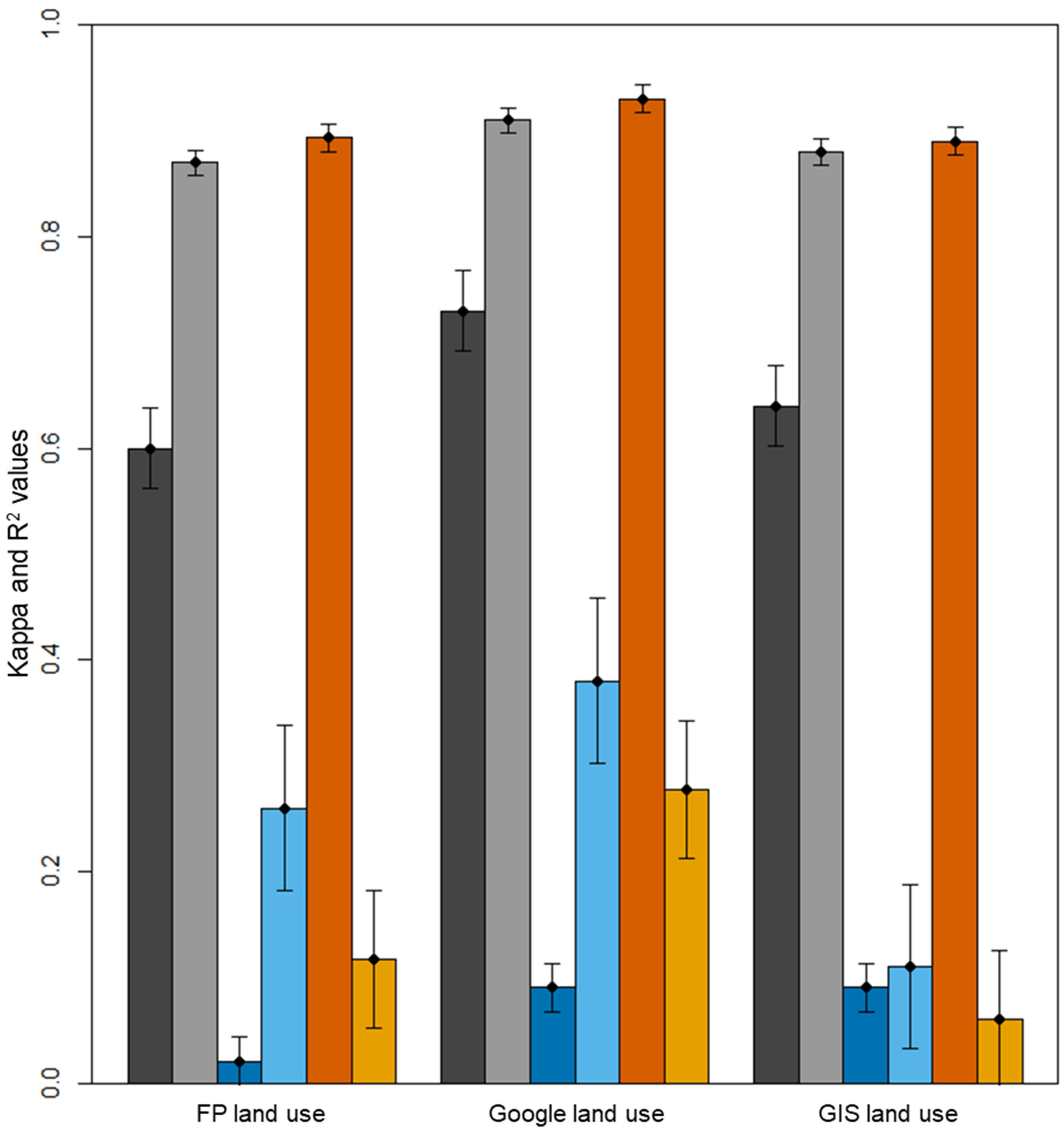


Figure 6.2 Average model performances based on the three-fold cross validation of the three land use sets (based on Table E11) with their standard errors; (unweighted Kappa training), (weighted Kappa training), (unweighted Kappa testing), (weighted Kappa testing), (R^2 training), (R^2 testing).

Table 6.1 Selected key variables influencing the ecological water quality for the three land use sets with their number of presences in the three-fold cross validation models, each variable significantly contributed ($p < 0.05$) to the models.

Variables	# presences in three-fold CV FP land use set	# presences in three-fold CV Google land use set	# presences in three-fold CV GIS land use set
Total P		2x	
DO		2x	
Turbidity		2x	
Chloride		2x	
Nitrate-N	2x	2x	
Shading	3x	2x	3x
Variation in width		2x	
Erosion	2x		
Bank profile	2x		2x
Flow variation	2x	2x	2x
Bank shape	2x		2x
Bank slope	2x		
Sediment angularity	3x		3x
Land use	3x	2x	

6.3.3 Sensitivity analysis

Here, the effects of different land use types (Fig. 6.3), total P, nitrate-N, turbidity and DO concentrations with different land use categories (Fig. 6.4-6.5 and E12-E13) are presented; whereas the effects of different flow variation, bank profile, shading and sediment angularity are shown in supporting information (Fig. E14-E17). The sensitivity analysis was performed on the key variables presented in Table 3 using fold 3, fold 2 and fold 1of the “FP land use”, the “Google land use” and the “GIS land use” models, respectively.

The outcomes of the models incorporating each of the three land use assessment methods varied. Models involving the “FP land use” did not result in a notable difference in the ecological water quality for forest, arable and orchard land uses; except for residential land use that resulted in a bad ecological water quality (Fig. 6.3). Models using the “Google land use” clearly distinguished forest from other categories: forest is related to a good ecological water quality, followed by pasture, residential and arable land uses. Models using the “GIS land use” distinguished all other categories from agriculture: all other categories will promote the ecological water quality better than agriculture.

Sensitivity analyses performed on total P concentration resulted in contradictory outcomes: the “FP land use” suggested an increasing ecological water quality (e.g. from 63 to 275 with forested land use, Fig. 6.4) with increasing total P concentration (from 0.5 to 5 mg/L), while the “Google land use” suggested the opposite. Sensitivity analyses performed by increasing nitrate-N concentrations from 0.2 to 2 mg/L also resulted in contradictory outcomes: the “FP land use” suggested a decreasing ecological water quality (e.g. from 63 to 9 with forested land use, Fig. 6.4), whereas the “Google land use” suggested an increasing ecological water quality (e.g. from 134 to 197 with forested land use). Sensitivity analyses performed on turbidity and DO concentrations gave similar results: increasing turbidity will decrease the ecological water quality, whereas increasing DO will increase the ecological water quality (Fig. 6.4-6.5 and Fig. E12-E13). Sensitivity analyses executed on selected key hydromorphological variables (Fig. E14-E17) also provided similar results using the three land use assessment methods: partial and complete shading supported a good ecological water quality, similar to moderate flow variation and flow variation due to construction. Higher BMWP-Col values were also supported by composite-but-not-trampled bank profile and rounded sediment.

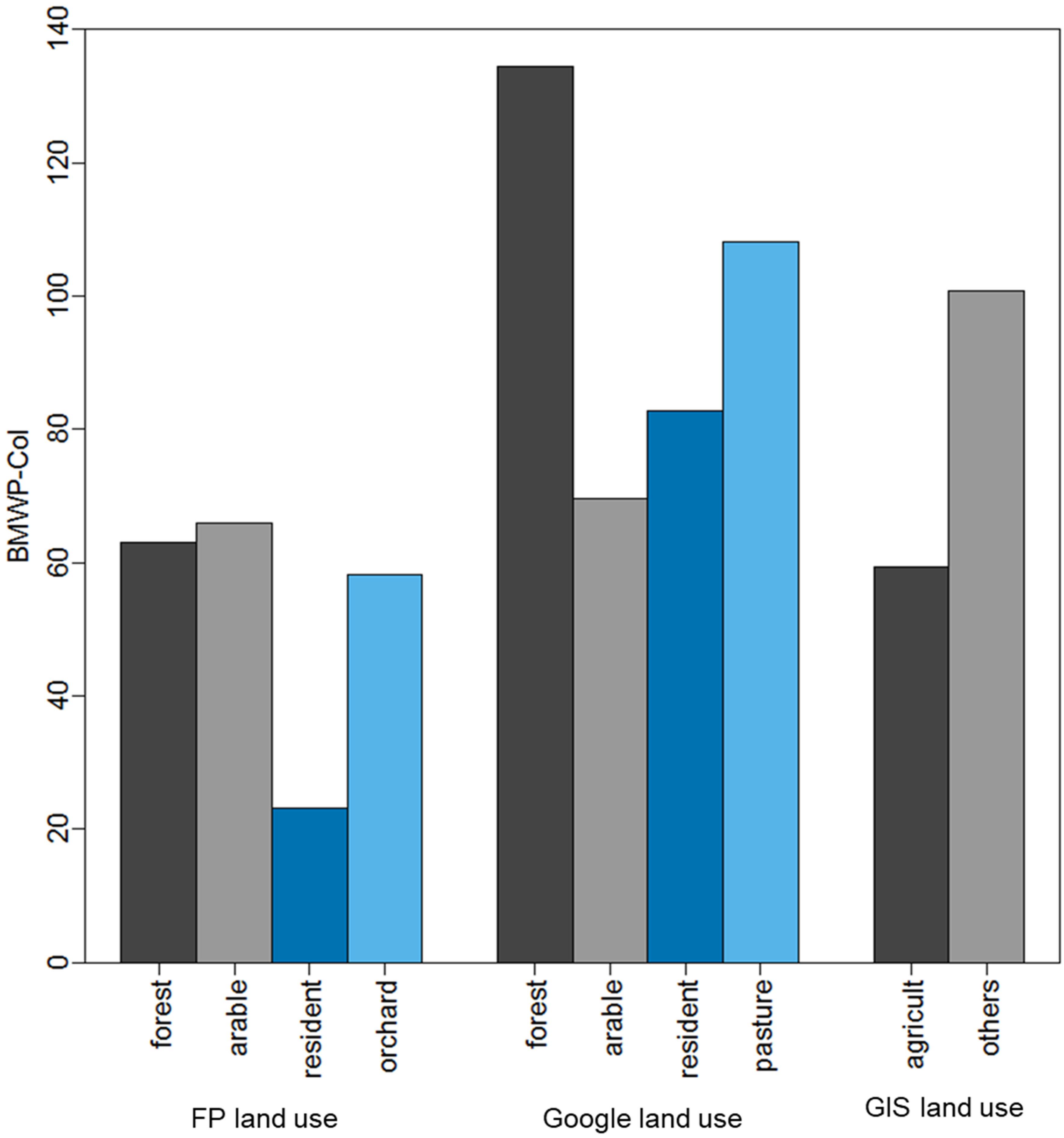


Figure 6.3 Sensitivity analysis showing the effects of varrying land use categories on the BMWP-Col under the condition that other variables were set constant to their “median” values; resident: residential, agricult: agriculture.

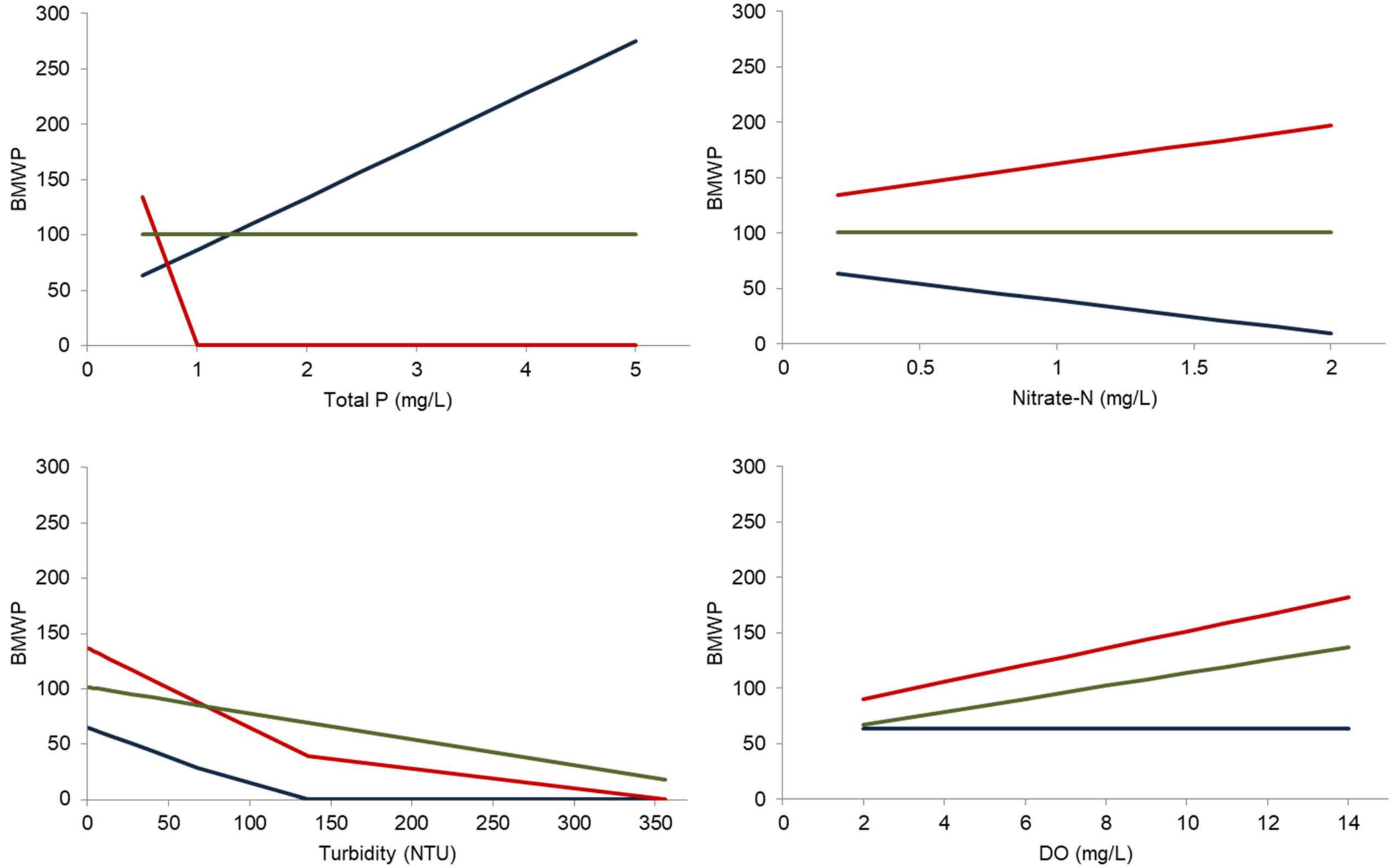


Figure 6.4 Sensitivity analysis showing the effect of total P, nitrate-N, turbidity and DO concentrations on the BMWP-Col: ___ (FP land use forest), ___ (Google land use forest), ___ (GIS land use all other categories).

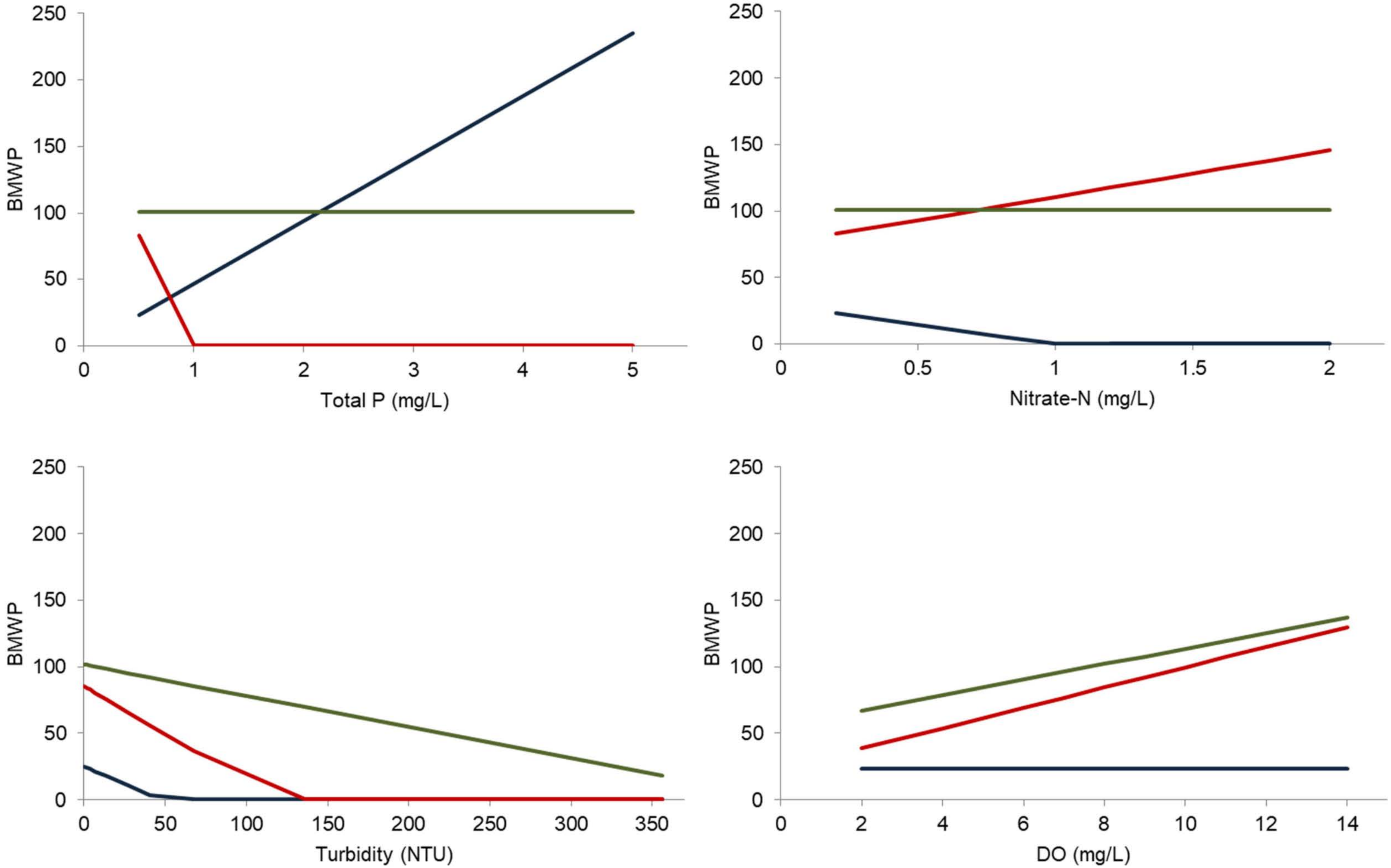


Figure 6.5 Sensitivity analysis showing the effect of total P, nitrate-N, turbidity and DO concentrations on the BMWP-Col: ___ (FP land use residential), ___ (Google land use residential), ___ (GIS land use all other categories).

6.4 Discussion

6.4.1 Comparison of local land use assessment methods in relation to ecological water quality

Models containing the “Google land use” had the best average coefficient of determination ($R^2 = 0.93$), Kappa (0.91 and 0.73 for weighted and unweighted Kappa, respectively) and p-value (< 0.05) compared to models containing the “FP land use” and the “GIS land use”. The models containing the “Google land use” on average consist of more explanatory variables compared to the models containing the “FP land use” and the “GIS land use”, thus better explain the relationship between the BMWP-Col and the environmental variables (Table E11).

The associated key variables for the “Google land use” models were mainly the chemical variables, whereas the associated key variables for the “FP land use” models were hydromorphological variables. The only chemical variable selected in both sets was nitrate-N concentration, whereas shading, variation in width and flow variation were the hydromorphological variables selected in both sets. However, it can be concluded that the results of both models containing the “FP land use” and the “Google land use” are complementing each other.

The sensitivity analysis using the “Google land use” set showed that sites surrounded by forest is associated with a good ecological water quality, whereas the same conclusion was not apparent for the sensitivity analysis performed using the “FP land use”. The category “residential” showed the most negative influence on the ecological water quality compared to “arable”, “orchard” and “pasture”. These results agree with the collected data. However, the sensitivity analysis performed using the “GIS land use” set shows opposite results to the collected data: the sensitivity analysis suggested that agriculture is negatively associated with the ecological water quality, whereas the collected data showed that sites surrounded by agriculture resulted in a better ecological water quality compared to all other categories.

The category and spatial coverage differences between the “FP land use” and the “Google land use” classifications might cause the different result. Both the “FP land use” and the “Google land use” classifications were divided into four categories, with three similar categories and only one difference: the “FP land use” has a category orchard, while the “Google land use” has a category pasture. Out of 120 sites, only 34 sites had a similar land use in both sets. Moreover, the “Google land use" covers a larger area compared to the "FP land use", therefore it provides more information on the land use within the direct vicinity of the sampling sites. Since aerial mapping of sampling sites provides a larger coverage than field observations, both assessment methods can be combined to obtain sufficient information regarding land use within direct vicinity of the sampling sites. The use of different scales for both methods can be maintained. Thus, the use of similar categories for both land use classifications is also suggested whenever possible, which will simplify the combination of both assessment methods. Nevertheless, both the "FP land use" and the "Google land use" can be combined if the "Google land use" information is updated within the same period as when field observations took place.

As already discussed by Chapman (1996) and Robinson *et al.* (2014), agriculture and residential areas are potential sources of nutrient enrichment in surface water which may cause eutrophication. Farming can also decrease oxygen concentration, modify stream channels and banks, change the type of riparian vegetation, increase erosion and sediment input (Robinson *et al.*, 2014) and increase turbidity (Turunen *et al.*, 2016). Indeed, in the present chapter a strong relation between residential and agriculture related activities and nutrient concentrations and hydromorphological variables (Fig. 6.5 and Fig. E12-E17) was found. Moreover, increasing human settlements and industrial activities showed clear negative effects on water quality leading to an increase of physico-chemical variables such as pH, temperature, conductivity, nitrate and phosphorus (Englert *et al.*, 2015; Younes-Baraille *et al.*, 2005; Yun and An, 2016).

Intensive and continuous grazing in pastured land have shown negative impacts on water quality (Raymond and Vondracek, 2011), a common practice in the Guayas river basin. Grazing animals can reduce riparian vegetation, modify stream channels and banks, increase runoff, erosion and sediment input (Trimble and Mendel, 1995), and transport nutrients into the water (Vondracek *et al.*, 2005). In this context, rotational grazing has been considered as an alternative system to lower the negative impacts of grazing animals. Rotational grazing allows the growth of vegetation up to a minimal height that is beneficial for the animals and provides shade to the water (Raymond and Vondracek, 2011). Therefore, replacing intensive and continuous grazing with rotational grazing can provide a better grazing management in the Guayas river basin and thus a better water quality.

An increasing demand of land for agriculture and domestic purposes in Ecuador enhances deforestation. Previous surveys showed an annual deforestation rate of 2.86% for 1989–2008 in South Ecuador only (Tapia-Armijos *et al.*, 2015). Forested areas are beneficial in maintaining a lower water temperature, pH, conductivity, turbidity and nutrient concentrations, providing food in the form of organic matter for aquatic organisms (Kasangaki *et al.*, 2008b; Townsend *et al.*, 1997), preventing pollutants from entering the streams, and enhancing in-stream processing of pollutants (Sweeney *et al.*, 2004). Deforestation will consequently lower water quality and ecosystem functioning (Tapia-Armijos *et al.*, 2015; Townsend *et al.*, 1997). Fortunately, efforts have been taken to better protect forested landscapes, to reforest clear-cut sites and to reduce deforestation in Ecuador (Bass *et al.*, 2010; REDD, 2011).

The “GIS land use” models suggested only hydromorphological variables as the key variables that determined the ecological water quality, and none of the chemical variables. Land use was not considered a key variable affecting the ecological water quality in these models. These results are in agreement with the studies by Park *et al.* (2011) but they are different from the study conducted by Rios and Bailey (2006). Generally, the “GIS land use” was expected to provide more information of local land use because it covers the largest area among the three local land use classifications; however, the “GIS land use” data might have been outdated. It is acknowledged that the data was published in 2012, while data collection might have been done much earlier than the publication time and several land use conversions might have been taking place. Moreover, the “GIS land use” classification was condensed into two categories, which might reduce the ability of the models to define the detailed impact of local land use. As suggested by Crétaz and Barten (2007) and Hansen *et al.* (2010), the size of riparian zone is around 100 m wide. Therefore in future studies, the use of a scale similar to the “Google land use” scale (100*100 m) is recommended.

GIS data for land use will provide more information for impact assessment of land use on ecological water quality within direct vicinity of the sampling sites than merely field observation. However, GIS data needed to determine the land use is not always available and updated in developing countries. Whenever possible, the inclusion of land use data retrieved from GIS data is necessary. However, when the

GIS data is not available, local land use assessed through field observation and aerial mapping are required. Nevertheless, the models involving the three land use classifications put forward similar key variables (shading, sediment angularity, bank profile, flow variation and bank shape) as explanatory variables for both scales within the direct vicinity and within 200 m distance from the sampling sites.

Land use influences the water quality of a small stream more than a large river, because it covers a larger proportion of the small stream's catchment area. Water quality degradation at a small stream will eventually influence the water quality of a larger river downstream (Walsh *et al.*, 2004). The Guayas river basin is composed of small streams, big rivers and a reservoir; therefore, larger-scale actions involving restoration at the source of environmental stressors will be more beneficial for water quality than merely end of pipe measures. Depending on the needs and the government's regulation, these can include best management practices in the watershed, dam removal, creation of wetlands and revegetation of riparian buffers (Palmer *et al.*, 2014; Palmer *et al.*, 2010; Smucker and Detenbeck, 2014; Walsh *et al.*, 2004). These larger-scale actions will improve water quality and aquatic habitat that will consequently enhance the presence and diversity of aquatic organisms (Smucker and Detenbeck, 2014).

6.4.2 Ecological water quality and environmental variables

The results show that ecological water quality was affected more by hydromorphological than physico-chemical variables, which can be seen from the number of hydromorphological variables selected as the key variables in the models (Table 6.1). Land use, shading, sediment angularity, total P, DO, turbidity, chloride, nitrate-N, variation in width, erosion, bank profile, flow variation, bank shape and bank slope affected the ecological water quality significantly ($p < 0.05$, Table E12).

Hydromorphological processes such as sediment flow, channel modification, dam construction, erosion and vegetation presence at an upstream location can affect downstream ecosystems (Poppe *et al.*, 2016). In addition, hydromorphological degradation of the river is often related to multiple stressors (Lorenz *et al.*, 2004). For example, morphological alteration of the river banks and river channel related to agricultural activities alters the microhabitat composition and affects community structure and taxa richness (Lorenz *et al.*, 2004; Turunen *et al.*, 2016). Those stable river banks are required by aquatic organisms and are important habitats to maintain a good water quality (Poppe *et al.*, 2016). Shading provided by riparian vegetation promotes an optimum water temperature and oxygen concentration required by aquatic organisms. Shading also provides particulate organic matter for macroinvertebrates (Lester and Boulton, 2008; Robinson *et al.*, 2014). The removal of riparian vegetation may result in channel incision that consequently destroys the aquatic habitat and reduces the presence of aquatic animals (Lester and Boulton, 2008). These hydromorphological conditions have shown their influence in affecting the ecological water quality in this current study; hence, agriculture activities in the Guayas river basin need to be organized by considering these variables.

The concentrations of total P and nitrate-N in the water generally increase as a result of fertilizer use (Robinson *et al.*, 2014) and the discharge of domestic and industrial wastewater. Phosphorus originating from agriculture mainly enters the water via erosion, since phosphorus has a higher affinity to soil compared to nitrogen. Phosphorus and nitrate are essential nutrients for aquatic organisms such as primary producers and macrophytes (Chapman, 1996). An increased nutrient concentration in the water will stimulate the growth of aquatic plants (Frankforter *et al.*, 2010; Hilton *et al.*, 2006) which consequently will increase the productivity of fish and other aquatic animals. However, high concentrations of either nutrient in the water can be harmful to organisms and can cause eutrophication (Chapman, 1996; Hilton *et al.*, 2006). This condition was observed at several sampling sites, where nutrient concentrations were relatively higher than other sites and the water was fully covered by algae or macrophytes.

The sensitivity analyses suggested contradictory results for total P and nitrate-N concentrations. It should be noted that since the concentrations of total P at most sampling sites could not be detected by the kits, the actual concentrations were unknown, despite the use of the lowest detection limit to replace all missing values in analyses. With this limitation, the models might not be able to define the exact impacts of total P concentration. Future water quality monitoring can use kits with a lower detection limit to be able to measure the accurate total P concentration and to evaluate its effect on the ecological water quality.

However, the correctly measured total P concentrations (minimum 0.8 mg/L and maximum 4.5 mg/L) were considered high compared to other water quality studies (e.g. 0.06-0.11 mg/L in the wetland area Abras de Mantequilla in the Guayas river basin (Alvarez-Mieles *et al.*, 2013) and 0.06-0.8 mg/L in a study by Hou *et al.* (2013)). Generally, phosphorus has high affinity towards sediment (Chapman, 1996; Paudel *et al.*, 2017), thus phosphorus concentration of the water would be relatively low. Hilton *et al.* (2006) discussed the possible reason for the high concentration of phosphorus in the water which is according to their research due to phosphorus saturated sediment. The sediment can no longer adsorb phosphorus, and macrophytes are also unable to uptake additional phosphorus. Consequently, the phosphorus concentration in the water increases (Hilton *et al.*, 2006). However, this study did not measure total P concentration of the sediment; thus, this assumption could not be tested. Future monitoring campaigns need to also measure total P concentrations of the sediment in order to understand its influence on the total P concentration of the water column. Moreover, both minimum and maximum concentrations were much higher than the guideline for surface waters (10 µg/L) from the Ministry of Environment of Ecuador (Ministerio del Ambiente del Ecuador - MAE, 2015). Therefore, water managers need to make effort to reduce total P concentration of the water. Similar to total P, not all nitrate-N concentrations could be detected by the kits that might have resulted in difficulty of the models to evaluate the effect of nitrate-N concentration on the ecological water quality (Damanik-Ambarita *et al.*, 2016a), as discussed in chapter 4 and 5.

6.5 Conclusions

The present chapter assessed which type of land use assessment method is most suitable to quantify the impact of local land use on the ecological water quality using three different assessment methods: the “FP land use”, the “Google land use” and the “GIS land use”. It was found that the “Google land use” had the best outcome. Models involving the “FP land use” and the “GIS land use” were more associated with hydromorphological variables, whereas models involving the “Google land use” were more associated with physico-chemical variables. A combination of field observations and GIS data can provide comprehensive land use data to the water management. However, when an updated “GIS land use” data is unavailable, combined information of the “FP land use” and the “Google land use” is sufficient to define local land use of the sampling sites. The use of similar scale to quantify the “Google land use” and the “GIS land use” is also recommended.

Chapter 7: General discussion

Abstract

Land use was suggested to having clear influence on the ecological water quality of the Guayas river basin. The overall results of this PhD studies selected land use as a key environmental variable affecting the ecological water quality in the Guayas river basin. The ecological water quality was threatened by intensive agriculture related activities and human settlement in residential areas. This chapter provides the findings of each chapter following the outline in chapter 1. This chapter also provides a general discussion on the methodologies applied in data collection and reliability, data analysis and use of macroinvertebrates; the take home message showing macroinvertebrate taxa and their presences at different land uses; and the recommendation for future water quality preservation and studies.

7.1 Introduction

Studies regarding the impact of land use change on ecological water quality are relatively rare in developing countries, most often due to the lack availability of land use data, poor methodology to assess land use impacts or limited water quality monitoring data (see chapter 2). In this PhD study, the impact of human activities on the ecological water quality of the Guayas river basin (Ecuador) based on land use data was evaluated. To do so, the ecological water quality based on macroinvertebrate community was quantified. Anthropogenic impacts on the ecological water quality were evaluated using physico-chemical, hydromorphological and land use data that were collected simultaneously during an integrated sampling campaign.

Multivariate statistics as well as ecological models were used to analyze the relationship between environmental variables and biotic index and to define the relationship between land use and aquatic macroinvertebrates. Multivariate analyses and regression analyses were used, which are two commonly-used modelling techniques in ecological water quality studies.

This chapter summarizes the results of the thesis, and provides a general discussion on the applied methodologies and the recommendations for future ecological water quality management and studies. The results are presented and discussed according to the scheme presented in chapter 1. The methodologies applied within the thesis are discussed regarding data collection and reliability, data analysis, and the use of macroinvertebrates to assess the ecological water quality. Lastly, a take home message and recommendations on how to preserve the ecological water quality and to perform future studies to overcome the limitations encountered in this PhD study are provided.

7.2 Quantitative analysis to infer relationship between land use and aquatic macroinvertebrates (Ch. 2)

Questions: Why is land use information often not included in ecological water quality studies? What is the best way to include land use information in ecological water quality studies? Are ecological models useful to quantify the relationship between land use and ecological water quality?

➔ By reviewing published scientific papers, it was clear that limited availability of (in particular local) land use information made it challenging to integrate land use information in ecological water quality studies. This result confirmed the first hypothesis regarding the lack availability of land use information. This was due to the fact that land use data is not regularly updated, not all countries have registered all of their land use, and in some cases the land use is only assessed when there is a specific sampling campaign taking place at the area. The second research hypothesis, i.e. that the relationship between land use and the ecological water quality is still quantified with insufficient methodologies, was also confirmed. For example, Alemneh *et al.* (2017) only assessed the information of local land use through field observation in defining the ecological water quality of Choke mountain catchment, while Feio *et al.* (2007) performed their bioassessment by considering only the local land use. It can be concluded that the inclusion of land use information from both local and regional scales in ecological water quality studies will provide better understanding on land use impacts on the ecological water quality. Whenever possible, combining field and online observations in obtaining land use information is recommended. Furthermore, the benefits of using models to define and quantify the relationship between land use and the ecological water quality also confirmed the last hypothesis. In this context, various model types can be selected based on the aim of the study and the nature of the data. Multivariate analyses, regression analyses and decision trees are the most commonly applied methods in ecological water quality studies. Despite their advantages, the use of ecological models also has its disadvantages. Some disadvantages are: there is no single model that can assess all environmental conditions, thus there is a need of pre-analyses to select appropriate models for an intended purpose. Another disadvantage is that models simplify variables selection. In the future, the goal should be to find solutions to counter the disadvantages of using ecological models.

7.3 Ecological water quality status of the Guayas river basin, Ecuador (Ch. 4)

Question: What is the current ecological water quality of the Guayas river basin, Ecuador?

➔ Two biotic indices to calculate the ecological water quality status of the Guayas river basin were used: the BMWP-Col (ranging from very bad to good) and the NLSMI (ranging from bad to reference). The ecological water quality of the Guayas river basin ranged from very bad (0) to good (168) according to the BMWP-Col and from bad (0) to reference (9.1) according to the NLSMI. Sites located at higher elevations (i.e. > 250 m) and tributaries of the Babahoyo river had high BMWP-Col (i.e. > 100) and NLSMI (i.e. > 6) values compared to sites located downstream and at the main channel of the Daule and Babahoyo rivers. Additionally, the ASPT was also calculated to further assess the suitability of the BMWP-Col. Nutrient concentrations within the sampling sites were generally lower than the detection limits of the kits, hence the actual concentrations were not quantified. These results rejected the first hypothesis that the ecological water quality was generally bad, since several sampling sites also had good ecological water quality. However, the research hypothesis regarding the impact of nutrients on the water quality could not be tested due to unavailability of correctly measured nutrient concentrations. The results also suggested that the BMWP-Col is more suitable to assess the ecological water quality of the Guayas river basin than the NLSMI. The main reason for this is that the NLSMI is type-specific and performs relatively well to assess the ecological water quality of small rivers located at an elevation lower than 250 m above sea level. Due to intensive agricultural practice in the region, the importance of nutrient and pesticide measurements in future monitoring campaign is emphasized.

7.4 Relationship between environmental variables and ecological water quality (Ch. 4 and Ch. 5)

Questions: How is the relationship between the presence of macroinvertebrate and physico-chemical variables? What are the key environmental variables affecting the ecological water quality? What river management actions can be proposed?

➔ Multivariate analyses (chapter 4) and regression analyses (chapter 5) were used to relate the BMWP-Col and environmental variables. Both techniques showed the influence of environmental variables on the ecological water quality. Flow velocity, sludge layer, chlorophyll a concentration, sediment type, conductivity and land use showed strong influence on the presence of macroinvertebrates, based on multivariate analyses (all variables had $p < 0.001$). Whereas regression analyses selected a set of hydromorphological and chemical variables (elevation, nitrate-N and chlorophyll a concentrations, sediment angularity, presence of logs and macrophytes, flow velocity, turbidity, bank shape, and land use; $p < 0.05$ except for chlorophyll a had $p = 0.064$) as key environmental variables affecting the BMWP-Col. All these results confirmed both hypotheses that macroinvertebrate composition was highly influenced by physico-chemical variables and that agricultural-related nutrient and land use were the key environmental variables influencing the ecological water quality. It is proposed that future management actions can focus on better management of land use at different elevations, to monitor nitrate-N and chlorophyll a concentrations, macrophyte presence, sediment transport and bank stability of the rivers.

7.5 Land use effects on ecological water quality (Ch. 6)

Questions: Which type of data collection is most suitable to classify land use and to quantify its effect on the ecological water quality? Which spatial scale is most appropriate when quantifying local land use change and its effect on the ecological water quality? Which environmental variables are associated with each type of data collection?

➔ Three methods and data sources were utilized to collect land use data: field protocols to assess land use within a stretch of 100×10 m, Google maps to assess land use for a stretch of 100×100 m, and GIS data to assess land use for a stretch of 200×200 m, all for the left and right banks of the sampling sites. Regression analyses performed on each land use methods and data sources and environmental variables suggested that land use was best quantified using Google maps. It is understandable because they provided relatively updated land use data (compared to GIS data) and larger area coverage than field protocols. This result rejected the hypothesis that field observation was the best technique in quantifying local land use. Effects of local land use on the ecological water quality were best assessed using land use recorded with Google maps ($R^2 = 0.93$, $p < 0.05$), thus confirming the second hypothesis that land use within direct vicinity to sampling sites was the most influencing land use on the ecological water quality. Moreover, models involving land use assessed using Google maps were associated mainly with physico-chemical variables, whereas models involving land use assessed using field protocols and GIS data were associated mainly with hydromorphological variables. These last results overruled the hypothesis that the three methods and data sources were associated with similar environmental variables. It can be suggested to combine field observation and the use of Google maps or GIS data to gather land use information for future ecological water quality studies. The use of similar scale to quantify land use from both Google maps and GIS data is also recommended.

7.6 Methodologies

7.6.1 Data collection and reliability

Typically, regular monitoring campaigns and ecological water quality studies will collect various physico-chemical, hydromorphological, land use and biological information of the rivers (Table 7.1). The number and type of variables to be collected depend on the aim of the study and available resources. Fortunately, protocols for data collection have been written and are publicly available to support ecological water quality studies and regular monitoring campaigns. This way, consistent data collection for all sampling sites and over the years can be guaranteed. Moreover, the protocols are relatively simple and easy to follow, with explanation on variables to be measured and how to measure them (Barbour *et al.*, 1999; Bartram and Ballance, 1996; Chapman, 1996). Nevertheless, practical challenges remain to deal with detailed and practical aspects of data collection, preservation and laboratory measurements.

Table 7.1 List of common physical, chemical and hydromorphological variables measured in ecological water quality monitoring (Barbour *et al.*, 1999; Bartram and Ballance, 1996; Chapman, 1996).

Physical variables	Chemical variables	Hydromorphological variables
Temperature	Nitrate-nitrogen	Valley form
Turbidity	Nitrite-nitrogen	Channel form
Suspended solids	Ammonium-nitrogen	Stream width
Woody debris	Total nitrogen	Flow variation
	Phosphate-phosphorus	Bank profile
	Total phosphorus	Erosion
	Chloride	Sludge layer
	Dissolved oxygen	Pool/riffle class
	Chemical oxygen demand	Bank shape
	Biological oxygen demand	Bank slope
	Chlorophyll a	Substrate type
	Conductivity	Land use
	pH	Water depth
	Pesticides	Channel alteration
		Elevation
		Shading
		Flow velocity

For this research, the sampling sites and their surroundings were assessed using a standard field protocol that was adapted from the Australian River Assessment System (AUSRIVAS) physical assessment protocol (Parsons *et al.*, 2002) and the United Kingdom and the Isle of Man River Habitat Survey (RHS) (Raven *et al.*, 1998). These two protocols have been used worldwide due to their detailed and broad range questionnaires, either in their original format or adapted. A similar protocol has been used to assess rivers and reservoir; since the information required to assess the rivers are also applicable to the reservoir. The use of the adapted field protocol was proven beneficial because it gives a relatively complete overview of site's surroundings.

Nutrient measurement was done using Hach-Lange®DR 3900 spectrophotometer kits having the detection limits of 1 mg/L, 0.5 mg/L, 0.23 mg/L, 0.015 mg/L and 0.015 mg/L for total N, total P, nitrate-N, nitrite-N and ammonium-N, respectively. As can be seen from Table 4.1 original data, the actual nutrient concentrations at many sites were below the detection limits, which resulted in missing values in the data. Similar problem was encountered in COD measurement that was also done using Hach-Lange®DR 3900 spectrophotometer kits having the detection limit of 5 mg/L. Two alternative strategies were implemented to handle this problem. Firstly, for analyses in chapter 4, all variables with missing values were discarded. This has caused an exclusion of six variables from analyses. Secondly, for analyses in chapter 5 and 6, all missing values due to concentrations below detection limits were replaced by the values of the respected detection limits. This way, all nutrient variables could be included in analyses. Despite the selection of nutrient (nitrate-N in chapter 5, total P and nitrate-N in chapter 6), the reliability of the models is questionable due to the use of assumed values for the nutrients.

The study also lacked information on several variables: biological oxygen demand (BOD), COD, pesticides and total P of sediment. The samples for BOD measurement were collected, but due to a human error, none of the samples could be measured. COD concentrations were not available for all sampling sites and thus the variable was excluded from analyses. Both variables are important in determining water quality, but unfortunately the study could not evaluate that. As discussed in chapter 6, total P concentration of the sediment might be necessary in understanding the total P concentration of the water. Assumption of phosphorus saturated sediment could not be tested since total P concentration of the sediment was not measured. Lastly, there was no pesticide measurement in this study. As discussed in chapter 4, pesticide might be an important variable in determining water quality in the area. However, it could not be assessed here as well.

The type of aquatic environment also determines the choice of sampling location and frequency of monitoring. For a regular monitoring campaign to assess the trend, rivers need to be monitored at least once a month for water quality variables due to their tendency to fluctuate, whereas particulate matter and biological variables can be monitored once a year. Concentration of particulate matter also fluctuates, but its trend can be assessed annually. However, monitoring frequency of particulate matter varies based on the objectives of the assessment. The aquatic biota have relatively long life span, therefore their monitoring can be done once a year. Monitoring campaigns might also need to be done at different seasons, since several variables might change with seasons. Some variables (e.g. temperature, oxygen concentrations, pH) have diel cycles especially during summer time, thus might require more frequent observations (Bartram and Ballance, 1996; Chapman, 1996). However, limited resources might only allow a one-time survey for a finite duration but not for a long term.

For the aims of this study, one sampling campaign was conducted to collect physico-chemical and biological variables from the Guayas river basin. The previous sampling campaign was only done in one wetland area under the WETwin project (Alvarez-Mieles *et al.*, 2013; Arias-Hidalgo *et al.*, 2013), thus this was the first sampling campaign for the entire Guayas river basin. Being the first sampling campaign, there was not enough preliminary information available regarding the river basin, including the range of nutrient concentrations. Due to intensive agriculture activities in the area, it was expected that nutrient concentrations would be high, and the used detection limits of the kits would be appropriate. Apparently this was not the case, and has caused limitations to the study, as discussed above. One time sampling campaign was considered enough for the study, because the objectives of the study were intended for current conditions and the collected data could already met the objectives.

The sampling campaign was done during the dry season to be able to capture extreme environmental conditions of the water (e.g. chemical variables). Besides, practical consideration to ensure access to all sampling sites was taken. As this was the first sampling campaign in the entire river basin, the logistics were considered easier in the dry season than in the rainy season. In the rainy season, several chemical variables may be diluted (such as nitrogen compounds) or elevated (such as turbidity) by rain events that may lead to additional practical challenges . Besides, higher water level during the rainy season will also hinder macroinvertebrate sampling at several locations. Overall, it is best to wait for few days after rain events before sampling the water, especially during the rainy season (Barbour *et al.*, 1999; Chapman, 1996). As argued by Helson and Williams (2013), biomonitoring is recommended to be done during the low flow of the dry season. Moreover, Hering *et al.* (2010) suggested a sampling restriction to one season to reduce natural variability in performing the assessment. However, since the data were collected only during the dry season and there were no data available for the rainy season, seasonal difference could not be assessed (see chapter 5). Furthermore, the observed data might not reflect general conditions of the whole year.

Besides the frequency and sampling time, the choice of the sampling sites is also important. The sampling sites were selected along a disturbance gradient from less-disturbed to heavily-disturbed locations including up- and down-stream locations of the river, the confluence of rivers, the tributaries and the reservoir. This site selection allows understanding of overall condition of the Guayas river basin as a whole, not only the rivers but also the reservoir. Even though no exact proportion was allocated for different land uses (i.e. agricultural, forested and residential), all sampling sites cover enough cases of each land use type. Moreover, a relevant proportion (a quarter of total number of sites) was allocated for reservoir.

The use of different scales to quantify land use gave a limitation to select the most appropriate method and scale. Indeed the variance of the scales is not high, but the influence of scale difference might be present. As discussed in chapter 6, the scales to quantify land use from field observations and Google maps can be maintained, but both classifications can be combined. Both observations can be classified into similar categories to ease the process in combining them. The scale to quantify land use from GIS can follow the scale for Google maps observation (i.e. 100*100 m). Moreover, the GIS data might have been outdated, and the results of analyses might be misleading. When there is no updated data available, the possibly outdated land use data can be used only as a comparison to data collected from field observations or Google maps.

7.6.2 Data analysis

Since environmental processes are generally complex, appropriate analysis to understand its processes is required. Often, there are a large number of variables to deal with, and there are various ways to analyze the data. Thus, ecological models are often used to perform statistical data analysis to analyze complex environmental relationships. Model selection depends on the nature of the data and the aim of the study (Crawley, 2007; Zuur *et al.*, 2007). To define the nature of the data, Zuur *et al.* (2010) has formulated a scheme for various data exploration techniques to be able to select an appropriate model (see chapter 2 and Fig. 2.2). However, not every dataset requires each step, because not all statistical techniques require all assumptions (Zuur *et al.*, 2010). For the aims of the study, scatterplots, boxplots, Pearson correlation (for chapter 4) and variance inflation factor (VIF, for chapter 5 and 6) were used.

In this study, multivariate analysis (correspondence analysis – CA) and regression analysis (general linear model – GLM) were used. Both model types were selected because they were most appropriate for the nature of the data and the objectives of the current research. Similar studies also showed that both models were being mostly used in ecological water quality studies integrating land use data (see chapter 2). Another reason for the selection of both model types is that they are relatively easy to use in defining the relationship between a continuous (i.e. the BMWP-Col) and environmental variables (Guisan and Zimmermann, 2000; Van Echelpoel *et al.*, 2015). The CA was applied first to determine the distribution of macroinvertebrate taxa. To understand the influence of environmental variables on taxa distribution, environmental variables were then fitted on the CA graph. CA is one of multivariate techniques that has proven its ability to deal with a large number of variables (Crawley, 2007; Zuur *et al.*, 2007), as in this case. The CA successfully explained the influence of flow velocity, sediment type, land use, conductivity, chlorophyll a and sludge layer on macroinvertebrate's composition at different locations (see chapter 4).

The GLM was performed to find key environmental variables that influenced the ecological water quality of the Guayas river basin as a whole. Besides determining variables such as elevation, flow velocity, land use and chlorophyll a as the key environmental variables influencing the ecological water quality (see chapter 5), the GLM also successfully determined Google maps observation as the most suitable technique in defining land use effects on the ecological water quality (see chapter 6). The use of drop1 command in R (R-Core-Team, 2013) was proven beneficial since it provides the AIC as a measure of model fitness that has helped in model selection process, while variable selection could be done based on the p-values of variables in the model. As explained by Tuffery (2011), data mining software such as R software combines inferential and predictive statistics, and thus provide more complex analyses. In this study, the GLM might not be able to evaluate the influence of nitrate-N and total P in the model. This problem was related to data reliability, as already discussed in previous section.

Various physico-chemical and hydromorphological variables were collected, which result in a large number of variables to be analyzed, and the selection of GLM that can deal with many variables. Since several variables were correlated, removal of correlated variables was necessary to reduce the complexity of the model. The cut off value of 3 based on VIF was used. However, the total number of variables to be included in the model was still large. To reduce model's complexity, expert knowledge can be applied to select variables to be included in the model, after VIF analysis. As already discussed in chapter 5, 6 and section 7.6.1, the limitation of the GLM was mainly due to data reliability. Therefore, the GLM was deemed appropriate to handle data with many variables as in this study, compared to other techniques (see chapter 5).

Despite the use of two different model types, the results corroborated each other, especially in selecting land use, flow velocity and chlorophyll a as key environmental variables (see chapter 4 and 5). For management purposes, confirming results from different techniques are beneficial in deciding which variables to focus on. The results suggested the robust use of multivariate and regression analyses in ecological water quality studies and the possibility of using different model types on the current dataset and study aims. Besides, both model types are relatively easy to use and therefore can be easily applied in developing countries. Moreover, both multivariate and regression analyses were performed in R, free software that is accessible in developing countries.

7.6.3 Use of macroinvertebrates

7.6.3.1 Macroinvertebrate sampling and identification

Macroinvertebrates are the most used indicator biota in biological monitoring and ecological water quality studies. They consist of long-lived diverse communities and habits and are easy to sample and identify (Bartram and Ballance, 1996; Chapman, 1996; De Pauw *et al.*, 2006). Besides, due to their long life-cycles, sampling at different seasons is not required and less frequent sampling such as once a year macroinvertebrate monitoring is sufficient (Barbour *et al.*, 1999; Chapman, 1996; De Pauw *et al.*, 2006). For the aims of this study, macroinvertebrate sampling was performed together with physico-chemical sampling during the dry season (see section 7.6.1 Data collection and reliability). Jacobsen and Encalada (1998) and

Jacobsen *et al.* (1997) also reported that taxa richness and abundance are higher in the dry season than in the rainy season. Therefore, more taxa could be observed in the dry season. Further, sampling during the dry season will provide information of worse conditions of the water when water quality variables are generally reaching their peak values. Thus, the influence of extreme values of water quality variables on the macroinvertebrates can be assessed.

Macroinvertebrate sampling has to be done in a proper way to ensure good representation of macroinvertebrate communities that are present in the water. During the sampling campaign, the macroinvertebrates were sampled following the procedure described by Gabriels *et al.* (2010) by collecting macroinvertebrates from all present habitats at the sampling sites (Barbour *et al.*, 1999; Gabriels *et al.*, 2010). Macroinvertebrate identification was done until family level, using established identification keys for European streams (De Pauw *et al.*, 1996) and South American streams (Domínguez and Fernández, 2009). As discussed by Chapman (1996), Barbour *et al.* (1999) and further reviewed and confirmed by Bailey *et al.* (2001) and Marshall *et al.* (2006), macroinvertebrate identification up to family level is sufficient for biotic index calculation and multivariate analyses, as in this study. Moreover, all procedure was done consistently for all samples (Barbour *et al.*, 1999).

7.6.3.2 Biotic index

The last thing to consider when using macroinvertebrate in ecological water quality studies is applying a suitable biotic index. Since Ecuador does not have its own biotic index, two available biotic indices were selected: the BMWP-Col (Alvarez, 2005) and the NLSMI (Helson and Williams, 2013) to calculate the ecological water quality of the Guayas river basin. As explained in section 3.3.1, the BMWP-Col was selected since it contained most of the encountered taxa; while the NLSMI was selected since it was developed in a country having a relatively similar environmental climate to Ecuador. The performance of the BMWP has been proven in other regions, and the use of its adapted version (the BMWP-Col) was deemed suitable. However, the NLSMI performed well only when it was applied on small streams at low elevations, but did not perform well on large streams or at higher elevations. Despite the fact that both indices were developed in neighboring countries with relatively similar environmental conditions with Ecuador, the NLSMI is proven to be highly type specific and therefore was not applicable for the whole Guayas river basin. Since the BMWP-Col performed relatively well and its suitable use was further confirmed by the ASPT, thus it was used for further analyses of the dataset (see chapter 4). These results confirmed the necessity of careful consideration for selecting the most suitable index to be used in a country or region away from the origin where the index was developed. These results also confirmed the robust applicability of the BMWP-Col for Ecuadorian rivers. Overall, the results confirmed the possibility of using available biotic indices to calculate ecological water quality of a water body instead of developing a new index. Indeed, when resources permit, developing a new biotic index specifically for a water body, a region or a country will be beneficial for the water body or the country. Especially since this study encountered several taxa but they were not included in the calculation of the BMWP-Col, e.g. Corbiculidae, Gerridae (Alvarez, 2005; Roldán Pérez, 2003) and Acari (that are not included in many other indices as well (Goldschmidt, 2016)). However, developing a new biotic index will require much effort, time and resources.

7.6.4 Recommendation

7.6.4.1 Take home message

The relationship between land use and macroinvertebrate communities in the Guayas river basin can be seen from the occurrence of macroinvertebrate taxa at different land use types (Fig. 7.1). Here, only the most prevalent taxa are shown. Several taxa (i.e. Baetidae, Chironomidae and Coenagrionidae) were present in all types of land use, while sensitive taxa (e.g. Leptoceridae and Leptophlebiidae) were present where the land use supports good ecological water quality (i.e. forest). The presence of Acari (also present in Chaguana river basin, Ecuador (Dominguez-Granda, 2007), in Costa Rica and Panama (Goldschmidt, 2016)), but not reported to be present in the Gilgel Gibe watershed, Southwest Ethiopia (Mereta *et al.*, 2012) indicated the difference in macroinvertebrate assemblage between Ecuador and other countries from different continent such as Ethiopia, which indicates Ecuador biodiversity.

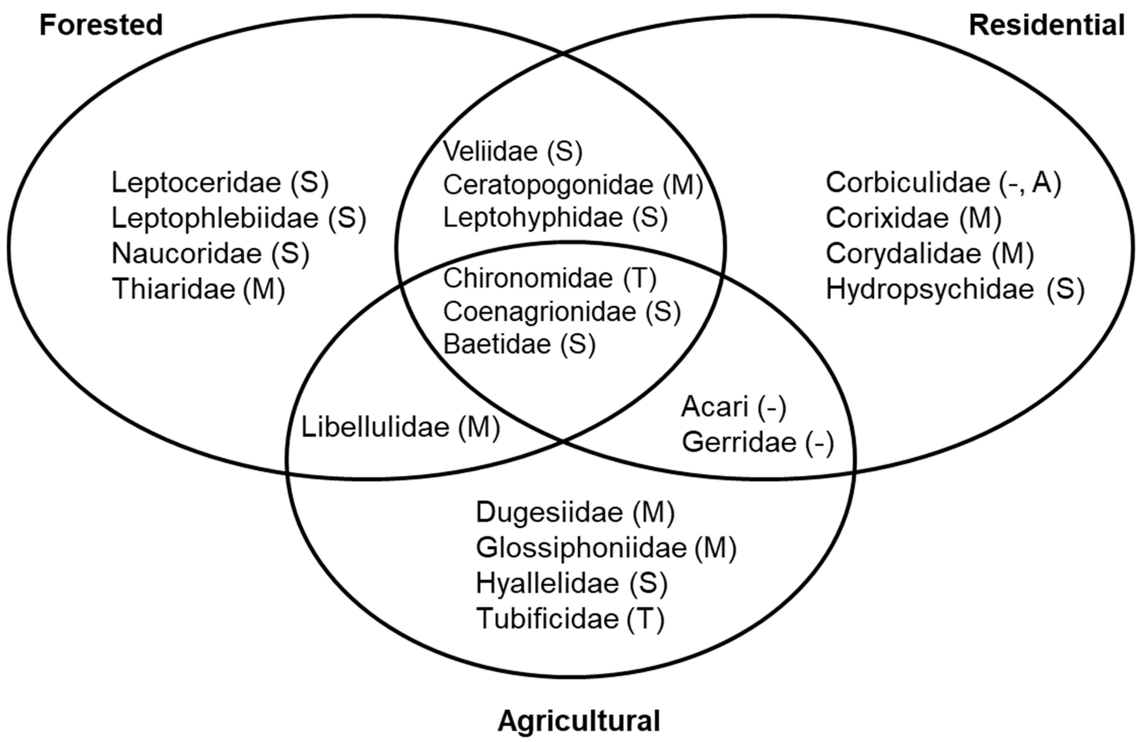


Figure 7.1 Land use types and the presence of most prevalent macroinvertebrate taxa observed in the Guayas river basin, Ecuador; A: alien taxa, M: moderate sensitivity, S: sensitive, T: tolerant, (-): no assigned tolerance value.

Further, the shift in taxa presence among land uses shows their preference toward a certain land use (Table 7.2 shows examples of taxa presences at different land uses observed in the Guayas river basin). The presence of sensitive taxa associated with forested land use can already indicate a good water quality, whereas an abundant presence of tolerant taxa associated with agricultural and residential areas can already indicate a bad ecological water quality. To maintain a good ecological water quality, the presence of sensitive macroinvertebrate taxa in the Guayas river basin need to be accommodated by managing the land use surrounding the water bodies (see section 5.4.1, 6.4 and specifically section 7.6.4.2 as recommendations for ecological water quality preservation).

Table 7.1 The presence of macroinvertebrate taxa at several sampling sites in the Guayas river basin; TS: tolerance score calculated for the BMWP-Col based on Alvarez (2005).

Taxa		FP land use												Google land use												GIS land use					
		Forest			Arable			Orchard			Residential			Forest			Arable			Pasture			Residential			Agriculture			All others		
	Site	35	16	27	2	3	26	1	32	79	11	20	65	35	16	19	26	85	111	119	81	47	20	11	65	49	54	97	98	75	16
	TS																														
Acari	-				p	p				p								p			p	p					p	p	p	p	
Aeshnidae	6			p														p												p	
Ampullariidae	6								p																						
Ancylidae	7						p										p	p													
Baetidae	7	p		p						p	p			p					p	p	p	p		p		p	p				
Belostomatidae	4																														
Blepharoceridae	10																														
Caenidae	6									p								p			p										
Calamoceratidae	8																														
Calopterygidae	7	p					p							p			p														
Cambaridae	-																	p													
Ceratopogonidae	5			p																		p					p				
Chaoboridae	3																														
Chironomidae	2	p	p	p	p	p	p	p	p	p	p		p	p	p	p	p	p	p	p	p	p		p	p	p	p	p	p	p	p
Coenagrionidae	7	p		p			p						p	p			p	p	p		p				p	p				p	
Corbiculidae	-			p	p						p		p						p	p				p	p				p		
Corixidae	5	p	p		p									p	p							p							p		p
Corydalidae	6		p		p										p	p				p								p			p
Coryphoridae	9		p		p										p	p															p
Crambidae	7	p	p		p		p							p	p		p			p								p			p
Culicidae	2	p		p										p								p				p					
Dixidae	7	p												p																	
Dryopidae	6																														
Dugesiidae	6									p						p		p		p	p									p	

Taxa		FP land use												Google land use												GIS land use					
		Forest			Arable			Orchard			Residential			Forest			Arable			Pasture			Residential			Agriculture			All others		
	Site	35	16	27	2	3	26	1	32	79	11	20	65	35	16	19	26	85	111	119	81	47	20	11	65	49	54	97	98	75	16
	TS																														
Dytiscidae	-				p																	p				p				p	
Elmidae	6		p		p										p	p															p
Empididae	4		p												p	p															p
Gerridae	-			p						p									p		p									p	
Glossiphoniidae	5			p								p				p		p			p		p								
Glossosomatidae	7																			p											
Gomphidae	9	p				p	p							p			p														
Gyrinidae	5																														
Hebridae	8																														
Helicopsychidae	8	p	p											p	p																p
Heteroceridae	-																														
Hyallelidae	7									p								p			p									p	
Hydrobiidae	7	p												p		p															
Hydrobioscidae	9																														
Hydrometridae	4																														
Hydrophilidae	3			p																		p					p				
Hydropsychidae	7	p	p		p									p	p	p				p								p	p		p
Hydroptilidae	8															p		p		p								p	p		
Lampyridae	10			p																											
Leptoceridae	8	p			p									p						p							p	p	p		
Leptohyphidae	7	p	p	p	p		p						p	p	p	p	p		p	p					p			p			p
Leptophlebiidae	9	p	p		p	p								p	p	p				p							p	p			p
Libellulidae	5	p		p	p					p				p				p	p	p	p	p				p	p		p	p	
Limoniidae	3			p																									p		
Lumbriculidae	-																														
Lymnaeidae	8	p		p										p							p										

Taxa		FP land use												Google land use												GIS land use					
		Forest			Arable			Orchard			Residential			Forest			Arable			Pasture			Residential			Agriculture			All others		
	Site	35	16	27	2	3	26	1	32	79	11	20	65	35	16	19	26	85	111	119	81	47	20	11	65	49	54	97	98	75	16
	TS																														
Macroveliidae	-	p												p																	
Megapodagrionidae	6						p										p			p								p			
Mesoveliidae	5												p												p		p			p	
Mysidae	-																									p					
Naucoridae	8	p	p							p			p	p	p	p				p	p				p				p		p
Chordodidae	10					p																									
Nereidae	-																														
Noteridae	4								p																						
Notonectidae	5									p												p								p	
Ocypodidae	-											p											p								
Odontoceridae	10	p												p																	
Oligoneuriidae	9																														
Palaemonidae	8																									p					
Perlidae	10		p												p	p												p			p
Philopotamidae	9		p		p	p									p	p				p								p	p		p
Physidae	3	p		p										p								p									
Planorbidae	8	p						p						p																p	
Platystictidae	9					p																									
Pleidae	6																					p									
Polycentropodidae	9																			p									p		
Polymitarcidae	10															p															
Psephenidae	10	p	p											p	p	p															p
Ptilodactylidae	10		p												p																p
Scirtidae	4																														
Simuliidae	7				p															p								p			
Sphaeriidae	8								p																						

Taxa		FP land use												Google land use												GIS land use					
		Forest			Arable			Orchard			Residential			Forest			Arable			Pasture			Residential			Agriculture			All others		
	Site	35	16	27	2	3	26	1	32	79	11	20	65	35	16	19	26	85	111	119	81	47	20	11	65	49	54	97	98	75	16
	TS																														
Staphylinidae	6																											p			
Stratiomyidae	3	p		p										p																	
Tabanidae	5		p												p							p									p
Thiaridae	5	p		p				p	p			p	p	p					p				p		p	p		p	p		
Trichodactylidae	-																														
Tubificidae	1		p						p		p				p	p		p			p			p		p					p
Veliidae	7	p	p	p	p		p							p	p	p	p					p					p	p			p

7.6.4.2 Recommendation for ecological water quality preservation and future studies

The results showed a strong association between macroinvertebrates and environmental variables. The ecological water quality decreased when environmental conditions deteriorated, and vice versa. These results confirmed the need of ecological water quality preservation in order to maintain aquatic ecosystem services in the future and to protect Ecuador biodiversity. Land use was considered the main cause of environmental disturbances occurring in the water (Garnier *et al.*, 2013; Pilgrim *et al.*, 2014; Tu, 2009).

The same situation occurred in the Guayas river basin, where local land use was selected as a key variable affecting the ecological water quality. Other selected key variables such as flow velocity, sediment type, sludge layer, turbidity, conductivity, chlorophyll a and nitrate-N concentrations of the water are generally influenced by land use occurring surrounding the water bodies. It is also understandable that the ecological water quality is deteriorating due to intensive agricultural activities within the Guayas river basin.

Thus, the first recommendation is to regulate agricultural activities throughout the Guayas river basin. Intensive agricultural activities such as banana and rice production together with cattle farming have been taking place and occupy the land until the edge of the rivers (Arias-Hidalgo *et al.*, 2013; Flood, 2000; Seo *et al.*, 2010). Surface runoff and cattle movements fasten nutrient and sediment transport from the land into the water, and without proper management, ecological water quality will deteriorate even faster. Agricultural activities can be managed by regulating the use of fertilizers as one source of nutrient enrichment in the water. The use of pesticides also needs to be regulated (will be discussed further in recommendation for pesticides). Besides, replacing intensive and continuous grazing into rotational grazing can be a good alternative for cattle farming in the Guayas river basin. Rotational grazing can lower the negative impacts of grazing animals on the ecological water quality and provides shade to the water (Raymond and Vondracek, 2011) (see chapter 6).

The second recommendation to local government is to use buffer zones at the riparian area. Since buffer zones are defined and protected by laws or set to maintain riparian vegetation (Crétaz and Barten, 2007), local government can manage the zones as required; therefore buffer zones can lower the impacts of land use surrounding the water bodies. The size of riparian buffer can vary from 10 to 90 m wide from the water body (Crétaz and Barten, 2007; Hansen *et al.*, 2010) or even wider, depending on the need and land availability. Local government might need to compensate farmers and local residents in acquiring land for the buffer zones, which might not be affordable for the Guayas' governmental budget. To solve this problem, discussion among government, residents and other stakeholders could help achieve this project with less cost than what was originally foreseen. However, cost analysis will not be addressed here.

Thirdly, local government needs to monitor the key variables selected within this PhD study. Since the sampling campaign performed for this study was the first sampling campaign in the entire Guayas river basin, the data collected for this study can be used as the baseline data. The regular monitoring campaign needs to collect physico-chemical, hydromorphological and biological variables simultaneously, following the methodologies applied within this study. As already discussed in section 7.6.1, the use of kits having lower detection limits for nutrient compounds (i.e. total P, total N, nitrate-N, nitrite-N, ammonium-N) is endorsed. When a new location outside the Guayas river basin is going to be sampled, it is recommended to check available information on the range of physico-chemical concentrations in order to use appropriate detection limits. The measurement of total P concentration from the sediment is also recommended. Further, measurement of BOD and COD is endorsed.

Performing monitoring campaign during the rainy season besides the dry season is also suggested. Since the sampling campaign in this study was done during the dry season, the seasonal difference of ecological water quality cannot be studied. Due to safety and accessibility reasons, indeed it might not be possible to sample all sampling sites as in the dry season. However, it will provide useful information regarding macroinvertebrate's ability in dealing with different environmental conditions and the extent of environmental difference between the rainy and the dry seasons. Since the sampling campaign performed for this study was done under the VLIR Ecuador Biodiversity Network project (the project is still ongoing), it is important to continue the monitoring campaign beyond the project's duration. Also, the monitoring campaign needs to be done at all sites that were sampled in this study. When resources permit, it will be beneficial to enlarge/add monitored locations and sites. The regular monitoring can be used to determine how effective the applied activities are in preserving the ecological water quality. It is possible that after some time, environmental variables that need to be monitored are changed due to improved environmental conditions. The regular monitoring can then be adjusted based on the outcome of each monitoring campaign.

Next recommendation is the application of urban best management practices (BMPs) (Palmer *et al.*, 2014; Palmer *et al.*, 2010; Smucker and Detenbeck, 2014; Walsh *et al.*, 2004). Residential was defined to have influenced the ecological water quality most negatively, compared to other land uses. Urban BMPs include provision of wastewater treatment facilities and detention ponds for storm water management. To date, the Guayas river basin does not have wastewater treatment facility yet, which means that wastewater is discharged directly to the rivers without being treated. By treating the wastewater prior to being discharged into the rivers will reduce nutrient concentrations and other pollutants from entering the water (Von Sperling and Chernicharo, 2002; Younes-Baraille *et al.*, 2005). Whereas the detention ponds can reduce peak discharges during storm or heavy rain events. Peak discharge attenuation will reduce erosion and sediment transport into the water which will consequently protect aquatic habitat quality (Crétaz and Barten, 2007).

For future studies, the use of combined land use observation (field and either Google maps or GIS data observations) to define land use effect on the ecological water quality is suggested. The use of Google maps or GIS data can enlarge area coverage of land use data, thus provides more information of the land use surrounding the rivers. Land use data obtained from field observations and Google maps can be combined into one classification, thus both observations can be classified using similar categories to ease their combination. The land use data obtained from field observation can also be used as a separate validation data in analysis. Furthermore, since this PhD work only studied the effects of local land use, the effects of catchment or regional land use could not be defined. Therefore, if catchment or regional land use is available, it is recommended to use both local and catchment or regional land use scales to determine the extent of land use effects on the ecological water quality.

The use of the BMWP-Col to calculate the ecological water quality studies of the Guayas river basin is recommended, when a specific index for the area is not yet available. Future studies can also analyze the reservoir, the up- and down-stream parts of the rivers separately. This was not done in this study (see chapter 5), but it is possible that taxa composition would change in different seasons and due to land use change. In case missing values cannot be avoided in the future and completing them is necessary, available methods in handling missing data can be used.

Lastly, monitoring pesticides concentration of the water is suggested (see chapter 4), which was due to limited resources, could not be measured for this PhD study. However, possible influence of pesticides on the aquatic macroinvertebrates is acknowledged. Since the Guayas river basin has used a large amount of pesticides (number 14 of the world's largest intensive pesticide users) for agricultural purposes (Caceres *et al.*, 2002; FAO, 2011; Horgan *et al.*, 2014; Matamoros, 2004), it is likely that pesticides have also influenced the ecological water quality. Especially because pesticides (26 different pesticides) have been observed during 2016 sampling campaign at the Guayas river basin (Deknock, 2017), thus pesticides monitoring is clearly important in future ecological water quality monitoring programs.

7.7 General conclusions

This PhD study evaluated land use effects on the ecological water quality of the Guayas river basin, as a case study of a developing country facing intensive agricultural and urbanization activities. Being the first sampling campaign performed in the entire Guayas river basin, this study stands as a starting point for future ecological water quality monitoring in the area.

The ecological water quality decreases along a gradient of anthropogenic disturbance and elevations (Fig. 7.2 and 7.3). Land use was selected as a key variable affecting the ecological water quality, and its influence can be linked with other key environmental variables that are generally related to land use (e.g. nutrients and hydromorphological variables). As expected, forested areas are associated with a good water quality, whereas residential and agricultural areas are associated with a bad water quality (Fig. 7.4). Since the actual concentrations of nutrients were not available in this study (values of the detection limits were used instead), nutrient measurement is highly important in future monitoring to determine the actual influence of nutrients on the ecological water quality. Especially since total

P concentration was much higher than the guideline of the government, regulating its concentration in the future is important.

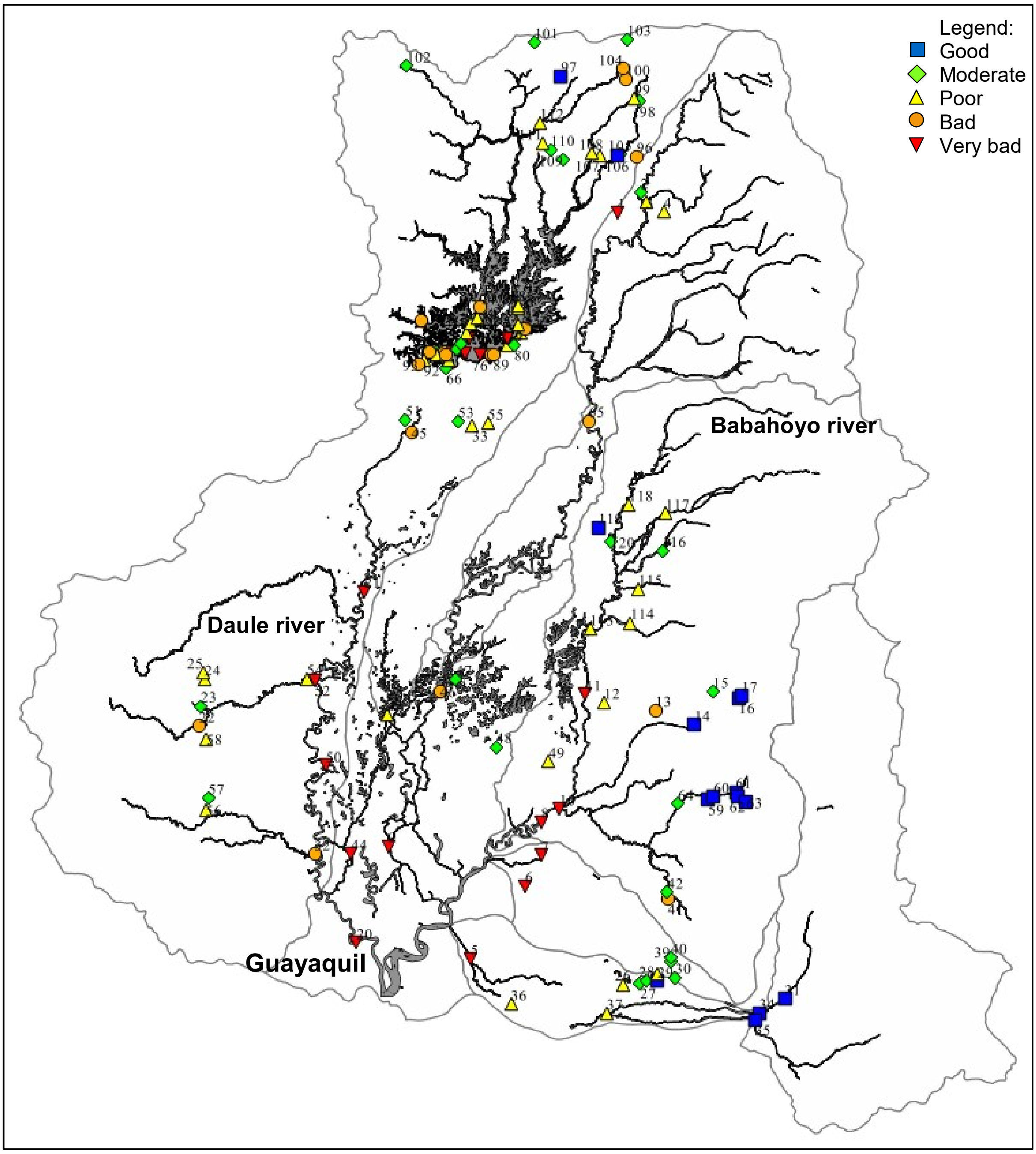


Figure 7.2 Sampling sites in the Guayas river basin with indication of the ecological water quality based on the BMWP-Col ranging from good to bad, as shown in the legend.

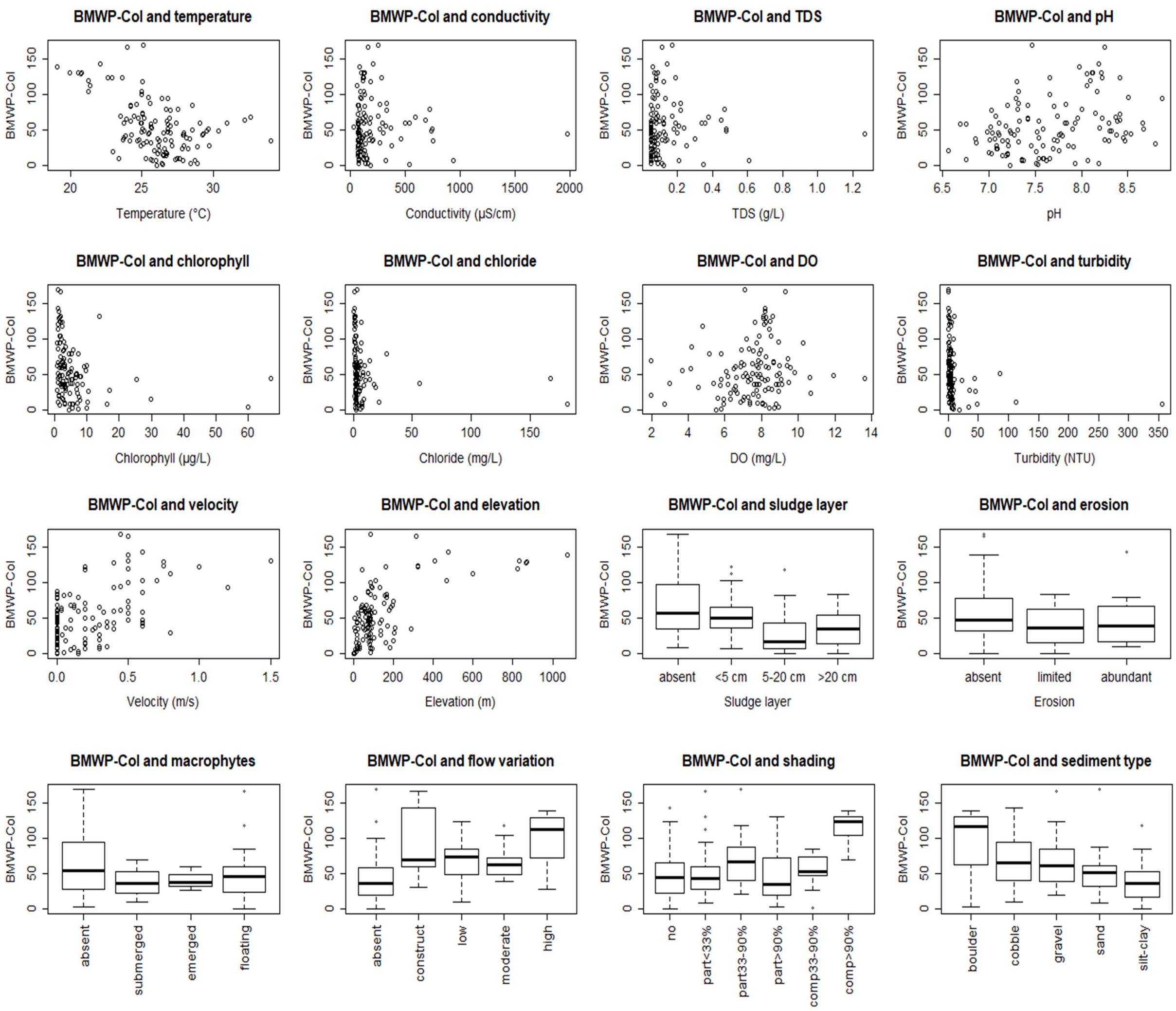


Figure 7.3 Plots showing the distribution of the data for physico-chemical variables in relation to BMWP-Col for 120 sampling sites in the Guayas river basin, classification of categorical variables are based on Table B1; part: partly, comp: completely.

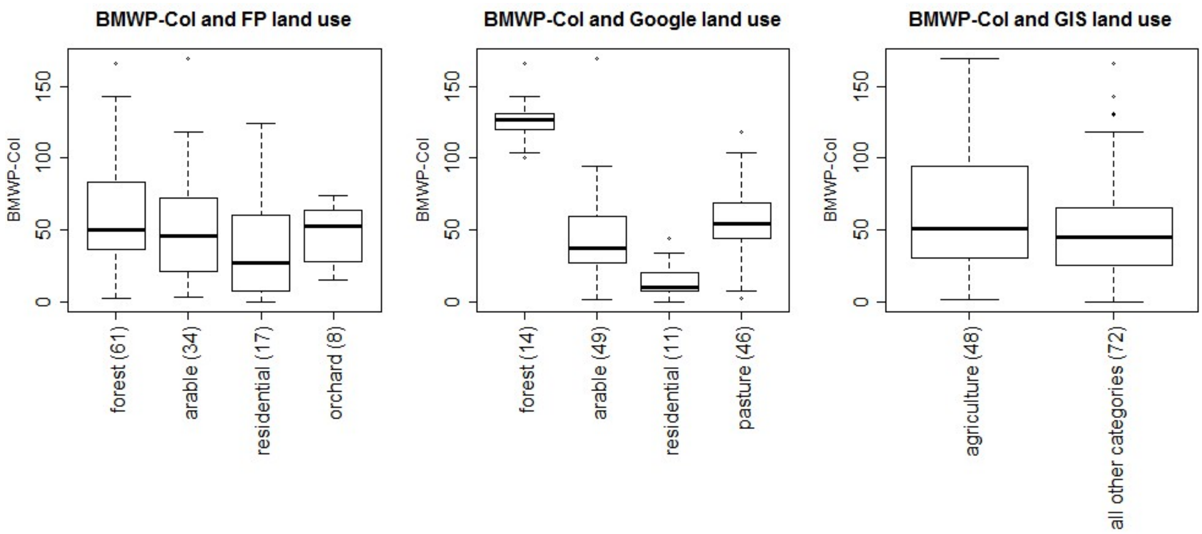


Figure 7.4 Boxplots showing the ecological water quality of 120 sampling sites in the Guayas river basin in relation to three land use assessment methods and sources, number of observations is shown in brackets.

The relationship between land use and the ecological water quality was better understood when various methods and sources of land use data collection was utilized; than using mainly field observational data (Fig. 7.5). The combination of observations via field, remote sensing and other sources in obtaining land use information at local or riparian scale was a novel approach for future monitoring campaigns. Furthermore, the use of macroinvertebrate data and biotic index such as the BMWP-Col together with environmental variables to assess the ecological water quality was beneficial and can be performed as a standard monitoring procedure in the future. Last but not least, ecological models are fast scientific techniques to define the relationship between land use and the ecological water quality.

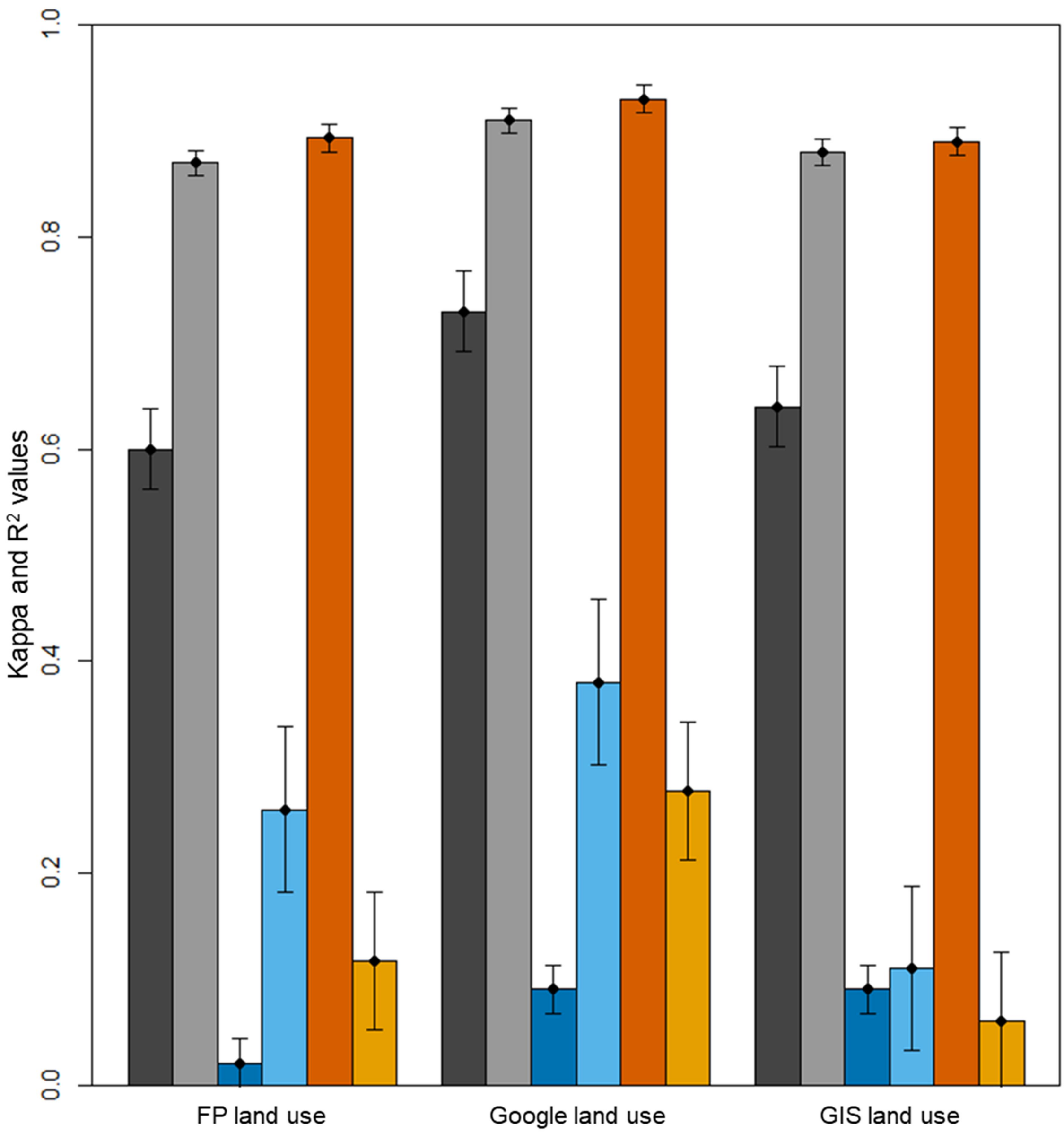


Figure 7.5 Average model performances based on the three-fold cross validation of the three land use sets (based on Table E11) with their standard errors; ▬ (unweighted Kappa training), ▬ (weighted Kappa training), ▬ (unweighted Kappa testing), ▬ (weighted Kappa testing), ▬ (R^2 training), ▬ (R^2 testing).

A – Supporting information for chapter 2

Table

Table A1 Negative effects of anthropogenic activities on different aspects of aquatic ecosystems, adapted from Carr and Neary (2008).

Impacts	Activities					
	Agriculture	Urban	Forestry	Hydropower generation and water storage	Mining	Industries
Sedimentation	√	√	√	√	√	√
Eutrophication	√	√	√	√	√	√
Thermal pollution	√	√	√	√	√	√
Dissolved oxygen		√		√	√	√
Acidification					√	√
Microbial contamination	√	√				
Salinization	√	√				√
Metal pollution	√	√		√	√	√
Bio toxins					√	√
Organic compounds	√	√	√			√
Micronutrient depletion				√		

Table A2 Countries of studies, spatial scale and temporal aspects of land use data in the ecological water quality studies.

Country		Spatial scale			Temporal	Scenario
Developed	Developing	Local or riparian	Catchment/ regional	Combined		
31	8	21	7	11	2	5

Table A3 Observation methods utilized to acquire land use data, based on the selected papers; RS: remote sensing, obs: observation.

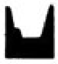
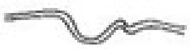
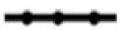
RS	Field obs.	GIS	National database	National data + GIS / RS	National data + field obs.	Field obs. + RS/GIS	RS + GIS
10	5	9	6	4	2	2	1

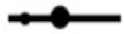
B – Supporting information for chapter 3

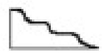
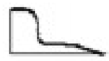
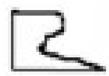
Table

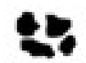
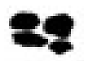
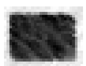
Table B1 Definition of categorical variables assessed in 120 sampling sites, modified from AUSRIVAS (Parsons *et al.*, 2002) and RHS (Raven *et al.*, 1998).

No	Variables	Categories	Definition
1	FP land use	1.forest	land covered by high density of trees, includes primary, secondary and tertiary forests.
		2.arable	land used for agriculture or farm (eg. Maize)
		3.residential	land used for residential houses
		4.orchard	land used for fruits production (eg. Cacao, banana, mango)
2	Shading	0.No shading	no shading at the sampling sites
		1.partly shaded, limited stretch < 33 %	less than 33 % of the sampling site is partly shaded
		2.partly shaded, longer stretch 33-90 %	About 33-90 % of the sampling site is partly shaded
		3.partly shaded, whole stretch > 90 %	More than 90 % of the sampling site is partly shaded
		4.completely shaded, limited stretch < 33 %	less than 33 % of the sampling site is completely shaded
		5.completely shaded, longer stretch 33-90 %	About 33-90 % of the sampling site is completely shaded
		6.completely shaded, whole stretch > 90 %	More than 90 % of the sampling site is completely shaded
3	Type of macrophytes cover[a]	0.No macrophyte	macrophytes are absent
		1.Interrupted	macrophytes are not sharing a common border at more than one intersection
		2.Contigous	macrophytes are sharing a common border at more than one intersection
4	Main macrophytes	0.absent	macrophytes are not present
		1.submerged macrophytes	Macrophytes rooted in the bottom substrate with vegetative parts predominantly immerse

No	Variables	Categories	Definition
		2.emerged macrophytes	Macrophytes rooted in the bottom substrate with vegetative parts emerging above the water surface
		3.floating macrophytes	macrophytes with roots, if present, hang on water surface
5	Valley form	1.Canyon	
		2.V-shaped valley	
		3.Trough	
		4.Meander valley	
		5.U-shaped valley	
		6.Plain floodplain	
		7.no bank	macroinvertebrates were collected from macrophytes, away from the bank
6	Channel form	1.Meandering	
		2.Braided	
		3.Anabranching	
		4.Sinuate	
		5.Constrained (natural)	
		6.Constrained (artificial)	
		7.no bank	macroinvertebrates were collected from macrophytes, away from the bank
7	Variation in width	0	data collected at the reservoir
		1	
		2	
		3	

No	Variables	Categories	Definition
		4	
		5	
8	Extent of erosion	0.absent	erosion is not present
		1.limited	less than 30 % is eroded
		2.abundant	more than 30 % is eroded
9	Bank profile	1.Vertical	
		2.steep (> 45°)	
		3.gradually not trampled	
		4.composite not trampled	
		5.no bank	macroinvertebrates were collected from macrophytes, away from the bank
10	Variation in flow	0.absent	no variation in flow
		1.at human constructions	flow is varied at human constructions
		2.low	variation in flow is less than 20 %
		3.moderate	variation in flow is about 20-50 %
		4.high	variation in flow is more than 50 %
11	Sludge layer	0.absent	sludge layer is absent
		1.< 5 cm	sludge is accumulated for less than 5 cm
		2.5-20 cm	sludge is accumulated about 5-20 cm
		3.> 20 cm	sludge is accumulated for more than 5 cm
	Dead wood		similar categories and definition for twigs, branch, logs
12	- twigs d < 3cm	0.Absent	dead wood is not present
13	- branch 3-30 cm	1.Limited	presence of dead wood is less than 5 %
14	- logs d > 30 cm	2.Abundant	presence of dead wood is more than 5 %

No	Variables	Categories	Definition
15	Pool/Riffle class	1.Class 1	Pool-riffle pattern is (nearly) pristine: extensive sequences of pools and riffles
		2.Class 2	Pool-riffle pattern is well developed: high variety in pools and riffles
		3.Class 3	Pool-riffle pattern is moderately developed: variety in pools and riffles but locally
		4.Class 4	Pool-riffle pattern is poorly developed: low variety in pools and riffles
		5.Class 5	Pool-riffle pattern is absent: uniform pool-riffle pattern
		6.Class 6	Pool-riffle pattern is absent due to structural changes: uniform pool-riffle pattern due to reinforced bank and bed structures
16	Bank shape	0.no bank	macroinvertebrates were collected from macrophytes, away from the bank
		1.concave	
		2.convex	
		3.stepped	
		4.wide lower bench	
		5.undercut	
17	Bank slope	0.no bank	macroinvertebrates were collected from macrophytes, away from the bank
		1.vertical	80-90° bank sloping
		2.steep	60-80° bank sloping
		3.moderate	30-60° bank sloping
		4.low	10-30° bank sloping
		5.flat	less than 10° bank sloping
18	Bed compaction	0.invisible	bed is not visible
		1.tightly packed	array of sediment sizes overlapping, tightly packed and very hard to dislodge

No	Variables	Categories	Definition
		2.packed	array of sediment sizes overlapping, tightly packed but can be dislodged moderately
		3.moderate compaction	array of sediment sizes little overlapping, some packing but can be dislodged moderately
		4.low compaction (1)	limited range of sediment sizes, little overlapping, some packing and structure but can be dislodged very easily
		5.low compaction (2)	loose array of fine sediments, no overlapping, no packing and structure, and can be dislodged very easily
19	Sediment matrix	1.Bedrock	formation of bedrock
		2.Open framework	0-5 % fine sediment, high availability of interstitial spaces
		3.Matrix filled contact	5-32 % fine sediment, moderate availability of interstitial spaces
		4.Framework dilated	32-60 % fine sediment, low availability of interstitial spaces
		5.matrix dominated	more than 60 % fine sediment, interstitial spaces virtually absent
20	Sediment angularity	1.very angular	
		2.angular	
		3.sub-angular	
		4.rounded	
		5.well rounded	
		6.cobble, pebble and gravel fractions not present	
21	Main sediment type	1.boulder	sediment composed of substrates with diameter larger than 256 mm
		2.cobble	sediment composed of substrates with diameter about 64-256 mm
		3.gravel	sediment composed of substrates with diameter about 2-64 mm
		4.sand	sediment composed of substrates with diameter about 0.062-2 mm

No	Variables	Categories	Definition
		5.silt & clay	sediment composed of substrates with diameter about 0.24-62 μm
22	Depth class	1.0-10 cm	Water depth is 0-10 cm
		2.10-50 cm	Water depth is 10-50 cm
		3.50-100 cm	Water depth is 50-100 cm
		4. > 100 cm	Water depth is > 100 cm

Table B2 Classification of land use assessment methods and source assessed at 120 sampling sites.

Variables	Categories	Definition
FP land use	forest	land covered by a high density of trees, includes primary, secondary and tertiary forests
	arable	land used for agriculture or farm (e.g., maize)
	residential	land used for residential houses
	orchard	land used for fruits production (e.g., cacao, banana, mango)
Google land use	forest	land covered by a high density of trees, includes primary, secondary and tertiary forests
	arable	land used for agriculture or farm (e.g., maize)
	residential	land used for residential houses
	pasture	land covered by grass and used for livestock
GIS land use	conservation and protection	government's conserved and protected land, covered by high density of trees (forest)
	livestock-conservation and protection	combined government's conserved and protected land and livestock
	agriculture	land used for agriculture or farm and fruits (eg. maize, cacao, banana)
	mix agroforestry	land used for agriculture, livestock and forest
	mix uses	land used for agriculture, livestock, forest and residential
	residential	land used for residential houses
	livestock	land used for livestock

Table B3 Observed land use category of sampling sites based on the three methods and sources; agric: agriculture, l-stock: livestock, for: forest, urb: urban, conserv: conservation, prot: protection.

Sampling site	FP land use	Google Land use	GIS land use 7 classes	GIS land use 2 classes
1	orchard	arable	agriculture	agriculture
2	arable	pasture	agriculture	agriculture
3	arable	pasture	agriculture	agriculture
4	arable	arable	agric., l-stock, for., urb.	all other categories
5	residential	residential	agriculture	agriculture
6	residential	arable	agriculture	agriculture
7	residential	residential	agriculture	agriculture
8	arable	arable	agriculture	agriculture
9	residential	arable	agriculture	agriculture
10	residential	residential	anthropogenic, urban	all other categories
11	residential	residential	anthropogenic, urban	all other categories
12	arable	arable	livestock	all other categories
13	forest	arable	agriculture	agriculture
14	residential	forest	agriculture	agriculture
15	arable	forest	agriculture	agriculture
16	forest	forest	agric., l-stock, for., urb.	all other categories
17	forest	forest	agric., l-stock, for., urb.	all other categories
18	forest	arable	l-stock, conserv., prot.	all other categories
19	forest	arable	agric., l-stock, for., urb.	all other categories
20	residential	residential	anthropogenic, urban	all other categories
21	arable	residential	anthropogenic, urban	all other categories
22	arable	arable	agric., l-stock, for.	all other categories
23	forest	pasture	agriculture	agriculture
24	forest	pasture	agric., l-stock, for.	all other categories
25	residential	pasture	agric., l-stock, for.	all other categories
26	arable	arable	agriculture	agriculture
27	forest	arable	agriculture	agriculture
28	forest	arable	agriculture	agriculture
29	arable	arable	agriculture	agriculture
30	arable	arable	agriculture	agriculture
31	forest	forest	conserv., prot.	all other categories
32	orchard	arable	agriculture	agriculture
33	arable	pasture	agric., l-stock, for., urb.	all other categories
34	forest	forest	agriculture	agriculture
35	forest	forest	agric., l-stock, for., urb.	all other categories
36	residential	arable	agriculture	agriculture
37	forest	arable	agriculture	agriculture
38	arable	arable	agriculture	agriculture
39	residential	arable	agriculture	agriculture
40	forest	arable	agriculture	agriculture
41	arable	arable	agriculture	agriculture

Sampling site	FP land use	Google Land use	GIS land use 7 classes	GIS land use 2 classes
42	residential	arable	agriculture	agriculture
43	residential	arable	agric., l-stock, for., urb.	all other categories
44	arable	arable	agriculture	agriculture
45	arable	residential	agric., l-stock, for., urb.	all other categories
46	arable	residential	agric., l-stock, for., urb.	all other categories
47	arable	pasture	agric., l-stock, for., urb.	all other categories
48	arable	pasture	agric., l-stock, for., urb.	all other categories
49	arable	arable	agriculture	agriculture
50	arable	residential	agriculture	agriculture
51	forest	pasture	livestock	all other categories
52	residential	arable	agric., l-stock, for., urb.	all other categories
53	forest	pasture	agric., l-stock, for., urb.	all other categories
54	arable	arable	agriculture	agriculture
55	forest	pasture	agriculture	agriculture
56	forest	pasture	conserv., prot.	all other categories
57	forest	pasture	agric., l-stock, for., urb.	all other categories
58	forest	pasture	agriculture	agriculture
59	forest	forest	agriculture	agriculture
60	forest	forest	agriculture	agriculture
61	forest	forest	agriculture	agriculture
62	forest	forest	agriculture	agriculture
63	forest	forest	agriculture	agriculture
64	residential	forest	anthropogenic, urban	all other categories
65	residential	residential	anthropogenic, urban	all other categories
66	forest	pasture	agric., l-stock, for., urb.	all other categories
67	forest	pasture	agric., l-stock, for., urb.	all other categories
68	forest	pasture	agric., l-stock, for., urb.	all other categories
69	forest	pasture	agric., l-stock, for., urb.	all other categories
70	arable	pasture	agric., l-stock, for., urb.	all other categories
71	forest	pasture	agric., l-stock, for., urb.	all other categories
72	forest	pasture	agric., l-stock, for., urb.	all other categories
73	forest	pasture	agric., l-stock, for., urb.	all other categories
74	forest	pasture	agric., l-stock, for., urb.	all other categories
75	forest	pasture	livestock	all other categories
76	forest	pasture	agric., l-stock, for., urb.	all other categories
77	forest	pasture	agric., l-stock, for., urb.	all other categories
78	forest	pasture	agric., l-stock, for., urb.	all other categories
79	orchard	pasture	agric., l-stock, for., urb.	all other categories
80	forest	pasture	agric., l-stock, for., urb.	all other categories
81	orchard	pasture	agric., l-stock, for., urb.	all other categories
82	forest	pasture	agric., l-stock, for., urb.	all other categories
83	forest	pasture	agric., l-stock, for., urb.	all other categories
84	forest	pasture	agric., l-stock, for., urb.	all other categories
85	arable	arable	agric., l-stock, for., urb.	all other categories
86	forest	arable	agric., l-stock, for., urb.	all other categories
87	arable	pasture	agric., l-stock, for., urb.	all other categories

Sampling site	FP land use	Google Land use	GIS land use 7 classes	GIS land use 2 classes
88	forest	arable	agric., l-stock, for., urb.	all other categories
89	forest	pasture	agric., l-stock, for., urb.	all other categories
90	forest	arable	agric., l-stock, for., urb.	all other categories
91	forest	arable	agric., l-stock, for., urb.	all other categories
92	forest	arable	agric., l-stock, for., urb.	all other categories
93	forest	arable	agric., l-stock, for., urb.	all other categories
94	forest	arable	agric., l-stock, for., urb.	all other categories
95	forest	arable	agric., l-stock, for., urb.	all other categories
96	arable	arable	agriculture	agriculture
97	arable	pasture	agriculture	agriculture
98	arable	pasture	agric., l-stock, for., urb.	all other categories
99	forest	arable	agric., l-stock, for., urb.	all other categories
100	forest	arable	agric., l-stock, for., urb.	all other categories
101	orchard	pasture	livestock	all other categories
102	forest	pasture	agric., l-stock, for., urb.	all other categories
103	orchard	pasture	agric., l-stock, for., urb.	all other categories
104	orchard	arable	agric., l-stock, for., urb.	all other categories
105	forest	pasture	agriculture	agriculture
106	forest	forest	agriculture	agriculture
107	forest	arable	agriculture	agriculture
108	forest	arable	agriculture	agriculture
109	arable	pasture	agric., l-stock, for., urb.	all other categories
110	arable	pasture	agric., l-stock, for., urb.	all other categories
111	arable	arable	agriculture	agriculture
112	orchard	pasture	livestock	all other categories
113	arable	residential	agriculture	agriculture
114	forest	arable	agriculture	agriculture
115	forest	arable	agric., l-stock, for., urb.	all other categories
116	forest	pasture	agriculture	agriculture
117	residential	arable	agriculture	agriculture
118	arable	arable	agric., l-stock, for., urb.	all other categories
119	arable	pasture	agric., l-stock, for., urb.	all other categories
120	forest	arable	agric., l-stock, for., urb.	all other categories

C – Supporting information for chapter 4

Figures

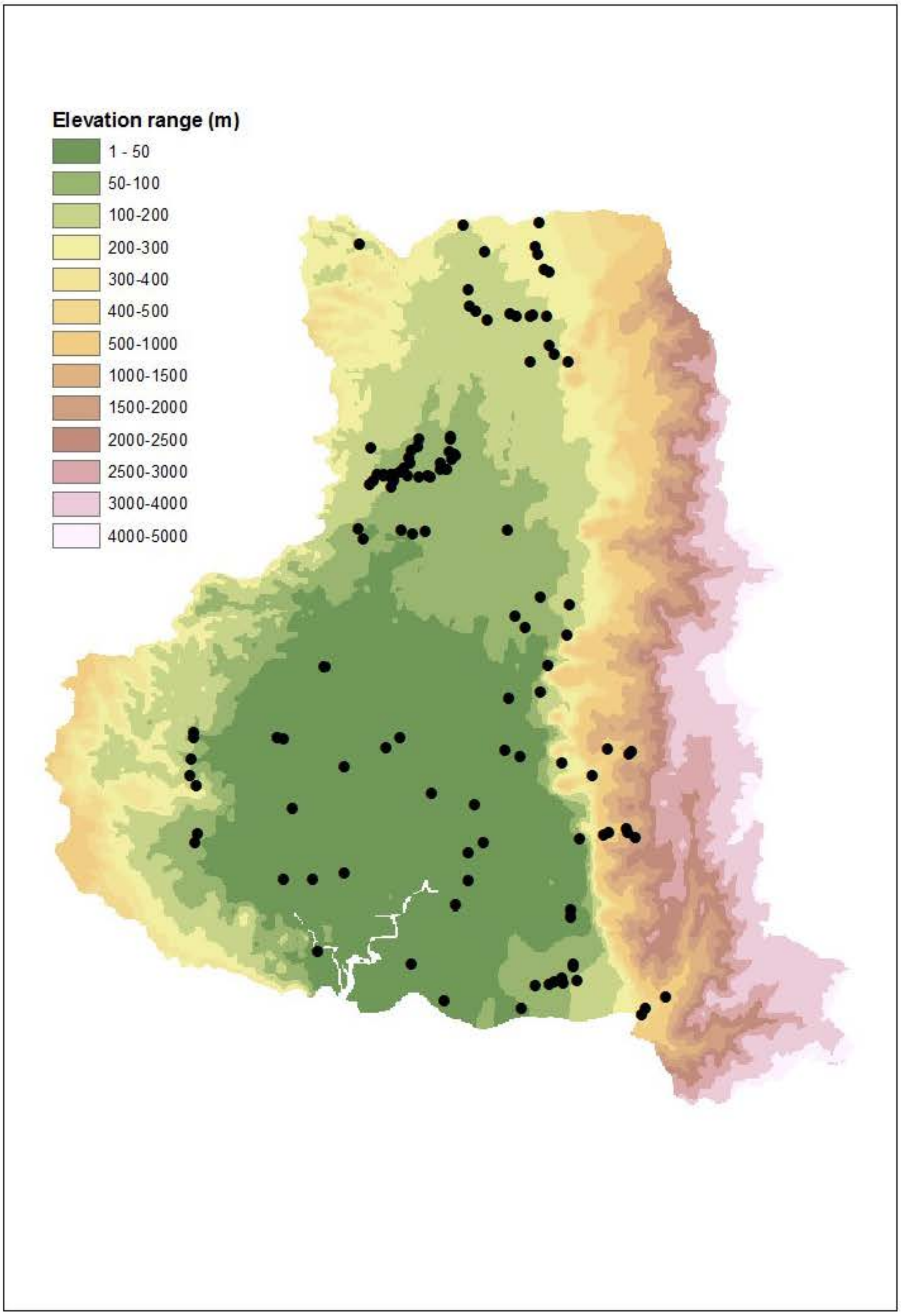


Figure C1 Elevation map of the 120 sampling sites in the Guayas river basin, different colors indicating elevation gradients.

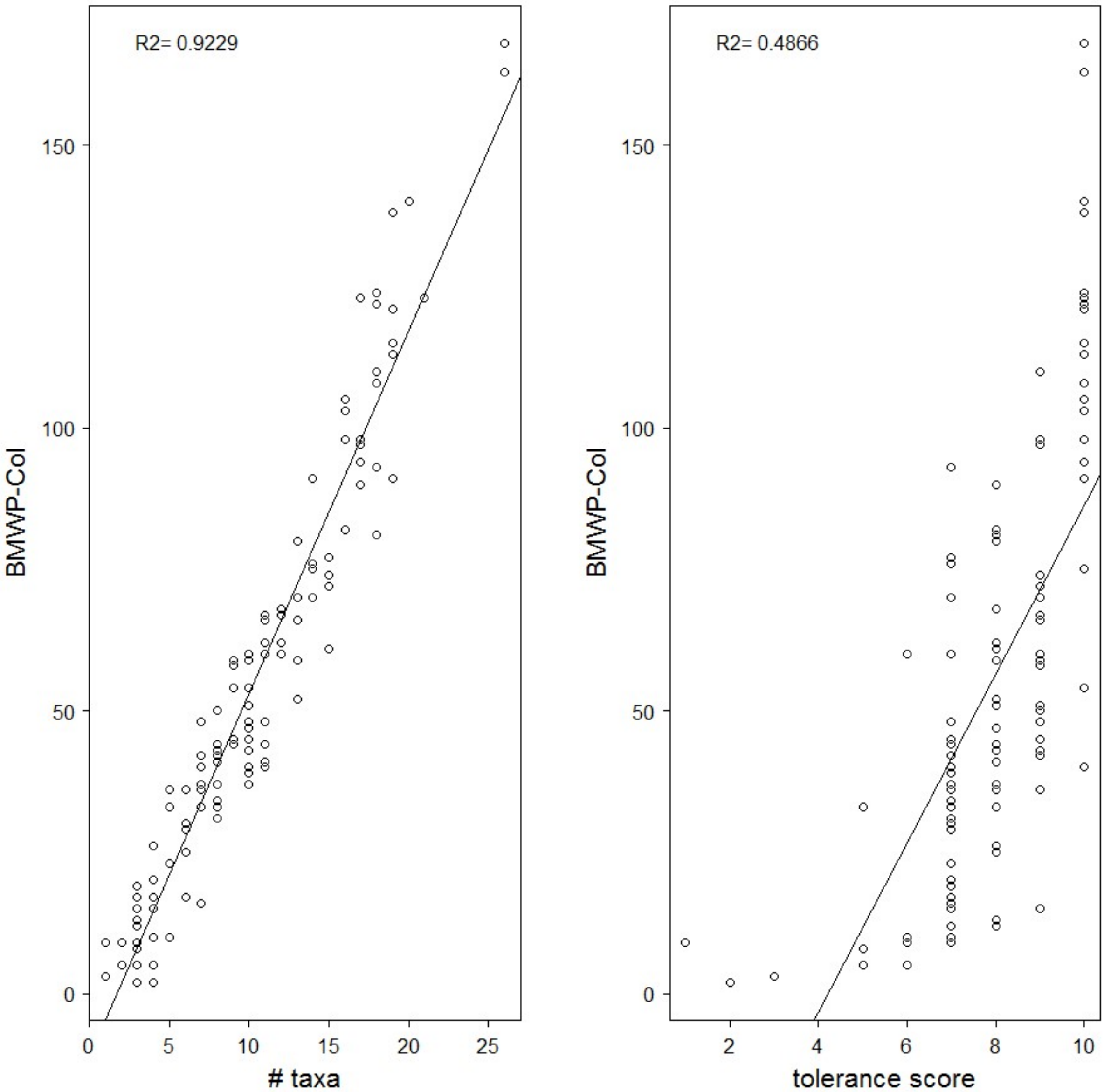


Figure C2 Plots of BMWP-Col in relation to the number of macroinvertebrates taxa (left) and their tolerance score (right).

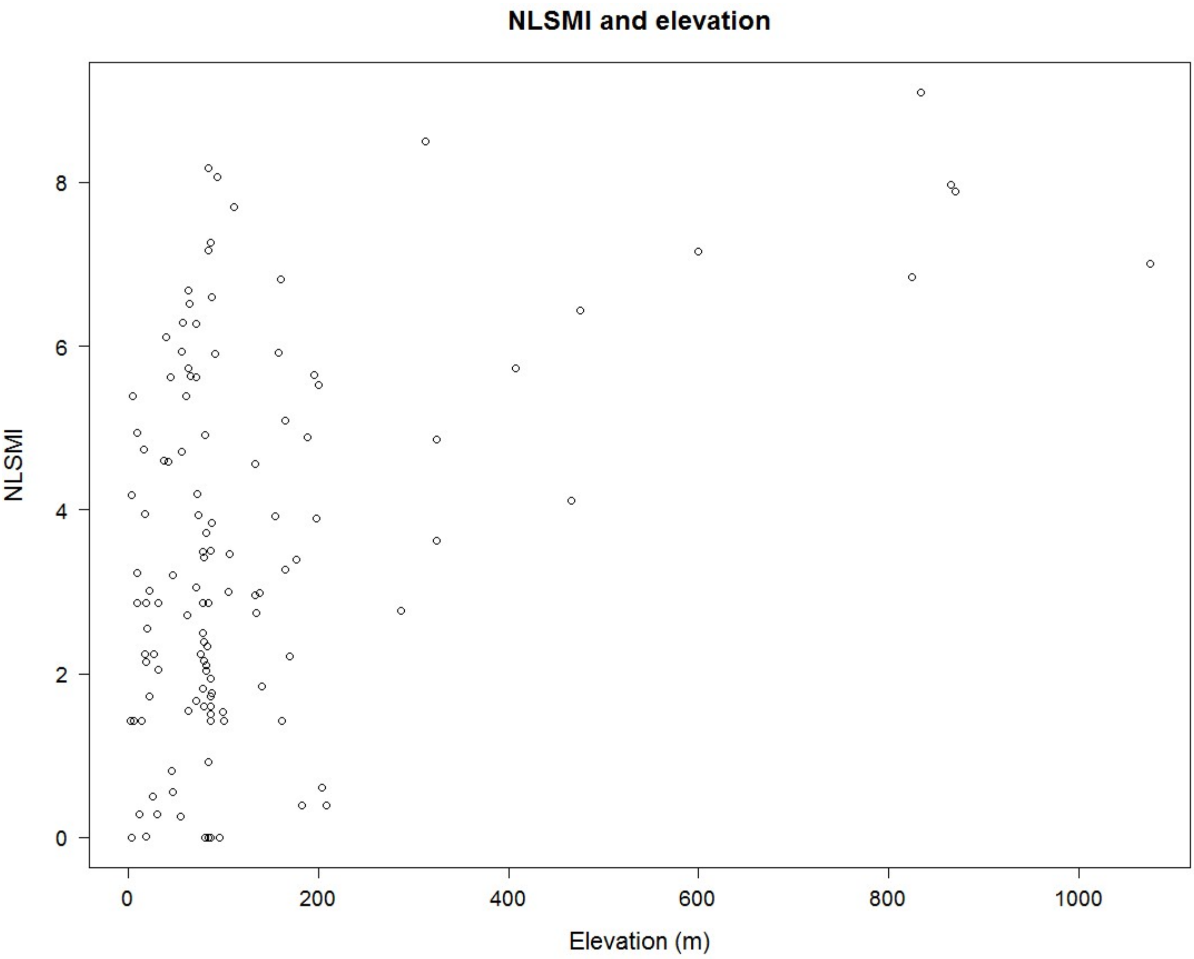


Figure C3 NLSMI in relation to elevation for 120 sampling sites.

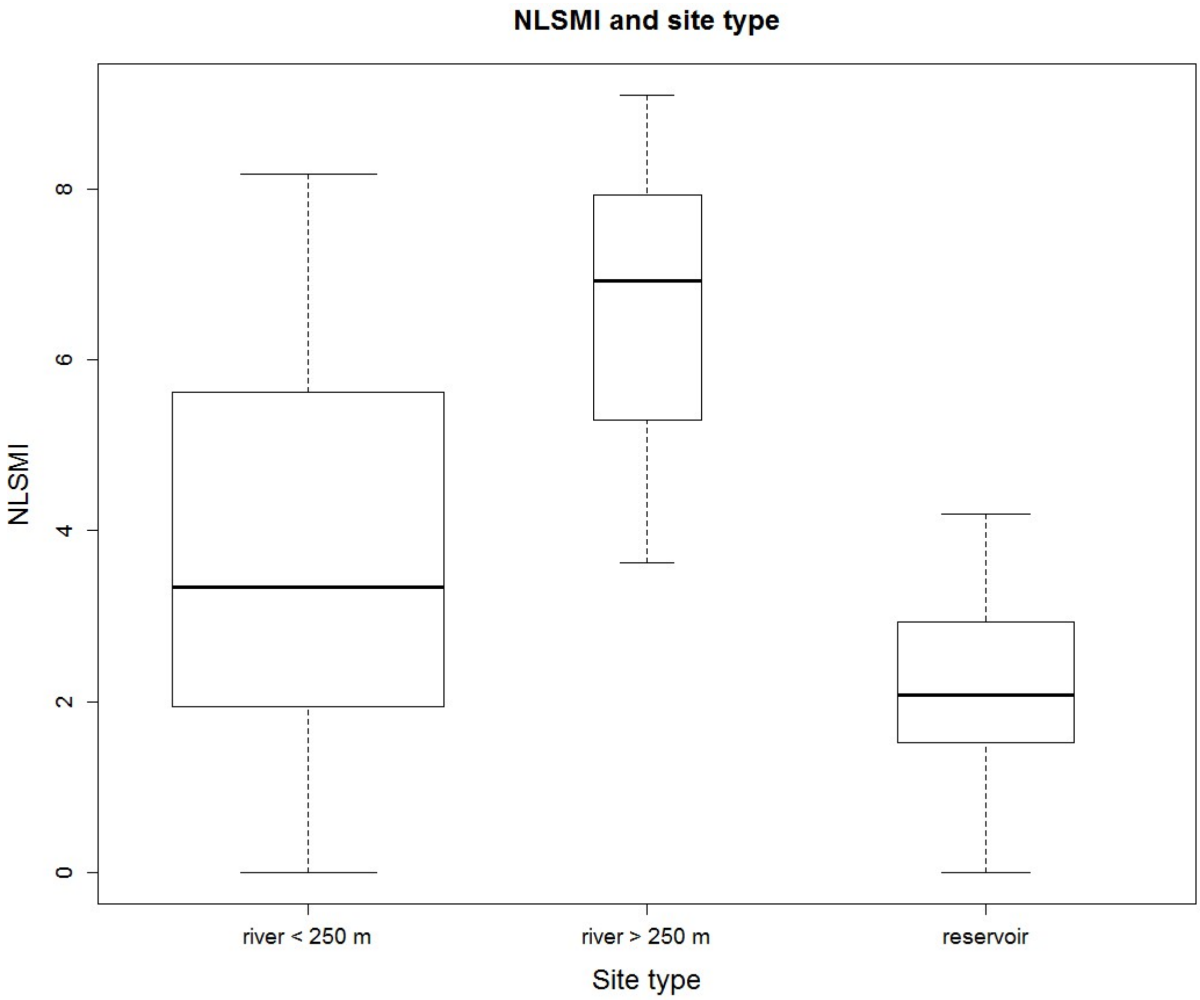


Figure C4 Boxplots of the NLSMI for 120 sampling sites classified in three groups according to the type of site.

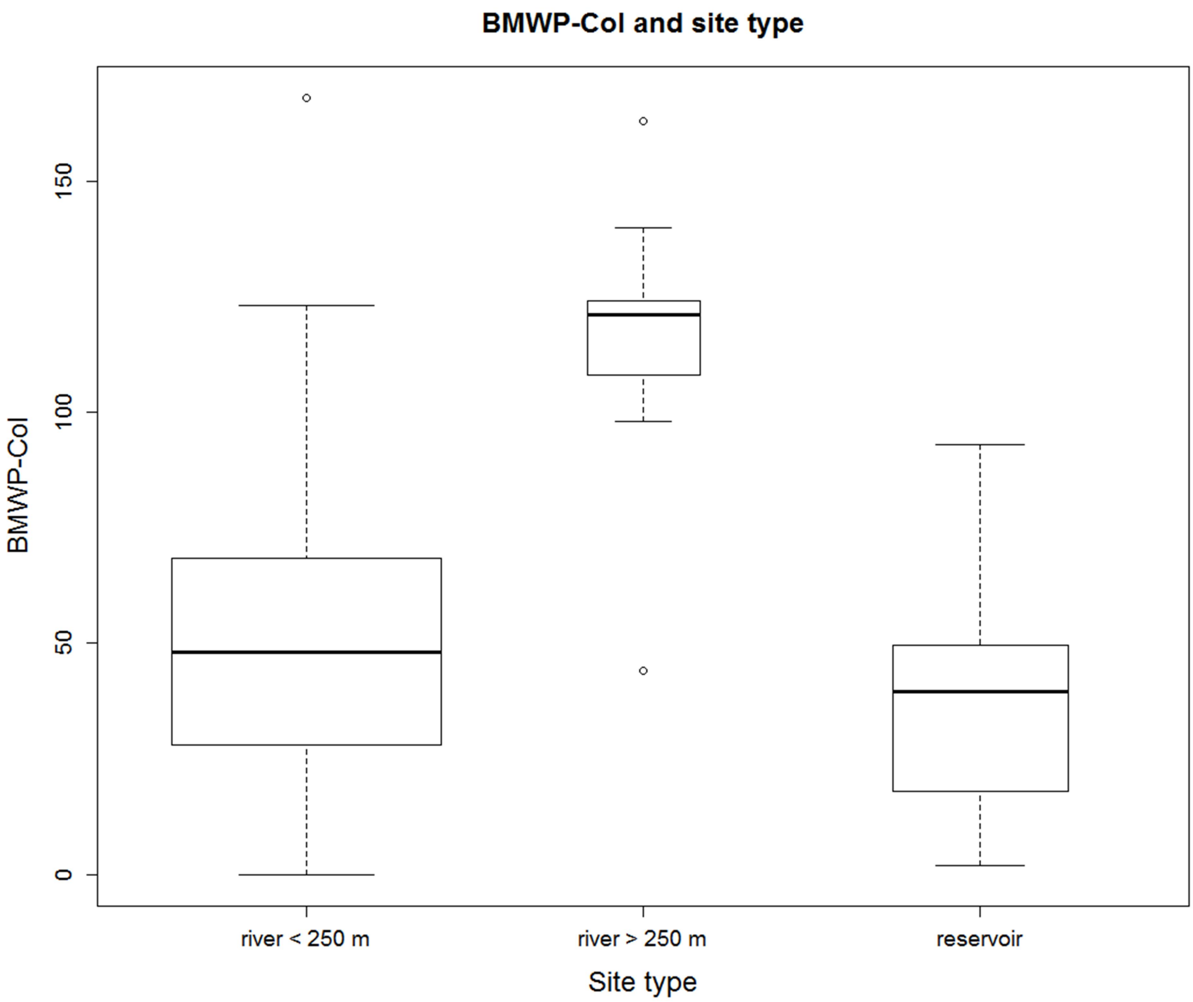


Figure C5 Boxplots of the BMWP-Col for 120 sampling sites classified in three groups according to the type of site.

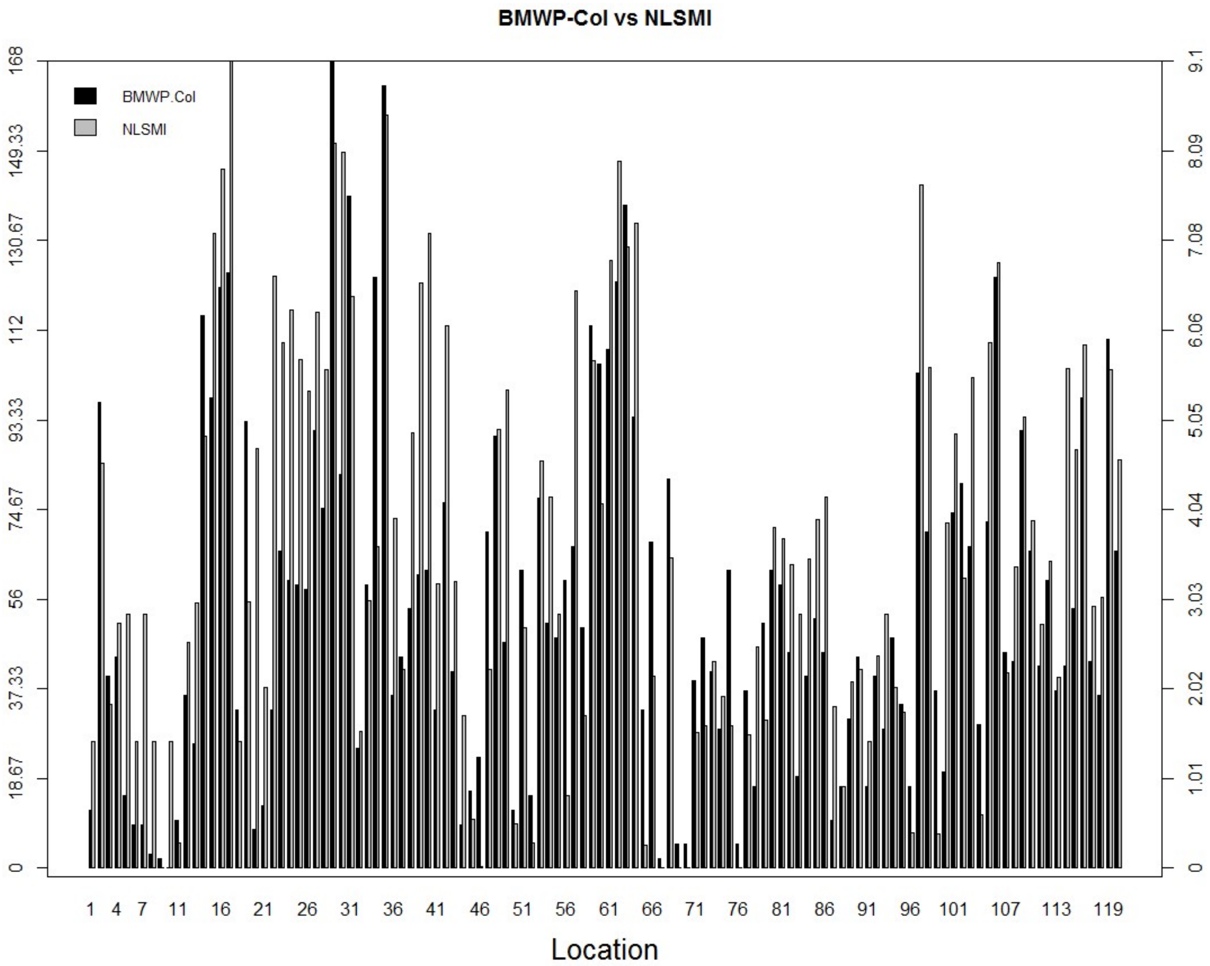


Figure C6 Correlation barplot between BMWP-Col and NLSMI for 120 sampling sites.

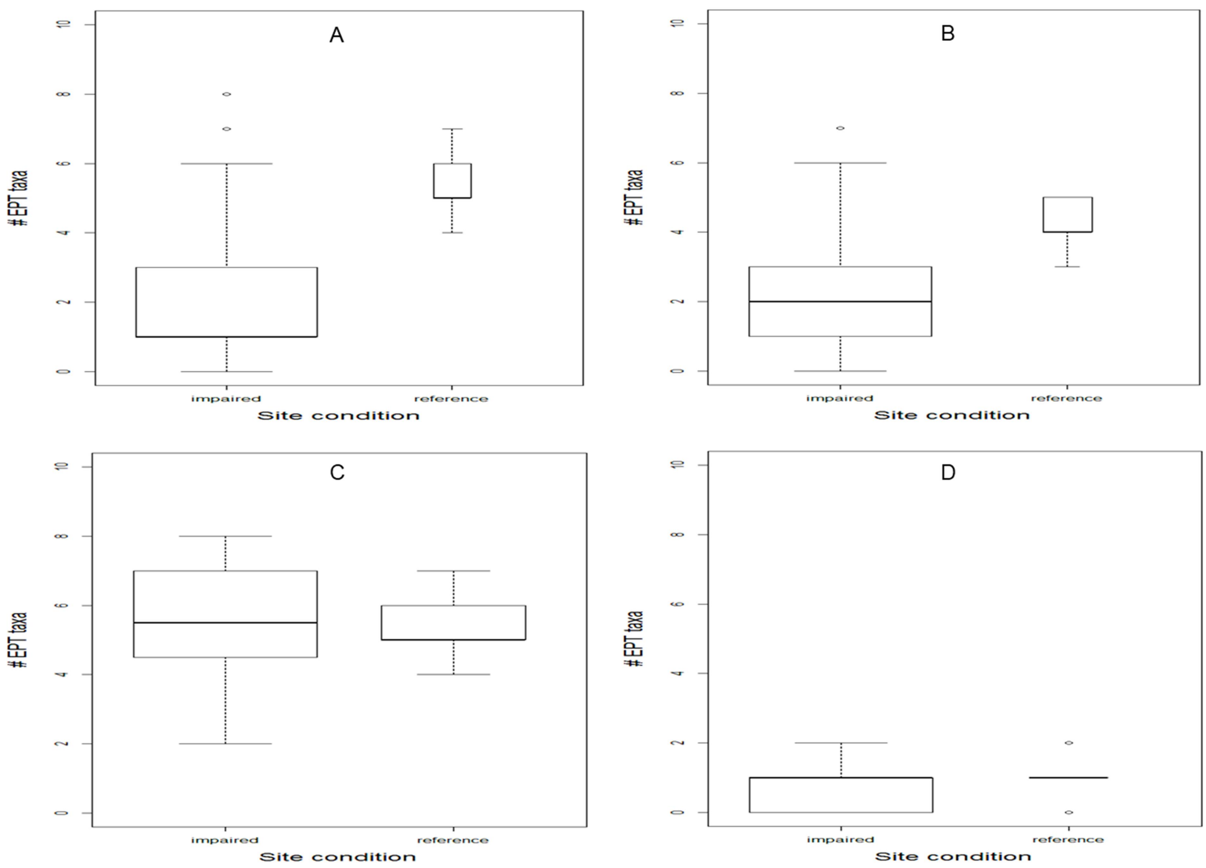


Figure C7 Boxplots of number of Ephemeroptera Plecoptera Trichoptera (EPT) taxa in relation to site condition for 120 sampling sites (A), for sampling sites located at the elevation lower than 250 m (B), elevation higher than 250 m (C), and at the reservoir (D). The width of the boxes is proportional to the number of observations.

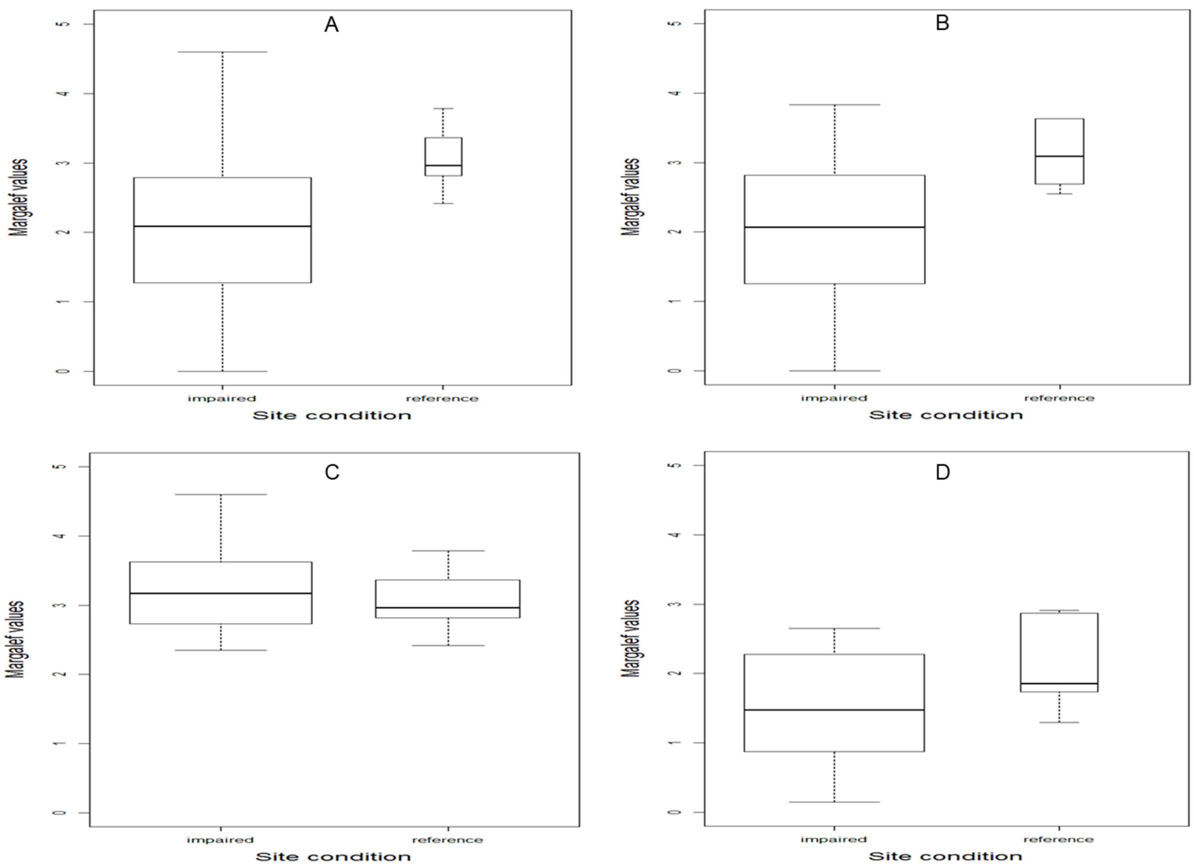


Figure C8 Boxplots of Margalef's index in relation to site condition for 120 sampling sites (A), for sampling sites located at the elevation lower than 250 m (B), elevation higher than 250 m (C), and at the reservoir (D). The width of the boxes is proportional to the number of observations.

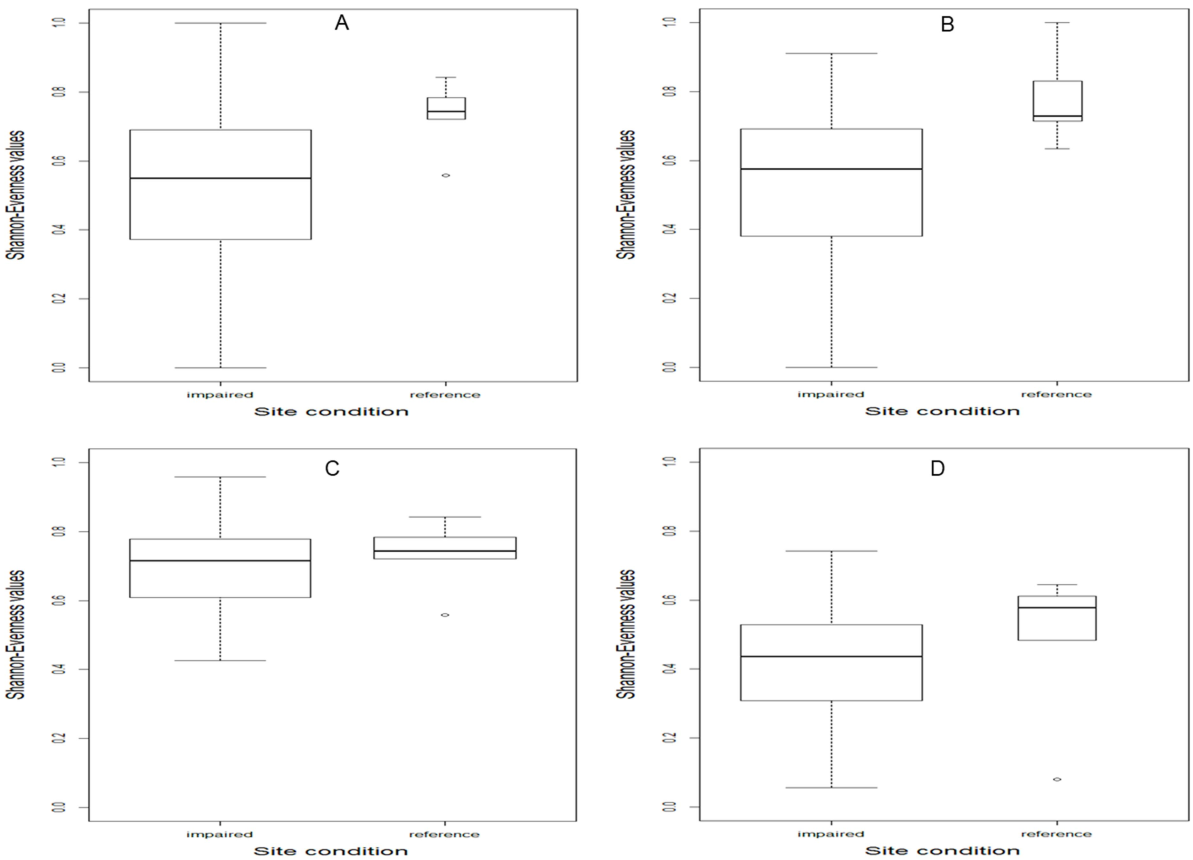


Figure C9 Boxplots of Shannon-Wiener Evenness index in relation to site condition for 120 sampling sites (A), for sampling sites located at the elevation lower than 250 m (B), elevation higher than 250 m (C), and at the reservoir (D). The width of the boxes is proportional to the number of observations.

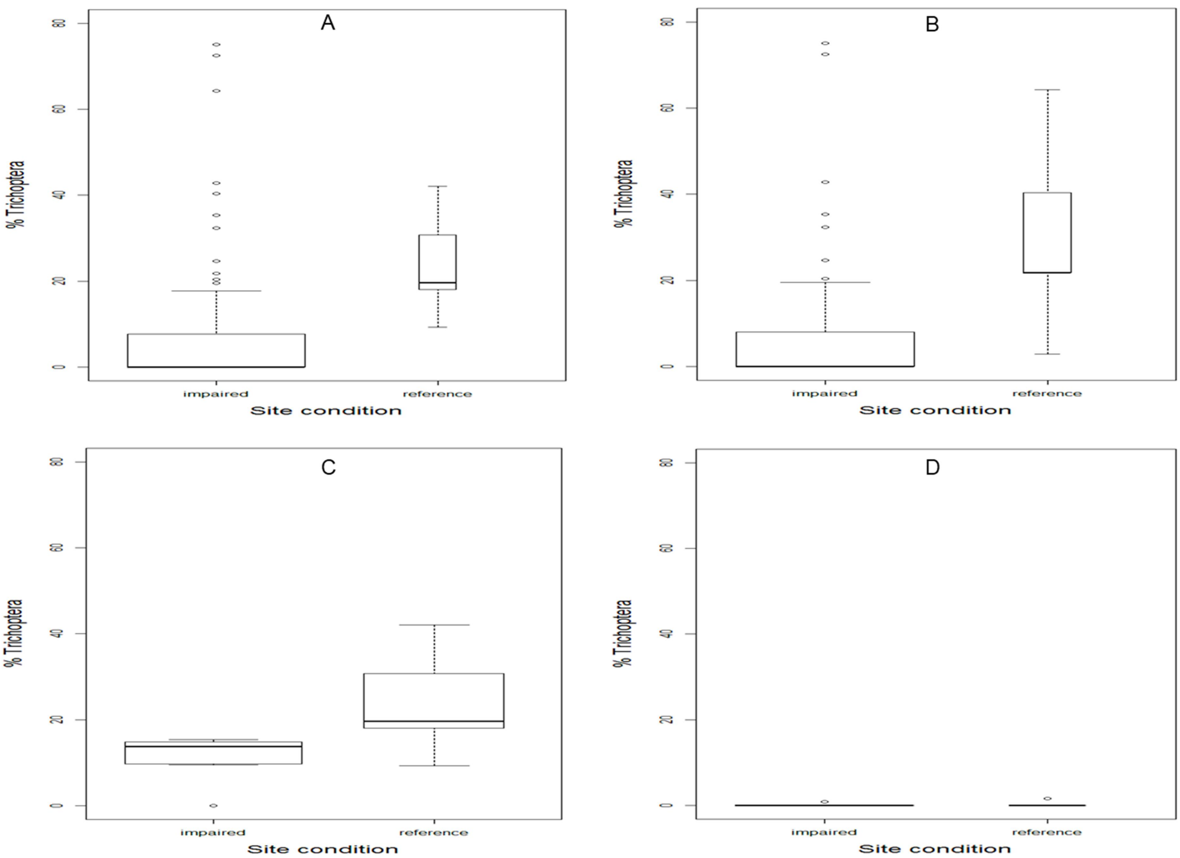


Figure C10 Boxplots of percentage of Trichoptera taxa in relation to site condition for 120 sampling sites (A), for sampling sites located at the elevation lower than 250 m (B), elevation higher than 250 m (C), and at the reservoir (D). The width of the boxes is proportional to the number of observations.

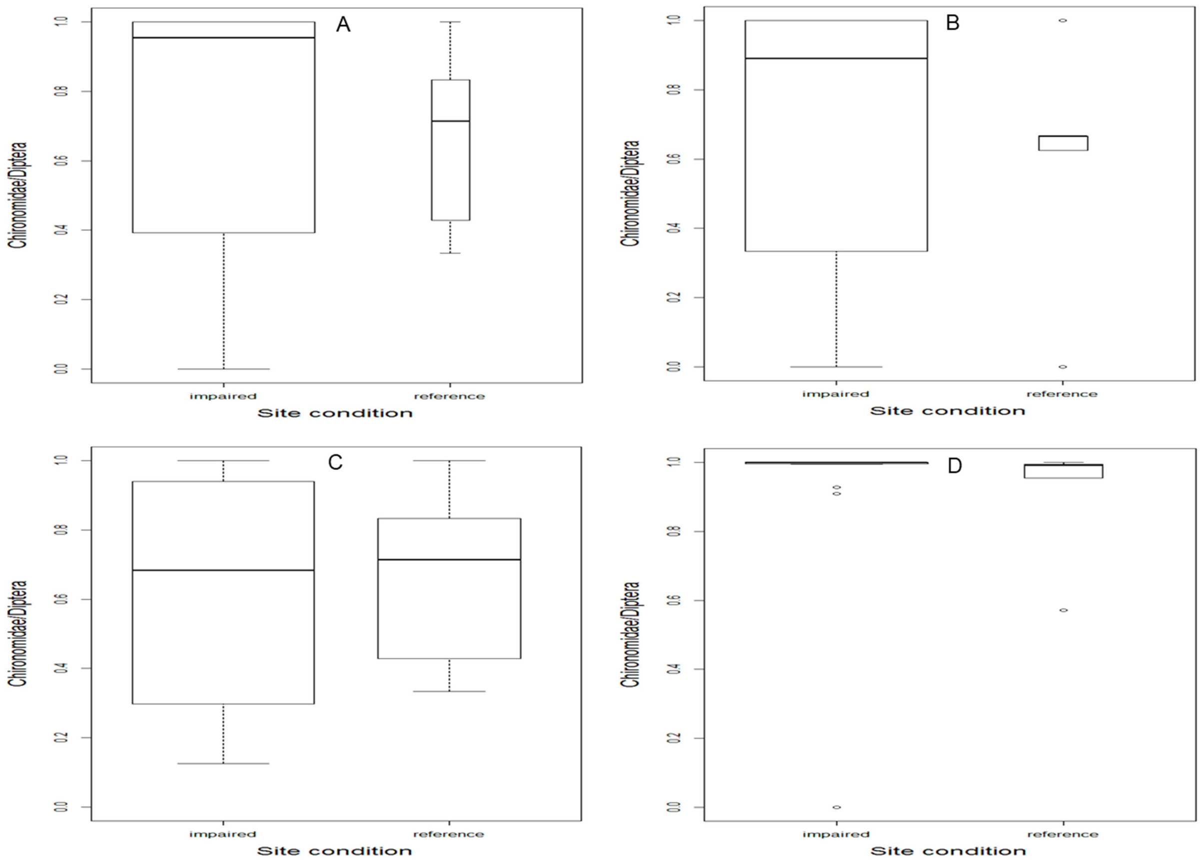


Figure C11 Boxplots of Chironomidae/Diptera individuals ratio in relation to site condition for 120 sampling sites (A), for sampling sites located at the elevation lower than 250 m (B), elevation higher than 250 m (C), and at the reservoir (D). The width of the boxes is proportional to the number of observations.

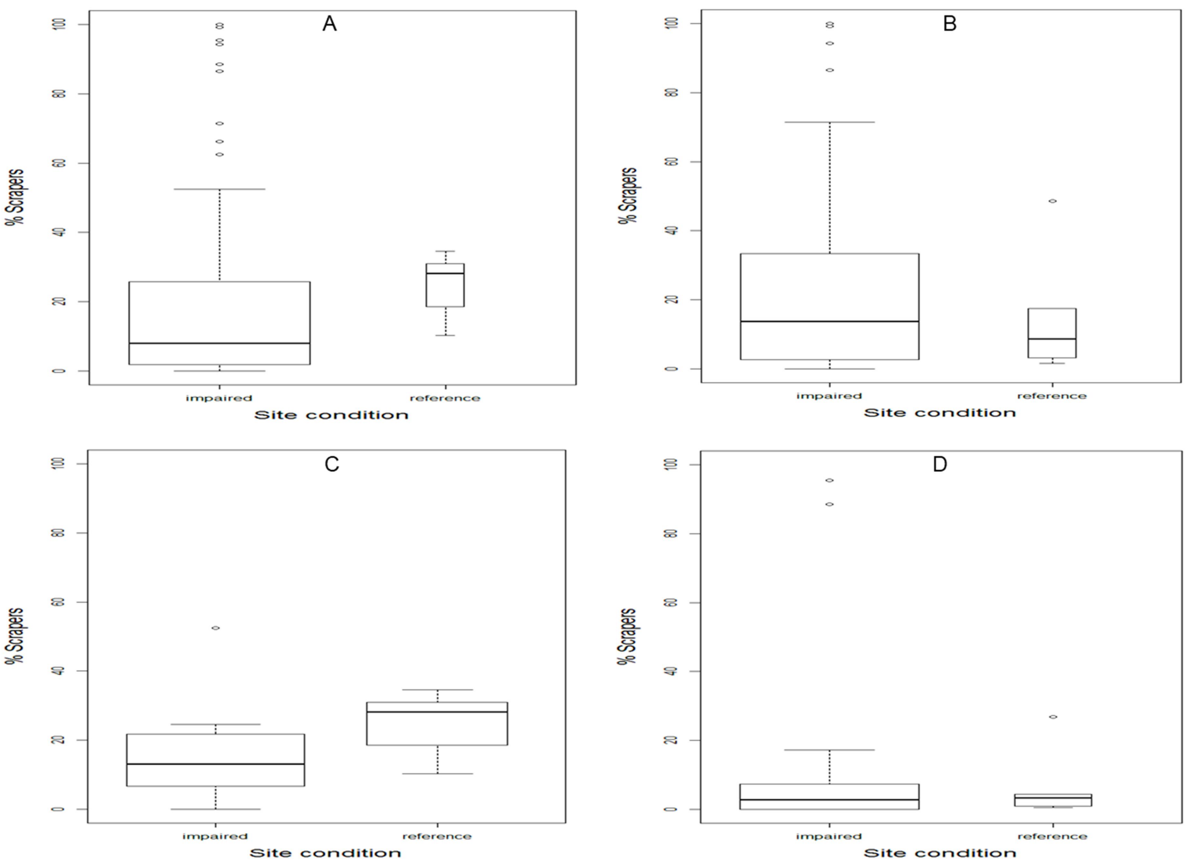


Figure C12 Boxplots of percentage of scrapers in relation to site condition for 120 sampling sites (A), for sampling sites located at the elevation lower than 250 m (B), elevation higher than 250 m (C), and at the reservoir (D). The width of the boxes is proportional to the number of observations.

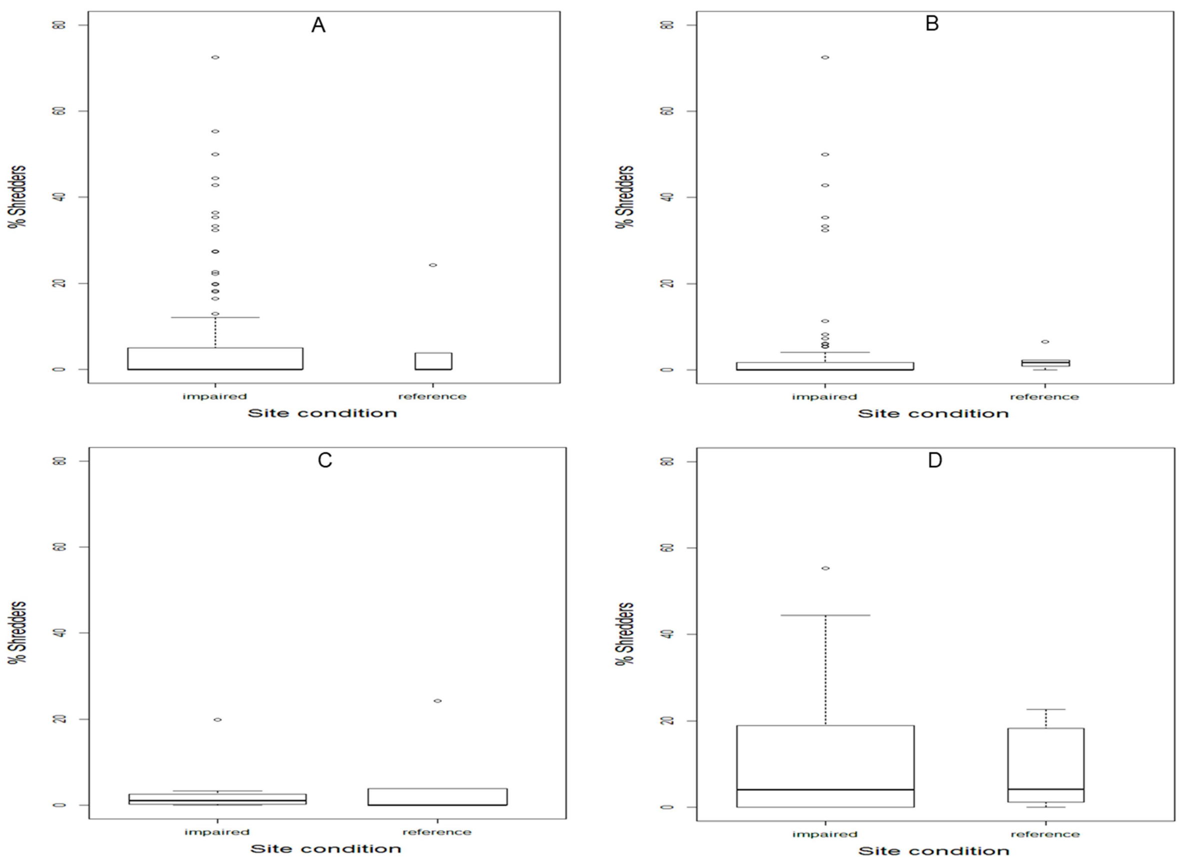


Figure C13 Boxplots of percentage of shredders in relation to site condition for 120 sampling sites (A), for sampling sites located at the elevation lower than 250 m (B), elevation higher than 250 m (C), and at the reservoir (D). The width of the boxes is proportional to the number of observations.

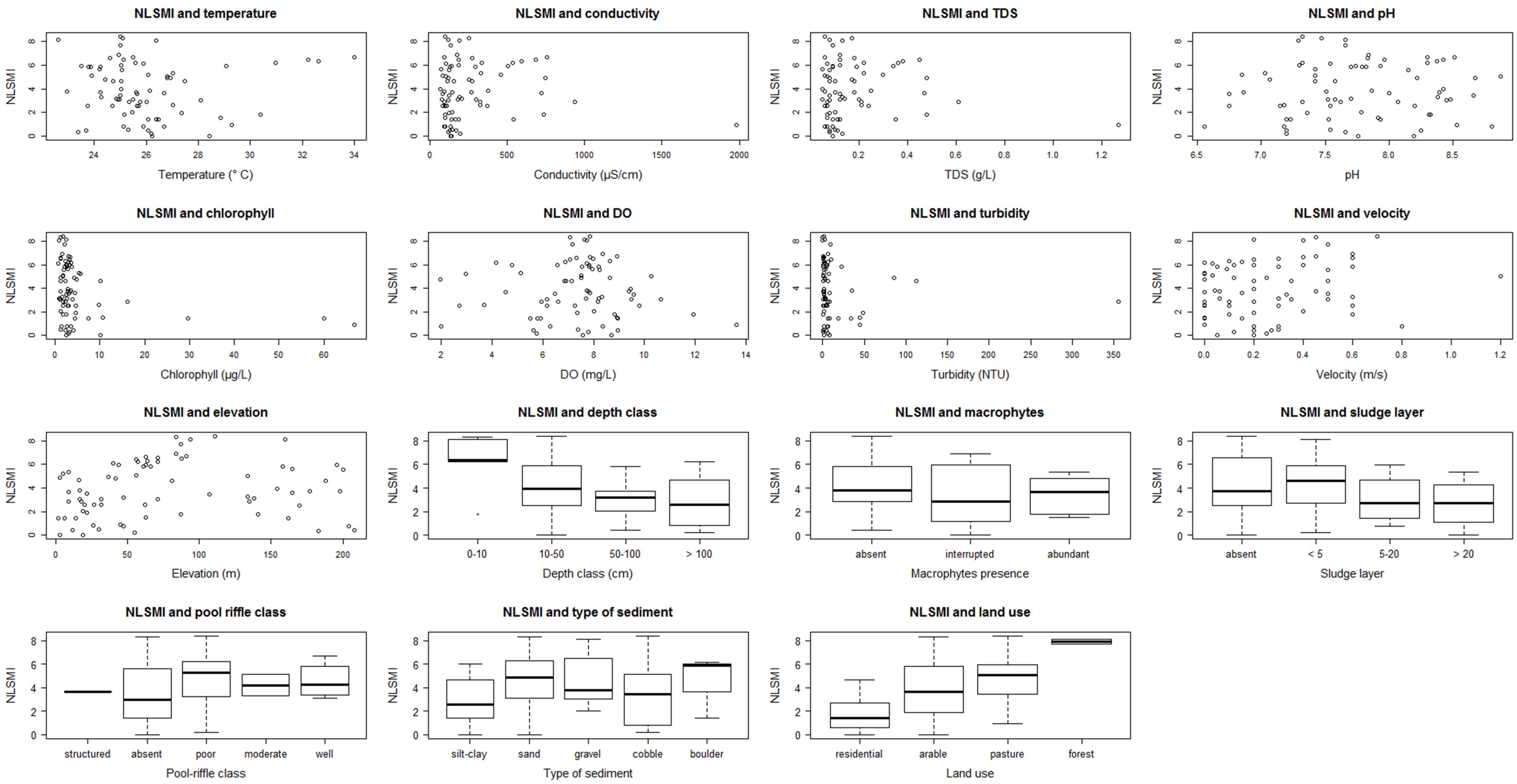


Figure C14 Plots showing the distribution of the data for physico-chemical variables in relation to NLSMI for rivers lower than 250 m. The classification of depth class, presence of macrophytes, sludge layer, pool-riffle class, type of sediment and land use is based on Table B1.

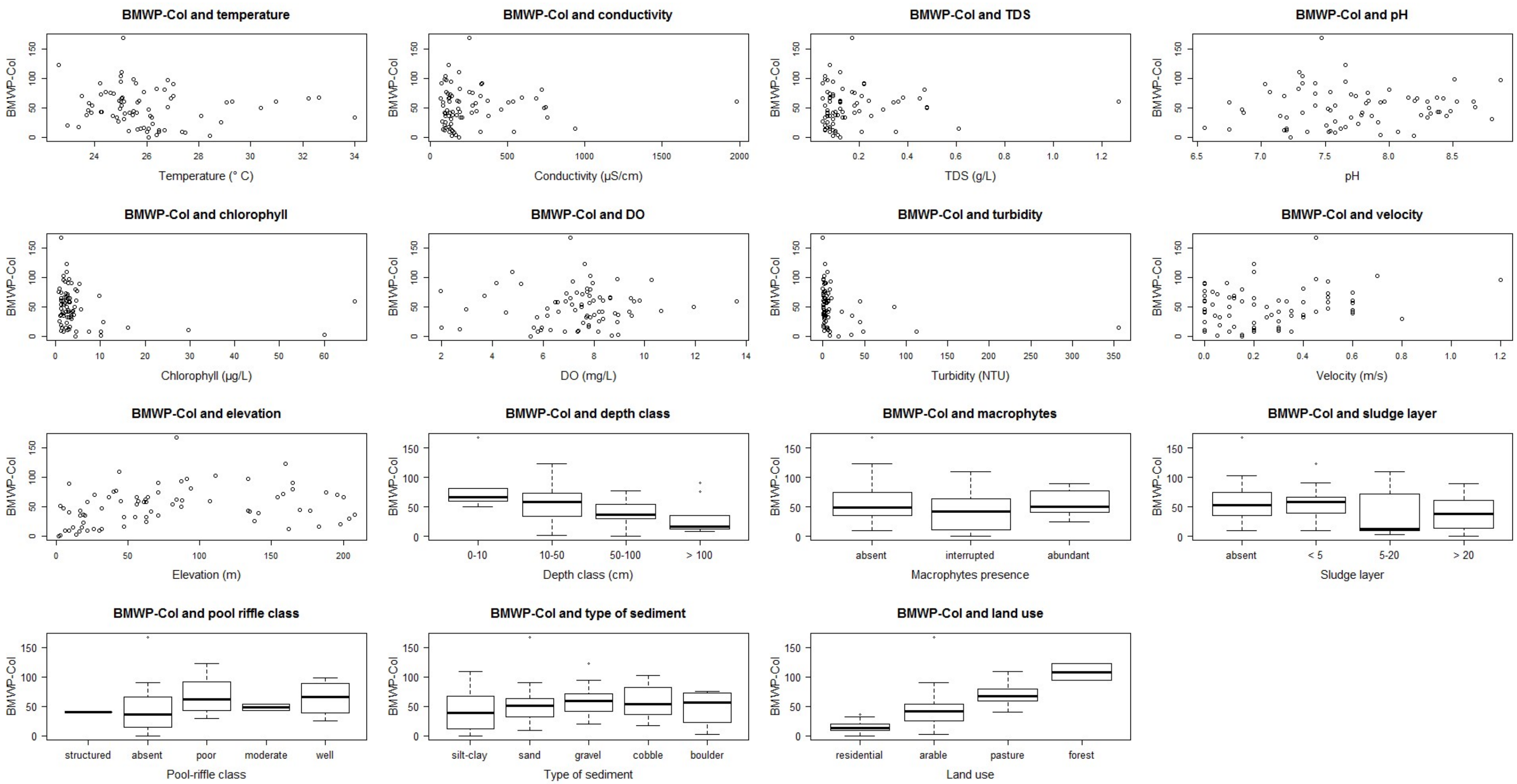


Figure C15 Plots showing the distribution of the data for physico-chemical variables in relation to BMWP-Col for rivers lower than 250 m. The classification of depth class, presence of macrophytes, sludge layer, pool-riffle class, type of sediment and land use is based on Table B1.

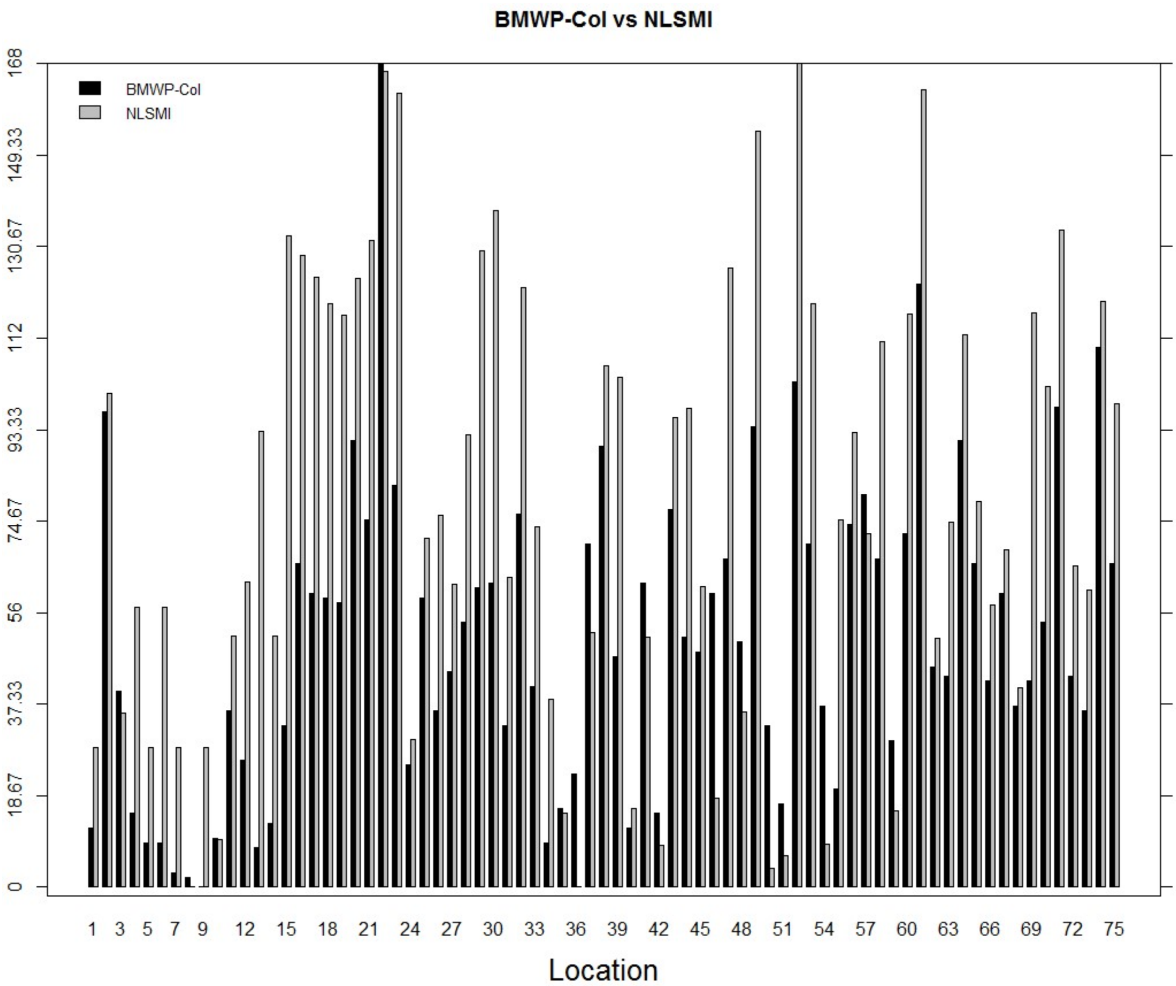


Figure C16 Correlation barplot between BMWP-Col and NLSMI for rivers lower than 250 m.

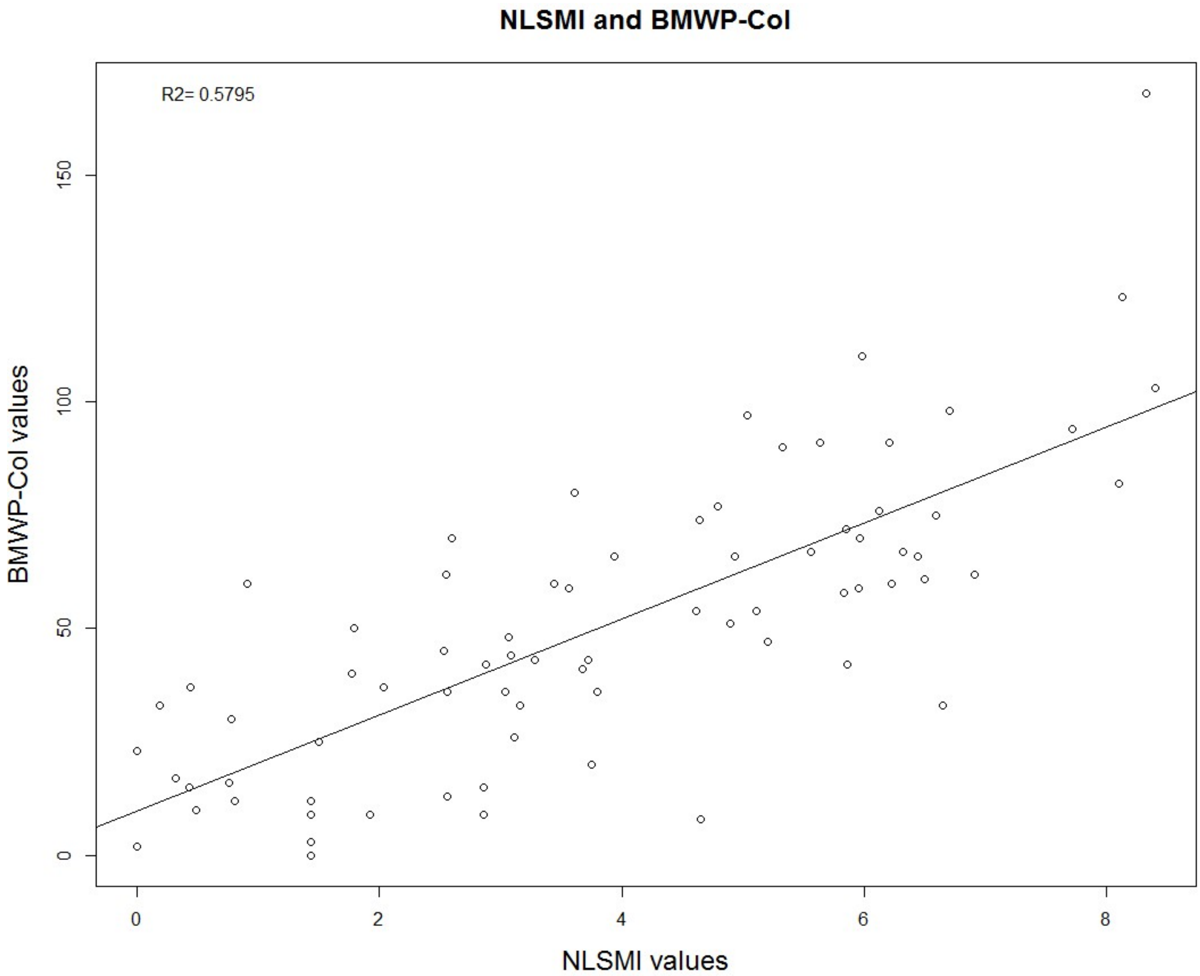


Figure C17 Correlation between BMWP-Col and NLSMI for rivers lower than 250 m, $p < 0.001$.

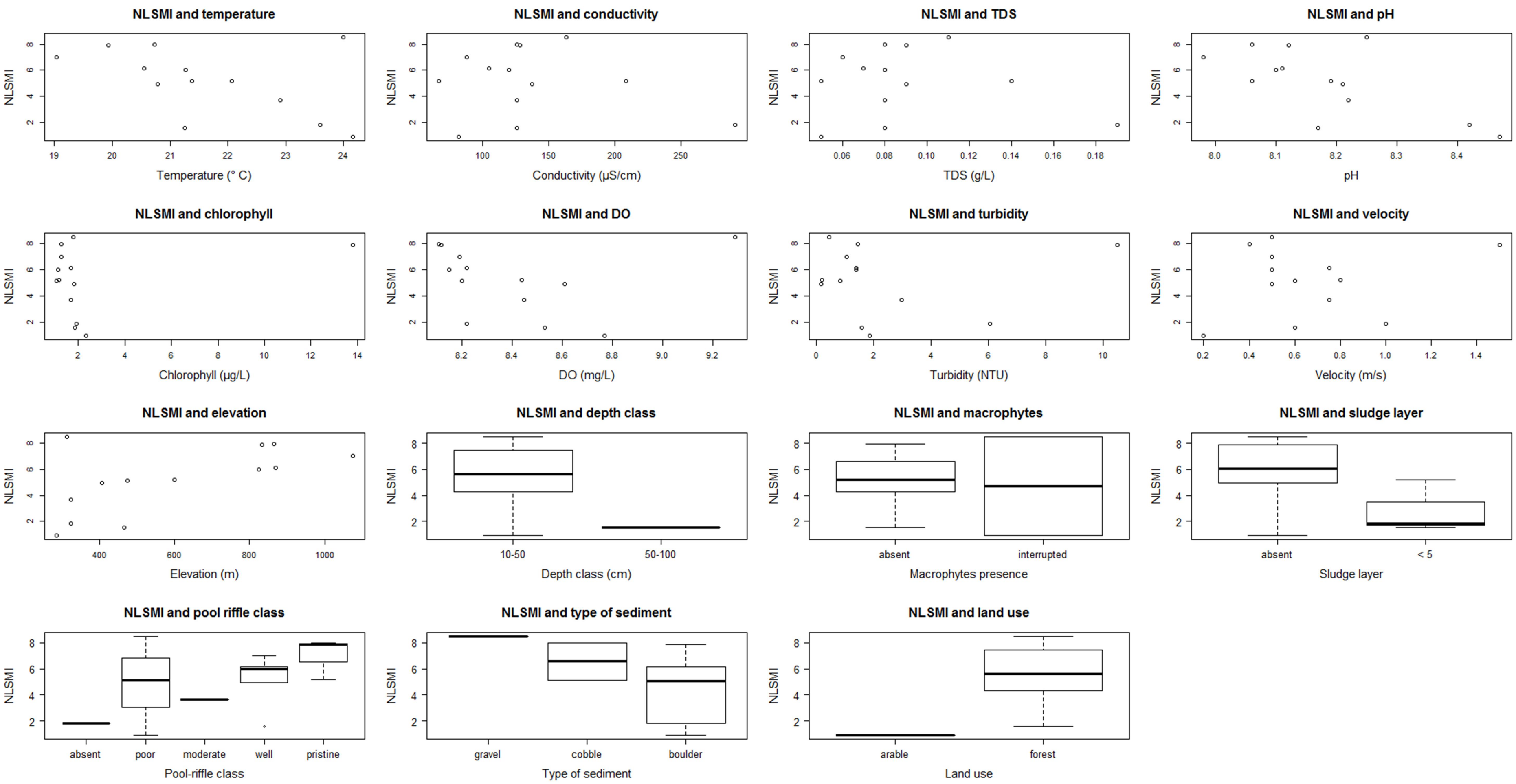


Figure C18 Plots showing the distribution of the data for physico-chemical variables in relation to NLSMI for rivers higher than 250 m. The classification of depth class, presence of macrophytes, sludge layer, pool-riffle class, type of sediment and land use is based on Table B1.

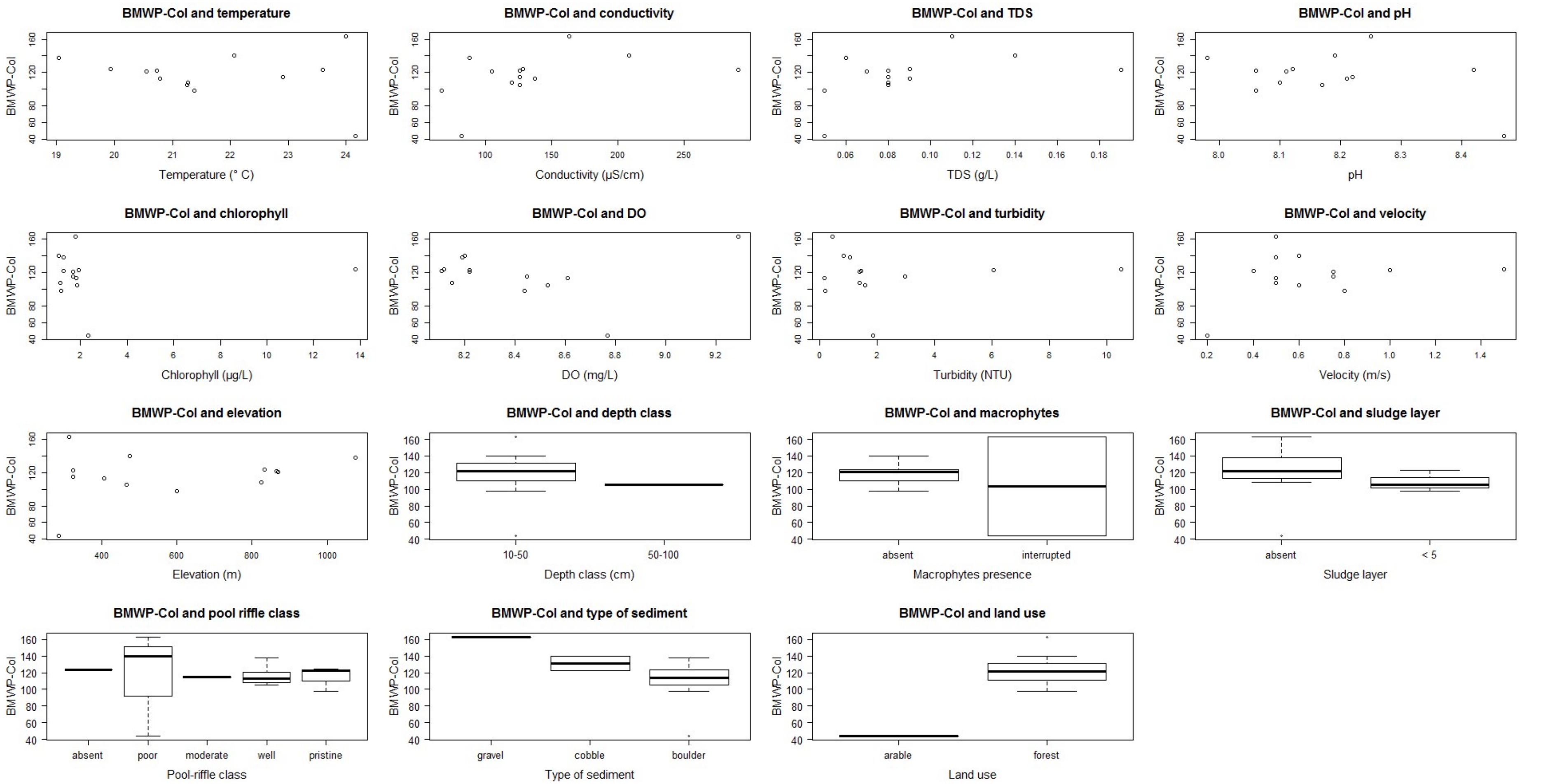


Figure C19 Plots showing the distribution of the data for physico-chemical variables in relation to BMWP-Col for rivers higher than 250 m. The classification of depth class, presence of macrophytes, sludge layer, pool-riffle class, type of sediment and land use is based on Table B1.

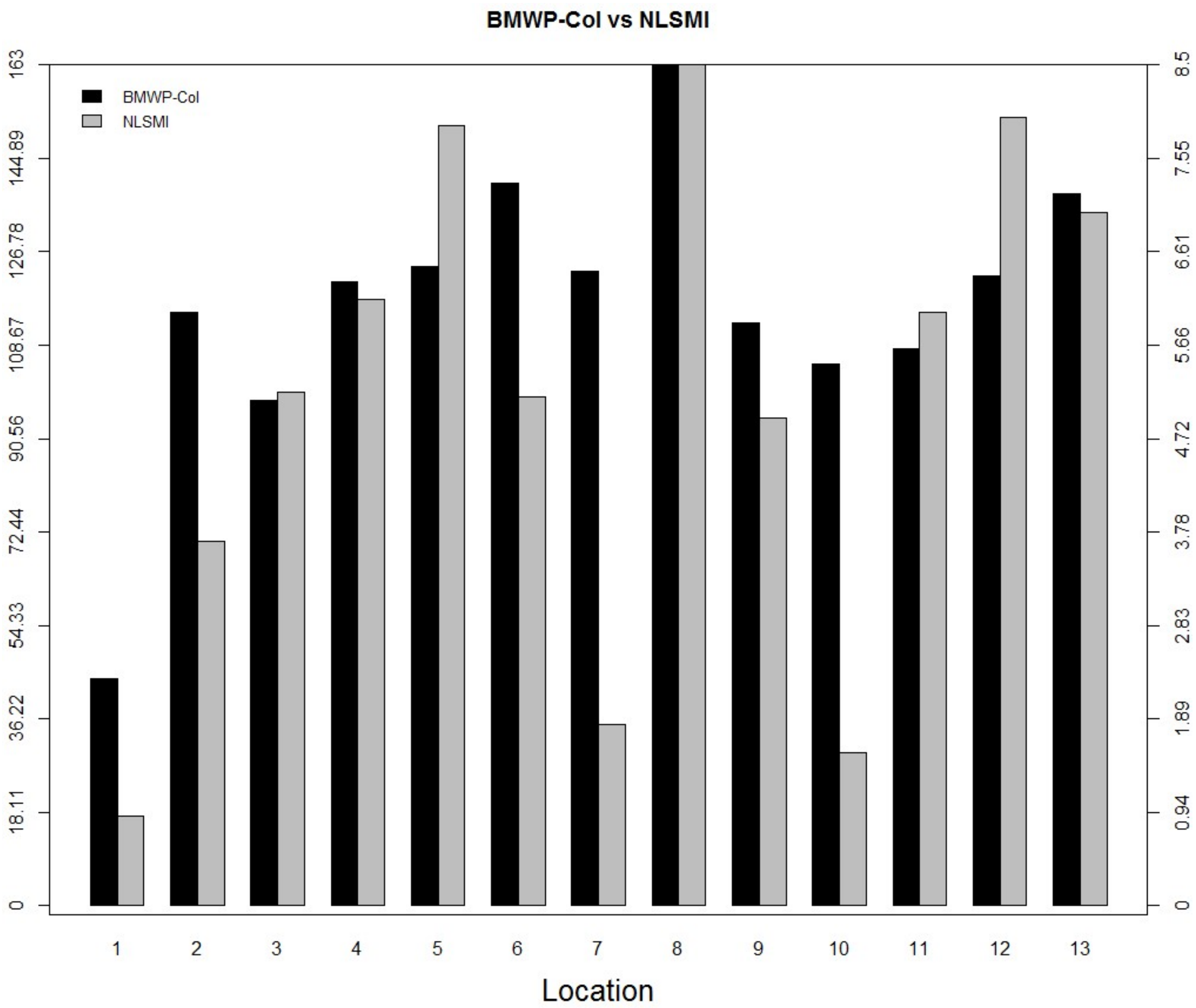


Figure C20 Correlation barplot between BMWP-Col and NLSMI for rivers higher than 250 m.

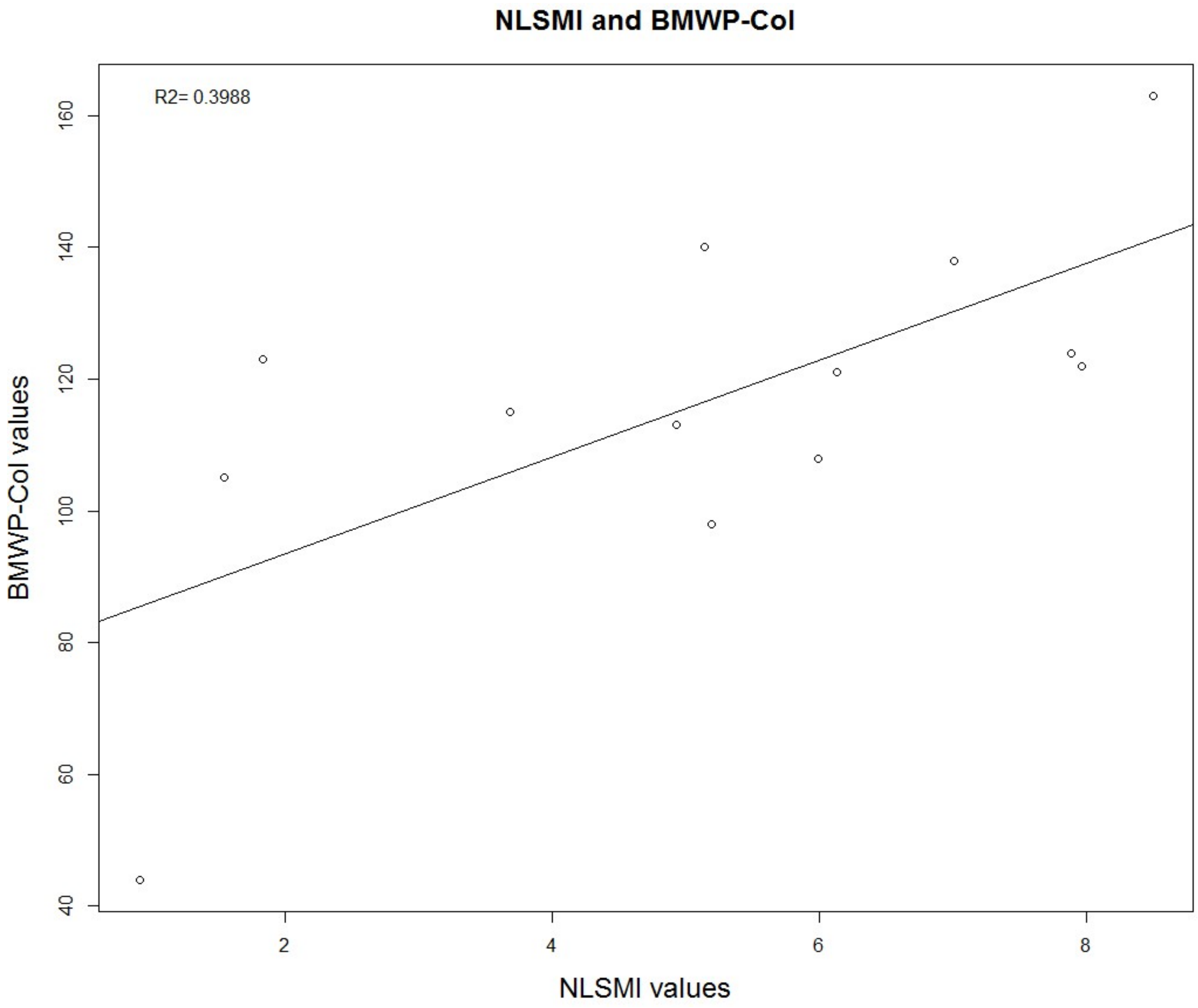


Figure C21 Correlation between BMWP-Col and NLSMI for rivers higher than 250 m, p = 0.01.

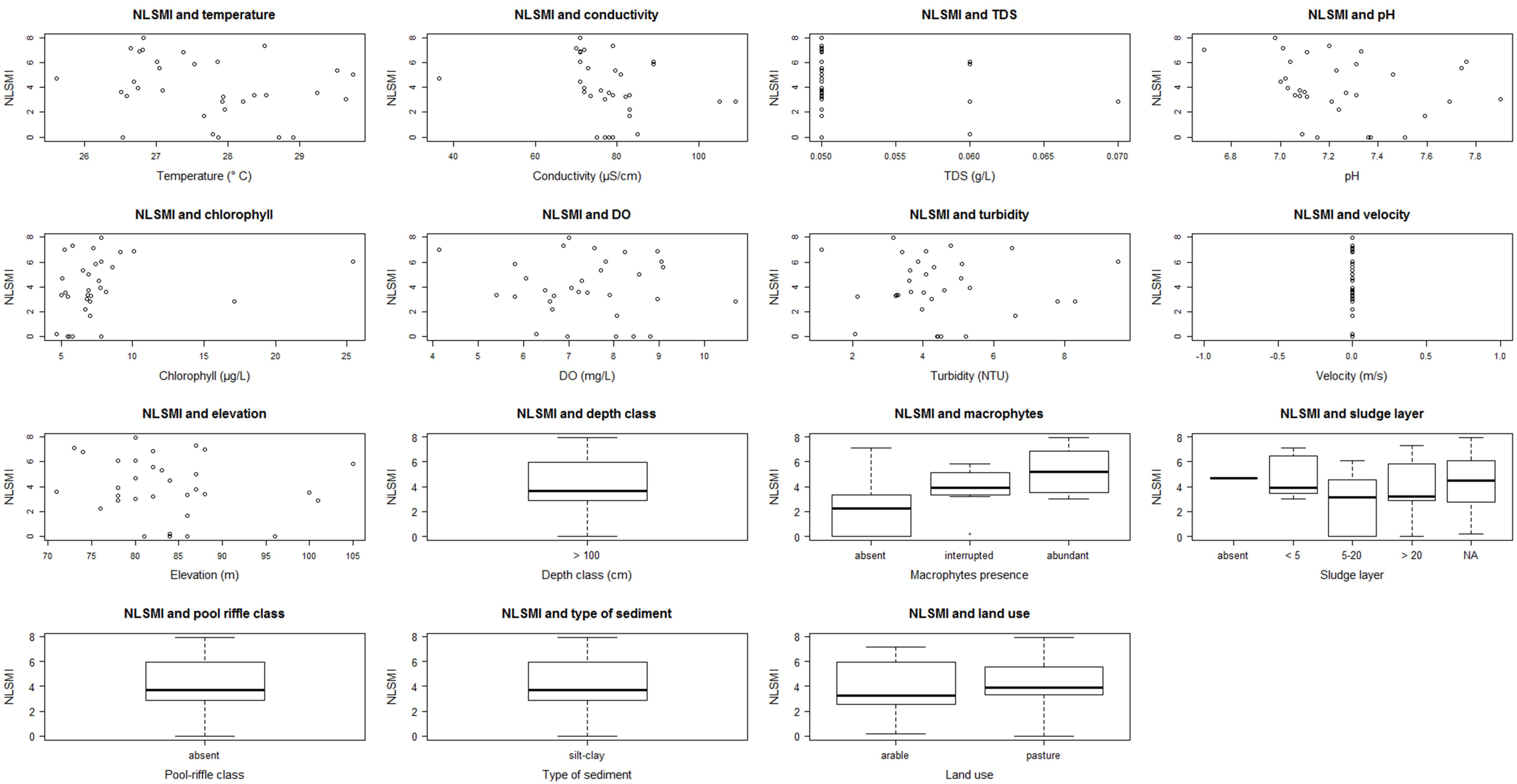


Figure C22 Plots showing the distribution of the data for physico-chemical variables in relation to NLSMI for reservoir. The classification of depth class, presence of macrophytes, sludge layer, pool-riffle class, type of sediment and land use is based on Table B1.

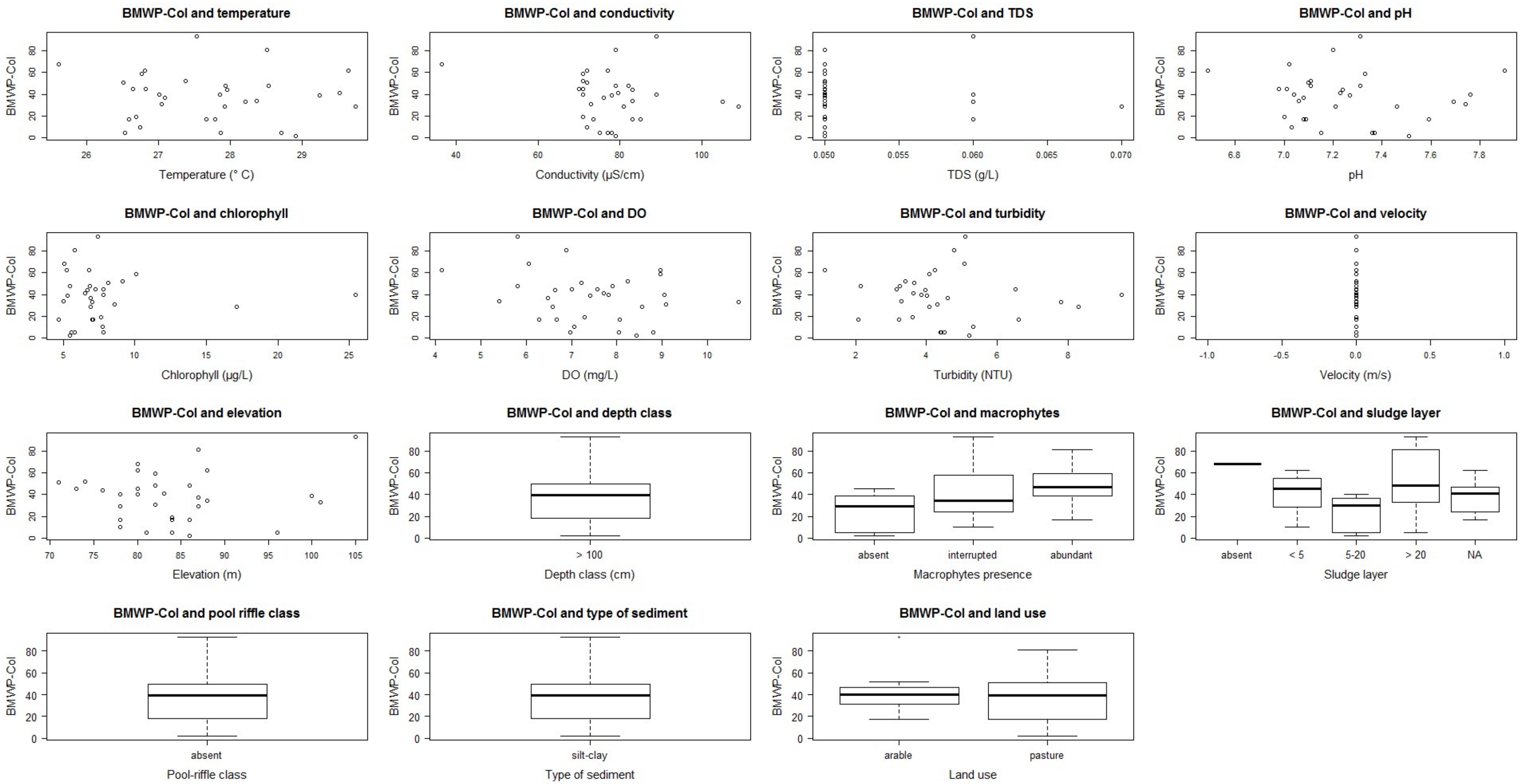


Figure C23 Plots showing the distribution of the data for physico-chemical variables in relation to BMWP-Col for reservoir. The classification of depth class, presence of macrophytes, sludge layer, pool-riffle class, type of sediment and land use is based on Table B1.

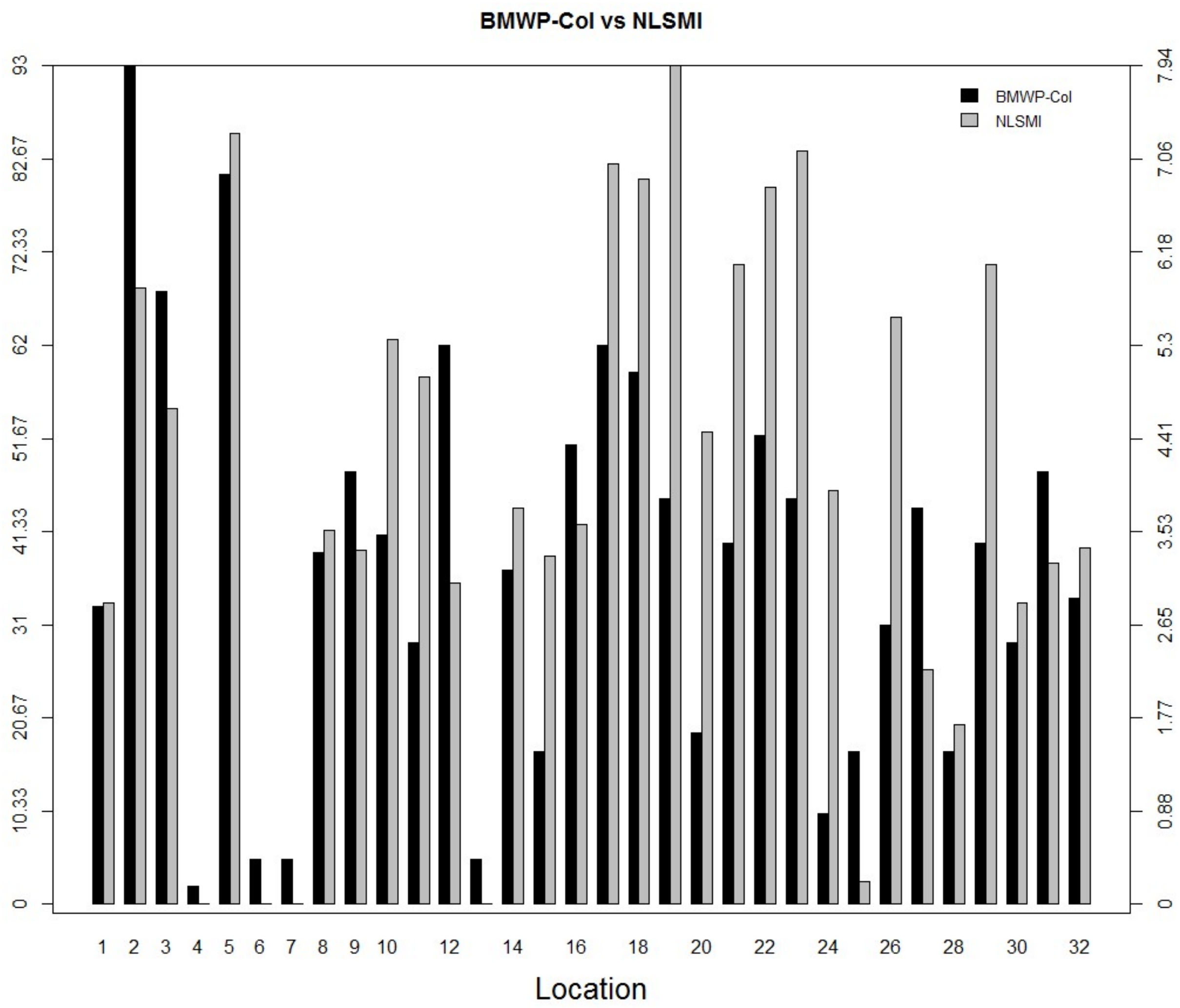


Figure C24 Correlation barplot between BMWP-Col and NLSMI for reservoir.

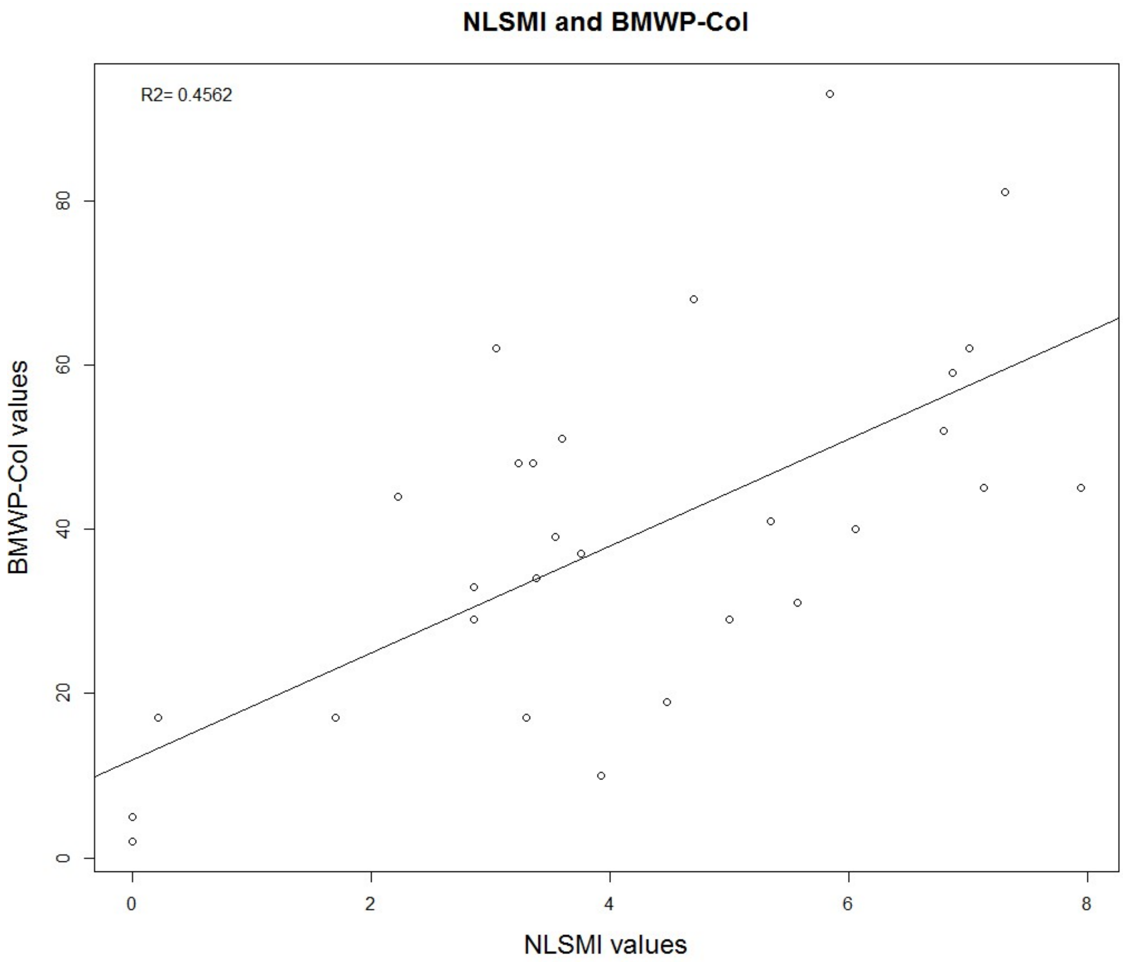


Figure C25 Correlation between BMWP-Col and NLSMI for reservoir, $p < 0.001$.

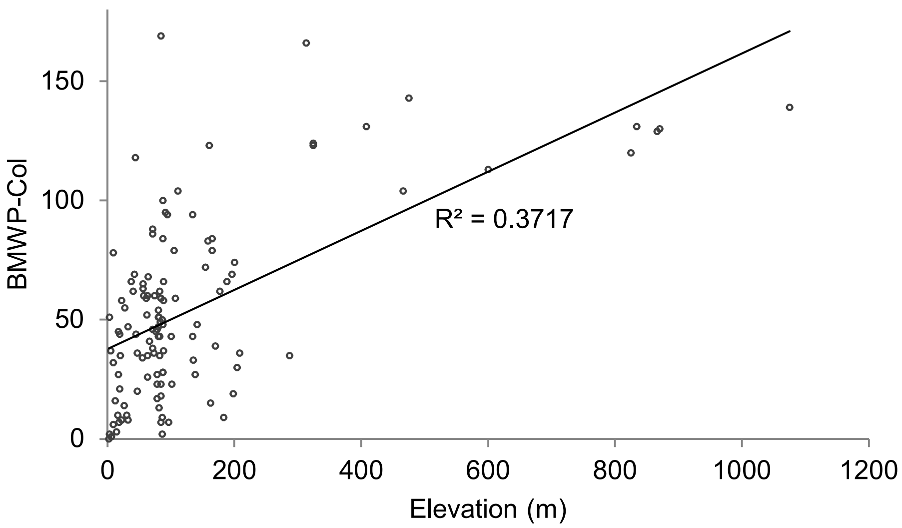


Figure C26 The ecological water quality expressed as the BMWP-Col of sites at different elevations.

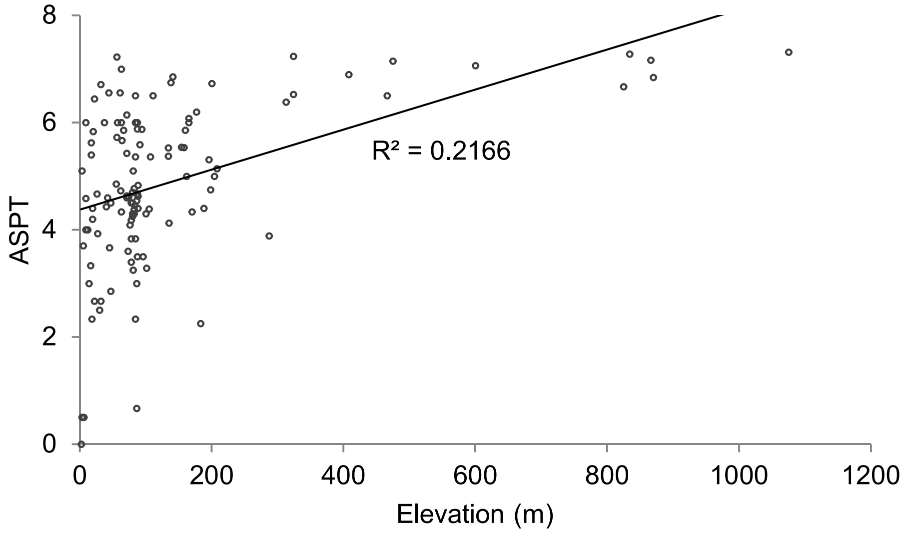
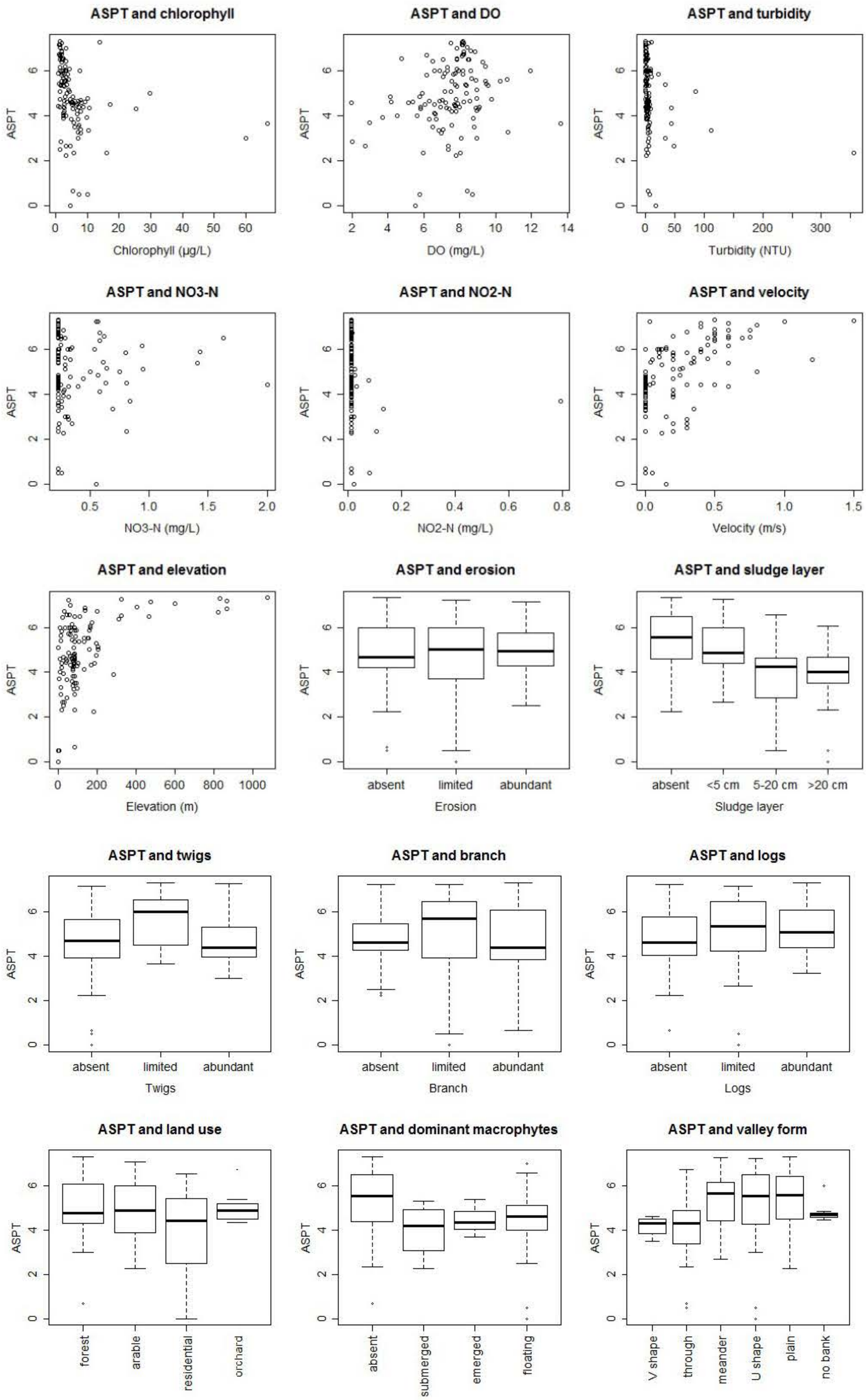


Figure C27 The ecological water quality expressed as the ASPT of sites at different elevations.


ASPT and chlorophyll
ASPT
Chlorophyll (µg/L)
ASPT and DO
DO (mg/L)
ASPT and turbidity
Turbidity (NTU)
ASPT and NO3-N
NO3-N (mg/L)
ASPT and NO2-N
NO2-N (mg/L)
ASPT and velocity
Velocity (m/s)
ASPT and elevation
Elevation (m)
ASPT and erosion
absent
limited
abundant
Erosion
ASPT and sludge layer
absent
<5 cm
5-20 cm
>20 cm
Sludge layer
ASPT and twigs
Twigs
ASPT and branch
Branch
ASPT and logs
Logs
ASPT and land use
forest
arable
residential
orchard
ASPT and dominant macrophytes
absent
submerged
emerged
floating
ASPT and valley form
V shape
through
meander
U shape
plain
no bank

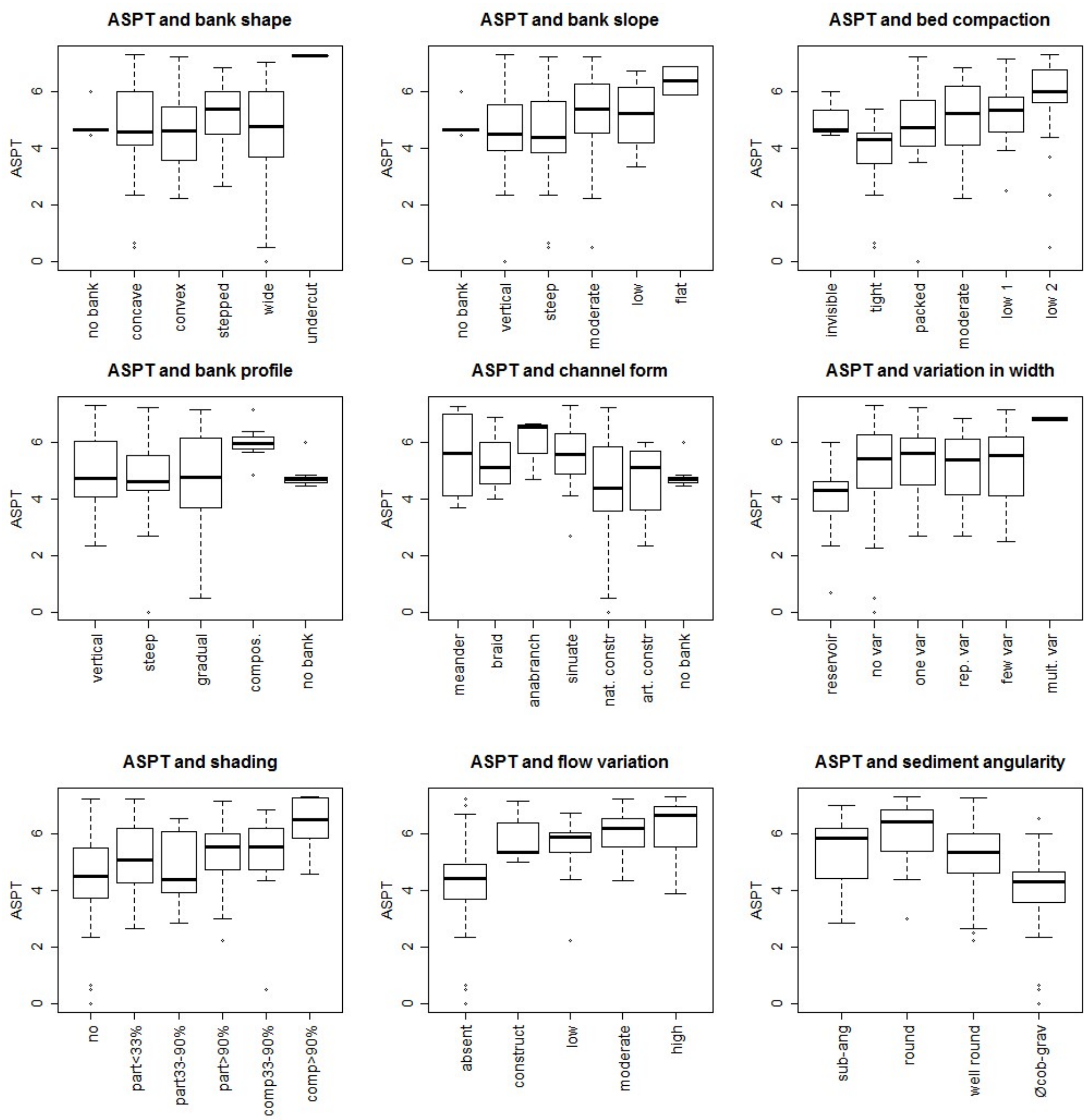


Figure C28 Plots showing the distribution of the data for physico-chemical variables in relation to ASPT, classification of categorical variables are based on Table B1; compos: composite, nat: natural, art: artificial, constr: construction, var: variation, part: partly, comp: completely, ang: angular, cob-grav: cobble-pebble-gravel.

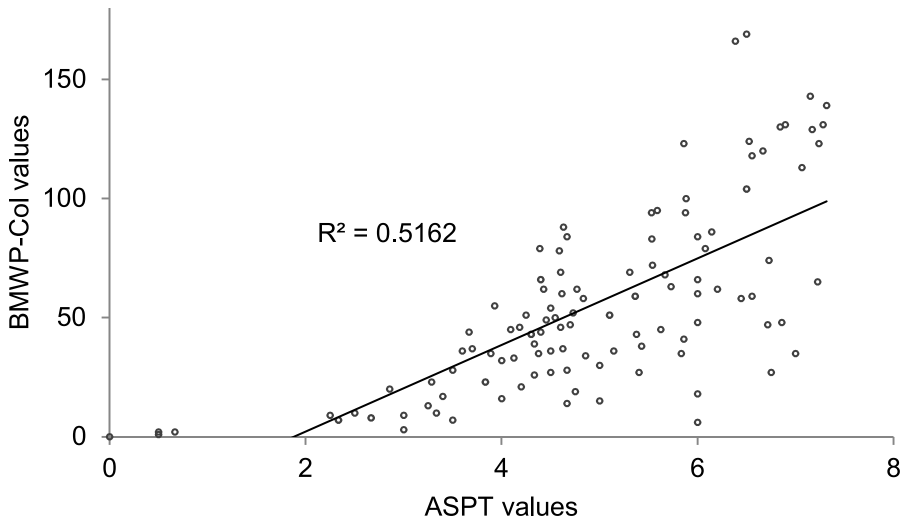
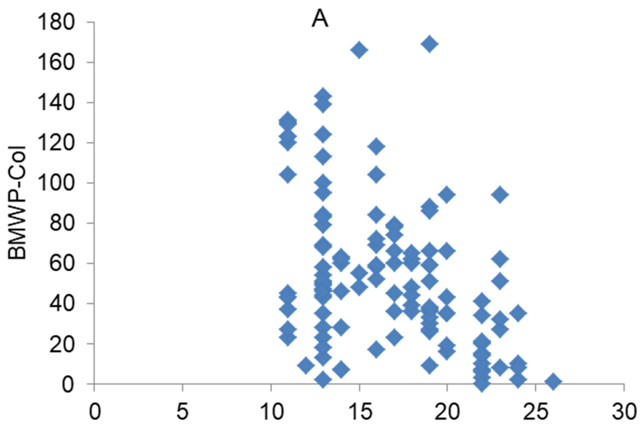


Figure C29 Correlation between BMWP-Col and ASPT and its coefficient of determination.



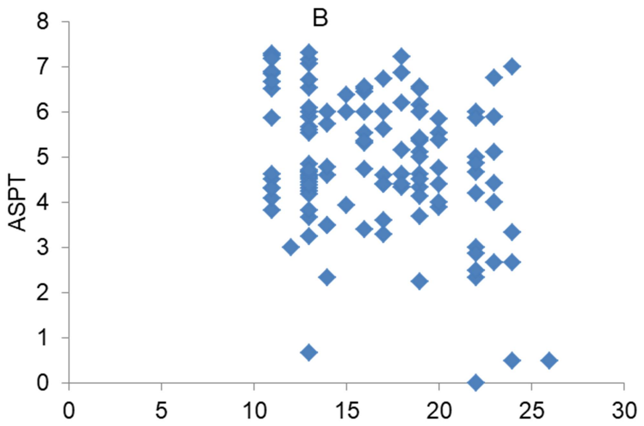


Figure C30 Habitat disturbance score in relation with BMWP-Col (A) and ASPT (B).

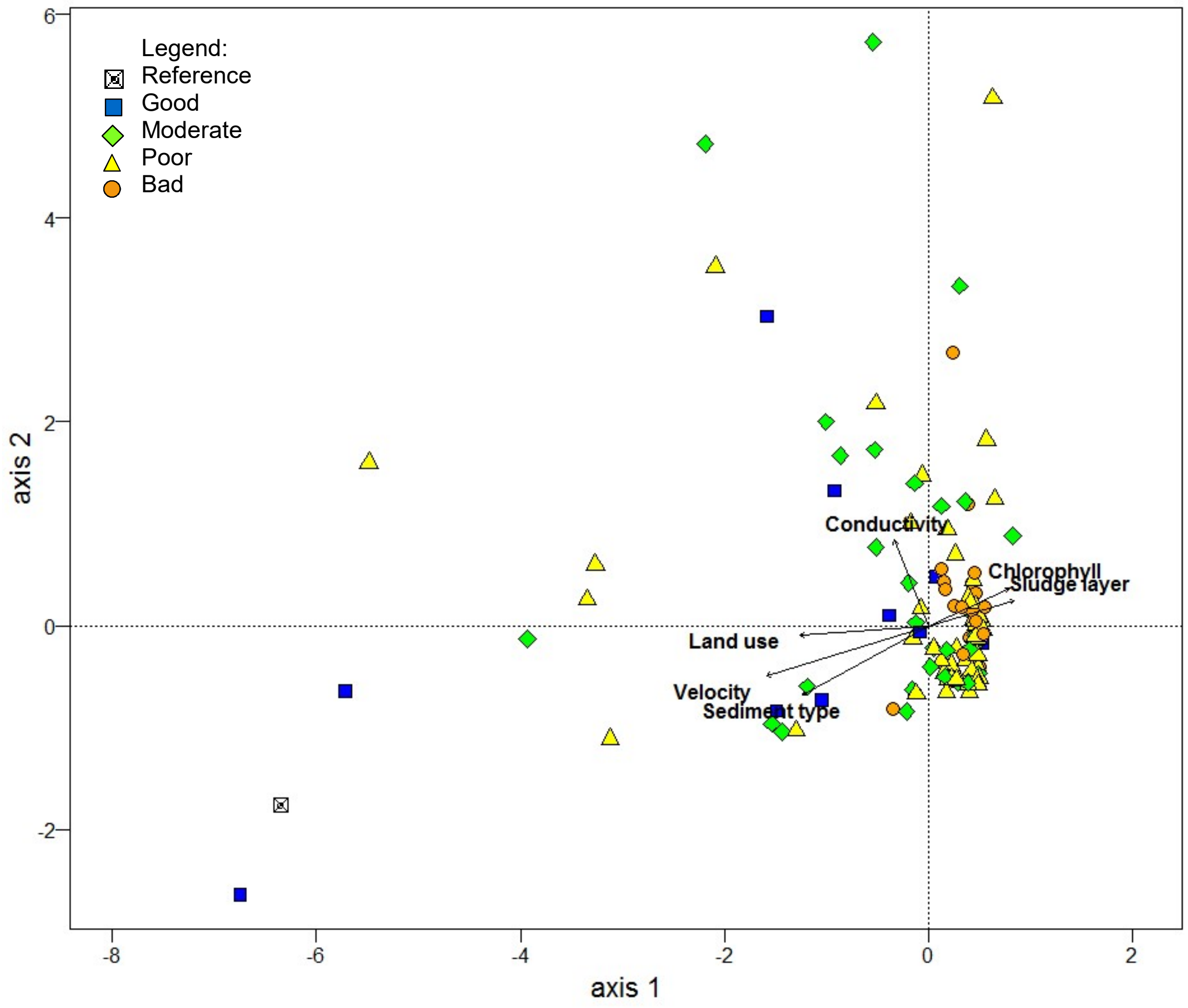


Figure C31 Correspondence analysis of taxa abundance (83 taxa) and fitted environmental variables in with indication of the ecological water quality of 119 sampling sites expressed as NLSMI ranging from reference to bad, as shown in the legend.

Tables

Table C1 Classification of water quality based on the BMWP-Col, NLSMI and oxygen Prati indices.

BMWP-Col		**NLSMI**		**Oxygen Prati**	
Values	**Category**	**Values**	**Category**	**Values**	**Category**
>100	Good	>8	Reference	>8	Heavily polluted
61 – 100	Moderate	6 – 8	Good	4 – 8	Polluted
36 – 60	Poor	4 – 6	Moderate	2 – 4	Moderately polluted
16 – 35	Bad	2 – 4	Poor	1 – 2	Acceptable
0 - 15	Very bad	<2	Bad	0 – 1	Unpolluted

Table C2 Habitat disturbance criteria and scoring list, adapted from Barbour *et al.* (1999), Hruby (2004), USEPA (2002) and Mereta *et al.* (2013).

Disturbance		**Score = 1**	**Score = 2**	**Score = 3**	**Score = 4**
Habitat alteration	Grazing	Minimal grazing	Moderate grazing	Intensive grazing	
	Vegetation removal	< 10 % vegetation removal	10 - 50 % of vegetation removal	> 50 % vegetation removal	
	Tree plantation	No tree plantation or plantation at > 50 m	Tree plantation at < 50 m but not in the wetland	Tree plantation in the wetland	
	Grading	no grading	grading near the wetland	grading within the wetland	
	Filling	No filling	Filling near the wetland	Filling in the wetland	
Land scape	Land use	forested	pasture	arable	residential
	Farming	No farming	Less intensive farming	Intensive farming	
	Soil mining	No soil mining	Soil mining > 50 m	Soil mining in the wetland or < 50 m	
	Wastewater discharge	No wastewater discharge into the river	Treated wastewater discharge into the river	Untreated wastewater discharge into the river	
Hydrological modification	Water inlet	natural source	berm	dam	
	Draining and water abstraction	no draining nor abstraction	Draining nearby < 50 m	Draining in the wetland	

Table C3 List of all families, tolerance scores, number of presences in the samples and functional feeding group (FFG) encountered in the Guayas river basin.

Taxa	BMWP-Col tolerance score	# presence	FFG
Acari	0	56	predator
Aeshnidae	6	11	predator
Ampullariidae	9	6	scraper
Ancylidae	6	13	scraper
Baetidae	7	64	scraper
Belostomatidae	5	6	predator
Blepharoceridae	10	2	scraper
Caenidae	7	12	collector-gatherer
Calamoceratidae	10	2	shredder
Calopterygidae	7	14	predator
Cambaridae	0	7	collector-gatherer
Ceratopogonidae	3	22	collector-gatherer
Chaoboridae	0	2	predator
Chironomidae	2	100	collector-gatherer
Chordodidae	10	1	parasite
Coenagrionidae	7	50	predator
Corbiculidae	0	22	collector-filterer
Corixidae	7	28	predator
Corydalidae	6	11	predator
Coryphoridae	0	7	scraper
Crambidae	5	13	shredder
Culicidae	2	14	collector-gatherer
Dixidae	7	2	collector-gatherer
Dryopidae	7	2	scraper
Dugesiidae	0	33	parasite
Dytiscidae	9	13	predator
Elmidae	6	20	scraper
Empididae	4	4	collector-gatherer
Gerridae	8	26	predator
Glossiphoniidae	3	28	predator
Glossosomatidae	7	2	scraper
Gomphidae	10	17	predator
Gyrinidae	9	1	predator
Hebridae	8	1	predator
Helicopsychidae	8	14	scraper
Heteroceridae	0	1	scraper
Hyallelidae	7	26	shredder
Hydrobiidae	8	7	scraper
Hydrobioscidae	9	3	predator
Hydrometridae	3	1	predator
Hydrophilidae	3	13	predator
Hydropsychidae	7	31	collector-filterer
Hydroptilidae	7	12	collector-gatherer

Taxa	BMWP-Col tolerance score	# presence	FFG
Lampyridae	10	1	predator
Leptoceridae	8	27	shredder
Leptohyphidae	7	52	collector-gatherer
Leptophlebiidae	9	30	collector-gatherer
Libellulidae	6	55	predator
Limoniidae	3	14	collector-gatherer
Lumbriculidae	0	2	collector-gatherer
Lymnaeidae	4	10	scraper
Macroveliidae	0	4	predator
Megapodagrionidae	6	6	predator
Mesoveliidae	5	15	predator
Mysidae	0	1	scraper
Naucoridae	7	33	predator
Nereidae	0	1	-
Noteridae	4	3	predator
Notonectidae	7	17	predator
Ocypodidae	0	2	shredder
Odontoceridae	10	2	collector-gatherer
Oligoneuriidae	10	1	collector-filterer
Palaemonidae	8	3	scraper
Perlidae	10	12	predator
Philopotamidae	9	17	collector-filterer
Physidae	3	9	scraper
Planorbidae	5	8	scraper
Platystictidae	0	11	predator
Pleidae	8	7	predator
Polycentropodidae	9	6	collector-gatherer
Polymitarcidae	9	1	collector-gatherer
Psephenidae	10	17	scraper
Ptilodactylidae	10	4	shredder
Scirtidae	7	3	collector-gatherer
Simuliidae	8	8	collector-filterer
Sphaeriidae	4	1	collector-filterer
Staphylinidae	6	3	predator
Stratiomyidae	4	7	collector-gatherer
Tabanidae	5	6	predator
Thiaridae	5	36	scraper
Trichodactylidae	0	1	shredder
Tubificidae	1	29	collector-gatherer
Veliidae	8	40	predator

Table C4 Mean, minimum, maximum and standard deviation of continuous variables measured in the Guayas river basin for rivers lower than 250 m. Lowest detection limit is the lowest concentration detectable by the kit, % missing values in comparison with the number of sampling sites.

Variable	Mean	Min	Max	Std. deviation	Lowest detection limit	% missing values
Temperature (° C)	26.0	22.6	34.0	2.1	-	-
Conductivity (µS/cm)	263	71	1981	281	-	-
Total dissolved solids (g/L)	0.17	0.05	1.27	0.18	-	-
pH	7.7	6.6	8.9	0.5	-	-
Chlorophyll a (µg/L)	5.2	0.7	66.8	10.5	-	-
Dissolved oxygen (mg/L)	7.4	2.0	13.6	2.0	-	-
Turbidity (NTU)	13.4	0.0	355.6	44.1	-	-
Chemical oxygen demand (mg/L)	18.9	5.2	117.6	15.8	5	17
Total nitrogen (mg/L)	1.8	1.0	7.7	1.7	1	81
Total phosphorus (mg/L)	2.7	0.8	4.5	2.6	0.5	97
Nitrate-nitrogen (mg/L)	0.6	0.2	2.0	0.4	0.23	37
Nitrite-nitrogen (mg/L)	0.1	0.0	0.8	0.2	0.015	85
Ammonium-nitrogen (mg/L)	0.3	0.0	8.8	1.1	0.015	1
Flow velocity (m/s)	0.3	0.0	1.2	0.2	-	-
Elevation (m)	78	2	208	62	-	-

Table C5 Mean, minimum, maximum and standard deviation of continuous variables measured in the Guayas river basin for rivers higher than 250 m. Lowest detection limit is the lowest concentration detectable by the kit, % missing values in comparison with the number of sampling sites.

Variable	Mean	Min	Max	Std. deviation	Lowest detection limit	% missing values
Temperature (° C)	21.7	19.0	24.2	1.6	-	-
Conductivity (µS/cm)	136	67	291	59	-	-
Total dissolved solids (g/L)	0.09	0.05	0.19	0.04	-	-
pH	8.2	8.0	8.5	0.1	-	-
Chlorophyll a (µg/L)	2.5	1.1	13.8	3.4	-	-
Dissolved oxygen (mg/L)	8.4	8.1	9.3	0.3	-	-
Turbidity (NTU)	2.3	0.2	10.5	2.9	-	-
Chemical oxygen demand (mg/L)	9.9	6.3	13.9	3.1	5	69
Total nitrogen (mg/L)	-	-	-	-	1	100

Variable	Mean	Min	Max	Std. deviation	Lowest detection limit	% missing values
Total phosphorus (mg/L)	-	-	-	-	0.5	100
Nitrate-nitrogen (mg/L)	0.5	0.3	0.6	0.2	0.23	77
Nitrite-nitrogen (mg/L)	-	-	-	-	0.015	100
Ammonium-nitrogen (mg/L)	0.0	0.0	0.1	0.0	0.015	-
Flow velocity (m/s)	0.7	0.2	1.5	0.3	-	-
Elevation (m)	590	287	1075	270	-	-

Table C6 Mean, minimum, maximum and standard deviation of continuous variables measured in the Guayas river basin for reservoir. Lowest detection limit is the lowest concentration detectable by the kit, % missing values in comparison with the number of sampling sites.

Variable	Mean	Min	Max	Std. deviation	Lowest detection limit	% missing values
Temperature (° C)	27.7	25.6	29.7	1.0	-	-
Conductivity (µS/cm)	78	37	109	12	-	-
Total dissolved solids (g/L)	0.05	0.05	0.07	0.01	-	-
pH	7.3	6.7	7.9	0.3	-	-
Chlorophyll a (µg/L)	7.7	4.7	25.4	3.9	-	-
Dissolved oxygen (mg/L)	7.4	4.1	10.7	1.3	-	-
Turbidity (NTU)	4.5	1.1	9.5	1.8	-	-
Chemical oxygen demand (mg/L)	16.6	9.0	23.4	5.6	5	44
Total nitrogen (mg/L)	1.1	1.0	1.4	0.2	1	88
Total phosphorus (mg/L)	-	-	-	-	0.5	100
Nitrate-nitrogen (mg/L)	0.3	0.2	0.6	0.1	0.23	81
Nitrite-nitrogen (mg/L)	0.1	0.0	0.1	0.0	0.015	94
Ammonium-nitrogen (mg/L)	0.1	0.0	0.6	0.1	0.015	6
Flow velocity (m/s)	0.0	0.0	0.0	0.0	-	-
Elevation (m)	84	71	105	8	-	-

Table C7 Overview of water quality studies in middle and south American rivers based on macroinvertebrates, with the indication of the methodology being locally developed (*), optimized (**), or used.

Author(s) and year	Country of study	Assessment methodology
Morpurgo (1996)	Brazil	Saprobic index, Indice Biotico Esteso (IBE)
Astorga et al. (1997)	Costa Rica	Belgian Biotic Index (BBI), Indice Biologique Global (IBG), Biological Monitoring Working Party Score (BMWP), Average Score per Taxon (ASPT), Proposed Costa Rican Biotic Index ((P)CRBI)*
Jacobsen (1998)	Ecuador	BMWP
Ometo et al. (2000)	Brazil	Diversity index
Rodrigues Capítulo et al. (2001)	Argentina	Biotic Index for Pampean rivers and streams (IBPAMP)*, Diversity index, Chandler score, BMWP, Index for Macroinvertebrates of Pampean Rivers (IMRP)
Marques and Barbosa (2001)	Brazil	Multivariate methods
Buss et al. (2002)	Brazil	Multivariate methods, Diversity index
Fenoglio et al. (2002)	Nicaragua	IBE, Family Biotic Index (FBI), Diversity indices
Weigel et al. (2002)	Mexico	Multimetric index*
Iannacone et al. (2003)	Peru	Diversity index
Figueroa et al. (2003)	Chile	FBI
Roldán Pérez (2003)	Colombia	BMWP/Col*, ASPT/Col*
Jacobsen (2003)	Ecuador	Diversity index
de Drago et al. (2004)	Paraguay	Diversity index
Soldner et al. (2004)	Dominican Republic	BMWP

Author(s) and year	Country of study	Assessment methodology
Neri et al. (2005)	Brazil	Diversity index
Callisto et al. (2005)	Brazil	Diversity index
Paredes et al. (2005)	Peru	BMWP
Silveira et al. (2005)	Brazil	Diversity index, BMWP-ASPT
Paggi et al. (2006)	Argentina	Diversity index, IMRP
Ayres-Peres et al. (2006)	Brazil	Diversity index
Bond et al. (2006)	Mexico	Diversity index
Umana-Villalobos and Springer (2006)	Costa Rica	BMWP
Moya et al. (2007)	Bolivia	Multimetric index*
Henriques-de-Oliveira et al. (2007)	Brazil	Diversity index
Figueroa et al. (2007)	Chile	IBE, BMWP, Family Biotic Index (IBF), Stream Invertebrate Grade Number Average Level (SIGNAL)
Albertoni et al. (2007)	Brazil	Diversity index
Baptista et al. (2007)	Brazil	Multimetric Index for Serra dos Orgaos (SOMI)*, BMWP-CETEC
Buckup et al. (2007)	Brazil	Diversity indices
Lopez-Hernandez et al. (2007)	Mexico	WQI, (Extended Biotic Index) EBI
Furstenberger et al. (2008)	Brazil	Saprobic index, BMWP
Jacobsen and Marin (2008)	Bolivia	BMWP, ASTP, FBI
Miserendino et al. (2008)	Argentina	Diversity index, Biotic Monitoring Patagonian Streams (BMPS)
Ocon et al. (2008)	Argentina	Diversity index, IBPamp, IMRP
Buss and Borges (2008)	Brazil	BMWP

Author(s) and year	Country of study	Assessment methodology
Stein et al. (2008)	Costa Rica	BMWP-CR
Mugnai et al. (2008)	Brazil	Índice Biótico Estendido - Instituto Oswaldo Cruz (IBE-IOC)*, IBE
Mugnai et al. (2008)	Brazil	IBE-IOC**, IBE
Mancilla et al. (2009)	Chile	Diversity index
Cordova et al. (2009)	Chile	ChFBI
Acosta et al. (2009)	Ecuador & Peru	Calidad Ecológica de Ríos Altoandinos (CERA), ABI
Carvajal et al. (2009)	Colombia	Diversity index
Correa-Araneda et al. (2010)	Chile	Biotic Family Index (ChIBF)
Hepp et al. (2010)	Brazil	Diversity index
García-Alzate et al. (2010)	Colombia	Diversity index
Oliveira and Nessimian (2010)	Brazil	Diversity index
Bieger et al. (2010)	Brazil	FBI, BMWP
Pinilla (2010)	Colombia	Biotic Indices of Communities (BI)*, Limnological Conditions Index (LICOI)*
Chalar et al. (2011)	Uruguay	Trophic State Index for Benthic Invertebrates (TSI-BI)*
Suriano et al. (2011)	Brazil	Diversity index, BMWP
Ferreira et al. (2011)	Brazil	Benthic Multimetric Index (BMI)*, BMWP-CETEC
Mugnai et al. (2011)	Brazil	IBE-IOC
Moya et al. (2011)	Bolivia	Multimetric Index*
Dos Santos et al. (2011)	Bolivia & Argentina	*Yungas Biotic Index based on 4 taxa* (IBY-4)*, BMWP, ASPT
Oliveira et al. (2011)	Brazil	Guapiac, u-Macau

Author(s) and year	Country of study	Assessment methodology
		Multimetric Index (GMMI)*
Dominguez-Granda et al. (2011)	Ecuador	BMWP, ASPT, BMWP/Col, BMWP/CR, SIGNAL, Nepalese Biotic Score (NEPBIOS), BMWP^{THAI}, South African Scoring System (SASS), SASS5, FBI, IBMWP, Diversity indices
Barbola et al. (2011)	Brazil	Diversity index
Guevara Mora (2011)	Costa Rica	BMWP-CR
Baptista et al. (2011)	Brazil	Piabanha-Paquequer-Preto Multimetric Index (PPPMI), Diversity index
Goncalves and de Menezes (2011)	Brazil	BMWP, BMWP-ASPT, Hilsenhoff
		Family Biotic Index (HFBI)
Couceiro et al. (2012)	Central Amazon, Brazil	Multimetric index*
Armendariz et al. (2012)	Argentina	Diversity index
Gomez et al. (2012)	Argentina	Index of Biotic Integrity for the Río de la Plata (IBIRP)*
Fierro et al. (2012)	Chile	Hilsenhoff's Index, modified FBI
Ocon and Rodrigues Capítulo (2012)	Argentina	IBPamp
Alvial et al. (2012)	Chile	ChBMWP, ChIBF
Villamarin et al. (2013)	Ecuador & Peru	Índice Multimétrico del Estado Ecológico para Ríos Altoandinos
		(IMEERA)*
Holguin-Gonzalez et al. (2013a)	Ecuador	Biotic Integrity Index using aquatic invertebrates (IBIAP)
Holguin-Gonzalez et al. (2013b)	Colombia	BMWP-Col,

Author(s) and year	Country of study	Assessment methodology
Helson and Williams (2013)	Panama	Neotropical Low-land Stream Multimetric Index (NLSMI)*
Barba-Alvarez et al. (2013)	Mexico	Hilsenhoff Biotic Index (HBI)
Alvarez-Mieles et al. (2013)	Ecuador	BMWP/Col
Baptista et al. (2013)	Brazil	SOMI, Serra da Bocaina Multimetric Index (MISB)*
Rizo-Patron et al. (2013)	Costa Rica	BMWP –CR
Molozzi et al. (2013)	Brazil	Diversity indices
Valle et al. (2013)	Brazil	IBE-IOC
Rosa et al. (2013)	Brazil	Diversity index
Trama and Mejía Marcacuzco (2013)	Peru	Diversity index
Sobczak et al. (2013)	Brazil	Diversity index, BMWP
Forero-Cespedes et al. (2013)	Colombia	BMWP/Col, WQI
Piñón Flores et al. (2014)	Mexico	Index of Biotic Integrity for Aquatic Macroinvertebrate Associations (IBIAMA)
Rios-Touma et al. (2014)	Ecuador & Peru	Andean Biotic Index (ABI)**, BMWP
Martinez-Rodriguez and Pinilla-A (2014)	Colombia	Biotic Index of Pollution (BIP), Biotic Integrity Index of Macroinvertebrates (BIIM)
Forero et al. (2014)	Colombia	Índice de Calidad Ecológica con base en macroinvertebrados acuáticos para la cuenca del rio Negro (ICE RN-MAE)*
Calderon et al. (2014)	Argentina	Sierra of San Luis Macroinvertebrates Biotic Index (MBI)
Gutierrez-Fonseca and Lorion (2014)	Costa Rica	BMWP-CR

Author(s) and year	Country of study	Assessment methodology
Reyes-Morales and Springer (2014)	Guatemala	BMWP/Atitlán index
Uherek and Pinto Gouveia (2014)	Brazil	BMWP
Forio et al. (2015)	Ecuador	BMWP/Col
Rocha et al. (2015)	Brazil	BMWP, Water Quality Index (WQI)**
Melo et al. (2015)	Brazil	Multimetric index
Dedieu et al. (2015)	French Guiana	*Indice Biotique Macroinvertébrés de Guyane* (IBMG)*
Nguyen et al. (2015)	Ecuador	Diversity index
Selvanayagam and Abril (2015)	Ecuador	Diversity indices, ABI, BMWP

References for Table C7

Acosta, R., Rios, B., Rieradevall, M., Prat, N., 2009. Proposal for an evaluation protocol of the ecological quality of Andean rivers (CERA) and its use in two basins in Ecuador and Peru. Limnetica 28(1), 35-64.

Albertoni, E.F., Prellvitz, L.J., Palma-Silva, C., 2007. Macroinvertebrate fauna associated with Pistia stratiotes and Nymphoides indica in subtropical lakes (south Brazil). Braz. J. Biol. 67(3), 499-507.

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

Alvial, I.E., Tapia, D.H., Castro, M.J., Duran, B.C., Verdugo, C.A., 2012. Analysis of benthic macroinvertebrates and biotic indices to evaluate water quality in rivers impacted by mining activities in northern Chile. Knowl. Manag. Aquat. Ecosyst.

Armendariz, L., Ocon, C., Capitulo, A.R., 2012. Potential responses of oligochaetes (Annelida, Clitellata) to global changes: Experimental fertilization in a lowland stream of Argentina (South America). Limnologica 42(2), 118-126.

Astorga, Y., de Pauw, N., Coto, J., Persoone, G., Castillo, L.E., Solis, J., Beyst, B., Lambert, V., Amparado, R., van Wichelen, T., 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.

Ayres-Peres, L., Sokolowicz, C.C., Santos, S., 2006. Diversity and abundance of the benthic macrofauna in lotic environments from the central region of Rio Grande do Sul state, Brazil. Biota Neotrop. 6.

Baptista, D.F., Buss, D.F., Egler, M., Giovanelli, A., Silveira, M.P., 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.

Baptista, D.F., de Souza, R.S.G., Vieira, C.A., Mugnai, R., Souza, A.S., de Oliveira, R.B.S., 2011. Multimetric index for assessing ecological condition of running waters in the upper reaches of the Piabanha-Paquequer-Preto Basin, Rio de Janeiro, Brazil. Zoologia 28(5), 619-628.

Baptista, D.F., Henriques-Oliveira, A.L., Oliveira, R.B.S., Mugnai, R., Nessimian, J.L., Buss, D.F., 2013. Development of a benthic multimetric index for the Serra da Bocaina bioregion in Southeast Brazil. Braz. J. Biol. 73(3), 573-583.

Barba-Alvarez, R., De la Lanza-Espino, G., Contreras-Ramos, A., Gonzalez-Mora, I., 2013. Aquatic insects indicators of water quality in Mexico: study cases, Copalita, Zimatan and Coyula rivers, Oaxaca. Rev. Mex. Biodivers. 84(1), 381-383.

Barbola, I.F., Moraes, M.F.P.G., Anazawa, T.M., Nascimento, E.A., Sepka, E.R., Polegatto, C.M., Milleo, J., Schuhli, G.S., 2011. Evaluation of the aquatic macroinvertebrate community as a tool for monitoring a reservoir in the Pitangui river basin, Parana, Brazil. Iheringia Ser. Zool. 101(1-2), 15-23.

Bieger, L., Carvalho, A.B.P., Strieder, M.N., Maltchik, L., Stenert, C., 2010. Are the streams of the Sinos River basin of good water quality? Aquatic macroinvertebrates may answer the question. Braz. J. Biol. 70(4), 1207-1215.

Bond, J.G., Novelo-Gutierrez, R., Ulloa, A., Rojas, J.C., Quiroz-Martinez, H., Williams, T., 2006. Diversity, abundance, and disturbance response of odonata associated with breeding sites of Anopheles pseudopunctipennis (Diptera : Culicidae) in southern Mexico. Environ. Entomol. 35(6), 1561-1568.

Buckup, L., Bueno, A.A.R., Bond-Buckup, G., Casagrande, M., Majolo, F., 2007. The benthic macroinvertebrate fauna of highland streams in southern Brazil: composition, diversity and structure. Rev. Bras. Zool. 24(2), 294-301.

Buss, D.F., Baptista, D.F., Silveira, M.P., Nessimian, J.L., Dorville, L.F.M., 2002. Influence of water chemistry and environmental degradation on macroinvertebrate assemblages in a river basin in south-east Brazil. Hydrobiologia 481(1-3), 125-136.

Buss, D.F., Borges, E.L., 2008. Application of Rapid Bioassessment Protocols (RBP) for benthic macroinvertebrates in Brazil: Comparison between sampling techniques and mesh sizes. Neotrop. Entomol. 37(3), 288-295.

Calderon, M.R., Gonzalez, P., Moglia, M., Gonzales, S.O., Jofre, M., 2014. Use of multiple indicators to assess the environmental quality of urbanized aquatic surroundings in San Luis, Argentina. Environ. Monit. Assess. 186(7), 4411-4422.

Callisto, M., Goulart, M., Barbosa, F.A., Rocha, O., 2005. Biodiversity assessment of benthic macroinvertebrates along a reservoir cascade in the lower Sao Francisco river (northeastern Brazil). Braz. J. Biol. 65(2), 229-240.

Carvajal, J.J., Moncada, L.I., Rodriguez, M.H., Perez, L.D., Olano, V.A., 2009. Characterization of Aedes albopictus (Skuse, 1894) (Diptera:Culicidae) larval habitats near the Amazon River in Colombia. Biomedica 29(3), 413-423.

Chalar, G., Arocena, R., Pacheco, J.P., Fabian, D., 2011. Trophic assessment of streams in Uruguay: A Trophic State Index for Benthic Invertebrates (TSI-BI). Ecol. Indic. 11(2), 362-369.

Cordova, S., Gaete, H., Aranguiz, F., Figueroa, R., 2009. Water quality assessment in the Limache stream (central Chile), using bioindicators and bioassays. Lat. Am. J. Aquat. Res. 37(2), 199-209.

Correa-Araneda, F., Rivera, R., Urrutia, J., De Los Rios, P., Contreras, A., Encina-Montoya, F., 2010. Effects of an urban zone on the benthonic macroinvertebrate community of a fluvial ecosystem in southern Chile. Limnetica 29(2), 183-194.

Couceiro, S.R.M., Hamada, N., Forsberg, B.R., Pimentel, T.P., Luz, S.L.B., 2012. A macroinvertebrate multimetric index to evaluate the biological condition of streams in the Central Amazon region of Brazil. Ecol. Indic. 18, 118-125.

de Drago, I.E., Marchese, M., Wantzen, K.M., 2004. Benthos of a large neotropical river: spatial patterns and species assemblages in the Lower Paraguay and its floodplains. Arch. Hydrobiol. 160(3), 347-374.

Dedieu, N., Clavier, S., Vigouroux, R., Cerdan, P., Céréghino, R., 2015. A Multimetric Macroinvertebrate Index for the Implementation of the European Water Framework Directive in French Guiana, East Amazonia. River Res. Appl.

Dominguez-Granda, L., Lock, K., Goethals, P.L.M., 2011. Using multi-target clustering trees as a tool to predict biological water quality indices based on benthic macroinvertebrates and environmental parameters in the Chaguana watershed (Ecuador). Ecol. Inform. 6(5), 303-308.

Dos Santos, D.A., Molineri, C., Reynaga, M.C., Basualdo, C., 2011. Which index is the best to assess stream health? Ecol. Indic. 11(2), 582-589.

Fenoglio, S., Badino, G., Bona, F., 2002. Benthic macroinvertebrate communities as indicators of river environment quality: an experience in Nicaragua. Rev. Biol. Trop. 50(3-4), 1125-1131.

Ferreira, W.R., Paiva, L.T., Callisto, M., 2011. Development of a benthic multimetric index for biomonitoring of a neotropical watershed. Braz. J. Biol. 71(1), 15-25.

Fierro, P., Bertran, C., Mercado, M., Pena-Cortes, F., Tapia, J., Hauenstein, E., Vargas-Chacoff, L., 2012. Benthic macroinvertebrate assemblages as indicators of water quality applying a modified biotic index in a spatio-seasonal context in a coastal basin of Southern Chile. Rev. Biol. Mar. Oceanogr. 47(1), 23-33.

Figueroa, R., Palma, A., Ruiz, V., Niell, X., 2007. Comparative analysis of biotic indexes used to evaluate water quality in a Mediterranean river of Chile: Chillan River, VIII Region. Rev. Chil. Hist. Nat. 80(2), 225-242.

Figueroa, R., Valdovinos, C., Araya, E., Parra, O., 2003. Benthic macroinvertebrates as indicators of water quality of southern Chile rivers. Rev. Chil. Hist. Nat. 76(2), 275-285.

Forero-Cespedes, A.M., Reinoso-Florez, G., Gutierrez, C., 2013. Water quality assessment of the Opia River (Tolima-Colombia), using macroinvertebrates and physicochemical parameters. Caldasia 35(2), 371-387.

Forero, L.C., Longo, M., Ramírez R., J.J., Chalar, G., 2014. Índice de calidad ecológica con base en macroinvertebrados acuáticos para la cuenca del río Negro (ICERN-MAE), Colombia. Rev. Biol. Trop. 62(2).

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

Furstenberger, C.B., Tomotake, M.E.M., Rodrigues, P.R.P., Moro, R.S., Quinaia, S.P., Caparica, R., Camargo, M., de Lima, A.G., 2008. Bioindication of water quality in flood and subnormal drainage periods of Pinhaozinho River, Guarapuava, Parana, Brazil. Aplicacoes Da Ficologia: Anais Do Xi Congresso Brasileiro De Ficologia E Simposio Latino-Americano Sobre Algas Nocivas 30, 149-166.

García-Alzate, C.A., Román-Valencia, C., Taphorn, D.C., Gonzalez, M.I., 2010. Physicochemical and biological characterization of the Roble river, Upper Cauca, western Colombia. Rev. Mus. Argent. Cienc. Nat. 12, 5-16.

Gomez, N., Licursi, M., Bauer, D.E., Ambrosio, E.S., Capitulo, A.R., 2012. Assessment of Biotic Integrity of the Coastal Freshwater Tidal Zone of a Temperate Estuary of South America through Multiple Indicators. Estuaries Coasts 35(5), 1328-1339.

Goncalves, F.B., de Menezes, M.S., 2011. A comparative analysis of biotic indices that use macroinvertebrates to assess water quality in a coastal river of Parana state, southern Brazil. Biota Neotrop. 11(4), 27-36.

Guevara Mora, M., 2011. Aquatic insects and water quality in Penas Blancas watershed and reservoir. Rev. Biol. Trop. 59(2), 635-654.

Gutierrez-Fonseca, P.E., Lorion, C.M., 2014. Application of the BMWP-Costa Rica biotic index in aquatic biomonitoring: sensitivity to collection method and sampling intensity. Rev. Biol. Trop. 62 (2), 275-289.

Helson, J.E., Williams, D.D., 2013. Development of a macroinvertebrate multimetric index for the assessment of low-land streams in the neotropics. Ecol. Indic. 29, 167-178.

Henriques-de-Oliveira, C., Baptista, D.F., Nessimian, J.L., 2007. Sewage input effects on the macroinvertebrate community associated to Typha domingensis Pers in a coastal lagoon in southeastern Brazil. Braz. J. Biol. 67(1), 73-80.

Hepp, L.U., Milesi, S.V., Biasi, C., Restello, R.M., 2010. Effects of agricultural and urban impacts on macroinvertebrates assemblages in streams (Rio Grande do Sul, Brazil). Zoologia 27(1), 106-113.

Holguin-Gonzalez, J.E., Boets, P., Alvarado, A., Cisneros, F., Carrasco, M.C., Wyseure, G., Nopens, I., Goethals, P.L.M., 2013a. Integrating hydraulic, physicochemical and ecological models to assess the effectiveness of water quality management strategies for the River Cuenca in Ecuador. Ecol. Modell. 254, 1-14.

Holguin-Gonzalez, J.E., Everaert, G., Boets, P., Galvis, A., Goethals, P.L.M., 2013b. Development and application of an integrated ecological modelling framework to analyze the impact of wastewater discharges on the ecological water quality of rivers. Environ. Model. Softw. 48, 27-36.

Iannacone, J., Mansilla, J., Ventura, K., 2003. Macroinvertebrados en las lagunas de Puerto Viejo en Lima, Perú. Ecol. Apl. 2(1), 116-124.

Jacobsen, D., 1998. The effect of organic pollution on the macroinvertebrate fauna of Ecuadorian highland streams. Arch. Hydrobiol. 143(2), 179-195.

Jacobsen, D., 2003. Altitudinal changes in diversity of macroinvertebrates from small streams in the Ecuadorian Andes. Arch. Hydrobiol. 158(2), 145-167.

Jacobsen, D., Marin, R., 2008. Bolivian Altiplano streams with low richness of macroinvertebrates and large diel fluctuations in temperature and dissolved oxygen. Aquat. Ecol. 42(4), 643-656.

Lopez-Hernandez, M., Ramos-Espinosa, M.G., Carranza-Fraser, J., 2007. Multimetric analyses for assessing pollution in the Lerma river and Chapala lake, Mexico. Hidrobiologica 17(1), 17-30.

Mancilla, G., Valdovinos, C., Azocar, M., Henriquez, M., Figueroa, R., 2009. Multimetric Approach to Water Quality Evaluation of Basins with Diferent Levels of Antropic Perturbation. Interciencia 34(12), 857-864.

Marques, M.M., Barbosa, F., 2001. Biological quality of waters from an impacted tropical watershed (middle Rio Doce basin, southeast Brazil), using benthic macroinvertebrate communities as an indicator. Hydrobiologia 457, 69-76.

Martinez-Rodriguez, M.D., Pinilla-A, A.A., 2014. Assessing the water quality of three wetlands in the Department of Cesar, Colombia, through aquatic macroinvertebrates associated with Eichhornia crassipes (Pontederiaceae). Caldasia 36(2), 305-321.

Melo, S., Stenert, C., Dalzochio, M.S., Maltchik, L., 2015. Development of a multimetric index based on aquatic macroinvertebrate communities to assess water quality of rice fields in southern Brazil. Hydrobiologia 742(1), 1-14.

Miserendino, M.L., Brand, C., Di Prinzio, C.Y., 2008. Assessing urban impacts on water quality, benthic communities and fish in streams of the Andes Mountains, Patagonia (Argentina). Water Air Soil Pollut. 194(1-4), 91-110.

Molozzi, J., Salas, F., Callisto, M., Marques, J.C., 2013. Thermodynamic oriented ecological indicators: Application of Eco-Exergy and Specific Eco-Exergy in capturing environmental changes between disturbed and non-disturbed tropical reservoirs. Ecol. Indic. 24, 543-551.

Morpurgo, M., 1996. Confronto fra Indice Saprobico e Indice Biotico Esteso. Biol. Ambient. 2(3), 30-36.

Moya, N., Hughes, R.M., Dominguez, E., Gibon, F.M., Goitia, E., Oberdorff, T., 2011. Macroinvertebrate-based multimetric predictive models for evaluating the human impact on biotic condition of Bolivian streams. Ecol. Indic. 11(3), 840-847.

Moya, N., Tomanova, S., 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.

Mugnai, R., Buss, D.F., Oliveira, R.B., Sanfins, C., Carvalho, A.L., Baptista, D.F., 2011. Application of the biotic index IBE-IOC for water quality assessment in wadeable streams in south-east Brazil. Acta Limnol. Bras. 23, 74-85.

Mugnai, R., Oliveira, R.B., Carvalho, A.D., Baptista, D.F., 2008. Adaptation of the Indice Biotico Esteso (IBE) for water quality assessment in rivers of Serra do Mar, Rio de Janeiro State, Brazil. Trop. Zool. 21(1), 57-74.

Neri, D.B., Kotzian, C.B., Siegloch, A.E., 2005. Composição de Heteroptera aquáticos e semi-aquáticos na área de abrangência da U.H.E. Dona Francisca, RS, Brasil: fase de pré-enchimento. Iheringia Ser. Zool. 95(4), 421-429.

Nguyen, T.H.T., Boets, P., Lock, K., Damanik-Ambarita, M.N., Forio, M.A.E., Sasha, P., Dominguez-Granda, L.E., Hoang, T.H.T., Everaert, G., Goethals, P.L.M., 2015. Habitat suitability of the invasive water hyacinth and its relation to water quality and macroinvertebrate diversity in a tropical reservoir. Limnologica 52, 67-74.

Ocon, C.S., Capitulo, A.R., Paggi, A.C., 2008. Evaluation of zoobenthic assemblages and recovery following petroleum spill in a coastal area of Rio de la Plata estuarine system, South America. Environ. Pollut. 156(1), 82-89.

Ocon, C.S., Rodrigues Capítulo, A., 2012. Assessment of water quality in temperate-plain streams (Argentina, South America) using a multiple approach. Ecol. Austral. 22, 81-91.

Oliveira, A.L.H., Nessimian, J.L., 2010. Spatial distribution and functional feeding groups of aquatic insect communities in Serra da Bocaina streams, southeastern Brazil. Acta Limnol. Bras. 22(4), 424-441.

Oliveira, R.B.S., Baptista, D.F., Mugnai, R., Castro, C.M., Hughes, R.M., 2011. Towards rapid bioassessment of wadeable streams in Brazil: Development of the Guapiacu-Macau Multimetric Index (GMMI) based on benthic. Ecol. Indic. 11(6), 1584-1593.

Ometo, J.P.H.B., Martinelli, L.A., Ballester, M.V., Gessner, A., Krusche, A.V., Victoria, R.L., Williams, M., 2000. Effects of land use on water chemistry and macroinvertebrates rates in two streams of the Piracicaba river basin, south-east Brazil. Freshw. Biol. 44(2), 327-337.

Paggi, A.C., Ocon, C., Tangorra, M., Capitulo, A.R., 2006. Response of the zoobenthos community along the dispersion plume of a highly polluted stream in the receiving waters of a large river (Rio de la Plata, Argentina). Hydrobiologia 568, 1-14.

Paredes, C., Iannacone, J., Alvariño, L., 2005. Uso de macroinvertebrados bentónicos como bioindicadores de la calidad de agua en el río Rímac, Lima-Callao, Perú. Rev. Colomb. Entomol. 31, 219-225.

Pinilla, G., 2010. An index of limnological conditions for urban wetlands of Bogota city, Colombia. Ecol. Indic. 10(4), 848-856.

Piñón Flores, M.A., Pérez Munguía, R.M., Torres García, U., Pineda López, R., 2014. Integridad biótica de la microcuenca del Río Chiquito, Morelia Michoacán, México, basada en el ensamblaje de macroinvertebrados acuáticos. Rev. Biol. Trop. 62(2), 219-229.

Reyes-Morales, F., Springer, M., 2014. [Effect of sampling effort on taxa richness of aquatic macroinvertebrates and the BMWP/Atitlan index]. Rev. Biol. Trop. 62 (2), 291-301.

Rios-Touma, B., Acosta, R., Prat, N., 2014. The Andean Biotic Index (ABI): revised tolerance to pollution values for macroinvertebrate families and index performance evaluation. Rev. Biol. Trop. 62 (2), 249-273.

Rizo-Patron, F., Kumar, A., Colton, M.B.M., Springer, M., Trama, F.A., 2013. Macroinvertebrate communities as bioindicators of water quality in conventional and organic irrigated rice fields in Guanacaste, Costa Rica. Ecol. Indic. 29, 68-78.

Rocha, F.C., Andrade, E.M., Lopes, F.B., 2015. Water quality index calculated from biological, physical and chemical attributes. Environ. Monit. Assess. 187(1).

Rodrigues Capítulo, A., Tangorra, M., Ocón, C., 2001. Use of benthic macroinvertebrates to assess the biological status of Pampean streams in Argentina. Aquat. Ecol. 35(2), 109-119.

Roldán Pérez, G., 2003. Bioindicación de la calidad del agua en Colombia : propuesta para el uso del método BMWP/Col, 1st ed. Editorial Universidad de Antioquia, Medellín, Colombia.

Rosa, B.F.J.V., Dias-Silva, M.V.D., Alves, R.G., 2013. Composition and Structure of the Chironomidae (Insecta: Diptera) Community Associated with Bryophytes in a First-Order Stream in the Atlantic Forest, Brazil. Neotrop. Entomol. 42(1), 15-21.

Selvanayagam, M., Abril, R., 2015. Water Quality Assessment of Piatua River Using Macroinvertebrates in Puyo, Pastaza, Ecuador. Am. J. Life Sci. 3(3), 167-174.

Silveira, M.P., Baptista, D.F., Buss, D.F., Nessimian, J.L., Egler, M., 2005. Application of biological measures for stream integrity assessment in south-east Brazil. Environ. Monit. Assess. 101(1-3), 117-128.

Sobczak, J.R.S., Valduga, A.T., Restello, R.M., Cardoso, R.I., 2013. Conservation unit and water quality: the influence of environmental integrity on benthic macroinvertebrate assemblages. Acta Limnol. Bras. 25, 442-450.

Soldner, M., Stephen, I., Ramos, L., Angus, R., Wells, N.C., Grosso, A., Crane, M., 2004. Relationship between macroinvertebrate fauna and environmental variables in small streams of the Dominican Republic. Water Res. 38(4), 863-874.

Stein, H., Springer, M., Kohlmann, B., 2008. Comparison of two sampling methods for biomonitoring using aquatic macroinvertebrates in the Dos Novillos River, Costa Rica. Ecol. Eng. 34(4), 267-275.

Suriano, M.T., Fonseca-Gessner, A.A., Roque, F.O., Froehlich, C.G., 2011. Choice of macroinvertebrate metrics to evaluate stream conditions in Atlantic Forest, Brazil. Environ. Monit. Assess. 175(1-4), 87-101.

Trama, F.A., Mejía Marcacuzco, J.A., 2013. Biodiversidad de macroinvertebrados bentónicos en el sistema de cultivo de arroz en el sector muñuela margen derecho en Piura, Perú. Ecol. Apl. 12(2).

Uherek, C.B., Pinto Gouveia, F.B., 2014. Biological Monitoring Using Macroinvertebrates as Bioindicators of Water Quality of Maroaga Stream in the Maroaga Cave System, Presidente Figueiredo, Amazon, Brazil. Int. J. Ecol. 2014(1-2), 1-7.

Umana-Villalobos, G., Springer, M., 2006. Environmental variation in the Grande de Terraba river and some of its tributaries, south Pacific of Costa Rica. Rev. Biol. Trop. 54, 265-272.

Valle, I.C., Buss, D.F., Baptista, D.F., 2013. The influence of connectivity in forest patches, and riparian vegetation width on stream macroinvertebrate fauna. Braz. J. Biol. 73(2), 231-238.

Villamarin, C., Rieradevall, M., Paul, M.J., Barbour, M.T., Prat, N., 2013. A tool to assess the ecological condition of tropical high Andean streams in Ecuador and Peru: The IMEERA index. Ecol. Indic. 29, 79-92.

Weigel, B.M., Henne, L.J., Martinez-Rivera, L.M., 2002. Macroinvertebrate-based index of biotic integrity for protection of streams in west-central Mexico. J. N. Am. Benthol. Soc. 21(4), 686-700.

D – Supporting information for chapter 5

Figures

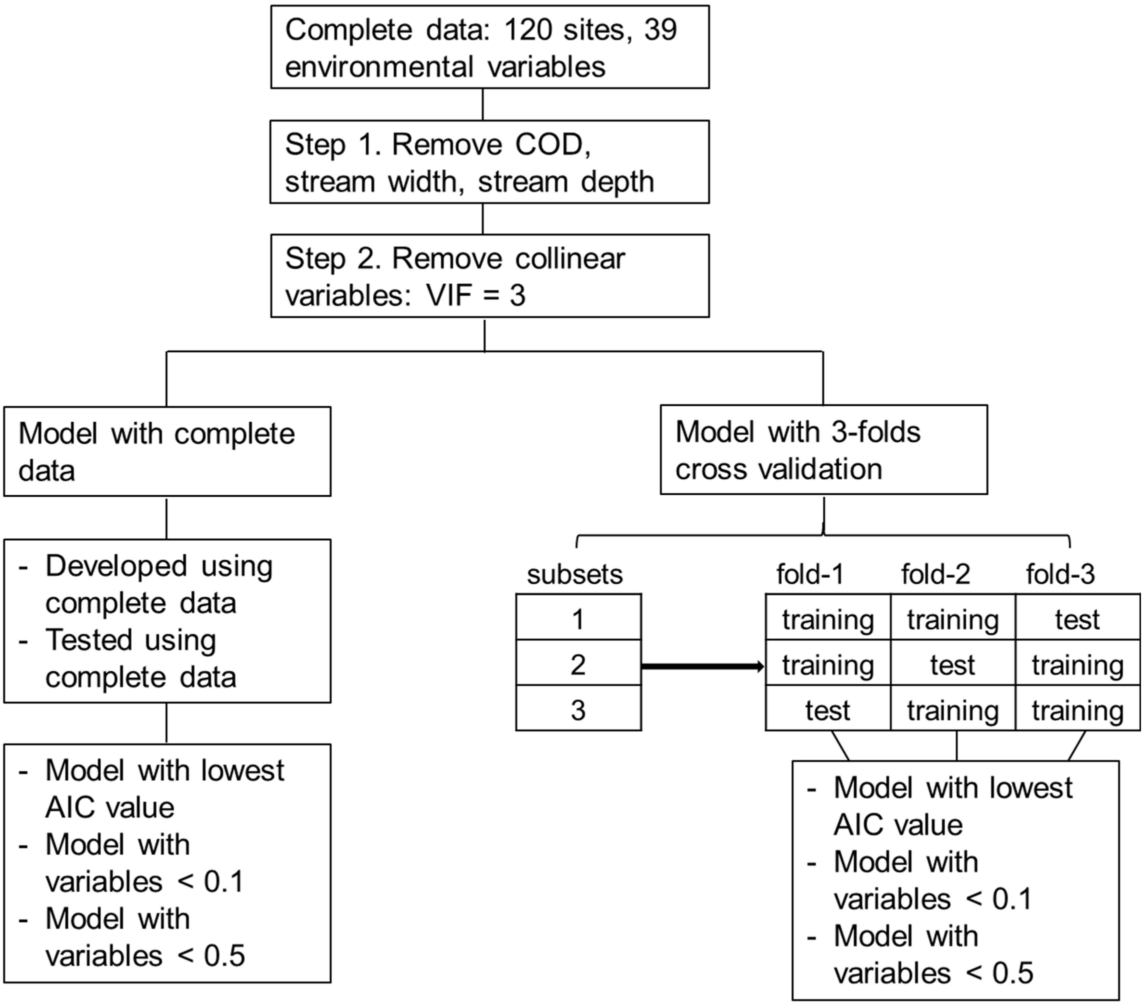
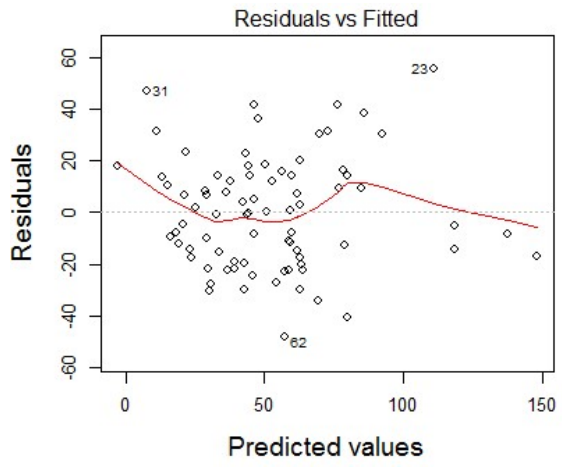
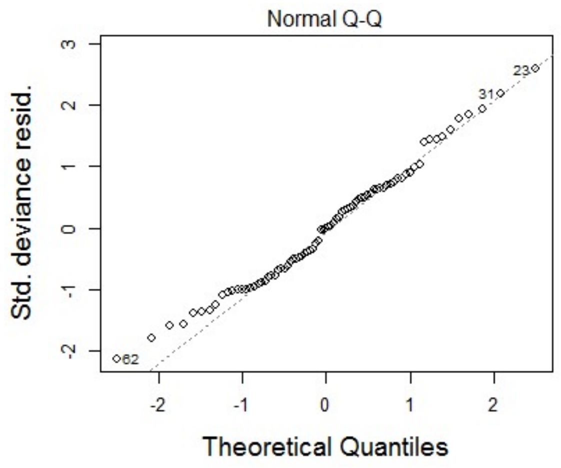
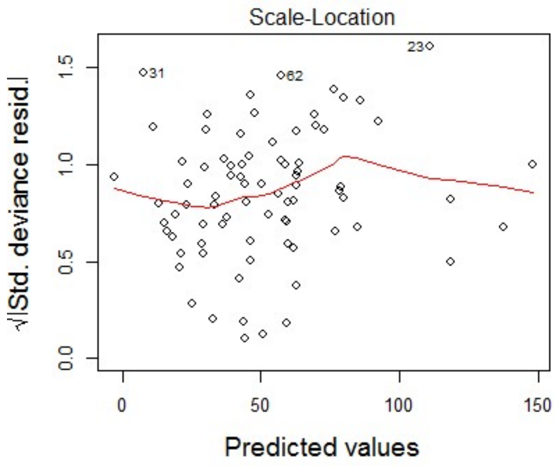


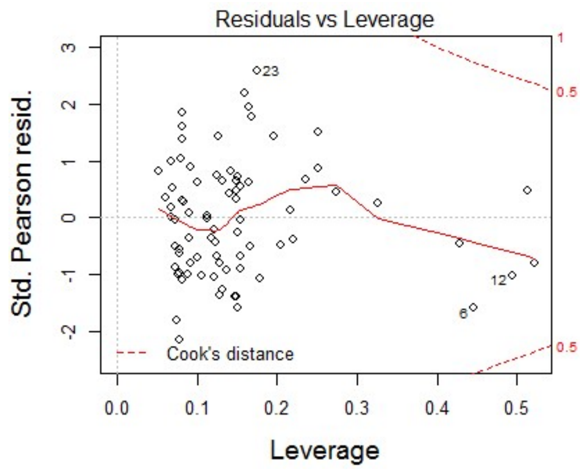
Figure D1 Scheme for model development and criteria for the final models.





A B





C D

Figure D2 Residuals plots of model based on folds training set 1 + 2 with lowest AIC, (A) residuals versus fitted values; (B) QQ-plot for normality; (C) scaled residuals versus fitted values; (D) standardized residuals versus leverage.

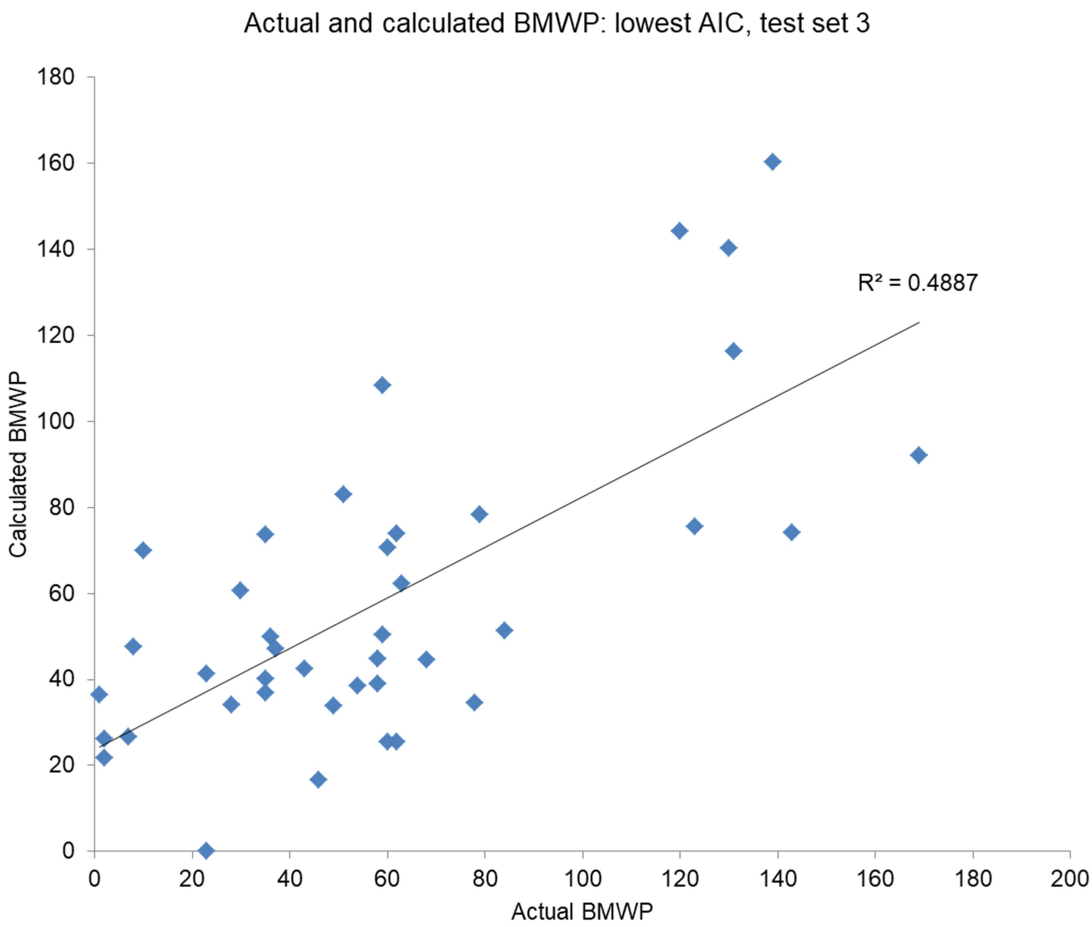


Figure D3 Validation of model based on folds test set 3 with lowest AIC.

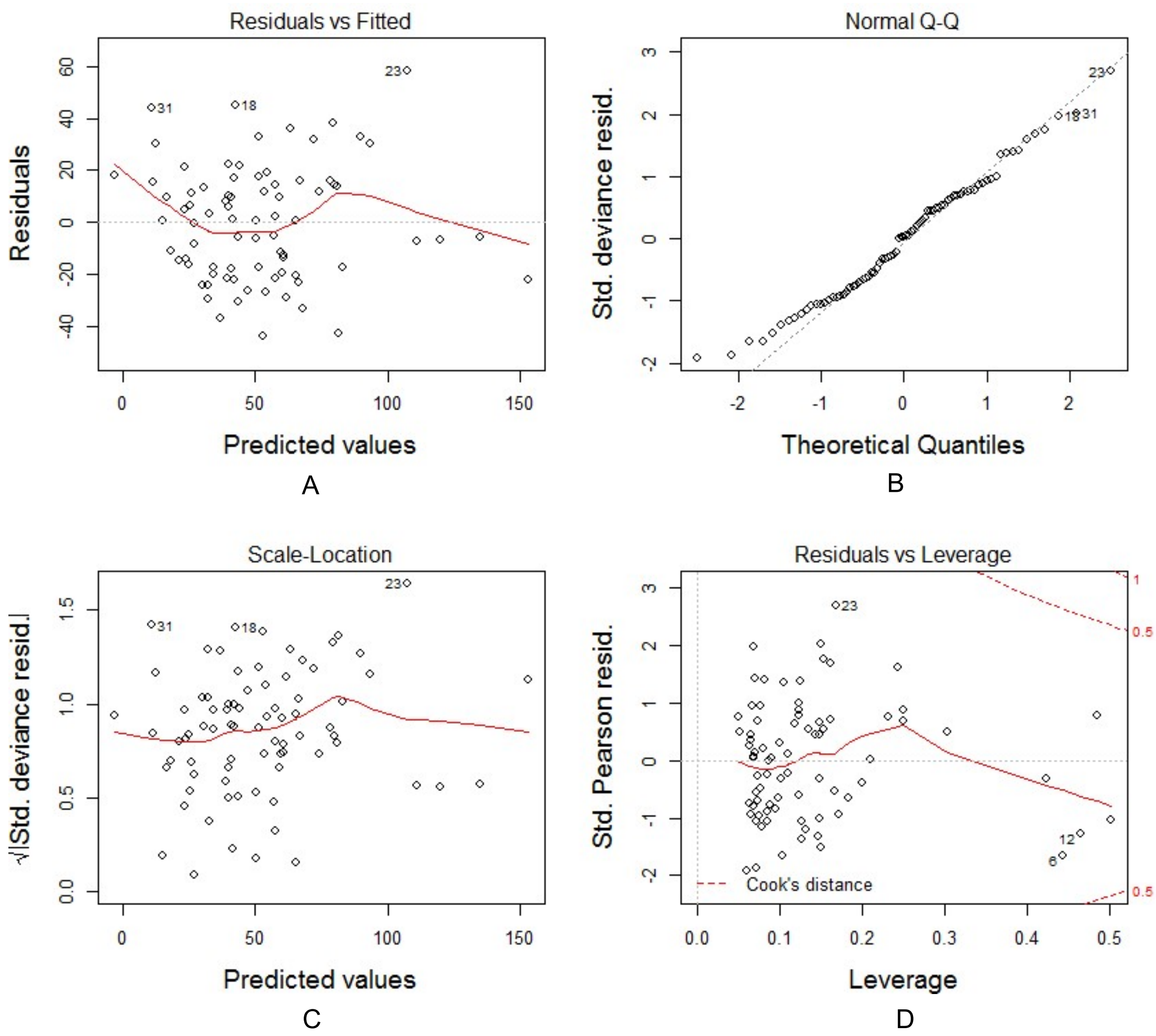


Figure D4 Residuals plots of model based on folds training set 1 + 2 with input variables significant at $p < 0.1$, (A) residuals versus fitted values; (B) QQ-plot for normality; (C) scaled residuals versus fitted values; (D) standardized residuals versus leverage.

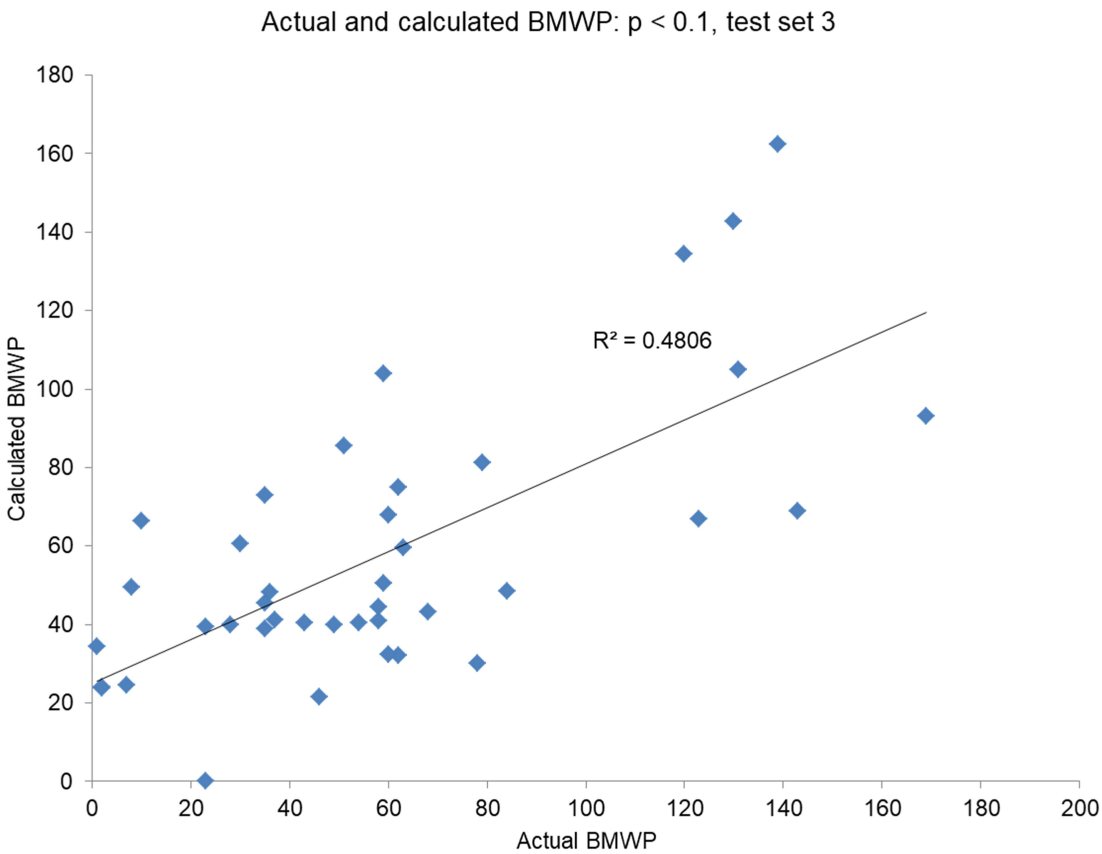


Figure D5 Validation of model based on folds test set 3 with input variables significant at $p < 0.1$.

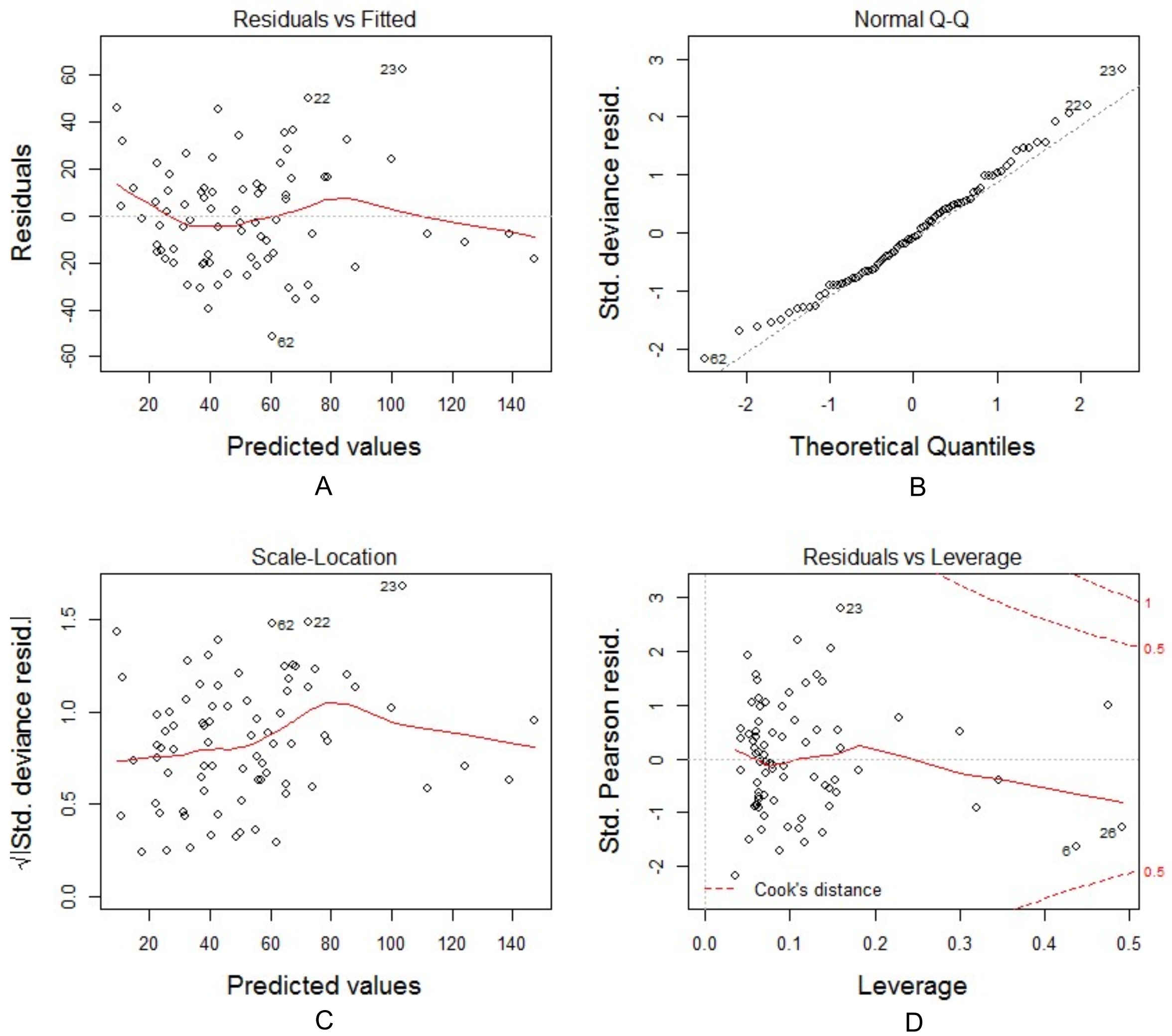


Figure D6 Residuals plots of model based on folds training set 1 + 2 with input variables significant at $p < 0.05$, (A) residuals versus fitted values; (B) QQ-plot for normality; (C) scaled residuals versus fitted values; (D) standardized residuals versus leverage.

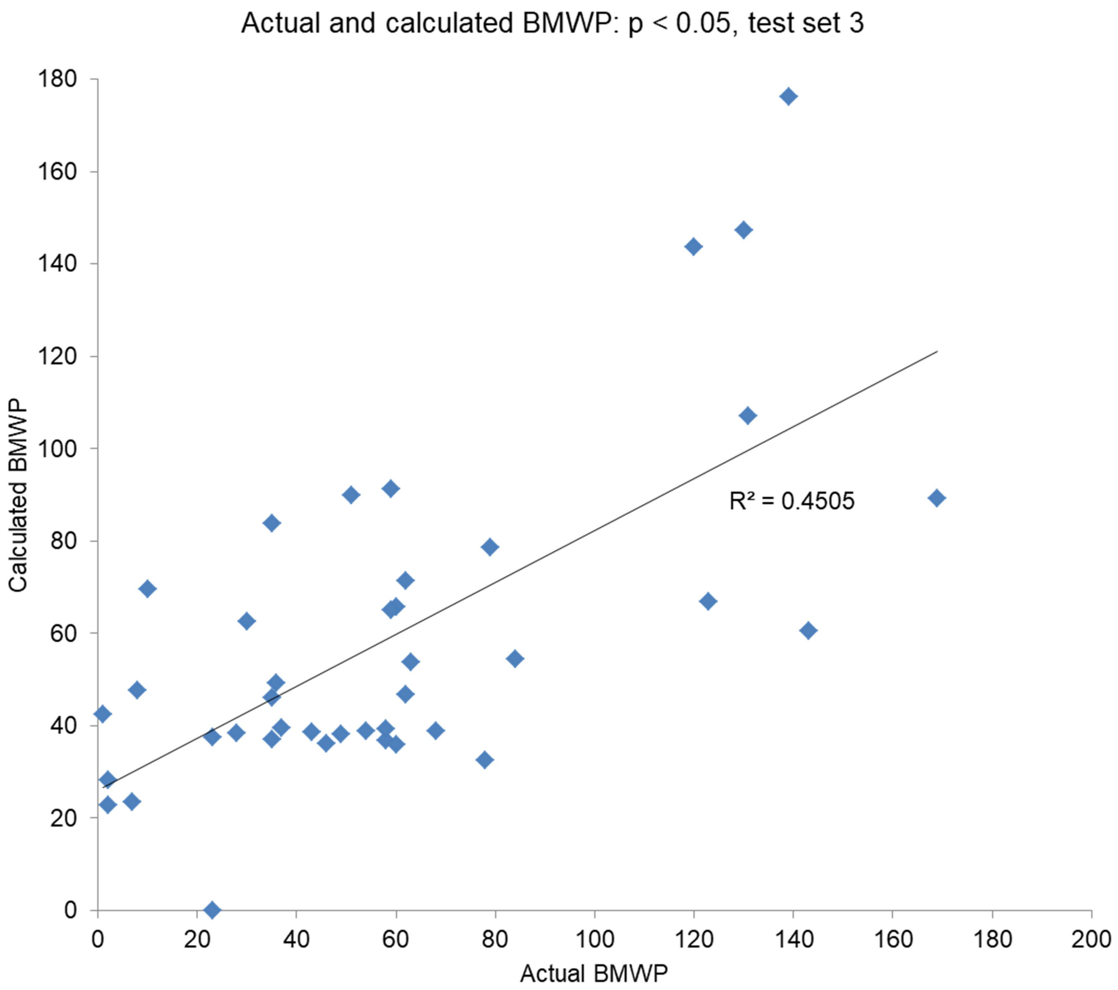


Figure D7 Validation of model based on folds test set 3 with input variables significant at $p < 0.05$.

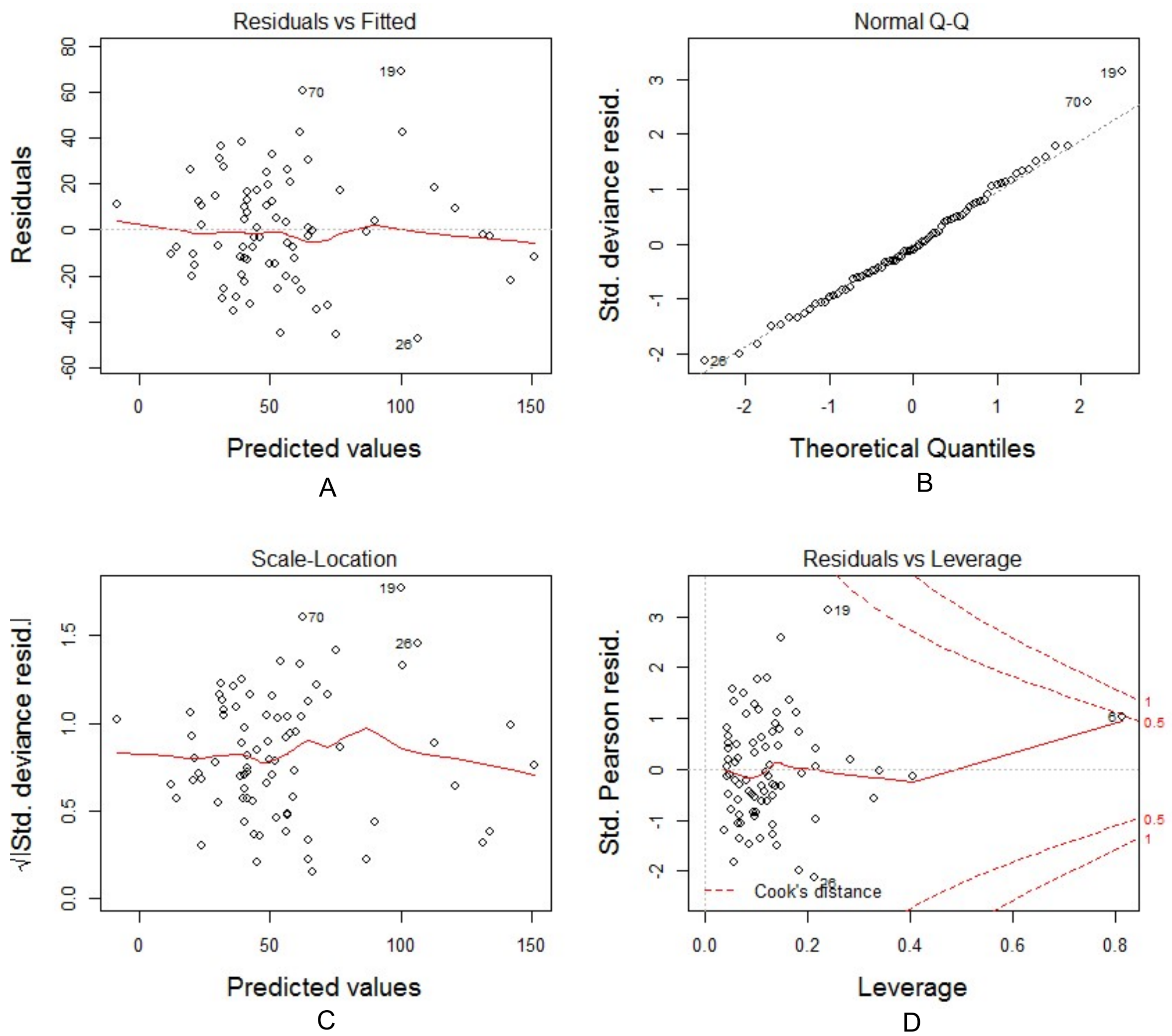


Figure D8 Residuals plots of model based on folds training set 1 + 3 with lowest AIC, (A) residuals versus fitted values; (B) QQ-plot for normality; (C) scaled residuals versus fitted values; (D) standardized residuals versus leverage.

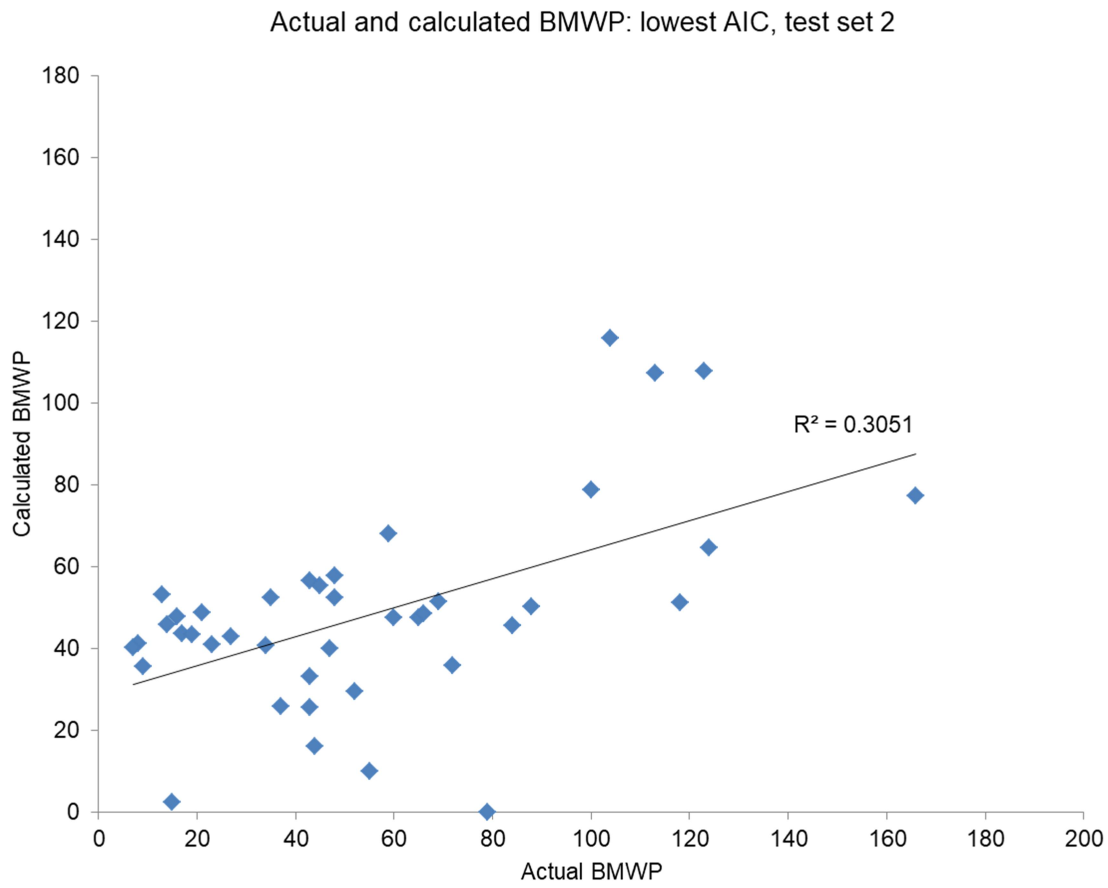


Figure D9 Validation of model based on folds test set 2 with lowest AIC.

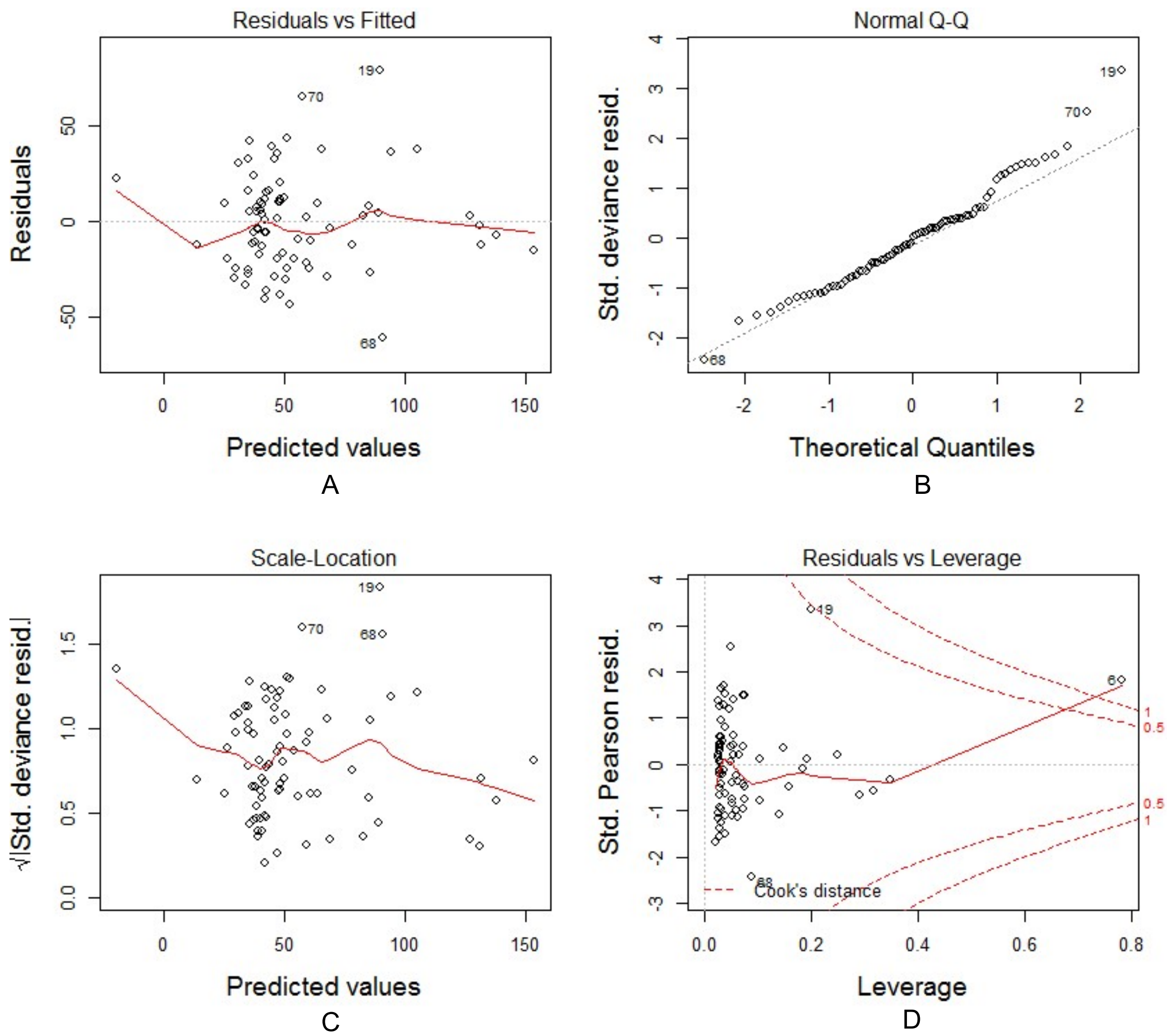


Figure D10 Residuals plots of model based on folds training set 1 + 3 with input variables significant at $p < 0.1$, (A) residuals versus fitted values; (B) QQ-plot for normality; (C) scaled residuals versus fitted values; (D) standardized residuals versus leverage.

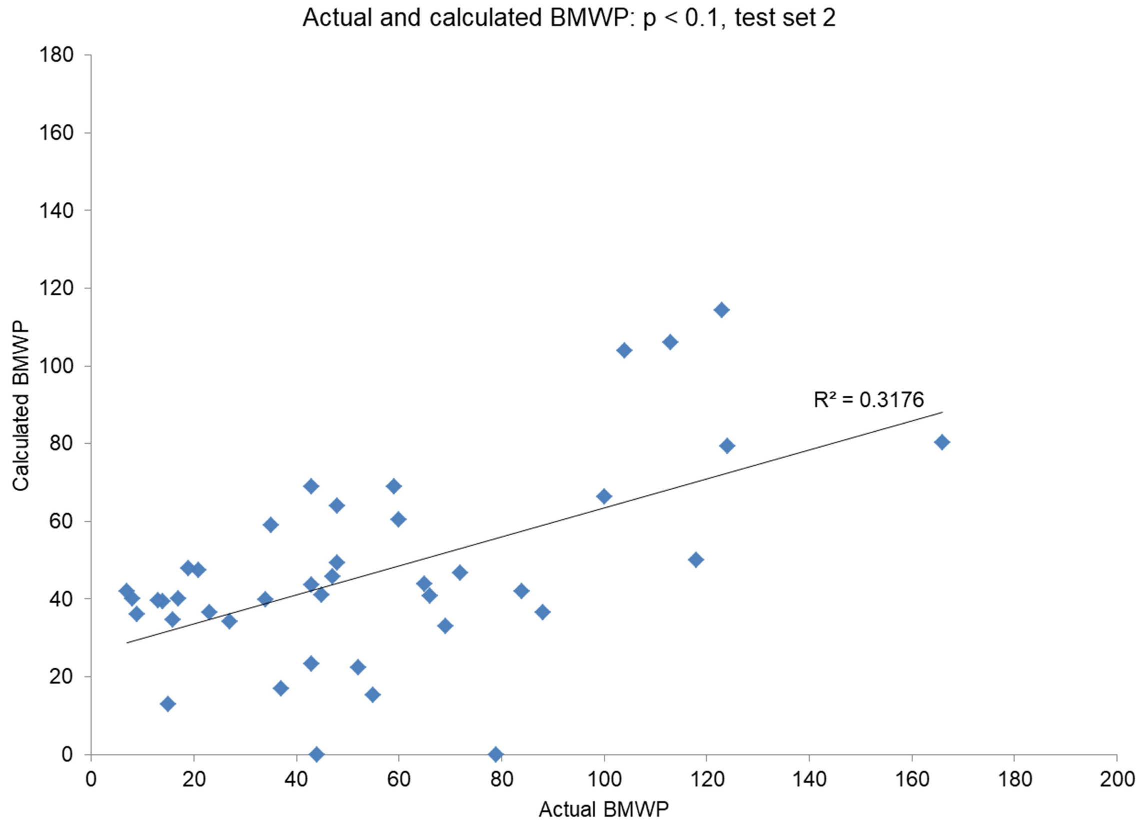


Figure D11 Validation of model based on folds test set 2 with input variables significant at $p < 0.1$.

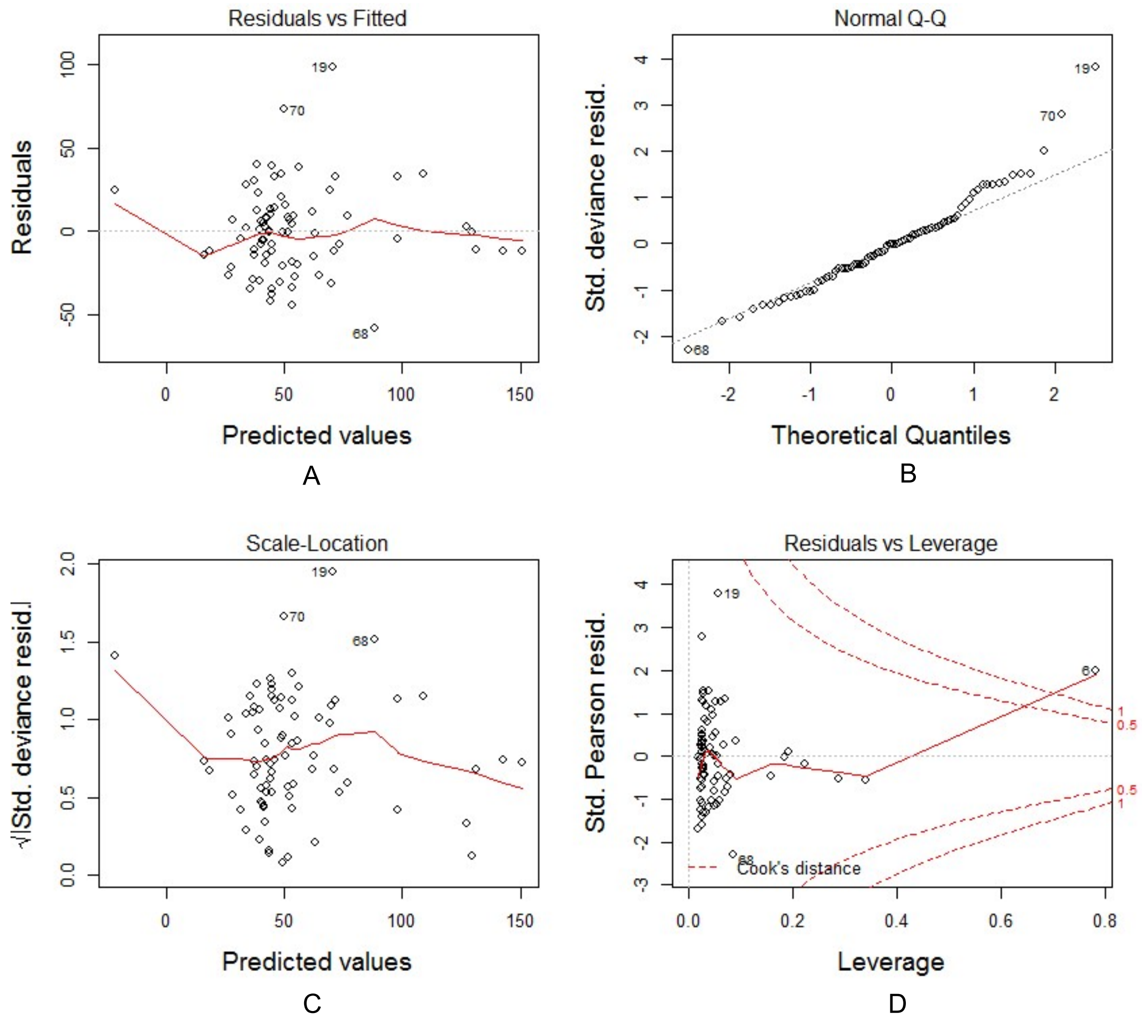


Figure D12 Residuals plots of model based on folds training set 1 + 3 with input variables significant at $p < 0.05$, (A) residuals versus fitted values; (B) QQ-plot for normality; (C) scaled residuals versus fitted values; (D) standardized residuals versus leverage.

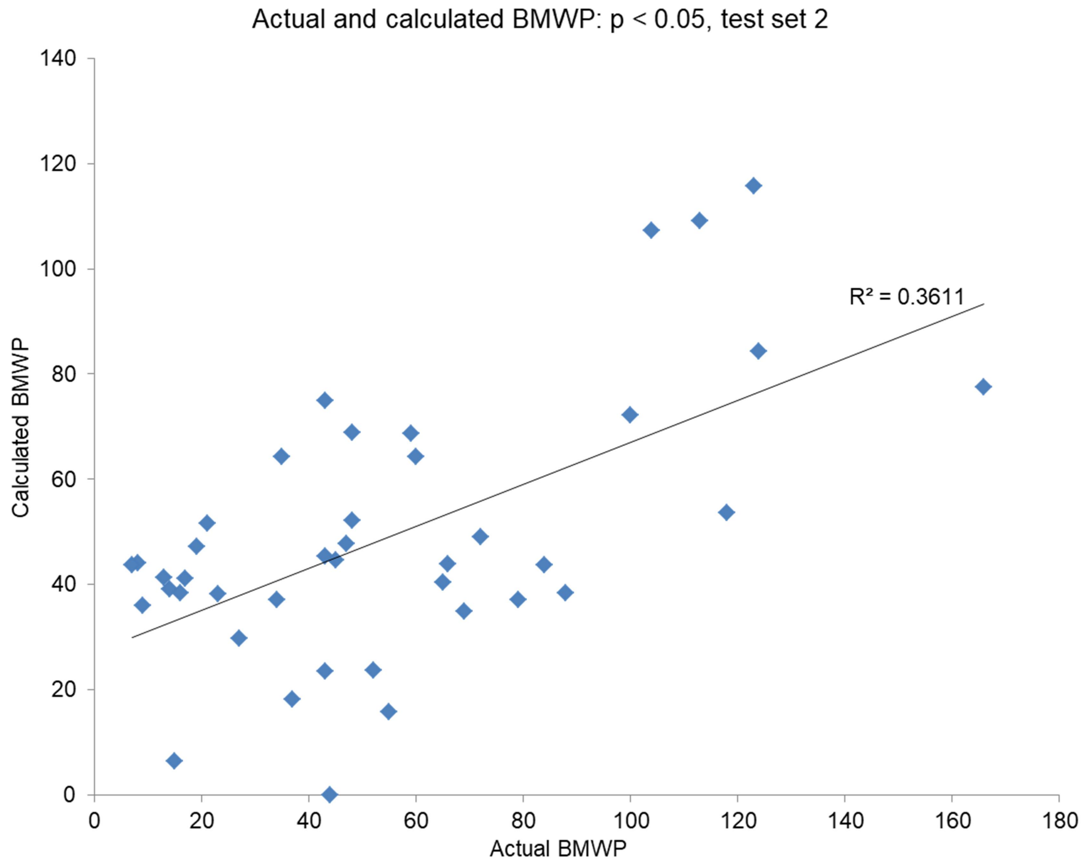


Figure D13 Validation of model based on folds test set 2 with input variables significant at $p < 0.05$.

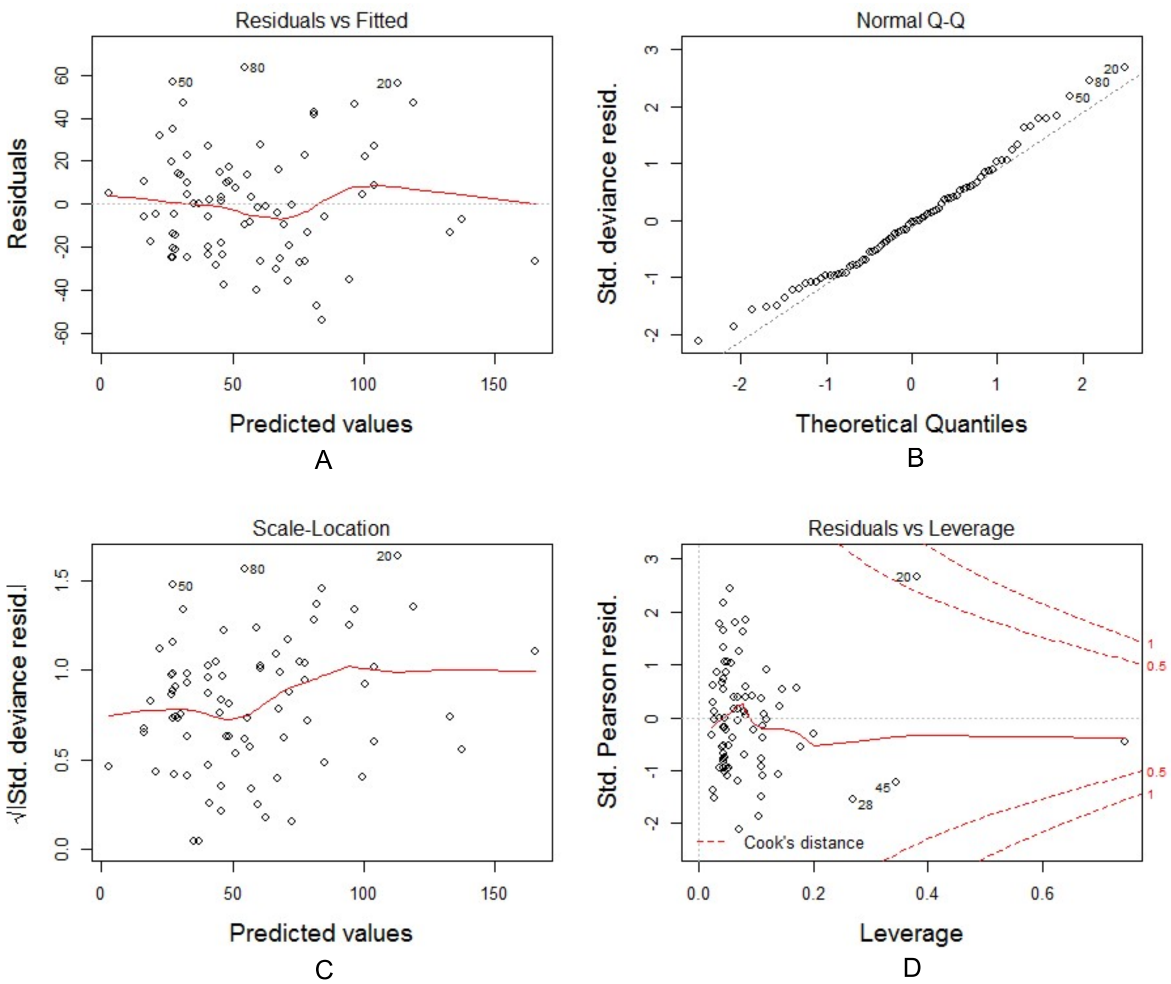


Figure D14 Residuals plots of model based on folds training set 2 + 3 with lowest AIC and input variables significant at $p < 0.1$, (A) residuals versus fitted values; (B) QQ-plot for normality; (C) scaled residuals versus fitted values; (D) standardized residuals versus leverage.

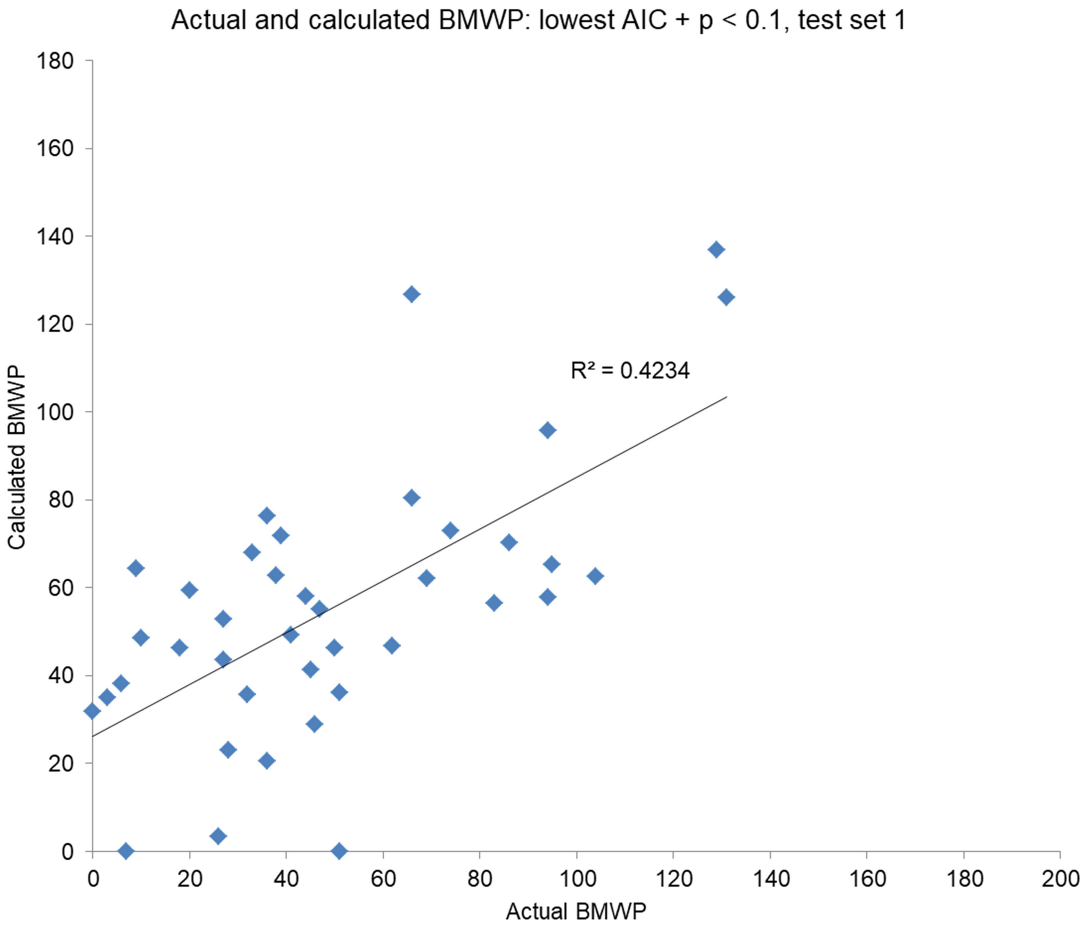


Figure D15 Validation of model based on folds test set 1 with lowest AIC and input variables significant at $p < 0.1$.

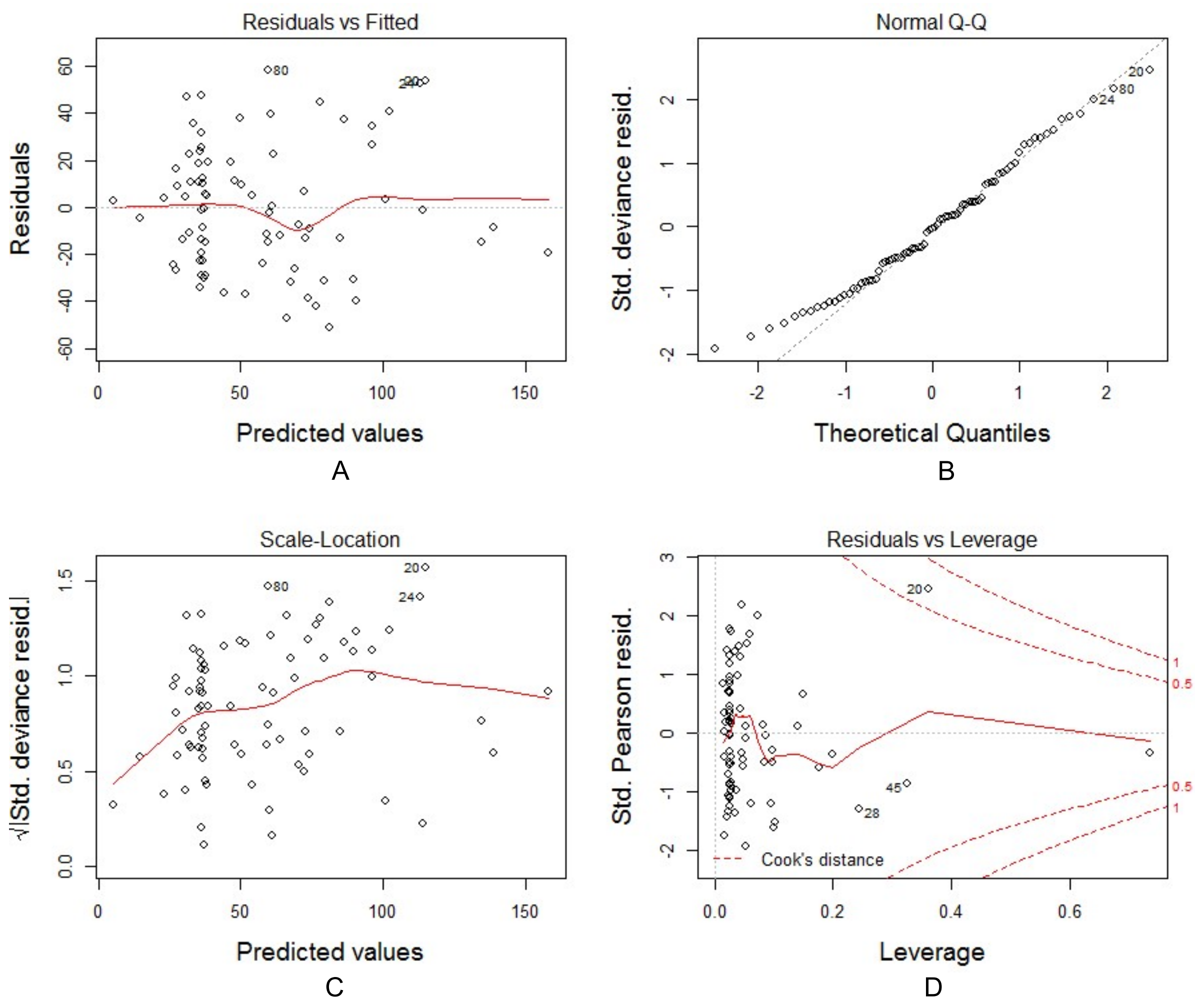


Figure D16 Residuals plots of model based on folds training set 2 + 3 with input variables significant at $p < 0.05$, (A) residuals versus fitted values; (B) QQ-plot for normality; (C) scaled residuals versus fitted values; (D) standardized residuals versus leverage.

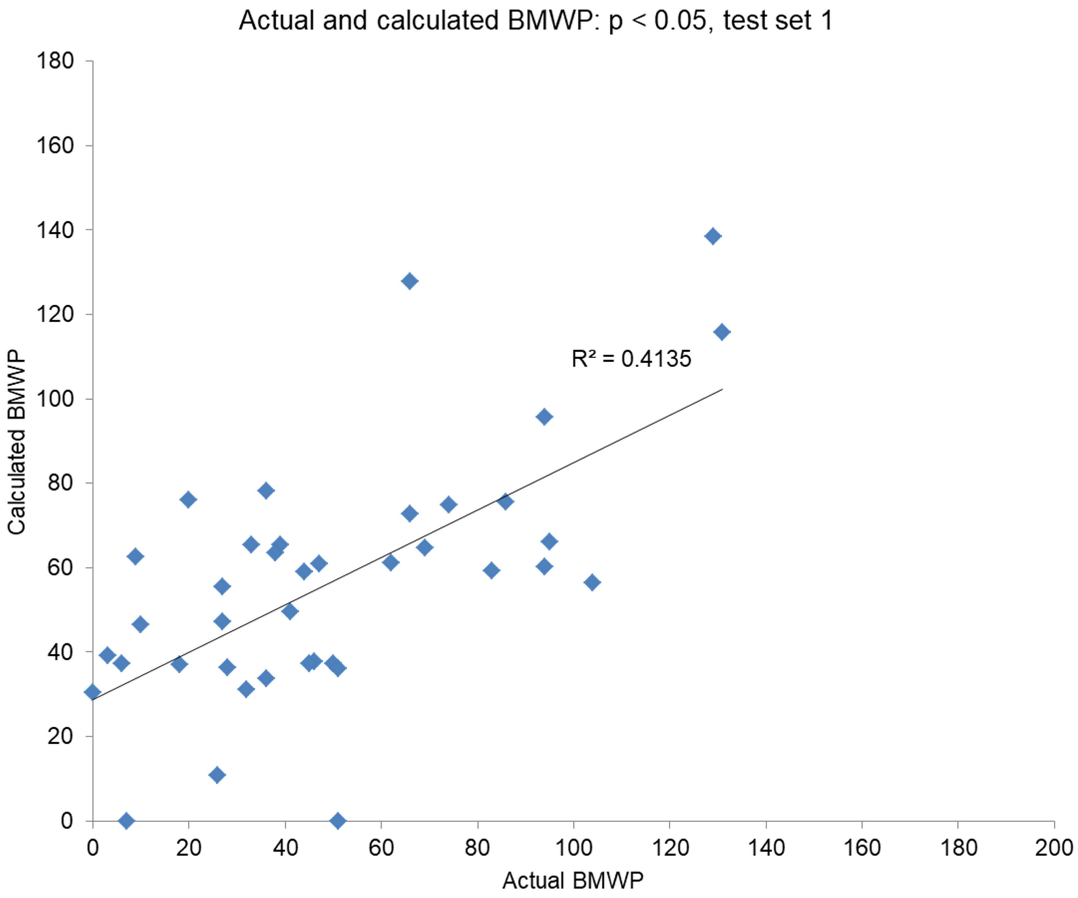


Figure D17 Validation of model based on folds test set 1 with input variables significant at $p < 0.05$.

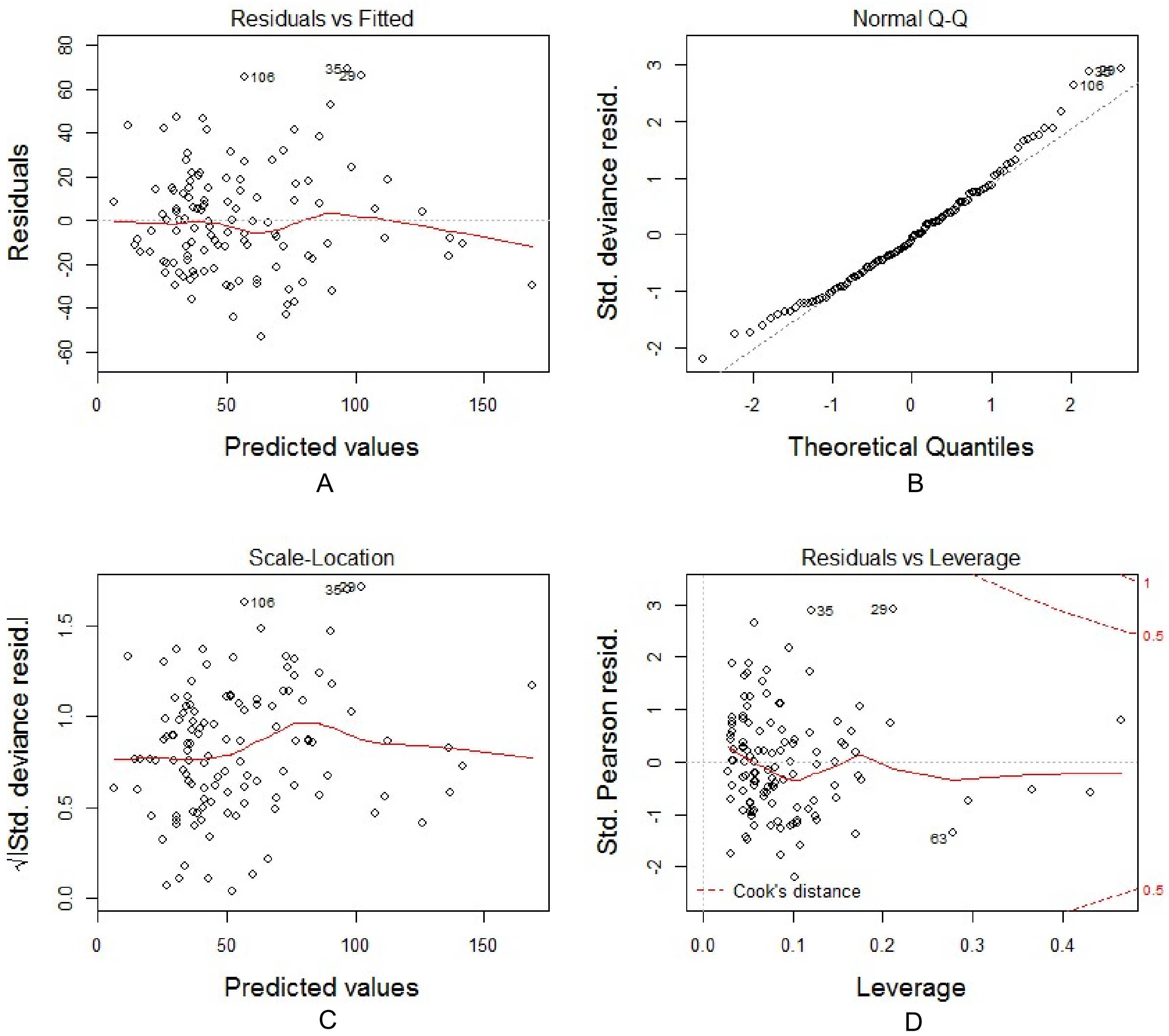


Figure D18 Residuals plots of model with complete data set and lowest AIC, (A) residuals versus fitted values; (B) QQ-plot for normality; (C) scaled residuals versus fitted values; (D) standardized residuals versus leverage.

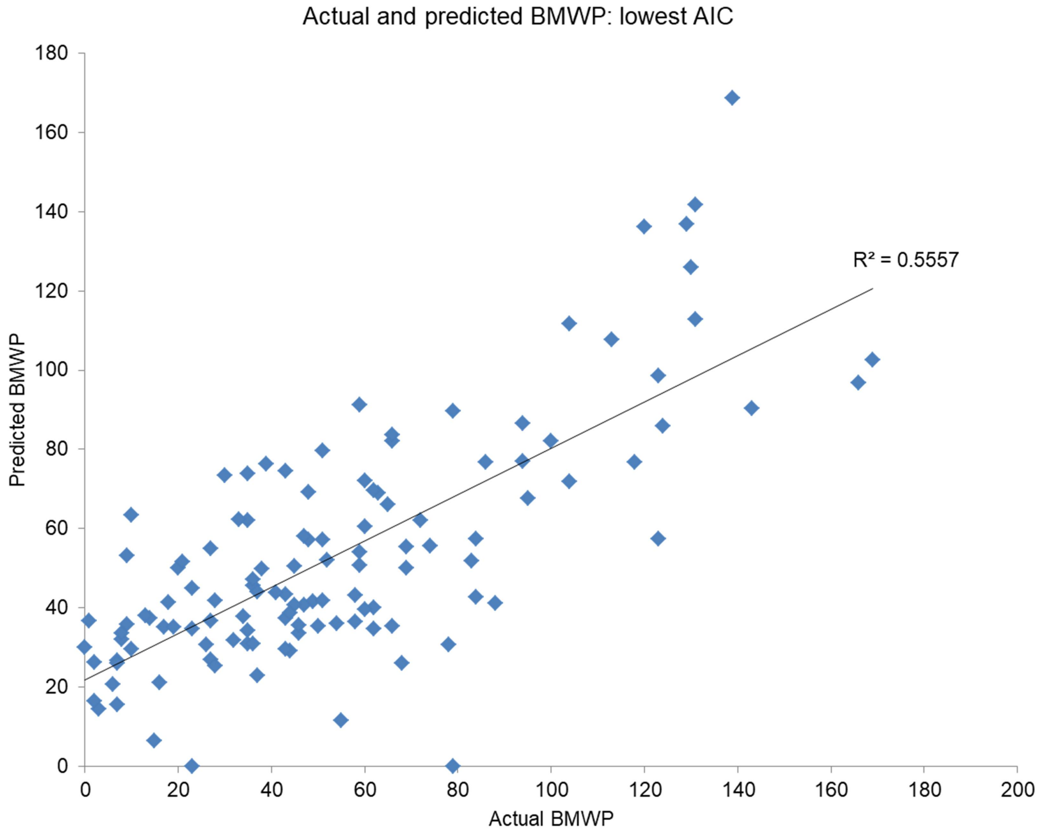


Figure D19 Validation of model with complete data set and lowest AIC.

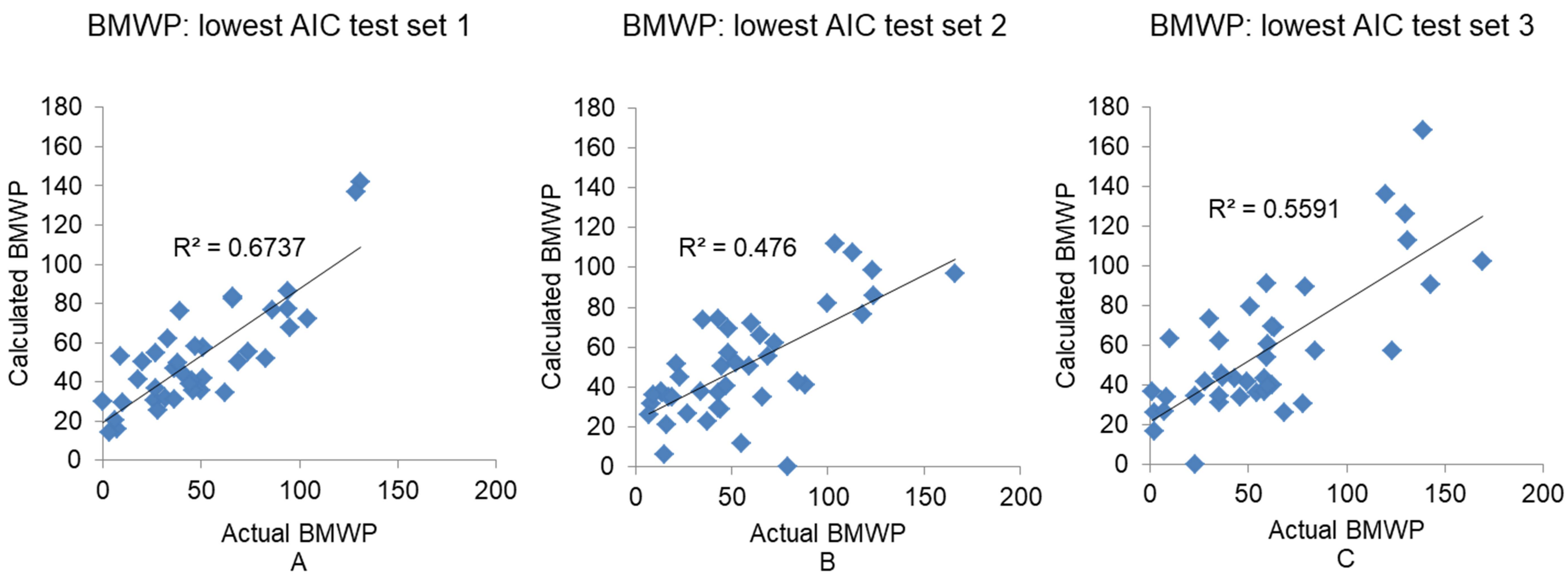


Figure D20 Validation of model with complete data set and lowest AIC on three folds, (A) for test set 1; (B) for test set 2; (C) for test set 3.

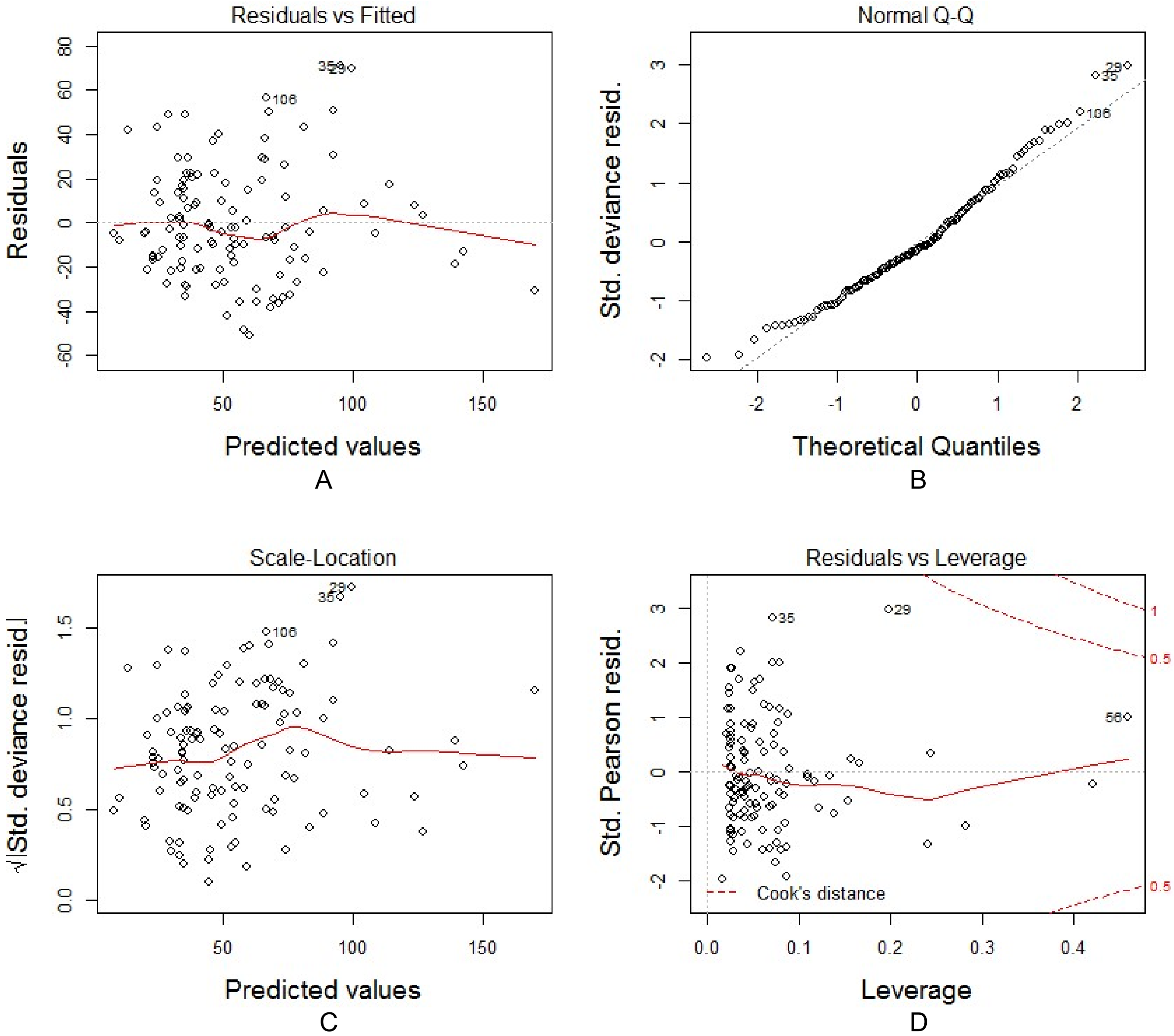


Figure D21 Residuals plots of model with complete data set and input variables significant at $p < 0.1$, (A) residuals versus fitted values; (B) QQ-plot for normality; (C) scaled residuals versus fitted values; (D) standardized residuals versus leverage.

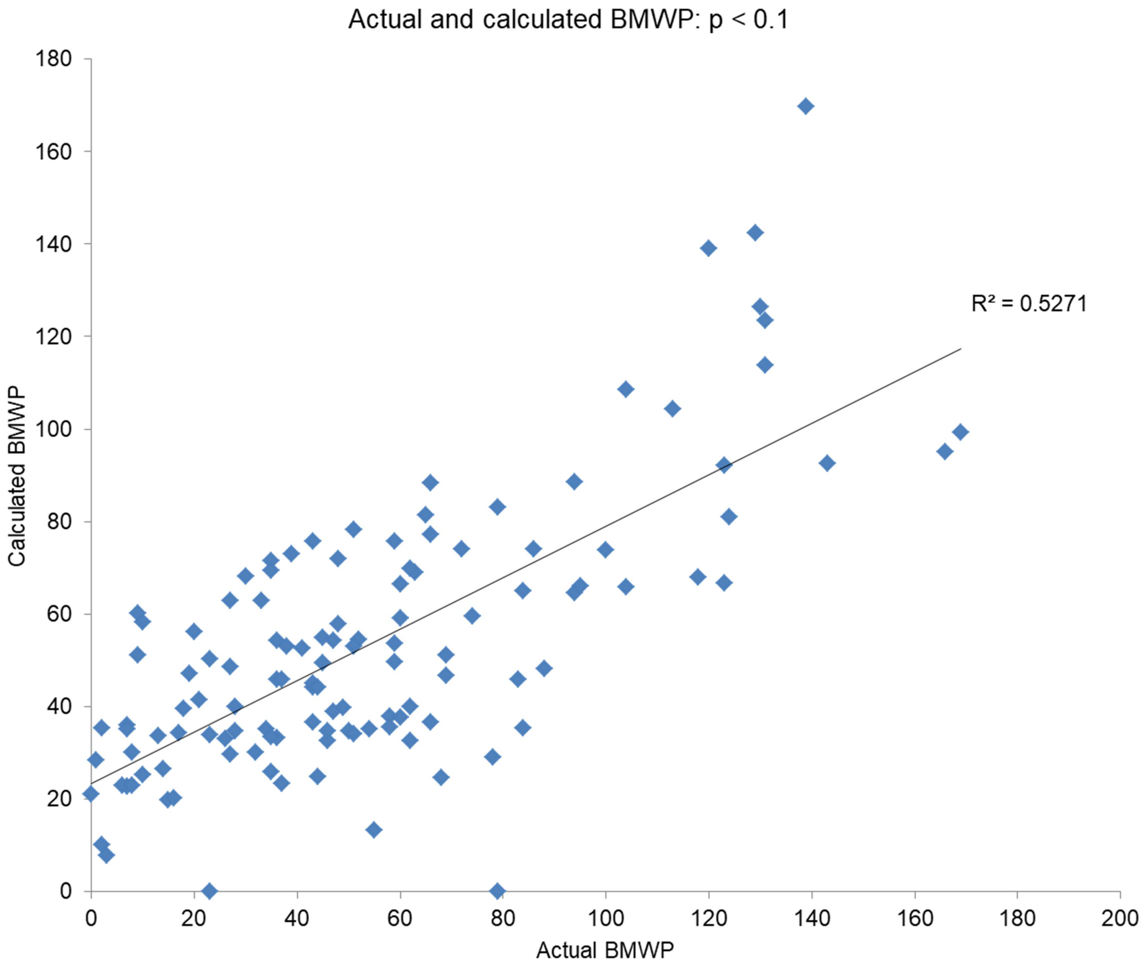


Figure D22 Validation of model with complete data set and input variables significant at $p < 0.1$.

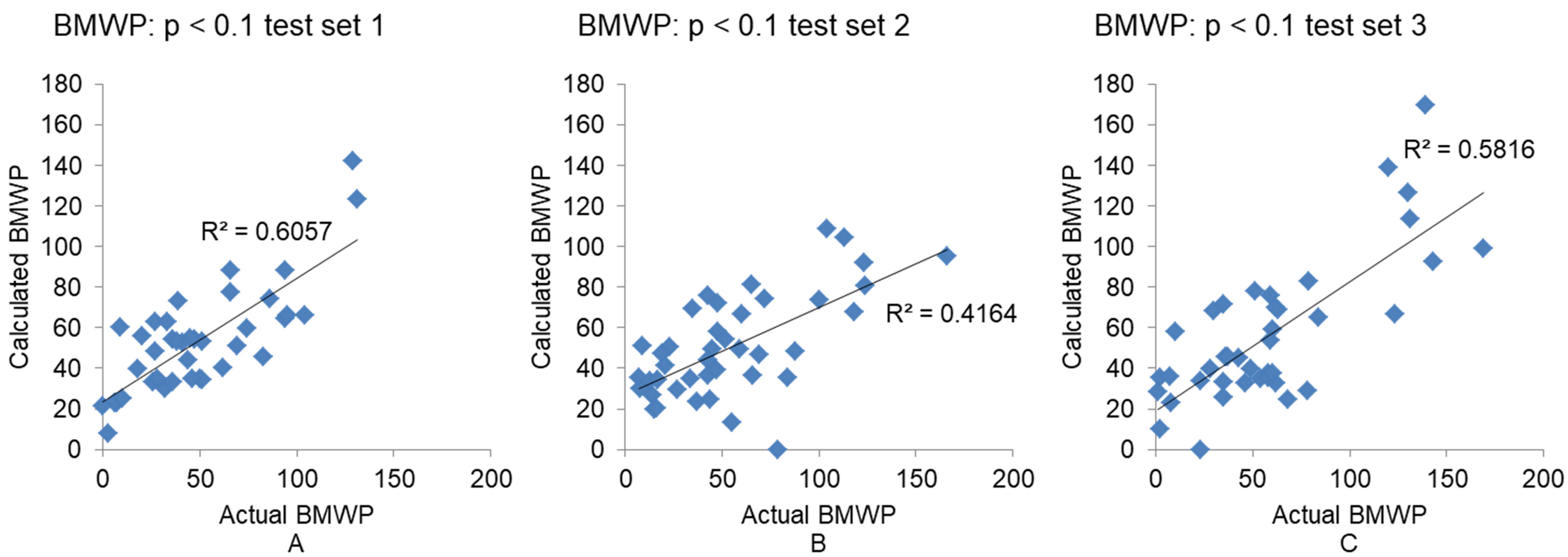


Figure D23 Validation of model with complete data set and input variables significant at $p < 0.1$ on three folds, (A) for test set 1; (B) for test set 2; (C) for test set 3.

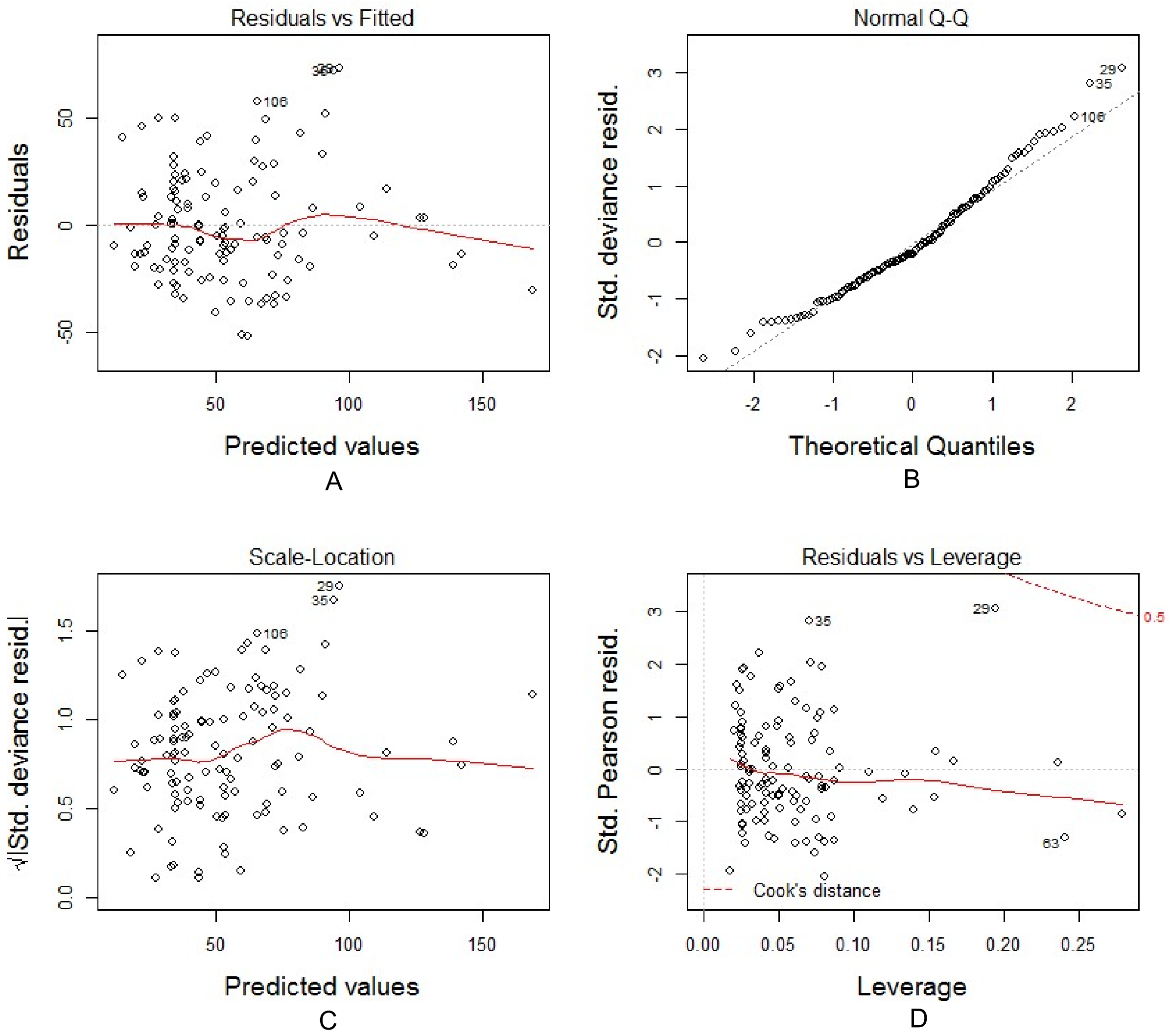


Figure D24 Residuals plots of model with complete data set and input variables significant at p < 0.05, (A) residuals versus fitted values; (B) QQ-plot for normality; (C) scaled residuals versus fitted values; (D) standardized residuals versus leverage.

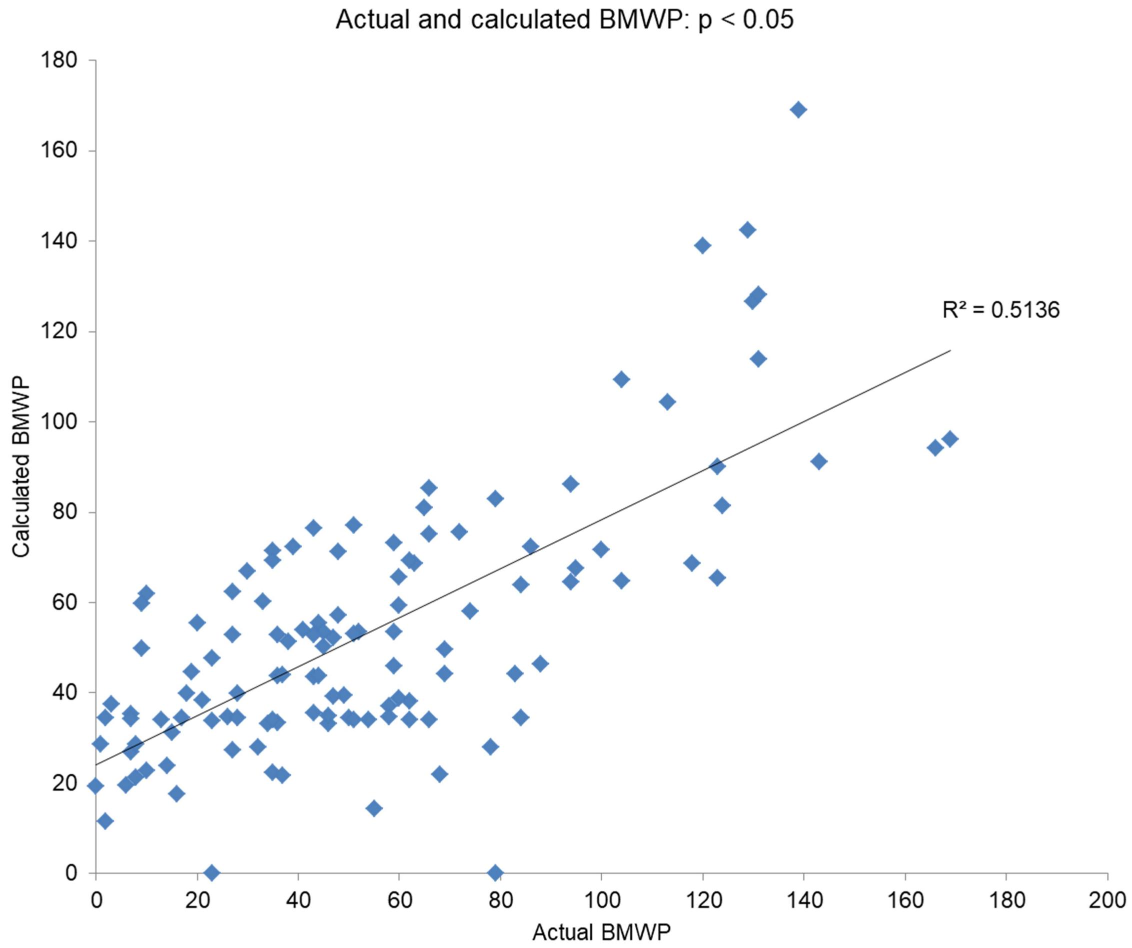


Figure D25 Validation of model with complete data set and input variables significant at $p < 0.05$.

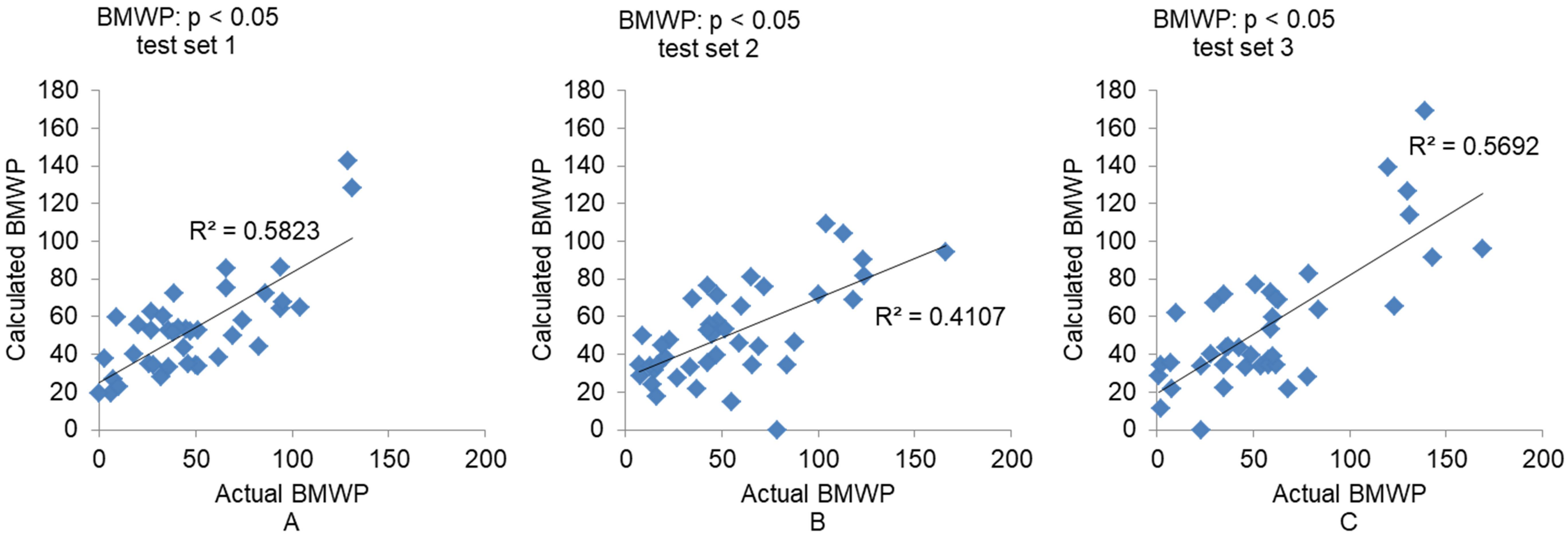


Figure D26 Validation of model with complete data set and input variables significant at $p < 0.05$ on three folds, (A) for test set 1; (B) for test set 2; (C) for test set 3.

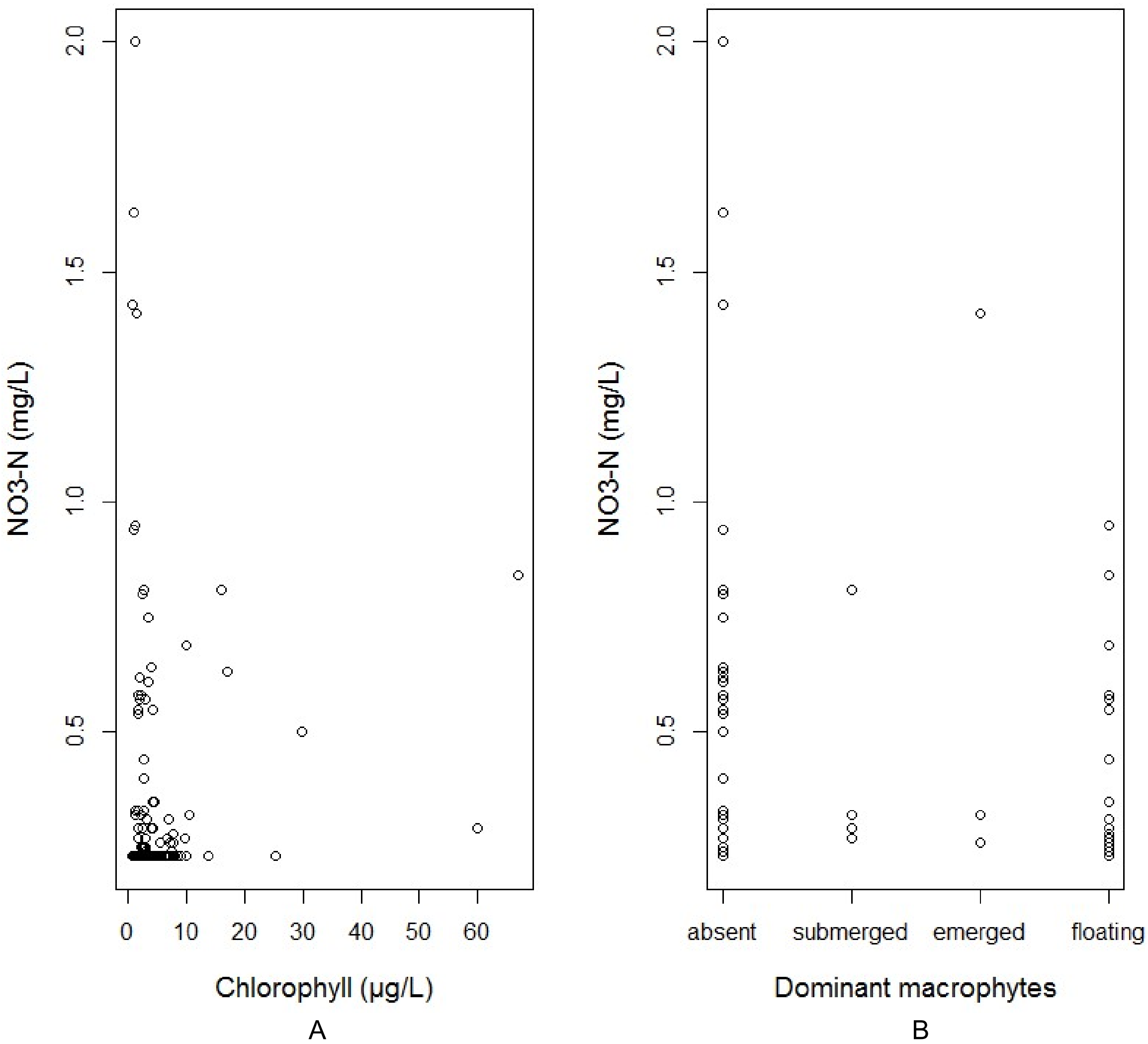


Figure D27 Relationship between nitrate-N and chlorophyll a (A), and between nitrate-N and dominant macrophytes (B).

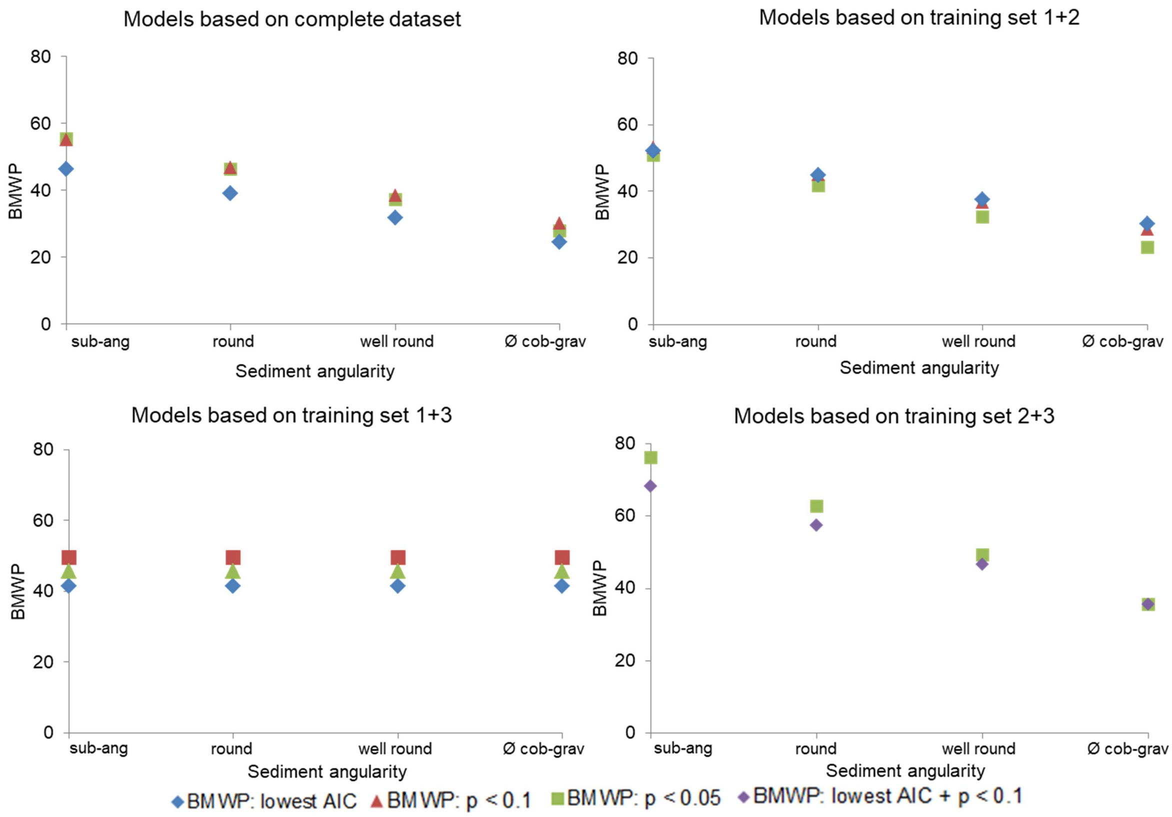


Figure D28 The impact of different types of sediment angularity on the ecological water quality expressed as BMWP-Col for model from different folds. The values used in the analysis were based on Table D8.

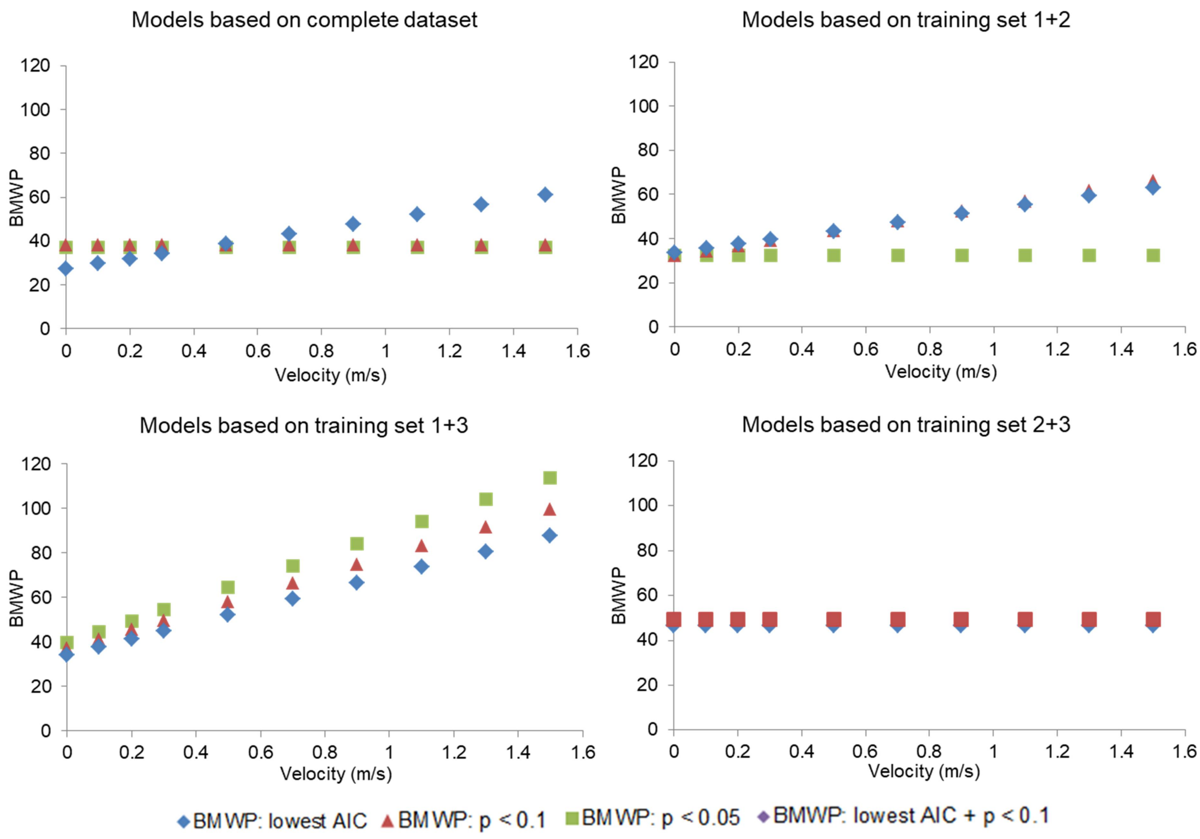


Figure D29 The impact of different flow velocity on the ecological water quality expressed as BMWP-Col for model from different folds. The values used in the analysis were based on Table D8.

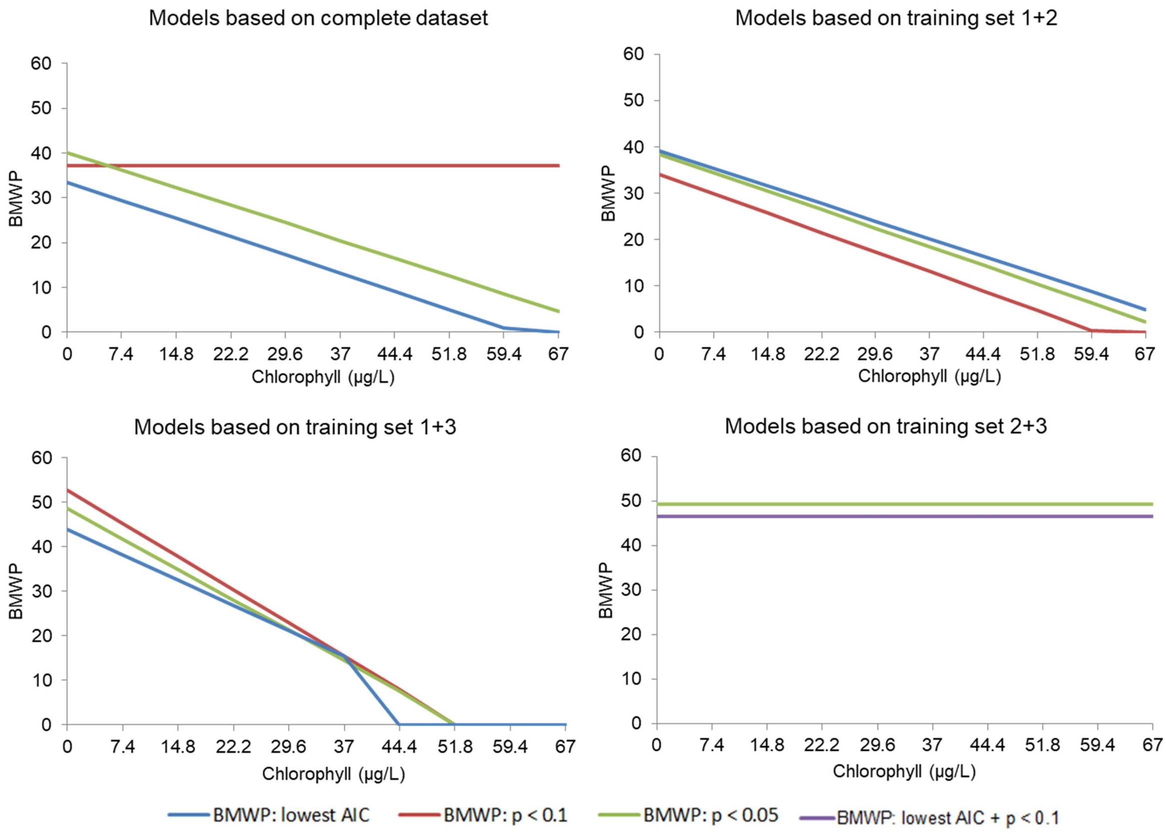


Figure D30 The impact of different chlorophyll a concentrations on the ecological water quality expressed as BMWP-Col for model from different folds. The values used in the analysis were based on Table D8.

Tables

Table D1 Variables' selection for three folds: showing variable with the highest *p*-value in the model together with the AIC of each model.

Model with Training set 1 + 2	Model with Training Set 1 + 3	Model with Training Set 2 + 3
Bed compaction, *p* = 0.97954, AIC = 754.92	Valley form, *p* = 0.90422, AIC = 757.05	Bed compaction, *p* = 0.95311, AIC = 773.14
Bank profile, *p* = 0.93177, AIC = 752.92	Shading, *p* = 0.8122, AIC = 755.07	Bank slope, *p* = 0.93633, AIC = 771.15
Sludge layer, *p* = 0.83762, AIC = 750.93	Turbidity, *p* =0.753, AIC = 753.16	NO2, *p* = 0.92297, AIC = 769.16
Valley form, *p* = 0.8139, AIC = 748.99	Bank profile, *p* = 0.73639, AIC = 751.3	Sludge layer, *p* = 0.86893, AIC = 767.17
Branch, *p* = 0.68334, AIC = 747.07	Sediment angularity, *p* = 0.55891, AIC = 749.46	Velocity, *p* = 0.75993, AIC = 765.21
Bank shape, *p* = 0.63615, AIC = 745.3	Variation in flow, *p* = 0.50215, AIC = 747.93	Chlorophyll, *p* = 0.782108, AIC = 763.34
DO, *p* = 0.57422, AIC = 743.6	Main macrophytes, *p* = 0.4821, AIC = 746.54	Branch, *p* = 0.672055, AIC = 761.44
Turbidity, *p* = 0.57365, AIC = 742.02	Sludge layer, *p* = 0.531783, AIC = 745.2	Erosion, *p* = 0.528448, AIC = 759.68
NO2, *p* = 0.50653, AIC = 740.43	Width variation, *p* = 0.521583, AIC = 743.71	Channel form, *p* = 0.485696, AIC = 758.2
Channel form, *p* = 0.43673, AIC = 739	Erosion, *p* = 0.3957, AIC = 742.24	Bank profile, *p* = 0.553846, AIC = 756.82
Width variation, *p* = 0.31746, AIC = 737.77	DO, *p* = 0.294835, AIC = 741.15	Land use, *p* = 0.359555, AIC = 755.26
Twigs, *p* = 0.231301, AIC = 737.01	Bed compaction, *p* = 0.3629, AIC = 740.52	Main macrophytes, *p* = 0.336407, AIC = 754.31
Shading, *p* = 0.204631, AIC = 736.76	Channel form, 0.4092, AIC = 739.54	Width variation, *p* = 0.336441, AIC = 753.44
Bank slope, *p* = 0.18118, AIC = 736.7	NO_2, *p* = 0.383788, AIC = 738.36	Variation in flow, *p* = 0.184036, AIC = 752.55
Velocity, *p* = 0.099362, AIC = 736.83	Twigs, *p* = 0.31435, AIC = 737.27	DO, *p* = 0.233515, AIC = 752.65
Main land use, *p* = 0.160108, AIC = 738.01	Branch, *p* = 0.141006, AIC = 736.45	Bank shape, *p* = 0.224147, AIC = 752.31
Chlorophyll, *p* = 0.043185, AIC = 738.28	Logs, *p* = 0.112273, AIC = 736.96	Logs, *p* = 0.147498, AIC = 752.01
Main macrophytes, *p* = 0.040888, AIC = 740.93	Main land use, *p* = 0.12452, AIC = 737.83	Shading, *p* = 0.25308, AIC = 752.4
NO3, *p* = 0.109585, AIC = 743.61	Bank slope, *p* = 0.10978, AIC = 738.48	Twigs, *p* = 0.06822, AIC = 751.86
Logs, *p* = 0.131943, AIC = 755.67	NO_3, *p* = 0.05478, AIC = 739.31	Valley form, *p* = 0.103873, AIC = 753.54
Variation in flow, 0.046539, AIC = 756.14	Chlorophyll, *p* = 0.022016, AIC = 749.99	NO_3, *p* = 0.006893, AIC = 754.42

Table D2 Ranking of importance of input variables in the models with 3-folds cross validation, based on the *p*-values (the *p*-values are given between brackets).

Variables	Variables' Ranking										
	First	**Second**	**Third**	**Fourth**	**Fifth**	**Sixth**	**Seventh**	**Eighth**	**Ninth**	**Tenth**	**Eleventh**
Training set 1 + 2											
Lowest AIC	Elevation (0.001)	Main macrophytes (0.013)	NO_3^--N (0.024)	Angularity (0.027)	Logs (0.044)	Land use (0.048)	Erosion (0.048)	Chlorophyll (0.064)	Flow variation (0.067)	Velocity (0.151)	Bank slope (0.181)
$p < 0.1$	Elevation (0.003)	Main macrophytes (0.012)	Angularity (0.013)	Logs (0.016)	NO_3^--N (0.019)	Erosion (0.048)	Chlorophyll (0.050)	Flow variation (0.067)	Land use (0.085)	Velocity (0.099)	
$p < 0.05$	Elevation (<0.001)	Angularity (0.005)	Flow variation (0.009)	Erosion (0.014)	Logs (0.016)	Main macrophytes (0.017)	NO_3^--N (0.020)	Chlorophyll (0.043)			
Training set 1 + 3											
Lowest AIC	Elevation (<0.001)	Velocity (0.015)	Bank shape (0.036)	Logs (0.042)	Land use (0.063)	NO_3^--N (0.066)	Chlorophyll (0.072)	Bank slope (0.081)	Branch (0.141)		
$p < 0.1$	Elevation (<0.001)	Velocity (0.004)	Bank shape (0.024)	Chlorophyll (0.033)	NO_3^--N (0.055)						
$p < 0.05$	Elevation (<0.001)	Velocity (<0.001)	Bank shape (0.020)	Chlorophyll (0.022)							
Training set 2 + 3											
Lowest AIC	Elevation (<0.001)	Angularity (0.002)	NO_3^--N (0.003)	Turbidity (0.005)	Valley form (0.056)	Twigs (0.068)					
$p < 0.1$	Elevation (<0.001)	Angularity (0.002)	NO_3^--N (0.003)	Turbidity (0.005)	Valley form (0.056)	Twigs (0.068)					
$p < 0.05$	Elevation (<0.001)	Angularity (<0.001)	Turbidity (0.004)	NO_3^--N (0.007)							
# occurrence											
Nine times	Elevation										
Three times		Angularity,	Bank shape								

Variables	**Variables' Ranking**										
	First	**Second**	**Third**	**Fourth**	**Fifth**	**Sixth**	**Seventh**	**Eighth**	**Ninth**	**Tenth**	**Eleventh**
		velocity									
Two times		Main macrophytes	NO_3^--N	Logs, chlorophyll	Logs, NO_3^--N		Chlorophyll	Chlorophyll		Velocity	
One time			Angularity, flow variation, turbidity	Angularity, erosion, turbidity, NO_3^--N	Land use, valley form	Land use, erosion, main macrophytes, NO_3^--N, twigs	Erosion, NO_3^--N	Flow variation, bank slope	Flow variation, land use, branch		Bank slope

Table D3 Predictive performances of models with 3-folds cross validation: showing the number of input variables that construct the models and the coefficient of determination (R^2) values of training and testing sets.

Models	# Input Variables	R^2	
		Training Set	**Testing Set**
Fold 1 lowest AIC	11	0.57	0.49
Fold 1 $p < 0.1$	10	0.56	0.48
Fold 1 $p < 0.05$	8	0.54	0.45
Fold 2 lowest AIC	9	0.62	0.31
Fold 2 $p < 0.1$	5	0.56	0.32
Fold 2 $p < 0.05$	4	0.54	0.36
Fold 3 lowest AIC	6	0.55	0.42
Fold 3 $p < 0.1$	6	0.55	0.42
Fold 3 $p < 0.05$	4	0.52	0.41
Average lowest AIC	9	0.58	0.41
Average $p < 0.1$	7	0.56	0.41
Average $p < 0.05$	5	0.53	0.41

Table D4 Variables' selection for model with complete data set: showing variable with the highest *p*-value in the model and the AIC of each model.

Variable with Highest *p*-Value	Model's AIC	Remarks
Bank profile, $p = 0.895125$	1131.9	
Sludge layer, $p = 0.89966$	1129.9	
Channel form, $p = 0.882337$	1127.9	
Erosion, $p = 0.781556$	1126	
NO2, $p = 0.753035$	1124	
Branch, $p = 0.730042$	1122.2	
DO, $p = 0.704733$	1120.3	
Variation in width, $p = 0.576093$	1118.5	
Bed compaction, $p = 0.582336$	1116.9	
Bank slope, $p = 0.36968$	1115.2	
Valley form, $p = 0.353982$	1114.1	
Main land use, $p = 0.204995$	1113.1	
Shading, $p = 0.18378$	1113	
Turbidity, $p = 0.184818$	1113	
Logs, $p = 0.112776$	1112.9	Lowest AIC, $R^2 = 0.56$
Main macrophytes, $p = 0.10811$	1113.7	
Velocity, $p = 0.13639$	1114.5	
Chlorophyll, $p = 0.06668$	1115	Significant at $p < 0.1$, $R^2 = 0.53$
Twigs, $p = 0.01639$	1116.6	Significant at $p < 0.05$, $R^2 = 0.51$

Table D5 Variables' ranking of importance based on the *p*-values for models with complete data set, the *p*-values are given between brackets.

	First	Second	Third	Fourth	Fifth	Sixth	Seventh	Eighth	Ninth	Tenth
Lowest AIC	Elevation (<0.001)	NO_3^--N (0.004)	Bank shape (0.004)	Angularity (0.013)	Flow variation (0.027)	Chlorophyll (0.056)	Twigs (0.067)	Velocity (0.073)	Main macrophytes (0.087)	Logs (0.113)
$p < 0.1$	Elevation (<0.001)	NO_3^--N (0.003)	Angularity (0.004)	Twigs (0.011)	Flow variation (0.011)	Bank shape (0.013)	Chlorophyll (0.067)			
$p < 0.05$	Elevation (<0.001)	Angularity (0.001)	NO_3^--N (0.005)	Flow variation (0.006)	Bank shape (0.013)	Twigs (0.016)				

Table D6 Median, minimum and maximum values for sensitivity analysis of models based on complete data set.

Variables	Unit	Median	Min	Max
Chlorophyll	(µg/L)	3.1	0.7	66.8
Nitrate-N	(mg/L)	0.2	0.2	2.0
Turbidity	(NTU)	3.4	0.0	355.6
Velocity	(m/s)	0.2	0.0	1.5
Elevation	(m)	82	2	1075
Main macrophytes		0	0	3
Variation in flow		0	0	4
Twigs		0	0	2
Logs		0	0	2
Bank shape		2	0	5
Sediment angularity		5	3	6
Valley form		5	2	7
Main land use		1	1	4
Bank slope		2	0	5
Branch		1	0	2
Erosion		0	0	2
BMWP-Col		47	0	169

E – Supporting information for chapter 6

Figures

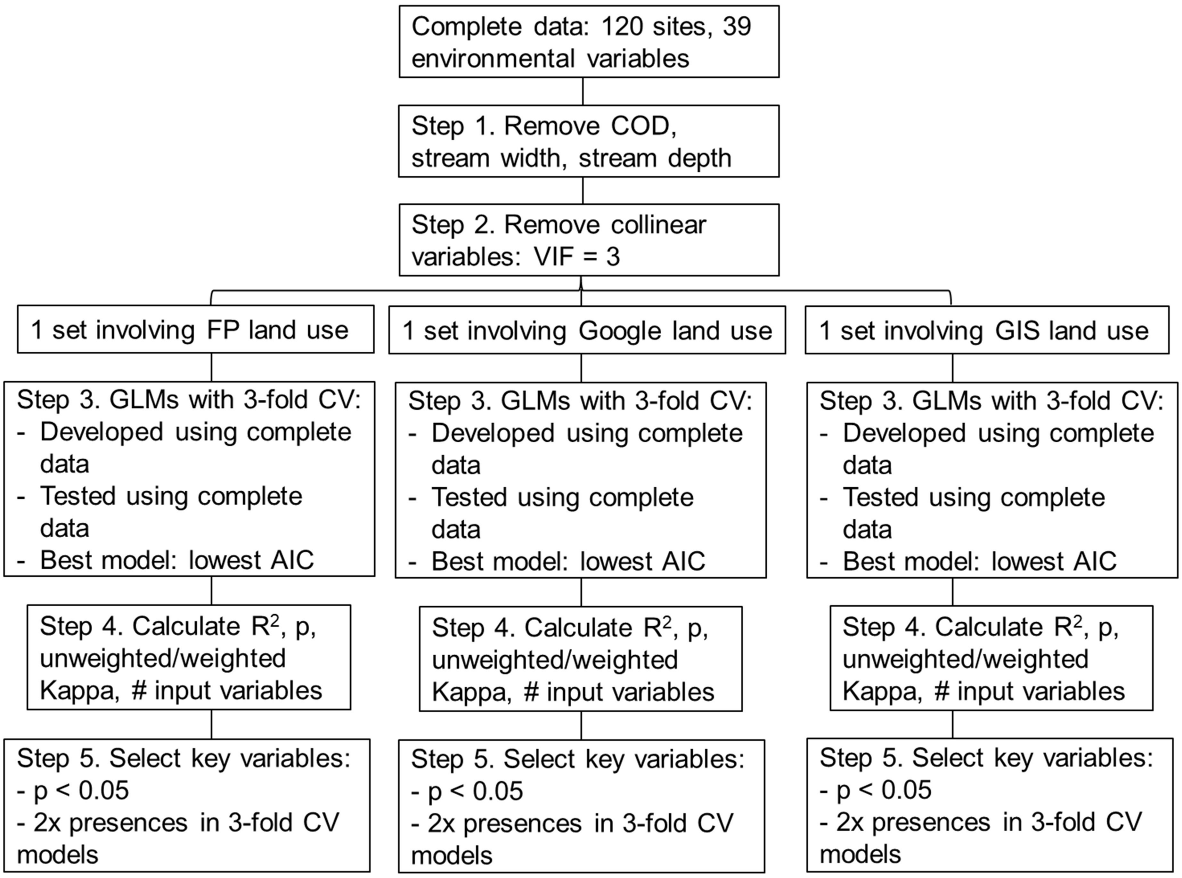
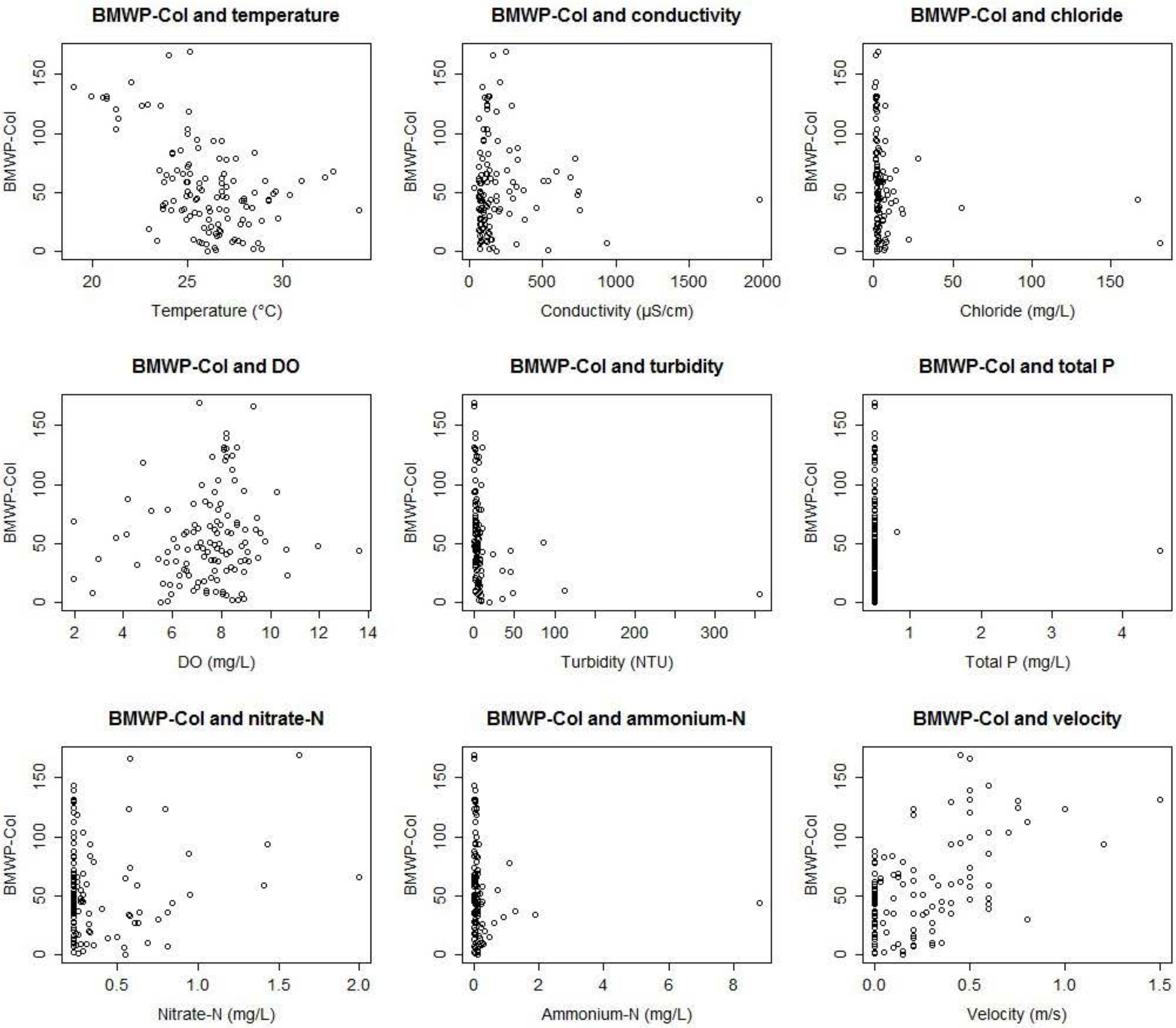
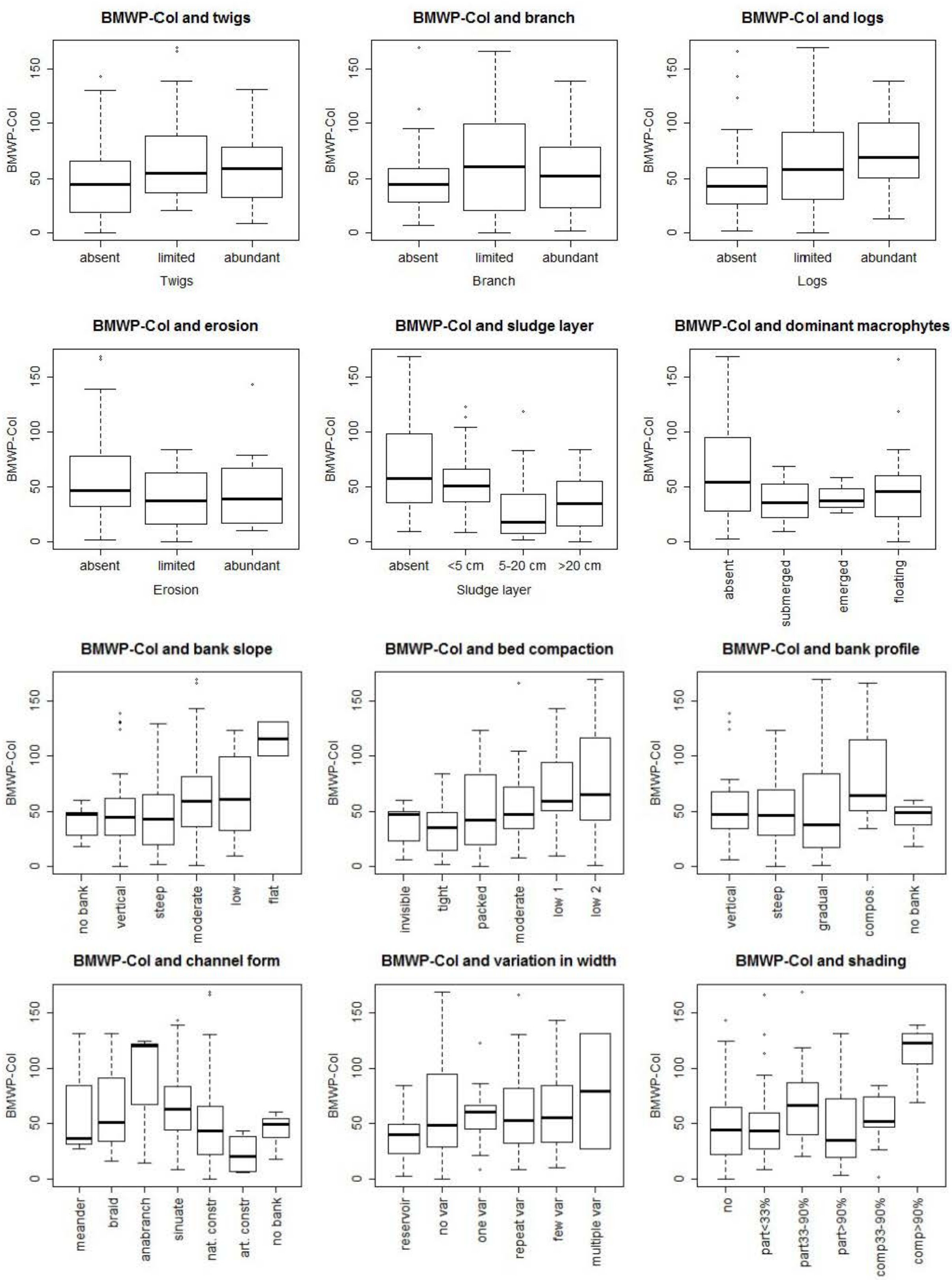


Figure E1 Schematic procedures of model development, best model selection and key variables selection.


BMWP-Col and temperature
BMWP-Col
Temperature (°C)
BMWP-Col and conductivity
Conductivity (μS/cm)
BMWP-Col and chloride
Chloride (mg/L)
BMWP-Col and DO
DO (mg/L)
BMWP-Col and turbidity
Turbidity (NTU)
BMWP-Col and total P
Total P (mg/L)
BMWP-Col and nitrate-N
Nitrate-N (mg/L)
BMWP-Col and ammonium-N
Ammonium-N (mg/L)
BMWP-Col and velocity
Velocity (m/s)


BMWP-Col and twigs
BMWP-Col
absent limited abundant
Twigs
BMWP-Col and branch
absent limited abundant
Branch
BMWP-Col and logs
absent limited abundant
Logs
BMWP-Col and erosion
absent limited abundant
Erosion
BMWP-Col and sludge layer
absent <5 cm 5-20 cm >20 cm
Sludge layer
BMWP-Col and dominant macrophytes
absent submerged emerged floating
BMWP-Col and bank slope
no bank vertical steep moderate low flat
BMWP-Col and bed compaction
invisible tight packed moderate low 1 low 2
BMWP-Col and bank profile
vertical steep gradual compos. no bank
BMWP-Col and channel form
meander braid anabranch sinuate nat. constr art. constr no bank
BMWP-Col and variation in width
reservoir no var one var repeat var few var multiple var
BMWP-Col and shading
no part<33% part33-90% part>90% comp33-90% comp>90%
0 50 100 150

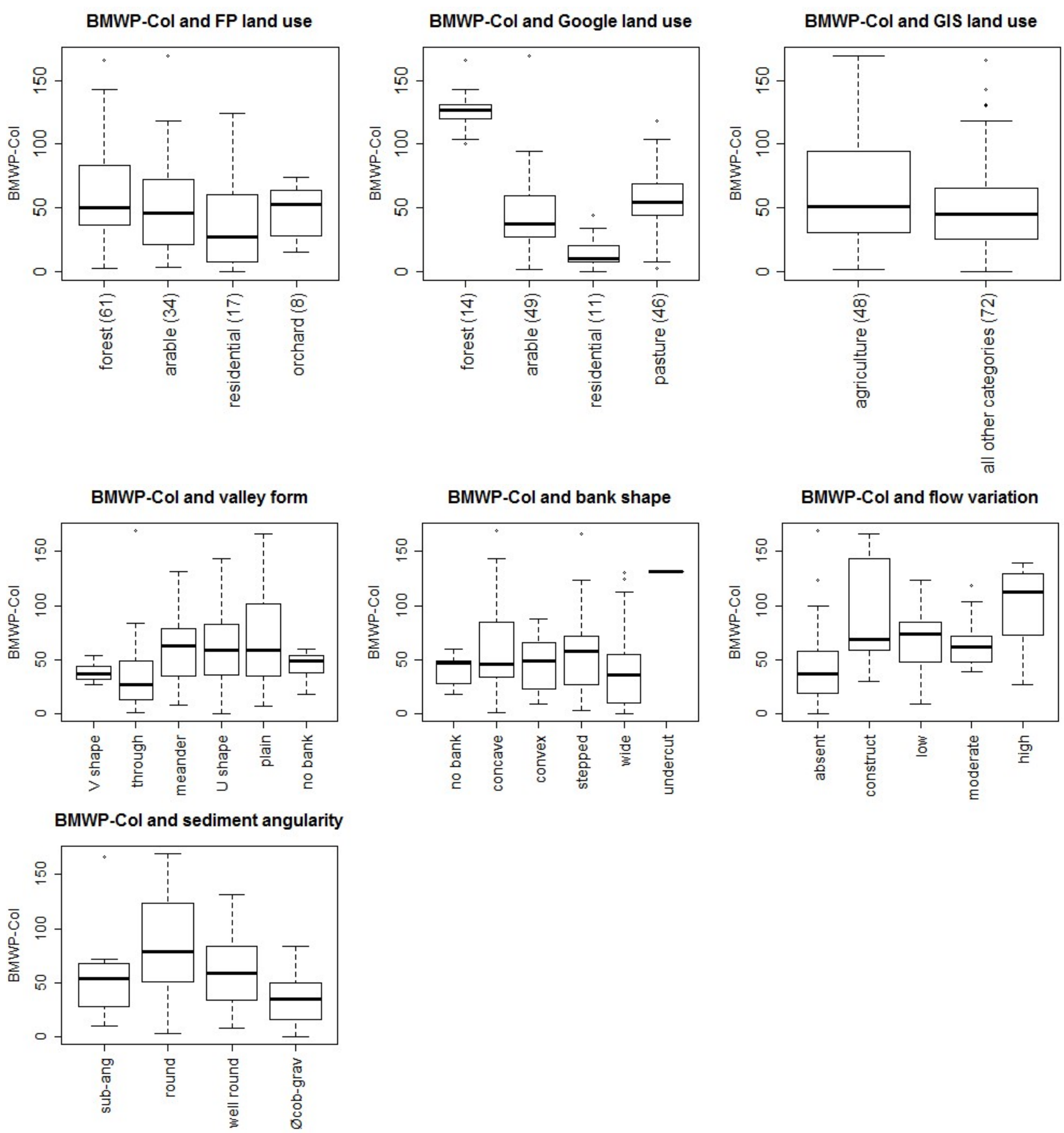


Figure E2 Plots showing the distribution of the data for physico-chemical variables in relation to BMWP-Col for 120 sampling sites. The classification of categorical variables is based on Table B1; compos: composite, nat: natural, art: artificial, constr: construction, var: variation, part: partly, comp: completely, ang: angular, cob-grav: cobble-pebble-gravel.

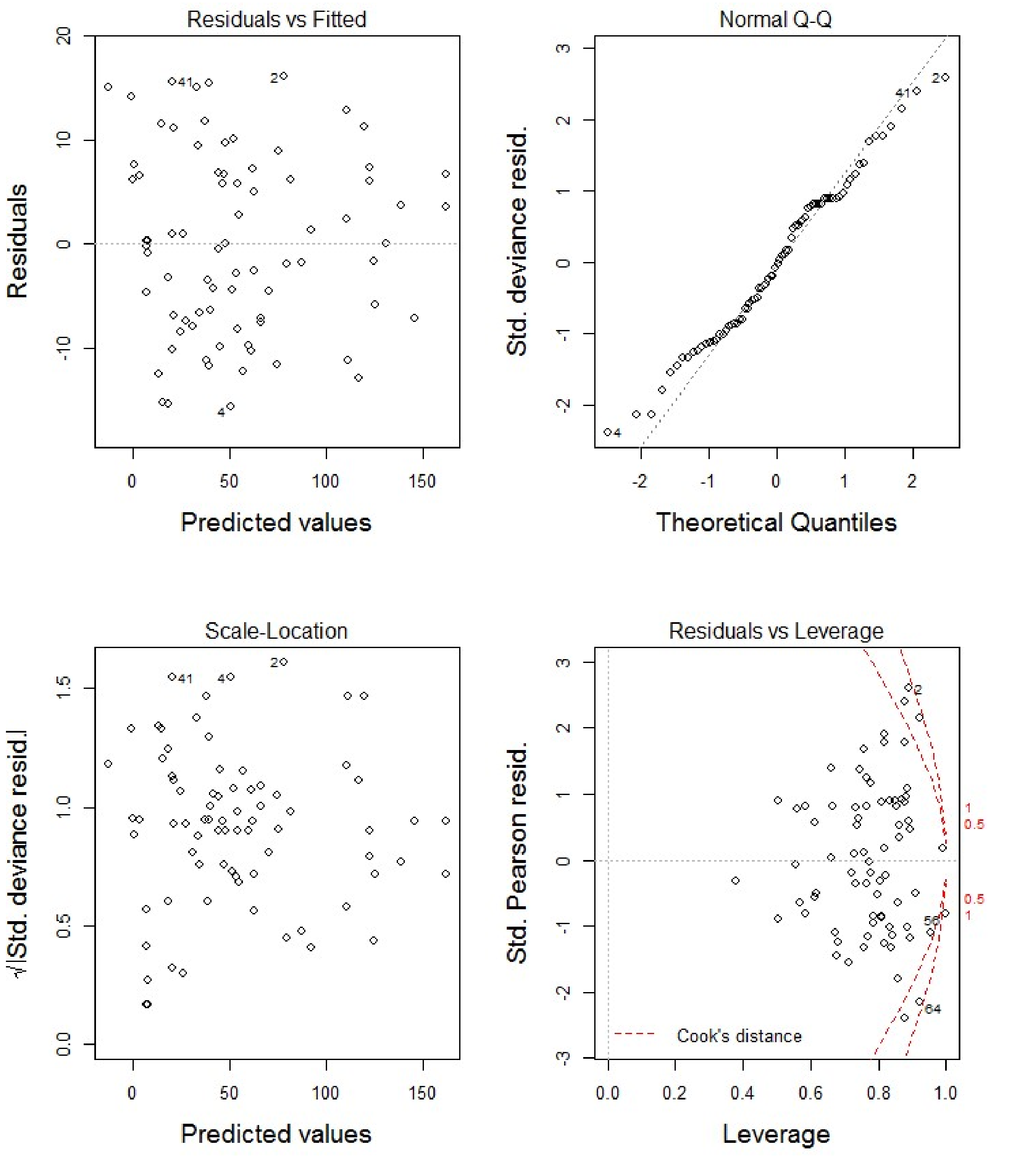


Figure E3 Residuals plots of model based on three-fold cross validation of fold1 with FP land use set.

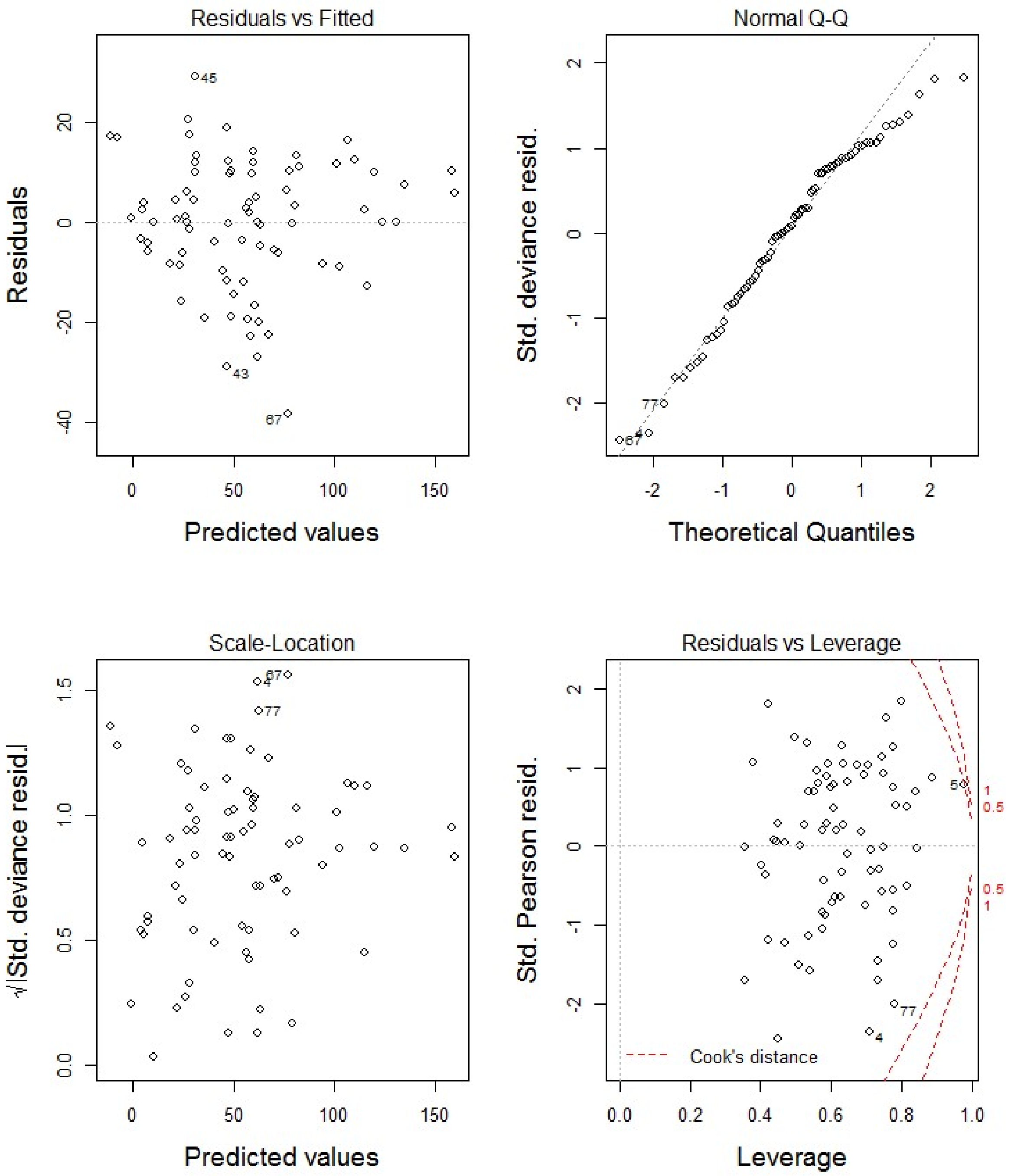


Figure E4 Residuals plots of model based on three-fold cross validation of fold2 with FP land use set.

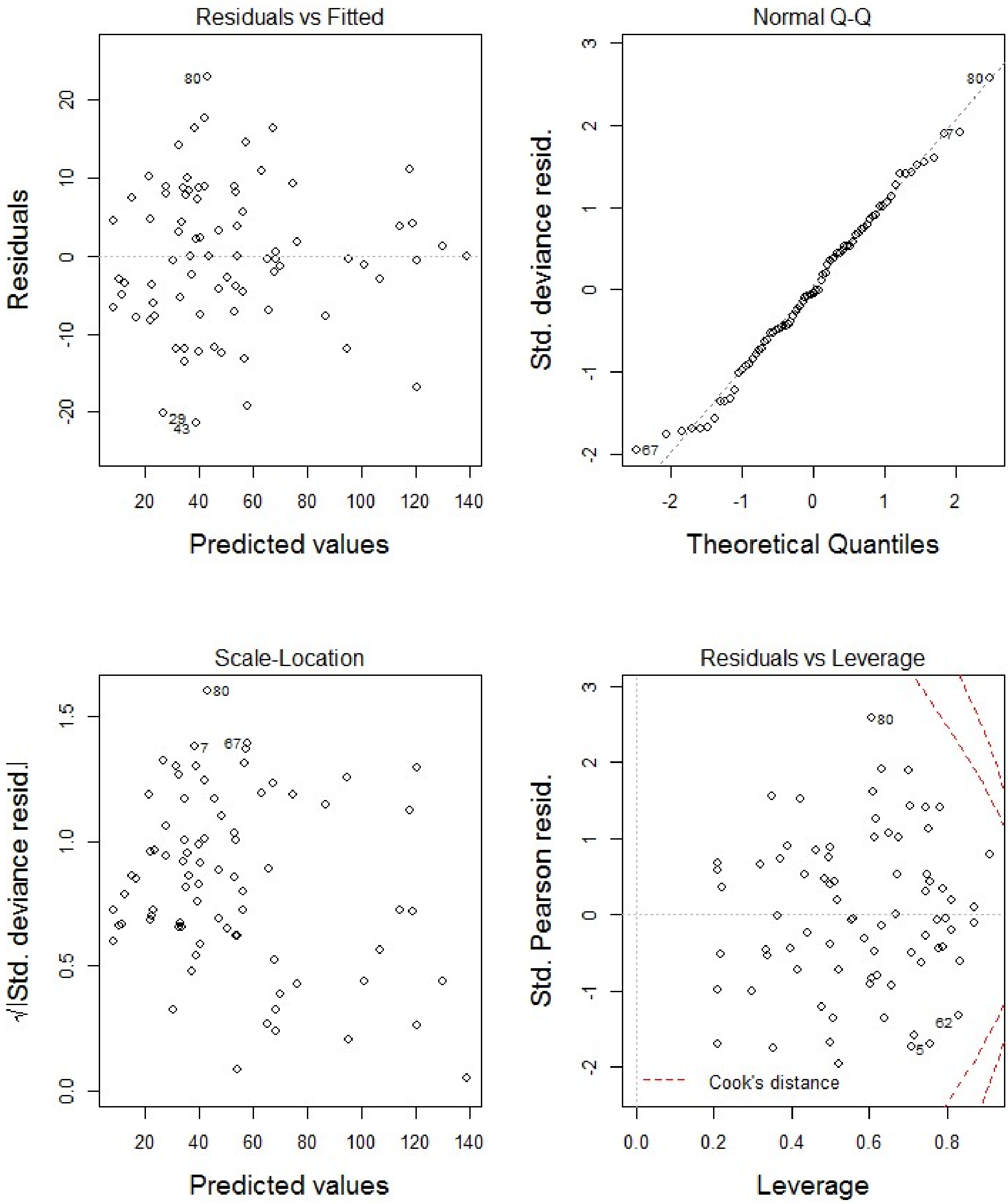


Figure E5 Residuals plots of model based on three-fold cross validation of fold3 with FP land use set.

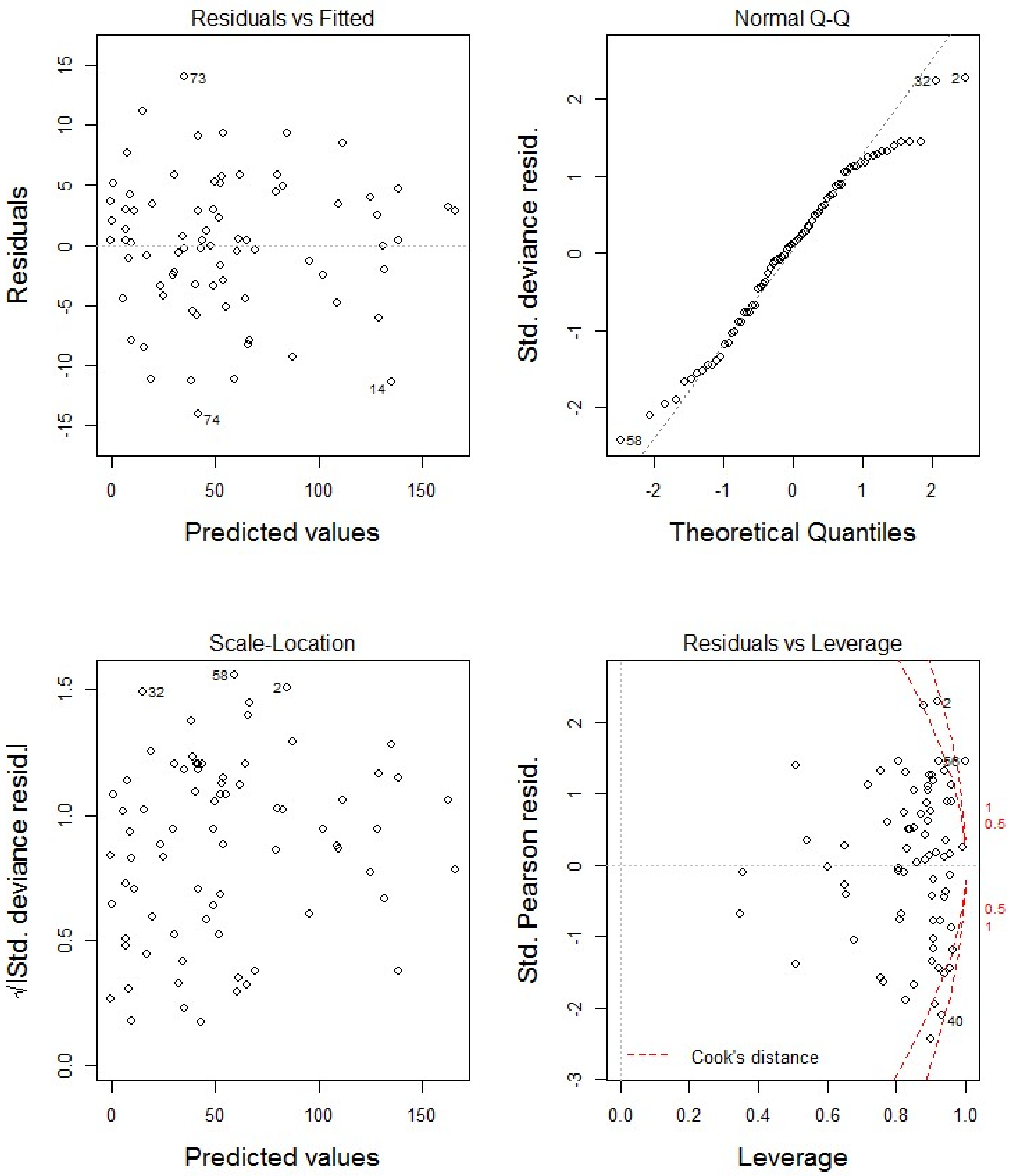


Figure E6 Residuals plots of model based on three-fold cross validation of fold1 with Google land use set.

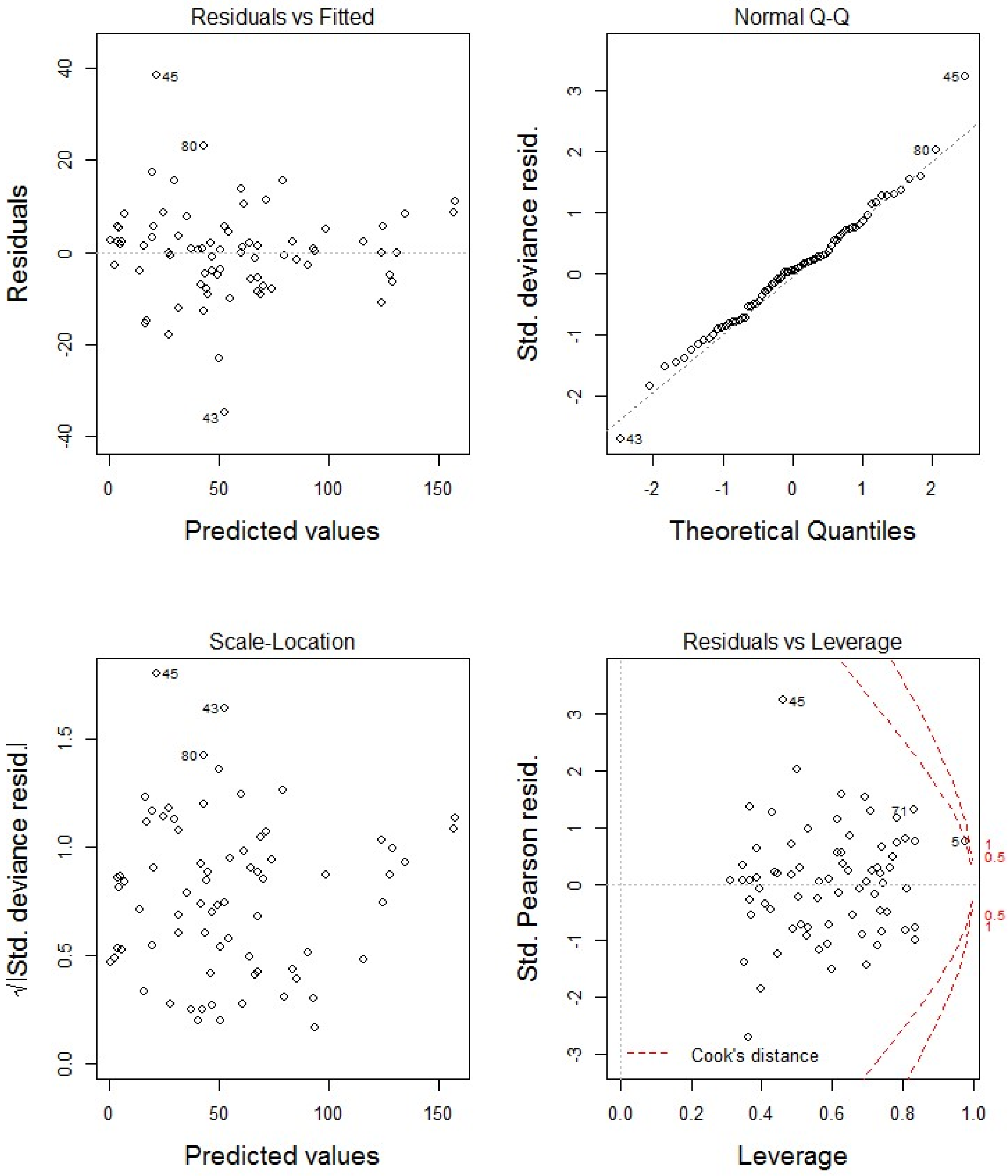


Figure E7 Residuals plots of model based on three-fold cross validation of fold2 with Google land use set.

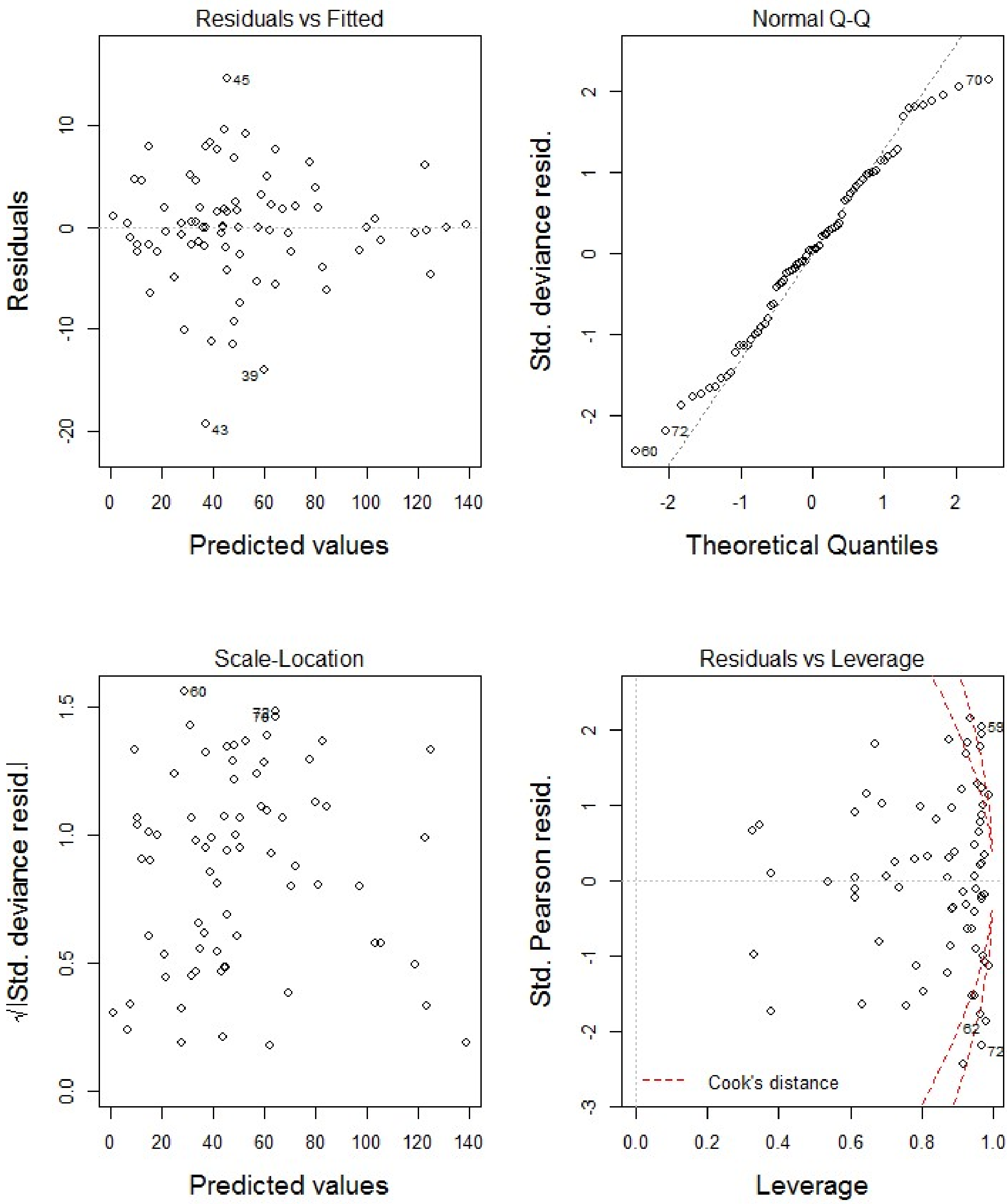


Figure E8 Residuals plots of model based on three-fold cross validation of fold3 with Google land use set.

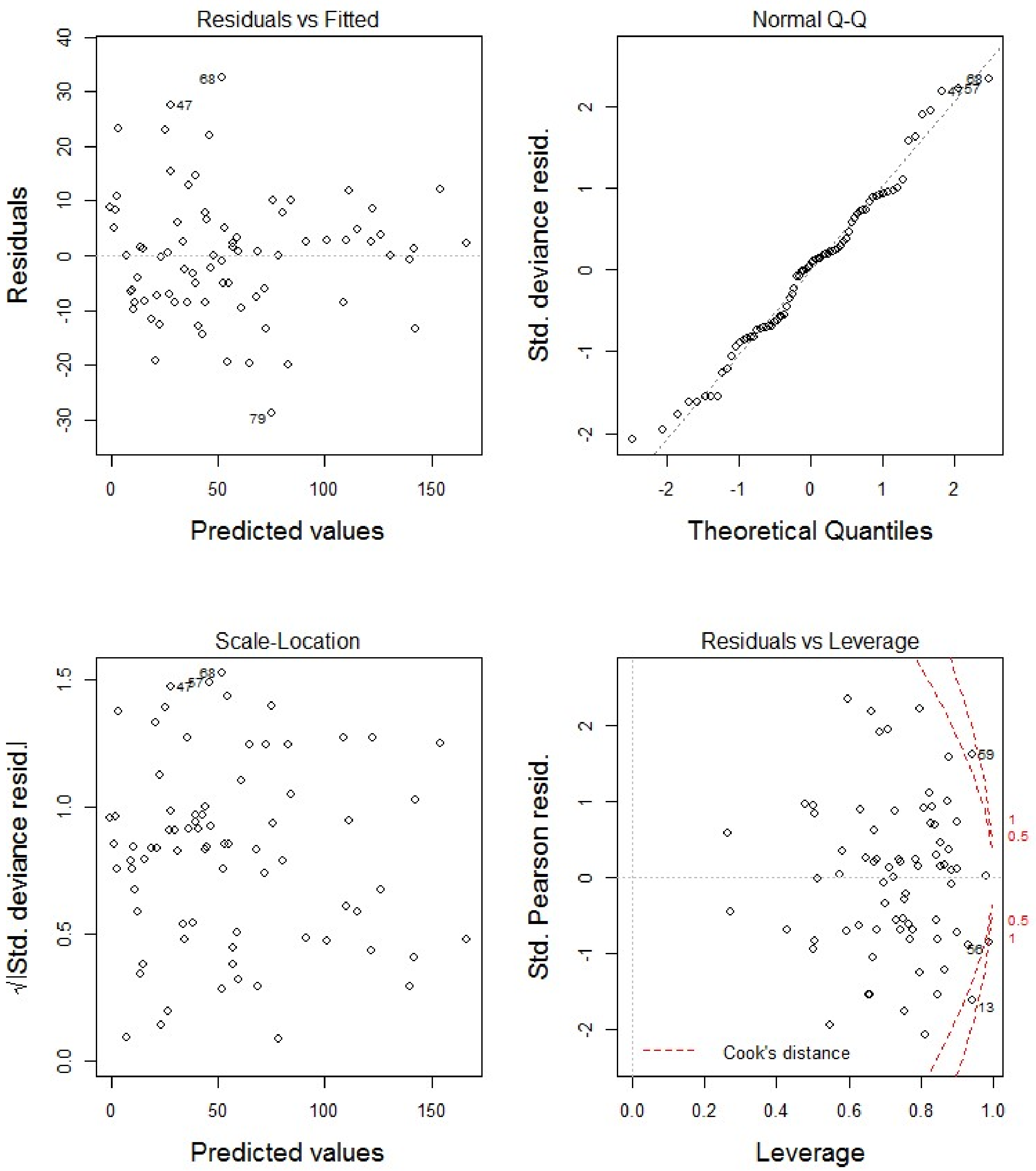


Figure E9 Residuals plots of model based on three-fold cross validation of fold1 with GIS land use set.

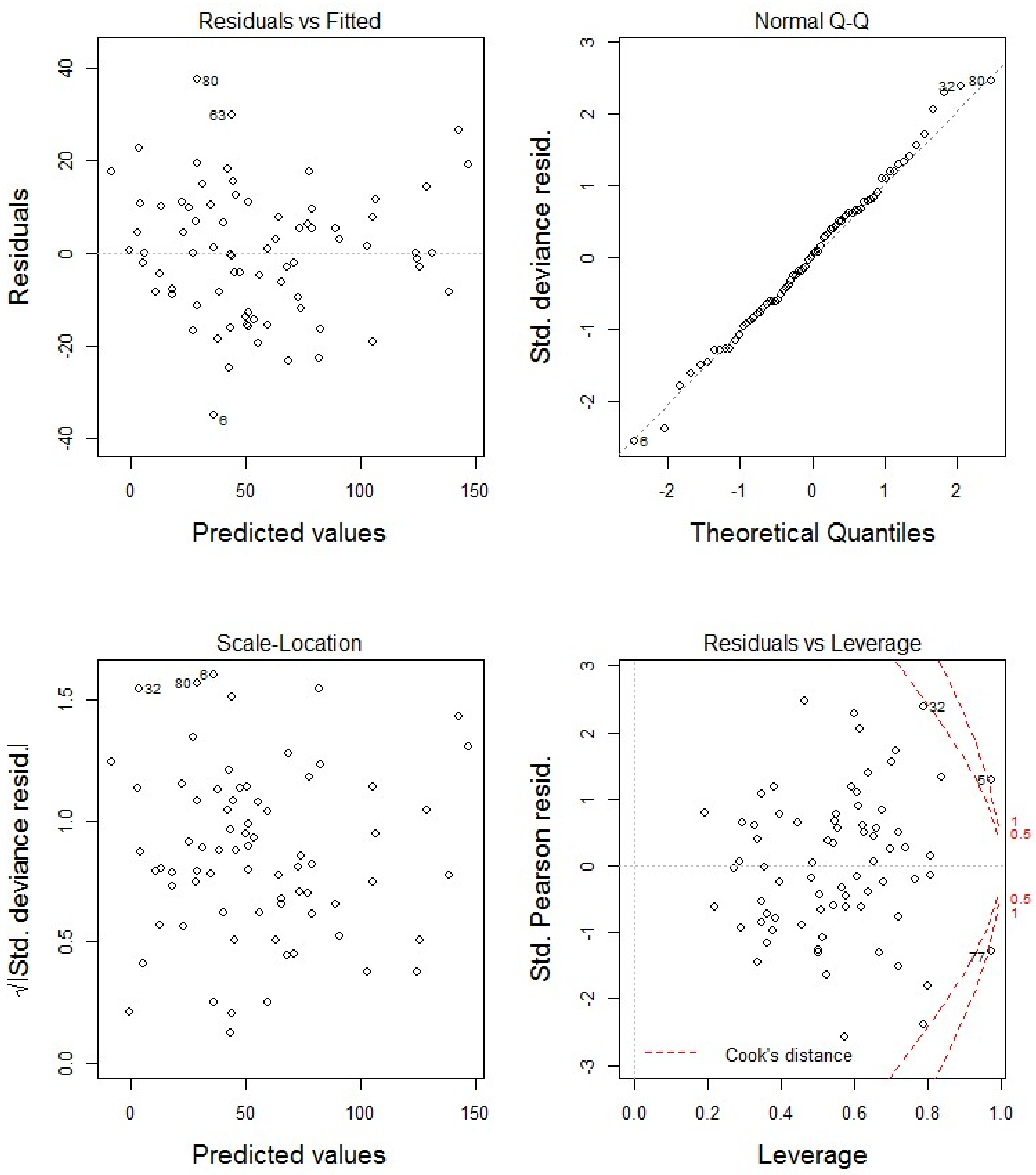


Figure E10 Residuals plots of model based on three-fold cross validation of fold2 with GIS land use set.

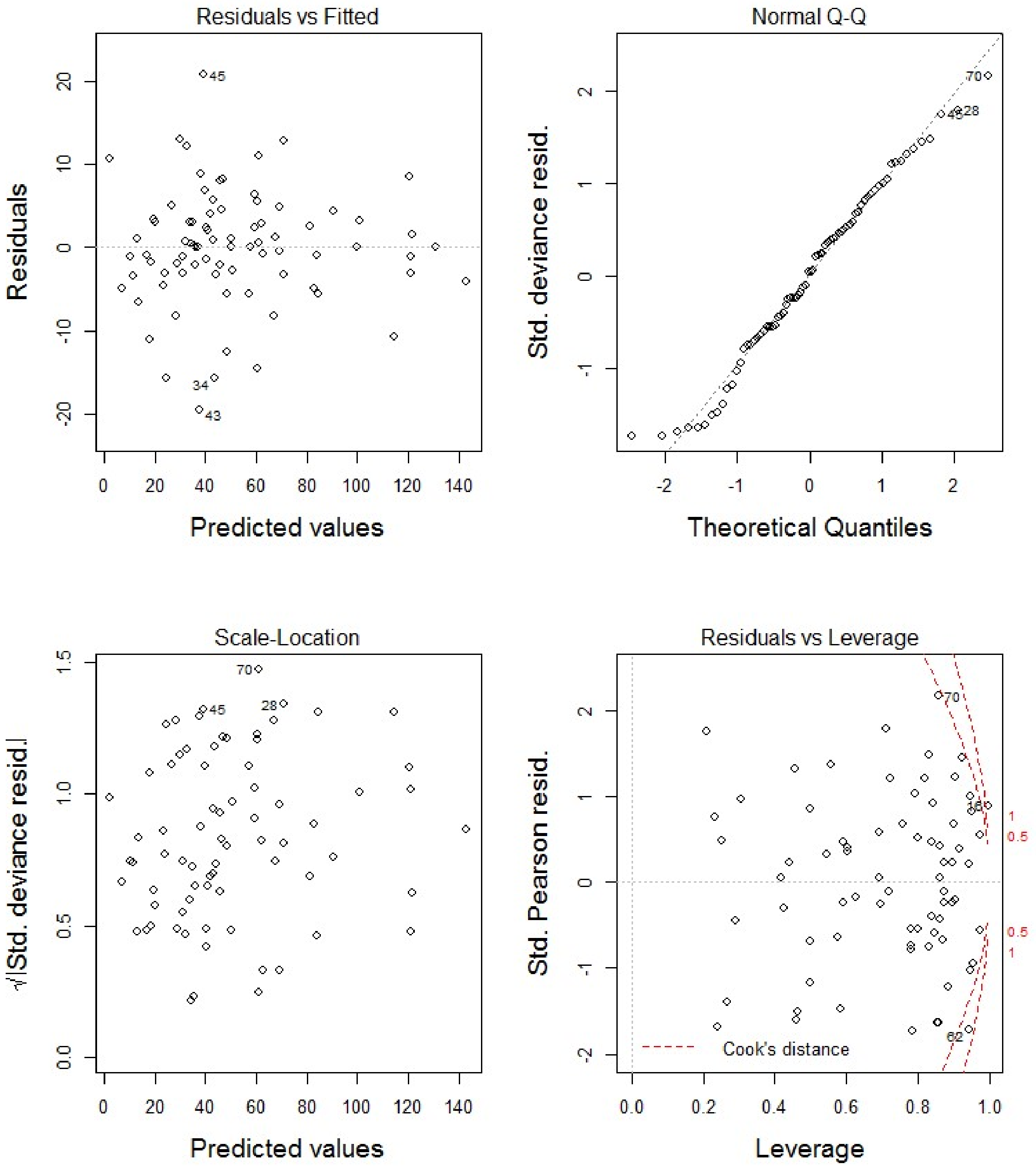


Figure E11 Residuals plots of model based on three-fold cross validation of fold3 with GIS land use set.

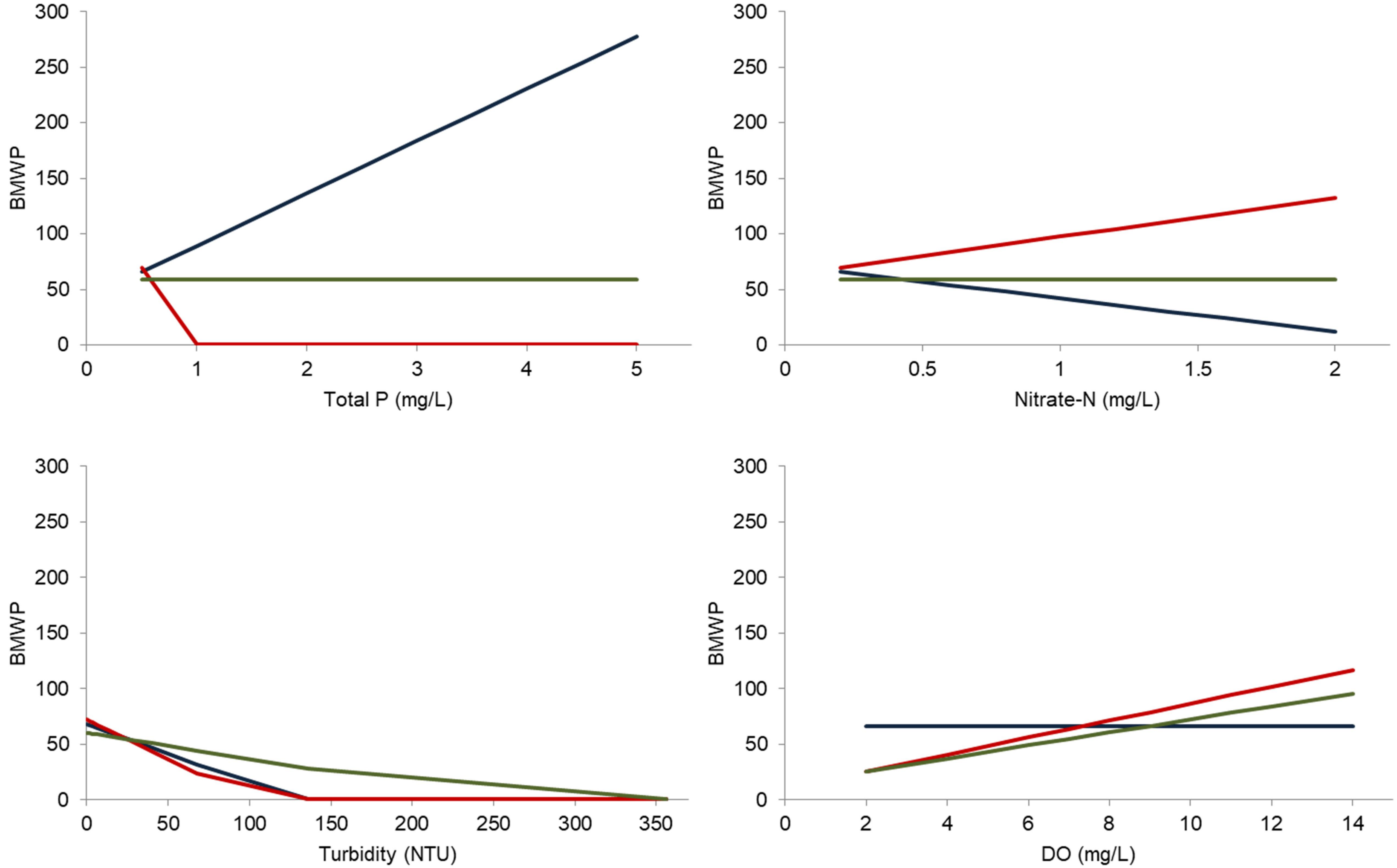


Figure E12 Sensitivity analysis showing the effect of total P, nitrate-N, turbidity and DO concentrations on the BMWP-Col: —— (FP land use arable), —— (Google land use arable), —— (GIS land use agriculture).

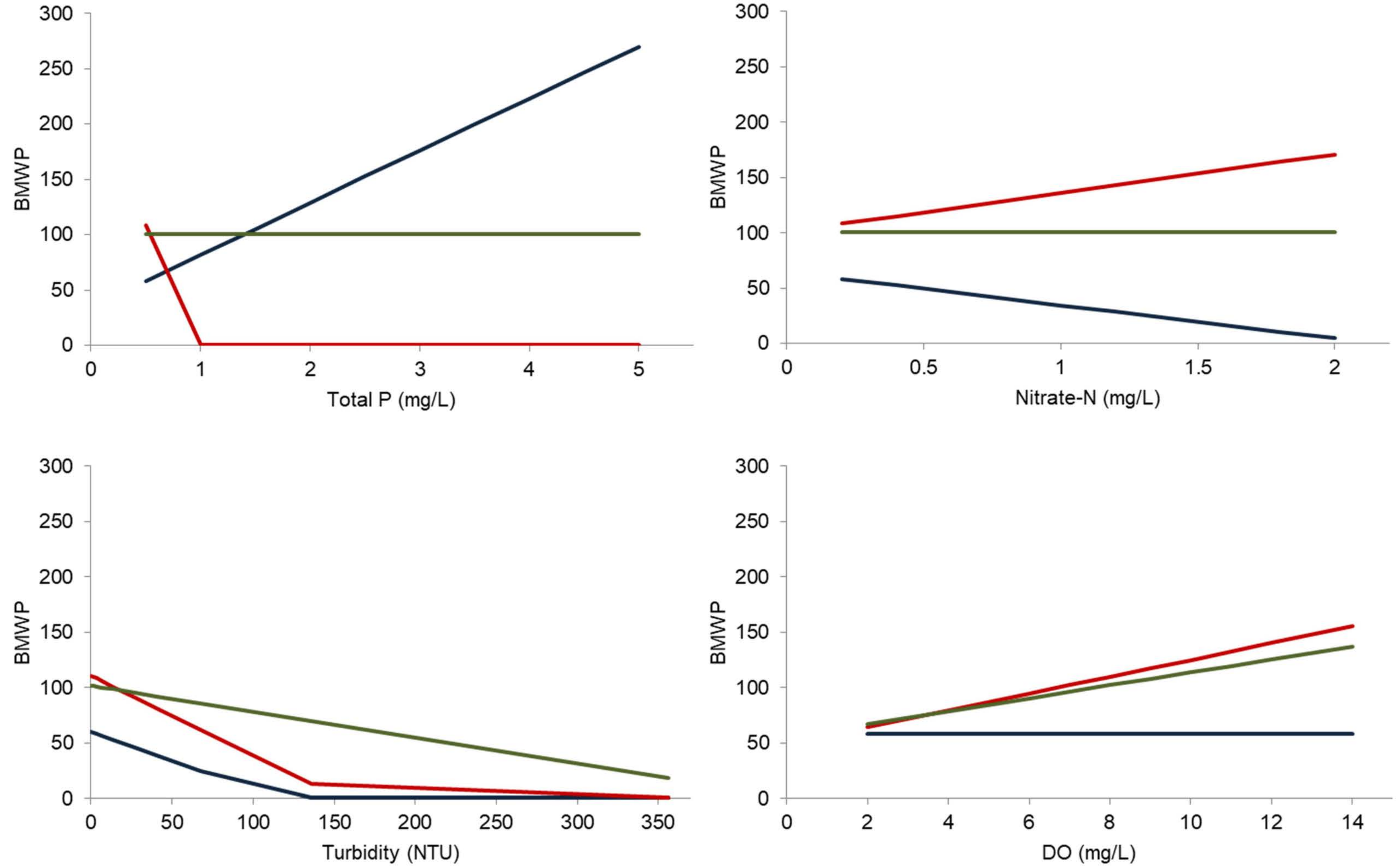


Figure E13 Sensitivity analysis showing the effect of total P, nitrate-N, turbidity and DO concentrations on the BMWP-Col: (FP land use orchard), (Google land use pasture), (GIS land use all other categories).

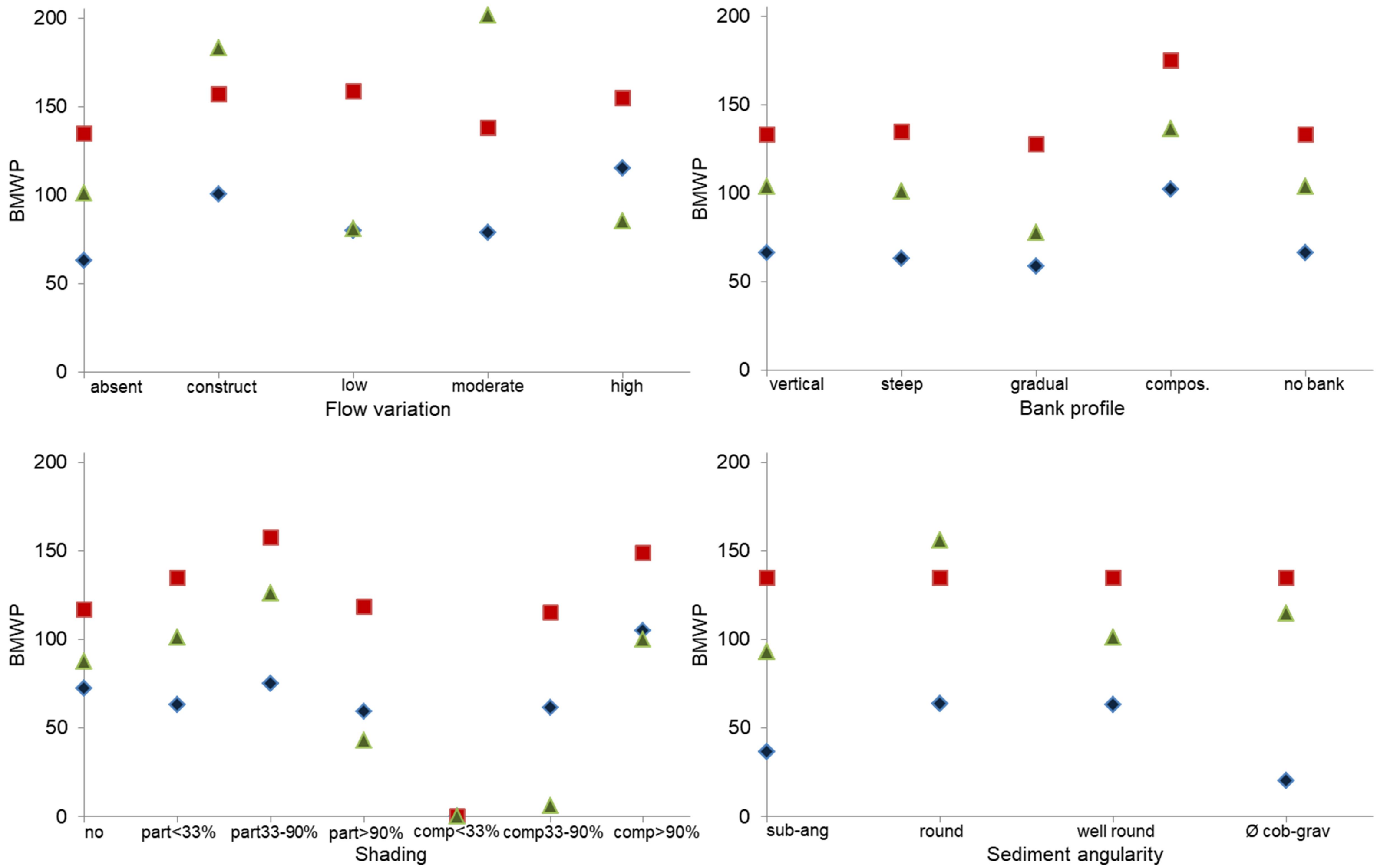


Figure E14 Sensitivity analysis showing the effects of changing flow variation, bank profile, shading and sediment angularity on the BMWP-Col: ◆ (FP land use forest), ■ (Google land use forest), ▲ (GIS land use all other categories). The classification of categorical variables is based on Table B1; construct: construction, compos: composite, part: partly, comp: completely, sub-ang: sub angular, cob-grav: cobble-pebble-gravel.

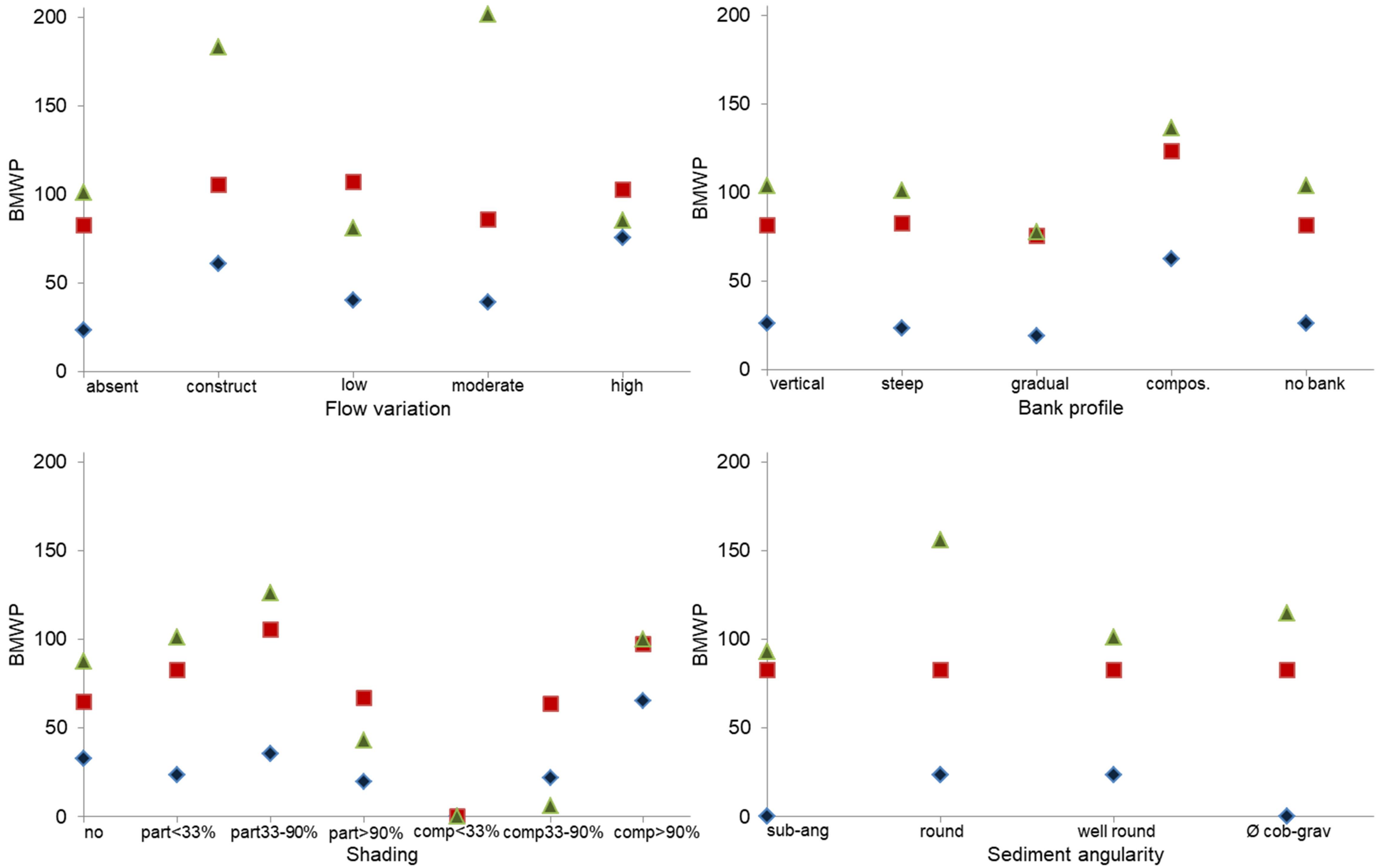


Figure E15 Sensitivity analysis showing the effects of changing flow variation, bank profile, shading and sediment angularity on the BMWP-Col: ◆ (FP land use residential), ■ (Google land use residential), ▲ (GIS land use all other categories). The classification of categorical variables is based on Table B1; construct: construction, compos: composite, part: partly, comp: completely, sub-ang: sub angular, cob-grav: cobble-pebble-gravel.

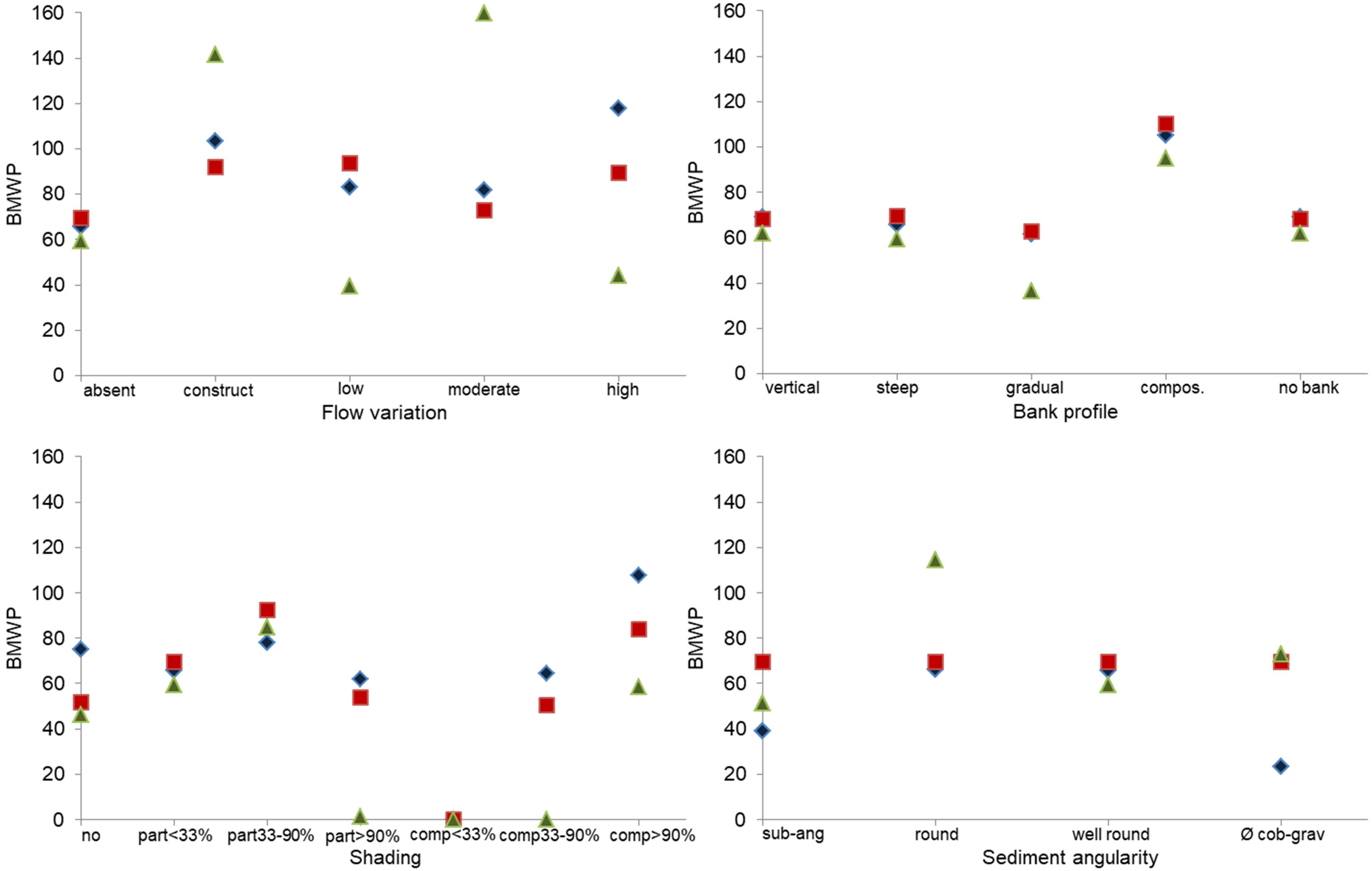


Figure E16 Sensitivity analysis showing the effects of changing flow variation, bank profile, shading and sediment angularity on the BMWP-Col: ◆ (FP land use arable), ■ (Google land use arable), ▲ (GIS land use agriculture). The classification of categorical variables is based on Table B1; construct: construction, compos: composite, part: partly, comp: completely, sub-ang: sub angular, cob-grav: cobble-pebble-gravel.

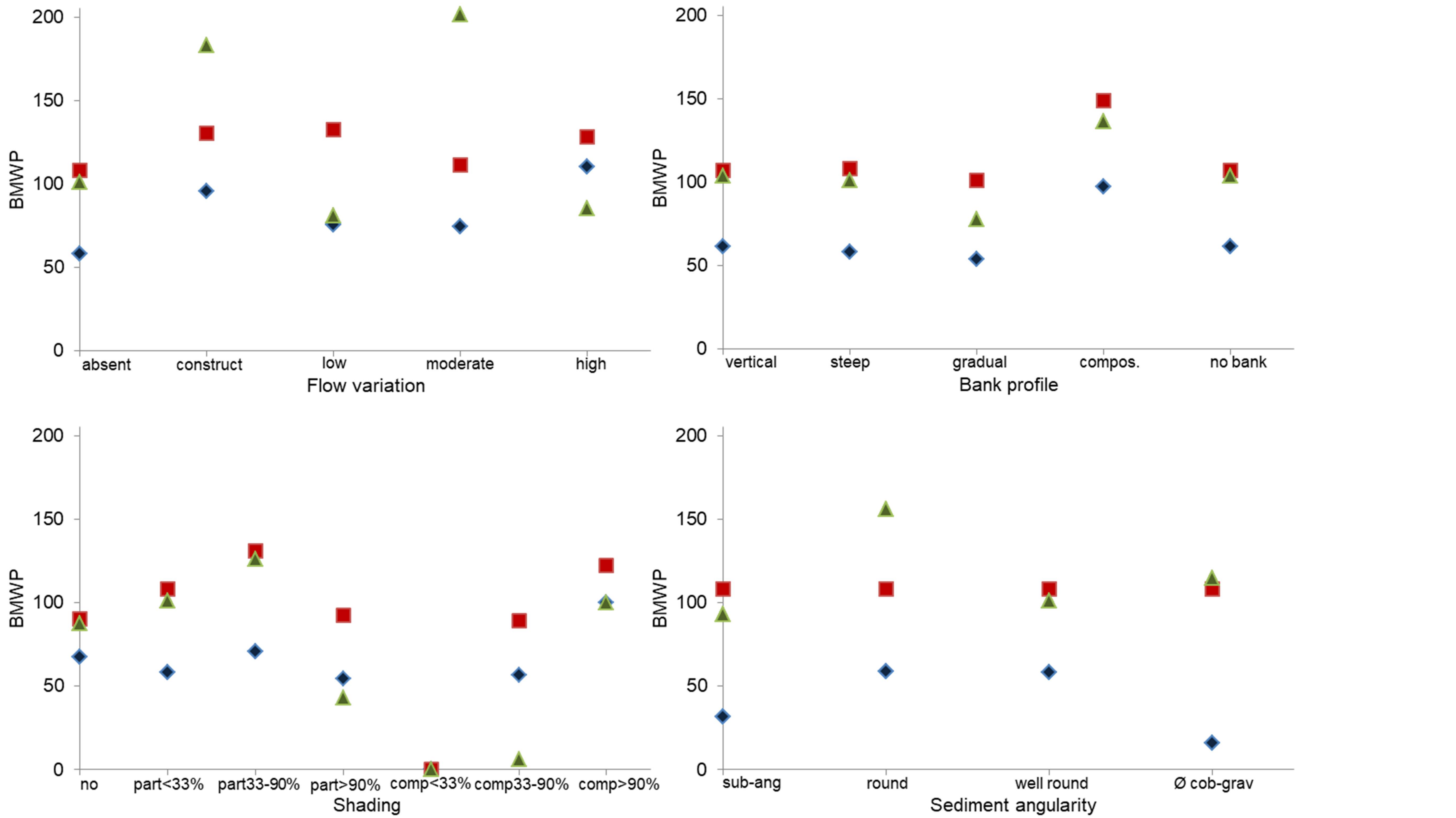


Figure E17 Sensitivity analysis showing the effects of changing flow variation, bank profile, shading and sediment angularity on the BMWP-Col: ◆ (FP land use orchard), ■ (Google land use pasture), ▲ (GIS land use all other categories). The classification of categorical variables is based on Table B1; construct: construction, compos: composite, part: partly, comp: completely, sub-ang: sub angular, cob-grav: cobble-pebble-gravel.

Tables

Table E1 Site's classification based on FP land use and Google land use.

FP land use	Google land use			
	Forest	Arable	Residential	Pasture
Forest	16, 17, 31, 34, 35, 59-63, 106	13, 18, 19, 27, 28, 37, 40, 86, 88, 90-95, 99, 100, 107, 108, 114, 115, 120	-	23, 24, 51, 53, 55-58, 66-69, 71-78, 80, 82-84, 89, 102, 105, 116
Arable	15	4, 8, 12, 22, 26, 29, 30, 38, 41, 44, 49, 54, 85, 96, 111, 118	21, 45, 46, 50, 113	2, 3, 33, 47, 48, 70, 87, 97, 98, 109, 110, 119
Residential	14, 64	6, 9, 36, 39, 42, 43, 52, 117	5, 7, 10, 11, 12, 65	25
Orchard	-	1, 32, 104	-	79, 81, 101, 103, 112

Table E2 Site's classification based on FP land use and GIS land use.

FP land use	GIS land use	
	Agriculture	All other categories
Forest	13, 23, 27, 28, 34, 37, 40, 55, 58-63, 105-108, 114, 116	16-19, 24, 31, 35, 51, 53, 56, 57, 66-69, 71-78, 80, 82-84, 86, 88-95, 99, 100, 102, 115, 120
Arable	2, 3, 8, 15, 26, 29, 30, 38, 41, 44, 49, 50, 54, 96, 97, 111, 113	4, 12, 21, 22, 33, 45-48, 70, 85, 87, 98, 109, 110, 118, 119
Residential	5-7, 9, 14, 36, 39, 42, 117	10, 11, 20, 25, 43, 52, 64, 65
Orchard	1, 32	79, 81, 101, 103, 104, 112

Table E3 Site's classification based on Google land use and GIS land use.

Google land use	GIS land use	
	Agriculture	All other categories
Forest	14, 15, 34, 59-63, 106	16, 17, 31, 35, 64
Arable	1, 6, 8, 9, 13, 26-30, 32, 36-42, 44, 49, 54, 96, 107, 108, 111, 114, 117	4, 12, 18, 19, 22, 43, 52, 85, 86, 88, 90-95, 99, 100, 104, 115, 118, 120
Residential	5, 7, 50, 113	10, 11, 20, 21, 45, 46, 65
Pasture	2, 3, 23, 55, 58, 97, 105, 116	24, 25, 33, 47, 48, 51, 53, 56, 57, 66-84, 87, 89, 98, 101-103, 109, 110, 112, 119

Table E4 Variables removal based on VIF values involving FP land use.

Variables	Removed due to high correlation	Included in model development
Nitrite-N	Highly collinear	
TDS	Highly collinear	
Total N	Highly collinear	
Abundance macrophytes	Highly collinear	
Pool/riffle class	Highly collinear	
pH	Highly collinear	
Main sediment	Highly collinear	
Ammonium-N	Highly collinear	
Bed compaction	Highly collinear	
Sediment matrix	Highly collinear	
Chlorophyll	Highly collinear	
Temperature	Highly collinear	
Conductivity		Yes
Chloride		Yes
DO		Yes
Turbidity		Yes
Total P		Yes
Nitrate-N		Yes
Velocity		Yes
Shading		Yes
Main macrophytes		Yes
Valley form		Yes
Channel form		Yes
Variation in width		Yes
Erosion		Yes
Bank profile		Yes
Variation in flow		Yes
Sludge layer		Yes
Twigs		Yes
Branch		Yes
Logs		Yes
Bank shape		Yes
Bank slope		Yes
Sediment angularity		Yes
FP land use		Yes

Table E5 Variables removal based on VIF values involving Google land use.

Variables	Removed due to high correlation	Included in model development
Nitrite-N	Highly collinear	
TDS	Highly collinear	
Total N	Highly collinear	
Abundance macrophytes	Highly collinear	
Pool/riffle class	Highly collinear	
pH	Highly collinear	
Main sediment	Highly collinear	
Ammonium-N	Highly collinear	
Conductivity	Highly collinear	
Sediment matrix	Highly collinear	
Chlorophyll	Highly collinear	
Temperature		Yes
Chloride		Yes
DO		Yes
Turbidity		Yes
Total P		Yes
Nitrate-N		Yes
Velocity		Yes
Shading		Yes
Main macrophytes		Yes
Valley form		Yes
Channel form		Yes
Variation in width		Yes
Erosion		Yes
Bank profile		Yes
Variation in flow		Yes
Sludge layer		Yes
Twigs		Yes
Branch		Yes
Logs		Yes
Bank shape		Yes
Bank slope		Yes
Bed compaction		Yes
Sediment angularity		Yes
Google land use		Yes

Table E6 Variables removal based on VIF values involving GIS land use.

Variables	Removed due to high correlation	Included in model development
Total P	Highly collinear	
TDS	Highly collinear	
Total N	Highly collinear	
Abundance macrophytes	Highly collinear	
Pool riffle	Highly collinear	
pH	Highly collinear	
Main sediment	Highly collinear	
Nitrite-N	Highly collinear	
Conductivity	Highly collinear	
Sediment matrix	Highly collinear	
Chlorophyll	Highly collinear	
Temperature		Yes
Chloride		Yes
DO		Yes
Turbidity		Yes
Nitrate-N		Yes
Ammonium-N		Yes
Velocity		Yes
Shading		Yes
Main macrophytes		Yes
Valley form		Yes
Channel form		Yes
Varwidth		Yes
Erosion		Yes
Bank profile		Yes
Varflow		Yes
Sludge layer		Yes
Twigs		Yes
Branch		Yes
Logs		Yes
Banks shape		Yes
Banks slope		Yes
Bed compaction		Yes
Sediment angularity		Yes
GIS land use		Yes

Table E7 Median, minimum and maximum values of variables used for sensitivity analyses.

Variables	Unit	Median	Min	Max
Temperature	° C	26.0	19.0	34.0
Conductivity	μS/cm	123	37	1981
Chloride	mg/L	2.5	0.5	181.7
DO	mg/L	7.8	2.0	13.6
Turbidity	NTU	3.4	0.0	355.6
Total P	mg/L	0.5	0.5	4.5
Nitrate-N	mg/L	0.2	0.2	2.0
Ammonium-N	mg/L	0.1	0.02	8.8
Velocity	m/s	0.2	0.0	1.5
Shading		2	1	7
Erosion		1	1	3
Main macrophytes		1	1	4
Variation in flow		1	1	5
Variation in width		2	1	6
Twigs		1	1	3
Branch		2	1	3
Logs		1	1	3
Bank profile		2	1	5
Bank shape		3	1	6
Bank slope		3	1	6
Sludge layer		2	1	4
Bed compaction		4	1	6
Sediment angularity		5	3	6
Valley form		5	2	7
Channel form		5	1	7
FP land use		1	1	4
Google land use		3	1	4
GIS land use		1	1	2
BMWP-Col		47	0	169

Table E8 Variable's selection in model development for three-fold cross validation involving FP land use.

Model with fold1	Model with fold2	Model with fold3
Twigs, AIC = 688.3	Bank slope, AIC = 726.42	Velocity, AIC = 688.53
Conductivity, AIC = 684.42	Total P, AIC = 721.64	Channel form, AIC = 686.53
Sludge layer, AIC = 682.54	Bank profile, AIC = 719.7	Logs, AIC = 685.09
Logs, AIC = 686.23	Branch, AIC = 721.14	Varwidth, AIC = 681.09
Bank shape, AIC = 688.4	DO, AIC = 720.92	DO, AIC = 678.59
Chloride, AIC = 692.39	Valley form, AIC = 719.7	Branch, AIC = 676.66
Velocity, AIC = 690.77	Velocity, AIC = 723.39	Twigs, AIC = 676.53
Channel form, AIC = 691.85	Sediment angularity, AIC = 726.27	Conductivity, AIC = 677.48
DO, AIC = 698.15	Turbidity, AIC = 732.2	Erosion, AIC = 677.23
Main mac, AIC = 697.94	Erosion, AIC = 731.65	Chloride, AIC = 679.11
Varwitdh, AIC = 700.89	Twigs, AIC = 732.67	Valley form, AIC = 680.79
	Chloride, AIC = 737.83	Nitrate-N, AIC = 688.06

Table E9 Variable's selection in model development for three-fold cross validation involving Google land use.

Model with fold1	Model with fold2	Model with fold3
Sludge layer, AIC = 631.14	Sludge layer, AIC = 693.74	Channel form, AIC = 635.87
Bed compaction, AIC = 625.34	Bed compaction, AIC = 690.03	Velocity, AIC = 639.81
Temperature, AIC = 637.66	Sediment angularity, AIC = 687.5	Temperature, AIC = 637.93
Bank slope, AIC = 639.64	Valley form, AIC = 684.95	Shading, AIC = 635.99
Twigs, AIC = 654.43	Velocity, AIC = 680.89	Nitrate-N, AIC = 638.05
Logs, AIC = 660.94	Temperature, AIC = 679.47	Sludge layer, AIC = 636.92
Channel form, AIC = 670.79	Branch, AIC = 678.56	DO, AIC = 642.04
Main mac, AIC = 683.2	Erosion, AIC = 680.14	Bank shape, AIC = 643.58
Erosion, AIC = 688.03	Bank slope, AIC = 684.77	Valley form, AIC = 649.77
Velocity, AIC = 695.38	Bank profile, AIC = 691.73	Twigs, AIC = 659.05
	Channel form, AIC = 696.92	
	Varwidth, AIC = 698.74	

Table E10 Variable's selection in model development for three-fold cross validation involving GIS land use.

Model with fold1	Model with fold2	Model with fold3
Logs, AIC = 722.57	Sludge layer, AIC = 736.15	Ammonium-N, AIC = 662.56
Velocity, AIC = 719.45	Velocity, AIC = 732.45	Logs, AIC = 660.56
Temperature, AIC = 717.67	GIS land use, AIC = 730.52	Twigs, AIC = 657.62
Nitrate-N, AIC = 716.43	Ammonium-N, AIC = 728.62	Turbidity, AIC = 655.3
Main mac, AIC = 715.21	Valley form, AIC = 726.85	DO, AIC = 653.95
Chloride, AIC = 716.45	Erosion, AIC = 727.73	Velocity, AIC = 653.61
Sludge layer, AIC = 714.61	Bank slope, AIC = 725.03	Temperature, AIC = 654.4
Branch, AIC = 718.36	Bank profile, AIC = 722.83	Channel form, AIC = 654.82
Valley form, AIC = 722.82	Branch, AIC = 719.94	GIS land use, AIC = 663.65
DO, AIC = 728.08	Bed compaction, AIC = 720.13	Sludge layer, AIC = 662.73
Varwidth, AIC = 727.66	Varwidth, AIC = 723.08	Chloride, AIC = 674.13
Ammonium-N, AIC = 730.83	Channel form, AIC = 724.27	Valley form, AIC = 675.76
Channel form, AIC = 748.66	Nitrate-N, AIC = 724.91	
	DO, AIC = 752.98	

Table E11 Model performance with three-fold cross validation: showing the number of input variables that construct the models, the Kappa performed on BMWP-Col classes, coefficient of determinations (R^2) and p-values of training and testing sets.

Models	# input variables	Kappa on class: un/weighted		R^2		p-value	
		Training set	Testing set	Training set	Testing set	Training set	Testing set
FP land use fold1	21	0.65/0.90	-0.01/0.20	0.91	0.02	< 0.05	0.33
FP land use fold2	18	0.55/0.84	0.07/0.37	0.85	0.18	< 0.05	< 0.05
FP land use fold3	17	0.60/0.88	0.05/0.46	0.92	0.15	< 0.05	< 0.05
Average FP land use	19	0.60/0.87	0.02/0.26	0.89	0.12	<0.05	0.11
Google land use fold1	23	0.76/0.92	0.03/0.43	0.94	0.31	< 0.05	< 0.05
Google land use fold2	18	0.66/0.88	0.11/0.36	0.88	0.43	< 0.05	< 0.05
Google land use fold3	24	0.78/0.94	0.14/0.36	0.97	0.09	< 0.05	0.07
Average Google land use	22	0.73/0.91	0.09/0.38	0.93	0.28	< 0.05	< 0.05
GIS land use fold1	18	0.57/0.88	0.16/0.20	0.89	0.06	< 0.05	0.12
GIS land use fold2	16	0.59/0.85	0.01/-0.15	0.83	0.03	< 0.05	0.33
GIS land use fold3	19	0.75/0.92	0.09/0.29	0.95	0.09	< 0.05	0.07
Average GIS land use	18	0.64/0.88	0.09/0.11	0.89	0.06	< 0.05	0.17

Table E12 Variable's ranking for each fold with different land use assessment methods; only p-value of significant variable is given within brackets.

Variables	**FP land use**			**Google land use**			**GIS land use**		
	Fold1	**Fold2**	**Fold3**	**Fold1**	**Fold2**	**Fold3**	**Fold1**	**Fold2**	**Fold3**
Temperature				20		20		1 (< 0.05)	18
Conductivity		1 (< 0.05)	16						
Chloride	19	4 (< 0.05)	13	13 (< 0.05)	8 (< 0.05)	21		3 (< 0.05)	9
DO	14			10 (< 0.05)	10 (< 0.05)	14	14	10	
Turbidity	15	6 (< 0.05)	11	7 (< 0.05)	9 (< 0.05)	15	6	6 (< 0.05)	
Total P	11 (< 0.05)		10	3 (< 0.05)	2 (< 0.05)	10			
Nitrate-N	1 (< 0.05)	9 (< 0.05)	12	2 (< 0.05)	5 (< 0.05)	13		13	13
Ammonium-N							10		
Velocity	13	16		8 (< 0.05)		19			19
Shading	6 (< 0.05)	5 (< 0.05)	7 (< 0.05)	4 (< 0.05)	11 (< 0.05)	18	3 (< 0.05)	2 (< 0.05)	4 (< 0.05)
Main macrophytes	18	15	4 (< 0.05)	15	6 (< 0.05)	9		9	8
Valley form	12	18	15	5 (< 0.05)		23	15		17
Channel form	16	13		14	14 (< 0.05)	24	11	8	15
Varwidth	8 (< 0.05)	11		6 (< 0.05)	12 (< 0.05)	6	12	12	12
Erosion	9 (< 0.05)	8 (< 0.05)	14	11 (< 0.05)	16	3	7		10
Bank profile	5 (< 0.05)		5 (< 0.05)	16	13 (< 0.05)	1	5 (< 0.05)		3 (< 0.05)
Varflow	3 (< 0.05)	12	6 (< 0.05)	12 (< 0.05)	15 (< 0.05)	4	1 (< 0.05)	11	6 (< 0.05)
Sludge layer	21	14	8 (< 0.05)			11	18		16
Twigs		17	17	22	3 (< 0.05)	16	8	7 (< 0.05)	
Branch	7 (< 0.05)			21	18	7	17	16	14
Logs	17	2 (< 0.05)		18	7 (< 0.05)	17		14	
Bank shape	20	3 (< 0.05)	3 (< 0.05)	19	4 (< 0.05)	12	13	5 (< 0.05)	5 (< 0.05)
Bank slope	10 (< 0.05)		2 (< 0.05)	17	17	5	16		2 (< 0.05)
Bed compaction				23		22	9	15	7

Variables	**FP land use**			**Google land use**			**GIS land use**		
	Fold1	**Fold2**	**Fold3**	**Fold1**	**Fold2**	**Fold3**	**Fold1**	**Fold2**	**Fold3**
Angularity	4 (< 0.05)	7 (< 0.05)	1 (< 0.05)	9 (< 0.05)		8	2 (< 0.05)	4 (< 0.05)	1 (< 0.05)
FP land use	2 (< 0.05)	10 (< 0.05)	9 (< 0.05)						
Google land use				1 (< 0.05)	1 (< 0.05)	2			
GIS land use							4 (< 0.05)		11

References

Abouali, M., Nejadhashemi, A.P., Daneshvar, F., Woznicki, S.A., 2016. Two-phase approach to improve stream health modeling. Ecol. Inform. 34, 13-21.DOI: 10.1016/j.ecoinf.2016.04.009.

Alahuhta, J., Virtala, A., Hjort, J., Ecke, F., Johnson, L.B., Sass, L., Heino, J., 2017. Average niche breadths of species in lake macrophyte communities respond to ecological gradients variably in four regions on two continents. Oecologia 184(1), 219-235.DOI: 10.1007/s00442-017-3847-y.

Alemneh, T., Ambelu, A., Bahrndorff, S., Mereta, S.T., Pertoldi, C., Zaitchik, B.F., 2017. Modeling the impact of highland settlements on ecological disturbance of streams in Choke Mountain Catchment: Macroinvertebrate assemblages and water quality. Ecol. Indic. 73, 452-459.DOI: 10.1016/j.ecolind.2016.10.019.

Allan, E., Manning, P., Alt, F. *et al.*, 2015. Land use intensification alters ecosystem multifunctionality via loss of biodiversity and changes to functional composition. Ecol. Lett. 18(8), 834-843.DOI: 10.1111/ele.12469.

Alvarez-Cabria, M., Gonzalez-Ferreras, A.M., Penas, F.J., Barquin, J., 2017. Modelling macroinvertebrate and fish biotic indices: From reaches to entire river networks. Sci. Total Environ. 577, 308-318.DOI: 10.1016/j.scitotenv.2016.10.186.

Alvarez-Mieles, G., Irvine, K., Griensven, A.V., Arias-Hidalgo, M., Torres, A., Mynett, A.E., 2013. Relationships between aquatic biotic communities and water quality in a tropical river-wetland system (Ecuador). Environ. Sci. Policy 34, 115-127.DOI: 10.1016/j.envsci.2013.01.011.

Alvarez, A.L.F., 2005. Metodologia para la evaluacion de los macroinvertebrados acuaticos como indicadores de los recursos hidrobiologicos. Instituto Alexander von Humboldt. http://biblioteca.humboldt.org.co/cgi-bin/koha/opac-detail.pl?biblionumber=6014&query_desc=su%3A%22Bioindicadores%22. Accessed 30 June 2014.

Arias-Hidalgo, M., Villa-Cox, G., Griensven, A.V., Solorzano, G., Villa-Cox, R., Mynett, A.E., Debels, P., 2013. A decision framework for wetland management in a river basin context: The "Abras de Mantequilla" case study in the Guayas River Basin, Ecuador. Environ. Sci. Policy 34, 103-114.DOI: 10.1016/j.envsci.2012.10.009.

Arimoro, F.O., Odume, O.N., Uhunoma, S.I., Edegbene, A.O., 2015. Anthropogenic impact on water chemistry and benthic macroinvertebrate associated changes in a southern Nigeria stream. Environ. Monit. Assess. 187(2).DOI: ARTN 14 10.1007/s10661-014-4251-2.

Armitage, P.D., Moss, D., Wright, J.F., 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 Res. 17(3), 333-347.DOI: 10.1016/0043-1354(83)90188-4.

Arriaga, L., 1989. The Daule-Peripe dam project, urban development of Guayaquil and their impact on shrimp mariculture, in: Olsen S., Arriaga L. (Eds.), A sustainable shrimp mariculture industry for Ecuador. Coastal Resources Center, University of Rhode Island Narragansett, RI.

Austin, M.P., 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. Ecol. Model. 157(2-3), 101-118.DOI: Pii S0304-3800(02)00205-3, 10.1016/S0304-3800(02)00205-3.

Bailey, R.C., Norris, R.H., Reynoldson, T.B., 2001. Taxonomic resolution of benthic macroinvertebrate communities in bioassessments. J. N. Am. Benthol. Soc. 20(2), 280-286.DOI: 10.2307/1468322.

Baillie, B.R., Neary, D.G., 2015. Water quality in New Zealand's planted forests: a review. N. Z. J. For. Sci. 45.DOI: Artn 7 10.1186/S40490-015-0040-0.

Bainbridge, Z.T., Brodie, J.E., Faithful, J.W., Sydes, D.A., Lewis, S.E., 2009. Identifying the land-based sources of suspended sediments, nutrients and pesticides discharged to the Great Barrier Reef from the Tully-Murray Basin, Queensland, Australia. Mar. Freshw. Res. 60(11), 1081-1090.DOI: 10.1071/Mf08333.

Ballance, R., 1996. Chapter 7 - Physical and chemical analyses, in: Bartram J., Ballance R. (Eds.), Water Quality Monitoring - A Practical Guide to the Design and Implementation of Freshwater Quality Studies and Monitoring Programmes. Published on behalf of United Nations Environment Programme and the World Health Organization

Baltazar, D.E.S., Magcale-Macandog, D., Tan, M.F.O., Zafaralla, M.T., Cadiz, N.M., 2016. A River Health Status Model Based on Water Quality, Macroinvertebrates and Land Use for Niyugan River, Cabuyao City, Laguna, Philippines. J. Environ. Sci. Manag. 19(2), 38-53.

Baptista, D.F., Henriques-Oliveira, A.L., Oliveira, R.B.S., Mugnai, R., Nessimian, J.L., Buss, D.F., 2013. Development of a benthic multimetric index for the Serra da Bocaina bioregion in Southeast Brazil. Braz. J. Biol. 73(3), 573-583.

Barber, L.B., Paschke, S.S., Battaglia, W.A., Douville, C., Fitzgerald, K.C., Keefe, S.H., Roth, D.A., Vajda, A.M., 2017. Effects of an Extreme Flood on Trace Elements in River Water-From Urban Stream to Major River Basin. Environ. Sci. Technol. 51(18), 10344-10356.DOI: 10.1021/acs.est.7b01767.

Barbour, M.T., Gerritsen, J., Snyder, B.D., Stribling, J.B., 1999. Rapid bioassessment protocols for use in streams and wadeable rivers: periphyton, benthic macroinvertebrates and fish, 2nd ed. EPA 841-B-99-002. U.S. Environmental Protection Agency, Office of Water, Washington, D.C.

Barton, D.R., 1996. The use of Percent Model Affinity to assess the effects of agriculture on benthic invertebrate communities in headwater streams of southern Ontario, Canada. Freshw. Biol. 36(2), 397-410.DOI: 10.1046/j.1365-2427.1996.00053.x.

Bartram, J., Ballance, R., 1996. Water quality monitoring : a practical guide to the design and implementation of freshwater quality studies and monitoring programmes, 1st ed. E&FN on behalf of UNEP/WHO, London, UK.

Bass, M.S., Finer, M., Jenkins, C.N., Kreft, H., Cisneros-Heredia, D.F., McCracken, S.F., Pitman, N.C.A., English, P.H., Swing, K., Villa, G., Di Fiore, A., Voigt, C.C., Kunz, T.H., 2010. Global Conservation Significance of Ecuador's Yasuni National Park. Plos One 5(1).DOI: ARTN e8767 10.1371/journal.pone.0008767.

Beals, E.W., 1984. Bray-Curtis Ordination - an Effective Strategy for Analysis of Multivariate Ecological Data. Adv. Ecol. Res. 14, 1-55.DOI: 10.1016/S0065-2504(08)60168-3.

Beasley, G., Kneale, P., 2002. Reviewing the impact of metals and PAHs on macro invertebrates in urban watercourses. Prog. Phys. Geogr. 26(2), 236-270.DOI: 10.1191/0309133302pp334ra.

Beisel, J.N., Usseglio-Polatera, P., Thomas, S., Moreteau, J.C., 1998. Stream community structure in relation to spatial variation: the influence of mesohabitat characteristics. Hydrobiologia 389(1-3), 73-88.DOI: 10.1023/A:1003519429979.

Bellucci, C.J., Becker, M., Beauchene, M., 2011. Characteristics of Macroinvertebrate and Fish Communities From 30 Least Disturbed Small Streams in Connecticut. Northeast. Nat. 18(4), 411-444.DOI: 10.1656/045.018.0402.

Bennetsen, E., Gobeyn, S., Goethals, P.L.M., 2016. Species distribution models grounded in ecological theory for decision support in river management. Ecol. Model. 325, 1-12.DOI: 10.1016/j.ecolmodel.2015.12.016.

Berger, E., Haase, P., Kuemmerlen, M., Leps, M., Schafer, R.B., Sundermann, A., 2017. Water quality variables and pollution sources shaping stream macroinvertebrate communities. Sci. Total Environ. 587, 1-10.DOI: 10.1016/j.scitotenv.2017.02.031.

Berk, R.A., 2008. Statistical learning from a regression perspective. Springer Verlag, New York, NY.

Bertone, E., Sahin, O., Richards, R., Roiko, A., 2016. Extreme events, water quality and health: A participatory Bayesian risk assessment tool for managers of reservoirs. J. Clean. Prod. 135, 657-667.DOI: 10.1016/j.jclepro.2016.06.158.

Birk, S., Bonne, W., Borja, A., Brucet, S., Courrat, A., Poikane, S., Solimini, A., van de Bund, W.V., Zampoukas, N., Hering, D., 2012. Three hundred ways to assess Europe's surface waters: An almost complete overview of biological methods to implement the Water Framework Directive. Ecol. Indic. 18, 31-41.DOI: 10.1016/j.ecolind.2011.10.009.

Bixby, R.J., Cooper, S.D., Gresswell, R.E., Brown, L.E., Dahm, C.N., Dwire, K.A., 2015. Fire effects on aquatic ecosystems: an assessment of the current state of the science. Freshw. Sci. 34(4), 1340-1350.DOI: 10.1086/684073.

Blanchette, M.L., Pearson, R.G., 2013. Dynamics of habitats and macroinvertebrate assemblages in rivers of the Australian dry tropics. Freshw. Biol. 58(4), 742-757.DOI: 10.1111/fwb.12080.

Bolstad, P.V., Swank, W.T., 1997. Cumulative impacts of landuse on water quality in a southern Appalachian watershed. J. Am. Water Resour. Assoc. 33(3), 519-533.DOI: 10.1111/j.1752-1688.1997.tb03529.x.

Bonada, N., Rieradevall, M., Dallas, H., Davis, J., Day, J., Figueroa, R., Resh, V.H., Prat, N., 2008. Multi-scale assessment of macroinvertebrate richness and composition in Mediterranean-climate rivers. Freshw. Biol. 53(4), 772-788.DOI: 10.1111/j.1365-2427.2007.01940.x.

Borbor-Cordova, M.J., Boyer, E.W., Mcdowell, W.H., Hall, C.A., 2006. Nitrogen and phosphorus budgets for a tropical watershed impacted by agricultural land use: Guayas, Ecuador. Biogeochemistry 79(1-2), 135-161.DOI: 10.1007/s10533-006-9009-7.

Boulton, A.J., Scarsbrook, M.R., Quinn, J.M., Burrell, G.P., 1997. Land-use effects on the hyporheic ecology of five small streams near Hamilton, New Zealand. N. Z. J. Mar. Freshw. Res. 31(5), 609-622.

Brogna, D., Michez, A., Jacobs, S., Dufrene, M., Vincke, C., Dendoncker, N., 2017. Linking Forest Cover to Water Quality: A Multivariate Analysis of Large Monitoring Datasets. Water 9(3).DOI: Artn 176 10.3390/W9030176.

Brown, L.R., May, J.T., Wulff, M., 2012. Associations of Benthic Macroinvertebrate Assemblages with Environmental Variables in the Upper Clear Creek Watershed, California. West. North Am. Nat. 72(4), 473-494.

Bruno, D., Belmar, O., Sanchez-Fernandez, D., Guareschi, S., Millan, A., Velasco, J., 2014. Responses of Mediterranean aquatic and riparian communities to human pressures at different spatial scales. Ecol. Indic. 45, 456-464.DOI: 10.1016/j.ecolind.2014.04.051.

Bucker, A., Sondermann, M., Frede, H.G., Breuer, L., 2010. The influence of land-use on macroinvertebrate communities in montane tropical streams - a case study from Ecuador. Fundam. Appl. Limnol. 177(4), 267-282.DOI: 10.1127/1863-9135/2010/0177-0267.

Bussi, G., Janes, V., Whitehead, P.G., Dadson, S.J., Holman, I.P., 2017. Dynamic response of land use and river nutrient concentration to long-term climatic changes. Sci. Total Environ. 590, 818-831.DOI: 10.1016/j.scitotenv.2017.03.069.

Caceres, T., Ying, G.G., Kookana, R., 2002. Sorption of pesticides used in banana production on soils of Ecuador. Aust. J. Soil Res. 40(7), 1085-1094.DOI: 10.1071/Sr02015.

Cairns, J., Dickson, K.L., 1971. Simple Method for Biological Assessment of Effects of Waste Discharges on Aquatic Bottom-Dwelling Organisms. J. Water Pollut. Control Fed. 43(5), 755-772.

Camargo, J.A., Alonso, A., Salamanca, A., 2005. Nitrate toxicity to aquatic animals: a review with new data for freshwater invertebrates. Chemosphere 58(9), 1255-1267.DOI: 10.1016/j.chemosphere.2004.10.044.

Capítulo, A.R., Tangorra, M., Ocón, C., 2001. Use of benthic macroinvertebrates to assess the biological status of Pampean streams in Argentina. Aquat. Ecol. 35, 109-119.

Carlisle, D.M., Hawkins, C.P., 2008. Land use and the structure of western US stream invertebrate assemblages: predictive models and ecological traits. J. N. Am. Benthol. Soc. 27(4), 986-999.DOI: 10.1899/07-176.1.

Carlisle, D.M., Meador, M.R., 2007. A biological assessment of streams in the eastern united states using a predictive model for macroinvertebrate assemblages. J. Am. Water Resour. Assoc. 43(5), 1194-1207.DOI: 10.1111/j.1752-1688.2007.00097.x.

Carr, G.M., Neary, J.P., 2008. Water quality for ecosystem and human health, 2nd edition. United Nations Environment Programme (UNEP) Global Environment Monitoring System (GEMS) / Water Programme, Ontario, Canada.

Carvalho, L., Cortes, R., Bordalo, A.A., 2011. Evaluation of the ecological status of an impaired watershed by using a multi-index approach. Environ. Monit. Assess. 174(1-4), 493-508.DOI: 10.1007/s10661-010-1473-9.

Castillo, L.E., Ruepert, C., Solis, E., 2000. Pesticide residues in the aquatic environment of banana plantation areas in the north Atlantic zone of Costa Rica. Environ. Toxicol. Chem. 19(8), 1942-1950.DOI: 10.1897/1551-5028(2000)019<1942:Pritae>2.3.Co;2.

CELEC, 2013. Revista 25 años de la presa Daule-Peripa. CELEC EP-Hidronacion, Guayaquil, Ecuador. https://www.celec.gob.ec. Accessed 31 July 2015.

Cereghino, R., Giraudel, J.L., Compin, A., 2001. Spatial analysis of stream invertebrates distribution in the Adour-Garonne drainage basin (France), using Kohonen self organizing maps. Ecol. Model. 146(1-3), 167-180.DOI: 10.1016/S0304-3800(01)00304-0.

Chandler, J.R., 1970. A biological approach to water quality management. Water Pollut. Control 69, 415-422.

Chapman, D., 1996. Water quality assessments - A guide to use of biota, sediments and water in environmental monitoring, 2nd ed. E&FN Spon on behalf of UNESCO/WHO/UNEP, London, UK.

Chaudhary, A., Pourfaraj, V., Mooers, A.O., 2018. Projecting global land use-driven evolutionary history loss. Divers. Distrib. 24(2), 158-167.DOI: 10.1111/ddi.12677.

Clapcott, J.E., Goodwin, E.O., Snelder, T.H., Collier, K.J., Neale, M.W., Greenfield, S., 2017. Finding reference: a comparison of modelling approaches for predicting macroinvertebrate community index benchmarks. N. Z. J. Mar. Freshw. Res. 51(1), 44-59.DOI: 10.1080/00288330.2016.1265994.

Colin, N., Maceda-Veiga, A., Flor-Arnau, N., Mora, J., Fortuno, P., Vieira, C., Prat, N., Cambra, J., de Sostoa, A., 2016. Ecological impact and recovery of a Mediterranean river after receiving the effluent from a textile dyeing industry. Ecotoxicol. Environ. Saf. 132, 295-303.DOI: 10.1016/j.ecoenv.2016.06.017.

Collins, A.L., Walling, D.E., 2007. Fine-grained bed sediment storage within the main channel systems of the Frome and Piddle catchments, Dorset, UK. Hydrolog. Process. 21(11), 1448-1459.DOI: 10.1002/hyp.6269.

Compin, A., Cereghino, R., 2007. Spatial patterns of macroinvertebrate functional feeding groups in streams in relation to physical variables and land-cover in Southwestern France. Landsc. Ecol. 22(8), 1215-1225.DOI: 10.1007/s10980-007-9101-y.

Cortes, R.M.V., Hughes, S.J., Pereira, V.R., Varandas, S.D.P., 2013. Tools for bioindicator assessment in rivers: The importance of spatial scale, land use patterns and biotic integration. Ecol. Indic. 34, 460-477.DOI: 10.1016/j.ecolind.2013.06.004.

Courtonne, J.Y., Longaretti, P.Y., Alapetite, J., Dupre, D., 2016. Environmental Pressures Embodied in the French Cereals Supply Chain. J. Ind. Ecol. 20(3), 423-434.DOI: 10.1111/jiec.12431.

Crawley, M.J., 2007. The R book. Wiley, Chichester, England.

Crétaz, A.L.d.l., Barten, P.K., 2007. Land use effects on streamflow and water quality in the northeastern United States. CRC, Boca Raton, Fla.

Cunha, E.J., Montag, L.F.D., Juen, L., 2015. Oil palm crops effects on environmental integrity of Amazonian streams and Heteropteran (Hemiptera) species diversity. Ecol. Indic. 52, 422-429.DOI: 10.1016/j.ecolind.2014.12.024.

da Silva, M.V.D., Rosa, B.F.J.V., Alves, R.G., 2015. Effect of mesohabitats on responses of invertebrate community structure in streams under different land uses. Environ. Monit. Assess. 187(11).DOI: Artn 714 10.1007/S10661-015-4926-3.

Dahm, V., Hering, D., 2016. A modeling approach for identifying recolonisation source sites in river restoration planning. Landsc. Ecol. 31(10), 2323-2342.DOI: 10.1007/s10980-016-0402-x.

Dalgaard, P., 2008. Introductory statistics with R, 2nd ed. Springer, New York, USA.

Damanik-Ambarita, M.N., Everaert, G., Forio, M.A.E., Nguyen, T.H.T., Lock, K., Musonge, P.L.S., Suhareva, N., Dominguez-Granda, L., Bennetsen, E., Boets, P., Goethals, P.L.M., 2016a. Generalized Linear Models to Identify Key Hydromorphological and Chemical Variables Determining the Occurrence of Macroinvertebrates in the Guayas River Basin (Ecuador). Water 8(7).DOI: Artn 297 10.3390/W8070297.

Damanik-Ambarita, M.N., Lock, K., Boets, P., Everaert, G., Nguyen, T.H.T., Forio, M.A.E., Musonge, P.L.S., Suhareva, N., Bennetsen, E., Landuyt, D., Dominguez-Granda, L., Goethals, P.L.M., 2016b. Ecological water quality analysis of the Guayas river basin (Ecuador) based on macroinvertebrates indices. Limnologica.DOI: 10.1016/j.limno.2016.01.001.

Davies, S.P., Jackson, S.K., 2006. The biological condition gradient: A descriptive model for interpreting change in aquatic ecosystems. Ecol. Appl. 16(4), 1251-1266.DOI: 10.1890/1051-0761(2006)016[1251:Tbcgad]2.0.Co;2.

de Morais, L., de Oliveira Sanches, B., Sanches, B.D.O., Kaufmann, P.R., Hughes, R.M., Molozzi, J., Callisto, M., 2017. Assessment of disturbance at three spatial scales in two large tropical reservoirs. J. Limnol. 76(2), 240-252.DOI: 10.4081/jlimnol.2016.1547.

De Pauw, N., Gabriels, W., Goethals, P., 2006. River monitoring and assessment methods based on macroinvertebrates, in: Ziglio G., Siligardi M., Flaim G. (Eds.), Biological monitoring of rivers: applications and perspectives. John Wiley & Sons Ltd, West Sussex, England.

De Pauw, N., Van Damme, D., Bij De Vaate, A., 1996. Manual for macroinvertebrate identification and water quality assesment, Integrated programme for implementation of the recommended transnational monitoring strategy for the Danube river basin, CEC PHARE/TACIS project. University of Ghent, Ghent, Belgium.

De Pauw, N., Vanhooren, G., 1983. Method for Biological Quality Assessment of Watercourses in Belgium. Hydrobiologia 100, 153-168.

Dedecker, A.P., Goethals, P.L.M., D'Heygere, T., Gevrey, M., Lek, S., De Pauw, N., 2005. Application of artificial neural network models to analyse the relationships between Gammarus pulex L. (Crustacea, Amphipoda) and river characteristics. Environ. Monit. Assess. 111(1-3), 223-241.DOI: 10.1007/s10661-005-8221-6.

Deknock, A. (2017). Occurrence and distribution of agricultural pesticides in the freshwater environment of the Guayas river basin (Ecuador). Master of Science thesis, Ghent University, Ghent, Belgium.

Dominguez-Granda, L. (2007). Indices based on macroinvertebrate communities for assessment of the quality of the Chaguana river in Ecuador. PhD Thesis, Ghent University, Ghent, Belgium.

Dominguez-Granda, L., Lock, K., Goethals, P.L.M., 2011a. Application of classification trees to determine biological and chemical indicators for river assessment: case study in the Chaguana watershed (Ecuador). J. Hydroinform. 13(3), 489-499.DOI: 10.2166/hydro.2010.082.

Dominguez-Granda, L., Lock, K., Goethals, P.L.M., 2011b. Using multi-target clustering trees as a tool to predict biological water quality indices based on benthic macroinvertebrates and environmental parameters in the Chaguana watershed (Ecuador). Ecol. Inform. 6(5), 303-308.DOI: 10.1016/j.ecoinf.2011.05.004.

Domínguez, E., Fernández, H.R., 2009. Macroinvertebrados bentónicos sudamericanos : sistemática y biología. Fundación Miguel Lillo, Tucumán, Argentina.

Donner, A., 1982. The Relative Effectiveness of Procedures Commonly Used in Multiple-Regression Analysis for Dealing with Missing Values. Am. Stat. 36(4), 378-381.DOI: 10.2307/2683092.

Dudgeon, D., Arthington, A.H., Gessner, M.O., Kawabata, Z.I., Knowler, D.J., Leveque, C., Naiman, R.J., Prieur-Richard, A.H., Soto, D., Stiassny, M.L.J., Sullivan, C.A., 2006. Freshwater biodiversity: importance, threats, status and

conservation challenges. Biol. Rev. 81(2), 163-182.DOI: 10.1017/S1464793105006950.
Einheuser, M.D., Nejadhashemi, A.P., Sowa, S.P., Wang, L.Z., Hamaamin, Y.A., Woznicki, S.A., 2012. Modeling the effects of conservation practices on stream health. Sci. Total Environ. 435, 380-391.DOI: 10.1016/j.scitotenv.2012.07.033.
Ellison, C.A., Skinner, Q.D., Hicks, L.S., 2009. Assessment of Best-Management Practice Effects on Rangeland Stream Water Quality Using Multivariate Statistical Techniques. Rangel. Ecol. Manag. 62(4), 371-386.DOI: 10.2111/08-026.1.
Englert, D., Zubrod, J.P., Schulz, R., Bundschuh, M., 2015. Variability in ecosystem structure and functioning in a low order stream: Implications of land use and season. Sci. Total Environ. 538, 341-349.DOI: 10.1016/j.scitotenv.2015.08.058.
Epele, L.B., Miserendino, M.L., 2015. Environmental Quality and Aquatic Invertebrate Metrics Relationships at Patagonian Wetlands Subjected to Livestock Grazing Pressures. Plos One 10(10).DOI: ARTN e0137873 10.1371/journal.pone.0137873.
Erb, K.H., Haberl, H., Jepsen, M.R., Kuemmerle, T., Lindner, M., Muller, D., Verburg, P.H., Reenberg, A., 2013. A conceptual framework for analysing and measuring land-use intensity. Curr. Opin. Environ. Sustain. 5(5), 464-470.DOI: 10.1016/j.cosust.2013.07.010.
Erba, S., Pace, G., Demartini, D., Di Pasquale, D., Dorflinger, G., Buffagni, A., 2015. Land use at the reach scale as a major determinant for benthic invertebrate community in Mediterranean rivers of Cyprus. Ecol. Indic. 48, 477-491.DOI: 10.1016/j.ecolind.2014.09.010.
EU, 1998. EC - Drinking Water Directive - DWD - 98/83/EC. http://www.wise-rtd.info/en/info/ec-drinking-water-directive-dwd-9883ec. Accessed 7 July 2015.
European Commission, 2000. Directive 2000/60/EC of the European Parliement and of the Council establishing a framework for Community action in the field of water policy, OJ L 327.
Everaert, G., De Neve, J., Boets, P., Dominguez-Granda, L., Mereta, S.T., Ambelu, A., Hoang, T.H., Goethals, P.L.M., Thas, O., 2014. Comparison of the Abiotic Preferences of Macroinvertebrates in Tropical River Basins. Plos One 9(10).DOI: ARTN e108898 10.1371/journal.pone.0108898.
Everaert, G., Pauwels, I.S., Boets, P., Verduin, E., de la Haye, M.A.A., Blom, C., Goethals, P.L.M., 2013. Model-based evaluation of ecological bank design and management in the scope of the European Water Framework Directive. Ecol. Eng. 53, 144-152.DOI: 10.1016/j.ecoleng.2012.12.034.
Everaert, G., Pauwels, I.S., Goethals, P.L.M., 2010. Development of data-driven models for the assessment of macroinvertebrates in rivers in Flanders. Int. Congr. Environ. Model. Softw. Soc. (iEMSs).
FAO, 2011. The state of the world's land and water resources for food and agriculture (SOLAW) - Managing systems at risk. http://www.fao.org/nr/solaw/the-book/en/. Accessed 4 August 2014.
FAOSTAT, 2017. FAOSTAT database on land use. Food and Agriculture Organization of the United Nations - FAO. http://www.fao.org/faostat/en/#data/EL/visualize. Accessed 11 October 2017.

Feio, M.J., Norris, R.H., Graca, M.A.S., Nichols, S., 2009. Water quality assessment of Portuguese streams: Regional or national predictive models? Ecol. Indic. 9(4), 791-806.DOI: 10.1016/j.ecolind.2008.09.012.

Feio, M.J., Reynoldson, T.B., Ferreira, V., Graca, M.A.S., 2007. A predictive model for freshwater bioassessment (Mondego River, Portugal). Hydrobiologia 589, 55-68.DOI: 10.1007/s10750-006-0720-0.

Ferreira, A.R.L., Fernandes, L.F.S., Cortes, R.M.V., Pacheco, F.A.L., 2017. Assessing anthropogenic impacts on riverine ecosystems using nested partial least squares regression. Sci. Total Environ. 583, 466-477.DOI: 10.1016/j.scitotenv.2017.01.106.

Ferreira, C.S.S., Walsh, R.P.D., Costa, M.D., Coelho, C.O.A., Ferreira, A.J.D., 2016. Dynamics of surface water quality driven by distinct urbanization patterns and storms in a Portuguese peri-urban catchment. J. Soil. Sediment. 16(11), 2606-2621.DOI: 10.1007/s11368-016-1423-4.

Ferreira, W.R., Paiva, L.T., Callisto, M., 2011. Development of a benthic multimetric index for biomonitoring of a neotropical watershed. Braz. J. Biol. 71(1), 15-25.

Fierro, P., Bertran, C., Mercado, M., Pena-Cortes, F., Tapia, J., Hauenstein, E., Caputo, L., Vargas-Chacoff, L., 2015. Landscape composition as a determinant of diversity and functional feeding groups of aquatic macroinvertebrates in southern rivers of the Araucania, Chile. Lat. Am. J. Aquat. Res. 43(1), 186-200.DOI: 10.3856/vol43-issue1-fulltext-16.

Fisher, P., Comber, A.J., Wadsworth, R., 2005. Land Use and Land Cover: Contradiction or Complement, in: Fisher P., Unwin D.J. (Eds.), Re-presenting GIS. John Wiley & Sons, Ltd, Chichester, England, UK.

Flood, D., 2000. Ecuador 2000 Methodology, Antecedents and brief methodologic description of III the Farming National Census. http://www.fao.org/fileadmin/templates/ess/ess_test_folder/World_Census_Agriculture/Country_info_2000/Methodology/ECU_ENG_MET1_2000.pdf. Accessed 4 August 2014.

Flügel, E., 2004. Microfacies of carbonate rocks : analysis, interpretation and application. Springer, Berlin ; New York.

Forero-Cespedes, A.M., Reinoso-Florez, G., Gutierrez, C., 2013. Water quality assessment of the Opia River (Tolima-Colombia), using macroinvertebrates and physicochemical parameters. Caldasia 35(2), 371-387.

Forio, M.A.E., Landuyt, D., Bennetsen, E., Lock, K., Nguyen, T.H.T., Damanik-Ambarita, M.N., Musonge, P.L.S., Boets, P., Everaert, G., Dominguez-Granda, L., Goethals, P.L.M., 2015. Bayesian belief network models to analyse and predict ecological water quality in rivers. Ecol. Modell. 312, 222-238.DOI: 10.1016/j.ecolmodel.2015.05.025.

Forio, M.A.E., Mouton, A., Lock, K., Boets, P., Tien, N.T.H., Damanik-Ambarita, M.N., Musonge, P.L.S., Dominguez-Granda, L., Goethals, P.L.M., 2017. Fuzzy modelling to identify key drivers of ecological water quality to support decision and policy making. Environ. Sci. Policy 67, 58-68.DOI: 10.1016/j.envsci.2016.12.004.

Fornaroli, R., Cabrini, R., Sartori, L., Marazzi, F., Vracevic, D., Mezzanotte, V., Annala, M., Canobbio, S., 2015. Predicting the constraint effect of environmental characteristics on macroinvertebrate density and diversity using quantile regression mixed model. Hydrobiologia 742(1), 153-167.DOI: 10.1007/s10750-014-1974-6.

Frankforter, J.D., Weyers, H.S., Bales, J.D., Moran, P.W., Calhoun, D.L., 2010. The relative influence of nutrients and habitat on stream metabolism in agricultural streams. Environ. Monit. Assess. 168(1-4), 461-479.DOI: 10.1007/s10661-009-1127-y.
Frappart, F., Bourrel, L., Brodu, N., Salazar, X.R., Baup, F., Darrozes, J., Pombosa, R., 2017. Monitoring of the Spatio-Temporal Dynamics of the Floods in the Guayas Watershed (Ecuadorian Pacific Coast) Using Global Monitoring ENVISAT ASAR Images and Rainfall Data. Water 9(1).DOI: Artn 12 10.3390/W9010012.
Futter, M.N., Hogbom, L., Valinia, S., Sponseller, R.A., Laudon, H., 2016. Conceptualizing and communicating management effects on forest water quality. Ambio 45, S188-S202.DOI: 10.1007/s13280-015-0753-6.
Gabriels, W., Lock, K., De Pauw, N., Goethals, P.L.M., 2010. Multimetric Macroinvertebrate Index Flanders (MMIF) for biological assessment of rivers and lakes in Flanders (Belgium). Limnologica 40(3), 199-207.DOI: 10.1016/j.limno.2009.10.001.
Garcia, E.A., Pettit, N.E., Warfe, D.M., Davies, P.M., Kyne, P.M., Novak, P., Douglas, M.M., 2015. Temporal variation in benthic primary production in streams of the Australian wet-dry tropics. Hydrobiologia 760(1), 43-55.DOI: 10.1007/s10750-015-2301-6.
Garnier, J., Brion, N., Callens, J., Passy, P., Deligne, C., Billen, G., Servais, P., Billen, C., 2013. Modeling historical changes in nutrient delivery and water quality of the Zenne River (1790s-2010): The role of land use, waterscape and urban wastewater management. J. Mar. Syst. 128, 62-76.DOI: 10.1016/j.jmarsys.2012.04.001.
Gerebizza, E., 2009. The Daule Peripa project: Italy's responsibilities in Ecuador's illegitimate debt. Campagna per la Riforma della Banca Mondiale - CRBM, Rome, Italy. http://eurodad.org/uploadedfiles/whats_new/reports/the%20daule%20peripa%20project_crbm.pdf. Accessed 23 November 2016.
Gerth, W.J., Li, J., Giannico, G.R., 2017. Agricultural land use and macroinvertebrate assemblages in lowland temporary streams of the Willamette Valley, Oregon, USA. Agric. Ecosyst. Environ. 236, 154-165.DOI: 10.1016/j.agee.2016.11.010.
Goethals, P., De Pauw, N., 2001. Development of a concept for integrated ecological river assessment in Flanders, Belgium. J. Limnol. 60(1), 7-16.DOI: 10.4081/jlimnol.2001.s1.7.
Goethals, P.L.M., 2013. Sustainability of water quality and ecology: Easier said than defined and implemented. Sustain. Water Qual. Ecol. 1-2, 1-2.
Goethals, P.L.M., Dedecker, A.P., Gabriels, W., Lek, S., De Pauw, N., 2007. Applications of artificial neural networks predicting macroinvertebrates in freshwaters. Aquat. Ecol. 41(3), 491-508.DOI: 10.1007/s10452-007-9093-3.
Goldschmidt, T., 2016. Water mites (Acari, Hydrachnidia): powerful but widely neglected bioindocators - a review. Neotrop. Biodivers. 2(1), 12-25.DOI: 10.1080/23766808.2016.1144359.
Goss, C.W., Goebel, P.C., Sullivan, S.M.P., 2014. Shifts in attributes along agriculture-forest transitions of two streams in central Ohio, USA. Agric. Ecosyst. Environ. 197, 106-117.DOI: 10.1016/j.agee.2014.07.026.
Govenor, H., Krometis, L.A.H., Hession, W.C., 2017. Invertebrate-Based Water Quality Impairments and Associated Stressors Identified through the US

Clean Water Act. Environ. Manag. 60(4), 598-614.DOI: 10.1007/s00267-017-0907-3.
Greenacre, M., Primicerio, R., 2013. Multivariate Analysis of Ecological Data. Fundación BBVA, Bilbao, Spain.
Greenwood, M.J., Booker, D.J., 2015. The influence of antecedent floods on aquatic invertebrate diversity, abundance and community composition. Ecohydrol. 8(2), 188-203.DOI: 10.1002/eco.1499.
Grubaugh, J.W., Wallace, J.B., Houston, E.S., 1996. Longitudinal changes of macroinvertebrate communities along an Appalachian stream continuum. Can. J. Fish. Aquat. Sci. 53(4), 896-909.DOI: 10.1139/cjfas-53-4-896.
Guhathakurta, S., 2005. Telling Stories with Models: Reflecting on Land Use and Ecological Trends in the San Pedro Watershed, in: Fisher P., Unwin D.J. (Eds.), Re-presenting GIS. John Wiley & Sons, Ltd, Chichester, England, UK.
Guisan, A., Lehmann, A., Ferrier, S., Austin, M., Overton, J.M.C., Aspinall, R., Hastie, T., 2006. Making better biogeographical predictions of species' distributions. J. Appl. Ecol. 43(3), 386-392.DOI: 10.1111/j.1365-2664.2006.01164.x.
Guisan, A., Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. Ecol. Model. 135(2-3), 147-186.DOI: 10.1016/S0304-3800(00)00354-9.
Guneralp, B., Lwasa, S., Masundire, H., Parnell, S., Seto, K.C., 2018. Urbanization in Africa: challenges and opportunities for conservation. Environ. Res. Lett. 13(1).DOI: Artn 015002 10.1088/1748-9326/Aa94fe.
Guo, C.B., Park, Y.S., Liu, Y., Lek, S., 2015. Toward a new generation of ecological modelling techniques: Review and bibliometrics. Advanced Modelling Techniques Studying Global Changes in Environmental Sciences 27, 11-44.DOI: 10.1016/B978-0-444-63536-5.00002-8.
Guse, B., Kail, J., Radinger, J., Schroder, M., Kiesel, J., Hering, D., Wolter, C., Fohrer, N., 2015. Eco-hydrologic model cascades: Simulating land use and climate change impacts on hydrology, hydraulics and habitats for fish and macroinvertebrates. Sci. Total Environ. 533, 542-556.DOI: 10.1016/j.scitotenv.2015.05.078.
Han, J.H., Paek, W.K., An, K.G., 2016. Efficiency comparisons of fish sampling gears for a lentic ecosystem health assessments in Korea. J. Asia-Pac. Biodivers.DOI: 10.1016/j.japb.2016.10.001.
Hansen, B., Reich, P., Lake, P.S., Cavagnaro, T. (2010) Minimum width requirements for riparian zones to protect flowing waters and to conserve biodiversity: a review and recommendations - With application to the State of Victoria *Report to the Office of Water, Department of Sustainability and Environment*: Monash University.
Harding, J.S., Benfield, E.F., Bolstad, P.V., Helfman, G.S., Jones, E.B.D., 1998. Stream biodiversity: The ghost of land use past. Proc. Natl. Acad. Sci. U. S. A. 95(25), 14843-14847.DOI: 10.1073/pnas.95.25.14843.
Hawkes, H.A., 1998. Origin and development of the Biological Monitoring Working Party score system. Water Res. 32(3), 964-968.DOI: 10.1016/S0043-1354(97)00275-3.
Hawkins, C.P., Norris, R.H., Hogue, J.N., Feminella, J.W., 2000. Development and evaluation of predictive models for measuring the biological integrity of streams. Ecol. Appl. 10(5), 1456-1477.DOI: 10.1890/1051-0761(2000)010[1456:Daeopm]2.0.Co;2.

Hawkins, C.P., Yuan, L.L., 2016. Multitaxon distribution models reveal severe alteration in the regional biodiversity of freshwater invertebrates. Freshw. Sci. 35(4), 1365-1376.DOI: 10.1086/688848.

Helsel, D.R., Hirsch, R.M., 1992. Statistical methods in water resources. Elsevier, Amsterdam ; New York.

Helson, J.E., Williams, D.D., 2013. Development of a macroinvertebrate multimetric index for the assessment of low-land streams in the neotropics. Ecol. Indic. 29, 167-178.DOI: 10.1016/j.ecolind.2012.12.030.

Hering, D., Borja, A., Carstensen, J., Carvalho, L., Elliott, M., Feld, C.K., Heiskanen, A.S., Johnson, R.K., Moe, J., Pont, D., Solheim, A.L., Van De Bund, W., 2010. The European Water Framework Directive at the age of 10: A critical review of the achievements with recommendations for the future. Sci. Total Environ. 408(19), 4007-4019.DOI: 10.1016/j.scitotenv.2010.05.031.

Hering, D., Feld, C.K., Moog, O., 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.DOI: 10.1007/s10750-006-0087-2.

Hilsenhoff, W.L., 1987. An Improved Biotic Index of Organic Stream Pollution. Gt. Lakes Entomol. 20(1), 31-39.

Hilsenhoff, W.L., 1988. Rapid field assessment of organic pollution with a family-level biotic index. J. N. Am. Benthol. Soc. 7(1), 65-68.

Hilton, J., O'Hare, M., Bowes, M.J., Jones, J.I., 2006. How green is my river? A new paradigm of eutrophication in rivers. Sci. Total Environ. 365(1-3), 66-83.DOI: 10.1016/j.scitotenv.2006.02.055.

Hoang, T.H., Lock, K., Mouton, A., Goethals, P.L.M., 2010. Application of classification trees and support vector machines to model the presence of macroinvertebrates in rivers in Vietnam. Ecol. Inform. 5(2), 140-146.DOI: 10.1016/j.ecoinf.2009.12.001.

Holguin-Gonzalez, J.E., Everaert, G., Boets, P., Galvis, A., Goethals, P.L.M., 2013. Development and application of an integrated ecological modelling framework to analyze the impact of wastewater discharges on the ecological water quality of rivers. Environ. Model. Softw. 48, 27-36.DOI: 10.1016/j.envsoft.2013.06.004.

Holomuzki, J.R., Biggs, B.J.F., 2003. Sediment texture mediates high-flow effects on lotic macroinvertebrates. J. N. Am. Benthol. Soc. 22(4), 542-553.DOI: 10.2307/1468351.

Hook, S.E., Kroon, F.J., Metcalfe, S., Greenfield, P.A., Moncuquet, P., McGrath, A., Smith, R., Warne, M.S.J., Turner, R.D., McKeown, A., Westcott, D.A., 2017. Global transcriptomic profiling in barramundi (Lates calcarifer) from rivers impacted by differing agricultural land uses. Environ. Toxicol. Chem. 36(1), 103-112.DOI: 10.1002/etc.3505.

Horgan, F.G., Felix, M.I., Portalanza, D.E., Sanchez, L., Rios, W.M.M., Farah, S.E., Wither, J.A., Andrade, C.I., Espin, E.B., 2014. Responses by farmers to the apple snail invasion of Ecuador's rice fields and attitudes toward predatory snail kites. Crop Prot. 62, 135-143.DOI: 10.1016/j.cropro.2014.04.019.

Hou, D.K., He, J., Lu, C.W., Sun, Y., Zhang, F.J., Otgonbayar, K., 2013. Effects of Environmental Factors on Nutrients Release at Sediment-Water Interface and Assessment of Trophic Status for a Typical Shallow Lake, Northwest China. Sci. World J.DOI: Artn 716342 10.1155/2013/716342.

Hrodey, P.J., Sutton, T.M., Frimpong, E.A., Simon, T.P., 2009. Land-use Impacts on Watershed Health and Integrity in Indiana Warmwater Streams. Am. Midl. Nat. 161(1), 76-95.DOI: 10.1674/0003-0031-161.1.76.

Hruby, T., 2004. Washington State wetland rating system for eastern Washington - Revised. Washington State Department of Ecology Publication # 04-06-15.

Hubbart, J.A., Kellner, E., Kinder, P., Stephan, K., 2017. Challenges in aquatic physical habitat assessment: Improving conservation and restoration decisions for contemporary watersheds. Chall. 8(31), 11.DOI: 10.3390/challe8020031.

Hughes, S.J., Cabral, J.A., Bastos, R., Cortes, R., Vicente, J., Eitelberg, D., Yu, H.R., Honrado, J., Santos, M., 2016. A stochastic dynamic model to assess land use change scenarios on the ecological status of fluvial water bodies under the Water Framework Directive. Sci. Total Environ. 565, 427-439.DOI: 10.1016/j.scitotenv.2016.04.153.

Hughes, S.J., Santos, J.M., Ferreira, M.T., Caraca, R., Mendes, A.M., 2009. Ecological assessment of an intermittent Mediterranean river using community structure and function: evaluating the role of different organism groups. Freshw. Biol. 54(11), 2383-2400.DOI: 10.1111/j.1365-2427.2009.02253.x.

Hutchens, J.J., Schuldt, J.A., Richards, C., Johnson, L.B., Host, G.E., Breneman, D.H., 2009. Multi-scale mechanistic indicators of Midwestern USA stream macroinvertebrates. Ecol. Indic. 9(6), 1138-1150.DOI: 10.1016/j.ecolind.2009.01.001.

Ibrahim, J.G., Chen, M.H., Lipsitz, S.R., Herring, A.H., 2005. Missing-data methods for generalized linear models: A comparative review. J. Am. Stat. Assoc. 100(469), 332-346.DOI: 10.1198/016214504000001844.

Iniguez-Armijos, C., Leiva, A., Frede, H.G., Hampel, H., Breuer, L., 2014. Deforestation and Benthic Indicators: How Much Vegetation Cover Is Needed to Sustain Healthy Andean Streams? Plos One 9(8).DOI: ARTN e105869 10.1371/journal.pone.0105869.

Jackson, L.J., Trebitz, A.S., Cottingham, K.L., 2000. An introduction to the practice of ecological modeling. Biosci. 50(8), 694-706.DOI: 10.1641/0006-3568(2000)050[0694:Aittpo]2.0.Co;2.

Jacobsen, D., Encalada, A., 1998. The macroinvertebrate fauna of Ecuadorian highland streams in the wet and dry season. Arch. Hydrobiol. 142(1), 53-70.

Jacobsen, D., Schultz, R., Encalada, A., 1997. Structure and diversity of stream invertebrate assemblages: the influence of temperature with altitude and latitude. Freshw. Biol. 38(2), 247-261.DOI: 10.1046/j.1365-2427.1997.00210.x.

Jayawardana, J.M.C.K., Gunawardana, W.D.T.M., Udayakumara, E.P.N., Westbrooke, M., 2017. Land use impacts on river health of Uma Oya, Sri Lanka: implications of spatial scales. Environ. Monit. Assess. 189(4).DOI: Artn 192 10.1007/S10661-017-5863-0.

Jun, Y.C., Kim, N.Y., Kwon, S.J., Han, S.C., Hwang, I.C., Park, J.H., Won, D.H., Byun, M.S., Kong, H.Y., Lee, J.E., Hwang, S.J., 2011. Effects of land use on benthic macroinvertebrate communities: Comparison of two mountain streams in Korea. Ann. Limnol.-Int. J. Limnol. 47, S35-S49.DOI: 10.1051/limn/2011018.

Kairo, K., Timm, H., Haldna, M., Virro, T., 2012. Biological Quality on the Basis of Macroinvertebrates in Dammed Habitats of Some Estonian Streams, Central -

Baltic Europe. Int. Rev. Hydrobiol. 97(6), 497-508.DOI: 10.1002/iroh.201111530.
Kang, S.M., Seager, R., 2013. Croll revisited: Why is the Northern hemisphere warmer than the Southern hemisphere?
Karr, J.R., 1991. Biological Integrity - a Long-Neglected Aspect of Water-Resource Management. Ecol. Appl. 1(1), 66-84.DOI: 10.2307/1941848.
Kasangaki, A., Chapman, L.J., Balirwa, J., 2008a. Land use and the ecology of benthic macroinvertebrate assemblages of high-altitude rainforest streams in Uganda.
Kasangaki, A., Chapman, L.J., Balirwa, J., 2008b. Land use and the ecology of benthic macroinvertebrate assemblages of high-altitude rainforest streams in Uganda. Freshw. Biol. 53(4), 681-697.DOI: 10.1111/j.1365-2427.2007.01925.x.
Kehoe, L., Kuemmerle, T., Meyer, C., Levers, C., Vaclavik, T., Kreft, H., 2015. Global patterns of agricultural land-use intensity and vertebrate diversity. Divers. Distrib. 21(11), 1308-1318.DOI: 10.1111/ddi.12359.
Kidd, H., James, D.R., 1991. The Agrochemicals handbook, third ed. Royal Society of Chemistry, Information Services, Cambridge, England.
Kincheloe, J.W., Wedemeyer, G.A., Koch, D.L., 1979. Tolerance of Developing Salmonid Eggs and Fry to Nitrate Exposure. Bull. Environ. Contam. Toxicol. 23(4-5), 575-578.
Kindt, R., Coe, R., 2005. Tree diversity analysis : a manual and software for common statistical methods for ecological and biodiversity studies. World Agroforestry Centre, Nairobi, Kenya.
Knee, K.L., Encalada, A.C., 2014. Land Use and Water Quality in a Rural Cloud Forest Region (Intag, Ecuador). River Res. Appl. 30(3), 385-401.DOI: 10.1002/Rra.2634.
Kuemmerle, T., Erb, K., Meyfroidt, P. *et al.*, 2013. Challenges and opportunities in mapping land use intensity globally. Curr. Opin. Environ. Sustain. 5(5), 484-493.DOI: 10.1016/j.cosust.2013.06.002.
LADA (2008) Mapping Land Use Systems at global and regional scales for Land Degradation Aassessment Analysis. In Nachtergaele F., Petri M. (Eds.), *LADA Technical report n.8, version 1.1*.
Lanz, B., Dietz, S., Swanson, T., 2018. The Expansion of Modern Agriculture and Global Biodiversity Decline: An Integrated Assessment. Ecol. Econ. 144, 260-277.DOI: 10.1016/j.ecolecon.2017.07.018.
Larras, F., Coulaud, R., Gautreau, E., Billoir, E., Rosebery, J., Usseglio-Polatera, P., 2017. Assessing anthropogenic pressures on streams: A random forest approach based on benthic diatom communities. Sci. Total Environ. 586, 1101-1112.DOI: 10.1016/j.scitotenv.2017.02.096.
Lee, B.Y., Park, S.J., Paule, M.C., Jun, W., Lee, C.H., 2012. Effects of Impervious Cover on the Surface Water Quality and Aquatic Ecosystem of the Kyeongan Stream in South Korea. Water Environ. Res. 84(8), 635-645.DOI: 10.2175/106143012X13373550426878.
Lee, F., Simon, K.S., Perry, G.L.W., 2017. Increasing agricultural land use is associated with the spread of an invasive fish (Gambusia affinis). Sci. Total Environ. 586, 1113-1123.DOI: 10.1016/j.scitotenv.2017.02.101.
Leps, M., Tonkin, J.D., Dahm, V., Haase, P., Sundermann, A., 2015. Disentangling environmental drivers of benthic invertebrate assemblages: The role of spatial

scale and riverscape heterogeneity in a multiple stressor environment. Sci. Total Environ. 536, 546-556.DOI: 10.1016/j.scitotenv.2015.07.083.

Lester, R.E., Boulton, A.J., 2008. Rehabilitating agricultural streams in Australia with wood: A review. Environ. Manag. 42(2), 310-326.DOI: 10.1007/s00267-008-9151-1.

Li, L., Fassnacht, F.E., Storch, I., Burgi, M., 2017. Land-use regime shift triggered the recent degradation of alpine pastures in Nyanpo Yutse of the eastern Qinghai-Tibetan Plateau. Landsc. Ecol. 32(11), 2187-2203.DOI: 10.1007/s10980-017-0510-2.

Liu, H., Bu, H.M., Liu, G.H., Wang, Z.X., Liu, W.Z., 2015. Effects of surrounding land use on metal accumulation in environments and submerged plants in subtropical ponds. Environ. Sci. Pollut. Res. 22(23), 18750-18758.DOI: 10.1007/s11356-015-5067-5.

Lock, K., Goethals, P.L.M., 2013. Habitat suitability modelling for mayflies (Ephemeroptera) in Flanders (Belgium). Ecol. Inform. 17, 30-35.DOI: 10.1016/j.ecoinf.2011.12.004.

Lock, K., Goethals, P.L.M., 2014. Predicting the occurrence of stoneflies (Plecoptera) on the basis of water characteristics, river morphology and land use. J. Hydroinform. 16(4), 812-821.DOI: 10.2166/hydro.2013.188.

Lofgren, S., Grandin, U., Stendera, S., 2014. Long-term effects on nitrogen and benthic fauna of extreme weather events: Examples from two Swedish headwater streams. Ambio 43, 58-76.DOI: 10.1007/s13280-014-0562-3.

Lorenz, A., Hering, D., Feld, C.K., Rolauffs, P., 2004. A new method for assessing the impact of hydromorphological degradation on the macroinvertebrate fauna of five German stream types. Hydrobiologia 516(1), 107-127.DOI: 10.1023/B:Hydr.0000025261.79761.B3.

Lowrance, R., Altier, L.S., Newbold, J.D., Schnabel, R.R., Groffman, P.M., Denver, J.M., Correll, D.L., Gilliam, J.W., Robinson, J.L., Brinsfield, R.B., Staver, K.W., Lucas, W., Todd, A.H., 1997. Water quality functions of riparian forest buffers in Chesapeake Bay watersheds. Environ. Manag. 21(5), 687-712.DOI: 10.1007/s002679900060.

Malmqvist, B., Maki, M., 1994. Benthic Macroinvertebrate Assemblages in North Swedish Streams - Environmental Relationships. Ecography 17(1), 9-16.DOI: 10.1111/j.1600-0587.1994.tb00072.x.

Maloney, K.O., Weller, D.E., 2011. Anthropogenic disturbance and streams: land use and land-use change affect stream ecosystems via multiple pathways. Freshw. Biol. 56(3), 611-626.DOI: 10.1111/j.1365-2427.2010.02522.x.

Mandaville, S.M., 2002. Benthic Macroinvertebrates in Freshwaters- Taxa Tolerance Values, Metrics, and Protocols. Soil & Water Conservation Society of Metro Halifax (Project H-1). 128.

Manfrin, A., Bombi, P., Traversetti, L., Larsen, S., Scalici, M., 2016. A landscape-based predictive approach for running water quality assessment: A Mediterranean case study. J. Nat. Conserv. 30, 27-31.DOI: 10.1016/j.jnc.2016.01.002.

Mantyka-Pringle, C.S., Martin, T.G., Moffatt, D.B., Linke, S., Rhodes, J.R., 2014. Understanding and predicting the combined effects of climate change and land-use change on freshwater macroinvertebrates and fish. J. Appl. Ecol. 51(3), 572-581.DOI: 10.1111/1365-2664.12236.

Marshall, J.C., Steward, A.L., Harch, B.D., 2006. Taxonomic resolution and quantification of freshwater macroinvertebrate samples from an Australian

dryland river: The benefits and costs of using species abundance data. Hydrobiologia 572, 171-194.DOI: 10.1007/s10750-005-9007-0.
Matamoros, D. (2004). Predicting river concentrations of pesticides from banana plantations under data-poor conditions. PhD Thesis, Ghent University, Ghent, Belgium.
McDonald, R.I., Weber, K.F., Padowski, J., Boucher, T., Shemie, D., 2016a. Estimating watershed degradation over the last century and its impact on water-treatment costs for the world's large cities. Proc. Natl. Acad. Sci. U. S. A. 113(32), 9117-9122.DOI: 10.1073/pnas.1605354113.
McDonald, R.I., Weber, K.F., Padowski, J., Boucher, T., Shemie, D., 2016b. Estimating watershed degradation over the last century and its impact on water-treatment costs for the world's large cities. Proc. Natl Acad. Sci. U. S. A. 113(32), 9117-9122.DOI: 10.1073/pnas.1605354113.
Mereta, S.T., Boets, P., Bayih, A.A., Malu, A., Ephrem, Z., Sisay, A., Endale, H., Yitbarek, M., Jemal, A., De Meester, L., Goethals, P.L.M., 2012. Analysis of environmental factors determining the abundance and diversity of macroinvertebrate taxa in natural wetlands of Southwest Ethiopia. Ecol. Inform. 7(1), 52-61.DOI: 10.1016/j.ecoinf.2011.11.005.
Mereta, S.T., Boets, P., De Meester, L., Goethals, P.L.M., 2013. Development of a multimetric index based on benthic macroinvertebrates for the assessment of natural wetlands in Southwest Ethiopia. Ecol. Indic. 29, 510-521.DOI: 10.1016/j.ecolind.2013.01.026.
Merriam, E.R., Petty, J.T., Merovich, G.T., Fulton, J.B., Strager, M.P., 2011. Additive effects of mining and residential development on stream conditions in a central Appalachian watershed. J. N. Am. Benthol. Soc. 30(2), 399-418.DOI: 10.1899/10-079.1.
Meyer, M.D., Davis, C.A., Dvorett, D., 2015. Response of Wetland Invertebrate Communities to Local and Landscape Factors in North Central Oklahoma. Wetlands 35(3), 533-546.DOI: 10.1007/s13157-015-0642-6.
Milliman, J.D., Lee, T.Y., Huang, J.C., Kao, S.J., 2017. Impact of catastrophic events on small mountainous rivers: Temporal and spatial variations in suspendedand dissolved-solid fluxes along the Choshui River, central western Taiwan, during typhoon Mindulle, July 2-6, 2004. Geochim. Cosmochim. Acta 205, 272-294.DOI: 10.1016/j.gca.2017.02.015.
Ministerio de Agricultura Ganadería Acuacultura y Pesca - MAGAP, 2015. Mapas provinciales. Ministerio de Agricultura, Ganadería, Acuacultura y Pesca (MAGAP), Ecuador. http://geoportal.agricultura.gob.ec/mapas-provinciales. Accessed 25 April 2016.
Ministerio del Ambiente del Ecuador - MAE, 2015. Reforma del libro VI del texto unificado de legislación secundaria. No 061. Ministerio del Ambiente del Ecuador (MAE), Quito, Ecuador.
Molina, M.C., Roa-Fuentes, C.A., Zeni, J.O., Casatti, L., 2017. The effects of land use at different spatial scales on instream features in agricultural streams. Limnologica 65, 14-21.DOI: 10.1016/j.limno.2017.06.001.
Moreno, P., Franca, J.S., Ferreira, W.R., Paz, A.D., Monteiro, I.M., Callisto, M., 2009. Use of the BEAST model for biomonitoring water quality in a neotropical basin. Hydrobiologia 630(1), 231-242.DOI: 10.1007/s10750-009-9796-7.
Mouton, A.M., Dedecker, A.P., Lek, S., Goethals, P.L.M., 2010. Selecting Variables for Habitat Suitability of Asellus (Crustacea, Isopoda) by Applying Input

Variable Contribution Methods to Artificial Neural Network Models. Environ. Modell Assess. 15(1), 65-79.DOI: 10.1007/s10666-009-9192-8.
Mustow, S.E., 2002. Biological monitoring of rivers in Thailand: use and adaptation of the BMWP score. Hydrobiologia 479(1), 191-229.DOI: 10.1023/A:1021055926316.
Mwedzi, T., Bere, T., Mangadze, T., 2016. Macroinvertebrate assemblages in agricultural, mining, and urban tropical streams: implications for conservation and management. Environ. Sci. Pollut. Res. 23(11), 11181-11192.DOI: 10.1007/s11356-016-6340-y.
Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., Kent, J., 2000. Biodiversity hotspots for conservation priorities. Nature 403(6772), 853-858.DOI: 10.1038/35002501.
National Academy of Sciences, National Academy of Engineering (1972) Water Quality Criteria. United States Environmental Agency, Washington, D.C.
Nguyen, T.H.T. (2017). Ecological impact assessment of hydropower generation in river systems. PhD thesis, Ghent University, Ghent, Belgium.
Nguyen, T.H.T., Boets, P., Lock, K., Damanik-Ambarita, M.N., Forio, M.A.E., Sasha, P., Dominguez-Granda, L.E., Hoang, T.H.T., Everaert, G., Goethals, P.L.M., 2015. Habitat suitability of the invasive water hyacinth and its relation to water quality and macroinvertebrate diversity in a tropical reservoir. Limnologica 52, 67-74.DOI: 10.1016/j.limno.2015.03.006.
Nirmalakhandan, N., 2002. Modeling tools for environmental engineers and scientists. CRC Press, Boca Raton, Fla.
O'Toole, C., Donohue, I., Moe, S.J., Irvine, K., 2008. Nutrient optima and tolerances of benthic invertebrates, the effects of taxonomic resolution and testing of selected metrics in lakes using an extensive European data base. Aquat. Ecol. 42(2), 277-291.DOI: 10.1007/s10452-008-9185-8.
Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Henry, M., Stevens, H., Wagner, H., 2013. vegan: community ecology package, R package version 2.0-10. http://CRAN.R-project.org/package=vegan. Accessed 27 July 2014.
Oliveira, R.B.S., Baptista, D.F., Mugnai, R., Castro, C.M., Hughes, R.M., 2011. Towards rapid bioassessment of wadeable streams in Brazil: Development of the Guapiacu-Macau Multimetric Index (GMMI) based on benthic. Ecol. Indic. 11(6), 1584-1593.DOI: 10.1016/j.ecolind.2011.04.001.
Paillex, A., Reichert, P., Lorenz, A.W., Schuwirth, N., 2017. Mechanistic modelling for predicting the effects of restoration, invasion and pollution on benthic macroinvertebrate communities in rivers. Freshw. Biol. 62(6), 1083-1093.DOI: 10.1111/fwb.12927.
Paliy, O., Shankar, V., 2016. Application of multivariate statistical techniques in microbial ecology. Mol. Ecol. 25(5), 1032-1057.DOI: 10.1111/mec.13536.
Palmer, M.A., Hondula, K.L., Koch, B.J., 2014. Ecological Restoration of Streams and Rivers: Shifting Strategies and Shifting Goals. Annu. Rev. Ecol. Evol. Syst. 45, 247-+.DOI: 10.1146/annurev-ecolsys-120213-091935.
Palmer, M.A., Menninger, H.L., Bernhardt, E., 2010. River restoration, habitat heterogeneity and biodiversity: a failure of theory or practice? Freshw. Biol. 55, 205-222.DOI: 10.1111/j.1365-2427.2009.02372.x.
Park, S.R., Lee, H.J., Lee, S.W., Hwang, S.J., Byeon, M.S., Joo, G.J., Jeong, K.S., Kong, D.S., Kim, M.C., 2011. Relationships between land use and multi-

dimensional characteristics of streams and rivers at two different scales. Ann. Limnol.-Int. J. Limnol. 47, S107-S116.DOI: 10.1051/limn/2011023.
Parsons, M., Thoms, M., Norris, R., 2002. Australian River Assessment System: AusRivAS Physical Assessment Protocol, Monitoring River Health Initiative Technical Report no. 22, Commonwealth of Australia and University of Canberra, Canberra.
Paudel, B., Weston, N., O'Connor, J., Sutter, L., Velinsky, D., 2017. Phosphorus Dynamics in the Water Column and Sediments of Barnegat Bay, New Jersey. J. Coast. Res., 60-69.DOI: 10.2112/SI78-006.1.
Pearson, C.E., Ormerod, S.J., Symondson, W.O.C., Vaughan, I.P., 2016. Resolving large-scale pressures on species and ecosystems: propensity modelling identifies agricultural effects on streams. J. Appl. Ecol. 53(2), 408-417.DOI: 10.1111/1365-2664.12586.
Pietron, J., Chalov, S.R., Chalova, A.S., Alekseenko, A.V., Jarsjo, J., 2017. Extreme spatial variability in riverine sediment load inputs due to soil loss in surface mining areas of the Lake Baikal basin. Catena 152, 82-93.DOI: 10.1016/j.catena.2017.01.008.
Pilgrim, C.M., Mikhailova, E.A., Post, C.J., Hains, J.J., 2014. Spatial and temporal analysis of land cover changes and water quality in the Lake Issaqueena watershed, South Carolina. Environ. Monit. Assess. 186(11), 7617-7630.DOI: 10.1007/s10661-014-3953-9.
Poff, N.L., Bledsoe, B.P., Cuhaciyan, C.O., 2006. Hydrologic variation with land use across the contiguous United States: Geomorphic and ecological consequences for stream ecosystems. Geomorphol. 79(3-4), 264-285.DOI: 10.1016/j.geomorph.2006.06.032.
Poppe, M., Kail, J., Aroviita, J., Stelmaszczyk, M., Gielczewski, M., Muhar, S., 2016. Assessing restoration effects on hydromorphology in European mid-sized rivers by key hydromorphological parameters. Hydrobiologia 769(1), 21-40.DOI: 10.1007/s10750-015-2468-x.
Prati, L., Pavanell.R, Pesarin, F., 1971. Assessment of surface water quality by a single index of pollution. Water Res. 5(9), 741-&.DOI: 10.1016/0043-1354(71)90097-2.
R-Core-Team, 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/. Accessed 1 April 2014.
Raapysjarvi, J., Hamalainen, H., Aroviita, J., 2016. Macrophytes in boreal streams: Characterizing and predicting native occurrence and abundance to assess human impact. Ecol. Indic. 64, 309-318.DOI: 10.1016/j.ecolind.2016.01.014.
Raper, E., Davies, S., Perkins, B., Lamb, H., Hermanson, M., Soares, A., Stephenson, T., 2015. Ecological conditions of ponds situated on blast furnace slag deposits located in South Gare Site of Special Scientific Interest (SSSI), Teesside, UK. Environ. Geochem. Health 37(3), 545-556.DOI: 10.1007/s10653-014-9672-x.
Raven, P.J., Fox, P., Everard, M., Holmes, N.T.H., Dawson, F.H., 1997. River habitat survey: A new system for classifying rivers according to their habitat quality, in: Boon P.J., Howell D.L. (Eds.), Freshwater Quality: Defining the Indefinable? pp. 215-234.
Raven, P.J., Holmes, N.T.H., Dawson, F.H., Fox, P.J.A., Everard, M., Fozzard, I.R., Rouen, K.J., 1998. River Habitat Quality: the physical character of rivers and streams in the UK and Isle of Man, River Habitat Survey Report No. 2.

Raymond, K.L., Vondracek, B., 2011. Relationships among rotational and conventional grazing systems, stream channels, and macroinvertebrates. Hydrobiologia 669(1), 105-117.DOI: 10.1007/s10750-011-0653-0.

REDD, 2011. REDD in Ecuador. The REDD Desk. http://theredddesk.org/countries/ecuador. Accessed 24 November 2016.

Revelle, W., 2016. psych: Procedures for psychological, psychometric and personality research, Northwestern University, Evanston, Illinois, USA, R package version 1.6.9.

Rezende, R.S., Santos, A.M., Henke-Oliveira, C., Goncalves, J.F., 2014. Effects of spatial and environmental factors on benthic a macroinvertebrate community. Zoologia 31(5), 426-434.DOI: 10.1590/S1984-46702014005000001.

Rios-Touma, B., Acosta, R., Prat, N., 2014. The Andean Biotic Index (ABI): revised tolerance to pollution values for macroinvertebrate families and index performance evaluation. Rev. Biol. Trop. 62 (2), 249-273.

Rios-Touma, B., Encalada, A.C., Fornells, N.P., 2011. Macroinvertebrate Assemblages of an Andean High-Altitude Tropical Stream: The Importance of Season and Flow. Int. Rev. Hydrobiol. 96(6), 667-685.DOI: 10.1002/iroh.201111342.

Rios-Touma, B., Prescott, C., Axtell, S., Kondolf, G.M., 2015. Habitat Restoration in the Context of Watershed Prioritization: The Ecological Performance of Urban Stream Restoration Projects in Portland, Oregon. River Res. Appl. 31(6), 755-766.DOI: 10.1002/rra.2769.

Rios, S.L., Bailey, R.C., 2006. Relationship between riparian vegetation and stream benthic communities at three spatial scales. Hydrobiologia 553, 153-160.DOI: 10.1007/s10750-005-0868-z.

Robillard, C.M., Kerr, J.T., 2017. Assessing the shelf life of cost-efficient conservation plans for species at risk across gradients of agricultural land use. Conserv. Biol. 31(4), 837-847.DOI: 10.1111/cobi.12886.

Robinson, C.T., Schuwirth, N., Baumgartner, S., Stamm, C., 2014. Spatial relationships between land-use, habitat, water quality and lotic macroinvertebrates in two Swiss catchments. Aquat. Sci. 76(3), 375-392.DOI: 10.1007/s00027-014-0341-z.

Rocchini, D., Petras, V., Petrasova, A., Chemin, Y., Ricotta, C., Frigeri, A., Landa, M., Marcantonio, M., Bastin, L., Metz, M., Delucchi, L., Neteler, M., 2017a. Spatio-ecological complexity measures in GRASS GIS. Comput. Geosci. 104, 166-176.DOI: 10.1016/j.cageo.2016.05.006.

Rocchini, D., Petras, V., Petrasova, A., Horning, N., Furtkevicova, L., Neteler, M., Leutner, B., Wegmann, M., 2017b. Open data and open source for remote sensing training in ecology. Ecol. Inform. 40, 57-61.DOI: 10.1016/j.ecoinf.2017.05.004.

Rockstrom, J., Steffen, W., Noone, K. *et al.*, 2009. A safe operating space for humanity. Nature 461(7263), 472-475.DOI: 10.1038/461472a.

Roder, A., Propper, M., Stellmes, M., Schneibel, A., Hill, J., 2015. Assessing urban growth and rural land use transformations in a cross-border situation in Northern Namibia and Southern Angola. Land Use Policy 42, 340-354.DOI: 10.1016/j.landusepol.2014.08.008.

Roldán Pérez, G., 2003. Bioindicación de la calidad del agua en Colombia : propuesta para el uso del método BMWP/Col, 1. ed. Editorial Universidad de Antioquia, Medellín, Colombia.

Romero, R.D., Ceneviva-Bastos, M., Baviera, G.H., Casatti, L., 2013. Community structure of aquatic insects (Ephemeroptera, Plecoptera, and Trichoptera) in Cerrado streams of Paraguay, Parana, and Sao Francisco river basins. Biota Neotrop. 13(1), 97-107.

Rosenberg, D.M., Resh, V.H., 1993. Freshwater biomonitoring and benthic macroinvertebrates. Chapman & Hall, New York.

Roth, G.W., 2017. Crop rotations and conservation tillage. College of Agricultural Sciences, The Pennsylvania State University. https://extension.psu.edu/crop-rotations-and-conservation-tillage. Accessed 9 January 2018.

Sanchez, G.M., Nejadhashemi, A.P., Zhang, Z., Woznicki, S.A., Habron, G., Marquart-Pyatt, S., Shortridge, A., 2014. Development of a socio-ecological environmental justice model for watershed-based management. J. Hydrol. 518, 162-177.DOI: 10.1016/j.jhydrol.2013.08.014.

Scherr, S.J., McNeely, J.A., 2008. Biodiversity conservation and agricultural sustainability: towards a new paradigm of 'ecoagriculture' landscapes. Philos. Trans. R. Soc. B 363(1491), 477-494.DOI: 10.1098/rstb.2007.2165.

Schmalz, B., Kuemmerlen, M., Kiesel, J., Cai, Q., Jahnig, S.C., Fohrer, N., 2015. Impacts of land use changes on hydrological components and macroinvertebrate distributions in the Poyang lake area. Ecohydrol. 8(6), 1119-1136.DOI: 10.1002/eco.1569.

Schuwirth, N., Dietzel, A., Reichert, P., 2016. The importance of biotic interactions for the prediction of macroinvertebrate communities under multiple stressors. Funct. Ecol. 30(6), 974-984.DOI: 10.1111/1365-2435.12605.

Seo, S.N., McCarl, B.A., Mendelsohn, R., 2010. From beef cattle to sheep under global warming? An analysis of adaptation by livestock species choice in South America. Ecol. Econ. 69(12), 2486-2494.DOI: 10.1016/j.ecolecon.2010.07.025.

Sheldon, F., Peterson, E.E., Boone, E.L., Sippel, S., Bunn, S.E., Harch, B.D., 2012. Identifying the spatial scale of land use that most strongly influences overall river ecosystem health score. Ecol. Appl. 22(8), 2188-2203.

Shmueli, G., 2010. To Explain or to Predict? Stat. Sci. 25(3), 289-310.DOI: 10.1214/10-STS330.

Shrestha, M.K., Recknagel, F., Frizenschaf, J., Meyer, W., 2017. Future climate and land uses effects on flow and nutrient loads of a Mediterranean catchment in South Australia. Sci. Total Environ. 590, 186-193.DOI: 10.1016/j.scitotenv.2017.02.197.

Silva, D.M.L., Camargo, P.B., Mcdowell, W.H., Vieira, I., Salomao, M.S.M.B., Martinelli, L.A., 2012. Influence of land use changes on water chemistry in streams in the State of Sao Paulo, southeast Brazil. An. Acad. Bras. Cienc. 84(4), 919-930.

Slevers, M., Hale, R., Morrongiello, J.R., 2017. Do trout respond to riparian change? A meta-analysis with implications for restoration and management. Freshw. Biol. 62(3), 445-457.DOI: 10.1111/fwb.12888.

Smucker, N.J., Detenbeck, N.E., 2014. Meta-Analysis of Lost Ecosystem Attributes in Urban Streams and the Effectiveness of Out-of-Channel Management Practices. Restor. Ecol. 22(6), 741-748.DOI: 10.1111/rec.12134.

Solomatine, D.P., Ostfeld, A., 2008. Data-driven modelling: some past experiences and new approaches. J. Hydroinform. 10(1), 3-22.DOI: 10.2166/hydro.2008.015.

Strauch, A.M., MacKenzie, R.A., Giardina, C.P., Bruland, G.L., 2015. Climate driven changes to rainfall and streamflow patterns in a model tropical island hydrological system. J. Hydrol. 523, 160-169.DOI: 10.1016/j.jhydrol.2015.01.045.

Strayer, D.L., Beighley, R.E., Thompson, L.C., Brooks, S., Nilsson, C., Pinay, G., Naiman, R.J., 2003. Effects of land cover on stream ecosystems: Roles of empirical models and scaling issues. Ecosyst. 6(5), 407-423.DOI: 10.1007/s10021-002-0170-0.

Strehmel, A., Schmalz, B., Fohrer, N., 2016. Evaluation of Land Use, Land Management and Soil Conservation Strategies to Reduce Non-Point Source Pollution Loads in the Three Gorges Region, China. Environ. Manag. 58(5), 906-921.DOI: 10.1007/s00267-016-0758-3.

Sueyoshi, M., Ishiyama, N., Nakamura, F., 2016. beta-diversity decline of aquatic insects at the microhabitat scale associated with agricultural land use. Landsc. Ecol. Eng. 12(2), 187-196.DOI: 10.1007/s11355-015-0283-1.

Sundermann, A., Leps, M., Leisner, S., Haase, P., 2015. Taxon-specific physico-chemical change points for stream benthic invertebrates. Ecol. Indic. 57, 314-323.DOI: 10.1016/j.ecolind.2015.04.043.

Sweeney, B.W., Bott, T.L., Jackson, J.K., Kaplan, L.A., Newbold, J.D., Standley, L.J., Hession, W.C., Horwitz, R.J., 2004. Riparian deforestation, stream narrowing, and loss of stream ecosystem services. Proc. Natl. Acad. Sci. USA 101(39), 14132-14137.DOI: 10.1073/pnas.0405895101.

Tapia-Armijos, M.F., Homeier, J., Espinosa, C.I., Leuschner, C., de la Cruz, M., 2015. Deforestation and Forest Fragmentation in South Ecuador since the 1970s-Losing a Hotspot of Biodiversity. Plos One 10(9).DOI: ARTN e0133701 10.1371/journal.pone.0133701.

Tchakonte, S., Ajeagah, G.A., Camara, A.I., Diomande, D., Tchatcho, N.L.N., Ngassam, P., 2015. Impact of urbanization on aquatic insect assemblages in the coastal zone of Cameroon: the use of biotraits and indicator taxa to assess environmental pollution. Hydrobiol. 755(1), 123-144.DOI: 10.1007/s10750-015-2221-5.

Teillard, F., Anton, A., Dumont, B., Finn, J.A., Henry, B., Souza, D.M., Manzano, P., Mila i Canals, L., Phelps, C., Said, M., Vijn, S., White, S., 2016. A review of indicators and methods to assess biodiversity - Application to livestock production at global scale. Livestock Environment Assessment and Performance (LEAP) Partnership. FAO, Rome, Italy.

Terrado, M., Sabater, S., Chaplin-Kramer, B., Mandle, L., Ziv, G., Acuna, V., 2016. Model development for the assessment of terrestrial and aquatic habitat quality in conservation planning. Sci. Total Environ. 540, 63-70.DOI: 10.1016/j.scitotenv.2015.03.064.

The Biodiversity Group. (2016). Ecuadorian Biodiversity Project, https://biodiversitygroup.org/documenting-biodiversity-ecuador/, accessed 12 October 2017

Thornhill, I., Batty, L., Death, R.G., Friberg, N.R., Ledger, M.E., 2017. Local and landscape scale determinants of macroinvertebrate assemblages and their conservation value in ponds across an urban land-use gradient. Biodivers. Conserv. 26(5), 1065-1086.DOI: 10.1007/s10531-016-1286-4.

Thuiller, W., 2003. BIOMOD - optimizing predictions of species distributions and projecting potential future shifts under global change. Glob. Change Biol. 9(10), 1353-1362.DOI: 10.1046/j.1365-2486.2003.00666.x.

Townsend, C.R., Arbuckle, C.J., Crowl, T.A., Scarsbrook, M.R., 1997. The relationship between land use and physicochemistry, food resources and macroinvertebrate communities in tributaries of the Taieri River, New Zealand: A hierarchically scaled approach. Freshw. Biol. 37(1), 177-191.DOI: 10.1046/j.1365-2427.1997.00151.x.

Trimble, S.W., Mendel, A.C., 1995. The Cow as a Geomorphic Agent - a Critical Review. Geomorphol. 13(1-4), 233-253.DOI: 10.1016/0169-555x(95)00028-4.

Tu, J., 2009. Combined impact of climate and land use changes on streamflow and water quality in eastern Massachusetts, USA. J. Hydrol. 379(3-4), 268-283.DOI: 10.1016/j.jhydrol.2009.10.009.

Tuan, P.T., Dung, M.T., Duc, P.T., Trang, H.M., Khai, N.M., Thuy, P.T., 2016. Industrial water mass balance as a tool for water management in industrial parks. Water Resour. Ind. 13, 14-21.DOI: 10.1016/j.wri.2016.04.001.

Tuffery, S., 2011. Data mining and statistics for decision making. Wiley, Chichester, West Sussex ; Hoboken, NJ.

Turner, B.L., Meyer, W.B., Skole, D.L., 1994. Global Land-Use Land-Cover Change - Towards an Integrated Study. Ambio 23(1), 91-95.

Turunen, J., Muotka, T., Vuori, K.M., Karjalainen, S.M., Raapysjarvi, J., Sutela, T., Aroviita, J., 2016. Disentangling the responses of boreal stream assemblages to low stressor levels of diffuse pollution and altered channel morphology. Sci. Total Environ. 544, 954-962.DOI: 10.1016/j.scitotenv.2015.12.031.

United States Army Corps of Engineers - USACE, 1998. Water resources assessment of Ecuador. US Army Corps of Engineers (USACE). http://www.sam.usace.army.mil/Portals/46/docs/military/engineering/docs/WRA/Ecuador/Ecuador%20WRA%20English.pdf. Accessed 27 June 2014.

United States Environmental Protection Agency - USEPA, 2012. Stream flow. in: Water: monitoring & assessment. United States Environmental Protection Agency (USEPA), Washington D.C., United States. http://water.epa.gov/type/rsl/monitoring/vms51.cfm. Accessed 5 September 2014.

UNSD, 2017. Population by national and/or ethnic group, sex and urban/rural residence. United Nations Statistics Division (UNSD). http://data.un.org/Data.aspx?d=POP&f=tableCode:26. Accessed 25 April 2017.

USEPA, 1986. Quality Criteria for Water, EPA 440/5-86-001, United States Environmental Agency - USEPA, Washington, DC.

USEPA, 2002. Methods for Evaluating Wetland Condition: Developing Metrics and Indexes of Biological Integrity. Office of Water, U.S. Environmental Protection Agency, Washington, DC. EPA-822-R-02-016.

Usio, N., Nakagawa, M., Aoki, T., Higuchi, S., Kadono, Y., Akasaka, M., Takamura, N., 2017. Effects of land use on trophic states and multi-taxonomic diversity in Japanese farm ponds. Agric. Ecosyst. Environ. 247, 205-215.DOI: 10.1016/j.agee.2017.06.043.

Vaclavik, T., Lautenbach, S., Kuemmerle, T., Seppelt, R., 2013. Mapping global land system archetypes. Glob. Environ. Ch. 23(6), 1637-1647.DOI: 10.1016/j.gloenvcha.2013.09.004.

Van den Brink, P.J., Alexander, A.C., Desrosiers, M., Goedkoop, W., Goethals, P.L., Liess, M., Dyer, S.D., 2011. Traits-based approaches in bioassessment and ecological risk assessment: strengths, weaknesses, opportunities and threats. Integr. Environ. Assess. Manag. 7(2), 198-208.DOI: 10.1002/ieam.109.

van der Zanden, E.H., Levers, C., Verburg, P.H., Kuemmerle, T., 2016. Representing composition, spatial structure and management intensity of European agricultural landscapes: A new typology. Landsc. Urban Plan. 150, 36-49.DOI: 10.1016/j.landurbplan.2016.02.005.

Van Echelpoel, W., Boets, P., Landuyt, D., Gobeyn, S., Everaert, G., Bennetsen, E., Mouton, A., Goethals, P.L.M., 2015. Species distribution models for sustainable ecosystem management. Advanced Modelling Techniques Studying Global Changes in Environmental Sciences 27, 115-134.DOI: 10.1016/B978-0-444-63536-5.00008-9.

Van Sickle, J., Baker, J., Herlihy, A., Bayley, P., Gregory, S., Haggerty, P., Ashkenas, L., Li, J., 2004. Projecting the biological condition of streams under alternative scenarios of human land use. Ecol. Appl. 14(2), 368-380.DOI: 10.1890/02-5009.

Vannote, R.L., Minshall, G.W., Cummins, K.W., Sedell, J.R., Cushing, C.E., 1980. River Continuum Concept. Can. J. Fish. Aquat. Sci. 37(1), 130-137.DOI: 10.1139/F80-017.

Verissimo, H., Neto, J.M., Teixeira, H., Franco, J.N., Fath, B.D., Marques, J.C., Patricio, J., 2012. Ability of benthic indicators to assess ecological quality in estuaries following management. Ecol. Indic. 19, 130-143.DOI: 10.1016/j.ecolind.2011.06.014.

Verkaik, I., Vila-Escale, M., Rieradevall, M., Baxter, C.V., Lake, P.S., Minshall, G.W., Reich, P., Prat, N., 2015. Stream macroinvertebrate community responses to fire: are they the same in different fire-prone biogeographic regions? Freshw. Sci. 34(4), 1527-1541.DOI: 10.1086/683370.

Villamarin, C., Rieradevall, M., Paul, M.J., Barbour, M.T., Prat, N., 2013. A tool to assess the ecological condition of tropical high Andean streams in Ecuador and Peru: The IMEERA index. Ecol. Indic. 29, 79-92.DOI: 10.1016/j.ecolind.2012.12.006.

Von Sperling, M., Chernicharo, C.A.L., 2002. Urban wastewater treatment technologies and the implementation of discharge standards in developing countries. Urban Water 4, 105-114.

Vondracek, B., Blann, K.L., Cox, C.B., Nerbonne, J.F., Mumford, K.F., Nerbonne, B.A., Sovell, L.A., Zimmerman, J.K.H., 2005. Land use, spatial scale, and stream systems: Lessons from an agricultural region. Environ. Manag. 36(6), 775-791.DOI: 10.1007/s00267-005-0039-z.

Waite, I.R., 2014. Agricultural disturbance response models for invertebrate and algal metrics from streams at two spatial scales within the U.S. Hydrobiologia 726(1), 285-303.DOI: 10.1007/s10750-013-1774-4.

Walsh, C.J., Leonard, A.W., Ladson, A.R., Fletcher, T.D., 2004. Urban stormwater and the ecology of streams. Cooperative Research Centre for Freshwater Ecology and Cooperative Research Centre for Catchment Hydrology, Canberra, Australia.

Wang, R.Z., Xu, T.L., Yu, L.Z., Zhu, J.J., Li, X.Y., 2013. Effects of land use types on surface water quality across an anthropogenic disturbance gradient in the upper reach of the Hun River, Northeast China. Environ. Monit. Assess. 185(5), 4141-4151.DOI: 10.1007/s10661-012-2856-x.

Weigel, B.M., 2003. Development of stream macroinvertebrate models that predict watershed and local stressors in Wisconsin. J. N. Am. Benthol. Soc. 22(1), 123-142.DOI: 10.2307/1467982.

Weirich, S.R., Silverstein, J., Rajagopalan, B., 2011. Effect of average flow and capacity utilization on effluent water quality from US municipal wastewater treatment facilities. Water Res. 45(14), 4279-4286.DOI: 10.1016/j.watres.2011.06.002.

Wen, T., Sheng, S., An, S.Q., 2016. Relationships between stream ecosystem properties and landscape composition at multiple spatial scales along a heavily polluted stream in China: Implications for restoration. Ecol. Eng. 97, 493-502.DOI: 10.1016/j.ecoleng.2016.10.028.

Wilkins, P.M., Cao, Y., Heske, E.J., Levengood, J.M., 2015. Influence of a forest preserve on aquatic macroinvertebrates, habitat quality, and water quality in an urban stream. Urban Ecosyst. 18(3), 989-1006.DOI: 10.1007/s11252-015-0464-6.

Witten, I.H., Frank, E., 2005. Data mining : practical machine learning tools and techniques, 2nd ed. Morgan Kaufman, Amsterdam ; Boston, MA.

Woodiwiss, F.S., 1964. The biological system of stream classification used by the Trent River Board. Chem. Ind.(11), 443-447.

Woznicki, S.A., Nejadhashemi, A.P., Abouali, M., Herman, M.R., Esfahanian, E., Hamaamin, Y.A., Zhang, Z., 2016. Ecohydrological modeling for large-scale environmental impact assessment. Sci. Total Environ. 543, 274-286.DOI: 10.1016/j.scitotenv.2015.11.044.

Wright, R.F., Couture, R.M., Christiansen, A.B., Guerrero, J.L., Kaste, O., Barlaup, B.T., 2017. Effects of multiple stresses hydropower, acid deposition and climate change on water chemistry and salmon populations in the River Otra, Norway. Sci. Total Environ. 574, 128-138.DOI: 10.1016/j.scitotenv.2016.09.044.

Wyzga, B., Amirowicz, A., Oglecki, P., Hajdukiewicz, H., Radecki-Pawlik, A., Zawiejska, J., Mikus, P., 2014. Response of fish and benthic invertebrate communities to constrained channel conditions in a mountain river: Case study of the Biala, Polish Carpathians. Limnologica 46, 58-69.DOI: 10.1016/j.limno.2013.12.002.

Wyzga, B., Amirowicz, A., Radecki-Pawlik, A., Zawiejska, J., 2009. Hydromorphological Conditions, Potential Fish Habitats and the Fish Community in a Mountain River Subjected to Variable Human Impacts, the Czarny Dunajec, Polish Carpathians. River Res. Appl. 25(5), 517-536.DOI: 10.1002/rra.1237.

Yang, L., Bai, X., Hu, Y., 2017. Comparison between the linear model and k-nearest neighbor method for predicting macroinvertebrate assemblages in a city river in Beijing, China. Appl. Ecol. Environ. Res. 16(1), 387-406.DOI: 10.15666/aeer/1601_387406.

Yang, Y.Y., Toor, G.S., 2017. Sources and mechanisms of nitrate and orthophosphate transport in urban stormwater runoff from residential catchments. Water Res 112, 176-184.DOI: 10.1016/j.watres.2017.01.039.

Yates, A.G., Bailey, R.C., Schwindt, J.A., 2006. No-till cultivation improves stream ecosystem quality. J. Soil Water Conserv. 61(1), 14-19.

Younes-Baraille, Y., Garcia, X.F., Gagneur, J., 2005. Impact of the longitudinal and seasonal changes of the water quality on the benthic macroinvertebrate assemblages of the Andorran streams. C. R. Biol. 328(10-11), 963-976.DOI: 10.1016/j.crvi.2005.09.004.

Yun, Y.J., An, K.G., 2016. Roles of N:P Ratios on Trophic Structures and Ecological Stream Health in Lotic Ecosystems. Water 8(1).DOI: Artn 22 10.3390/W8010022.

Zhang, Y.X., Dudgeon, D., Cheng, D.S., Thoe, W., Fok, L., Wang, Z.Y., Lee, J.H.W., 2010. Impacts of land use and water quality on macroinvertebrate communities in the Pearl River drainage basin, China. Hydrobiologia 652(1), 71-88.DOI: 10.1007/s10750-010-0320-x.

Zuur, A.F., 2009. Mixed effects models and extensions in ecology with R. Springer, New York, NY.

Zuur, A.F., Ieno, E.N., Elphick, C.S., 2010. A protocol for data exploration to avoid common statistical problems. Methods Ecol. Evol. 1(1), 3-14.DOI: 10.1111/j.2041-210X.2009.00001.x.

Zuur, A.F., Ieno, E.N., Smith, G.M., 2007. Analysing ecological data. Springer, New York, USA.

Summary

High population growth especially since the 20th century has required extra provision of housing, water and food through agriculture and industry. Consequently, land use conversion from natural land such as forest to agricultural and urban cannot be avoided. Together with land use conversion, streams and rivers have been modified to support urban and agricultural development. This land use conversion from natural land to agricultural and urban effects ecological water quality and decreases ecosystem services (**chapter 1**). However, studies regarding land use effects on ecological water quality are still lacking in developing countries such as the Guayas river basin, Ecuador. This work stands as the starting point of an ecological water quality study where land use is integrated in the analyses. To do so, this PhD study aims to: (1) investigate why is land use information often not included in ecological water quality studies and investigate the use of ecological models to quantify the relationship between land use and the ecological water quality; (2) investigate current ecological water quality status of the Guayas river basin, Ecuador; (3) investigate key environmental variables affecting the ecological water quality; (4) investigate which type of land use data gathering that works best to quantify land use effect on the ecological water quality.

Based on reviewed scientific papers, land use information was often not included in ecological water quality studies because of the lack availability of land use information (**chapter 2**). To gain broad understanding of land use effect on the ecological water quality, an inclusion of land use information from local or riparian and catchment scales are required. Whenever possible, a combination of field observations and other sources in obtaining land use information is recommended. Furthermore, statistical analyses and models such as multivariate analyses, regression analyses and decision trees can be used to perform analyses in defining the relationship between land use and the ecological water quality.

As part of this research, an integrated sampling campaign was conducted at the end of the dry season of 2013. The sampling campaign was performed to collect biological (macroinvertebrate) and environmental (physico-chemical and hydromorphological) variables of 120 sampling sites at the Guayas river basin, Ecuador. In total, 39 environmental variables were collected.

To assess the current ecological water quality of the Guayas river basin, two biotic indices were calculated (**chapter 4**): the Biological Monitoring Working Party adapted for Colombia (BMWP-Col) and the Neotropical Low-land Stream Multimetric Index (NLSMI). The ecological water quality of the Guayas river basin ranged from very bad (0) to good (168) according to the BMWP-Col and from bad (0) to reference (9.1) according to the NLSMI. Nutrient concentrations were generally lower than the detection limits of nutrient-measuring kits, therefore nutrient concentrations could not be quantified at most of sampling sites). The results also suggested that the BMWP-Col is more suitable to assess the ecological water quality of the Guayas river basin than the NLSMI because the NLSMI is river-type-specific for small streams located at an elevation lower than 250 m above sea level.

The key environmental variables affecting the ecological water quality of the Guayas river basin were investigated in **chapter 4 and 5**. This was done by investigating the relationship between the presence of macroinvertebrate (BMWP-Col) and environmental variables using multivariate analyses (**chapter 4**) and regression analyses (**chapter 5**). Flow velocity, sludge layer, chlorophyll a concentration, sediment type, conductivity and land use showed strong influence on the distribution of macroinvertebrate taxa, based on multivariate analyses (all variables had $p < 0.001$). Whereas regression and sensitivity analyses selected a set of hydromorphological and chemical variables (elevation, nitrate-N and chlorophyll a concentrations, sediment angularity, presence of logs and macrophytes, flow velocity, turbidity, bank shape, and land use; $p < 0.05$ except for chlorophyll a had $p = 0.064$) as key environmental variables affecting the BMWP-Col. These results confirmed the influence of physico-chemical variables on macroinvertebrate presence and that agriculture-related variables and land use were the key environmental variables influencing the ecological water quality.

To assess which type of land use data gathering that works best to quantify land use effect on the ecological water quality, three methods and sources were used to collect local land use data: field protocols to assess land use within a stretch of 100×10 m, Google maps to assess land use for a stretch of 100×100 m, and GIS data to assess land use for a stretch of 200×200 m, all for the left and right banks of the sampling sites. Regression analyses were performed on each land use method

or source and environmental variables (**chapter 6**). The results suggested that the effect of local land use was best quantified using Google maps ($R^2 = 0.93$, $p < 0.05$). Moreover, models involving land use assessed using Google maps were associated mainly with physico-chemical variables, whereas models involving land use assessed using field protocols and GIS data were associated mainly with hydromorphological variables.

Samenvatting

De wereldpopulatiegroei vraagt extra voorziening van woningen, water en voedsel via landbouw en industrie. Conversie van natuurlijke gebieden (e.g. bossen) tot landbouwgronden en geürbaniseerde gebieden kan bijgevolg niet vermeden worden. Daarnaast worden waterstromen en rivieren structureel gemodificeerd ten voordele van verstedelijking en ontwikkeling van de landbouw. Deze modificaties hebben een impact op de ecologische waterkwaliteit wat op zich leidt tot een verminderd aanbod van ecosysteemdiensten (**hoofdstuk 1**). Onderzoek naar de invloed van landconversie op de ecologische waterkwaliteit in ontwikkelingslanden, is vrij beperkt en wordt in dit doctoraat bestudeerd voor het Guayas rivierbekken in Ecuador met als doelstellingen: (1) nagaan waarom landgebruik vaak niet wordt opgenomen in ecologische waterkwaliteitsstudies met een focus op het belang van ecologische modellen om de relatie tussen landgebruik en ecologische waterkwaliteit te kwantificeren; (2) de huidige ecologische waterkwaliteitsstatus van het Guayas rivierbekken bestuderen; (3) de sleutelvariabelen van ecologische waterkwaliteit bepalen; (4) onderzoeken welk manier van verzamelen van landgebruiksdata het best het effect van landgebruik op de ecologische waterkwaliteit beschrijft.

Omwille van een gebrek aan beschikbare data over landgebruik is het vaak niet mogelijk dit op te nemen in ecologische waterkwaliteitsstudies (**hoofdstuk 2**). Om ten volle het effect van landgebruik op de ecologische waterkwaliteit te begrijpen, is inclusie van landgebruiksdata op lokale en regionale schaal onontbeerlijk. Indien mogelijk wordt een combinatie van veld-observaties en andere bronnen aangeraden. Daarenboven kunnen statistische analyses en ecologische modellen zoals multivariate analyses, regressieanalyses en beslissingsbomen gebruikt worden in ecologische waterkwaliteitsstudies om de relatie tussen landgebruik en ecologische waterkwaliteit te achterhalen.

Als onderdeel van dit onderzoek werd een intensieve staalnamecampagne uitgevoerd tijdens het einde van het droge seizoen in 2013 in Ecuador. Data over biologie (macroinvertebraten), fysicochemie en hydromorfologie werden verzameld op 120 staalnamelocaties in het Guayas rivierbekken, Ecuador. In totaal werden er 39 milieuvariabelen opgemeten.

Om de huidige ecologische waterkwaliteit te evalueren werden er twee biotische indices berekend (**hoofdstuk 4**): de Biological Monitoring Working Party aangepast aan Colombia (BMWP-Col) en de Neotropical Low-land Stream Multimetrix Index (NLSMI). De ecologische waterkwaliteit van het Guayas rivierbekken varieerde tussen zeer slecht (0) en goed (168) volgens de BMWP-Col en tussen slecht (0) en referentiewaarde (9.1) volgens de NLSMI. Nutriëntenconcentraties waren over het algemeen lager dan de detectielimieten van de kits gebruikt voor nutriëntenmetingen, waardoor de nutriëntenconcentraties niet bepaald konden worden in de meeste sample locaties. De resultaten tonen ook aan dat de BMWP-Col een meer geschikte maat is om de ecologische waterkwaliteit te evalueren dan de NLSMI daar NLSMI rivier-specifiek bedoeld is voor kleine stromen gesitueerd op minder dan 250m boven zeeniveau.

De sleutelvariabelen die de ecologische waterkwaliteit van het Guayas rivierbekken het meest beïnvloeden, worden besproken in **hoofdstuk 4** en **hoofdstuk 5**. De relatie tussen de aanwezigheid van macroinvertebraten (a.d.h.v. BMWP-Col) en milieuvariabelen werd bestudeerd door middel van multivariate analyses (**hoofdstuk 4**) en regressieanalyses (**hoofdstuk 5**). De multivariate analyses onthulden dat stroomsnelheid, sliblaag, chlorofyl concentratie, sediment type, conductiviteit en landgebruik een significante invloed hadden op de distributie van macroinvertebraten taxa (alle p-waarden < 0.001). Uit de regressie- en gevoeligheidsanalyses bleek echter dat hoogteligging, nitraat-N, chlorofyl a concentratie, de hoekigheid van het sediment, aanwezigheid van boomstammen en macrofyten, stroomsnelheid, turbiditeit, vorm van de oeverbank en landgebruik de voornaamste variabelen waren die de BMWP-Col beïnvloedden (alle p-waarden $<$ 0.05 behalve chlorofyl a, $p = 0.064$). Deze resultaten bevestigen de significante invloed van fysisch-chemische variabelen op de macroinvertebratengemeenschap en onderstreepten het belang van landbouw-gerelateerde variabelen en landgebruik voor de ecologische waterkwaliteit in het Guayas rivierbekken.

Om te onderzoeken welke manier van verzamelen van landgebruiksdata het best het effect van landgebruik op de ecologische waterkwaliteit simuleert, werden drie methoden en bronnen gebruikt: veldprotocols die het landgebruik evalueerden binnen een gebied van 100 m x 10 m, Google maps (100 m x 100 m) en GIS data

(200 m x 200 m). Dit werd gedaan voor zowel de linkeroever als de rechteroever van de sites. Regressieanalyses werden verricht tussen elke methode of bron en de gemeten milieuvariabelen (**hoofdstuk 6**). De resultaten toonden aan dat het effect van landgebruik op de ecologische waterkwaliteit het best werd gekwantificeerd door middel van Google maps (R^2 = 0.93, $p < 0.05$). Daarnaast zijn modellen met informatie over landgebruik, gebruik makend van Google maps, voornamelijk gelinkt met fysisch-chemische variabelen terwijl modellen met landgebruiksdata verkregen via veldprotocols en GIS data meer gelinkt waren met hydromorfologische variabelen.

Curriculum vitae

Personal information

Name	Minar Naomi
Family name	Damanik Ambarita
Date of birth	21 September 1978
Place of birth	Pematangsiantar, North Sumatera, Indonesia
Nationality	Indonesian
E-mail	minarnaomi.damanikambarita@ugent.be; minardamanik@yahoo.com

Education

2013-present	Doctoral studies in the Applied Biological Sciences, Ghent University, Belgium Thesis: Ecological impact assessment of land use on river systems Promoters: Prof. Dr. ir. Peter Goethals, Dr. ir. Gert Everaert
2011-2013	Master of Science in Environmental Sanitation, Ghent University, Belgium Thesis: Monitoring and modelling of the ecological water quality of the Cuenca rivers in Ecuador. Promoter: Prof. Dr. ir. Peter Goethals
1997-2002	Bachelor in Agriculture Engineering, Bogor Agricultural University (IPB), Indonesia Thesis: Unit hydrograph change as the result of land conservation at Upper Ciliwung river basin, West Java. Promoter: Prof. Dr. ir. H. Soedodo Hardjoamidjojo, MSc.

Work experience

Jan 2009-Sep 2011	*Disaster Management Officer*, **German Red Cross (GRC)**, Jakarta
Apr 2007-Jun 2008	*District Liaison (PMI Training and Support) Senior Officer*, **International Federation of Red Cross and Red Crescent Societies (IFRC)**, Yogyakarta

Oct 2006-Mar 2007	*Training and PMI Support Early Recovery Senior Officer*, **International Federation of Red Cross and Red Crescent Societies (IFRC)**, Yogyakarta
Sep 2005-Sep 2006	*Water and Habitat Field Officer*, **International Committee of the Red Cross (ICRC)**, Jakarta, covering Indonesia
Apr-Aug 2005	*Tracing Field Officer*, **International Committee of the Red Cross (ICRC)**, Jakarta, covering Indonesia
Oct 2004-Jan 2005	*Community Development Trainee*, **World Vision Indonesia**, Jakarta and Sumba (East Nusa Tenggara)
Aug-Sep 2004	*Management Trainee*, **Hoka Hoka Bento Restaurant**, Jakarta
Jul-Aug 2001	*Internship* with focus on plant's irrigation usage, **Vegetables Research Institute**, Lembang (Bandung, West Java)

Workshops and training

Dec 2015-Apr 2016	Communication, negotiation and conflict handling skills courses, Ghent
Mar 2016	Poster presentation at Belmundo Water Talks, Ghent
Jan 2016	Project management and leadership foundation courses, Ghent
May, Nov-Dec 2015	Creative thinking and personal effectiveness courses, Ghent
Feb-May 2015	Advanced academic English writing skills and conference skills courses, Ghent
Oct-Nov 2014	Writing and reviewing scientific literature course, Ghent
Oct 2012	Water & climate day seminar, Ghent
Jul 2011	Lessons learned on tsunami awareness and education materials adaptation & development, Jakarta
May 2011	Climate change and community based disaster workshop, Dhaka
Dec 2010	National conference on community based disaster risk education, Jakarta
Nov 2009	Monitoring and evaluation workshop, Bengkulu
Mar 2009	SPHERE training, Semarang
Jan 2009	Red Cross orientation, Jakarta
May 2007	Assessment training, Salatiga
Sep 2006	Working in ICRC course, Jakarta
Nov 2005	Leadership in strategic health communication workshop, Banda Aceh

Organization

2011-2013	Students' representative of Master of Science in Environmental Sanitation, UGhent
1999-2000	Service Division Coordinator of Christian Special Service Commission, UK-PMK, IPB
1997-2002	Member of Agriculture Engineering Students Association (Himateta), IPB
1997-2002	Member of Christian Special Service Commission, UK-PMK, IPB

Publications

Minar Naomi Damanik Ambarita, Gert Everaert, Peter L.M. Goethals. (2018). Ecological models to infer the quantitative relationship between land use and the aquatic macroinvertebrate community. Water 10(184). DOI: 10.3390/w10020184.

Marie Anne Eurie Forio, Ans Mouton, Koen Lock, Pieter Boets, Thi Hanh Tien Nguyen, **Minar Naomi Damanik Ambarita**, Peace Liz Sasha Musonge, Luis Elvin Dominguez Granda, Peter L.M. Goethals. (2017). Fuzzy modelling to identify key drivers of ecological water quality to support decision and policy making. Environmental Science and Policy 67, 58-68. DOI: 10.1016/j.envsci.2016.12.004.

Minar Naomi Damanik Ambarita, Gert Everaert, Marie Anne Eurie Forio, Thi Hanh Tien Nguyen, Koen Lock, Peace Liz Sasha Musonge, Natalija Suhareva, Luis Elvin Dominguez Granda, Elina Bennetsen, Pieter Boets, Peter L.M. Goethals. (2016). Generalized Linear Models to Identify Key Hydromorphological and Chemical Variables Determining the Occurrence of Macroinvertebrates in the Guayas River Basin (Ecuador). Water 8(7). DOI: Artn 297 10.3390/W8070297.

Minar Naomi Damanik Ambarita, Koen Lock, Pieter Boets, Gert Everaert, Thi Hanh Tien Nguyen, Marie Anne Eurie Forio, Peace Liz Sasha Musonge, Natalija Suhareva, Elina Bennetsen, Dries Landuyt, Luis Elvin Dominguez Granda, Peter L.M. Goethals. (2016). Ecological water quality analysis of the Guayas river basin (Ecuador) based on macroinvertebrates indices. Limnologica 57, 27-59. DOI: 10.1016/j.limno.2016.01.001.

Marie Anne Eurie Forio, Dries Landuyt, Elina Bennetsen, Koen Lock, Thi Hanh Tien Nguyen, **Minar Naomi Damanik Ambarita**, Peace Liz Sasha Musonge, Pieter Boets, Gert Everaert, Luis Elvin Dominguez Granda, Peter L.M. Goethals. (2015). Bayesian belief network models to analyse and predict ecological water quality in rivers. Ecological Modelling. 312, 222-238. DOI: 10.1016/j.ecolmodel.2015.05.025.

Thi Hanh Tien Nguyen, Pieter Boets, Koen Lock, **Minar Naomi Damanik Ambarita**, Marie Anne Eurie Forio, Peace Liz Sasha Musonge, Luis Elvin Dominguez Granda, Thu Huong Thi Hoang, Gert Everaert, Peter L.M. Goethals. (2015). Habitat suitability of the invasive water hyacinth and its relation to water quality and macroinvertebrate diversity in a tropical reservoir. Limnologica 52, 67-74. DOI: 10.1016/j.limno.2015.03.006.

Minar Naomi Damanik Ambarita, Pieter Boets, Thi Hanh Tien Nguyen, Marie Anne Eurie Forio, Gert Everaert, Koen Lock, Peace Liz Sasha Musonge, Natalija Suhareva, Elina Bennetsen, Sacha Gobeyn, Tuan Long Ho, Luis Elvin Dominguez Granda, Peter L.M.

Goethals. Impact assessment of local land use on ecological water quality of the Guayas river basin (Ecuador). Ecological Informatics (under review).

Grants

2013-2017	Scholarship from the Special Research Fund of Ghent University (BOF) for PhD at Ghent University
2011-2013	Scholarship from Flemish Interuniversity Cooperation (VLIR-UOS) for Master of Science in Environmental Sanitation, Ghent University